UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Electronic and Computer Science

Cross-layer Impact Analysis and a Novel Security Architecture for Cyber-Physical Power System

by

Xiaomeng Feng
ORCiD: 0000-0002-0821-1385

A thesis for the degree of
Doctor of Philosophy

Supervisor:
Dr. Leonardo Aniello

May 2025

To my family and friends for their unwavering support.

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Electronic and Computer Science

Doctor of Philosophy

Cross-layer Impact Analysis and a Novel Security Architecture for Cyber-Physical Power System

by Xiaomeng Feng

With the increasing interdependency of advanced information and communication technologies, power systems are undergoing a rapid transition to cyber-physical power systems (CPPS). This interdependency introduces cross-layer cyber threats that propagate their effects from the cyber layer to the physical layer, disrupting power system operations and potentially causing widespread blackouts. This research investigates two cyber challenges affecting CPPS security from two aspects: (1) degraded communication quality of service (QoS), which compromises data availability, and (2) false data injection attacks (FDIAs), which target data integrity.

Degraded QoS poses a critical cross-layer threat to CPPS by disrupting the timely and accurate transmission of control signals or measurements. Such disruptions undermine key functions such as frequency, voltage regulation in a cross-layer fashion. To address this, a novel technique is proposed, comprising (1) a CPPS model for quantitatively analyzing the cross-layer impact of resource allocation on physical states, specifically frequency, voltage, and (2) a multi-objective optimization framework to develop an optimal resource allocation strategy that minimizes disruptions to physical state regulation while enhancing QoS. The proposed strategy achieves a 13.74% reduction in frequency deviation and a 4.57% reduction in voltage deviation in the test system.

Another type of cyberattack, FDIAs, also pose critical cross-layer threats to CPPS by targeting data integrity. By compromising multiple measurement devices and cooperatively manipulating their measurements, FDIAs can construct stealthy attack vectors that evade residue-based bad data detection (BDD), mislead power system state estimation (PSSE), and ultimately cause market instability and economic losses. With the increasing integration of electricity markets and carbon trading markets, the cross-layer threats posed by FDIAs are further exacerbated due to additional vulnerabilities in energy price calculation mechanisms. Traditional approaches that assess economic risks based solely on electricity markets are no longer sufficient. This

research represents the first effort to extend the investigation of economic risks induced by FDIAs beyond the electricity market, incorporating the impacts of carbon emission costs. Simulations reveal an economic risk increase of up to 201.61 ($/MWh) on a certain transmission line in the PJM test system, compared with the traditional risks assessment only considering electricity costs.

Following the economic risk analysis of FDIA, this research further investigates mitigation strategies by disrupting its stealthiness, which depends on their capability of propagating across the system and manipulating a sufficient number of measurements. To address this, this research introduces the concept of zero-trust architecture (ZTA) and develops a novel security architecture based on a micro-segmentation technique. This technique divides measuring devices into finer security segments, restricting lateral attack propagation within the cyber layer while reducing FDIA stealthiness in the physical layer. To optimize the micro-segmentation strategy, a cyber-physical-BDD-enhancement-metric and a Graph Attention Network (GAT) combined with a reinforcement learning (RL) algorithm are proposed, evaluating the technique's effectiveness in enhancing BDD detection capability and mitigating the impact of FDIAs. Simulations demonstrate a significant improvement in the BDD detection rate against FDIAs, increasing from 5.23% to 94.02% with the proposed technique.

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published in the provided List of Publications in Section 1.6

Signed:..... Date:.....

Acknowledgements

I would like to express my heartfelt gratitude to everyone who supported me throughout my PhD journey. First and foremost, I am deeply thankful to my supervisors, Dr. Leonardo Aniello and Prof. Shiyan Hu. Your guidance, encouragement, and insightful feedback have been invaluable in helping me refine my ideas and elevate my research to new heights.

I am also incredibly grateful to my colleagues and friends in the Cyber Security research team. The collaborative environment, shared insights, and resources were essential to the successful completion of my work. A special thanks to Nawful for your thoughtful suggestions during the challenging times of the pandemic, and to BooJoong for your valuable feedback during my milestone viva sessions. I am also deeply thankful to my friends, Xin Liu, Xin Huang, Betul, and Enze Yi, for your continuous support and encouragement throughout this journey.

Finally, I dedicate this thesis to my family, especially my husband, Xinrui Liu, whose unwavering love, patience, and belief in me have been a constant source of strength. Your support has kept me motivated through the toughest moments. To my parents, your endless encouragement and love have been my greatest blessing, and I am forever grateful to have you in my life.

Contents

List of Figures

List of Tables

List of Abbreviations

ICT	Information and Communication Technology
CPPS	Cyber-Physical Power System
SCADA	Supervisory Control and Data Acquisition
IDS	Intrusion Detection System
MTD	Moving Target Defence
PSSE	Power System State Estimation
ML	Machine Learning
ZTA	Zero Trust Architecture
QoS	Quality of Service
FDIA	False Data Injection Attack
DDOS	Distributed Denial-of-Service
D2D	Device-to-Device Communication
DG	Distributed Generation
AGC	Automatic Generation Control
DA, RT	Day-Ahead and Real-Time Markets
LMP	Locational Marginal Pricing
C&E	Integrated Carbon and Electricity Markets
BDD	Bad Data Detection
PMU	Phasor Measurement Unit
AC, DC	Alternating Current, Direct Current
D-FACTS	Distributed Flexible AC Transmission Systems
MRAS	Model Reference Adaptive Search
AW-MRAS	Adaptive Weight Model Reference Adaptive Search
H-MADDPG	Hierarchical Multi-Agent Deep Deterministic Policy Gradient
GAT	Graph Attention Network
RL	Reinforcement Learning
CPS	Cyber-Physical Systems
NSF	National Science Foundation
WLS	Weighted Least Squares
OFDMA	Orthogonal Frequency Division Multiple Access
MDP	Markov Decision Process
DDPG	Deep Deterministic Policy Gradient

MADDPG	Multi-Agent Deep Deterministic Policy Gradient
ZT	Zero Trust
PEP	Policy Enforcement Point
PDP	Policy Decision Point
PA	Policy Administration
PE	Policy Enforcement
NIST	National Institute of Standards and Technology
EMS	Energy Management System
RTU	Remote Terminal Unit
NUE	Normal Users Equipment
SINR	signal-to-Interference-and-Noise Ratio
FIFO	First In, First Out
MSE	Mean Squared Error
SGD	Stochastic Gradient Descent
LSTM	Long Short-Term Memory
Adam	Adaptive Moment Estimation
SIS	Susceptible-Infected-Susceptible

List of Symbols

$\mathbf{x}$	State variables of the physical system (e.g., voltage, frequency)
$\mathbf{z}$	Measurement collected by sensors
$\hat{\mathbf{x}}_{\text{bad}}$	Estimated state variables obtained via PSSE
$\hat{\mathbf{z}}$	Estimated measurements
$\mathbf{a}$	Attack vector injected into the measurement z
$\mathbf{z}_a$	Compromised measurements
$\mathbf{H}$	Measurement matrix
$\mathbf{K}$	Gain matrix used in the BDD detection algorithm
$\mathbf{r}$	Residual vector for BDD
$\mathbf{e}$	Independent noise vector
$u[k], u'[k]$	Control input and compromised control input received by the system
$z[k], z'[k]$	Measurement and the measurement received by the system
$x[k]$	State vector of the system at time step k
$y[k]$	Output vector of the system at time step k
k	Discrete time step index
$\mathbf{W}$	Diagonal matrix
$\mathbf{I}$	Identity matrix
τ	Delay time
τ_r	Predetermined threshold for BDD
N	Total number of system buses
N_x	Total number of states for PSSE
$\mathcal{N}$	Set of system buses
M	Total number of measurements for $\mathbf{z}$
N_{dg}	Total number of distributed generators
$\mathcal{N}_{dg}$	Set of distributed generators
L	Total number of power lines
$\mathcal{L}$	Set of power lines
F_l	Transmission flow at line l
F_l^{max}, F_l^{min}	Upper and lower transmission capacity at line l
P_{g_i}	Generation of bus (MW)
$P_{g_i}^{max}, P_{g_i}^{min}$	Upper and lower generation capacity for bus i

$P_{g_i}^{DA}, P_{g_i}^{RT}$	Generation capacity in DA, RT market
$\hat{P}_{g_i}, \hat{P}_{d_i}$	Estimation of generated power and load
$\hat{E}^{co}$	Estimation of trading carbon emission rights
ΔE^{co}	Adjustment of trading carbon emission rights
$\Delta P_{g_i}^{DA}, \Delta P_{g_i}^{RT}$	Hypothetical incremental generation of i in DA, RT market
ΔP_{d_i}	Dispacth loads
$\lambda^{DA}, \lambda^{RT}$	LMP in DA, RT market
$C_i(\cdot)$	Generation cost function for bus i
γ^{DA}, γ^{RT}	Lagrange multiplier for power balance constraint in DA, RT market
R^{Profit}	Attack profits
v_a	Power trading volume
η_l, ζ_l	Lagrange multipliers for transmission capacity
μ_i^+, μ_i^-	Lagrange multipliers for generator capacity constraint
ι_j	Lagrange multipliers for carbon emission constraints
$\iota_{j,att}$	Lagrange multipliers for carbon emission constraints under FDIAs
$\eta_{l,att}^+, \zeta_{l,att}^-$	Lagrange multipliers for transmission capacity under FDIAs
Payoff_l^{DA},Payoff_l^{RT}	Attack profits in DA and RT market
I^{max}, I^{min}	Upper and lower price for carbon emission rights
E^{max}, E^{min}	Upper and lower limits of carbon emission rights
GSF_{l-i}	Generation shift factor matrix.
CSF_{l-i}	Carbon emission shift factor matrix.
L_{d2d}	Total number of D2D links
$\mathcal{L}_{d2d}$	Set of D2D links
I_{cha}	Total number of channels
$\mathcal{I}_{cha}$	Set of channels
$\xi_{l,i}$	Index of D2D link j assigned to channel i
$\mathbf{p}^d$	Set of transmit power of D2D links
$p_{l,i}^d$	Transmit power of D2D link j via channel i
p_i^c	Transmit power for NUE i on channel i
$p_{\text{max},l}^d$	Maximum transmit power of D2D link l
B^c	Bandwidth of each channel
$\gamma_{l,i}$	SINR for D2D l in channel i
$g_{l',l}^d$	Channel gain from D2D l' transmitter to D2D l receiver
$g_{i,l}^c$	Channel gain from NUE i transmitter to D2D l receiver
$d_{a,b}$	Distance between the transmitter a to receiver b
$P_{s,l}$	Probability of a successful packet transmission of D2D l
$\mathcal{R}_l$	Throughput for D2D link l
E_l	Expected throughput for D2D link l
T_l	Sum of expected throughput for D2D link l
$\mathbf{A}$	Adjacency matrix of communication network among devices
a_{ij}	Element in $\mathbf{A}$

δ	Nodal voltage phase angle
ω_{sys}	System nominal angular frequency of the microgrid
ω_z	Angular frequency of DG z
v_z	Voltage amplitude of DG z
m_ω, m_v	Droop coefficient of the primary control for frequency, voltage
Ω_z,Φ_z	Control input
S_l^w,S_l^v	Sent variables for frequency, voltage
B_l^w,B_l^v	Buffer variables for frequency, voltage
R_l^w,R_l^v	Receive variables for frequency, voltage
η^1,η^2	Low-pass time constant of power filter
L^p	Size of one transmission packet
$\mathcal{P}^c$	Total power consumption
$\mathcal{R}^c$	Total throughput
γ_i^c	SINR for NUE i on channel i
γ_{th}^i	SINR requirement for NUE i on channel i
P^*	Rated active power of DG j
Q^*	Rated reactive power of DG j
k_ω^1	Frequency recovery coefficient
k_v^1	Voltage magnitude recovery coefficient
k_ω^2	Active power sharing recovery coefficient
k_v^2	Reactive power sharing recovery coefficient
ω_z^*	Nominal angular frequency
$\mathcal{S}$	Set of states s in RL
$\mathcal{A}$	Set of actions α in RL
N_s	Total number of states s in RL
N_α	Total number of actions α in RL
$P(s_{t+1} \mid s_t, \alpha_t)$	Probability from s_t to s_{t+1} after taking α_t
$r(s_t, \alpha_t)$	Immediate reward from s_t after taking a_t
$\alpha_t = u(s_t)$	Deterministic policy function
$\alpha_t = \pi(\cdot \mid s_t)$	Stochastic policy function
$V^\pi(s)$	State-value function, represents expected cumulative reward from s based on π
$Q^\pi(s, \alpha)$	Action-value function, represents expected cumulative reward from s after taking a based on π
$\pi^*(\cdot)$	Optimal policy function
$\mathcal{U}(\cdot)$	Leader's objective function
$\mathcal{G}(\cdot)$	Follower's objective function
x^A	Leader's strategy & Solution to Leader's objective function
x^B	Follower's strategy & Solution to Follower's objective function
x^{A*}	Optimal leader's strategy & Optimal solution to leader's objective function

x^{B*}	Optimal follower's strategy & Optimal solution to follower's objective function
Ω^A, Ω^B	Set of leader's and follower's strategies
s^{RT}	Power system states in RT market
r^A, r^B	Rewards for leaders, followers
T	Incidence matrix for arbitrage attack
$\pi(\cdot)$	Policy function
K	Total number of security segments
$\mathbf{B}$	Security segment matrix
$\mathbf{B}_\mathrm{i}$	Indicator matrix for the i-th security segment
$\mathbf{a}'$	Reconstructed attack vector under micro-segmentation
$\mathcal{K}$	Set of security segments
$\mathcal{N}_k$	Set of measuring devices in the k-th security segment
$\mathcal{P}'$	Updated infection probability distribution under micro-segmentation
ρ^*	Steady-state infection probability vector for all nodes in the network
L_{cp}	Combined cyber-physical-BDD-Enhancement metric
L_p	Physical metric for evaluating micro-segmentation effectiveness
L_c	Cyber metric representing lateral spreading capability reduction
$\mathbf{G}^0$	Original network topology matrix
$\mathbf{G}'$	Updated network topology matrix after micro-segmentation
$\Delta\mathbf{G}$	Change in topology matrix caused by micro-segmentation
β	Infection rate parameter in the SIS dynamics model
$\boldsymbol{\mathcal{B}}^{\boldsymbol{\pi}}$	Segmentation scheme dividing devices into security segments
$\mathbf{v}$	Sequence of security group representations from the encoder
$\mathbf{p}$	Probability of segmenting devices into specific security groups
$\mathbf{1}$	Vector of ones
ρ_i, ρ_i'	Infection probability, updated infection probability after segmentation
$\mathbf{h}^A, \mathbf{h}^B$	Intermediate outputs of the actor-network
$\hat{\mathbf{h}}^A, \hat{\mathbf{h}}^B$	Normalized intermediate outputs of the actor-network
$\mathbf{a}^A, \mathbf{a}^B$	Actions output by the actor-network for attackers and operators
$\theta^A_{\mu,t}, \theta^B_{\mu,t}$	Actor-network parameters for attackers and operators at time t
$\theta^A_{q,t}, \theta^B_{q,t}$	Critic-network parameters for attackers and operators at time t
η^A, η^B	Learning rates for attackers and operators
d_K	Dimension of the intermediate output for attackers, $d_K = 3L$
d'_K	Dimension of the intermediate output for operators, $d'_K = 2N + 1$
$F_k^{\min}, F_k^{\max}$	Minimum and maximum transmission flow limits
$E_{\max,i}$	Maximum carbon emission rights for bus i
s_t^{RT}	States of power systems at time t
s_t^A, s_t^B	States for attackers, operators at time t
r^A, r^B	Rewards for attackers and operators
Q_t^A, Q_t^B	Q-value functions for attackers and operators

J_t^A, J_t^B	Objective functions for attackers and operators
L_{pre}	Loss function for pre-training
N_{pre}	Total number of samples in the pre-training dataset
$u_{sim}(\cdot)$	Simulator function outputting the optimal economic dispatch states
$\mathrm{y}, \hat{\mathrm{y}}$	Actual and predicted output from the LSTM network
a_{ini}^B	Initial action of the follower during pre-training
$\mathrm{Norm}(\cdot)$	Normalization operator
$\tanh(\cdot)$	Hyperbolic tangent activation function
$\|\cdot\|_2$	ℓ_2-norm operator, defined as $\|\|\mathbf{x}\|\|_2 = \sqrt{\sum_i x_i^2}$
$\circ$	Hadamard (element-wise) product operator
$\mathrm{diag}(\cdot)$	Diagonal matrix constructed from a vector
ρ	Parameter for updating target network weights
$\hat{v}_{a,t}^{RT}, \Delta v_{a,t}^{RT}$	Predicted and adjusted trading volume of attackers
$\circ$	Hadamard (element-wise) product operator
$\mathbb{E}(\cdot)$	Expected function

Chapter 1

Introduction

With the rapid advancement of Information and Communication Technology (ICT), traditional power systems have evolved further beyond smart grids into sophisticated Cyber-Physical Power Systems (CPPS), distinguished by their deep interdependence between cyber and physical domains. While this transformation has substantially enhanced the efficiency and intelligence of power infrastructures, it also introduces unforeseen and multifaceted security threats. In particular, certain threats, defined as cross-layer threats in this research, despite originating in the cyber layer, have the capability to propagate across layers and compromise physical components, leading to highly destructive impacts in CPPSs. This research explores both offensive and defensive dimensions, exploring the cross-layer security threats inherent to CPPS.

1.1 Motivation

The evolution of power systems has progressed from traditional infrastructures to smart grids and, ultimately, to CPPS, reflecting significant advancements in efficiency, technology, and security. To begin with, traditional power systems rely on centralized generation and unidirectional electricity flow, with security concerns primarily focused on physical infrastructure, such as equipment failures, vandalism, and natural disasters [1]. The transition to smart grids, driven by advancements in ICT, enabled traditional power systems to utilize bidirectional data flow, enhancing efficiency and reliability [2, 3]. However, this digital integration also introduced new vulnerabilities, including cyber threats and communication disruptions. As the interaction between cyber and power systems becomes increasingly interdependent, CPPS has evolved to replace the smart grid, representing a tightly integrated system that combines physical and digital infrastructures [4].

CPPS typically includes power equipment, sensor networks, actuators, communication units, and computing facilities, with two critical data flows: measurements and control commands. Measurements, such as voltage, active power, and current, are collected by sensors from power equipment and transmitted to the computation system. This system processes these measurements to generate control commands, which are subsequently dispatched to actuators for the regulation of power equipment [5, 6]. The secure and stable operation of CPPS heavily relies on the availability and integrity of these data flows. By leveraging advanced information technologies, including real-time data analytics, artificial intelligence, and autonomous decision-making, CPPS enhances system performance, resilience, and adaptability [7]. However, the growing interdependence between cyber and physical layers in CPPS also introduces significant security challenges. Unlike traditional power components, which are often insulated from external access, cyber components are interconnected with external networks. This connectivity makes CPPS vulnerable to cyberattacks, thereby resulting in undesired power communication interruptions and even blackouts [8].

For example, on 23 December 2015, a synchronized and coordinated cyber-attack compromised three Ukrainian regional electric power distribution companies, resulting in power outages affecting approximately 225,000 customers for several hours [9]. In this case, the malicious attackers delivered BlackEnergy 3 malware via spear phishing emails and were granted an initial access vector to the internal network (e.g., SCADA system). Afterward, the virtual private network credentials of the authorized users were successfully stolen for further penetration and destruction. These malicious events reveal the critical risks associated with cross-layer propagation within CPPS. Such cross-layer threats are particularly severe due to their capability to propagate to the physical layer, intensifying their impact. Specifically, certain threats in CPPS are no longer confined to traditional data security risks; their potential for cross-space propagation must also be considered, as they can cause failures in power equipment. For instance, degraded QoS within communication networks can disrupt the transmission of control signals, thereby impairing the operation of actuators in power plants. Therefore, addressing the issue of cross-space risk propagation has become an urgent priority.

However, when confronting cross-layer threats within CPPS, a universal solution remains absent. Each specific threat requires tailored defensive measures. Some researchers focus on data processing and analysis in the control center to enable effective state estimation and decision-making under emergency conditions. Other research highlights the protection of data transmissions, employing techniques such as Intrusion Detection Systems (IDS) [10], secure communication protocols, and network segmentation [11] to safeguard data transmission pathways and prevent unauthorized access. In addition, some advanced control algorithms [12, 13] are

proposed to maintain system stability and operational robustness while mitigating the impact of data corruption and malicious attacks. Moreover, researchers are developing strategies such as edge measurement device protection [14] and admittance-adjustment-based moving target defense (MTD) [15]. Although each existing defensive approach has shown feasibility and effectiveness in addressing specific threat scenarios, each approach has inherent limitations, leaving certain vulnerabilities unaddressed.

More critically, the complexity of components and functionalities in CPPS poses significant challenges for investigating cross-layer threats. The tight integration of numerous information devices (e.g., Internet of Things (IoT) devices [16]) and physical assets (e.g., renewable energy technologies [17]) has significantly heightened the complexity of CPPS. This large-scale expansion renders CPPS highly dynamic and nonlinear, complicating modeling and prediction efforts. Consequently, monitoring and real-time security decision-making face considerable challenges [18]. In addition, the cross-layer interdependence in CPPS enhances the integration between market mechanisms and power functionalities, such as economic dispatch. This integration, driven by mechanisms such as carbon trading and renewable energy credits alongside low-carbon policies, increases the economic sensitivity of power nodes to market fluctuations [19], thereby introducing novel economic vulnerabilities. Overall, the interdependence within CPPS expands the attack surface, complicates real-time monitoring and security responses, and introduces novel vulnerabilities from other operational functions, which pose significant challenges for future security measures.

Confronting the aforementioned challenges, several innovative techniques have gained attention to address security issues of CPPS based on their unique characteristics. Some researchers have initiated investigations into coordinated security defense strategies that simultaneously leverage integrated state data and feedback from both the cyber and physical layers [20]. These defense mechanisms are specifically designed to mitigate cross-space risk propagation. For instance, in power system state estimation (PSSE), operators not only detect and eliminate the falsified measurements that do not conform to physical system operation rules but also evaluate the potential impact of seemingly legitimate measurements on generator dispatch. This evaluation involves assessing whether such measurements might lead to system overloads or economic consequences by compromising the operation of generators in a cross-layer manner. In addition, machine learning (ML) techniques have been widely adopted for monitoring, threat detection, and system control within CPPS [21]. These techniques excel at processing vast amounts of data and can provide real-time defensive decision-making strategies following model training. Furthermore, the expanding attack surfaces and lateral threat propagation have increasingly challenged the traditional defensive capabilities of CPPS. To address these challenges, many researchers have turned to Zero Trust Architecture (ZTA) [22]

as a promising approach for cross-layer security defense. Although the application of ZTA in CPPS has only been preliminarily explored, its potential has been extensively demonstrated in other systems. ZTA hinders lateral movement within networks, ensuring that even if an attacker gains access to a specific resource, further access to additional resources is effectively restricted. These innovative defense strategies, which integrate coordinated approaches, machine learning, and zero trust principles, represent significant advancements in addressing the unique vulnerabilities of CPPS. However, current research remains in its infancy, with many of these approaches remaining under initial investigations.

1.2 Overview of Degraded QoS and False Data Injection Attacks

While many defensive approaches and innovative techniques have been developed to explore potential solutions for these vulnerabilities in CPPS from different aspects, two representative cross-layer threats, degraded QoS and FDIAs, remain critical cross-layer threats that require further investigation. This research aims to explore these two cross-layer threats and their corresponding countermeasures within CPPS. This section provides an overview of existing research on these two threats, beginning with a brief introduction of the mechanisms underlying degraded quality of service (QoS) and False data injection attacks (FDIAs), with further details in Section 2.2.2 and 2.2.3, respectively. It then offers a concise review of their impacts and associated defensive approaches.

