

## Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems Enhancing Resilience of Power Grid Operational Technologies

Presekal, A.

**DOI**

[10.4233/uuid:6a657c5b-adb5-474c-933e-f7166e7b9544](https://doi.org/10.4233/uuid:6a657c5b-adb5-474c-933e-f7166e7b9544)

**Publication date**

2025

**Document Version**

Final published version

**Citation (APA)**

Presekal, A. (2025). *Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems: Enhancing Resilience of Power Grid Operational Technologies*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:6a657c5b-adb5-474c-933e-f7166e7b9544>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems

Enhancing Resilience of Power Grid Operational Technologies



Alfan Presekai

# **ADVANCED PERSISTENT THREAT DETECTION AND CORRELATION FOR CYBER-PHYSICAL POWER SYSTEMS**

ENHANCING RESILIENCE OF POWER GRID  
OPERATIONAL TECHNOLOGIES





# **ADVANCED PERSISTENT THREAT DETECTION AND CORRELATION FOR CYBER-PHYSICAL POWER SYSTEMS**

**ENHANCING RESILIENCE OF POWER GRID  
OPERATIONAL TECHNOLOGIES**

## **Dissertation**

for the purpose of obtaining the degree of doctor  
at Delft University of Technology,  
by the authority of the Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,  
Chair of the Board for Doctorates  
to be defended publicly on  
Wednesday 21 May 2025 at 17.30

**Alfan PRESEKAL**

Master of Science in Computing,  
Imperial College London, United Kingdom  
born in Blora, Indonesia.

The dissertation has been approved by the promotor.

Promotors: Prof.dr. P. Palensky

Copromotor: Dr. A.I. Ștefanov

Composition of the doctoral committee:

Rector Magnificus,

Prof.dr. P. Palensky,

Dr. A.I. Ștefanov,

Chairperson

Delft University of Technology, *promotor*

Delft University of Technology, *copromotor*

*Independent members:*

Prof.dr. G. Smaragdakis,

Prof.dr. J. M. Maza Ortega,

Prof.dr. M. Gibescu,

Dr. J. Hong,

Dr. G. Yang,

Prof.dr. ir P. Bauer,

Delft University of Technology

University of Seville, Spain

Utrecht University

University of Michigan-Dearborn, USA

Technical University of Denmark

Delft University of Technology (*reserve member*)

This research was funded in part by the Designing Systems for Informed Resilience Engineering (DeSIRE) Program High Tech for a Sustainable Future (HTSF) of the 4TU Center for Resilience Engineering (4TU.RE) the Netherlands and the Horizon Europe project Co-operative Cyber Protection for Modern Power Grids (COCOON) with Grant Agreement No. 101120221.



**Keywords:** Advanced Persistent Threats, Artificial Intelligence, Cyber-Physical Power System, Cyber Security, Operational Technology, Smart Grids, Spatio-Temporal Correlation

**Style:** TU Delft House Style, with modifications by Moritz Beller  
<https://github.com/Inventitech/phd-thesis-template>

**Printed by:** ridderprint.nl

**Design & cover:** Krisna Sahwono & Alfian Presekal

**Email & website:** [alfan@ieee.org](mailto:alfan@ieee.org) & [a.presekal.com](http://a.presekal.com)

Copyright © 2025 by A. Presekal

ISBN (Paperback/softback): 978-94-6384-786-5

ISBN (Digital version): 978-94-6518-054-0

An electronic version of this dissertation is available at:  
<https://repository.tudelft.nl/>.

# CONTENTS

<b>Contents</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xiii</b>
<b>Summary</b>	<b>xv</b>
<b>Samenvatting</b>	<b>xix</b>
<b>Ringkasan</b>	<b>xxiii</b>
<b>List of Acronyms</b>	<b>xxvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background and Motivation . . . . .	2
1.2 Thesis Objective and Research Questions. . . . .	3
1.3 Thesis Outlines and Contributions . . . . .	5
<b>2 Cyber Security of Power Systems</b>	<b>11</b>
2.1 Power Grid Digitalization . . . . .	12
2.2 Taxonomy of Cyber Attacks on Power Grids . . . . .	13
2.2.1 Phishing . . . . .	13
2.2.2 Malware . . . . .	14
2.2.3 Network-based Attacks . . . . .	18
2.2.4 Man-in-the-Middle Attacks. . . . .	19
2.2.5 Denial of Service Attacks. . . . .	23
2.2.6 Host-based Attacks. . . . .	24
2.3 Impacts of Cyber Attacks on Power Grids . . . . .	26
2.3.1 Cascading Failures . . . . .	27
2.3.2 Impacts Analysis . . . . .	28
2.4 Vulnerabilities of Power Grid Operational Technologies . . . . .	28
2.4.1 Communication Protocol Vulnerabilities . . . . .	29
2.4.2 Software Application Vulnerabilities . . . . .	32
2.5 Secure Communication Protocols . . . . .	33
2.6 Network Security Controls . . . . .	37
2.6.1 Firewalls . . . . .	37
2.6.2 Intrusion Detection and Prevention Systems . . . . .	38
2.7 Cyber-Physical Power System Co-Simulation and Cyber Range . . . . .	39
2.7.1 Cyber-Physical Power System Co-Simulation . . . . .	39
2.7.2 Cyber Range for Cyber-Physical Power System . . . . .	44
2.8 Conclusion. . . . .	46

<b>3</b>	<b>Advanced Persistent Threat Kill Chain for Cyber-Physical Power Systems</b>	<b>47</b>
3.1	Introduction . . . . .	48
3.2	Advanced Persistent Threats on Cyber-Physical Power System. . . . .	52
3.2.1	APT Characteristics . . . . .	52
3.2.2	APTs in Information Technology Systems . . . . .	53
3.2.3	APTs in Cyber-Physical Systems . . . . .	54
3.2.4	APTs in Cyber-Physical Power Systems . . . . .	54
3.2.5	APT Characteristics in Cyber-Physical Power Systems . . . . .	55
3.3	Advanced Cyber-Physical Power System Kill Chain . . . . .	57
3.3.1	Attack Preparation . . . . .	59
3.3.2	Initial Engagement and IT System Access . . . . .	60
3.3.3	Main Cyber Attack on IT-OT Systems . . . . .	61
3.3.4	Engagement with Physical System and System Recovery Impedi- ment . . . . .	62
3.3.5	Power System Cascading Failures and Blackouts . . . . .	62
3.3.6	Social Impacts and Restoration . . . . .	66
3.4	Advanced Persistent Threats on Power Grids Case Studies . . . . .	67
3.4.1	Ukrainian Power Grid Cyber Attack 2015 . . . . .	68
3.4.2	Ukrainian Power Grid Cyber Attack 2016 . . . . .	71
3.4.3	Ukrainian Power Grid Cyber Attack 2022 . . . . .	73
3.4.4	Experimental Cyber Attack . . . . .	73
3.4.5	Result and Discussion . . . . .	81
3.5	Conclusion and Recommendations . . . . .	83
<b>4</b>	<b>Attack Graph Model For Cyber-Physical Power System Using Hybrid Deep Learning</b>	<b>85</b>
4.1	Introduction . . . . .	86
4.2	Anomaly Detection and Attack Graph Model. . . . .	90
4.2.1	Cyber-Physical System Model . . . . .	91
4.2.2	Traffic Dispersion Graph . . . . .	92
4.2.3	Graph Convolutional Long Short-Term Memory . . . . .	93
4.2.4	Time Series Anomaly Detection . . . . .	95
4.2.5	Attack Graph Model . . . . .	96
4.2.6	Forensic Graph Model . . . . .	98
4.3	Experimental Results. . . . .	99
4.3.1	Experimental Setting. . . . .	99
4.3.2	Network Traffic Prediction . . . . .	100
4.3.3	Anomaly Detection . . . . .	102
4.3.4	Attack Graph Generation and Analysis. . . . .	105
4.3.5	Forensic Graph Generation and Analysis. . . . .	106
4.4	Conclusion. . . . .	109
<b>5</b>	<b>Spatio-Temporal Advanced Persistent Threats Detection and Correlation</b>	<b>111</b>
5.1	Introduction . . . . .	112
5.2	CPPS and Cyber Threat Model . . . . .	116
5.2.1	Cyber-Physical Power System Model. . . . .	116

5.2.2	Cyber Threat Model for Cyber-Physical Power System . . . . .	117
5.3	Spatio-Temporal Anomaly Detection, Correlation and Prediction. . . . .	118
5.3.1	Semi-Supervised Learning for Deep Packet Inspection . . . . .	118
5.3.2	Cyber-Physical System Integration Matrix . . . . .	119
5.3.3	APT Spatio-Temporal Correlation . . . . .	120
5.3.4	Resilient Associative Method. . . . .	122
5.3.5	CPPS Log Comparator . . . . .	123
5.4	Experimental Results. . . . .	124
5.4.1	Experimental Setup . . . . .	125
5.4.2	Semi-Supervised Deep Packet Inspection for OT Anomaly Detec- tion . . . . .	126
5.4.3	Spatio-Temporal Anomaly Correlation and Prediction . . . . .	128
5.4.4	Detection for Zero Day Attack Scenarios. . . . .	130
5.5	Conclusion. . . . .	130
<b>6</b>	<b>Intrusion Detection System using Semi-Supervised Learning and Similar- ity Clustering</b>	<b>133</b>
6.1	Introduction . . . . .	134
6.2	Semi-Supervised Intrusion Detection System for Digital Substation . . . .	136
6.2.1	Digital Substation Architecture and Cyber Threats . . . . .	137
6.2.2	Traffic Distance Similarity Vectors in Digital Substations. . . . .	137
6.2.3	Hybrid Semi-Supervised Intrusion Detection System. . . . .	139
6.2.4	Digital Substation Traffic State Transition Model. . . . .	139
6.3	Experimental Results and Discussion. . . . .	140
6.3.1	Experimental Setting and Dataset . . . . .	140
6.3.2	Traffic Payload Distance Similarities . . . . .	142
6.3.3	Traffic Interarrival Distance Similarities . . . . .	143
6.3.4	Hybrid Semi-Supervised Classifier . . . . .	144
6.3.5	Digital Substation State Transition Analysis . . . . .	146
6.4	Conclusion. . . . .	148
<b>7</b>	<b>Conclusion and Discussion</b>	<b>151</b>
7.1	Conclusion. . . . .	151
7.2	Discussion and Future Research . . . . .	154
	<b>Bibliography</b>	<b>157</b>
	<b>List of Publications</b>	<b>183</b>
	<b>Acknowledgments</b>	<b>185</b>
	<b>Author Biography</b>	<b>187</b>



# LIST OF FIGURES

1.1	Thesis outline. . . . .	6
2.1	Taxonomy of cyber attacks on power systems and ICS [1]. . . . .	13
2.2	Lifecycle of a software vulnerability [2]. . . . .	32
2.3	Mapping of OSI layers, cyber attacks and mitigation techniques [3]. . . . .	34
2.4	Summary of secure protocol research and classification [3]. . . . .	35
2.5	Comparison of communication network simulators for CPS modelling [3]. . . . .	40
2.6	Comparison of real hardware, virtual machines, and container-based system [3]. . . . .	41
2.7	CPS architecture in CRoF at TU Delft [3]. . . . .	42
2.8	CPPS co-simulation architecture with Windows and Ubuntu servers. . . . .	43
2.9	Servers allocation for CPPS co-simulation. . . . .	43
2.10	Data flow diagram from CPPS co-simulation . . . . .	44
2.11	CPS and cyber range architecture of CRoF at TU Delft [3]. . . . .	45
2.12	Blue and red team tools for power grid IT-OT systems in CRoF at TU Delft [3]. . . . .	45
3.1	Advanced Cyber-Physical power system kill chain framework [4]. . . . .	58
3.2	Flowchart representing sequences of APTs on power grids according to the ACPPS kill chain [4]. . . . .	59
3.3	Cyber attack on Ukrainian power grids 2015 [4]. . . . .	69
3.4	Cyber attack on Ukrainian power grids 2016 [4]. . . . .	72
3.5	Cyber-physical co-simulation experimental setup to analyze the impact of cyber attacks on the power system. The simulated power grid is based on the IEEE 39 bus system with 27 Substations (Sub). The cyber attack impact on power grid started from the malicious opening of the circuit breaker at Bus 2 in Substation 2 [4]. . . . .	75
3.6	Traffic comparison from normal and DoS traffic [4]. . . . .	77
3.7	Box plot comparison from normal and DoS traffic [4]. . . . .	77
3.8	Graph visualization from attack on substation 2 [4]. . . . .	78
3.9	Bus voltage magnitude [4]. . . . .	78
3.10	Generator rotor angles [4]. . . . .	79
3.11	Rate of Change of Frequency on Generator Generator 4 and Generator 8 [4]. . . . .	79
3.12	Change of load active power for load 03, 04, 18, 23, 26, and 28 [4]. . . . .	79
3.13	Cascading impact visualization [4]. . . . .	80
4.1	Cyber kill chain stages and impacts [5]. . . . .	86
4.2	Abstraction layers of SDN architecture [5]. . . . .	88

4.3	Attack graph creation using CyResGrid method [5]. . . . .	90
4.4	Cyber-physical system model of the power grid with IT-OT communication networks [5]. . . . .	91
4.5	Traffic Dispersion Graph (TDG) processes [5]. . . . .	93
4.6	Traffic dispersion graph of 27 substations [5]. . . . .	93
4.7	CyResGrid – hybrid deep learning model [5]. . . . .	95
4.8	CyResGrid – hybrid deep learning model [5]. . . . .	97
4.9	Forensic graph ( <i>FGraph</i> ) model [5]. . . . .	99
4.10	Comparison of real and predicted traffic under normal conditions [5]. . . . .	101
4.11	Histogram of real and predicted traffic under normal conditions [5]. . . . .	101
4.12	Statistical comparison of real (r) and predicted traffic (p) [5]. . . . .	102
4.13	Comparison of throughput between real and predicted OT traffic for sneaky network scanning cyber attack scenario [5]. . . . .	102
4.14	Dataset for time series classification [5]. . . . .	103
4.15	ROC comparison of the deep learning-based TSC [5]. . . . .	104
4.16	ROC comparison of the hybrid deep learning-based TSC [5]. . . . .	105
4.17	Attack graph maps to identify and visualize cyber attack locations [5]. . . . .	106
4.18	Digital substation experimental setup for OT traffic generation [6]. . . . .	107
4.19	Statistical comparison between normal, predicted, and attack or anomalous traffic for data A, B, and C [6]. . . . .	108
4.20	ROC comparison for data A, B, and C [6]. . . . .	109
4.21	Forensic graphs for normal traffic and anomalous traffic [6]. . . . .	110
5.1	Cyber-physical system model of the power grid with IT/OT communication networks [7]. . . . .	114
5.2	Integrated processes for spatio-temporal anomaly detection, correlation, and prediction [7]. . . . .	115
5.3	Cyber-physical system integration matrix [7]. . . . .	120
5.4	Vector database query search strategy with KNN [7]. . . . .	122
5.5	Circuit breaker log comparator function [7]. . . . .	123
5.6	CPPS Adjacency Matrix [7]. . . . .	125
5.7	Statistical box plot from normal traffic and cyber attacks [7]. . . . .	127
5.8	OT traffic images representation and the result of Gaussian Mixture with partial labelling for each protocols [7]. . . . .	127
5.9	Proportion of labeled data impact on the MSE for all protocols [7]. . . . .	128
5.10	ROC curve from all traffic categories [7]. . . . .	129
5.11	Sample EGC-LSMT Prediction Results from case 76 and 218 [7]. . . . .	130
5.12	Search algorithms computation time comparison [7]. . . . .	131
6.1	Digital substation architecture. . . . .	136
6.2	Summary of the proposed hybrid semi-supervised intrusion detection systems. . . . .	136
6.3	Digital substation cyber-physical system state transition. . . . .	140
6.4	Digital substation co-simulation architecture with HIL. . . . .	141
6.5	Comparison of packet size statistical characteristics from data A, B, and C in box plots. . . . .	142



6.6	scatter plot based on the CNN ( $\alpha$ ) and Chebyshev distance ( $\beta$ ) for GOOSE normal and spoofing, and other cyber attacks. . . . .	143
6.7	Characteristic of interarrival time from sample data A, B, and C. The top row plots show probability density distribution with the normal distribution curve of interarrival time. The bottom row plots show Fast Fourier Transform (FFT) amplitude for the traffic interarrival time. . . . .	144
6.8	3D scatter plots for data A, B, and C with the blue nodes represent normal GOOSE and red nodes represent anomalous GOOSE. . . . .	145
6.9	Comparison of average value of $\alpha, \beta, \gamma$ under normal, fault, and spoofing traffic. . . . .	146
6.10	3D scatter plots for traffic under normal, spoofing, and faults. . . . .	147
6.11	3D scatter plots for traffic transition from normal operation, faults (1 and 2), reclosure (3), and spoofing (4). . . . .	147
6.12	Time series plot from the simulated events representing value of traffic vector $\Phi = \langle \alpha, \beta, \gamma \rangle$ , circuit breaker status, current (kA) and voltage (kV). .	148



# LIST OF TABLES

1.1	Cyber Attacks Targeting Cyber-Physical Power Systems . . . . .	2
2.1	Comparison of malware capabilities . . . . .	14
2.2	History of malware involved in major ICS cyber-related incidents . . . . .	15
2.3	Summary of known cyber attacks on power grids and their impact . . . . .	27
2.4	Summary of network security control applications. . . . .	37
2.5	Cyber-physical system models for power systems research. . . . .	39
3.1	Comparison of ACPPS stages with other kill chain frameworks. . . . .	51
3.2	Comparison of APT attacks and conventional cyber attacks. . . . .	52
3.3	Cyber attacks targeting IT systems. . . . .	53
3.4	Cyber attacks targeting cyber-physical systems. . . . .	54
3.5	Cyber attacks targeting cyber-physical power systems. . . . .	55
3.6	Comparison of APT attacks in IT Systems, General CPS, and CPPS. . . . .	56
3.7	Impacts Comparison of Attacks in IT System, General CPS, and CPPS. . . . .	56
3.8	Advanced Cyber-physical system kill chain stages in Ukraine cyber attack 2015. . . . .	69
3.9	Advanced cyber-physical system kill chain stages in Ukraine cyber attack 2016. . . . .	72
3.10	Advanced cyber-physical system kill chain stages in simulated cyber attack. . . . .	76
3.11	Summary of ACPPS Kill Chain Implementation for Real Cyber Attack in Ukraine 2015 and 2016, and Experimental Cyber Attacks. . . . .	82
3.12	Comparison of ACPPS Kill with Other Frameworks for Cyber Attack Stages Identification. . . . .	82
4.1	Cyber Attack Scenarios . . . . .	103
4.2	Performance Comparison of Anomaly Detection Methods . . . . .	104
4.3	Summary of Network Traffic Data . . . . .	107
5.1	CPS Breaker Log Comparator Categories . . . . .	124
5.2	MSE Scores Comparison of Graph-Based Spatio-Temporal Deep Learning Models . . . . .	129
5.3	Performance Comparison of Graph-Based Spatio-Temporal Deep Learning Models with Sequential and Neural Network Filters . . . . .	130
6.1	Comparison of GOOSE Traffic Data . . . . .	141
6.2	Performance Comparison of Clustering Methods . . . . .	145



# SUMMARY

Power grids are experiencing a digital transformation through the integration of information and communication technologies, such as the Internet of Things (IoT), big data, and Artificial Intelligence (AI). All of these technologies enhance the operational efficiency and intelligence of power grids. However, digitalization introduces new vulnerabilities in power systems, highlighting the urgent necessity of strengthening cyber resilience to protect the stability and security of power grids against emerging threats. Incidents in the real world, including the Ukrainian power grid cyber attacks in 2015, 2016, and 2022, illustrate the imminent threat presented by cyber adversaries utilizing Advanced Persistent Threat (APT). In contrast to traditional cyber attacks, APTs exhibit more sophisticated techniques portrayed to their characteristics, including stealthy tactics, prolonged persistence, and exploitative use of zero-day vulnerabilities. Due to the characteristics of APTs, traditional cyber security measures are inadequate for addressing these challenges. Motivated by these challenges, this thesis is focused on APT detection and correlation on cyber-physical power systems.

This thesis starts with an investigation of cyber security in power grids, which is crucial to formulating effective mitigation strategies. It provides an in-depth analysis of the cyber threat landscape, system vulnerabilities, state-of-the-art mitigation techniques, and cyber attack modeling within cyber-physical power systems. Building on this foundation, the thesis proposes an advanced cyber-physical power system kill chain to overcome the limitations of existing frameworks for identifying cyber attack stages. The focus of this research is to address APTs on power grids by targeting three key challenges of APTs (1) stealthiness, (2) prolonged persistence, and (3) zero-day attack.

*APTs stealthiness:* APTs are highly stealthy, utilizing advanced methods to evade detection. They frequently employ concealed attack processes, hiding behind legitimate traffic to evade traditional security measures such as Intrusion Detection System (IDS) and firewalls. The stealthy nature of APTs is characterized by infinitesimal anomalies and insignificant changes compared to legitimate traffic. Consequently, it presents scientific challenges regarding the necessity for highly sensitive anomaly detection systems.

*APTs prolonged persistence:* APTs are specifically designed to evade detection for prolonged durations, spanning months or even years. To address this challenge, it is crucial to analyze anomaly correlation over prolonged periods. However, the state-of-the-art anomaly detection techniques for cyber-physical power systems are predominantly formulated as standalone solutions, addressing either power system anomalies or cyber anomalies independently. These methodologies are constrained in scope as they concentrate on identifying individual anomaly instances without addressing the interdependencies and correlations among multiple anomalies across systems in long-terms. Consequently, they are insufficient for detecting and correlating the lateral movement of anomalies triggered by APTs. The main scientific challenge at this point lies in the detection of low-frequency, timely non-

deterministic, and stealthy anomalies, which often evade traditional detection techniques due to their evasive and intermittent nature.

*APTs zero-day attacks:* Adversaries often implement zero-day attacks that leverage a vulnerability in the communication protocols, software, or hardware that remains undisclosed to the vendor or the public. Conventional security systems, such as antivirus software, firewalls, and IDS, rely heavily on signature-based detection, where known attack patterns are identified using predefined signatures. However, a zero-day attack is a new cyber attack without a known signature or any preliminary information. From a scientific standpoint, the challenge of detecting zero-day attacks resulted from the need for anomaly-based detection methods. These methods operate independently from pre-existing attack signatures or known threat patterns. Unlike signature-based methods, which rely on prior knowledge, zero-day detection must focus on identifying anomalies from baseline system behavior.

Based on the discussion above, this thesis addresses the aforementioned challenges of APTs by proposing novel hybrid deep learning models using graph-based deep learning and semi-supervised learning. In particular, the major contributions of this thesis are summarized as follows.

*Cyber-Physical Power System Model and Advanced Cyber-Physical Power System Kill Chain:* This thesis provides a comprehensive investigation of cyber security in power systems, with a focus on the evolving cyber threat landscape, system vulnerabilities, and existing mitigation strategies. It introduces a cyber-physical power system model that incorporates a cyber range, designed to simulate both attacks and defenses within a high-fidelity environment. Additionally, it proposes an Advanced Cyber-Physical Power System (ACPPS) Kill Chain framework. The ACPPS Kill Chain identifies the APT characteristics that are unique to power systems. It defines and examines the cyber-physical APT stages spanning from the initial phases of infiltration to cascading failures and a power system blackout. These elements together are critical for defining and implementing effective mitigation strategies against APTs on power systems.

*Attack Graph Model:* For addressing stealthy APTs, this thesis proposes an attack graph model that leverages a Software Defined Networking (SDN) for online situational awareness of cyber attacks. The model utilizes a hybrid deep learning approach combining Graph Convolutional Long Short-Term Memory (GC-LSTM) and Convolutional Neural Network (CNN) to classify Operational Technology (OT) network traffic as either anomalous or normal. This advanced deep learning model can detect infinitesimal traffic anomalies, significantly reducing false positives and false negatives. Furthermore, the proposed method can accurately identify the specific location of OT anomalies in near real-time, offering enhanced responsiveness in detecting and mitigating APTs.

*APT Spatio-Temporal Correlation:* To address the prolonged persistence of APTs, this thesis proposes an APT spatio-temporal correlation strategy. This approach utilizes a Cyber-Physical System Integration Matrix (CPSIM) and an Enhanced Graph Convolutional Long Short-Term Memory (EGC-LSTM) model. The CPSIM matrix constructs a topological correlation between cyber and physical system anomalies in Cyber-Physical Power System (CPPS), while the EGC-LSTM model applies spatio-temporal correlation to predict subsequent anomalies caused by the lateral movement of APTs. Together, these methods offer a comprehensive solution for correlating APT activities in both spatial and temporal domains, enabling effective prediction of lateral movements and enhancing the overall

defense against persistent cyber threats.

*Semi-Supervised IDS for Digital Substations:* To address zero-day attacks, this thesis introduces a semi-supervised intrusion detection system specifically designed for digital substations. The detection methodology leverages both traffic payload and interarrival time, which is compiled into a vector that represents the behavioral characteristics of OT traffic. To enhance the classification performance between normal and anomalous traffic, the approach incorporates frequency domain characterization of interarrival times using the FFT and the Kolmogorov-Smirnov test. The semi-supervised classifier is implemented through a combination of Self-Organizing Maps (SOM) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), enabling the identification of zero-day attacks while addressing the challenges posed by imbalanced datasets. This integrated framework effectively improves the detection accuracy and robustness of the IDS, providing a comprehensive solution for identifying zero-day attacks in digital substations.





# SAMENVATTING

Energiesystemen ondergaan een digitale transformatie door de integratie van informatie- en communicatietechnologieën, zoals het Internet of Things (IoT), big data en Kunstmatige Intelligentie (AI). Al deze technologieën verbeteren de operationele efficiëntie en intelligentie van energienetwerken. Digitalisering introduceert echter nieuwe kwetsbaarheden in energiesystemen, wat de dringende noodzaak voor versterkte cyberweerbaarheid benadrukt om zo de stabiliteit en veiligheid van energienetwerken tegen opkomende bedreigingen te beschermen. Recente incidenten, waaronder de cyberaanvallen op het Oekraïense elektriciteitsnet in 2015, 2016 en 2022, illustreren de dreiging van cybertegenstanders die gebruikmaken van Advanced Persistent Threats (APTs). In tegenstelling tot traditionele cyberaanvallen vertonen APT's geavanceerdere technieken die worden gekenmerkt door hun onzichtbare tactieken, langdurige aanwezigheid en het uitbuiten van zero-day kwetsbaarheden. Vanwege deze kenmerken van APT's zijn traditionele cyberbeveiligingsmaatregelen ontoereikend om deze uitdagingen te adresseren. Gemotiveerd door deze uitdagingen richt dit proefschrift zich op de detectie en correlatie van APT's in cyber-fysieke energiesystemen.

Dit proefschrift begint met een onderzoek naar cyberbeveiliging in energienetwerken, wat cruciaal is voor het formuleren van effectieve mitigatiestrategieën. Het biedt een diepgaande analyse van het cyberdreigingslandschap, systeemkwetsbaarheden, de meest geavanceerde mitigatietechnieken en de modellering van cyberaanvallen in cyber-fysieke energiesystemen. Op basis van deze fundering introduceert het proefschrift een geavanceerde kill chain voor cyber-fysieke energiesystemen, om de beperkingen van bestaande raamwerken bij het identificeren van fasen van cyberaanvallen te overwinnen. De focus van dit onderzoek is gericht op drie belangrijke uitdagingen van APT's: (1) zichtbaarheid, (2) langdurige aanwezigheid en (3) zero-day aanvallen.

*Zichtbaarheid van APT's:* APT's zijn zeer sterk verborgen en maken gebruik van geavanceerde methoden om detectie te vermijden. Ze maken vaak gebruik van obscure aanvalsmethoden en verschuilen zich achter legitiem netwerkverkeer om traditionele beveiligingsmaatregelen zoals Intrusion Detection Systems (IDS) en firewalls te omzeilen. APT's worden gekenmerkt door minimale afwijkingen van legitiem netwerkverkeer. Dit vormt wetenschappelijke uitdagingen vanwege de noodzaak voor extreem gevoelige anomaliedetectiesystemen.

*Langdurige aanwezigheid van APT's:* APT's zijn specifiek ontworpen om detectie gedurende lange periodes, vaak maanden of zelfs jaren, te vermijden. Om deze uitdaging aan te pakken is het essentieel om anomalieën over langere tijdsperiodes te analyseren en te correleren. De huidige state-of-the-art technieken voor anomaliedetectie in cyber-fysieke energiesystemen worden echter meestal als standalone-oplossingen geformuleerd en richten zich uitsluitend op cyber- of fysieke anomalieën. Deze methoden zijn beperkend, omdat ze zich concentreren op afzonderlijke anomalieën zonder rekening te houden met de interdependenties en correlaties tussen meerdere anomalieën over langere periodes.

Hierdoor zijn ze onvoldoende om de laterale beweging van anomalieën, veroorzaakt door APT's, te detecteren en te correleren. De belangrijkste wetenschappelijke uitdaging op dit punt is het detecteren van weinig voorkomende, tijdgevoelige en moeilijk voorspelbare anomalieën.

*Zero-day aanvallen van APT's:* Tegenstanders voeren vaak zero-day aanvallen uit die misbruik maken van een kwetsbaarheid in communicatieprotocollen, software of hardware die niet bekend is bij de fabrikant of het publiek. Traditionele beveiligingssystemen zoals antivirussoftware, firewalls en IDS vertrouwen sterk op signature-based – met andere woorden: de vingerafdruk of blauwdruk van een anomalie – detectie. Een zero-day aanval is echter een nieuwe cyberaanval zonder bekende vingerafdruk of voorafgaande informatie. Wetenschappelijk gezien vereist de detectie van zero-day aanvallen anomaliegericht detectiemethoden die onafhankelijk zijn van bestaande aanvalsblauwdrukken of bekende dreigingspatronen.

Op basis van bovenstaande bespreking behandelt dit proefschrift de genoemde uitdagingen van APT's door een voorstel voor nieuwe hybride deep learning-modellen, met behulp van op grafen-gebaseerde deep learning en semi-gesuperviseerde leermethoden. De belangrijkste bijdragen van dit proefschrift zijn als volgt samengevat:

*Cyber-Physical Power System Model en Advanced Cyber-Physical Power System Kill Chain:* Deze thesis biedt een uitgebreide analyse van cyberbeveiliging in energiesystemen, met een focus op het steeds veranderende cyberdreigingslandschap, systeemkwetsbaarheden en bestaande mitigatiestrategieën. Het introduceert een cyber-fysiek energiesysteemmodel dat een cyber-range omvat, ontworpen om zowel aanvallen als verdedigingen in een hoog-realistische omgeving te simuleren. Daarnaast wordt een Advanced Cyber-Physical Power System Kill Chain (ACPPS kill chain) raamwerk voorgesteld. De ACPPS kill chain identificeert de kenmerken van Advanced Persistent Threats (APT) die specifiek zijn voor energiesystemen. Het definieert en onderzoekt de cyber-fysieke APT-fasen, variërend van de initiële infiltratie tot cascaderende storingen en een volledige stroomuitval. Deze elementen zijn essentieel voor het definiëren en implementeren van effectieve mitigatiestrategieën tegen APT's op energiesystemen.

*Aanvalsdiaagrammodel:* Om verborgen APT's aan te pakken, stelt deze thesis een aanvalsdiaagrammodel voor dat gebruikmaakt van een Software Defined Network (SDN) voor online situationeel bewustzijn van cyberaanvallen. Het model maakt gebruik van een hybride deep learning-aanpak die Graph Convolutional Long Short-Term Memory (GC-LSTM) en Convolutional Neural Networks (CNN) combineert om netwerkverkeer in Operationele Technologie (OT) te classificeren als abnormaal of normaal. Dit geavanceerde deep learning model kan uiterst kleine verkeersanomalieën detecteren, waardoor het aantal false positives en false negatives aanzienlijk wordt verminderd. Bovendien kan de voorgestelde methode de specifieke locatie van OT-anomalieën bijna real-time identificeren, wat de reactietijd bij het detecteren en mitigeren van APT's aanzienlijk verbetert.

*APT Spatio-Temporal Correlation:* Om de langdurige aanwezigheid van APT's aan te pakken, stelt deze thesis een APT ruimte-tijd correlatiestrategie voor. Deze aanpak maakt gebruik van een Cyber-Physical System Integration Matrix (CPSIM) en een Enhanced Graph Convolutional Long Short-Term Memory (EGC-LSTM) model. De CPSIM-matrix legt een topologische correlatie vast tussen cyber- en fysieke systeemanomalieën in Cyber-Physical Power Systems (CPPS). Het EGC-LSTM-model past ruimtelijke-temporele correlatie toe om

daaropvolgende anomalieën te voorspellen die worden veroorzaakt door de laterale beweging van APT's. Deze methoden samen bieden een uitgebreide oplossing voor het correleren van APT-activiteiten in zowel ruimtelijke als temporele domeinen, waardoor effectieve voorspellingen van laterale bewegingen mogelijk worden en de algehele verdediging tegen aanhoudende cyberdreigingen wordt versterkt.

*Semi-gecontroleerde IDS voor digitale onderstations:* Om zero-day-aanvallen aan te pakken, introduceert deze thesis een semi-supervised intrusion detection system (IDS) dat speciaal is ontworpen voor digitale onderstations. De detectiemethode maakt gebruik van zowel de pakketinhoud als de tijd tussen pakketoverdrachten, die worden verwerkt tot een vector die de gedragskenmerken van OT-verkeer vertegenwoordigt. Om de classificatieprestaties tussen normaal en abnormaal verkeer te verbeteren, wordt een frequentiedomeinkarakterisering van de tussentijden toegepast met behulp van de Fast Fourier Transform (FFT) en de Kolmogorov-Smirnov-test. De semi-supervised classificatie wordt geïmplementeerd met een combinatie van Self-Organizing Maps (SOM) en Density-Based Spatial Clustering of Applications with Noise (DBSCAN), waardoor zero-day-aanvallen kunnen worden geïdentificeerd en de uitdagingen van onevenwichtige datasets worden aangepakt. Dit geïntegreerde raamwerk verbetert de detectienauwkeurigheid en robuustheid van de IDS, en biedt een uitgebreide oplossing voor het identificeren van zero-day-aanvallen in digitale onderstations.



# RINGKASAN

Sistem tenaga listrik sedang mengalami transformasi digital. Integrasi antara teknologi informasi dan komunikasi, seperti *Internet of Things* (IoT), big data, hingga kecerdasan buatan (Artificial Intelligence/AI) meningkatkan efisiensi operasional dan kecerdasan jaringan listrik. Namun, proses digitalisasi ini juga memperkenalkan kerentanan baru di dalam sistem tenaga listrik. Sehingga, diperlukan peningkatan ketahanan siber untuk melindungi stabilitas dan keamanan jaringan listrik dari ancaman yang muncul. Beberapa insiden di dunia nyata, termasuk serangan siber terhadap jaringan listrik di Ukraina pada tahun 2015, 2016, dan 2022, menggambarkan ancaman nyata dari pihak-pihak yang menggunakan strategi *Advanced Persistent Threats* (APTs).

Berbeda dengan serangan siber konvensional, APT menunjukkan teknik yang lebih canggih sesuai dengan karakteristiknya, termasuk taktik tersembunyi, persistensi jangka panjang, dan eksploitasi terhadap *zero-day vulnerabilities*. Karena karakteristik APT tersebut, langkah-langkah keamanan siber tradisional tidak cukup memadai untuk mengatasi tantangan ini. Berangkat dari tantangan ini, tesis ini fokus untuk mendeteksi dan mengorelasikan APT dalam sistem siber-fisik.

Tesis ini dimulai dengan mengeksplorasi keamanan siber pada jaringan listrik, yang berguna untuk merumuskan strategi mitigasi yang efektif. Tesis ini memberikan analisis mendalam tentang lanskap ancaman siber, kerentanan sistem, teknik mitigasi terkini, dan pemodelan serangan siber dalam sistem siber-fisik. Berdasarkan landasan tersebut, tesis ini mengusulkan *advanced cyber-physical power system (ACPPS) kill chain* untuk mengatasi keterbatasan kerangka analisa yang ada di dalam proses untuk mengidentifikasi tahapan serangan siber. Fokus penelitian ini adalah untuk menangani APT pada jaringan listrik dengan menargetkan tiga tantangan utama APT, yaitu (1) tersembunyi, (2) mempunyai persistensi jangka panjang, dan (3) menggunakan serangan *zero-day*.

*Tersembunyi*: APTs sifatnya relatif tersembunyi dan sukar dideteksi. APT kerap menggunakan metode canggih untuk menghindari deteksi. APT sering kali menggunakan proses serangan yang disamarkan dan bersembunyi di balik komunikasi yang sah untuk menghindari sistem keamanan konvensional seperti *Intrusion Detection Systems* (IDS) dan *firewall*. Sifat tersembunyi APT ini ditandai dengan anomali yang sangat kecil dan perubahan yang tampaknya tidak signifikan dibandingkan komunikasi normal. Oleh karena itu, hal ini menimbulkan tantangan ilmiah mengenai kebutuhan terhadap sistem deteksi anomali yang sangat sensitif.

*Mempunyai Persistensi Jangka Panjang*: APT dirancang khusus untuk menghindari deteksi selama periode waktu yang lama, yang dalam hal ini bisa berlangsung berbulan-bulan hingga bertahun-tahun. Untuk menghadapi tantangan ini, penting untuk menganalisis korelasi anomali dalam jangka waktu yang lama. Namun, teknik deteksi anomali terkini untuk sistem siber-fisik sebagian besar dirumuskan sebagai solusi terpisah, hanya menangani anomali pada sistem fisik atau anomali siber secara terpisah. Metodologi tersebut menjadi terbatas, karena hanya berfokus untuk mendeteksi anomali individual tanpa mempertim-

bangkan korelasi dan ketergantungan antar anomali dalam jangka panjang. Akibatnya, mereka tidak cukup mampu mendeteksi dan menghubungkan pergerakan lateral dari anomali yang dipicu oleh APT. Tantangan ilmiah utama terletak pada deteksi anomali yang frekuensinya rendah, waktunya tidak dapat diidentifikasi secara deterministik, dan bersifat tersembunyi, yang sering kali luput dari teknik deteksi konvensional.

*Rentan Terhadap Serangan Zero-Day:* Serangan siber sering kali menggunakan serangan *zero-day* dengan memanfaatkan kerentanan dalam protokol komunikasi, perangkat lunak, atau perangkat keras yang belum diketahui oleh produsennya atau publik secara umum. Sistem keamanan konvensional seperti *antivirus*, *firewall*, dan IDS sangat bergantung pada deteksi berbasis *signature*, di mana pola serangan yang diketahui diidentifikasi menggunakan *signature* yang telah ditentukan sebelumnya. Namun, serangan *zero-day* adalah serangan siber baru tanpa *signature* yang diketahui atau informasi awal apapun. Dari sudut pandang ilmiah, tantangan dalam mendeteksi serangan *zero-day* berasal dari kebutuhan untuk metode deteksi berbasis anomali. Metode ini bekerja tanpa bergantung pada *signature* serangan yang sudah ada atau pola ancaman yang dikenal. Berbeda dengan metode berbasis *signature*, deteksi *zero-day* perlu berfokus untuk mengenali anomali dari karakteristik dan perilaku sistem dasar.

Berdasarkan pembahasan di atas, tesis ini menghadapi tantangan APT tersebut dengan mengusulkan model *deep learning* hibrida baru menggunakan pendekatan *graph-based deep learning* dan *semi-supervised learning*. Secara khusus, kontribusi utama tesis ini dirangkum sebagai berikut:

*Model Sistem Siber-Fisik dan Advanced Cyber-Physical Power System Kill Chain:* Tesis ini memberikan penyelidikan komprehensif terhadap keamanan siber dalam sistem kelistrikan, dengan fokus pada evolusi lanskap ancaman siber, kerentanan sistem, dan strategi mitigasi yang ada. Tesis ini memperkenalkan model sistem siber-fisik yang mencakup *cyber range*, yang dirancang untuk menyimulasikan serangan dan pertahanan dalam lingkungan yang sangat mirip dengan kenyataan. Selain itu, tesis ini mengusulkan kerangka kerja *Advanced Cyber-Physical Power System Kill Chain (ACPPS kill chain)*, yang mengidentifikasi karakteristik APT yang unik pada sistem kelistrikan. Kerangka ini mendefinisikan dan menguji tahapan APT siber-fisik mulai dari fase awal infiltrasi, propagasi gangguan, hingga gangguan total. Semua elemen ini sangat penting dalam mendefinisikan dan menerapkan strategi mitigasi yang efektif terhadap APT pada sistem siber fisik kelistrikan.

*Model Attack Graph:* Untuk menangani APT yang tersembunyi, tesis ini mengusulkan model *attack graph* yang memanfaatkan *Software Defined Network (SDN)* untuk menghadirkan kesadaran situasional serangan siber secara langsung. Model ini menggunakan pendekatan *deep learning* hibrida yang menggabungkan *Graph Convolutional Long Short-Term Memory (GC-LSTM)* dan *Convolutional Neural Networks (CNN)* untuk mengklasifikasikan lalu lintas jaringan *Operational Technology (OT)* sebagai anomali atau normal. Model *deep learning* ini mampu mendeteksi anomali komunikasi yang sangat kecil, sehingga secara signifikan mengurangi *false positive* dan *false negative*. Selain itu, metode ini dapat mengidentifikasi lokasi spesifik anomali OT secara langsung, sehingga memberikan respons yang lebih cepat dalam mendeteksi dan mengatasi APT.

*Korelasi Spatio-Temporal APT:* Untuk menghadapi tantangan persistensi APT dalam jangka panjang, tesis ini mengusulkan strategi korelasi spatio-temporal APT. Pendekatan ini menggunakan *Cyber-Physical System Integration Matrix (CPSIM)* dan model *Enhanced Graph*

*Convolutional Long Short-Term Memory (EGC-LSTM)*. Matriks CPSIM membentuk korelasi topologis antara anomali sistem siber dan fisik dalam *Cyber-Physical Power Systems (CPPS)*, sementara model EGC-LSTM menerapkan korelasi spasial-temporal untuk memprediksi anomali berikutnya akibat pergerakan lateral dari APT. Kedua metode ini memberikan solusi menyeluruh untuk mengorelasikan aktivitas APT dalam domain spasial dan temporal, memungkinkan prediksi pergerakan lateral secara efektif dan meningkatkan pertahanan terhadap ancaman siber yang persisten.

*Semi-Supervised IDS untuk Digital Substation*: Untuk mengatasi serangan *zero-day*, tesis ini memperkenalkan sistem IDS *semi-supervised* yang dirancang khusus untuk *digital substation*. Metodologi deteksi ini memanfaatkan informasi lalu lintas dan waktu antar kedatangan (*interarrival time*), yang dikompilasi ke dalam vektor yang merepresentasikan karakteristik perilaku lalu lintas OT. Untuk meningkatkan performa klasifikasi antara lalu lintas normal dan anomali, pendekatan ini mengintegrasikan karakterisasi domain frekuensi dari waktu antar kedatangan menggunakan *Fast Fourier Transform (FFT)* dan uji *Kolmogorov-Smirnov*. Klasifikasi *semi-supervised* diimplementasikan melalui kombinasi *Self-Organizing Maps (SOM)* dan *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*, memungkinkan identifikasi serangan *zero-day* sambil mengatasi tantangan data tidak seimbang. Kerangka kerja terintegrasi ini secara efektif meningkatkan akurasi dan ketahanan deteksi IDS, memberikan solusi menyeluruh untuk mendeteksi serangan *zero-day* pada *digital substation*.





# LIST OF ACRONYMS

<b>ACPPS</b>	Advanced Cyber-Physical Power System
<b>AD</b>	Active Directory
<b>AI</b>	Artificial Intelligence
<b>AMI</b>	Advanced Metering Infrastructure
<b>APDU</b>	Application Protocol Data Unit
<b>API</b>	Application Programming Interface
<b>APT</b>	Advanced Persistent Threat
<b>AUC</b>	Area Under the Curve
<b>BCU</b>	Bay Control Unit
<b>BMU</b>	Best Matching Unit
<b>C2</b>	Command and Control
<b>CB</b>	Circuit Breaker
<b>CT</b>	Current Transformer
<b>CNN</b>	Convolutional Neural Network
<b>COMTRADE</b>	COMmon format for TRAnsient Data Exchange
<b>CPPS</b>	Cyber-Physical Power System
<b>CPS</b>	Cyber-Physical System
<b>CPSIM</b>	Cyber-Physical System Integration Matrix
<b>CRoF</b>	Control Room of the Future
<b>CSIRT</b>	Computer Security Incident Response Team
<b>CVE</b>	Common Vulnerabilities and Exposures
<b>CyResGrid</b>	Cyber Resilient Grid
<b>DBSCAN</b>	Density-Based Spatial Clustering of Applications with Noise
<b>DDoS</b>	Distributed Denial of Services
<b>DLL</b>	Dynamic Link Library

<b>DNP3</b>	Distributed Network Protocol 3
<b>DNS</b>	Domain Name System
<b>DoS</b>	Denial of Services
<b>DPI</b>	Deep Packet Inspection
<b>DSO</b>	Distribution System Operator
<b>ED</b>	Euclidean Distance
<b>EGC-LSTM</b>	Enhanced Graph Convolutional Long Short-Term Memory
<b>ENTSO-E</b>	European Network of Transmission System Operators for Electricity
<b>FACTS</b>	Flexible Alternating Current Transmission System
<b>FCN</b>	Fully Convolutional Neural Network
<b>FDI</b>	False Data Injection
<b>FFT</b>	Fast Fourier Transform
<b>FGraph</b>	Forensic Graph
<b>GC-LSTM</b>	Graph Convolutional Long Short-Term Memory
<b>GCN</b>	Graph Convolutional Network
<b>GConvGRU</b>	Graph Convolutional Gated Recurrent Unit
<b>Gmean</b>	Geometric Mean
<b>GMM</b>	Gaussian Mixture Model
<b>GNN</b>	Graph Neural Network
<b>GOOSE</b>	Generic Object Oriented Substation Event
<b>GPS</b>	Global Positioning System
<b>GRU</b>	Gated Recurrent Unit
<b>HD</b>	Hamming Distance
<b>HIL</b>	Hardware in the Loop
<b>HMAC</b>	Hash-based Message Authentication Code
<b>HMI</b>	Human Machine Interface
<b>HNSW</b>	Hierarchical Navigable Small World
<b>HTTP</b>	Hypertext Transfer Protocol
<b>HTTPS</b>	Hypertext Transfer Protocol Secure

---

<b>ICCP</b>	Inter-Control Center Communications Protocol
<b>ICMP</b>	Internet Control Message Protocol
<b>ICS</b>	Industrial Control System
<b>IDPS</b>	Intrusion Detection and Prevention System
<b>IDS</b>	Intrusion Detection System
<b>IED</b>	Intelligent Electronic Devices
<b>IoT</b>	Internet of Things
<b>IP</b>	Internet Protocol
<b>IT</b>	Information Technology
<b>KDT</b>	K Decision Tree
<b>KM</b>	K Means
<b>KNN</b>	K-Nearest Neighbors
<b>LRM</b>	Long Range Memory
<b>LSASS</b>	Local Security Authority Subsystem Service
<b>LSH</b>	Locality Sensitive Hashing
<b>LSTM</b>	Long Short-Term Memory
<b>MAC</b>	Message Authentication Code
<b>MITM</b>	Man-in-the-Middle
<b>MLP</b>	Multi-Layer Perceptron
<b>MU</b>	Merging Unit
<b>MMS</b>	Manufacturing Message Specification
<b>MSE</b>	Mean Square Error
<b>NGF</b>	Next-Generation Firewall
<b>NN</b>	Neural Network
<b>OLTC</b>	On-Load Tap Changer
<b>OPC</b>	Open Platform Communication
<b>OPC UA</b>	Open Platform Communication Unified Architecture
<b>OS</b>	Operating System
<b>OSI</b>	Open Systems Interconnection

<b>OSINT</b>	Open Source Intelligence
<b>OT</b>	Operational Technology
<b>PDC</b>	Phasor Data Concentrator
<b>PLC</b>	Programmable Logic Controller
<b>PMU</b>	Phasor Measurement Unit
<b>PNR</b>	Point of No Return
<b>PPDU</b>	Presentation Protocol Data Unit
<b>RMS</b>	Root Mean Square
<b>RNN</b>	Recurrent Neural Network
<b>ROC</b>	Receiver Operating Characteristic
<b>ROCOF</b>	Rate of Change of Frequency
<b>RPC</b>	Remote Procedure Call
<b>RQ</b>	Research Question
<b>RSA</b>	Rivest-Shamir-Adleman
<b>RTDS</b>	Real-Time Digital Simulator
<b>RTU</b>	Remote Terminal Units
<b>SCADA</b>	Supervisory Control And Data Acquisition
<b>SDN</b>	Software Defined Networking
<b>SE</b>	State Estimation
<b>SIEM</b>	Security Information and Event Management
<b>SIS</b>	Safety Instrument System
<b>SMB</b>	Server Message Block
<b>SMTP</b>	Simple Mail Transfer Protocol
<b>SNTP</b>	Simple Network Time Protocol
<b>SOC</b>	Security Operations Center
<b>SOM</b>	Self-Organizing Maps
<b>SPDU</b>	Session Protocol Data Unit
<b>SQL</b>	Structured Query Language
<b>STATCOM</b>	Static Synchronous Compensator

---

<b>SV</b>	Sampled Values
<b>SVC</b>	Static VAR Compensator
<b>TCP</b>	Transmission Control Protocol
<b>TCP/IP</b>	Transmission Control Protocol/Internet Protocol
<b>TDG</b>	Traffic Dispersion Graph
<b>TLS</b>	Transport Layer Security Protocol
<b>TPDU</b>	Transport Protocol Data Unit
<b>TPP</b>	Traffic Pre-Processing
<b>TSC</b>	Time Series Classification
<b>UFLS</b>	Under Frequency Load Shedding
<b>UDP</b>	User Datagram Protocol
<b>UPS</b>	Uninterruptible Power Supply
<b>URL</b>	Uniform Resource Locator
<b>USB</b>	Universal Serial Bus
<b>UVLS</b>	Under Voltage Load Shedding
<b>VM</b>	Virtual Machine
<b>VPN</b>	Virtual Private Network
<b>WAN</b>	Wide Area Network
<b>WAMPAC</b>	Wide Area Monitoring Protection and Control



# 1

## INTRODUCTION

*Power grids are undergoing a digital transformation, integrating advanced technologies to enhance operational efficiency. However, this digitalization introduces vulnerabilities, especially from Advanced Persistent Threats (APTs), which employ stealthy tactics, prolonged persistence, and zero-day attacks. These sophisticated threats surpass the capabilities of traditional cyber security measures. Motivated by these challenges, this thesis focuses on addressing APTs in cyber-physical power systems by overcoming three key challenges, i.e., detecting stealthy anomalies, correlating prolonged anomalies, and identifying zero-day attacks. This thesis introduces a Cyber-Physical Power System Model and an Advanced Cyber-Physical Power System Kill Chain, which map out APT stages from infiltration to system blackout. To detect stealthy APTs, it proposes an attack graph model using hybrid deep learning techniques, enabling real-time anomaly detection with minimal false positives. For prolonged persistence, the study proposes a spatio-temporal correlation strategy with a correlation matrix and deep learning, predicting lateral anomaly movements across systems. Addressing zero-day attacks, a semi-supervised Intrusion Detection System (IDS) is developed, integrating traffic payload analysis, frequency characterization, and unsupervised clustering to detect anomalies. This thesis contributes innovative methodologies to enhance APT detection and correlation in power grids, ensuring their security and resilience in the face of evolving cyber threats.*

## 1.1 BACKGROUND AND MOTIVATION

Power grids are facing a major digital transformation. Advanced technologies are incorporated to enhance the monitoring, control, and intelligence of power grids, such as Operational Technologies (OTs), the Internet of Things (IoT), big data, and Artificial Intelligence (AI). These innovations are essential to build future power systems with enhanced operational efficiency, intelligence, and resilience. However, this digitalization also exposes power grids to notable cyber security challenges. The integration of digital technologies expands the attack surfaces, introducing new vulnerabilities and threats that possibly jeopardize the security and stability of power systems. The imminence of these threats has been demonstrated by recent cyber attacks on power grids. This situation highlights the necessity of enhancing the cyber resilience of power grids to safeguard their digital transformation.

Table 1.1 shows the incidents of cyber attacks impacting cyber-physical power systems. These incidents prove that cyber security incidents related to power grids are already present across the world. On December 23, 2015 cyber attacks were conducted on the power grid in Ukraine that resulted in power outages, which affected 225,000 customers [8]. More sophisticated cyber attacks on the Ukrainian power grid followed on December 17, 2016 resulting in a power outage in the distribution network where 200 MW of load was unsupplied [9]. On March 9, 2020 it was reported that the Information Technology (IT) network of the European Network of Transmission System Operators for Electricity (ENTSO-E) had been compromised in a cyber intrusion [10]. In late 2022, a cyber attack targeting the Ukrainian power grids was reported, with evidence pointing to the involvement of the Sandworm hacker group [11]. The adversaries intend to open the victim's substation circuit breakers, resulting in unplanned power outages. Such advanced cyber attacks conducted by powerful adversaries are a real threat to the security of the modern society. Cyber attacks on power systems can initiate cascading failures and result in a catastrophic blackout, ending up in a doomsday scenario. The power outage can disrupt other critical infrastructures, including water supply, gas networks, telecommunication, transportation, and healthcare services. Without electricity, hospitals and other critical services are severely affected. A disruption of service may lead to financial loss, damages, chaos, or even a loss of lives.

Table 1.1: Cyber Attacks Targeting Cyber-Physical Power Systems

Attack Cases	Year Impacts
European system operator malware [12]	2003 Loss of control in distribution substations for over three days
Aurora experimental cyber attack [13]	2007 Physical damage to power system generator
USB-drive malware in power plant [12]	2012 Three weeks restart delay to power plant
Ukrainian power grid cyber attack 2015 [8]	2015 Power outage affecting 225,000 customers for 6 hours
Ukrainian power grid cyber attack 2016 [9]	2016 200 MW of load was unsupplied
ENTSO-E cyber intrusion [10]	2020 Undisclosed impact
RedEcho malware intrusion [14]	2020 Two hours power outage
ReverseRat malware [15]	2021 Intrusion on power system operator
KA-SAT attack [16]	2022 Disruption on German windfarm satellite communications
Ukrainian power grid cyber attack 2022 [11]	2022 Power outage

Mitigation of cyber attacks on power grids has been extensively studied in recent years. Nonetheless, the majority of the existing research is focused on the identification of cyber



attacks on power grids under False Data Injection (FDI) attack scenarios. These scenarios focus on analyzing power system measurements to identify anomalies in power grids [17–24]. However, in the real-world cyber attacks on power grids reported in [8, 9, 11] are beyond FDI attacks. The real cyber attacks employ Advanced Persistent Threat (APT) tactics that unfold through multiple stages before achieving their final objective of causing a power outage. Compared to conventional cyber attacks, mitigating APTs is substantially more challenging due to their characteristics, including stealthy tactics, prolonged persistence, and exploitative use of zero-day vulnerabilities. Therefore, mitigation strategies for addressing cyber attacks on power grids must systematically take into consideration the characteristics of APTs to deliver the most viable solution.

Driven by the critical aforementioned challenges, this thesis intends to advance the cyber security and cyber resilience of power grids. The underlying philosophy of this thesis lies in understanding the characteristics of power grid OT and APT cyber attacks. By leveraging a comprehensive understanding of both aspects, this thesis proposes APT mitigation strategies by leveraging the distinctive characteristics of power grids' OT communication traffic, which differs from IT traffic. The proposed methods aim to provide power system operators with a near real-time situational awareness, enabling the detection and localization of system-wide anomalies with precision. Furthermore, the proposed methods allow for the early identification of APT-related anomalies, facilitating preemptive mitigation measures before such threats escalate to cause significant operational disruptions or adverse impacts to power systems.

## 1.2 THESIS OBJECTIVE AND RESEARCH QUESTIONS

The primary objective of this thesis is as follows.

Develop APT detection and correlation methods for cyber-physical power systems by considering APT characteristics, including stealthiness, prolonged persistence, and zero-day vulnerabilities.

Based on the primary research objective, this thesis formulates several Research Question (RQ) as follows.

**RQ 1:** *How to unveil the characteristics of cyber attacks on power grids by considering the cyber attack taxonomy, impacts, power grid OT vulnerabilities, and network security controls? How to design a high-fidelity model of the cyber attack on power grids?*

- Review of the existing cyber attacks on power grids and their impacts.
- Formulate cyber attack taxonomy on power grids.
- Identify communication protocol and software vulnerabilities of power grids.
- Identify the state-of-the-art network security controls of power grids.
- Design a cyber-physical power system co-simulation by considering attack and detection capabilities.

## 1

**RQ 2:** *How to identify the stages of APT targeting cyber-physical power systems that incorporate IT and OT attack stages while integrating the complex operational conditions of power systems?*

- Review and analyze the existing methodologies for identifying stages of cyber attacks on power grids.
- Formulate an ACPPS kill chain framework to identify stages of APT targeting cyber-physical power systems that incorporate IT and OT attack stages while integrating the complex operational conditions of power systems.
- Implement and validate the ACPPS kill chain for the cyber attack use cases on power grids.

**RQ 3:** *How to detect stealthy APTs on power grids with minimum anomalies and insignificant changes compared to legitimate traffic?*

- Review of the state-of-the-art methods for cyber attack detection in power systems.
- Develop Software Defined Networking (SDN)-based traffic monitoring for power grid OT networks in digital substations, wide area networks, and control centers.
- Design a deep learning model to learn from spatio-temporal characteristics of OT traffic throughput using Graph Convolutional Long Short-Term Memory (GC-LSTM).
- Design a throughput time series classifier using hybrid deep learning of GC-LSTM and Convolutional Neural Network (CNN) for detecting infinitesimal anomalies and minimizing false positive rates.
- Design an attack graph model for pinpoint anomaly locations in near real-time for raising operator situational awareness.
- Validate the effectiveness of the proposed method and perform benchmarking with the state-of-the-art time series classifier deep learning models.

**RQ 4:** *How can cyber and physical anomalies caused by APTs in cyber-physical power systems be effectively detected and correlated, specifically in the context of prolonged attacks with non-deterministic temporal anomaly instances?*

- Review of the state-of-the-art methods for cyber-physical and spatio-temporal correlation.
- Formulate a cyber-physical system integration matrix (CPSIM) to implement topological correlation of cyber and physical system anomalies in CPPS.
- Design enhanced GC-LSTM with sequential and neural network filters to predict anomalies lateral movement resulting from APT attacks.
- Design a resilient associative method based on vector databases and K-Nearest Neighbors (KNN) to identify unknown attack propagation patterns.

- Formulate a CPPS log comparator to verify and differentiate between physical power system anomalies caused by cyber attacks and physical power system disturbances.
- Validate the effectiveness of the proposed methods and perform benchmarking with the state-of-the-art deep learning spatio-temporal models and vector search algorithms.

**RQ 5:** *How to detect APT's zero-day attacks in power systems considering behavioral characteristics of OT communication traffic and limited preliminary knowledge about the attacks?*

- Perform experiments and collect data from a cyber attack on a digital substation using Hardware in the Loop (HIL) setup.
- Analyze the characteristics of OT traffic in digital substations.
- Review of the state-of-the-art methods for semi-supervised and unsupervised anomaly detection methods.
- Design a frequency domain interarrival time traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov for digital substation traffic for OT traffic.
- Formulate a novel traffic distance similarity vector based on OT traffic quantitative parameters from traffic payload and interarrival time.
- Design a novel hybrid semi-supervised classification model based on a Self-Organizing Map (SOM) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) for zero-day attack detection.
- Validate the effectiveness of the proposed methods and perform benchmarking with the state-of-the-art unsupervised clustering algorithms.

## 1.3 THESIS OUTLINES AND CONTRIBUTIONS

The thesis outline is depicted in Fig. 1.1. Chapter 2 presents the cyber security of the power system, and Chapter 3 presents the ACPPS kill chain. Chapters 4, 5, and 6 address the challenges of APTs in power grids based on the APT's characteristics. Chapter 4 addresses stealthy attack detection using the attack graph model. Chapter 5 addresses APTs' long-term correlation using spatio-temporal APTs detection and correlation. Chapter 6 addresses zero-day attacks using semi-supervised IDS. Chapters 4 and 5 present wide-area system anomaly detection, and chapter 6 presents localized anomaly detection in digital substations. Finally, chapter 7 presents the conclusions and potential future research directions.

The contributions associated with every chapter and RQ of this thesis are summarized as follows:

### 1. Chapter 2: Cyber Security of Power System (RQ1)

This chapter addresses the 1st challenge of understanding the characteristics of cyber security on power grids. This chapter provides a comprehensive overview of cyber security in power systems, focusing on the challenges and solutions in

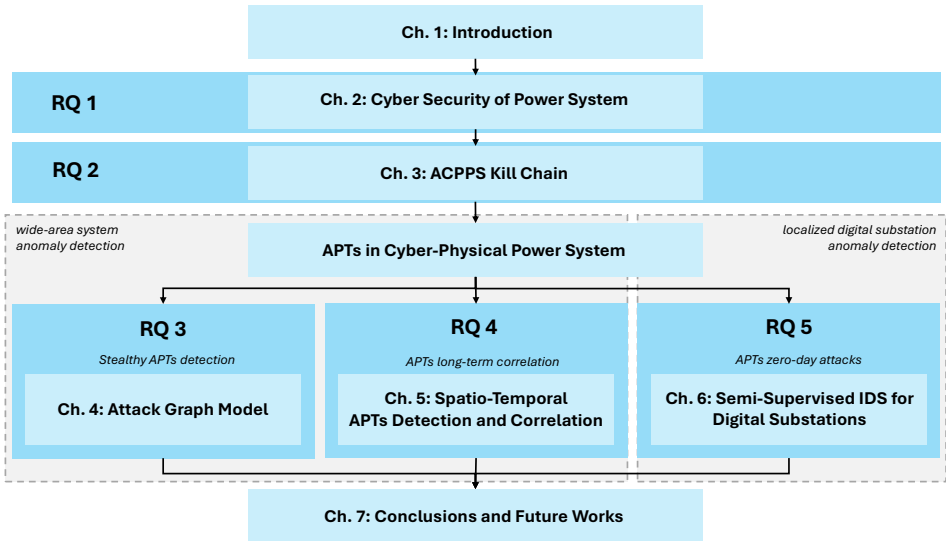


Figure 1.1: Thesis outline.

protecting modern power grids. It begins by exploring the digitalization of power grids, emphasizing the increasing reliance on interconnected OT and IT, which has expanded the attack surface for cyber threats. Subsequently, this chapter discussed a taxonomy of cyber attacks which categorizes the types of attacks that target power grids, followed by a discussion of their impact on system stability and reliability. This chapter then delves into the vulnerabilities of power grids' OT, with specific attention to communication protocols and software application vulnerabilities. To address these risks, the chapter discusses the role of secure communication protocols designed to protect OT in power systems and highlights state-of-the-art network security control measures that enhance the defense against cyber threats. Finally, it presents a cyber-physical model for simulating cyber attacks and testing the resilience of power grids in a controlled environment, offering a critical foundation for developing robust cyber security strategies.

The cyber-physical power system model is composed of a simulation of the power system as well as an IT/OT simulation. DlgSILENT PowerFactory is used for the simulation of the power system, i.e., IEEE 39-bus. The power system model provides circuit breaker status and measurement data of active and reactive powers, voltages, and currents from busbars, lines, and generators. The implementation of Open Platform Communication Unified Architecture (OPC UA) facilitates the interfacing of data exchange between power grids and IT/OT simulation. The implementation of the IT/OT architecture is carried out through the application of Mininet. Each host in the IT/OT network, e.g., merging units, intelligent electronic devices, network switches, routers, databases, etc., are implemented in Mininet using containers. Every container incorporates a tailored application for IT/OT host operations, such as the acquisition and transmission of measurement data, control setpoints, database access,

and so forth. The current implementation of Cyber-Physical System (CPS) comprises 27 substations and 210 hosts. A unique application has been tailored for each host to replicate the CPS of power grid components. At present, the simulation of all 27 substations runs on 50,000 lines of code on 26 Virtual Machine (VM).

*Chapter 2 Contributions:*

- (1) A comprehensive overview of cyber security in power systems, focusing on the cyber threat landscape, vulnerabilities, and existing mitigation.
- (2) A cyber-physical power system model for simulating cyber attacks and detection in a high-fidelity cyber range environment as part of TU Delft Control Room of the Future (CRoF) technology center.

## 2. Chapter 3: Advanced Cyber-Physical Power System Kill Chain (RQ 2)

The cyber attacks on the Ukrainian power grid demonstrate the APT's real impact on power systems. However, existing research has not yet covered a thorough investigation of APT stages on CPPS and their consequences on power system operation. Several frameworks exist to analyze APT stages in IT systems. Currently, the analysis of cyber attacks on power grids is primarily performed using the cyber kill chain [25], CPS kill chain [26], MITRE ATT&CK ICS [27], and SANS ICS [28]. These frameworks are heavily focused on the cyber stages of the attacks and briefly cover their impact. However, they don't cover the impact of cyber attacks on the operation of the physical power system. According to our literature review, there is no framework that provides a comprehensive analysis of APT stages in CPPS, including the integrated IT/OT communication networks and impact on power grid operation, affecting the system stability and causing cascading failures and a blackout. Therefore, this chapter proposes an in-depth analysis of the capabilities of APTs on CPPS, considering the integration of the IT/OT system and its impact on power system operation. Subsequently, this chapter defines the characteristics of APTs on CPPS and proposes the novel Advanced Cyber-Physical Power System (ACPPS) kill chain framework. The ACPPS kill chain identifies the APT characteristics that are unique to power systems. It defines and examines the cyber-physical APT stages spanning from the initial phases of infiltration to cascading failures and a power system blackout. The proposed ACPPS kill chain is validated with real-world APT attacks on the power grid in Ukraine in 2015 and 2016, and cyber-physical simulations.

*Chapter 3 Contribution:*

Advanced Cyber-Physical Power System Kill Chain framework for identification of APT stages on power grids.

## 3. Chapter 4: Attack Graph Model For Cyber-Physical Power System using Hybrid Deep Learning (RQ 3)

This chapter proposes the first known Software Defined Network-based online cyber

attack situational awareness method, i.e., Cyber Resilient Grid (CyResGrid). It is specifically designed for anomaly detection using communication traffic throughput in OT networks for stealthy cyber attacks during the early stages of the cyber kill chain, e.g., network reconnaissance. Therefore, CyResGrid aids operators to locate and identify power system-wide cyber attacks in near real-time through an attack graph map. The CyResGrid is integrated with a hybrid deep learning model to classify the OT network traffic throughput as anomalous or normal. The model combines GC-LSTM and a deep convolutional network to detect OT network anomalies caused by cyber attacks. It outperforms existing state-of-the-art deep learning-based time series classifiers [29, 30], as indicated by Geometric mean and F1 scores. To achieve this, the proposed method uses GC-LSTM for traffic normalization. Subsequently, to detect the anomaly, it uses an optimized CNN hyperparameters through Bayesian optimization. Based on the network throughput monitoring and anomaly detection, the proposed method creates an attack graph map of power system-wide cyber attacks, in near real-time.

*Chapter 4 Contributions:*

- (1) Cyber Resilient Grid (CyResGrid) is an SDN-based online cyber attack situational awareness method.
- (2) A hybrid deep learning of GC-LSTM and CNN to classify the OT network traffic throughput as anomalous or normal.

#### 4. Chapter 5: Spatio-temporal Advanced Persistent Threat Detection and Correlation For Cyber-Physical Power Systems Using Enhanced GC-LSTM (RQ 4)

This chapter proposes a novel spatio-temporal APT detection, correlation, and prediction in cyber-physical power systems. It allows power system operators to locate system-wide anomalies in near real-time from control centers and mitigate APTs early before they cause adverse impacts. At substations and control centers, distributed semi-supervised Deep Packet Inspection (DPI) classifiers monitor OT communication traffic using SDN-enabled switch. The traffic observation points communicate with the SDN controller at the control center to construct a cyber anomaly graph. This is generated based on the DPI classification results using a Traffic Dispersion Graph (TDG) with SDN [5]. The power system graph is constructed based on the energized power lines in accordance with the status of Circuit Breaker (CB) [31, 32]. The cyber-physical anomaly graph is input into a Cyber-Physical System Integration Matrix (CPSIM) for spatio-temporal correlation. Subsequently, an EGC-LSTM model with sequential and neural network filters is used to predict APTs in CPPS. Furthermore, to identify zero-day APT patterns, this chapter proposes a resilient associative method based on vector databases and KNN. The method employs a CPPS log comparator function to verify and differentiate between circuit breakers opened by operators, faults, and cyber attacks.

*Chapter 5 Contributions:*

- (1) A novel semi-supervised deep packet inspection method for OT communication network traffic utilizing the OT homogeneous characteristics based on CNN and Hamming Distance vector.
- (2) A CPSIM that constructs a topological correlation of cyber and physical system anomalies in CPPS.
- (3) Enhanced Graph Convolutional Long Short-Term Memory (EGC-LSTM) method with sequential and neural network filters to predict subsequent anomalies resulting from APT attacks.

## 5. Chapter 6: Intrusion Detection System for Digital Substations using Semi-Supervised Learning and Traffic Distance Similarity Clustering (RQ 5)

This chapter proposes a novel frequency domain interarrival time traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov. This method enhanced statistical-based methods that are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. Compared to statistical-based interarrival time, the combination of Fast Fourier Transform and Kolmogorov-Smirnov is able to improve the accuracy by 26% and F1 score by 41%. This thesis also proposes a novel traffic distance similarity vector of operational technology communication traffic. The vector is derived from the packet payload and interarrival time. The vector quantified the packet payload based on the convolutional neural network and Chebyshev distance and quantified the packet interarrival time using Fast Fourier Transform and Kolmogorov-Smirnov. Finally, to identify the anomalies in digital substations, this thesis proposes a novel hybrid semi-supervised model based on the self-organizing map and DBSCAN. The hybrid combination of them aims to improve classification performance and address the imbalanced dataset. Results indicate that the proposed method can identify zero-day attacks and achieve accuracy and F1 above 95%.

*Chapter 6 Contributions:*

- (1) A frequency domain interarrival time traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov for improving performance classification of normal and anomalous traffic.
- (2) A traffic distance similarity vector of OT communication traffic derived from the packet payload and interarrival time.
- (3) A hybrid semi-supervised classification model based on self-organizing map and DBSCAN to identify zero-day attacks and improve the classification performance of imbalanced datasets.





