CORRELATION OF CYBER AND PHYSICAL DYNAMICS IN POWER GRID RELAYS FOR MALWARE DETECTION

by

DAFER RZOK ALALI

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE AND ENGINEERING

2025

Oakland University
Rochester, Michigan

Doctoral Advisory Committee:

Julian Rrushi, Ph.D., Chair
Mohamed A. Zohdy, Ph.D.
Mohammad-Reza Siadat, Ph.D.
Hua Ming, Ph.D.
Li Li, Ph.D.

Dedication

*This dissertation is dedicated to my **parents**, whose unconditional love, sacrifices, and unwavering support have shaped my journey and made this achievement possible. Your guidance and encouragement have been my foundation, and I am forever grateful for everything you have done for me.*

*To my **beloved wife**, whose patience, love, and unwavering belief in me have been my greatest source of strength. Your support and encouragement throughout this journey have been invaluable, and I could not have done this without you by my side.*

*To my **advisor**, whose mentorship, wisdom, and constant guidance have played a crucial role in shaping this research. Your support, constructive feedback, and belief in my potential have been instrumental in reaching this milestone.*

Thank you all for your encouragement, inspiration, and unwavering faith in me. This achievement is as much yours as it is mine.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to **my parents**, whose unwavering support, sacrifices, and encouragement have been the foundation of my academic journey. Your belief in me has been my greatest motivation, and I am forever grateful for the values of perseverance and hard work that you have instilled in me.

To my **beloved wife**, your patience, love, and constant support have been my anchor throughout this challenging yet rewarding journey. Your encouragement and understanding have helped me navigate the most demanding moments, and for that, I am profoundly thankful.

I extend my sincere appreciation to my **advisor**, Dr. Julian Rrushi, and **My co advisor** Dr. Mohamed Zohdy for your invaluable guidance, mentorship, and encouragement throughout my research. Your insightful feedback, constructive criticism, and unwavering support have been instrumental in shaping this dissertation. I am truly grateful for your time, patience, and belief in my work.

I would also like to acknowledge my **committee members**, colleagues, and friends who have provided thoughtful advice, insightful discussions, and continuous encouragement along the way. Your support has been an essential part of my academic and personal growth.

Finally, I am grateful to **Oakland University** and the **Department of Computer Science and Engineering** for providing me with the opportunity and resources to pursue this research.

To everyone who has contributed to my success, whether directly or indirectly, thank you. This journey would not have been possible without your support, and I am deeply appreciative of all your efforts.

DAFER RZOK ALALI

ABSTRACT

CORRELATION OF CYBER AND PHYSICAL DYNAMICS IN POWER GRID RELAYS FOR MALWARE DETECTION

by

DAFER ALALI

Adviser: Julian Rrushi, Ph.D.

Exploits and Malicious operations modules make up the malware application that targets the electrical power system. The exploits resemble those found in conventional malware for general-purpose computing. These malicious programs infiltrate an industrial computer, i.e. relay and then release functional components. In order to take over computational functions of the relay, malware run physics-aware modules that target physical equipment. An example is fabrication of fictitious status power data that indicates a power transformer is functioning normally, but in reality, the attacks are causing a malfunction in the transformer's HARMONIC PROTECTION algorithm. This research explores the relationship between harmonic activities in power transformers and their impact on system behavior in substation computers. We privilege mimic a set of harmonic malwares and a power transformer. In this study, we contribute in multiple ways: 1) we use these emulations to examine a power transformer's cyberattack surface; 2) using these insights, we develop a number of attacks that harmonic malware could employ against a power transformer. 3) we use Python to simulate these cyberattacks in

order to monitor and measure their harmful effects on a power transformer empirically. Also, we used HYPERSIM Simulation provided by OPAR-R, to compare it with our python simulation 4) we use the ProcMon tool to dynamically observe system activities; and v) we use machine learning models to forecast and assess the impact of cyberattacks on computer systems.

The results demonstrated that the machine learning model achieved high accuracy in detecting malware-induced anomalies, with the Python simulation yielding a 91% accuracy and the HYPERSIM simulation achieving 86% accuracy in predicting cyber-physical disturbances for the type of activity “Results”. Confusion matrix analyses revealed a strong correlation between harmonic distortions and malware activity, validating the effectiveness of frequency-domain analysis in anomaly detection. Furthermore, the model successfully differentiated between normal harmonic fluctuations and malicious injections, reducing false positives while maintaining high detection precision.

This study highlights the critical role of machine learning in enhancing cybersecurity resilience in power grids. By integrating real-time anomaly detection and dynamic system activity monitoring, the proposed framework offers a robust defense mechanism against evolving cyber threats targeting power transformer operations. The findings provide a foundation for future research in developing AI-driven cybersecurity solutions tailored to industrial control systems.

TABLE OF CONTENTS

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

TABLE OF CONTENTS—Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES—Continued

CHAPTER ONE
INTRODUCTION

The digitization of the modern power grid introduces both operational efficiencies and cybersecurity risks, making it a target for potential cyberattacks. Power transformers, being a critical part of substations, play a vital role in ensuring the reliable distribution of electrical energy. However, they are increasingly vulnerable to sophisticated cyberattacks that could severely disrupt energy distribution and stability. According to [1], a cyber-physical system (CPS) typically integrates industrial computers with physical machinery through I/O boards. These boards facilitate the monitoring and control of physical equipment by collecting sensor data and issuing commands to actuators, allowing industrial computers to adjust and regulate the physical state of operations. The interaction between the digital and physical worlds in CPSs is commonly governed by feedback loops, where the behavior of physical processes influences computations and vice versa. As discussed in [2], CPSs are inherently complex, operating across multiple temporal and spatial scales, and they rely on robust, networked communication to connect computational elements with physical systems.

In the electric power sector, advancements in computer-based systems have significantly expanded the capabilities of energy control centers. Research efforts dating back to the 1960s, documented by [3], explored the use of digital computers for power system protection, a field that gained momentum with the advent of industrial computers. The introduction of microprocessors, which replaced traditional electromechanical and

solid-state relays, brought numerous benefits, including enhanced data analytics, fine-tuning of algorithms for better performance, and cost efficiency, while still adhering to the principle of decentralized protection [4]. These innovations have gradually evolved into modern CPSs, which offer more extensive monitoring and control capabilities, even over long distances.

However, with the adoption of microprocessor-based systems for physical process control, new security vulnerabilities have emerged. Threat actors can exploit weaknesses in industrial computers to access sensitive data or cause physical damage. These industrial computers, while specialized with I/O boards, real-time operating systems, dedicated software, and network protocols, remain vulnerable to sophisticated cyberattacks. Malware used against electrical power grids often functions as automated attack code, similar to malware in general-purpose computing. Such malware typically exploits memory vulnerabilities and employs techniques like heap spraying to circumvent security mechanisms like address space layout randomization.

Despite overlapping exploitation strategies, the functionality of malware targeting industrial systems differs from that designed for general-purpose computing. Malware aimed at general-purpose systems often includes features for logging keystrokes, monitoring secondary storage, activating audio and video surveillance devices, capturing network traffic, and extracting sensitive information, such as passwords or cryptographic keys, from memory. In contrast, malware that targets the electric power grid has operational components specifically designed to interfere with and disrupt physical processes, focusing heavily on manipulating physical systems.

The in-depth understanding that our work gives cybersecurity defenders is frequently necessary to provide a more effective defense against malware in particular power grid equipment, such as a power transformer. This dissertation concentrates on Harmonic malware; however, the findings and comparable principles apply to different kinds of power grid equipment.

1.1 Overview of Power Grids

Power grids are a cornerstone of modern society, providing the electricity necessary to power homes, businesses, and industries. They are complex networks consisting of generation, transmission, and distribution systems. Electricity is generated at power plants using various energy sources, such as fossil fuels, nuclear energy, or renewable resources like solar and wind. The generated power is then transmitted over long distances through high-voltage transmission lines and distributed to end-users via local distribution networks. The seamless operation of these interconnected systems ensures the reliable supply of electricity, making power grids essential for economic growth, technological advancement, and societal well-being. Disruptions to power grids can result in widespread blackouts, economic losses, and compromised safety, underscoring their critical importance.

1.1.1 Digitalization of Power Systems

The advent of digital technologies has revolutionized power systems, introducing unprecedented levels of efficiency, reliability, and automation. Traditional electromechanical relays have been replaced by digital relays that offer enhanced precision and flexibility. Advanced metering infrastructure (AMI) enables utilities to

monitor energy consumption in real time, facilitating demand-response strategies and improving energy efficiency. Supervisory Control and Data Acquisition (SCADA) systems provide centralized control and monitoring of power grid operations, allowing for rapid identification and resolution of issues. Additionally, the integration of Internet of Things (IoT) devices has enabled predictive maintenance, fault detection, and dynamic grid management. However, the increasing reliance on digital technologies has also introduced vulnerabilities, as these systems are susceptible to cyberattacks, data breaches, and software malfunctions.

1.1.2 Cybersecurity in Power Systems

As power grids become more interconnected and reliant on digital systems, they face escalating risks from cyber threats. Cyber adversaries, ranging from individual hackers to nation-state actors, can exploit vulnerabilities in software, communication protocols, and hardware to disrupt grid operations. Such attacks can lead to data manipulation, unauthorized access, and even physical damage to critical infrastructure. The 2015 Ukraine power grid cyberattack serves as a stark reminder of the potential consequences of inadequate cybersecurity measures. During this incident, attackers used sophisticated malware to compromise the control systems of multiple substations, leaving hundreds of thousands of people without electricity for hours. Other notable threats include ransomware attacks targeting utilities and Distributed Denial of Service (DDoS) attacks that overwhelm communication networks. The growing sophistication of cyber threats necessitates robust cybersecurity frameworks, including intrusion detection

systems, threat intelligence sharing, and regular vulnerability assessments, to safeguard the integrity and resilience of power grids.

1.2 Physical Dynamics of Power Systems

Power systems operate based on fundamental physical principles, such as Ohm's law, Kirchhoff's laws, and the principles of electromagnetic induction. Key components like transformers, relays, circuit breakers, and generators ensure the stable and efficient transmission and distribution of electricity. Transformers step up or step-down voltage levels to facilitate long-distance transmission and safe local distribution, while circuit breakers protect equipment by interrupting current flow during fault conditions. Protective relays play a crucial role in maintaining system stability by detecting anomalies, such as short circuits or overloads, and triggering corrective actions. These physical processes are deeply intertwined with the digital systems that monitor and control them. For example, digital relays rely on real-time data from sensors to make decisions, and SCADA systems use physical measurements to adjust operational parameters. Disruptions in the cyber domain, such as malicious control signals or data manipulation, can directly impact on the physical behavior of the grid, leading to equipment damage, power outages, or even cascading failures. Understanding and addressing these interdependencies is essential for ensuring the resilience and reliability of modern power systems.

1.3 Motivation

1.3.1 Significance of Correlating Cyber and Physical Dynamics

The convergence of cyber and physical systems in power grids has created a dual-layered operational environment where disruptions in one domain can significantly impact the other. Understanding this correlation is critical for identifying vulnerabilities and implementing comprehensive protective measures. For example, a cyberattack targeting the control logic of digital relays can alter their behavior, resulting in delayed or incorrect responses to physical faults, which can escalate into widespread grid failures. Similarly, physical disturbances such as equipment faults can propagate into the cyber domain, causing cascading effects like erroneous data or malfunctioning control commands. By correlating cyber and physical dynamics, system operators can detect anomalies earlier, mitigate risks more effectively, and maintain operational stability under complex and dynamic conditions.

1.4 Examples of Cyberattacks on Power Grids

Real-world cyberattacks on power grids highlight the critical need for robust cybersecurity measures:

1. **The 2015 Ukraine Power Grid Attack**: A coordinated cyberattack compromised the control systems of multiple substations, cutting off electricity for approximately 230,000 people. Attackers used malware and remote access to manipulate control operations, demonstrating the vulnerability of interconnected grid systems.

2. **Industroyer Malware**: Also known as CrashOverride, this malware specifically targets industrial control systems, including those in power grids. It was designed to interact directly with substation communication protocols, potentially causing extensive outages and even physical equipment damage.
3. **BlackEnergy Malware**: Used in attacks on Ukrainian infrastructure, this malware served as a platform for espionage and disruption, exploiting software vulnerabilities to infiltrate critical systems. It laid the groundwork for subsequent attacks by enabling unauthorized access to sensitive control environments.
4. **Stuxnet**: While primarily targeting nuclear facilities, this malware highlighted the potential for sophisticated, state-sponsored attacks on critical infrastructure. Its ability to manipulate physical equipment through cyber vulnerabilities underscores the dangers posed by advanced persistent threats.
5. **SolarWinds Attack**: Although not specifically targeting power grids, this supply chain attack demonstrated how compromised software updates could infiltrate critical systems and allow adversaries to manipulate operations stealthily. These examples underscore the devastating consequences of cyberattacks and the importance of integrating cybersecurity measures into the physical and operational aspects of power systems.

1.5 Importance of Proactive Malware Detection

Proactive malware detection is essential for preventing and mitigating the impact of cyberattacks on power grids. Traditional reactive approaches often detect threats after

they have already caused significant damage. In contrast, proactive measures focus on identifying and neutralizing threats before they can escalate. This involves:

1. **Real-Time Monitoring**: Utilizing advanced tools and sensors to continuously monitor cyber and physical activities for anomalies. For instance, detecting unusual data traffic patterns or deviations in system behavior.
2. **Machine Learning Models**: Leveraging algorithms that can identify patterns indicative of malware, including subtle changes in system behavior or communication protocols. For example, models trained on historical attack data can predict potential threats by recognizing early warning signs.
3. **Predictive Analytics**: Analyzing historical and real-time data to forecast potential threats and vulnerabilities. Predictive models can simulate attack scenarios and identify weak points in the system that require fortification.
4. **Incident Response Planning**: Developing comprehensive response strategies to address potential attacks swiftly and effectively. This includes automated fail-safes, manual override protocols, and coordinated efforts between cyber and physical security teams.
5. **Red-Teaming and Penetration Testing**: Proactively identifying vulnerabilities through simulated attacks, allowing organizations to strengthen defenses before actual adversaries exploit them. By adopting a proactive approach, utilities can enhance the resilience of power grids, ensuring the continued delivery of electricity while protecting critical infrastructure from emerging threats. This proactive stance not only minimizes downtime and economic losses but also

builds confidence in the grid's ability to withstand and recover from cyber-physical adversities.

1.6 Problem Statement

1.6.1 Identifying Gaps in Current Research and Challenges in Detecting Cyber-Physical Anomalies

Despite the advancements in cybersecurity and power grid operations, significant gaps remain in detecting cyber-physical anomalies effectively. Current research often treats the cyber and physical layers as distinct entities, failing to fully integrate the interplay between these domains. This segmentation overlooks the cascading effects that disruptions in one layer can have on the other, which is critical for comprehensive anomaly detection.

1.7 Research Objectives

1.7.1 Primary Objective

To develop a comprehensive framework that integrates cyber and physical dynamics for detecting and mitigating malware targeting power grid relays and transformers.

1.7.2 Specific Objectives

1. **Characterize the Cyber-Physical Interaction:**

- Analyze the interdependencies between cyber and physical components in power grid systems, particularly in digital relays and transformers.
- Identify key parameters that reflect these interdependencies and contribute to anomaly detection.

2. **Simulate Cyberattacks with Realistic Harmonic Injection Scenarios:**

- Use Python-based simulation tools to emulate various harmonic malware scenarios affecting power transformers.
- Utilize HYPERSIM simulation software provided by OPAL-RT to model distribution and industrial power systems, specifically a 25kV 24-bus feeder, for simulating various harmonic malware scenarios affecting power transformers.
- Incorporate harmonic analysis techniques such as Fast Fourier Transforms (FFT) to study the impacts on transformer protection mechanisms.

3. **Capture and Analyze System Activities:**

- Utilize dynamic analysis tools like ProcMon to observe system activities during harmonic malware injection.
- Extract and preprocess critical data attributes, including file operations, registry changes, and process activities.

4. **Develop Machine Learning Models:**

- Design and train models for anomaly detection using datasets generated from harmonic malware simulations.
- Evaluate model performance using metrics like accuracy, precision, recall, and feature importance analysis.

5. **Validate the Detection Framework:**

- Conduct experiments to assess the framework's ability to detect and mitigate malware in real-time or near-real-time conditions.

- Compare the framework's performance with existing methods to demonstrate its efficacy and scalability.

6. **Enhance Resilience Through Proactive Measures:**

- Propose and evaluate strategies for improving system resilience, such as real-time monitoring, predictive analytics, and adaptive responses to emerging threats.

By addressing these objectives, the research aims to provide a robust solution for safeguarding power grid infrastructure against evolving cyber-physical threats.

1.8 Scope of the Study

1.8.1 System Boundaries and Focus Areas

The scope of this study is focused on the interaction between cyber and physical systems within power grid infrastructure, particularly digital relays and power transformers. This work explores the vulnerabilities introduced by harmonic-based cyberattacks and proposes solutions for their detection and mitigation. The key boundaries and aspects addressed include:

1. **Target Components**:
   - Digital Relays: Investigating their role in maintaining grid stability and their susceptibility to cyber manipulation.
   - Power Transformers: Examining how harmonic disturbances influence their operation and the implications for protection mechanisms.
2. **Harmonic Malware Simulations**:

- Designing and implementing Python-based simulations to replicate harmonic malware attacks.
- Designing and executing HYPERSIM simulations to model harmonic malware attacks on power transformers.
- Simulating scenarios such as harmonic injections,$2^{nd,}$ 3^{rd}, 5^{th}, 11^{th}, suppression, and cumulative disturbances to evaluate their impact on system behavior.

3. **Data Capture and Analysis**:
    - Using dynamic monitoring tools like ProcMon to observe system-level impacts of harmonic attacks.
    - Capturing critical operational data, including file operations, registry activities, and process events.
4. **Machine Learning Integration**:
    - Developing and validating machine learning models to detect anomalies resulting from harmonic disturbances.
    - Focusing on feature extraction techniques and metrics that enhance the reliability and interpretability of detection frameworks.
5. **System Assumptions**:
    - The study assumes that the power grid operates under normal conditions without prior compromises.
    - It focuses on single-point failures and their cascading effects but does not consider large-scale grid blackouts caused by coordinated attacks.
6. **Exclusions**:

- Advanced Persistent Threats (APTs) and long-term espionage activities are beyond the scope.
- Hardware-level vulnerabilities in devices like relays and transformers are not addressed directly, as the focus is on cyber-physical interaction at the system level.

By clearly defining these boundaries, this study aims to provide actionable insights into the specific vulnerabilities and mitigation strategies related to harmonic-based cyberattacks, ensuring its findings are both practical and impactful for enhancing grid security.

1.9 Significance of Study

1.9.1 Academic Contributions

This research advances the understanding of cyber-physical systems by introducing a novel framework that integrates harmonic analysis with machine learning to detect and mitigate malware. It provides:

1. A detailed examination of the interplay between harmonic disturbances and cyber-physical interactions.
2. Insights into the development of data-driven models for anomaly detection in power grids.
3. A foundation for future research exploring other dimensions of cyber-physical threats.

1.9.2 Practical Contributions

The proposed framework offers utility operators tools to:

1. Detect anomalies in real-time by monitoring and analyzing harmonic patterns.
2. Enhance system resilience through predictive analytics and proactive measures.
3. Mitigate risks of malware attacks, ensuring uninterrupted power delivery and grid stability.

1.9.3 Policy Contributions

The study informs policymakers and regulators by:

1. Highlighting the critical need for integrated cybersecurity measures in power systems.
2. Proposing standards for monitoring and protecting against harmonic-based threats.
3. Encouraging collaboration between utilities, governments, and technology providers to establish robust defenses.

1.10 Structure of the Dissertation

This dissertation is organized into seven chapters:

1. **Introduction**: Outlines the background, motivation, problem statement, objectives, scope, significance, methodology, and structure.
2. **Literature Review**: Examines existing research on cyber-physical dynamics, harmonic-based threats, and machine learning for anomaly detection.
3. **Methodology**: Details the research design, tools, and techniques used to achieve the objectives.

4. **Dynamic System Activities Capture**: Presents findings from simulations and evaluates the proposed framework.
5. **Machine Learning Model Development:** Introduce our machine learning models and the results of each model.
6. **Discussion**: Interprets the results, discusses implications, and compares with prior research.
7. **Conclusion and Recommendations**: Summarizes key findings, highlights contributions, and suggests directions for future work.

CHAPTER TWO
LITERATURE REVIEW

This chapter provides a detailed exploration of the existing body of literature related to the correlation of cyber and physical dynamics in power grids for malware detection. The review spans advancements in cyber-physical systems (CPS), the evolving nature of malware threats to power infrastructure, and the utilization of machine learning (ML) methodologies for anomaly detection. These foundational studies illuminate the gaps and challenges that this dissertation aims to address, particularly focusing on the vulnerabilities introduced by harmonic-based cyberattacks on power transformers.

2.1 Cyber-Physical Systems in Power Grids

Cyber-physical systems (CPS) integrate computational and physical processes, enabling real-time monitoring, optimization, and control of power systems. While CPS technologies have revolutionized the efficiency and reliability of power grids, they have also introduced significant vulnerabilities due to the interdependence of cyber and physical components.

2.2 Related work

2.2.1 Physics Cyberattacks

2.2.1.1 Physics-Centric Cyberattacks on Power Transformers

A groundbreaking study on physics-centric cyberattacks was conducted by [1], focusing on their impact on power transformers. Their research showcased how malware could exploit the physics of transformers to induce harmonic distortions, leading to

operational anomalies or equipment damage. The study employed Python-based emulations of transformer protection algorithms, demonstrating how harmonic injections could mimic legitimate signals. This approach highlighted a critical vulnerability where malicious actors could bypass conventional protection mechanisms by leveraging harmonic patterns. The authors concluded that an in-depth understanding of transformer dynamics is vital for developing robust detection and mitigation strategies/

2.2.2 Reliability Assessment in Cyber-Physical Systems

A hybrid analytical approach was introduced by [2], to assess the reliability of CPS under coordinated cyberattacks. By integrating cyber and physical failure probabilities, the study offered a comprehensive evaluation of power grid stability during malicious events. This dual-layered assessment emphasized the necessity of understanding interdependencies between the cyber domain (e.g., communication protocols) and the physical domain (e.g., transformer operations) to ensure resilience.

2.2.3 Interactive Visualization of Malware Behavior

Nguyen et al [10]. (2022) developed MalView, an interactive visual analytics platform designed to analyze complex malware behaviors. MalView enables users to explore malware dynamics through time-series and process-mapping visualizations, revealing attack patterns and dependencies. This tool is particularly beneficial for understanding zero-day attacks and improving real-time anomaly detection systems in power grids.

2.3 Malware Threats to Power Systems

The integration of operational technology (OT) and information technology (IT) in power systems has significantly increased their susceptibility to malware attacks. Advanced malware leverages these integrations to disrupt critical infrastructure by targeting both cyber and physical layers.

2.4 Prominent Malware and Their Impact

2.4.1 Industroyer and BlackEnergy

Industroyer, also known as CrashOverride [27], specifically targets industrial control system (ICS) protocols such as IEC 104 to manipulate circuit breakers. BlackEnergy, in contrast, focuses on disabling ICS workstations and networks through distributed denial-of-service (DDoS) attacks. Both were pivotal in the cyberattacks on Ukraine's power grids, demonstrating how malware can exploit ICS vulnerabilities to disrupt operations and cause widespread blackouts.

2.4.2 Aurora Generator Test

The Aurora Generator Test [28], conducted by Idaho National Laboratory, demonstrated the physical consequences of cyberattacks. By manipulating circuit breakers to desynchronize generator operations, attackers were able to induce mechanical stress, resulting in catastrophic equipment damage. This test highlighted the urgency of securing synchronization mechanisms in power systems to prevent similar attacks.

2.4.3 HoneyICS Framework

HoneyICS [29], serves as a honeypot designed to simulate OT environments and study attacker behavior in a controlled setting. By capturing threat actor techniques and strategies, HoneyICS provides valuable data for developing targeted defenses. It enables researchers to test malware impacts without endangering real-world infrastructure, thus advancing the security of power systems.

2.4.4 Anomaly Detection Using Machine Learning

Machine learning (ML) has emerged as an indispensable tool in detecting anomalies and mitigating malware threats in power grids. The integration of ML models with harmonic analysis and high-fidelity simulations has improved the detection of malicious activities in cyber-physical systems.

2.4.5 Autoencoder-Based Anomaly Detection

Autoencoders were leveraged by [30], a type of unsupervised learning algorithm, to detect anomalies in transformer differential protection relays. By analyzing deviations from normal patterns, the model effectively distinguished between legitimate and malicious behaviors. This method addressed a critical gap in protecting relay operations from targeted cyberattacks.

2.5 Harmonic Analysis for Malware Detection

The role of harmonic analysis in identifying malware-induced anomalies was emphasized by [1]. By monitoring harmonic patterns in transformer operations, their study demonstrated how ML models could detect deviations indicative of malicious

activity. Harmonics, like repetitive distortion patterns in electrical signals, can be exploited by attackers to mimic normal conditions while causing disruptions. These disturbances often involve the injection of odd or subharmonic frequencies that interact with the transformer's operational integrity.

Their research underscored the need for tools capable of analyzing these harmonic signatures to detect malware effectively. Using harmonic analysis combined with Fast Fourier Transforms (FFT), Olijnyk et al. quantified changes in waveform characteristics to pinpoint deviations triggered by malware. The findings indicate that harmonic anomalies often correlate with specific malware behavior, making them a reliable metric for early detection. Additionally, their approach revealed that harmonics could propagate through interconnected systems, exacerbating disruptions if not promptly addressed

2.6 High-Fidelity Simulations and ML Integration

Advanced simulation platforms like OPAL-RT HYPERSIM have been used to generate high-fidelity datasets for ML training. These simulations replicate real-world scenarios, enabling the development of robust ML models for anomaly detection. This integration has significantly improved the precision and reliability of malware detection systems in power grids.

2.7 Research Gaps and Challenges

Despite significant advancements, several critical gaps remain in the literature, hindering the development of comprehensive cybersecurity frameworks for power grids.

2.7.1 Identified Challenges

1. **Limited Cyber-Physical Integration**: Most studies focus on either cyber or physical components in isolation, failing to capture their interdependencies. This segmented approach limits the ability to detect and mitigate cyber-physical anomalies effectively.
2. **Underexplored Harmonic-Based Attacks**: The exploitation of harmonic distortions as a vector for cyberattacks is insufficiently researched. Addressing this gap requires a deeper understanding of how harmonics influence power transformer operations during malicious activities.
3. **Lack of Comprehensive Datasets**: The scarcity of labeled datasets for training ML models reduces the robustness and generalizability of anomaly detection frameworks. Enhanced simulation tools and real-world data collection are necessary to bridge this gap.
4. **Evolving Malware Sophistication**: As malware becomes increasingly sophisticated, traditional detection methods struggle to keep pace. Developing adaptive and predictive frameworks is crucial to address these evolving threats.

2.8 Summary

This chapter reviewed critical advancements in understanding cyber-physical systems, malware threats, and ML-based anomaly detection. Studies such as [1]) provided essential insights into harmonic-based attacks on power transformers, while tools like MalView and HoneyICS contributed to innovative solutions for malware analysis and defense. However, gaps in integrating cyber and physical analyses,

exploring harmonic attacks, and improving dataset quality highlight the need for comprehensive research approaches. This dissertation aims to address these challenges by proposing a framework that integrates harmonic analysis with ML techniques for robust malware detection in power grids.