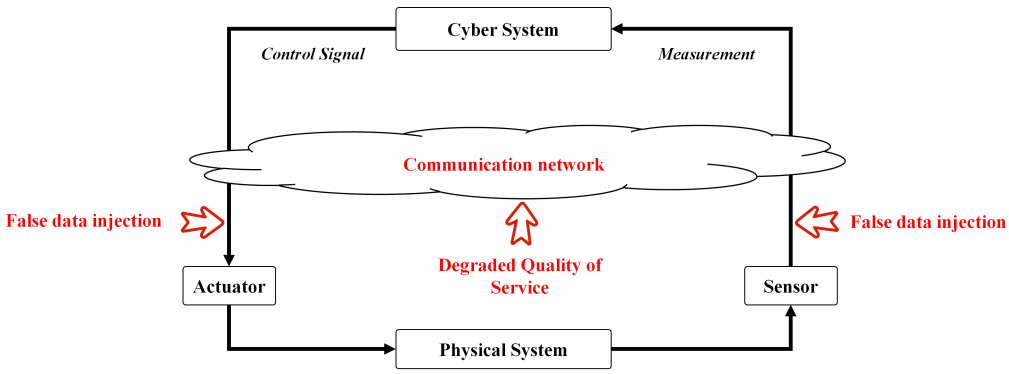


FIGURE 1.1: Simplification of data flow and cross-layer threats in CPPS.

Generally, as shown in Figure 1.1, the data flow in CPPS primarily comprises two key components: control signals, generated by the cyber system and transmitted through actuators to regulate the physical system, and measurements, collected by sensors and sent to the cyber system for decision-making.

1.2.1 Overview of Degraded Quality of Service

Degraded QoS typically refers to communication network issues, such as limited bandwidth, communication delays and packet loss, which can disrupt the transmission of measurements and control signals. Such disruptions can result in erroneous control signals, potentially causing cross-layer impacts on critical physical processes such as frequency regulation, voltage regulation, and load balancing, thereby compromising the overall stability and reliability of the power system. For example, during the 2015 Ukraine cyberattack, in addition to targeting the primary power systems, a DDoS was also launched to flood customer service lines, disrupting communication with citizens regarding the outages. This communication breakdown hindered the capability of the control center to respond promptly to customer inquiries, amplifying confusion during the crisis.

These cases underscore the potential threats posed by degraded QoS. Generally, QoS requirements for ensuring the secure and stable operation of CPPS vary according to specific functionalities and applications. Demand response management, for instance, can tolerate high latencies, ranging from 500 milliseconds to several minutes [23], as it does not require immediate adjustments. In contrast, real-time monitoring and control systems, such as SCADA, require lower latency and higher reliability to ensure immediate responses to system changes. For example, voltage and frequency regulation secondary control loops in microgrid, which are tasked with restoring equilibrium after disturbances, are particularly sensitive to delays. Even millisecond-level delays can disrupt the delicate balance of power systems, leading to instability [24]. Hence, the cross-space impacts caused by degraded QoS are particularly critical in high-demand systems, such as microgrids, where precise and timely communication is essential for maintaining stability.

To address the challenges posed by degraded QoS, some existing research has focused on designing advanced control algorithms aimed at improving CPPS resilience against communication latency [25]. However, the effectiveness of such algorithms is typically ensured only when QoS degradation is minor. When the degradation becomes severe and exceeds the thresholds of these algorithms, their effectiveness diminishes significantly, particularly under extremely constrained communication resources, such as excessively long time delays and extremely low throughput. For instance, when centralized base stations are compromised due to natural disasters or cyberattacks, bandwidth limitations can lead to data packet delays or losses, adversely affecting control algorithms and physical system stability, particularly in isolated microgrids. Beyond improving control algorithms, other research efforts have aimed at mitigating QoS degradation by enhancing communication network architectures. These approaches include optimal sampling time strategies [26], routing selection mechanisms [27], and communication resource allocation techniques [27, 28].

However, these approaches often prioritize enhancing data transmission performance while often neglecting the broader cross-layer impacts on the stability of physical systems. Advancements in wireless communications for microgrids, particularly D2D techniques, have shown promise for supporting remote microgrids by optimizing subcarrier allocation and transmit power [29]. However, only a few studies [28, 30] have integrated QoS improvements with physical system stability considerations in D2D communication strategies. Literature [28] incorporates the tolerable latency constraints of physical state variables as additional constraints when optimizing D2D resource allocation strategies. Alternatively, literature [30] pre-allocates optimal D2D communication resources for controllers to ensure reliable and timely data transmission.

1.2.2 Overview of False Data Injection Attack

FDIAs involve attackers injecting bad data into measurements. Such falsified measurements can be transferred to the cyber layer components, such as the control center, ultimately damaging the accuracy of the PSSE and leading to incorrect control signals. As one of the first known cyberattacks targeting SCADA systems, Stuxnet, discovered in 2010 [31], reveals the potential vulnerabilities that attackers can exploit to manipulate measurements. In the past, the false data can be detected by the residue-based traditional (BDD) referring to the constraints of the power flow equation. However, a type of completely stealthy (unobservable) FDIA has been first proposed in [32], which can bypass the residue-based BDD on the premise that attackers can gain sufficient information about network topology and branch parameters. While such stealthy FDIAs are effective at evading detection, they might not cause destructive cross-layer impacts on the CPPS. Therefore, researchers have further investigated the balance between the stealthiness of FDIAs with their potential damage.

A specialized form of FDIA, known as observability attacks, has been proposed [33] and further refined [34, 35], where strategically designed attack vectors render the control center incapable of distinguishing between unobservable states and malicious activities. Unlike purely stealthy attacks, observability attacks deliberately manipulate the PSSE in the control center, misleading it into making erroneous control decisions. For instance, by injecting false data, attackers can deceive the control center into perceiving an increase in certain loads. This misleading can result in incorrect dispatch instructions, causing transmission line overloads and jeopardizing the stability of the power system. The ability of observability attacks to actively mislead system operations underscores their significant threat, enabling economic exploitation or operational disruptions while evading detection.

In addition to the implication of FDIAs to damage stability, another type of FDIAs, profit-oriented FDIA, has been widely explored [36, 37, 38, 39], which can manipulate energy pricing and execute arbitrage strategies [40]. Literature [36] proposes the profit-oriented FDIA model, which demonstrates that attackers can make profits by using a buy-low and sell-high locational marginal price (LMP) strategy in the day-ahead (DA) and real-time (RT) market, respectively. Nowadays, the interdependence in CPPS enhances the integration between market mechanisms and economic dispatch, introducing novel economic vulnerabilities utilized by profit-oriented FDIAs. As global focus on environmental protection grows, carbon trading mechanisms have been incorporated into integrated C&E markets and LMP calculation mechanisms. The LMP calculation is transitioning from being solely based on fossil fuel costs in the generation system to considering charges tied to carbon emissions exceedances, also known as carbon-aware LMP [41, 42, 19, 43]. In this context, fluctuations in carbon prices can lead to variations in energy prices, i.e., LMP, thereby disrupting supply-demand dynamics and destabilizing electricity market operations. For instance, a sharp rise in carbon prices may compel high-emission power plants to reduce output, resulting in supply shortages and increased electricity costs. In addition, the inherent characteristics of electricity markets make them particularly susceptible to cyberattacks, such as data manipulation, which can undermine market integrity and propagate cascading disruptions into the carbon trading system. In extreme weather events, such as the 2019 Texas event, deregulation was implemented to encourage competition in response to a sharp increase in electricity demand, which provides arbitrage opportunities for attackers to make profits.

Although FDIAs are highly stealthy, their successful launch requires satisfying two key conditions: (i) sufficient knowledge of the structural network and parameters to construct a stealthy attack vector, and (ii) access to a sufficient number of measurements for collaborative manipulation. Due to the aforementioned cross-layer threats induced by FDIAs, various defensive approaches have been proposed for the detection, localization, and mitigation of FDIAs by obstructing these two preconditions, which can be categorized into several approaches [44]: securing measurement sensors [14], implementing moving target defense strategies [45], employing temporally and spatially relevant detection mechanisms, and adopting data-driven approaches. Although these methods can significantly enhance the accuracy of BDD and mitigate the impact of FDIAs, their practical implementation remains challenging due to the high costs associated with deploying Phasor Measurement Units (PMUs) and Distributed Flexible AC Transmission Systems (D-FACTS). Moreover, the design of such security techniques requires stringent conditions to eliminate the existence of stealthy attack vectors. For instance, literature [46] highlights that protecting all buses in the IEEE-14 system would necessitate

perturbing approximately 61.9% of transmission lines, which is economically infeasible given the substantial costs associated with D-FACTS deployment.

1.3 Research Gap

In a microgrid with a D2D communication network, a typical CPPS, degraded QoS due to limited communication resources can disrupt the stability of frequency and voltage regulation in a cross-layer fashion. As discussed in **overview 1.2.1**, most research has explored optimal D2D communication allocation strategies to mitigate degraded QoS on data transmission, with the primary objective of maximizing QoS. In this context, QoS metrics are insufficient, as the stability of the power system is also a critical factor for optimizing allocation strategies. A superior communication strategy regarding QoS does not always result in superior physical stability. Although some advanced works consider both QoS and its cross-layer impact on physical stability, they typically treat QoS as an intermediate objective when analyzing the cross-layer effects of communication resource allocation strategies on physical stability. Consequently, they either optimize communication reliability alone or focus solely on physical stability. **In summary, this research gap underscores the necessity for designing co-optimization strategies for D2D communication resource allocation that simultaneously enhances QoS and mitigates its cross-layer influences on frequency and voltage regulation, thereby improving overall system stability under conditions of degraded QoS.**

As discussed in **literature review 1.2.2**, the integration of electricity and carbon markets introduces new economic risks to CPPS, exposing the limitations of traditional vulnerability assessments against FDIAs. Specifically, the integration of carbon emission costs has significantly increased the complexity of Locational Marginal Pricing (LMP) calculations. This added complexity introduces constraints for managing carbon cost exceedances, rendering traditional node vulnerability assessments ineffective. Furthermore, it creates opportunities for attackers to inject false data, manipulate LMP calculations, disrupt carbon market transactions, and compromise the fairness and security of single-market systems. Existing research has primarily focused on vulnerabilities in traditional electricity markets, often overlooking the specific challenges that arise from carbon market integration. **The research gap lies in the fact that traditional vulnerability assessment approaches for FDIAs fail to identify the economic risks posed by the integration of the electricity and carbon markets, which requires a new vulnerability assessment approach to assess the economic risks of transmission lines and guide defensive approaches.**

As discussed in **literature review 1.2.2**, the principle of FDIA lies in designing a completely stealthy attack vector capable of bypassing the BDD. However,

constructing such attack vectors requires attackers to penetrate a sufficient number of measuring devices and manipulate their measurements cooperatively. Once attackers can lateral spreading across the measuring devices in the sensor network, most existing non-data driven defensive approaches such as securing measurement sensors and MTD are not always feasible. This limitation arises from two primary challenges. First, the implementation of these defense techniques is impractical due to the high cost and limited scalability of deploying PMUs and D-FACTS. Second, the effectiveness of these defense techniques relies on satisfying stringent constraints to obtain the objective that there exists no stealthy attack vectors. For example, literature [46] indicates that in order to protect all buses in IEEE-14 system, almost 61.9% transmission lines must be perturbed, which is unrealistic due to the high costs of D-FACTS. **The research gap lies in the urgent need for a scalable cybersecurity defense mechanism to disrupt the lateral movement of FDIAs, thereby undermining the stealthiness of FDIAs.**

1.4 Research Objective and Contribution

The interdependency between cyber and power components in CPPS has intensified the risks posed by cyber threats. While some threats may initially manifest as localized or minor anomalies at the cyber layer, their cross-layer propagation to the power system can lead to large-scale instability or even blackouts. To mitigate the cross-layer propagation of threats in CPPS, our research investigates two representative cross-layer risks: Degraded QoS and FDIAs, and develops corresponding countermeasures. Specifically, we address these challenges by focusing on three key research objectives: **1) developing a D2D communication resource allocation strategy to simultaneously mitigate QoS degradation and its cross-layer disruptions to physical state regulation**, **2) assessing the novel economic vulnerabilities of power nodes under FDIA induced by carbon constraints**, and **3) designing a micro-segmentation technique to enhance the detection rate of FDIAs and mitigate their cross-layer impacts**. In alignment with these objectives, this research makes the following contributions:

- **Contributions to Research Objective 1:**

  In a microgrid with D2D communication network, a typical CPPS, degraded QoS due to limited communication resources can disrupt the stability of frequency and voltage regulation in a cross-layer fashion. To address this challenge, a novel technique is proposed, comprising (1) a CPPS model for quantitatively analyzing the cross-layer impact of resource allocation on physical states, specifically frequency and voltage, and (2) a multi-objective optimization framework to develop an optimal resource allocation strategy that

minimizes disruptions to physical state regulation while enhancing QoS. Instead of analyzing all variables in the system, this CPPS model extracts interdependent cyber and physical states to reduce system complexity. Based on this model, a multi-objective optimization problem is formulated to identify the optimal communication resource allocation strategy, balancing QoS in communication and microgrid stability. To efficiently tackle this optimization problem, a weight-adjusted model reference adaptive search (AW-MRAS) algorithm is proposed, which significantly reduces the search space by leveraging the unique characteristics of the CPPS model. Compared to state-of-the-art strategies that only optimize QoS indices, the proposed strategy achieves a 13.74% and 4.57% reduction in frequency and voltage deviations, respectively, with only a minor compromise in QoS performance. In addition, the AW-MRAS algorithm demonstrates superior performance in balancing population diversity and convergence when compared to five other multi-objective optimization algorithms.

- **Contributions to Research Objective 2:**

  An attack model based on the Stackelberg game is proposed, making the first attempt to analyze the threats introduced by carbon emissions in the integrated carbon-electricity market. In this model, the attackers act as leaders, leveraging multi-transaction arbitrage and FDIAs to maximize attack profits, while the operators act as followers, calculating the LMP in response to the attackers' strategies. This interaction can also be viewed as an optimization problem, where the optimal attack strategies are derived using the proposed Hierarchical Multi-Agent Deep Deterministic Policy Gradient (H-MADDPG) algorithm. Building upon the previously identified most threatening attack strategy, a novel vulnerability assessment framework is developed for each power node, focusing on arbitrage opportunities driven by carbon cost considerations. The effectiveness of the proposed attack model and the H-MADDPG algorithm is evaluated against other algorithms using a modified version of the PJM test system. This framework provides crucial insights into the vulnerabilities of low-carbon initiatives and offers practical guidance for designing corresponding defensive measures.

- **Contributions to Research Objective 3:**

  A novel security architecture based on micro-segmentation technique is proposed, which restricts lateral attack propagation, and reduces the stealthiness of FDIAs. In addition, its effectiveness against FDIAs is proven under the direct current (DC) model. The optimization of micro-segmentation strategy is formulated as a multi-objective optimization problem, leveraging a cyber-physical-BDD-enhancement metric and a Graph Attention Network (GAT) combined with a reinforcement learning (RL) algorithm to improve its

effectiveness. Within this framework, a GAT-based feature extraction algorithm is introduced to capture cross-layer characteristics of both cyber and power components. Simulations demonstrate that the proposed micro-segmentation technique significantly enhances the detection rate of residual-based BDD against FDIAs, increasing from 5.23% to 94.02%, and highlights the effectiveness of the GAT+RL optimization algorithm, which considerably outperforms state-of-the-art algorithms in computing time while maintaining solution quality.

1.5 Structure of Thesis

The rest of this thesis is organized as follows.

- Chapter 2 provides the foundational background for this research. It begins by introducing the core concepts and hierarchical structure of CPPS, detailing the interconnections among its layers and components. Subsequently, this chapter explores the implications of cross-layer threats in CPPS and examines the existing fundamental defensive approaches. In addition, it illustrates the mechanisms and implications of two representative cross-layer threats within the scope of this research: degraded QoS and FDIAs, respectively. Finally, it illustrates the techniques adopted in this research, including ZTA and machine learning methodologies.
- Chapter 3 presents an optimal D2D communication resource allocation strategy designed to mitigate the impact of QoS and its cross-layer effects on microgrid stability. Section 3.2 introduces the system model. Section 3.3 presents the proposed CPPS model, focusing on the cross-layer impacts of D2D communication allocation strategies on microgrid stability. A joint multi-objective optimization problem is formulated to minimize degraded QoS disruptions and their cross-layer effects on frequency and voltage regulation. Section 3.4 details the AW-MRAS algorithm developed to optimize the proposed D2D strategy. Section 3.5 provides the simulation results and analysis, while Section 3.6 concludes the chapter by summarizing the key findings. The content of this chapter corresponds to the work detailed in **Publication 4**.
- Chapter 4 designs a novel economic vulnerability assessment framework for power nodes, incorporating the unique vulnerabilities induced by low-carbon policies. Section 4.2 introduces the profit-oriented FDIA and its properties. Section 4.3 discusses the Stackelberg-based FDIA model and its components, including multi-transaction strategies and LMP with carbon considerations. Section 4.4 introduces the optimization of the Stackelberg-Game-Based attack

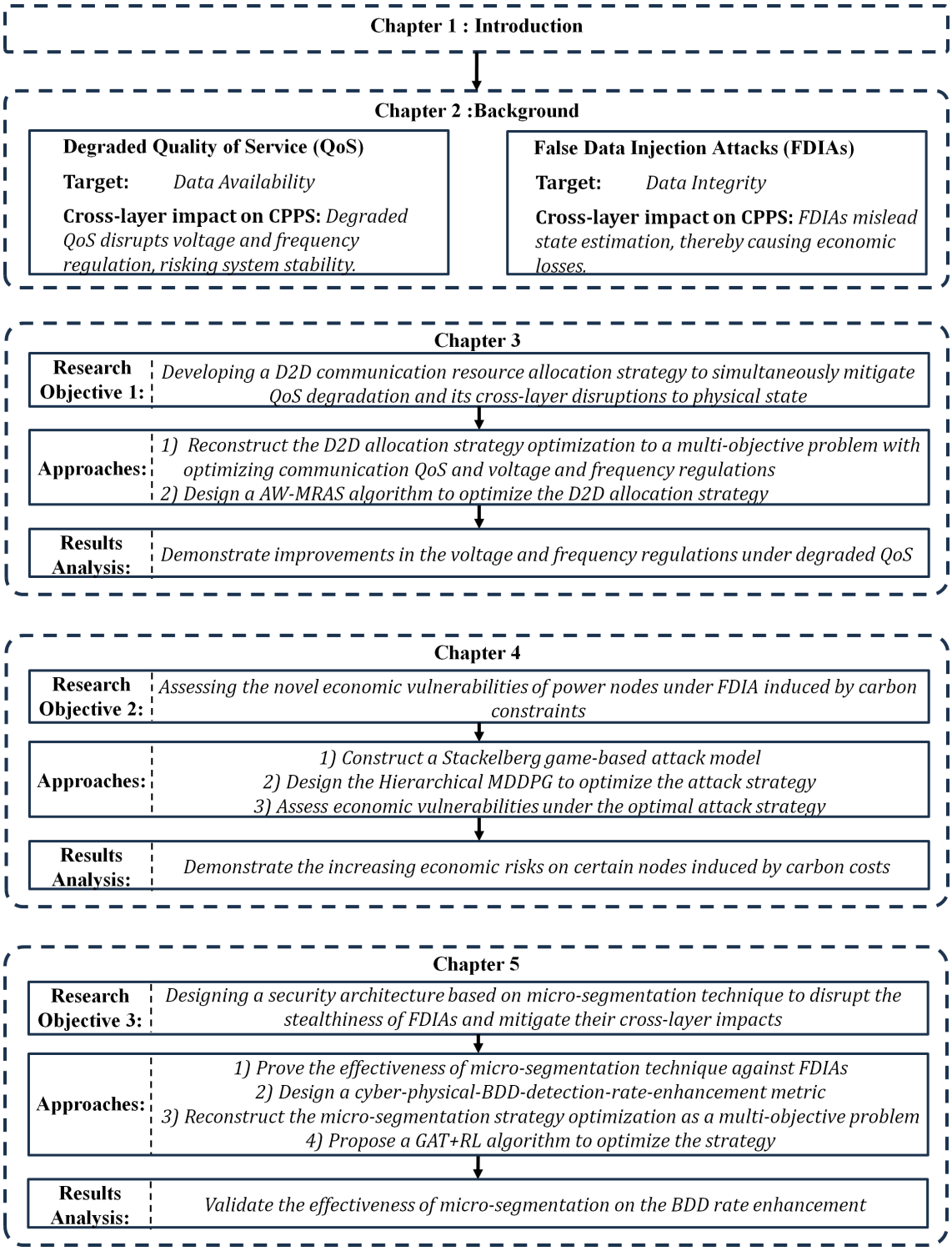


FIGURE 1.2: Structure of the thesis

using hierarchical multi-agent reinforcement learning, focusing on LMP pre-training and constrained action search. Section 4.5 examines the economic vulnerability of transmission lines under the optimal attack strategy. Section 4.6 presents the simulation results, including the evaluation of the proposed attack strategy, the validation of the H-MADDPG algorithm, and the analysis of

arbitrage opportunities. Finally, Section 4.7 concludes the chapter by highlighting the main findings. The content of this chapter corresponds to the work detailed in **Publication 2 and Publication 3**.

- Chapter 5 proposes an innovative micro-segmentation technique to enhance the detection rate of residual-based BDD against FDIAs. Section 5.2 introduces the PSSE and the stealthiness of FDIAs. Section 5.3 explores the implementation of the proposed micro-segmentation technique in CPPS and demonstrates its effectiveness against FDIAs. Section 5.4 formulates the optimal micro-segmentation strategy as a combinatorial optimization problem and introduces a cyber-physical-BDD-enhancement-metric to evaluate the impact of the strategy on lateral spreading capability and BDD detection probability. Section 5.5 describes the GAT+RL algorithm developed to solve this optimization problem. Section 5.6 provides simulation results and analysis, while Section 5.7 concludes the chapter with a summary of the findings. The content of this chapter corresponds to the work detailed in **Publication 5 and 1**.
- Chapter 6 outlines the conclusions and future works.