## 2

## CYBER SECURITY OF POWER SYSTEMS

*The digitalization of power grids, driven by ITs and OTs, has enhanced monitoring, control, and intelligence, improving sustainability, affordability, and resilience. However, it also introduces cyber security vulnerabilities and threats. This chapter explores the complex landscape of cyber security in power systems, presenting a taxonomy of attacks with their impacts on system stability and reliability. The chapter underscores the severe impacts of cyber attacks, ranging from equipment damage and load loss to cascading failures that could result in catastrophic blackouts. To address these challenges, a CPPS co-simulation model is introduced, enabling the simulation of cyber attacks and testing of mitigation strategies in a controlled environment. The model also supports cyber range applications, allowing for the evaluation of defense mechanisms and incident response strategies. This chapter concludes by emphasizing the urgent need for robust, adaptive defense frameworks and comprehensive mitigation techniques to ensure the cyber resilience of power grids in an increasingly digitalized and interconnected power systems.*

---

This chapter is partly based on the book chapters [1] A. Presekal et al. "Cyber Attacks on Power Systems [1]," and [2] A. Presekal et al. "Anomaly Detection and Mitigation in Cyber-Physical Power Systems based on Hybrid Deep Learning and Attack Graphs [3]," in *Cyber-Physical Power Systems: Challenges and Solutions by AI/ML, Big Data, Blockchain, IoT, and Information Theory Paradigms*, IEEE-Wiley Press, Feb. 2025.

## 2.1 POWER GRID DIGITALIZATION

Power grids are undergoing a fast-paced process of digitalization for enhanced monitoring and control capabilities and grid intelligence. Infrastructure and participants are enhanced and supported by ITs. Further integration of digital technologies is vital for the development of the future power grid, e.g., next generation OTs, IoT, digital substations, artificial intelligence, and big data analytics. All this is expected to increase sustainability, affordability, and resilience of the power system. The latter, however, is also challenged by all these new elements. Opening up the energy system to everyone by means of ITs requires careful considerations with regard to data privacy and information security in general. This combined with the trend towards distributed renewable generation, electrification of virtually all aspects of our lives, and easy market participation for all energy system participants, the cyber security and resilience requirements of the power grid become even more critical. The increased digitalization raises questions, especially with regard to vulnerabilities, threats, and cyber secure operation of the power system. It is well recognized that IT/OT systems are vulnerable to cyber attacks. Furthermore, the combination of heterogeneous, co-existing smart and legacy technologies generates significant vulnerabilities and security challenges. With respect to security of supply and reliability of the future energy system provision, special attention is needed for new vulnerabilities and threats that come with digitalization. Accordingly, cyber resilience aspects are critical for a further power grid digitalization.

Examples of cyber security incidents related to power grids already exist around the world. On December 23, 2015 cyber attacks were conducted on the power grid in Ukraine that resulted in power outages, which affected 225,000 customers. More sophisticated cyber attacks on the Ukrainian power grid followed on December 17, 2016 resulting in a power outage in the distribution network where 200 MW of load was unsupplied. On March 9, 2020 it was reported that the IT network of the ENTSO-E had been compromised in a cyber intrusion. Fortunately, the compromised IT network was not connected to any operational electric transmission system. However, this indicates that interconnected power grids may become targets. Such laborious cyber attacks conducted by powerful adversaries are a real threat to the security of the modern society. Cyber attacks on power systems can initiate cascading failures and result in a catastrophic blackout, ending up in a doomsday scenario especially if it is considered that the world may experience a global crisis such as a pandemic. The power outage can disrupt the entire energy chain including water supply, heating, and gas networks. Without power, hospitals and other critical services are severely affected. A disruption of service may lead to financial loss, damages, chaos, or even a loss of lives.

This chapter presents an investigation for understanding the characteristics of power grid cyber security. This chapter discusses a taxonomy of cyber attacks which categorizes the types of attacks that target power grids, followed by a discussion of their impact on system stability and reliability. This chapter then delves into the vulnerabilities of power grids' OT, with specific attention to communication protocols and software application vulnerabilities. To address these risks, the chapter discusses the role of secure communication protocols designed to protect OT communication networks in power systems and highlights state-of-the-art network security control measures that enhance the defense against cyber threats. Finally, it presents a cyber-physical co-simulation model for simulat-

ing attacks and testing the resilience of power grids in a controlled environment, offering a critical foundation for developing robust cyber security strategies.

## 2.2 TAXONOMY OF CYBER ATTACKS ON POWER GRIDS

There are many attack techniques that can potentially be deployed to specifically target power grids. In this section, we classify such types of attacks into six categories, i.e., phishing, malware, network-based attacks, man-in-the-middle, host-based attacks, and denial of service. Fig. 2.1 shows the taxonomy of cyber attacks on power systems and Industrial Control System (ICS). We delve in depth into each cyber attack category targeting industrial control systems and power grids in the following subsections.

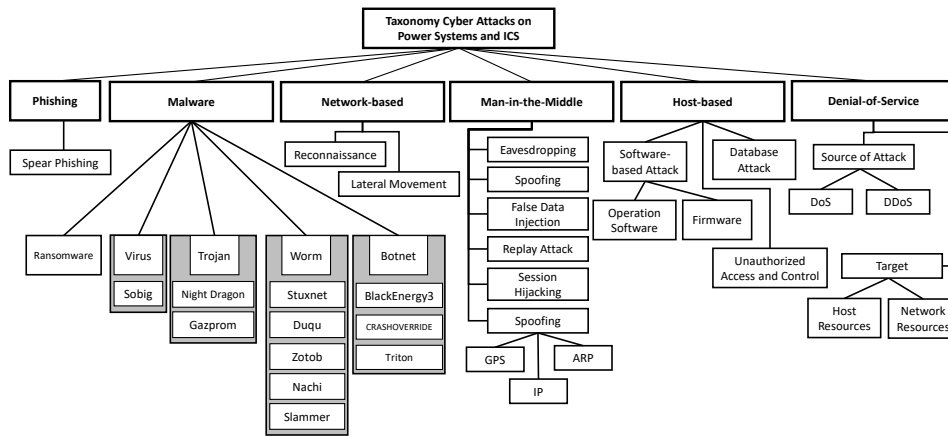


Figure 2.1: Taxonomy of cyber attacks on power systems and ICS [1].

### 2.2.1 PHISHING

Phishing is a type of social engineering attack exploiting human factors. Phishing attacks aim to obtain sensitive information or data and deliver malware inside the corporate IT network, thereby often serving as entry points in the delivery mechanism of sophisticated cyber attacks. In phishing, attackers pretend to be a trustworthy source, persuading the victims to carry out certain actions. These could be opening an email, accessing Uniform Resource Locator (URL) links, downloading files, and unknowingly providing sensitive information. In this section, our discussion is focused on email phishing attacks. Emails represent an important individual identity on The Internet and also carry organizational significance as formal communication channels. Due to these reasons, emails have become the most dominant phishing media. Attackers can arbitrarily send phishing emails to any email address, aiming to persuade recipients to carry out further actions, as described above. These actions may be followed by malicious outcomes such as information breaches, malware installation, and financial losses.

Spear phishing is a variant of phishing, which targets a specific individual or organization. Before launching a spear phishing attack, an attacker gathers detailed information

about the target. Information gathering can be done via The Internet or social engineering. Based on preliminary information, the attacker crafts the phishing email content to be relevant and strongly related to the target, prompting a higher success rate compared to arbitrary phishing. In general, state-sponsored attackers or cyber criminals are the ones behind spear phishing attacks. Consequently, spear phishing serves as an important entry point for advanced persistent threats. For example, in Ukraine 2015, the attackers pretended to be from the Ukrainian Ministry of Energy, which was strongly related to the targeted distribution system operators. Hence, the spear phishing campaign was successful in deceiving the Distribution System Operator (DSO) employees to access the malicious email attachments, resulting in the installation of malware.

Protection against phishing attacks can be achieved by strong organizational policies and corporate IT security. This can include verifying and screening email address sources, URL links, and file attachments in emails. Through such measures, malicious emails can then be flagged and filtered, serving as the initial line of defense against phishing attacks. However, the most critical defense method is raising user awareness. With proper awareness and training, users can recognize phishing emails.

### 2.2.2 MALWARE

Malware is a piece of malicious code that exploits software vulnerabilities in a targeted system. The occurrence of malware-related attacks has been on the rise, especially over the last decade [33]. There are various types of malware, e.g., ransomware, virus, Trojan, worm, and botnet, examples of which are presented in Fig. 2.1. Advanced types of malware such as worms and botnets, owing to their intelligence and controllability, are gaining popularity as modern cyber attack vectors. Furthermore, given their ability to propagate and spread independently, they have become the cyber weapons of choice targeting industrial control systems. Therefore, in this section, we focus on the most well-known examples of worms and botnets involved in cyber attacks on ICS and power grids, i.e., Stuxnet, BlackEnergy, CRASHOVERRIDE, and Triton. Table 2.1 shows an overview of the comparison of the capabilities between most commonly found types of malware in ICS-related cyber incidents.

Table 2.1: Comparison of malware capabilities

No.	Capabilities	Viruses	Trojans	Worms	Botnets
1	Self-replication	✓		✓	
2	Backdoor		✓		✓
3	Remote control		✓		✓
4	Network spread			✓	✓

Table 2.2 shows a brief history of malware involved in ICS cyber-related incidents from 1982 to 2017. It can be seen that worms were popular in the early 2000s due to their capabilities to spread across the network. However, an attacker cannot fully control a worm, leading to an uncontrollable spread. This may result in loss of stealth and unexpected exposure and discovery. For example, the Stuxnet worm was originally created only to target Iranian nuclear reactors. Unfortunately, it ended up spreading across the world [34], leading to its discovery in 2010. As a result, botnets have become the malware of choice

in recent years. A botnet can be controlled remotely and acts as a vector for advanced persistent threats [35, 36].

Table 2.2: History of malware involved in major ICS cyber-related incidents

No.	Target	Type of Malware (Name)	Year
1	SCADA gas pipeline, Siberia [3]	Trojan horse	1982
2	SCADA gas pipeline, Russia [3]	Trojan horse (Gazprom)	1999
3	Web service RTU, PLC [37]	Worm	2002–2003
4	SCADA database [37]	Worm (Slammer)	2003
5	SCADA petrochemical plant [37]	Worm (Nachi)	2003
6	Train signaling system, USA [3]	Virus (Sobig)	2003
7	SCADA automotive manufacturing plants [38]	Worm (Zotob)	2005
8	Energy companies [39]	Trojan horse (Night Dragon)	2009
9	SCADA nuclear centrifuge [40]	Worm (Stuxnet)	2010
10	SCADA reconnaissance [3]	Worm (Duqu)	2011
11	SCADA steel mill, Germany [41]	Botnet	2014
12	SCADA power grid, Ukraine [42]	Botnet (BlackEnergy3)	2015
13	SCADA power grid, Ukraine [43]	Botnet (CRASHOVERRIDE)	2016
14	SCADA petrochemical plant, Saudi Arabia [44]	Botnet (Triton)	2017

## STUXNET

Stuxnet is a worm discovered in June 2010, designed specifically to target Programmable Logic Controller (PLC) in industrial control systems. It exploited unprecedented zero-day vulnerabilities present in the Microsoft Windows operating system, Siemens STEP7 PLC software, and Remote Procedure Call (RPC) server mechanism. Stuxnet also employed the rootkit technique to hide from commercial antivirus software. It was later identified as a state-sponsored and developed cyber weapon to target the uranium enrichment facilities in Iran. Detailed technical studies about Stuxnet are presented in [34, 45].

Stuxnet spread through three main mechanisms. The initial entry point was via Universal Serial Bus (USB) flash drives in computers located within the Supervisory Control And Data Acquisition (SCADA) system. When a USB drive is plugged into a computer, the Windows operating system will execute autorun.inf or Windows .lnk as an autorun mechanism. Exploiting this vulnerability, Stuxnet copied itself onto the computer's hard drive. The second spreading mechanism was through Windows network shares, wherein Stuxnet had the capability to duplicate itself into a shared folder on the same network. Through this mechanism, Stuxnet exploited the Server Message Block (SMB), popularly known as Samba, which is a protocol for file and folder sharing. The third spreading mechanism exploited vulnerabilities present in Siemens' software, i.e., Siemens WinCC and Siemens STEP7. Stuxnet identified and compromised access to computers that ran Siemens WinCC. Furthermore, it also replicated itself into WinCC computers via an Structured Query Language (SQL) injection command. Siemens STEP7 is an application used to program Siemens PLCs. Hence, through STEP7, Stuxnet spread into PLC devices as well. If a computer was running Siemens STEP7, Stuxnet modified the Windows Dynamic Link Library (DLL) and associated executable files. With this infection, Stuxnet added malicious code into the PLC devices, allowing complete control over them. In the reported incidents, it caused centrifuges to spin abnormally, while blindsiding operators.

Taking a closer look at its structure, Stuxnet mainly consists of two modules, i.e., user-mode and kernel-mode. The user-mode module has four functions: (i) searching function for specific targets, (ii) privilege escalation, (iii) malicious code injection into PLC, and (iv) installation kernel-mode. It is interesting to note, the malware also had a time limit date set to June 24th, 2012. Hence, it was functional only until that specific date. The kernel-mode made Stuxnet work on the lower system levels, well below the application level of the user-mode. By implementing the kernel-mode, Stuxnet was maliciously launched during every Windows bootup process. This system-level execution made it more persistent and impervious to antivirus protection. The main functionality of Stuxnet was defined in its code. However, it was also designed to remotely communicate to the command and control server. This mechanism allowed adversaries to remotely update the malware. However, this remote update was never performed, in order to limit the uncontrollable spread of Stuxnet outside of its designated target.

### **BLACKENERGY**

BlackEnergy is a malware, initially identified in 2007 as an Hypertext Transfer Protocol (HTTP)-based botnet for Distributed Denial of Services (DDoS) attacks. It identified and targeted multiple file extensions, including Microsoft Office, Java, and executable files. More specifically, BlackEnergy injected malicious files into the Windows System32 folder on a target Windows machine. However, its most infamous use was as a weaponized malware during the cyber attack on the power grid in Ukraine, in December 2015.

In this attack, BlackEnergy3 exploited the SMB protocol to propagate across the IT/OT networks. SMB was used as the attack vector due to its capabilities to bypass typical firewalls. Through the infected SMB, BlackEnergy3 replicated across hosts in the IT network, delivering malicious payloads. In addition, it was accompanied by an RPC to serve as a backdoor module. This was done to establish a connection between the infected hosts and the attacker's Command and Control (C2) server. RPC allowed the malware to receive commands from the remote C2 server, and relay critical information back. Such advanced remote capabilities allowed adversaries to perform early reconnaissance using its many plugins during the attack. These plugins were programmed with many functionalities, such as executing file operations, i.e., enumerate, execute, download, and overwrite, stealing credentials, discovering networks, and self-destructing. The self-destruction function of BlackEnergy3 was executed via KillDisk. Furthermore, BlackEnergy3 also contained information regarding the current time and location. This allowed it to run malicious activities during non-peak hours, e.g., at midnight. Overall, with all these capabilities, BlackEnergy3 has been proven to be a vicious tool for cyber attacks on power grids. Further detailed investigation of BlackEnergy3 is presented in [46].

### **CRASHOVERRIDE**

CRASHOVERRIDE or Industroyer was the root cause of the cyber attack on the power grid in Ukraine in December 2016. Its many features such as backdoors, intrusion, and reconnaissance strategies are quite similar to BlackEnergy3. However, according to its code level investigations, there is no strong connection between the two malware. It is very likely that CRASHOVERRIDE was a new type of malware, specifically created for the Ukraine 2016 attack [47]. It mainly comprised of three components, i.e., backdoor, payload, and launcher.

The backdoor of CRASHOVERRIDE can be further subclassified into the main backdoor and the additional backdoor. The main backdoor served as the main controller, connecting to a remote C2 server via Hypertext Transfer Protocol Secure (HTTPS). Remote commands were encapsulated as HTTPS traffic, while source and destination addresses for establishing connectivity to C2 server were hardcoded. This clearly shows that CRASHOVERRIDE was created on purpose to specifically target the Ukrainian power grid. Using the backdoor, attackers could define a specific time to activate the malware allowing them to perform a multitude of actions. This includes remote control for process execution, switch execution into a specific user account, download file from C2 server, copy files, start and stop services, change registry values, and execute shell commands. The additional backdoor was set up as contingency, in case of a failure of the main backdoor. It was deployed in the form of a Trojan file disguised as a Notepad executable file. This additional backdoor had a different configuration and connected to a different C2 server.

The payload component described the payload for specific protocols such as IEC 101, IEC 104, IEC 61850, and Open Platform Communication (OPC). Hence, in order to craft malicious packets, an adversary must possess proper knowledge about power grid communication and automation standards. These payloads are typically stored using the DLL file extension on the Windows operating system. Exploiting such mechanisms, adversaries saved critical operational data related to network configurations such as Internet Protocol (IP) addresses and running protocols inside .ini extension files.

In order to carry out the attack, adversaries executed the launcher module that triggered the malicious payload execution and packet delivery into targeted substations, based on predefined protocols. These malicious packets were expected to open circuit breakers and cause a blackout. Finally, to remove all traces of the attack, the launcher module also executed the data wiper function. Data wipers changed the registry value on the Windows operating system to make it unbootable and also deleted files on the infected computers. Subsequently, the data wiper triggered process termination causing the operating system to crash. Besides these three main components, CRASHOVERRIDE possessed additional capabilities such as port scanner and Denial of Services (DoS). The port scanner identified open ports on the target's IP address. On the other hand, the DoS tool sent malicious packets to Siemens SIPROTEC devices to make them unresponsive. Such capabilities were expected to increase the severity of the cyber attack. However, the attack did not work as expected. One of the probable reasons was the hardcoded nature of CRASHOVERRIDE that made it less flexible.

### **TRITON**

Triton is a botnet that targeted the Safety Instrument System (SIS) from Schneider Electric at a Saudi Arabian petrochemical processing plant in 2017 [44]. SIS is an automated mechanism in ICS to prevent operational failures and protect from hazards such as fires and explosions. Hence, Triton's objective was to disrupt SIS functioning to allow the potential occurrence of catastrophic incidents. This malware was an advanced persistent threat as almost all of its operations were carried out in a stealthy manner. Investigations shows that Triton is arguably the stealthiest malware targeting ICS to date. Hence, it is very fortunate that it was uncovered. The Triton attack was exposed because the adversaries made a mistake in triggering the safety system mechanism, thereby shutting down the



entire ICS. Otherwise, it is estimated that Triton would have probably remained undetected with potentially catastrophic consequences.

Triton employed social engineering techniques to gain access to the ICS network. The plant operators received or downloaded a `trilog.exe` file. This file pretended to be a legitimate Schneider Triconex SIS application. The executable file served as a vector to initiate the cyber attack. The malicious file then injected the Triton payload into the memory of the Triconex SIS controller. In addition, it also injected two files `inject.bin` and `imain.bin` to the SIS devices. `Inject.bin` contained payload data for the attack, and `imain.bin` became a backdoor for allowing remote execution. To carry out such a major attack, adversaries had good knowledge of how the SIS system worked. By exploiting the payload of the SIS protocols and executing remote control, attackers conducted a coordinated attack on the SIS protection systems. There were three possible attacks performed by Triton. The first was to shutdown the SIS process itself. The second and third were to reprogram and persistently maintain the SIS in an unsafe state [48]. It is worth mentioning that there is no detailed technical information available regarding Triton's attack process. Nonetheless, Pinto et al. investigated the technical mechanism of the Triton attack process using a replicated attack environment [44] where some of the attack stages were simulated based on assumptions.

### 2.2.3 NETWORK-BASED ATTACKS

The communication network is the backbone for data exchange between connected hosts in IT/OT systems. Thereby, a successful cyber attack can potentially target multiple aspects of the communication network, including physical connections, device information, and protocols in use. Active network reconnaissance and lateral movement are the two most common network-based attacks.

#### NETWORK RECONNAISSANCE

Network reconnaissance is the process of discovering information related to the computer network such as connected hosts, network topology, protocols, applications, and services running on the IT/OT network. It can also be used to discover vulnerabilities. Attackers typically employ the Internet Control Message Protocol (ICMP) to identify active hosts connected to the communication network. Based on a prediction of the range of active IP addresses in the network, an attacker launches a ping ICMP scan using tools such as `tcpdump` or `nmap`. The scan provides a list of active host IP addresses that responded to the ping message. This can be mitigated by filtering ICMP packets and discarding them. Another variant of the attack involves attackers using a Transmission Control Protocol (TCP) scan by launching TCP sync packets to the list of IP addresses. Active hosts respond with an acknowledge packet. However, for this attack one also needs to consider the number of active ports, which is time and resource intensive. Hence, TCP scanning is more challenging. In any case, as a result of network reconnaissance, attackers can obtain a list of active host IP addresses connected to the network.

From the list of active hosts, attackers can obtain further details such as the running services by scanning for active ports. For example, if port 80 is open, it may imply that the host is running an HTTP service/webserver. If port 25 is open, then the host is probably running a Simple Mail Transfer Protocol (SMTP) mail server. Taking this further, attackers



can also find a detailed version of the running applications through host fingerprinting and obtain potential vulnerabilities. Therefore, in a cyber attack scenario on power grids, network reconnaissance plays an important role in discovering the target hosts and protocols on the IT/OT systems. For example, in Ukraine 2015, attackers successfully identified vulnerabilities in the Microsoft active directory server, paving the way for subsequent access to the OT systems in the control center through login credential theft. Similarly, in Ukraine 2016, attackers successfully detected active protocols such as IEC 101, IEC 104, and IEC 61850 used for communication within substations.

### **LATERAL MOVEMENT**

Lateral movement is the attack process of progressively propagating throughout the targeted communication network. It starts from the most vulnerable host, serving as the entry point, moving through multiple hosts to reach the final OT target. This attack typically exploits user login credentials to access various hosts and move laterally within the IT/OT network. This technique is a common mechanism, often found in advanced persistent threats. Consequently, lateral movement was heavily employed in the cyber attacks targeting the power grid in Ukraine in 2015 and 2016. In 2015, the attackers' final objective through lateral movement was to access the SCADA system in the control center and cause a blackout. IT/OT network segmentation is one solution to mitigate the lateral movement threat. However, as seen in Ukraine repeatedly, despite network segmentation, the attacks were still successful. This is because network segmentation alone cannot guarantee a complete protection against advanced persistent threats. Hence, power grid operators must complement network segmentation with additional security controls such as next-generation firewalls and Intrusion Detection and Prevention System (IDPS) to minimize the threat of lateral movement.

### **2.2.4 MAN-IN-THE-MIDDLE ATTACKS**

Man-in-the-middle is a type of cyber attack classified based on the location of the adversaries. In this attack, adversaries are located between two or more hosts, allowing them to maliciously observe and be involved in their communication traffic. In this section, we focus on potential Man-in-the-Middle (MITM) attacks targeting power grids, which are classified into five categories, i.e., eavesdropping, spoofing, FDI, replay attacks, and session hijacking.

#### **EAVESDROPPING**

Eavesdropping, also known as sniffing or snooping, is an attack where adversaries intercept information transmitted over the network. The main objective of eavesdropping is to intercept and gather information about the contents of the transmitted data. In comparison to other attacks, adversaries seek to only observe and not change legitimate communication. To mitigate the threat of eavesdropping, encrypted protocols can be used for communication. Due to encryption, adversaries or third parties cannot easily decipher the contents of the transmitted data. However, in a real SCADA system, most of the dataflows are unencrypted. Moreover, communications between SCADA end devices such as station control systems, Remote Terminal Units (RTU), protection relays, and Merging Unit (MU) mainly work based on a broadcast communication mechanism. For example, broadcast

communication through the Distributed Network Protocol 3 (DNP3) protocol is widely adopted in SCADA communications [49]. Such broadcast mechanisms are susceptible to eavesdropping and sniffing attacks. Adversaries can easily intercept communications by gaining unauthorized access to the OT system and exploiting the vulnerabilities of the broadcast uncommunication mechanism. Therefore, eavesdropping plays an important role in an advanced persistent threat, especially during the reconnaissance stage. Hence, it can be used as a stealthy mechanism to gather network intelligence through passive means [50]. Valli et al. present a study about eavesdropping attacks targeting smart grids in [51]. The eavesdropping is mainly focused on the Advanced Metering Infrastructure (AMI). AMI establishes communications through wireless channels between a smart grid operator and smart metering devices. Adversaries may exploit the vulnerabilities present in wireless networks to intercept communications and capture transmitted data. Research also shows that eavesdropping can lead to privacy concerns for smart grid users, as seen in [52, 53].

### SPOOFING

Spoofing is an active attack where adversaries pretend to be legitimate entities and disrupt normal communications. Such attacks can be realized through many forms of spoofing such as emails, website URLs, text messages, Global Positioning System (GPS), and IP addresses. The most widely researched spoofing attack targeting power grids is based on GPS spoofing of Phasor Measurement Unit (PMU) data. There are multiple studies in this direction, as discussed in [54–56]. PMUs provide magnitudes and phase angles of fundamental power system parameters such as voltages and currents, using a common time source for synchronization [57]. Hence, GPS spoofing attacks, mainly targeting the timing signals used for synchronization, may lead to distortion of observed PMU data, which includes phase angle errors [58]. There are two types of GPS signals widely in use for civilian and military applications. GPS signals for military purposes are encrypted, while civilian ones are not. Typical PMUs for power grids function based on civilian GPS signals. This may allow adversaries to spoof GPS signals by exploiting the lack of encryption and using a portable device without a direct access to the power grid communication network. Currently, there are two approaches to mitigate GPS spoofing attacks. The first is through GPS spoofing signal detection using parameter such as signal to noise ratio [59] and the second is via anomaly detection in power system measurements.

Another commonly reported type of spoofing attack on power systems is IEC 61850 spoofing. The IEC 61850 standard is a modern power system communications standard used for substation automation and protection in digital substations. It enables information exchange through different communication protocols, of which two are of utmost importance. The Generic Object Oriented Substation Event (GOOSE) and Sampled Values (SV) protocols are used to communicate critical substation events and measurements within a substation, respectively. Although it provides increased benefits, IEC 61850 is not cyber secure. Due to strict operational constraints and timing requirements for power system protection schemes, the standard does not implement any encryption. This makes it particularly vulnerable to packet sniffing and spoofing type of attacks. Such types of spoofing attacks are well reported and have been investigated extensively in literature [60, 61]. Multiple vulnerabilities and exploits, specifically targeting GOOSE and SV protocols are widely discussed in [62–64].

The premise of all these discussions is similar. Due to the lack of encryption in IEC 61850, an attacker with access to the substation communication infrastructure can wreak havoc. By carefully monitoring IEC 61850 traffic via the process and station buses, it is possible to craft spoofed GOOSE packets that can maliciously open circuit breakers. When such spoofed packets are sent to a target relay, it is tricked into opening the circuit breaker. This is successful as the packet is made to appear to be originating from the station or a bay controller, within the same substation. It is also possible to inhibit protection functionality of relays due to spoofing of SV measurement data, causing protection equipment to not operate during a critical fault conditions [65]. The spoofing attack causes a relay to get blocked from further operations due to multiple concurrent input SV streams. With the target protection device blocked, other relays in the system may trip during fault conditions. Such types of spoofing attacks can have disastrous consequences for power system operation and stability. A well targeted spoofing attack can not only compromise but also disable equipment and components within a digital substation. Subsequently, this can instigate major system instabilities, and may even induce cascading failures, due to the sudden loss of multiple components. In a doomsday scenario, attackers may trigger a system-wide collapse, i.e., a blackout, by compromising critical digital substations, leading to catastrophic damages. Nonetheless, such types of spoofing attacks can be mitigated by adopting proper cyber security measures, as discussed in [66].

IEC 62351-6 is a standard that specifically addresses cyber security of IEC 61850. It recommends an additional field to the GOOSE and SV data payloads for security-related information. This field contains a Rivest-Shamir-Adleman (RSA) based digital signature to ensure payload integrity. Through this mechanism, sending and receiving Intelligent Electronic Devices (IED) are clearly identified and it becomes impossible to manipulate the payload. Similarly, the standard also recommends usage of Hash-based Message Authentication Code (HMAC) using cryptographic algorithms such as SHA-256 to ensure data integrity of GOOSE and SV frames. Such techniques can prevent spoofing and sniffing attacks. However, the suggested use of the digital signatures and security based on RSA and HMAC algorithms comes with associated costs. For protection applications where a 4 ms or lower response time is strictly required, such measures are unsuitable. This is because, encryption and decryption are computationally demanding [66, 67]. Furthermore, the usage of RSA and HMAC-based authentication keys for IEDs and equipment necessitates a dedicated key management infrastructure within the digital substation. Hence, such security mechanisms have not gained widespread use, yet.

### FALSE DATA INJECTION

The most extensively researched type of cyber attack on power systems is the false data injection (FDI) attack. An FDI attack operates under the assumption that attackers have access to the station control systems and RTUs in substations and/or the SCADA master in the control center. Consequently, they can inject falsified SCADA measurements, maliciously introducing correlated and consistent power flow measurements into State Estimation (SE), aiming to mislead system operators. Nowadays, SE is an integral tool in the energy management system for contingency analysis, security-constrained optimal power flow, and pricing calculation algorithms. The critical nature of SE highlights the importance of making it accurate and secure for power system operation. However, as discussed above, the SCADA system is vulnerable to FDI attacks. In [68], Liu et al. introduced a class of

FDI attacks that can perturb the estimated states without being detected by the safeguard scheme within the SE process. The interesting part of such attack is that the adversary is assumed to have the knowledge of the targeted power system including the power network topology and parameters, and thus can exploit such knowledge to systematically generate multiple FDIIs on power flow measurements [69]. It has been illustrated that such FDI attacks can bring potential economic damages by manipulating the nodal price of market operations [70] or even physical impact such as a line overload [71]. The FDI attack may seem difficult to conduct as the adversary needs to be equipped with enough knowledge of the target power system and vast attack resources to manipulate multiple measurement data channels. However, the complexity and functionalities of malware in recent cyber incidents on industrial control systems provide credible means to realize the FDI attack [72]. Notably, in addition to FDI on state estimation, recent research also considers attack scenarios where other critical applications, e.g., automatic generation control, are targeted [73]. In addition, studies on power system vulnerability analysis have been carried out to explore how FDI attacks can achieve the desired targets with incomplete system knowledge or very few attack resources, using both static and dynamic (time-variant) FDI strategies [74, 75]. Detection and mitigation techniques at the physical layer of the power system are proposed in [76, 77] based on both model-based and data-driven detectors from control-theoretic domains or machine learning areas.

### REPLAY ATTACKS

A replay attack is a variant of the man-in-the-middle attack where attackers record communication traffic and replay it to mimic legitimate entities. Pidikiti et al. investigated replay attacks on the SCADA system exploiting IEC 101 and IEC 104 protocols in [78]. These SCADA protocols were originally created without cyber security considerations. Nevertheless, these protocols do implement a packet checksum mechanism to prevent replay attacks to a certain extent. However, the size of the checksum is small and limited by the packet frame size and bandwidth. This condition leads to unreliable checksums to ensure data integrity. Therefore, this vulnerability can be exploited to launch replay attacks on SCADA systems. On the other hand, replay attacks in IT systems are more common and usually prevented by using authentication and secure session mechanisms. For example, a countermeasure to replay attacks was proposed using Kerberos authentication protocol [79]. This protocol would force the network hosts to authenticate themselves. After a successful authentication, a secure session is established between hosts. Such a session is typically valid only for a limited period of time preventing the reuse of session information. However, adoption of such prevention mechanisms in OT systems is challenging. For example, in IEC 101 and IEC 104, the limited packet frame size makes it difficult to add more data to improve protocol security. In addition to the aforementioned authentication mechanisms at the cyber layer, research efforts have also been undertaken to study the replay attack from the perspective of the physical power system layer. Such research is focused on detection methods to secure the control process of the SCADA system in power grids [80]. However, it is to be mentioned, there could still exist sufficient conditions under which plausible replay attacks may remain stealthy irrespective of the detection mechanism used. This is applicable even to a control-theoretic approach wherein the attacker has access to all the necessary data channels and executes the replay attack at a suitable time [81].

### SESSION HIJACKING

Communication sessions are interactive information exchanges between two or more networked devices for a limited time duration. Typical session establishment is initiated through authentication between hosts via secure protocols. Therefore, a session hijacking attack aims to bypass these protocols, allowing adversaries to circumvent authentication mechanisms and gain unauthorized access to legitimate communications. Kleinmann et al. presented a study on session hijacking in SCADA systems by exploiting Modbus protocol [82]. Modbus was originally designed only for serial communications between field devices in substations. To improve its flexibility, it was later upgraded to implement TCP. This modification allowed Modbus to work on Ethernet connections using IP addresses providing more data faster. The session establishment in Transmission Control Protocol/Internet Protocol (TCP/IP) works based on a three-way handshake mechanism. However, TCP/IP is widely known to be susceptible to cyber attacks including session hijacking [83] and thereby compromising Modbus as well. Besides session hijacking at the protocol level, hijacking can also be conducted through web applications or Human Machine Interface (HMI) of SCADA systems. A successful session hijacking attack allows adversaries to assume the identity of the compromised devices / users and provides unauthorized access and control of the OT system. Burgers et al. presented a session hijacking case study and mitigation techniques for SCADA [84]. This research focuses on session hijacking via web-based applications, which use login authentication for session establishment.

### 2.2.5 DENIAL OF SERVICE ATTACKS

Denial-of-Service (DoS) is a cyber attack with the objective of preventing legitimate access for users / networked devices to specific system resources such as network connections, computing capabilities, and application services. The term distributed denial-of-service refers to a coordinated DoS attack originating from multiple, distributed sources to increase attack severity and prevent tracking and identification of attackers' origin. A single DoS attack can be mitigated by blocking the sole attack source. Conversely, for a DDoS attack, blocking all attack sources is challenging, making its mitigation difficult. DoS attacks can further be classified into bandwidth depletion and resource depletion attacks [85]. The bandwidth depletion DoS attack aims to overload the bandwidth capacity of a target communication network. This can be achieved by either directly flooding the communication channel with bogus traffic or via third parties, which send multiple legitimate requests at the same time in an amplification attack. As a consequence, in either instance, legitimate communication traffic is affected, which significantly reduces the overall network performance. The resource depletion attack aims to overwhelm the target's resource usage, e.g., computing resources of a targeted host, by exploiting protocols and known response mechanisms. For example, as previously mentioned, TCP/IP implements a three-way handshake mechanism, allowing two hosts to initiate communication with a preliminary request and response mechanism. Adversaries may exploit this mechanism by sending a multitude of malicious requests to the targeted host. Consequently, the targeted host is kept busy responding to all malicious requests, leading to the disruption of a proper response to legitimate ones.

DoS attacks can target SCADA systems of power grids. Studies about SCADA susceptibility to DoS attacks are reported in [86, 87]. Petrovic et al. demonstrated DoS attacks

on SCADA systems using OPNET communication network simulator [86]. The attacks significantly reduced SCADA network throughput and processing capabilities, directly affecting power system monitoring and control. Similarly, Kalluri et al. demonstrated a DoS attack exploiting IEC 104 protocol used in substations, affecting the processing and communication capabilities of RTUs [87]. Carcano et al. demonstrated a resource exhaustion attack targeting IEC 62351 [88], highlighting its cyber security shortcomings. In summary, the DoS attack is a potential threat against data availability in power grids, as it prevents successful communication of measurements and controls. Attackers can either jam the SCADA communication channels or compromise field devices and prevent them from communicating data. They may also attack the routing protocols or flood the network with bogus traffic [89]. DoS attacks on power systems may be modelled to analytically study the impact of data absence on power system monitoring and control. By properly designing DoS attack sequences, attackers can corrupt the normal operation of controllers and consequently impact power system stability [90, 91]. Mitigation techniques are discussed in [92].

### 2.2.6 HOST-BASED ATTACKS

A host-based attack as the name suggests is an attack targeting various hosts in IT/OT systems, such as SCADA servers and HMIs, databases, application servers, station control systems, RTUs, protection relays, and merging units. In this section, we classify host-based attacks into three categories, i.e., software-based, database, and unauthorised access and control attacks.

#### SOFTWARE-BASED ATTACKS

Software-based attacks on power grids exploit vulnerabilities present in software used in IT/OT systems such as SCADA and energy management systems. Usually, the software applications and security controls in OT systems inherit the same vulnerabilities present in regular IT systems. The main issue is that software and security controls of such IT systems may be patched, and their vulnerabilities may be mitigated more often than in OT systems. It is more difficult to update the OT systems of critical infrastructures as this process can affect the normal operation of physical facilities. A disruption of service such as electricity supply to customers may result in regulatory penalties and financial loss. Furthermore, extensive commissioning is needed after each update process that prolongs the voluntary outage for maintenance.

Most SCADA system solutions provided by vendors were developed before the emergence of cyber security concerns [93]. Software vulnerabilities in SCADA systems can be classified into three categories, i.e., improper input validation, software or source code, and resources control vulnerabilities [94]. As a result of input validation vulnerabilities, SCADA software is susceptible to modification attacks such as data injection and buffer overflow. SCADA source codes have also been found to contain improper security mechanisms and vulnerabilities such as the null pointer dereference vulnerability [95]. Resources control vulnerabilities are strongly related to software updates and patches. Corporate IT security typically pushes software updates and operating system patches over the IT network. However, SCADA software updates and patching in control centers and substations are more



difficult to implement. This is due to the blend of state-of-the-art and legacy end devices, in addition to continuous operational requirements of the SCADA system in production.

In August 2020, nineteen software vulnerabilities were exposed by JSOF, an Israeli cyber security company. These vulnerabilities, dubbed Ripple20, affected ICS devices using the proprietary Treck TCP/IP stack software libraries. Two of the most severe vulnerabilities are related to TCP/IP tunneled packet fragmentation [96] and Domain Name System (DNS) packet decompression mechanisms [97]. The Treck software library has widely been adopted in IoT networked devices by several vendors across a whole range of industries including manufacturing, healthcare, and power grids. It is a cause for serious concern, as shown in [97], that a specific payload injection could remotely turn off an Uninterruptible Power Supply (UPS) device. Therefore, we can infer that Ripple20 is a real-world example of challenges pertaining to updates and security of software in industrial control systems, further complicated by global supply chains.

### **DATABASE ATTACKS**

A database is an essential element of the SCADA system as it stores real-time information from substations along with user access credentials. Zhu et al. categorize a database attack as an important cyber attack vector targeting SCADA systems [98]. Most common databases work based on SQL. Thus, one of the popular attacks targeting databases is SQL injection. This attack exploits input handling of the database system. When a database cannot correctly parse and handle inputs, it may lead to database access violations and illegitimate manipulation. In the worst-case scenario, with the breached confidential database information, adversaries can gain unauthorized control of the SCADA master. Consequently, databases have proven to be an important attack element in real-world cyber attacks on power grids. For example, in Ukraine 2015 attack, adversaries gained access to the control center using stolen credentials from the Windows Active Directory (AD) database [97]. AD is one of the most critical applications since early 2000 as it offers flexibility and interoperability of service authentication and authorization. However, a breach in security measures of AD can lead to a breach of the entire system since AD serves as the central authentication and authorization point. There are many publicly available tools to exploit AD security. For example, Mimikatz can be used to exploit AD hashes and Kerberos ticketing mechanism. Nonetheless, there are counter measures to prevent cyber attacks targeting AD. One of the options is the application of Microsoft Credential Guard. Credential Guard is a virtualization-based isolation technology for Local Security Authority Subsystem Service (LSASS) which prevents attackers from stealing credentials and prevents hash attacks. Another option is the implementation of tiered (multi-level) administrator models. The tiered admin model can prevent attackers from gaining top level privileges in an AD. Another common practice to prevent AD breaches is to implement secure credential policies. For example, a user has to change passwords periodically and use strong combination of characters. Multi-level or two factor authentication mechanisms through mobile phone messages and emails can also be applied to increase the overall system access security.

### **UNAUTHORIZED ACCESS AND CONTROL**

Access authorization typically uses an authentication mechanism applied to secure hosts, software, and web services. Unauthorized access occurs when an adversary gains access

to the system without legitimate credentials. Hence, unauthorized access and control can be achieved if attackers circumvent the authentication mechanisms. There are many techniques to achieve this objective such as credential theft using a keylogger, database breaches, brute force attacks, and buffer overflows. It is also possible to gain unauthorized access using penetration testing tools such as Metasploit. This exploits system vulnerabilities by injecting malicious payloads on the target system. The most basic form of unauthorized access is achieved through the guest (non-administrator) mode. However, in this mode attackers' options are limited. Thereby, to increase attack severity, attackers can perform privilege escalation and become administrators allowing them complete control over the compromised system. SCADA systems typically employ Windows-based operating systems. However, Windows is vulnerable to unauthorized access attacks. Thus, Windows operating systems in IT/OT systems must be regularly updated and protected with firewalls and antiviruses. Researchers have also proposed solutions to prevent unauthorized access in SCADA systems. Taylor et al. proposed a SCADA authentication technique using a custom key distribution mechanism [99] applicable to DNP3 protocol, to prevent unauthorized access and control. Similarly, Vaidya et al. proposed an authentication and authorization for substation level communications [100]. This method implements multi-level authentication and uses public key certificates to authenticate and authorize access to the substation automation system. Other approaches to prevent and reduce the risk of unauthorized access and control include measures such as securing the host operating systems and implementing security perimeters and IDPS.

## 2.3 IMPACTS OF CYBER ATTACKS ON POWER GRIDS

Cyber attacks on power grids are considered high impact, low frequency disturbances with a wide range of effects. These could include, but are not limited to, equipment damages, loss of load, and power system instability. In the worst case, sophisticated cyber attacks may also cause system-wide cascading failures, leading to a blackout. Hence, this section discusses the various potential impacts of cyber attacks on power grids, ranging from component to system level. Table 2.3 summarizes the known cyber attacks on power grids and their impact. Four of the attacks shown in Table 2.3 are real, except the Aurora attack. The Aurora project was an experimental cyber attack that led to the physical destruction of a 2 MW synchronous generator. This was mainly done as a demonstration to raise awareness about cyber security and associated threats. The significant cyber attack on power grids, so far, is the Ukraine 2015 attack. It has been confirmed that this attack directly led to a power outage, affecting over a quarter-million customers for a duration of over six hours. Besides the real-world examples of cyber attacks on power grids, research has also been carried out to investigate the potential impacts of cyber attacks on power system operation [101–103]. Such empirical studies discuss various cyber attack scenarios and associated effects. A doomsday scenario would entail a cyber induced cascading failure culminating in a complete blackout. Hence, the subsequent subsection firstly provides an overview of the cascading failure mechanism, followed by various cyber attack scenarios and their impact analysis, as reported in the literature.



Table 2.3: Summary of known cyber attacks on power grids and their impact

No.	Attack	Year	Category	Impact
1	Malware infection of SCADA system, Europe [104]	2003	Service disruption	Loss of management functions in substations for three days
2	Aurora experimental cyber attack, USA [105]	2007	Physical damage	Damage to a 2 MW synchronous generator
3	Power plant malware infection, USA [104]	2012	Service disruption	3-week restart delay of power plant
4	Cyber attack on power grid, Ukraine [35]	2015	Service disruption	Outage affecting 225,000 customers for 6 hours
5	Cyber attack on power grid, Ukraine [43]	2016	Service disruption	Outage in distribution network, 200 MW unsupplied load

### 2.3.1 CASCADING FAILURES

Any major power system blackout is preceded by the phenomenon of cascading failures. A cascading failure, as the name suggests, is a successive failure of power system elements that can lead to a complete system collapse, i.e., a blackout. Most cascading failures are initiated by one or a set of multiple related events. These can include line flashovers, protection maloperation, human error, etc. Historically, most of these events tend to be caused by a combination of equipment failures, e.g., ageing equipment, environmental conditions, and human factors. Depending on the operating state of the system and severity of the initiating events, the entire power system may enter an emergency state. Without proper control actions or remedial measures, the system is highly vulnerable to further cascading effects. In such a case, various outcomes are possible. One such outcome, commonly observed in historical blackouts such as Italy 2003 and USA 2003 [106] is as follows. Due to the initial set of events, overloading of parallel transmission lines occurs, to account for power redistribution. Eventually, these lines are also overloaded beyond their limits and start tripping, initiating a cascading process of transmission line disconnections. After a certain time, the effect of these outages is felt on system dynamics. Transient instability can occur in a matter of a few seconds due to the large disturbances. Generators may lose synchronism due to sudden loss of transmission lines. This will also affect system voltages, causing major voltage drops. Consequently, in the case of heavy system loading, voltage stability problems may also arise. An inability to meet growing reactive power demands can eventually result in a voltage collapse. If left unchecked, such dynamic phenomena can result in islanding, i.e., formation of smaller clusters in the system with a mismatch of supply and demand. Ultimately, the power system reaches a so-called point of no return [107]. From this point onwards, the entire cascading process is rapid, involving loss of multiple generator units and loads. This domino effect is uncontrollable, culminating in a blackout. It is worth mentioning here that such a sequence of events is based on historical cascading failures and blackouts, involving physical system events. In case of a well targeted and coordinated cyber attack, the effects can be severely magnified. As shown in [108, 109] cyber attacks targeting bulk power systems may initiate cascading failures. A sophisticated cyber attack can target multiple substations, disconnecting many lines and tampering with control setpoints. As a result, power system instability and associated phenomena discussed above may be induced much faster. Consequently, in comparison to

previous blackouts, the point of no return may be reached much sooner in case of cyber attacks.

### 2.3.2 IMPACTS ANALYSIS

2

As previously mentioned, the physical impact of cyber attacks on power grids is wide-ranging. There are many empirical studies reported in literature, covering these impacts. The most reported consequence of a cyber attack on power grid infrastructure is equipment damage. The Aurora experiment is a good real-world example of such possible effects. The attack demonstrated how rapid opening and closing of a generator's circuit breaker cause an out of phase reconnection and permanent equipment damage. Along similar lines, [109] extensively discusses switching attacks on generators. This work clearly highlights how sophisticated cyber attacks can not only disconnect generators and cause equipment damages, but also initiate cascading failures. By applying a fast, switching attack on a generator's main circuit breaker, transient instability can be induced, destabilizing the entire power grid in a matter of a few seconds. Other possible impacts include damage to equipment such Flexible Alternating Current Transmission System (FACTS) devices and On-Load Tap Changer (OLTC) through setpoint modification [110, 111]. These devices are critical in ensuring voltage stability, and such equipment damage can trickle down and affect the entire power system. Loss of load is another commonly reported result of cyber attacks on power grids. If a cyber attack affects the system frequency, automatic measures such as load shedding are undertaken to preserve system integrity. Additionally, switching or data modification attacks can directly lead to loss of load [112]. The worst possible outcome, however, is that of cyber induced cascading failures and a blackout. As discussed in [108], a cyber attack on multiple substations in any power system may lead to a blackout. Cyber attacks targeting specific grid components or equipment can impact power system stability. For example, targeting voltage control mechanisms such as Static VAR Compensator (SVC) and Static Synchronous Compensator (STATCOM) can severely affect voltage stability. By carrying out data modification or MITM attacks, as stated in [110, 111] reactive power compensation is severely affected. As a result, voltages throughout the system can be influenced. Sustained under voltages can lead to emergency load shedding, and in the worst case, a voltage collapse. As part of a coordinated effort, such attacks can induce system-wide cascading failures and even a blackout.

## 2.4 VULNERABILITIES OF POWER GRID OPERATIONAL TECHNOLOGIES

OT systems have a longer lifecycle than traditional IT systems. IT cyber security is not covered in this study, as successful cyber attacks aim to gain access and extract digital information, meaning confidentiality is a major concern. On the other hand, by compromising and manipulating measurements or control commands, the attackers can cause physical damage, endangering both the operational and human safety of an ICS. In a recent report by Dragos Inc. published in 2022, the number of critical vulnerabilities on the whole spectrum of industrial systems has exploded [113]. Focusing on CPPS, the standards based on which the OT architecture of these systems are designed often lack cyber security considerations [114, 115]. Additionally, as the life cycle of the power system OT equipment

is long, newly developed hardware will be implemented partially on substations, while vulnerable devices will still be present. Two classes of vulnerabilities are identified: the communication protocols used in CPPS and the software vulnerabilities.

### 2.4.1 COMMUNICATION PROTOCOL VULNERABILITIES

The standard communication protocols in power systems aim to connect the various industrial equipment and metering infrastructure to the local control systems and the control center. These industrial protocols are utilized to define the communication between the devices in the CPPS. Field devices such as RTUa, IEDs, and MUs are connected through communication nodes and links to the local SCADA applications for local monitoring and control. Currently, most protocols are IP-based, using TCP or User Datagram Protocol (UDP) packets for flexibility of implementation.

The communication between sensors, meters, IEDs, and the SCADA or HMI of a substation is realized with protocols such as Modbus, DNP3, and IEC 61850. Modbus is a well-known example of a standard communications protocol that was adopted across a wide area of industries, including power systems. It is implemented between programmable logic controllers and HMIs, utilizing the master-slave configuration [116]. The devices can also be configured to run these roles in parallel. The Modbus protocol works on two types of communications, serial line and TCP over Ethernet. Notable cyber security vulnerabilities of this protocol are derived from the lack of security applications, authentication mechanisms, encryption, and integrity validation [117]. Adversaries can intercept existing Modbus sessions and replicate the session by analyzing the traffic, making it susceptible to MITM attacks. Additionally, Modbus can be made more scalable by embedding its frame in a TCP packet. As a result, Modbus over TCP inherits cyber security issues of TCP [118]. Moreover, the application of Modbus over TCP does not properly implement the TCP message checksum. Hence, it can be easily intercepted and compromised by a spoofing attack.

DNP3 is a popular communications protocol used in power systems for SCADA operations. It was introduced to support communications between the control center and substations. The original objective of DNP3 was to transmit small-sized data packets using serial RS232 for relatively short-distance point-to-point communications. The protocol implements four layers from the Open Systems Interconnection (OSI) model, i.e., physical, data link, transport, and application. The latter variants of DNP3 were extended to work using TCP and UDP packets over Ethernet. While DNP3 is more reliable than Modbus, it also consists of vulnerabilities, making it prone to spoofing and distributed DoS attacks. In [119], 28 possible attack vectors that could target the DNP3 protocol were identified. Attacks targeting DNP3 can be categorized into the process of interception, interruption, modification, and fabrication. The protocol is also vulnerable to MITM attacks like packet sniffing and spoofing attacks. These attacks can result in three types of impact: 1) loss of confidentiality, 2) loss of awareness, or 3) loss of control. Loss of confidentiality happens when the attacker successfully intercepts the communication. Loss of awareness occurs when the control center does not obtain precise and trustful information. The most critical type of attack is the loss of control, wherein attackers can take unauthorized control of the system.

IEC 61850 is a modern power system communications standard for substation automa-

tion and protection, which allows information exchange through several communication protocols, including GOOSE, SV, and Manufacturing Message Specification (MMS) [114]. Compared to Modbus and DNP3, IEC 61850 Ethernet-based communications provide larger bandwidth. In this standard, power system communication is mapped into TCP/IP packets sent over Ethernet and can be applied for local and wide-area communication [120]. For example, Routable IEC 61850 was implemented through data encapsulation into TCP packets [121]. This mechanism facilitates the expansion of IEC 61850 communication beyond the boundaries of a single substation, enabling its routing across more extensive areas of the OT networks. IEC 61850 is used to exchange control and measurement packets for local communication within a substation, between substations, as well as between substations and the control center. Although IEC 61850 is defined by its high scalability and increased functionalities, cyber security vulnerabilities are identified. IEC 61850 GOOSE and SV protocols are identified as vulnerable to spoofing attacks [65, 122, 123]. Additionally, a critical issue is derived from the problem that encryption measures are difficult to apply due to the low latency requirements (3-4 milliseconds) for power system protection applications.

The most commonly used standard for establishing communication between the substations and the control center is IEC 60780-5, as well as IEC 61850. From this category, IEC 60870-5-101 and IEC 60870-5-104 are the most widely used protocols. IEC 101 is a protocol for basic supervisory control and data acquisition, while IEC 104 has increased performance due to its network and transport layers, in addition to the application layer protocol [115]. IEC 104 can provide network access to IEC 101 using TCP/IP. Due to the vulnerabilities in the TCP/IP stack, common vulnerabilities are: i) messages are transmitted in plain text [124], and ii) lack of authentication mechanism. Thus, cyber actors can perform MITM attacks by sending malicious control commands or connecting to the network.

To address cyber security issues of the aforementioned standards, IEC 62351 is proposed, with a goal to enhance confidentiality, integrity, and authenticity. As such, many of the parts of this new standard are based on IEC 61850 and 60870. IEC 62351 is still in the process of development. It provides new definitions related to cyber security, like role-based access control, key management, and security architecture [125]. Some identified vulnerabilities, described in [126], enable cyber actors to perform replay attacks using GOOSE and SV. One vulnerability is related to time exchange based on the simple network time protocol. Security enhancements proposed in IEC 62351 rely on the implementation of cryptographic measures, such as RSA cryptography. However, entire packets are not encrypted using RSA, and only the protocol data unit is encrypted. This situation allows attackers to modify the unencrypted parts. Despite the implementation of IEC 62351, modification of packet counters and time stamps allows attackers to launch GOOSE and SV-based attacks. IEC 62351 prevents the manipulation of traffic by implementing a Message Authentication Code (MAC), which is a hash-based authentication to validate the integrity of data. Still, the attackers can modify time-related information using the vulnerabilities of the data encryption standard that MAC relies on. Consequently, this can violate the rules of packet processing, thereby triggering DoS conditions. Besides the possible theoretical exploits of IEC 62351, there is also an example of a real attack demonstration. Carcano et al. demonstrated cyber attacks targeting SCADA networks running IEC 62351 [127].

Additional protocols that are employed for measurement and communication in the power system are IEEE C37.118, IEEE C37.247, and Inter-Control Center Communications

Protocol (ICCP). IEEE C37.118 defines a mechanism for real-time exchange of synchrophasor data and messaging formats, message types, and content [128]. The messaging format, as well as the requirements for data transfer, are defined by the standard. It uses a universal time source as a reference and time stamps based on the GPS to provide time-synchronized measurements of power system parameters from different locations. This feature is crucial for implementing Wide Area Monitoring Protection and Control (WAMPAC) applications [129]. Cyber security considerations are mainly due to the lack of security mechanisms like authentication and encryption [130]. The lack of the aforementioned security mechanisms enables cyber attackers to manipulate the data packets for MITM attacks or to inject forged ones. These vulnerabilities can be utilized to target wide-area applications, enabling cyber actors to indirectly target the physical power system.

The C.37.247 defines the Phasor Data Concentrator (PDC) operation, which is used to synchronize, process, and transmit the data collected from individual PMUs [131]. Synchrophasor measurements from the distributed PMUs are received in real-time by the PDCs, which are used to shape the data in a single data stream and transmit it to higher levels. PDCs are used for timestamp alignment, filtering corrupted or false data, and maintaining a record of data. Regarding cyber security, the communication links used for data transmission to/from PDCs are not encrypted. Malicious actors can utilize the lack of encryption, vulnerabilities that enable access to the configuration and programming software of the devices, and lack of traffic control to launch DoS, MITM, and time delay attacks [132, 133].

Finally, ICCP, also known as IEC 60870-6/TASE.2, is a data exchange protocol commonly used for communication between independent system operators, regional transmission operators, generators, control centers, and utilities over the Wide Area Network (WAN) [134]. It facilitates communication between two control centers based on a client-server model. The vulnerabilities of this protocol are identified in [135] and enable cyber actors to attack the process control data, the ICCP servers, and server operating systems. These vulnerabilities mainly exploit the lack of security mechanisms like encryption and authentication. Without those two security properties, many possible exploitations can be performed over ICCP. Studies showed that security mechanisms can be applied to ICCP [136]. It can be encrypted and authenticated as secure ICCP. Implementation of secure ICCP relies on public-key cryptography.

In summary, the protocols employed by CPPS exhibit vulnerabilities to cyber attacks as a result of inadequate implementation of cyber security measures. The security measures employed in the field of IT communication protocols primarily rely on the implementation of cryptographic techniques. Nevertheless, the implementation of cryptography in CPPS is challenging due to the stringent demands for high availability and low latency. In [137], the author provided evidence that CPPS face difficulties when attempting to integrate cryptography into their systems. This is primarily attributed to the substantial amount of computational time that cryptographic processes demand. Although cryptographic algorithms like 2048-bit RSA and 1024-bit DSA are considered robust, the processing time of cryptographic operations respectively entailed a total of 61.04 milliseconds and 14.90 milliseconds. Due to limited time availability, this situation led to the utilization of cryptographic techniques that offer reduced security and computational requirements or, in many instances, the complete absence of cryptographic measures.

### 2.4.2 SOFTWARE APPLICATION VULNERABILITIES

In existing electrical power systems, software technologies for OT are mainly related to SCADA systems. In the future, this will be integrated with various software functionalities, i.e., energy management systems and advanced distribution management systems. SCADA is a control system architecture that consists of interconnected devices controlled by OT software. The main challenge for a SCADA software system is the regular software updates. Most of the existing software was created before cyber security issues became a major concern [138]. In [139], three groups of SCADA software vulnerabilities are identified: 1) improper input validation, 2) resource control, and 3) software code. SCADA software is susceptible to input value modification attacks such as buffer overflow and data injection. With regard to the code itself, OT systems tend to be less secure. This is because OT software is designed for high availability requirements, with less consideration of regular updates and security mechanisms.

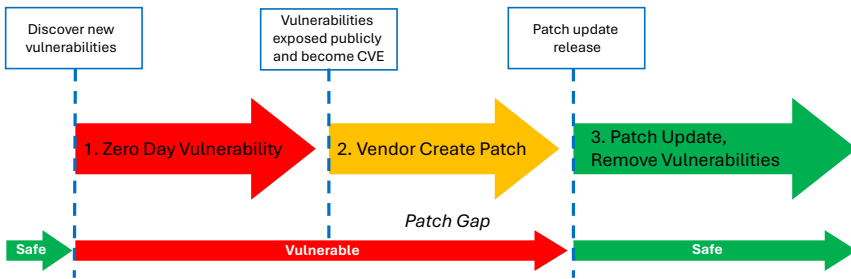


Figure 2.2: Lifecycle of a software vulnerability [2].

Resource control vulnerabilities are mainly related to software updates and patch control mechanisms. Vulnerable software that may have been deployed in the field must be updated and patched to eliminate vulnerabilities. However, software updates and patches in OT are challenging and can potentially disturb system operations. Fig. 2.2 shows the lifecycle of a software vulnerability that is applicable to SCADA software. In principle, software vulnerabilities will always exist. When limited parties identify these vulnerabilities, it becomes stage one, i.e., a zero-day vulnerability. The zero-day vulnerability is dangerous when exposed by adversaries. In the second stage, information is exposed publicly, and software vendors create software updates and patches to address the vulnerability. After this, it is no longer considered a zero-day vulnerability. For example, MITRE's Common Vulnerabilities and Exposures (CVE) lists all the vulnerabilities in SCADA-related applications [140]. A patch update is released by the vendor to address specific vulnerabilities. However, as previously mentioned, patch updates in the SCADA system are quite challenging, and they may be deployed in remote locations [141, 142]. Hence, vulnerabilities in SCADA software are likely to stay present and unaddressed during the software lifecycle.

An example of the potential impact of the software code vulnerabilities is the Ripple20. In June 2020, nineteen software vulnerabilities were discovered by JSOF, a cyber security firm. These vulnerabilities affect devices using the Treck Inc. TCP/IP stack software library. The vulnerabilities are based on the exploitation of TCP packet fragmentation,



tunneling mechanism [143], and DNS decompression mechanism [97]. Many networked devices widely use this software library for the TCP/IP stack across a plethora of industries, including SCADA, offices, healthcare, etc. By exploiting the vulnerabilities, adversaries can disrupt the functioning of the devices. An investigation in [97] shows an example of a malicious payload that can successfully switch off a UPS device remotely. These vulnerabilities are a significant problem since it is difficult to update software or firmware in embedded devices. Ripple20 is a real-world example of difficulties performing software updates for SCADA devices and systems. Legacy SCADA systems can also be integrated with energy management systems in the future power grid for various advantages.

Future power grid software like energy management systems and advanced distribution management systems can help achieve more intelligent grid operations. The operation of the power grid is not only dependent on human operators but also on smart and intelligent systems. The software can be in the form of AI applications. The implementation of smart software or AI will advance the digitalization of the overall grid. However, the cyber security aspects cannot be overlooked. Adversarial machine learning is one such major potential threat that can fool the AI-based system. In [144] and [145], the authors show how adversarial machine learning can have adverse effects on the operation of smart software systems. Such adversarial machine learning may become a new type of threat to power system software systems in the near future.

## 2.5 SECURE COMMUNICATION PROTOCOLS

In order to successfully mitigate the threat of cyber attacks on power grids, it is important to first understand the relationship between computer networking and cyber security. Fig. 2.3 presents the mapping between communication network layers and associated cyber threats and countermeasures, based on the well-known OSI seven-layer and TCP/IP four-layer models. The seven-layer OSI abstraction explains the flow of data in computer networks as bits in the physical layer, frames in the data link layer, packets in the network layer, Transport Protocol Data Unit (TPDU) in the transport layer, Session Protocol Data Unit (SPDU) in the session layer, Presentation Protocol Data Unit (PPDU) in the presentation layer, and finally as Application Protocol Data Unit (APDU) in the application layer. SCADA communications typically uses APDUs to deliver the payloads, i.e., measurements and controls. Information exchange and delivery is done either through network layer or data link layer. Layer 2 communication is limited to the confines of a substation where the data is exchanged as a frame. Meanwhile, layer 3 communication is used for communication between the substations and control center. Layer 3 communication uses the TCP/IP stack and network routing mechanisms to deliver information.

Fig. 2.3 also shows the attack types for each layer of the OSI model and its associated countermeasures. The physical layer is prone to attacks such as sniffing and signal jamming. A suitable solution to protect layer 1 is by using physical security such as physical protection of cable connections.

Information exchange at layer 2 uses physical addresses to identify hosts. This is typically implemented at substations, employing a broadcast mechanism for information delivery. Due to this situation, layer 2 communication is prone to spoofing attacks. Attackers can observe all communication traffic in the network and mimic legitimate traffic to launch a spoofing attack. On the other hand, layer 3 communication works based on IP

addresses. Unlike layer 2, the network layer is a closed-loop communication from source to destination using IP addresses and routing mechanisms. This form of communication is typically used between substations and the control center through a WAN. However, layer 3 is vulnerable to man-in-the-middle attacks. Attackers can perform IP spoofing to mimic legitimate IP addresses for a successful MITM attack. Layer 4 is the transport layer that defines communication protocols. Attacks on this layer mainly exploit protocol operations. For example, TCP sync mechanism can be exploited to launch a DoS sync flood attack. In order to protect layers 2, 3, and 4, security mechanisms such as network firewalls and intrusion detection and prevention systems can be applied.

TCP/IP 4 Layers	OSI 7 Layers	Implementation	Attack Types	Attack Countermeasures
Application	Layer 7: Application (APDU)	Modbus, IEC 61850, IEC 104, DNP3	Application Exploit, SQL Injection	Antivirus, Host-based Firewall, Data Encryption, Secure Coding
	Layer 6: Presentation (PPDU)	Data Formatting, Compression	Phishing	
	Layer 5: Session (SPDU)	Interhost Communication, Authentication, Ports	Session Hijacking	
Transport	Layer 4: Transport (TPDU)	TCP, UDP	Protocol Exploitation, DoS, Reconnaissance	Network Firewall, Intrusion Detection and Prevention System
Network	Layer 3: Network (Packet)	IP Addresses	Man-in-The-Middle Attack	
Network Interface	Layer 2: Data Link (Frame)	MAC Addresses, Ethernet	Spoofing	
	Layer 1: Physical (Bit)	Physical Connection, Cable, Wireless, Signal	Sniffing, Jamming	Physical Security

Figure 2.3: Mapping of OSI layers, cyber attacks and mitigation techniques [3].

For power system communication, typically only layer 7 from the upper layers is used wherein the APDU stores traffic payload. Layers 5 and 6 are typically not used. This is due to the limitation of advanced security implementations in the application layers of power system communications. It is difficult to implement cryptographical techniques to secure power system communications due to the increased latencies. SCADA communication in a power system requires low latency and high rates of data exchange. Hence, communications in the power system are unencrypted and less secure in order to provide a better communication performance. Due to these limitations, cyber security of power system communication has become a vital issue. This chapter discusses secure protocols and security controls for power grids.

There are many standard protocols that have been deployed for power grid operations. However, the implementation of secure communication protocols poses a challenge in OT systems, owing to the high-availability requirement. Consequently, security protocols have been identified to be critical areas requiring significant improvement [146]. We identified five approaches to improve the security of OT communication protocols. The first mechanism is achieved through altering the pre-existing protocols. The second approach involves the integration of established legacy power grid protocols with existing protocols that offer enhanced security measures. The third mechanism is achieved by



developing a brand-new protocol. The fourth mechanism pertains to the enhancement of key exchange, while the fifth mechanism involves the integration of the protocol with blockchain technology. Fig. 2.4 summarizes the secure OT protocol research directions.

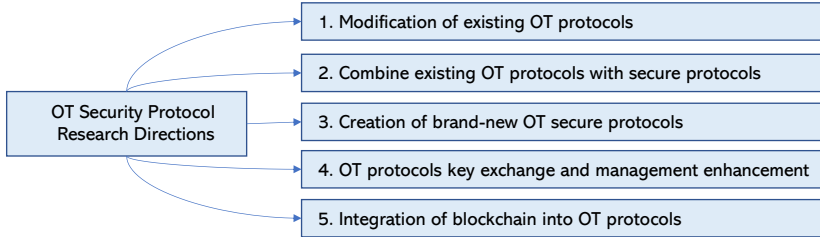


Figure 2.4: Summary of secure protocol research and classification [3].

The first mechanism proposed an alteration of the existing protocols. The authors in [147] carried out a study utilizing formal methods to examine potential authentication vulnerabilities present in DNP3. Upon the identification of vulnerabilities, the authors subsequently suggested the implementation of security enhancements for the DNP3 secure authentication broadcast [49]. The conventional implementation of DNP3 employs a broadcast mechanism for the purpose of verifying the authenticity of communication that is transmitted between the master and remote station. The default broadcast mechanism sends information arbitrarily without a well-defined mechanism. This mechanism may lead to potential vulnerabilities like a man-in-the-middle attack, modification, replay, and injection attacks. The research in [49] proposes a modification of the DNP3 secure authentication broadcast message and checks the validity of the established connection. The proposed solution improves the efficiency and enhances the resiliency of DNP3 broadcast messages against man-in-the-middle attacks. In [148], the authors describe the Secure DNP3 protocol with additional authentication mechanism for enhancing communication integrity. An authentication challenge is issued by the slave when the master station requests a “write” message. The master station sends an authentication response. The slave confirms with acknowledgment and response messages. At this stage, it is inferred that the master station is recognized as a trustworthy and legitimate entity. The authors in [148] also present the security enhancement of ICCP through the utilization of digital certificates to improve communication integrity.

In the second direction, there is already research being done with the intention of using a combination of existing protocols to put the approach into practice. Authors in [149] proposed the utilization of Modbus communication via Transport Layer Security Protocol (TLS) Protocol to create a secure communication channel. The Modbus protocol is a conventional communication standard utilized in power grid systems that lacks security mechanisms. Meanwhile, TLS is considered a broadly adopted mechanism for facilitating secure communication through the use of encrypted data. The proposed mechanism involves the encapsulation and encryption of Modbus information within a TLS packet. The aforementioned mechanism necessitates the process of encapsulating and subsequently de-encapsulating data. Therefore, this approach shows that it is possible to implement power grid communications utilizing pre-existing security protocols.

Instead of modifying existing protocols, the third direction is to create new protocols and standards. An example of a new standard is OPC UA, which replaces the previous versions of OPC through the integration of cryptographic and authentication mechanisms [148]. Another example is IEC 62351 which aims to mitigate cyber security concerns in current protocols via the implementation of cryptographic techniques [123]. Nevertheless, the deployment of cryptographic techniques presents several obstacles. One of the foremost challenges is related to the distribution of keys. Therefore, it comes to the fourth approach using key exchange and management enhancement. Key exchange and management have been identified as a challenge in the SCADA system [150]. Numerous key exchange and management schemes have been suggested to enhance the security of SCADA communication. However, a comprehensive solution to this issue cannot be achieved through a silver bullet solution. The proposed solutions inevitably entail a trade-off between real-time availability and security. The authors in [151] propose a scheme for the pre-distribution of SCADA network keys. The secret key is transmitted over the untrusted network using a pre-distributed matrix-based key. Each device generates unique keys using an algorithm for key generation based on a preliminary matrix reference. This mechanism prevents a man-in-the-middle attack against the key. Unfortunately, if attackers successfully compromise a device, they may still be able to circumvent the secure communication process.

The fifth proposed solution for enhancing security in power grid communications involves the implementation of blockchain technology. Data in the blockchain is stored in the form of a chain of information to preserve integrity [152]. The authors in [153] present diverse potential applications of blockchain technology in the context of power systems. The primary purpose of blockchain technology is to enhance credibility and safeguard the confidentiality of transactions within the energy sector. The proposal of utilizing blockchain technology to enhance the security of message exchange protocols in ICS was proposed in [154]. It is anticipated that blockchain technology will enhance the mechanisms for protocol identification, methods for authentication, and chain of encrypted information. This type of scenario could be appropriate for limited message transmissions. Nevertheless, the communication traffic of power grids primarily comprises telemetry and measurement data that exhibit a high volume of traffic. Therefore, the implementation of blockchain remains challenging and there is currently no practical implementation of blockchain to improve the security of power grid communication protocols.

To summarize, the implementation of the first and second mechanisms represents a straightforward approach to promptly enhance the security of power grid communication protocols. These solutions exhibit a high degree of elegance in addressing deficiencies pertaining to data encryption and authentication in legacy power grid protocols. Nevertheless, these mechanisms may lack reliability due to the absence of inherent security within the protocols. The fourth and fifth mechanisms have the potential to serve as alternative solutions for augmenting the key exchange and authentication aspects of the protocol. Nevertheless, similar to the aforementioned alternatives, these approaches are not inherently incorporated within the existent protocols. Therefore, the third mechanism has the potential to emerge as a viable alternative for enhancing protocol security over a longer time frame. New security standards, e.g., IEC 62351, provide guidelines and requirements for implementing security measures to protect the operation and data exchange within OT systems, including protection against cyber threats and unauthorized access. Unfor-

unately, the implementation of new protocols is a time-intensive process. Moreover, the implementation of new protocols does not always guarantee high reliability and security. For instance, in [155], it was demonstrated that IEC 62351 is still susceptible to resource exhaustion attacks.

## 2.6 NETWORK SECURITY CONTROLS

Security controls are a set of measures and mechanisms that are put in place to ensure the protection of information systems from potential threats, vulnerabilities, and unauthorized access. Security controls have been devised with the purpose of reducing potential hazards and guaranteeing the confidentiality, integrity, and accessibility of both data and resources. This section discusses the state-of-the-art research conducted on network security controls for power grids, which are divided into two categories, i.e., firewalls and IDPS. The summary of network security controls is provided in Table 2.4.

Table 2.4: Summary of network security control applications.