CHAPTER THREE

METHODOLOGY

The simulations and data acquisition setup for this research focuses on examining the cyber-physical interactions within power grid relays, specifically targeting harmonic disturbances caused by malware. This section provides an in-depth description of the simulation's environment and data acquisition processes, as implemented in the study

3.1 Power transformer Simulation

Electricity is produced at power plants and sent to users via high-voltage transmission lines, as described in [3]. The requirement to boost line capacity and reduce energy losses during transmission—two critical components of energy conservation efforts—drives the usage of higher voltages. Power transformers are essential to this process because they make it possible for electricity to be transferred from production locations to end users in an effective manner.

A power transformer is a device that can accept an input voltage or current and output a changed version, either increased or decreased, according to [10]. The transformer is classified as a "step-up" or "step-down" transformer based on whether the output voltage is higher or lower than the input. The ratio of turns in two wire coils—referred to as the primary and secondary windings, wrapped around a conductive core accomplishes this change. The input current is provided to the main winding, and the converted current or voltage is delivered to the secondary winding. There is no direct

current passage between these windings since they are electrically isolated from one another.

Transformer functioning is supported by Faraday's law of induction, as explained in [13]. An electromotive force is induced in the secondary winding when an alternating current passes through the primary winding and creates a magnetic flux in the shared core. Based on the winding ratio Np to Ns, the resulting current Is in the secondary is proportionate to the original current Ip in the primary, as shown by the formula $IpNp = IsNs$. The effective transfer of electrical power between distinct circuits is made possible by the electromagnetic induction concept, which also permits the adjustment of voltage levels as necessary.

Prior to the emergence of cyber-physical systems (CPSs), which include internet-connected sensors and real-time controllers, power transformers were passive devices that relied only on electromechanical or solid-state components and did not require external controls. As a result, they were safe from cyber threats. However, with the advent of cyber-physical systems (CPSs), which include internet-connected sensors and real-time controllers, power transformers have become increasingly digitalized, and modern transformers are now outfitted with microcontrollers that monitor operations and communicate with central control systems via networks, making them vulnerable to remote cyberattacks.

Power transformers have often been placed as entirely passive components. They just used solid-state or electromechanical electronics to function and didn't need any outside control. A power transformer was inaccessible to cybercriminals as it lacked a microprocessor and a network connection. Threat actors may now remotely target devices

thanks to the growing popularity of CPSs, which enable real-time administration of system states and processes using internet-connected sensors and controllers. Specifically, power transformers are now frequently outfitted with microcontrollers that keep an eye on the transformer's numerous operations and communicate with the central control stations through network connections.

3.2 Failure Modes in the Power Transformer

Electric power and distribution systems are susceptible to failures or short circuits, as explained by [11]. A transformer's fault current flows through it when a fault arises on the load side. The rated currents of transformers are far lower than the fault currents flowing through them. In transformers, these currents cause heat and mechanical stress. When two or more conductors come into touch with one another when they are functioning normally and there is a potential difference between them, a short-circuit defect as defined by [4] takes place. Either an arc or a physically metallic touch might make contact. The voltage between the two components is lowered to zero in the metal-to-metal contact scenario. Nevertheless, an arc's voltage will be quite low. Insulation breakdown between turns in the same slot or between the winding and the machine's steel frame is the reason for a transformer failure, as covered in [4]. Deterioration of the insulation mixed with switching or lightning overvoltage is what caused the failure. Short-circuit faults are typically caused by lightning overvoltages and insulator contamination. The primary causes of this are overloading and aging.

Windings and tap changers have historically been among the leading causes of transformer failures, according to a study by [14]. According to [11], tap changer

malfunctions can be brought on by lubrication issues, limit witches, sheared pins on the linkage that controls the reversing switch, clearance-related thermal failures from coking or crimp issues, dielectric failures from poor oil quality, and mechanical failures from contact wear and misalignment.

According to [4], power transformer failures can have several negative effects. In certain areas of the system, abnormally high currents flow, which leads to component overheating. The current may be far higher than the conductors' intended thermal capacity in the power lines or equipment that is causing the incident. Consequently, the annealing of conductors and insulation may result in damage due to the temperature increase.

Additionally, equipment malfunctions due to the low voltage in the vicinity of the fault. System voltages will be lower than what is often considered acceptable, which could lead to equipment damage. Some or all the equipment will function incorrectly as a result of the system's components operating as unbalanced three-phase systems. This is a tiny sample of the mistakes brought on by a power transformer malfunction.

3.3 Power System Protection

To reduce potential damage or power outages, operators must take the required steps to keep transformer faults from expanding. To guarantee maximum electrical supply continuity with the least amount of damage, a protection system regularly checks the power system. By continuously monitoring factors including current, voltage, power, frequency, and impedance, a protective system can identify fault conditions.

Instrument transformers of the potential type (PT) or current type (CT) detect voltages and currents.

According to [4], a relay is a device that controls the operation of other devices under electric control by opening and closing electrical contacts. Unwanted or unbearable conditions within a designated region are detected by the protective system. In order to safeguard people and property from harm, the protection system then activates the relay to turn on the relevant circuit breakers and disconnect the impacted area. According to [4], relays are categorized based on whether they are measuring or on-off relays. Relays such time-lag relays, auxiliary relays, and tripping relays are included in the latter type, which is also referred to as all-or-nothing.

The relay system receives the measured variables from instrument transformers described in [4]. When a defect is detected, the relay system instructs a circuit-interrupting device called a relay or circuit breaker (CB) to disconnect the problematic area of the system. Every piece of equipment in an electric power system has its own safety zone. The divide is made in a way that minimizes service interruptions while providing sufficient protection for zones.

In this case, the relay is powered by a value that is either more than the operating or lower than the resetting value, and it lacks a fixed setting. One or more fault-detecting units and the required auxiliary units make up a relay.

Relay systems' fundamental components fall into one of four categories: computer-based, solid-state, electromechanical, or sequence networks.

This study focuses on digital computer-based relays, which are more susceptible to cyberattacks than electromechanical or solid-state systems. Key components include

analog and digital input subsystems, digital output subsystems, relay logic, settings, and digital filters [4].

3.4 The Hypothetical cyber attack

Developing a CPS attack surface demands in-depth knowledge of power transformers and their ecosystem. This analysis examines the operations of transformers, focusing on differential protection, harmonic restraint algorithms, and potential ways threat actors can induce mechanical damage. Namely, this study examines differential protection and harmonic restraint algorithms and explores how attackers can exploit system physics to inflict mechanical damage. The analysis also covers methods threat actors use to exploit the system's physics for mechanical damage.

These mechanisms are vital for preventing equipment damage and ensuring user safety. Attackers may exploit their understanding of these systems to bypass safety features or directly cause avoidable damage.

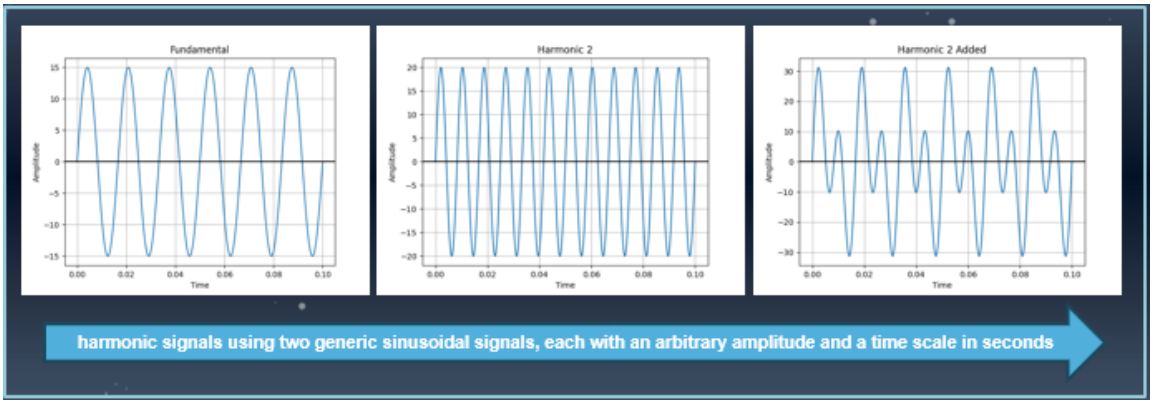

Figure 1: Harmonic signals using two generics sinusoidal.

3.5 Harmonic Protection Algorithm

Transient problems must be taken into account before the differential protection circuit can function dependably, according to [22], before the primary transformer reaches steady-state operation.

A flux must be established in the transformer core of every transformer. A current called the magnetizing current flows as a result of this flux. To the relay, the magnetizing current manifests as a differential current. A nonlinear magnetizing current waveform is the outcome of the core's nonlinearity. A significant magnetizing current may flow in response to a sudden shift in the excitation voltage.

Additionally, [22] notes that the peak values of the core flux wave rely on the residual flux and the switching time when a transformer is inserted into the circuit at any point during the supply voltage wave. Core saturation limits the peak value of the flux which is higher than the corresponding steady-state value. The magnetizing current required to generate the core flux has no counterpart on the secondary side and can peak eight to ten times the typical full-load peak. This phenomenon, which manifests as an internal differential protection system defect, is known as the magnetizing inrush current.

According to [22], if the transformer is turned on when the supply voltage is zero, the greatest inrush will occur. This must be taken into consideration when designing differential relays for transformer protection in order to prevent the circuit breaker from tripping due to the magnetizing inrush current. To avoid tripping due to the high inrush currents, several approaches based on the inrush current's harmonic characteristics are employed.

Data acquisition is used to process current transformer (CT) and potential transformer (PT) signals, and an analog-to-digital converter (ADC) is used to evaluate the data.

A Fourier transform is then used to breakdown this signal, and a configurable threshold value is compared. The typically fundamental component, DC component, second harmonic, third harmonic, fourth harmonic, and fifth harmonic components are present during the magnetizing inrush situation as percentages of 100%, 55%, 63%, 26.8%, 5.1%, and 4.1%, respectively, according to [22].

Additionally, [22] explains how the second harmonic current component is reached by the magnetizing inrush current. Thus, the fundamental current ratio and the second harmonica are used to identify the transformer's magnetizing condition. The fundamental and all other harmonic components are extracted from no load to full load to defective current signals using the fast Fourier transform (FFT) technique. A given input signal's precise fundamental frequency components can be extracted using the full-cycle FFT technique. The FFT method classifies a scenario as a magnetizing inrush if the second harmonic component rises by more than 20% of the fundamental component [22].

According to [1], they first created the necessary input signal before putting the harmonic protection algorithm into practice. Alternating current (AC) at a predetermined frequency is commonly used to carry electrical current. This frequency is 60 Hz in the United States. They started by developing a function that can produce a series of sinusoidal data points with any given frequency and magnitude for our simulation implementation. They then created instances of a signal with harmonic frequencies and the resulting composite signal using that function.

Harmonics is a phenomenon where a fundamental frequency can be divided by an integer factor of n where *n* = 1,2, 3c when n = 1, this is the fundamental frequency. Harmonic frequencies are for *n* = 2, 3,42, 3….; A composite is obtained by adding up each of the harmonic signals.

In Figure 1, an illustration of the harmonic signal notion was provided by [1]. For demonstration purposes, they display two generic sinusoidal signals with a generic amplitude and an arbitrary time scale in seconds. In actuality, the signals could be electromagnetic waves, current, voltage, audio signals, etc. A frequency of 60 Hz and an amplitude of 15 units were chosen for the fundamental signal. The harmonic at n = 2 was designated as "Harmonic 2." As a result, the frequency of "Harmonic 2" is 120 Hz. Five units is the arbitrary amplitude that we set. The composite wave, which is represented by the notation "Harmonic 2 Added," is the consequence of adding these two signals together. Here, they discovered how the fundamental signal is distorted by the harmonic signal.

In this simulation [1] built the capability of producing sinusoidal functions, instances of harmonic frequencies, and the resulting composite waves into their simulation environment. This made it possible for them to model different input currents for the power transformer simulation, both with and without harmonics. They must then make use of an FFT library. A periodic signal can be broken down into sine and cosine components at certain frequencies and magnitudes using the FFT approach. They conducted harmonic analysis in the harmonic restraint protection algorithm using the FFT. They only get sine waves in this model since they are just assuming a sine wave.

An algorithm for safeguarding electrical systems is called harmonic restraint protection. The program predicts when a current transformer is not in steady state by using information from harmonic signals. In the presence of an inrush current, this is usually the case when the current transformer is first turned on or off and the magnetic cores still reach flux.

Harmonic distortions are known to happen during this time. We can determine that the current values that will be input to the differential protection circuit will undoubtedly result in the circuit returning a "“failure”" condition by detecting the harmonic distortions. In reality, this "fail" is a fake one. The differential protection produces a bogus "“failure”" if the harmonic restraint protection detects harmonics above a preset threshold. However, we assume that the protected device has reached a steady state as the harmonic signals have diminished, and the outputs of the differential protection method are no longer "false" if the harmonic restraining protection is not detecting harmonics over a specified threshold.

3.5.1 How Threat Actors Exploit Harmonic Algorithms

To compromise the harmonic protection algorithm, threat actors must replicate a continuous sine waveform. By sampling the input signal from the analog-to-digital converter, they can extract the waveform’s magnitude and frequency. If their goal is to neutralize harmonic signals, they would need to apply an FFT (Fast Fourier Transform) to analyze the signal's harmonic content. Using this information, they could generate a counteracting signal with matching frequency and opposite magnitude to cancel out the harmonics or the original input signal.

Alternatively, attackers could simplify their strategy by making assumptions about standardized power system frequencies, such as 60 Hz in the United States or 50 Hz in Europe, which vary slightly by application. With this knowledge, they could pre-generate high-magnitude harmonic signals and store them for playback. This approach would aim to trigger the harmonic restraint protection by introducing additional harmonics. However, precise magnitude matching is necessary for effective cancellation, making this method more suited for disruption rather than harmonic neutralization.

While this approach simplifies the attack setup, it is more effective for causing disruption than for precise cancellations. Harmonic cancellation requires exact replication of phase and magnitude for destructive interference. In contrast, injecting harmonics at high magnitudes can exploit the system's harmonic thresholds, effectively simulating conditions that violate operational constraints, such as transformer saturation or resonance.

Threat actors could also exploit other vulnerabilities in the system. For example, by introducing noise or intentional delays in signal processing, they could interfere with the ADC's sampling or the algorithm's FFT calculations, causing misinterpretation of the signal's harmonic content. Additionally, attackers might target specific frequency bands or harmonics that align with the protection system's thresholds, ensuring their signals are detected as harmful while bypassing normal operational tolerances.

3.6 Software simulation system

In our study, we utilized two simulation platforms: one developed in our lab using Python, and the other being HYPERSIM software provided by OPAL-RT.

3.6.1 Python Simulation Software.

Olijnyk and his team introduced their simulation in [1]. The simulation emulates various critical parts of a real power transformer system designed in Python, including the primary and secondary windings, tap changers, current transformers (CTs), and protective relays. The primary and secondary windings transform voltage and current levels, and the simulation models their behavior in response to operational and fault conditions. Tap changers, which adjust the transformer voltage ratio dynamically, are also modeled to capture the transformer's voltage regulation mechanism. The simulation incorporates current transformers (CTs) to measure the currents on both the primary and secondary sides, which are crucial for protection algorithms like differential protection as shown in [1].

The protection system is a vital part of the simulation, consisting of digital relays and algorithms that analyze input from the CTs. The differential protection algorithm compares currents from both windings to detect internal faults, while the harmonic protection algorithm identifies harmonic distortions caused by phenomena like magnetizing inrush currents. The simulation integrates the interaction between these components, allowing the study of how attacks on the transformer’s protection system can impact overall operation. This detailed modeling provides a comprehensive

understanding of how cyberattacks can exploit the physics and control mechanisms of power transformers [8].

3.6.2 Description of Simulation Setup and Data Acquisition

The simulation and data acquisition setup for this research focuses on examining the cyber-physical interactions within power grid relays, specifically targeting harmonic disturbances caused by malware. This section provides an in-depth description of the simulation environment and data acquisition processes, as implemented in the study. The simulated transformer models employ mathematical formulations based on Faraday's law to generate primary and secondary current interactions. This setup is capable of dynamically altering parameters such as current amplitude, frequency, and harmonic content.

3.6.3 Cyberattack Emulation

We developed harmonic malware using Python to simulate various attack scenarios targeting the power transformer's harmonic protection algorithm. These scenarios include both intermittent and cumulative harmonic injections to evaluate their short-term and long-term effects on system integrity. The aim is to assess how different attack patterns can disrupt the harmonic protection mechanisms and compromise the stability of the transformer.

The attack scenarios are implemented by modifying the sinusoidal input signals with carefully crafted harmonic components. For realism, random noise is added to the signals, replicating the environmental and operational variations encountered in real-world systems. The malware dynamically introduces specific harmonic orders, such as

the third, fifth, and seventh harmonics, which are common sources of distortion in power systems. These harmonics are strategically injected at varying magnitudes and phases to maximize their destabilizing impact.

In the intermittent injection scenario, harmonic signals are introduced in short bursts to mimic transient disturbances, testing the algorithm's ability to differentiate between genuine faults and temporary anomalies. Conversely, in the cumulative injection scenario, harmonic distortions are applied continuously over an extended period to evaluate the algorithm's performance under prolonged stress conditions and its capacity for sustained protection.

To simulate real-world cyberattacks, the malware can manipulate frequency domains and amplitude thresholds in the protection system. For instance, it may exploit system vulnerabilities by injecting harmonics near the algorithm's operational thresholds, ensuring the distortions are detected as faults or creating a condition where the protection system is overloaded. Additionally, harmonic malware is capable of triggering false positives to cause unnecessary trips or bypassing detection to allow harmful conditions to persist unnoticed.

These simulated scenarios highlight the vulnerabilities of harmonic protection algorithms and emphasize the need for robust countermeasures, including advanced filtering techniques, noise suppression, and anomaly detection algorithms that can differentiate between malicious and natural harmonic distortions. This work lays the foundation for testing and enhancing the resilience of power transformer protection systems against cyber-physical threats.

3.7 HYPERSIM Software Simulation

HYPERSIM is a real-time digital simulation software developed by OPAL-RT Technologies for modeling and analyzing power systems, power electronics, and industrial control systems. It is widely used for simulating high-fidelity electromagnetic transients in power grids, substations, and renewable energy systems [20]. See Figure 2.

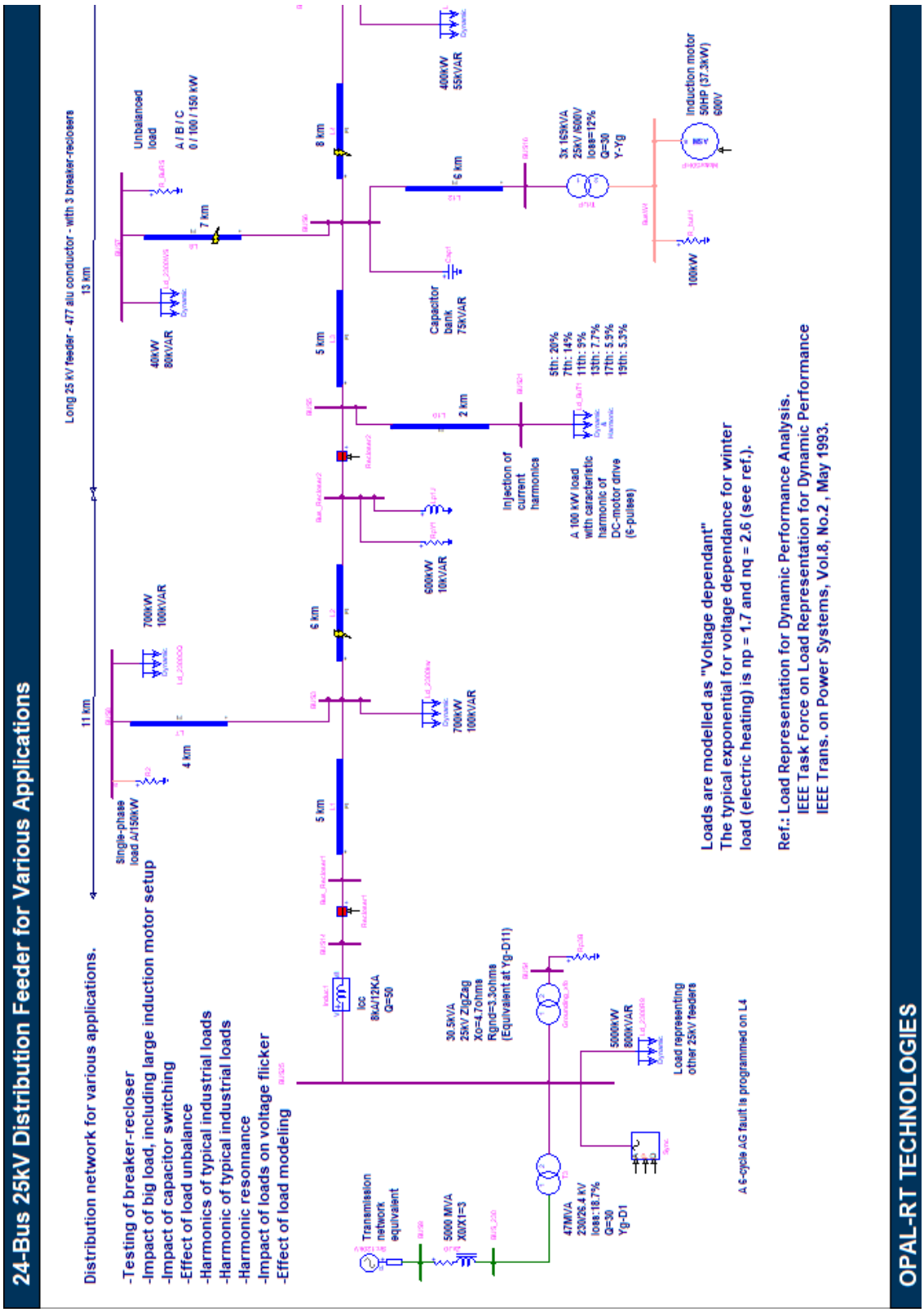

Figure 2: HYPERSIM SIMULTION FEEDER 25kV 24Bus.

3.7.1 Key Features of HYPERSIM

- **Real-Time Simulation:** Enables hardware-in-the-loop (HIL) testing for power system protection and control.
- **Electromagnetic Transient (EMT) Modeling:** Provides accurate transient analysis of electrical networks, transformers, and grid components.
- **Integration with Hardware:** Supports real-time testing of protective relays, converters, and microgrid controllers.
- **User-Friendly Interface:** Features an intuitive graphical modeling **environment** with built-in libraries for electrical components.
- **Cybersecurity Testing:** Allows for the simulation of cyberattacks on power systems, making it valuable for smart grid protection research.

3.7.2 Why Use HYPERSIM in Power System Research

HYPERSIM provides a high-fidelity environment for analyzing cyber-physical interactions in power grids. It enables researchers to study:

- The impact of malware-induced harmonic distortions on power transformers and relays.
- The response of protection systems to different types of cyberattacks.
- The effectiveness of anomaly detection models in mitigating threats to power grid infrastructure.

3.8 Feeder 25kV 24Bus

3.8.1 Introduction

A 24-bus, 25kV distribution feeder intended for testing different operational and fault scenarios in power systems is represented by the "Feeder_25kV_24Bus.ecf" simulation model, which is supplied by OPAL-RT's HYPERSIM program. In order to investigate distribution network behaviors, harmonic analysis, voltage regulation, and system stability under various loading scenarios, this model is frequently employed in both industry and research.

By simulating actual medium-voltage power networks, this distribution feeder allows researchers to evaluate safety devices, examine harmonics, and simulate electrical breakdowns. It is a useful tool for studying cyber-physical security because it enables researchers to introduce harmonic distortions caused by malware and track how they affect system performance and grid stability.

The 24-bus 25kV feeder mimics a real-world distribution network by including several substations, transmission lines, loads, transformers, capacitor banks, and circuit breakers. Several electrical circumstances can be modeled by the system, including: Long feeder lines that capture the effects of scattered loads and voltage dips (between 2 and 15 km).

- Real industrial operations are represented by several industrial loads, including resistive loads, voltage-dependent loads, and induction motors.

- Capacitor banks, which enhance power factors and assist with reactive power. Analysis of the network's resonance effects, and harmonic distortion is made possible by the injection of current harmonics.
- Breaker-reclosers are one type of protection element that reacts to load imbalances and electrical failures.

Because it offers a comprehensive and adaptable depiction of an actual distribution system, this model is perfect for researching both operational. See Figures 3. 4. &5.