1.6 List of Publications

1.6.1 Intended Submissions

1. **Xiaomeng Feng**, Leonardo Aniello, and Shiyan Hu. "Zero Trust Architecture to Mitigate False Data Injection Attack in Cyber-Physical Power Systems." *Intended to submit to IEEE Transactions on Industrial Informatics*. **[Chapter 5]**
2. **Xiaomeng Feng**, Leonardo Aniello, and Shiyan Hu. "A Vulnerability Assessment of Economic Risks in Carbon-Electricity Integrated Trading Systems." *Intended to submit to Applied Energy*. **[Chapter 4]**

1.6.2 Published

3. **Xiaomeng Feng**, Leonardo Aniello, and Shiyan Hu. "Multi-Transaction and Carbon-Aware Strategies for Profit-Oriented FDIAs on Electricity Trading Systems." *2024 IEEE Power & Energy Society General Meeting (PESGM)*, Seattle, WA, USA, 2024, pp. 1–5, doi: 10.1109/PESGM51994.2024.10688834. **[Chapter 4]**
4. **Xiaomeng Feng**, Shiyan Hu. "Co-Optimization of Microgrid Secondary Control and Communication QoS: A Cross-Layer Perspective in Cyber–Physical System." *IEEE Transactions on Industrial Informatics*, doi: 10.1109/TII.2024.3452755. **(2024 Impact Factor: 11.7) [Chapter 3]**

5. **Xiaomeng Feng**, Shiyan Hu. "Cyber-Physical Zero Trust Architecture for Industrial Cyber-Physical Systems." *IEEE Transactions on Industrial Cyber-Physical Systems*, vol. 1, pp. 394–405, 2023, doi: 10.1109/TICPS.2023.3333850. **[Chapter 5]**

6. **Xiaomeng Feng**, Shiyan Hu. "Zero Trust Architecture for Cyber-Physical Power System Security Based on Machine Learning." *AI Embedded Assurance for Cyber Systems*, Springer, 2023.doi: https://doi.org/10.1007/978-3-031-42637-7_5. **[Chapter 5]**

7. **Xiaomeng Feng**, Yang Liu, and Shiyan Hu. "Machine Learning for Cyber-Physical Power System Security." *Machine Learning for Embedded System Security*, Springer, 2022. doi:10.1007/978-3-030-94178-9_4. **[Chapter 4]**

1.6.3 Contributing Author Works

8. **Qiuye Sun**, **Bingyu Wang**, **Xiaomeng Feng** *et al.* "Small-signal stability and robustness analysis for microgrids under time-constrained DoS attacks and a mitigation adaptive secondary control method." *Science China Information Sciences*, vol. 65, 162202, 2022. doi: 10.1007/s11432-021-3290-3. (27 citations, **(2024 Impact Factor: 7.275)**) **[Preliminary work for subsection 3.2.2]**

Chapter 2

Background

2.1 Basics of Cyber-Physical Power System

2.1.1 Concept of Cyber-Physical Power System

Cyber-Physical Systems (CPS) are engineered systems that integrate computational elements with physical processes for seamless interaction, emphasizing their interdependence. This term originated around 2006 by Helen Gill at the National Science Foundation (NSF) in the United States [47]. CPSs have unique features including real-time monitoring, feedback loops, robust connectivity, and autonomous decision-making using sensors and actuators.

The transition from general Cyber-Physical Systems to CPPS represents a specialized application of CPS principles within the energy sector. This evolution involves integrating traditional power system components, such as generators and transformers, with advanced computational technologies to enhance monitoring, control, and optimization capabilities [4]. While the deep interconnection of digital networks facilitates secure and stable power system operation, this interdependency makes CPPS more vulnerable to cyber threats. Therefore, cybersecurity becomes a critical concern for ensuring the reliability and resilience of modern power systems.

2.1.2 Hierarchical Structure for Cyber-Physical Power System

Generally, the cyber-physical power model is composed of the physical power layer and the corresponding cyber layer, including the computation layer, the communication layer and the power control layer. The data flow and the interconnection between different layers are displayed in Figure 2.1 [4].

The physical power layer encompasses key power infrastructure and components, including systems for generation, transmission, and distribution, as well as distributed generators, transformers, power lines, loads, and breakers. To model these power components and network, many existing works have discussed mathematical equations and accurate simulations [16, 48, 6].

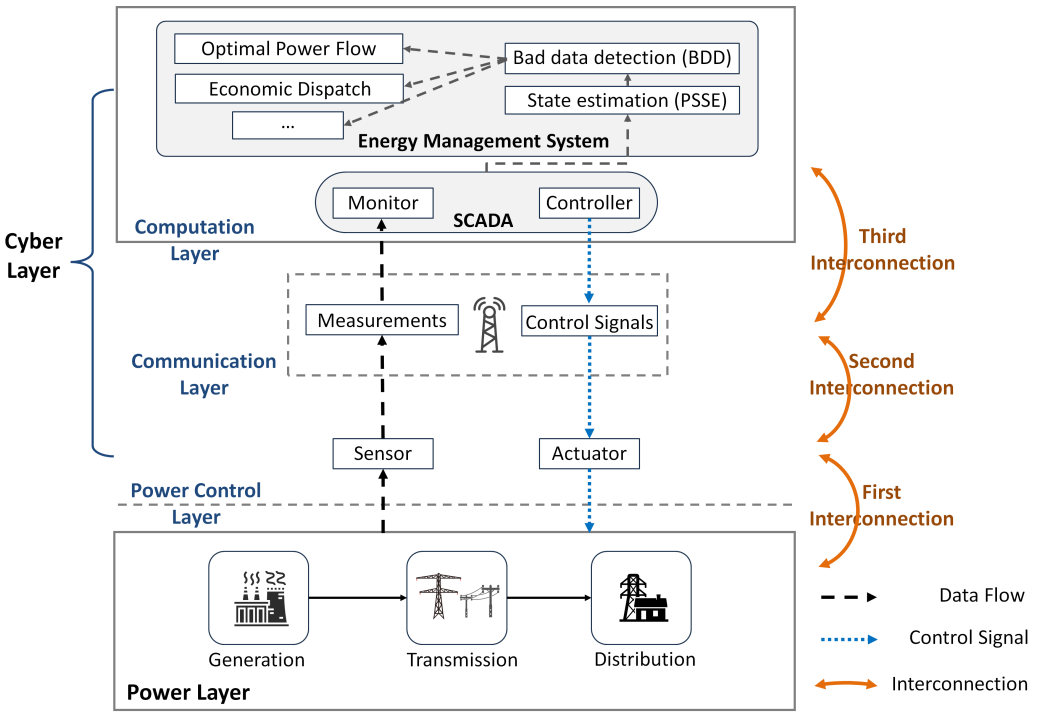


FIGURE 2.1: Hierarchical structure for cyber-physical power system.

The power control layer, including measuring devices, local controllers, sensors, actuators and other terminal devices with similar functionalities, is responsible for direct interaction. Measuring devices, such as PMU and phasor data concentrators (PDCs), sample real-time power measurements such as voltage, current, circuit breakers status and frequency. Correspondingly, after receiving the control signals from the control center, the local controllers such as generator controllers and breakers can directly launch the control signals to regulate the state operating of the power components.

The communication layer is responsible for exchanging information between the control layer and the computation layer, which is composed of devices such as switches, routers, and so on. Generally, measuring devices collect real-time power measurements and then send them into the EMS located in the power computation layer through the communication layer. Correspondingly, after selecting and analyzing these measurements, the EMS can estimate the operating state of the power system and launch control signals into the local controllers transmitted by the communication layer. Such information exchange can be wired or wireless depending

on requirements. The quality of service is the main index to evaluate the performance of the power communication infrastructures, which can also impact the physical system in a cross-layer fashion.

The computation layer can be viewed as the brain of the framework of CPPS, which is responsible for monitoring, controlling and making decisions to ensure the security and stability of the operating of the power system, including SCADA system, and the EMS. Through receiving numerous measuring data for the control layer via the communication infrastructure, the computation layer decides on which control signal to be executed in the following time step.

2.2 Cross-Layer Threats in Cyber-Physical Power Systems

CPPSs are characterized by a high degree of interdependence between their cyber and physical layers, wherein impacts in one layer can significantly cascade to the other. This interdependence creates novel cross-layer threats, which refer to vulnerabilities that originate in the cyber layer and propagate to the physical layer, resulting in blackouts and substantial economic losses.

In this section, we classify cross-layer threats in CPPSs and explore the concepts and impacts of two representative cross-layer threats: degraded QoS and FDIAs, providing a foundational understanding for the subsequent chapters.

2.2.1 Classification of Cross-layer Threats

These threats are analyzed based on their origins into two primary categories [49]: non-malicious threats, such as performance degradation caused by communication delays or data loss, which arise due to unintentional system inefficiencies, and malicious threats, including sophisticated and targeted attacks designed to exploit system vulnerabilities.

2.2.1.1 Non-Malicious Threats

Non-malicious threats in CPPS arise from natural causes, system limitations, and operational factors rather than intentional attacks. These threats pose significant challenges to system reliability and performance, which potentially result in gradual performance degradation or unexpected failures triggered by various internal and external factors. Human Errors can occur at any stage of system operation. For instance, an operator may incorrectly configure system parameters or input erroneous data, resulting in incorrect control signals that disrupt the functionality of the power

system. Similarly, maintenance personnel may inadvertently damage critical infrastructure during repair or upgrade processes, leading to system downtime and operational interruptions.

In addition to human errors, degraded QoS in communication networks can pose significant cross-layer threats. As highlighted in [50], degraded QoS compromises the ability to monitor and control corresponding power devices effectively. Although power devices might operate normally in the short term based on predefined electrical operating rules, they are prone to failure under abnormal conditions such as overload.

2.2.1.2 Malicious Threats

Generally, the existing works on the cyberattacks in CPPS mainly focus on eavesdropping, data manipulation and traffic abnormity [51, 52]. Different from other systems, the power system has its own characteristics and physical constraints, upon which the malicious attackers will adjust existing attacking methods. This chapter illustrates three cyberattacks which are widely analyzed in the research field in the power system, to emphasize the process which these cyberattacks impact the power components in a cross-layer fashion. However, the process and mechanism of the cyberattacks are not discussed; they are beyond the scope of this research.

- Eavesdropping attack

  Eavesdropping attack is a type of passive attack, in which the malicious attackers observe the confidential information of component nodes, such as location, public key, private key and password. Subsequently, the observed information can be utilized by the succeeding attacks to penetrate to the terminal measuring devices and control devices in the power control layer, thereby inducing the abnormality or even deteriorations in CPPS.

- False data injection attack

  The principle of FDIA is to manipulate the data collected/stored at the terminal power components or the control center to impact the state estimation. For example, it can cause estimation errors of power system states through manipulating bus load and line flow measurements simultaneously. Consequently, the control center may launch the falsified control signals, which further propagates the falsified load distribution strategies in a cascaded fashion [53]. The traditional data modification methods in cyber systems mainly focus on deceiving the detection mechanisms. In contrast, the FDI attack in CPPS targets on state estimation, thereby causing damage to the power grid. Due to this reason, cyberattack are identified as a critical underlying threat to the power system.

- Distributed denial of service (DDoS)

  In CPPS, DDoS interferes the communication functionalities of cyber devices through broadcasting invalid packages or jamming communication channels with random noise. In the power system with hierarchical control architecture, the timeliness of information exchange should be rigorously guaranteed to ensure the performance of the controllers. In this case, the objective of DDoS attack is to affect the timeliness and success rate of data exchange of the cyber components, which subsequently induces time delay [54] and packet loss [55]. Concrete examples of DDoS are depicted in [56].

2.2.2 Basics of Degraded Quality of Service: Mechanisms and Impacts

This section briefly outlines the potential threats to data flows within CPPS posed by degraded QoS to and examines how these compromised data flows can propagate across layers, ultimately impacting the physical system.

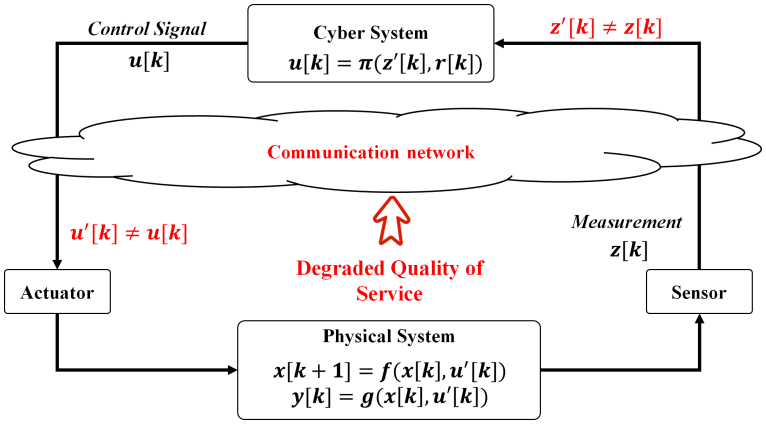


FIGURE 2.2: A simplified illustration of QoS impact on CPPS.

2.2.2.1 Mechanisms of Degraded Quality of Service

The critical data flow in CPPS is discussed in Fig. 2.2, with the physical system described by a difference equation and the cyber system mainly based on the control algorithm and communication network to transmit the data flows. The physical system is expressed using the following difference equations:

$$\begin{aligned} x[k+1] &= f\left(x[k], u'[k]\right), \\ y[k] &= g\left(x[k], u'[k]\right), \end{aligned} \tag{2.1}$$

where $x[k]$ represents the state vector of the physical system at time step k, such as voltage or frequency. $u[k]$ denotes the original control signal, while $u'[k]$ denotes the control signal received by the physical system after transmission through the

communication network, which may be corrupted due to degraded QoS. $y[k]$ represents the output vectors, which is derived from the state vector $x[k]$ and $u'[k]$. In this subsection, $z[k]$ represents the measurement, while and $z'[k]$ denotes the measurement received by the cyber system after transmission through the communication network, which may be corrupted due to degraded QoS. The functions $f(\cdot)$ and $g(\cdot)$ describe the functions of the system dynamics and the relationship between the system state and output.

The control algorithm, implemented within the cyber system, is defined as

$$u[k] = \pi(z'[k], r[k]), \tag{2.2}$$

where $u[k]$ is the control signal generated by the controller at time step k. This signal is calculated based on the current system state $x[k]$ and the reference signal $r[k]$, which represents the desired target state, such as a reference voltage or frequency level. The function $\pi(\cdot)$ represents the control policy that determines the high-performance control action required to maintain system stability and achieve the desired performance.

The communication network is responsible for transmitting control signals and measurements. However, as illustrated in Fig. 2.2, the degraded QoS might compromise the reliability of the communication network by introducing packet loss, data delays, and other related disruptions. Consequently, deviations arise in the transmitted signals and measurements. Specifically, the control signal received by the physical system, $u'[k]$, deviates from the intended signal, $u[k]$, such that $u'[k] \neq u[k]$. Similarly, the measurement received by the controller, $z'[k]$, differs from the actual output, $z[k]$, resulting in $z'[k] \neq z[k]$.

2.2.2.2 Impacts of Degraded Quality of Service

The interaction between the physical and cyber systems highly relies on ensuring accurate and effective control signals[57]. As previously discussed, deviations in measurement $z'[k]$-$z[k]$ might disrupt the feedback loop and impair the controller's ability to generate accurate control signals. Therefore, the inaccurate $u'[k]$ can damage the secure and stable operations of physical systems.

For example, the primary impact of degraded QoS is manifested in communication delays. The delay in measurements can be expressed as

$$z'[k] = z[t - \tau], \tag{2.3}$$

where $z'[k]$ represents the delayed measurements transferred via the communication network. τ represents the delay time. Such deviations in $z'[k]$ can misestimate the

system state in the cyber system, thereby introducing additional errors into the generated control signals, denoted as $u'[k]$-$u[k]$. Furthermore, the degraded QoS in the communication network can also delay the transmission of control signals, which can be

$$u'[k] = u[k - \tau], \tag{2.4}$$

where $u'[k]$ denotes the control signal received by the physical system compromised by time delays. As a result, $u'[k]$ leads to inaccurate adjustments by the actuator to the physical system. These issues compromise the stability and performance of the physical system, posing a significant cross-layer threat to its reliable operation.

2.2.3 Basics of False Data Injection Attacks: Mechanisms and Impacts

This section briefly outlines the potential threats to data flows within CPPS posed by FDIAs and examines how these compromised data flows can propagate across layers, ultimately impacting the physical system.

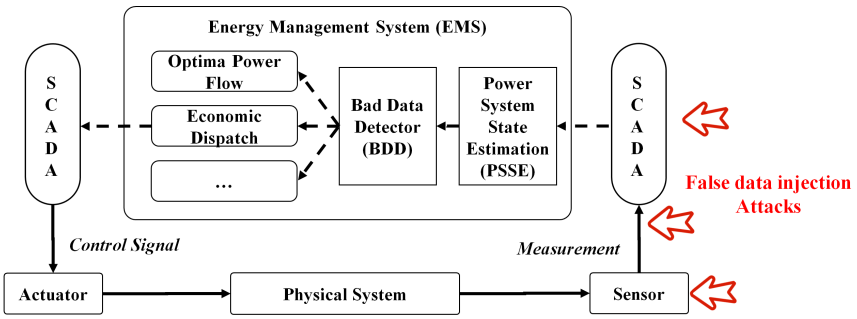


FIGURE 2.3: A simplified illustration of FDIAs impact on CPPS.

Although the deep interaction between the industrial infrastructures and the information and communication technology considerably facilitates the control and monitor of CPPS, it still poses new attack surfaces for intruders. Especially for the edge physical devices, they represent highly attractive targets to attacks due to the limited defense resources and complex interconnections. In this scenario, the FDIA, which aims to penetrate the edge measuring devices and falsify the measurements, has drawn much attention. Generally, power measurements are utilized to estimate the states of the power system. Through falsifying the measurements, FDIA can cause errors in the PSSE, and thus mislead the grid operators to take the actions.

2.2.3.1 Mechanisms of False Data Injection Attack

- DC State Estimation and bad data detection

  Power system state estimation aims to estimate state variables with the measurements collected by the SCADA, which is a significant function in

maintaining the secure operation of the power system. Define that N' power system state variables $\mathbf{x} = (x_1, x_2, \cdots, x_{N'})^T$ are evaluated M ($N' \ll M$) measurements $\mathbf{z} = (z_1, z_2, \cdots, z_M)^T$, where $N' = N - 1$. It yields $\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}$, where $\mathbf{H} \in \mathbb{R}^{M \times N'}$ is the measurement matrix and $\mathbf{e} \in \mathbb{R}^M$ is the independent noise. When $\boldsymbol{e}$ is the independent noise, the estimated state is $\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z} \triangleq \mathbf{K}\mathbf{z}$. $\mathbf{W} = \text{diag}\left(\sigma_i^{-2}, 0\right)$ is the diagonal matrix, where σ_i^2 is the variance of the measurement errors associated with the i-th measurement. In this DC model, the state variables $\mathbf{x}$ are $\{\theta_1, \theta_2, \ldots, \theta_{N'}\}$ and the estimated states $\hat{\mathbf{x}} = \{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{N'}\}$.

To localize and identify the abnormal measurements, the BDD based on measurement residual is widely utilized to localize and remove the abnormal measurements. The residual is $\mathbf{r} = (\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{z}$, where $\mathbf{I}$ is the identity matrix. To bypass the BDD, the measurements are required to satisfy the following condition:

$$\|\mathbf{r}\|_2 = \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\| < \tau_r, \tag{2.5}$$

where τ is the predetermined threshold. $||\cdot||_2$ is the $L2$ norm. To simplify notation, $||\cdot||_2$ is replaced by $||\cdot||$ in the following.

- False Data Injection Attack

  An emerging attack model, false data injection has been widely explored in many existing work, which can bypass the BDD through constructing a type of completely stealthy attack vector [32]. This type of BDD-bypassing attack vector is designed as $\mathbf{a} = \mathbf{H}\mathbf{c}$, where $\mathbf{c} \in \mathbb{R}^{N'}$ is an arbitrary vector. Without regarding to the measurement noise $\mathbf{e}$, such an attacker vector can satisfy the condition in Eqn. (2.5) in the following form [32]:

$$\begin{aligned} \|\mathbf{r}\| &= \|\mathbf{z}_\mathbf{a} - \mathbf{H}\hat{\mathbf{x}}_{\text{bad}}\| \\ &= \left\|\mathbf{z} + \mathbf{a} - \mathbf{H}\left(\hat{\mathbf{x}} + \left(\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{W}\mathbf{a}\right)\right\| \\ &= \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}} + (\mathbf{H}\mathbf{c} - \mathbf{H}\mathbf{c})\| = \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\| \leq \tau_r, \end{aligned} \tag{2.6}$$

  where $\mathbf{z}_\text{a} = \mathbf{z} + \mathbf{a}$ is compromised measurement and $\hat{\mathbf{x}}_{\text{bad}}$ is the estimated results calculated from $\mathbf{z}_\text{a}$.

  Lemma 2.1. *[32] If the attack vector can satisfy $\mathbf{a} \in \text{col}(\mathbf{H})$, and then it can bypass the BDD.*

2.2.3.2 Profit-oriented FDIA and its Economic Impact

- Two-Settlement electricity market

  Two-settlement electricity markets are organized by independent system operators (ISOs) or regional transmission organizations (RTOs), which consist of

DA and RT markets [58]. Generally, the DA unit dispatches and DA LMPs are computed by solving unit commitment and DA economic dispatch. With these DA LMPs, ISOs facilitate the aggregation of offers and bids between generators and load aggregators, thereby economically clearing the market.

DA Market Model:

$$\begin{aligned}
&\underset{P_{g_i}^{DA}}{\text{minimize}} \quad \sum_{i=1}^{N} C_i(P_{g_i}^{DA}) \\
&\text{subject to } \sum_{i=1}^{N} P_{g_i}^{DA} = \sum_{j=1}^{N} P_{d_j}^{DA}, \\
&P_{g_i}^{\max} \le P_{g_i}^{DA} \le P_{g_i}^{\min}, \quad \forall i \in \mathcal{N}, \\
&F_l^{\min} \le \sum_{i=1}^{N} \mathrm{GSF}_{li}(P_{g_i}^{DA} - P_{d_i}^{DA}) \le F_l^{\max}, \forall l \in \mathcal{L}.
\end{aligned} \tag{2.7}$$

To solve the optimal generator output P_g^{DA}, Eqn. (2.7) is reformulated into an unconstrained Lagrangian relaxation problem to facilitate the solution process. The DA LMP at bus i models the effect of incremental load on the fuel cost function [58].

RT Market Model:

Due to unforeseen load variations, the real-time power states P_g^{RT}, P_d^{RT} diverge from the optimal estimated states $\hat{P}g$, $\hat{P}d$ based on the RT economic dispatch. The purpose of the RT market is to provide compensatory adjustments for the differences in optimal estimation between DA and RT markets, as formulated by

$$\underset{\Delta P_{g_i}^{RT}}{\text{minimize}} \quad \sum_{i}^{N} C_i(\Delta P_{g_i}^{RT} + \hat{P}_{g_i}) \tag{2.8}$$

$$\text{subject to } \sum_{i=1}^{N} \Delta P_{g_i}^{RT} = \sum_{j=1}^{N} \Delta P_{d_j}^{RT}, \tag{2.8a}$$

$$\Delta P_{g_i}^{\max} \le \Delta P_{g_i}^{RT} \le \Delta P_{g_i}^{\min}, \quad \forall i \in \mathcal{N}, \tag{2.8b}$$

$$\sigma F_l^{\min} \le \sum_{i=1}^{N} \mathrm{GSF}_{li}(\Delta P_{g_i}^{RT} - \Delta P_{d_i}^{RT}) \le \sigma F_l^{\max}, \forall l \in \mathcal{L}, \tag{2.8c}$$

where σ is a small positive factor, typically around 0.001 [59]. $\Delta P_{g_i}^{\max}$ and $\Delta P_{g_i}^{\min}$ are generally set as 0.1 MWh and -2 MWh, respectively [59]. To solve this optimization, Lagrange multipliers are established for each constraint to formulate the Lagrange function. The multipliers for the constraint (2.8c) can be used to calculate the RT LMP, reflecting the effect of incremental load on the cost function, as shown in

$$\lambda_i^{RT} = \gamma^{RT} - \sum \mathrm{GSF}_{li}\eta_l^{RT} + \sum \mathrm{GSF}_{li}\zeta_l^{RT}. \tag{2.9}$$

- Profit-Oriented FDIA in Electricity Market

Attackers can exploit electricity markets by executing arbitrage attacks and FDIAs. Specifically, arbitrage attacks enable attackers to sell electricity at high prices and purchase it at low prices. Under normal conditions, LMPs in DA and RT markets are predictable. However, attackers can strategically manipulate estimated power states through FDIAs. Such manipulation can induce line congestion, distort LMPs, and widen the price gap between DA and RT markets, ultimately increasing the attack profits.