Security Control	Methods	Protocols	References
Firewall	Packet filtering	DNP3	[156]
		Modbus	[157]
	Next Generation Firewall / Deep Packet Inspection	Not specified	[158]
		IEC 104	[124, 159, 160]
		Not specified	[161, 162]
IEC 104		[163, 164]	
IDPS	Signature-based	Modbus	[165–167]
		DNP3	[165, 168]
		Siemens S7	[169]
		IEC 61850	[170–173]
	Anomaly-based and AI-based	IEEE C37.118	[174]
		Not specified	[175–192]
		IEC 104	[193]
		DNP3	[194–196]

### 2.6.1 FIREWALLS

The firewall was initially designed to operate predominantly through conventional IT systems. However, the implementation of a firewall is also a viable measure for enforcing security controls for power grids. In [156], a proposal was made for a Linux-based firewall modification intended for use in power grid applications. The Linux operating system features a firewall application that is configured through the implementation of iptables rules. Iptables enables the user to designate IP address origin and destination, port, and packet type for inclusion in either a blacklist or whitelist reference. Furthermore, the study suggests the utilization of an extra 32 bits of header data derived from the DNP3 protocol. The decision to filter is made using 32 bits of information extracted from DNP3 packets. In [157], another variant with a comparable filtering mechanism was proposed for the Modbus protocol. In general, implementing security measures based on firewalls represents a straightforward approach to safeguarding communication networks for power

grids. The firewall operates on predetermined rules that are hardcoded, and subsequently applies these rules to filter packets accordingly. Unfortunately, a firewall is considered inadequate for dealing with advanced cyber attacks. By utilizing advanced methods of attack, adversaries may circumvent the static firewall rules.

Another type of firewall known as Next-Generation Firewall (NGF) is equipped with the capacity to perform DPI. DPI enables NGF to not only inspect the header information of a packet, but also to inspect the contents and contextual information of the packet payload. Several studies have suggested the utilization of DPI applications for enhancing security measures in power grids. For instance, the DPI application for IEC 104 protocol is researched in [124, 159, 160] and other OT protocols in [158]. NGF exhibits superior performance when compared to traditional packet filtering firewalls. Prior knowledge of the traffic is a prerequisite for NGF to effectively execute traffic classification and filtering. Consequently, NGF exhibits limitations in its ability to identify anomalies from new types of cyber attacks.

## 2.6.2 INTRUSION DETECTION AND PREVENTION SYSTEMS

IDPS is a security mechanism that was specifically developed to identify and counteract any malicious actions or unauthorized entry attempts that may occur within an IT/OT system. The operational mechanism involves the monitoring of network traffic, system events, and user activities with the aim of detecting potential security breaches or policy violations. In general, there exist two primary classifications of IDPS, namely signature-based and anomaly-based.

A signature-based IDPS operates by utilizing a predetermined set of information, i.e., signatures for known cyber attacks, for classifying the network traffic. Numerous studies have been carried out related to the utilization of signature-based IDPS in various power systems-related communication protocols. These include IEC 104 [163, 164], Modbus [165, 166, 197], DNP3 [165, 168], Siemens S7 [169], IEC 61850 [170–173], and IEEE C37.118 [174]. Additionally, certain implementations have been developed for carrying out general OT protocols as described in [161, 162].

An alternative type of IDPS runs through the application of anomaly detection techniques. Rather than depending on pre-defined attack signatures, this approach establishes a standard baseline for typical behavior for the network, systems, and users' activities. The system continuously observes network traffic and system events, seeking out any deviations or anomalies from the normal pattern. An alert is generated if an activity or behavior deviates significantly from what is considered normal. An anomaly-based IDPS is an effective method for detecting previously unseen or zero-day attacks and advances attack techniques.

Statistical analysis, expert systems, and AI are three techniques that can be employed to identify an anomaly. In recent years, the AI-based technique gained more attention. In general, AI-based methods can be subdivided into machine learning and deep learning. Prior studies have proposed the application of machine learning techniques for IDPS in power grids. The vast majority of the research focuses on IDPS in general and does not address any specific OT protocols [175–189, 191, 192]. Some of them also implement anomaly-based IDPS for specific protocols, e.g., IEC 104 [198], IEC 61850 [193].

Deep learning is a subset of machine learning that involves more complex neural

network layers and higher computing demands. Some of the popular deep learning models include CNN, Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Graph Neural Network (GNN). In [195], the authors proposed IDPS based on deep learning to classify DNP3 traffic. The traffic is classified into four categories, i.e., normal, DoS attack, unsolicited attack, and cold restart attack. Another example, CNN-based attack detection for DNP3 protocol was proposed in [194]. More deep learning-based IDPS examples are provided in Table 2.4. Although deep learning requires more computational resources, it outperformed traditional machine learning in terms of performance. As a result, the majority of IDPS research in recent years has focused on applications of deep learning.

## 2.7 CYBER-PHYSICAL POWER SYSTEM CO-SIMULATION AND CYBER RANGE

### 2.7.1 CYBER-PHYSICAL POWER SYSTEM CO-SIMULATION

A power grid is an example of critical infrastructure that requires a high level of availability. Conducting experiments on actual power grids is a challenging task owing to their stringent operational requirements. Therefore, CPS modeling and simulation are essential components of the research. The utilization of CPS modeling and simulation provides significant importance in the domain of power system resilience research. Many survey papers concerning the current state of the art in smart grid modeling can be found in [199–202]. This section focuses on CPS models with cyber security capabilities. The CPS modeling framework comprises two primary components, i.e., the power systems and IT/OT systems. Table 2.5 provides a summary of the CPS model simulators utilized in power systems. There are many power system simulators currently available, including but not limited to Real-Time Digital Simulator (RTDS), OPAL-RT, Typhoon HIL, DigSILENT PowerFactory, GridLab-D, OpenDSS, Siemens PSS/E, Homer, Cymdist, PSAT, and MATPOWER. Numerous communication network simulators are also available, including NS-2, NS-3, OPNET, OMNeT++, NetSim, NeSSi, DeterLab, and Mininet. Therefore, there are numerous potential combinations of power systems and communication network simulators for the purpose of modeling the cyber-physical power system.

Table 2.5: Cyber-physical system models for power systems research.

Cyber-Physical System	Power System Simulator	IT-OT Simulator	Protocols
TASSCS [203]	Software Based	OPNET	DNP3, IEC 61850, OPC UA
SCADASim [204]	Software Based	OMNeT++	DNP3, Modbus
Washington State University [205]	RTDS	Mininet, Core	IEC 61850, Modbus, DNP3
DeterLab [206, 207]	Software Based	Virtual Machine	-
ISAAC [208]	RTDS	Real Hardware	IEC 61850, IEEE C37.118, DNP3
SCEPTRE [209]	PyPower, OpenDSS, Power-World	Virtual Machine	-

According to the state-of-the-art literature review [199–202], RTDS has emerged as the preeminent simulator for power systems. RTDS is a computational tool that enables the simulation of power systems in real-time, allowing for the accurate representation of the

dynamic behavior of these systems in synchronization with the actual system time. This capability is important in the context of testing and validating control systems, protection schemes, and other applications that require timely execution. In the meantime, for IT/OT communication networks, the majority of organizations are moving toward adopting a virtual environment that is based on Virtual Machines. Over the past ten years, there has been a rise in alternative communication network simulators for the CPS model of power grids, including OPNET [210–212], OMNET++ [213, 214], and NS2/NS3 [215]. Nevertheless, the fidelity of these simulators is inferior when contrasted with the virtual environment.

In summary, the communication network simulators utilized for CPS modeling of power grids can be classified into four different categories. They are 1) code/script-based, 2) software-based, 3) virtualization-based, and 4) real hardware implementation. Fig. 2.5 displays the clustering and categorization for each respective category. In Fig. 2.5, each category is evaluated according to its scalability and level of fidelity. It would be preferable for the CPS model to have higher scalability as well as fidelity. The most realistic and least scalable form of simulation is real hardware. The most scalable simulators, meanwhile, are code-based simulators. Code-based simulators enable the simulation of a network at a large scale. However, the code-based needs to specify what constitutes communication and it requires to manually specify each type of communication functionality in the code. Furthermore, unlike in a real system, the communication process is not natural. The subsequent category pertains to simulators that are based on software. The low scalability and low fidelity of these particular simulators leave it a less desirable alternative. Considering the aforementioned factors, it is very likely that the optimal choice for simulation would be based on virtualization.

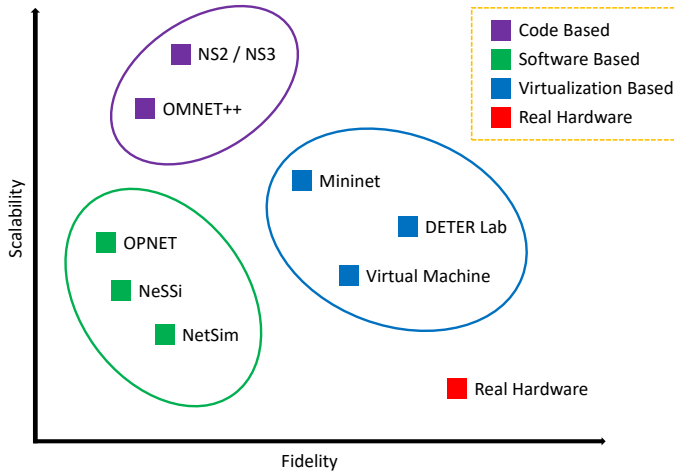


Figure 2.5: Comparison of communication network simulators for CPS modelling [3].

A VM-based simulator is likely to provide an environment that is nearly identical to that of real hardware. It also can be more scalable than real hardware through hardware virtualization techniques using hypervisor. For instance, DETERLab is classified as a VM because it consists of a cluster of VMs. The other option is Mininet, an operating-system

level virtualization, which works based on the Linux namespace over containerization. In contrast to VMs, containers employ virtualization to encapsulate the Operating System (OS) and application dependencies, thereby allowing for the sharing of the host OS kernel across multiple containers. In summary, it can be concluded that the most suitable alternatives for communication network simulation are those based on VM and container technologies, as they offer an optimal equilibrium between scalability and high fidelity.

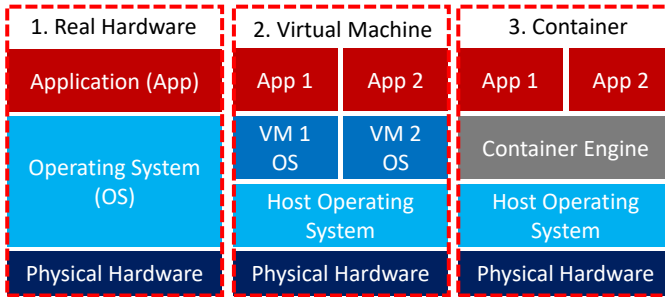


Figure 2.6: Comparison of real hardware, virtual machines, and container-based system [3].

The differences between an application running on actual physical hardware, virtual machine, and containerization are illustrated in Fig. 2.6. When compared to actual hardware, VM allows us to run applications in a more isolated manner within the operating system. This feature enables users to simulate a greater number of virtual environments within the IT/OT network. However, as illustrated in Fig. 2.6, the VM was required to install the guest operating system on top of the host operating system. The scenario involving the stacking of operating systems is known to significantly consume a substantial amount of resources. To address this challenge, operating system level virtualization through containerization applications such as Docker and Linux-based namespace have experienced an increase in popularity in the past few years [216]. One of the reasons for this is that they are able to deploy applications directly on top of the host operating system by utilizing an isolation mechanism, which optimizes the utilization of available resources. In addition, the utilization of containers enables users to emulate a greater number of hosts and larger networks in comparison to VMs. Due to the aforementioned factors, operating system virtualization solutions may become the most suitable network communication simulation tool for modeling power grid CPS. However, the current implementation of power grid CPS models developed through containerization is limited. It is likely that the number of implementations will increase in the near future, which will align with the development of virtualization technology.

Fig. 2.7 depicts an example of CPS co-simulation architecture implemented in Control Room of the Future (CRoF) technology center at Delft University of Technology. It is composed of a simulation of the power system as well as an IT/OT simulation. DiGSILENT PowerFactory and RTDS are used for the simulation of the power system, i.e., IEEE 39-bus. The power system model provides circuit breaker status and measurement data of active and reactive powers, voltages, and currents from busbars, lines, and generators. The implementation of OPC UA facilitates the interfacing of data exchange between power



grids and IT/OT simulation. The implementation of the IT/OT architecture is carried out through the application of Mininet. Each host in the IT/OT network, e.g., merging units, intelligent electronic devices, network switches, routers, databases, etc., are implemented in Mininet using containers. Every container incorporates a tailored application for IT/OT host operations, such as the acquisition and transmission of measurement data, control setpoints, database access, and so forth. The current implementation of CPS comprises of 27 substations and 210 hosts. A unique application has been tailored for each host to replicate the CPS of power grid components. At present, the simulation of all 27 substations runs on 50,000 lines of code on 26 VMs.

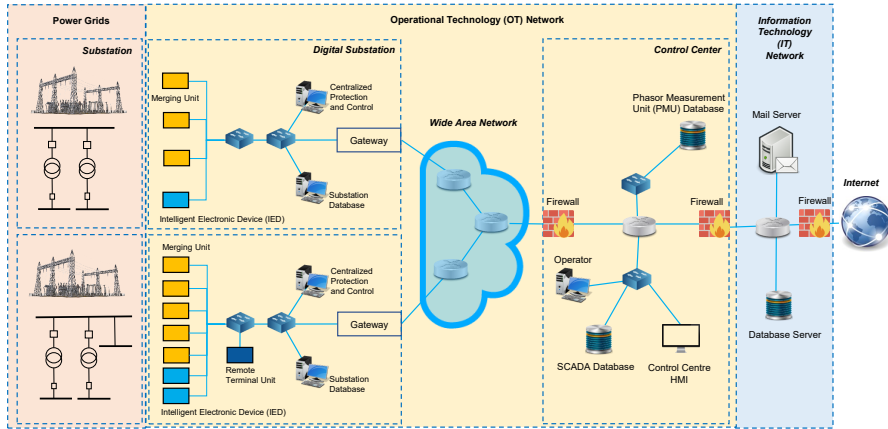


Figure 2.7: CPS architecture in CRoF at TU Delft [3].

In the CPS co-simulation architecture, two distinct types of virtual servers, Windows and Ubuntu, play important roles to ensure the seamless execution of various applications in different operating system. The architecture of these servers, as depicted in Fig. 2.8, demonstrates their specific functions and configurations. Windows VMs are tailored to support Windows-based applications critical to the co-simulation environment, including PowerFactory for power system simulation, the PostgreSQL database for data management, and Power BI for data visualization. On the other hand, Ubuntu virtual machines are dedicated to running open-source tools that are equally essential to the co-simulation process. Specifically, Ubuntu VMs host the OPC UA, which facilitates secure and reliable data exchange between devices, and Mininet, a network emulator used to simulate complex communication networks in the co-simulation environment. This separation of tasks ensures that each VM operates optimally within its respective domain. Fig. 2.9 shows further details the allocation of these virtual machines within the CPS co-simulation, showcasing how the resources are distributed to achieve an integrated and balanced computational environment. This allocation strategy highlights the importance of utilizing specialized virtual environments to address the diverse requirements of power system and IT/OT network simulations, ensuring a comprehensive co-simulation environment.

In the CPS co-simulation model, the virtual network in Mininet is designed to represent specific devices within digital substations, such as IEDs and MUs, with each Mininet host



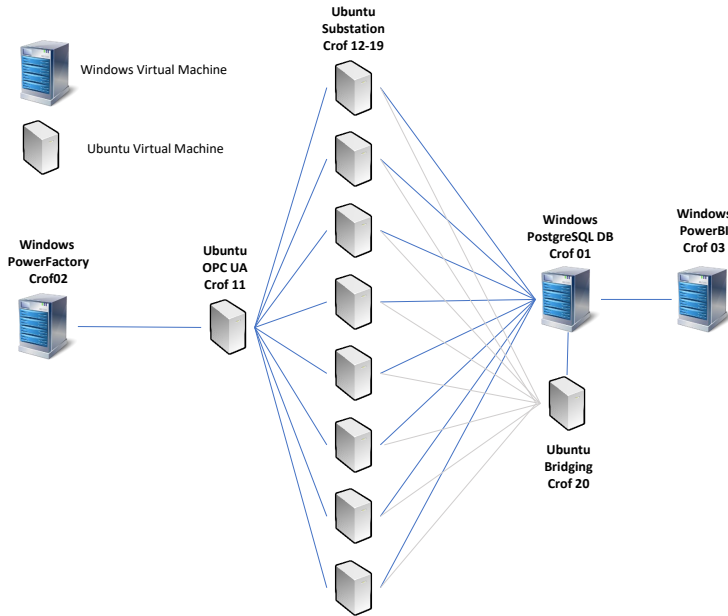


Figure 2.8: CPPS co-simulation architecture with Windows and Ubuntu servers.

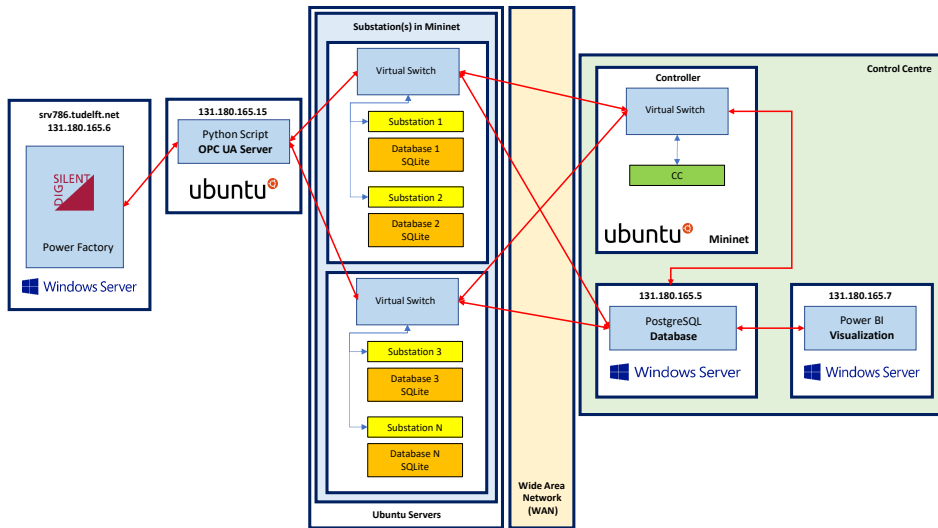


Figure 2.9: Servers allocation for CPPS co-simulation.

simulating a specific device. While Mininet is inherently capable of creating virtualized networks and hosts, additional customization is required to ensure the co-simulation functions effectively. To address this, a customized application is added to each Mininet

host, which employs socket programming to enable communication, acting as both a sender and receiver. This application facilitates seamless data exchange between PowerFactory (representing the power system) and the control center through the Mininet-based digital substation. The data flow within the CPS co-simulation is illustrated in Fig. 2.10, which demonstrates how information is transmitted between components. The exchanged data, represented as  $X$ , includes parameters such as active power, reactive power, generator set points, and circuit breaker statuses. These parameters are categorized into two groups:  $X_{res}$  and  $X_{ctrl}$ .  $X_{res}$  refers to the results, which represent actual values derived from the PowerFactory simulation, while  $X_{ctrl}$  represents control signals that allow changes to be made to parameters within the PowerFactory simulation. This structured approach ensures robust and efficient communication and data management within the co-simulation environment.

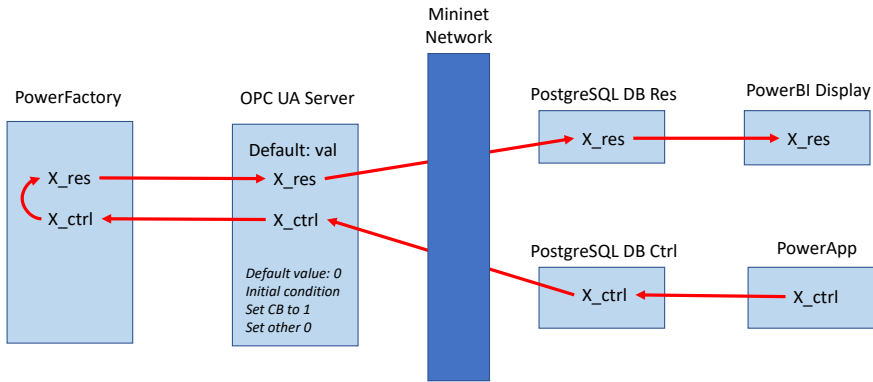


Figure 2.10: Data flow diagram from CPPS co-simulation

## 2.7.2 CYBER RANGE FOR CYBER-PHYSICAL POWER SYSTEM

Cyber ranges have emerged as a prevalent approach for evaluating defense mechanisms and simulating potential attack strategies in the domain of cyber attack and defense simulations [217]. Typically, cyber ranges have been predominantly utilized in the environment of IT systems. In order to align with forthcoming power grid operations, it is essential that CPS models possess cyber range capabilities to enable investigation and assessment of future power grid cyber security.

In accordance with the CPS model depicted in Fig. 2.7, a cyber range is envisioned to be incorporated into CROF technology center at Delft University of Technology. Fig. 2.11 depicts the CPS and cyber range architecture, enabling blue and red teams experiments. The blue team is typically responsible for safeguarding an organization's IT/OT assets and infrastructure, serving as the internal security team or defenders [218]. Their responsibility entails upholding the security posture of both the IT/OT systems and networks. The blue team has several key objectives, e.g., system monitoring, defending, incident response, and cyber security assessment. Meanwhile, the red team plays the offensive or adversarial role in the cyber range exercise [218]. The red team conducts realistic cyber attacks and attempts to

get past the organization's security controls. The main goals of the red team are penetration testing, vulnerability analysis, reporting and providing security recommendations.

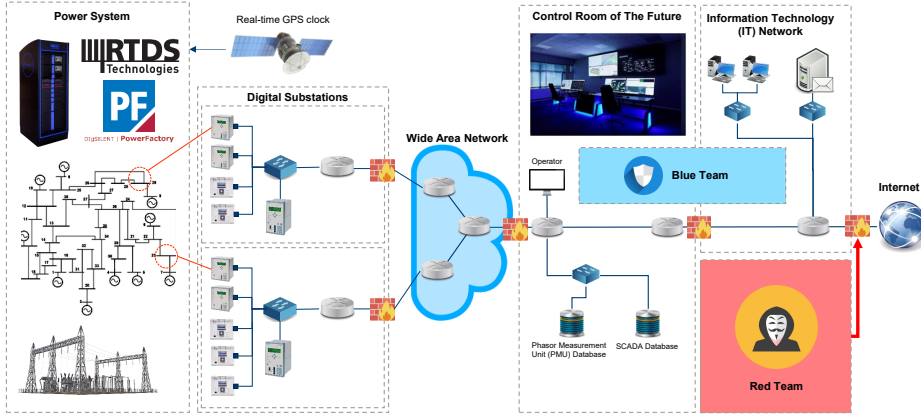


Figure 2.11: CPS and cyber range architecture of CRoF at TU Delft [3].

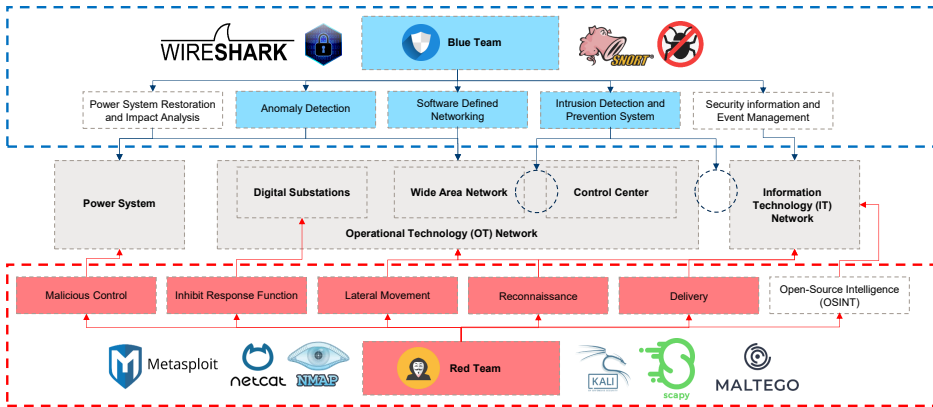


Figure 2.12: Blue and red team tools for power grid IT-OT systems in CRoF at TU Delft [3].

Fig. 2.12 presents the envisioned cyber range architecture with blue and red team's instruments for the power grid IT/OT systems in CRoF technology center at TU Delft. The blue team employs multiple applications to ensure the secure operation of the power system. These applications include Security Information and Event Management (SIEM), intrusion detection and prevention systems, SDN, impact analysis and defense against cascading failures, and power system restoration. Contrariwise, the red team employs cyber attack tools to execute Open Source Intelligence (OSINT), payload delivery, IT/OT reconnaissance, lateral movement, response function inhibition, and malicious control. The red and blue teams are engaged in a cyber range competition to evaluate the capabilities of power system operators and Computer Security Incident Response Team (CSIRT) to mitigate the impact of cyber attacks on power grid operation.

This thesis contributes to the development of CRoF cyber range in the highlighted components in Fig. 2.12. In the red team part, this thesis contributes to the malicious control, inhibit response function, lateral movement, reconnaissance, and payload delivery. In the blue team part, this thesis contributes to the anomaly detection, SDN, and IDPS. The cyber range is implemented within the CPS architecture depicted in Fig. 2.7 using Mininet. In this implementation, each Mininet host represents an OT device in substations and control centers. Within these hosts, individual OT devices are modeled to represent their real-world functionalities, enabling a high-fidelity environment for cyber security experiments. These simulated OT devices can be targeted by red team activities during cyber security exercises. For instance, red team members can perform reconnaissance attacks using tools like Nmap to probe the simulated OT network and identify active hosts, vulnerabilities, and other critical network details. Such activities show the methods employed by adversaries in real-world scenarios, allowing for practical and realistic testing. At the same time, the blue team is equipped with the capability to monitor network traffic within the OT network through SDN features integrated into the Mininet environment. This monitoring is crucial for detecting and analyzing communication traffic, providing insights into potential attacks. The ability to observe traffic patterns and identify anomalies forms the foundation for developing and testing IDPS. By leveraging the SDN-based monitoring capabilities, the cyber range facilitates a comprehensive approach to evaluating and improving the resilience of CPPS against cyber attacks.

## 2.8 CONCLUSION

Power grids are undergoing a fast-paced process of digitalization, opening up the energy system to everyone by means of ITs. However, future grid digitalization will require careful considerations with regard to data privacy and cyber security. It is now well recognized that IT/OT systems are vulnerable to cyber attacks. Hence, cyber resilience requirements of the power grid are more critical than ever before. The complexity of cyber attacks on power systems is likely to increase. To improve the cyber resilience of power grids, it is needed to identify potential threats and IT/OT system vulnerabilities, classify and review major types of cyber attacks on power grids, analyze their impact on system operation and stability, and develop mitigation techniques. Hence, this chapter provided the state-of-the-art and essential knowledge of threats and cyber attacks on power systems. It reviewed major cyber attacks on power grids and industrial control systems and provided a detailed taxonomy of cyber attacks. The most common security controls implemented in power grids include antiviruses, firewalls, network segmentation, and intrusion detection systems. Worryingly, even these security control mechanisms may be outdated or insufficient. Consequently, cyber attacks on power grids exploiting various threat vectors can have a catastrophic impact on system operation. This chapter provided indicative simulation results of such a hypothetical cyber attack scenario. Results show that sophisticated attacks may not only cause loss of load, but also induce cascading failures resulting in a blackout. Therefore, the urgent need of the hour is to develop comprehensive defense, mitigation, and incident response techniques to enhance power grid cyber resilience. Furthermore, this chapter also investigates a cyber-physical power system model. The cyber-physical power system model is composed of co-simulations of the power system as well as an IT/OT simulation. The co-simulation enables cyber attacks simulation and detection in power grids.

## 3

## 3

# ADVANCED PERSISTENT THREAT KILL CHAIN FOR CYBER-PHYSICAL POWER SYSTEMS

*Power systems are undergoing rapid digitalization. This introduces new vulnerabilities and cyber threats in future Cyber-Physical Power Systems (CPPS). Some of the most notable incidents include the cyber attacks on the power grid in Ukraine in 2015, 2016, and 2022, which employed Advanced Persistent Threat (APT) strategies that took several months to reach their objectives and caused power outages. This highlights the urgent need for an in-depth analysis of APTs on CPPS. However, existing frameworks for analyzing cyber attacks, i.e., MITRE ATT&CK ICS and Cyber Kill Chain, have limitations in comprehensively analyzing APTs in CPPS environments. To address this gap, we propose a novel Advanced Cyber-Physical Power System (ACPPS) kill chain framework. The ACPPS kill chain identifies the APT characteristics that are unique to power systems. It defines and examines the cyber-physical APT stages spanning from the initial phases of infiltration to cascading failures and a power system blackout. The proposed ACPPS kill chain is validated with real-world APT attacks on the power grid in Ukraine in 2015 and 2016, and cyber-physical simulations.*

### 3.1 INTRODUCTION

Cyber-Physical Power Systems (CPPS) are critical infrastructures undergoing rapid digitalization. Grid digitalization enhances monitoring and control capabilities, as well as intelligence and advanced analytics. Yet, it also introduces new vulnerabilities and cyber threats, which increase the risk of cyber attacks on future CPPS. For instance, some of the most notable incidents include the cyber attacks on the power grid in Ukraine in 2015 and 2016, which employed complex Advanced Persistent Threat (APT) strategies that took several months to reach their objectives and caused power outages. In December 2015, a coordinated cyber attack affected the Ukrainian power grid making it inoperable for several hours [8]. Adversaries initiated the cyber attack from the Information Technology (IT) network segment. The attack began with a spear phishing email campaign directed at power system operators. Using a weaponized Microsoft Excel file enclosed in the phishing emails, adversaries were able to infect the targets with the BlackEnergy3 malware. From there, they established access to the Operational Technology (OT) network controlling the electricity distribution system. In this instance, the cyber attack was not discovered until the attackers took control of the Supervisory Control And Data Acquisition (SCADA) system via remote desktop sessions and disconnected power lines from the grid. The attack caused power outages that affected seven 110 kV and twenty-three 25 kV substations. This incident is acknowledged as the first cyber attack in the world to cause a power outage. Adversaries carried out a second attack on Ukraine's power grid in 2016 [219], which resulted in a lower degree of success and impact in comparison to the incident that occurred in 2015. However, the attackers were successful in implementing more sophisticated attack methods using malware by exploiting vulnerabilities in the SCADA communication protocols. In October 2022, Sandworm malware disrupted the OT systems in the Ukrainian power grid, leading to a power outage [11]. These cyber attacks brought attention to the fact that the adversaries possessed a comprehensive understanding of the vulnerabilities present in power system OT networks. This awareness implies they have the potential to inflict even more catastrophic impacts in future attacks. Furthermore, the examples serve to demonstrate the pressing nature of cyber attacks on power systems, necessitating in-depth analysis capabilities of APTs on CPPS and proactive detection and mitigation techniques.

Due to the aforementioned cyber incidents, cyber security research for power grids is gaining more attention. Ideally, cyber attacks on power systems are detected and mitigated in their earliest stages of attack to avoid disastrous outcomes. However, most research is focused on detecting the physical impact of cyber attacks on power systems [200, 220] based on anomalies in physical power system measurements, e.g., False Data Injection (FDI). Detection in CPPS based on the physical impact is only valid in the later stages of a cyber attack. In the initial stages, the majority of attacks operate in cyberspace without affecting the physical system. Consequently, the physical impact-based detection is insufficient, and CPPS must incorporate IT-OT anomaly detection. A study in [221] demonstrates the significance of both cyber and physical components for detecting attacks on Cyber-Physical Systems (CPS). The study examines several cyber attack scenarios, including Denial of Services (DoS) and replay attacks. Nevertheless, the attack scenarios do not correspond to APTs on power grids, e.g., cyber attacks on the Ukrainian power grid in 2015, 2016, and 2022.

The cyber attacks in Ukraine indicate the involvement of APTs in targeting the power grid. APT is a type of complex cyber attack that is orchestrated by well-funded and well-organized adversaries to obtain critical information from its target and inflict damage to the infrastructure [222]. The cyber attacks on the Ukrainian power grid demonstrate the APT's real impact on power systems. However, the existing cyber security framework has not yet covered a thorough investigation of APT stages on CPPS and their consequences on power system operation.

In [223] and [224], the authors use a cyber kill chain framework to analyze the stages of cyber attack in power systems, which was originally proposed in [25]. The cyber kill chain was initially proposed to identify stages of cyber attack in the IT system. Therefore, it does not provide any stages related to the power system. In [225], the stages of cyber attacks in power grids were analyzed using MITRE ATT&CK ICS [27]. In [226], the stages of cyber attacks in power grids were analyzed using SANS ICS [28]. The MITRE ATT&CK ICS and SANS ICS frameworks provide a more comprehensive stage analysis compared to the cyber kill chain. These frameworks incorporate stages that are associated with the physical process of the Industrial Control System (ICS). However, both of them do not include the physical process associated with the power system, i.e., cascading failure and point of no return.

According to our literature review in [25–28, 222, 227, 228], there is no framework that provides a comprehensive analysis of APT stages in CPPS. Therefore, in this research work, we provide an in-depth analysis of the capabilities of APTs on CPPS, considering the integration of the IT-OT system and its impact on power system operation. We define the characteristics of APTs on CPPS and propose the first Advanced Cyber-Physical Power System (ACPPS) kill chain framework that defines and examines the cyber-physical APT stages on power grids. It offers comprehensive attack stages for a thorough analysis of APTs on power systems that cause cascading failures and a blackout. Table 3.1 summarizes the comparison of existing frameworks with ACPPS Kill Chain. This table highlights the novelties of the ACPPS kill chain in comparison to other frameworks. The proposed ACPPS kill chain is validated by cyber attack case studies using cyber-physical simulations in the time domain on the IEEE 39-bus test system.

Several frameworks exist to analyze APT stages in IT systems. Currently, the analysis of cyber attacks on power grids is primarily performed using the cyber kill chain [25], CPS kill chain [26], MITRE ATT&CK ICS [27], and SANS ICS [28]. These frameworks are heavily focused on the cyber stages of the attacks and briefly cover their impact. However, they don't cover the impact of cyber attacks on the operation of the physical system. According to our literature review, there is no framework that provides a comprehensive analysis of APT stages in CPPS, including the integrated IT-OT communication networks and impact on power grid operation, affecting the system stability and causing cascading failures and a blackout. Therefore, in this work, we provide in-depth analysis capabilities of APTs on CPPS considering the IT-OT system integration and impact on power system operation. We define the characteristics of APTs on CPPS and propose the first Advanced Cyber-Physical Power System (ACPPS) kill chain framework that defines and examines the cyber-physical APT stages on power grids. It offers comprehensive attack stages for a thorough analysis of APTs on power systems that cause cascading failures and a blackout. The proposed ACPPS kill chain is validated by cyber attack case studies using cyber-physical simulations

in the time domain on the IEEE 39-bus test system.

The key contributions of this research are as follows:

1. We define the characteristics of APTs on cyber-physical power systems, which are different compared to APTs in IT systems and general CPS.
2. We propose the first ACPPS kill chain framework. ACPPS defines and examines the cyber-physical APT stages on power grids that cause cascading failures and a blackout. This novel kill chain framework offers more comprehensive attack stages for a thorough analysis of APTs on power systems and early-stage mitigation compared to the current frameworks reported in the literature [25–28, 222, 227, 228].
3. We conduct a comprehensive analysis of how ACPPS kill chain is applied to analyze real-world cyber attacks. The case studies include the actual attacks on the Ukrainian power grids in 2015, 2016, and 2022 based on publicly available information. In addition, an experimental case study is also presented to provide a comprehensive impact analysis in time domain of how cyber attacks on the IEEE 39-bus test system cause cascading failures and a blackout.

The chapter is structured as follows. Section I is the introduction. Section II describes the characteristics of APTs on CPPS and compares them with the characteristics of APTs in IT and CPS. Section III proposes the ACPPS kill chain framework, and Section IV provides the case study and experimental results. Section V presents the conclusions of the work.



Table 3.1: Comparison of ACPPS stages with other kill chain frameworks.

Stages	Sub-Stages	[226]	[227]	[228]	[229]	[222]	[223]	[224]	ACPPS Kill Chain
A. Attack Preparation	1. External Reconnaissance	✓	✓	✓	✓	✓	✓	✓	✓
	2. Weaponization	✓	✓	✓	✓	✓	✓	✓	✓
	3. Delivery		✓	✓	✓	✓	✓	✓	✓
B. Initial Engagement	4. Exploit	✓	✓	✓	✓	✓	✓	✓	✓
	5. Privilege Escalation	✗	✗	✓	✓	✓	✓	✗	✓
	6. Credential Access	✗	✗	✓	✓	✓	✓	✗	✓
	7. Defense Evasion	✗	✗	✓	✓	✓	✓	✗	✓
	8. Establish Foothold	✗	✓	✓	✓	✗	✓	✓	✓
C. Main Attack Phases	9. Internal Reconnaissance	✗	✗	✗	✓	✓	✓	✓	✓
	10. Lateral Movement	✗	✗	✗	✓	✓	✓	✓	✓
	11. Collection		✗	✗	✓	✗	✓	✓	✓
	12. Exfiltration	✗	✗	✗	✓	✗	✓	✓	✓
D. Physical System Engagement	13. Inhibit Response Function and Impair Process Control	✗	✓	✓	✓	✗	✗	✓	✓
	14. Unauthorized Control on OT System	✗	✓	✓	✗	✗	✗	✓	✓
E. Power System Impacts	15. Cyber Attack Impacts Power System Operation	✗	✗	✗	✗	✗	✗	✗	✓
	16. Induced Power System Events	✗	✗	✗	✗	✗	✗	✗	✓
	17. Operator and Automated Remedial Action	✗	✗	✗	✗	✗	✗	✗	✓
	18. Slow Cascading Failure	✗	✗	✗	✗	✗	✗	✗	✓
	19. Point of No Return	✗	✗	✗	✗	✗	✗	✗	✓
	20. Fast Cascade and System-Wide Collapse	✗	✗	✗	✗	✗	✗	✗	✓
	21. Blackout	✗	✗	✗	✗	✗	✗	✗	✓
F. Social Impacts and Recovery	22. Social Impacts	✗	✗	✗	✗	✗	✗	✗	✓
	23. OT Recovery and Power System Restoration	✗	✗	✗	✗	✗	✗	✗	✓

## 3.2 ADVANCED PERSISTENT THREATS ON CYBER-PHYSICAL POWER SYSTEM

### 3.2.1 APT CHARACTERISTICS

The APT terminology was introduced as a name for intrusion activities of APT1 that was discovered by Mandiant in [230]. This intrusion carries out sophisticated and long-term attacks against a variety of targets, including government agencies, defense contractors, and technology companies, primarily in the United States and Canada. The definition of APT has shifted over time to refer to sophisticated adversaries who target critical information with the intention to covertly profit from the stolen information [231].

APT implements traditional cyber attack techniques in an organized manner. However, compared to traditional cyber attacks, APT is different. In [232], the authors identify different characteristics among them. Traditional attacks are typically conducted by individuals who are not well organized. The motive for traditional attacks is to obtain financial benefits or personal satisfaction. Meanwhile, in APTs, the adversaries are more well-organized and well-resourced. APTs target specific organizations, e.g., governmental institutions and commercial enterprises. In terms of attack techniques and strategies, APTs are more persistent in establishing a foothold in the target.

The National Institute of Standards and Technology (NIST) identifies three characteristics of APTs [233]. First, APTs pursue their goals in a systematic way over a prolonged period of time. Second, APTs are able to adapt to the efforts that defenders make to endure security control measures. And finally, APTs are determined to establish a foothold and maintain the level of interaction with the targeted system to carry out their final objectives.

Table 3.2: Comparison of APT attacks and conventional cyber attacks.

Parameters	APT Attacks	Conventional Cyber Attacks
Actors	Well-organized adversaries	Individual, small group
Attack resources	Resourceful of tools and funding	Limited resources
Motivation	Political, cyber warfare, competition	Financial benefit, hacktivism, personal satisfaction
Target	Governments and enterprise	Mainly individual or organization
Attack technique	Novel/advanced attack techniques	Common attack techniques
Duration	Long term	Single run, short duration
Adaptation	Requires adaptation before final objective	Doesn't require adaptation, objective directly met
Mitigation	Hard to mitigate with security controls	Can be prevented with typical security controls

In [222], the authors identified three requirements to categorize a cyber attack as an APT. The first requirement is that the attack is hard to prevent, even by implementing multiple security controls. The second is that adversaries must adapt to the targeted system over time. If such adaptability is not necessary for the adversaries, it could mean that the defense system is not properly implemented. For targets with advanced security measures, adaptability will allow adversaries to learn about the targeted system's operation, thereby increasing the likelihood of successful attacks. The third requirement is that the adversary exhibits novel attack techniques not commonly implemented in a typical cyber attack. These requirements clearly distinguish APTs from conventional cyber attacks. With these requirements, it will be hard for individual adversaries to perform such sophisticated attacks. Therefore, in general, APTs are conducted by well-organized adversaries with

a considerable number of resources. Table 3.2 summarizes the comparison of APTs and traditional cyber attacks based on [222, 232, 233].

### 3.2.2 APTs in INFORMATION TECHNOLOGY SYSTEMS

IT is a diverse set of technological tools that are used to transmit, store, share, and exchange information. In IT systems, the information is predominantly in the form of digital data. Therefore, APTs in the IT system primarily aim to get access to and exfiltration of digital information. In [231], the author identified that the main objective of the APT attacks is for data exfiltration. Data is a valuable asset for governments and enterprises that potentially can benefit adversaries. Data can be defined as a new form of valuable capital [234]. In [235], the authors identified that data has social and economic value. Therefore, the exfiltration of sensitive data potentially can lead to social and economic impacts.

Table 3.3: Cyber attacks targeting IT systems.

Attack Cases	Year	Impacts
Titan Rain [236]	2003	Espionage cyber attack that led to a data breach
Sykipot Attacks [237]	2006	Sykipot malware stealing intellectual property data
Estonia Attack [238]	2007	DDoS attack led to inaccessible official website
GhostNet [239]	2009	Cyber espionage for stealing confidential information
Shadows [240]	2009	Cyber espionage for stealing confidential information
Operation Aurora [241]	2009	Cyber espionage for stealing confidential information
Night Dragon [242]	2009	Cyber espionage for stealing confidential information
APT1 [231]	2013	Cyber espionage for stealing confidential information
Adobe Data Breach [243]	2013	Data breach on 39 million Adobe software users
Yahoo Data Breach [244]	2013	Data breach on 3 billion Yahoo users
Sony Pictures Hacks [245]	2014	Data breach of Sony Pictures confidential information
OPM Data Breach [246]	2015	Data breach on US Office of Personal Management (OPM)
Uber Data Breach [247]	2016	Data breach on 57 million Uber users
WannaCry [248]	2017	Ransomware encrypted user data causing the data to be inaccessible
Petya/NotPetya [249]	2017	Ransomware encrypted user data causing the data to be inaccessible
Marriott Data Breach [250]	2018	Data breach on Marriott hotel data
RockYou [251]	2021	Data breach on 8.4 billion passwords

APT typically target the IT systems of organizations or individuals that have access to valuable information or resources. Examples of these types of organizations include government agencies, financial institutions, and large corporations. Table 3.3 shows an example of APT attacks targeting IT systems. These attacks are potentially carried out by adversaries that are technologically advanced and have access to significant resources, such as actors representing nation-states or organized criminal groups. The impacts of the attacks include data breaches, inaccessible resources, and system operation disturbance. In summary, the impacts of APTs in IT systems lead to digital or cyber impacts and do not directly affect the physical world.

### 3.2.3 APTs IN CYBER-PHYSICAL SYSTEMS

The CPS terminology refers to a system that can interact with humans through a wide variety of components. This system will possess integrated computational and physical functionalities. CPS is able to interact with the physical world and expand its capabilities through computation, communication, and control [252]. In contrast to conventional IT systems, CPSs exhibit distinct characteristics owing to their ability to interface with the physical world through sensors and actuators [42]. Consequently, the CPS also can affect the physical environment through its actuators.

Table 3.4: Cyber attacks targeting cyber-physical systems.

Attack Cases	Year	Impacts
Siberian Pipeline [253]	1982	Trojan attack led to Siberian pipeline explosion
Salt River Project [253]	1994	Disruption on water treatment facility
Gazprom [253]	1999	Trojan attack led to disruption of gas flow controller
Maroochy Water System [254]	2000	Unintended release of up to 1 million liters of sewage
CSX Transportation [254]	2003	Worm infection led to disruption in railway signaling system
Slammer worm [254]	2003	Worm disabled safety monitoring in nuclear power plant
Zotob Worm [255]	2005	Disruption of manufacturing SCADA system
Stuxnet Worm [256]	2010	Disruption of controllers in Iranian nuclear reactor facility
Steel Mill [257]	2014	Breakdown of the control system in steel mill
Triton [155]	2017	Adversaries took remote control of industrial control system
Colonial Pipeline [258]	2021	Ransomware disrupted the operation of gas pipeline controllers
Ukraine critical infrastructure [259]	2022	Compromised critical infrastructure including nuclear power plant

Considering the aforementioned physical properties of CPS, APT attacks on CPS can impact the physical environment. Table 3.4 summarizes the recorded cyber attacks targeting industrial control systems and their impacts. In general, the impacts can be classified into two categories, i.e., disruption of operation and physical impacts. The differentiation between the impacts of attacks on CPS and IT systems is evident when comparing Table 3.3 and Table 3.4. Any attack on CPS will not only result in the loss of data, but it also has the potential to lead to disastrous events in the real world, e.g., flooding, explosion, or a blackout.

### 3.2.4 APTs IN CYBER-PHYSICAL POWER SYSTEMS

In addition to the aforementioned attacks on CPS, there are APTs that target CPPS. Table 3.5 summarizes the attacks on CPPS. In 2003, there was the first reported cyber attack through a malware infection in the SCADA system of a European power grid operator. This caused a loss of energy management-related functionality in several distribution substations for three days [104]. Amongst all the attacks in Table 3.5, the most notable ones are the cyber attacks in Ukraine in 2015, 2016, and 2022. More detailed discussions on Ukraine's power grid cyber attacks are given in section IV.

As described in Table 3.5, attacks on CPPS can lead to a physical impact. The impacts of cyber incidents on power system operations are classified into four categories, i.e., (i) impact on physical equipment, (ii) impact on the OT communication network, (iii) impact

on energy management system applications, and iv) impact on data/information [10]. The attacks with an impact in the first category are the most severe cyber-physical system attacks, e.g., [8, 219, 260]. This type of attack can directly cause power outages or damage to insulation, power plants, and transformers. The remaining categories mainly affect the monitoring and control capabilities of the power grid, which may indirectly also result in a blackout. Nevertheless, these non-physical impacts are correlated with the initial phase of cyber attacks, which leads to a more severe impact in later stages. Besides the direct impact on the physical and digital elements in power grids, there is also the risk of complex cascading effects on the power system.

Table 3.5: Cyber attacks targeting cyber-physical power systems.

Attack Cases	Year	Impacts
European system operator malware infection [104]	2003	Loss of control in distribution substations for over three days
Aurora experimental cyber attack [260]	2007	Physical damage to power system generator
USB-drive malware in power plant [104]	2012	Three weeks restart delay to power plant
Ukrainian power grid cyber attack 2015 [8]	2015	Power outage affecting 225,000 customers for 6 hours
Ukrainian power grid cyber attack 2016 [219]	2016	200 MW of load was unsupplied
ENTSO-E cyber intrusion [10]	2020	Undisclosed impact
RedEcho malware intrusion [261]	2020	Two hours power outage
ReverseRat malware [262]	2021	Intrusion on power system operator
KA-SAT attack [263]	2022	Disruption on German windfarm satellite communications
Ukrainian power grid cyber attack 2022 [11]	2022	Power outage

### 3.2.5 APT CHARACTERISTICS IN CYBER-PHYSICAL POWER SYSTEMS

In this subsection, we identify the characteristics of APTs targeting CPPS and compare them with the APTs on IT systems and CPSs. Table 3.6 summarizes the characteristics of each category. These characteristics are evaluated based on the APT attack cases in the previous subsections. There are six characteristic categories, i.e., motivation, targeted asset, attack techniques, direct and indirect impacts, and responses. In general, a CPPS has similar characteristics as a CPS, with certain notable differences, i.e., attack techniques, impacts, and response. Our proposed criteria are based on the foundation of power system operation.

An advanced attack technique on CPPS was presented in [219] through the SCADA protocol exploit. Although this attack was unsuccessful, other adversaries have already shown their advanced understanding of CPPS operational communication aspects. However, in this attack, adversaries did not show sufficient knowledge of power system operation. Therefore, the impacts of the attack could be mitigated, and the operator could perform immediate system recovery. In the future, adversaries may have substantial knowledge of the power system operation. Instead of only performing reconnaissance on SCADA communication, adversaries may also gather information from power system operations, for example, by obtaining critical information about power system components, load profiles, and physical vulnerabilities of the power grid. Using this information, adversaries

can optimize their timing and strategies to maximize the attack impact on power system operation, e.g., cause system instability, cascading failures, and a blackout.

Table 3.6: Comparison of APT attacks in IT Systems, General CPS, and CPPS.

	IT	General CPS	CPPS
<b>Motivation</b>	Financial gain, espionage, or hacktivism	Conflict of interest, cyber war	Conflict of interest, cyber-physical war
<b>Targeted assets</b>	Data	Cyber-physical system process	Power system operation
<b>Attack techniques</b>	Phishing, intrusion, malware-based, or ransomware	Specialized techniques based on industrial control operation	Specialized techniques based on power system operation
<b>Direct impacts</b>	Data losses, data breach	System disruption, physical damage	Disruption, power outage, physical damage
<b>Indirect impacts</b>	Financial losses, reputational damage, political implication	Financial losses, reputational damage, political implication	Financial losses, reputational damage, political implication
<b>Response</b>	Patching vulnerabilities, monitoring network traffic, and restoring data from backups	Combination cyber and physical response	Combination cyber and physical response; need to consider power system state for restoration.

Table 3.7: Impacts Comparison of Attacks in IT System, General CPS, and CPPS.

Categories	Impacts	IT	CPS	CPPS
Digital impact	Application / service disruption	✓	✓	✓
	Communication disruption	✓	✓	✓
	Data loss	✓	✓	✓
	Data breach	✓	✓	✓
Physical impact	Physical operation disruption	×	✓	✓
	Physical damage	×	✓	✓
Power system impact	Local power system disruption	×	×	✓
	Wide-area power system instability	×	×	✓
	Cascading failures	×	×	✓
Indirect impact	Financial loss	✓	✓	✓
	Reputation damage	✓	✓	✓
	Political implication	✓	✓	✓

Furthermore, CPPS has more specific impact categories compared to CPS. The comparison of the impact among IT, CPS, and CPPS is summarized in Table 3.7. We identified three impact levels in CPPS, including local power disruption, wide-area power system instability, and wide-area cascading failure. A local outage happens when a particular power grid element is disconnected from the main grid. In general, this does not affect the main power grid. However, during power system instability, an attack may cause a wide-area power system to become unstable for a relatively short period of time. In

this case, there is no significant impact on the power system operation. This impact can be handled through dynamic response from power system operators. Cascading failure impacts occur when the power system cannot recover to its normal operational state. This situation is also indicated by power system instability. However, the remedial actions in the system are not sufficient to tackle this condition. Therefore, the power system will reach a Point of No Return (PNR), followed by cascading failures that lead to a wide-area power outage or even a total blackout [264]. After reaching a PNR, the power system restoration requires considerable time and effort [265]. A more detailed impact of cyber attacks on power systems is discussed in section III.

From the above review, we summarize the following key takeaways from the characteristics of APTs on CPPS:

1. Unlike traditional APTs targeting IT systems, adversaries do not only focus their attention on the cyber components of the CPPS. Adversaries have the potential to target specific components and specific times within power system operation to maximize the impact of their attacks. This can be accomplished by gathering information about critical elements of the power system, power system operational conditions, load profiles, and other relevant information.
2. The impacts on the power system can be further categorized into more detailed and complex stages when compared to those observed in general CPS. These stages include local power outages, power system instability, cascading failures, and a blackout. This classification shows the wider spectrum of attack impact from APTs in CPPS.
3. The restoration process on the power system is complex and cannot be performed by simply restarting the system. The recovery needs to consider many factors, such as black start generation units, load state, generator condition, interconnectors, and the condition of the neighboring power grids. The restoration process for a wide-area power system from a blackout can take several days and even weeks. It is performed incrementally through sequential remedial actions for each electrical substation, power plant, and area/region.

### 3.3 ADVANCED CYBER-PHYSICAL POWER SYSTEM KILL CHAIN

The cyber kill chain is a framework for cyber security investigations and defenses based on intelligence. It is derived from a military model that was initially developed to identify, prepare for the attack, engage, and destroy a target. The cyber kill chain is a method that can be utilized to comprehend better, anticipate, recognize, and fight APTs [25]. The cyber kill chain framework has seven stages, which correspond to the typical phases of a cyber attack. These stages are reconnaissance, weaponization, delivery, exploitation, installation, Command and Control (C2), and actions and objectives. All stages in the cyber kill chain primarily affect the cyber elements of a system and culminate in the action and objective phases. In addition, this framework does not cover the subsequent impact on the physical elements of a system. Consequently, the cyber kill chain framework is inappropriate for identifying APT stages in CPS.

An attempt to cover the physical layer of a cyber-physical system was presented in [26] through the form of a CPS kill chain. This framework is an extension of its predecessor, and it does so by introducing the perturbation of control and physical objectives. However, this framework is lacking in specific stages of the attack and cannot capture the whole process of APT stages. More detailed attack stages on the CPS were proposed by MITRE in [27]. This framework suggested adding three additional categories for the final stages, i.e., impacts, inhibit response function, and inhibit process control. However, within the MITRE framework, the impact part does not cover a comprehensive assessment of the physical system operation in CPPS. Therefore, in this research, we propose an ACPPS kill chain framework to provide a comprehensive definition and analysis of APT attack stages in CPPS.

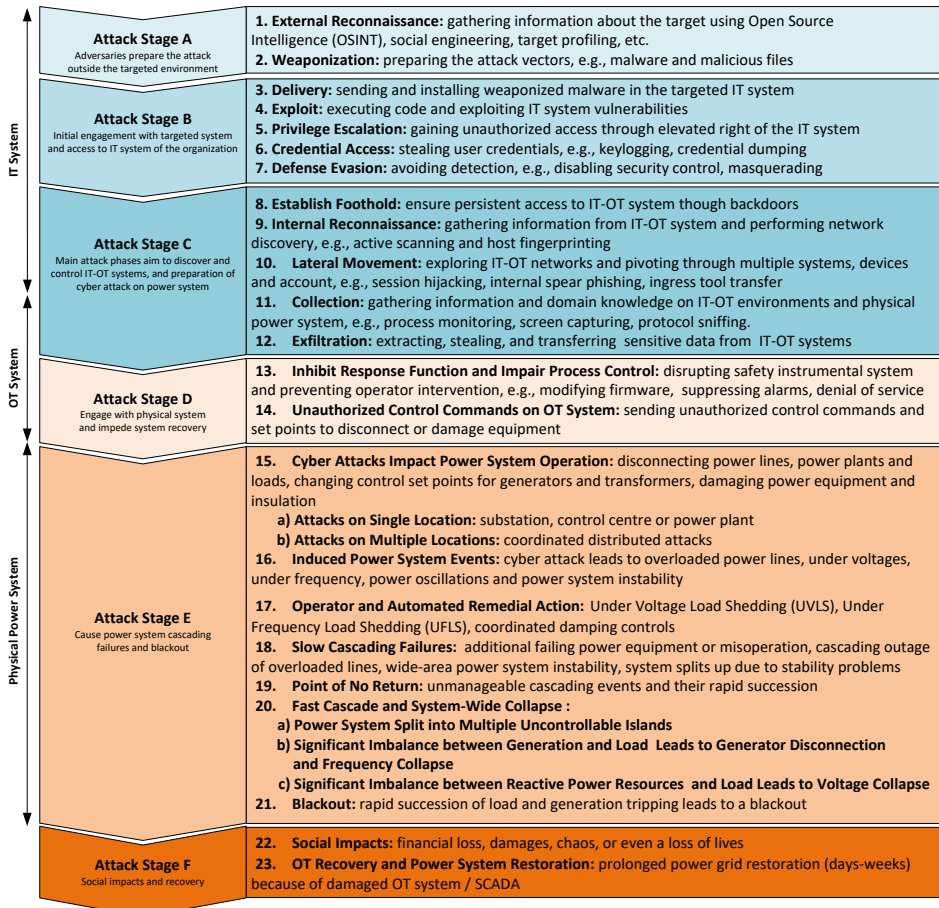


Figure 3.1: Advanced Cyber-Physical power system kill chain framework [4].

Compared to other frameworks, the ACPPS kill chain provides more detailed stages of cyber attacks on power grids. The comparison between the ACPPS stages with other



kill chain frameworks is presented in Table 3.1. We divide the CPPS attack process and impact on power system operation into six stages (A to F). Each attack stage is comprised of a number of sub-stages that are representative of the different attack techniques. Fig. 3.1 provides a summary of the stages and sub-stages involved in the process. The existing stages of a cyber attack that have been identified in other frameworks [25–28, 222, 227, 228] are incorporated into the ACPPS kill chain in stages A, B, C, and D, respectively. The ACPPS kill chain proposes new sub-stages for the impact stages in E and F. A detailed step-by-step breakdown of the ACPPS framework development, including theoretical justifications for each stage, is provided in the following subsection. In the following subsection, the summary of all ACPPS kill chain stages is depicted in Fig. 3.1, and the flowchart illustrates the transitions between stages depicted in Fig. 3.2.

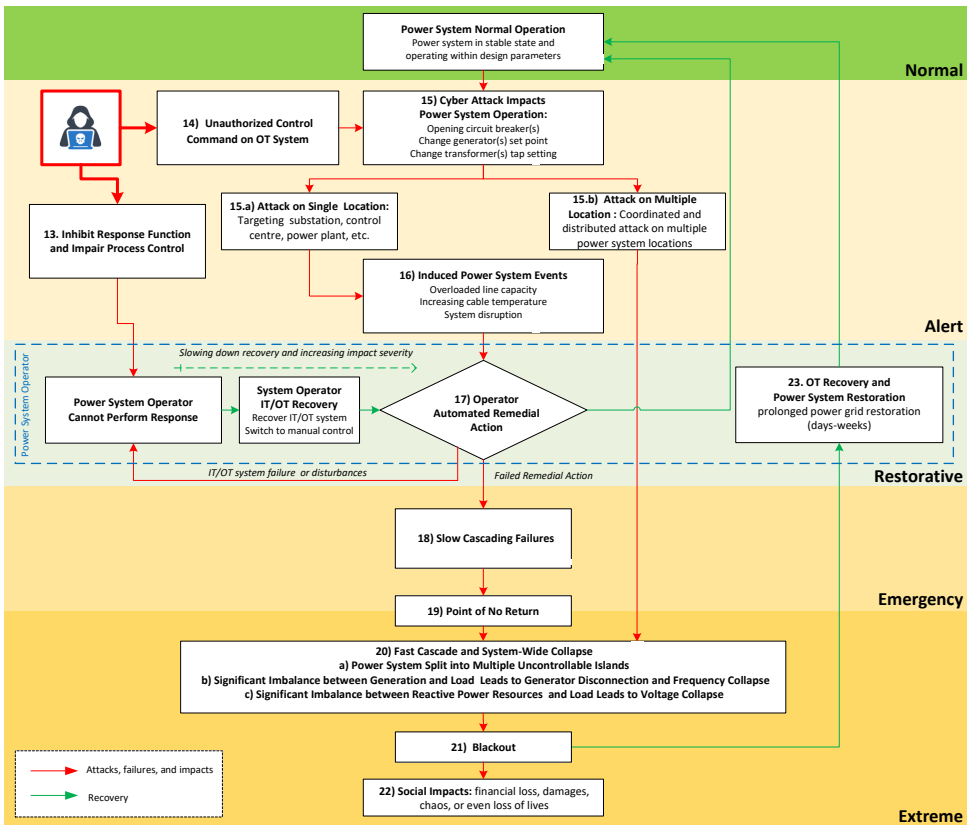


Figure 3.2: Flowchart representing sequences of APTs on power grids according to the ACPPS kill chain [4].

### 3.3.1 ATTACK PREPARATION

The first stage of a cyber attack is attack preparation. It is during this stage that adversaries are preparing for the attack. The attack preparation stage has two sub-stages, namely external reconnaissance and weaponization. In external reconnaissance, the attacker

gathers information about the target system from outside. This information is used to determine target vulnerabilities and strategize the attack on the target. Network scanning, web scraping, social engineering, and other information-gathering tactics are all examples of reconnaissance techniques. The external reconnaissance is also associated with Open Source Intelligence (OSINT), where adversaries collect and analyze open-source data related to the target [266]. Initial information gathering is crucial for adversaries to profile the target and determine the next stages of a cyber attack.

The subsequent phase of preparation is weaponization, in which adversaries prepare the tools for a cyber attack. One of the scenarios for weaponization involves software that has been identified in [267] and [268]. The purpose of weaponized software is to enable an attacker to carry out the actions they desire, such as stealing sensitive information, disrupting operations, or taking control of a target system. This purpose is accomplished by transforming the software into malicious software, also known as malware. Another study identified a variant of weaponization, such as weaponization based on artificial intelligence [269, 270]. The weaponized tools that were prepared in the early stage of the cyber attack are utilized in the later stages of the attack.

### 3.3.2 INITIAL ENGAGEMENT AND IT SYSTEM ACCESS

Fig. 3.1 shows that the second stage in the ACPPS kill chain is the initial attack. It happens when the adversaries perform initial engagement with the targeted system and access the IT system of the organization. In comparison to the previous stage, adversaries in this stage begin to engage and interact with the target that is initiated through delivery. The adversaries prepare the weaponized file and then deliver it to the target. Phishing is the most prevalent technique for cyber attack delivery. The objective of a phishing attack is to convince the victim to open a malicious email or website that delivers the weaponized payload. Phishing attacks frequently employ social engineering techniques to make the message appear legitimate and try to convince the target to click on the malicious link. A phishing attack can be broken down into a variety of different techniques, such as email, clickjacking, cross-site scripting (XSS), drive-by-download, JavaScript obfuscation, and malicious advertisement [65]. Phishing attack primarily focuses on taking advantage of the target's lack of awareness in order to successfully install malicious payloads on the system. This strategy serves as the cyber attack's entry point to infiltrate the targeted system.

Following successful infiltration through the delivery sub-stage, the exploits sub-stage takes advantage of a vulnerability in software, hardware, or a system to execute malicious code and gain unauthorized access. The information on exploits can be obtained from publicly accessible data sources, including the exploit database [271] and the MITRE common vulnerability exposure [272]. An adversary may carry out an exploit with a focus on a zero-day vulnerability in order to increase the likelihood of a successful attack. Zero-day vulnerability is a type of vulnerability that the vendor is either unaware of or has not yet patched [273]. Vulnerability is a crucial component in the process of performing system exploits, as it determines the methodology that is used to carry out the exploit. For example, Ripple20 vulnerability described in CVE-2020 11896 allows for the remote execution of code in SCADA devices [143].

There are two sub-stages that adversaries use to obtain unauthorized authority, i.e., privilege escalation and credential access. Privilege escalation refers to the process of

acquiring higher levels of access authority or permissions on a system than those initially granted to the users. This can be accomplished through the use of a variety of methods, such as exploiting software vulnerabilities and hooking. In order to elevate privileges, adversaries exploit software vulnerabilities by taking advantage of a programming error in a program, service, or operating system. Meanwhile, hooking is a technique used by adversaries to take advantage of Application Programming Interface (API) functions, which allows them to elevate privileges and redirect calls for execution. Normally, security permission levels should restrict those malicious activities. Due to privilege escalation, however, adversaries can circumvent these restrictions. Once an attacker has gained access to higher levels of the system, they have a greater chance of being able to perform unauthorized actions, steal sensitive data, or cause damage to the system. Apart from privilege escalation, credential access aims to gain access to a legitimate username, password, or other authentication credentials. Credentials can be retrieved via brute force, password cracking, exploiting vulnerabilities, etc. With privilege escalation and credential access, adversaries can acquire administrative control over a targeted system.

Defense evasion refers to a tactic employed by an adversary to circumvent or undermine security measures in the interest of avoiding detection or analysis. This can be accomplished through a variety of methods, including the obfuscation of malware code, the utilization of encryption in order to conceal malicious traffic, or the manipulation of security tools and monitoring systems. Defense evasion is a core pillar of APTs and other types of sophisticated cyber attacks. Through the use of defense evasion, it is possible for adversaries to go unnoticed by the application of a security system, such as a firewall or an intrusion detection and prevention system.

### 3.3.3 MAIN CYBER ATTACK ON IT-OT SYSTEMS

The aforementioned initial attack phase is followed by the main attack phase, during which the adversaries gain substantial authority to accomplish their objectives. In the main attack stage, the sub-stages involve establishing a foothold, internal reconnaissance, lateral movement, collecting information, command and control settings, and exfiltration.

During the process of establishing a foothold, the adversaries install a backdoor so that they can have persistent and sustained access to the target. Techniques for establishing a foothold include any access or configuration changes made to protect their illegal activity and maintain a foothold on systems. This may involve replacing or hijacking legitimate code, firmware, and other system files, as well as modifying the system's boot process. A backdoor serves as an entry point in a compromised system that enables adversaries to bypass security controls. An adversary with a backdoor can perform internal reconnaissance or discovery on the targeted system. The cyber attack techniques typically implemented in this sub-stage are network sniffing and enumeration, operating system fingerprinting, and remote system discovery. Furthermore, internal reconnaissance gives potential attackers the chance to gather information about the IT-OT system's behavior, such as its network topology, security protection, running applications, and so on, in order to formulate their final attack strategy.

Once an attacker has established a foothold in a network, they may attempt to gain broader access to other components in the IT-OT system. To achieve this objective, adversaries engage in lateral movement sub-stage. In the context of a cyber attack, the term

lateral movement refers to the process by which an adversary moves from one compromised system to another within a network. This sub-stage is used by the adversary to pivot to the next point in the environment, thereby positioning themselves closer to the ultimate objective. From lateral movement and internal, adversaries identify the location of the final stage of cyber attack.

During the main stage of the attack, the adversary may conduct information collection and exfiltration. Collection refers to the methods that adversaries employ to obtain domain knowledge from the targeted IT-OT system. The techniques implemented in collection sub-stages are process monitoring, screen capturing, and protocol sniffing. The collection is critical for the planning and execution of attacks in CPPS. Exfiltration is the process of performing an unauthorized transfer of data from a compromised system or network to an external location controlled by adversaries. Meanwhile, information collection aims to obtain valuable information from the system that is being targeted. These two steps have a strong connection to data breaches because they expose valuable information to third parties who are not authorized [274]. This type of attack becomes the ultimate objective of a typical high-profile cyber attack targeting businesses and government institutions. Nevertheless, this type of attack does not cause any direct physical impact.

### 3.3.4 ENGAGEMENT WITH PHYSICAL SYSTEM AND SYSTEM RECOVERY IMPEDIMENT

In stage D, adversaries start to perform direct engagement with a physical system from the OT system. In this stage, adversaries inhibit response functions and impair process control. The first sub-stage aims to impede system recovery before executing the final attack. The potential techniques implemented in this sub-stage are firmware modification, alarm suppression, blocking legitimate communication, data destruction, force system restart or shutdown, and DoS.

Finally, adversaries execute the final stage of the cyber attack through unauthorized control commands on the OT system. In these sub-stages, adversaries are granted the ability to exert control over the system by sending the specified control commands. This attack has the potential to have repercussions for the physical system. For instance, in [260], the Aurora experiment demonstrated how a cyber attack could be used to maliciously control a generator. The experiment demonstrates that a 2 MW synchronous generator can be physically destroyed by malicious control. Another illustration of command and control is the attack on the Ukrainian power grid, in which the adversaries took control of the SCADA interface and opened circuit breakers for power lines [8]. The command and control stages of cyber attacks rarely happen, but when they do, they have the potential to cause significantly more damage than other types of non-physical cyber attacks.

### 3.3.5 POWER SYSTEM CASCADING FAILURES AND BLACKOUTS

Adverse, unmanaged power system events and disturbances have the potential to result in cascading failures, ultimately leading to the collapse of the entire power grid. The root causes of such events are 1) deterioration and aging of power system equipment, 2) insufficient time to take decisive and adequate corrective actions, and 3) a lack of adequate automated and coordinated controls to take swift and decisive measures [275]. The occurrence of cyber attacks on power grids [8, 219, 276] has raised concerns about the

potential for such attacks to instigate the final three root causes mentioned earlier. In [275], the authors comprehensively investigated the analysis of power system impacts caused by single or multiple events. However, cyber attack factors were not incorporated into the events themselves. Therefore, in this subsection, the ACPPS kill chain incorporates an analysis of the possible effects of a cyber attack on the power system. The ACPPS kill chain classifies the impact stages into seven sub-stages, namely (i) cyber attacks impact on power system operations, (ii) induced power system events, (iii) operator and automated remedial actions, (iv) slow cascading failures, (v) point of no return, (vi) fast cascade and power system-wide collapse, and (vii) blackout. Fig. 3.2 presents a comprehensive overview of the flowchart depicting the various sub-stages involved in assessing the impacts of cyber attacks on the power system.

### **CYBER ATTACKS IMPACT ON POWER SYSTEM OPERATIONS**

After adversaries have targeted the physical power system, unauthorized control commands may affect the physical components. The physical effects include but are not limited to, the disconnection of power lines, power plants, and loads, the modification of generator and transformer control set points, and the damage to power equipment and insulation. The disconnection of the power system components can cause widespread power outages and disruptions in the electrical supply. By manipulating the control set points for generators and transformers, adversaries can disrupt the stability and control of the power system, potentially causing voltage and frequency fluctuations. Cyber attacks may involve the destruction of power equipment and insulation. Attacks on critical infrastructure components, such as transformers and generators, can compromise their integrity and result in costly physical damage that necessitates repairs or replacements. Insulation damage can cause faults and short circuits, exacerbating power system disruption.

The cyber attack on power systems may vary with single or multiple targeted locations. Attacks on a single location target a specific facility within the power system infrastructure. For example, an attacker may focus on a substation, control center, or power plant. Meanwhile, attacks on multiple locations are coordinated and distributed attacks that aim to target multiple facilities within the power system simultaneously. The attackers orchestrate a synchronized assault on various points of the infrastructure to maximize the impact and spread the disruption across a wider area. The flowchart of Fig. 3.2 compares single (15.a) and multiple (15.b) location attacks on the power system. As depicted in Fig. 3.2, multiple location attacks can directly cause wide-area system collapse.

### **INDUCED POWER SYSTEM EVENTS**

The cyber attacks impact on power system operation has the potential to induce subsequent power system events. Initial impacts initiate a chain of undesirable events that disrupt the power system's normal operation and stability. The induced power system events include overloaded power lines, under voltages, under frequency, power oscillations, and power system instability. Overloaded power lines can cause thermal stress, increased line losses, and even transmission infrastructure damage or failure. Under voltages occur when the power system's voltage levels fall below the normal operating range. Under voltages can result in issues such as decreased efficiency of electrical equipment, malfunctioning of sensitive electronic devices, and diminished performance of motors and other loads. Under frequency events refer to instances when the frequency of the Alternating Current (AC)

power system drops below the standard operating frequency. Under-frequency conditions can impact power system stability and functionality, and affect the performance of time-sensitive equipment such as motors and generators. Extended under-frequency events can result in cascading failures and widespread power outages if not promptly resolved. Power oscillations are uncontrolled and irregular fluctuations in power flows within a system. The oscillations may cause system destabilization, equipment strain, and voltage and frequency instabilities. Power oscillations can degrade power quality and reliability. Overall, induced power system events resulting from a cyber attack can have severe consequences for the stability, reliability, and safety of the power grid.