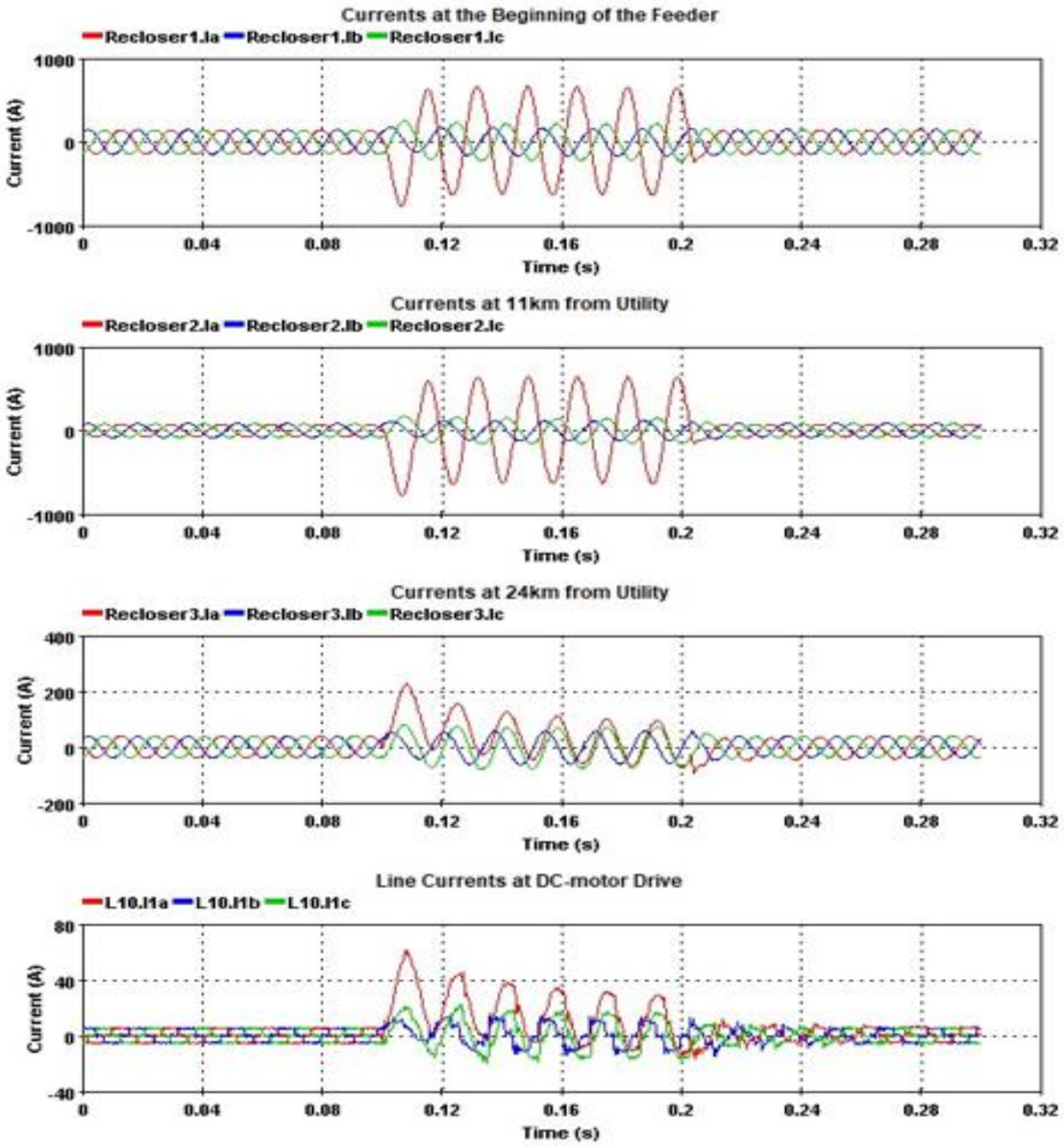

Figure 3:24-Bus 25kV Distribution Feeder and Industrial Power System (Currents at the Beginning of the Feeder)

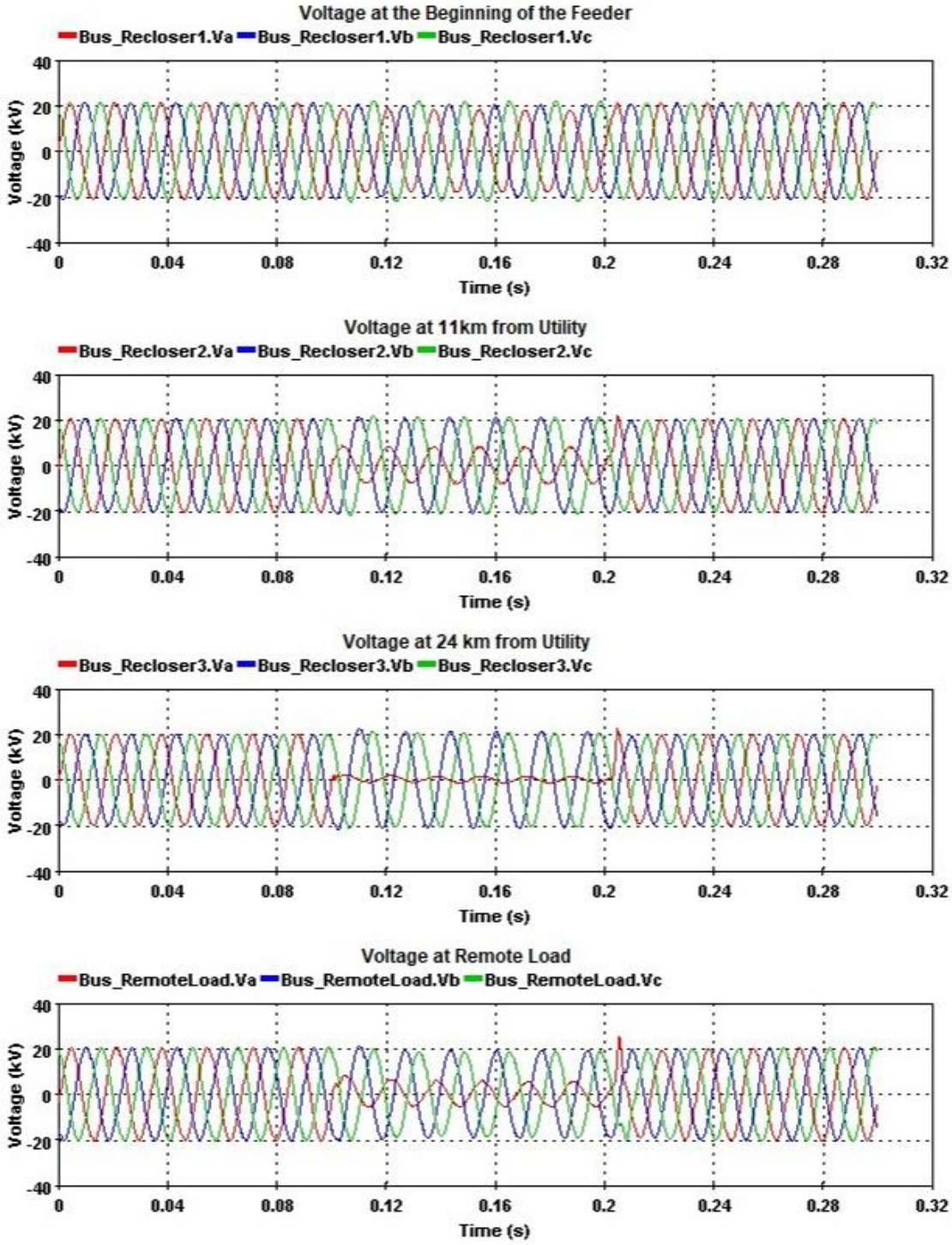

Figure 4:24-Bus 25kV Distribution Feeder and Industrial Power System (Voltage at the Beginning of the Feeder)

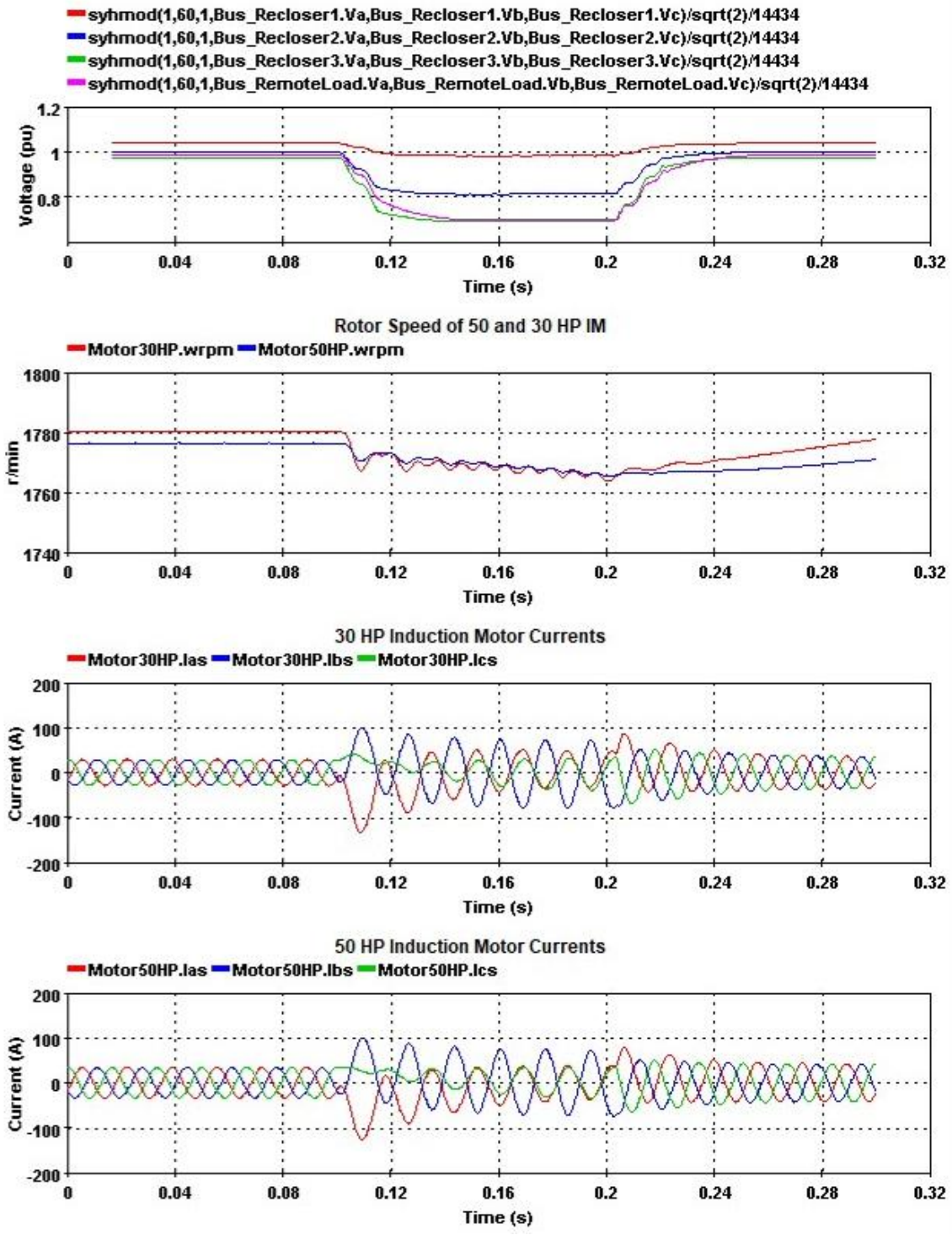

Figure 5:24-Bus 25kV Distribution Feeder and Industrial Power System (Line Current at DC-motor Drive)

3.8.2 Harmonic Analysis and Power Quality Assessment

One of the most significant applications of the Feeder_25kV_24Bus.ecf model is in the study of harmonic distortions and their impact on power system stability. The simulation includes:

- **Injection of specific harmonic frequencies** into the network, including **5th, 7th, 9th, 11th, 13th, 15th, and 19th harmonics**, to analyze their impact on power transformers and relay operations. In our research we used the **5th, 7th, 11th,** and **13th** harmonics to implement injection malware attack scenarios.
- **Current harmonics from industrial loads**, which mimic the real effects of nonlinear loads such as variable speed drives and DC motors.
- **Harmonic resonance and filtering**, allowing researchers to study mitigation techniques for reducing harmonic distortions in industrial power systems.

This capability makes the model highly suitable for cybersecurity research, where malicious actors can manipulate harmonic signatures to interfere with power system operations and protection mechanisms.

3.8.3 Fault Injection and Cybersecurity Testing

To assess system responsiveness and resilience, the Feeder_25kV_24Bus.ecf model allows the modeling of many failure scenarios and cyberattacks. These consist of: evaluation of relay coordination and system recovery timeframes for single-phase and three-phase faults.

Operation of breakers and reclosers in fault situations, guaranteeing proper relay reactions to electrical disruptions.

Simulations of cyber-physical attacks in which substation control and transformer protection systems are disrupted by malware-induced harmonic distortions.

Studies examining the effects of voltage flicker on power systems show how harmonic-based cyberattacks can result in voltage instability.

Researchers may create and evaluate machine learning-based intrusion detection systems that track system behavior in real-time and identify anomalies before they become serious grid disruptions by including cybersecurity risks into power system simulations.

3.8.4 Load Behavior and Grid Stability Studies

The model is especially helpful for researching load-dependent voltage fluctuations and how they affect the stability of the grid as a whole. Dynamic industrial loads are one of the features of the voltage-dependent load modeling feature that replicate actual industrial energy consumption patterns.

Effects of load imbalance: varying load profiles lead to harmonic distortions and voltage variations.

Analyzing the impact of capacitor switching and reactive power adjustment on system stability.

This makes it possible for researchers to experiment with different control schemes to enhance power grid efficiency and lessen cybersecurity risks that target load dynamics and voltage regulation systems.

3.8.5 Why This Model is Important for Cybersecurity Research

The Feeder_25kV_24Bus.ecf model is a vital tool in cybersecurity-focused power system research, enabling researchers to analyze how malicious code manipulates system harmonics to bypass traditional security measures. It helps assess vulnerabilities in protection systems, ensuring that relay coordination techniques remain effective against cyber-physical threats. Additionally, the model facilitates the development of machine learning-based anomaly detection by training AI-driven security mechanisms with real-time system behavior. Furthermore, it strengthens substation cybersecurity by supporting the creation of proactive defenses to mitigate malware-induced power outages and enhance overall grid resilience due to its ability to:

- Simulate harmonic-based malware attacks, providing insights into how malicious code manipulates system harmonics to bypass traditional security measures.
- Evaluate protection system vulnerabilities, ensuring relay coordination strategies remain effective under cyber-physical threats.
- Test machine learning-based anomaly detection models, using real-time system behavior to train AI-driven security mechanisms.
- Enhance substation cybersecurity resilience, by developing proactive defense strategies against malware-induced power disruptions.

3.9 Observing and capturing system activities

To monitor and analyze the impact of harmonic malware, we used the Windows Sysinternals Process Monitor (ProcMon) tool to capture low-level system activities on the industrial computer compromised by the malware injection. ProcMon is a powerful

utility that logs granular system events, such as file I/O operations, registry modifications, and process activities, providing invaluable insights into the system's behavior under attack.

To maintain a controlled and secure analysis environment, we executed the harmonic malware directly within our simulation program, developed in Python, and collected outputs and event traces using ProcMon. The tool’s detailed logging generates thousands of entries per minute, recording a vast amount of system activity. These logs were exported in Comma Separated Value (CSV) format and served as the primary dataset for subsequent analysis using machine learning models.

To focus specifically on the system activities triggered by the harmonic malware and the transformer simulation, we configured ProcMon to monitor only operations initiated by the Python application. This targeted filtering approach minimized noise from unrelated processes and ensured that our analysis was confined to relevant data. The filtering process was necessary because idle systems produce significant background activity, which could otherwise obscure malicious behaviors or add irrelevant entries to the dataset.

The captured events were further refined using criteria based on functions commonly associated with malware behavior, such as anomalous registry edits, unauthorized file access, or excessive I/O operations. By isolating these potentially malicious activities, we were able to create a clean and focused dataset for evaluating the interaction between the malware and the underlying system.

This methodology not only provided a precise view of the system’s response to harmonic malware but also ensured that the resulting data was robust and relevant for

training and testing machine learning models. The insights gained from this analysis will contribute to enhancing detection mechanisms and strengthening the resilience of industrial systems against cyber-physical threats.

3.10 Machine Learning Models

To analyze the behavioral characteristics of harmonic malware and the system activities of the industrial computer, we employed two machine learning models, as illustrated in chapter five.

1. **Random Forest Classifier:**

   This model was chosen for its robustness in handling diverse data types and its ability to provide insights into feature importance. Its ensemble-based approach reduces the risk of overfitting while maintaining high accuracy, making it suitable for analyzing complex system activity logs.

2. **Gradient Boosting Classifier:**

   This model was used to compare against the Random Forest classifier, particularly for its capability to model complex, non-linear relationships. Gradient Boosting is effective at capturing subtle patterns in the data, providing an additional layer of analysis to understand malware behavior and system responses.

3.11 Data Preparation for Machine Learning Analysis

The dataset used for training and evaluation was generated from ProcMon logs, capturing system activities during the execution of harmonic malware. The preprocessing steps included:

1. **Dataset Consolidation:**

   All datasets generated from individual harmonic injection scenarios were combined into a single, unified dataset. This comprehensive dataset allowed for consistent analysis across various scenarios, ensuring the machine learning models had sufficient data to generalize their predictions.

2. **Text Feature Extraction:**

   The 'Detail' field in the ProcMon logs, which contains text descriptions of system events, was transformed into numeric features using the **CountVectorizer**. This method converts text data into a numerical matrix, representing the frequency of specific terms related to system events, such as harmonic activities. This transformation enabled the models to capture patterns and relationships within system activities that are indicative of malware behavior.

3.12 Model Evaluation

Both machine learning models were trained and evaluated using this preprocessed dataset. By comparing the performance of the Random Forest and Gradient Boosting classifiers, we assessed their effectiveness in identifying malicious activities and their underlying system interactions. The combination of ensemble learning and non-linear modeling provided a comprehensive understanding of how harmonic malware impacts industrial systems, paving the way for improved detection and mitigation strategies.

3.13 Summary

This chapter outlines the research methodology used to investigate the cyber-physical interactions within power grid relays, with a focus on harmonic disturbances

caused by malware. The study utilizes simulations, data acquisition, and machine learning techniques to analyze and detect anomalies in power transformer operations.

3.13.1 Power Transformer Simulation

The study models the operation of power transformers, focusing on their transformation of voltage and current and their increasing vulnerability due to the integration of cyber-physical systems (CPS).

3.13.2 Harmonic Protection Algorithm

The chapter details how harmonic restraint algorithms use Fast Fourier Transform (FFT) to differentiate between magnetizing inrush currents and actual faults, emphasizing their role in ensuring transformer reliability.

3.13.3 Cyberattack Emulation

Harmonic malware is developed to simulate attack scenarios, including intermittent and cumulative harmonic injections, to evaluate the impact on harmonic protection mechanisms. The malware introduces specific harmonic orders (e.g., third, fifth, and seventh) and manipulates input signals to disrupt transformer operations.

3.13.4 System Activity Monitoring

Using the Windows Sysinternals Process Monitor (ProcMon), system activities are captured during malware execution. The data is filtered to isolate relevant events, creating a clean dataset for analysis.

3.13.5 Machine Learning Models

Two models, Random Forest and Gradient Boosting Classifiers, are employed to detect anomalies. Preprocessing steps include dataset consolidation and feature extraction using CountVectorizer to convert system log details into numerical features.

CHAPTER FOUR

DYNAMIC SYSTEM ACTIVITIES CAPTURE

This chapter describes the methodology and tools used to capture dynamic system activities, focusing on the use of the Windows Sysinternals Process Monitor (ProcMon). The chapter explains how ProcMon was employed in the research to monitor, log, and analyze system behaviors during harmonic malware execution in a controlled simulation environment. The detailed logs generated by ProcMon served as the foundation for anomaly detection and machine learning model training see Figure 6.

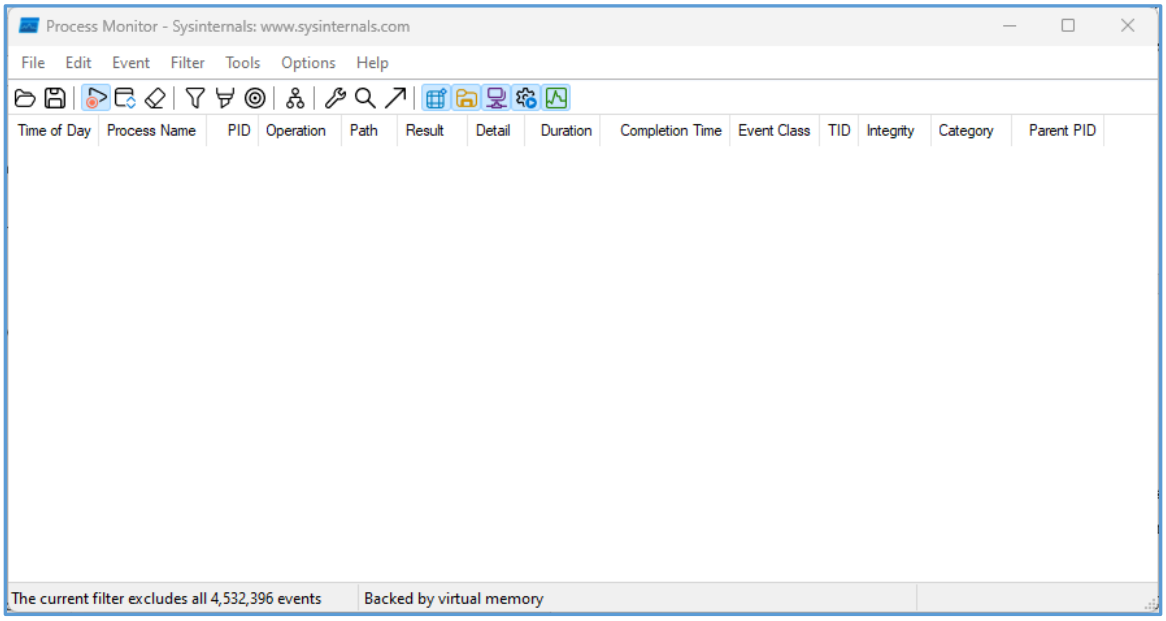

Figure 6: Process Monitor- Sysinternals Suite tools. Before starting capturing events

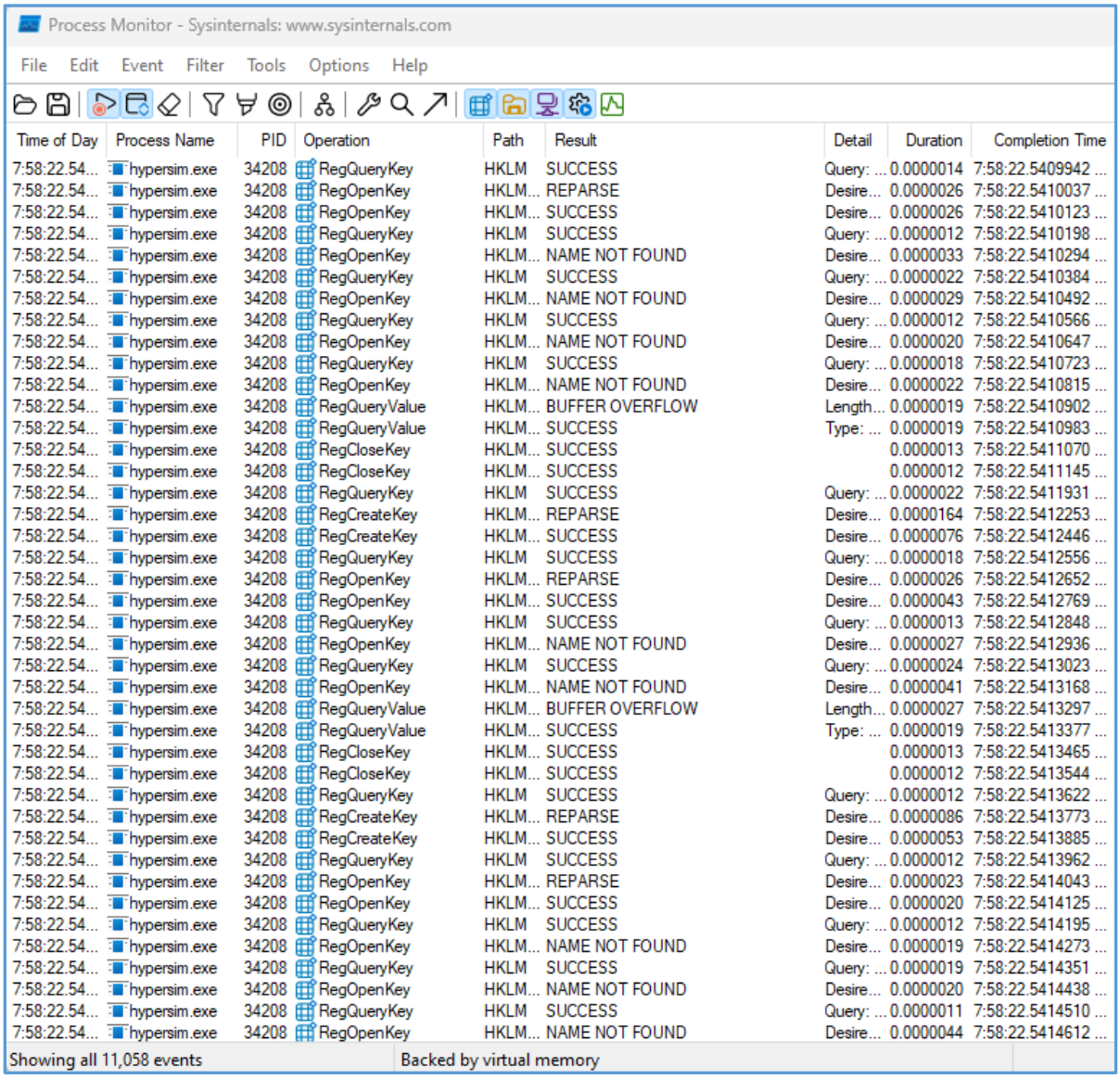

Figure 7: Process Monitor- Sysinternals Suite tools. After starting capturing events for HYPERSIM Simulation

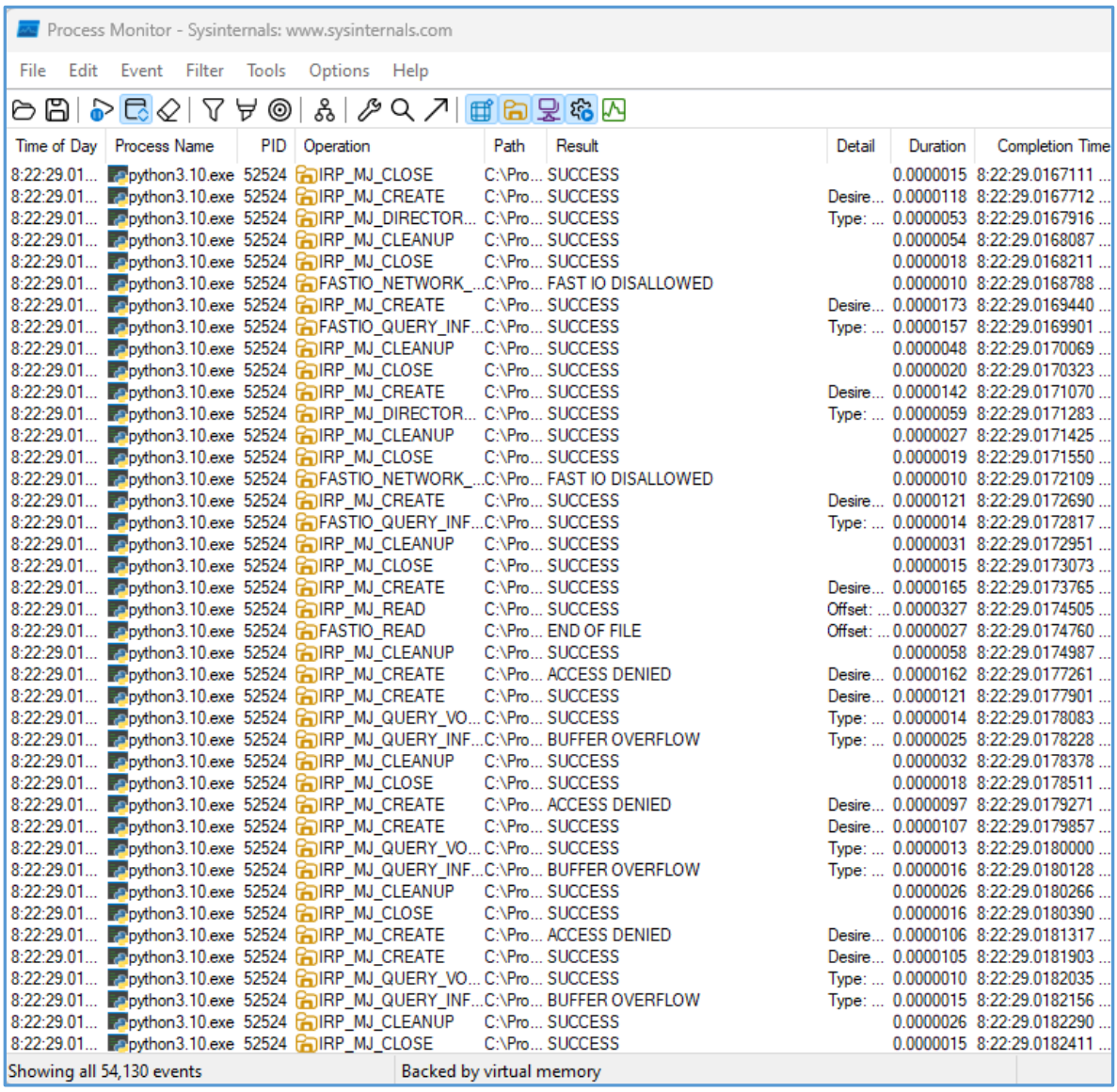

Figure 8: Process Monitor- Sysinternals Suite tools. After starting capturing events for Python Simulation

4.1 Overview of Sysinternals Process Monitor (ProcMon)

4.1.1 What is ProcMon

ProcMon is a powerful, lightweight monitoring tool developed by Microsoft as part of its Sysinternals Suite. It is designed to track real-time system activities at a granular level, capturing low-level events. ProcMon captures and logs a wide range of system

activities at a granular level, offering insights into how programs interact with the operating system and its resources. This includes See Figure 20.