The process of this attack is summarised as [58]:

1. Engage in virtual power transactions, purchasing and selling a certain amount of virtual power v_a at bus i and j at price λ_i^{DA} and λ_j^{DA}, respectively.

2. Launch FDIA to the measurements and then manipulate the nodal LMPs in the RT market.

3. Buy and sell the same amount power v_a on the bus j and i at price λ_j^{RT} and λ_i^{RT}, respectively.

The attack profits can be given as

$$R^{Profit} = \left(\lambda_i^{RT} - \lambda_j^{RT} + \lambda_j^{DA} - \lambda_i^{DA}\right) \cdot v_a. \tag{2.10}$$

2.3 Defensive Approaches in Cyber-Physical Power Systems

Various defensive mechanisms have been developed to address cross-layer threats in CPPS [60]. Based on their implementation approaches, these mechanisms can be categorized into two approaches: Layer-specific defense approach and machine learning-based defense approach. Each approach provides unique capabilities for safeguarding CPPS against cross-layer threats.

2.3.1 Overview of Defensive Approaches

- Computation Layer Defense Approaches

  The computation layer defensive approaches play a critical role in enhancing real-time data processing and advanced analytics to enable effective anomaly and decision-making, particularly under emergency conditions and in response to cyber threats. Real-time data processing is pivotal for effective anomaly detection, especially integrated with advanced machine learning algorithms. By exploiting periodic data patterns, it can significantly improve the speed and accuracy of anomaly detection. In addition, the computation layer functionalities further support decision-making by generating actionable insights from processed data, which are crucial for rapid and effective risk

mitigation during emergencies. The computation layer defensive approaches play a critical role in enhancing real-time data processing and advanced analytics to enable effective anomaly and decision-making, particularly under emergency conditions and in response to cyber threats. Real-time data processing is pivotal for effective anomaly detection, especially integrated with advanced machine learning algorithms. By exploiting periodic data patterns, it can significantly improve the speed and accuracy of anomaly detection. In addition, the computation layer functionalities further support decision-making by generating actionable insights from processed data, which are crucial for rapid and effective risk mitigation during emergencies.

- Communication Layer Defense Approaches

  Some researchers highlights communication layer defensive approaches, employing techniques such as Intrusion Detection Systems (IDS) [10], secure communication protocols, and network segmentation [11] to safeguard data transmission pathways and prevent unauthorized access. IDS is an effective approach for continuously monitoring network traffic to identify and prevent malicious activities, such as DdoS attacks and data breaches. Literature [10] proposes a novel IDS tailored for cybersecurity of IEC 61850 based substations. Unlike the traditional IDS, the proposed IDS integrates physical knowledge, protocol specifications, and logical behavior to provide a comprehensive and effective solution that is able to mitigate various cyberattacks in CPPS. Secure communication protocols also serve as a protection to sensitive data protecting critical systems like AMI and SCADA. This is crucial for maintaining the confidentiality and integrity of information exchanged between components in CPPS, including smart meters and control systems. In addition, network segmentation can effectively restrict malware and prevent lateral movement across critical infrastructure systems. For example, literature [11] proposes a network segmentation approach to distribute the trust nodes in SCADA systems to enhance the resilience of CPPS. An emerging security isolation architecture, known as ZTA, further enhances security by enforcing strict access controls and isolating critical cyber components into distinct security zones. This emerging approach shows significant potential for securing CPPS by mitigating risks through effective isolation and verification mechanisms.

- Control Layer Defense Approaches

  Control layer defenses enhance the robustness and stability of systems by utilizing advanced control algorithms to mitigate cross-layer threats. For example, model-based strategies, such as Weighted Least Squares (WLS) and Kalman Filters [12], are utilized for state estimation to detect FDIAs. In addition, predictive and robust control methods are used to mitigate communication delays by compensating for time-lagged feedback or event-trigger feedback [13].

- Physical Layer Defense Approaches

  Physical layer defenses prioritize safeguarding critical infrastructure and equipment from physical threats such as natural disasters and unauthorized access. To mitigate cross-layer impacts in CPPS, protecting edge measurement devices is a key strategy [14], as it prevents tampering with physical measurements. Another effective approach is a novel moving target defense [15], which dynamically alters physical network parameters, such as admittance, to prevent attackers from accurately obtaining power network parameters, thereby reducing the likelihood of successful attacks.

2.4 Microgrid with Device-to-Device Communication

To analyze the impact of degraded QoS on cross-space voltage and frequency regulation, this section introduces the foundational concepts related to a representative CPPS. Specifically, it focuses on microgrid secondary Control with D2D communication. The discussion covers the D2D communication allocation strategy and microgrid secondary control.

2.4.1 Device-to-Device Communication

Cellular networks are an essential communication tool for smart grids, which is typically used to support real-time data exchange among power communication components such as smart meters, sensors, and control centers. However, their limitations, including resource allocation limitations, high latency, and scalability challenges during peak usage, hinder their capability to meet the requirements for the secure operations of smart grids [61].

Existing technologies such as Qrthogonal Frequency Division Multiple Access (OFDMA) and D2D communication present potential solutions to these challenges [29]. OFDMA optimizes resource allocation by dividing the spectrum into sub-channels, thereby reducing latency and improving bandwidth utilization to support multiple simultaneous communications. When Integrated with OFDMA, D2D cellular-assisted underlay D2D communication enables direct device-to-device interactions without routing via base stations, further reducing latency. However, such D2D transmissions are susceptible to interference from concurrent cellular transmissions, which can significantly impact the reliability of data exchange, particularly between RTUs and NUEs as discussed in this research. Therefore, some research [62, 63] has explored optimal strategies for allocating communication resources, including the set of binary variables indicating channel assignments ξ and transmission power $\mathbf{P}^d$, with the multi-objectives of maximizing the sum capacity of

all D2D pairs and minimizing power consumption. In addition, these optimal strategies aim to satisfy constraints related to power, bandwidth, and interference. This optimization is generally modeled as

$$\begin{aligned} \max_{\mathbf{P}^d, \xi} \quad & \frac{\sum_{l=1}^{L_{d2d}} \sum_{i=1}^{I_{cha}} \xi_{l,i} B^c \log_2 (1+\gamma_{l,i})}{\sum_{l=1}^{L_{d2d}} \sum_{i=1}^{I_{cha}} \xi_{l,i} p_{l,i}^d}, \\ \text{s.t.} \quad & \sum_{l=1}^{L_{d2d}} \xi_{l,i} \leq 1, \quad \forall i \in \{1, \ldots, I_{cha}\}, \\ & 0 \leq p_{l,i}^d \leq P_{\max}^d, \qquad \forall l \in \{1, \ldots, L_{d2d}\}, \end{aligned} \tag{2.11}$$

where the objective function is defined as the ratio of the sum capacity of all D2D pairs to the total power consumption. $\sum_{l=1}^{N_{d2d}} \sum_{i=1}^{I_{cha}} \xi_{l,i} p_{l,i}^d$ represents the total power consumption of all D2D links. $p_{l,i}^d$ is the transmission power of D2D link l on channel i, which is constrained within the feasible limits $[0, P_{\max}^d]$. $\xi_{l,i} = 1$ represents the D2D l is assigned to channel i; otherwise, $\xi_{l,i} = 0$ indicates that l is not assigned to i. $\sum_{l=1}^{L_{d2d}} \sum_{i=1}^{I_{cha}} \xi_{l,i} B^c \log_2(1+\gamma_{l,i})$ represents the total throughput of all D2D links, where B^c is the bandwidth of each channel. $B^c \log_2(1+\gamma_{l,i})$ is the capacity of a D2D link l, which depends on their signal-to-interference-and-noise-ratio (SINR). The SINR for D2D link l is defined as

$$\gamma_{l,i} = \log_2 \left(1 + \frac{g_{l,i} P_{l,i}^d}{\sigma^2 + I_l} \right), \tag{2.12}$$

where I_l represents the interference experienced by D2D l. $I_l = \sum_{l'=1, l' \neq l}^{L_{d2d}} \xi_{l',i} g_{l',l} p_{l',i}^d$, where $g_{l',l}$ is the channel gain from D2D link l' transmitter to and link l receiver. σ^2 is the noise power.

The optimization problem aims to achieve trade-offs in D2D communication, ensuring efficient resource allocation while maintaining communication quality and minimizing energy usage.

2.4.2 Microgrid with Secondary Control

This section provides a detailed discussion of the widely implemented hierarchical microgrid structure, as presented in [64].

2.4.2.1 Primary Droop Control System

Generally, the primary control is leveraged to achieve power sharing through the droop controller. For the i-th distributed generator, the droop controller in the

primary frequency control system is modelled as [65]

$$\begin{aligned} \omega_i &= \omega^* - m_i \left(p_i - p_i^*\right) + \Phi_i \left(k\right), \\ v_i &= v^* - n_i \left(q_i - q_i^*\right) + \Omega_i \left(k\right), \end{aligned} \tag{2.13}$$

where ω_i is the nodal frequency, ω^* is the rated frequency, p_i is the output active power, p_i^* is the rated active output power, v_i is the voltage, v_i^* is the rated voltage, q_i is the output reactive power and q_i^* is the rated reactive output power. m_i and n_i are the $P-\omega$ and the $Q-V$ droop coefficients, respectively. ϕ_i and Ω are the auxiliary power variable that are utilized to eliminate frequency deviation resulting from the droop controller.

2.4.2.2 Secondary Frequency Control System

The frequency and voltage deviation from the rated frequency and voltage subjected to the primary droop control can be eliminated through the secondary control [65], which is

$$\begin{aligned} \Phi_i \left(k+1\right) =& \Phi_i \left(k\right) - hk_1^{\omega} \left(\omega \left(k\right) - \omega^*\right) - hk_2^{\omega} \sum_{j \in N_{dg}} a_{ij} \left(\frac{p_i \left(k\right)}{p_i^*} - \frac{p_j \left(k\right)}{p_j^*}\right), \\ \Omega_i \left(k+1\right) =& \Omega_i \left(k\right) - hk_1^{v} \left(v_i \left(k\right) - v^*\right) - hk_2^{v} \sum_{j \in N_{dg}} a_{ij} \left(\frac{q_i \left(k\right)}{q_i^*} - \frac{q_j \left(k\right)}{q_j^*}\right), \end{aligned} \tag{2.14}$$

where k_1^{ω} and k_1^{v} are the frequency recovery coefficients for frequency and voltage, respectively. Similarly, k_2^{ω} and k_2^{v} are the active power sharing recovery coefficient for frequency, and reactive power sharing for voltage, respectively. a_{ij} is the weight of the communication edge between distributed generator i and distributed generator j. h is the control time instant and N_{dg} is the number of distributed generators.

2.4.3 Power Flow Analysis

The power flow analysis in power system is responsible for guaranteeing the stability of the system, which is widely modelled as the solution of nonlinear algebraic equations [66]

$$\begin{aligned} p_i &= \sum_{j=1}^{N} \left|v_i\right| \left|v_j\right| \left(G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}\right), \\ q_i &= \sum_{j=1}^{N} \left|v_i\right| \left|v_j\right| \left(G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}\right), \end{aligned} \tag{2.15}$$

where δ is the phase angle. $Y_{ij} = G_{ij} + B_{ij}$ is the inductive admittance for the line between bus i and bus j. G and B represent the conductance and susceptance, respectively.

This section introduces the concept of a microgrid with D2D communication as a representative CPPS for analyzing the impact of optimal allocation strategies on voltage and frequency regulation. Microgrids are fundamental components of CPPS, which maintains power quality by ensuring stable voltage and frequency levels. In addition, D2D wireless communication is an emerging power communication technology, offering enhanced flexibility and efficiency in CPPS.

2.5 Basics of Reinforcement Learning

RL serves as an effective heuristic approach for multi-objective optimization, offering significant advantages in security approach development. By interacting with the environment, RL algorithms autonomously learn and adapt to evolving threats, enabling dynamic strategy adjustments to improve security approaches. In addition, well-trained RL models improve the efficiency of incident response and enable the simulation of adversarial scenarios, optimizing strategies through attack-defense modeling. Therefore, this section introduces the foundational concepts of RL and its emerging applications in CPPS compared with other machine learning approaches.

2.5.1 Concept of Reinforcement Learning

RL is a powerful machine learning framework where an agent learns to make sequential decisions and receive feedback by interacting with the environment with the aim to maximize cumulative rewards. The process is formalized as a Markov Decision Process (MDP), which can be represented by a 4-triple $(\mathcal{S}, \mathcal{A}, P, r)$, as shown in Fig. 2.4.

- $\mathcal{S} = \{s_I, s_{II}, ..., s_{N_s}\}$ is the set of all states of the power system with $s \in \mathcal{S}$.
- $\mathcal{A} = \{\alpha_I, \alpha_{II}, ..., \alpha_{N_\alpha}\}$ is the set of attacking actions and each state S has a relevant set of available actions $\mathcal{A}(s)$ with $\alpha \in \mathcal{A}$.
- $P(s_{t+1}|s_t, \alpha_t)$ denotes the probability when the state changes from s_t to s_{t+1} after taking action α.
- $r(s, \alpha)$ denotes the immediate reward that the system starts from state s by adopting an action α.

The objective of the agent is to derive an optimal policy that maps states to actions to maximize the expected cumulative reward, often referred to the return as

$$\mathbb{E}\left[\sum_{k=0}^{\infty}\gamma^k r_{t+k} \mid s_t = s, a_t = a\right], \tag{2.16}$$

where $\gamma \in [0, 1)$ represents the discount factor to balance the future reward and the immediate rewards.

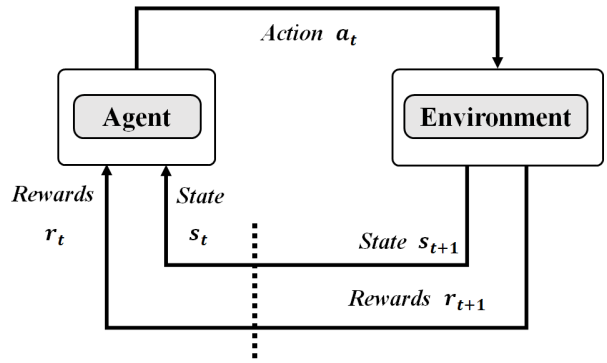


FIGURE 2.4: Basic of reinforcement learning.

Policy Definition: Two main types of policies are analyzed, including deterministic policy $a_t = \mu(s_t)$ and stochastic policy. Specifically, a **deterministic policy** assigns a unique action to each state, expressed as: $a_t = \mu(s_t)$, where μ is the policy function and s_t is the state at time t. Comparatively, a **stochastic policy** defines a probability distribution over possible actions for each state, rather than selecting a single deterministic action. It is expressed as: $a_t \sim \pi(\cdot|s_t)$, where π represents the probability of selecting action a_t given the state s_t.

Value functions and optimization: The **state-value function** $V^\pi(s)$ represents the expected return starting from state s and following policy π, which is defined as $V^\pi(s) = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty}\gamma^k r(s_k, a_k)\middle| s_0 = s\right]$, where $\mathbb{E}_\pi$ is the expected value under policy π. Comparatively, the **action-value function** $Q^\pi(s, a)$ represents the expected return starting from state s, taking action a, and following policy π, which is $Q^\pi(s, a) = \mathbb{E}_\pi\left[r(s, a) + \gamma V^\pi(s')\right]$. The optimal policy π^* maximizes the function as $\pi^*(a|s) = \arg\max_a Q^*(s, a)$, where $Q^*(s, a)$ is the optimal action-value function.

Policy and Value-Based Methods: There are two types of approaches to policy improvement. **Policy optimization methods** directly optimize the policy $\pi_\theta(a|s)$ parameterized by θ using gradient-based approaches as $\nabla_\theta J(\pi_\theta) = \mathbb{E}_\pi\left[\nabla_\theta \log \pi_\theta(a \mid s) Q^\pi(s, a)\right]$. **Value-based methods** use Bellman equations to iteratively compute value functions as $V^\pi(s) = \mathbb{E}_\pi\left[R(s, a) + \gamma V^\pi(s')\right]$.

2.5.2 Application of Deep Reinforcement Learning in Combinatorial Optimization

Many combinatorial optimization problems are difficult to solve within an efficient polynomial time. Hence, the heuristic algorithm is favorable in practical scenarios, since they can usually get a feasible solution faster compared with the approximate or exact algorithms. One of the new trends in these years to solve combinatorial optimization is to leverage the reinforcement learning technique [67].

Take the novel architecture "end-to-end neural combinatorial optimization pipeline"[68] as an example. The dataset corresponding to a specific optimization problem can processed by machine learning algorithms to produce embeddings for various instances. Afterward, the trained model is used to approximate and extract information for this specific optimization problem. The process of searching for the optimal solution can be viewed as a Markov decision process, where the agent can search for an optimal action sequence (i.e., solution) to optimize the reward (i.e., objective).

A combinatorial optimization problem can be described using the following formula:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & h_j(x) = 0, \quad j = 1, \ldots, n, \end{aligned} \tag{2.17}$$

where the objective function $f(x)$ defines the goal of the optimization problem, such as minimizing costs or maximizing efficiency. The decision variable x is chosen from the solution space $\mathcal{X}$, which represents all possible discrete combinations or arrangements. The constraints $g_i(x) \leq 0$ and $h_j(x) = 0$ restrict the feasible region by enforcing inequality and equality conditions, respectively. Since the solution space $\mathcal{X}$ is discrete, combinatorial optimization problems are often computationally challenging as the size of $\mathcal{X}$ increases, leading to exponential growth in potential solutions.

The relationship between combinatorial optimization and reinforcement learning is given in Table. 2.1.

2.5.3 Application of Hierarchical Multi-agent Deep Deterministic Policy Gradient Algorithm in Stackelberg Game

Chapter 4 proposes a H-MADDPG algorithm to address multi-agent problems involving attackers and operators within the Stackelberg game framework. H-MADDPG facilitates the optimization of policies through the interactions between

TABLE 2.1: Mapping combinatorial optimization to reinforcement learning.

Combinatorial Optimization	Reinforcement Learning
Objective function $f(x)$	Reward function $R(s, a)$
Solution space $\mathcal{X}$	Set of all states $\mathcal{S}$
Decision variable x	Set of all actions $\mathcal{A}$
Constraint conditions g/h	Invalid action penalties and constraints
Optimal solution x^*	Optimal policy $\pi^*(a\|s)$
Optimization process	Policy-Based/Value-Based Approaches

attackers and operators. This section introduces the key concepts of H-MADDPG and Stackelberg game theory to facilitate a better understanding.

Stackelberg Game: A Stackelberg game is a strategic framework characterized by a hierarchical leader-follower relationship. Within this model, the leader acts first by making an initial decision, which is subsequently observed by the follower. The follower then optimizes their strategy in response to the reader's actions. This sequential decision-making process enables the leader to influence the follower's actions, thereby establishing a strategic advantage [69], which is formulated as

$$\begin{aligned} \left(x^{A*}, x^{B*}\right) &= \arg \max_{(x^A, x^B) \in \Omega^A \times \Omega^B} \mathcal{U}\left(x^A, x^{B*}\right) \\ \text{s.t.} \quad x^{B*} &= \arg \max_{x^B \in \Omega^B} \mathcal{L}\left(x^A, x^B\right), \end{aligned} \tag{2.18}$$

where the leader solves $x^{A*} = \arg\max_{x^A \in \Omega^A} \mathcal{U}\left(x^A, x^{B*}\right)$ with the leader's objective function $\mathcal{U}\left(x^A, x^B\right)$. Comparatively, the follower solves $x^{B*} = \arg\max_{x^B \in \Omega^B} \mathcal{L}\left(x^A, x^B\right)$ utilizing the follower's objective function $\mathcal{L}\left(x^A, x^B\right)$. The solution $\left(x^{A*}, x^{B*}\right)$ forms the Stackelberg equilibrium, where the leader and follower's strategies are mutually optimal.

Hierarchical Multi-Agent Deep Deterministic Policy Gradient: The Actor-Critic architecture is a type of RL that integrates policy optimization and value estimation. The Actor selects actions based on the current policy to maximize expected rewards, while the Critic evaluates these actions by estimating the value and generating feedback to the Actor. The Deep Deterministic Policy Gradient (DDPG) is an off-policy algorithm suitable for continuous action spaces based on the actor-critic framework. Building on DDPG, the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) adapts it for multi-agent environments, where each agent has its own actor network, and a centralized critic evaluates joint actions to optimize strategies in the presence of other agents [70].

The H-MADDPG algorithm extends the MADDPG to address hierarchical agent structures, particularly in contexts such as Stackelberg games. By optimizing policies

across multiple levels of agents, H-MADDPG facilitates behaviors in complex environments, allowing for strategic interactions.

2.5.4 Emergence of Reinforcement Learning in Defensive Approaches in Cyber-Physical Power Systems

The growth of CPPS has generated vast datasets, posing challenges in analysis, insight extraction, and security [21]. ML has emerged as a critical tool in addressing CPPSs security concerns, primarily leveraging three approaches: supervised learning, GANs, and RL [21]. Supervised learning excels at extracting patterns from historical data, making it highly effective for anomaly detection and threat identification. When labeled datasets are insufficient, GANs can generate labeled data, addressing data scarcity. In addition, RL offers a distinct advantage by learning directly from interactions with the environment, without relying on pre-existing datasets or models. This adaptability enables systems to develop rapid defensive decision-making strategies and facilitate real-time responses to evolving cyber threats.

Supervised learning leverages labeled datasets to train models, enabling them to learn the relationship between input features and output labels. This approach is effective in anomaly detection in CPPSs, including detecting FDIAs)and identifying abnormal traffic patterns. Literature [71] formulates FDIA detection as a supervised learning classification problem, leveraging observed measurements as features and incorporating prior system knowledge to address sparsity constraints. Building on this, [72] incorporates semi-supervised, online learning, and fusion algorithms within a generic attack construction framework for diverse attack scenarios.

The limited availability of sufficient samples and labeled data from power system operation poses a significant challenge for many ML-based approaches in CPPS. To address this problem, an unsupervised learning approach, generative adversary network (GAN), has been increasingly applied in CPPS security since it can supplement sufficient samples utilizing generative learning. As contrast to solely relying on sampling data for the power system state estimation, GAN has been used in [73] to fill in missing measurements caused by cyberattacks or failures, improving data completeness in power system state estimation. Moreover, [74] judiciously combines traditional power models with GANs to improve data reliability. In this approach, the compromised measurements generated by the power model, induced by FDIAs, are subsequently repaired by the GAN to achieve high accuracy.

Reinforcement learning focuses on developing optimal policies through interactions with the environment, leveraging states, actions, rewards, and policies. In CPPS, many defensive decision-making problems can be transformed as sequential decision-making tasks, which are challenging to solve within polynomial time. To

address this challenge, RL models these tasks as MDPs, optimizing action sequences (i.e., solutions) to maximize rewards (i.e., objectives) through interactions with the environment. Once the model is trained, RL enables the optimization of defensive decision-making strategies during inference time, making it well-suited for immediate security responses [67].