### **OPERATOR AND AUTOMATED REMEDIAL ACTIONS**

When a power system is subjected to a major disturbance, operator and automated remedial actions are usually undertaken to mitigate the impact of the event. These actions aim to maintain system stability, prevent widespread outages, and minimize the impact of disruptive events. One method of remedial action is Under Voltage Load Shedding (UVLS). When the voltage levels in the power system drop below a certain threshold, UVLS is employed to shed or disconnect certain loads to alleviate the strain on the system. By shedding non-critical loads, UVLS helps to restore and maintain voltage levels within an acceptable range. This action prevents voltage collapse, reduces the risk of equipment damage, and ensures a stable and reliable power supply. Under Frequency Load Shedding (UFLS) is another similar action. In the event of a decrease in the frequency of the power system, UFLS is activated to shed predetermined loads. By shedding certain loads, UFLS reduces the demand on the system, allowing it to recover and stabilize frequency levels. UFLS helps to prevent frequency collapse, maintain system integrity, and avoid widespread power outages. Coordinated damping controls are a set of automated measures employed to dampen power oscillations and stabilize the power system. Power oscillations can occur due to disturbances or imbalances within the grid. Coordinated damping controls utilize various techniques, such as adjusting generator excitation, modifying power system stabilizer settings, or implementing supplementary control signals. These actions aim to counteract power oscillations, enhance system stability, and improve the dynamic response of the power grid. By implementing operator and automated remedial actions such as UVLS, UFLS, and coordinated damping controls, power system operators can respond to critical events or disturbances. Nevertheless, as illustrated in Fig. 3.2, these corrective measures do not always result in favorable outcomes. Adversaries have the potential to impede remedial action, resulting in elevated undesirable consequences for the power system.

Furthermore, electrical power systems are protected by a variety of automated protection schemes. These protection schemes are implemented on critical components of the power grid, such as generators, transformers, busbars, and power lines. In case of an event such as a short circuit, these schemes aim to isolate the affected component or area on time, thereby safeguarding and ensuring that the equipment will not be stressed or destroyed and the rest of the system will not destabilize. To achieve these goals, a variety of protection schemes are implemented and need to be coordinated. This difficult task is essential as each part of the system needs to be covered by multiple protection schemes. This is done to ensure that multiple operational aspects are addressed, e.g., frequency and voltage protection for generators, and to provide proper coverage in case of maloperation



of one device, e.g., distance protection for power lines. Maloperation or improper tuning of the protection relays is also an issue that can occur. Potential maloperation of relays needs to be considered in the protection coordination design. However, as shown in past blackout in North America in 2003 [277], if the settings are not properly tuned, the protection can be triggered to trip power lines and generators.

### **SLOW CASCADING FAILURES**

When the remedial action fails, power system events from the previous stage enter the emergency state and lead to slow cascading failure. Cascading failures take place as a consequence of vulnerabilities in interconnected infrastructures [278]. This is a direct consequence of the complex system interactions and interdependencies in electrical power grids. Slow cascading failures are related to additional failing power equipment or maloperation, cascading outage of overloaded lines, wide-area power system instability, system splits up due to stability problems. When there are numerous instances of power equipment failures or operational mistakes within a power system, slow cascading failures may occur. These malfunctions or failures may involve switches, transformers, generators, or other crucial components. Maloperation, such as incorrect settings or human errors in controlling the power system, can also contribute to cascading failures. In the North American blackout mentioned above, the distance protection of power lines was tripped, as the low voltages and overload currents were confused for uncleared fault. This was done as the measured impedance of certain transmission lines fell in Zone 3 of the distance relay. As a result, the disconnection of additional elements led to the continuation of the slow cascading failure propagation. Slow cascading failures can result in the instability of a wide-area power system. This instability refers to a loss of balance between power generation and consumption, which results in voltage and frequency deviations that exceed acceptable limits. In some instances, the instability caused by slow cascading failures can result in the disconnection of power system regions. This occurs when network stability issues become severe enough to cause a separation between network segments. The split can disconnect certain regions from the power supply and disrupt the system's overall operation.

### **POINT OF NO RETURN (PNR)**

The power system has the potential to transition from an emergency state to an extreme state, which occurs upon surpassing the PNR. The PNR represents a crucial point in a cascading failure scenario that occurs within a power system. The phenomenon of PNR encompasses a range of intricate and dynamic events including but not limited to 1) the overloading of transmission lines, 2) disconnections of generators, 3) variations in frequency, 4) instabilities in voltage, and 5) loss of synchronism [74]. The propagation of cascading failures is significantly influenced by each of these distinct physical phenomena. At this stage, the situation becomes increasingly difficult to manage, and the sequence of events rapidly accelerates beyond control. After the PNR, the power system enters into a fast cascade and system-wide collapse. In a cascading failure, the power system may divide into uncontrollable islands. This results in system fragmentation and the emergence of isolated regions or islands with an unreliable power supply. The split may result from significant disruptions and failures that impede regular electricity transmission throughout the network.

### FAST CASCADE AND POWER SYSTEM-WIDE COLLAPSE

A fast cascade is associated with a significant imbalance between the power generation capacity and the power demand in the system. Insufficient generation capacity can cause generator overload, instability, and disconnection from the system due to high demand. The disconnection may lead to an imbalance and subsequent frequency collapse, characterized by a rapid drop in system frequency beyond the acceptable operating range. Fast cascades are also linked to the inability of a stressed power system to maintain its voltage levels in the safety margins. Reactive power is essential for voltage stability in power systems. An imbalance between reactive power resources and demand, as well as limited capability of transferring the necessary reactive power to the loads, can result in voltage collapse in multiple areas of the system. Soon after the fast cascade, a power system suffers a blackout.

### BLACKOUT

A blackout is a complete and unexpected loss of electrical power over a large area, typically affecting many customers or an entire region. In [279], the authors identified the preliminary stages before the blackout, i.e., the contingency condition, power system problems, protection system trips, and system separation. There are many major power system blackouts happened in the past. For example, the blackout in Italy 2003 [280], North America 2003 [277], Europe 2006 [281], India 2012 [282], Türkiye 2015 [283], Ukraine 2015 [8], and United Kingdom 2019 [284]. Those blackouts were triggered by various factors, and only one of them was triggered by a cyber attack. In [285], the authors identified that cyber attacks can accelerate cascading failures and blackouts. Therefore, it is necessary to identify cyber attacks on power grids in the early stage to avoid a blackout and more severe impacts.

### 3.3.6 SOCIAL IMPACTS AND RESTORATION

APTs have the potential to bring a variety of consequences, including economic losses [286]. In addition, cyber attacks on a physical system, i.e., attacks on electrical power grids, potentially can have more severe repercussions. The reason for this is because electrical power grids are considered to be a part of the nation's critical infrastructure. Power system blackouts can lead to wide-area of social consequences, including financial loss, damages, chaos, or even a loss of lives. Power supplies are essential to the functioning of fundamental necessities, such as hospitals, transportation networks, and communication networks [287]. Disruptions of the power grid can have severe consequences for general safety, public health, and the economy. According to research, power outages can have a societal cascading impact, i.e., an increase in the mortality rate [288], disruptions in transportation [289], and an impact on the economy [290]. Furthermore, these indirect impacts have the potential to be politically utilized and become the objective of cyber warfare, as demonstrated by the attack on the Iranian nuclear facility [256] and the conflict between Ukraine and Russia [291].

Upon the culmination of a cyber attack, the system operator endeavors to restore the power system to its normal operational state through OT recovery and power system restoration. The primary objective of OT recovery is to reinstate the OT infrastructure responsible for monitoring and controlling the power system. The process entails the identification and removal of any malicious software, the repair or substitution of compro-



mised hardware, and the reinstatement of the soundness and operability of the OT systems. This process requires a thorough investigation to understand the extent of the attack, the vulnerabilities that were exploited, and the impact on the power system's operational capabilities. Power system restoration involves bringing the entire power system back to its normal functioning state after a blackout. It is a complex optimization problem, which involves advanced coordination across the affected area, as the grid is gradually restored. As power plants are reconnected to the grid, and the loads need to be gradually restored, the communication network plays an important role. In the case of cyber attacks, the validity of the communication system may be compromised. As a result, restoration actions may be further hindered by unresponsive control systems. Due to power system complexity [292], to fully recover the power system requires a complex restoration process within days or weeks. For example, in the North America 2003 blackout that affected 50 million customers, the full restoration process took 48 hours [277]. Another blackout in Italy in 2003 affected 60 million customers and took 12 hours to fully restore the power system [280]. Both of the blackouts were caused by a disruption of the physical power system. The Ukrainian power grids blackout in 2015 was triggered by a cyber attack. This blackout affected 225,000 customers, and took 6 hours to restore the power grids [8]. Additionally, in the Ukrainian power grid attack in 2016, the malware that was utilized contained a module able to launch DoS attack on the IEDs of the targeted substation [219]. Although unsuccessful, the attack aimed to make the IEDs unresponsive to remote commands from the system operators, and could delay the restoration of the affected system.

### 3.4 ADVANCED PERSISTENT THREATS ON POWER GRIDS CASE STUDIES

The digitalization of the electrical power grid has simultaneously introduced the possibility of cyber attacks on electrical power grids as an imminent threat. The repercussions of such sophisticated forms of cyber attack, like the APTs, are to be worried about. They are high-impact, low-frequency events with a wide range of ramifications. It should be noted, however, that only a small number of actual cyber attacks have been recorded as deliberately targeting power grids. However, these attacks have shown that they are capable, and they have given us a glimpse of the potentially disastrous consequences.

In this section, three case studies of cyber attack on power grids, including the real cyber attacks in Ukraine 2015 and 2016, and a hypothetical cyber attack scenario. The analysis of the real cyber attacks is based on the reports in [8, 11, 219]. Those reports analyzed the stages of cyber attacks in Ukrainian power grids in 2015 and 2016 using the cyber kill chain [25] and SANS Industrial Control System Cyber Kill Chain framework [28]. Compared to the other reports, our survey provides a more detailed cyber attack stage identification and analysis based on the ACPPS kill chain.

In addition to that, we also present a hypothetical cyber attack scenario that was experimented with using a co-simulation testbed of CPPS. The experimental cyber attack is needed because of the limited available information on the physical impact of Ukraine's power grid attacks in 2015 and 2016. Therefore, using the experimental scenarios, we simulate a more detailed physical impact using CPPS co-simulation. To simulate the cyber attack scenarios, we implement CPPS co-simulation, which consists of power system simulation

and OT network simulation environment. The power system simulation is implemented using DIgSILENT PowerFactory, and the OT network simulation is implemented using Mininet. Both DIgSILENT PowerFactory and Mininet have been recognized as prominent power system co-simulation tools. These tools have been implemented in many CPPS testbeds for industrial and academic purposes [293]. The DIgSILENT PowerFactory is capable of simulating the power system in real-time using the Root Mean Square (RMS) dynamic model. Meanwhile, the Mininet implements operating-system-level virtualization, which allows the implementation of real communication protocols and cyber attacks. A more detailed discussion of all cyber attack case studies is provided in the following sub-section.

### 3.4.1 UKRAINIAN POWER GRID CYBER ATTACK 2015

On December 23, 2015, at 15:30 local time, there was a cyber attack on the Ukrainian power grids. This was the first instance of a cyber attack on power systems that was reported, and it resulted in a power outage. The attackers were successful in compromising three different Distribution System Operators (DSOs) SCADA systems, which allowed them to disconnect a total of seven 110 kV substations and twenty-three 35 kV substations from the distribution network. The attack was successful and resulted in a power outage that impacted a total of 225,000 customers [8].

Fig. 3.3 depicts the system that governs the operation of power grids. This system includes the IT network, the OT network, and the physical power grids. The IT network is used to assist the operation of the power grid in various ways, including the administration of resources and assets and office operations. From a topological perspective, the IT network is connected to both the Internet and the OT network. On the other hand, they are isolated from one another by means of network segmentation and other forms of security control, such as a firewall. The goal of network segmentations and security controls is to strengthen the network's security measures by separating the different segments of the network. In the OT network, digital substations are responsible for collecting measurement data from the substation bays and station control systems. This data includes the phase angle, active and reactive power, as well as voltage and current magnitude. Afterward, the measurements are transmitted to the control center in order to provide centralized wide-area monitoring. Despite the fact that various cyber security safeguards have been included in the operation of power grids, the cyber attack that happened in 2015 in the Ukrainian power grid highlighted that the system remains vulnerable and can be compromised.

Table 3.8 gives an overview of the method that was taken during a cyber attack in Ukraine in 2015. Accompanying Table 3.8, Fig. 3.3 depicts the locations of the different stages of a cyber attack, each of which is denoted by a number within a red circle. The adversaries apply APT attack strategies to achieve their goals within a few months prior to executing their true objective. The adversaries first gain access to the system through the office network and then go on to the control center and the substations with the intention of triggering a blackout. The majority of the time, the adversary's operations remain undiscovered while operating in a stealthy mode until the last phases of attacks are carried out.

Before launching the attack, in the preparation stage, adversaries gather the profile of the target to prepare spear phishing. Spear phishing is accompanied by substantial

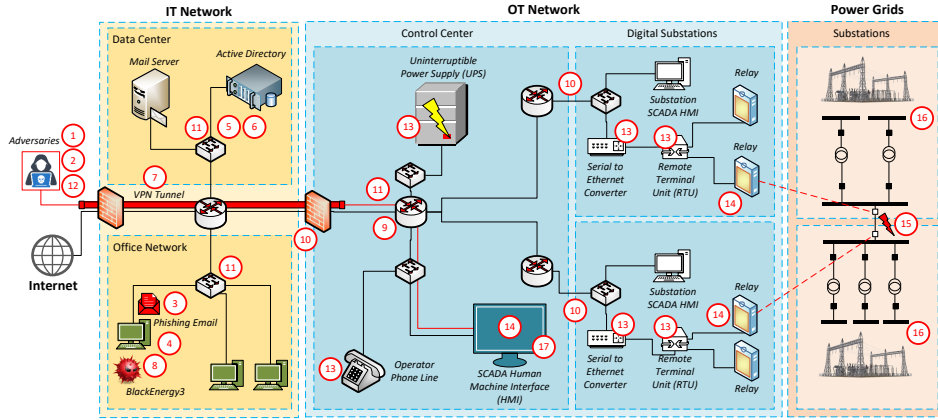


Figure 3.3: Cyber attack on Ukrainian power grids 2015 [4].

Table 3.8: Advanced Cyber-physical system kill chain stages in Ukraine cyber attack 2015.

Stages	Sub-stages	Ukrainian power grids cyber attack in 2015 processes
A	1. External Reconnaissance	Target profiling and gathering information about Ukrainian DSOs
	2. Weaponization	Prepare BlackEnergy3 malware and malicious Microsoft Excel files
	3. Delivery	Send spear phishing email to DSOs operator pretending as an email from Ukraine Ministry of Energy with malicious Microsoft Excel attachment
B	4. Exploit	An operator opens the file and exploits Microsoft Excel macro's vulnerability to get remote access
	5. Privilege Escalation	Gain unauthorized access to the database
	6. Credential Access	Get credentials from the active directory
	7. Defense Evasion	Avoid firewall detection through VPN tunnels
C	8. Establish Foothold	Using a backdoor and VPN to maintain its presence
	9. Internal Reconnaissance	Gather information from the OT network system
	10. Lateral Movement	Moving between office network, data center, control center, digital substation
	11. Collection	Collect information from IT and OT systems
D	12. Exfiltration	Send the data to the C2 server
	13. Inhibit Response Function and Impair Process Control	Telephony denial of service and disable UPS to de-energized control center
	14. Unauthorized Control Commands on OT System	Control SCADA HMI remotely to switch off the breaker
	15. Cyber Attacks Impact Power System Operation	The breaker on the power grid switched off leading to a power outage
E	16. Induced Power System Events	De-energized several substations
	17. Operator and Automated Remedial Action	Operator's initial attempts to recover the power grids failed because of inhibited response function and impaired process control
	18. Slow Cascading Failures	Impact on distribution system operator, did not cause cascading failures and full blackout
	19. Point of No Return	Not applicable
F	20. Fast Cascade and System-Wide Collapse	Not applicable
	21. Blackout	Power outages affect 225,000 power grid customers
	22. Social Impacts	Disruption of energy supply during the winter and financial losses
	23. OT Recovery and Power System Restoration	Operator recovers through manual mode within 6 hours

information about the target to craft the email and deceive the target. Before sending the phishing, adversaries also prepare malicious Excel files and BlackEnergy3 malware during the weaponization. The attack entered the IT network through spear phishing. The attackers specifically targeted three DSOs while impersonating officials from Ukraine's Ministry of Energy. The malicious emails, which appeared to come from reliable sources, included a weaponized Microsoft Excel file attachment. The attackers used macro vulnerabilities in Microsoft Excel to install malware on the DSO IT network. A macro is a program that uses Visual Basic Application (VBA) scripts to automate tasks in Microsoft Excel. When the recipients enabled the macro, the VBA script was executed, resulting in the installation of BlackEnergy3 malware on the computer.

Following successful installation, the malware established a connection to the attackers' remote C2 server via IP address 5.149.254.114 and port number 80. The malware was specifically designed to communicate to the remote IP address and port number controlled by adversaries. The connection via port 80 appeared to be innocuous. This is due to the fact that port 80 is a commonly used port for website traffic via Hypertext Transfer Protocol (HTTP). Therefore, the attackers could remotely control the BlackEnergy3 malware through its connection to the C2 server.

BlackEnergy3 included some functionalities, i.e., network scanner, file stealer, password stealer, key logger, screenshot capturer, and network discovery. The BlackEnergy3 malware roles were substantial for the initial and main attack stages. The malware discovered information about the IT network configuration, such as network segments, network topology, hosts connected, and so on, by using the network scanner and network discovery modules. With BlackEnergy3, attackers could also use a key logger to steal passwords, steal files, and capture screenshots of the targeted computers. This information was critical for preparing the subsequent attack stages. The attackers sent all of the collected data directly to the remote C2 server.

The attackers discovered a vulnerable active directory server during the internal reconnaissance stage, which became a breach point of the IT/OT system. AD is a database service for IT networked system operations that runs on Windows. An AD manages a user's access permissions by serving as a central authentication and authorization authority for managed accounts, hosts, and services. With centralized authentication and authorization rather than segregated services, AD makes IT system operations easier and more flexible. Active directory databases also store usernames, passwords, and information about hosts and services. As a result, the AD is a critical point for authentication and security. The attackers in the Ukraine 2015 cyber attack compromised the AD server to gain login credentials to the majority of hosts in the IT and OT network. Subsequently, this access is utilized by the adversaries to perform lateral movement through the IT and OT network, including the control center.

After gaining access to the control center, the attackers established a Virtual Private Network (VPN) connection from one of the control center's computers to a remote location on the Internet. VPN enabled the attackers to gain access to the targeted computer via tunnel and encrypted connections. Instead of using a static C2 server with port 80, VPN allowed attackers to access the computer from anywhere on the Internet. Furthermore, the VPN enabled the attackers to avoid detection by firewalls and conceal their true locations. At this point, the attackers had gained complete control and were ready to launch the final

attack. The attackers, however, remained undetected, carrying out additional actions to amplify the impact of the cyber attack on the distribution network.

For increasing attack severity, adversaries also impaired legitimate process control and inhibited response function. The impairment of legitimate process control performed through substation device firmware modification wiped out the hard disk and disabled UPS. The adversaries gained access to substations and compromised RTUs and serial-to-Ethernet converters. A serial-to-Ethernet converter connects substation Ethernet communications, e.g., IEC 104, to control center serial communications, e.g., IEC 101. These devices depend on firmware for controlling their processes. The attackers also created malicious firmware and replaced the legitimate firmware in RTUs and serial-to-Ethernet converters, causing them to be inoperable upon reboot. In addition, the malicious firmware prevented grid operators from remotely controlling the substations to perform recovery. KillDisk was used by the attackers to erase hard drives in the control center computers, causing them to be unbootable. KillDisk is a part of BlackEnergy3 malware module that deletes data, registry entries, and system configuration. To prolong the system recovery, adversaries also disabled UPS in the control center, causing them to be inoperative during the blackout. For the inhibit response function, adversaries performed telephony denial of service to make the operator in the control center unable to get information from the outside. After finishing all preparation stages, adversaries launched the final attack on December 23rd by opening the circuit breaker, causing an instant power outage. Cascading impact on the power system was not reported in this attack, but this attack caused a power outage affecting wide area distribution within 6 hours. As a result, the attackers successfully carried out one of the two most advanced cyber attacks on power systems to date.

### 3.4.2 UKRAINIAN POWER GRID CYBER ATTACK 2016

On December 17, 2016, at 23:53 local time, a second cyber attack on Ukraine's power grid took place. This incident was the first publicly reported cyber attack that employed customized malware to target power systems. The malicious software used in the 2016 attack was named CRASHOVERRIDE or Industroyer. The attack had an effect on the SCADA system at the transmission level, and it was directed at a single 330 kV substation as its target. Because of the attack, the distribution network power outage resulted in a total load of 200 MW unable to be supplied. The attack that took place in 2016 was significantly more advanced in terms of its technique than the one that took place in 2015. Fortunately, the damage from this attack was considerably less than the previous one. In [11], an extensive study of this attack is offered and discussed. Within this sub-section, we carried out a review of the attack that took place in 2016 employing the ACPPS kill chain. Table 3.9 provides a summary of the different stages of the attack according to the ACPPS kill chain. In addition, the study was accompanied by an illustration of the part of the system that was being targeted, which can be found in Fig. 3.4.

During a few prior months, the adversaries effectively acquired control of the compromised hosts in the control center by utilizing techniques associated with APTs. After that, in the later stages, the adversaries opened circuit breakers and transferred malicious payloads to substations by exploiting SCADA protocol vulnerabilities. On December 17, 2016, at 23:53 local time, these attacks got underway. The attacks were directed at the SCADA system at the transmission level. A single 330 kV/110 kV/10 kV substation was

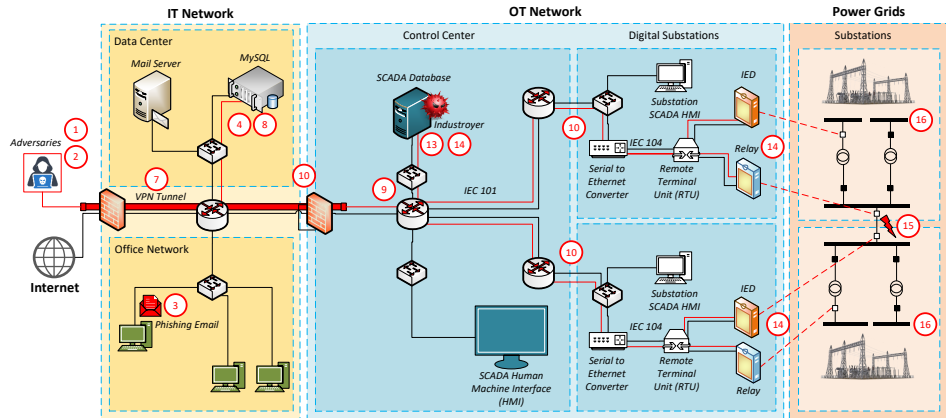


Figure 3.4: Cyber attack on Ukrainian power grids 2016 [4].

Table 3.9: Advanced cyber-physical system kill chain stages in Ukraine cyber attack 2016.

Stages	Sub-stages	Ukrainian power grids cyber attack in 2016 processes
A	1. External Reconnaissance	Gather information about the targeted system and identifying potential protocol used in power system operation
	2. Weaponization	Prepare specific malware to exploit specific SCADA protocols, i.e., IEC 101, IEC 104, IEC 61850, and Open Platform Communication (OPC)
B	3. Delivery	Phishing email targeting Ukrainian power grids operator in January 2016
	4. Exploit	Exploit MySQL server vulnerability to gain full control to the server
	5. Privilege Escalation	No information available
	6. Credential Access	No information available
	7. Defense Evasion	VPN tunnel used to bypass the firewall and remotely access the compromised servers
C	8. Establish Foothold	Maintain adversaries presence using compromised MySQL server
	9. Internal Reconnaissance	Malware used to facilitate internal reconnaissance
	10. Lateral Movement	Compromised MySQL server used as pivot point for lateral movement
	11. Collection	No information available
D	12. Exfiltration	No information available
	13. Inhibit Response Function and Impair Process Control	The wiper in CRASHOVERRIDE module removed files relating to ICS operations to prevent instantaneous controller recovery and power system restoration
E	14. Unauthorized Control Commands on OT System	Launched the control command attack on December 17, 2016 at 23:53 local time
	15. Cyber Attacks Impact Power System Operation	The breaker on the power grid switched off lead to power outage
	16. Induced Power System Events	The attacks affected the SCADA system at the transmission level focusing on a single 330 kV / 110 kV / 10 kV substation, resulting in a distribution-level outage
	17. Operator and Automated Remedial Action	No information available
	18. Slow Cascading Failures	Impact on distribution system operator, did not caused cascading failures and full blackout
F	19. Point of No Return	Not applicable
	20. Fast Cascade and System-Wide Collapse	Not applicable
	21. Blackout	Power system outage
	22. Social Impacts	No information available
	23. OT Recovery and Power System Restoration	The disruption was recovered by system operator using manual mode

the focus of the attacks, which resulted in an outage at the distribution level. As soon as they became aware of the irregularity in the system, operators reacted swiftly to the attack by switching the controls to manual mode. Fortunately, the operator was able to recover, making this attack ineffective, but it did demonstrate how an advanced protocol exploit might be used to launch an attack.

The overall impact of the cyber assault in 2016 was significantly less due to a number of different factors. The fact that the malicious payload injections did not work properly was the primary cause. It is likely that this was brought about by a manually coded attack technique that did not perform as intended [219]. The attackers should have approached the development of the protocol payload module in a methodical manner and provided it with the relevant testing environment, such as real SCADA devices. Nevertheless, such instruments are not widely available and are exclusively utilized by operators of industrial systems. As a result, the developed malware failed to operate in its intended function. Yet, this attack successfully demonstrates a sophisticated APTs with weaponized malware and a deep understanding of the targeted power system. These kinds of attacks have the potential to become more prevalent in the future, which would have a significant adverse impact on the infrastructure of the power grids.

### 3.4.3 UKRAINIAN POWER GRID CYBER ATTACK 2022

In late 2022, a cyber attack targeting the Ukrainian power grids was reported, with evidence pointing to the involvement of the Sandworm hacker group [11]. The adversaries employ the OT-level Living off the Land (LotL) technique, which intends to open the victim's substation circuit breakers, resulting in an unplanned power outage. Compared to the attacks in 2015 and 2016, there is not much information available regarding the attack processes. One reason is that the adversaries employed anti-forensic techniques to hinder the forensic investigation of the attack processes.

The cyber attack started in June 2022 before leading to a disruptive event on October 10 and 12, 2022. There is no available information on how the intruder accessed the OT system. The attack exploits the vulnerabilities of MicroSCADA applications to launch malicious control commands. The malicious control commands successfully open the circuit breakers causing a power outage. However, the detailed impact on the power system is unknown. Based on the available information, the cyber attack process can primarily be categorized using stages C and D of the ACPPS kill chain.

### 3.4.4 EXPERIMENTAL CYBER ATTACK

This subsection discusses an experimental case study involving an example of a transmission grid IT-OT network. The topology of the IT-OT network and the transmission power system is depicted in Fig. 3.5. The power system is modeled and simulated using DiGSILENT PowerFactory, while the cyber system is emulated through Mininet. It is an open-source network emulator that allows users to create a virtual network topology using software-defined networking (SDN) [294]. Furthermore, it implements operating-system-level virtualization based on the Linux namespace containerization. This allows Mininet to emulate larger communication networks in comparison to typical virtual machines. The emulated IT-OT network consists of 27 user-defined substations, 118 measurement devices, and over 800 data points for the entire simulated power system. SCADA device



functionality within the network is implemented through custom Python scripts.

The simulated power grid runs a Root Mean Square (RMS) simulation of the IEEE 39-bus test system. The power grid simulation provides time-domain measurement data from substation bays, e.g., buses, lines, and generators, in the form of active and reactive power, voltage, and current measurements. All measurement data is then sent from the substation to the control center. The data is also stored in local databases located within substations and the control center. Such a cyber-physical experimental setup allows us to study the impact of cyber attacks on the power system.

3

In the preparation stage of the attack, the adversaries conduct reconnaissance to gather information about the target, including email addresses, potential operational protocols of the system, etc. This initial information is then used to prepare weaponized cyber attack tools for the later stage of the attack. Subsequently, a weaponized file is then delivered to the target via phishing emails. It serves as an entry point and backdoor for the attackers to the targeted system.

From the entry point, the attackers exploit vulnerabilities in the MySQL SCADA database server in the control center. The vulnerability exploit allows the attackers to gain administrator privileges and perform credential system theft. To circumvent firewall detection and secure direct access, attackers enable VPN access from the external network to the compromised MySQL server. Therefore, the compromised MySQL server acts as a central attack location during the main attack stages.

From the compromised MySQL server, attackers maintain their presence and launch the next stages of the attack. This includes internal reconnaissance, lateral movement, data collection, and exfiltration. During these main stages, the attackers learn the operating protocols used for communications between the control center and substations. Through the protocol exploits, the attackers spoof the legitimate control command for opening circuit breakers, as demonstrated in [65]. As shown in Fig. 3.5, this malicious command subsequently results in the opening of three circuit breakers in substation 2. Under normal circumstances, the system operator would quickly recognize the situation and take immediate corrective action, i.e., the circuit breakers will be closed. In this scenario, however, in addition to the spoofing attack, the attackers also launch a DoS attack against the substation. Fig. 3.6 shows the traffic comparison in the substation gateway and relay before and after the DoS attack at  $t=10$  s. Before and after the DoS, there is a significant change in substation network traffic. Fig. 3.7 depicts a statistical summary of the traffic data, including the minimum, median, maximum, first quartile, and third quartile. The box plot also indicates the variability, spread, and skewness of the data. Meanwhile, the circles in the plot indicate the outlier data. Furthermore, the attack location is visualized in Fig. 3.8. In graph-based visualization, red nodes represent the anomalous traffic due to DoS traffic, and blue nodes represent the normal ones. This graph-based anomaly visualization is presented based on our previous research in [5]. Due to the DoS attack, timely corrective actions are made more difficult because commands sent from the control center do not reach the substation in a timely manner. As a result, system stability is affected, leading to cascading failures and a power outage. Table 3.10 summarizes the simulated cyber attack stages based on the ACPPS kill chain.



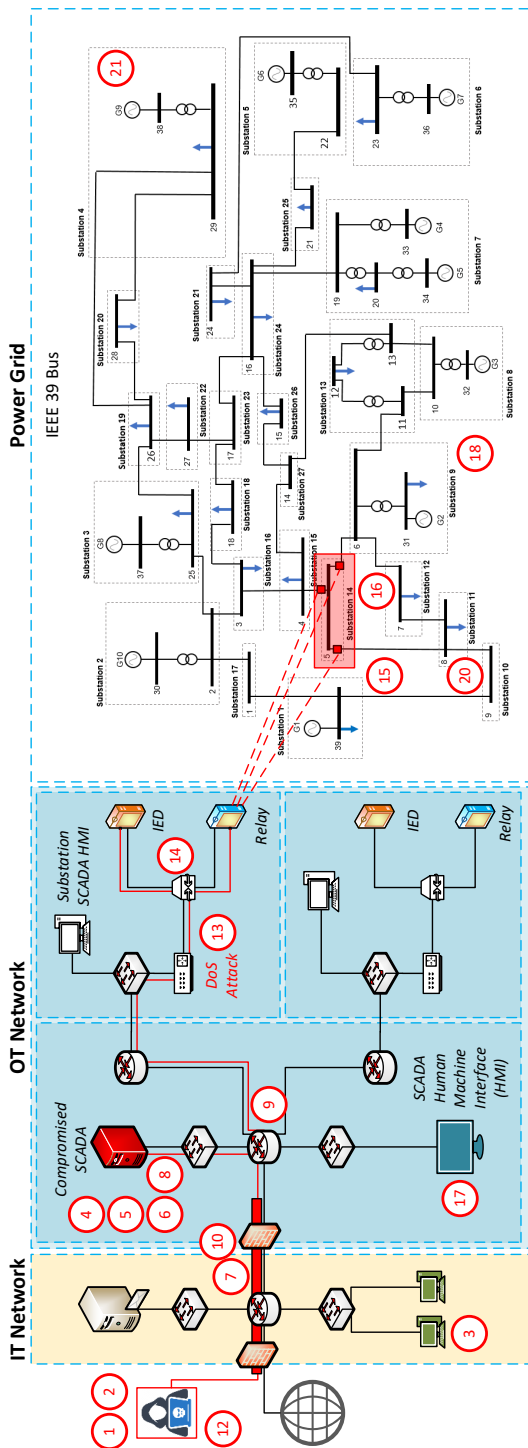


Figure 3.5: Cyber-physical co-simulation experimental setup to analyze the impact of cyber attacks on the power system. The simulated power grid is based on the IEEE 39 bus system with 27 Substations (Sub). The cyber attack impact on power grid started from the malicious opening of the circuit breaker at Bus 2 in Substation 2 [4].

Table 3.10: Advanced cyber-physical system kill chain stages in simulated cyber attack.

Stages	Sub-stages	Relation to experimental cyber attack on power grid processes
A	1. External Reconnaissance	Gather information about the targeted system and identify potential communication protocols used in power system operations
	2. Weaponization	Prepare a malicious script based on reverse shell and a payload for the circuit breaker attack
	3. Delivery	Phishing email targeting grid operator
B	4. Exploit	Exploit MySQL server vulnerability to gain full control to the server
	5. Privilege Escalation	Gain admin privilege of SCADA MySQL database
	6. Credential Access	Perform credential theft of username and password accounts from compromised SCADA database
	7. Defense Evasion	VPN tunnel is used to bypass the firewall and remotely access the compromised servers using open source VPN
C	8. Establish Foothold	Maintain presence using compromised MySQL server
	9. Internal Reconnaissance	Perform reconnaissance using publicly available tools, i.e., Nmap, Tshark
	10. Lateral Movement	Compromised MySQL server serves as a pivot point that allows adversaries to compromise other host within the network
	11. Collection	From internal reconnaissance phase, adversaries collect samples of communication packets between control center and substations
D	12. Exfiltration	Send the sample protocol to the external remote host
	13. Inhibit Response Function and Impair Process Control	To prevent the remedial actions, adversaries also launch a DoS attack using hping3 on the OT communication network
	14. Unauthorized Control Commands on OT System	Cyber attack is executed to launch the spoofed payload with open circuit breaker command. The malicious control command open circuit breakers sent in the substation 2
E	15. Cyber Attacks Impact Power System Operation	At time $t = 10$ s circuit breaker on lines 01-02, 02-03, and 02-25 maliciously disconnected. Synchronous generator G10 disconnects from the main grid
	16. Induced Power System Events	At $t + 7.184$ s over frequency protection of synchronous generator G10 trips
	17. Operator and Automated Remedial Action	Operator unable to perform system recovery due to DoS attacks
	18. Slow Cascading Failures	<ul style="list-style-type: none"> <li><math>t + 15.111</math> to <math>t + 15.233</math> s : Lines 08-09 and 25-26 in vicinity of attacked substation are tripped by distance protection</li> <li><math>t + 15.87</math> to <math>t + 17.583</math> s : Generators G8 and G9 tripped due to Rate of Change of Frequency (ROCOF) interface protection and disconnected. System is now heavily affected</li> </ul>
	19. Point of No Return	Power system enter to critical state and difficult to return to initial condition.
F	20. Fast Cascade and System-Wide Collapse	<ul style="list-style-type: none"> <li><math>t + 18.87</math> to <math>t + 19.072</math> s : Under frequency load shedding activated. Loads in affected area are shed by 6.7%</li> <li><math>t + 19.1</math> s : Line 16-17 trips on distance protection. This line is the tie-link between two areas</li> <li><math>t + 19.139</math> to <math>t + 19.879</math> s : Under frequency load shedding activated. All loads shed by 6.5%. <math>t + 20.621</math> s : Line 03-04 trips on distance protection. The affected area is completely isolated from the rest of the grid</li> <li><math>t + 20.887</math> to <math>t + 31.941</math> s : Frequency of system is still below permissible limits. Under frequency load shedding activated in steps of 5.9% and 7%. The rest of the system stabilizes</li> </ul>
	21. Blackout	The attack leads to a partial blackout with 12 busbars being de-energized and a loss of 2285 MW load (37% of total load)
	22. Social Impacts	The power outage causing disruption on wide area electricity consumer. Consequently, the power outage will disrupt service in various area including healthcare, transportation, communication, etc.
	23. OT Recovery and Power System Restoration	To recover the OT system, the root cause of cyber attack need to be neutralized

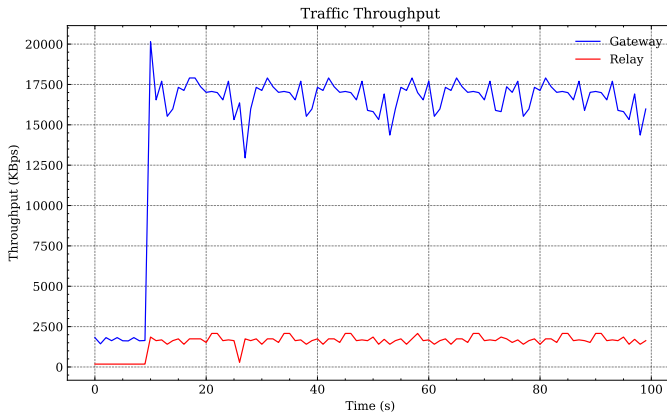


Figure 3.6: Traffic comparison from normal and DoS traffic [4].

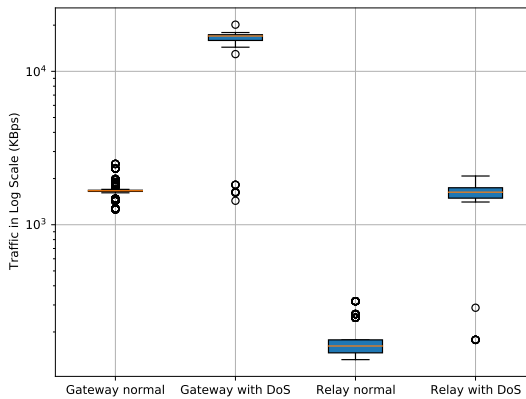


Figure 3.7: Box plot comparison from normal and DoS traffic [4].

The aforementioned cyber attack leads to the malicious disconnection of lines 01-02, 02-03, and 02-25, as well as a DoS attack that inhibits the system operator's capacity to carry out remedial measures. Consequently, the prolonged cyber attack affects the stability of the power system. After this cyber-induced contingency, the system becomes unstable due to the loss of three transmission lines and the resulting disconnection of a major generating unit. As can be shown in Fig. 3.9, for a prolonged period after the attack, the system is intact, but due to the imbalance between generation and consumption, voltage instability occurs. This can be seen by the oscillations in the voltage magnitudes measured in the buses of the system. Due to this imbalance and the limited capacity of the transmission lines in the vicinity of the attack location to support the power flows, multiple lines are disconnected by distance relays operating on sustained under voltages and over currents. A similar phenomenon was also observed during the 2003 cascading failures and blackouts in North America [295]. A critical line tripped because of incorrect operation of zone 3

distance protection, which exacerbated the domino effect, contributed to the spread of the cascading phenomenon, and ultimately resulted in a widespread power outage [277].

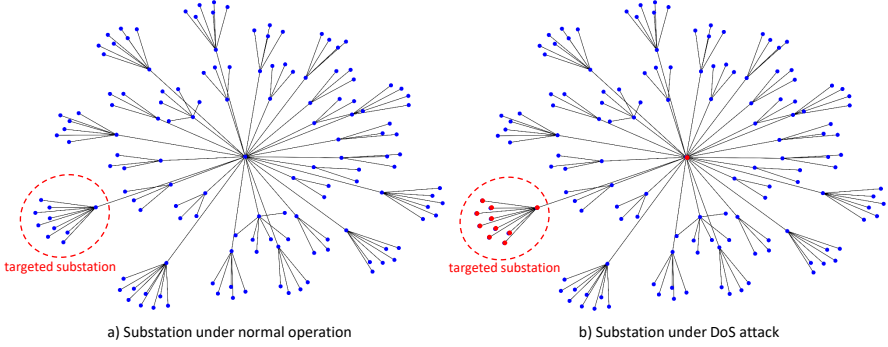


Figure 3.8: Graph visualization from attack on substation 2 [4].

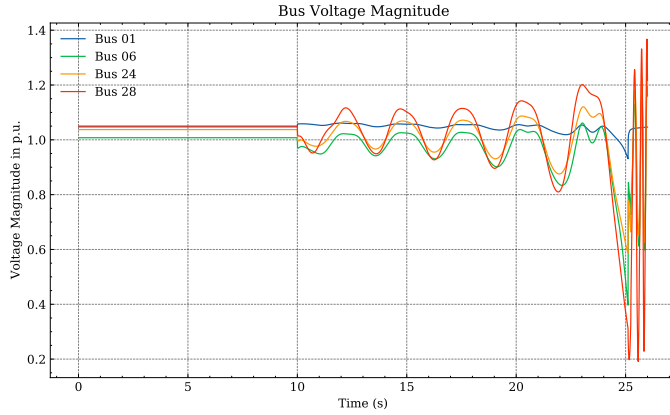


Figure 3.9: Bus voltage magnitude [4].

As a result of multiple line disconnections, oscillations are observed between the generators in the affected area and the rest of the system. This is observed in Fig. 3.10, where it can be seen that generator 08 is oscillating against generators 04 and 05. The resulting instability and the absence of remedial actions cause two generators, namely G8 and G9, to be tripped by their interface protection due to the high ROCOF condition. As seen in Fig. 3.13, generator 08 is exceeding the ROCOF setting of 2 Hz/s for over 500 milliseconds, which is the protection setting. Now, due to the loss of generation, system frequency starts plummeting, and emergency load shedding is activated to preserve system integrity. This is illustrated in Fig. 3.12. Ultimately, the cyber attack led to a partial blackout with 12 busbars being de-energized and a loss of load amounting to 2285 MW. In Fig. 3.9- Fig. 3.12, cyber attack action that opens the breaker executed at time  $\tau=10$  s. The impacts of cyber attacks are shown after  $\tau$ .

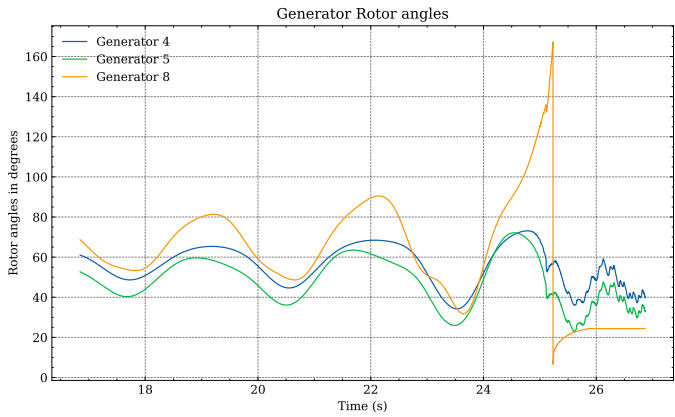


Figure 3.10: Generator rotor angles [4].

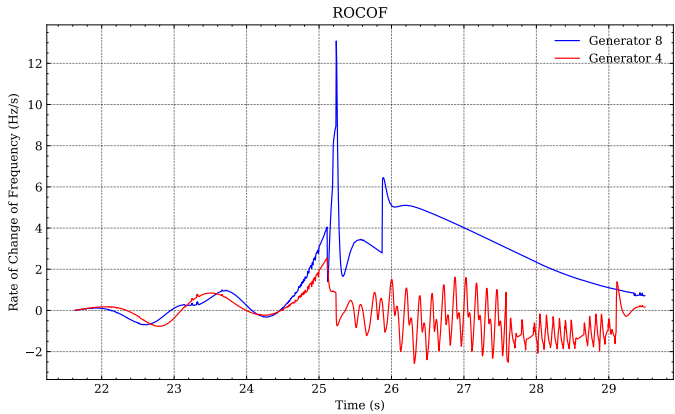


Figure 3.11: Rate of Change of Frequency on Generator Generator 4 and Generator 8 [4].

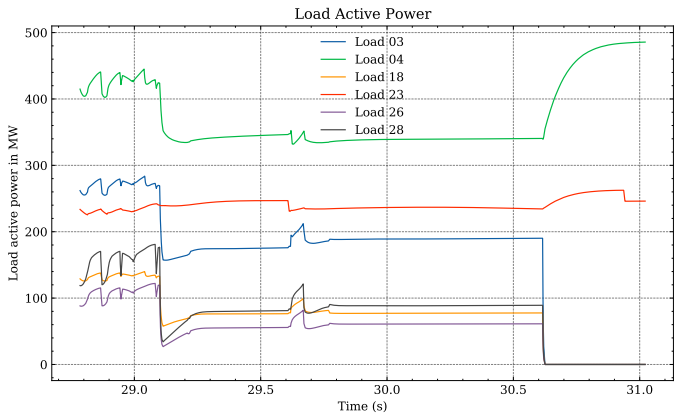


Figure 3.12: Change of load active power for load 03, 04, 18, 23, 26, and 28 [4].

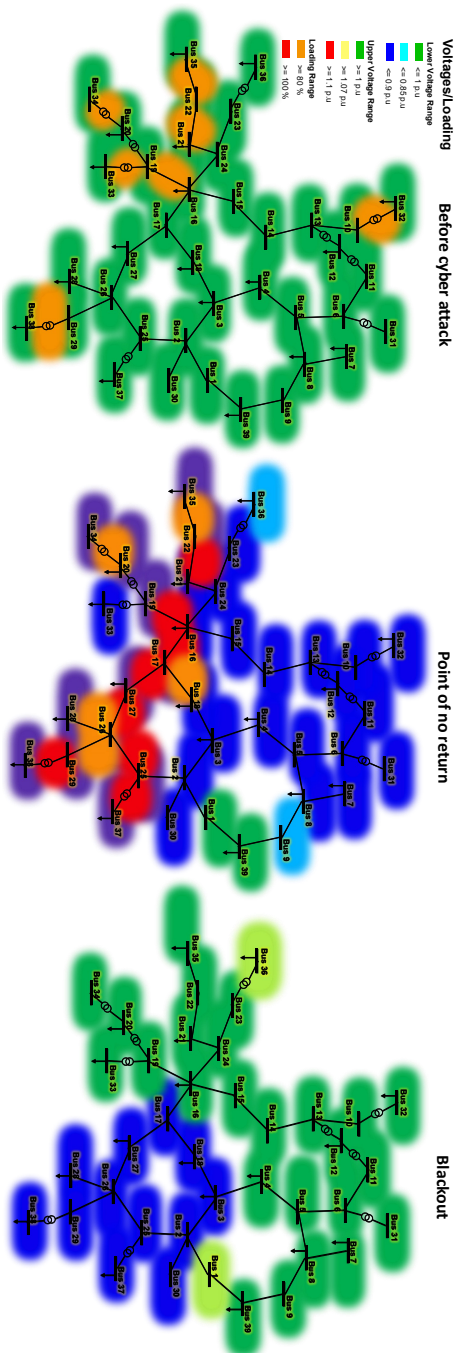


Figure 3.13: Cascading impact visualization [4].

The entire power system comparison before and after the cyber attack and the propagation of cascading events are shown in Fig. 3.13. The left image shows the heatmap of the voltage magnitudes and lines loading before the cyber attack simulation ( $\tau - 1$  s), while the right depicts the outcome after the cyber attack ( $\tau + 50$  s). The middle image depicts the state of the system at the point of no return. As it can be seen, most of the system is stressed, with very low voltage levels and overloaded transmission lines. From the simulated cyber attack, although it was only executed from a single compromised substation, the impact of the attack is not local. This simulation shows that cyber attacks on power grids can cause wide-area cascading impacts.

### 3.4.5 RESULT AND DISCUSSION

Based on the aforementioned cyber attack case studies, the ACPPS kill chain is able to identify all stages of cyber attacks in power grids. Table 3.11 summarizes all case studies mapping into all stages of the ACPPS kill chain. ACPPS kill chain can identify more granular stages of cyber attacks. In the Ukraine 2015 and 2016 case studies, the ACPPS kill chain was unable to identify some stages. The reason is that there is no available information associated with the particular stages. For example, neither real case of a cyber attack provided any information related to the cascading failure or point of no return. These stages required information related to the power system measurement, and the available reports in [8, 11, 219] only provided information related to the impact on IT and OT systems. In the experimental attack, we are able to capture power system measurement data from the co-simulation. Therefore, the experimental case study provides a more comprehensive analysis of cyber attacks and their impact on the CPPS.

Compared to other cyber attack stage identification frameworks, the ACPPS kill chain provides more detailed stages. Table 3.12 summarizes the comparison of the ACPPS kill chain with other frameworks. The ACPPS kill chain refers to the MITRE ATT&CK ICS framework as the most comprehensive cyber attack stage identification for the IT and OT systems. Therefore, from the stages in the IT/OT network, the ACPPS kill chain provides the same quantity of stages as the MITRE ATT&CK ICS. During the physical stages, the ACPPS kill chain proposed seven new stages associated with the cyber attack's impact on the power system. The ACPSS kill chain also proposed new stages associated with secondary impact and recovery. Overall, there are 23 stages in the ACPPS kill chain, which provides nine new stages that are unavailable in other frameworks. Therefore, by enhancing the granularity of cyber attack stages by a factor of 0.64, the ACPPS kill chain improves the effectiveness of identifying cyber attack stages in CPPS.

Despite providing a more detailed stage analysis of cyber attacks on CPPS, the ACPPS kill chain remains inadequate in identifying all stages of an attack, primarily due to the absence of power system data, as indicated in Table 3.11. A significant factor contributing to this limitation is the insufficient integration among IT, OT, and physical power systems. The lack of seamless integration creates silos that hinder comprehensive monitoring and analysis. Fortunately, the ongoing digitalization of power systems presents a promising solution to this problem. The convergence of IT, OT, and physical power systems through digitalization facilitates real-time data sharing and interoperability, which are crucial for holistic situational awareness and more effective threat detection. This integration not only promises to fill existing data gaps but also provides a unified platform for implementing

advanced security measures that can preemptively identify and mitigate sophisticated cyber threats. Consequently, the ACPPS kill chain could be refined into an innovative framework for analyzing cyber attacks on CPPS, providing a more comprehensive representation of the attack stages.

Table 3.11: Summary of ACPPS Kill Chain Implementation for Real Cyber Attack in Ukraine 2015 and 2016, and Experimental Cyber Attacks.

Stages	Sub-Stages	Ukraine 2015	Ukraine 2016	Experimental Attack
<b>A. Attack Preparation</b>	1. External Reconnaissance	✓	✓	✓
	2. Weaponization	✓	✓	✓
<b>B. Initial Engagement</b>	3. Delivery	✓	✓	✓
	4. Exploit	✓	✓	✓
	5. Privilege Escalation	✓	NA	✓
	6. Credential Access	✓	NA	✓
	7. Defense Evasion	✓	✓	✓
<b>C. Main Attack Phases</b>	8. Establish Foothold	✓	✓	✓
	9. Internal Reconnaissance	✓	✓	✓
	10. Lateral Movement	✓	✓	✓
	11. Collection	✓	✓	✓
	12. Exfiltration	✓	✓	✓
<b>D. Physical System Engagement</b>	13. Inhibit Response Function and Impair Process Control	✓	✓	✓
	14. Unauthorized Control on OT System	✓	✓	✓
<b>E. Power System Impacts</b>	15. Cyber Attack Impacts Power System Operation	✓	✓	✓
	16. Induced Power System Events	✓	✓	✓
	17. Operator and Automated Remedial Action	✓	✓	✓
	18. Slow Cascading Failure	✓	✓	✓
	19. Point of No Return	NA	NA	✓
	20. Fast Cascade and System-Wide Collapse	NA	NA	✓
	21. Blackout	✓	✓	✓
<b>F. Social Impacts and Recovery</b>	22. Social Impacts	✓	NA	✓
	23. OT Recovery and Power System Restoration	✓	✓	✓

✓ = Available; NA = Not Available

Table 3.12: Comparison of ACPPS Kill with Other Frameworks for Cyber Attack Stages Identification.

Frameworks	Number of Stages			
	IT/OT	Physical	Secondary Impact and Recovery	Total All Stages
Cyber Kill Chain [226]	7	0	0	7
CPS Kill Chain [227]	5	2	0	7
MITRE ATT&CK ICS [228]	12	2	0	14
SANS ICS [229]	5	0	0	5
ACPPS Kill Chain	12	9	2	23

The current version of the ACPPS kill chain only provides the stages of cyber attack in CPPS and does not quantitatively assess every stage of the attack. For future work, it is possible to provide quantitative stages identification on the ACPPS kill chain. One potential solution is to integrate AI with the ACPPS kill chain. This is aligned with the state-of-the-art AI application for cyber security applications [296–299]. With the comprehensive quantitative matrices and AI application, the ACPPS kill chain can be used to classify anomalous events into specific stages in the ACPPS kill chain. Furthermore, it is also possible to predict the potential impact of cyber attacks on power systems to avoid severe impacts.



### 3.5 CONCLUSION AND RECOMMENDATIONS

In this research, APTs to CPPS were investigated. This research presents three parts of contributions.

In the first part, the chapter identified and compared the characteristics of APT attacks in IT, CPS, and CPPS. We define the characteristics of APTs on CPPS, which are different compared to APTs in IT systems and general CPS.

In the second part, we propose a novel ACPPS kill chain framework. ACPPS defines and examines the cyber-physical APT stages on power grids that cause cascading failures and a blackout. This novel kill chain framework offers more comprehensive attack stages for a thorough analysis of APTs on power systems and early-stage mitigation compared to the current frameworks reported in the literature.

In the third part, this manuscript provides an in-depth analysis of actual and experimental cyber attacks on power grids. The in-depth analysis is performed based on the proposed ACPPS kill chain on Ukraine's 2015, 2016, and 2022 cyber attacks and experimental scenarios.

Overall, this manuscript's contribution is by enhancing state-of-the-art research comprehension of APTs targeting CPPS. It achieves this by introducing a novel framework of ACPPS kill chain for analyzing these threats and providing practical insights through both real-world and experimental case studies. Through these contributions, this work aims to stimulate further research and development endeavors focused on improving the resilience of CPPS in response to evolving cyber threats. It is important to note that we are currently living in a world where artificial intelligence plays an increasing role. Therefore, there is an opportunity for future research on integrating AI with the ACPPS kill chain. The ACPPS kill chain navigates the stages of cyber attacks in CPPS while the AI helps to classify anomalous events into the associated ACPPS kill chain stages. This integration will provide a comprehensive solution for cyber attack mitigation in the earlier stage of the ACPPS kill chain and prevent more severe impacts. In addition, in the current version, the ACPPS kill chain does not provide quantitative metrics to evaluate APTs in CPPS. With AI integration, the ACPPS can provide more comprehensive quantitative evaluation metrics to predict and mitigate cascading failures and points of no return.



## 4

## 4

# ATTACK GRAPH MODEL FOR CYBER-PHYSICAL POWER SYSTEM USING HYBRID DEEP LEARNING

*Electrical power grids are vulnerable to cyber attacks, as seen in Ukraine in 2015 and 2016. However, existing attack detection methods are limited. Most of them are based on power system measurement anomalies that occur when an attack is successfully executed at the later stages of the cyber kill chain. In contrast, the attacks on the Ukrainian power grid show the importance of system-wide, early-stage attack detection through communication-based anomalies. Therefore, in this chapter, we propose a novel method for online cyber attack situational awareness that enhances the power grid resilience. It supports power system operators in the identification and localization of active attack locations in Operational Technology (OT) networks in near real-time. The proposed method employs a hybrid deep learning model of Graph Convolutional Long Short-Term Memory (GC-LSTM) and a deep convolutional network for time series classification-based anomaly detection. It is implemented as a combination of software defined networking, anomaly detection in communication throughput, and a novel attack graph model. Results indicate that the proposed method can identify active attack locations, e.g., within substations, control center, and WAN, with an accuracy above 96%. Hence, it outperforms existing state-of-the-art deep learning-based time series classification methods.*

4.1 INTRODUCTION

Cyber attacks on power grids are high-impact and low-frequency disturbances with a wide range of consequences. These could include but are not limited to, equipment damage, loss of load, and power system instability. In the worst-case scenario, cyber attacks and advanced persistent threats may cause system-wide cascading failures and a blackout. Therefore, cyber attacks on power grids are severe threats and have already been identified in the real world. For example, on December 23, 2015, a cyber attack was conducted on the power grid in Ukraine that resulted in a power outage, affecting 225,000 customers [8]. A more sophisticated cyber attack followed on December 17, 2016, resulting in a power outage in the distribution network, where 200 MW of load was left unsupplied [9]. The attackers employed several attack strategies and steps to achieve their objectives. These can be mapped with the seven stages of the cyber kill chain for an in-depth analysis of such an advanced persistent threat, i.e., reconnaissance, weaponization, delivery, exploitation, installation, command and control, and action on objectives [25] as depicted in Fig. 4.1. However, existing detection methods for cyber attacks on power grids are limited. Most of them are based on power system measurement anomalies that occur when an attack is successfully executed at the later stages of the cyber kill chain, e.g., false data injection [17–24]. In contrast, in the aforementioned cyber attacks in Ukraine, the cyber kill chain lasted for more than six months between the reconnaissance and command and control stages. The latter caused power outages in a matter of minutes [8, 9, 219]. Hence, this highlights the urgency of timely early-stage attack detection through Information Technology-Operational Technology (IT-OT) system anomalies. Physical measurement-based anomaly detection is only valid for later stages in the cyber kill chain, i.e., command and control and actions on objectives. Therefore, in this research, we propose an early-stage anomaly detection method for OT systems. It is implemented in the control center to detect cyber attacks at the early stages of the cyber kill chain, based on throughput anomalies in OT communication traffic power system wide.

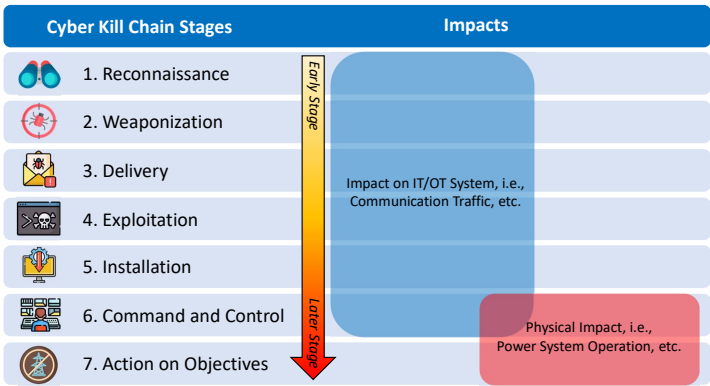


Figure 4.1: Cyber kill chain stages and impacts [5].

Cyber attack detection on power grids have been extensively studied in recent years. Nonetheless, the majority of the existing research is focused on the identification of cyber

attacks on power grids under False Data Injection (FDI) attack scenarios. These scenarios focus on analyzing power system measurements to identify anomalies in power grids [17–24]. However, in the real-world cyber attacks on power grids reported in [8, 9, 219] adversaries did not perform FDI attacks. Instead, in the early stages of the cyber kill chain, attackers targeted the IT-OT communications. Therefore, in this research, we omit power system measurements under FDI attack scenarios and focus on the OT communication traffic anomalies.

There are four major methods reported in the literature for power grid communication traffic anomaly detection, i.e., signature-based [300], sequence-based [301], rule-based [302–304], and machine learning-based [305–307]. Recent research shows that machine learning-based methods are gaining increased attention and provide superior performance for anomaly detection [308–310]. Therefore, in this work, we focus on machine learning-based communication traffic anomaly detection. Our proposed model is based on a semi-supervised learning. It does not use signatures, sequences nor rules for detection and classification. The proposed model classifies OT network traffic into two categories, i.e., normal and anomalous, based on the network traffic throughput. Previous research in this area is discussed in [305, 307]. In [305], the authors used labeled communication packets from UNSW-NB15 and IDE2012/16 datasets as inputs to predict the Distributed Denial of Services (DDoS) attacks. Meanwhile, in [307], the authors use traffic data logs from Snort to create a sequence-based anomaly detection technique. However, both machine learning implementations do not use traffic throughput data, which is our research focus. Furthermore, the vast majority of machine learning-based anomaly detection methods only focus on IT systems [308–311]. Even though the IT and OT systems of a utility are integrated, the traffic characteristics are distinct. The network traffic in OT systems is generated from automated processes with deterministic and homogenous behavior, whilst the IT system traffic consists of user-generated data with a stochastic behavior [312]. Hence, the implementation of traffic-based anomaly detection for OT systems is fundamentally different from that of IT systems.

Amongst the machine learning-based traffic anomaly detection methods, most recent works use deep learning models that provide a better performance [309, 313]. In [314], the authors propose a deep reinforcement learning-based method for traffic flow matching control. They focus on detection of DDoS attacks that systematically trigger considerable anomalies in traffic throughput. Therefore, this method is not suitable to detect infinitesimally small changes in OT network traffic throughput, e.g., caused by stealthy attacks [314]. In [315], the authors used Convolutional Neural Network (CNN) for communication traffic classification. However, the CNN method cannot detect unknown cyber attacks because it depends on preliminary traffic data for the training. To address this gap, instead of using specific labeled data for each attack category, we use the quantitative anomaly. The quantitative anomaly detection uses the throughput of the OT communication traffic. The throughput is quantified as a time series to generate a unique waveform pattern as shown in [316–318]. Therefore, instead of classifying specific attack types or sequences, in this work we classify the time series traffic flow into two categories, i.e., normal and anomalous. In other related work, time series-based anomaly detection and classification were studied in [29, 30, 319, 320]. The state-of-the-art Time Series Classification (TSC) methods are based on deep learning models, as described in [29, 30]. However, based

on our experiments, they do not perform well in the detection of stealthy attacks due to infinitesimally small changes in the traffic throughput. Additionally, these methods do not perform well due to imbalanced data that is indicated in their F1 and Geometric Mean (Gmean) scores. Therefore, to address these challenges, we propose a novel hybrid deep learning model for anomaly detection in power grid OT network traffic. The hybrid model uses Graph Neural Networks (GNN), Long Short-Term Memory (LSTM), and CNN. It employs unsupervised learning to learn the complex behavior of OT network traffic throughput and supervised learning to classify the OT traffic.

GNN-based deep learning models have been implemented for various applications, e.g., residential load forecasting [321], detection of false data injection [322], road traffic prediction [323], and road traffic anomaly detection [324]. LSTM has been used to detect anomalies in Supervisory Control and Data Acquisition (SCADA) systems [325]. This method can detect anomalies based on temporal features of time series data. CNN has been proposed to detect anomalies in power system data [326]. It has advantages in learning spatial features and correlations of the datasets. In this research, we propose the application of a Graph-Convolutional Long Short-Term Memory (GC-LSTM) to preprocess the data of OT network traffic and generate traffic predictions. The output from the GC-LSTM is then used as an input for the CNN-based time-series classification. We generate an attack graph to identify in near real-time the active cyber attack locations in the power grid.

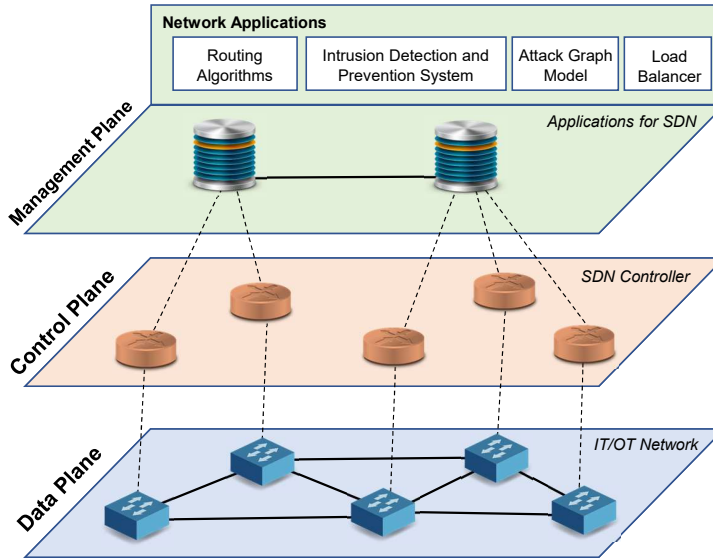


Figure 4.2: Abstraction layers of SDN architecture [5].

The attack graph provides topological information on the possible attack paths for a specific cyber attack on a given network. Hence, the attack graph is an important method to identify vulnerabilities in the system [327]. The knowledge about the attack path is also crucial to prevent and mitigate cyber attacks. At current, the attack graphs are mostly constructed based on vulnerability information obtained from network elements [328, 329].

This type of attack graph is not flexible, because it heavily depends on system vulnerability data. However, in this research, we propose an alternative attack graph map generation model, based on the online traffic monitoring in the OT networks of power grids. This is made possible through the wide deployment of an emergent technology, i.e., Software Defined Networking (SDN). SDN is a networking paradigm based on network virtualization and segregation of data and control planes [330]. In the SDN architecture, as seen in Fig. 4.2, there are three abstraction layers present, i.e., data plane, control plane, and management plane. The data plane represents locations of conventional communication networks, while control plane provides controllability over the data plane. Additionally, the management plane in SDN allows the deployment of network applications, e.g., attack graph model. Although SDN is an emergent paradigm in the field of computer networking, earlier research has investigated its implementation in cyber-physical power systems [331–335]. Earlier research has used SDN for anomaly detection based on traffic flow information [314, 336]. However, these works are not designed to detect anomalies triggered by cyber attacks in OT networks. In this research, we use SDN to monitor the network traffic in real-time, originating from the data plane of the OT Wide Area Network (WAN) for power systems. In summary, a critical examination of related state-of-the-art methods for communication traffic anomaly detection reveals the following. (1) Existing SDN applications for cyber-physical systems are not focused on cyber security of OT networks [314, 332–336]. (2) They are solely based on packet flow rules [336]. (3) They overlook the cyber kill chain and do not address any type of stealthy cyber attacks [314, 336].

The scientific contributions of this research are as follows:

1. To the best knowledge of the authors, we propose the first known SDN-based online cyber attack situational awareness method, i.e., Cyber Resilient Grid (CyResGrid). It is specifically designed for anomaly detection using communication traffic throughput in OT networks for stealthy cyber attacks during the early stages of the cyber kill chain, e.g., network reconnaissance. Therefore, CyResGrid aids operators to locate and identify power system-wide cyber attacks in near real-time through an attack graph map.
2. We propose a hybrid deep learning model to classify the OT network traffic throughput as anomalous or normal. The model combines GC-LSTM and a deep convolutional network to detect OT network anomalies caused by cyber attacks. It outperforms existing state-of-the-art deep learning-based time series classifiers [29, 30], as indicated by Gmean and F1 scores. To achieve this, we use GC-LSTM for traffic normalization. Subsequently, to detect the anomaly, we design a deep convolutional network by tuning the hyperparameters through Bayesian optimization. Based on the network throughput monitoring and anomaly detection, we create an attack graph map of power system-wide cyber attacks, in near real-time.
3. As there is a strong need for synthetic Cyber-Physical System (CPS) datasets for research [53], we create the first synthetic dataset of OT communication traffic throughput, which is generated through a cyber-physical power system model. To the best of our knowledge, the majority of the existing datasets are not suitable for cyber security [337–343]. A cyber-physical system dataset was proposed in

[344, 345] for intrusion detection. However, the OT traffic data is only in the form of signature-based logs without detailed traffic information [344, 345]. Therefore, in this research, we employ a CPS model of the power grid consisting of the physical system and associated OT communication networks. The model is used to co-simulate the power grid and OT network, from substations up to the control center. It also has cyber range capabilities to simulate various cyber attack scenarios. Based on this model, we generate a synthetic dataset of OT communication traffic throughput for cyber-physical power system operation under cyber attacks.

The chapter is structured as follows. Section I is the introduction and Section II describes the methodology proposed in this research, including cyber-physical system model, Traffic Dispersion Graph (TDG), GC-LSTM, TSC for anomaly detection, and the attack graph model. Section III provides the experimental results. Section IV presents the conclusions and future work.

## 4

## 4.2 ANOMALY DETECTION AND ATTACK GRAPH MODEL

In this section, the proposed methods for anomaly detection and attack graph modeling are introduced. Furthermore, we also elaborate on the cyber-physical model that serves as the basis for the aforementioned methods. Fig. 4.3 summarizes the methodology of anomaly detection and attack graph creation. The method consists of four steps as follows.

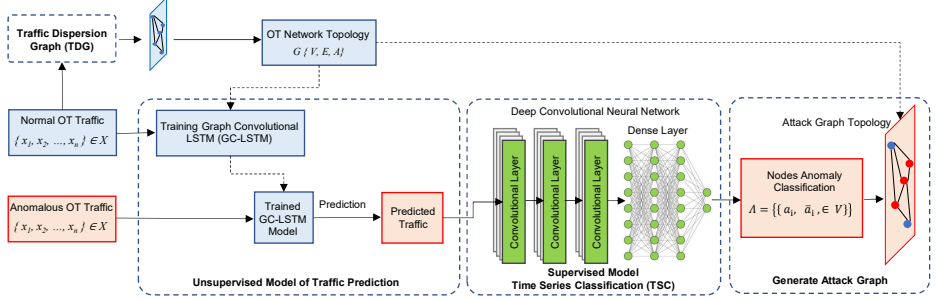


Figure 4.3: Attack graph creation using CyResGrid method [5].

**Step 1:** GC-LSTM training and TDG. The normal OT traffic is used to train the GC-LSTM model for traffic prediction. The process generates a trained GC-LSTM model. Additionally, the normal OT traffic is used to generate the OT network topology using a TDG.

**Step 2:** Deep CNN training. The trained GC-LSTM model is used to predict the OT network traffic. The prediction is then used to train a Deep Convolutional Neural Network for TSC. This process generates a trained Deep CNN model for OT traffic classification.

**Step 3:** Online node classification. This step monitors the online OT traffic as input for node classification. The trained GC-LSTM and Deep CNN are used sequentially to classify the nodes as normal or anomalous.



*Step 4: Attack graph generation.* The node classification results from step 3 in conjunction with OT graph data from step 1 are used to generate the attack graph visualization. A more detailed explanation of the method in each step is provided in the following subsections.

#### 4.2.1 CYBER-PHYSICAL SYSTEM MODEL

Detailed CPS models are needed for research on cyber security of power grids. They are used to simulate the power systems along with their associated IT-OT communication networks and cyber events. The state-of-the-art in smart grid modeling and simulations is discussed in [200, 346–351]. Hence, as part of our CPS model, we perform a co-simulation of the power grid and IT-OT systems, as depicted in Fig. 4.4.