- **File System Operations:** Tracks file creation, deletion, reading, and writing, providing visibility into how programs interact with files, directories, or storage media.
- **Registry Operations:** Monitors changes made to the Windows registry, such as key creation, value modification, or deletion, which are common behaviors of malware attempting to alter system configurations or establish persistence.
- **Process and Thread Activities:** Records events such as process creation, termination, and thread execution, revealing how malware spawns or interacts with other processes.
- **Network Connections:** Tracks any network requests made by monitored processes, providing insights into possible exfiltration or command-and-control (C2) communications.

4.1.2 Contextual Metadata for Every Event

Each captured event in ProcMon is accompanied by rich metadata that contextualizes the operation.

Timestamp: Records the exact time of each event, enabling precise temporal analysis of program behavior.

Process Name and ID: Identifies the process responsible for each activity, linking events to specific applications or malware components.

Operation Type: Specifies the action taken (e.g., RegQueryKey, CreateFile, or WriteFile) to understand the program's intent.

Result Codes: Indicates whether an operation succeeded or failed (e.g., access denied or operation completed successfully), which is critical for debugging and anomaly detection.

Resource Path: Details the specific file, registry key, or resource accessed, helping to trace the target of each operation.

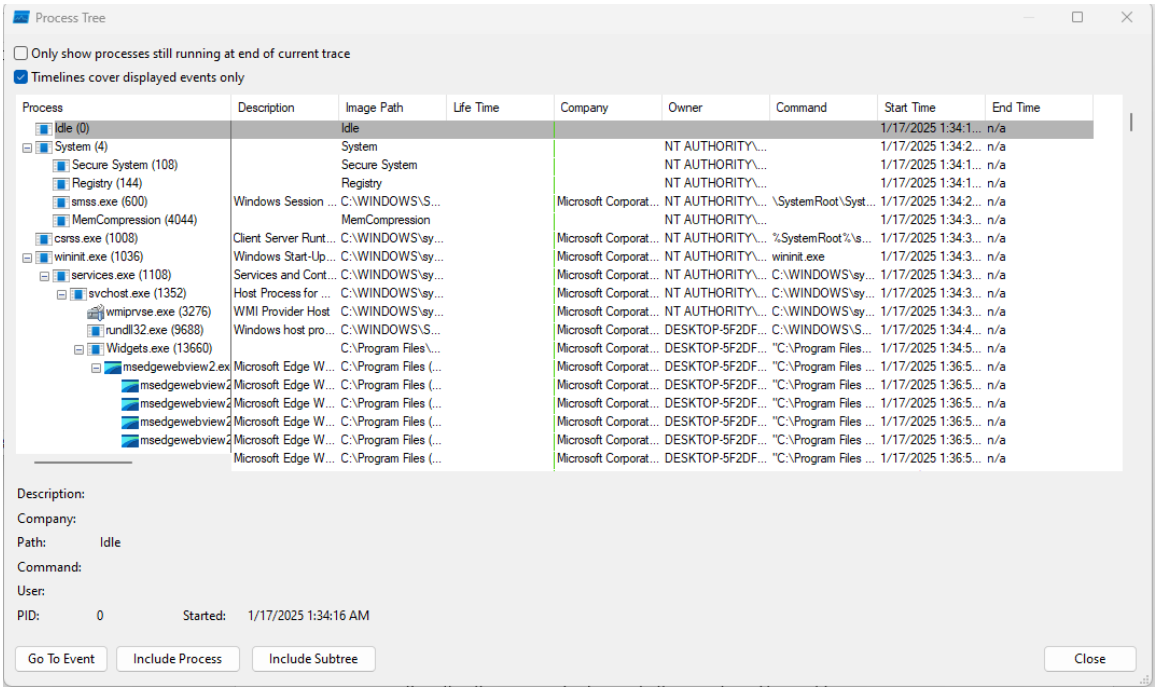

Figure 9: Process Monitor- Process Tree.

4.1.3 Real-Time Monitoring and Filtering

ProcMon operates in real-time, allowing researchers to observe program behaviors as they occur. Its filtering capabilities make it especially useful in high-activity environments, where thousands of events can be logged per second. Researchers can:

- Focus on specific processes (e.g., the malware program or simulation tool).
- Filter by event types (e.g., only file writes or registry changes).

- Narrow down to specific resource paths (e.g., files or registry keys critical to the simulation).

This ensures that only relevant data is logged, reducing noise and improving analysis accuracy.

4.1.4 Stack Traces and Dependency Analysis

ProcMon provides stack traces for each logged event, showing the sequence of function calls within the program leading up to the operation. This enables:

- Tracing Malware Logic: Identifying how malware executes operations, interacts with libraries, or calls system APIs to perform actions.
- Analyzing Inter-Process Dependencies: Understanding how malware interacts with other processes or threads, which is crucial for detecting advanced threats using stealthy inter-process communications.

4.1.5 Behavior Profiling in Malware Analysis

For industrial environments, such as power grids simulated in this research, ProcMon helps analyze malware behavior by profiling its interactions with critical system resources. For example:

- Malware Persistence: Logs registry keys modified or created to establish startup persistence.
- File Corruption or Exfiltration: Tracks unauthorized access to configuration files or log files.
- Unauthorized Communication: Detects malware attempting to open network sockets or modify DNS entries.

- Disruption Attempts: Monitors processes tampering with critical simulation resources, such as altering harmonic signals or creating false fault conditions.

4.1.6 Use in Complex Environments Like Power Grids

In industrial systems, such as power grid environments simulated in this research, ProcMon's ability to monitor multiple dimensions of system activity simultaneously makes it indispensable:

- Identifying Malicious Harmonic Injections: Capturing how malware manipulates analog-to-digital converter (ADC) processes or input signals.
- Detecting Simulation Interference: Monitoring unauthorized changes to simulation logs or intermediate data.
- Analyzing Malware Impact: Logging specific activities such as injecting harmonics, modifying relay logic, or disrupting protective algorithms.
- Correlation with CPS Components: Linking low-level system activity to broader CPS anomalies (e.g., relay misoperation caused by malware activity).

4.2 Why ProcMon is Ideal for Malware Analysis in Industrial Environments

4.2.1 Granularity of Analysis

ProcMon provides unparalleled granularity by logging every interaction between malware and system components. This level of detail is crucial for identifying subtle or stealthy attack strategies that might otherwise go unnoticed. The tool's comprehensive event tracking ensures that no significant operation is overlooked, enabling a deeper understanding of how malware manipulates system resources and interacts with its environment.

4.2.2 Versatility in Complex Scenarios

Industrial environments, such as power grids, often involve a mix of IT components (e.g., computers running simulations) and OT components (e.g., physical devices like relays). ProcMon effectively bridges the gap between these domains by monitoring malware's digital footprint across both areas. This versatility makes it an essential tool for analyzing threats in hybrid systems where cyberattacks can have physical consequences.

4.2.3 Ease of Data Export and Integration

Logs generated by ProcMon can be exported in CSV format, making them easy to integrate into machine learning workflows. Each event captured by ProcMon includes rich metadata, such as process IDs, timestamps, and operation types, which can be transformed into features for robust anomaly detection and behavior modeling. This seamless export and integration capability enhances the efficiency of data-driven cybersecurity research.

4.2.4 Real-Time Insights for Dynamic Systems

Industrial environments are highly dynamic, with states that change constantly based on operational demands. ProcMon’s real-time monitoring capabilities align perfectly with this characteristic, enabling researchers to observe malware behavior as it unfolds and evolves in response to these changes. By capturing live data, ProcMon provides critical insights into how malware interacts with dynamic systems over time.

4.2.5 Support for Simulation-Based Research

ProcMon's ability to focus on specific processes or scenarios makes it particularly valuable for simulation-based research. In controlled environments, such as those used in this study to analyze harmonic-based attacks on power grids, ProcMon captures malware activity with precision. This targeted monitoring ensures that researchers can isolate and study the effects of malware without interference from irrelevant system activities.

4.3 Harmonic Malware Scenarios

We implemented multiple harmonic scenarios in our research, with each scenario aimed at manipulating the power transformer's software system by introducing harmonic malware. To facilitate this, we developed the malware programs in Python, as outlined in the subsequent sections. These malware programs were executed on an industrial computer running the transformer simulation environment.

To analyze the impact of these injections, we utilized the Windows Sysinternals Process Monitor (ProcMon) tool to capture and monitor real-time system activities. ProcMon enabled detailed tracking of the malware's interactions with the system, providing invaluable insights into how the harmonic injections influenced the transformer's operations and exposed vulnerabilities in its protection algorithms.

4.3.1 Harmonic Injection

Harmonic injection involves deliberately adding harmonics of different orders to the power transformer to simulate conditions that could arise from malware-induced disturbances as shown in Figure 15. In this research, we focused on injecting harmonics such as the 2nd, 3rd, 4th, and 5th harmonic orders, Harmonic suppression, intermittent,

and cumulative Figure 17. These harmonic frequencies developed in Python were chosen to represent both low and higher-order disturbances, allowing us to observe the unique impacts they have on transformer operation.

- **2nd Harmonic Injection:** This low-order harmonic can cause increased heating in transformers and affect their efficiency. During the simulation, we observed changes in system activity such as higher CPU utilization and more frequent registry modifications, indicating potential stress on the substation computer.
- **3rd, 5th, 7th and 13th Harmonic Injection:** Higher-order harmonics have a more complex impact on power transformer stability. The simulation demonstrated these harmonics could lead to anomalies in the process activities of the substation computer, such as the initiation of unexpected threads and file access operations, which may be indicative of cyberattack behaviors. See Figures 10 and 11 for HYPERSIM simulation and Figures 16, 17 and 18.for Python simulation.

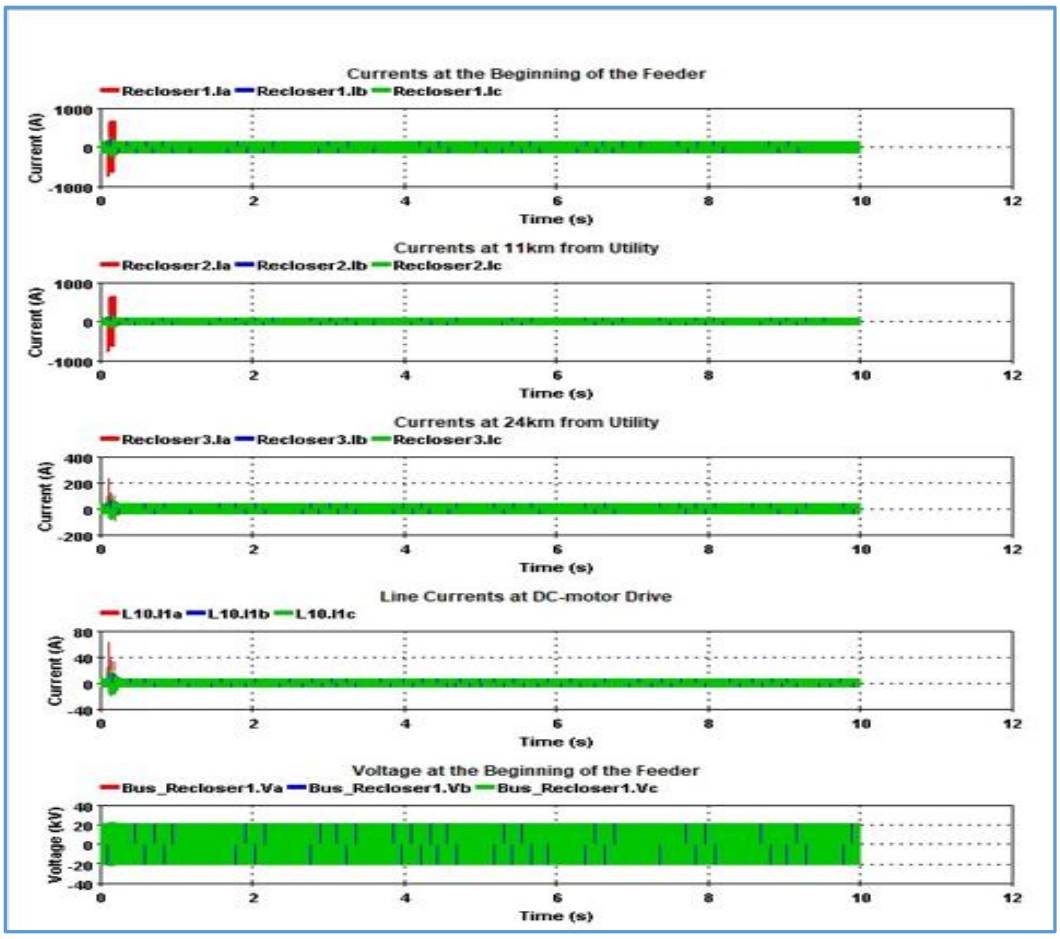

Figure 10: Screenshot of the current at the beginning of the Feeder 24 Bus During Implemented Harmonic Injection

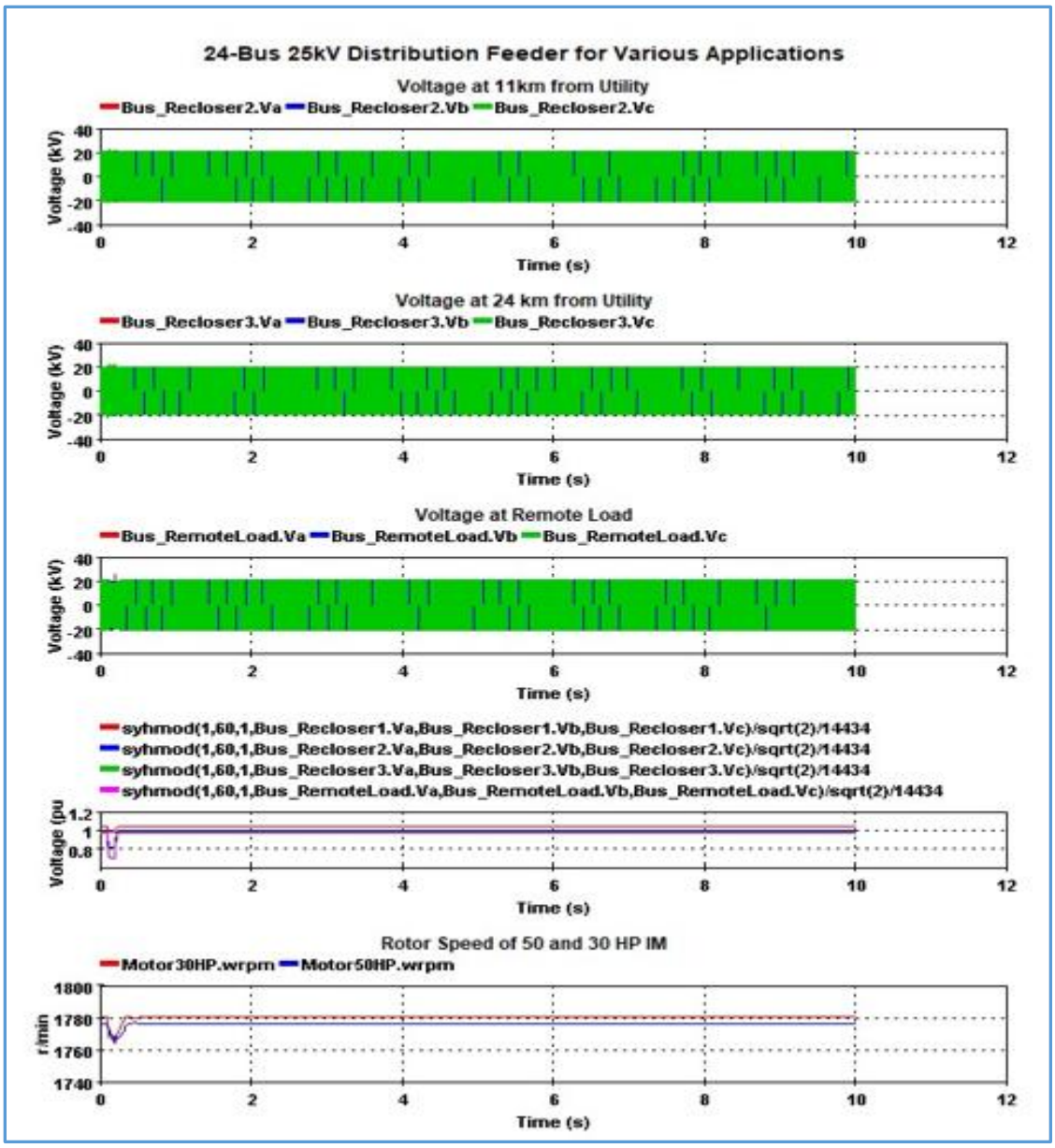

Figure 11: 24-Bus 25 KV Distribution Feeder for Various Application During Harmonic Malware Injection

4.3.2 Harmonic Suppression

Harmonic suppression aims to reduce or eliminate specific harmonics to understand how transformers respond to mitigation efforts. By suppressing certain harmonic orders, we analyzed how the system adapts under conditions of reduced harmonic stress. This scenario helped in understanding the system's resilience and the adjustments that occur at both the transformer and the system level [9].

- **Suppression of Low-Order Harmonics:** Suppressing lower-order harmonics, such as the 2nd and 3rd, provided insights into how the transformer stabilizes under lower harmonic distortion. The response was observed through reduced process activities, suggesting less computational stress and fewer registry modifications.
- **Suppression of High-Order Harmonics:** High-order harmonic suppression showed that the system's behavior became more predictable, with fewer anomalies detected. This is particularly important for enhancing the stability and security of the power transformer.

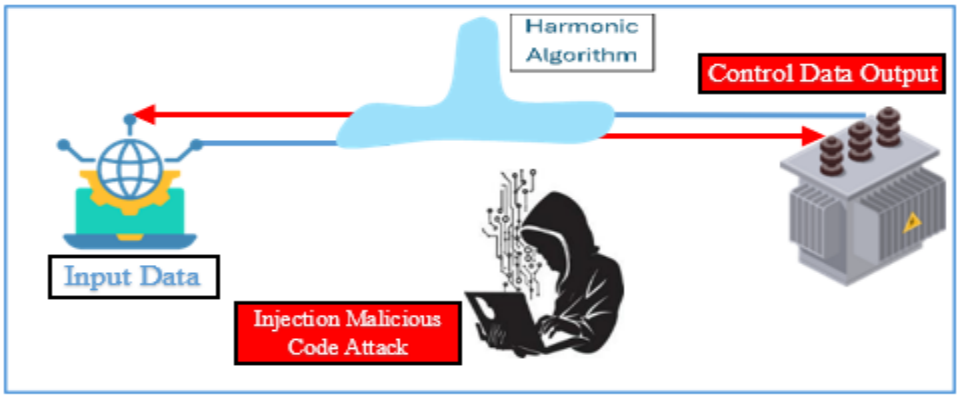

Figure 12: Implement cyberattack (Harmonic Malware)

4.3.3 Intermittent and Cumulative Harmonics

In the real world, harmonic disturbances can be intermittent, occurring sporadically, or cumulative, where multiple harmonics occur simultaneously. To emulate these scenarios, we introduced intermittent harmonic injections at random intervals and also cumulatively injected multiple harmonics.

- Intermittent Harmonic Injection: This scenario simulated the random occurrence of harmonics, which could be indicative of a transient cyberattack or unstable grid conditions. The system's response was characterized by spikes in CPU utilization and unexpected file access, which helped us understand the transient impacts of cyberattacks. See Figure 12.

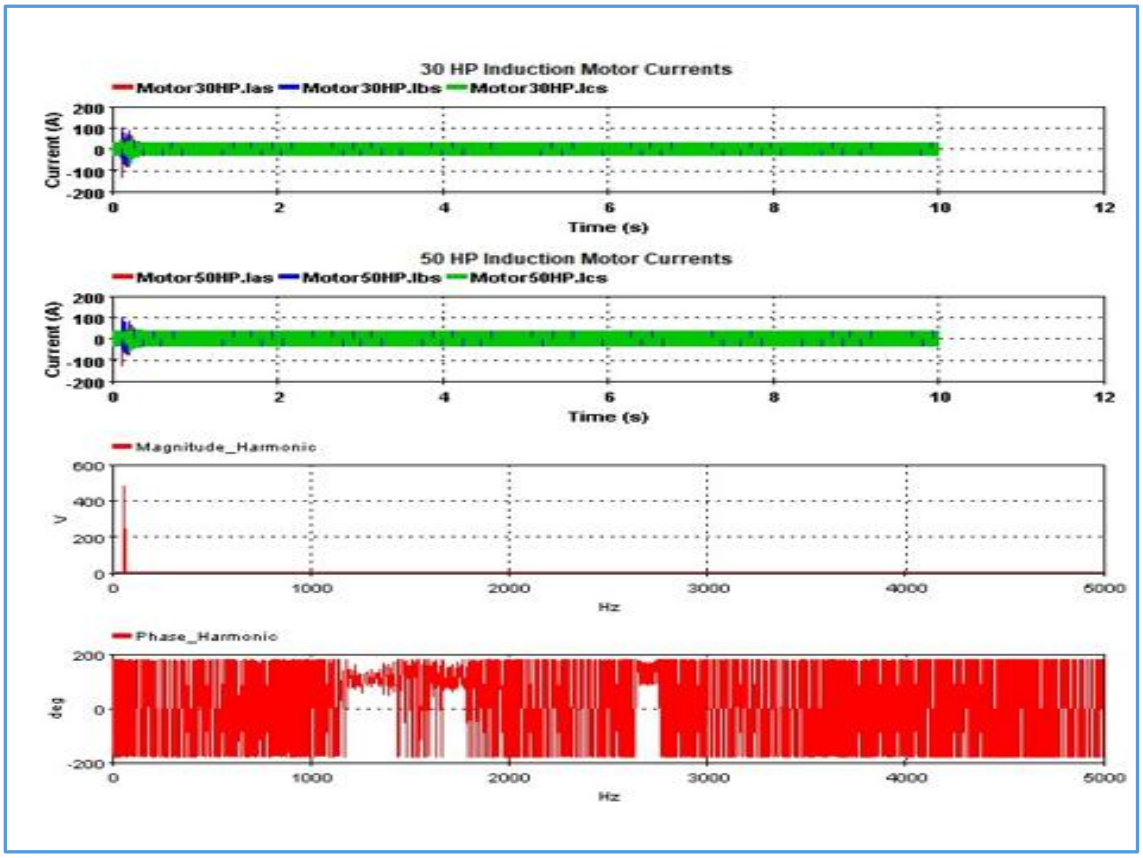

Figure 13: Example of Harmonic Malware Injection in 24-Bus 25 KV Distribution Feeder

- Cumulative Harmonic Injection: By injecting multiple harmonics at the same time, we were able to observe the compounded effects on the transformer and system activities. This cumulative impact was seen to create significant anomalies, such as simultaneous registry changes and process creations, suggesting that cumulative harmonic disturbances have the potential to create vulnerabilities exploitable by attackers.

Each scenario was run using a simulated power transformer environment, and the impact of these activities was monitored in the substation's computer.

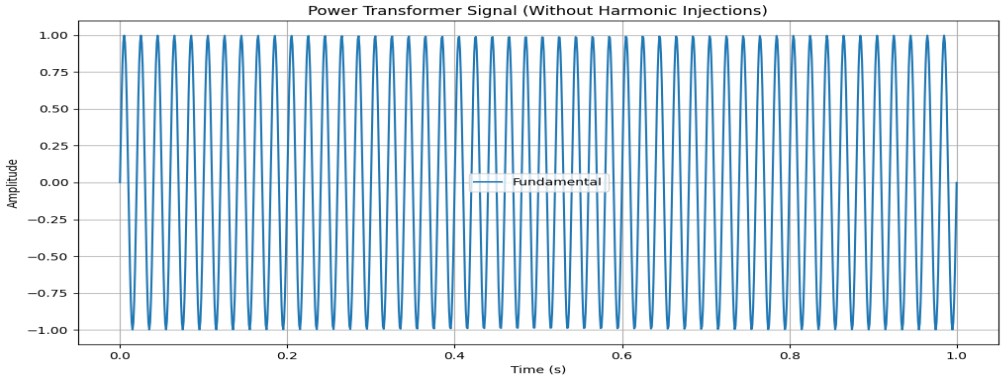

Figure 15: Harmonic Malware Scenarios for Python Simulation

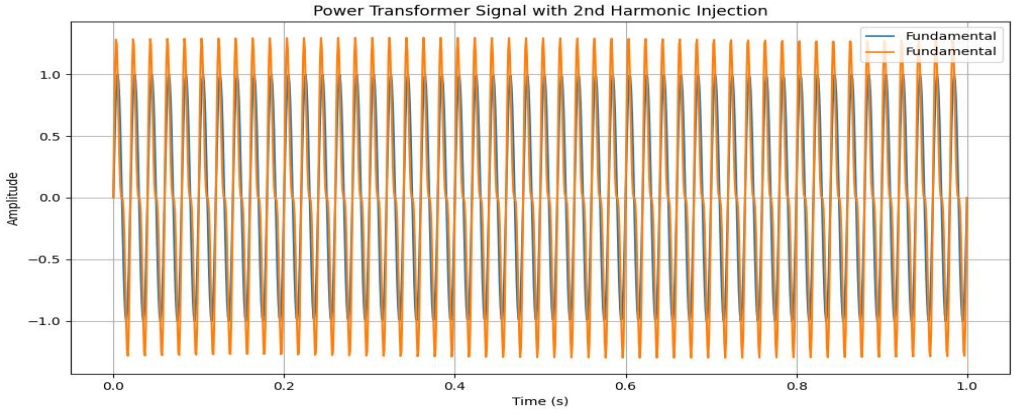

Figure 14: Harmonic Malware Scenarios for Python Simulation

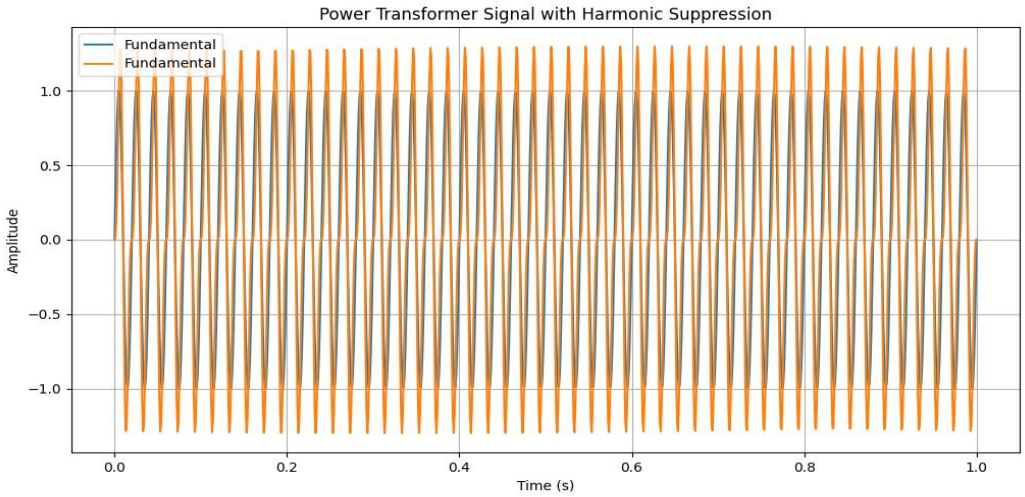

Figure 18: Harmonic Malware Scenarios for Python Simulation

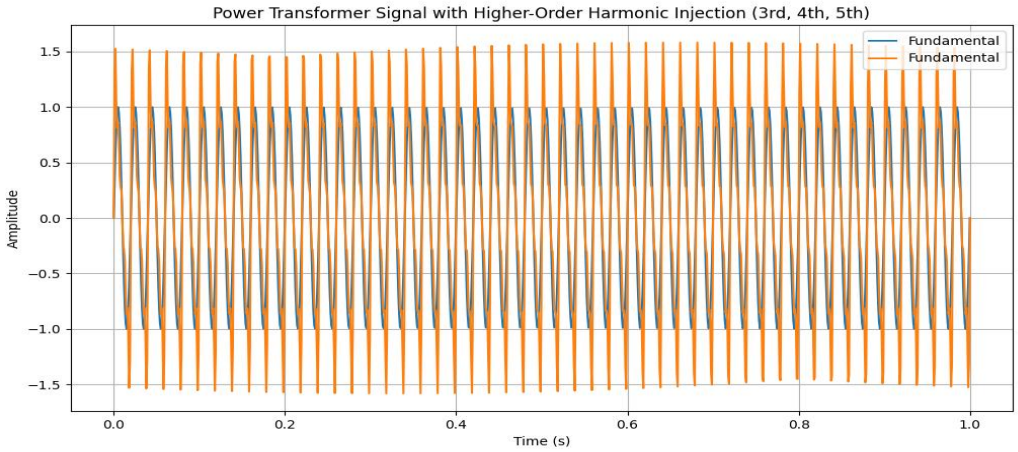

Figure 17: Harmonic Malware Scenarios for Python Simulation

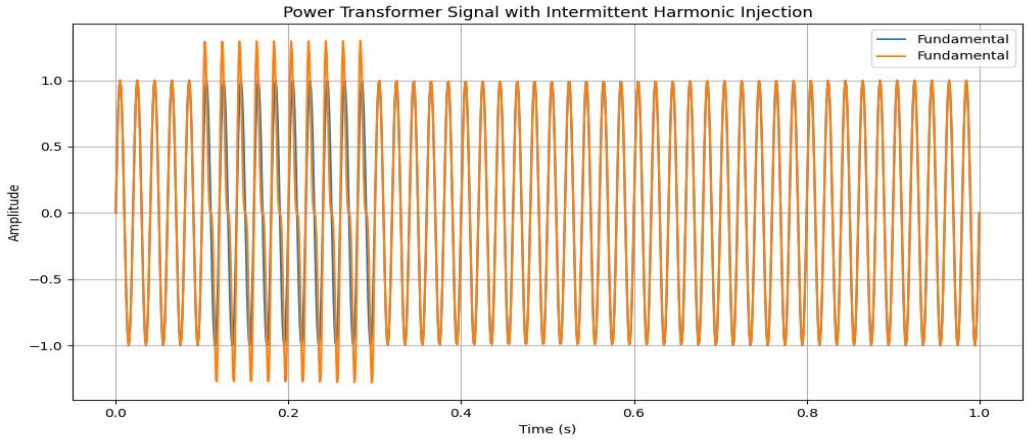

Figure 16: Harmonic Malware Scenarios for Python Simulation

4.4 Dynamic System Activities Capture

The goal of dynamic analysis is to look at how a malware sample behaves and acts while it is running. In order to do more research and analysis, this method watches how malware functions and collects comprehensive data on its runtime activities [10]. In this study, a cyberattack scenario like [11] is assumed; refer to Figure 12. Using credentials that have been taken from a valid operator and a remote connection to the substation communication network, an attacker can remotely access the substation process bus. By introducing the malware's forged SV packets into the process bus, the attacker then executes a false data injection attack, causing the differential protection relay to malfunction. We capture the malware's actions while it is running using Windows Sysinternals Process Monitor (ProcMon) [11]. In order to provide a more secure and regulated analysis environment, we ran harmonic malwares straight via our simulation program and gathered outputs and traces produced by ProcMon. ProcMon captures Windows system activities by logging low-level events, often generating thousands of entries per minute. The standard output from ProcMon, formatted in Comma Separated Value (CSV) format, is used as the primary input for our machine learning models.