Generally, due to limited resources, attackers can launch attacks in a finite number of times to do as much damage as possible [75], which can be considered as an optimization problem. On the basis, intelligent attackers can be viewed as an agent to adjust its attack actions to optimize the attacking rewards. An approach based on Q-learning is proposed in [76] to optimize the attack sequences against the topology attack. Literature [77] utilizes the Markov decision process to simulate the risk propagation process of a CPPS, which seeks to maximize deviation from the estimated states by optimizing the attack sequence and targets. Moreover, from the perspective of defenders, the effectiveness of a defense strategy is highly associated with the dynamic attack-defense interactions instead of purely focusing on the intentions from attackers. Literature [78] model the DDoS attack-defense architecture in the real-time energy market as a Markov decision process, where attackers and defenders alter the target link selection sequentially to achieve their objectives. The attackers aim to maximize the market price decline from the true to the depressed value, whilst the defenders aim to decrease it. Literatures [79, 80, 81] have modeled the process of attack-defense interaction under the FDI attack scenario as a zero-sum stochastic game. The Nash equilibrium strategy for both defenders and attackers is solved and utilized to guide the defense resource allocation.

2.6 Basics of Zero Trust Architecture

In this section, the concept of ZTA, the framework of ZTA, and the existing techniques for ZTA are discussed, respectively.

2.6.1 Concept of Zero Trust Architecture

A ZTA leverages zero trust principles and security philosophy to deploy industrial workflows and infrastructure. Conventional network security architecture divides the single and interconnected network into smaller network zones, which are isolated through virtual firewalls. According to the predefined level of trust of each network zone, the network resources are determined whether granted or denied permission to access, upon which this conventional architecture is always provided with a strong defense perimeter. In other words, once the hackers have penetrated into the network

zone through achieving the trust level of this network zone, almost all the resources in this zone will be vulnerable.

ZTA addresses this issue through focusing on resources, asset and entities, instead of network zones, since for one resource, its specific location in network is no longer deemed as the main factor to evaluate the security. The main assumption of zero trust (ZT) is that there is no implicit trust granted to assets or user accounts based only on their physical or network location or based on asset ownership. ZTA is responsible for the novel network trends which consist of terminal devices, cloud computing technique, bringing your own devices and so on, since these devices and techniques are not located in an exact perimeter of the network. In addition, the ZTA pays more attention to protecting resources, such as workflows, services, network assets, database and so on, as the location of both subject and resource in the network is no longer deemed as the main factor to the evaluate security of the resource [82].

To establish a zero trust network requires the following five essential assertions:

- The entire network is assumed to be hostile at any time.
- The potential threats may occur inside the network or outside the network every time.
- Network locality is not sufficient for deciding trust in a network.
- To access network, each user, device and asset must be authenticated and authorized.
- Numerous sources of data are used to calculate and update the access assignment policies.

2.6.2 Logical Components of Zero Trust Architecture

As shown in Figure 2.5, there are five main components for ZTA, including subjects, policy enforcement point (PEP), resources, policy decision point (PDP), and other supplements [22]. Subjects are defined as the user or any device which request being granted access the resources. A policy decision point is used to deciding to grant or deny access to the resources. In this way, the communication between the resource being requested and the subject will be established or terminated. PDP consist of two components: policy engine (PE) and policy administrator (PA). PE is responsible for making decisions and PA is responsible for communication management. In this framework, the concept of resource represents the resources which is requested by the subjects. PEP is responsible for forwarding the request information for PDP, and direct control the communication between the subjects and the requested resources. In addition, PEP can also monitor the incoming network traffic between the resources

and the subjects. Finally, supplement is responsible for offering reference information, such as network traffic logs, threat intelligence information and so on, to the PE. By referring to this information, PE can make decisions more accurately and correctly, thereby improving the security of the overall system.

Generally, while a subject request to access one resource, it can first issue a request to access this resource. Subsequently, the PEP will receive this request and then transmit it into PDP, which decides whether to grant this request or not. This decision made in PDP will be transmitted into PEP, which will establish or terminate the communication referring to the request.

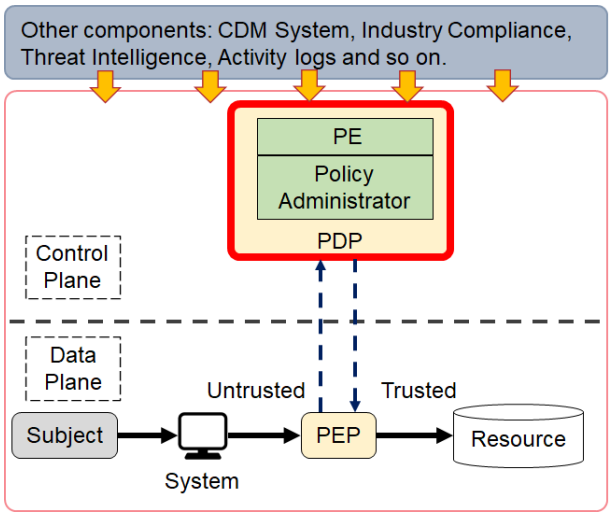


FIGURE 2.5: Core zero trust logical components.

2.6.3 Variations of Zero Trust Architecture Techniques

There are various approaches for designing ZTA for workflows through using the existing software and hardware as possible, which vary in using ZT components and the main policy rules. Even though each method is on the basis for the tenets of ZT, they may leverage one or two ZT components as their core policies. Generally, one complete ZT solution will include elements of all three methods, consisting of enhanced identity governance–driven, logical micro-segmentation, and network-based segmentation. The main determination for which method to apply is the requirement for minimizing the fundamental changes to the current workflows.

- Enhanced Identity Governance:

  As the core component of policy engine, the improved identity governance approach is to assign the dynamical access mainly based on the identity of actors. For example, if one subject was not requesting access for the resource, it would not be assigned any access authorization. The core objective is to access the least privileges granted the set of subjects for accessing the resources.

- Micro-Segmentation

  The principle of this approach is to design a unique network segment to support the functionalities in ZTA, which is isolated by one gateway security part. The main challenge of this approach is the deployment cost and large changes while it is deployed in the novel system.

- Software-Defined Perimeter

  The above approaches are based on concepts from Software Defined Networks, where PA works as the network controller which can launch and reconfigure the network according to the decisions launched from the PE. The subjects can then request the accessing authentication through PEP, which is controlled by the PA part.

This research investigates the impact of two representative cross-layer threats in CPPS: QoS and FDIAs. To analyze these threats and their impacts for CPPS, this chapter provides a literature review covering the definitions and impacts of QoS and FDIAs, examines existing defense mechanisms, identifies research gaps, and presents the preliminaries of the relevant technologies employed in the following chapters.

2.6.4 Applications of ZTA in Industrial Systems

In recent years, an emerging security technique known as ZTA has gained more attention in industrial systems. Referring to the National Institute of Standards and Technology (NIST) report [22], the key principle for ZTA is continuous authentication and dynamic access control. In May 2021, the U.S. president issued an executive order on strengthening the Nation's Cybersecurity [83], which guides agencies to implement ZTAs. In February 2022, the OMB Acting Director issued [84] that federal government agencies are required to meet five Zero-Trust goals in strategy [85] until the end of 2024. According to Okta's report "The State of Zero Trust Security 2022" [86], 55% of the organizations surveyed had a Zero-Trust plan and the majority (97%) intended to have one in the future. These incidents imply that ZTA became mainstream while replacing the legacy network and security standards.

In fact, the increasing attack surfaces and lateral spreading of industrial systems have challenged the capabilities of traditional CPS security architectures [87]. Take the CPPS as an example. In a CPPS, the majority of power components are controlled by various cyber systems such as SCADA and EMS. Any vulnerability in these cyber systems can serve as a gateway for attackers to penetrate. The reason is that the traditional perimeter-based security architecture has trust zones where users are granted redundant privileges to access other resources. In case the weakest endpoint

is compromised, attackers can utilize it to spread laterally and further penetrate the physical system in a cross-layer fashion, thereby causing large damage.

The key solution to the aforementioned issues is continuous authentication and making access control policies for each access request inside the perimeter, which can be achieved by the standard ZTA. The main principle of the standard ZTA is that every network component cannot be trustworthy without authenticating and authorizing, thereby restricting the lateral movement [82]. Many researchers have investigated designing and implementing ZTA and its variants to several industrial architectures, including the IoT platform, power IoT [88, 89], Cloud platform [90, 91, 92] and Smart Healthcare [93, 94], to enhance or provide security posture. Meanwhile, many organizations have already implemented ZTA into their infrastructure in practice, including Google and Microsoft. In addition, while research into ZTA is still ongoing, a majority of organizations and corporations have already implemented ZTA in their existing security architectures in practice, as shown in TABLE. 2.2.

TABLE 2.2: The development of ZTA in industrial infrastructures and corporations.

2014	Google: BeyondCorp program is initiated to use ZTA in internal network infrastructure [95].
2019	Microservices: Designing eZTrust, a network-independent ZT perimeterization [96].
2020	Netskope: A cloud-based ZT network access is issued to general availability [97].
	Microsoft: A ZT maturity model is published [98].
2021	5G Edge Computing: Constructing a dual-layer ZTA to improve 5G multi-access edge computing [99].
	Power IoT: Designing the power IoT security defensive architecture based on ZT principles [88].
	GitLab: A roadmap for ZTA is built [100].
2022	Cloud Computing: Comparing novel requirement-specific features for ZT cloud networks [92].
	Virtual Power Plants: Illustrating ZTA flow in prosumer requests to plant operator [89].
2024	Palo Alto: Using ZT with Palo-Alto Networks [101].

These cases suggest that ZTA has the potential to develop from a novel concept to a widely used architecture to secure industrial architecture security in the future.

Chapter 3

Cross-Layer Optimization of Microgrid Secondary Control and Communication QoS for Cyber-Physical Power Systems

.

3.1 Introduction

As the advanced information technologies are gradually deployed in CPPS, the power system is increasingly vulnerable to malicious attacks. In addition, extreme weather conditions might disrupt the normal operation of CPPS, thereby resulting in undesired power communication interruptions and even blackouts [8]. For instance, in 2012, over 8 million customers were left without power supplies and $70 billion losses in total when Hurricane Sandy damaged energy infrastructure in the United States [102]. Similarly, in 2015, three Ukrainian energy companies were damaged by a sophisticated cyber-attack, leading to power outages for several hours impacting around 225,000 customers [9]. Therefore, enhancing the communication reliability and power system stability under abnormal situations is a crucial requirement for the CPPS.

3.1.1 Overview

To address the challenges posed by degraded QoS, many existing works have designed advanced control algorithms and achieved success in making the CPPS

intelligent against communication latency [25]. However, such control algorithms require high QoS of communication with low latency and high throughput but they have limited communication resource. Specifically, in the control processing, the sensor will send the sampled data to a controller to calculate the control output, which is then forwarded to the actuators. For example, when the centralized base stations that support the communication resources are damaged by the natural disasters or the cyberattacks, the limited bandwidth cannot satisfy the requirement of each transmission link. In this case, the data packets in transmission are dropped off or delayed, the performance of the control algorithms, and even the stability of the physical system will be impacted, especially for island Microgrids away from the main grid.

One might think that degraded QoS could be mitigated by communication technologies in the cyber layer before the interrupted or falsified data packets are transmitted to controllers or actuators. These techniques can be classified into three categories, namely, optimal sampling time selection [26], optimal routing selection [27], and optimal communication resource allocation [27, 28]. These works aim to mitigate the cyber failures through enhancing the QoS of data transmission, such as maximizing throughput and minimizing time delay. Literature [26] designs the control period as a state variable to analyze the input-to-state stability of this closed-loop system, to mitigate the influence of the time delay. In addition, to improve the performance of smart grids with hierarchical control framework in the face of possible cyber-physical coupling failures, literature [27] proposes an optimal routing strategy to maximize the accessibility of data transmission. However, these two techniques are more applicable as pre-event methods since they use wired communication, in contrast to post-event methods, which requires wireless communication.

For wireless communication, a popular D2D communication technique can be leveraged as a support for remote Microgrids away from main base stations [30]. The existing works aim to improve the QoS of communication through designing the optimal communication resource strategy to assign subcarriers and transmit power of users [29]. According to our knowledge, there are several works focusing on using D2D communication in industrial control systems, considering the cross-layer impact from the communication system on the industrial physical systems. [28] designs the optimal communication resource allocation strategy that not only minimizes power consumption but also, for the first time, introduces an optimal sampling approach to ensure the cross-layer impact of this strategy on the QoS requirement of the control subsystem. In addition, [30] designs a cyber-physical coordinative mitigation framework that pre-allocates the communication resource to minimize the time delay while also analysing the impact of this time delay on the frequency control of the power system within the control algorithm. [103] proposes a post-disaster restoration

framework that achieves the goal of load recovery by simultaneously designing D2D communication and active distribution networks.

Although these works consider both communication reliability and its cross-layer impact on physical stability, they leverage communication reliability as the intermediate objective when analyzing the cross-layer impact of communication resource allocation strategy on physical stability. That is, they either optimize the communication reliability alone or optimize the physical stability alone. However, a superior communication strategy in terms of QoS does not always lead to superior physical stability. For example, for certain insignificant measurements, low communication delays do not necessarily lead to a well-performing control algorithm. As above mentioned, in addition to the QoS, the power system stability needs to be considered as a key factor when designing communication resource strategies, as such a strategy can impact physical stability in a cross-layer fashion. However, it is challenging to consider multiple factors simultaneously in optimization strategies. As indicated in [104], even simultaneously optimizing voltage regulation and frequency regulation can lead to complex Pareto-optimal solutions.

3.1.2 Contribution

This chapter explores the CPPS structure including a microgrid and a wireless communication network with limited bandwidth as the communication network to support its hierarchical control architecture framework to regulate frequency and voltage in failure situations. In the extreme scenario, communication facilities may be compromised, wherein the QoS which supports the information exchange in the distributed control is degraded. In this case, the performance of distributed secondary control will be impacted, resulting in the frequency and voltage deviation. To mitigate such cross-layer impact induced by degraded QoS, this chapter designs a D2D communication resource allocation strategy that can minimize both QoS disruptions and their cross-layer impacts on microgrid stability. The main contributions of this chapter are as follows.

1. A specialized CPPS model is proposed to evaluate the cross-layer impact of the communication resource allocation strategy on the physical states. This model extracts the interdependent cyber and physical states instead of all the variables, to reduce the complexity of this joint system.

2. A joint multiobjective optimization problem is formulated to determine the optimal communication resource allocation strategy (i.e., channel selection and transmit power allocation) to find the tradeoff between the QoS in communication and the Microgrid stability, simultaneously. In addition, an

AW-MRAS algorithm is designed to solve the proposed problem, which can prune the search space based on the unique characteristics of our CPPS model.

3. Compared with the state-of-the-art strategy for the single objective problem that only optimizes the QoS index, our proposed optimal strategy results in 13.74% and 4.57% decrease in frequency and voltage deviation with a slight compromise in the QoS index, respectively. In addition, the proposed AW-MRAS algorithm has superiority in balancing population diversity and convergence, compared with four multiobjective optimization algorithms.

3.1.3 Structure of Chapter

The remainder of this chapter is organized as follows. Section 3.2 outlines the system model. Section 3.3 introduces the proposed CPPS model, analyzing the cross-layer impacts of D2D communication allocation strategies on microgrid stability. A joint multi-objective optimization problem is formulated to derive the optimal D2D communication strategy that minimizes degraded QoS disruptions and their cross-layer impacts on frequency and voltage regulation. Section 3.4 presents the AW-MRAS algorithm designed to optimize the proposed D2D strategy. Section 3.5 provides the simulation results and analysis. Finally, Section 3.6 concludes this chapter with a summary of findings.

3.2 System Model

Generally, the QoS for the information exchange between local controllers in the hierarchical control [64] of an islanded Microgrid can be supported by the wireless communication or wired communication. Such communication technology, however, might become unstable in regard to the extreme scenario. In this case, the bandwidth resource, which is previously used to support other communication services such as the mobile communication, can be multiplexed by the power services. Considering such a scenario, a structure of CPPS is designed as shown in Fig. 3.1, including a Microgrid with a set of DGs and a wireless communication network with a set of D2D communication links and normal cell users [105, 28]. To simplify the expression, this chapter defines the set of DGs and D2D links as $\mathcal{N}_{dg} = \{1, 2, \ldots, N_{dg}\}$, and the $\mathcal{L}_{d2d} = \{1, 2, \ldots, L_{d2d}\}$, respectively. Each DG is deployed with the sensor, actuator communicating with local controllers and remote terminal units (RTU). The information exchange between the RTUs of DG z and its adjacent DGs is transmitted via D2D links in the communication network. Users in communication network are classified into two types: Normal users equipments (NUE) i and D2D link users l, where $l = (j, z) \in \mathcal{L}_{d2d}$, $j \in \mathcal{N}_{dg}$ and $z \in \mathcal{N}_{dg}$. $l = (j, z)$ supports the D2D link

transmitted from RTU of DG j to the RTU of its adjacent DG z. The total bandwidth dedicated to users is divided into I_{cha} channels in the set $\mathcal{I}_{cha} = \{1, 2, \ldots, I_{cha}\}$ and each channel is assumed to pre-allocate to one NUE. These channels allocated to NUEs are multiplexed by D2D links. Note that wireless network handles all information exchange, including the D2D links between RTUs of adjacent DGs (supporting the secondary control), and intercommunications among NUEs. In details, the system model can be described in two aspects:

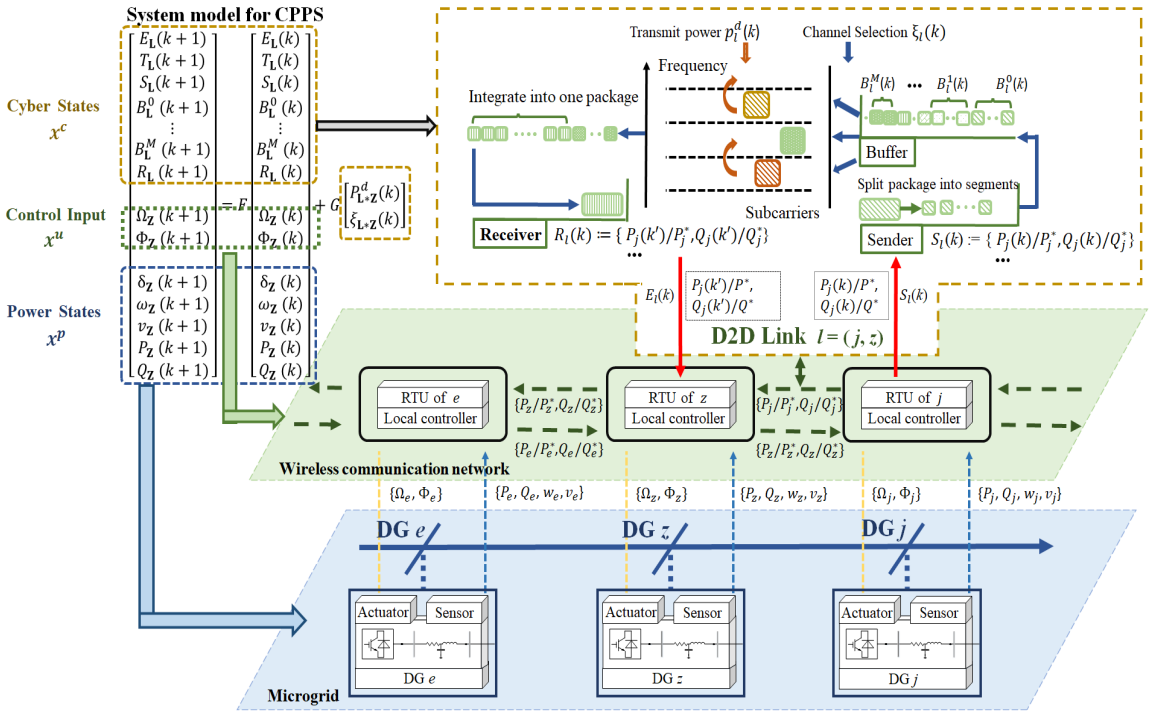


FIGURE 3.1: The topology of the CPPS system model.

3.2.1 The Cyber Layer

The SINR for the D2D link $l = (j, z)$ on channel i (supporting the information exchange from RTU j to RTU z) is determined as

$$\gamma_l(k) = \frac{\sum_{i=1}^{\mathcal{I}_{cha}} \xi_{l,i} p_{l,i}^d(k) g_l^d}{\sigma_0^2 + \sum_{i=1}^{\mathcal{I}_{cha}} \xi_{l,i} \left(p_i^c g_{i,l}^c + \sum_{l'=1, l' \neq l}^{\mathcal{L}} \xi_{l',i} p_{l',i}^d(k) g_{l',l}^d \right)}, \tag{3.1}$$

where σ_0 is the Gaussian noise. p_i^c is the transmits power for NUE i on channel i. g_l^d is channel gains for D2D link l, and $g_{i,l}^c$ and $g_{l',l}^d$ represent channel gains from NUE i transmitter to D2D link l receiver, and from the D2D link l' transmitter to D2D link l receiver, which can be modeled as $g_{i,l}^c = {d_{i,l}}^{-3} u \left(g_{l',l}^d = {d_{l',l}}^{-3} u \right)$, respectively. $d_{a,b}$ is the distance between the transmitter a to receiver b and u is the loss factor.

Given an SINR, based on the well-known Shannon formula, the transmission rates, also termed as throughput, for D2D link $l = (j, z)$ on can be modeled as

$$\mathcal{R}_l(k) = B^c \log_2 \left(1 + \gamma_l(k)\right), \tag{3.2}$$

where B^c is the bandwidth of each channel. Generally, due to various uncertainties in data communication, the $\mathcal{R}_l(k)$ is the throughput in an ideal scenario, which will impact the QoS of data communication for the control systems.

3.2.2 The Physical Layer

The microgrid system model adopted in this section is widely used among experimentalists, which consists a set of DGs and the dynamics of each DG z can be given as [65]

$$\begin{aligned}
\delta_z(k+1) &= \delta_z(k) + h \cdot (\omega_z(k) - \omega_{sys}) \\
\omega_z(k+1) &= \omega_z(k) + h \cdot m_\omega \eta^1 (P_{ez}(k) - P_z(k)) + \Omega_z(k+1) - \Omega_z(k) \\
v_z(k+1) &= v_z(k) + h \cdot m_v \eta^2 (Q_{ez}(k) - Q_z(k)) + \Phi_z(k+1) - \Phi_z(k) \\
P_z(k+1) &= P_z(k) + h \cdot \eta^1 \left(P_{ez}(k) - P_z(k)\right) \\
Q_z(k+1) &= Q_z(k) + h \cdot \eta^2 \left(Q_{ez}(k) - Q_z(k)\right),
\end{aligned} \tag{3.3}$$

where η^1 and η^2 are the low-pass time constant of power filter [106]. R_l is the received variable by local controller of z from the RTU j via the communication network. Ω_z and Φ_z are the control input, which is calculated by controllers, relying on the QoS in the cyber layer.

$$\begin{aligned}
P_{ez}(k) &= \sum_{j=1}^{N} |v_z||v_j|(G_{zj}\cos(\delta_{zj}(k)) + B_{zj}\sin(\delta_{zj}(k)), \\
Q_{ez}(k) &= \sum_{j=1}^{N} |v_z||v_j|(G_{zj}\cos(\delta_{zj}(k)) - B_{zj}\sin(\delta_{zj}(k)).
\end{aligned} \tag{3.4}$$

where $\delta_{zj}(k) = \delta_z(k) - \delta_j(k)$ and Y_{zj} is the inductive admittance for the inductive line between DG i and DG j.

3.3 Joint Optimization for D2D Communication Resource Allocation Strategy

Based on the system model, this section presents a CPPS model for quantitatively analyzing the cross-layer impact of resource allocation on frequency and voltage, and

proposes a multi-objective optimization framework to minimize physical state disruptions while improving QoS.