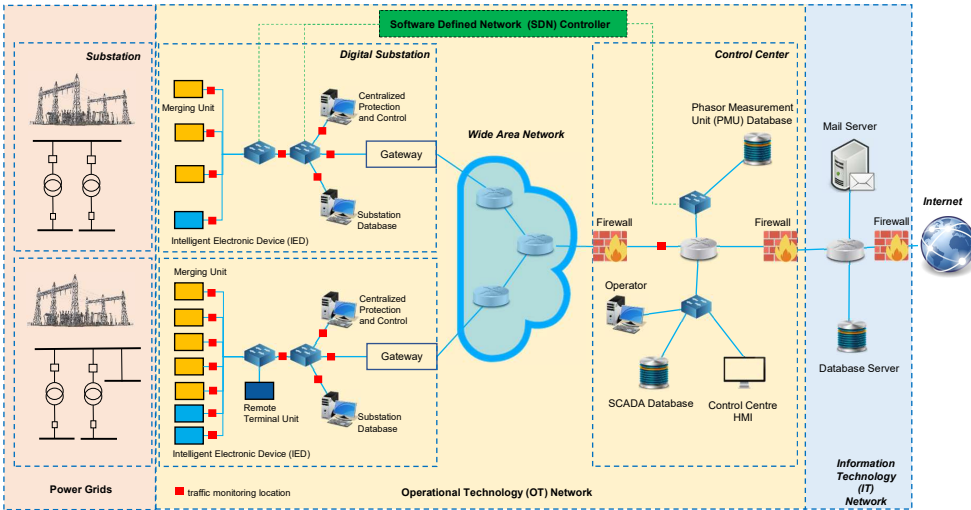


Figure 4.4: Cyber-physical system model of the power grid with IT-OT communication networks [5].

The CPS model provides time-domain measurement data from substation bays, e.g., buses, lines, and generators, in the form of active and reactive power, voltage, and current measurements. All measurement data is then delivered from the substation to the control center via a WAN as SCADA telemetry. The SCADA data is also stored in local databases located in substations and the control center. For the cyber system, every node in the OT network is emulated using operating system-level virtualization. The network connectivity between substations, WAN, and control center is realized through network virtualization and SDN. With this configuration, the developed CPS architecture can model and simulate realistic OT network traffic for the power system.

The OT network is modeled based on custom functions for every device in the communication network. The measurement devices represent components, such as Merging Units (MUs), Remote Terminal Units (RTUs), and Intelligent Electronic Devices (IEDs). These devices perform data acquisition from the power grid, with a SCADA sampling rate of one sample per second. Legitimate control commands from the control center modify the

set points for power grid controllers in real-time. For example, a control command can set a circuit breaker to open or close, set values for voltage, and active power set points of generator automatic voltage regulators and governors. The measurement values and control set points are communicated across the OT network using Transmission Control Protocol/Internet Protocol (TCP/IP) packets.

The CPS model is integrated with SDN capability that creates network virtualization using virtual switches. Based on Fig. 4.4, the OT and IT networks are present in the data plane layer of the SDN. Meanwhile, the control and management plane are represented by the SDN controller. Network virtualization allows the SDN controller to monitor and control traffic and run custom network applications. Fig. 4.4 depicts how the SDN controller is applied to the typical SCADA architecture. SDN improves the OT network monitoring and control by collecting OT communication traffic reports in the control center. The traffic observation points are visualized as red squares, which are distributed across the substations and control center. Using these points, we observe real-time OT network traffic from the control center to detect traffic anomalies for each observation location and create a power system wide attack graph.

## 4

#### 4.2.2 TRAFFIC DISPERSION GRAPH

The TDG is an analytical model for communication traffic monitoring and analysis. The core idea for TDG is derived from the social behavior of hosts in a network [352]. Therefore, the flow of OT network traffic is analyzed based on the interactions between all hosts in the communication network. Based on this analysis, information related to communication sources and destinations is extracted. Furthermore, TDG represents nodal information using graph structures. Every host in a network is represented by a single node in a graph. On the other hand, communication between hosts is represented by connectivity between nodes, i.e., graph edges. Fig. 4.5. shows the TDG generation processes. Firstly, information on the IP address source and destination from flowing packets in the network is in the collected information table. Information about the path between two IP addresses is added based on prior knowledge of the network topology. The information in the table is then used to create an individual flow graph. Finally, all individual graph is converged into a dispersion graph which provides an overall topology of the network.

The TDG has previously been used to analyze communication network patterns. For example, a research proposed an application of TDG for anomaly detection based on the degree distribution values of a graph [353]. In our research, the CyResGrid method uses TDG to generate graph structures of the power system OT network. This includes a graphical representation of the OT network topology between the control center and substations. The anomalous nodes in the graph are then detected based on OT network traffic anomalies. In our model, the CPS topology of a power grid possesses a tree-like network structure. Fig. 4.6 illustrates the TDG of the OT network that is used in our model, containing a total of 27 substations and one control center. Every substation consists of OT devices, e.g., MUs, IEDs, RTUs, etc., and a communication gateway, e.g., router/firewall, that communicates with the control center.

In this research, the nodes represent traffic observation locations, while edges represent communication links between nodes. The traffic observation locations are situated in the Ethernet ports of virtual SDN switches that are directly connected to a host. All

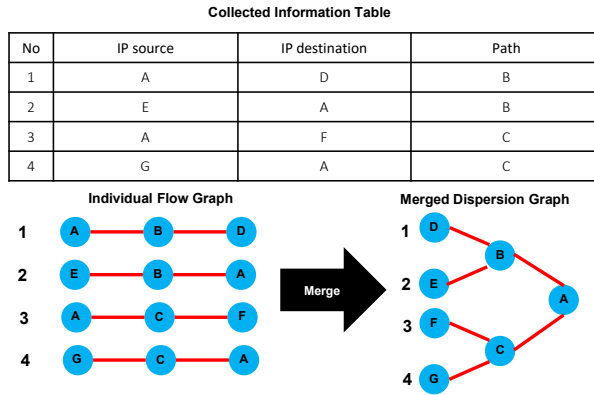


Figure 4.5: Traffic Dispersion Graph (TDG) processes [5].

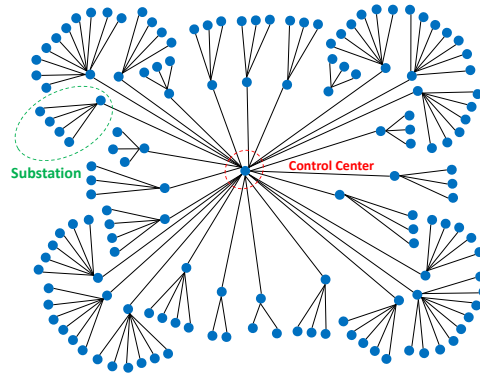


Figure 4.6: Traffic dispersion graph of 27 substations [5].

measurement data from each substation is sent to the control center via SCADA protocols, e.g., IEC 104 and DNP3. Thereby, this traffic flow allows the control center to gain a complete overview of the entire OT network. Using observation locations in the control center, the dispersion graph determines the nodes that actively communicate measurements. Also, the dispersion graph can determine unusual behavior, i.e., when a node is not sending measurement data or sending an abnormal quantity of traffic. In this research, anomaly detection works based on the total volume of observed network traffic, i.e., throughput, measured in KiloBytes per second (KBps). Furthermore, the dispersion graph can also identify unknown nodes with unidentified or unknown sources and destinations of IP Addresses or MAC Addresses.

### 4.2.3 GRAPH CONVOLUTIONAL LONG SHORT-TERM MEMORY

GC-LSTM aims to learn the traffic behavior of the OT network. Two machine learning models are applied in GC-LSTM, i.e., Graph Convolutional Network (GCN) and LSTM. GCN processes the OT network topological information expressed as a graph, along

with localized features from neighboring communication nodes in the spatial domain. Subsequently, LSTM performs temporal learning based on time-series data of observed OT network traffic. The combination of GCN and LSTM has the advantage of learning from both the spatial and temporal domains. Various applications using graph-based spatial and temporal models were proposed in [321–324]. In this research, we propose a novel method for nodal feature prediction based on communication network topology and features of neighboring nodes. CyResGrid proposes an innovative application of GC-LSTM to model the OT network traffic of the power system. It uses a hybrid combination of unsupervised and supervised models for OT traffic anomaly detection. The former is based on GC-LSTM which learns the complex behavior of OT network data and topology. Subsequently, the GC-LSTM generates traffic for the supervised predictions of the TSCs. The OT traffic model is then integrated with deep convolutional network-based TSC to generate an attack graph based on observed anomalies in the communication network traffic.

## 4

The graph structure of the OT network topology serves as the main input for GC-LSTM method. This graph structure is obtained from the TDG. It can be represented as  $G = (V, E)$  where  $G$  is the graph,  $V$  represents the vertices/nodes and  $E$  represents the edges/links. The connection between the nodes in the graph is represented by the adjacency matrix  $A$ . Elements of the adjacency matrix are represented by  $A_{i,j}$  where  $i$  and  $j$  represent the node index numbers, such that  $A_{i,j} = 1$  when two nodes are connected, and  $A_{i,j} = 0$  otherwise.

The GCN function is used to obtain the nodal features as described in (4.1). GCN operates based on the Hadamard product multiplication ( $\odot$ ) of the weight matrix ( $W_{gcn}$ ), adjacency matrix ( $A$ ), and node features from the observed traffic data ( $X_t$ ). The adjacency matrix captures information related to the OT network topology. The adjacency matrix ( $A$ ) is added with the identity matrix ( $I$ ) to form a modified adjacency matrix ( $\hat{A}$ ). The data set ( $X_t$ ) is represented as a time series, where the equation considers the single time instant ( $t$ ) and total number of time observations,  $T$ . The node feature matrix ( $X$ ) contains individual nodal information ( $x_i$ ), where the total number of nodes is represented by ( $n$ ). The equation also considers the number of hops from a communication node to neighboring nodes, i.e.,  $k$  as an exponent of , as explained in [323, 354]. This research uses the maximum number of hops between each substation and the control center being two, i.e.,  $k = 2$ .

After obtaining the spatial features from the graph convolutional operation, LSTM is then used to analyze the temporal / time-series features. The LSTM functions and processes inside an LSTM cell are described in Eqs. 4.2-4.7. There are six main sub-equations in the LSTM process, including the forget gate ( $f_t$ ), input gate ( $i_t$ ), output gate ( $o_t$ ), internal cell state ( $c'_t$ ), transferable cell state ( $c_t$ ), and hidden state ( $h_t$ ). The previously calculated nodal features output (GCNtk) serves as the input for the LSTM cell.

In this work, we consider each substation to have unique characteristics. Given the communication network traffic data from all nodes that are present in a substation as ( $X$ ), Algorithm 1 describes how an independent process is performed for each substation to provide the independent set GC-LSTM models for every substation ( $s_i$ ). During the training process, this output is compared with the real OT traffic data ( $X_{t+1}$ ) to update the weight values in GCN and LSTM. The final output of LSTM predicts the OT traffic in corresponding nodes represented by ( $h_t$ ) in (4.1).

$$GCN_t^k \leftarrow (W_{gc} \odot \hat{A}) X_t \quad (4.1)$$

$$f_t = \sigma (W_f GCN_t^k + U_f h_{t-1} + b_f) \quad (4.2)$$

$$i_t = \sigma (W_i GCN_t^k + U_i h_{t-1} + b_i) \quad (4.3)$$

$$o_t = \sigma (W_o GCN_t^k + U_o h_{t-1} + b_o) \quad (4.4)$$

$$\tilde{c}_t = \tanh (W_c GCN_t^k + U_c h_{t-1} + b_c) \quad (4.5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (4.6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (4.7)$$

#### 4.2.4 TIME SERIES ANOMALY DETECTION

TSC for anomaly detection was studied in [29, 30, 319, 320]. In this research, we propose a new method using TSC to detect anomalies in the OT communication network traffic throughput for power systems. As a benchmark, we focus on state-of-the-art deep learning-based anomaly detection techniques, i.e., ResNets [355], Inception [30], Fully Convolutional Neural Network (FCN) [356], and Multi-Layer Perceptron (MLP) [357]. Meanwhile, in our research, we propose CyResGrid; a hybrid of method for unsupervised and supervised OT traffic anomaly detection. The unsupervised learning application for time series data was studied in [358]. We specifically use an unsupervised GC-LSTM model to learn the complex behavior of OT network data and topology. Subsequently, the GC-LSTM generates traffic predictions as inputs to TSCs.

$$y_i^l = \text{ReLU} \left( \sum_{j=1}^{m-1} w_j y_{(i)}^{l-1} + b \right) \quad (4.8)$$

$$x^* = \arg \max_x f(x) \quad (4.9)$$

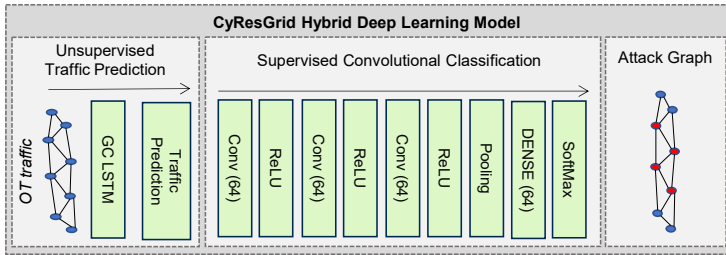


Figure 4.7: CyResGrid – hybrid deep learning model [5].

We propose a supervised deep convolutional neural network for TSC-based anomaly detection. The deep convolutional network is based on a multi-layer one-dimensional convolutional with the ReLU activation function as shown in (4.8). In (4.8), we consider the number of layers ( $l$ ), filter size ( $m$ ), weight ( $w$ ), and bias ( $b$ ). This model is trained

to optimize the performance of classification based on the previous GC-LSTM output. To formulate our hybrid deep learning model, we perform hyperparameter tuning based on the number of layers, filters, and kernel size. Bayesian optimization [359] is used to optimize the deep learning model. The objective function maximizes the deep learning performance as described in (4.9). Bayesian optimization works based on the surrogate model and acquisition function. The surrogate model is a Gaussian process that quantifies the uncertainty of the unobservable region. To achieve the optimum value of the objective function, we use the Expected Improvement (*EI*) as the acquisition function. Bayesian optimization performs iterations to obtain a function with the best performance. From the iterative process, we obtain the best performing deep convolutional network that has 3 layers, 64 filters, and 3 kernel sizes. Fig. 4.7 shows the architecture of CyResGrid hybrid deep learning model that consists of a GC-LSTM layer, three layers of convolutional neural network, and one layer of fully connected neural network (dense).

4

#### 4.2.5 ATTACK GRAPH MODEL

An attack graph is a method to model CPS vulnerabilities and potential exploits. Since a successful exploit of a vulnerability may lead to a partial or even a total failure of the CPS, an attack graph is an important tool for vulnerability analysis and mitigation strategies. Meanwhile, in a communication network, there are many hosts that may become vulnerable. As a result, the cyber security of the entire CPS cannot only rely on the security of a single host. Therefore, it is important to locate and identify all vulnerable nodes/hosts in a communication network as a set of potential threats in the CPS. Subsequently, in this research, we propose the observation and analysis of anomalous OT traffic behavior to detect nodes potentially compromised by cyber attacks. The information regarding anomalous nodes is then used to construct an online attack graph in near real-time for the entire OT network of the power grid.

Algorithm 1 explains the process of attack graph generation. The OT network traffic ( $X$ ) is the input for the algorithm. The network traffic from each substation ( $X_n$ ) is used to predict the OT traffic using GC-LSTM. The GC-LSTM model provides a set of traffic predictions ( $h_t$ ) as outputs. The output from the prediction is then used as input for the TSC-based CNN. The time series-based anomaly detection is performed for each node ( $a$ ) in  $V$ . The classifier labels each node as anomalous or normal based on the input OT traffic prediction. This information is then used to construct the attack graph.

$$\Lambda = \{\{a_i, \bar{a}_i, \in V\}\} \quad (4.10)$$

$$\Lambda = \{\{a_i, \bar{a}_i, \in V\}, \{u_i \notin V\}\} \quad (4.11)$$

There are two types of attack graphs as described through equations (4.10) and (4.11). The attack graph type I in (4.10) is constructed based on prior knowledge of the OT network topology and node classification results. Meanwhile, the attack graph type II in (4.11) considers unidentified nodes based on the TDG. There are two elements of attack graph ( $\Lambda$ ) type I as indicated in (4.10), i.e., normal nodes ( $a_i$ ), and anomalous nodes ( $\bar{a}_i$ ). Both of the nodes are elements of the known nodes ( $V$ ). In contrast, attack graph ( $\Lambda$ ) type II as indicated in (11) contains one extra element of unidentified nodes ( $u_i$ ). The unidentified

nodes are considered as anomalous since these nodes are not elements of the known nodes ( $V$ ).

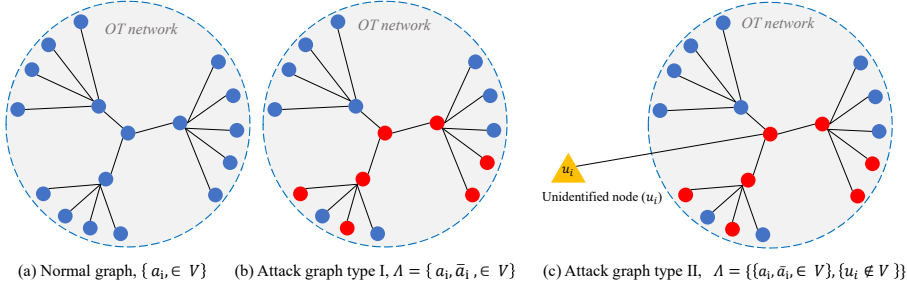


Figure 4.8: CyResGrid – hybrid deep learning model [5].

---

**Algorithm 1** CyResGrid Attack Graph Generation
 

---

**Input:**  $S\{s_1, s_2, \dots, s_n\}$ ;  $X \in s_i$ : Substations traffic data  
 $\{x_1, x_2, \dots, x_n\} \in X$ : Nodes traffic data

**Output:**  $\Lambda = \{\{a_i, \bar{a}_i, \in V\}\}$ : Nodes classification as attack graph

---

**for** each substation  $s_i$  in  $S$  **do**

**for**  $t = 1$  to  $T$  **do**

        Traffic prediction

$$GCN_t^k \leftarrow (W_{gc} \odot \hat{A}^k)X\{x_1, x_2, \dots, x_n\}_t$$

$$h_t, c_t = \text{LSTM}(X\{x_1, x_2, \dots, x_n\}_t, GCN_t^k, h_{t-1}, c_{t-1})$$

**for** each node  $a$  in  $V$  **do**

            Node classification

$$\bar{a}_i = \sum_{j=1}^{m-1} w h_{i(j)}^{-1} + b$$

**return**  $\Lambda = \{\{a_i, \bar{a}_i, \in V\}\}$

---

Fig. 4.8 depicts an example comparison of attack graph representations of the OT network under normal network traffic conditions in Fig. 4.8(a) and anomalous traffic in Figs. 4.8(b) and 4.8(c). The anomalous network traffic conditions are determined based on observed abnormal node behavior shown in red. Subsequently, these nodes are combined to form an attack graph ( $\Lambda$ ). There are three elements in the attack graph, i.e., normal nodes ( $a_i$ ), anomalous nodes ( $\bar{a}_i$ ), and unidentified nodes ( $u_i$ ). The attack graph type I from Fig. 4.8(b) only classifies nodes as anomalous based on observed traffic from all known nodes. This notion is represented by a set of attack graphs ( $\Lambda$ ) and described through (4.10). On the other hand, the attack graph type II in Fig. 4.8(c) also considers all unidentified nodes for the classification of anomalous behavior, as described in (11). The unidentified nodes ( $u_i$ ) are determined based on unknown sources or destinations address obtained



from the TDG. The unknown nodes ( $u_i$ ) are assumed to indicate an active cyber attack, originating from an unlisted host in the known OT network ( $V$ ).

#### 4.2.6 FORENSIC GRAPH MODEL

The aforementioned attack graph model is implemented based on the SDN for enabling wide area traffic monitoring in near real time. However, this implementation is restricted by the limited adoption of SDN in the present power system [360]. Therefore, as an alternative, this section proposed a novel forensic graph model (*FGraph*). The *FGraph* analyze OT traffic throughput based on packet historical data. The methods of the *FGraph* is the same with the attack graph. However, instead of using real time network traffic, it is using historical traffic captured using Wireshark as a .pcap file. Based on the historical traffic, the *FGraph* the traffic as a normal or anomalous.

4

Network forensics pertains to the acquisition, preservation, and scrutiny of network data with the aim of identifying unauthorized access and conducting subsequent inquiries [361]. It is a crucial component of network security, as it enables organizations to quickly detect and respond to cyber threats. Network administrators typically employ network traffic analysis tools to perform network traffic forensics, which involves capturing and analyzing traffic data in real-time or from historical traffic logs. These tools aid in detecting network anomalies, such as abnormal traffic patterns or unauthorized access attempts, that may suggest security breaches or malware infections. *Wireshark*, *Tshark*, *Snort*, and *tcpdump* are well-known software tools for network traffic analysis. These tools can capture network traffic data and provide a comprehensive analysis of the data, including the source and destination of the traffic, traffic type, and any detected anomalies or suspicious activities.

One of the methods to perform a deeper forensic analysis is through network forensic data visualization [362]. A matrix-based visualization from network forensic data was presented in [363]. The authors show the visualization summary of network data, e.g., IP addresses, ports, NetFlow payloads, entropy of source and destination IP, etc. The visualizations help to facilitate network traffic analysis and pinpoint anomalies within the network. An alternative method to visualize the network traffic data is using a TDG. The TDG is an analytical framework utilized for the purpose of observing and evaluating communication traffic. The fundamental concept behind TDG is interactions between hosts within a network [352]. Moreover, TDG employs graph structures to represent nodal information. Each individual node in a graph represents an individual host within a network. Conversely, the transmission of information among hosts is denoted by the inter-connectivity of nodes, i.e., graph edges. Previously, the TDG was utilized to analyze communication network patterns. For instance, studies in [353] proposed an application of TDG for anomaly detection, based on graph information from network traffic. As shown in Fig. 4.9, in this research, we use TDG to generate a network graph topological representation from recorded OT traffic data.

Besides the aforementioned TDG, we also implement Traffic Pre-Processing (TPP) in the model for the historical packets. This extracts information from the packets, i.e., nodes, edges, and time series traffic throughput. Algorithm 2 summarizes the pseudocode of both TDG and TPP. The input for the proposed algorithm is historical traffic packets ( $P$ ) captured using Wireshark or Tshark. TDG processes the OT traffic to extract Graph information ( $G$ ) from the packets, including vertices/nodes ( $V$ ), edges ( $E$ ), and the adjacency matrix



(A). Meanwhile, TPP aims to convert the packets into time series throughput data for each node ( $X$ ). The extracted graph ( $G$ ) and time series throughput ( $X$ ) serve as input for the subsequent forensic graph stages.

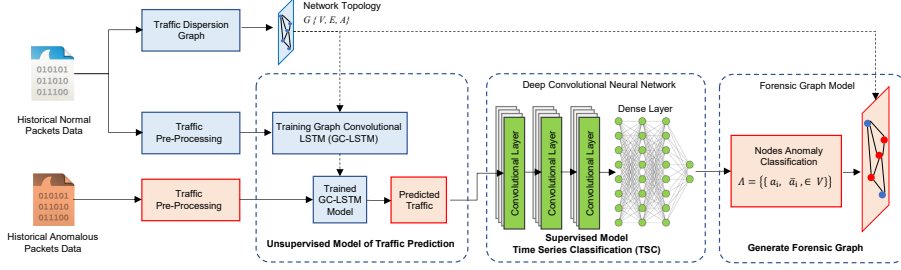


Figure 4.9: Forensic graph ( $FGraph$ ) model [5].

4

---

#### Algorithm 2 TDG and TPP Processes

---

**Input:**  $P$ : Historical communication traffic packets  
**Output:**  $G = \{V, E, A\}$ : Graph with nodes, edges, and adjacency  
 $\{x_1, x_2, \dots, x_t\} \in X$ : Time series throughput data

---

*TDG iteration for each packet  $p$  in  $P$*

```

for  $p$  in  $P$  do
    if  $v$  not in  $G\{V\}$  then
        add  $v$  to  $V$ 
    if  $e$  not in  $G\{E\}$  then
        add  $e$  to  $E$ 

```

*TPP throughput extraction iteration for each time  $t$  in  $T$*

```

for  $t$  in  $T$  do
    for  $v$  in  $G\{V\}$  do
         $x_v^t = \sum x_v$ 

```

**return**  $G = \{V, E, A\}$  and  $\{x_1, x_2, \dots, x_t\} \in X$

---

## 4.3 EXPERIMENTAL RESULTS

### 4.3.1 EXPERIMENTAL SETTING

All experiments in this research are conducted using the previously discussed CPS model of the power grid represented in Fig. 4.4. The power system is simulated in real-time using a Root Mean Square (RMS) dynamic model of the IEEE 39-bus test system in DiGSILENT

PowerFactory. The CPS model employs OPC UA implemented through Python to interface the time domain simulation of the power grid and emulated OT communication network. The OT network emulation is based on Mininet<sup>1</sup>, which uses the operating-system-level virtualization. The entire emulated OT network runs on 10 virtual servers and consists of 27 user-defined substations, 118 measurement devices, and over 800 data points for the entire simulated power system. SCADA device functionality within the OT network is realized through custom Python code. Therefore, we generate SCADA traffic from substations and the control center. All OT network traffic is captured using the Linux *bwm-ng*<sup>2</sup> tool and used as the main dataset for this research. The OT network traffic is measured in KBps. The observed OT network traffic data under nominal operating conditions is used to train the GC-LSTM model.

We collect OT network traffic data during various cyber attack scenarios. Two types of cyber attacks are considered, i.e., DDoS and active reconnaissance, i.e., OT network scanning. The DDoS attack is launched to target multiple substations and aims to disrupt the power system operation with a malicious increase of the OT network traffic loading. To this end, we use the well-known Syn Flood cyber attack vector that exploits vulnerabilities in the TCP/IP packets to target network hosts [364]. This attack vector is chosen as it can flood the OT network and cause the targeted hosts to crash. The DDoS attack is executed using the Linux *hping3*<sup>3</sup> tool. The second examined cyber attack scenario is based on OT network scanning. This attack aims to enumerate active hosts within the OT network. Network scanning targets IP addresses and ports within a specified range. It is typically performed during reconnaissance at the early stages of a cyber attack kill chain. In this work, we conduct a six-level network scanning using *nmap*, i.e., paranoid, sneaky, polite, normal, aggressive, and insane. The first two scanning levels are stealthy and used to evade intrusion detection systems [365]. The scanning intensity determines the number of packets delivered to the network. For all cyber attack scenarios and simulations, we collect the observed OT network traffic data into a labeled dataset for deep learning applications.

### 4.3.2 NETWORK TRAFFIC PREDICTION

In this research, The GC-LSTM model is implemented using the PyTorch<sup>4</sup> library. The training of the GC-LSTM model is performed using the simulated OT network traffic dataset. This dataset consists of operational data for 27 substations, resulting in a total of 146 columns and  $25 \times 10^4$  rows. The number of columns represents the total number of traffic observation points in the OT network. On the other hand, the number of rows in the dataset represents the temporal observations. The sampling rate for all observations is 1 sample/second. Therefore, the dataset for normal OT traffic is collected for a total duration of  $25 \times 10^4$  seconds. The training was performed using a computer with the following specifications: Intel® Xeon® CPU 3.60GHz, 64 GB of RAM, and an NVIDIA Quadro RTX 4000 graphics processing unit. During the training process, the OT observation points are further classified for each individual substation to create 27 independent models of traffic predictions. The total training time for all 27 substations is 26.5 hours.

<sup>1</sup> <https://mininet.org/>

<sup>2</sup> <https://linux.die.net/man/1/bwm-ng>

<sup>3</sup> <https://linux.die.net/man/8/hping3>

<sup>4</sup> <https://pytorch.org/>

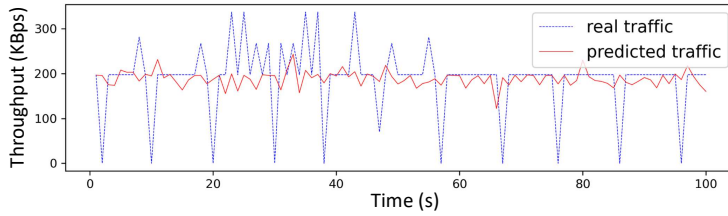


Figure 4.10: Comparison of real and predicted traffic under normal conditions [5].

Fig. 4.10 shows the comparison of the real OT traffic under normal conditions and GC-LSTM predicted traffic in node 2, substation 7. The observed traffic rate is around 197 KBps. However, occasionally, the real OT traffic slightly increases or drops to zero but we cannot consider this situation as an anomaly. In distributed communication systems, the zero-value and variability happen because of the latency and delay that lead to variations in the packet arrival time. These factors are common phenomena for distributed communications, which have been studied in [80]. The zero value in Fig. 9 represents zero in Fig. 13. On average, the observed OT traffic data contains 3.6% of zeroes.

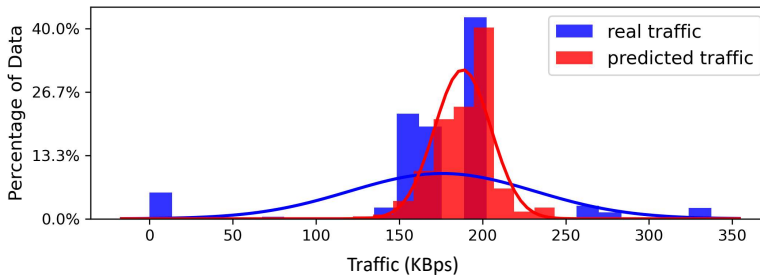


Figure 4.11: Histogram of real and predicted traffic under normal conditions [5].

Fig. 4.11 presents the histogram and probability distribution of the real and predicted OT traffic in node 2, substation 7. Fig. 4.11 shows that the predicted OT traffic is more concentrated. We also compare the normal and predicted OT traffic for nodes 1 to 5 in substation 7 as represented in Fig. 4.12. The box plot in Fig. 4.12 shows the statistical summary from the traffic data including the minimum, median, maximum, first quartile, and third quartile. The box plot also indicates the variability, spread, and skewness of the data. The circles in the plot indicate the outlier data. Based on the plots in Fig. 4.10-Fig. 4.12, the predicted OT traffic has a more concentrated value and fewer outliers compared to the real data. Therefore, the GC-LSTM performs as a filter to normalize and reduce the variability and outliers traffic.

Fig. 4.13 shows the comparison of the real and predicted OT traffic during a sneaky cyber attack. The cyber attack triggers a higher spike in OT traffic. The time series-based anomaly detection is then expected to distinguish the spikes due to traffic variability and cyber attacks. Therefore, the GC-LSTM-based prediction is important to normalize the OT traffic and reduce data variability on the predicted traffic. This is then used to improve the

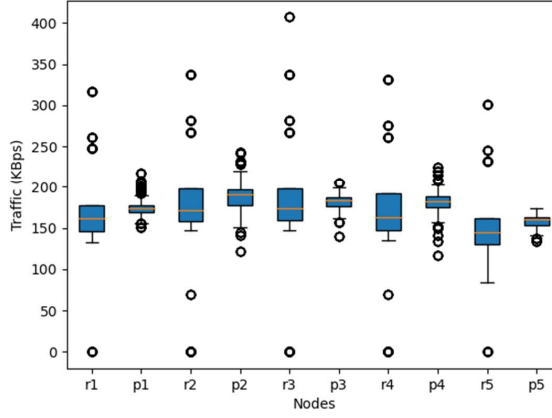


Figure 4.12: Statistical comparison of real (r) and predicted traffic (p) [5].

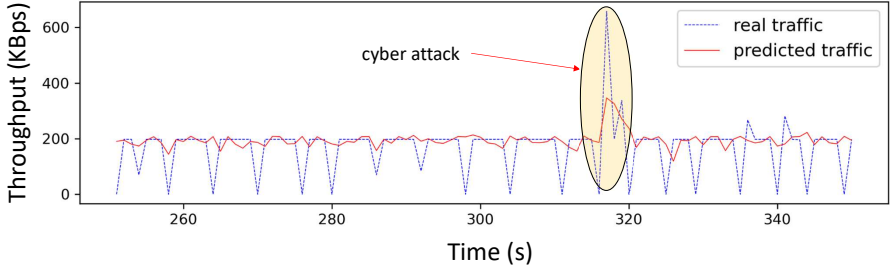


Figure 4.13: Comparison of throughput between real and predicted OT traffic for sneaky network scanning cyber attack scenario [5].

anomaly detection accuracy of TSC.

### 4.3.3 ANOMALY DETECTION

To perform anomaly detection on the OT traffic, we generate a dataset with network traffic ( $X$ ) and labels ( $L$ ) for univariate TSC. This is depicted in Fig. 4.14. Each column ( $x_n$ ) in the observed traffic data has one associated label column ( $l_n$ ). A label value of zero corresponds to the normal operation, while one represents anomalous OT traffic. We simulate two types of cyber attacks to generate anomalous traffic, i.e., DDoS and OT network scanning during the reconnaissance stage of the cyber kill chain. The attack scenarios are summarized in Table 4.1. There are nine variations in the intensity of the communication network scanning amongst the scenarios. In total, the cyber attacks run for 345,000 seconds, and data is collected every second to create the dataset, as represented in Fig. 13, from  $t = 1$  until  $t = 345,000$ . This dataset is then used to train 70% and test 30% of the TSC algorithm.

Using the same generated dataset, we compare our proposed CyResGrid method with four state-of-the-art deep learning-based TSC techniques for anomaly detection, i.e., ResNets [355], Inception [30], FCN [356], and MLP [357]. These deep learning mod-

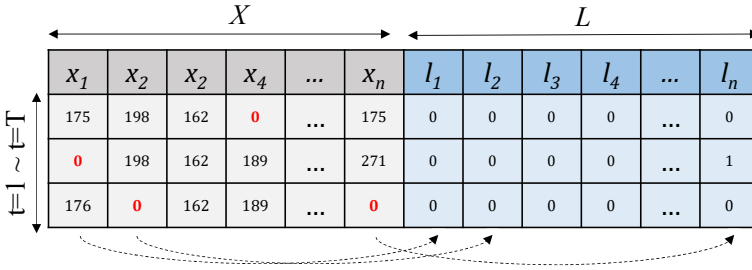


Figure 4.14: Dataset for time series classification [5].

Table 4.1: Cyber Attack Scenarios

Attack Type	Intensity	Tool	Time (s)	Duration (s)
DDoS	High	<i>hping3</i>	30,000	30,000
	Medium	<i>hping3</i>	30,000	30,000
	Low	<i>hping3</i>	30,000	30,000
Reconnaissance	Paranoid	<i>nmap</i>	75,000	75,000
	Sneaky	<i>nmap</i>	50,000	50,000
	Polite	<i>nmap</i>	40,000	40,000
	Normal	<i>nmap</i>	30,000	30,000
	Aggressive	<i>nmap</i>	30,000	30,000
	Insane	<i>nmap</i>	30,000	30,000

els are chosen as they address the general time series classification problem and are not domain-specific. This makes them suitable for benchmarking and comparison of various TSC methods. Additionally, we also combine them with the proposed GC-LSTM method and test their performances, as summarized in Table 6.2.

In Table 6.2, we classify the cyber attacks into two scenarios. The first is for all combined attacks, i.e., no. 1-9, and the second only focuses on stealthy attack scenarios, i.e., paranoid and sneaky attacks no. 10-16. We consider the test dataset as imbalanced because, for the combined attacks, only 6.4% of the data is labeled as an anomaly. Meanwhile, for the stealthy attacks, only 2.7% of the data is labeled as an anomaly. Therefore, to evaluate the anomaly detection performance, we use as metrics the Gmean in Equation (4.12) [366] and F1 score in Equation (4.13) [367, 368]. From Table 6.2, it is clearly seen that for the combined attack scenario, CyResGrid provides the best performance with the highest scores in the Area Under the Curve (AUC), accuracy, G mean, and F1. Meanwhile, for the stealthy attack dataset, we ignore the MLP method due to its lower performance. For this scenario, Inception seems to provide the best AUC and accuracy. However, its true positive rate is significantly low. Furthermore, its F1 and G mean score are amongst one the lowest. Therefore, we can still conclude that CyResGrid provides the most balanced performance, even for stealthy attack detection.

$$G_{\text{mean}} = \sqrt{\text{true positive rate} \cdot \text{true negative rate}} \quad (4.12)$$

$$F1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

(4.13)

Table 4.2: Performance Comparison of Anomaly Detection Methods

No	Methods	AUC	TN	FP	FN	TP	Accuracy	F1	G mean	t(s)
Combined attack scenarios										
1	ResNet	0.849	82.27	11.32	3.49	2.92	85.19	28.29	15.50	633
2	Inception	0.961	93.50	0.20	4.10	2.31	95.71	51.76	14.68	976
3	FCN	0.955	88.16	5.43	3.92	2.49	90.65	34.76	14.81	1016
4	MLP	0.758	72.22	21.37	4.86	1.55	73.77	10.55	10.57	113
5	GC-LSTM + ResNet	0.974	93.29	0.31	3.27	3.14	96.42	63.77	17.12	1056
6	GC-LSTM + Inception	0.976	92.10	1.49	3.35	3.06	95.16	55.87	16.79	1409
7	GC-LSTM + FCN	0.972	92.28	1.30	3.68	2.73	95.01	52.26	15.87	1342
8	GC-LSTM + MLP	0.937	93.40	0.19	6.13	0.28	93.68	8.14	5.12	765
9	CyResGrid	0.984	93.47	0.13	3.42	2.99	96.45	65.03	17.16	714
Stealthy attack scenarios										
10	ResNet	0.863	86.94	12.02	0.96	0.08	87.02	1.26	2.69	91
11	Inception	0.9887	98.93	0.02	1.04	0.0004	98.93	0.09	0.22	224
12	FCN	0.9833	87.82	11.13	1.01	0.02	87.85	0.47	1.58	240
13	GC-LSTM + ResNet	0.9524	89.93	0.02	0.95	0.09	90.02	1.87	2.92	226
14	GC-LSTM + Inception	0.9489	89.96	0.99	0.95	0.10	90.05	1.87	2.92	303
15	GC-LSTM + FCN	0.9491	89.96	0.99	0.95	0.10	90.05	1.87	2.92	304
16	CyResGrid	0.9243	91.15	7.81	0.94	0.111	91.25	2.32	3.08	138

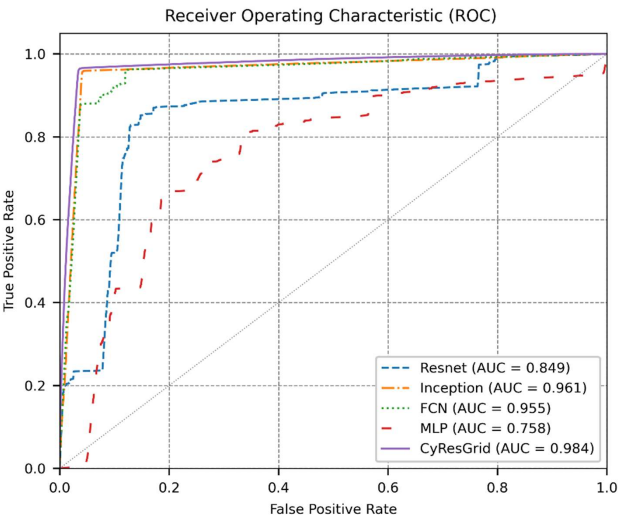


Figure 4.15: ROC comparison of the deep learning-based TSC [5].

Table 6.2 also indicates that GC-LSTM hybrid models can significantly improve the performance of deep learning-based classification, as indicated in row number 5, 6, 7, 8, 13, 14, and 15. The performance comparisons are also shown in Figs. 4.15 and 4.15. The Receiver Operating Characteristic (ROC) curve shows the performance of the classifier. The hybrid classification integrated with GC-LSTM provides improved result, as seen in Fig. 4.16, in comparison to the one without GC-LSTM in Fig. 4.15. According to Figs. 4.7-

4.10, the actual OT traffic data is noisier compared to the predicted one. This condition leads to better anomaly detection using the hybrid model, as described above.

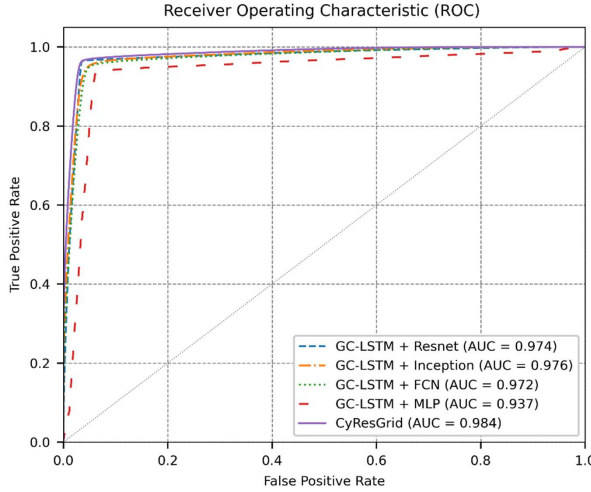


Figure 4.16: ROC comparison of the hybrid deep learning-based TSC [5].

#### 4.3.4 ATTACK GRAPH GENERATION AND ANALYSIS

As discussed in previous section, the attack graph is modeled by comparing the normal and anomalous OT traffic. The result of this comparison is then used to determine the nodal abnormality. The attack graph classifies nodes into two categories, i.e., normal and anomalous. Anomalous nodes ( $\hat{a}_i$ ) are indicated by red, while normal nodes ( $a_i$ ) are highlighted in blue.

Fig. 4.17 illustrates the entire attack graph map for online cyber attack identification and visualization. Fig. 4.17 (a) depicts OT network scanning, originating from the control center to an OT device in substation 7. Consequently, this leads to the control center, substation 7 gateway, and targeted OT device to be flagged as anomalous, as shown in red. Fig. 4.17 (b) depicts a DDoS attack targeting substations 1-7 that originates from the control center. The DDoS attack on multiple substation targets triggers widespread traffic anomalies in substations 1-7, as indicated in red. It is considerably easier to detect a DDoS attack, as it results in notably increased OT network traffic volume, in comparison to a network scanning attack. Fig. 4.17 (c) and (d) depict attack graphs for cyber attacks originating from other sources than the control center. In Fig. 4.17 (c), we highlight OT network scanning performed by a compromised OT device located in substation 7. The scanning attacks lead to all nodes in substation 7 being classified as anomalous, except the router gateway. This scenario is explained as a local cyber attack that occurs in a substation. Finally, Fig. 4.17 (d) shows OT network scanning by an unidentified node, as indicated by an orange triangle. The attack source is classified as unidentified because it is not included on the list of known nodes in the OT network.

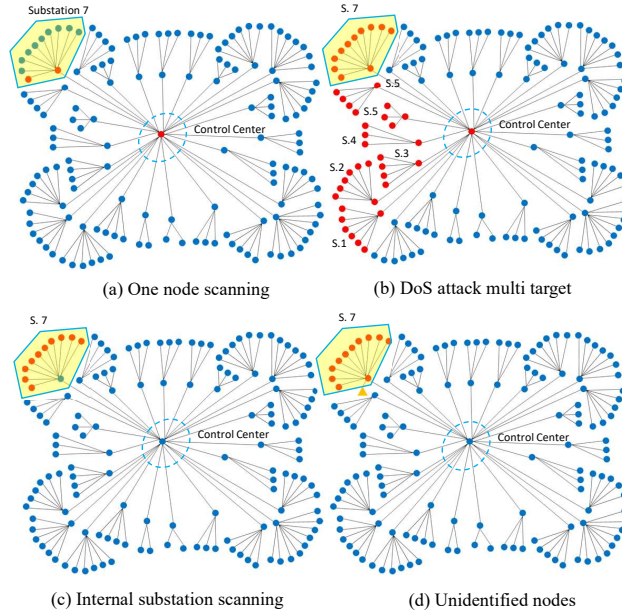


Figure 4.17: Attack graph maps to identify and visualize cyber attack locations [5].

### 4.3.5 FORENSIC GRAPH GENERATION AND ANALYSIS

#### EXPERIMENTAL HARDWARE-IN-THE-LOOP SETTING

Fig. 4.18 depicts the Hardware-in-the-Loop (HIL) configuration utilized for performing the Forensic Graph (FGraph) implementation. A Real-Time Digital Simulator (RTDS) is used to model the physical power system, while IEC 61850 communication is realized between the RTDS and Intelligent Electronic Devices (IEDs) through a network switch. The IEDs comply with the IEC 61850 standard, enabling Generic Object Oriented Substation Event (GOOSE) messaging and Sampled Values (SV) for measurements. During normal operation, the RTDS sends packets to IEDs periodically. However, under cyber attack scenarios, the packet rate varies. More details on the cyber attack vector are provided in [65, 122]. Based on the co-simulation setup and cyber attack scenarios, we collect OT network traffic data for later analysis using FGraph.

#### COMPARISON WITH OPEN DATASETS

Other than the aforementioned experimental set up, in this work, we also analyze multiple open datasets, i.e., IEC 61850 [369] and DAPT 2020 [370]. In [369], the authors provide communication data from a digital substation based on IEC 61850 standard. The dataset provides OT communication traffic data under normal, disturbance, and cyber attack scenarios. Normal data is derived from normal traffic with and without variable loading. The disturbance scenarios include busbar protection, breaker failure protection, and UFLS. The cyber attack scenarios cover Denial of Service, GOOSE spoofing, merging unit measurement spoofing, circuit breaker Boolean value injection, and replay attack.



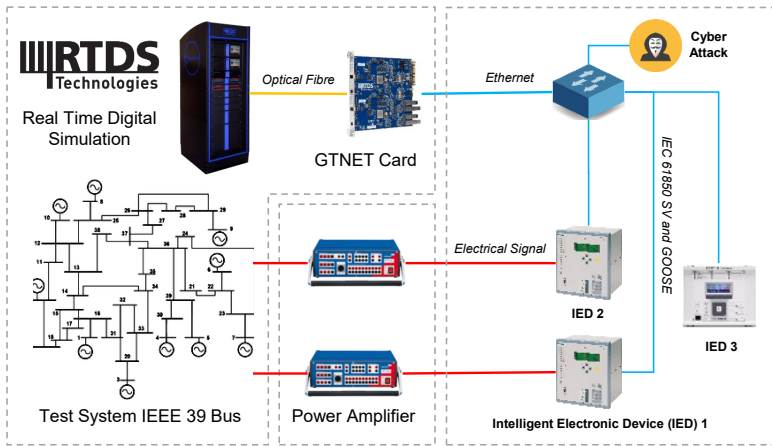


Figure 4.18: Digital substation experimental setup for OT traffic generation [6].

In [370], the authors generate data based on normal and Advance Persistent Threat (APT) traffic for a duration of 5 days. The scenarios implement various stages of cyber attack kill chain, including vulnerability scanning, exploitation, establishing a foothold, privilege escalation, etc. The experiments incorporate red team and blue team tools, e.g., Metasploit and Snort. The NetFlow data collected from the experiment within 5 days includes source, destination, flow duration, flow bytes, etc. However, the provided NetFlow CSV data is not suitable for our proposed method of TDG and TCC. Therefore, in this work, we use the provided raw original source of packet data in .pcap format.

### NETWORK TRAFFIC ANALYSIS

Table 4.3 summarizes the network traffic data from the experimental HIL (A), IEC 61850 dataset (B) [369], and APT dataset (C) [370]. Data A and B originate from the substation models within a local network, which primarily transmits layer 2 broadcast messages using MAC addresses. Meanwhile, data C is dominated by layer 3 communication using IP addresses. Data C also indicates that the network is segregated into private and public networks. Additionally, this data has the most accumulated packet history of 5 days, with a total size of 17 GB.

Table 4.3: Summary of Network Traffic Data

Parameters	A	B	C
No of Nodes	85	103	786
No of Edges	198	246	821
Traffic duration	30 minutes	150 minutes	5 days
Total packet size	50 MB	100 MB	17 GB

All the aforementioned data is then processed using the forensic graph generation model. The GC-LSTM generates traffic predictions that serve as a normalization filter.

Fig. 4.19 depicts a statistical comparison as box plots between normal, predicted, and attack traffic for all 3 cases. As shown in Fig. 4.19, normal traffic also contains outliers, indicated by red dots. These outliers can affect classification performance and result in increased false positives. Meanwhile, in the predicted traffic, the outliers are significantly reduced. Therefore, GC-LSTM helps to improve the classification accuracy of the CNN time series classifier.

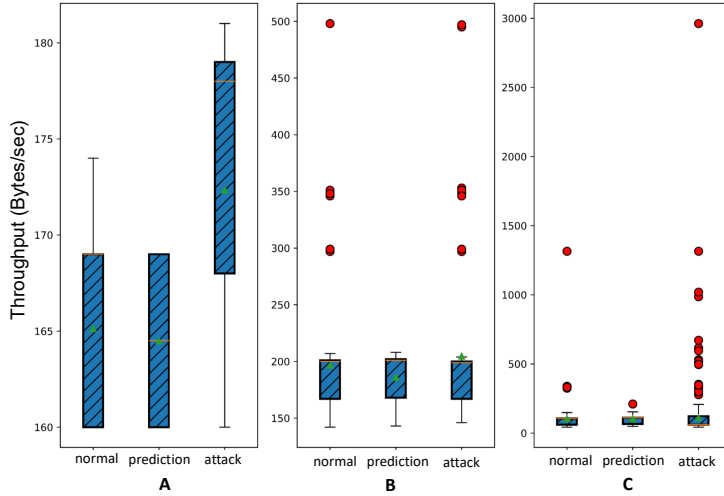


Figure 4.19: Statistical comparison between normal, predicted, and attack or anomalous traffic for data A, B, and C [6].

#### ANOMALY DETECTION AND FORENSIC GRAPH

The anomaly detection is performed based on Time Series Classification (TSC) using CNN. TSC classifies the traffic throughput as normal or anomalous. Fig. 4.20 shows the performance comparison for each dataset using the Receiver Receiver Operating Characteristic (ROC) curve. Dataset A provides the best result with an AUC score 0.819, followed by datasets B and C. Results for dataset C show the worst performance as the data contains more noise compared to the other two datasets.

Fig. 4.21 shows the forensic graph plot for normal and anomalous traffic. The blue node represents normal traffic, while the red one represents anomalous traffic. Fig. 4.19 a, b, and c show the graph representation from normal traffic, while the others show the graph under attack scenarios. The cyber attack scenarios include GOOSE replay attack, reconnaissance, data manipulation, and foothold establishment. The graph comprises nodes that store data pertaining to the source and destination IP addresses or MAC addresses, as outlined in the TDG references [352, 353]. Results from the TDG show the ability to identify anomalous nodes within the network by tracing them back to their respective IP or MAC address. The operator utilizes these particular IP or MAC addresses to identify the root causes of the traffic anomaly. These IP and MAC addresses can potentially be associated with a compromised host or a host that has been targeted by an attack.

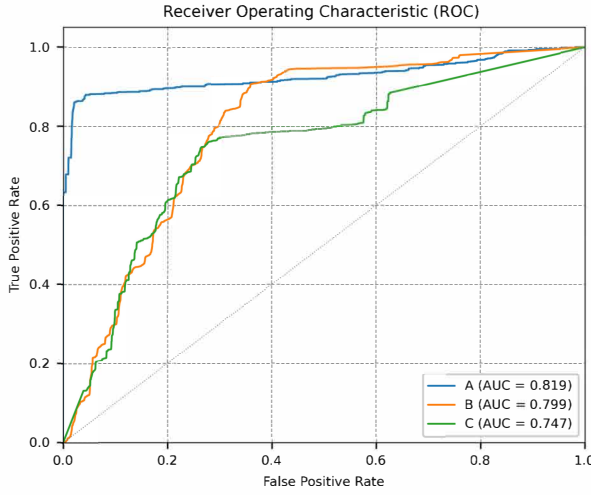


Figure 4.20: ROC comparison for data A, B, and C [6].

#### FORENSIC GRAPH RESULT AND ANALYSIS

Based on the conducted experiments, datasets A and B provide better anomaly detection performance, in comparison to dataset C. This is because the first two datasets contain homogenous OT traffic. Meanwhile, dataset C is IT traffic that has more heterogeneous characteristics. This characteristic is also shown in Fig. 4.19. Therefore, FGraph is more suitable for throughput anomaly detection in OT networks.

Compared to the attack graph, the performance of FGraph is lower because the FGraph input consists of packets captured with Wireshark. Other research has already identified problems related to Wireshark time inaccuracies [371, 372]. The Wireshark packet timestamp is inaccurate because it does not reflect the actual packet arrival or departure time. In particular, it is dependent on the time necessary for the kernel to process the arriving packets and access the clock. Regardless of this limitation, FGraph can serve as an alternative solution for graph-based forensic analysis in power grid OT communication networks. Although the performance is lower than attack graph, FGraph has more advantages due to its flexible implementation, as it does not require the deployment of SDN in the OT network. In addition, FGraph aims to avoid the degradation of the OT communication performance. Furthermore, with the recorded historical OT traffic, Security Operations Center (SOC) can perform thorough analyses of the packet payloads to avoid false positives.

## 4.4 CONCLUSION

With the ever-increasing threat of cyber attacks on power grids, it is now crucial to improve attack detection capabilities in OT systems. In this work, we proposed CyResGrid, a hybrid model of GC-LSTM and a deep convolutional network for anomaly detection in OT communication networks for power grids. It helps power system operators to localize and identify cyber attacks in near real-time. GC-LSTM creates OT traffic predictions based on the spatial and temporal features of the input data. Through its predictions,

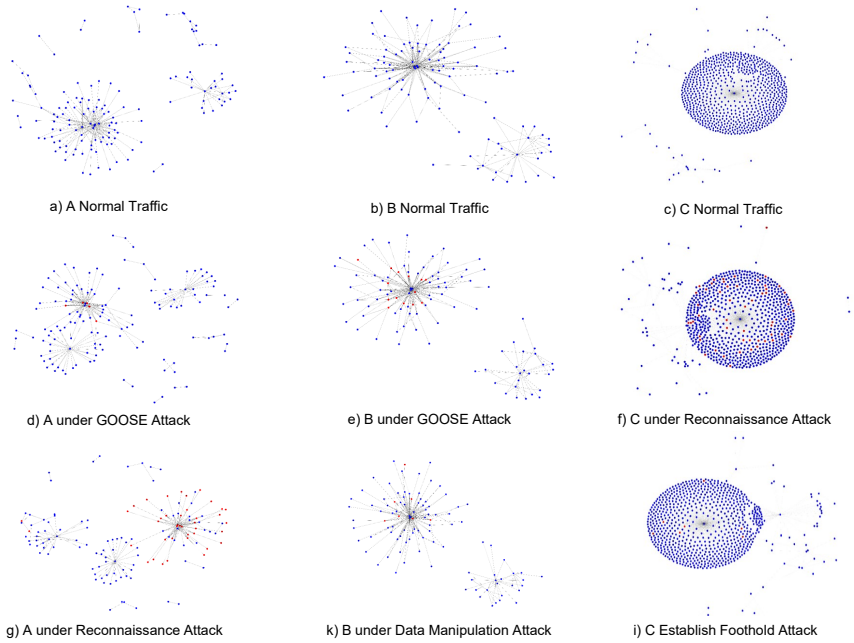


Figure 4.21: Forensic graphs for normal traffic and anomalous traffic [6].

the data variability and outliers are reduced. GC-LSTM also serves as a mechanism to improve the anomaly detection performance of TSCs. Furthermore, the deep convolutional network in CyResGrid is designed based on the hyperparameter tuning using Bayesian optimization. Hence, CyResGrid outperforms the state-of-the-art deep learning-based TSC. It provides the best detection performance, with the highest accuracy of 96.45%, F1 score of 65.03%, and G mean of 17.16%, and the lowest false positive rate of 0.13%. Additionally, for stealthy cyber attack scenarios, i.e., paranoid and sneaky attacks, CyResGrid provides the best performance indicated by the highest F1 score of 2.32% and G mean score of 3.08%. Other methods seem to provide higher accuracy and AUC. However, they have a lower performance to detect anomalies as indicated by the lower TP, F1, and G mean scores. This classification is then used to generate an attack graph that serves as an online tool for power system operators to identify and localize active cyber attacks in OT networks of power systems.

In a future work, we will focus on augmenting the proposed CyResGrid method with prevention capabilities, in addition to the existing detection features. Subsequently, it can be integrated with an intrusion detection and prevention system. The developed method is equally applicable to different OT networks and CPS topologies, besides other cyber attack vectors, such as malware-based and privilege escalation attacks. Moreover, the performance of the detection algorithm can further be improved to detect more variations of cyber attacks with infinitesimally small changes to OT network traffic intensity and frequency of occurrences.

## 5

# SPATIO-TEMPORAL ADVANCED PERSISTENT THREATS DETECTION AND CORRELATION

5

*Electrical power grids are vulnerable to cyber attacks, as seen in Ukraine in 2015, 2016, and 2022. These cyber attacks are classified as Advanced Persistent Threats (APTs) with potential disastrous consequences such as a total blackout. However, state-of-the-art intrusion detection systems are inadequate for APT detection owing to their stealthy nature and long-lasting persistence. Furthermore, they are ineffective as they focus on individual anomaly instances and overlook the correlation between attack instances. Therefore, this research proposes a novel method for spatio-temporal APT detection and correlation for cyber-physical power systems. It provides online situational awareness for power system operators to pinpoint system-wide anomaly locations in near real-time and preemptively mitigate APTs at an early stage before causing adverse impacts. We propose an Enhanced Graph Convolutional Long Short-Term Memory (EGC-LSTM) by using sequential and neural network filters to improve APT detection, correlation, and prediction. Control center and substation communication traffic is used to determine cyber anomalies using semi-supervised deep packet inspection and software-defined networking. Power grid circuit breaker status is used to determine physical anomalies. Cyber-physical anomalies are correlated in cyber-physical system integration matrix and EGC-LSTM. The EGC-LSTM outperforms existing state-of-the-art spatio-temporal deep learning models, achieving the lowest mean square error.*

## 5.1 INTRODUCTION

Cyber-Physical Power Systems (CPPS) are critical infrastructures that have been targeted by a growing number of cyber attacks in recent years. Some of the notable cyber attacks on power grids are the cyber attacks in Ukraine in 2015 [8, 9], 2016 [219], and 2022 [11]. These incidents highlight the imminent threat of cyber attacks on power grids, which had patterns resembling to APTs. The detection of APTs poses significant challenges owing to their stealthy nature and long-lasting persistence [373]. The majority of the existing research on APTs focuses on individual anomaly instances and overlooks the correlation between them [373–376]. Those studies have highlighted the necessity of anomaly correlation, but there is still a shortage of research in this area. Furthermore, according to literature studies, existing research on APTs only focuses on the cyber system [373–376] and omits the APTs on Cyber-Physical Systems (CPS).

The literature review lists four main methods for detecting power grid communication traffic anomalies, i.e., signature-based [377], sequence-based [378], rule-based [379, 380], and machine learning-based [5, 381]. Machine learning-based anomaly detection methods have gained popularity due to their superior performance [382, 383]. However, machine learning models need large amounts of data to learn and perform well. Meanwhile, cyber attack data in CPPS is scarce [382], especially for zero-day attacks. Given this constraint, a fully supervised machine learning model may not be the best option. Therefore, in this research, we employ semi-supervised DPI to identify anomalies in OT communication traffic of CPPS. The technique leverages the advantages of the homogeneous characteristics of OT network traffic generated from automated processes [384].

Semi-supervised classifiers can be constructed by combining CNN and Hamming Distance (HD). CNN usually solves supervised classification problems, i.e., intrusion detection [385]. Integration of the CNN classifier with the HD addresses data dependency in supervised learning. The HD application for distance metric learning has been proposed in [386]. CNN and HD generate Gaussian Mixture Model (GMM) vectors for semi-supervised classification with partial labeling. This GMM classification strategy improves classifier robustness with scarce labeled data [387–389]. Therefore, this classification method is suitable for zero-day attacks.

Along with a semi-supervised classifier for anomaly instance detection, system-wide monitoring is needed to correlate anomalies and track APT propagation. The state-of-the-art system-wide intrusion detection graphs are only focused on cyber anomalies and omit physical anomalies [5, 276, 390]. Meanwhile, as demonstrated in [391], combining cyber and physical anomalies would provide better cyber attack detection on CPS. Our literature review shows that cyber and physical system-wide anomaly detections in power systems are not integrated. Existing methods track anomalies using cyber graphs [5] and power system graphs [392, 393]. In [394], the authors proposed Long Range Memory (LRM) to correlate anomalies and use this knowledge to predict future attack trends. In [395], the authors proposed an AI generative model for addressing limited OT traffic and estimating the CPPS vulnerabilities and potential intrusion likelihood based on anomaly correlation. Align with our research objectives, these works highlight the necessity of spatial and temporal correlation for cyber attacks mitigation. Therefore, anomalies must be integrated and correlated to provide a system-wide visibility for spatio-temporal APT events.

Spatio-temporal correlation for APTs can determine the correlation of the anomalies

based on spatial and temporal data. Spatio-temporal correlation based on a dynamic heterogeneous graph network has been proposed to detect and correlate APTs in [396]. Graph representation and natural language processing were used to detect spatio-temporal APTs [397]. However, these APT spatio-temporal correlation methods only use IT system logs and are insufficient for the CPPS. Spatio-temporal graph modelling was proposed in [398] to correlate spatial and temporal features from sensor network measurement data. This research only focused on sensor measurement anomalies and did not consider cyber anomaly detection. According to the literature review, graph-based methods are used to develop state-of-the-art spatial correlations, and temporal machine learning models like RNN, Gated Recurrent Unit (GRU), and LSTM are used to build temporal correlations. The LSTM is the most advanced temporal model and performs best. However, LSTM has limitations when it comes to long-term memory preservation [399]. Therefore, the LSTM is not optimal for capturing the temporal correlation of APTs with non-deterministic temporal windows.

In this research, we propose a novel spatio-temporal APT detection, correlation, and prediction in cyber-physical power systems. It allows power system operators to locate system-wide anomalies in near real-time from control centers and mitigate APTs early before they cause adverse impacts. At substations and control centers, distributed semi-supervised DPI classifiers monitor OT communication traffic using SDN-enabled switch. The summary of the proposed architecture is presented in Fig. 5.1. They communicate with the SDN controller at the control center to construct a cyber anomaly graph. This is generated based on the DPI classification results using a TDG with SDN [5]. The power system graph is constructed based on the energized power lines in accordance with the status of CBs [31, 32]. The cyber-physical anomaly graph is input into a CPSIM for spatio-temporal correlation. Subsequently, an Enhanced EGC-LSTM model with sequential and neural network filters is used to predict APTs in CPPS. Furthermore, to identify zero-day APT patterns, we propose a resilient associative method based on vector databases and KNN. The method employs a CPPS log comparator function to verify and differentiate between circuit breakers opened by operators, faults, and cyber attacks. The overall processes from the proposed methods are presented in Fig. 5.2. The scientific contributions of this chapter are summarized as follows:

1. We propose a novel semi-supervised deep packet inspection method for OT communication network traffic utilizing the OT homogeneous characteristics. The method uses a combination of CNN and Hamming distance to generate vectors. The method identifies zero-day attacks by utilizing semi-supervised clustering on the baseline OT traffic vectors using a Gaussian mixture model with partial labeling. In addition, the proposed method is also integrated with software defined networking and traffic dispersion graph to facilitate power system-wide OT communication traffic monitoring in the control center and substations.
2. We propose a cyber-physical system integration matrix that constructs a topological correlation of cyber and physical system anomalies in CPPS. Control center and substation OT communication network traffic is used to construct a cyber anomalies graph. The circuit breaker status is used to construct a power system graph. The

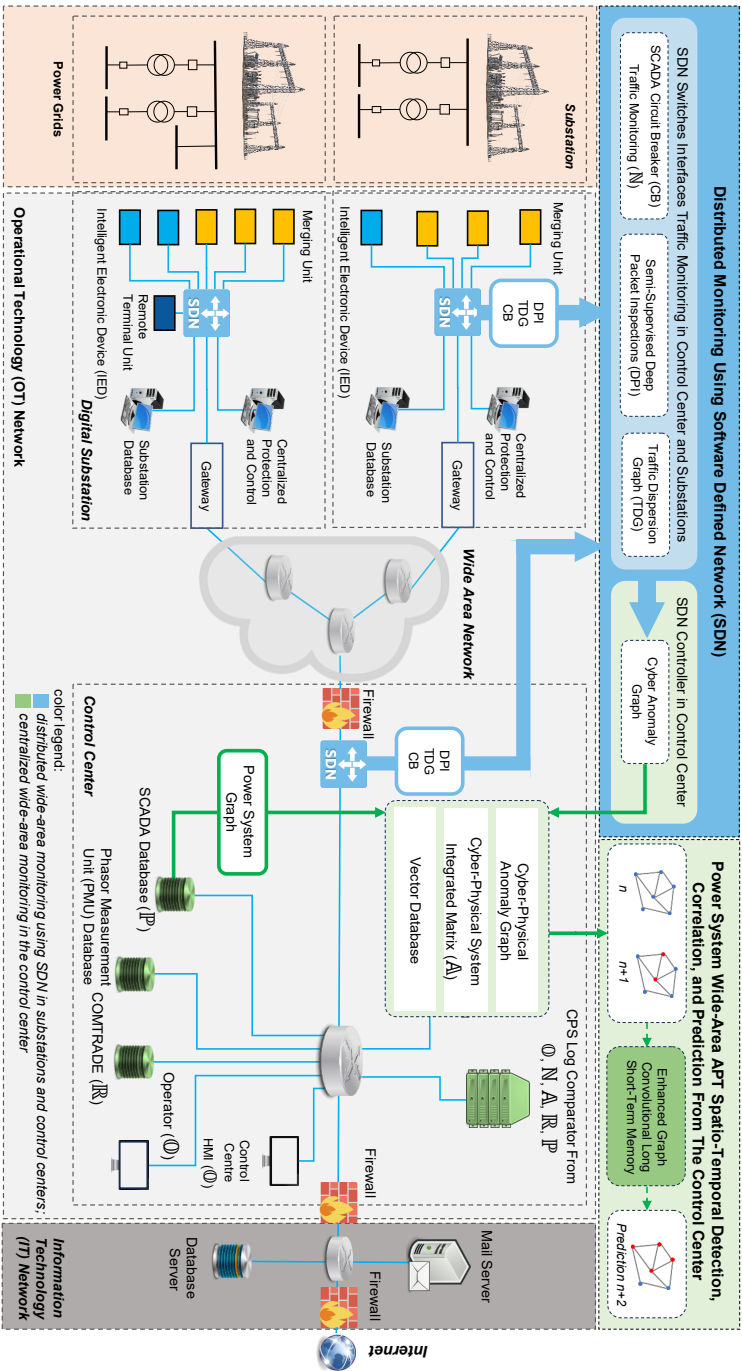


Figure 5.1: Cyber-physical system model of the power grid with IT/OT communication networks [7].



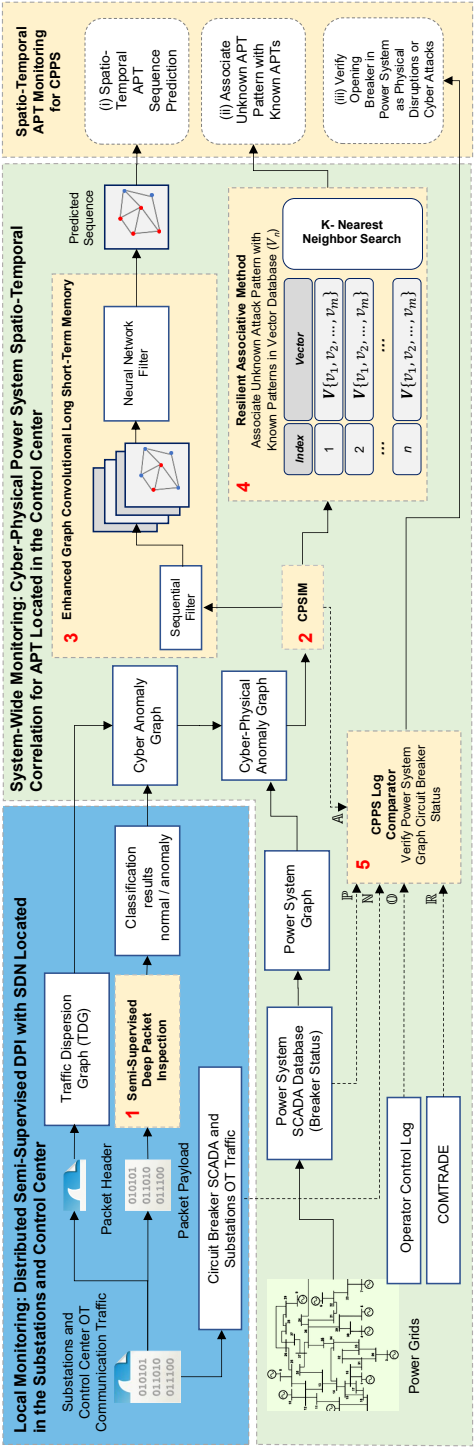


Figure 5.2: Integrated processes for spatio-temporal anomaly detection, correlation, and prediction [7].

CPSIM serves as the primary data for the APT spatio-temporal correlation and prediction processes.

3. We propose a novel EGC-LSTM model with sequential and neural network filters to predict subsequent anomalies resulting from APT attacks. The proposed EGC-LSTM uses the Sequential and Neural Network filter to minimize the Mean Square Error (MSE). Standalone implementation of the Sequential and Neural Network (NN) filter reduces the MSE by 31% and 35%, respectively. Meanwhile, the integration of both filters reduces MSE by 97%.
4. We propose a resilient associative method based on vector databases and KNN to improve the resilience of EGC-LSTM for detection of zero-day attack scenarios. The vector database of CPSIM allows the proposed model to associate zero-day attack scenarios with the known attacks using the KNN search.
5. We propose a CPPS log comparator to correlate CPPS information, i.e., operator activities, OT communication network traffic, COMmon format for TRANSient Data Exchange (COMTRADE) information from protective relays, power system CB status, and CPSIM. The log comparator enables system operators to verify and differentiate between physical power system anomalies caused by cyber attacks and physical power system disturbances.

The chapter is structured as follows. Section II explains the CPPS and cyber threat model. Section III describes the method for spatio-temporal anomaly detection, correlation, and prediction. Section IV provides the experimental results. Section V presents the conclusions and future work.

## 5.2 CPPS AND CYBER THREAT MODEL

### 5.2.1 CYBER-PHYSICAL POWER SYSTEM MODEL

CPPS models are essential for conducting research on power system cyber security. Therefore, we model the power system integrated with IT/OT communication networks as depicted in Fig. 5.1. The CPPS model incorporates SDN functionality to establish OT communication network virtualization through SDN switches in the substations and control center and SDN controller in the control center. SDN has three abstraction layers, i.e., data plane, control plane, and management plane. The data plane forwards the OT network traffic, which is controlled by the control plane. In the management plane, SDN allows the deployment of custom network applications. The model is built based on our previous research in [5]. Compared to the previous research, we improved the CPPS model with new SDN management and control functionalities, i.e., monitor the traffic of SCADA measurements and CB status, collect the summary from the semi-supervised DPI, and deploy TDG. The CPPS model is used to compute time-domain simulations and generate measurement data from substation bays, such as busbars, power lines, transformers, and generators. This data includes measurements of active and reactive power, voltage, current, and circuit breaker status. The measurements are communicated from the substations to the control center via a wide area communication network as SCADA telemetry. The SCADA data is kept in local databases situated within substations as well as the control center. The CPPS

architecture emulates the OT communication network traffic for power system monitoring and control.