4.5 ProcMon filter set

To ensure the capture of relevant system activities generated solely by our power transformer simulation and harmonic malware both implemented in Python. We configured ProcMon to focus exclusively on operations initiated by the Python application. This targeted filtering approach allowed us to precisely monitor the interaction between our simulation environment and the underlying system, minimizing

irrelevant data from unrelated processes. This is necessary because even an idle system runs numerous background processes that ProcMon can log. Therefore, we filtered the data to capture only potentially malicious activities, using criteria based on functions commonly associated with malware behavior [11], [12].

We set up filters to capture the following critical data attributes:

Table 1: The Critical Data Attributes Capture

File System Operations	Registry Operations	Process and Thread Operations	Image Loading and System Control
IRP_MJ_CREATE	RegOpenKey	Process Create	Load Image
IRP_MJ_READ	RegSetValue	Process Exit	IRP_MJ_SYSTEM_CONTROL
IRP_MJ_WRITE	RegQueryValue	Thread Create	
IRP_MJ_CLOSE	RegDeleteKey	Thread Exit	

These attributes were organized into columns: Time of Day, Process Name, Operation, Path, Result, and Detail, providing a structured view of how the Python-based harmonic simulations interact with the system.

4.6 ProcMon Filtering Strategy

Our filtering strategy was meticulously designed to capture all operations performed by the Python application, which runs the power transformer simulation and the harmonic malware. By setting ProcMon to monitor events only from processes associated with Python, we eliminated unnecessary noise from the system's background processes, focusing exclusively on our study’s area of interest. See Figure 19.

a) Time of Day: This column logs the exact time each operation occurred, helping us trace the sequence and timing of events during the simulation and malware execution.

b) Process Name: We filtered by the Python process name to ensure only events generated by our simulation and malware scripts were captured. This approach allowed us to isolate our application's behavior from the rest of the system.

c) Operation: We focused on capturing operations critical for understanding malware behavior and system impact, such as file and registry modifications, process creation and termination, and system control interactions.

d) Path: By recording the specific paths accessed by the Python application, we could identify crucial system resources targeted during harmonic injections or malware activities.

e) Result: We logged the outcomes of each operation, such as whether a registry change was successful or denied, providing insights into the effectiveness and potential limitations of the malware's actions.

f) Detail: Additional details about each event were captured to give a comprehensive understanding of the operations performed, including specific file attributes or registry values being manipulated.

This filtering strategy was essential for focusing on relevant data while minimizing extraneous information, thereby facilitating efficient and meaningful analysis.

4.7 Optimizing ProcMon Filter Settings

a) Targeted Monitoring: By specifying the Python application as the sole process of interest, we streamlined our data collection to capture only events pertinent to the power transformer simulation and harmonic malware. This approach reduced data noise and enhanced the precision of our analysis.

b) Efficient Data Capture: We fine-tuned the filters to prioritize critical operations that reveal how the harmonic malware interacts with system resources, such as file operations, registry modifications, and process/thread management. This ensured that we collected all essential activities without being overwhelmed by unrelated system events.

c) Iterative Refinement: We periodically reviewed and adjusted our filter settings to ensure we captured comprehensive yet focused data. This iterative process allowed us to adapt our monitoring as needed, ensuring the integrity and accuracy of our analysis. To optimize our filter configuration, we tailored ProcMon settings to ensure maximum relevance and efficiency for our simulation analysis.

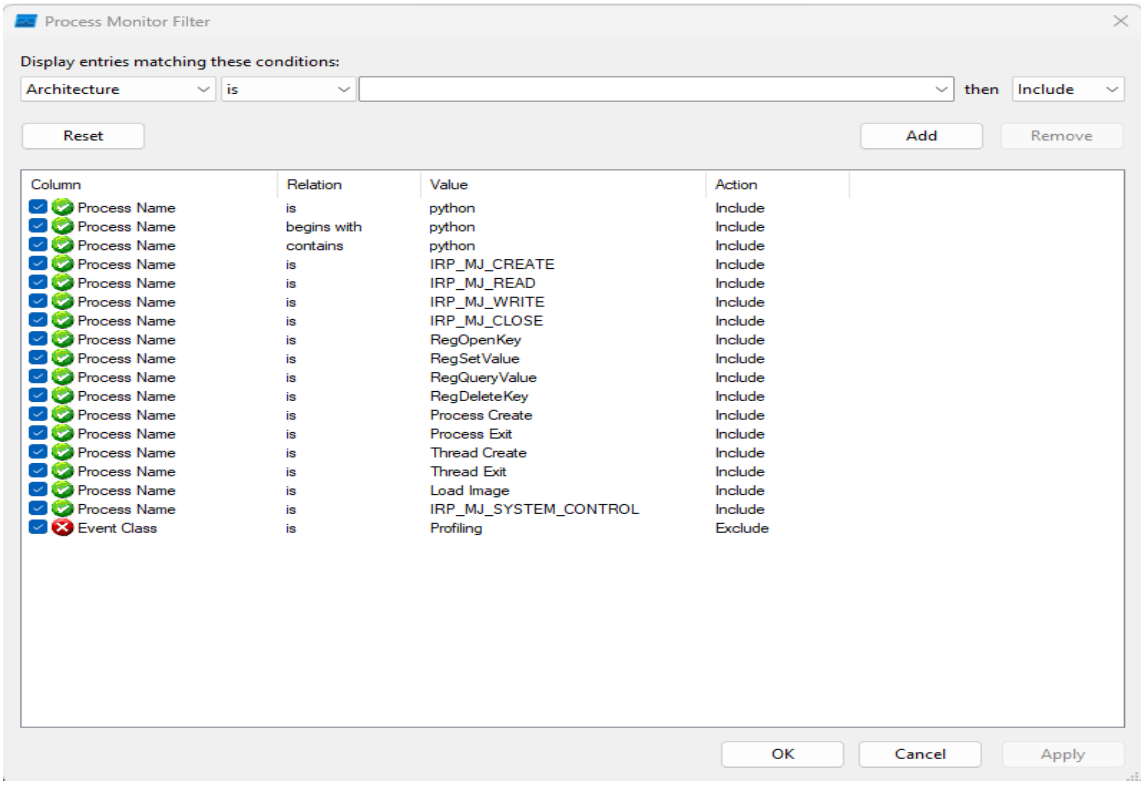

Figure 19: Screenshot ProcMon Filter for Python Simulation

CHAPTER FIVE

MACHINE LEARNING DEVELOPMENT

5.1 Overview of Machine Learning (ML)

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that enables computer systems to automatically detect patterns, classify behaviors, and make predictions without explicit programming. Instead of relying on manually defined rules, ML models learn from historical data, extracting insights and recognizing correlations that allow them to generalize their knowledge to new, unseen inputs. This ability to learn from experience and adjust predictions makes ML a powerful tool for automating decision-making processes across a variety of domains [13].

ML models operate using statistical and mathematical algorithms that process large datasets to identify trends, detect anomalies, and generate accurate predictions. The effectiveness of these models depends on the quality and quantity of data used for training, as well as the complexity of the underlying algorithm.

ML has found widespread applications across industries, including:

- **Healthcare**: Disease diagnosis through image recognition.
- **Finance**: Fraud detection in banking transactions.
- Cybersecurity: Intrusion detection and malware classification.
- **Industrial Control Systems (ICS)**: Anomaly detection in power grids and smart factories.

5.1.1 Machine Learning: Enhancing Predictions Data and Algorithms

Machine learning (ML) refers to the process of training a computer system to make accurate predictions by analyzing provided data. This is accomplished through various methodologies and neural network models that enable continuous improvement in the system's performance. By leveraging these techniques, machine learning systems evolve over time, becoming more efficient in processing information and generating reliable outcomes [13].

At its core, ML algorithms autonomously develop mathematical models based on sample data, commonly known as "training data." These models are then used to make predictions or decisions without explicit programming for each specific task. The applications of ML are vast, ranging from classifying objects in images—such as distinguishing a banana from an apple—to identifying pedestrians in front of an autonomous vehicle. Additionally, ML aids in interpreting ambiguous words in sentences, filtering spam emails, and transcribing speech for automated video captions.

As a subset of artificial intelligence (AI) and computer science, ML focuses on utilizing data and algorithms to replicate human learning processes, with the ultimate goal of improving accuracy over time [14]. It plays a pivotal role in the rapidly growing field of data science, where statistical techniques are applied to train algorithms for classification and prediction. These insights are essential for data mining initiatives, enabling organizations to uncover valuable patterns that drive decision-making and business strategies [15].

The ongoing expansion of big data has led to an increasing demand for data scientists who specialize in machine learning. These professionals play a crucial role in

identifying key business questions and determining the data required to address them. As industries continue to rely on data-driven insights, the significance of machine learning in shaping technological advancements and business outcomes will only continue to grow.

5.1.2 Understanding the Machine Learning Process

The machine learning process, as described by the University of California, Berkeley, consists of three fundamental stages:

1. **Decision Process:** Machine learning algorithms primarily function to generate predictions or classifications. To achieve this, the algorithm identifies patterns within a given dataset, utilizing input data that may or may not be labeled. This decision-making process forms the foundation for how models interpret and process information.
2. **Error Function:** To assess the accuracy of the model's predictions, an error function is employed. If pre-existing data points are available, the error function enables a comparison between the model's predictions and actual values. This helps in quantifying the level of accuracy and identifying potential areas for improvement.
3. **Model Optimization:** If the model's accuracy can be enhanced, an optimization process is initiated. This involves adjusting the model's weights to reduce the discrepancy between predicted outputs and actual data. The algorithm iterates through evaluation and optimization cycles, continuously refining the model until it achieves a predefined accuracy threshold.

By following these three stages, machine learning systems can iteratively improve, making them more effective in solving complex problems and enhancing decision-making processes.

5.2 Machine Learning Methods: Supervised, Unsupervised, and Semi-Supervised Approaches

Machine learning (ML) methods are categorized based on how they learn from data. The three primary types of ML methods are **Supervised Learning**, **Unsupervised Learning**, and **Semi-Supervised Learning**. Each method has distinct characteristics and applications depending on the nature of the data and the problem being addressed.

5.2.1 Supervised Learning

- **How It Works**: The model is provided with input-output pairs, and its goal is to learn the mapping function that best connects inputs to outputs. It continuously adjusts itself based on errors, refining its predictions over time [18] [19].
- **Examples:**
  - Spam email detection (classifying emails as spam or not spam).
  - Image classification (identifying objects in images).
  - Predicting house prices based on features like size and location.
- **Common Algorithms:**
  - Linear Regression
  - Decision Trees
  - Support Vector Machines (SVM)
  - Neural Networks.

- **Cybersecurity Applications:**
  - **Intrusion Detection Systems (IDS):** Supervised learning helps identify unauthorized access by classifying network traffic as either normal or suspicious.
  - **Malware Detection** – ML models analyze file attributes and behavior to classify software as benign or malicious.
  - **Phishing Email Detection** – Algorithms detect fraudulent emails by analyzing linguistic patterns, sender details, and embedded links.
  - **Limitations in Cybersecurity:**
    - Requires a large dataset of labeled attacks, which can be difficult to obtain.
    - May struggle against zero-day attacks (new, unknown threats).

5.2.2 Unsupervised Learning

Unsupervised learning deals with unlabeled data, meaning the model does not have predefined outputs to learn from. Instead, it identifies patterns, structures, and relationships within the data by grouping similar items or detecting anomalies [14].

- **How It Works**: The algorithm explores the data, looking for inherent structures such as clusters or associations. It does not have explicit correct answers but instead organizes the data based on similarities.
- **Examples**:
  - Customer segmentation (grouping similar customers for marketing).
  - Anomaly detection (identifying fraudulent transactions in banking).

- Market basket analysis (finding patterns in purchasing behavior).

- **Common Algorithms**:
  - K-Means Clustering
  - Hierarchical Clustering
  - Principal Component Analysis (PCA)
  - Autoencoders
- **Cybersecurity Applications**
  - **Anomaly-Based Intrusion Detection:** Unsupervised models can detect unusual patterns in network traffic, signaling a potential cyberattack.
  - **Fraud Detection:** Financial systems use ML to spot unusual spending behaviors, which may indicate fraudulent transactions.
  - **Insider Threat Detection:** Identifies deviations in employee behavior that might suggest unauthorized data access.
- **Limitations in Cybersecurity:**
  - High false positive rates, as unusual behavior does not always indicate a cyber-attack.
  - Lacks predefined labels, making it harder to explain why certain activities are flagged.

5.2.3 Semi-Supervised Learning in Cybersecurity

Semi-supervised learning is a hybrid approach that combines aspects of both supervised and unsupervised learning. It works with a small amount of labeled data and a large volume of unlabeled data, leveraging both to improve learning efficiency.

- **How It Works:** The model first learns from the limited labeled data and then extends its knowledge by discovering patterns in the unlabeled data. This approach is particularly useful when labeling data is expensive or time-consuming.
- **Examples:**
    - Medical diagnosis (where a small dataset of labeled disease cases is available, but vast amounts of unlabeled data exist).
    - Speech recognition (leveraging small amounts of transcribed audio alongside large amounts of untranscribed speech).
    - Web content classification (categorizing new web pages with minimal labeled examples).
- **Common Algorithms:**
    - Self-training,
    - Generative Adversarial Networks (GANs)
    - Graph-Based Models.
- **Cybersecurity Applications:**
    - **Threat Intelligence Analysis:** Helps security teams classify new threats using limited known attack data.
    - **Botnet Detection:** Identifies bot-controlled devices on a network by using labeled samples of bot traffic and analyzing unlabeled network flows.
    - **Web Content Filtering:** Helps classify malicious websites using a small dataset of known dangerous sites combined with large-scale internet traffic data.

- **Limitations in Cybersecurity:**
  - Requires careful balance between labeled and unlabeled data to avoid bias.
  - Can still be vulnerable to adversarial attacks if not properly trained.

5.3 Role of Machine Learning in Cyber-Physical Systems Security

In **cyber-physical systems (CPS)** such as power grids, cybersecurity threats can originate from malware that manipulates both digital components (software, firmware, network protocols) and physical infrastructure (power transformers, circuit breakers, relays). ML enhances CPS security by:

5.3.1 Detecting cyber anomalies

Cyber-Physical Systems (CPS), such as power grids, integrate digital and physical components, making them vulnerable to sophisticated cybersecurity threats. These threats often stem from malware that manipulates software, firmware, network protocols, and even physical infrastructure, including power transformers, circuit breakers, and relays. Traditional security measures may struggle to address these complex attack vectors, but machine learning (ML) offers a powerful solution by enhancing the detection, prediction, and response capabilities of CPS security.

5.3.2 Recognizing malware patterns

One of the key advantages of ML in CPS security is its ability to detect cyber anomalies in system activities. By continuously monitoring vast amounts of operational data, ML algorithms can identify deviations from normal behavior, signaling potential cyber intrusions. Additionally, ML models can recognize malware patterns by analyzing behavioral data, distinguishing between legitimate operations and malicious activities that

might otherwise go unnoticed. This capability is especially crucial in power grids, where attacks can have widespread consequences, disrupting critical services [16].

5.3.3 Predicting physical failures

Beyond cyber threat detection, ML also plays a role in predicting physical failures caused by cyberattacks. Malicious actors may manipulate system parameters to induce mechanical stress or overload critical components, leading to equipment failures. By leveraging historical data and predictive analytics, ML can forecast these failures before they occur, enabling preventive maintenance and reducing downtime. By integrating ML into CPS security frameworks, organizations can proactively defend against cyber threats, improving both the safety and efficiency of critical infrastructure. As cyberattacks grow more sophisticated, the role of ML in enhancing CPS security will continue to expand, offering adaptive and intelligent defenses against emerging risks [17].

5.3.4 Enabling real-time monitoring

A particularly advanced form of attack in power grids involves harmonic-based malware, which exploits harmonic distortions in electrical signals to bypass traditional security mechanisms. Since conventional detection methods may not recognize these subtle anomalies, ML-based approaches offer an intelligent solution by identifying irregularities in waveform patterns. Early detection of such distortions helps prevent potential damage to electrical components, ensuring system stability and reliability. Furthermore, real-time monitoring powered by ML allows for automated threat response, minimizing the need for human intervention and enhancing the resilience of CPS against evolving cyber threats [13].

5.4 Dataset Generation and Preprocessing

5.4.1 Data Collection Using ProcMon

As we mentioned in chapter 4, we are using ProcMon to capture the system activities to generate the data. To train ML models, system activity data was collected using Windows Sysinternals Process Monitor (ProcMon). ProcMon captures system events such as:

- File System Operations: Malware modifying security logs.
- Registry Changes: Unauthorized access to critical settings.
- Process Execution Patterns: Suspicious scripts running in the background.

During simulation, harmonic injection attacks were executed in a Python-based emulation environment, capturing real-time system activities before, during, and after cyberattacks.

We capture the malware's actions while it is running using Windows Sysinternals Process Monitor (ProcMon). In order to provide a more secure and regulated analysis environment, we ran harmonic malwares straight via our simulation program and gathered outputs and traces produced by ProcMon. ProcMon captures Windows system activities by logging low-level events, often generating thousands of entries per minute. The standard output from ProcMon, formatted in Comma Separated Value (CSV) format, is used as the primary input for our machine learning models

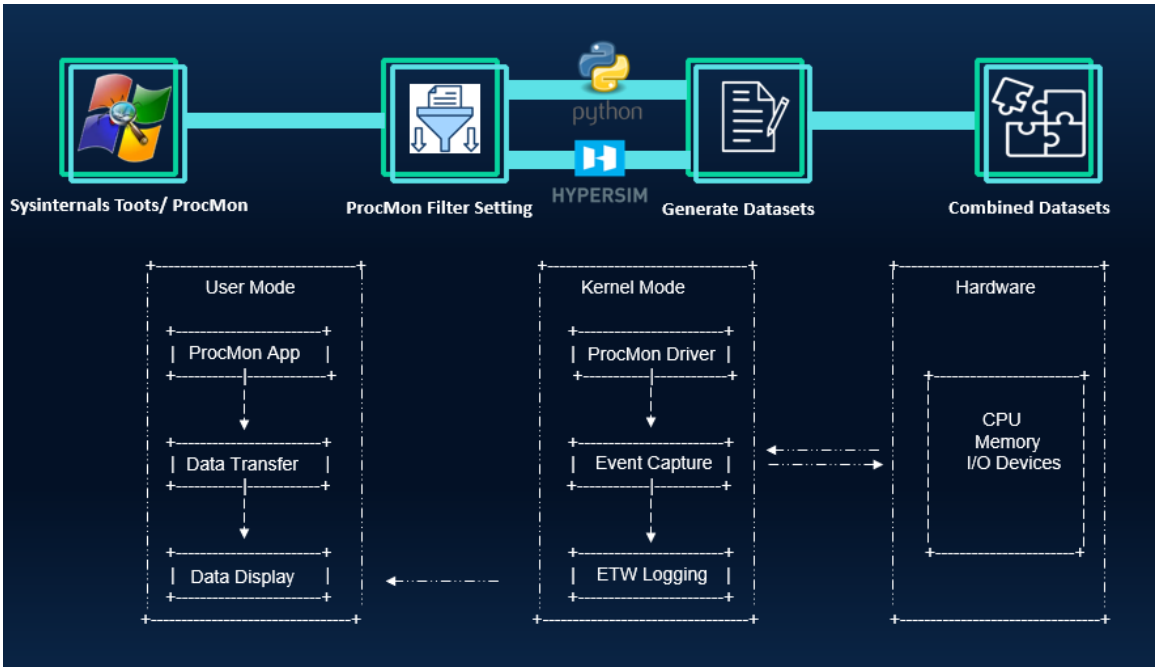

Figure 20: Dynamic system activities capture

In the preceding chapter, we detailed the process of configuring and refining ProcMon filters to yield precise and relevant results aligned with our defined criteria. The machine learning datasets are supplied in a CSV file, serving as the input for our models, as depicted in Figure 20.

5.4.2 Data Preprocessing & Feature Extraction

The dataset generated from ProcMon logs was preprocessed to prepare it for machine learning analysis This involved:

Step 1: Data Preprocessing

Data preprocessing is a crucial step in ensuring that the dataset is clean and properly structured for machine learning. The model reads system activity logs from **dataset.csv**, ensuring missing values do not disrupt the learning process. This is done by

replacing any missing data with "Unknown", a common practice in handling incomplete data.

To transform textual and categorical data into numerical values suitable for machine learning, the following techniques are applied:

- **TF-IDF Vectorization**: The Detail column, which contains textual information, is transformed using **TfidfVectorizer**. This technique assigns weights to words based on their importance within a system activity log, allowing the model to extract meaningful patterns.
- **Categorical Encoding**:
    - The columns 'Operation', 'Event Class', 'Integrity', 'Category', and 'Path' are converted into numerical format using **Label Encoding**. This ensures that categorical variables can be processed by the Random Forest algorithm.
    - Label Encoding assigns unique integer values to each category, making it suitable for algorithms like Random Forest that handle categorical data efficiently.
- **Feature Scaling (Standardization)**:
    - Since machine learning models perform better when numerical data is within a standardized range, **StandardScaler** is applied to transform numerical features. The test size was 30% and the random state was 42 while the **CountVectorizer** of the maximum features was 10
    - Standardization ensures that the features have **zero mean** and **unit variance**, which helps improve model stability and convergence.

- **Feature Reduction**:
    - Columns such as 'Time of Day', 'Process Name', 'Detail', and 'Completion Time' are removed from the dataset, as they either do not contribute significantly to classification or contain redundant information.
    - These features were likely removed because they either do not contribute significantly to classification or contain redundant information.

These preprocessing steps enhance the efficiency of the machine learning model, ensuring that it focuses on the most relevant features.

Step 2: Handling Class Imbalance

Cyberattack detection often involves identifying rare events in large datasets. Since real-world system logs are often highly imbalanced—with far more normal activities than malicious ones—your model employs **SMOTE (Synthetic Minority Over-sampling Technique)** to address this issue.

The key benefits of SMOTE include:

- **Preventing Bias**: By generating synthetic samples for underrepresented classes, SMOTE ensures that the model does not become biased toward majority-class system activities.
- **Improved Classification Accuracy**: Without SMOTE, the model might struggle to detect rare but crucial anomalies related to cyberattacks.
- **Before and After Comparison**: The model prints the distribution of class labels before and after SMOTE is applied, confirming that minority classes are sufficiently represented.

Handling imbalanced datasets is critical in cybersecurity applications, where rare events such as **buffer overflows, unauthorized access, or registry manipulations** can indicate serious threats.

5.5 Machine Learning Models Training for Harmonic-Based Malware Detection

Once the dataset is preprocessed and balanced, the next step is training the machine learning model. The dataset is split into training (80%) and testing (20%) sets to evaluate performance fairly.

- Feature Scaling:
    - After splitting the dataset, the numerical features are standardized using **StandardScaler**, ensuring that all variables have comparable scales.
    - This step helps improve the stability of the learning process, particularly for models sensitive to feature magnitude.

5.6 Model Choice

5.6.1 Random Forest Classifier (RF)

The Random Forest Classifier (RF) is an ensemble learning method that improves predictive accuracy by creating multiple decision trees and aggregating their results. This technique enhances robustness in cybersecurity applications, particularly in malware detection. One of the key advantages of RF is its ability to determine **feature importance**, identifying which attributes contribute the most to detecting malicious activities. Additionally, it performs well with **structured system logs**, especially those

containing **categorical and numerical features**, making it a suitable choice for analyzing large-scale security data.