3.3.1 CPPS Model: Cross-Layer Impact of D2D Communication Strategies on Microgrid

In this section, the relationship between the throughput $\mathcal{R}_l(k)$ in Eqn. (3.2) in the cyber layer and the control input Ω_z and Φ_z in Eqn. (3.3) in the physical layer is formulated. Based on this relationship, the joint CPPS structure including the Microgrid with the distributed secondary control via the communication network can be modeled as

$$\begin{bmatrix} x^c_{L_{d2d}}(k+1) \\ x^u_{N_{dg}}(k+1) \\ x^p_{N_{dg}}(k+1) \end{bmatrix} = F \begin{bmatrix} x^c_{L_{d2d}}(k) \\ x^u_{N_{dg}}(k) \\ x^p_{N_{dg}}(k) \end{bmatrix} + G \begin{bmatrix} P^d_{L_{d2d}\times I_{cha}}(k) \\ \xi_{L_{d2d}\times I_{cha}}(k) \end{bmatrix}, \tag{3.5}$$

where $x^c_l := [E_l, T_l, S_l, B^0_l, \cdots B^M_l, R_l]$ is defined as the cyber states for D2D link l, and $x^p_l := [\delta_z, \omega_z, v_z, P_z, Q_z]$ is the physical states. $S_l = [S^\omega_l, S^v_l]$, $B_l = [B^\omega_l, B^v_l]$ and $R_l = [R^\omega_l, R^v_l]$ are the transmitted variable, the buffer variable and received variable, respectively. In addition, $\delta_z, \omega_z, v_z, P_z$ and Q_z are the state variables of the power system. In this structure, the interaction between the cyber and physical layer relies on the secondary control input $x^u := [\Omega, \Phi]$. Specifically, at time step k, physical states for DG j are measured by the sensors, and then transmitted to the local controller and RTU. Next, these state variables can be transmitted from RTU j to the RTU and local controller of z. Afterwards, the control input x^u can be calculated according to these state variables, and then transmitted to actuators to regulate the frequency and voltage. $\mathbf{P}^d := [p^d_{l,i}\,(l \in \mathcal{L}_{d2d}, i \in \mathcal{I}_{cha})]$ is defined as the transmit power allocation matrix, where p^d is the transmit power for D2D links on channels.
$\xi := [\xi_{l,i}\,(l \in \mathcal{L}_{d2d}, i \in \mathcal{I}_{cha})]$ denotes the channel selection matrix. k is the discrete time step.

Define P_s as the probability of a successful packet transmission, which is related with the SINR [107] and given as

$$P_{s,l}(k) = \left(1 - \frac{1}{2}\,\mathrm{erfc}(\sqrt{\gamma_l(k)})\right)^{L^p}, \tag{3.6}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function. Hence, the expected throughput can be reformulated as

$$E_l(k) = P_{s,l}(k) \cdot \mathcal{R}_l(k). \tag{3.7}$$

Afterwards, $T_l(k)$, the sum of throughput over time from k_0 to k, is given by

$$T_l(k+1) = T_l(k) + E_l(k). \tag{3.8}$$

The transmission process of D2D link l in communication network, is shown at the right side of Fig. 3.1. During each period, the calculated state variables P/P^* and Q/Q^* are reformulated as the sent variable S_l^ω and transmitted to the send buffer before accessing the communication network. It means that the value of $S_l = [S_l^w, S_l^v]$ is updated within each period, which is given as

$$\begin{aligned} S_l^w(k+1) &= \sum_{j=1}^{N_{dg}} b_{lj} \cdot \frac{P_j(k+1)}{P_j^*}, \\ S_l^v(k+1) &= \sum_{j=1}^{N_{dg}} b_{lj} \cdot \frac{Q_j(k+1)}{Q_j^*}, \end{aligned} \tag{3.9}$$

where b_{lj} is the incidence matrix, where b_{lj} =1 if $l = (j, \cdot)$; otherwise, $b_{lj} = 0$. The content of S_l is encapsulated as data packets, which are then split into several segments and queued in the buffer. Afterwards, these segments are transmitted via D2D links in communication network to the receiver of RTU z, where they are integrated into the original packet.

The buffer is used to temporarily record the latest transmission data with limited storage capacity, inside which the storage is divided into M blocks. Each block holds one transmitted data S_l, which is updated as the new data are collected and calculated. Using FIFO, at time step k, once $S_l(k)$ is inserted into the block queue and there are no empty blocks, $S_l(k-M)$ is dropped due to the fixed number of blocks. The buffer variable $B_l = [B_l^\omega, B_l^v]$ is modelled as

$$\begin{aligned} B_l^{\omega,m}(k+1) &= \begin{cases} \varepsilon \cdot ((k-m+2) \cdot L^p - T(k+1)) \cdot B_l^{\omega,m-1}(k), & \text{if } m \neq 0 \\ S_l^\omega(k+1), & \text{if } m = 0 \end{cases} \\ B_l^{v,m}(k+1) &= \begin{cases} \varepsilon \cdot ((k-m+2) \cdot L^p - T(k+1)) \cdot B_l^{v,m-1}(k), & \text{if } m \neq 0 \\ S_l^v(k+1), & \text{if } m = 0 \end{cases} \end{aligned} \tag{3.10}$$

where ε is the Step function and $B_l^m \neq 0$ means that block m contains one data packet and $B_l^m = 0$ means that block m is empty). Without loss of generality, Eqn. (3.10) can deal with buffer overflow and empty buffer according to FIFO strategy.

Afterwards, the RTU z can receive this packet and the receive variable $R_l = [R_l^\omega, R_l^v]$ can be modelled as

$$
\begin{aligned}
R_l^{\omega}(k+1) = \sum_{m=0}^{M} B_l^{\omega,m}(k+1)(\varepsilon \cdot (T_l(k+1) - (k-m) \cdot L^p) \\
- \varepsilon \cdot (T_l(k+1) - (k+1-m) \cdot L^p), \\
R_l^{v}(k+1) = \sum_{m=0}^{M} B_l^{v,m}(k+1)(\varepsilon \cdot (T_l(k+1) - (k-m) \cdot L^p) \\
- \varepsilon \cdot (T_l(k+1) - (k+1-m) \cdot L^p).
\end{aligned}
\tag{3.11}
$$

Then, the control input $x^u = [\Omega, \Phi]$ can be calculated using R_l, which are obtained as

$$
\begin{aligned}
\Omega_z(k+1) &= \Omega_z(k) - h \cdot k_{\omega}^1 (\omega_z(k) - \omega_z^*) - h \cdot k_{\omega}^2 \sum_{z \in \mathcal{N}_{dg}, l \in \mathcal{L}_{d2d}} a_{zl} \left(\frac{P_z(k)}{P_z^*} - R_l^{\omega}(k) \right), \\
\Phi_z(k+1) &= \Phi_z(k) - h \cdot k_{v}^1 (v_z(k) - v_z^*) - h \cdot k_{v}^2 \sum_{z \in \mathcal{N}_{dg}, l \in \mathcal{L}_{d2d}} a_{zl} \left(\frac{Q_z(k)}{Q_z^*} - R_l^{v}(k) \right),
\end{aligned}
\tag{3.12}
$$

where Ω_z and Φ_z are also termed as the auxiliary power variable whose derivative is delivered to the primary control from the secondary control, which is used to compensate for the frequency deviation induced by the droop controller. a_{zl} is the incidence matrix of the DG z, and D2D link l transmitted from its adjacent DGs.

In normal operating scenario, at time step k, the RTU z can receive the latest information $P/P^*(k)$ and $Q/Q^*(k)$ transmitted from adjacent RTUs, and then calculate control input $\Omega_z(k)$ and $\Phi_z(k)$, when $R_l(k)$ is equal to $S_l(k)$. However, in the extreme scenario, the reliability of information exchange for distributed control cannot be ensured, i.e., the latest information $P/P^*(k)$ and $Q/Q^*(k)$ cannot be received by RTU z at time step k. In this case, $R_l(k)$ will be kept as the last received variable $S_l(k')$, which will be used to calculate the control input $\Omega_z(k)$ and $\xi_z(k)$, $k' \leq k$. Take note that $S_l(k')$ is out of date, and thus the $\Omega_z(k)$ and $\xi_z(k)$ are not accurate. Therefore, the performance of control method will be impacted.

To further illustrate the relationships between these variables, a simple example is given in Fig. 3.2, where the time interval between time step k and $k+1$ is defined as h. At time step k^*=3, the $R_l(k^*)$ is expected to be equal to $S_l(k^*)$. However, due to the uncertainty of communication in the extreme scenario, assume that $T_l(k^*) \in [2L^p, 3L^p)$, where L^p is the size of one transmission data which stores one piece of complete exchange information. In this case, only two pieces of complete exchange information are successfully transmitted, and received by the RTU z. That is, the current received transmission variable $R_l(k^*)$ is equal to the value of $S_l(k'=2)$. In this case, the control input $\Omega_z(k^*=3)$ and $\xi_z(k^*=3)$ are calculated using the last received calculated state variables $P/P^*(k'=2)$ and $Q/Q^*(k'=2)$, which will mitigate the performance of control method. To solve this problem, the AW-MRAS algorithm is designed, combined with a search space reduction technique.

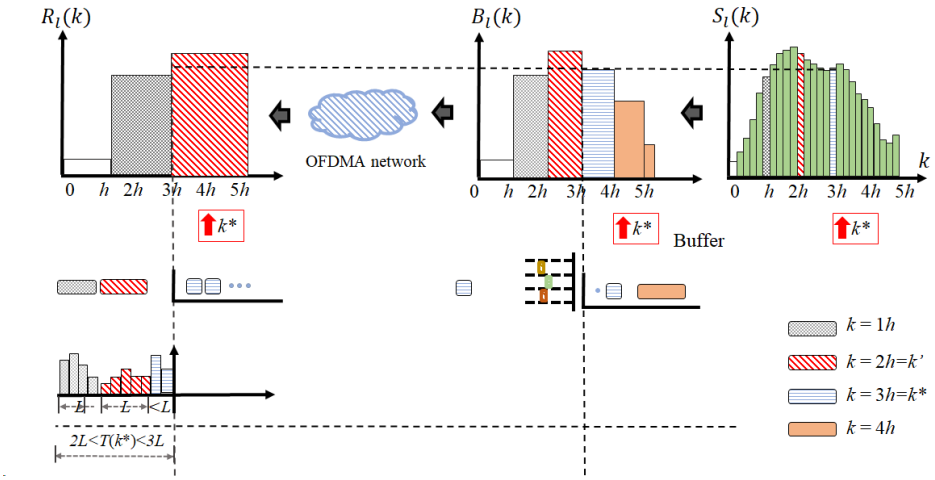


FIGURE 3.2: A simple example to illustrate the transmission variables.

3.3.2 Joint Optimization for D2D Communication Allocation Strategy

In this section, a joint optimization problem is formulated to obtain a tradeoff between the QoS in communication, and the stability in Microgrid. In this joint optimization problem, the QoS index is defined as the energy-efficient of radio resource allocation, which takes both the maximization of the communication system capacity and the minimization of transmit power into consideration at the same time [62, 63]. In addition, the Microgrid stability is defined as the frequency and voltage deviation [108]. Correspondingly, there are three objectives of the proposed formulation: Obtaining the optimal resource allocation strategy with the maximal energy efficiency f_1^c in the cyber layer, and the minimal frequency deviation f_2^ω and voltage deviation f_3^v subject to the secondary control in the physical layer. Based on the power balance and the stability in the Microgrid, and the QoS and the channel constraints in

communication network, this multiobjective problem is stated as

$$
\begin{aligned}
\max_{\mathbf{P}^d,\xi} \quad & f_1^c(\mathbf{P}^d,\xi) = \frac{\mathcal{R}^c(\mathbf{P}^d,\xi)}{\mathcal{P}^c(\mathbf{P}^d,\xi)} \\
\min_{\mathbf{P}^d,\xi} \quad & f_2^{\omega}(\mathbf{P}^d,\xi) = \sum_{k=1}^{T}\sum_{z=1}^{N_{dg}}((\omega_z(k)-\omega^*)) \\
\max_{\mathbf{P}^d,\xi} \quad & f_3^{v}(\mathbf{P}^d,\xi) = \sum_{k=1}^{T}\sum_{z=1}^{N_{dg}}((v_z(k)-v_z^*)) \\
\text{s.t.} \quad & C_1 : \sum_{i=1}^{I_{cha}}\xi_{l,i} \le 1, \qquad \forall\, l \in \mathcal{L}_{d2d}, \\
& C_2 : \xi_{l,i} \in \{0,1\}, \qquad \forall\, l \in \mathcal{L}_{d2d}, \forall\, i \in \mathcal{I}_{cha}, \\
& C_3 : 0 \le p_{l,i}^d < p_{l,i}^{\max}, \quad p_{l,i}^{\max} = \min\left\{ p_{\max}^d, \frac{p_i^c}{\gamma_{th}^i g_{l,i}^d} - \frac{\sigma_0^2}{g_{l,i}^d} \right\}, \\
& C_4 : \underline{V_z} \le V_z(k) \le \overline{V_z}, \\
& C_5 : \underline{P_z} \le P_z(k) \le \overline{P_z},
\end{aligned}
\tag{3.13}
$$

where the energy efficiency metric f_1^c is defined as the ratio of the total throughput $\mathcal{R}$ and the total power consumption $\mathcal{P}$, given in the following Eqn. (3.14) and Eqn. (3.15), respectively. In addition, constrains C_1 to C_3 ensure the QoS of NUE user i and the D2D link l, respectively. C_4 and C_5 represents the power requirements of the Microgrid. γ_{th}^i is the signal noise ratio requirement for the NUE i on the channel i, which satisfies $\gamma_{th}^i \le \gamma_i^c = \frac{p_i^c g_i^c}{\sigma_0^2 + \sum_{l=1}^{\mathcal{L}} \xi_{l,i} p_{l,i}^d g_{l,i}^d}$. The voltage and active power of each DG are required to be inside the allowable voltage margin and active power margin, respectively. Allowable lower and upper bounds of the voltage and active power are defined as $\underline{V_z}$ and $\overline{V_z}$, and $\underline{P_z}$ and $\overline{P_z}$, respectively.

$$
\mathcal{R}^c\left(\mathbf{P}^d,\xi\right) = \sum_{i=1}^{I_{cha}} \mathcal{R}_i^c + \sum_{i=1}^{I_{cha}}\sum_{l=1}^{L_{d2d}} \xi_{l,i}\mathcal{R}_l, \tag{3.14}
$$

$$
\mathcal{P}^c\left(\mathbf{P}^d,\xi\right) = \frac{1}{2}\left(\sum_{i=1}^{I_{cha}} p_i^c + p_{BS}\right) + \sum_{i=1}^{I_{cha}}\sum_{l=1}^{L_{d2d}} \xi_{l,i}\epsilon_l p_{l,i}^d, \tag{3.15}
$$

where $\mathcal{R}_i^c = B^c \log_2\left(1+\gamma_i^c\right)$ is the throughput of NUE i. p_{BS} is the power consumption of the base station, and ϵ is amplifier inefficiency for D2D link l.

Obviously, the feasible solution that can optimize three objective functions for the proposed multiobjective optimization problem simultaneously cannot be obtained [109]. Hence, the concept of the Pareto solution set can be used for this problem.

Definition 3.1. Assume that in a multiobjective optimization problem, $f_i(x)$, $i \in 1,2,3,\cdots,k$ are all to be minimized. In this case, the solution x' can be dominated by another feasible solution x^* if the following conditions can be satisfied:

$f_i(x^*) \leq f_i(x')$ for all $i \in 1,2,3,\cdots,N$;

$f_i(x^*) < f_i(x')$ for at least one $j \in 1,2,3,\cdots,N$.

x^* is defined as the Pareto optimal solution if no other solutions can dominate x^*.

3.4 AW-MRAS Algorithm for Optimal D2D Communication Allocation Strategy

To solve the aforementioned joint optimization problem, the AW-MRAS algorithm is designed, combined with a search space reduction technique. The optimization problem (3.13) is a multiobjective mixed integer nonlinear problem with high complexity, which indicates that the feasible solution set of this problem comprises discrete and nonlinear forms, and requires to balance the weights of different objectives. To solve this problem, the AW-MRAS algorithm combined with a search space reduction technique is designed, which includes a decomposition-based algorithm, and an advanced stochastic optimization technique, MRAS algorithm. The decomposition-based algorithm shows its capability for balancing different objectives with complicated Pareto sets [112]. The MRAS is verified to be effective for solving such an optimization problem with almost no characteristics, such as differentiability and convexity [113]. In addition, the search space reduction technique is designed to accelerate the optimization. This technique can prune the search space wherein the solutions are dominated by other nondominated optimal solutions, based on the characteristics for this proposed CPPS model.

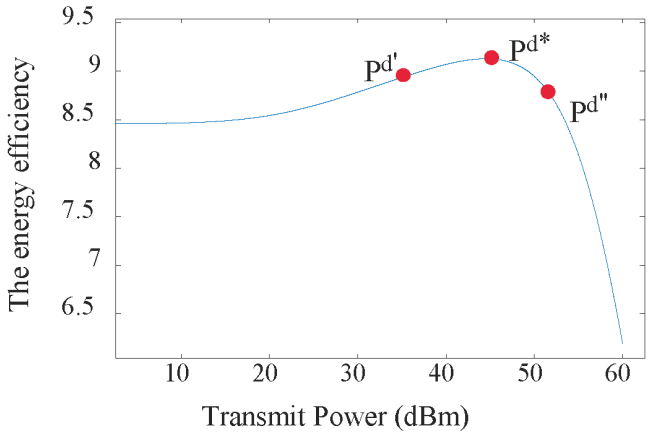


FIGURE 3.3: The correlation between the total energy efficiency and the transmit power for one D2D link in the proposed CPPS.

Algorithm 1: AW-MRAS with search space reduction technique

1. Initialization: Set the initial p.d.f $f(x;\theta_0) \geq 0$ and the sample size S_0. $\lambda \in (0,1]$, ρ, $\varepsilon \geq 0$. Set k=0, K = N_k, $t_k = 0$, $\Phi_R = \varnothing$ and initialize a set of normalized weight vectors $w_0 = (w_0^1, w_0^2, w_0^3)$.
2. Sampling: Generate samples $\{x_k^i\}_{i=1}^S$ from $\tilde{f}(\cdot;\theta_k) = (1-\lambda)f(\cdot;\theta_k) + \lambda f(\cdot;\theta_0)$
3. Sorting: $\{x_k^i\}_{i=1}^S$ = fast-non-dominated-sort ($\{x_k^i\}_{i=1}^S$) based on the fast nondominated sorting approach referring to [110].
4. Pruning:

foreach $i < S$ **do**

Judging whether $x_k^i \in \Phi_R$. If yes, remove this infeasible solution x_k^i; else, continue.

foreach *Communication link in sample* x_k^i **do**

Calculating the new peak point and the infeasible set Ξ according to *Lemma* 3.2.

Adding Ξ into Φ_R .

If the number of remaining generated samples exceeds S_0 or the loop times $t_k < T$, then go to step 5 and t_k is set as 0; else, $t_k = t_k + 1$ go to step Sampling.

5. Evaluating Functions: Evaluate objective functions in the candidate solutions $\{x_k^i\}_{i=1}^S$, and set $\hat{\gamma}_k$ as the sample $(1-\rho)-$ quantile function of $\{\mathcal{J}(x_k^i)\}_{i=1}^S$
6. Reconstructing Sampling set: To reduce the code complexity, an updated sampling set S'' is reconstructed to estimate the parameter θ_{k+1}. The details of constructing S'' can be seen in Appendix A.1.1.
7. Updating: Update parameters by estimating the parameter θ from the sampling set S'' [111]. I is the indicator function.

$$\theta_{k+1} = \arg\max_{\theta\in\Theta} \int_{\mathcal{X}} [\mathcal{S}(\mathcal{J}(x))]^k I_{\{\mathcal{J}(x)\geq\tilde{\gamma}_{k+1}\}} \ln f(x,\theta)\nu(dx),$$

$$w_{k+1}^j = \frac{w_k^j \times (f_k^{j,\max} - f_k^{j,\min})}{\sum_{j=1}^{|w|}(w_k^j \times (f_k^{j,\max} - f_k^{j,\min}))}, j = 1,2,\ldots,|w|.$$

8. Stopping: If $k >= N_k$, then stop; else, set $k := k+1$ and then go to the step Sampling.

3.4.1 Search Space Reduction Technique

The characteristics of the proposed CPPS model are utilized to prune the invalid search space for solving Enq. (3.13). Specifically, with the fixed channel matrix and transmit powers on other channels, the total energy efficiency f_1^c varies as transmit power for a D2D link l on channel i increases, as depicted in Fig. 3.3. Here, $p_{l,i}^{d'}$ and $p_{l,i}^{d''}$ are on the two sides of the peak point $p_{l,i}^{d*}$, which is proved in A.1.3. In addition, given limited communication resources, both frequency deviation f_2^{ω} and voltage deviation f_3^{v} decrease as transmit power increases. This is because as transmit power increases, the total throughput increases, enhancing the quality of information exchange in the

distributed secondary control. This analysis leads to a lemma that underscores the unique features of our CPPS model in solving problems (3.13).

Lemma 3.2. *Given Ξ is a subset of feasible solutions of problem (3.13). If there exists at least one solution $D = \{\mathrm{P}^d, \xi\} \in \Xi$ with element $p^d_{l,i}$ in matrix P^d satisfying $p^d_{l,i} \leq p^{d*}_{l,i}$, and there exists one solution D^Δ with $p^{d^\Delta}_{l,i} > p^{d*}_{l,i}$. In this case, D is not a solution for (3.13), and Ξ is thus not a subset of the Pareto solution set.*

Proof: This lemma is proved by contradiction. Assume that D is a solution in the Pareto solution set for (3.13). In this case, there exists a solution $D^\Delta = \{\mathrm{P}^{d^\Delta}, \xi^\Delta\}$ with the same parameters as D except for the elements $p^{d^\Delta}_{l,i} > p^{d*}_{l,i}$, where the total energy efficiency with $p^{d^\Delta}_{l,i}$ is not lower than $p^d_{l,i}$. In other words, $f^c_1(D) \leq f^c_1(D^\Delta)$. In addition, since $p^{d^\Delta}_{l,i} \geq p^d_{l,i}$, the total throughput of D^Δ is higher than D. By increasing the total throughput, the second and third objectives of the formulated problem are decreasing and thus $f^w_2(D) \geq f^w_2(D^\Delta)$, and $f^v_3(D) \geq f^v_3(D^\Delta)$. Therefore, D^Δ dominates D, and the previous assumption that D is a solution of the Pareto solution set for the problem (3.13) is contradicted.

Hence, according to this *Lemma*, the search space Ξ with $D = \{\mathrm{P}^d, \xi\}$ can be dynamically pruned when there exists the entry $p^d_{n,i}$ of channel matrix P^d, where $p^d_{n,i} \leq p^{d*}_{n,i}$.

3.4.2 Weight Adjustment Technique for Single-Objective Reformulation

To simplify solving the multi-objective optimization problem (3.13), it is transformed into a single-objective function by assigning weights to each objective and aggregating them.

$$x^* = \arg\min_{x \in \mathcal{X}} \ \mathcal{J}(x), \tag{3.16}$$

where x^* is the non-dominated solution, consisting of decision variables $\mathbf{P}^d$ and ξ. $\mathcal{X}$ is the space of solutions and a nonempty compact set. Eqn. (3.16) is the reformulated optimization from the $\mathcal{J}(x) = \sum_i^{|w|} w^i \cdot f_i(x)$, where $f = (-f^c_1, f^\omega_2, f^v_3)$ and $w = (w^1, w^2, w^3)$ is the weight vector, which is updated in Algorithm 1.