The OT communication network consists of customized functionalities for each OT device within the communication network. The measurement devices include MU, RTU, and IED. These devices collect data from the power grid using SCADA with a sampling rate of one sample per second. The control center uses control commands to dynamically adjust the set points for power grid controllers in real-time. For example, control commands are used to either open or close circuit breakers for power lines, and change set points for automatic voltage regulators and governors. The measurement values and control set points are communicated across the OT network using TCP/IP packets.

### 5.2.2 CYBER THREAT MODEL FOR CYBER-PHYSICAL POWER SYSTEM

A cyber threat model is a systematic representation of potential security threats and an analysis of the techniques and pathways that attackers may employ to exploit communication network vulnerabilities. In this research, the cyber threat model is constructed based on the cyber attacks on the Ukrainian power grid in 2015 [9], 2016 [219], and 2022 [11]. These attacks resemble APT's strategies from the early phase of the intrusions until the power outages in the later stages. In the Ukrainian power grid attack in 2015, the adversary used spear phishing emails as an attack vector against the distribution system operators. The phishing emails contained a Microsoft Excel file attachment that was infected with the BlackEnergy3 malware. Subsequently, adversaries performed stealthy operations in the IT and OT communication networks while preparing for the final phase of the attack. During the early attack phase, the adversaries conducted several malicious activities to intrude from the IT communication network into the control center and substations, i.e., reconnaissance, exploit, lateral movement, firmware modification, and command and control. These activities inevitably caused anomalies in the IT/OT communication network traffic. However, the absence of an early detection mechanism rendered these activities imperceptible to the distribution system operators.

In addition to the aforementioned threat posed by external adversaries, there is also the possibility of an attack from internal actors, known as an insider threat. Insider threats have different characteristics from external threats. The external threat required a lateral movement to reach its final objective in a timely fashion. These scenarios provide an opportunity for the early identification of external threats. However, insider threat potentially has direct access to the substations and control centers and has the potential to cause an immediate severe impact. There is also a possibility when the external and insider threats are combined into more sophisticated and coordinated attack scenarios. However, modelling the insider threat behavior and integration with the test simulation has been identified as a notable challenge for anomaly detection [400]. Therefore, in this research, we omit the insider threat constraint in our CPPS threat model.

Using the aforementioned CPPS co-simulation, our research simulates the early phase of a cyber attack in the simulated substations and control center, which includes reconnaissance, command injection, and malware traffic. The normal and anomalous communication traffic is then used to train the semi-supervised deep packet inspection. Using a traffic dispersion graph, anomaly detection also tracks the sources and destinations of anomalous packets. The graph representation of anomalies is able to track lateral movement processes

within the OT network from the entry point to the end device that has direct control of the power grid components. This information is also combined with the CB status of the power lines and transformers to track anomalies in power system-wide. Subsequently, the cyber and physical anomalies are used for the spatio-temporal anomaly correlation and prediction.

### 5.3 SPATIO-TEMPORAL ANOMALY DETECTION, CORRELATION AND PREDICTION

The cyber-physical power system architecture integrating the power grid, IT/OT communication networks, and SDN is depicted in Fig. 5.1. The WAN monitoring is enabled based on data collected in near real-time at substations and control center, i.e., OT communication network traffic and CB status. Fig. 5.2 shows the integrated processes of the proposed method for spatio-temporal anomaly detection, correlation, and prediction. Their implementation in CPS is represented in Fig. 5.1. The OT communication traffic is monitored locally in all substations and control center on the SDN-enabled switches. The OT traffic is classified using semi-supervised DPI to determine whether an individual packet is normal or anomalous. This information is combined with TDG to generate and update in near real-time as a system-wide cyber anomaly graph. A power system graph is updated in near real-time based on the CB status. It is combined with the cyber anomaly graph into CPSIM. The EGC-LSTM runs continuously to predict subsequent anomalies according to the input from the last four anomalies in CPSIM. To identify zero-day attacks, the resilient associative method associates the zero-day CPSIM with the known CPSIM scenarios using a KNN-based search on the vector database. The CPPS log comparator runs in near real-time to verify the CB status and distinguish between a CB opened by cyber attacks and physical power system disturbances.

The proposed method provides three main results, i.e., (i) spatio-temporal APT detection, correlation and sequence prediction, (ii) identification of zero-day attacks, and (iii) identification of circuit breakers opened by cyber attacks. A detailed description of the proposed method and corresponding processes are provided in the following subsections.

#### 5.3.1 SEMI-SUPERVISED LEARNING FOR DEEP PACKET INSPECTION

The DPI uses supervised CNN, HD, and semi-supervised learning based on GMM with partial labeling. The CNN model performs supervised classification for packet payload from OT communication network into normal and anomalous. The packet payload is converted into a 2-Dimensional (2D) data representation as an image. Eq. (5.1) shows the convolution function from the 2D CNN layer. The  $*$  denotes convolution operation,  $f$  is the filter size  $m \times n$  and,  $g$  is the input data size  $i, j$ . Bayesian optimization [37] is used to optimize the CNN model. Between the convolutional layers, CNN uses Rectified Linear Unit (ReLU) activation function and pooling layer. An activation function is applied to introduce non-linearity into the model, and pooling layers are used to reduce the size. In the end part of CNN layers, there are flattening processes and fully connected neural networks. The flattening layer converts the 2D data into one-dimensional data. The fully connected neural network allows weight adjustment during the training to produce the best-fit output. After the fully connected neural networks, a SoftMax function in (5.2) is

used to classify the output according to a probability distribution.

Typical neural networks use SoftMax to determine the class categories. However, in our model, we use the SoftMax probability score to generate a vector  $\Phi$ . The vector  $\Phi$  in (5.4) is generated from the combination of SoftMax ( $\sigma$ ) in (5.2) and Hamming Distance (IH) in (5.3). The IH computes the distance between the average normal traffic payload in OT and other traffic payloads. For normal traffic, the hamming distance will be close to zero ( $\text{IH} \approx 0$ ). Meanwhile, anomalous traffic tends to result in a larger IH score. The combination of outputs from CNN and IH is then used to generate 2D vectors  $\Phi$  for semi-supervised learning.

GMM with partial labeling is used as a semi-supervised learning technique to classify the 2D vectors  $\Phi$ . It uses both labeled and unlabeled data to fit the model. The presence of labeled data facilitates the learning process by enhancing the accuracy of the estimation of GMM parameters. Eq. (5.5) shows the equation for GMM probability where  $\Phi$  is the 2D vector data points,  $K$  is the number of Gaussian distributions in the mixture,  $\pi_k$  represents the mixing coefficient, and  $N$  is the probability density function. The presence of a limited number of labels in the data denotes the presence of a limited dataset that can be used for anomaly detection. Unlabeled data represents the zero-day attack traffic. By employing this semi-supervised learning strategy, our model can effectively identify zero-day attack traffic.

5

$$(f * g)(i, j) = \sum_m \sum_n f(m, n) \cdot g(i - m, j - n) \quad (5.1)$$

$$\sigma = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (5.2)$$

$$\text{IH} = \frac{1}{N} \sum_{i=1}^N |a_i - b_i| \quad (5.3)$$

$$\Phi = \langle \text{IH}, \sigma \rangle \quad (5.4)$$

$$\text{GMM}_p(\Phi) = \sum_{k=1}^K \pi_k N(\Phi \mid \mu_k, \Sigma_k) \quad (5.5)$$

### 5.3.2 CYBER-PHYSICAL SYSTEM INTEGRATION MATRIX

In order to integrate the cyber and physical components of the CPPS as an integrated graph, we construct the CPSIM. CPSIM is formed by combining the adjacency of the OT network topology ( $c$ ) with the power system topology ( $p$ ). Fig. 5.3 shows the representation of CPSIM integration. The cyber adjacency matrix shown in the blue area is represented by  $c_{ij}$ , where  $i$  and  $j$  indicate the element of the matrix. Meanwhile, the power system adjacency matrix shown in the red area is represented by  $p_{uv}$ , where  $u$  and  $v$  indicate the element of the matrix. The CPSIM  $A_{c+p}$  is a combination of cyber and physical elements with dimensions  $A_{i+u, j+v}$ . Other than the cyber and physical elements, we introduce a connection matrix in the yellow area represented by  $x$ . This area represents the functional connectivity between cyber and physical systems. For example, a node in the cyber element is able to change the physical state of a node in the physical element. The connection matrix is constructed based on the prior information of control function configuration

from cyber into the physical system. All information from cyber and physical topology and connectivity information are integrated into a single adjacency of CPSIM.

The adjacency matrix from CPSIM ( $A$ ) serves as a main reference for the entire cyber and physical system state in CPSIM ( $\hat{A}$ ). In the CPSIM, the anomalous elements are indicated by  $A_{i+u,j+v} = 1$ , and  $A_{i+u,j+v} = 0$  otherwise. This reference is then used to track anomalies in both cyber and power systems. Our model identifies the power system graph by analyzing the energized lines based on breaker status information. When the circuit breaker is closed, it indicates a normal condition (0) in the CPSIM. Alternatively, when the breaker is in the open position, it indicates a potential anomaly state (1) in the CPSIM. Meanwhile, to identify the anomaly on the OT network, we use the semi-supervised DPI and TDG. The TDG is utilized to determine the location of OT communication traffic anomalies. The TDG utilizes graph structures to depict nodal information. Every node in a graph represents a distinct host in the communication network. The transfer of information between hosts is shown by the interconnectedness of nodes, specifically, the edges of a graph [353]. By combining DPI and TDG, we identify the location of anomalies in the OT communication networks. In the CPSIM record, the cyber anomalous element is recorded as  $A_{i,j} = 1$  in the CPSIM, and  $A_{i,j} = 0$  otherwise.

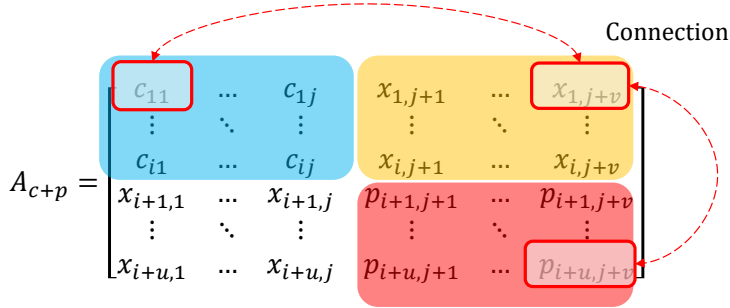


Figure 5.3: Cyber-physical system integration matrix [7].

### 5.3.3 APT SPATIO-TEMPORAL CORRELATION

Methods exist in the literature for spatio-temporal correlation, i.e., Graph Convolutional Gated Recurrent Unit (GConvGRU) [401], Temporal Graph Convolutional Network (TGCN) [402], and GC-LSTM [41]. These methods can capture the spatial correlation of data using graph convolution. However, they are not the best-fit solution to achieve optimal performance in terms of temporal correlation for APTs. The APTs exemplify non-deterministic temporal characteristics and typically endure extended time intervals between attack stages. Meanwhile, the time-series models, i.e., GRU and LSTM, have the limitation to address long-term temporal correlation [399]. Therefore, in this research we propose an EGC-LSTM to address the issue arising from the spatio-temporal correlation of APTs. There are three main improvements in the EGC-LSTM, i.e., Bayesian optimization, sequential filter, and NN filter. Bayesian optimization aims to optimize the GC-LSTM model architecture [359]. The sequential filter is implemented to reduce the recorded data from CPSIM. The filter

selectively saves CPSIM data that have distinct values compared to the most recent data in CPSIM, instead of saving all data indiscriminately. This mechanism enables the GC-LSTM to prioritize the detection of anomaly changes instead of analyzing the entire data stream. This mechanism improves the temporal correlation performance. Algorithm 1 shows the pseudocode of the sequence filter algorithm. This algorithm aims to address the APT non-deterministic temporal windows of anomaly records in the CPSIM.

---

**Algorithm 3** Sequence Filter
 

---

**Input:**  $A = [ ]$  : Initialize the log storage for CPSIM  
 $n = 0$  : Nodes traffic data  
 $CPSIM_t$  : CPSIM matrix stream for every time  $t$   
**Output:**  $A = [CPSIM_1, \dots, CPSIM_n]$  : CPSIM matrix log

---

Iteration for every CPSIM stream

**if**  $CPSIM_t \neq A[n]$  **then**

└

$A[n+1] == CPSIM_t$

$n = n + 1$  **else**

└

*continue*

**return**  $A = [CPSIM_1, \dots, CPSIM_n]$

---

The output from the sequence filter serves as input for GC-LSTM that combines Graph Convolutional Network (GCN) and LSTM. The GCN function is utilized to extract the nodal characteristics from CPPS elements in CPSIM A as described in Eq. (5.6). GCN operates based on the Hadamard product multiplication ( $\odot$ ) of the weight matrix ( $W_{gcn}$ ), adjacency matrix ( $A$ ), and node features from CPSIM A. The adjacency matrix ( $A$ ) is augmented with the identity matrix ( $I$ ) to create a modified adjacency matrix ( $\tilde{A}$ ). The equation incorporates the number of hops from a communication node to neighboring nodes, denoted as  $k$ . The temporal features are processed using LSTM subsequent to the acquisition of the spatial features through the GCN. The LSTM input originated from the last four CPSIM anomalies to predict subsequent anomalies in near-real-time. The operations performed within an LSTM cell are described in Eqs. (5.7 - 5.12). There are six main sub-equations in the LSTM process, including the forget gate ( $f_t$ ), input gate ( $i_t$ ), output gate ( $o_t$ ), internal cell state ( $c'_t$ ), transferable cell state ( $c_t$ ), and hidden state ( $h_t$ ). The predicted output from EGC-LSTM ( $o_t$ ) serves as an input for the NN filter in Eq. (5.13). This EGC-LSTM and NN filter are optimized using a Bayesian optimization for hyperparameters tuning. The NN filter transforms the EGC-LSTM output into a binary value of 0 or 1 to enhance the prediction performance of the EGC-LSTM.

$$GCN_t^k \leftarrow (W_{gcn} \odot \tilde{A}^k) \mathbf{A} \quad (5.6)$$

$$f_t = \sigma((W_f GCN_t^k) + (U_f h_{t-1}) + b_f) \quad (5.7)$$

$$i_t = \sigma((W_i GCN_t^k) + (U_i h_{t-1}) + b_i) \quad (5.8)$$

$$o_t = \sigma((W_o GCN_t^k) + (U_o h_{t-1}) + b_o) \quad (5.9)$$

$$c'_t = \tanh((W_c GCN_t^k) + (U_c h_{t-1}) + b_c) \quad (5.10)$$

$$c_t = (f_t \odot c_{t-1}) + (i_t \odot c'_t) \quad (5.11)$$

$$h_t = o_t \odot \tanh(c_t) \quad (5.12)$$

$$EGC-LSTM = f(W_o + b) \quad (5.13)$$

### 5.3.4 RESILIENT ASSOCIATIVE METHOD

To enhance the resilience of EGC-LSTM in handling new or unknown APT patterns, we propose a resilient associative method by performing a KNN search on a vector database. Vector databases are specifically designed to store and handle vector data, which consists of data points defined by arrays or lists of values [403]. Fig. 5.4 represents the vector database search strategy with KNN. Vector database  $\mathbf{A}_n$  represents known historical anomalies recorded in the CPSIM matrix. There are numerous potential combinations of anomalous events ( $\mathbf{U}$ ) that may not be included in the existing data  $\mathbf{A}_n$ , i.e., zero-day attacks. Therefore, to address event detection for zero-day attacks, the KNN algorithm is implemented to search for the most similar pattern from known data. By implementing this strategy, the model can identify zero-day anomalies ( $\mathbf{U}$ ) by associating this anomaly with the known one ( $\mathbf{A}_n$ ).

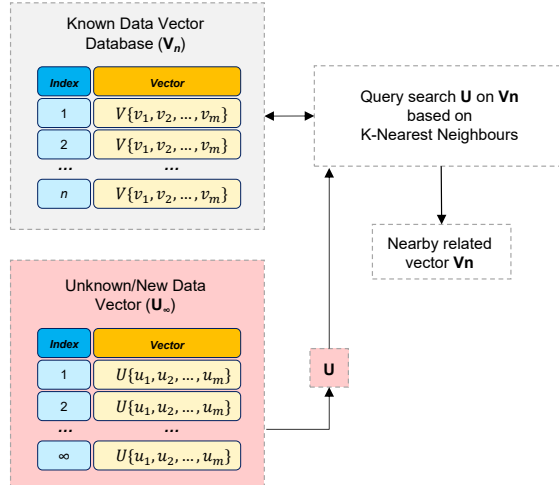


Figure 5.4: Vector database query search strategy with KNN [7].



### 5.3.5 CPPS LOG COMPARATOR

We introduce an innovative circuit breaker log comparator as a multi-log anomaly detection system for CPPS that specifically targets CB-related events. The breaker log comparator function is utilized to compare recorded log activities from CPSIM anomaly record (A), CB of SCADA and substations traffic (N), operator control log (O), power system SCADA database CB status (P), and relay COMTRADE (R). Fig. 5.5 shows the diagram of the breaker log comparator. Data A originates from the spatio-temporal data of anomalies in CPSIM. Log N is generated by the observed CB control traffic in the OT network. Log O is produced through authorized control operations recorded by the power system operator. Log P denotes the current state of the CB in the physical power grid. Log R represents the relay COMTRADE that records transient events in the power system. COMTRADE data has been used to identify electrical disturbances in power systems based on transient waveforms [404]. The breaker log comparator utilizes the COMTRADE to differentiate anomalies in the power system caused by physical system events and cyber attacks.

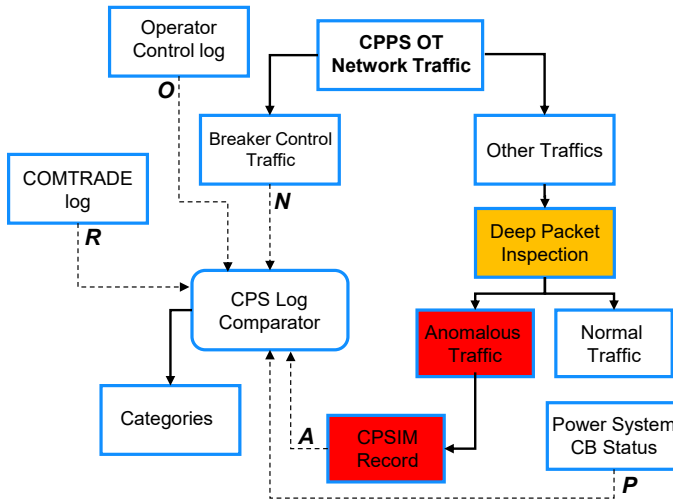


Figure 5.5: Circuit breaker log comparator function [7].

Table 5.1 presents the log comparison categories corresponding to the five types of data logs. The value of 1 indicates the presence of an activity log or anomaly, while the value of 0 represents the normal operating condition. During normal operation, all log parameters are indicated as 0. Otherwise, elements of CPSIM (A) records are indicated with 1. The anomalies in CPSIM are subsequently compared with other logs to identify different scenarios, i.e., system operator performing control actions, physical disturbances, activation of protection relays, and cyber attacks. For physical disturbances, the COMTRADE data serves as a primary indicator. Meanwhile, CPSIM (A) serves as a primary indicator for cyber attacks.

There are four cyber attack scenarios. The first scenario is when the adversaries compromise the legitimate operator's control workstation. Therefore, the operator control log (O) will be indicated by 1. The second scenario is characterized by the adversaries

executing a spoofed remote control and disguising themselves as a legitimate operator. It does not originate from the legitimate operator's control workstation and is indicated by control log O. The spoofed controls originate from a compromised device in the OT communication network that sends malicious breaker control commands. The third attack scenario occurs when the adversaries compromise a device in the Bay Control Unit (BCU). Compared to the previous scenarios, this attack does not provide an indicator of a breaker control command in the network. This is possible due to the position of the BCU devices that have a direct connection with the power grids. The fourth attack scenario is a cyber anomaly, which refers to a situation where the cyber anomaly is recorded in CPSIM (A) and does not have any impact on the power system. This category corresponds to the reconnaissance phase in the cyber kill chain.

Table 5.1: CPS Breaker Log Comparator Categories

R	O	A	N	P	Causes	Categories
0	0	0	0	0	-	Normal operation
1	0	0	0	1	Physical disturbances	Direct response in BCU
1	0	0	1	1		Coordinated protection
0	1	0	1	1	Operator	Operator control action
0	1	1	1	1	Cyber attacks	Compromised operator control
0	0	1	1	1		Spoofing attack
0	0	1	0	1		Compromised BCU device
0	0	1	0	0		Preliminary kill chain stages

## 5.4 EXPERIMENTAL RESULTS

In this section, the experimental results of the proposed methods are presented. This section provides an overview of the experimental setup, including the cyber-physical power system co-simulation setup and dataset. Subsequently, two main results are presented in this section. First, the semi-supervised deep packet inspection is presented to quantitatively assess the capability of zero-day detection. The experimental results demonstrate the effectiveness of the proposed method in identifying unknown attacks despite having a limited amount of training data. This solution intends to address the problem of the limited availability of the dataset acquired from the OT communication traffic of power grids. Second, the spatio-temporal anomaly correlation and prediction demonstrate the prediction performance for subsequent anomalies resulting from APT attacks. The experimental results present the superiority of the proposed EGC-LSTM in comparison with the state-of-the-art graph spatio-temporal deep learning models. A more detailed explanation is provided in the following subsections.

### 5.4.1 EXPERIMENTAL SETUP

The experiments in this work are performed using the CPPS model of the power grid represented in Fig. 5.1. The power system is simulated in real-time using a Root Mean Square (RMS) dynamic model of the IEEE 39-bus test system in DlgSILENT PowerFactory. The CPS model employs OPC UA to establish a connection between the time domain simulation of the power grid and simulated IT/OT communication networks. The OT network emulation utilizes Mininet<sup>1</sup> deployed on a total of 10 virtual servers. It consists of 27 user-defined substations, 118 measurement devices, and over 800 data points to emulate the OT communication network of IEEE 39-bus test power system. The CPPS model comprises a total of 185 nodes, consisting of 146 OT nodes and 39 physical nodes from the IEEE 39-bus system. Fig. 5.6 depicts the adjacency matrix representing the connection between 185 nodes of CPPS that are associated with CPSIM. The nodes that are connected are represented by the value 1 in the adjacency matrix, whereas the nodes that are not connected are represented by the value 0. The blue area corresponds to the OT adjacency matrix, while the red area corresponds to the IEEE 39-bus adjacency matrix. The cyber-physical control region illustrates the functional connectivity between the OT and power system, coupling the cyber and physical systems together.

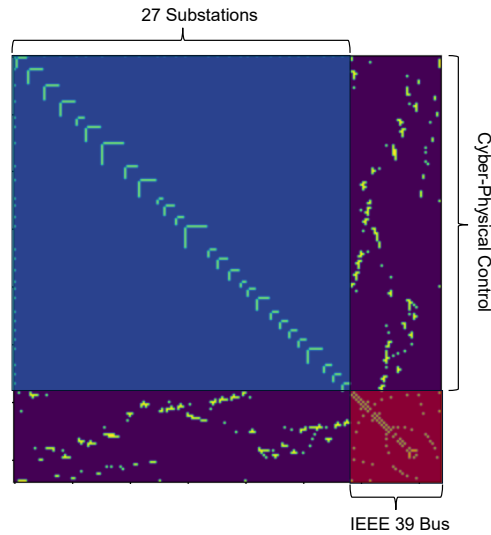


Figure 5.6: CPPS Adjacency Matrix [7].

The SCADA functionalities in the OT communication network are achieved by implementing customized Python code on each Mininet host. The OT devices include MUs, RTUs, IEDs, database server, gateway, human machine interfaces, and control center. The measurement values and control set points are communicated across the OT network using TCP/IP packets. In this research, we focus on the control traffic associated with CBs control. As shown in Fig. 5.1, the CPPS model is integrated with the SDN application to monitor

<sup>1</sup> <https://mininet.org/>

OT traffic payload using SDN-enabled switch interfaces. With this capability, the CPPS model performs traffic monitoring in substations and control center.

From the CPPS co-simulation, the experiment collects two types of data. The first type of data is network traffic from the OT communication network collected as .pcap files. The traffic from multiple locations contains the source and destination addresses. This is processed using a traffic dispersion graph. In addition, the packet payload is classified as normal or anomalous using semi-supervised deep packet inspection. The second type of data is the CB status collected from the DigSILENT PowerFactory simulation, which represents the status of the power system. Subsequently, both the cyber and physical system data are combined into the CPPS dataset in the CPSIM. The historical wide-area CPPS data from the CPSIM is used as input parameters for the EGC-LSTM. The EGC-LSTM model is implemented using the PyTorch<sup>2</sup> and Pytorch Geometric<sup>3</sup>. Based on this information, the EGC-LSTM performs spatio-temporal anomaly correlation and prediction. In this experiment, the EGC-LSTM model combines graph convolution and LSTM with hidden state vector parameters with the size of 32. Subsequently, the EGC-LSTM uses the ReLU as an activation function.

## 5

### 5.4.2 SEMI-SUPERVISED DEEP PACKET INSPECTION FOR OT ANOMALY DETECTION

Our research uses OT traffic generated from the CPPS simulation to evaluate the performance of DPI. The simulation in Mininet produced TCP/IP traffic in the OT network. In the CPPS model, we test several cyber attacks scenarios, i.e., DoS, network scanning, exploits, and malwares. In addition to the OT traffic generated by our simulation, we verify the model's performances by using open OT traffic datasets, i.e., IEC 61850 [6], Routable IEC 61850 [405], IEC 104 [406], DNP 3 [407], and Modbus [408]. Furthermore, we also incorporate samples of cyber attack datasets [409] and Industroyer malware traffic samples [410]. A total of 7.71 GB of .pcap data is collected from the OT traffic samples for the evaluation of semi-supervised DPI. Fig. 5.7 depicts the statistical distribution of packet size using a box plot across several OT traffic categories. Overall, the average size of the normal OT traffic from various protocols is 118.599 bytes, and 304.735 bytes for cyber attacks. In order to handle the size of the OT traffic, we use a 16x16 convolutional input with a total capacity of 256 bytes. When the traffic exceeds 256 bytes, the extra bytes are discarded. Conversely, when the traffic is less than 256 bytes, the remaining spaces are filled with zeros. The top part of Fig. 5.8 shows the image representation from each tested OT traffic. This 2D data is used as input for the supervised CNN algorithms.

The outputs from CNN and HD generate vectors for GMM with partial labelling. The GMM is implemented using the Scikit-learn<sup>4</sup> library. The bottom section of Fig. 5.8 depicts the result from GMM with partial labelling for all protocols. K represents the number of classes in the GMM classifier. This parameter value is decided based on the number of classes in the tested dataset. The GMM uses the probability density function of multivariate Gaussian distribution with a full covariance matrix. The GMM implementation does not utilize an explicit distance metric, e.g., Euclidean Distance (ED). GMM uses Mahalanobis

<sup>2</sup> <https://pytorch.org/>

<sup>3</sup> <https://pytorch-geometric.readthedocs.io/>

<sup>4</sup> <https://scikit-learn.org/stable/>

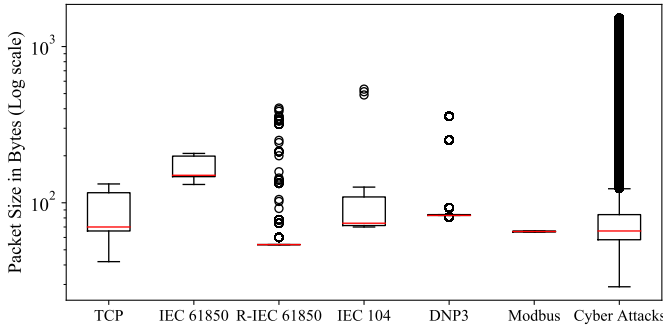


Figure 5.7: Statistical box plot from normal traffic and cyber attacks [7].

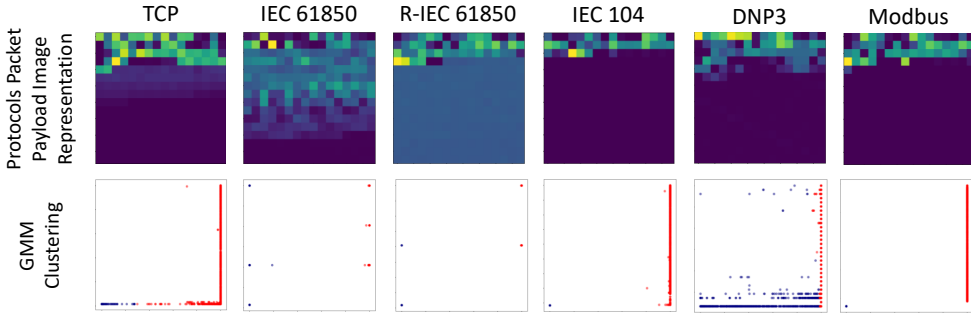


Figure 5.8: OT traffic images representation and the result of Gaussian Mixture with partial labelling for each protocols [7].

distance to calculate a point's likelihood for a given Gaussian component. The Mahalanobis distance quantifies the spatial separation between an individual data point and a given probability distribution.

In our scenarios, we create a pair of two classes from the normal/baseline OT protocols with cyber attack traffic. In this experiment, we evaluate the performance of the GMM with partial labelling with different proportion of training and test data. The labeled data proportion selection is carried out by running GMM with a variation of the labeled data proportion between 1% and 30%. As shown in Fig. 5.9, the majority of the tested dataset only required less than 5% labelled data to achieve the minimum MSE. However, for the DNP3, it required 19% labelled data to achieve the best performance. This is because of the DNP3 characteristics, which has more data variation, as shown in Fig. 5.8. In addition, compared to other datasets, the DNP3 dataset has a substantial amount of DNP3-modified attack packets, which resemble the normal DNP3 traffic. Therefore, the clustering plot of DNP3 is different from that of the other protocols. Consequently, to achieve the best performance for all datasets, the experiment incorporates a 20% proportion of labeled data.

The x-axis represents the probability scores from the CNN, and the y-axis represents the HD scores. The red dots indicate the cyber attack traffic that is associated with a higher CNN anomaly probability and HD score close to one. Conversely, the blue dots indicate

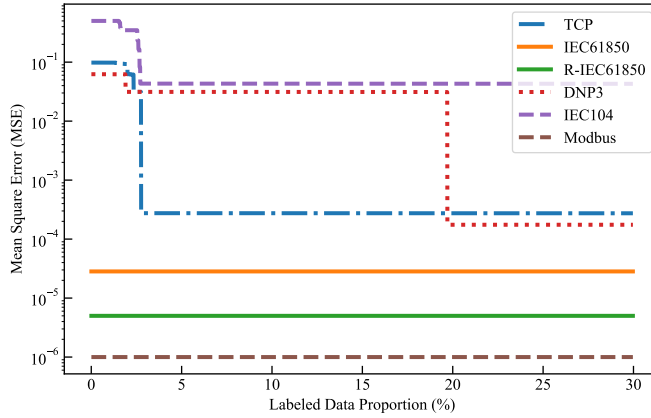


Figure 5.9: Proportion of labeled data impact on the MSE for all protocols [7].

## 5

normal OT traffic that has a lower probability of CNN anomalies and a low HD score nearing zero. Fig. 5.10 depicts the ROC curve from all traffic categories and AUC score. This plot indicates that the semi-supervised DPI and GMM with partial labelling provides a good classification performance.

### 5.4.3 SPATIO-TEMPORAL ANOMALY CORRELATION AND PREDICTION

The anomaly detection result generated from the semi-supervised DPI is further processed using TDG and recorded in CPSIM, together with the CB's status retrieved from the power grid. During instances of cyber attacks, the CPSIM matrix will deviate from its normal state (all zeroes). As the attack progresses, particular elements of CPSIM are shifting to a value of 1. In the matrix, the value 1 corresponds to traffic anomaly or an open CB in the power grid. Based on this constraint, the transition on the CPSIM matrix will be varied depending on the cyber attack scenarios and location.

This research performs 220 cyber attacks scenarios with variation of location and methods. These scenarios serve as primary data to evaluate the performance of EGC-LSTM. In addition, we also perform benchmarking with the state-of-the-art graph-based spatio-temporal deep learning models, i.e., GConvLSTM and GConGRU [401], TGCN [402], and GC-LSTM [354]. Table 5.2 shows the performance comparison of the tested models based on MSE. A smaller MSE indicates superior prediction results. Our proposed strategy with sequence and NN filter reduces the MSE for all models. Table 5.3 shows the average performance comparison of the five original graph-based spatio-temporal deep learning models and their variants with Algorithm 1 Sequential and NN filter. Table 5.3 quantifies the impact of the implementation of Seq. filter, NN filter, and combination of Seq. filter and NN filter. The standalone implementation of the Sequential and NN filter reduces the MSE by 31% and 35%, respectively. Meanwhile, the integration of both filters reduces MSE by 97%. The best MSE of 0.0003 is achieved in the proposed EGC-LSTM that implements Bayesian optimization, sequence filter, and NN filter. Besides reducing the MSE, as shown in Table 5.3, the filters also increase the computing time by 1.5 3.7 %.

Fig. 5.11 depicts the sample prediction result from EGC-LSTM for cases 76 and 218.

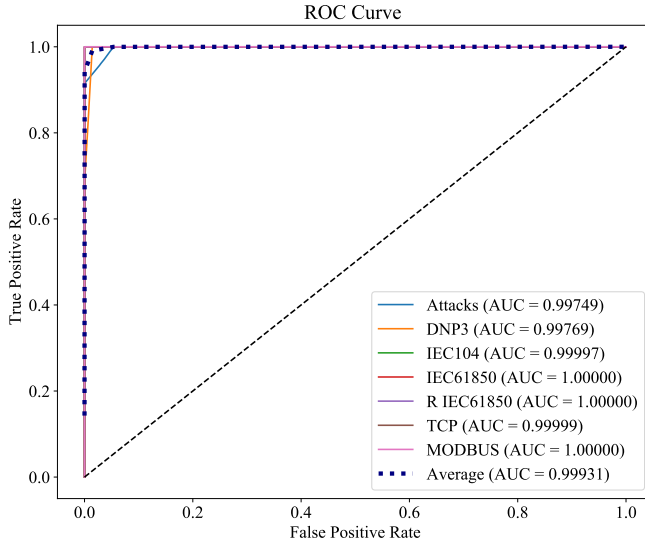


Figure 5.10: ROC curve from all traffic categories [7].

Table 5.2: MSE Scores Comparison of Graph-Based Spatio-Temporal Deep Learning Models

Combinations	Performance Parameters	Tested Models				
		GConvLSTM	GConvGRU	TGCN	GC-LSTM	EGC-LSTM
Original	MSE $\pm$ SDev	0.055 $\pm$ 0.017	0.054 $\pm$ 0.018	0.052 $\pm$ 0.021	0.045 $\pm$ 0.017	<b>0.035<math>\pm</math>0.014</b>
	Time $\pm$ SDev	528 $\pm$ 112	371 $\pm$ 91	253 $\pm$ 67	304 $\pm$ 84	<b>307<math>\pm</math>79</b>
+ Seq. Filter	MSE $\pm$ SDev	0.034 $\pm$ 0.012	0.039 $\pm$ 0.013	0.037 $\pm$ 0.018	0.034 $\pm$ 0.016	<b>0.021<math>\pm</math>0.011</b>
	Time $\pm$ SDev	532 $\pm$ 118	381 $\pm$ 97	259 $\pm$ 64	308 $\pm$ 81	<b>309<math>\pm</math>80</b>
+ NN Filter	MSE $\pm$ SDev	0.027 $\pm$ 0.011	0.037 $\pm$ 0.014	0.026 $\pm$ 0.011	0.043 $\pm$ 0.019	<b>0.023<math>\pm</math>0.012</b>
	Time $\pm$ SDev	542 $\pm$ 145	378 $\pm$ 94	258 $\pm$ 66	311 $\pm$ 81	<b>323<math>\pm</math>72</b>
+ Seq. Filter NN Filter	MSE $\pm$ SDev	0.0026 $\pm$ 0.0012	0.0011 $\pm$ 0.0007	0.0016 $\pm$ 0.008	0.0019 $\pm$ 0.0009	<b>0.0003<math>\pm</math>0.0002</b>
	Time $\pm$ SDev	549 $\pm$ 127	385 $\pm$ 99	257 $\pm$ 69	313 $\pm$ 83	<b>324<math>\pm</math>84</b>

In case 76, the cyber attack started from substation 9 and compromised merging unit 9.1. This MU has the capability to control the CB of the power line between Bus 6 and Bus 7. During the state  $n$ , EGC-LSTM can predict incoming events in  $n+1$  before they actually happen. In state  $n+1$ , the method can predict the circuit breaker that will be affected after the power line between Bus 6 and 7 is disconnected. For case 218, the cyber attack is starting from substation 27. Compared to case 76, this scenario shows different highlighted anomalous locations. In the majority of cases, the breaker opening attack will not trigger other breakers to open. However, in cases 76 and 218, the opening a few breakers will trigger more breakers to open due to protection schemes implemented in the IEEE-39 bus model.

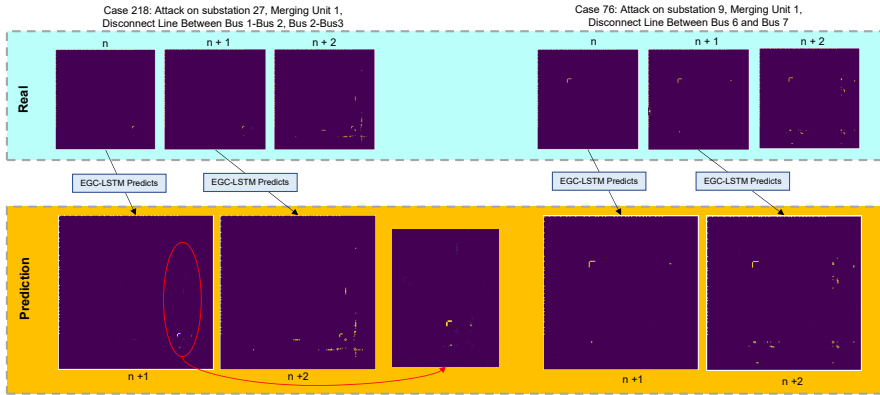


Figure 5.11: Sample EGC-LSMT Prediction Results from case 76 and 218 [7].

Table 5.3: Performance Comparison of Graph-Based Spatio-Temporal Deep Learning Models with Sequential and Neural Network Filters

Parameter	Original	Seq. Filter	NN Filter	Seq. and NN Filter
Average MSE	0.0482 (+ 0%)	0.033 (- 31%)	0.0312 (- 35%)	0.0015 (- 97%)
Average Time (ms)	352.6 (+ 0%)	357.8 (+ 1.5%)	362.4 (+ 2.8%)	365.6 (+ 3.7%)

#### 5.4.4 DETECTION FOR ZERO DAY ATTACK SCENARIOS

Considering the complexity of CPPS topology, there are various possibilities of cyber attacks scenarios. To address this concern, we implement the resilient associative method of vector database search using KNN as depicted in Fig. 5.4. To evaluate this method, we generate 20 new scenarios and test several vector search strategies, i.e., KNN [411], ED, K Decision Tree (KDT) [412], Hierarchical Navigable Small World (HNSW) [413], K Means (KM) [414], and Locality Sensitive Hashing (LSH) [415]. Fig. 5.12 shows a computation time comparison with variety of data quantities. Compared to the tested methods, KNN provides the most stable computational performance. Methods such as KDT, HNSW, and LSH provide a faster search time. However, these methods need preliminary computation to preprocess or generate the hash map. Therefore, these methods may not be suitable for fast pace changing data. Based on the search evaluation, the tested methods find the most related scenarios from the known scenarios. This strategy will serve as a resilient mechanism in identifying new possible APT cyber attack scenarios, i.e., zero-day attacks.

### 5.5 CONCLUSION

With the growing threat of cyber attacks on power grids, it is now more critical than ever to strengthen the attack detection capabilities in OT communication networks. It is important to note that, from 2024 onward, we will be living in a world where AI plays an increasing role alongside the advancement of AI models, i.e., deep learning, physic-informed, and generative AI models. In this context, our research aligns with this trend by proposing AI-based spatio-temporal APTs detection, correlation, and prediction in



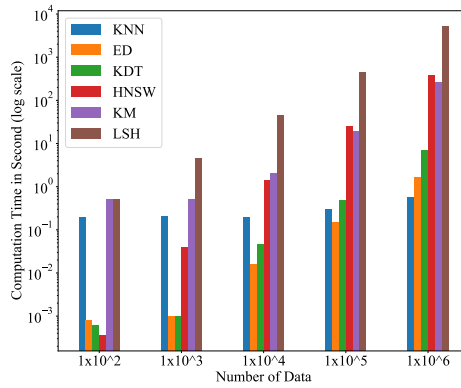


Figure 5.12: Search algorithms computation time comparison [7].

power systems. The implementation of deep learning for intrusion detection systems is becoming increasingly crucial to address the sophisticated and evolving nature of APTs. The proposed methods comprise of semi-supervised DPI, CPSIM, EGC-LSTM, a resilient associative method, and CPPS log comparator. The EGC-LSTM outperforms the state-of-the-art graph-based spatio-temporal deep learning model with the lowest MSE score of 0.0003. The proposed methods are also capable of identifying zero-day attacks, locating anomalous elements in the CPPS, and predicting the potential impact of anomalies. In contrast to most research that emphasizes the physical anomalies that occur during the later stages of a cyber attack on power systems, the proposed methods have the potential to detect cyber attacks during the early phases of the cyber kill chain. In addition, AI-based intrusion detection provides an online situational awareness for power system operators to pinpoint system-wide anomaly locations in near real-time and preemptively mitigate APTs at an early stage before causing adverse impacts.

In this work, the methods primarily focus on cyber anomalies that originate from external threat actors. The external APT required a lateral movement to reach its final objective in a timely fashion. These scenarios provide an opportunity for the early identification of the APT. However, there is also possibly an insider threat that can cause an immediate impact on the CPPS. Currently, the insider threat constraint is omitted from our objectives. Therefore, insider threat detection can become a potential future research direction, along with external threat detection. Furthermore, to enhance the methodology, it is essential to conduct comprehensive testing of the OT communication traffic by including a wider range of APT scenarios.



## 6

# INTRUSION DETECTION SYSTEM USING SEMI-SUPERVISED LEARNING AND SIMILARITY CLUSTERING

6

*Cyber attacks on power grids are imminent and potentially have a severe impact, as evidenced by the cyber attacks in Ukraine in 2015, 2016, and 2022. In response to this challenge, machine learning-based Intrusion Detection Systems (IDS) have become more prevalent as a potential mitigation owing to their alignment with the latest advances in artificial intelligence. However, existing anomaly detection methods for power grid OT are often inadequate, as they primarily focus on detecting power grid physical anomalies at the later attack stages and suffer from the scarcity of available data for supervised machine learning. To address these limitations, we propose a novel semi-supervised IDS specifically for digital substations of the power system. The proposed detection method identifies the distinctive distance similarity of digital substation OT communication traffic using Convolutional Neural Network and Chebyshev distance of packet payloads, and Kolmogorov-Smirnov of packets interarrival time using Fast Fourier Transform amplitude. Subsequently, these traffic features are combined into a vector and classified using a novel hybrid semi-supervised SOM and Density-Based Spatial Clustering of Applications with Noise DBSCAN. Results indicate that the proposed method is can identify zero-day attacks and achieve accuracy and F1 above 95%.*

## 6.1 INTRODUCTION

Cyber-Physical Power Systems (CPPS) are critical infrastructures that have experienced an increasing number of cyber attacks in recent years. In December 2015, a highly coordinated cyber attack had a major impact on the Ukrainian power grid, resulting a power outage for several hours [8]. In 2016, another cyber attack on Ukraine's power grid delivered a lower degree of impact compared with the 2015 attack [219]. The Ukrainian power grid also experienced a power outage in October 2022 due to the disruption caused by the Sandworm malware in its Operational Technology (OT) [11]. These incidents emphasize the imminent threat of cyber attacks on power grids.

Intrusion Detection Systems (IDSs) have emerged as a prominent solution for detecting anomalies in power grids [416–419], with some of them are focusing on the power grids digital substations [300, 418–422]. Recent research highlights a growing interest in machine learning-based IDSs due to their superior performances [422–424]. However, machine learning models are predominantly based on supervised learning, which requires large quantities of data to train effectively and achieve optimal performance. This requirement often contrasts with the limited availability of data [423], especially for zero-day attacks. Considering this limitation, a fully supervised machine learning model may not be the optimal choice. Therefore, in this study, we implement semi-supervised learning strategy to detect anomalies in digital substations of power grids and overcome the challenge of having limited data available. The technique leverages the advantages of the homogeneous characteristics of OT network traffic parameters generated from automated processes of machine-to-machine communications [312].

There are many parameters of network traffic that can be utilized as input features for an IDS [425]. Among various parameters, some IDSs quantify the traffic parameters as a distance similarity for anomaly detection [426]. In [7], anomaly detection was performed using the quantified distance similarity of packet payload. However, the distance similarities derived from packet payloads are inadequate for mitigating spoofing attacks due to insignificant differences in packet anomalies when compared to legitimate packets. As an alternative from payload-based anomaly detection, traffic interarrival time parameters have been extensively studied and implemented for IDS applications [427–430]. Traffic interarrival time is a relevant approach considering the homogeneous characteristics of digital substation traffic. Therefore, in this research, we proposed hybrid distance similarity parameters of digital substation network traffic based on the combination of packet payload and packet interarrival time.

Digital substation traffic under normal operating conditions is used as a reference point for anomaly detection. Based on the reference point, the packet payloads are quantified using CNN and Chebyshev distance. The application of CNN and Chebyshev distance for the intrusion detection system has been proven separately in [385, 422] and [431, 432]. The traffic interarrival time statistical features, i.e., mean and standard deviation, have been used as input features for anomaly detection in [427–430, 433]. However, these statistical features are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. Therefore, in this research, we introduce a novel packet interarrival time signature based on FFT and Kolmogorov-Smirnov. FFTs have been implemented for anomaly detection owing to their ability to identify anomalies within both the time and frequency domains of data [394, 434, 435]. Meanwhile, the Kolmogorov-

Smirnov method has been implemented for anomaly detection based on the statistical features of the data [436, 437]. In this research, the FFT converts the interarrival times of traffic into FFT amplitudes. The FFT amplitudes of a particular traffic is compared to the FFT amplitudes of a baseline traffic using the Kolmogorov-Smirnov in order to calculate the p-value, which indicates the statistical differences between the two. Subsequently, the parameters from CNN, Chebyshev distance, and Kolmogorov-Smirnov p-value are integrated into a three-dimensional vector representing traffic distance similarities.

The vectors representing the normal and anomalous traffic signatures are then used as input for semi-supervised classification. In this research, we proposed a classifier based on Self-Organizing Map (SOM) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The application of SOM and DBSCAN for the unsupervised intrusion detection system has been proven independently in [438–440]. Instead of using the clustering model independently, we proposed a novel hybrid classifier model for improving the classifier performances. A hybrid machine learning model is an emerging approach for improving the stand-alone models. With a hybrid model, the implementation of an algorithm can be strengthened through the advantages of other algorithms [441]. Because of some partially labeled data included in this process, we consider our hybrid unsupervised method as a semi-supervised one. SOM can be implemented for unsupervised classification to reduce data dimensionality and complexity [438, 439]. Subsequently, the DBSCAN is implemented to enhance the complex data clustering process. The hybrid combination of SOM and DBSCAN aims to improve classification performance for normal and anomalous traffic in digital substations.

The scientific contributions of this research are summarized as follows:

1. We propose a novel frequency domain interarrival time traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov. This method enhances the statistical-based methods that are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. Compared to statistical-based interarrival time, the combination of Fast Fourier Transform and Kolmogorov-Smirnov is able to improve the accuracy by 26% and F1 score by 41 %.
2. We propose a novel traffic distance similarity vector of operational technology communication traffic. The vector is derived from the packet payload and interarrival time. The vector quantifies the packet payload based on the Convolutional Neural Network and Chebyshev distance, and packet interarrival time using Fast Fourier Transform and Kolmogorov-Smirnov.
3. We propose a novel hybrid semi-supervised classification model based on Self-Organizing Map and DBSCAN. The hybrid combination of them aims to improve clustering performance and address the imbalanced dataset. Results indicate that the proposed method is able to identify zero-day attacks and achieve accuracy and F1 above 95
4. We propose a digital substation state transition model based on historical records of traffic distance similarity vectors and power system measurements. The historical state transition is visualized and analyzed for discriminating against anomalies due to cyber-attacks and physical disturbances. The visualization helps power system

operators track the state transition of digital substation traffic and discriminate traffic anomalies due to faults, reclosure, and spoofing attacks.

The chapter is structured as follows. Section II describes the method for semi-supervised IDS for digital substations. Section III presents the experimental result, and section IV presents the conclusion of the research.

## 6.2 SEMI-SUPERVISED INTRUSION DETECTION SYSTEM FOR DIGITAL SUBSTATION

This section explains the implementation of IDS for the power system digital substation and the methodology for the proposed hybrid semi-supervised IDS for the digital substation. Fig. 6.1 summarizes the model architecture of a digital substation, and Fig. 6.2 summarizes the overall architecture of the proposed methods. A more detailed explanation of the digital substation, proposed method, and corresponding processes are provided in the following subsections.

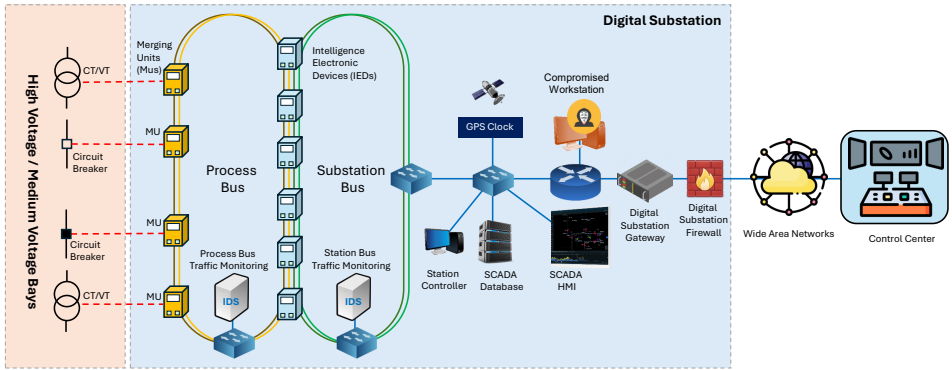


Figure 6.1: Digital substation architecture.

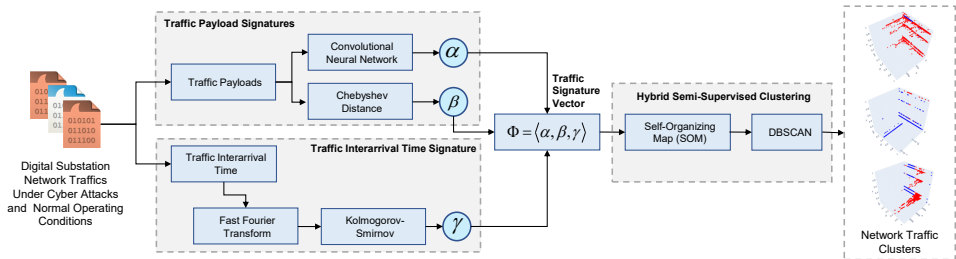


Figure 6.2: Summary of the proposed hybrid semi-supervised intrusion detection systems.

### 6.2.1 DIGITAL SUBSTATION ARCHITECTURE AND CYBER THREATS

The power grid's OT communication network consists of several network segments, including the control center, WAN, and digital substation [5, 7]. There are two primary network elements in the digital substation, i.e., the process bus and station bus, depicted in Fig. 6.1. The process bus connects the high/medium-voltage equipment in the field to the protection, control, and monitoring systems. In the process bus, MUs digitize the analog signals from the current transformers (CTs) and voltage transformers (VTs). These signals are communicated with the Intelligence Electronic Devices (IEDs) in the station bus. The IEDs monitor the status of circuit breakers, transformers, and other assets, and they also act on protective relays to isolate faults. Additional systems are present in the digital substation environment, e.g., Firewall, Gateway, GPS clock, SCADA database, and HMI.

Digital substations retain a critical role due to their ability to directly control physical power grids. As shown in the Ukrainian incidents [8, 219], the attacks originated from the IT network segment but only resulted in a disruption once it successfully reached the digital substation. In the Ukrainian power grid cyber attack in 2016, adversaries exploited the vulnerabilities of OT protocols in the digital substation to launch a malicious command for opening circuit breakers [219]. In [65], the authors demonstrated a spoofing attack in a simulated digital substation that successfully exploited the IEC61850 GOOSE packet for maliciously opening circuit breakers. These attacks show that existing digital substation communication lacks security measures.

In this research, we model a GOOSE spoofing attack on a digital substation. We assume that through lateral movement in the early stage of the cyber kill chain, the adversaries successfully compromise a workstation in the digital substation. Subsequently, adversaries launch cyber attacks from the compromised workstation. Workstations for traffic monitoring are deployed in the process bus and substation bus to identify the spoofing attack. These workstations can observe all traffic in the process and station buses via span port. Our proposed IDS for digital substation is implemented using the monitoring workstations.

### 6.2.2 TRAFFIC DISTANCE SIMILARITY VECTORS IN DIGITAL SUBSTATIONS

OT traffic poses a distinct characteristic in comparison to IT traffic. IT traffic typically originates from the human user with non-deterministic behavior, which makes it more heterogeneous. Meanwhile, OT traffic is generated by automated processes and machine-to-machine communications and is characterized by a greater level of homogeneity. While the OT traffic tends to be homogenous, this traffic remains poses a certain degree of variations due to traffic seasonality. Therefore, relying solely on statistical characteristics for OT traffic characterization is inadequate. In this research, we proposed a novel OT traffic distance similarity vector based on packet payloads and packet interarrival time.

Packet payload contains data that communicate between nodes in a network. This work characterized the packet payload using CNN and Chebyshev distance. The CNN classified the traffic into two classes, i.e., normal and anomalous. The packet payload from the nominal operating condition serves as a normal class, and the payload from other packets serves as an anomalous class. Considering the possibility of unknown payloads from a zero-day attack, in this research, we used a one percent anomalous payload for the

supervised CNN training. Eq. (6.1) shows the convolutional operation with  $x$  as input data from the packet payload,  $w$  represents the convolutional kernel, and  $b$  represents bias. The  $c_{in}$  represents the input dimension,  $k$  represents the kernel size, and  $c$  and  $j$  represent the index for the input and kernel. The output from the convolution operation ( $z$ ) serves as an input for the sigmoid function ( $\sigma$ ) in Eq. (6.2). The CNN operation in Eq. (6.1) and Eq. (6.2) produced a vector variable alpha ( $\alpha$ ). Eq. (6.3) shows the Chebyshev distance equation to obtain a vector variable  $\beta$ . This process is based on the input from the packet payload  $x$  and the average value of the normal traffic payload ( $n$ ).

$$z = \sum_{j=0}^{k-1} \sum_{c=0}^{c_{in}-1} x[i+j, c] \cdot w[j+c] + b \quad (6.1)$$

$$\alpha = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (6.2)$$

$$\beta = D_{\text{Chebyshev}}(x, n) = \max |x_i - n_i| \quad (6.3)$$

$$T_i^N = t_{i+1} - t_i \quad (6.4)$$

$$P = \sum_{i=0}^{N-1} \left| T[i] \cdot e^{-j \frac{2\pi}{N} k N} \right|^2 \quad (6.5)$$

$$D = \max_i |EDF_F(i) - EDF_{P_B}(i)| \quad (6.6)$$

$$\lambda = \left( \frac{N_P \cdot N_{P_B}}{N_P + N_{P_B}} + 0.12 + \frac{0.11}{\sqrt{N_P + N_{P_B}}} \right) \cdot D \quad (6.7)$$

$$\gamma = 1 - p_{\text{value}} = 1 - \left( 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 \lambda^2} \right) \quad (6.8)$$

$$\Phi = \langle \alpha, \beta, \gamma \rangle \quad (6.9)$$

Packet interarrival time ( $T$ ) is calculated based on the individual packet arrival time ( $t$ ) for all packet quantities ( $N$ ) depicted in Eq. (6.4). The interarrival time serves as an input to calculate FFT amplitude ( $P$ ) in Eq. (6.5). The parameters  $N$  in Eq. (6.5) represent the sliding window size for all interarrival data ( $N$ ). The FFT amplitude ( $P$ ) serves as an input for the KS equations in Eqs. (6.6, 6.7, 6.8). In Eq. (6.6), the maximum absolute difference ( $D$ ) between the two Empirical Distribution Functions ( $EDF$ ) is based on the calculated FFT amplitude ( $P$ ) and the baseline FFT amplitude ( $P_B$ ). The value of  $D$  is then used to calculate  $\lambda$ , which represents an asymptotic approximation formula in Eq. (6.7). Subsequently, the  $\lambda$  is then used to calculate vector variables  $\gamma$  based on  $p_{\text{value}}$  in Eq. (6.8). Finally, all vector variables are combined into a traffic distance similarity vector ( $\Phi$ ) depicted in Eq. (6.9).



### 6.2.3 HYBRID SEMI-SUPERVISED INTRUSION DETECTION SYSTEM

The traffic characterization vector ( $\Phi$ ) serves as input for the hybrid semi-supervised IDS. In this research, we proposed a hybrid sequential clustering process using SOM and DBSCAN. Initially, SOM trains a model ( $g$ ) to find the *Best Matching Unit (BMU)* depicted in Eq. (6.10). The best model is represented by the weight ( $w_g$ ) with the closest to the vector  $\Phi$ . The model with the BMU ( $g$ ) is used to perform a mapping from the three-dimensional vector  $\Phi$  into a two-dimensional vector  $\Omega$  depicted in Eq. (6.11). The SOM application aims to reduce data dimensionality and complexity into a lower-dimensional space while preserving the higher-dimensional structure of the data. Therefore, despite the reduction in dimensionality, the new vector  $\Omega$  inherits the features of the payload and interarrival time traffic characteristics.

$$g^* = \arg \min_g \|\Phi - w_g\|_2 \quad (6.10)$$

$$\Phi \rightarrow \Omega_{g^*} \quad (6.11)$$

$$C_q = \text{DBSCAN}(\Omega) \quad (6.12)$$

$$C_q \xrightarrow{c_B} C_{q=2} \quad (6.13)$$

The output vector  $\Omega$  from SOM is used as an input for the DBSCAN clustering algorithm shown in Eq. (6.12). The DBSCAN classifies the data into clusters ( $C$ ) with the number of clusters  $q$ . The proposed IDS is intended to categorize the traffic characteristics into two classes, i.e., normal and anomalous. However, the DBSCAN algorithm can potentially generate more than two clusters, which does not align with the intended outcome of two classes expected by the IDS. In this case, the number of DBSCAN clusters ( $q$ ) is reduced into two cluster categories through a mapping depicted in Eq. (6.13). The mapping is performed based on the baseline data cluster from the normal traffic ( $c_B$ ). Therefore, if a cluster predominantly contains values associated with a normal class in partially labeled data, it will be mapped into a normal class. Otherwise, it will be mapped to an anomalous class. Because of some partially labeled data included in this process, we consider our hybrid method as a semi-supervised clustering. Based on the aforementioned steps, the hybrid SOM and DBSCAN process is then used to classify traffic distance similarity vectors into two classes: normal and anomalous traffic.

### 6.2.4 DIGITAL SUBSTATION TRAFFIC STATE TRANSITION MODEL

To comprehensively analyze the cyber-physical system state of digital substations, we collect historical information on traffic distance similarity vectors and power system measurements depicted in Fig. 6.3. The collected information is then analyzed on a state transition model based on a time series plot and a 3-dimensional vector plot. The time series plot presents time series data from distance similarity vectors and power system measurements. Using the time series plot, this method can show the anomaly correlation between digital substation traffic and power system measurements. Meanwhile, the 3-dimensional vector plot presents the distance similarity vector transition in the digital

substation. These visualizations aim to help power system operators track the state transition of digital substation traffic and discriminate traffic anomalies due to faults, reclosures, and spoofing attacks.

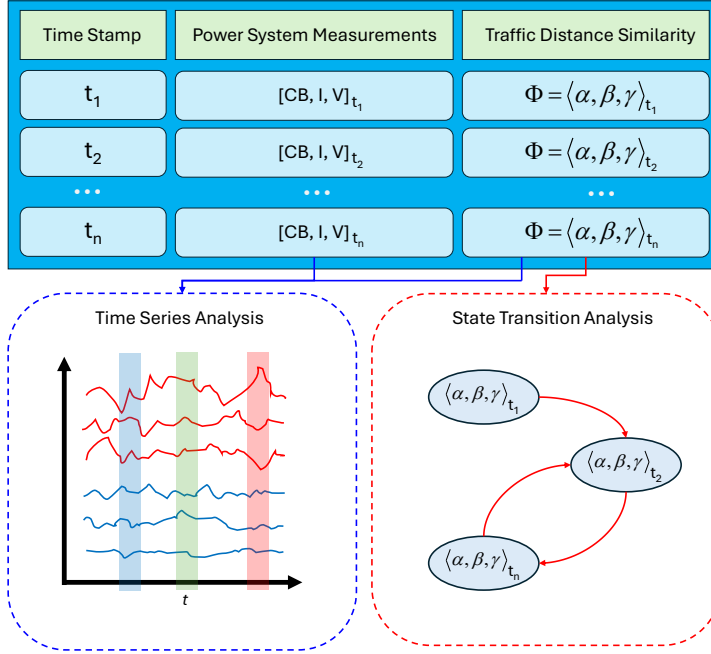


Figure 6.3: Digital substation cyber-physical system state transition.

## 6.3 EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we present the experimental results of the proposed methods, including a detailed overview of the dataset used, findings from traffic payload and interarrival time distance similarity, the performance of a hybrid semi-supervised IDS, and state transition analysis. The following subsections provide a more detailed explanation of the experimental results.

### 6.3.1 EXPERIMENTAL SETTING AND DATASET

Fig. 6.4 depicts the architecture of a HIL digital substation utilized in our experiments. The digital substation simulation consists of a power system simulator and OT networks. The power system simulations were implemented using a RTDS. The OT communications are implemented using several IEDs and computers. The IEDs implement the IEC 61850 standard, which ensures that the devices can utilize GOOSE messaging and employ SV for measurements. In this research, we are focusing on the GOOSE protocol, which represents control functionality to open or close circuit breakers in the simulated power system. In the simulated cyber attack, the adversaries gained control of a computer in the substation and

utilized it to carry out the cyber attack. From the computer, the adversary performs several cyber attack scenarios, i.e., network scanning, simulating malware traffic, and GOOSE spoofing attacks. During the attack, the IDS monitors the network traffic from the span port of a switch in the HIL setup. The monitored traffic is used for the implementation of the proposed hybrid semi-supervised IDS.

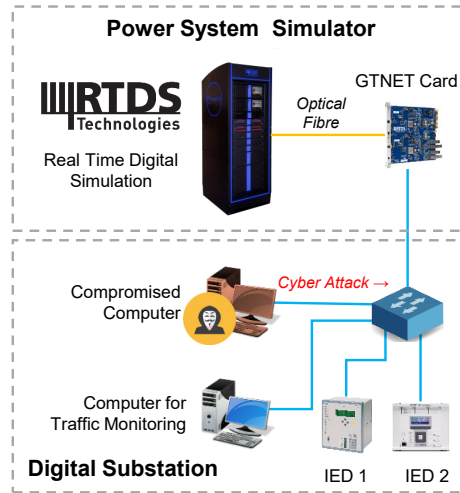


Figure 6.4: Digital substation co-simulation architecture with HIL.

Table 6.1: Comparison of GOOSE Traffic Data

Data	No. of IEDs	No. of Packets	$\Sigma$ .pcap File Size
A	2	27,259	5.31 MB
B[369]	18	328,017	56.5 MB
C[442]	9	1,048,576	895 MB

The experiment utilizes data derived from the simulation conducted in the HIL simulation setup depicted in Fig. 6.4. This data is denoted as data A and includes the normal operating conditions of IEC61850 GOOSE traffic from two IEDs depicted in Fig. 6.3. Other than data A, we also use other GOOSE attack data from the IEC61850 Security [369] and Power Duck [442]. In this study, these two data are denoted as data B and C. The data A, B, and C have similarities in the GOOSE communication protocols for both normal and anomaly conditions. Table 6.1 summarizes the comparison of all data based on the number of IEDs, the number of GOOSE packets, and the total all .pcap file size. All data primarily represents the GOOSE traffic under normal and anomaly conditions. Therefore, in this research, we focus on the characterization of the GOOSE under normal and anomaly conditions. Other than the GOOSE traffic data, we incorporate data from other types of cyber attacks, including DoS and network reconnaissance. The DoS and reconnaissance data are collected from the attack simulation using tools hping3 and Nmap. In addition, we

also incorporate sample attack traffic from publicly available data in [443]. The total size for all cyber attacks is 1.5 GB.

### 6.3.2 TRAFFIC PAYLOAD DISTANCE SIMILARITIES

The traffic from the three datasets has different average sizes. The average size from the data A, B, and C are 169.58, 165.31, and 217.02 Bytes. Fig. 6.5 summarizes the traffic size statistical characteristics in a box plot. From all data, the maximum traffic size is 250 Bytes, originating from data C. The cyber attack data has an average size of 304.74 Bytes. The different sizes of cyber attacks and the GOOSE lead to inconsistencies in the size of the payload characterization processes. Therefore, in the experiment, we decided to standardize the packet size to 256 Bytes to cover all possible GOOSE data sizes. When cyber attack traffic exceeds 256 bytes, the extra bytes are discarded. Conversely, when the GOOSE or cyber attack traffic is less than 256 bytes, the remaining spaces are filled with zeros. This process ensures the consistency of the traffic payload characterization using CNN and Chebyshev distance. The Chebyshev distance calculation is implemented using Scipy<sup>1</sup> library, and the CNN is implemented using Tensorflow<sup>2</sup> library.

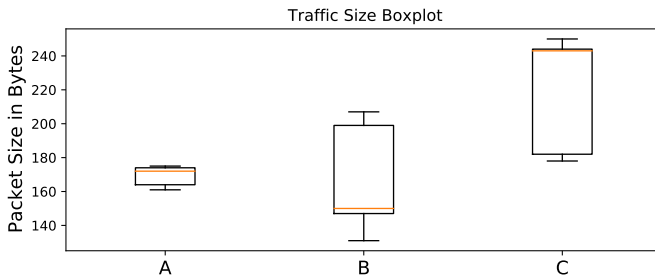


Figure 6.5: Comparison of packet size statistical characteristics from data A, B, and C in box plots.

Alternative distance measurement methods for anomaly detection exist beyond Chebyshev distance [431, 432]. In the experiment, we also evaluate alternative distance measurement methodologies, including Minkowski [444], Euclidean [445], Hamming [446], and Jaccard distances [447]. Our experiment evaluates these distance methods across all datasets utilizing feature selection through a Random Forest [448] and ANOVA test [449]. Upon assessing five distance metrics, the average ranks are as follows: Chebyshev (1.92), Minkowski (2.19), Euclidean (2.32), Hamming (4.74), and Jaccard (4.83). The overall result shows that Chebyshev poses the best average rank. Based on this result, our method uses Chebyshev as the best distance parameter.

Fig. 6.6 shows the scatter plot based on the CNN ( $\alpha$ ) and Chebyshev distance ( $\beta$ ) for three types of traffic samples. The plot shows that other types of cyber attacks are significantly different from the GOOSE packets with a higher  $\alpha$  and  $\beta$  score. However, the GOOSE normal and GOOSE spoofing are relatively the same with zero values for  $\alpha$  score and scattered values for  $\beta$  score. Therefore, the  $\alpha$  and  $\beta$  scores are unable to

<sup>1</sup> <https://scipy.org/>

<sup>2</sup> <https://www.tensorflow.org/>

discriminate between GOOSE normal and GOOSE spoofing. This is because the spoofing attack originated from the normal one, and the payloads from both packet categories exhibit similarities. Based on the plot, it appears that using CNN and Chebyshev distance could be effective in discriminating cyber attacks in digital substations. Consequently, CNN and Chebyshev distance are potentially sufficient to discriminate against other types of cyber attacks in digital substations including traffic from zero-day attack cases. However, these parameters are insufficient to discriminate between normal and spoofing GOOSE. Therefore, more features are needed to identify the distinction between them.

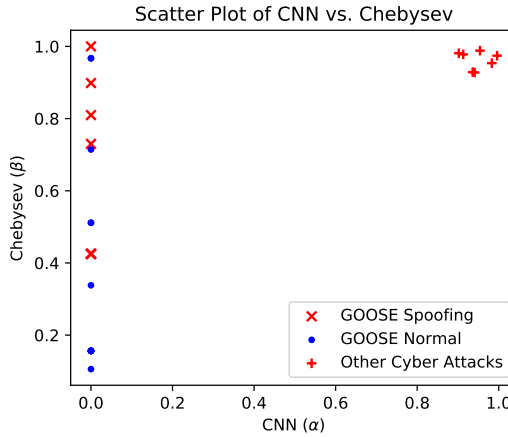


Figure 6.6: scatter plot based on the CNN ( $\alpha$ ) and Chebyshev distance ( $\beta$ ) for GOOSE normal and spoofing, and other cyber attacks.

### 6.3.3 TRAFFIC INTERARRIVAL DISTANCE SIMILARITIES

Packet interarrival time is the duration between the arrivals of consecutive data packets in a network. It is an essential measure to analyze the patterns of network traffic, enabling network performance evaluation, congestion, and the effectiveness of data transfer. In this research, we use traffic interarrival time to capture the communication signature and identify traffic anomalies. Fig. 6.7 depicts the plot for visualizing traffic characteristics from data A, B, and C. The blue area and blue line represent normal traffic. Otherwise, the red area and red line represent anomalous traffic.

The top row in Fig. 6.7 shows the probability density distribution with the normal distribution curve for interarrival time. The red and blue area represents the probability density distribution in a bar chart. Meanwhile, the red and blue lines represent the bell curve of the data normal distribution. Based on the plot, the interarrival time of the GOOSE attack tends to be smaller than the normal one. The bottom row in Fig. 6.7 depicts the FFT amplitude for the traffic interarrival time. The processes of obtaining the FFT power amplitudes are based on the Eq. (6.4) and Eq. (6.5). Subsequently, this information is then used to obtain  $\gamma$  value based on the Kolmogorov-Smirnov process depicted in Eqs. (6.6, 6.7, 6.8).