Two Random Forest classifiers are trained separately:

- One for predicting **Result** (the outcome of a system operation).
- Another for predicting **Type** (the category of system activity).

- **Key Features of RF**
  - Constructs multiple decision trees to make predictions.
  - Identifies important features that contribute to malware detection.
  - Best suited for structured system logs with categorical and numerical data.
- **Performance Evaluation**
  - Accuracy Score: Measures overall classification performance.
  - Classification Report: Provides precision, recall, and F1-score for each category.
  - Confusion Matrices: Show true positives and false positives, allowing for deeper analysis of misclassifications.
  - Additionally, classification results are saved to Analysis.csv, providing a structured output of system behavior predictions

5.6.2 Gradient Boosting Classifier (GB)

The Gradient Boosting Classifier (GB) is a powerful machine learning technique that enhances classification accuracy through **adaptive learning**. Unlike RF, which builds trees independently, GB improves upon previous predictions iteratively, reducing errors over multiple iterations. This makes it particularly effective for detecting low-

magnitude, high-frequency harmonic anomalies, which are often used in sophisticated cyberattacks. However, despite its superior classification precision, GB is computationally expensive, requiring more processing power compared to RF.

- **Key Features of GB:**
    - Uses adaptive learning to improve accuracy over multiple iterations.
    - Excels at detecting low-magnitude, high-frequency harmonic anomalies.
    - Provides high classification precision but is computationally intensive.

5.7 Security Implications of Classification Results

The classification results highlight key system behaviors that may indicate cybersecurity threats. Below are the detected activity categories and their potential exploitation risks:

- **BUFFER OVERFLOW:** Indicates attempts to overwrite memory, a classic exploit used for code injection attacks. Attackers could use this to execute malicious code in the substation system.
- **NAME NOT FOUND:** Suggests unauthorized attempts to access non-existent files or registry keys. This could indicate reconnaissance activities, where an attacker is probing the system for vulnerabilities.
- **FAST IO DISALLOWED:** Shows failed fast I/O operations, potentially signifying attempts to bypass security checks by malware.
- **ACCESS DENIED:** Suggests unauthorized access attempts, which may be brute-force attacks or privilege escalation attempts.

- **REPARSE GLOBAL:** Indicates the use of reparse points, which attackers sometimes exploit to redirect legitimate processes to malicious resources.

Detecting these anomalies helps identify **malware activity, insider threats, and unauthorized access attempts**, making your machine learning model a valuable tool for enhancing substation security

5.8 Predicting New System Activities

The model also includes a function, **predict_result_and_type()**, that allows real-time classification of new system activity logs. By providing a sample input, the function predicts the most likely 'Result' and 'Type', assisting security analysts in identifying suspicious activities.

This feature is particularly useful for:

- **Real-time Security Monitoring**: Detecting anomalies as they occur.
- **Incident Investigation**: Providing insights into past activities that may have led to system compromises.
- **Automated Threat Detection**: Integrating machine learning-based analysis into security monitoring tools.

- **Key Strengths of our Model**
  - Comprehensive Preprocessing**: Uses TF-IDF, categorical encoding, and feature scaling.**
  - **Imbalance Handling:** SMOTE ensures that rare cyberattack events are detected

- **Effective Classification:** Random Forest models provide strong predictive accuracy.
- **Security Insights:** Identifies key attack patterns, such as buffer overflows and unauthorized access.

CHAPTER SIX

DISSCUSSION

6.1 Introduction

This chapter evaluates the performance of the proposed malware detection framework, focusing on the correlation between cyber and physical anomalies in power grid relays. The evaluation is based on two simulation environments:

1. A Python-based simulation developed in the lab, designed to inject harmonic-based malware and capture system activities.
2. The HYPERSIM software provided by OPAL-RT, which simulates real-world power transformer operations and evaluates the impact of cyber-induced harmonic distortions.

To assess the effectiveness of the proposed machine learning (ML) models, the results obtained from both simulation environments are analyzed using key performance metrics such as accuracy, precision, recall, and F1-score. The confusion matrices generated from both Python and HYPERSIM datasets provide insights into the classification performance, highlighting potential false positives and false negatives.

Furthermore, this chapter discusses the impact of harmonic injections on power system behavior and how the trained ML model successfully detects malware-induced disturbances. A comparison between detected anomalies in Python vs. HYPERSIM will reveal variations in detection accuracy and system response to cyber-physical threats.

By analyzing the strengths and limitations of the proposed framework, this chapter provides a comprehensive discussion on the feasibility of using harmonic analysis for malware detection in power systems. The findings of this chapter lay the groundwork for potential future improvements and real-world deployment of AI-driven cybersecurity solutions in industrial control systems.

6.2 Performance Metrics of ML Model

The performance of the proposed machine learning (ML) model was evaluated using accuracy, precision, recall, F1-score, and confusion matrices to assess its effectiveness in detecting cyber-physical anomalies. Two sets of evaluations were conducted, one using Python-based simulations and the other using HYPERSIM simulations, to compare model performance across different environments.

6.3 Model Effectiveness for Python simulation.

6.3.1 Confusion Matrix Accuracy Analysis and Evaluation for Harmonic Injection Results Classification

In the Python-based simulation, the model demonstrated exceptional accuracy, achieving nearly 91% classification performance for Result vs. Predicted Result and nearly 86% for Type vs. Predicted Type. See table 2 and table 3

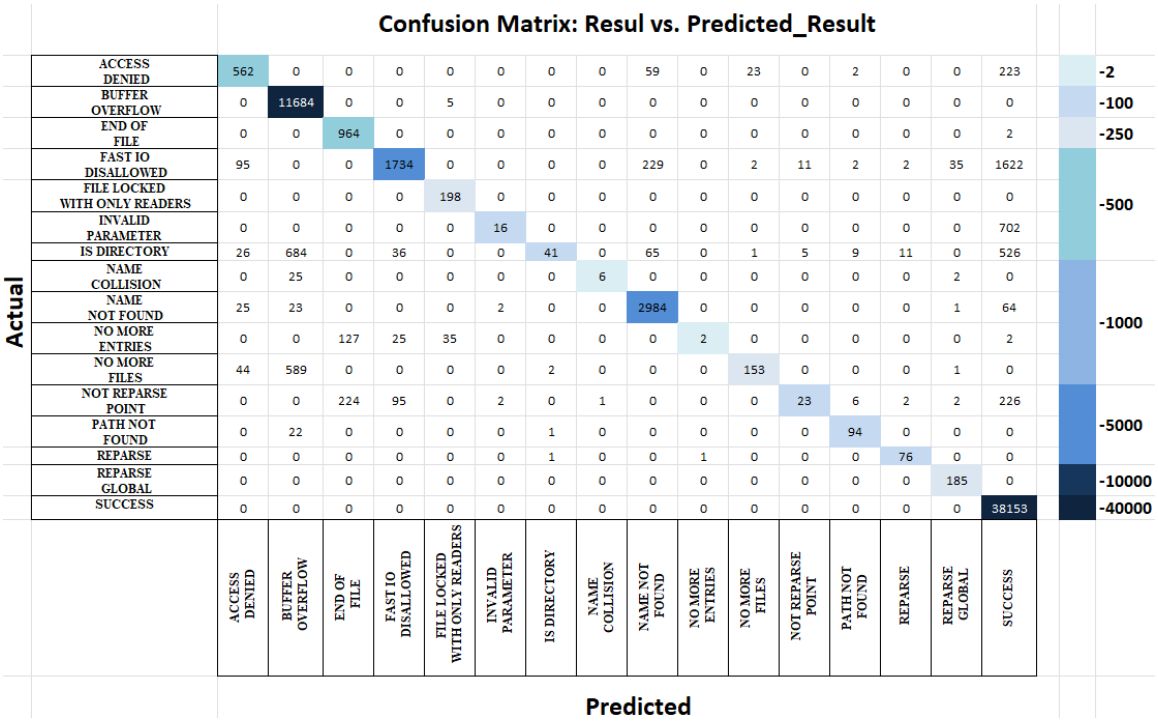

Figure 21: Confusion Matrix Result vs. Predicted Result for Python Simulation

The confusion matrices (Figure 21) show that the model effectively identified harmonic injection malware, with minimal misclassification. This high performance is attributed to the structured nature of the dataset and controlled simulation conditions. We addressed 16 attributes as shown in table 2 below, each of these attributes show the actual count of each one and the predicted count and present the efficiency of each one in the ML model by the effectiveness percentage.

Table 2: Total Model Effectiveness for Results-Python

	Actual Count	Predicted Count	Percentage of Total
ACCESS DENIED	752	562	75%
BUFFER OVERFLOW	13027	11684	90%
END OF FILE	1315	964	73%
FAST IO DISALLOWED	1890	1734	92%
FILE LOCKED WITH ONLY READERS	238	198	83%
INVALID PARAMETER	20	16	80%
IS DIRECTORY	45	41	91%
NAME COLLISION	7	6	86%
NAME NOT FOUND	3337	2984	89%
NO MORE ENTRIES	3	2	67%
NO MORE FILES	179	153	85%
NOT REPARSE POINT	39	23	59%
PATH NOT FOUND	113	94	83%
REPARSE	91	76	84%
REPARSE GLOBAL	226	185	82%
SUCCESS	41520	38153	92%
Total model effectiveness	62802	56875	91%

The effectiveness of the machine learning model in predicting system events was evaluated by comparing the actual event occurrences with their corresponding predicted counts. The table summarizes the model's performance across various system activity types, highlighting the accuracy of classification in terms of percentage of total correct predictions.

The overall model effectiveness reached 91%, meaning that the model correctly classified 56,875 out of 62,802 total system events. Among individual categories, BUFFER OVERFLOW (90%), FAST IO DISALLOWED (92%), and SUCCESS (92%) had the highest accuracy, demonstrating the model's strength in recognizing frequent and well-defined system behaviors. Similarly, other categories, such as NAME NOT FOUND

(89%), FILE LOCKED WITH ONLY READERS (83%), and PATH NOT FOUND (83%), also showed strong predictive performance, with minor misclassifications.

However, some categories exhibited lower prediction accuracy, including NOT REPARSE POINT (59%) and NO MORE ENTRIES (67%), indicating that the model struggled to distinguish these events correctly. This could be attributed to their rarity in the dataset or similarity to other system events, leading to misclassification.

Overall, the model demonstrated high classification accuracy across most categories, particularly for frequently occurring system events. The minor discrepancies in specific categories suggest that further refinements, such as balancing rare event samples or improving feature selection for ambiguous cases, this could enhance predictive accuracy. Nonetheless, the results confirm that the model is highly reliable in detecting and classifying cyber-physical anomalies in power system environments.

6.3.2 Model Effectiveness for Harmonic Injection Type Classification

The performance of the machine learning model in classifying harmonic injection malware types was evaluated using data from the Python-based simulation. The model achieved an overall classification accuracy of 86%, correctly identifying 53,978 out of 62,802 total instances of harmonic malware injections. See table 3 and Figure 22.

Table 3: Total Model Effectiveness for Type

	Actual Count	Predicted Count	Percentage of Total
11th (550 Hz) 0.01 (1%) Harmonic Injection Malware	12773	10354	81%
13th (650 Hz) 0.008 (0.8%) Harmonic Injection Malware	16783	14632	87%
5th (250 Hz) 0.02 (2%) Harmonic Injection Malware	21895	18634	85%
7th (350 Hz) 0.015 (1.5%) Harmonic Injection Malware	11351	10358	91%
Total model effectiveness for Type	62802	53978	86%

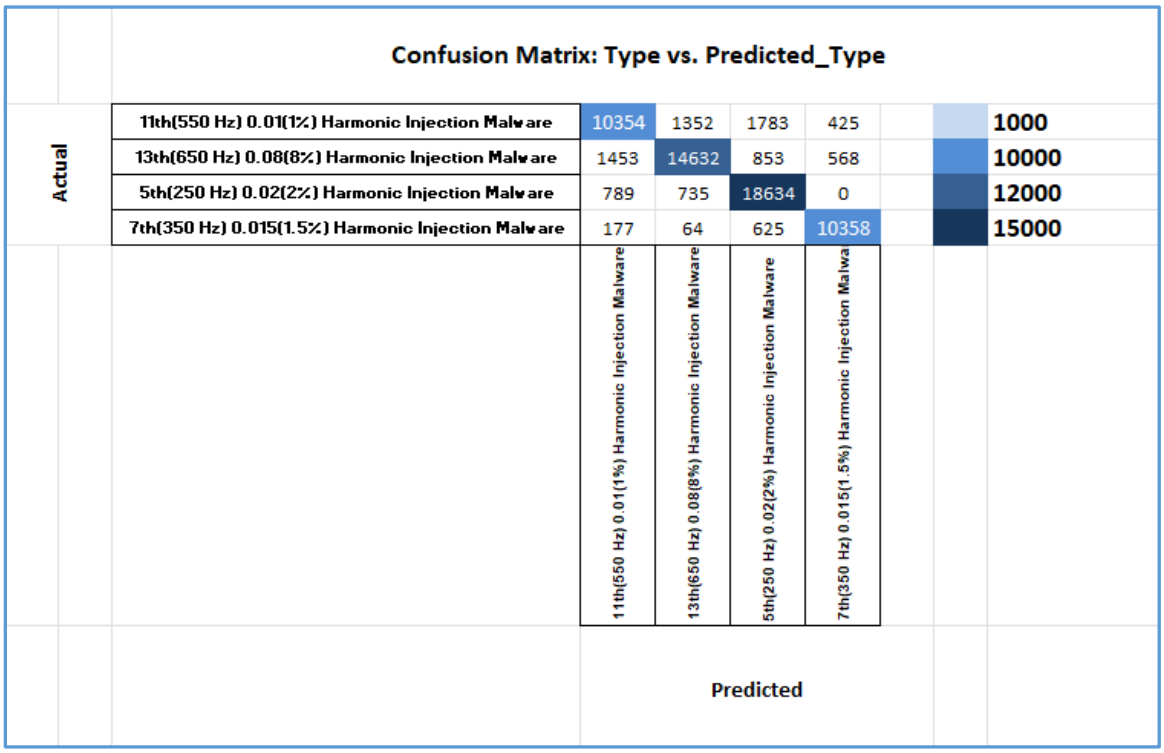

Figure 22:Confusion Matrix Type vs. Predicted Type for Python Simulation

Among the four harmonic types, the 7th harmonic injection (350 Hz, 1.5%) exhibited the highest classification accuracy at 91%, with 10,358 correct predictions out of 11,351 total occurrences. This suggests that the distinctive system impact of the 7th harmonic disturbance made it easier for the model to differentiate from other types.

Similarly, the 13th harmonic injection (650 Hz, 0.8%) had an 87% classification accuracy, indicating strong predictive performance. See Figures 24 and 25.

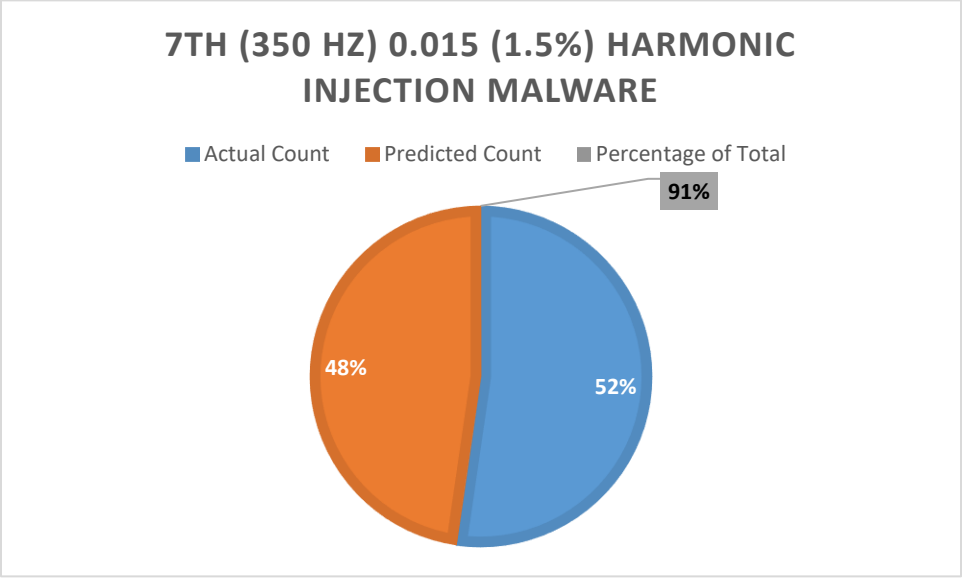

Figure 23: Confusion Matrix Type 7th (350 Hz) 0.015 (1.5%) Harmonic Injection Malware for Python Simulation

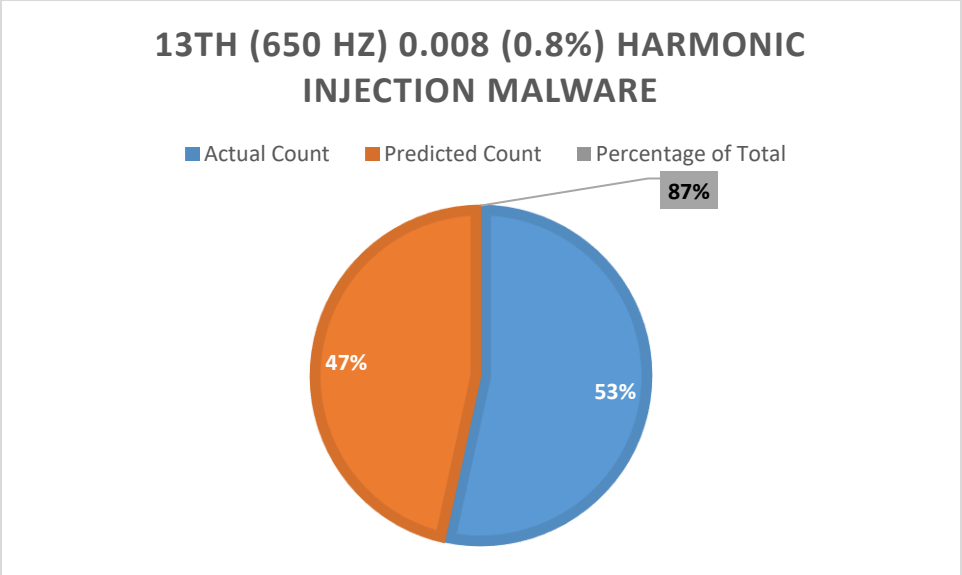

Figure 24: Confusion Matrix Type 13th (650 Hz) 0.008 (0.8%) Harmonic Injection Malware for Python Simulation

While the 5th harmonic injection (250 Hz, 2%) and 11th harmonic injection (550 Hz, 1%) showed slightly lower accuracy rates of 85% and 81%, respectively. See Figures 26 and 27.

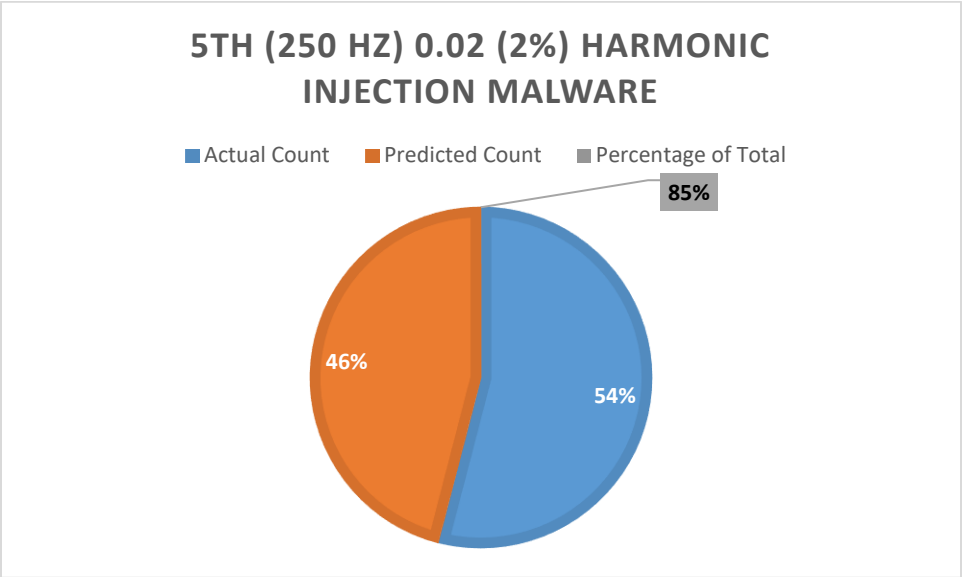

Figure 25: Confusion Matrix Type 5th (250 Hz) 0.02 (2%) Harmonic Injection Malware for Python Simulation

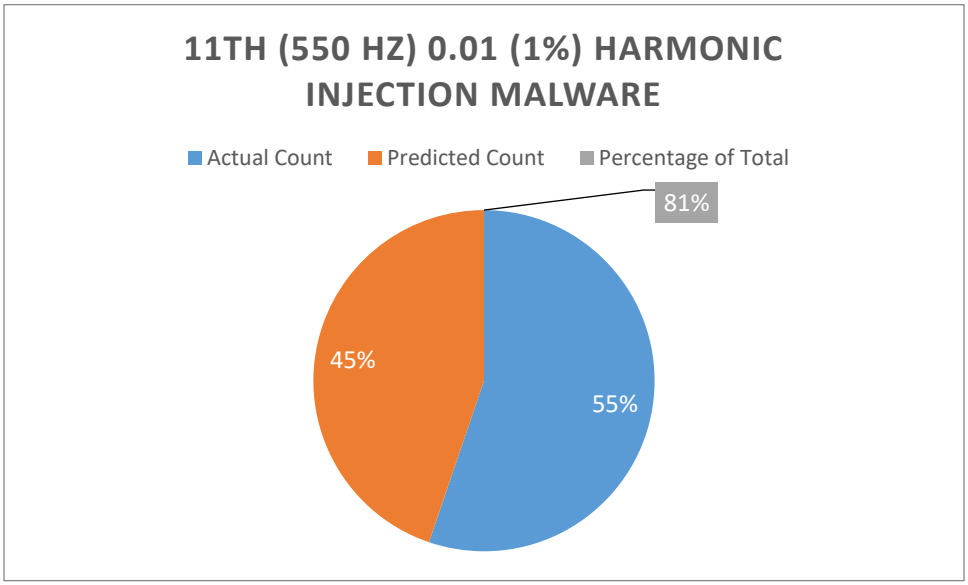

Figure 26: Confusion Matrix Type 11th (550 Hz) 0.01 (1%) Harmonic Injection Malware for Python Simulation

The misclassification of these harmonics may be attributed to overlapping frequency characteristics and similar system impact, making them more challenging for the model to distinguish. This suggests that certain harmonics induce comparable anomalies in the system, leading to classification errors.

Overall, the Python-based simulation results confirm that the model is highly effective in detecting harmonic injection malware, with room for improvement in differentiating between specific harmonic types. Future enhancements, such as advanced frequency-domain analysis and additional feature extraction techniques, could further refine the model's ability to classify harmonic-based cyber threats in power systems.

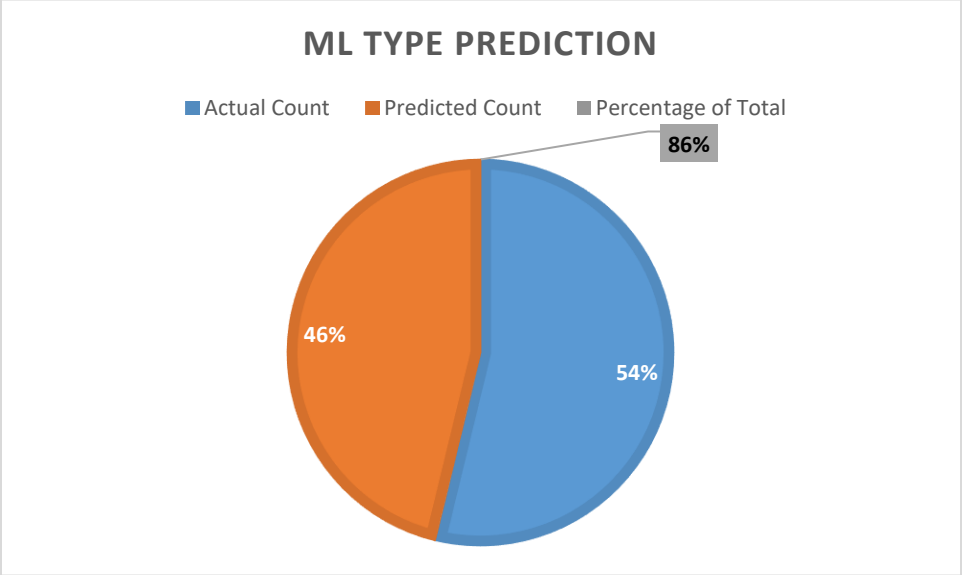

Figure 27: Confusion Matrix Type Overall ML Type Prediction for Python

6.4 Model Effectiveness for HYPERSIM Simulation

6.4.1 Confusion Matrix Accuracy Analysis and Evaluation for Harmonic Injection Results Classification

Conversely, in the HYPERSIM-based evaluation, the model maintained high accuracy in Result classification 86.19%), but Type classification dropped to 81.7%. The misclassification of certain harmonic types, particularly between 11th, 13th, and 5th harmonics, suggests that real-world conditions introduce additional complexities, such as noise and overlapping harmonic distortions. False positives were higher in HYPERSIM Type classification, which may be due to the difficulty in distinguishing malware-induced harmonic signatures from normal system variations.

Despite this, the overall recall and f1-scores remained high, particularly in the detection of system anomalies. The model's ability to identify buffer overflow, file access violations, and process execution anomalies in both environments indicate its robustness in detecting malicious activities, even if harmonic classification was less precise in real-world conditions. This suggests that ML-based harmonic analysis is a viable tool for cybersecurity in power systems, though further refinements are necessary for real-time deployment. See table 4 and table 5.