Here, MRAS is used to solve the reformulated problem (3.16), including two phases [113]. The first is to generate candidate solutions based on a specified probabilistic model. The second is to update parameters for the previous probabilistic model under the premise of evaluation functions based on previous candidate solutions. Define $\{g_k\}$ as the parameterized sequence of the reference distributions, which is updated by

$$g_{k+1}(x) = \frac{\mathcal{S}(\sum_i^{|w|} w^i f_i(x)) g_k(x)}{\int_{\mathcal{X}} \mathcal{S}(\sum_i^{|w|} w^i f_i(x)) g_k(x) v(\mathrm{d}x)}, \ \forall x \in \mathcal{X}, \tag{3.17}$$

where $\mathcal{S}(\cdot) : \mathbb{R} \to \mathbb{R}^+$ is a positive strictly decreasing function to make sure Eqn. (3.17) is a valid distribution. v represents the counting measure defined on $\mathcal{X}$ [111]. By minimizing the KL divergence between the sampling distribution $f(x;\theta)$ and the referenced distribution $g(x)$, the parameters of the sampling distribution can be updated. By weighting the reference distribution $g_{k+1}(x)$ based on the objective function $\mathcal{J}$, Eqn. (3.17) improves the expected performance as follows:

$$\begin{aligned}\mathbb{E}_{g_{k+1}}[\mathcal{S}(\textstyle\sum_i^{|w|} w^i f_i(x))] &= \int g_{k+1}(x)\mathcal{S}(\textstyle\sum_i^{|w|} w^i f_i(x))\mathrm{d}x \\ &= \frac{\int \mathcal{S}(\sum_i^{|w|} w^i f_i(x))^2 g_k(x)dx}{\int \mathcal{S}(\sum_i^{|w|} w^i f_i(x)) g_k(x)dx} \leq \mathbb{E}_{g_k}[\mathcal{S}(\textstyle\sum_i^{|w|} w^i f_i(x))].\end{aligned} \tag{3.18}$$

Referring to [113], the parameter θ_{k+1} is updated by

$$\theta_{k+1} = \arg\max_{\theta\in\Theta} E_{\theta_k}[\frac{[\mathcal{S}(\mathcal{J}(X))]^k}{f(X,\theta_k)} I_{\{\mathcal{J}(X)\geq\tilde{\gamma}_{k+1}\}} \ln f(X,\theta)]. \tag{3.19}$$

Intuitively, it is difficult to estimate the parameter θ_{k+1} in Eqn. (3.19) using existing optimization tools. To reduce the coding complexity and make the development process faster, the solution of Eqn. (3.19) can be converted into an equivalent problem.

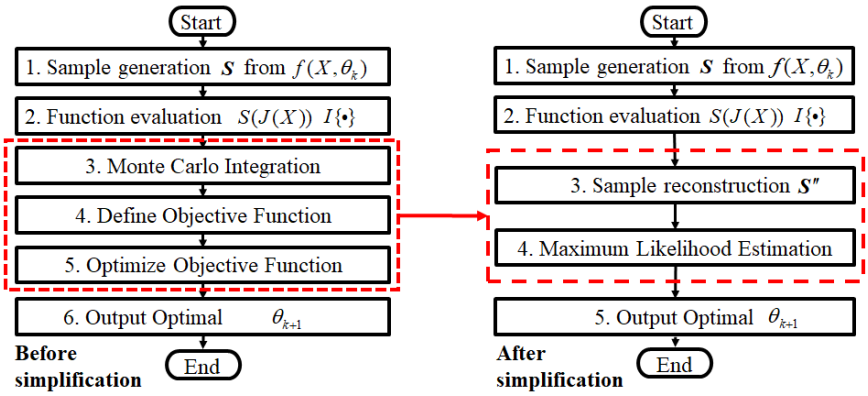


FIGURE 3.4: Simplication of code development.

Lemma 3.3. *The solution of Eqn. (3.19) is equivalent to estimate the parameter θ of p.d.f $f(x,\theta)$ from the reconstructed sample set S''.*

Given the reconstructed sampling set S'' in Appendix A.1.1, the equivalent problem is defined as

$$\theta_{k+1} = \arg\max_{\theta\in\Theta} \int_{\mathcal{X}} [\mathcal{S}(\mathcal{J}(x))]^k I_{\{\mathcal{J}(x)\geq\tilde{\gamma}_{k+1}\}} \ln f(x,\theta)\nu(\mathrm{d}x). \tag{3.20}$$

This equivalent problem, a well-known statistical task of estimating parameters of $f(x,\theta)$ via a sampling set, can be directly solved with existing tools in multiple

development platforms, thereby significantly reducing coding complexity in practice. Refer to A.1.2 for the proof of 3.3.

To provide a clear and concise illustration of the simplification process for the coding process of ***Algorithm 1***, the process is illustrated in Fig. 3.4.

3.4.3 Analysis of Computational Complexity

The computational complexity of the proposed AM-MRAS mainly consists of three parts: 1) Sorting, 2) Pruning, and 3) Calculating the weights. Sorting takes $O(T \cdot |S| \cdot log_2(|S|))$ to sort the non-dominated set, where $|S|$ is the number of generated solutions at one iteration and $T(T \ll |S|)$ is the predetermined maximum number of iterations. Pruning takes $O(T \cdot |S| \cdot n)$, where n is the number of decision variables. Calculating weights takes $O(2 \cdot |w| \cdot (|S| - 1))$, where $|w|$ is the number of the objectives.

Remark: In total, the computational complexity of AW-MRAS is $O(T \cdot |w| \cdot |S| \cdot log_2(|S|))$ in the worst case.

3.5 Simulation Results and Analysis

In this section, the effectiveness of the communication resource allocation proposed strategy is assessed in a test CPPS structure, as shown in Fig. 3.5. Our algorithm was implemented in the PlatEMO v3.0 and Matpower v7.1 on an Intel Core i7 PC with 16 GB of memory. This structure consists of an islanded 33-bus Microgrid including 8 DGs with the hierarchical control, and a communication network supporting the fully distributed communication with 24 D2D links. The detailed parameters are presented in Table 3.1.

TABLE 3.1: Simulation parameters of DGs and the D2D communication.

Parameter	Value	Parameter	Value
Rated Frequency	50 Hz	Rated Active Power	33 kW
Nominal Voltage	220 V	Rated Reactive Power	9 kVAr
P-ω Droop Coeff	$4.5e-5\frac{Hz}{W}$	Q-V Droop Coeff	$1.8e-4\frac{V}{VAr}$
k_ω^1&k_ω^2	1000	k_V^1&k_V^2	100&0.1
Noise power σ_0	-125 dBm	Length of D2D links L	84 bit
Number of channels	18	Channel Bandwidth	180 KHz
Distance d	50-80 m	Maximum power $p^{\max}$	65 dBm
Transmit power p^c	30-50 dBm	SINR requirement γ_{th}	25 dBm

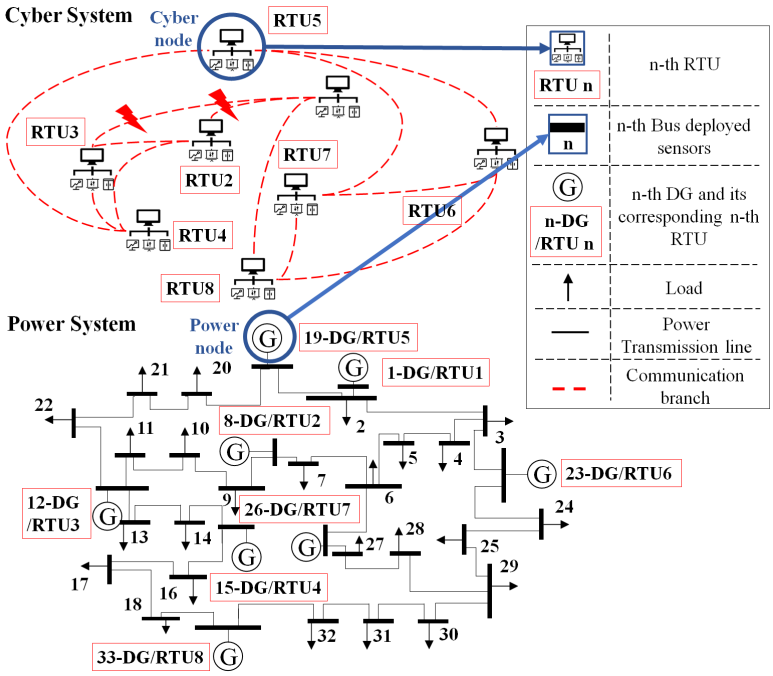


FIGURE 3.5: The CPPS including Microgrid and a communication network.

3.5.1 Evaluation of Optimal Resource Allocation Strategy with AW-MRAS Algorithm

In this section, the proposed AW-MRAS algorithm is used for solving the multiobjective optimization problem (3.13) to achieve the tradeoff between QoS in communication (i.e., the energy efficiency) and the Microgrid stability (i.e., frequency and voltage deviation). Among numerous possible combinations of decision variables for the test system, the Pareto solution set is obtained, wherein three nondominated solutions are selected as shown in Table 3.2, with the optimal objective f_1^c, f_2^w and f_3^v, respectively. Within this set, three critical non-dominated solutions (i.e., technique $A - C$) are selected to capture the characteristics of this set. Each solution within this set is selected for its optimal achievement in a single objective: technique A for f_1^c, technique B for f_2^w and technique C for f_3^v, respectively. technique A is selected for optimizing f_i^c disregarding f_2^w and f_2^v.

TABLE 3.2: Comparisons of different solutions in optimal solution set for the average results within 0.02s.

Strategy	Energy Efficiency f_1^c	Frequency deviation f_2^w	Voltage deviation f_3^v
A	285.51	0.0930004	4.90756
B	186.93	0.0930054	4.89726
C	265.26	0.0929835	4.92

In fact, the main objective of this chapter is to analyze the tradeoff between the energy efficiency and the deviation of the physical states. To intuitively analyze these two

parts, the problem (3.13) can be reformulated as a new multi-optimization problem with objective f_1^c and the reformulated objective $f_2' = f_2^\omega + \Gamma \cdot f_3^v$ with the weight value Γ.

3.5.2 Comparisons with Single-objective Optimization under Cyber and Physical Failure Situations

Due to the high interdependence of components in CPPS, potential failures in both communication network and the Microgrid in the extreme events can damage the operation of the entire system. As a result, this subsection conducts a comparative analysis between two algorithm solutions addressing Eqn. (3.13). The first solution, strategy D, is solved by the proposed AW-MRAS for the reformulated multi-objective optimization, targeting objectives f_1^c and f_2'. It is a non-dominated solution chosen from the Pareto optimal set, specifically selected for its maximal achievement in objective f_1. In contrast, the second solution, strategy E, is solved by MRAS for a single-objective optimization for optimizing QoS index (f_1^c). In the subsequent sections, techniques E and D are analyzed under cyber and physical failure scenarios, respectively.

3.5.2.1 Senario I (Cyber Contingency):

In a real-world scenario, external attackers can utilize exposed cyber vulnerabilities as targets to induce breakage in Microgrid in a cross-layer fashion. Hence, the effectiveness of the proposed strategy under cyber contingency is evaluated for the test system. Initially, the hierarchical control for the CPPS structure in Fig. 3.5 relies on sufficient bandwidth resource. At t = 1 s, malicious attackers send large number of invalid or illegitimate data packages via the communication network; as a result, D2D communication links between DG 1 and DG 2, and DG 1 and DG 3 are interrupted. During $1s \leq t \leq 1.2s$, a 1.5kW and -1kVar load is plugged in bus 3. In this case, a set of Pareto optimal solutions is generated by the proposed AW-MRAS algorithm, wherein the optimal strategy D with the maximal objective f_1 (i.e., energy efficiency) is selected. Comparatively, the optimal strategy E is obtained by the MRAS algorithm for the single optimization problem only considering the energy efficiency. The variations of the Microgrid with strategy D and strategy E are shown in Fig. 3.6, respectively.

The simulation results reveal that the proposed optimal strategy D outperforms the strategy E in frequency and voltage deviation, even though its energy efficiency is slightly compromised. Specifically, in the case with the strategy D, the observed frequency regulation and voltage regulation improvements are 13.74% (0.0041440%/ 0.003

6473%−1) and 4.57% (4.3198%/4.1310%−1) compared with the strategy E, respectively. Clearly, the output frequency and voltage of each DG reach to the reference value and keep close to each other under the strategy D in a shorter time, compared with these variations under the strategy E. In addition, considering QoS index, in the case with strategy D, the energy efficiency decrease is 0.34% compared with the strategy E. That is, the strategy D can effectively mitigate the deviation of physical states with limited communication resources, while maintaining satisfactory communication QoS. Comparatively, even though the optimal strategy E has the optimal energy efficiency, it cannot eliminate the deviation of the physical states. In conclusion, optimal strategy D is more suitable to be selected by operators during such a failure situation. It concludes that, within the context of our proposed multi-objective optimization framework, the solved strategy D offers a better trade-off between physical-layer stability and communication QoS compared to strategies E that focus solely on QoS under the cyber contingency scenario.

3.5.2.2 Scenario II (Power Emergency):

Extreme events might cause damage to transmission and distribution lines. Hence, the effectiveness of the proposed strategy under physical emergency is evaluated for the test system. From t = 0.5 s to 0.7 s, a 15kW and -10kVar load is plugged in bus 27 and the remaining loads in other buses are maintained. In this case, the Pareto optimal set can be obtained by our proposed AW-MARS algorithm, then the optimal strategy D corresponding to the maximal f_1^c can be selected in the Pareto set. Comparatively, the optimal strategy E is computed by the MRAS algorithm for the single optimization problem only considering the energy efficiency.

The variations of Microgrid with the optimal strategy D and the strategy E are shown in Fig. 3.7, respectively. The simulation results reveal that proposed strategy D has a better performance in frequency and voltage deviation, compared with strategy b, even though its energy efficiency is slightly compromised.

More specifically, in the case with the optimal strategy D, the observed frequency regulation and voltage regulation improvements are 6.23% (0.012369%/ 0.011644%−1) and 3.29% (3.9980%/3.8707%−1) compared with the strategy E, respectively. In addition, the strategy D reduce the energy efficiency by 0.78% compared with strategy E. Despite the fact that the strategy E has ideal energy efficiency, it has limited frequency and voltage regulation effectiveness. Overall, strategy D, as identified by the proposed multi-objective optimization framework, offers a more balanced trade-off between system stability and energy efficiency compared to strategies that optimize a single criterion, and is therefore the preferred choice for operators under physical emergency conditions.

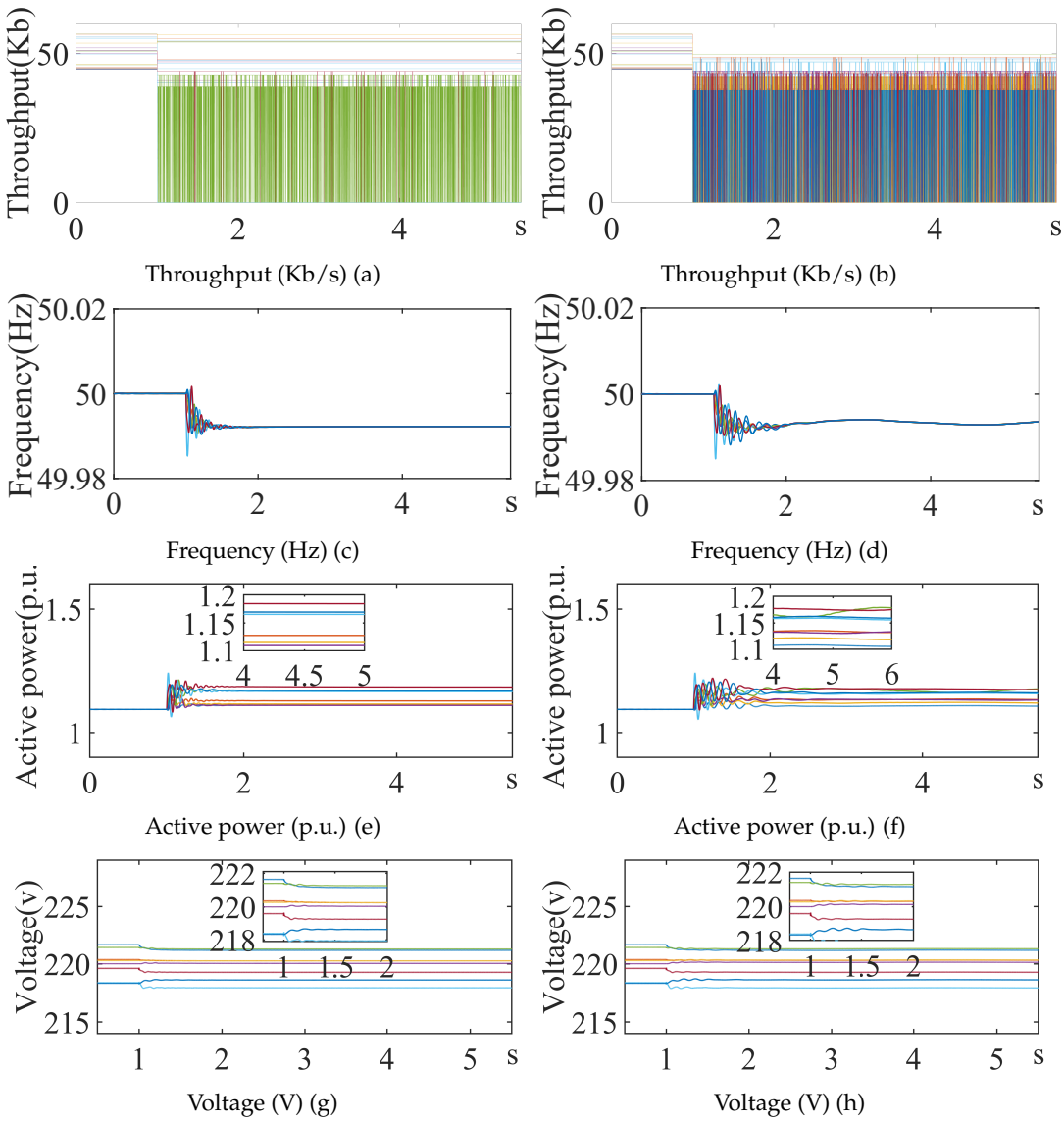


FIGURE 3.6: D2D throughput and microgrid Variations: (a)-(b) D2D throughput (Kb/s), (c)-(d) frequency (Hz), (e)-(f) active power (p.u.), and (g)-(h) voltage (V) under Strategy D and Strategy E, respectively in Scenario I.

3.5.3 Comparison with State-of-the-Art Multi-objective Algorithms

To validate the performance of the proposed AW-algorithm, we select the four multiobjective algorithms, namely, NSGA-II, NSGA-III [114], SPEA2, and MOPSO to compare with the proposed AW-MRAS algorithm. Each of them has run 30 time independently under the two scenarios to mitigate the impact of uncertainties. During each iteration, the number of generation samples is $|S|$=100, and the maximum number of generations is set as 10000. First, nondominated solutions are shown by the five compared algorithms. Then, the generational distance (GD), the hypervolume ratio (HV) and inverted generational distance (IGD) [115] are utilized to evaluate the convergency, diversity and proximity of the obtained solutions, respectively.

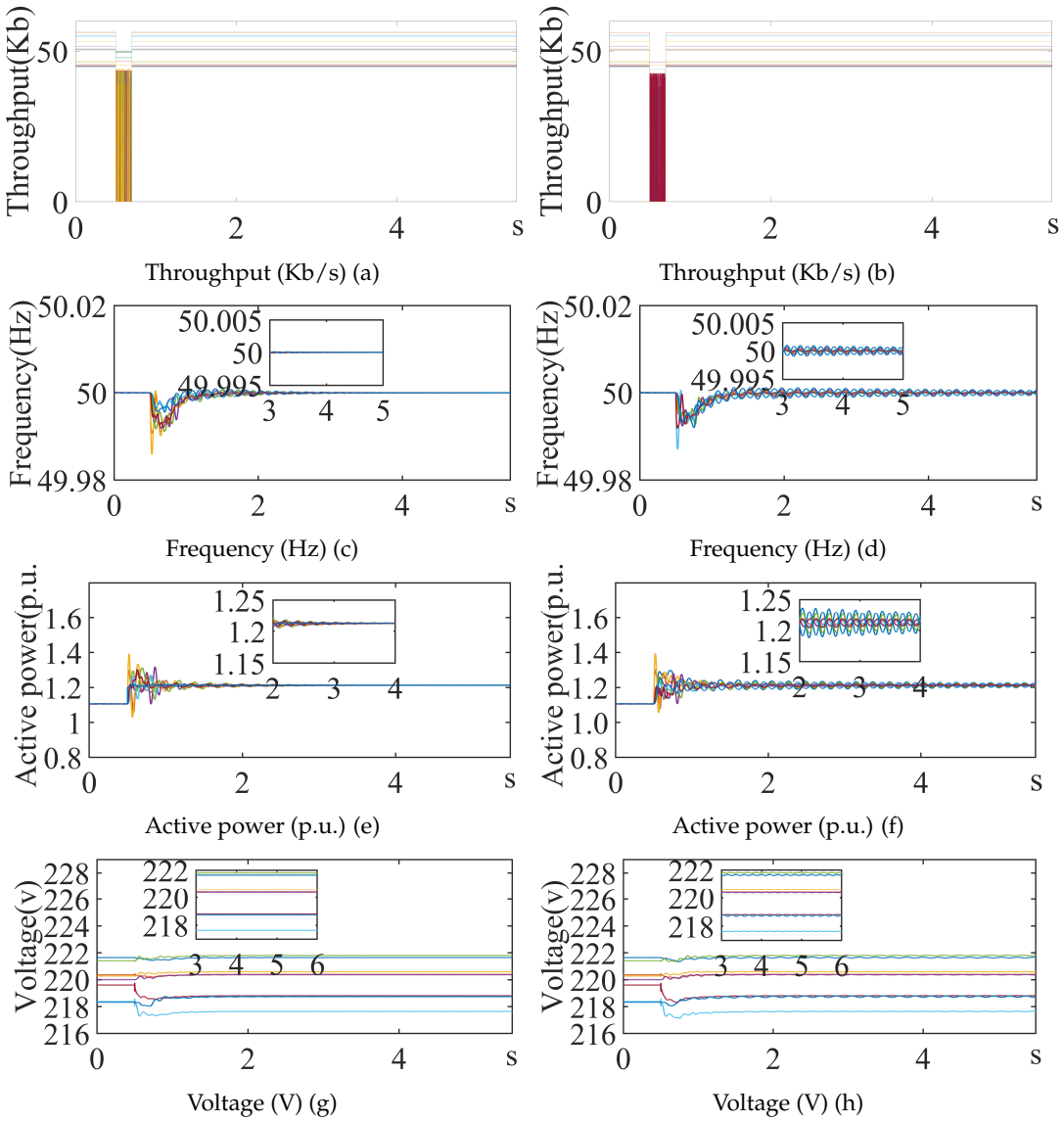
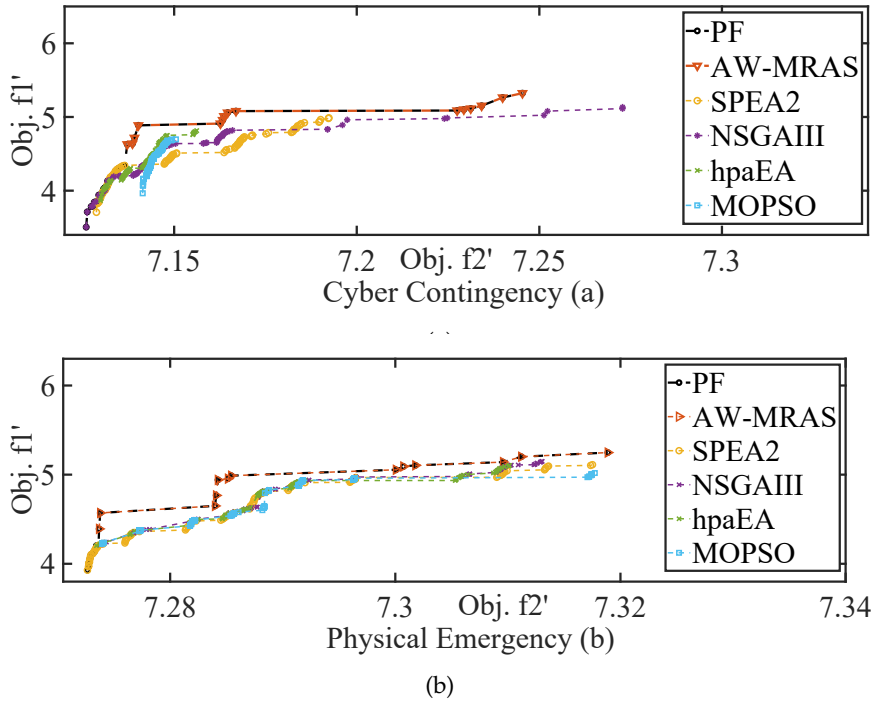


FIGURE 3.7: D2D throughput and microgrid Variations: (a)-(b) D2D throughput (Kb/s), (c)-(d) frequency (Hz), (e)-(f) active power (p.u.), and (g)-(h) voltage (V) under Strategy D and Strategy E, respectively in Scenario II.