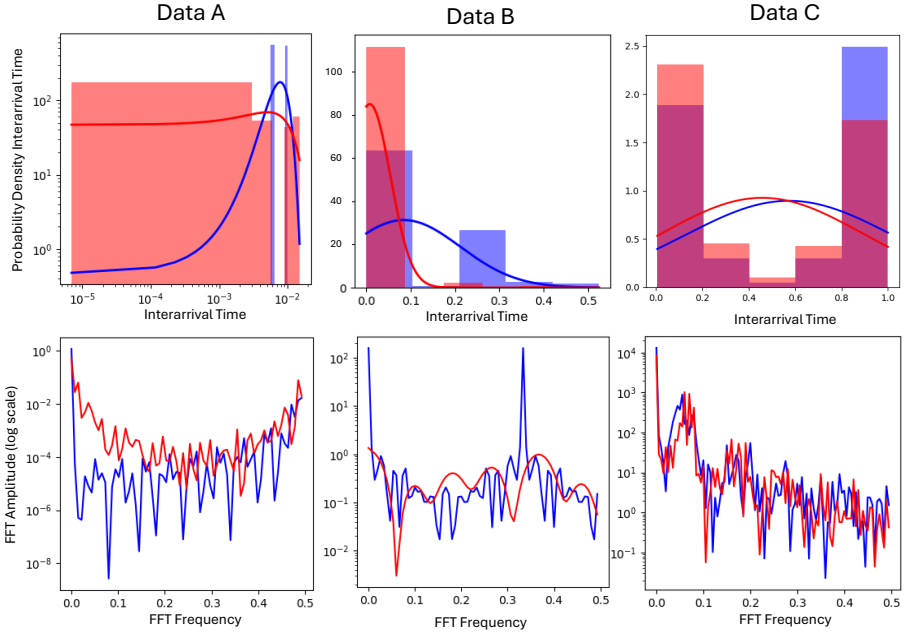


Figure 6.7: Characteristic of interarrival time from sample data A, B, and C. The top row plots show probability density distribution with the normal distribution curve of interarrival time. The bottom row plots show FFT amplitude for the traffic interarrival time.

### 6.3.4 HYBRID SEMI-SUPERVISED CLASSIFIER

The hybrid semi-supervised traffic classifier uses the vector  $\Phi$  as an input. The visualization of the 3D vector plot for data A, B, and C is depicted in Fig. 6.8. The blue node represents normal GOOSE traffic, and the red node represents anomalous GOOSE traffic. Every axis in the plot represents the vector element of  $\alpha$ ,  $\beta$ , and  $\gamma$  respectively. As shown in Fig. 6.8, the normal data is not always concentrated. There is also a possibility where the normal data is scattered, e.g., data B and C. The data B and C pose more scattered normal data due to more variability in the normal GOOSE data. For example, data B considers normal variable loading in the GOOSE traffic, which significantly makes the normal data plot more scattered. The plot in Fig. 6.8 also shows that the normal GOOSE traffic signature from different digital substations is unique. Therefore, we propose that this distinctive signature can serve as a reference point for comparing the anomalous GOOSE traffic and use this distance similarity parameter to perform a hybrid semi-supervised traffic classifier. Table 6.2 shows the performance comparison of several clustering methods including KNN [450], Agglomerative clustering [451], GMM [7, 452], SOM [438, 439], DBSCAN [440], and our proposed hybrid SOM and DBSCAN. The SOM is implemented using the Minisom<sup>3</sup> library, and DBSCAN is implemented using the Scikit-learn<sup>4</sup> library.

The top of Table 6.2 shows the performance benchmarking for input  $\psi = \langle \alpha, \beta, \tau \rangle$ ,

<sup>3</sup><https://github.com/JustGlowing/minisom>

<sup>4</sup><https://scikit-learn.org/stable/>

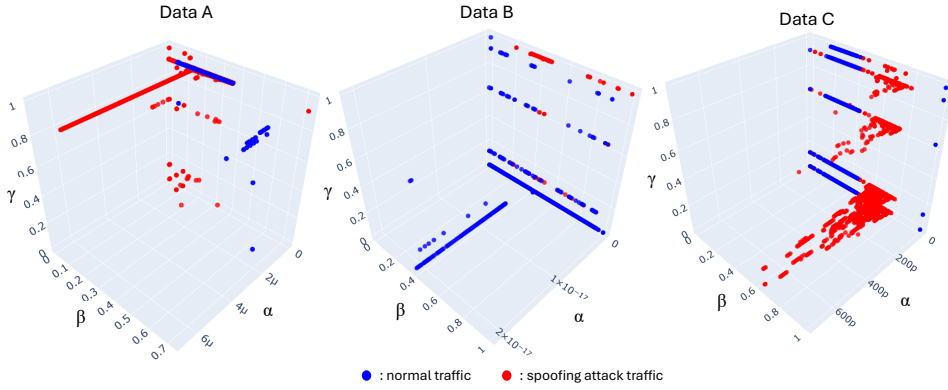


Figure 6.8: 3D scatter plots for data A, B, and C with the blue nodes represent normal GOOSE and red nodes represent anomalous GOOSE.

Table 6.2: Performance Comparison of Clustering Methods

Methods	Data A		Data B		Data C		Average	
	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1
Input $\psi = \langle \alpha, \beta, \tau \rangle$								
KNN	0.5665	0.4959	0.5689	0.4981	0.5755	0.5076	0.5703	0.5005
Agglomerative	0.5815	0.5814	0.5522	0.4543	0.5876	0.5024	0.5738	0.5127
GMM	0.5065	0.4935	0.5061	0.4831	0.5062	0.4867	0.5063	0.4876
SOM	0.8166	0.5649	0.8151	0.5543	0.8171	0.5837	0.8163	0.5676
DBSCAN	0.8459	0.4895	0.8319	0.4842	0.8403	0.4917	0.8394	0.4885
<b>SOM DBSCAN</b>	<b>0.6401</b>	<b>0.6283</b>	<b>0.6288</b>	<b>0.6184</b>	<b>0.6497</b>	<b>0.6363</b>	<b>0.6395</b>	<b>0.6277</b>
Input $\Phi = \langle \alpha, \beta, \gamma \rangle$								
KNN	0.5615	0.4595	0.5046	0.4176	0.7368	0.7052	0.6010	0.5274
Agglomerative	0.5477	0.4332	0.8899	0.7675	0.7541	0.7353	0.7353	0.6924
GMM	0.6011	0.5994	0.9394	0.7697	0.7697	0.7910	0.7701	0.7200
SOM	0.8471	0.7582	0.9977	0.9951	0.9999	0.9999	0.9482	0.9177
DBSCAN	0.8424	0.4819	0.9906	0.9086	0.9969	0.6653	0.9461	0.6853
<b>SOM DBSCAN</b>	<b>0.8903</b>	<b>0.8823</b>	<b>0.9942</b>	<b>0.9942</b>	<b>0.9976</b>	<b>0.9976</b>	<b>0.9607</b>	<b>0.9580</b>

where  $\tau$  represents Kolmogorov-Smirnov from interarrival time. Meanwhile, the bottom of Table 6.2 shows the performance benchmarking for the input  $\Phi = \langle \alpha, \beta, \gamma \rangle$ , where  $\gamma$  represents the Kolmogorov-Smirnov from the FFT amplitude of interarrival time. From the result in Table 6.2, the input  $\psi$  has an average accuracy of 0.6576 and an average F1 of 0.5307. For the input  $\Phi$  has an average accuracy of 0.8299 and an average F1 of 0.7452. Based on these results, implementation of the Kolmogorov-Smirnov from FFT amplitude of interarrival time improves the accuracy by 26% and the F1 score by 41%.

We evaluate the performance based on the accuracy and F1 score. According to the result, some methods provide relatively high accuracy. However, the accuracy metric can be misleading when dealing with imbalanced datasets, as it may produce high accuracy if the model only correctly predicts the majority class and ignores the minority class. Therefore, in order to evaluate the overall performance, we determined that accuracy and F1 scores were equally important. Out of all the methods in Table 6.2 for the input  $\Phi$ , SOM achieves the highest accuracy and F1 score as a standalone method, while DBSCAN comes in second. However, neither of them can provide the best performance compared to our proposed

method, the hybrid SOM and DBSCAN. On average, the hybrid method is able to result in an accuracy of 96% and F1 of 95% for input  $\Phi$ . In Table 6.2 for the input  $\psi$ , the hybrid SOM and DBSCAN do not provide the best accuracy. However, it increases the F1 score. From Table 6.2, overall, the hybrid is able to improve the F1 score. Therefore, the hybrid method is suitable for addressing an imbalanced dataset.

### 6.3.5 DIGITAL SUBSTATION STATE TRANSITION ANALYSIS

This section presents the state transition analysis of digital substations based on traffic distance similarity vectors and power system measurements. The analysis primarily highlights the state transition between traffic anomalies due to faults, reclosures, and spoofing attacks. In a digital substation, an anomaly of GOOSE traffic is also possibly caused by a fault in a power system. When a fault occurs in the power system, the IEDs detect the abnormal condition. Using the GOOSE protocol, these IEDs send an instantaneous trip command to the breaker. In this work, we also simulate GOOSE packets due to a fault in the power system. Fig. 6.9 depicts the comparison of the average value of  $\alpha, \beta, \gamma$  under normal, fault, and spoofing traffic. Fig. 6.10 shows the 3D plots of traffic under normal, fault, and spoofing. Based on these plots, the traffic from normal, fault, and spoofing show a distinct characteristic. We implemented the hybrid SOM DBSCAN to classify the data depicted in Fig. 6.8 into three classes. The result shows that the hybrid SOM DBSCAN can achieve an accuracy of 0.8889 and an F1 score of 0.8785. Although the result is not as good as the two-classes classification, this result indicated that our proposed method is also capable of discriminating fault and spoofing.

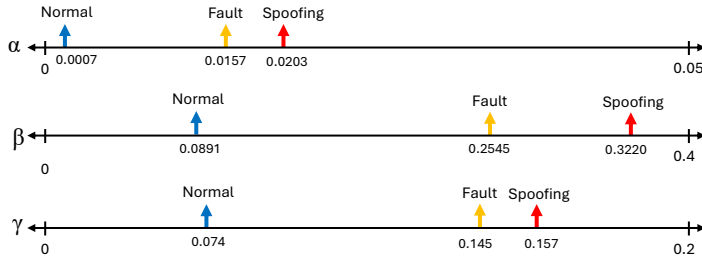
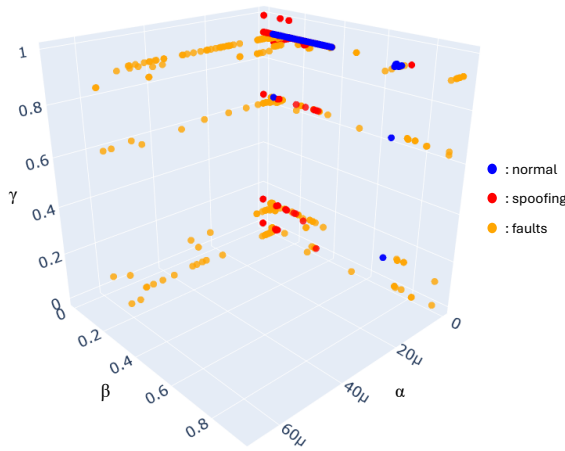


Figure 6.9: Comparison of average value of  $\alpha, \beta, \gamma$  under normal, fault, and spoofing traffic.

We conduct experiments for analyzing the performance of the methods under the state transition from the normal, faults, reclosure, and spoofing attack. The experiments are conducted using the HIL testbed, while the IEEE 5-bus system utilizes a real-time implementation and testing of virtualized controllers for software-defined IEC 61850 in digital substations. Fig. 6.11 and Fig. 6.12 show an overview of the transition from the simulated scenarios. Initially, the system runs under normal operating conditions. In step 1, at  $t=1.4$  s, a three-phase fault occurs in a connected transmission line. In step 2 at  $t=1.5$  s, the IED responds to the fault by sending the GOOSE trip command, based on the distance protection scheme. The trip command triggers a CB to open and change the CB status from 1 to 0. In a digital substation, the trip command is sent by a protection IED to CB circuit breakers when a fault is detected. This action isolates the fault, protects equipment, and



ensures system stability. After the fault is cleared, the IED aims to restore the system state to the normal operating condition. Therefore, in step 3 at  $t=5.7$  s, the IED sends reclosure commands to close the CB and change the CB status from 0 to 1. After the reclosure, as shown in Fig. 6.12, the system returns to the nominal operating conditions. In step 4, at  $t=9.1$  s, GOOSE spoofing packets are injected into the digital substation. The spoofing GOOSE traffic instructs to open CB. Consequently, this command led to the current anomaly. The measured current, obtained by the Current Transformer (CT) in the substation, is zero after the CB opens. Due to the direct connection of the substation to a generator, the voltage is equal to the generator output voltage.



6

Figure 6.10: 3D scatter plots for traffic under normal, spoofing, and faults.

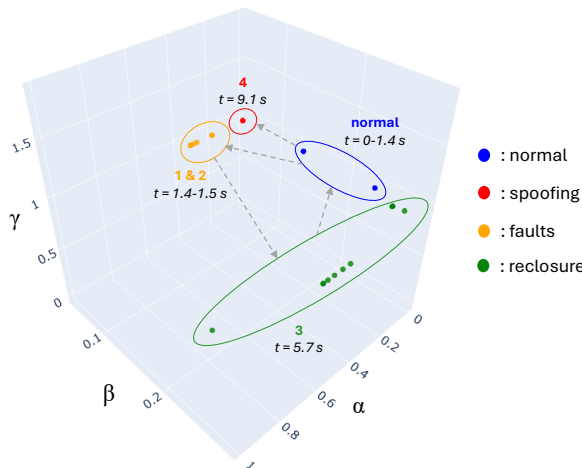


Figure 6.11: 3D scatter plots for traffic transition from normal operation, faults (1 and 2), reclosure (3), and spoofing (4).

Based on the experiment, we analyze the state transition of traffic distance similarities from different steps shown in Fig. 6.11. The plots show that every step poses different vector representation values. Therefore, it indicates that our proposed method can distinguish between normal, faults, reclosure, and spoofing attack. This capability is crucial for the power system operator to identify the anomalies due to fault or cyber attack. To improve the confidence of accuracy, the fault can also be detected through measurement anomaly, as shown in Fig. 6.12. Meanwhile, the confidence level of cyber attack detection can be improved using anomalies from other cyber attack traffic depicted in Fig. 6.6. Based on the cyber kill chain and the real cyber attack in Ukrainian power grids, the adversaries potentially trigger traffic anomalies in the early stage of the cyber kill chain. Therefore, the early stage anomaly detection in the early stage of the cyber kill chain can improve the confidence of anomaly detection in digital substations due to spoofing attacks.

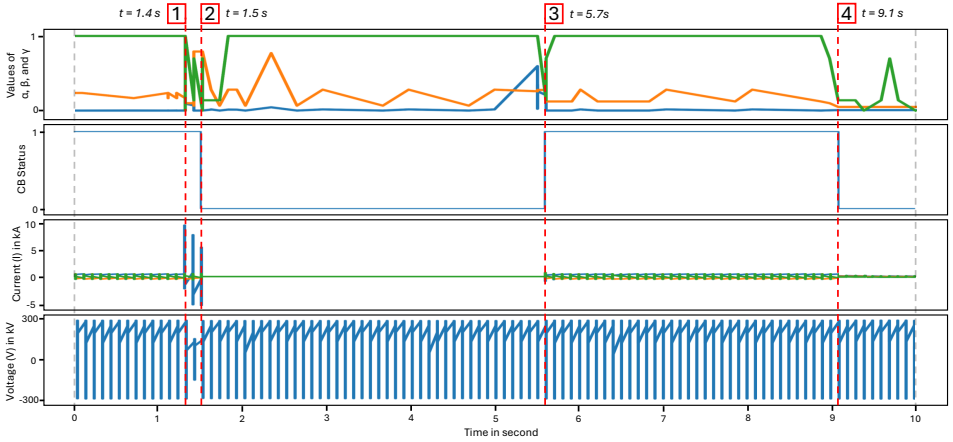


Figure 6.12: Time series plot from the simulated events representing value of traffic vector  $\Phi = \langle \alpha, \beta, \gamma \rangle$ , circuit breaker status, current (kA) and voltage (kV).

## 6.4 CONCLUSION

In light of the increasing threat that comes from cyber attacks targeting power grids, it is critical to enhance the capabilities for detecting such attacks in the power grids' OT systems. It is necessary to acknowledge that we are now living in a world where AI is playing an expanding role, particularly with the development of advanced AI models such as deep learning, physics-informed models, and generative AI models. However, existing AI-based IDS for power grids are limited in their ability to achieve optimal performance because of the limited availability and access to the data from the power grid's OT.

In this study, we introduced an innovative hybrid semi-supervised method to detect anomalies in power grid digital substations. The proposed method aims to address limited data availability for AI model training, especially for zero-day attacks. This method works based on the novel digital substation distance similarity vector, which consists of traffic payload and traffic interarrival time distance similarities. The experimental results

demonstrate that our hybrid SOM and DBSCAN algorithm outperforms other clustering methods, achieving accuracy and F1 score levels exceeding 95% respectively. In addition, the implementation of the Kolmogorov-Smirnov from FFT amplitude of interarrival time improves the accuracy by 26% and the F1 score by 41%. The method can also classify normal faults and spoofing with an accuracy of 0.8889 and an F1 score of 0.8785. In future works, our proposed methods can be enhanced by incorporating power system measurement data. This will enable a more thorough evaluation of the cyber-physical power system. In addition, the traffic distance similarity vector and hybrid semi-supervised model can also be implemented in other OT communication protocols beyond IEC61850.



## 7

## CONCLUSION AND DISCUSSION

### 7.1 CONCLUSION

This thesis addresses the challenges of the advanced persistent threats on cyber-physical power systems. To address the challenges, this thesis proposes APTs detection and correlation methods for CPPS by considering APT characteristics, including stealthy tactics, prolonged persistence, and exploitative use of zero-day vulnerabilities. The thesis is organized into five chapters with five RQs dedicated to addressing specific challenges. The detailed results and conclusions from every chapter and RQ are listed below.

**Chapter 2 (RQ 1):** *How to unveil the characteristics of cyber attacks on power grids by considering the cyber attack taxonomy, impacts, power grids OT vulnerabilities, and network security control? How to design a high-fidelity model of the cyber attack on power grids?*

Chapter 2 investigates the evolving correlation between technology and vulnerability in power systems, where the drive for efficiency through digitalization comes at the cost of increased attack surfaces for cyber threats. The investigation results provide a taxonomy of cyber attacks that are specific to power grids, categorizing a variety of attack types that threaten the stability and reliability of the system. This chapter also unveils vulnerabilities in power grid operational technology, specifically targeting vulnerabilities in communication protocols and software applications that may be susceptible to exploitation. To address these vulnerabilities, the implementation of secure communication protocols and advanced network security measures are highlighted as essential to improving the grid's resilience against cyber threats. Furthermore, chapter 2 also presents the design of a high-fidelity cyber-physical co-simulation model for power grids. The co-simulation model is integrated with cyber range to enable controlled simulations of cyber attacks. This CPS model integrates digital and physical dimensions, providing a basis for developing high-fidelity cyber-physical power system simulation. From this standpoint, the proposed model is fundamental for experimenting with unforeseen cyber threats to power grids. In summary, chapter 2 serves as a crucial foundation for comprehending the new domain of cyber attacks on power grids by examining its vulnerabilities, emerging solutions, and experimental design. The fundamental understanding serves as a compass for directing research towards viable methods for enhancing cyber security in power grids.

**Chapter 3 (RQ 2):** *How to identify stages of APT targeting cyber-physical power systems that incorporate IT and OT attack stages while integrating the complex operational conditions of power systems?*

Chapter 3 presents a novel Advanced Cyber-Physical Power System (ACPPS) kill chain for identifying APT stages in CPPS. The proposed ACPPS kill chain incorporates IT and OT attack stages and considers the complex operational conditions of power systems. There are six major stages in the ACPPS kill chain, including attack preparation, initial engagement, main attack, physical system engagement, power system impact, and social impact and recovery. These stages have more detailed substages which consider more detailed processes in IT/OT and power systems. The key contribution of the ACPPS kill chain lies in its profound integration of cyber and physical processes within the power system. This integration transcends traditional boundaries, emphasizing a systemic perspective where the cyber attack process in the IT/OT can affect power systems in various spectrums of attack impacts, including power system anomalies, point of no return, cascading failures, and wide area system backout. Therefore, the ACPPS kill chain enables power system operators to navigate the complex interaction of cyber attacks within a cyber-physical power system and facilitate comprehensive interaction with both digital and physical aspects of system resilience. This chapter marks a pioneering step toward integrating cyberspace and physical power systems into a unified cyber-physical system. It highlights a broader perspective on anomalies, considering both cyber and physical aspects. Additionally, it provides a comprehensive, staged analysis representing the interactions and processes between the cyber and physical components of power grids. This chapter reflects the changing paradigm of convergence, where cyber-physical systems transcend their duality to be regarded as a unified entity.

## 7

**Chapter 4 (RQ 3):** *How to detect stealthy APTs on power grids with minimum anomalies and insignificant changes compared to legitimate traffic?*

Chapter 4 presents CyResGrid to detect stealthy APTs in power grids with infinitesimal anomalies and insignificant deviations from legitimate traffic that implemented using a SDN. CyResGrid is specifically developed to identify cyber attacks during their early stages, such as network reconnaissance, by analyzing OT network traffic throughput. It enables operators to pinpoint and monitor system-wide attacks through an attack graph map that provides a topological visualization of potential threats in near real-time. The core of CyResGrid is a hybrid deep learning model, which classifies OT network traffic as normal or anomalous. This model integrates GC-LSTM with a CNN. The hybrid models provides superior accuracy in detecting anomalies caused by cyber attacks. The GC-LSTM initially normalizes traffic data, learning the patterns of OT network throughput. Then, a finely tuned CNN, optimized via Bayesian methods, detects the OT traffic anomalies. This combined model achieves high performance and reduces false positive rates. The results are indicated by enhanced Geometric mean and F1 scores, outperforming state-of-the-art deep learning-based classifiers. Furthermore, by continuously monitoring network throughput and mapping detected anomalies on an attack graph, CyResGrid empowers operators to respond swiftly and effectively addressing the challenge of stealthy APT detection in power grids. This innovative approach enhances real-time situational awareness, transforming it into a profound defense against APTs in power grids, particularly in the infinitesimal anomalies. Through the implementation of a highly sensitive detection system for in-

infinitesimal anomalies, the proposed method enhances understanding of the operational technology characteristics of power grids and enables the mitigation of stealthy APTs.

**Chapter 5 (RQ 4):** *How cyber and physical anomalies caused by APTs in cyber-physical power systems can be effectively detected and correlated, particularly in the context of prolonged attacks with non-deterministic temporal anomaly instances?*

Chapter 5 presents a novel spatio-temporal detection and prediction framework that enables power system operators to locate and respond to anomalies in near real-time. This method monitors OT communication traffic using distributed semi-supervised DPI classifiers connected to SDN-enabled switches. Based on this connected architecture, each observation point provides information to the SDN controller, constructing a cyber anomaly graph and a power system graph based on circuit breaker statuses. To correlate cyber and physical anomalies, a Cyber-Physical System Integration Matrix (CPSIM) is used, which enables spatio-temporal correlation and situational awareness. The EGC-LSTM model processes these correlations, with sequential and neural network filters predicting potential APTs propagation. This method extends standalone detection by integrating both cyber and physical components. A comprehensive understanding of complex cyber-physical systems is required for addressing anomalies in cyber-physical power systems. In such interconnected systems, cyber and physical elements cannot be fully separated as they are interdependent. Furthermore, the AI-based method also provides an online situational awareness for power system operators to pinpoint system-wide cyber-physical anomaly locations in near real-time and preemptively mitigate APTs at an early stage before causing adverse impacts.

**Chapter 6 (RQ 5):** *How to detect APT's zero-day attack in a power system considering behavioral characteristics of OT communication traffic and limited preliminary knowledge about the attacks?*

Chapter 6 presents detection strategies for APT's zero-day attack based on the characteristics of power system OT communication traffic. The OT traffic is characterized by a distance similarity vector derived from packet payload and interarrival time. The payload parameters are obtained using Chebyshev distance and CNN. The interarrival time parameter is obtained using novel time-frequency traffic characterization based on the Fast Fourier Transform and Kolmogorov-Smirnov. This time-frequency method enhanced statistical-based methods that are unable to adequately discriminate between normal and anomalous traffic due to the insignificant distinctions between them. The payload and interarrival time parameters are combined into distance similarity vectors as an input for clustering algorithms. Subsequently, a hybrid semi-supervised clustering based on SOM and DBSCAN is implemented to classify the traffic as normal or anomalous. The hybrid combination of them aims to improve classification performance and address the imbalanced datasets. The findings demonstrate that the proposed method effectively identifies zero-day attacks while enhancing accuracy and F1 score. This finding aligns with the needs for zero-day attack detection and the growing cyber threats to power grids. Our model is intended to mitigate potential future attacks, by providing a proactive strategy for protecting power grids. This ability to mitigate zero-day threats not only improves cyber threat intelligence but also substantially strengthens the resilience of power systems against unforeseen attacks.

## 7.2 DISCUSSION AND FUTURE RESEARCH

This thesis presents scientific advancements in anomaly detection for cyber-physical power systems by addressing critical limitations in state-of-the-art research. Traditional anomaly detection approaches primarily rely on signature-based methods and supervised learning models, which are constrained by their dependence on extensive data availability for generating signatures or training AI models. These limitations render them ineffective in detecting zero-day attacks. Additionally, current methods focus on detecting individual anomaly instances without considering the correlations between them, making them inadequate for addressing the nondeterministic nature of APTs. To overcome these challenges, this thesis introduces a novel approach to anomaly detection rooted in the fundamental characteristics of OT communication traffic in power grids. By leveraging an in-depth understanding of OT traffic, our proposed AI-based innovations are capable of identifying infinitesimal anomalies and detecting zero-day attacks. Furthermore, the thesis pioneers an innovative method for analyzing the spatio-temporal correlations of anomalies within cyber-physical power systems. These contributions empower power system operators to effectively mitigate the challenges posed by APTs, enhancing the overall resilience and security of modern power grids.

Future research built on this thesis should explore adaptive and real-time solutions for detecting and mitigating APTs in CPPS under increasingly complex and unpredictable conditions. One approach entails strengthening the proposed ACPSS kill chain model by integrating comprehensive insights into human factors, specifically the responses of power system operators, to improve situational awareness during cyber and physical anomalies. Future research should also take into account the dangers posed by insider threats, which are beyond the scope of this thesis. Furthermore, future research can conduct a more thorough investigation of CPPS, utilizing more power system measurements to capture the complexities of power system dynamics. These directions will contribute to enhancing the resilience and security of CPPS, equipping them to better withstand advanced cyber threats. Finally this thesis highlights the double-edged sword of technological advancement in power grids. While innovation and digitalization enhance operational efficiency, they also introduce new risks and challenges. Despite these vulnerabilities, advancing intelligent power grid technology is essential and must continue. Past cyber attacks on power grids should be viewed not merely as obstacles but as catalysts for system evolution, offering valuable lessons for resilience. As adversaries grow more sophisticated, operators must continually enhance their own capabilities, staying ahead with proactive strategies and adaptive defenses.

The relevance of this thesis is underscored by the evolution of cyber threats, which have extended far beyond data breaches in the digital realm to inflict tangible harm in the physical world. Critical infrastructures, particularly cyber-physical power systems, are increasingly vulnerable to cyber threats, which can lead to cascading failures, physical damages, and potentially threaten human life. These imminent threats have been exemplified by a series of cyber attacks on Ukrainian power grids in 2015, 2016, and 2022, where attackers successfully disrupted energy supplies and demonstrated the devastating impact of cyber attacks on critical infrastructure. These events underscore the growing urgency of enhancing cyber security in critical infrastructure, especially given the evolving geopolitical landscape where such attacks are leveraged as tools of modern warfare. In this context, the innovations



introduced in this thesis are not only timely but also essential. This thesis contributes to society by presenting early-stage attack detection for minimizing the impacts of cyber attacks on power grids. Ultimately, the findings of this thesis provide relevant solutions for protecting critical infrastructure, reinforcing its importance in addressing the pressing global issue of critical infrastructure security.



# BIBLIOGRAPHY

- [1] Alfán Presekal, Alexandru Ștefanov, Vetrivel Subramaniam Rajkumar, and Peter Palensky. Cyber attacks on power systems. In *Cyber-Physical Power Systems: Challenges and Solutions by AI/ML, Big Data, Blockchain, IoT, and Information Theory Paradigms*. IEEE-Wiley Press, USA, 2025.
- [2] Ioannis Semertzis, Alexandru Ștefanov, Alfán Presekal, Bas Kruimer, José Luis Rueda Torres, and Peter Palensky. Power system stability analysis from cyber attacks perspective. *IEEE Access*, 12:113008–113035, 2024.
- [3] Alfán Presekal, Alexandru Ștefanov, Vetrivel Subramaniam Rajkumar, and Peter Palensky. Anomaly detection and mitigation in cyber-physical power systems based on hybrid deep learning and attack graphs. In *Cyber-Physical Power Systems: Challenges and Solutions by AI/ML, Big Data, Blockchain, IoT, and Information Theory Paradigms*. IEEE-Wiley Press, USA, 2025.
- [4] Alfán Presekal, Alexandru Ștefanov, Vetrivel S Rajkumar, Ioannis Semertzis, and Peter Palensky. Advanced persistent threat kill chain for cyber-physical power systems. *IEEE Access*, 12:177746–177771, 2024.
- [5] Alfán Presekal, Alexandru Ștefanov, Vetrivel Subramaniam Rajkumar, and Peter Palensky. Attack graph model for cyber-physical power systems using hybrid deep learning. *IEEE Transactions on Smart Grid*, 14(5):4007–4020, 2023.
- [6] Alfán Presekal, Alexandru Ștefanov, Vetrivel Subramaniam Rajkumar, and Peter Palensky. Cyber forensic analysis for operational technology using graph-based deep learning. In *2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–7, 2023.
- [7] Alfán Presekal, Alexandru Ștefanov, Ioannis Semertzis, and Peter Palensky. Spatio-temporal advanced persistent threat detection and correlation for cyber-physical power systems using enhanced gc-lstm. *IEEE Transactions on Smart Grid*, 2024.
- [8] David E. Whitehead, Kevin Owens, Dennis Gammel, and Jess Smith. Ukraine cyber-induced power outage: Analysis and practical mitigation strategies. In *2017 70th Annual Conference for Protective Relay Engineers (CPRE)*, pages 1–8, 2017.
- [9] M. J. Assante, R. M. Lee, and T. Conway. Ics defense use case no. 6: Modular ics malware. Technical Report 6, Electricity Information Sharing and Analysis Center (E-ISAC), Washington, DC, USA, Aug. 2017.
- [10] ENTSO-E. Entso-e has recently found evidence of a successful cyber intrusion into its office network, Mar. 2020. Accessed: 2022-12-10.
- [11] Ken Proska, John Wolfram, Jared Wilson, Dan Black, Keith Lunden, Daniel Kapellmann Zafra, Nathan Brubaker, Tyler Mclellan, and Chris Sistrunk. Sandworm disrupts power in ukraine using a novel attack against operational technology. *Mandiant*: <https://www.mandiant.com/resources/blog/sandworm-disrupts-power-ukraine-operational-technology>, 2023.
- [12] Mutsuo Noguchi and Hirofumi Ueda. An analysis of the actual status of recent cyberattacks on critical infrastructures. *NEC Technical Journal, Special Issue Cybersecurity*, 12(2):19–24, 2019.
- [13] Doug Salmon, Mark Zeller, Armando Guzmán, Venkat Mynam, and Marcos Donolo. Mitigating the aurora vulnerability with existing technology. In *36th Annual western protection relay conference*, 2009.
- [14] Maj Gen PK Mallick. Chinese cyber exploitation in india’s power grid-is there a linkage to mumbai power outage? *Technical report, Strategic Study India,(India)*, 2021.

- [15] Sameer Patil. Assessing the efficacy of the west's autonomous cyber-sanctions regime and its relevance for india. In *Assessing the efficacy of the West's autonomous cyber-sanctions regime and its relevance for India: Patil, Sameer*. New Delhi, India: ORF, Observer Research Foundation, 2022.
- [16] Clémence Poirier. The war in ukraine from a space cybersecurity perspective. *ESPI Report*, 84, 2022.
- [17] Gaoqi Liang, Junhua Zhao, Fengji Luo, Steven R. Weller, and Zhao Yang Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 2017.
- [18] Ruilong Deng, Gaoxi Xiao, Rongxing Lu, Hao Liang, and Athanasios V. Vasilakos. False data injection on state estimation in power systems—attacks, impacts, and defense: A survey. *IEEE Transactions on Industrial Informatics*, 13(2):411–423, 2017.
- [19] Ahmed S. Musleh, Guo Chen, and Zhao Yang Dong. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 11(3):2218–2234, 2020.
- [20] Haftu Tasew Reda, Adnan Anwar, and Abdun Mahmood. Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts. *Renewable and Sustainable Energy Reviews*, 163:112423, 2022.
- [21] Ali Sayghe, Yaodan Hu, Ioannis Zografopoulos, XiaoRui Liu, Raj Gautam Dutta, Yier Jin, and Charalambos Konstantinou. Survey of machine learning methods for detecting false data injection attacks in power systems. *IET Smart Grid*, 3(5):581–595, 2020.
- [22] Hang Zhang, Bo Liu, and Hongyu Wu. Smart grid cyber-physical attack and defense: A review. *IEEE Access*, 9:29641–29659, 2021.
- [23] Usman Inayat, Muhammad Fahad Zia, Sajid Mahmood, Haris M Khalid, and Mohamed Benbouzid. Learning-based methods for cyber attacks detection in iot systems: A survey on methods, analysis, and future prospects. *Electronics*, 11(9):1502, 2022.
- [24] Ahmed S. Musleh, Haris M. Khalid, S. M. Muyeen, and Ahmed Al-Durra. A prediction algorithm to enhance grid resilience toward cyber attacks in wamcs applications. *IEEE Systems Journal*, 13(1):710–719, 2019.
- [25] Eric M Hutchins, Michael J Cloppert, Rohan M Amin, et al. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. *Leading Issues in Information Warfare & Security Research*, 1(1):80, 2011.
- [26] Adam Hahn, Roshan K Thomas, Ivan Lozano, and Alvaro Cardenas. A multi-layered and kill-chain based security analysis framework for cyber-physical systems. *International Journal of Critical Infrastructure Protection*, 11:39–50, 2015.
- [27] MITRE Corporation. Mitre attck for ics: Techniques, 2023. Accessed: 2023-12-12.
- [28] Michael J Assante and Robert M Lee. The industrial control system cyber kill chain. *SANS Institute InfoSec Reading Room*, 1(1):2, 2015.
- [29] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. Deep learning for time series classification: a review. *Data mining and knowledge discovery*, 33(4):917–963, 2019.
- [30] Hassan Ismail Fawaz, Benjamin Lucas, Germain Forestier, Charlotte Pelletier, Daniel F Schmidt, Jonathan Weber, Geoffrey I Webb, Lhassane Idoumghar, Pierre-Alain Muller, and François Petitjean. Inceptiontime: Finding alexnet for time series classification. *Data Mining and Knowledge Discovery*, 34(6):1936–1962, 2020.
- [31] Mauro Prais and Anjan Bose. A topology processor that tracks network modifications. *IEEE transactions on Power Systems*, 3(3):992–998, 1988.
- [32] Saioosh N Talukdar, Eleri Cardozo, and Ted Perry. The operator's assistant—an intelligent, expandable program for power system trouble analysis. *IEEE Transactions on power systems*, 1(3):182–187, 1986.

- [33] Panda Security. PandaLabs Annual Report 2017. Technical report, Panda Security, November 2017. Accessed: Jul. 13, 2023.
- [34] Nicolas Falliere, Liam O Murchu, Eric Chien, et al. W32. stuxnet dossier. *White paper, symantec corp., security response*, 5(6):29, 2011.
- [35] Defense Use Case. Analysis of the cyber attack on the ukrainian power grid. *Electricity information sharing and analysis center (E-ISAC)*, 388(1-29):3, 2016.
- [36] Robert M Lee, MJ Assante, and T Conway. Crashoverride: Analysis of the threat to electric grid operations. *Dragos Inc., March*, page 7, 2017.
- [37] Robert J Turk. Cyber incidents involving control systems. Technical report, Idaho National Lab.(INL), Idaho Falls, ID (United States), 2005.
- [38] Richard Derbyshire, Benjamin Green, Daniel Prince, Andreas Mauthe, and David Hutchison. An analysis of cyber security attack taxonomies. In *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 153–161. IEEE, 2018.
- [39] David Gewirtz. Night dragon: Cyberwar meets corporate espionage. *Journal of Counterterrorism & Homeland Security International*, 17(2), 2011.
- [40] Ralph Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE security & privacy*, 9(3):49–51, 2011.
- [41] Robert M Lee, Michael J Assante, and Tim Conway. German steel mill cyber attack. *Industrial Control Systems*, 30(62):1–15, 2014.
- [42] Anton Cherepanov and Robert Lipovsky. Blackenergy—what we really know about the notorious cyber attacks. *Virus Bulletin October*, 541, 2016.
- [43] Joe Slowik. Crashoverride: Reassessing the 2016 ukraine electric power event as a protection-focused attack. *Dragos, Inc*, 2019.
- [44] Alessandro Di Pinto, Younes Dragoni, and Andrea Carcano. Triton: The first ics cyber attack on safety instrument systems. *Proc. Black Hat USA*, 2018:1–26, 2018.
- [45] Paul Mueller and Babak Yadegari. The stuxnet worm. *Département des sciences de l’informatique, Université de l’Arizona. Recuperado de: <https://www2.cs.arizona.edu/~collberg/Teaching/466-566/2012/Resources/presentations/topic9-final/report.pdf>*, 2012.
- [46] Siddhant Shrivastava. Blackenergy-malware for cyber-physical attacks. *Singapore*, 74:115, 2016.
- [47] Anton Cherepanov. Win32/industroyer: A new threat for industrial control systems. *White paper, ESET (June 2017)*, 2017.
- [48] Blake Johnson, Dan Caban, Marina Krotofil, Dan Scali, Nathan Brubaker, and Christopher Glycer. Attackers deploy new ics attack framework “triton” and cause operational disruption to critical infrastructure. *Threat Research Blog*, 14:94, 2017.
- [49] Raphael Amoah, Seyit Camtepe, and Ernest Foo. Securing dnp3 broadcast communications in scada systems. *IEEE Transactions on Industrial Informatics*, 12(4):1474–1485, 2016.
- [50] Ibrahim Ali Ibrahim Diyeb, Anwar Saif, and Nagi Ali Al-Shaibany. Ethical network surveillance using packet sniffing tools: A comparative study. *International Journal of Computer Network and Information Security*, 11(7):12, 2018.
- [51] Craig Valli, Andrew Woodward, Clinton Carpena, Peter Hannay, Murray Brand, Reino Karvinen, and Christopher Holme. Eavesdropping on the smart grid. *SRI Security Research Institute*, 2012.
- [52] Huanhuan Yuan, Yuanqing Xia, Yuan Yuan, and Hongjiu Yang. Resilient strategy design for cyber-physical system under active eavesdropping attack. *Journal of the Franklin Institute*, 358(10):5281–5304, 2021.

- [53] Pardeep Kumar, Yun Lin, Guangdong Bai, Andrew Paverd, Jin Song Dong, and Andrew Martin. Smart grid metering networks: A survey on security, privacy and open research issues. *IEEE Communications Surveys & Tutorials*, 21(3):2886–2927, 2019.
- [54] Yawen Fan, Zhenghao Zhang, Matthew Trinkle, Aleksandar D Dimitrovski, Ju Bin Song, and Husheng Li. A cross-layer defense mechanism against gps spoofing attacks on pmus in smart grids. *IEEE Transactions on Smart Grid*, 6(6):2659–2668, 2014.
- [55] Ancheng Xue, Feiyang Xu, Jingsong Xu, Joe H Chow, Shuang Leng, and Tianshu Bi. Online pattern recognition and data correction of pmu data under gps spoofing attack. *Journal of Modern Power Systems and Clean Energy*, 8(6):1240–1249, 2020.
- [56] Paresh Risbud, Nikolaos Gatsis, and Ahmad Taha. Vulnerability analysis of smart grids to gps spoofing. *IEEE Transactions on Smart Grid*, 10(4):3535–3548, 2018.
- [57] Bernhard Sterzbach. Gps-based clock synchronization in a mobile, distributed real-time system. *Real-Time Systems*, 12(1):63–75, 1997.
- [58] Tianshu Bi, Jinrui Guo, Kai Xu, Li Zhang, and Qixun Yang. The impact of time synchronization deviation on the performance of synchrophasor measurements and wide area damping control. *IEEE Transactions on Smart Grid*, 8(4):1545–1552, 2016.
- [59] Fu Zhu, Amr Youssef, and Walaa Hamouda. Detection techniques for data-level spoofing in gps-based phasor measurement units. In *2016 international conference on selected topics in mobile & wireless networking (MoWNeT)*, pages 1–8. IEEE, 2016.
- [60] Maëlle Kabir-Querrec, Stéphane Mocanu, Jean-Marc Thiriet, and Eric Savary. A test bed dedicated to the study of vulnerabilities in iec 61850 power utility automation networks. In *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4. IEEE, 2016.
- [61] James G Wright and Stephen D Wolthusen. Stealthy injection attacks against iec61850’s goose messaging service. In *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2018.
- [62] Mohamad El Hariri, Tarek A Youssef, and Osama A Mohammed. On the implementation of the iec 61850 standard: Will different manufacturer devices behave similarly under identical conditions? *Electronics*, 5(4):85, 2016.
- [63] Tarek A Youssef, Mohamad El Hariri, Nicole Bugay, and OA Mohammed. Iec 61850: Technology standards and cyber-threats. In *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, pages 1–6. IEEE, 2016.
- [64] Nishchal Singh Kush, Ejaz Ahmed, Mark Branagan, and Ernest Foo. Poisoned goose: Exploiting the goose protocol. In *Proceedings of the Twelfth Australasian Information Security Conference (AISC 2014)[Conferences in Research and Practice in Information Technology, Volume 149]*, pages 17–22. Australian Computer Society, 2014.
- [65] Vetrivel Subramaniam Rajkumar, Marko Tealane, Alexandru Ștefanov, Alfian Presekal, and Peter Palensky. Cyber attacks on power system automation and protection and impact analysis. In *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 247–254. IEEE, 2020.
- [66] Junho Hong, Chen-Ching Liu, and Manimaran Govindarasu. Integrated anomaly detection for cyber security of the substations. *IEEE Transactions on Smart Grid*, 5(4):1643–1653, 2014.
- [67] Ahmed Elgargouri, Reino Virrankoski, and Mohammed Elmusrati. Iec 61850 based smart grid security. In *2015 IEEE International Conference on Industrial Technology (ICIT)*, pages 2461–2465. IEEE, 2015.
- [68] Yao Liu, Peng Ning, and Michael K Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, 14(1):1–33, 2011.

- [69] Kaikai Pan, André Teixeira, Milos Cvetkovic, and Peter Palensky. Cyber risk analysis of combined data attacks against power system state estimation. *IEEE Transactions on Smart Grid*, 10(3):3044–3056, 2018.
- [70] Liyan Jia, Jinsub Kim, Robert J Thomas, and Lang Tong. Impact of data quality on real-time locational marginal price. *IEEE Transactions on Power Systems*, 29(2):627–636, 2013.
- [71] Jingwen Liang, Lalitha Sankar, and Oliver Kosut. Vulnerability analysis and consequences of false data injection attack on power system state estimation. *IEEE Transactions on Power Systems*, 31(5):3864–3872, 2015.
- [72] Gaoqi Liang, Steven R Weller, Junhua Zhao, Fengji Luo, and Zhao Yang Dong. The 2015 ukraine blackout: Implications for false data injection attacks. *IEEE transactions on power systems*, 32(4):3317–3318, 2016.
- [73] Aditya Ashok, Pengyuan Wang, Matthew Brown, and Manimaran Govindarasu. Experimental evaluation of cyber attacks on automatic generation control using a cps security testbed. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2015.
- [74] Haftu Tasew Reda, Adnan Anwar, and Abdun Mahmood. Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts. *Renewable and Sustainable Energy Reviews*, 163:112423, 2022.
- [75] Siddharth Sridhar and Manimaran Govindarasu. Model-based attack detection and mitigation for automatic generation control. *IEEE Transactions on Smart Grid*, 5(2):580–591, 2014.
- [76] Kaikai Pan, Peter Palensky, and Peyman Mohajerin Esfahani. From static to dynamic anomaly detection with application to power system cyber security. *IEEE Transactions on Power Systems*, 35(2):1584–1596, 2019.
- [77] JQ James, Yunhe Hou, and Victor OK Li. Online false data injection attack detection with wavelet transform and deep neural networks. *IEEE Transactions on Industrial Informatics*, 14(7):3271–3280, 2018.
- [78] Durga Samanth Pidikiti, Rajesh Kalluri, RK Senthil Kumar, and BS Bindhumadhava. Scada communication protocols: vulnerabilities, attacks and possible mitigations. *CSI transactions on ICT*, 1(2):135–141, 2013.
- [79] Gagan Dua, Nitin Gautam, Dharmendar Sharma, and Ankit Arora. Replay attack prevention in kerberos authentication protocol using triple password. *arXiv preprint arXiv:1304.3550*, 2013.
- [80] Andreas Hoehn and Ping Zhang. Detection of replay attacks in cyber-physical systems. In *2016 American control conference (ACC)*, pages 290–295. IEEE, 2016.
- [81] Yilin Mo and Bruno Sinopoli. Secure control against replay attacks. In *2009 47th annual Allerton conference on communication, control, and computing (Allerton)*, pages 911–918. IEEE, 2009.
- [82] Amit Kleinmann, Ori Amichay, Avishai Wool, David Tenenbaum, Ofer Bar, and Leonid Lev. Stealthy deception attacks against scada systems. In *Computer Security: ESORICS 2017 International Workshops, CyberCPS 2017 and SECPRE 2017, Oslo, Norway, September 14-15, 2017, Revised Selected Papers 3*, pages 93–109. Springer, 2018.
- [83] Marco De Vivo, Gabriela O de Vivo, Roberto Koenek, and Germinal Isern. Internet vulnerabilities related to tcp/ip and t/tcp. *ACM SIGCOMM Computer Communication Review*, 29(1):81–85, 1999.
- [84] Willem Burgers, Roel Verdult, and Marko Van Eekelen. Prevent session hijacking by binding the session to the cryptographic network credentials. In *Secure IT Systems: 18th Nordic Conference, NordSec 2013, Ilulissat, Greenland, October 18-21, 2013, Proceedings 18*, pages 33–50. Springer, 2013.
- [85] Stephen M Specht and Ruby B Lee. Distributed denial of service: Taxonomies of attacks, tools, and countermeasures. In *PDCS*, pages 543–550, 2004.
- [86] Jasna D Markovic-Petrovic and Mirjana D Stojanovic. Analysis of scada system vulnerabilities to ddos attacks. In *2013 11th international conference on telecommunications in modern satellite, cable and broadcasting services (telsiks)*, volume 2, pages 591–594. IEEE, 2013.

- [87] Rajesh Kalluri, Lagineni Mahendra, RK Senthil Kumar, and GL Ganga Prasad. Simulation and impact analysis of denial-of-service attacks on power scada. In *2016 national power systems conference (NPSC)*, pages 1–5. IEEE, 2016.
- [88] Abdullah Albarakati, Chantale Robillard, Mark Karanfil, Marthe Kassouf, Mourad Debbabi, Amr Youssef, Mohsen Ghafouri, and Rachid Hadjidj. Security monitoring of iec 61850 substations using iec 62351-7 network and system management. *IEEE Transactions on Industrial Informatics*, 18(3):1641–1653, 2021.
- [89] An-Yang Lu and Guang-Hong Yang. Switched projected gradient descent algorithms for secure state estimation under sparse sensor attacks. *Automatica*, 103:503–514, 2019.
- [90] Sanjana Vijayshankar, Chin-Yao Chang, Kumar Utkarsh, Dylan Wald, Fei Ding, Sivasathya Pradha Balamurugan, Jennifer King, and Richard Macwan. Assessing the impact of cybersecurity attacks on energy systems. *Applied Energy*, 345:121297, 2023.
- [91] Kaikai Pan, Jingwei Dong, Elyas Rakhshani, and Peter Palensky. Effects of cyber attacks on ac and high-voltage dc interconnected power systems with emulated inertia. *Energies*, 13(21):5583, 2020.
- [92] Luca Schenato. To zero or to hold control inputs with lossy links? *IEEE Transactions on Automatic Control*, 54(5):1093–1099, 2009.
- [93] Dinesha Ranathunga, Matthew Roughan, Hung Nguyen, Phil Kernick, and Nickolas Falkner. Case studies of scada firewall configurations and the implications for best practices. *IEEE Transactions on Network and Service Management*, 13(4):871–884, 2016.
- [94] Darshana Upadhyay and Srinivas Sampalli. Scada (supervisory control and data acquisition) systems: Vulnerability assessment and security recommendations. *Computers & Security*, 89:101666, 2020.
- [95] Department of Homeland Security Office of Cybersecurity and Communication. NCCIC/ICS-CERT FY 2015 Annual Vulnerability Coordination Report. Technical report, Department of Homeland Security, 2015. Accessed: Jul. 15, 2023.
- [96] Moshe Kol and Shlomi Oberman. Cve-2020-11896 rce and cve-2020-11898 info leak. *JSOF Inc. White Paper*, pages 1–27, 2020.
- [97] M. Kol, A. Schon, and S. Oberman. CVE-2020-11901. White paper, JSOF Inc., Aug 2020. Accessed: Jul. 5, 2023.
- [98] Bonnie Zhu, Anthony Joseph, and Shankar Sastry. A taxonomy of cyber attacks on scada systems. In *2011 International conference on internet of things and 4th international conference on cyber, physical and social computing*, pages 380–388. IEEE, 2011.
- [99] Curtis R Taylor, Craig A Shue, and Nathanael R Paul. A deployable scada authentication technique for modern power grids. In *2014 IEEE International Energy Conference (ENERGYCON)*, pages 696–702. IEEE, 2014.
- [100] Binod Vaidya, Dimitrios Makrakis, and Hussein T Mouftah. Authentication and authorization mechanisms for substation automation in smart grid network. *IEEE Network*, 27(1):5–11, 2013.
- [101] Ren Liu, Ceeman Vellaithurai, Saugata S Biswas, Thoshitha T Gamage, and Anurag K Srivastava. Analyzing the cyber-physical impact of cyber events on the power grid. *IEEE Transactions on Smart Grid*, 6(5):2444–2453, 2015.
- [102] Gururaghav Raman, Bedoor AlShebli, Marcin Waniek, Talal Rahwan, and Jimmy Chih-Hsien Peng. How weaponizing disinformation can bring down a city’s power grid. *PloS one*, 15(8):e0236517, 2020.
- [103] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-art. *International Journal of Electrical Power & Energy Systems*, 99:45–56, 2018.
- [104] Mutsuo Noguchi and Hirofumi Ueda. An analysis of the actual status of recent cyberattacks on critical infrastructures. *NEC Technical Journal, Special Issue Cybersecurity*, 12(2):19–24, 2019.



- [105] Doug Salmon, Mark Zeller, Armando Guzmán, Venkat Mynam, and Marcos Donolo. Mitigating the aurora vulnerability with existing technology. In *36th Annual western protection relay conference*, 2009.
- [106] Göran Andersson, Peter Donalek, Richard Farmer, Nikos Hatziaargyriou, Innocent Kamwa, Prabhashankar Kundur, Nelson Martins, John Paserba, Pouyan Pourbeik, Juan Sanchez-Gasca, et al. Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance. *IEEE transactions on Power Systems*, 20(4):1922–1928, 2005.
- [107] Pouyan Pourbeik, Prabha S Kundur, and Carson W Taylor. The anatomy of a power grid blackout-root causes and dynamics of recent major blackouts. *IEEE Power and Energy Magazine*, 4(5):22–29, 2006.
- [108] Chee-Wooi Ten, Koji Yamashita, Zhiyuan Yang, Athanasios V Vasilakos, and Andrew Ginter. Impact assessment of hypothesized cyberattacks on interconnected bulk power systems. *IEEE Transactions on Smart Grid*, 9(5):4405–4425, 2017.
- [109] Shan Liu, Bo Chen, Takis Zourntos, Deepa Kundur, and Karen Butler-Purry. A coordinated multi-switch attack for cascading failures in smart grid. *IEEE Transactions on Smart Grid*, 5(3):1183–1195, 2014.
- [110] Bo Chen, Salman Mashayekh, Karen L Butler-Purry, and Deepa Kundur. Impact of cyber attacks on transient stability of smart grids with voltage support devices. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.
- [111] Bo Chen, Karen L Butler-Purry, Sruti Nuthalapati, and Deepa Kundur. Network delay caused by cyber attacks on svc and its impact on transient stability of smart grids. In *2014 IEEE PES General Meeting/Conference & Exposition*, pages 1–5. IEEE, 2014.
- [112] Anya Castillo, Bryan Arguello, Gerardo Cruz, and Laura Swiler. Cyber-physical emulation and optimization of worst-case cyber attacks on the power grid. In *2019 Resilience Week (RWS)*, volume 1, pages 14–18. IEEE, 2019.
- [113] ICS Dragos. Ot cybersecurity year in review 2022, 2023.
- [114] IEC TC57. Iec 61850: Communication networks and systems for power utility automation. *International Electrotechnical Commission Std*, 53:54, 2010.
- [115] International Electrotechnical Commission et al. Telecontrol equipment and systems-part 5-104: Transmission protocols-network access for iec 60870-5-101 using standard transport profiles. *International Electrotechnical Commission: Geneva, Switzerland*, 2006.
- [116] George Thomas. Introduction to the modbus protocol. *The Extension*, 9(4):1–4, 2008.
- [117] Roberto Nardone, Ricardo J Rodríguez, and Stefano Marrone. Formal security assessment of modbus protocol. In *2016 11th International conference for internet technology and secured transactions (ICITST)*, pages 142–147. IEEE, 2016.
- [118] Marco De Vivo, Gabriela O de Vivo, Roberto Koenke, and Germinal Isern. Internet vulnerabilities related to tcp/ip and t/tcp. *ACM SIGCOMM Computer Communication Review*, 29(1):81–85, 1999.
- [119] Samuel East, Jonathan Butts, Mauricio Papa, and Sujeet Sheno. A taxonomy of attacks on the dnp3 protocol. In *Critical Infrastructure Protection III: Third Annual IFIP WG 11.10 International Conference on Critical Infrastructure Protection, Hanover, New Hampshire, USA, March 23-25, 2009, Revised Selected Papers 3*, pages 67–81. Springer, 2009.
- [120] Christoph Brunner. Iec 61850 for power system communication. In *2008 IEEE/PES Transmission and Distribution Conference and Exposition*, pages 1–6. IEEE, 2008.
- [121] Suhail SM Hussain, Chen Yaohao, Muhammad M Roomi, Daisuke Mashima, and Ee-Chien Chang. An open-source framework for publishing/subscribing iec 61850 r-goose and r-sv. *SoftwareX*, 23:101415, 2023.

- [122] Vetrivel Subramaniam Rajkumar, Marko Tealane, Alexandru Ștefanov, and Peter Palensky. Cyber attacks on protective relays in digital substations and impact analysis. In *2020 8th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems*, pages 1–6. IEEE, 2020.
- [123] SM Suhail Hussain, Taha Selim Ustun, and Akhtar Kalam. A review of iec 62351 security mechanisms for iec 61850 message exchanges. *IEEE Transactions on Industrial Informatics*, 16(9):5643–5654, 2019.
- [124] Yi Yang, Kieran McLaughlin, Tim Littler, Sakir Sezer, Bernardi Pranggono, and HF Wang. Intrusion detection system for iec 60870-5-104 based scada networks. In *2013 IEEE power & energy society general meeting*, pages 1–5. Ieee, 2013.
- [125] Mathias Uslar, Michael Specht, Christian Dänekas, Jörn Trefke, Sebastian Rohjans, José M González, Christine Rosinger, Robert Bleiker, Christine Rosinger, and Mathias Uslar. Smart grid security: Iec 62351 and other relevant standards. *Standardization in Smart Grids: Introduction to IT-Related Methodologies, Architectures and Standards*, pages 129–146, 2013.
- [126] Maximilian Strobel, Norbert Wiedermann, and Claudia Eckert. Novel weaknesses in iec 62351 protected smart grid control systems. In *2016 IEEE International Conference on Smart Grid Communications (Smart-GridComm)*, pages 266–270. IEEE, 2016.
- [127] Andrea Carcano, Alessandro Di Pinto, and Younes Dragoni. The future of securing intelligent electronic devices using the iec 62351-7 standard for monitoring. Presented at Black Hat USA 2019, August 2019. Available at: <https://i.blackhat.com/USA-19/Thursday/us-19-Carcano-The-Future-Of-Securing-IED-Using-The-IEC62351-7-Standard-For-Monitoring.pdf>.
- [128] KE Martin, Gustavo Brunello, MG Adamiak, Galina Antonova, M Begovic, G Benmouyal, PD Bui, Heiko Falk, V Gharpure, A Goldstein, et al. An overview of the iec standard c37. 118.2—synchrophasor data transfer for power systems. *IEEE Transactions on Smart Grid*, 5(4):1980–1984, 2014.
- [129] Saghar Vahidi, Mohsen Ghafouri, Minh Au, Marthe Kassouf, Arash Mohammadi, and Mourad Debbabi. Security of wide-area monitoring, protection, and control (wampac) systems of the smart grid: A survey on challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 25(2):1294–1335, 2023.
- [130] Rafiullah Khan, Kieran McLaughlin, David Lavery, and Sakir Sezer. Analysis of iec c37. 118 and iec 61850-90-5 synchrophasor communication frameworks. In *2016 IEEE power and energy society general meeting (PESGM)*, pages 1–5. IEEE, 2016.
- [131] Ieee standard for phasor data concentrators for power systems. *IEEE Std C37.247-2019*, pages 1–44, 2019.
- [132] Kamakshi Prashadini Swain, Amit Tiwari, Ankush Sharma, Saikat Chakrabarti, and Amey Karkare. Comprehensive demonstration of man-in-the-middle attack in pdc and pmu network. In *2022 22nd National Power Systems Conference (NPSC)*, pages 213–217. IEEE, 2022.
- [133] Astha Chawla, Animesh Singh, Prakhar Agrawal, Bijaya Ketan Panigrahi, Bhavesh R Bhalja, and Kolin Paul. Denial-of-service attacks pre-emptive and detection framework for synchrophasor based wide area protection applications. *IEEE Systems Journal*, 16(1):1570–1581, 2021.
- [134] JT Robinson, T Saxton, A Vojdani, D Ambrose, G Schimmel, RR Blaesing, and R Larson. Development of the intercontrol center communications protocol (iccp)[power system control]. In *Proceedings of Power Industry Computer Applications Conference*, pages 449–455. IEEE, 1995.
- [135] Matthew Franz. Iccp exposed: assessing the attack surface of the utility stack. In *Proceedings of SCADA Security Scientific Symposium, Florida, USA*, 2007.
- [136] John T Michalski, Andrew Lanzone, Jason Trent, and Sammy Smith. Secure iccp integration considerations and recommendations. *SANDIA report*, 2007.
- [137] Adam Hahn. Operational technology and information technology in industrial control systems. *Cyber-security of SCADA and other industrial control systems*, pages 51–68, 2016.

- [138] Dinesha Ranathunga, Matthew Roughan, Hung Nguyen, Phil Kernick, and Nickolas Falkner. Case studies of scada firewall configurations and the implications for best practices. *IEEE Transactions on Network and Service Management*, 13(4):871–884, 2016.
- [139] Darshana Upadhyay and Srinivas Sampalli. Scada (supervisory control and data acquisition) systems: Vulnerability assessment and security recommendations. *Computers & Security*, 89:101666, 2020.
- [140] MITRE Corporation. Cve - scada vulnerabilities, 2024. Accessed: 2024-12-28.
- [141] Danielle Gonzalez, Fawaz Alhenaki, and Mehdi Mirakhorli. Architectural security weaknesses in industrial control systems (ics) an empirical study based on disclosed software vulnerabilities. In *2019 IEEE International Conference on Software Architecture (ICSA)*, pages 31–40. IEEE, 2019.
- [142] Geeta Yadav and Kolin Paul. Patchrank: Ordering updates for scada systems. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 110–117. IEEE, 2019.
- [143] Moshe Kol and Shlomi Oberman. Cve-2020-11896 rce and cve-2020-11898 info leak. *JSOF Inc. White Paper*, pages 1–27, 2020.
- [144] JD Tygar. Adversarial machine learning. *IEEE Internet Computing*, 15(5):4–6, 2011.
- [145] Huan Ying, Xuan Ouyang, Siwei Miao, and Yushi Cheng. Power message generation in smart grid via generative adversarial network. In *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 790–793. IEEE, 2019.
- [146] European Union Agency for Cybersecurity (ENISA). Ics scada dependencies, 2023. Accessed: 2023-08-28.
- [147] Raphael Amoah, Seyit Camtepe, and Ernest Foo. Formal modelling and analysis of dnp3 secure authentication. *Journal of Network and Computer Applications*, 59:345–360, 2016.
- [148] Eric D Knapp. *Industrial Network Security: Securing critical infrastructure networks for smart grid, SCADA, and other Industrial Control Systems*. Elsevier, 2024.
- [149] Matheus K Ferst, Hugo FM de Figueiredo, Gustavo Denardin, and Juliano Lopes. Implementation of secure communication with modbus and transport layer security protocols. In *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*, pages 155–162. IEEE, 2018.
- [150] Abdalhossein Rezai, Parviz Keshavarzi, and Zahra Moravej. Key management issue in scada networks: A review. *Engineering science and technology, an international journal*, 20(1):354–363, 2017.
- [151] TC Pramod, Kianoosh G Boroojeni, M Hadi Amini, NR Sunitha, and SS Iyengar. Key pre-distribution scheme with join leave support for scada systems. *International Journal of Critical Infrastructure Protection*, 24:111–125, 2019.
- [152] Zibin Zheng, Shaoan Xie, Hong-Ning Dai, Xiangping Chen, and Huaimin Wang. Blockchain challenges and opportunities: A survey. *International journal of web and grid services*, 14(4):352–375, 2018.
- [153] Tejasvi Alladi, Vinay Chamola, Joel JPC Rodrigues, and Sergei A Kozlov. Blockchain in smart grids: A review on different use cases. *Sensors*, 19(22):4862, 2019.
- [154] Ricardo Brandão. A blockchain-based protocol for message exchange in a ics network: student research abstract. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 357–360, 2020.
- [155] Alessandro Di Pinto, Younes Dragoni, and Andrea Carcano. Triton: The first ics cyber attack on safety instrument systems. *Proc. Black Hat USA*, 2018:1–26, 2018.
- [156] Jeyasingam Nivethan and Mauricio Papa. A linux-based firewall for the dnp3 protocol. In *2016 IEEE symposium on technologies for homeland security (HST)*, pages 1–5. IEEE, 2016.
- [157] Jeyasingam Nivethan and Mauricio Papa. On the use of open-source firewalls in ics/scada systems. *Information Security Journal: A Global Perspective*, 25(1-3):83–93, 2016.

- [158] Dong Li, Huaqun Guo, Jianying Zhou, Luying Zhou, and Jun Wen Wong. Scadawall: A cpi-enabled firewall model for scada security. *Computers & Security*, 80:134–154, 2019.
- [159] Justyna Chromik, Anne Remke, Boudewijn R Haverkort, and Gerard Geist. A parser for deep packet inspection of iec-104: A practical solution for industrial applications. In *2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks–Industry Track*, pages 5–8. IEEE, 2019.
- [160] Robin Sommer, Johanna Amann, and Seth Hall. Spicy: a unified deep packet inspection framework for safely dissecting all your data. In *Proceedings of the 32nd Annual Conference on Computer Security Applications*, pages 558–569, 2016.
- [161] Tiago Cruz, Luis Rosa, Jorge Proença, Leandros Maglaras, Matthieu Aubigny, Leonid Lev, Jianmin Jiang, and Paulo Simões. A cybersecurity detection framework for supervisory control and data acquisition systems. *IEEE Transactions on Industrial Informatics*, 12(6):2236–2246, 2016.
- [162] Shengyi Pan, Thomas H Morris, and Uttam Adhikari. A specification-based intrusion detection framework for cyber-physical environment in electric power system. *Int. J. Netw. Secur.*, 17(2):174–188, 2015.
- [163] Peter Maynard and Kieran McLaughlin. Towards understanding man-on-the-side attacks (mots) in scada networks. *arXiv preprint arXiv:2004.14334*, 2020.
- [164] Yi Yang, Keiran McLaughlin, Tim Littler, Sakir Sezer, and HF Wang. Rule-based intrusion detection system for scada networks. *IET Renewable Power Generation*, 2013.
- [165] Georgia Koutsandria, Vishak Muthukumar, Masood Parvania, Sean Peisert, Chuck McParland, and Anna Scaglione. A hybrid network ids for protective digital relays in the power transmission grid. In *2014 IEEE international conference on smart grid communications (SmartGridComm)*, pages 908–913. IEEE, 2014.
- [166] Niv Goldenberg and Avishai Wool. Accurate modeling of modbus/tcp for intrusion detection in scada systems. *international journal of critical infrastructure protection*, 6(2):63–75, 2013.
- [167] Abdulmohsen Almalawi, Adil Fahad, Zahir Tari, Abdullah Alamri, Rayed AlGhamdi, and Albert Y Zomaya. An efficient data-driven clustering technique to detect attacks in scada systems. *IEEE Transactions on Information Forensics and Security*, 11(5):893–906, 2015.
- [168] Hui Lin, Adam Slagell, Zbigniew Kalbarczyk, Peter W Sauer, and Ravishankar K Iyer. Semantic security analysis of scada networks to detect malicious control commands in power grids. In *Proceedings of the first ACM workshop on Smart energy grid security*, pages 29–34, 2013.
- [169] Amit Kleinmann and Avishai Wool. Automatic construction of statechart-based anomaly detection models for multi-threaded industrial control systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(4):1–21, 2017.
- [170] Oualid Koucham, Stéphane Mocanu, Guillaume Hiet, Jean-Marc Thiriet, and Frédéric Majorczyk. Efficient mining of temporal safety properties for intrusion detection in industrial control systems. *IFAC-PapersOnLine*, 51(24):1043–1050, 2018.
- [171] Hassan Lahza, Kenneth Radke, and Ernest Foo. Applying domain-specific knowledge to construct features for detecting distributed denial-of-service attacks on the goose and mms protocols. *International Journal of Critical Infrastructure Protection*, 20:48–67, 2018.
- [172] Yi Yang, Hai-Qing Xu, Lei Gao, Yu-Bo Yuan, Kieran McLaughlin, and Sakir Sezer. Multidimensional intrusion detection system for iec 61850-based scada networks. *IEEE Transactions on Power Delivery*, 32(2):1068–1078, 2016.
- [173] Hyunguk Yoo and Taeshik Shon. Novel approach for detecting network anomalies for substation automation based on iec 61850. *Multimedia Tools and Applications*, 74:303–318, 2015.
- [174] Shengyi Pan, Thomas Morris, and Uttam Adhikari. Developing a hybrid intrusion detection system using data mining for power systems. *IEEE Transactions on Smart Grid*, 6(6):3104–3113, 2015.

- [175] Lenhard Reuter, Oliver Jung, and Julian Magin. Neural network based anomaly detection for scada systems. In *2020 23rd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, pages 194–201. IEEE, 2020.
- [176] PS Chaithanya, S Priyanga, S Pravinraj, and VS Shankar Sriram. Sso-if: an outlier detection approach for intrusion detection in scada systems. In *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2019*, pages 921–929. Springer, 2020.
- [177] Leandros Maglaras, Tiago Cruz, Mohamed A Ferrag, and Helge Janicke. Teaching the process of building an intrusion detection system using data from a small-scale scada testbed. *Internet Technology Letters*, 3(1):e132, 2020.
- [178] Meir Kalech. Cyber-attack detection in scada systems using temporal pattern recognition techniques. *Computers & Security*, 84:225–238, 2019.
- [179] Shitharth Selvarajan, Masood Shaik, Sirajudeen Ameerjohn, and Sangeetha Kannan. Mining of intrusion attack in scada network using clustering and genetically seeded flora-based optimal classification algorithm. *IET Information Security*, 14(1):1–11, 2020.
- [180] Abdelouahid Derhab, Mohamed Guerroumi, Abdu Gumaie, Leandros Maglaras, Mohamed Amine Ferrag, Mithun Mukherjee, and Farrukh Aslam Khan. Blockchain and random subspace learning-based ids for sdn-enabled industrial iot security. *Sensors*, 19(14):3119, 2019.
- [181] Sara Tamy, Hicham Belhadaoui, Mahmoud Almostafa Rabbah, Nabila Rabbah, and Mounir Rifi. An evaluation of machine learning algorithms to detect attacks in scada network. In *2019 7th Mediterranean Congress of Telecommunications (CMT)*, pages 1–5. IEEE, 2019.
- [182] Jakapan Suaboot, Adil Fahad, Zahir Tari, John Grundy, Abdun Naser Mahmood, Abdulmohsen Almalawi, Albert Y Zomaya, and Khalil Drira. A taxonomy of supervised learning for idss in scada environments. *ACM Computing Surveys (CSUR)*, 53(2):1–37, 2020.
- [183] Majed Al-Asiri and El-Sayed M El-Alfy. On using physical based intrusion detection in scada systems. *Procedia Computer Science*, 170:34–42, 2020.
- [184] Izhar Ahmed Khan, Dechang Pi, Zaheer Ullah Khan, Yasir Hussain, and Asif Nawaz. Hml-ids: A hybrid-multilevel anomaly prediction approach for intrusion detection in scada systems. *IEEE Access*, 7:89507–89521, 2019.
- [185] Shamsul Huda, John Yearwood, Mohammad Mehedi Hassan, and Ahmad Almogren. Securing the operations in scada-iot platform based industrial control system using ensemble of deep belief networks. *Applied soft computing*, 71:66–77, 2018.
- [186] N Neha, S Priyanga, Suresh Seshan, R Senthilnathan, and VS Shankar Sriram. Sco-rnn: A behavioral-based intrusion detection approach for cyber physical attacks in scada systems. In *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2019*, pages 911–919. Springer, 2020.
- [187] Giulio Zizzo, Chris Hankin, Sergio Maffei, and Kevin Jones. Intrusion detection for industrial control systems: Evaluation analysis and adversarial attacks. *CoRR*, 2019.
- [188] Jun Gao, Luyun Gan, Fabiola Buschendorf, Liao Zhang, Hua Liu, Peixue Li, Xiaodai Dong, and Tao Lu. Lstm for scada intrusion detection. In *2019 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, pages 1–5. IEEE, 2019.
- [189] Jun Gao, Luyun Gan, Fabiola Buschendorf, Liao Zhang, Hua Liu, Peixue Li, Xiaodai Dong, and Tao Lu. Omni scada intrusion detection using deep learning algorithms. *IEEE Internet of Things Journal*, 8(2):951–961, 2020.
- [190] Mostofa Ahsan and Kendall E Nygard. Convolutional neural networks with lstm for intrusion detection. In *CATA*, volume 69, pages 69–79, 2020.