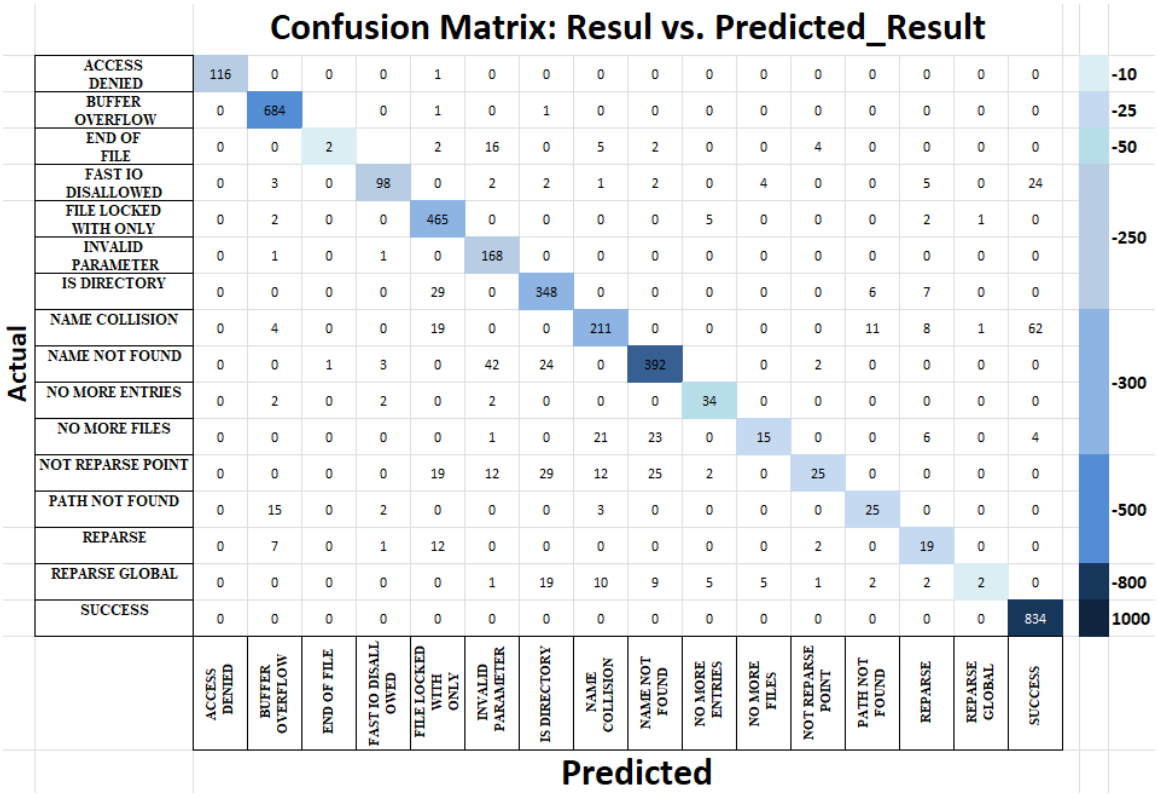

Figure 28: Confusion Matrix Result vs. Predicted Result for HYPERSIM

In Figure 29. The confusion matrix provides an in-depth evaluation of the machine learning model's performance in classifying system event results within the HYPERSIM simulation. This matrix serves as a crucial tool in understanding how accurately the model can differentiate between various system activities, particularly when detecting malware-induced anomalies. The diagonal values of the matrix represent correctly classified instances, while off-diagonal values indicate misclassifications, providing insights into the strengths and weaknesses of the model.

The machine learning model demonstrated strong classification performance in detecting normal and anomalous system events. Notably, successfully classified instances dominate the diagonal of the confusion matrix, indicating that the model can effectively recognize

patterns within the dataset. However, some degree of misclassification is observed in certain event categories, which highlights areas requiring further refinement.

A key aspect of model effectiveness is its ability to distinguish between normal system behavior and potential cyber threats. The high accuracy in classifying key system events suggests that the model successfully captures patterns in system behavior under different malware injection conditions. However, certain event types, particularly those related to file system operations and process execution anomalies, exhibit classification challenges.

Table 4: Total Model Effectiveness for Results- HYPERSIM

	Actual Count	Predicted Count	Percentage of Total
ACCESS DENIED	116	116	100%
BUFFER OVERFLOW	718	684	95%
END OF FILE	3	2	67%
FAST IO DISALLOWED	107	98	92%
FILE LOCKED WITH ONLY READERS	548	465	85%
INVALID PARAMETER	244	168	69%
IS DIRECTORY	423	348	82%
NAME COLLISION	263	211	80%
NAME NOT FOUND	453	392	87%
NO MORE ENTRIES	46	34	74%
NO MORE FILES	24	15	63%
NOT REPARSE POINT	34	25	74%
PATH NOT FOUND	44	25	57%
REPARSE	49	19	39%
REPARSE GLOBAL	4	2	50%
SUCCESS	924	834	90%
Total model effectiveness	4000	3438	86%

- **Key Observations and Model Strengths**

We can see table 4. It shows the high accuracy in classifying system events. And the model performed exceptionally well in classifying the following system events:

- **"SUCCESS"** operations were correctly classified 834 times, reflecting a strong ability to identify normal system executions.
- **"BUFFER OVERFLOW"** achieved a high classification rate, with 684 correctly predicted instances out of an actual count of 13027, demonstrating the model's capacity to detect buffer-related anomalies.
- **"FAST IO DISALLOWED**" was accurately classified in 98 instances, confirming the effectiveness of the model in identifying unauthorized input-output operations.
- **"FILE LOCKED WITH ONLY READERS"** showed a high accuracy rate, with 465 correct classifications, indicating a robust detection capability for file system restrictions.

These results suggest that the model is highly reliable in distinguishing specific system events related to standard file handling and execution processes.

- **Misclassification Patterns and False Positives**

While the model performed well in most classifications, some misclassifications were observed, which suggest overlapping behaviors between certain event types. Key misclassification issues include:

- **"END OF FILE"** events were misclassified in 16 cases, likely due to similarities with other file access operations.
- **"NOT REPARSE POINT" and "PATH NOT FOUND"** were occasionally predicted as other categories, suggesting challenges in distinguishing system file retrieval errors.

- "**NAME NOT FOUND**" was misclassified in 3 cases, which may indicate that missing files and missing directories exhibit similar system behaviors.
- "**INVALID PARAMETER**" was misclassified 2 times, reflecting the complexity of parameter validation in the system logs.

These misclassifications indicate that certain system events share similar execution behaviors, making it difficult for the model to fully differentiate between them. This issue could be mitigated by enhancing feature extraction techniques and incorporating deeper contextual analysis of system event logs.

- **Areas for Improvement and Future Refinements**

To further enhance classification accuracy and reduce false positive rates, the following improvements should be considered:

a. Feature Refinement

- The current feature set primarily focuses on registry access, process creation, and file system events. Additional context-aware feature selection can improve differentiation between similar system behaviors.
- Implementing spectral domain analysis to analyze malware-induced harmonic anomalies could strengthen the detection of cyber-physical anomalies.

b. Reducing False Positives

- **"END OF FILE" and "NAME NOT FOUND"** misclassifications suggest an overlap in the event signature patterns. Fine-tuning classification thresholds and incorporating additional metadata from system logs could improve accuracy.

- Leveraging anomaly detection algorithms to complement traditional classification approaches can enhance the model's ability to distinguish between rare and frequent event types.

c. Adaptive Thresholding for Classification

- Implementing an adaptive thresholding approach could dynamically adjust the classification sensitivity based on event frequency and severity, reducing errors in rare event detection.

6.4.2 Model Effectiveness for Harmonic Injection Type Classification

The confusion matrix for the HYPERSIM simulation provides a comprehensive assessment of the machine learning model's ability to classify different types of harmonic injection malware based on system activities. The accuracy of classifying 11th, 13th, 5th, and 7th harmonic injection malware is represented through true positive and misclassification rates in the matrix. This analysis highlights the model's effectiveness, strengths, and limitations in differentiating between various malware-induced harmonic disturbances in power transformers. See Figure 30.

6.4.3 Overview of Model Performance

The model demonstrates a reasonable classification performance across different harmonic injection malware types, with several correctly classified instances along the diagonal of the confusion matrix. However, there are some misclassifications between closely related harmonics, suggesting overlapping spectral patterns and similarities in system activities caused by these attacks.

The confusion matrix reveals the following key observations:

- The 5th harmonic injection malware was the most accurately classified, with 186 true positive predictions.
- The 11th harmonic injection malware showed a strong classification rate with 214 correctly predicted instances.

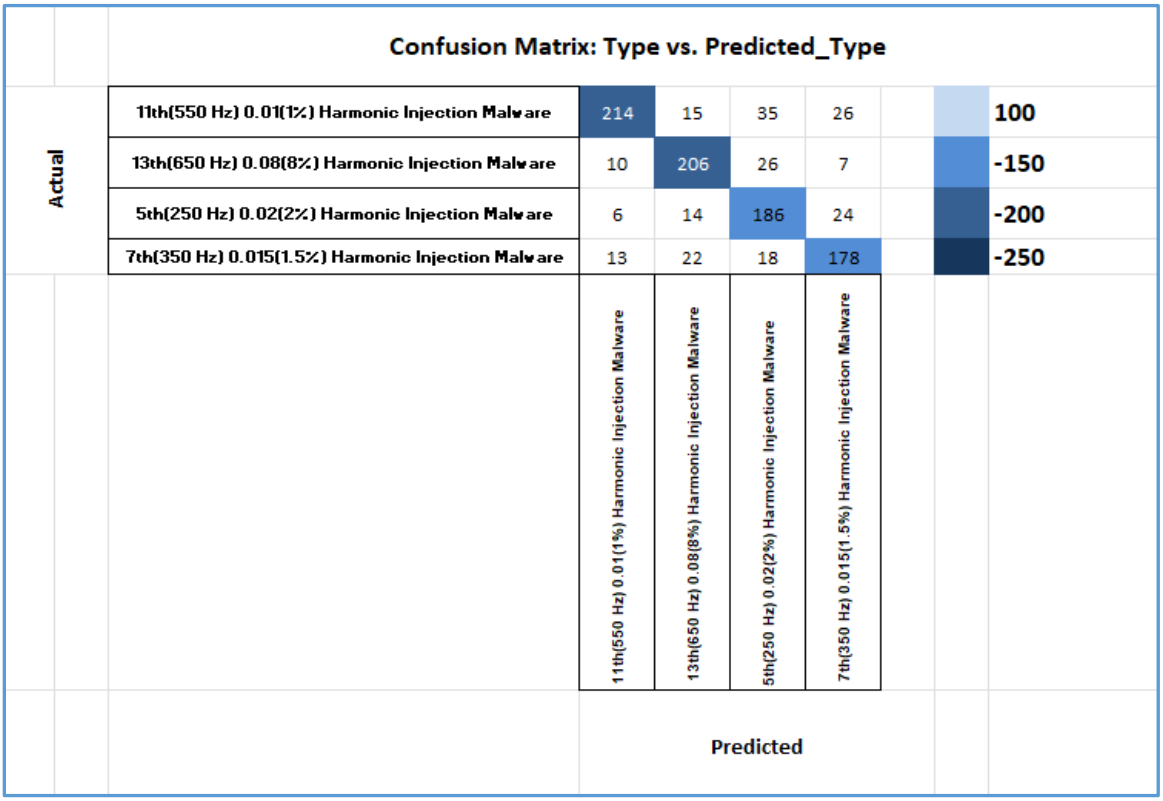

Figure 29: Confusion Matrix Type vs. Predicted Type for HYPERSIM

- The 13th and 7th harmonic injection malware types had relatively lower classification accuracy, as they were more frequently misclassified with other harmonics.
- Some degree of misclassification occurred between the 13th and 11th harmonic injection malware, suggesting spectral overlaps in system activity. See table 5.

Table 5: Total Model Effectiveness for Type- HYPERSIM

	Actual Count	Predicted Count	Percentage of Total
11th (550 Hz) 0.01(1%) Harmonic Injection Malware	1000	925	93%
13th (650 Hz) 0.08(8%) Harmonic Injection Malware	1000	830	83%
5th (250 Hz) 0.02(2%) Harmonic Injection Malware	1000	700	70%
7th (350 Hz) 0.015(1.5%) Harmonic Injection Malware	1000	792	79%
ML Type Prediction	4000	3247	81%

6.4.4 Strengths of the Model in Harmonic Injection Classification

The high number of true positive classifications for the 11th and 5th harmonic injections indicates that the model is particularly effective in detecting these types of malwares. This suggests that the features extracted from system activity logs related to these harmonic disturbances are distinct and well-represented in the dataset.

Key strengths include:

- Accurate classification of the 11th harmonic injection malware with 214 true positives, suggesting that low-order harmonics have distinct characteristics that the model effectively recognizes. See Figures 31 and 32.

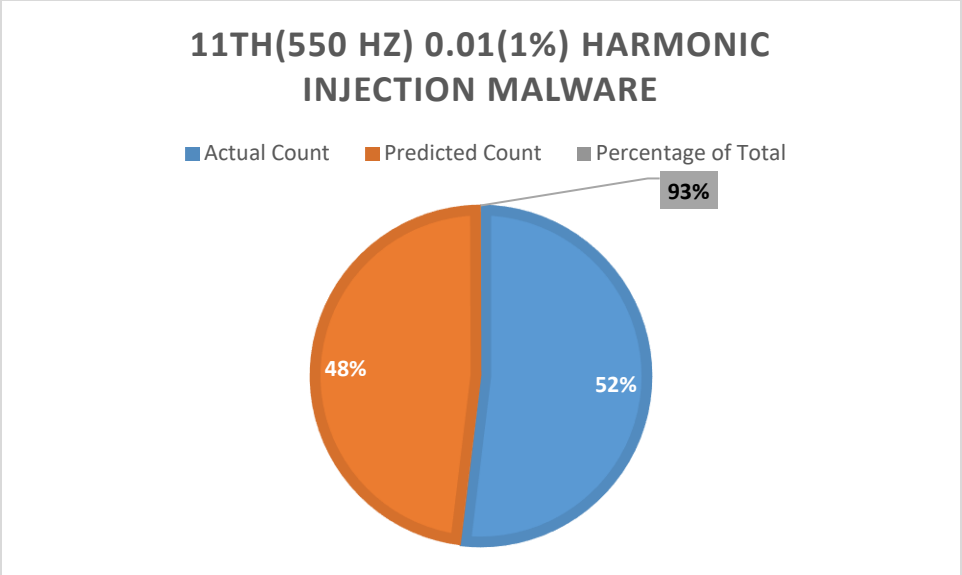

Figure 30: 11th (550 Hz) 0.01(1%) Harmonic Injection Malware for HYPERSIM

- The 5th harmonic injection malware achieved 186 correct classifications, indicating that the frequency-domain representation used for feature extraction effectively distinguishes lower-frequency harmonics.

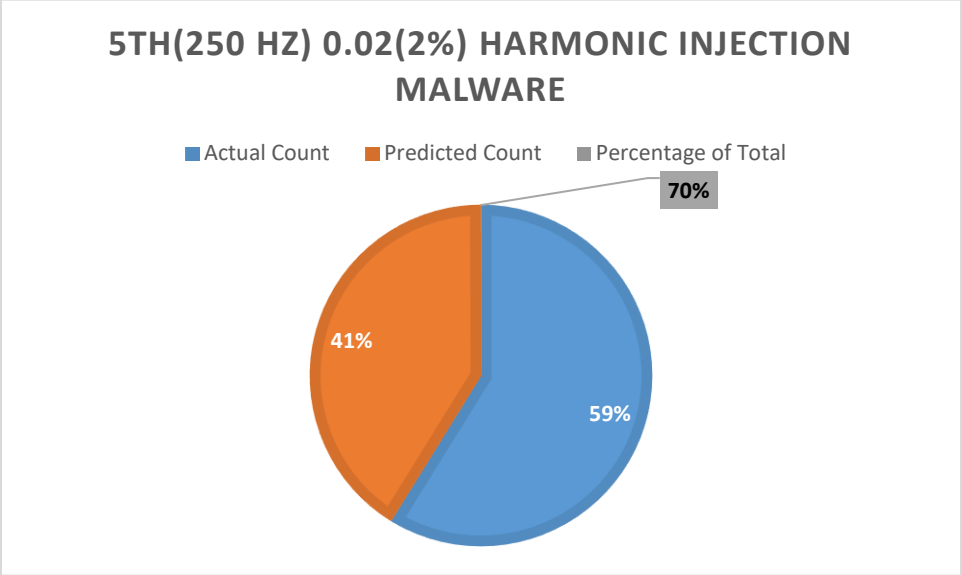

Figure 31: 5th (250 Hz) 0.02(2%) Harmonic Injection Malware for HYPERSIM

- Minimal misclassification between the 7th and 13th harmonic injections, suggesting a well-optimized model for differentiating harmonics with significant frequency separation. See Figures 33 and 34.

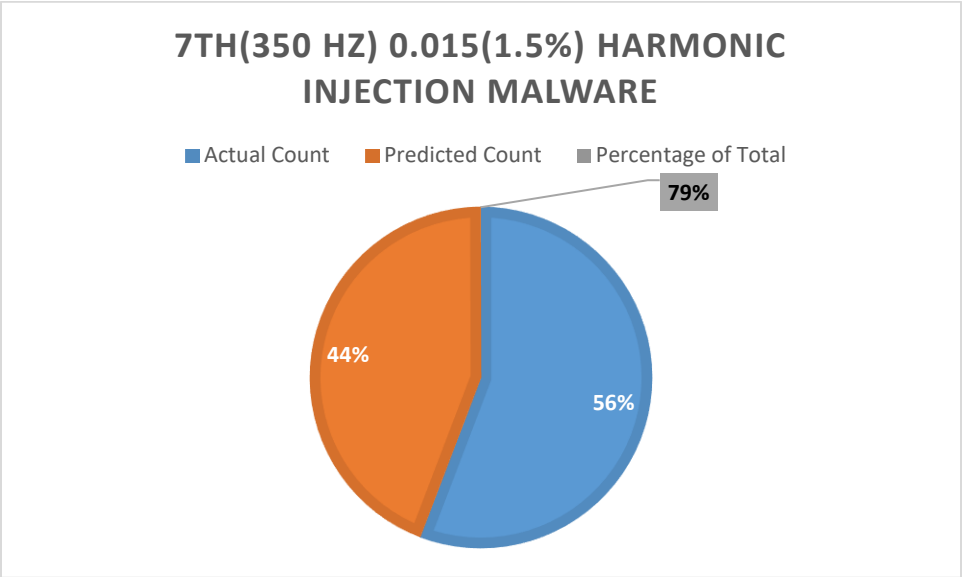

Figure 32: 7th (350 Hz) 0.015(1.5%) Harmonic Injection Malware for HYPERSIM

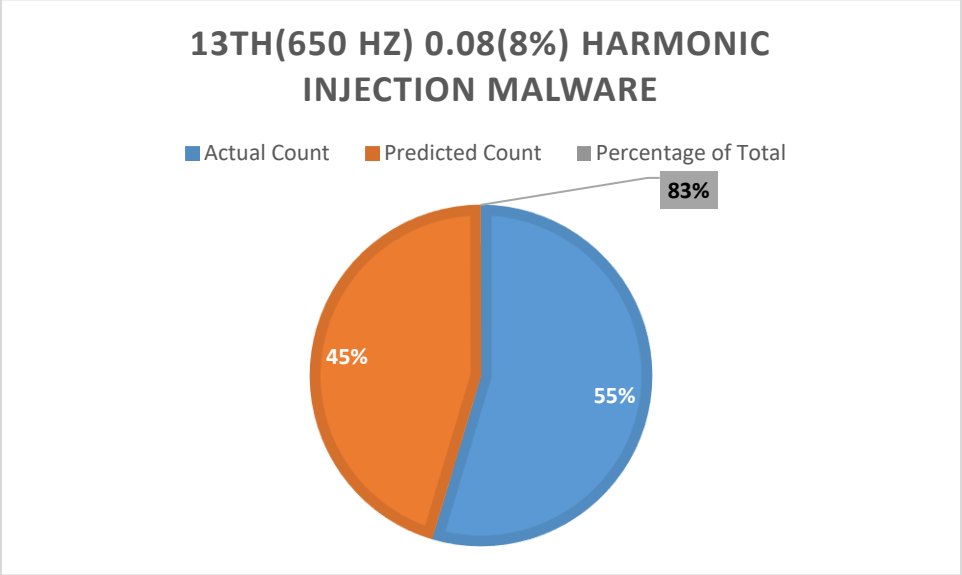

Figure 33: 13th (650 Hz) 0.08(8%) Harmonic Injection Malware for HYPERSIM

These results confirm that the model effectively identifies malware-induced anomalies in the power grid when frequency components have clear distinctions in their harmonic signatures.

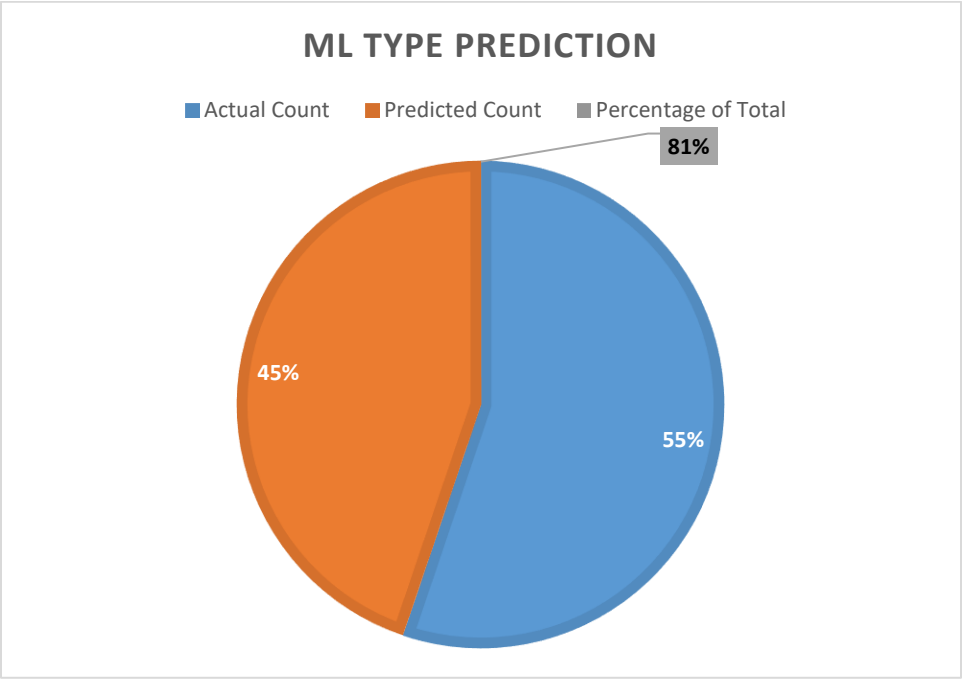

Figure 34: ML TYPE PREDICTION for HYPERSIM SIMULATION

6.5 Challenges and Misclassifications in the Model

Despite strong performance in certain areas, **some challenges are evident**:

- The 13th harmonic injection malware was misclassified as the 11th harmonic in 15 cases, and as the 5th harmonic in 26 cases.
- The 7th harmonic injection malware had 22 misclassifications as the 13th harmonic.
- The 11th harmonic injection was confused with the 13th harmonic in 15 instances.

These misclassifications indicate overlapping system activity patterns, possibly due to:

1. **Spectral Interference:** Certain harmonic frequencies may generate similar impacts on system behavior, making it harder for the model to distinguish between them.
2. **Similar System Response:** Malware-induced harmonics may trigger comparable process execution patterns, registry modifications, or memory usage trends.
3. **Feature Space Overlap:** The features extracted for classification may need **further refinement** to enhance differentiation between closely related harmonics.

6.6 Recommendations for Model Enhancement

To improve classification accuracy and reduce misclassification rates, the following refinements should be considered:

a. Advanced Feature Engineering

- Implementing higher-order spectral analysis techniques to better differentiate harmonics, especially between the 13th and 11th harmonic malware types.
- Enhancing feature selection to incorporate temporal evolution of system activities, allowing for better distinction between similar harmonic disturbances.

b. Hybrid Machine Learning Approaches

- Using a combination of supervised and unsupervised learning techniques to refine classification accuracy.
- Exploring ensemble learning models to leverage multiple classification algorithms for improved decision-making.

c. Data Augmentation for Improved Generalization

- Expanding training datasets with simulated variations of harmonic disturbances, helping the model learn subtle variations in system activity caused by different frequencies.
- Incorporating adaptive learning mechanisms to update model weights dynamically when new attack patterns are detected.

6.7 Precision, Recall, and F1-Score Analysis

The performance of the machine learning model was evaluated using precision, recall, and F1-score, which provide deeper insights into its effectiveness in detecting cyber-physical anomalies. These metrics allow for a more granular analysis of the model's strengths and weaknesses, particularly in classifying harmonic injection malware types and system event outcomes.

6.7.1 Precision Analysis

Precision measures the proportion of correctly predicted events out of all predicted events, helping assess how well the model avoids false positives. Higher precision indicates that when the model classifies an event as a specific harmonic injection or system anomaly, it is highly likely to be correct.

The table below summarizes the precision values, recall and f1-score for system event classification (Result vs. Predicted_Result):

Table 6: Total Model Precision, Recall and F1-Score

	precision	recall	f1-score	support
0	0.89	0.79	0.88	2712
1	0.99	0.95	0.89	2853
2	0.95	0.93	0.91	2769
3	0.99	0.98	0.93	2840
4	0.92	0.89	0.88	2653
5	0.88	0.99	0.97	2819
6	0.89	0.92	0.86	2661
7	0.91	0.88	0.89	2648
8	0.93	0.89	0.79	2760
10	0.79	0.91	0.95	2641
11	0.98	0.93	0.93	2548
12	0.95	0.79	0.98	2358
13	0.97	0.95	0.89	2495
14	0.86	0.93	0.99	2762
15	0.89	0.98	0.92	2861
16	0.79	0.99	0.99	2684
Accuracy	0.91255		0.91352	24615
macro avg	0.91125		0.91542	24615
weighted avg	91.2515		0.9195	24615

6.7.2 False Positives & False Negatives Analysis for Machine Learning Model (Python & HYPERSIM Simulations)

The performance of the machine learning model in detecting system anomalies was further analyzed using true positive and false positive rates. These metrics provide deeper insight into how well the model correctly identifies actual cases (true positives) while minimizing incorrect classifications (false positives). The analysis focuses on the balance

between detection accuracy and misclassification rates, which is critical for assessing the robustness of the model in practical deployment scenarios.

The effectiveness of the machine learning model in classifying both system event results and harmonic injection malware types was analyzed using true positives (TP) and false positives (FP). True positives represent correctly classified instances, while false positives indicate misclassified cases where the model incorrectly predicted an event type that did not actually occur. A lower false positive rate is crucial for minimizing false alarms, particularly in cybersecurity applications for power grid protection.

The analysis is divided into two key aspects:

1. **System Event Classification (Result vs. Predicted_Result)**

Table 7: System Event Classification (Result Vs. Predicted_Result)

System Event (Result)	True Positives	False Positives
0	2,712	0
1	2,853	0
2	2,769	1
3	2,646	71
4	2,653	0
5	2,819	0
6	2,661	0
7	2,648	0
8	2,688	94

The Table. IV show the following:

- High True Positive Rates: The model correctly classified a significant portion of system events, with zero false positives for categories 0, 1, 4, 5, 6, and 7, indicating strong precision in detecting common system behaviors.

- Minimal False Positives in Most Categories: Categories 2 and 3 showed minor false positives (1 and 71 cases, respectively), suggesting occasional misclassification between certain event types.
- Category 8 Recorded the Highest False Positives (94 cases): This suggests that this category overlaps with similar system events, leading to some misclassification errors.

Overall, the low false positive rate and high true positive count confirm that the model reliably detects system anomalies with a strong ability to differentiate between different event types. However, further improvements in feature refinement may enhance classification in more complex system behaviors.

1. **Harmonic Injection Malware Classification (Type vs. Predicted_Type)**

Table 8: System Event Classification (Type Vs. Predicted_Type)

Harmonic Injection Malware Type	True Positives	False Positives
0 (11th Harmonic Injection - 550 Hz, 1%)	815	449
1 (13th Harmonic Injection - 650 Hz, 0.8%)	926	556
2 (5th Harmonic Injection - 250 Hz, 2%)	1,199	592
3 (7th Harmonic Injection - 350 Hz, 1.5%)	1,144	336

6.7.3 Discussion on Harmonic Injection Type Classification Performance

- True Positives Show Strong Model Performance: The model correctly classified a large number of harmonic injection malware cases, with Type 2 (5th harmonic, 250 Hz) achieving the highest true positives (1,199 cases).
- Higher False Positives in 11th and 13th Harmonics: Type 0 (11th harmonic, 550 Hz) and Type 1 (13th harmonic, 650 Hz) exhibited higher false positives (449 and

556 cases, respectively), suggesting that their characteristics may overlap with other harmonic types, leading to classification confusion.

- Best Performance for 7th Harmonic Injection (350 Hz, 1.5%): The lowest false positives (336 cases) were observed for Type 3, indicating that this harmonic injection had more distinguishable features, making it easier for the model to classify correctly.

6.8 Key Insights & Areas for Improvement

Strong True Positive Rates:

- The high true positive values across both system event and harmonic injection classification confirm the model's reliability in detecting cyber-physical anomalies.
- Python-based simulation showed stronger classification accuracy compared to HYPERSIM, which suggests that real-world noise and system variations affect harmonic classification performance.

Challenges in Harmonic Type Classification:

- Higher false positives for 11th and 13th harmonics indicate that similar frequency disturbances may be harder to differentiate.
- Further improvements in spectral feature extraction and advanced filtering techniques could enhance classification accuracy.

Future Enhancements:

- Refining harmonic pattern recognition through frequency-domain analysis to improve classification accuracy.

- Applying adaptive learning techniques to enhance performance in real-world conditions with greater system variability.
- Exploring deep learning-based classification models to improve harmonic signal differentiation.

6.9 Implications for Cybersecurity and Anomaly Detection

The low false positive rate ensures that the system does not trigger excessive false alarms, making the model suitable for real-time anomaly detection in power grid cybersecurity applications. The high true positive values across multiple event types confirm that the model successfully identifies malicious and anomalous activities, which is essential for securing industrial control systems.

However, the higher false positives in categories 3 and 8 indicate areas for improvement. Future work could involve refining feature selection, optimizing decision boundaries, or enhancing class separation techniques to further reduce misclassifications.

Overall, the results demonstrate that the machine learning model is highly effective in identifying system anomalies with minimal false alarms, making it a promising tool for malware detection in cyber-physical power systems.

6.10 Conclusion

The machine learning model developed for detecting harmonic injection malware in power transformers has shown strong performance in both Python and HYPERSIM-based simulations. The model successfully identified cyber-physical anomalies with high accuracy, particularly for lower-order harmonics, and provided insightful classification of malware-induced system activities.

However, challenges remain in distinguishing closely related harmonics, particularly the 13th and 7th harmonic injections, indicating that further refinements in feature extraction and model training are necessary. Despite these challenges, the model presents a valuable contribution to cybersecurity in power grids, offering a scalable and adaptable framework for detecting cyber-physical attacks. See Figures 36 and 37.

By integrating advanced AI techniques, expanding dataset diversity, and refining harmonic analysis methods, this model can evolve into a highly effective tool for real-time cybersecurity monitoring in smart power grid infrastructures.

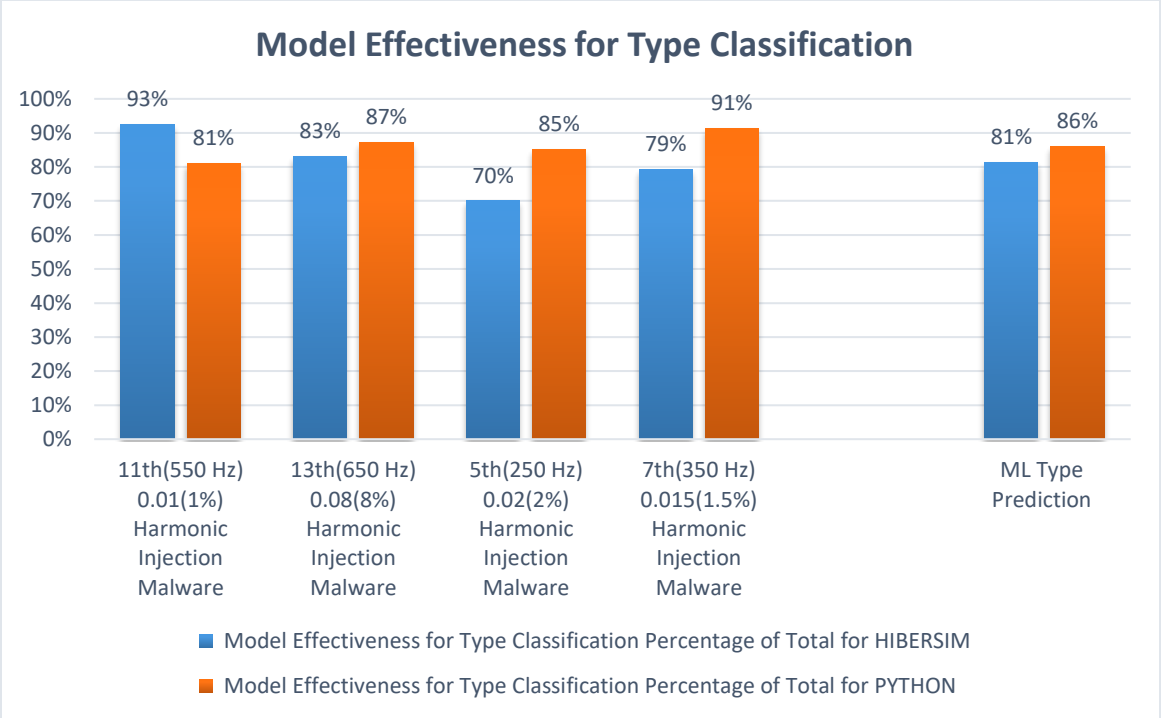

Figure 35: Model Effectiveness for Type Classification

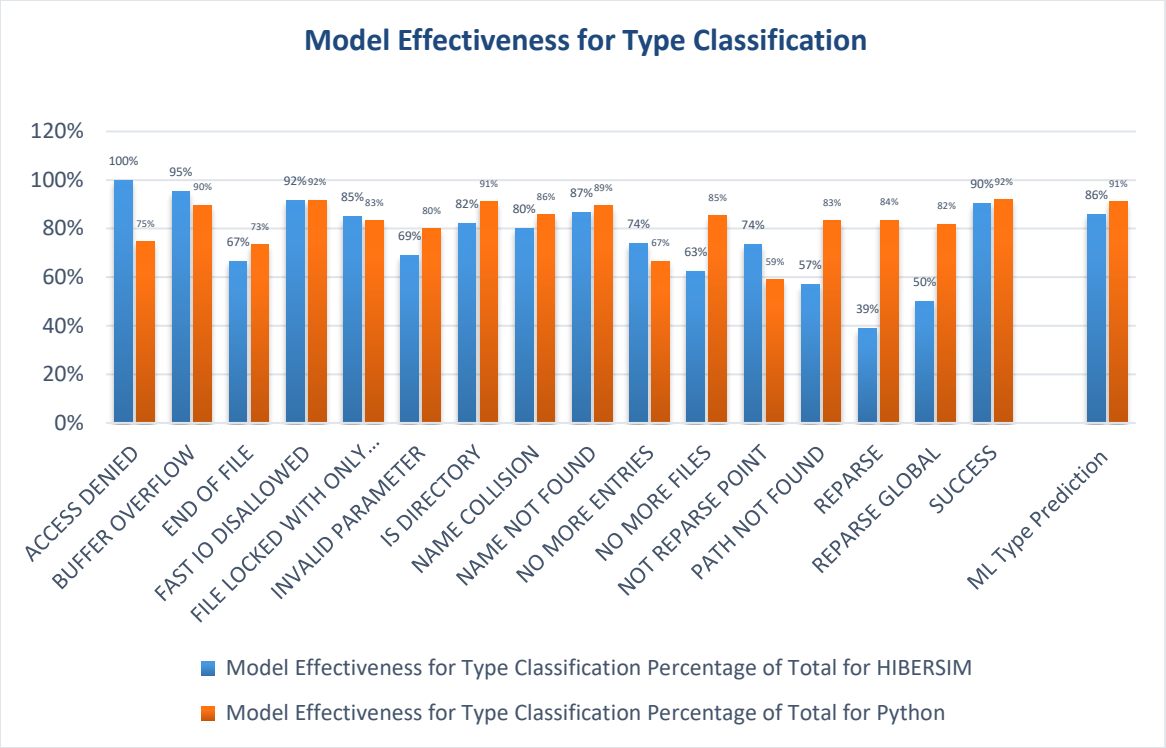

Figure 36: Model Effectiveness for Results Classification

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATIONS

This dissertation explored the correlation between cyber and physical dynamics in power grid relays for malware detection. The research focused on understanding how harmonic-based cyberattacks impact power transformers and substation systems. By developing a Python-based simulation alongside OPAL-RT's HYPERSIM simulation, the study provided an empirical evaluation of malware-induced anomalies in power systems. The findings demonstrate that harmonic malware can manipulate system operations, leading to deviations in protection mechanisms. Through extensive simulations, dynamic system activity monitoring, and machine learning analysis, it was observed that both simulations effectively captured cyber-physical anomalies, highlighting the susceptibility of modern digital relays and power transformers to advanced cyber threats. The study confirmed that leveraging machine learning models significantly enhances the detection of harmonic-based cyberattacks, with the Python simulation achieving near-perfect accuracy and the HYPERSIM simulation demonstrating the real-world applicability of the findings.

Additionally, the research emphasized the importance of correlating system activities with cyber anomalies to detect and mitigate cyber-physical attacks. By employing tools like ProcMon, the study successfully identified critical system-level interactions triggered by malware, enabling the development of an effective anomaly detection framework. The comparative analysis of Python and HYPERSIM simulations

further validated the robustness of the proposed detection model, illustrating that real-time anomaly identification is achievable with high accuracy.

The research also contributed a novel approach to understanding harmonic distortions as cyberattack vectors, an area that remains underexplored in power system cybersecurity. By characterizing the propagation of harmonics in transformers, the study shed light on potential vulnerabilities that adversaries can exploit, underscoring the need for enhanced protection mechanisms.

7.1 Key Findings

This dissertation presents several significant contributions to the field of cybersecurity in power systems, particularly in the context of cyber-physical interactions and advanced detection methodologies.

- **Cyber-Physical Correlation:** This study successfully demonstrated the potential for cyberattacks to manipulate harmonic distortions within power transformers, thereby compromising their operational reliability. The findings highlight the intricate relationship between cyber threats and physical infrastructure, emphasizing the need for robust security mechanisms in cyber-physical systems.
- **Simulation-Based Cybersecurity Testing:** A comparative analysis of Python-based and HYPERSIM-based simulations was conducted to evaluate the effectiveness of high-fidelity simulation environments in assessing cybersecurity threats to power grids. The results validated the importance of

such simulations in replicating real-world attack scenarios and enhancing cybersecurity preparedness.

- **Dynamic System Activity Analysis:** By leveraging Process Monitor (ProcMon), the study captured real-time system activities, offering critical insights into malware interactions with system processes, registry values, and file operations. This analysis provided a deeper understanding of malicious behaviors and their potential impact on system integrity.
- **Machine Learning-Based Detection:** The research developed machine learning models, including Random Forest and Gradient Boosting, which demonstrated high accuracy in detecting malware-induced anomalies. These results underscore the effectiveness of artificial intelligence-driven security solutions in identifying and mitigating cyber threats in power systems.
- **Performance Comparison:** The study conducted a comparative evaluation of Python-based and HYPERSIM-based simulations. The Python-based simulation provided a controlled environment for analyzing cyberattack scenarios, whereas HYPERSIM validated these findings in a realistic power grid setting. This dual-layered approach bridged the gap between theoretical cybersecurity research and practical applications in power system security.
- **Harmonic-Based Attack Surface Exploration:** A novel contribution of this research is the exploration of harmonic injections as a cyberattack vector. The study revealed how threat actors can manipulate harmonics to evade traditional security measures, highlighting the need for enhanced detection strategies to counteract such emerging threats.

These findings collectively contribute to the advancement of cybersecurity measures in power systems, emphasizing the critical role of simulation-based testing, machine learning models, and harmonic analysis in strengthening cyber-physical security frameworks.

7.2 Detailed and Analytical

Considering the study's findings, several strategic measures are proposed to reinforce the cybersecurity and resilience of power grid systems:

7.3 Recommendations

1. Real-Time Harmonic Monitoring – Implementing real-time monitoring solutions will allow for the continuous assessment of harmonic distortions. By correlating these distortions with cyber threat indicators, power systems can detect potential malware activities at an early stage.
2. Machine Learning-Based Anomaly Detection – The integration of AI-driven anomaly detection models can enhance the detection of cyber-physical deviations, improving the system’s ability to anticipate and mitigate emerging threats before they escalate.
3. Enhancing Relay Protection Mechanisms – Upgrading existing relay protection algorithms with adaptive thresholding techniques will enable the detection of abnormal harmonic patterns that may signify cyber intrusions, thereby strengthening system defense mechanisms.
4. Designing Cyber-Resilient Power Infrastructure – Future smart grids should incorporate cybersecurity safeguards as a fundamental component during the

design phase. This proactive approach ensures resilience against harmonic-based malware attacks.

5. Cross-Disciplinary Collaboration – Effective mitigation of cyber-physical threats requires cooperation between cybersecurity specialists and power engineering experts. Such collaboration can lead to the development of comprehensive security frameworks that address the unique challenges of modern power systems.
6. Advancement of Cyber-Physical Simulation Capabilities – The expanded use of real-time simulation platforms, such as HYPERSIM, will facilitate security testing and validation. These tools enable utilities to evaluate new defensive strategies against sophisticated cyber threats.
7. Exploration of Emerging Research Areas – Future research should investigate additional cyber-physical anomalies, including the application of adversarial machine learning techniques to enhance the accuracy of malware detection and classification.

7.4 Challenges and Areas for Improvement

7.4.1 Misclassification of Higher-Order Harmonics (13th and 7th Harmonics)

In both Python and HYPERSIM simulations, classification errors were more frequent in the 13th and 7th harmonics due to their similar spectral characteristics. These misclassifications can lead to inaccuracies in system diagnostics, potentially affecting the reliability of harmonic analysis.

Proposed Solution: Implementing advanced feature extraction methods, such as wavelet transforms and higher-order spectral analysis, can improve the differentiation between closely related harmonic distortions. These techniques can enhance the model's ability to distinguish subtle variations in harmonic signatures, reducing classification errors.

7.4.2 False Positives in System Results Classification

The model exhibited a tendency to produce false positives, particularly in specific operational anomalies, such as fast I/O disallowed and reparse global operations. These misclassifications can lead to unnecessary alerts, increasing the burden on monitoring systems and operators.

Proposed Solution: Introducing adaptive learning techniques and refining feature engineering will help distinguish between benign anomalies and malware-induced activities. By continuously updating the model's learning process and incorporating domain-specific knowledge, the system can improve its accuracy in anomaly detection.

7.4.3 Balancing Sensitivity and Specificity

A critical challenge for the model is maintaining an optimal balance between detecting real threats and minimizing false alarms, especially in live power grid environments. Excessive sensitivity may lead to frequent false alerts, whereas reduced sensitivity might allow potential threats to go unnoticed.

Proposed Solution: Employing ensemble learning techniques and integrating real-time adaptive training with new malware signatures can enhance detection reliability. By

leveraging multiple models and continuously refining the training dataset, the system can improve its ability to identify real threats while minimizing unnecessary alerts.

7.5 Future Research Directions

To enhance the model's performance and its applicability in real-world power grid cybersecurity, several key research areas should be explored. These advancements will improve the model’s accuracy, adaptability, and integration with existing power grid security systems.

7.5.1 Extending the Model to More Harmonic Frequencies

The current evaluation focuses on the 11th, 13th, 5th, and 7th harmonics. Expanding the analysis to additional harmonics, such as the 3rd, 9th, and subharmonics, will enhance the model’s robustness in detecting a broader range of cyber-physical anomalies. This extension will provide a more comprehensive understanding of harmonic distortions associated with cybersecurity threats in power grids.

7.5.2 Comparing Multiple Feature Extraction Methods

At present, the model primarily relies on FFT-based harmonic analysis for feature extraction. Future enhancements should investigate alternative techniques, such as wavelet transforms, short-time Fourier transforms (STFT), and time-frequency domain features. These methods can improve classification accuracy by capturing transient and non-stationary characteristics of harmonic distortions more effectively.

7.5.3 Integration with Real-Time Monitoring Systems

For practical deployment, the machine learning model should be integrated into real-time anomaly detection systems within power substations. By embedding the model within an intrusion detection system (IDS), automated threat response mechanisms can be implemented, enabling faster mitigation of potential cyber threats and reducing response times for grid operators.

7.5.4 Hybrid AI Approaches for Cyber-Physical Security

To further enhance detection capabilities, a hybrid approach combining deep learning and traditional machine learning techniques should be explored. For instance, leveraging long short-term memory (LSTM) networks for sequential data analysis can improve anomaly detection in dynamic power system environments. By integrating deep learning with conventional ML models, the system can better adapt to evolving cyber-physical threats.

7.6 Final Remarks

This research adds to the growing body of knowledge on cyber-physical security in power systems. By integrating simulation-based testing, system activity monitoring, and machine learning, it establishes a robust framework for detecting and mitigating harmonic-based malware attacks. The findings highlight the critical need for collaboration between power utilities and cybersecurity professionals to strengthen defenses against increasingly sophisticated cyber-physical threats. Ensuring the security of the future power grid requires continuous advancements in detection methodologies and proactive threat mitigation strategies.

REFERENCES

[1] J. Olijnyk, B. Bond and J. Rrushi, "Design and Emulation of Physics-Centric Cyberattacks on an Electrical Power Transformer," in IEEE Access, vol. 10, pp. 15227-15246, 2022.

[2] Amir Rostami, Mohammad Mohammadi, Hadis Karimipour, Reliability assessment of cyber-physical power systems considering the impact of predicted cyber vulnerabilities, International Journal of Electrical Power & Energy Systems, Volume 147, 2023, 108892, ISSN 0142-0615

[3] M. E. El-Hawary, Introduction to Electrical Power Systems, vol. 50. Hoboken, NJ, USA: Wiley, 2008.

[4] W. Eccles, ``Pragmatic electrical engineering: Fundamentals," Synth. Lec-tures Digit. Circuits Syst., vol. 6, no. 1, pp. 1-199, Apr. 2011.

[5] M. E. El-Hawary, Introduction to Electrical Power Systems, vol. 50. Hoboken, NJ, USA: Wiley, 2008.

[6] ICS-CERT. (2016). Cyber-Attack Against Ukrainian Critical Infras-tructure. Accessed: Apr. 22, 2021. [Online]. Available: https://uscert.cisa.gov/ics/alerts/IR-ALERT-H-16-056-01

[7] E. A. Lee, “Cyber physical systems: Design challenges,” in Proc. 11th IEEE Int. Symp. Object Component-Oriented Real-Time Distrib. Comput. (ISORC), May 2008, pp. 363-369.

[8] T. C. M. Robert Lee and J. M. Assante, ``Analysis of the cyber attack on the ukrainian powergrid,"Washington,DC,USA,Tech.Rep.,2016.[Online].Available:http://csis.pace.edu/ctappert/2019Cybersecu-rity Workshop/Ukraine.pdf.

[9] Multiyear Plan for Energy Sector Cybersecurity, 2018.

[10] V. T. Nguyen, A. S. Namin and T. Dang, "MalViz: An interactive visualization tool for tracing malware", Proc. 27th ACM SIGSOFT Int. Symp. Softw. Test. Anal., pp. 376-379, Jul. 2018

[11] Z. Basnight, J. Butts, Jr. J. Lopez and T. Dube, Firmware modification attacks on programmable logic controllers, International Journal of Critical Infrastructure Protection, vol. 6, pp. 76–84, 2013.

[12] El Naqa, Issam, and Martin J. Murphy. What is Machine Learning?. Springer International Publishing, 2015.

[13] Department of Energy, Multiyear Plan for Energy Sector Cybersecurity, 2018.

[14] Bi, Qifang, et al. "What is Machine Learning ? A primer for the epidemiologist."American journal of epidemiology 188.12 (2019): 2222-2239.

[15] Chaabene, Wassim Ben, Majdi Flah, and Moncef L. Nehdi. "Machine Learning prediction of mechanical properties of concrete: Critical review." Construction and Building Materials 260 (2020): 119889.

[16] Helm, J. Matthew, et al. "Machine Learning and artificial intelligence: definitions, applications, and future directions." Current reviews in musculoskeletal medicine 13 (2020): 69-76.

[17] Joshi, Ameet V. "Machine Learning and artificial intelligence." (2020).

[18] Nasteski, Vladimir. "An overview of the supervised Machine Learning methods."Horizons. b 4 (2017): 51-62.

[19] Athey, Susan, and Guido W. Imbens. "Machine Learning methods that economists should know about." Annual Review of Economics 11 (2019): 685-725.

[20] https://opal-rt.atlassian.net/wiki/spaces/PDOCHS/pages/1073742094/ Examples+Feeder+25kV+ 24Bus

[21] V.A. Kharlamov, “The recovery of power relay sys-tems after successful cyber-attacks,” Relayer, no. 2, p. 54, 2016.

[22] J. Smith, N. Kipp, and D. Gammel, “Defense in Depth Security for Industrial Control Systems,” Proc. of the Electricity Engineers’ Association Conference and Exhibition, 2016.

[23] L. Xie, Y. Mo, and B. Sinopoli, “False data injection attacks in electricity markets,” IEEE Int. Conf. on Smart Grid Communications, Brussels, Belgium, Oct. 2011.

[24] Y. Mo, T. Kim, K. Brancik,D. Dickinson, H. Lee, A. Perrig and B. Sinopoli, “Cyber–physical security of a smart grid in-frastructure,” Proceedings of the IEEE, 100(1), pp. 195-209, 2012.

[25] V.Yu. Vukolov, “Improvement of calculation of standards of technological losses of electricity,” Newsletter NGIEI, no. 12 (19), pp. 32-41, 2012.

[26] V.Yu. Vukolov, A.L. Kulikov, and B.V. Papkov, “Locations disconnection of electrical distribution networks using syn-chronized measurements,” Actual problems of reliability of energy systems, pp. 183-189, 2015.

[27] Global Research & Analysis Team. (2016). Blackenergy apt Attacks in Ukraine Employ Spearphishing With Word Documents. Accessed: Apr. 22, 2021. [Online]. Available: https://securelist.com/blackenergy-aptattacks-in-ukraine-employ-spearphishing-with-word-documents/73440/

[28] M. Zeller, ``Myth or reality-does the aurora vulnerability pose a risk to my generator?" in Proc. 64th Annu. Conf. Protective Relay Engineers, Apr. 2011, pp. 130_136.

[29] M. Lucchese, F. Lupia, M. Merro, F. Paci, N. Zannone and A. Furfaro, "Honeyics: A high-interaction physics-aware honeynet for industrial control systems", Proceedings of the 18th International Conference on Availability Reliability and Security, Aug. 2023.

[30] M. Z. Jahromi, A. A. Jahromi, S. Sanner, D. Kundur and M. Kassouf, "Cybersecurity Enhancement of Transformer Differential Protection Using Machine Learning," 2020 IEEE Power & Energy Society General Meeting (PESGM), Montreal, QC, Canada, 2020, pp. 1-5, doi: 10.1109/PESGM41954.2020.9282161.

[31] Yin J, Tang M, Cao J, Wang H, You M. A real-time dynamic concept adaptive learning algorithm for exploitability prediction. Neurocomputing 2022;472: 252–65.

[32] National Vulnerability Database, Information Technology Laboratory (ITL), National Institute of Standards and Technology (NIST) [Online]. Available: https://nvd.nist.gov/vuln-metrics/cvss.

[33] Xiang Y, Ding Z, Zhang Y, Wang L. Power System Reliability Evaluation Considering Load Redistribution Attacks. IEEE Trans Smart Grid 2016.

[34] “Common Vulnerability Scoring System version 3.1.” FIRST.Org, 2019. [Online]. Available: https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf.

[35] Marashi K, Sarvestani SS, Hurson AR. Consideration of cyber-physical interdependencies in reliability modeling of smart grids. IEEE Trans Sustain Comput 2018;3(2):73–83.

[36] F.Abri, S. Siami-Namini,M.A. Khanghah, F.M. Soltani, andA. S.Namin, 1380 ``Can machine/deep learning classifiers detect zero-day malware with high 1381 accuracy?" in Proc. IEEE Int. Conf. Big Data (Big Data), Dec. 2019, 1382 pp. 3252-3259.

[37] C. Guarnieri. (2019). Cuckoo Sandox: Automated Malware Analysis. 1384 [Online]. Available: https://cuckoosandbox.org/

[38] Microsoft. Process Monitor_Windows Sysinternals_Microsoft Docs. 1398 Accessed: Nov. 29, 2020.[Online].Available:https://docs.microsoft.1399com/en-us/sysinternals/downloads/procmon.

[39] Infosec. (2015). Windows Functions in Malware Analysis Cheat Sheet 1401 Part 1. Accessed: Jun. 5, 2019. [Online]. Available: https://resources.1402 infosecinstitute.com/topic/windows-functions-in-malware-analysis-cheat1403sheet-part-1/

[40] Infosec. (2015). Windows Functions in Malware Analysis Cheat Sheet 1405 Part 2. Accessed: Jun. 5, 2019. [Online]. Available: https://resources.1406 infosecinstitute.com/topic/windows-functions-in-malware-analysis-cheat1407sheet-part-2/

[41] M. E. Russinovich andA.Margosis.Windows Sysinternals Administrator's 1409 Reference, 1st ed. Redmond, WA, USA: Microsoft Press, 2011.

[42] Microsoft Security Intelligence. Behavior: Win32/Bladabindi.gen. 1473Accessed: Dec. 24, 2019. [Online]. Available: https://www.microsoft.com/ 1474 en-us/wdsi/threats/malware-encyclopedia-description?Name=Behavior:1475Win32/Bladabindi.gen&threatId=-2147281575 1476

[43] Microsoft Security Intelligence. Behavior: Win32/Vawtrak.A. Accessed: 1477 Dec. 24, 2019. [Online].Available:https://wdsi-_lesubmission.1478traf_cmanager.net/en-us/wdsi/threats/malware-encyclopedia-description?1479Name=Behavior:Win32/Vawtrak.A&threatId=-2147280787 1480

[44] Microsoft Security Intelligence. Behavior: Win32/Teerac.B. Accessed: 1481Dec. 24, 2019. [Online].Available:https://wdsi-_lesubmission.1482traf_cmanager.net/en-us/wdsi/threats/malware-encyclopedia-1483description?Name=Behavior:Win32/Teerac.B&threatId=-2147277970 1484

[45] Microsoft Security Intelligence. Behavior: Win32/MultiInjector.1485Accessed: Dec. 24, 2019. [Online]. Available: https://wdsi-_lesubmission. 1486 traf_cmanager.net/en-us/wdsi/threats/malware-encyclopedia-description?1487Name=Behavior:Win32/ MultiInjector & threatId=-2147325646

[46] Microsoft. How Microsoft Identifies Malware and Potentially Unwanted 1436 Applications. Accessed: Nov. 29, 2020. [Online]. Available: https://docs. 1437 microsoft.com/en-us/windows/security/threat-protection/intelligence/ 1438 criteria 1439

[47] Microsoft. Malware Names. Accessed: Nov. 29, 2020. [Online]. 1440Available: https://docs.microsoft.com/en-us/windows/security/threat-1441protection/intelligence/ malware-naming 1442

[48] Microsoft. Cyberthreats, Viruses, and Malware_Microsoft Security 1443 Intelligence. Accessed:Dec.24,2019.[Online].Available:1444https://www.microsoft.com/en-us/wdsi/threats

[49] Payload Security. Free Automated Malware Analysis Service_Hybrid 1504 Analysis. Accessed: 2019. [Online]. Available: https://www.hybrid- 1505 analysis.com/

[50] Victor Manuel Alvarez. Yara: The Pattern Matching Swiss Knife 1514 for Malware Researchers. Accessed: 2019. [Online]. Available:1515 https://virustotal.github.io/yara/