- [191] Tae-Young Kim and Sung-Bae Cho. Cnn-lstm neural networks for anomalous database intrusion detection in rbac-administered model. In *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part IV 26*, pages 131–139. Springer, 2019.
- [192] K Praanna, S Sruthi, K Kalyani, and A Sai Tejaswi. A cnn-lstm model for intrusion detection system from high dimensional data. *J. Inf. Comput. Sci*, 10(3):1362–1370, 2020.
- [193] Julian L Rrushi, Roy H Campbell, and U di Milano. Detecting attacks in power plant interfacing substations through probabilistic validation of attack-effect bindings. In *SCADA Security Scientific Symposium*. Citeseer, 2008.
- [194] Huan Yang, Liang Cheng, and Mooi Choo Chuah. Deep-learning-based network intrusion detection for scada systems. In *2019 IEEE Conference on Communications and Network Security (CNS)*, pages 1–7. IEEE, 2019.
- [195] Mustafa Altaha, Jae-Myeong Lee, Muhammad Aslam, and Sugwon Hong. Network intrusion detection based on deep neural networks for the scada system. In *Journal of Physics: Conference Series*, volume 1585, page 012038. IOP Publishing, 2020.
- [196] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.
- [197] Noam Erez and Avishai Wool. Control variable classification, modeling and anomaly detection in modbus/tcp scada systems. *International Journal of Critical Infrastructure Protection*, 10:59–70, 2015.
- [198] Chih-Yuan Lin and Simin Nadjm-Tehrani. Timing patterns and correlations in spontaneous {SCADA} traffic for anomaly detection. In *22nd international symposium on research in attacks, intrusions and defenses (RAID 2019)*, pages 73–88, 2019.
- [199] Mehmet Hazar Cintuglu, Osama A Mohammed, Kemal Akkaya, and A Selcuk Uluagac. A survey on smart grid cyber-physical system testbeds. *IEEE Communications Surveys & Tutorials*, 19(1):446–464, 2016.
- [200] Chih-Che Sun, Adam Hahn, and Chen-Ching Liu. Cyber security of a power grid: State-of-the-art. *International Journal of Electrical Power & Energy Systems*, 99:45–56, 2018.
- [201] Brij B Gupta and Tafseer Akhtar. A survey on smart power grid: frameworks, tools, security issues, and solutions. *Annals of Telecommunications*, 72:517–549, 2017.
- [202] Juan Montoya, Ron Brandl, Keerthi Vishwanath, Jay Johnson, Rachid Darbali-Zamora, Adam Summers, Jun Hashimoto, Hiroshi Kikusato, Taha Selim Ustun, Nayeem Ninad, et al. Advanced laboratory testing methods using real-time simulation and hardware-in-the-loop techniques: A survey of smart grid international research facility network activities. *Energies*, 13(12):3267, 2020.
- [203] Malaz Mallouhi, Youssif Al-Nashif, Don Cox, Tejaswini Chadaga, and Salim Hariri. A testbed for analyzing security of scada control systems (tasscs). In *ISGT 2011*, pages 1–7. IEEE, 2011.
- [204] Carlos Queiroz, Abdun Mahmood, and Zahir Tari. Scadasim—a framework for building scada simulations. *IEEE Transactions on Smart Grid*, 2(4):589–597, 2011.
- [205] Partha S Sarker, V Venkataramanan, D Sebastian Cardenas, A Srivastava, A Hahn, and Brian Miller. Cyber-physical security and resiliency analysis testbed for critical microgrids with ieee 2030.5. In *2020 8th workshop on modeling and simulation of cyber-physical energy systems*, pages 1–6. IEEE, 2020.
- [206] Jelena Mirkovic and Terry Benzel. Teaching cybersecurity with deterlab. *IEEE Security & Privacy*, 10(1):73–76, 2012.
- [207] Jelena Mirkovic, Terry V Benzel, Ted Faber, Robert Braden, John T Wroclawski, and Stephen Schwab. The deter project: Advancing the science of cyber security experimentation and test. In *2010 IEEE International Conference on Technologies for Homeland Security (HST)*, pages 1–7. IEEE, 2010.

- [208] Ibukun A Oyewumi, Ananth A Jillepalli, Philip Richardson, Mohammad Ashrafuzzaman, Brian K Johnson, Yacine Chakhchoukh, Michael A Haney, Frederick T Sheldon, and Daniel Conte de Leon. Isaac: The idaho cps smart grid cybersecurity testbed. In *2019 IEEE Texas Power and Energy Conference (TPEC)*, pages 1–6. IEEE, 2019.
- [209] Jay Johnson, Ifeoma Onunkwo, Patricia Cordeiro, Brian J Wright, Nicholas Jacobs, and Christine Lai. Assessing der network cybersecurity defences in a power-communication co-simulation environment. *IET Cyber-Physical Systems: Theory & Applications*, 5(3):274–282, 2020.
- [210] Malaz Mallouhi, Youssif Al-Nashif, Don Cox, Tejaswini Chadaga, and Salim Hariri. A testbed for analyzing security of scada control systems (tasscs). In *ISGT 2011*, pages 1–7. IEEE, 2011.
- [211] Bo Chen, Karen L Butler-Purpy, Ana Goulart, and Deepa Kundur. Implementing a real-time cyber-physical system test bed in rtds and opnet. In *2014 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2014.
- [212] Bo Chen, Nishant Pattanaik, Ana Goulart, Karen L Butler-Purpy, and Deepa Kundur. Implementing attacks for modbus/tcp protocol in a real-time cyber physical system test bed. In *2015 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)*, pages 1–6. IEEE, 2015.
- [213] Carlos Queiroz, Abdun Mahmood, and Zahir Tari. Scadasim—a framework for building scada simulations. *IEEE Transactions on Smart Grid*, 2(4):589–597, 2011.
- [214] Ammar Allaoua, Toufik Madani Layadi, Ilhami Colak, and Khaled Rouabah. Design and simulation of smart-grids using omnet++/matlab-simulink co-simulator. In *2021 10th International Conference on Renewable Energy Research and Application (ICRERA)*, pages 141–145. IEEE, 2021.
- [215] Ceeman B Vellaithurai, Saugata S Biswas, and Anurag K Srivastava. Development and application of a real-time test bed for cyber-physical system. *IEEE Systems Journal*, 11(4):2192–2203, 2015.
- [216] Junzo Watada, Arunava Roy, Raturaj Kadikar, Hoang Pham, and Bing Xu. Emerging trends, techniques and open issues of containerization: A review. *IEEE Access*, 7:152443–152472, 2019.
- [217] Muhammad Mudassar Yamin, Basel Katt, and Vasileios Gkioulos. Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. *Computers & Security*, 88:101636, 2020.
- [218] Yuri Diogenes and Erdal Ozkaya. *Cybersecurity-attack and defense strategies: Infrastructure security with red team and blue team tactics*. Packt Publishing Ltd, 2018.
- [219] Defense Use Case. Analysis of the cyber attack on the ukrainian power grid. *Electricity information sharing and analysis center (E-ISAC)*, 388(1-29):3, 2016.
- [220] Rajaa Vikhram Yohanandhan, Rajvikram Madurai Elavarasan, Premkumar Manoharan, and Lucian Mihet-Popa. Cyber-physical power system (cpps): A review on modeling, simulation, and analysis with cyber security applications. *IEEE Access*, 8:151019–151064, 2020.
- [221] Sangjun Kim, Kyung-Joon Park, and Chenyang Lu. A survey on network security for cyber-physical systems: From threats to resilient design. *IEEE Communications Surveys & Tutorials*, 24(3):1534–1573, 2022.
- [222] Adel Alshamrani, Sowmya Myneni, Ankur Chowdhary, and Dijiang Huang. A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 21(2):1851–1877, 2019.
- [223] Vivek Kumar Singh and Manimaran Govindarasu. Cyber kill chain-based hybrid intrusion detection system for smart grid. *Wide Area Power Systems Stability, Protection, and Security*, pages 571–599, 2021.
- [224] BoHyun Ahn, Taesic Kim, Jinchun Choi, Sung-won Park, Kuchan Park, and Dongjun Won. A cyber kill chain model for distributed energy resources (der) aggregation systems. In *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2021.
- [225] Amulya Sreejith and K Shanti Swarup. Mitre att&ck for smart grid cyber-security. In *Cyber-Security for Smart Grid Control: Vulnerability Assessment, Attack Detection, and Mitigation*, pages 59–73. Springer, 2024.



- [226] Neeraj Kumar Singh and Vasundhara Mahajan. Analysis and evaluation of cyber-attack impact on critical power system infrastructure. *Smart Science*, 9(1):1–13, 2021.
- [227] Atif Ahmad, Jeb Webb, Kevin C Desouza, and James Boorman. Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack. *Computers & Security*, 86:402–418, 2019.
- [228] Branka Stojanović, Katharina Hofer-Schmitz, and Ulrike Kleb. Apt datasets and attack modeling for automated detection methods: A review. *Computers & Security*, 92:101734, 2020.
- [229] Giuseppe Laurenza, Riccardo Lazzeretti, and Luca Mazzotti. Malware triage for early identification of advanced persistent threat activities. *Digital Threats: Research and Practice*, 1(3):1–17, 2020.
- [230] Mandiant Intelligence Center. Apt1: Exposing one of china’s cyber espionage units. *Mandiant.com*, 2013.
- [231] Eric Cole. *Advanced persistent threat: understanding the danger and how to protect your organization*. Newnes, 2012.
- [232] Ping Chen, Lieven Desmet, and Christophe Huygens. A study on advanced persistent threats. In *Communications and Multimedia Security: 15th IFIP TC 6/TC 11 International Conference, CMS 2014, Aveiro, Portugal, September 25-26, 2014. Proceedings 15*, pages 63–72. Springer, 2014.
- [233] Richard Kissel. *Glossary of key information security terms*. Diane Publishing, 2011.
- [234] Jathan Sadowski. When data is capital: Datafication, accumulation, and extraction. *Big data & society*, 6(1):2053951718820549, 2019.
- [235] Wendy Arianne Günther, Mohammad H Rezazade Mehrizi, Marleen Huysman, and Frans Feldberg. Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3):191–209, 2017.
- [236] James A Lewis. Computer espionage, titan rain and china. *Center for Strategic and International Studies-Technology and Public Policy Program*, 1, 2005.
- [237] Olivier Thonnard, Leyla Bilge, Gavin O’Gorman, Seán Kiernan, and Martin Lee. Industrial espionage and targeted attacks: Understanding the characteristics of an escalating threat. In *Research in Attacks, Intrusions, and Defenses: 15th International Symposium, RAID 2012, Amsterdam, The Netherlands, September 12-14, 2012. Proceedings 15*, pages 64–85. Springer, 2012.
- [238] Christian Czosseck, Rain Ottis, and Anna-Maria Talihärm. Estonia after the 2007 cyber attacks: Legal, strategic and organisational changes in cyber security. *International Journal of Cyber Warfare and Terrorism (IJCWT)*, 1(1):24–34, 2011.
- [239] Ronald Deibert, Rafal Rohozinski, A Manchanda, Nart Villeneuve, and Greg Walton. Tracking ghostnet: Investigating a cyber espionage network. *Tracking GhostNet: Investigating a cyber espionage network*, 2009.
- [240] Investigating Cyber Espionage. Shadows in the cloud. *Center for Strategic and International Studies-Technology and Public Policy Program*, pages 1–58, 2010.
- [241] Beth Binde, Russ McRee, and Terrence J O’Connor. Assessing outbound traffic to uncover advanced persistent threat. *SANS Institute. Whitepaper*, 16, 2011.
- [242] Global Energy Cyberattacks McAfee. Night dragon. *McAfee Foundstone Professional Services and McAfee Labs*, 2011.
- [243] BBC News. Adobe hack: At least 38 million accounts breached, Oct 2013. Accessed: 2023-10-22.
- [244] Nicole Perloth and Michael J. de la Merced. All 3 billion yahoo accounts were affected by 2013 attack, Oct 2017. Accessed: 2023-10-22.



- [245] Stephan Haggard and Jon R Lindsay. North korea and the sony hack: Exporting instability through cyberspace. 2015.
- [246] Stephanie Gootman. Opm hack: The most dangerous threat to the federal government today. *Journal of Applied Security Research*, 11(4):517–525, 2016.
- [247] Young B Choi. Organizational cyber data breach analysis of facebook, equifax, and uber cases. *International Journal of Cyber Research and Education (IJCRE)*, 3(1):58–64, 2021.
- [248] Saira Ghafur, Soren Kristensen, Kate Honeyford, Guy Martin, Ara Darzi, and Paul Aylin. A retrospective impact analysis of the wannacry cyberattack on the nhs. *NPJ digital medicine*, 2(1):98, 2019.
- [249] Sharifah Yaqoub A Fayi. What petya/notpetya ransomware is and what its remediations are. In *Information technology-new generations: 15th international conference on information technology*, pages 93–100. Springer, 2018.
- [250] Neil Daswani, Moudy Elbayadi, Neil Daswani, and Moudy Elbayadi. The marriott breach. *Big Breaches: Cybersecurity Lessons for Everyone*, pages 55–74, 2021.
- [251] Cybernews. Rocky2021: Largest password compilation of all time leaked, jun 2021. Accessed: 2023-10-25.
- [252] Radhakisan Baheti and Helen Gill. Cyber-physical systems. *The impact of control technology*, 12(1):161–166, 2011.
- [253] Bill Miller and Dale Rowe. A survey scada of and critical infrastructure incidents. In *Proceedings of the 1st Annual conference on Research in information technology*, pages 51–56, 2012.
- [254] Robert J Turk. Cyber incidents involving control systems. Technical report, Idaho National Lab.(INL), Idaho Falls, ID (United States), 2005.
- [255] Richard Derbyshire, Benjamin Green, Daniel Prince, Andreas Mauthe, and David Hutchison. An analysis of cyber security attack taxonomies. In *2018 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 153–161. IEEE, 2018.
- [256] James P Farwell and Rafal Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011.
- [257] Robert M Lee, Michael J Assante, and Tim Conway. Ics cp/pe (cyber-to-physical or process effects) case study paper–german steel mill cyber attack. *Sans ICS*, 2014.
- [258] John W Goodell and Shaen Corbet. Commodity market exposure to energy-firm distress: Evidence from the colonial pipeline ransomware attack. *Finance Research Letters*, 51:103329, 2023.
- [259] Microsoft Digital Security Unit. An overview of russia’s cyberattack activity in ukraine 2022. *Microsoft Special Report*, pages 1–20, 2022.
- [260] Doug Salmon, Mark Zeller, Armando Guzmán, Venkat Mynam, and Marcos Donolo. Mitigating the aurora vulnerability with existing technology. In *36th Annual western protection relay conference*, 2009.
- [261] Recorded Future. Redecho: Advanced persistent threat targeting the indian power sector, mar 2021. Accessed: 2022-12-10.
- [262] Recorded Future. Suspected pakistani actor compromises indian power company with new reverserat, 2021. Accessed: 2022-12-10.
- [263] Clémence Poirier. The war in ukraine from a space cybersecurity perspective. *ESPI Report*, 84, 2022.
- [264] Liang Che, Xuan Liu, Tao Ding, and Zuyi Li. Revealing impacts of cyber attacks on power grids vulnerability to cascading failures. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66(6):1058–1062, 2018.
- [265] Wei Sun, Chen-Ching Liu, and Li Zhang. Optimal generator start-up strategy for bulk power system restoration. *IEEE Transactions on Power Systems*, 26(3):1357–1366, 2010.

- [266] Michael Glassman and Min Ju Kang. Intelligence in the internet age: The emergence and evolution of open source intelligence (osint). *Computers in Human Behavior*, 28(2):673–682, 2012.
- [267] Per Larsen, Stefan Brunthaler, and Michael Franz. Automatic software diversity. *IEEE Security & Privacy*, 13(2):30–37, 2015.
- [268] Sydney Liles, Erin Poremski, and Samuel Liles. Fusion of malware and weapons taxonomies for analysis. *Journal of Information Warfare*, 14(1):75–83, 2015.
- [269] Muhammad Mudassar Yamin, Mohib Ullah, Habib Ullah, and Basel Katt. Weaponized ai for cyber attacks. *Journal of Information Security and Applications*, 57:102722, 2021.
- [270] Blessing Guembe, Ambrose Azeta, Sanjay Misra, Victor Chukwudi Osamor, Luis Fernandez-Sanz, and Vera Pospelova. The emerging threat of ai-driven cyber attacks: A review. *Applied Artificial Intelligence*, 36(1):2037254, 2022.
- [271] Offensive Security. Exploit database, 2024. Accessed: 2023-12-16.
- [272] MITRE Corporation. Common vulnerabilities and exposures (cve), 2024. Accessed: 2024-4-16.
- [273] Umesh Kumar Singh, Chanchala Joshi, and Dimitris Kanellopoulos. A framework for zero-day vulnerabilities detection and prioritization. *Journal of Information Security and Applications*, 46:164–172, 2019.
- [274] Long Cheng, Fang Liu, and Danfeng Yao. Enterprise data breach: causes, challenges, prevention, and future directions. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(5):e1211, 2017.
- [275] Pouyan Pourbeik, Prabha S Kundur, and Carson W Taylor. The anatomy of a power grid blackout-root causes and dynamics of recent major blackouts. *IEEE Power and Energy Magazine*, 4(5):22–29, 2006.
- [276] Cheng Wang and Hangyu Zhu. Wrongdoing monitor: A graph-based behavioral anomaly detection in cyber security. *IEEE Transactions on Information Forensics and Security*, 17:2703–2718, 2022.
- [277] Göran Andersson, Peter Donalek, Richard Farmer, Nikos Hatziaargyriou, Innocent Kamwa, Prabhashankar Kundur, Nelson Martins, John Paserba, Pouyan Pourbeik, Juan Sanchez-Gasca, et al. Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance. *IEEE transactions on Power Systems*, 20(4):1922–1928, 2005.
- [278] Richard G Little. Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. *Journal of Urban Technology*, 9(1):109–123, 2002.
- [279] Yuri V Makarov, Viktor I Reshetov, A Stroev, and I Voropai. Blackout prevention in the united states, europe, and russia. *Proceedings of the IEEE*, 93(11):1942–1955, 2005.
- [280] Final report of the investigation committee on the september 3 blackout in italy, July 2021. [Online]. Available: [http://ns2.rae.gr/old/cases/C13/italy/UCTE\\_rept.pdf](http://ns2.rae.gr/old/cases/C13/italy/UCTE_rept.pdf).
- [281] Janusz W Bialek. Why has it happened again? comparison between the ucte blackout in 2006 and the blackouts of 2003. In *2007 IEEE Lausanne Power Tech*, pages 51–56. IEEE, 2007.
- [282] Joshua J Romero. Blackouts illuminate india’s power problems. *IEEE spectrum*, 49(10):11–12, 2012.
- [283] Benjamin Schäfer and G Cigdem Yalcin. Dynamical modeling of cascading failures in the turkish power grid. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(9), 2019.
- [284] S. Wilde. 9 august 2019 power outage report, December 2022. [Online]. Available: [https://www.ofgem.gov.uk/sites/default/files/docs/2020/01/9\\_august\\_2019\\_power\\_outage\\_report.pdf](https://www.ofgem.gov.uk/sites/default/files/docs/2020/01/9_august_2019_power_outage_report.pdf).
- [285] Vetrivel Subramaniam Rajkumar, Alexandru Ștefanov, Alfán Presekal, Peter Palensky, and José Luis Rueda Torres. Cyber attacks on power grids: Causes and propagation of cascading failures. *IEEE Access*, 2023.

- [286] Christos Makridis and Benjamin Dean. Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities. *Journal of Economic and Social Measurement*, 43(1-2):59–83, 2018.
- [287] Cybersecurity and Infrastructure Security Agency (CISA). Energy sector - critical infrastructure sectors, 2024. Accessed: 2023-12-16.
- [288] G Brooke Anderson and Michelle L Bell. Lights out: impact of the august 2003 power outage on mortality in new york, ny. *Epidemiology*, 23(2):189–193, 2012.
- [289] Valentin R Melnikov, Valeria V Krzhizhanovskaya, Alexander V Boukhanovsky, and Peter MA Sloot. Data-driven modeling of transportation systems and traffic data analysis during a major power outage in the netherlands. *Procedia Computer Science*, 66:336–345, 2015.
- [290] Sinan Küfeoğlu et al. Economic impacts of electric power outages and evaluation of customer interruption costs. 2015.
- [291] Nadiya Kostyuk and Yuri M Zhukov. Invisible digital front: Can cyber attacks shape battlefield events? *Journal of Conflict Resolution*, 63(2):317–347, 2019.
- [292] Yutian Liu, Rui Fan, and Vladimir Terzija. Power system restoration: a literature review from 2006 to 2016. *Journal of Modern Power Systems and Clean Energy*, 4(3):332–341, 2016.
- [293] Rajaa Vikhram Yohanandhan, Rajvikram Madurai Elavarasan, Rishi Pugazhendhi, Manoharan Premkumar, Lucian Mihet-Popa, and Vladimir Terzija. A holistic review on cyber-physical power system (cpps) testbeds for secure and sustainable electric power grid–part–ii: Classification, overview and assessment of cpps testbeds. *International Journal of Electrical Power & Energy Systems*, 137:107721, 2022.
- [294] Rogério Leão Santos De Oliveira, Christiane Marie Schweitzer, Ailton Akira Shinoda, and Ligia Rodrigues Prete. Using mininet for emulation and prototyping software-defined networks. In *2014 IEEE Colombian conference on communications and computing (COLCOM)*, pages 1–6. Ieee, 2014.
- [295] John F Hauer, Navin B Bhatt, Kirit Shah, and Sharma Kolluri. Performance of "wams east" in providing dynamic information for the north east blackout of august 14, 2003. In *IEEE Power Engineering Society General Meeting, 2004.*, pages 1685–1690. IEEE, 2004.
- [296] Meaad Ali Khalaf and Amani Steiti. Artificial intelligence predictions in cyber security: Analysis and early detection of cyber attacks. *Babylonian Journal of Machine Learning*, 2024:63–68, 2024.
- [297] Abdulazeez Alsajri and Amani Steiti. Intrusion detection system based on machine learning algorithms:(svm and genetic algorithm). *Babylonian Journal of Machine Learning*, 2024:15–29, 2024.
- [298] Rasha Hameed Khudhur Al-Rubaye and AYÇA KURNAZ TÜRK BEN. Using artificial intelligence to evaluating detection of cybersecurity threats in ad hoc networks. *Babylonian Journal of Networking*, 2024:45–56, 2024.
- [299] Sara salman Qasim and Sarah Mohammed NSAIF. Advancements in time series-based detection systems for distributed denial-of-service (ddos) attacks: A comprehensive review. *Babylonian Journal of Networking*, 2024:9–17, 2024.
- [300] Chee-Wooi Ten, Junho Hong, and Chen-Ching Liu. Anomaly detection for cybersecurity of the substations. *IEEE Transactions on Smart Grid*, 2(4):865–873, 2011.
- [301] Qi Wang, Xingpu Cai, Yi Tang, and Ming Ni. Methods of cyber-attack identification for power systems based on bilateral cyber-physical information. *International Journal of Electrical Power & Energy Systems*, 125:106515, 2021.
- [302] Robert Mitchell and Ing-Ray Chen. Behavior-rule based intrusion detection systems for safety critical smart grid applications. *IEEE Transactions on Smart Grid*, 4(3):1254–1263, 2013.

- [303] Gregory M. Coates, Kenneth M. Hopkinson, Scott R. Graham, and Stuart H. Kurkowski. Collaborative, trust-based security mechanisms for a regional utility intranet. *IEEE Transactions on Power Systems*, 23(3):831–844, 2008.
- [304] Y. Yang, K. McLaughlin, S. Sezer, T. Littler, B. Pranggono, P. Brogan, and H. F. Wang. Intrusion detection system for network security in synchrophasor systems. In *IET International Conference on Information and Communications Technologies (IETICT 2013)*, pages 246–252, 2013.
- [305] Shan Ali and Yuancheng Li. Learning multilevel auto-encoders for ddos attack detection in smart grid network. *IEEE Access*, 7:108647–108659, 2019.
- [306] Mete Ozay, İñaki Esnaola, Fatos Tunay Yarman Vural, Sanjeev R. Kulkarni, and H. Vincent Poor. Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Networks and Learning Systems*, 27(8):1773–1786, 2016.
- [307] Manikant Panthi. Anomaly detection in smart grids using machine learning techniques. In *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, pages 220–222, 2020.
- [308] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.
- [309] Hongyu Liu and Bo Lang. Machine learning and deep learning methods for intrusion detection systems: A survey. *applied sciences*, 9(20):4396, 2019.
- [310] Arwa Aldweesh, Abdelouahid Derhab, and Ahmed Z Emam. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowledge-Based Systems*, 189:105124, 2020.
- [311] Preeti Mishra, Vijay Varadharajan, Uday Tupakula, and Emmanuel S. Pilli. A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys Tutorials*, 21(1):686–728, 2019.
- [312] Rafael RR Barbosa, Ramin Sadre, and Aiko Pras. Difficulties in modeling scada traffic: a comparative analysis. In *Passive and Active Measurement: 13th International Conference, PAM 2012, Vienna, Austria, March 12-14th, 2012. Proceedings 13*, pages 126–135. Springer, 2012.
- [313] Raghavendra Chalapathy and Sanjay Chawla. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*, 2019.
- [314] Trung V. Phan, Tri Gia Nguyen, Nhu-Ngoc Dao, Truong Thu Huong, Nguyen Huu Thanh, and Thomas Bauschert. Deepguard: Efficient anomaly detection in sdn with fine-grained traffic flow monitoring. *IEEE Transactions on Network and Service Management*, 17(3):1349–1362, 2020.
- [315] Ren-Hung Hwang, Min-Chun Peng, Chien-Wei Huang, Po-Ching Lin, and Van-Linh Nguyen. An unsupervised deep learning model for early network traffic anomaly detection. *IEEE Access*, 8:30387–30399, 2020.
- [316] Xiaohong Guan, Tao Qin, Wei Li, and Pinghui Wang. Dynamic feature analysis and measurement for large-scale network traffic monitoring. *IEEE Transactions on Information Forensics and Security*, 5(4):905–919, 2010.
- [317] Andreas Kind, Marc Ph. Stoecklin, and Xenofontas Dimitropoulos. Histogram-based traffic anomaly detection. *IEEE Transactions on Network and Service Management*, 6(2):110–121, 2009.
- [318] Kuai Xu, Zhi-Li Zhang, and Supratik Bhattacharyya. Internet traffic behavior profiling for network security monitoring. *IEEE/ACM Transactions On Networking*, 16(6):1241–1252, 2008.
- [319] Hu-Sheng Wu. A survey of research on anomaly detection for time series. In *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pages 426–431, 2016.

- [320] Kamran Shaukat, Talha Mahboob Alam, Suhui Luo, Shakir Shabbir, Ibrahim A Hameed, Jiaming Li, Syed Konain Abbas, and Umair Javed. A review of time-series anomaly detection techniques: A step to future perspectives. In *Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 1*, pages 865–877. Springer, 2021.
- [321] Weixuan Lin, Di Wu, and Benoit Boulet. Spatial-temporal residential short-term load forecasting via graph neural networks. *IEEE Transactions on Smart Grid*, 12(6):5373–5384, 2021.
- [322] Osman Boyaci, Mohammad Rasoul Narimani, Katherine R. Davis, Muhammad Ismail, Thomas J. Overbye, and Erchin Serpedin. Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks. *IEEE Transactions on Smart Grid*, 13(1):807–819, 2022.
- [323] Zhiyong Cui, Kristian Henrickson, Ruimin Ke, and Yinhai Wang. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *IEEE Transactions on Intelligent Transportation Systems*, 21(11):4883–4894, 2020.
- [324] Leyan Deng, Defu Lian, Zhenya Huang, and Enhong Chen. Graph convolutional adversarial networks for spatiotemporal anomaly detection. *IEEE Transactions on Neural Networks and Learning Systems*, 33(6):2416–2428, 2022.
- [325] Hansi Chen, Hang Liu, Xuening Chu, Qingxiu Liu, and Deyi Xue. Anomaly detection and critical scada parameters identification for wind turbines based on lstm-ae neural network. *Renewable Energy*, 172:829–840, 2021.
- [326] Sagnik Basumallik, Rui Ma, and Sara Eftekharijad. Packet-data anomaly detection in pmu-based state estimator using convolutional neural network. *International Journal of Electrical Power & Energy Systems*, 107:690–702, 2019.
- [327] Xinming Ou, Wayne F Boyer, and Miles A McQueen. A scalable approach to attack graph generation. In *Proceedings of the 13th ACM conference on Computer and communications security*, pages 336–345, 2006.
- [328] Kerem Kaynar and Fikret Sivrikaya. Distributed attack graph generation. *IEEE Transactions on Dependable and Secure Computing*, 13(5):519–532, 2016.
- [329] Seunghyun Yoon, Jin-Hee Cho, Dong Seong Kim, Terrence J. Moore, Frederica Free-Nelson, and Hyuk Lim. Attack graph-based moving target defense in software-defined networks. *IEEE Transactions on Network and Service Management*, 17(3):1653–1668, 2020.
- [330] Diego Kreutz, Fernando M. V. Ramos, Paulo Esteves Verissimo, Christian Esteve Rothenberg, Siamak Azodolmolky, and Steve Uhlig. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1):14–76, 2015.
- [331] Jun Wu, Shibo Luo, Shen Wang, and Hongkai Wang. Nles: A novel lifetime extension scheme for safety-critical cyber-physical systems using sdn and nfv. *IEEE Internet of Things Journal*, 6(2):2463–2475, 2019.
- [332] Yan Li, Yanyuan Qin, Peng Zhang, and Amir Herzberg. Sdn-enabled cyber-physical security in networked microgrids. *IEEE Transactions on Sustainable Energy*, 10(3):1613–1622, 2019.
- [333] Xu Zhang, Kefeng Wei, Lei Guo, Weigang Hou, and Jingjing Wu. Sdn-based resilience solutions for smart grids. In *2016 International Conference on Software Networking (ICSN)*, pages 1–5, 2016.
- [334] Ahmadrza Montazerolghaem and Mohammad Hossein Yaghmaee. Demand response application as a service: An sdn-based management framework. *IEEE Transactions on Smart Grid*, 13(3):1952–1966, 2022.
- [335] Mubashir Husain Rehmani, Fayaz Akhtar, Alan Davy, and Brendan Jennings. Achieving resilience in sdn-based smart grid: A multi-armed bandit approach. In *2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft)*, pages 366–371, 2018.
- [336] Peng Zhang, Fangzheng Zhang, Shimin Xu, Zuoru Yang, Hao Li, Qi Li, Huan Zhao Wang, Chao Shen, and Chengchen Hu. Network-wide forwarding anomaly detection and localization in software defined networks. *IEEE/ACM Transactions on Networking*, 29(1):332–345, 2021.

- [337] Xiangtian Zheng, Nan Xu, Loc Trinh, Dongqi Wu, Tong Huang, S Sivaranjani, Yan Liu, and Le Xie. A multi-scale time-series dataset with benchmark for machine learning in decarbonized energy grids. *Scientific Data*, 9(1):359, 2022.
- [338] Saleh Soltan, Alexander Loh, and Gil Zussman. A learning-based method for generating synthetic power grids. *IEEE Systems Journal*, 13(1):625–634, 2019.
- [339] Andreas Venzke, Daniel K Molzahn, and Spyros Chatzivasileiadis. Efficient creation of datasets for data-driven power system applications. *Electric Power Systems Research*, 190:106614, 2021.
- [340] Md Fazla Elahe, Min Jin, and Pan Zeng. Review of load data analytics using deep learning in smart grids: Open load datasets, methodologies, and application challenges. *International Journal of Energy Research*, 45(10):14274–14305, 2021.
- [341] Sakineh Tavakkoli, Jordan Macknick, Garvin A Heath, and Sarah M Jordaan. Spatiotemporal energy infrastructure datasets for the united states: A review. *Renewable and Sustainable Energy Reviews*, 152:111616, 2021.
- [342] Yassine Himeur, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira. Building power consumption datasets: Survey, taxonomy and future directions. *Energy and Buildings*, 227:110404, 2020.
- [343] M Naglic. Pmu measurements of ieee 39-bus power system model. *IEEE Dataport*, 2019.
- [344] Shengyi Pan, Thomas Morris, and Uttam Adhikari. Developing a hybrid intrusion detection system using data mining for power systems. *IEEE Transactions on Smart Grid*, 6(6):3104–3113, 2015.
- [345] Uttam Adhikari, Thomas Morris, and Shengyi Pan. Wams cyber-physical test bed for power system, cybersecurity study, and data mining. *IEEE Transactions on Smart Grid*, 8(6):2744–2753, 2017.
- [346] Adam Hahn, Aditya Ashok, Siddharth Sridhar, and Manimaran Govindarasu. Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid. *IEEE Transactions on Smart Grid*, 4(2):847–855, 2013.
- [347] Mehmet Hazar Cintuglu, Osama A. Mohammed, Kemal Akkaya, and A. Selcuk Uluagac. A survey on smart grid cyber-physical system testbeds. *IEEE Communications Surveys Tutorials*, 19(1):446–464, 2017.
- [348] Brij B Gupta and Tafseer Akhtar. A survey on smart power grid: frameworks, tools, security issues, and solutions. *Annals of Telecommunications*, 72:517–549, 2017.
- [349] Xin Zhou, Xiaodong Gou, Tingting Huang, and Shunkun Yang. Review on testing of cyber-physical systems: Methods and testbeds. *IEEE Access*, 6:52179–52194, 2018.
- [350] Muhammet Zekeriya Gunduz and Resul Das. A comparison of cyber-security oriented testbeds for iot-based smart grids. In *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*, pages 1–6, 2018.
- [351] Juan Montoya, Ron Brandl, Keerthi Vishwanath, Jay Johnson, Rachid Darbali-Zamora, Adam Summers, Jun Hashimoto, Hiroshi Kikusato, Taha Selim Ustun, Nayeem Ninad, et al. Advanced laboratory testing methods using real-time simulation and hardware-in-the-loop techniques: A survey of smart grid international research facility network activities. *Energies*, 13(12):3267, 2020.
- [352] Marios Iliofotou, Prashanth Pappu, Michalis Faloutsos, Michael Mitzenmacher, Sumeet Singh, and George Varghese. Network monitoring using traffic dispersion graphs (tdgs). In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pages 315–320, 2007.
- [353] Do Quoc Le, Taeyoel Jeong, H Eduardo Roman, and James Won-Ki Hong. Traffic dispersion graph based anomaly detection. In *Proceedings of the 2nd Symposium on Information and Communication Technology*, pages 36–41, 2011.
- [354] Jinyin Chen, Xueke Wang, and Xuanheng Xu. Gc-lstm: Graph convolution embedded lstm for dynamic network link prediction. *Applied Intelligence*, pages 1–16, 2022.

- [355] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.
- [356] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, 2015.
- [357] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.
- [358] Martin Längkvist, Lars Karlsson, and Amy Loutfi. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern recognition letters*, 42:11–24, 2014.
- [359] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.
- [360] Ahmadrza Montazerolghaem and Mohammad Hossein Yaghmaee. Demand response application as a service: An sdn-based management framework. *IEEE Transactions on Smart Grid*, 13(3):1952–1966, 2021.
- [361] Emmanuel S Pilli, Ramesh C Joshi, and Rajdeep Niyogi. Network forensic frameworks: Survey and research challenges. *digital investigation*, 7(1-2):14–27, 2010.
- [362] Vinicius Tavares Guimaraes, Carla Maria Dal Sasso Freitas, Ramin Sadre, Liane Margarida Rockenbach Tarouco, and Lisandro Zambenedetti Granville. A survey on information visualization for network and service management. *IEEE Communications Surveys & Tutorials*, 18(1):285–323, 2015.
- [363] Ronghua Shi, Mengjie Yang, Ying Zhao, Fangfang Zhou, Wei Huang, and Sheng Zhang. A matrix-based visualization system for network traffic forensics. *IEEE Systems Journal*, 10(4):1350–1360, 2015.
- [364] Reza Mohammadi, Reza Javidan, and Mauro Conti. Slicots: An sdn-based lightweight countermeasure for tcp syn flooding attacks. *IEEE Transactions on Network and Service Management*, 14(2):487–497, 2017.
- [365] Tomas Zitta, Marek Neruda, Lukas Vojtech, Martina Matejkova, Matej Jehlicka, Lukas Hach, and Jan Moravec. Penetration testing of intrusion detection and prevention system in low-performance embedded iot device. In *2018 18th International Conference on Mechatronics-Mechatronika (ME)*, pages 1–5. IEEE, 2018.
- [366] Ricardo Barandela, José Salvador Sánchez, Vicente Garcia, and Edgar Rangel. Strategies for learning in class imbalance problems. *Pattern Recognition*, 36(3):849–851, 2003.
- [367] Justin M Johnson and Taghi M Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of big data*, 6(1):1–54, 2019.
- [368] Bosung Kim, Youngjoong Ko, and Jungyun Seo. Novel regularization method for the class imbalance problem. *Expert Systems with Applications*, 188:115974, 2022.
- [369] Partha P Biswas, Heng Chuan Tan, Qingbo Zhu, Yuan Li, Daisuke Mashima, and Binbin Chen. A synthesized dataset for cybersecurity study of iec 61850 based substation. In *2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–7. IEEE, 2019.
- [370] Sowmya Myneni, Ankur Chowdhary, Abdulhakim Sabur, Sailik Sengupta, Garima Agrawal, Dijiang Huang, and Myong Kang. Dapt 2020-constructing a benchmark dataset for advanced persistent threats. In *Deployable Machine Learning for Security Defense: First International Workshop, MLHat 2020, San Diego, CA, USA, August 24, 2020, Proceedings 1*, pages 138–163. Springer, 2020.
- [371] André Felipe Silva Melo, Jose Miguel Riquelme-Dominguez, Francisco Gonzalez-Longatt, Jose L Rueda, and Peter Palensky. Sampled values rocof performance methodology breakdown. In *2022 IEEE International Conference on Environment and Electrical Engineering and 2022 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe)*, pages 1–5. IEEE, 2022.



- [372] Anne-Cécile Orgerie, Paulo Gonçalves, Matthieu Imbert, Julien Ridoux, and Darryl Veitch. Survey of network metrology platforms. In *2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet*, pages 220–225. IEEE, 2012.
- [373] Adel Alshamrani, Sowmya Myneni, Ankur Chowdhary, and Dijiang Huang. A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities. *IEEE Communications Surveys & Tutorials*, 21(2):1851–1877, 2019.
- [374] Atif Ahmad, Jeb Webb, Kevin C Desouza, and James Boorman. Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack. *Computers & Security*, 86:402–418, 2019.
- [375] Branka Stojanović, Katharina Hofer-Schmitz, and Ulrike Kleb. Apt datasets and attack modeling for automated detection methods: A review. *Computers & Security*, 92:101734, 2020.
- [376] BinHui Tang, JunFeng Wang, Zhongkun Yu, Bohan Chen, Wenhan Ge, Jian Yu, and TingTing Lu. Advanced persistent threat intelligent profiling technique: A survey. *Computers and Electrical Engineering*, 103:108261, 2022.
- [377] Chee-Wooi Ten, Junho Hong, and Chen-Ching Liu. Anomaly detection for cybersecurity of the substations. *IEEE Transactions on Smart Grid*, 2(4):865–873, 2011.
- [378] Qi Wang, Xingpu Cai, Yi Tang, and Ming Ni. Methods of cyber-attack identification for power systems based on bilateral cyber-physical information. *International Journal of Electrical Power & Energy Systems*, 125:106515, 2021.
- [379] Robert Mitchell and Ray Chen. Behavior-rule based intrusion detection systems for safety critical smart grid applications. *IEEE Transactions on Smart Grid*, 4(3):1254–1263, 2013.
- [380] Gregory M Coates, Kenneth M Hopkinson, Scott R Graham, and Stuart H Kurkowski. Collaborative, trust-based security mechanisms for a regional utility intranet. *IEEE Transactions on power systems*, 23(3):831–844, 2008.
- [381] Mete Ozay, Inaki Esnaola, Fatos Tunay Yarman Vural, Sanjeev R Kulkarni, and H Vincent Poor. Machine learning methods for attack detection in the smart grid. *IEEE transactions on neural networks and learning systems*, 27(8):1773–1786, 2015.
- [382] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.
- [383] Arwa Aldweesh, Abdelouahid Derhab, and Ahmed Z Emam. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowledge-Based Systems*, 189:105124, 2020.
- [384] Rafael RR Barbosa, Ramin Sadre, and Aiko Pras. Difficulties in modeling scada traffic: a comparative analysis. In *Passive and Active Measurement: 13th International Conference, PAM 2012, Vienna, Austria, March 12-14th, 2012. Proceedings 13*, pages 126–135. Springer, 2012.
- [385] Leila Mohammadpour, Teck Chaw Ling, Chee Sun Liew, and Alihossein Aryanfar. A survey of cnn-based network intrusion detection. *Applied Sciences*, 12(16):8162, 2022.
- [386] Mohammad Norouzi, David J Fleet, and Russ R Salakhutdinov. Hamming distance metric learning. *Advances in neural information processing systems*, 25, 2012.
- [387] Yu Zhou, Jianjun He, and Hong Gu. Partial label learning via gaussian processes. *IEEE transactions on cybernetics*, 47(12):4443–4450, 2016.
- [388] Heng-Chao Yan, Jun-Hong Zhou, and Chee Khiang Pang. Gaussian mixture model using semisupervised learning for probabilistic fault diagnosis under new data categories. *IEEE Transactions on Instrumentation and Measurement*, 66(4):723–733, 2017.



- [389] Miin-Shen Yang, Chien-Yo Lai, and Chih-Ying Lin. A robust em clustering algorithm for gaussian mixture models. *Pattern Recognition*, 45(11):3950–3961, 2012.
- [390] Tristan Bilot, Nour El Madhoun, Khaldoun Al Agha, and Anis Zouaoui. Graph neural networks for intrusion detection: A survey. *IEEE Access*, 11:49114–49139, 2023.
- [391] Sangjun Kim, Kyung-Joon Park, and Chenyang Lu. A survey on network security for cyber–physical systems: From threats to resilient design. *IEEE Communications Surveys & Tutorials*, 24(3):1534–1573, 2022.
- [392] Shimiao Li, Amritanshu Pandey, Bryan Hooi, Christos Faloutsos, and Larry Pileggi. Dynamic graph-based anomaly detection in the electrical grid. *IEEE Transactions on Power Systems*, 37(5):3408–3422, 2021.
- [393] Zhiwei Wang, Wei Jiang, Junjun Xu, Zhiqi Xu, Aihua Zhou, and Min Xu. Grid2vec: Learning node representations of digital power systems for anomaly detection. *IEEE Transactions on Smart Grid*, 2024.
- [394] Xing Ling, Yeonwoo Rho, and Chee-Wooi Ten. Predicting global trend of cybersecurity on continental honeynets using vector autoregression. In *2019 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pages 1–5. IEEE, 2019.
- [395] Zhiyuan Yang, Shipeng Zhang, Chee-Wooi Ten, Ting Liu, Xueyue Pang, and Hao Sun. Implementation of risk-aggregated substation testbed using generative adversarial networks. *IEEE Transactions on Smart Grid*, 14(1):677–689, 2022.
- [396] Peng Gao, Weiyong Yang, Haotian Zhang, Xingshen Wei, Hao Huang, Wang Luo, Zhimin Guo, and Yunhe Hao. Detecting unknown threat based on continuous-time dynamic heterogeneous graph network. *Wireless Communications and Mobile Computing*, 2022(1):7502294, 2022.
- [397] Chung-Kuan Chen, Si-Chen Lin, Szu-Chun Huang, Yung-Tien Chu, Chin-Laung Lei, and Chun-Ying Huang. Building machine learning-based threat hunting system from scratch. *Digital Threats: Research and Practice (DTRAP)*, 3(3):1–21, 2022.
- [398] Jingyu Yang and Zuogong Yue. Learning hierarchical spatial-temporal graph representations for robust multivariate industrial anomaly detection. *IEEE Transactions on Industrial Informatics*, 19(6):7624–7635, 2022.
- [399] Bryan Lim and Stefan Zohren. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, 379(2194):20200209, 2021.
- [400] Shuhan Yuan and Xintao Wu. Deep learning for insider threat detection: Review, challenges and opportunities. *Computers & Security*, 104:102221, 2021.
- [401] Youngjoo Seo, Michaël Defferrard, Pierre Vandergheynst, and Xavier Bresson. Structured sequence modeling with graph convolutional recurrent networks. In *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13-16, 2018, Proceedings, Part I* 25, pages 362–373. Springer, 2018.
- [402] Ling Zhao, Yujiao Song, Chao Zhang, Yu Liu, Pu Wang, Tao Lin, Min Deng, and Haifeng Li. T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE transactions on intelligent transportation systems*, 21(9):3848–3858, 2019.
- [403] James Jie Pan, Jianguo Wang, and Guoliang Li. Survey of vector database management systems. *The VLDB Journal*, 33(5):1591–1615, 2024.
- [404] Aaron J Wilson, Donald R Reising, Robert W Hay, Ray C Johnson, Abdelrahman A Karrar, and T Daniel Loveless. Automated identification of electrical disturbance waveforms within an operational smart power grid. *IEEE Transactions on Smart Grid*, 11(5):4380–4389, 2020.
- [405] Suhail SM Hussain, Chen Yaohao, Muhammad M Roomi, Daisuke Mashima, and Ee-Chien Chang. An open-source framework for publishing/subscribing iec 61850 r-goose and r-sv. *SoftwareX*, 23:101415, 2023.




- [406] Petr Matoušek, Ondřej Ryšavý, and Peter Grofčík. Ics dataset for smart grid anomaly detection, 2022.
- [407] Panagiotis Radoglou-Grammatikis, Vasiliki Kelli, Thomas Lagkas, Vasileios Argyriou, and Panagiotis Sarigiannidis. Dnp3 intrusion detection dataset, 2022.
- [408] ITI. ICS Security Tools. <https://github.com/ITI/ICS-Security-Tools>, 2024. Accessed: 2024-11-26.
- [409] Hyunjae Kang, Dong Hyun Ahn, Gyung Min Lee, Jeong Do Yoo, Kyung Ho Park, and Huy Kang Kim. Iot network intrusion dataset, 2019.
- [410] Netresec. Industroyer2 iec-104 analysis. <https://www.netresec.com/?page=Blog&month=2022-04&post=Industroyer2-IEC-104-Analysis>, 2022. Accessed: 2024-11-26.
- [411] Yong Zhang, Jiacheng Wu, Jin Wang, and Chunxiao Xing. A transformation-based framework for knn set similarity search. *IEEE Transactions on Knowledge and Data Engineering*, 32(3):409–423, 2018.
- [412] Jon Louis Bentley. K-d trees for semidynamic point sets. In *Proceedings of the sixth annual symposium on Computational geometry*, pages 187–197, 1990.
- [413] Yu A Malkov and Dmitry A Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE transactions on pattern analysis and machine intelligence*, 42(4):824–836, 2018.
- [414] Tapas Kanungo, David M Mount, Nathan S Netanyahu, Christine D Piatko, Ruth Silverman, and Angela Y Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis and machine intelligence*, 24(7):881–892, 2002.
- [415] Aristides Gionis, Piotr Indyk, Rajeev Motwani, et al. Similarity search in high dimensions via hashing. In *Vldb*, volume 99, pages 518–529, 1999.
- [416] Andrea Carcano, Alessio Coletta, Michele Guglielmi, Marcelo Masera, I Nai Fovino, and Alberto Trombetta. A multidimensional critical state analysis for detecting intrusions in scada systems. *IEEE Transactions on Industrial Informatics*, 7(2):179–186, 2011.
- [417] Patric Nader, Paul Honeine, and Pierre Beausery. Lp-norms in one-class classification for intrusion detection in scada systems. *IEEE Transactions on Industrial Informatics*, 10(4):2308–2317, 2014.
- [418] Junho Hong, Reynaldo F Nuqui, Anil Kondabathini, Dmitry Ishchenko, and Aaron Martin. Cyber attack resilient distance protection and circuit breaker control for digital substations. *IEEE Transactions on Industrial Informatics*, 15(7):4332–4341, 2018.
- [419] Ghada Elbez, Klara Nahrstedt, and Veit Hagenmeyer. Early attack detection for securing goose network traffic. *IEEE Transactions on Smart Grid*, 15(1):899–910, 2023.
- [420] Silvio E Quincozes, Célio Albuquerque, Diego Passos, and Daniel Mossé. A survey on intrusion detection and prevention systems in digital substations. *Computer Networks*, 184:107679, 2021.
- [421] Upeka Kanchana Premaratne, Jagath Samarabandu, Tarlochan S Sidhu, Robert Beresh, and Jian-Cheng Tan. An intrusion detection system for iec61850 automated substations. *IEEE Transactions on Power Delivery*, 25(4):2376–2383, 2010.
- [422] A Abu Nassar and WG Morsi. Detection of cyber-attacks and power disturbances in smart digital substations using continuous wavelet transform and convolution neural networks. *Electric Power Systems Research*, 229:110157, 2024.
- [423] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1):1–22, 2019.

- [424] Arwa Aldweesh, Abdelouahid Derhab, and Ahmed Z Emam. Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowledge-Based Systems*, 189:105124, 2020.
- [425] Ke He, Dan Dongseong Kim, and Muhammad Rizwan Asghar. Adversarial machine learning for network intrusion detection systems: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 25(1):538–566, 2023.
- [426] David J Weller-Fahy, Brett J Borghetti, and Angela A Sodemann. A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Communications Surveys & Tutorials*, 17(1):70–91, 2014.
- [427] Piotr Kokoszka, Hieu Nguyen, Haonan Wang, and Liuling Yang. Statistical and probabilistic analysis of interarrival and waiting times of internet2 anomalies. *Statistical Methods & Applications*, 29:727–744, 2020.
- [428] Xiangyu Kong, Yizhi Zhou, Yilei Xiao, Xuezhou Ye, Heng Qi, and Xiulong Liu. idetector: A novel real-time intrusion detection solution for iot networks. *IEEE Internet of Things Journal*, 2024.
- [429] Taki Eddine Toufik Djaidja, Bouziane Brik, Sidi Mohammed Senouci, Abdelwahab Boualouache, and Yacine Ghamri-Doudane. Early network intrusion detection enabled by attention mechanisms and rnns. *IEEE Transactions on Information Forensics and Security*, 2024.
- [430] Ning Wang, Yimin Chen, Yang Xiao, Yang Hu, Wenjing Lou, and Y Thomas Hou. Manda: On adversarial example detection for network intrusion detection system. *IEEE Transactions on Dependable and Secure Computing*, 20(2):1139–1153, 2022.
- [431] Gianmarco Baldini and Irene Amerini. Online distributed denial of service (ddos) intrusion detection based on adaptive sliding window and morphological fractal dimension. *Computer Networks*, 210:108923, 2022.
- [432] Xin Liu, Weitong Chen, Lu Peng, Dan Luo, Likai Jia, Gang Xu, Xiubo Chen, and Xiaomeng Liu. Secure computation protocol of chebyshev distance under the malicious model. *Scientific Reports*, 14(1):17115, 2024.
- [433] Liqun Yang, You Zhai, Yipeng Zhang, Yufei Zhao, Zhoujun Li, and Tongge Xu. A new methodology for anomaly detection of attacks in iec 61850-based substation system. *Journal of Information Security and Applications*, 68:103262, 2022.
- [434] Manuel Herrera, Yaniv Proselkov, Marco Perez-Hernandez, and Ajith Kumar Parlikad. Mining graph-fourier transform time series for anomaly detection of internet traffic at core and metro networks. *IEEE Access*, 9:8997–9011, 2021.
- [435] Longkai Sui and Yongguo Jiang. Argo data anomaly detection based on transformer and fourier transform. *Journal of Sea Research*, 198:102483, 2024.
- [436] Olayinka S Ohunakin, Emerald U Henry, Olaniran J Matthew, Victor U Ezekiel, Damola S Adelekan, and Ayodele T Oyeniran. Conditional monitoring and fault detection of wind turbines based on kolmogorov-smirnov non-parametric test. *Energy Reports*, 11:2577–2591, 2024.
- [437] Badreddine Cherkaoui, Mohammed-Alamine El Houssaini, Mohammed Kasri, Abderrahim Beni-Hssane, and Mohammed Erritali. Kolmogorov-smirnov based method for detecting black hole attack in vehicular ad-hoc networks. *Procedia Computer Science*, 236:177–184, 2024.
- [438] Xiaofei Qu, Lin Yang, Kai Guo, Linru Ma, Meng Sun, Mingxing Ke, and Mu Li. A survey on the development of self-organizing maps for unsupervised intrusion detection. *Mobile networks and applications*, 26:808–829, 2021.
- [439] Chathurika S Wickramasinghe, Kasun Amarasinghe, and Milos Manic. Deep self-organizing maps for unsupervised image classification. *IEEE Transactions on Industrial Informatics*, 15(11):5837–5845, 2019.


- [440] Shahneela Pitafi, Toni Anwar, I Dewa Made Widia, and Boonsit Yimwadsana. Revolutionizing perimeter intrusion detection: A machine learning-driven approach with curated dataset generation for enhanced security. *IEEE Access*, 2023.
- [441] Beatriz Flámia Azevedo, Ana Maria AC Rocha, and Ana I Pereira. Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Machine Learning*, pages 1–43, 2024.
- [442] Sven Zemanek, Immanuel Hacker, Konrad Wolsing, Eric Wagner, Martin Henze, and Martin Serror. Powerduck: a goose data set of cyberattacks in substations. In *Proceedings of the 15th Workshop on Cyber Security Experimentation and Test*, pages 49–53, 2022.
- [443] H. Kang et al. Iot network intrusion dataset. <https://ieee-dataport.org>, September 2019. Accessed: 2024-08-17.
- [444] Sohaib Asif et al. A fuzzy minkowski distance-based fusion of convolutional neural networks for gastrointestinal disease detection. *Applied Soft Computing*, 158:111595, 2024.
- [445] Zhiyuan Tan, Aruna Jamdagni, Xiangjian He, Priyadarsi Nanda, and Ren Ping Liu. Multivariate correlation analysis technique based on euclidean distance map for network traffic characterization. In *Information and Communications Security: 13th International Conference, ICICS 2011, Beijing, China, November 23-26, 2011. Proceedings 13*, pages 388–398. Springer, 2011.
- [446] Rajneesh Kumar Pandey and Tanmoy Kanti Das. Anomaly detection for industrial control networks using hamming distance. In *International Conference on Information Systems and Management Science*, pages 280–290. Springer, 2021.
- [447] Zhe Yao, Philip Mark, and Michael Rabbat. Anomaly detection using proximity graph and pagerank algorithm. *IEEE Transactions on Information Forensics and Security*, 7(4):1288–1300, 2012.
- [448] Md Al Mehedi Hasan, Mohammed Nasser, Shamim Ahmad, and Khademul Islam Molla. Feature selection for intrusion detection using random forest. *Journal of information security*, 7(3):129–140, 2016.
- [449] Pradip Dhal and Chandrashekhar Azad. A comprehensive survey on feature selection in the various fields of machine learning. *Applied Intelligence*, 52(4):4543–4581, 2022.
- [450] Yihua Liao and V Rao Vemuri. Use of k-nearest neighbor classifier for intrusion detection. *Computers & security*, 21(5):439–448, 2002.
- [451] Fokrul Alom Mazarbhuiya, Mohammed Y AlZahrani, and Lilia Georgieva. Anomaly detection using agglomerative hierarchical clustering algorithm. In *Information Science and Applications 2018: ICISA 2018*, pages 475–484. Springer, 2019.
- [452] Hongpo Zhang, Lulu Huang, Chase Q Wu, and Zhanbo Li. An effective convolutional neural network based on smote and gaussian mixture model for intrusion detection in imbalanced dataset. *Computer Networks*, 177:107315, 2020.

# LIST OF PUBLICATIONS

## JOURNAL ARTICLES PUBLISHED



1.  **A. Presekal**, A. Ştefanov, V. S. Rajkumar and P. Palensky, "Attack Graph Model for Cyber-Physical Power Systems Using Hybrid Deep Learning," in *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 4007-4020, Sept. 2023, doi:10.1109/TSG.2023.3237011
2.  **A. Presekal**, A. Ştefanov, I. Semertzis and P. Palensky, "Spatio-Temporal Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems using Enhanced GC-LSTM," in *IEEE Transactions on Smart Grid*, vol. 16, no. 2, pp. 1654-1666, March 2025, doi:10.1109/TSG.2024.3474039
3.  **A. Presekal**, A. Ştefanov, V. S. Rajkumar, I. Semertzis and P. Palensky, "Advanced Persistent Threat Kill Chain for Cyber-Physical Power Systems," in *IEEE Access*, vol. 12, pp. 177746-177771, 2024, doi: 10.1109/ACCESS.2024.3507386.
4. **A. Presekal** et al., "Cyber Security of HVDC Systems: A Review of Cyber Threats, Defense, and Testbeds," in *IEEE Access*, vol. 12, pp. 165756-165773, 2024, doi: 10.1109/ACCESS.2024.3490605.
5. Y. Liu, H. Xie, **A. Presekal**, A. Stefanov and P. Palensky, "A GNN-Based Generative Model for Generating Synthetic Cyber-Physical Power System Topology," in *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4968-4971, Nov. 2023, doi:10.1109/TSG.2023.3304134.
6. V. S. Rajkumar, A. Ştefanov, **A. Presekal**, P. Palensky and J. L. R. Torres, "Cyber Attacks on Power Grids: Causes and Propagation of Cascading Failures," in *IEEE Access*, vol. 11, pp. 103154-103176, 2023, doi: 10.1109/ACCESS.2023.3317695.
7. I. Semertzis, A. Ştefanov, **A. Presekal**, B. Kruimer, J. L. R. Torres and P. Palensky, "Power System Stability Analysis From Cyber Attacks Perspective," in *IEEE Access*, vol. 12, pp. 113008-113035, 2024, doi:10.1109/ACCESS.2024.3443061.

## JOURNAL ARTICLES UNDER REVIEW


1.  **A. Presekal**, A. Ştefanov, I. Semertzis, H. Goyel, and P. Palensky, "Semi-Supervised Intrusion Detection System for Digital Substation using Traffic Signatures," in *IEEE Transactions on Smart Grid*, under review.

2. **A. Presekal**, A. Ştefanov, V. S. Rajkumar, I. Semertzis and P. Palensky, "Spatio-Temporal Correlation of APT in Cyber-Physical Power System," in **IEEE Communication Survey and Tutorial**, under review.
3. Y. Liu, **A. Presekal**, P. Palensky, and A. Ştefanov. "SibGen: A Hybrid Generator for Digital Siblings of Cyber-Physical Power Systems," in **IEEE Transactions on Smart Grid**, under review.
4. N. Cibir, N. Kabbara, **A. Presekal**, I. Semertzis, V. S. Rajkumar, H. Goyel, P. Palensky, A. Ştefanov, Cyber-Physical Power System Dataset for Cyber Security of Digital Substation, in **Nature Scientific Data**, under review.



## BOOK CHAPTERS

1.  **A. Presekal**, A. Ştefanov, V. S. Rajkumar, Kaikai Pan and P. Palensky, "Cyber Attacks on Power Systems," *IEEE-Wiley Press Book: Smart Cyber-Physical Power Systems, Volume 1: Fundamental Concepts, Challenges, and Solutions*, Feb. 2025.
2.  **A. Presekal**, A. Ştefanov, V. S. Rajkumar and P. Palensky, "Anomaly Detection and Mitigation in Cyber-Physical Power Systems based on Hybrid Deep Learning and Attack Graphs," *IEEE-Wiley Press Book: Smart Cyber-Physical Power Systems, Volume 1: Fundamental Concepts, Challenges, and Solutions*, Feb. 2025.

## CONFERENCE PAPERS

1.  **A. Presekal**, A. Ştefanov, V. S. Rajkumar and P. Palensky, "Cyber Forensic Analysis for Operational Technology using Graph-Based Deep Learning," in **Proc. 2023 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)**, Glasgow, United Kingdom, 2023, pp. 1-6, doi: 10.1109/SmartGridComm57358.2023.10333922.
2. I. Semertzis, H. Goyel, V. S. Rajkumar, **A. Presekal**, A. Ştefanov and P. Palensky, "Towards Real-Time Distinction of Power System Faults and Cyber Attacks on Digital Substations Using Cyber-Physical Event Correlation," in **Proc. 2024 12th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)**, Hong Kong, China, 2024, pp. 1-6, doi: 10.1109/MSCPES62135.2024.10542753.
3. V. Rajkumar, M. Tealane, A. Ştefanov, **A. Presekal** and P. Palensky, "Cyber Attacks on Power System Automation and Protection and Impact Analysis," in **Proc. 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)**, The Hague, Netherlands, 2020, pp. 247-254, doi: 10.1109/ISGT-Europe47291.2020.9248840.

Notes:

-  Publications included in this thesis.
-  Selected as one of the top 5 outstanding articles among over a thousand papers published in the *IEEE Transactions on Smart Grid* journal.

# ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to Prof. Peter Palensky and Dr. Alex Ștefanov for selecting me from among hundreds of applicants for the PhD position in *Cyber-Resilient Power Grid Operational Technologies* in 2019. Working with Peter as my promotor has been a truly enriching experience. His kindness and approachable nature have been greatly appreciated. Despite his extremely busy schedule, even brief conversations with him, often involving short yet profound questions, have consistently provided me with valuable insights and helped me clearly identify my research's core challenges. I am equally grateful to Alex, who is my daily supervisor and copromotor. Collaborating with him has been a remarkable learning curve for me. Alex is a highly dedicated, hardworking, and detail-oriented individual. He went above and beyond the typical PhD supervisor, generously dedicating his time to research discussions and thorough reviews of manuscript drafts. Through this collaborative process, I learned the importance of paying close attention to every detail of my work. It was great for me, with Alex and Peter's supervision, that almost all of my research publications were going for review and were published very smoothly.

Secondly, I would like to express my sincere gratitude to the members of my doctoral committee: Prof. Georgios Smaragdakis (TU Delft), Prof. Jose Maria Maza Ortega (University of Seville, Spain), Prof. Madeleine Gibescu (Utrecht University), Dr. Junho Hong (University of Michigan Dearborn, USA), and Dr. Guangya Yang (Technical University of Denmark). I am sincerely grateful for the time and effort you invested in reviewing my PhD thesis, as well as for your constructive and supportive feedback. The quality of my dissertation has been substantially improved as a direct result of your insightful comments. I would also like to thank Prof. Lucas J. van Vliet, the Dean of the Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS), TU Delft, for becoming the Chair of my defense. When it comes to the development of my academic career, I have high hopes that I will have the opportunity to learn and collaborate with all committee members and the chair in the future.

I would like to highlight the incredible journey I had as part of the *Cyber Resilient Power Grids (CRPG)* group, led by Alex. CRPG research group aims to go beyond conventional research group boundaries. Traditionally, PhD students work individually, and their projects often come to an end once they graduate, contributing to a 'graveyard' of discontinued research. In contrast, our group embraces an unconventional approach by fostering continuous collaboration, ensuring that research efforts remain active and impactful beyond individuals. It was an especially remarkable experience for me and highly beneficial for preparing my academic career. Being part of this diverse and dynamic team was a truly memorable experience. Our group brought together individuals from various educational backgrounds and countries of origin, creating a rich environment for collaboration and cultural exchange. One of the most enjoyable aspects of our team was sharing food and traditions from around the world. Thanks to our diversity, we had the chance to enjoy culinary delights from India, Indonesia, China, Greece, Italy, Korea, Japan, Egypt, and many more. I have a remarkable memory of our morning meetings starting promptly at 8:30 a.m., often before the rest of the building had even come to life. These morning meetings reflected our team's dedication, discipline, and strong spirit of collaboration.

I would like to especially thank Vetrivel Rajkumar, my senior, who guided me from my very first day in Delft. I'm grateful to Yigu Liu for generously sharing his knowledge on writing scientific publications and to Ioannis Semertzis for helping me expand the Mininet code and deepen my understanding of power systems. Many thanks to Himanshu Goyal for leading and supporting me



throughout the COCOON project and to Nicola Cibir for helping debug code, and for our insightful discussions on digital substations. I also deeply appreciate May Myat Thwe for our fruitful discussion on digital twins and federated learning, Mohamed Radwan for discussions on power system graphs, and Shafiulla Syed for our thought-provoking conversations and Snickers supply. A special thank you to Simona Renzaglia for her invaluable assistance in addressing ICT and coding issues within our group. I'm also thankful to former CRPG members Steven Tan, Sjors Hijgenaar, Raifa Akkaoui, Mohsen Jorjani, Ali Abedi, Ali Mollaiee, and Sayed Mehran Hashemian. My appreciation extends as well to our guest PhD colleagues Sho Cremers, Agrippina Mwangi, Nadine Kabbara, and Sina Hassani. Special thanks to my former office roommate, Kutay Bölüt, whose often fully occupied office whiteboard with equations inspired me and helped me understand the Gaussian Mixture Model for my paper review. I also enjoyed great discussions with Nidarshan Veera Kumar, Chenguang Wang, and Aihui Fu, especially during the challenging early days of COVID-19. Thank you also to my fellow Indonesian friends in the department, Badzlin, Bagas, and Pricilia, and to all other members of IEPG and the ESE Department who made my time at TU Delft truly special.

Thanks for the warmth and support from Indonesian community in Delft and The Netherlands, especially for Ilham, Arry, Antra, Ani, Aga, Kitty, Nasikun, Sebrion, Atindriyo, Albert, Wicak, Gilang, Rifki, Dhoni, Jeje, Budi, Mamin, Reni, Emy, Bela, Dita, Tommy, Aldy, Perdana, Mikhta, Gagas, Aryo, Fazlur, Panji, Endy, Adibah, Yana, Ifan, Lucas, Widana, Syakal, Satria, Jeri, Fenno, Hanif, Arya, Fajril, Ranggo, Wildan, Aziiz, Aprisia, Ben, Gandhi, and many more from PPI and KMD community in Delft. Thank you also for the van Orange group across The Netherlands, Tofan, Habib, Tizar, Siddik, Rihan, Luis, Hasrul, Khalif, Teguh, and Jamal.

I would also like to express my sincere gratitude to my former supervisors, Prof. Chris Hankin and Dr. Naranker Dulay (Imperial College London), as well as Prof. Kalamullah Ramli (Universitas Indonesia), for their invaluable support and recommendation during my PhD application process. My heartfelt thanks also go to Prof. Riri Fitri Sari (Universitas Indonesia), who has supported and inspired me since my undergraduate studies, particularly for giving me the motivation to pursue higher education abroad. I am deeply grateful to all the professors and lecturers from the Computer Engineering and the Department of Electrical Engineering at Universitas Indonesia for their continuous support, which made it possible for me to pursue my PhD at TU Delft.

I am particularly grateful to my wife, Syafira Fitri Auliya. Her unwavering support and detailed feedback strengthened my personal development, as well as my character development. Both of us pursuing our Ph.D. while caring for our never-ending-curious toddler is not easy, but we did it seamlessly. I am capable of attaining the present state with your support. Additionally, I am grateful to my parents for their unwavering support and for keeping me in their thoughts and prayers. I am also thankful to my brother for his support.

*Delft, May 2025*  
*Alfan Presekala*



## AUTHOR BIOGRAPHY



**Alfian Presekhal** was born in Blora, on May 20, 1991. He spent the majority of his childhood in Blora and Sragen, Jawa Tengah, Indonesia. He obtained a Bachelor of Engineering (B.Eng.) degree in Computer Engineering from the Department of Electrical Engineering, Universitas Indonesia. During his undergraduate studies, he achieved international recognition as the world's Top 15 teams at the Microsoft Imagine Cup 2011 in New York, USA, after won as the first place in the national level. He was also named as the runner-up of the Most Outstanding Student Award (*Mawapres*) at Universitas Indonesia in 2012. During this time, he also served as Vice President of the IEEE Student Branch at Universitas Indonesia. From 2012 to 2013, Alfian received a full scholarship from the Japan Student Services Organization to participate in a young scientist

exchange program at the Institute Science of Tokyo (formerly Tokyo Institute of Technology). During this program, he completed his bachelor thesis on secure communication protocols with integrity checking, marking the beginning of his deep interest in cyber security research.

After earning his B.Eng. degree in 2014, he pursued a Master's degree in Secure Software Systems at the Department of Computing, Imperial College London, UK, supported by a full scholarship from the Indonesian Government's *LPDP* program. His master's thesis focused on cryptographic applications for seamless wireless authentication. Since 2016, Alfian has been a junior lecturer and researcher at the Department of Electrical Engineering, Universitas Indonesia. During this period, he also served as director of business unit of the department, IEEE Indonesia section web manager, and obtained several professional certifications in computer networking and cyber security from CISCO, CompTIA, and EC-Council. In 2019, he has been appointed as a tenured Assistant Professor in the same department.

In January 2020, he joined the Department of Electrical Sustainable Energy at Delft University of Technology as a Ph.D. researcher. He collaborates with the Cyber Resilient Power Grids (CRPG) group within the Intelligent Electrical Power Grids (IEPG) section and has contributed to the development of the Control Room of the Future (CRoF) technology center at the Electrical Sustainable Power Laboratory (ESP Lab). His research focuses on advancing strategies to enhance the cyber security and resilience of power grids, protecting them from cyber-attacks that could result in widespread power outages. In March 2025, one of his first-authored paper was awarded as one of the top five outstanding papers over the past three years among over a thousand publications in the prestigious IEEE Transactions on Smart Grid. He is also actively involved as a researcher in Horizon Europe projects, including HVDC-WISE and COCOON (Cooperative Cyber Protection for Modern Power Grids). He is currently completing his Ph.D. with research focus on *Advanced Persistent Threat Detection and Correlation for Cyber-Physical Power Systems*. His primary research interests include cyber security, cyber-physical systems, operational technology, and the application of artificial intelligence.

41 64 76 61 6E 63 65 64 20 50 65 72 73 69 73 74  
65 6E 74 20 54 68 72 65 61 74 20 44 65 74 65 63  
74 69 6F 6E 20 61 6E 64 20 43 6F 72 72 65 6C 61  
74 69 6F 6E 20 66 6F 72 20 43 79 62 65 72 2D 50  
68 79 73 69 63 61 6C 20 50 6F 77 65 72 20 53 79  
73 74 65 6D 73 0A 45 6E 68 61 6E 63 69 6E 67 20  
52 65 73 69 6C 69 65 6E 63 65 20 6F 66 20 50 6F  
77 65 72 20 47 72 69 64 20 4F 70 65 72 61 74 69  
6F 6E 61 6C 20 54 65 63 68 6E 6F 6C 6F 67 69 65  
73 0A 61 20 44 6F 63 74 6F 72 61 6C 20 64 69 73  
73 65 72 74 61 74 69 6F 6E 20 62 79 20 41 6C 66  
61 6E 20 50 72 65 73 65 6B 61 6C 0A 44 65 6C 66  
74 20 55 6E 69 76 65 72 73 69 74 79 20 6F 66 20  
54 65 63 68 6E 6F 6C 6F 67 79 20 2D 20 32 31 73  
74 20 4D 61 79 20 32 30 32 35 00 00 00 00 00 00