The Pareto results for these algorithms are shown in Fig. 3.8. Note that the proposed CPPS model is close to a realistic system, so it is difficult to obtain the true Pareto front. In this case, the nondominated solutions found by the five compared methods are viewed as the near Pareto set. This approximate approach to find nondominated solutions refers to [116], which is marked in Fig. 3.8.

It concludes that the proposed AW-MRAS algorithm can obtain most of the Pareto solutions under these two scenarios with the same times of iterations, even though the diversity of generation samples of AW-MRAS algorithm is not ideal under scenario I.

The GD, HV and IGD results obtained by the five algorithms after 30 independent runs under the two scenarios are illustrated in Fig. 3.9 (a) and Fig. 3.9 (b), respectively.



(b)

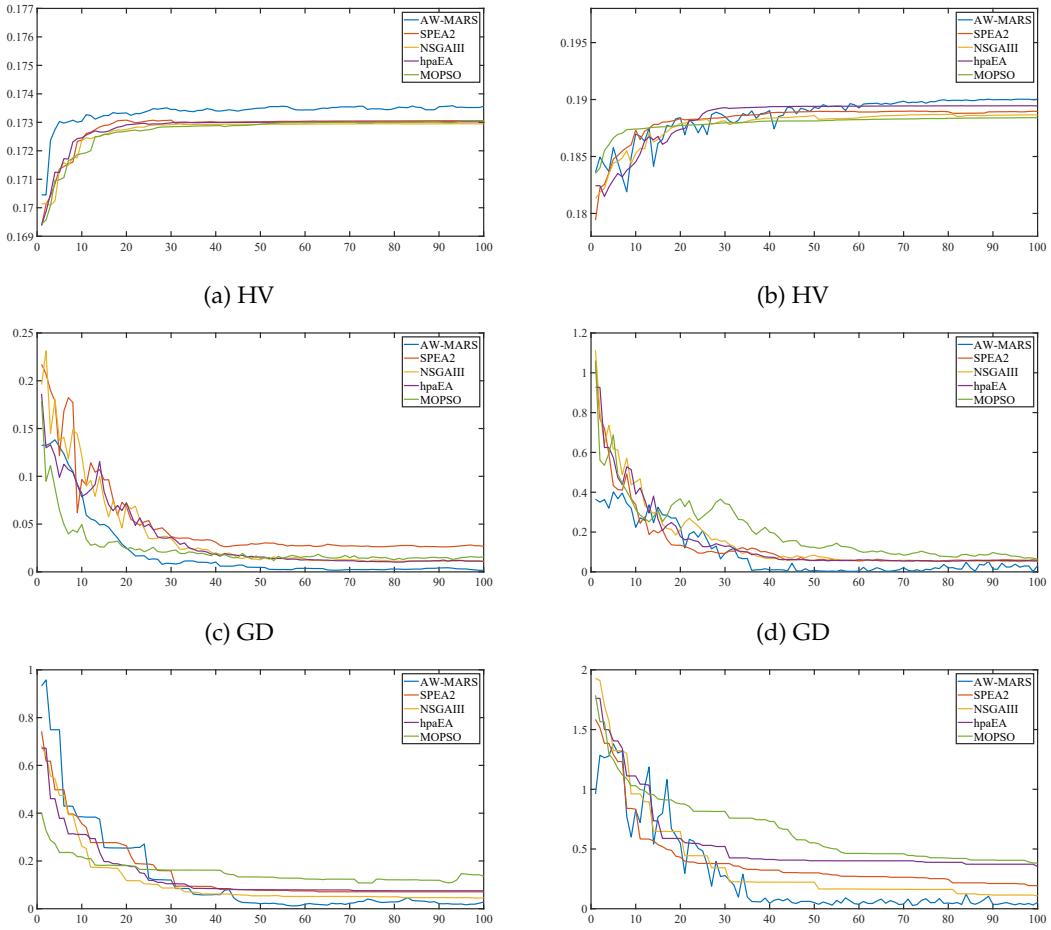
FIGURE 3.8: Pareto results of the compared algorithms under the reformulated objectives f_1^c and $f_2'(f_2^\omega + \Gamma \cdot f_3^v)$, tested in Scenario I (Cyber Contingency) and Scenario II (Physical Emergency)

The trend shows that these algorithms can converge quickly according to these metric values except for MOPSO.

Note that extreme events are unpredictable and destructive, so it seems that purely relying on the allocation strategy is not enough for regulating the deviation of physical states. Even though these deviations induced by extreme events cannot be entirely recovered, mitigating such deviations by adjusting the allocation strategy is also significant. This is because the operators can take more urgent treatments during the short period that the control approach still works.

3.6 Summary

This chapter presents a CPPS model developed to quantitatively assess the cross-layer impacts of communication resource allocation strategies on deviations in physical states. Based on this model, a joint multi-objective optimization problem is proposed to derive the optimal D2D communication resource allocation strategy, with the objectives of co-optimizing QoS and minimizing deviations in physical states. To solve this optimization problem, the AW-MRAS algorithm is proposed, which leverages the unique characteristics of the CPPS model to efficiently prune the search space, thereby enhancing computational efficiency. Simulation results demonstrate that the proposed D2D allocation strategy reduces frequency and voltage deviations



(a) HV (b) HV

(c) GD (d) GD

(e) IGD (f) IGD

FIGURE 3.9: Performance metrics: (a)-(b) present HV, (c)-(d) present GD, and (e)-(f) present IGD for Scenario I (Cyber Contingency) and Scenario II (Physical Emergency), respectively.

by 13.74% and 4.57%, respectively. In addition, the AW-MRAS algorithm demonstrates superior performance compared to state-of-the-art heuristic optimization algorithms, achieving an ideal solution quality in a reduced time. In conclusion, the proposed D2D communication resource allocation strategy effectively mitigates the degraded QoS disruptions and significantly minimizes frequency and voltage deviations within CPPS. By co-optimizing QoS and physical state deviations, the proposed strategy not only enhances the QoS but also ensures that physical systems safe and secure operation. These findings highlight the potential of integrating the proposed communication resource allocation techniques into CPPSs to improve system resilience and overall operational performance.

Chapter 4

A Vulnerability Assessment of Economic Risks in Carbon-Electricity Integrated Trading Systems

4.1 Introduction

The integration of information systems in electricity markets significantly enhances operational efficiency and facilitates market adjustments. However, this integration introduces cybersecurity vulnerabilities that can be exploited by profit-oriented FDIAs to manipulate energy pricing for arbitrage. In addition, in extreme weather, such as the 2019 Texas event, deregulation was implemented to encourage competition in response to a sharp rise in electricity demand, thereby inadvertently creating an unconstrained price gap. In real-world conditions, although extreme weather events are quite infrequent, FDIAs can replicate their impacts on power states by injecting attack vectors. Consequently, this incident underscores the necessity to address potential arbitrage risks within the electricity market.

4.1.1 Overview

Literature [36] proposes the profit-oriented FDIA model, which demonstrates that attackers can make profits by using a buy-low and sell-high LMP strategy in the DA and RT market, respectively. More specifically, adversaries can utilize FDIAs to manipulate LMP in RT market, creating the price gap (i.e., the difference between RT and DA LMP) on a specific virtual transaction. The profit function is then designed as

the price gap per unit trading volume. To clearly explore the attack profits, literature [37] has been conducted to quantify the relationship between the attack vector and the sensitivity of LMP calculation. Furthermore, many research refine the models of the profit function, encompassing features and additional information, such as limited resources due to compromised sensors, as detailed in literature [38]. Literature [39] extends the analysis by incorporating RT market clearing processes. This approach enables the anticipation of market clearing outcomes and facilitates the exploration of the interactions among market participants.

As discussed above, the final arbitrage by attackers is settled by the ISOs in two-settlement markets, which is determined by both the transaction volume and price gap. Our previous work initially investigated adjustable transaction volumes in profit-oriented FDIAs, executing multiple virtual transactions compared to single virtual transactions. This approach allowed for a more comprehensive analysis of the threat posed by attackers. In contrast to transaction volumes, which are actively manipulated by attackers, price gaps are calculated by system operators based on economic dispatch in real-time power states. The LMP calculation mechanism and the accuracy of power states can impact the LMP gap, consequently making it complex to analyze.

As global focus on environmental protection grows, carbon trading mechanisms have been incorporated into traditional electricity markets and LMP calculation mechanisms. Over the past decade, the Paris Agreement has been globally adopted, with its focus on controlling carbon emissions to mitigate the impacts of climate change. Notably, in 2021, the U.S. government issued a target of zero-carbon electricity by 2025 [117]. These instances underscore a marked shift in the LMP calculation: transitioning from being solely based on fossil fuel costs in the generation system to accounting for charges tied to carbon emissions exceedances, also known as carbon-aware LMP. In recent years, carbon-aware LMP mechanisms have been widely discussed to promote both low-carbon and economic operations. The research presented in [41, 42] incorporates the carbon emission cost factors into the economic dispatch problems, treating these as either emission constraints or objective functions. Based on the cap-and-trade principles of the carbon trading market, [19] incorporates carbon emissions within an inequality constraint. The literature [43] also presents a carbon-aware optimal power flow (COPF) model, crafted to derive an operational service pricing framework that encompasses both the Nodal Usage Charge (NUC) and the nodal carbon tax allowance price.

While many existing studies have explored the profit-oriented FDIAs and carbon-aware LMP in the integrated C&E market separately, the security challenges associated with carbon cost considerations remain unaddressed. Integrating carbon emission costs has significantly increased the complexity of LMP calculations, introducing additional constraints for managing carbon cost exceedances and

rendering traditional node vulnerability assessments ineffective. In addition, attackers can exploit these complexities by injecting false data to manipulate LMP calculations [118], disrupt carbon market transactions, and compromise the fairness and security of traditional single-market systems.

4.1.2 Contribution

To address this problem, this chapter conducts the first investigation into the security challenges of the integrated carbon-electricity market. This research emphasizes the novel arbitrage opportunities arising from carbon cost considerations and introduces an advanced framework for assessing the economic risks associated with each power node. This framework provides a foundation for developing robust defense strategies against these emerging threats. To address these problems. The main contributions in this chapter are as follows:

- This chapter proposes an attack model based on the Stackelberg game, which is the first to analyse the threats introduced by carbon emissions in the integrated C&E market, exploring arbitrage opportunities through FDIAs and multi-transaction strategies.
- Under this model, an H-MADDPG algorithm is designed to optimize the attack strategies, leveraging pre-training to enhance convergence and computational efficiency.
- Building upon the previously identified most threatening attack strategy, a novel vulnerability assessment framework is proposed to analyze economic risks induced by carbon emission considerations. The framework is demonstrated on a PJM test system, revealing an up to 201.61 ($/MWh) on a certain transmission line in the PJM test system, compared with the traditional risks assessment only considering electricity costs.

4.1.3 Structure of Chapter

The remainder of this chapter is organized as follows: Section 4.2 illustrates the profit-oriented FDIA and its properties. Section 4.3 illustrates the Stackelberg-based FDIA model and its components: multi-transaction strategies, and LMP with carbon considerations. Section 4.4 presents the optimization of the proposed Stackelberg-Game-Based attack using hierarchical multi-agent reinforcement learning, covering the LMP pre-training and constrained action search. Section 4.5 analyzes economic vulnerability analysis of transmission line under optimal attack strategy. Section 4.6 provides simulation results, including the performance of the proposed

attack strategy, validation of the H-MADDPG algorithm and evaluation of arbitrage opportunities. Finally, Section 4.7 concludes the chapter with key findings.

4.2 Attack Model

In the background section 2.2.3, the stealthiness of FDIA and its impact on PSSE are briefly discussed. This section briefly introduces the profit-oriented FDIA and its properties.

Attackers can exploit electricity markets by executing arbitrage attacks and FDIAs. Specifically, arbitrage attacks enable attackers to sell electricity at high prices and purchase it at low prices, as shown in Fig. 4.1 (a). Moreover, the integration of FDIAs allows attackers to falsify LMP predictions, resulting in additional and unpredictable profits [119], as depicted in Fig. 4.1 (b). Under normal conditions, LMPs in DA and RT markets are predictable. However, attackers can strategically manipulate estimated power states through FDIAs. Such manipulation can induce line congestion, distort LMPs, and widen the price gap between DA and RT markets, ultimately increasing the attack profits.

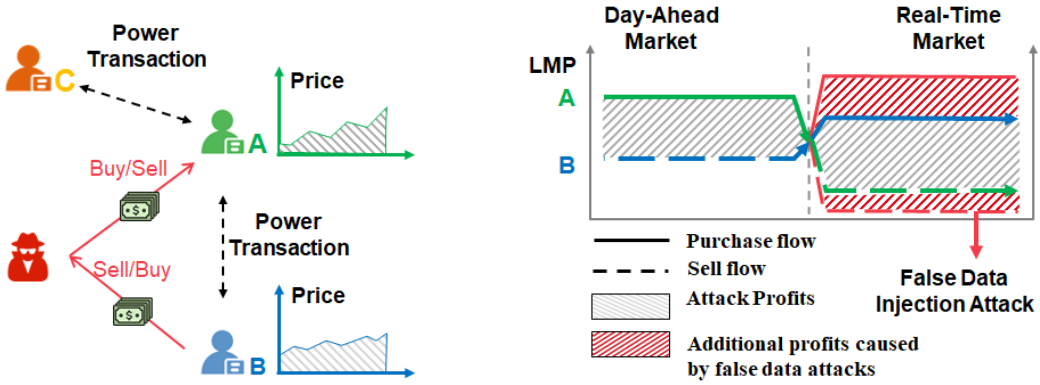


FIGURE 4.1: An example of attacker's participation in transactions between A and B: (a) Transaction flow and (b) Profit extraction.

The process of this attack is summarised as [58]:

1. Engage in virtual power transactions, purchasing and selling a certain amount of virtual power v_a at bus i and j at price λ_i^{DA} and λ_j^{DA}, respectively.

2. Launch FDIA to the measurements and then manipulate the nodal LMPs in the RT market.

3. Buy and sell the same amount power v_a on the bus j and i at price λ_j^{RT} and λ_i^{RT}, respectively.

The attack profits can be given as

$$R^{Profit} = \left(\lambda_i^{RT} - \lambda_j^{RT} + \lambda_j^{DA} - \lambda_i^{DA}\right) \cdot v_a. \tag{4.1}$$

4.3 Stackelberg-Game-based FDIA Model with Multi-transaction and Carbon-Aware Locational Marginal Price

A Stackelberg-based Profit-oriented FDIA is proposed to target the integrated the C&E trading system. It comprises two key components: (i) an attack model that formulates a profitable attack strategy by leveraging FDIAs and multi-transaction arbitrage, and (ii) an LMP calculator embedded within the optimal power flow function in EMS, which considers carbon emission costs.

4.3.1 Stackelberg-Game-Based FDIA Model

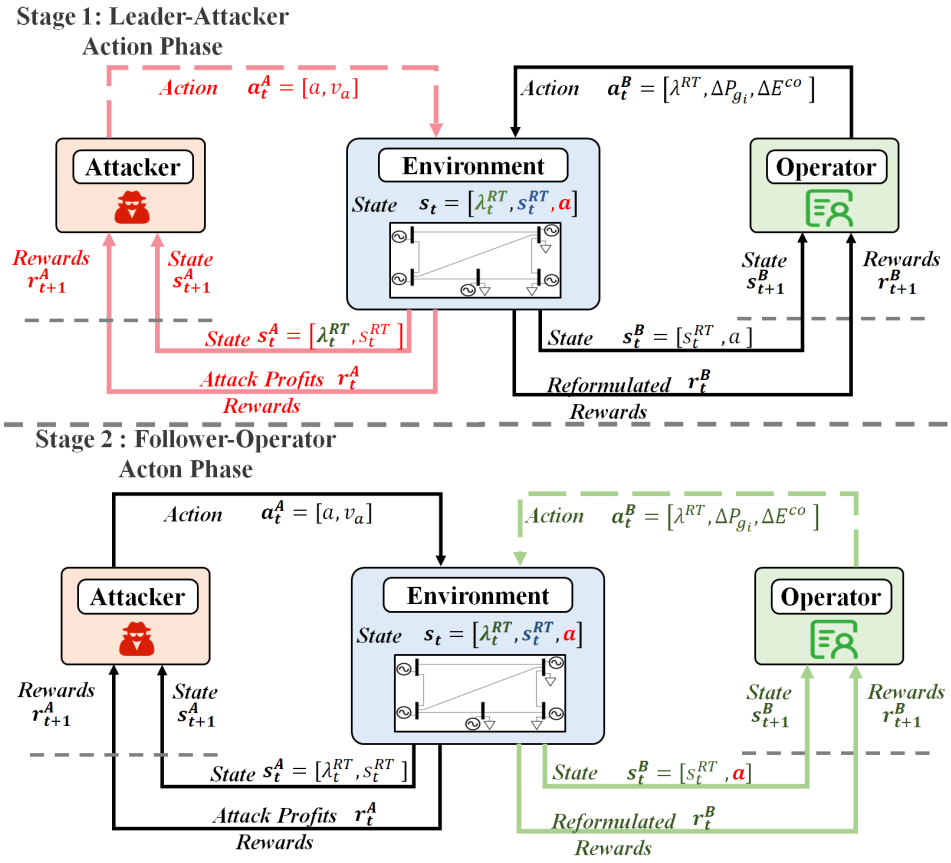


FIGURE 4.2: Leader-Follower stackelberg game.

This Stackelberg game is illustrated in Fig. 4.2. Attackers (i.e., leaders) maximize profits by optimizing their strategy, while system operators (i.e., followers) respond by minimizing costs and simultaneously updating RT market LMPs. These updated values are transmitted to attackers, facilitating iterative optimization in subsequent stages.

In the first stage, attackers, acting as leaders, utilize power states s^{RT} and LMP prices λ^{RT} to execute actions a and v_a. Here, a represents false data injected to system states, while v_a corresponds to purchased power volume. These actions are processed by the environment to calculate rewards $r^A = \mathcal{U}(\cdot)$, defined as the product of the purchased power volume v_a and the price gap $\lambda^{RT} - \lambda^{DA}$ by executing FDIAs a. The computed rewards are then used by attackers to adjust their strategies and determine subsequent actions.

In contrast, in the second stage, system operators, acting as followers, receive the false data a and power states s^{RT}. Operators aim to minimize operational costs by solving an economic dispatch optimization problem $\mathcal{G}(\cdot)$. The actions, including the RT LMP λ^{RT}, the generation output adjustment ΔP_{g_i}, and carbon emission right ΔE^{co}, yield rewards $r^B = \mathcal{G}(\cdot)$ through interaction with the environment. These rewards guide operators in refining their actions to maintain system stability and operational efficiency.

The Stackelberg equilibrium is defined as follows.

$$\begin{aligned}
&(x^{A*}, x^{B*}) = \arg \max_{(x^A, x^B) \in \Omega^A \times \Omega^B} \mathcal{U}(x^A, x^{B*}) \\
&\text{s.t.} \quad x^{B*} = \arg \max_{x^{B*} \in \Omega^B} \mathcal{L}(x^{A*}, x^{B*}), \\
&\Omega^A = \left\{ a_l, \hat{v}^{RT}_{a,l}, \Delta v^{RT}_{a,l} \mid (4.3a) - (4.3d) \right\}, \\
&\Omega^B = \left\{ \Delta P_{g_i}, \Delta E^{co}, \lambda^{RT}_i \mid (4.4a) - (4.4g) \right\}.
\end{aligned} \tag{4.2}$$

4.3.2 Multi-Transaction FDIA Model

In this section, the DC state estimation is discussed since a full AC model is not presently possible in the real world due to its complexity. Based on this, attackers can make profits from two aspects: (i) by compromising a set of measuring devices and manipulating corresponding measurements (i.e., FDIA) and (ii) by engaging in the virtual transaction, wherein they strategically design both the volume of trade and the selection of trading counterparties.

The attacker is assumed to have the following capabilities:

1. The attacker knows the underlying system model (i.e., network topology and power measurements), trading model and price model.

2. The attacker knows the estimated power states $\hat{P}g$ and $\hat{P}d$ given by the RT market.

Contrasting with the conventional attack profit function in (4.1), which calculates the profit from a single transaction between nodes i and j at a fixed trading amount v_a, our proposed model extends this model to calculate the attack profits from multiple transactions at adjustable trading amounts. Specifically, it aggregates the profits from all the potential target transactions $l \in \mathcal{L}$ for the attack. In addition, it assumes that attackers can adjust virtual demand and supply amounts in DA and RT markets, represented by $\hat{v}_a^{DA}$ and $\hat{v}_a^{RT}$, respectively. When $\hat{v}_a > 0$, it indicates attackers sell power, whereas when $\hat{v}_a < 0$, they buy power. The attack profit is modeled as

$$
\begin{aligned}
\max_{a,\hat{v}_a^{RT},\Delta v_a^{RT},\lambda^{RT}} \quad & \mathcal{U}(a,\hat{v}_a^{RT},\Delta v_a^{RT},\lambda^{RT}) && (4.3)\\
& = \sum_{l_{ij}\in\mathcal{L}} \lambda_i^{DA}(-\mathrm{T}_{il})v_{a,l}^{DA} + \lambda_j^{DA}(-\mathrm{T}_{jl})v_{a,l}^{DA} + \lambda_i^{RT}\mathrm{T}_{il}v_{a,l}^{RT} + \lambda_j^{RT}\mathrm{T}_{jl}v_{a,l}^{RT}\\
& = (\lambda^{DA})(-\mathrm{T})v_a^{DA} + (\lambda^{RT})\mathrm{T}v_a^{RT}\\
& = (\lambda^{RT}-\lambda^{DA})\mathrm{T}v_a^{RT} = (\lambda^{RT}-\lambda^{DA})(\mathrm{T}\hat{v}_a^{RT} + \mathrm{T}\Delta v_a^{RT})\\
& = (\lambda^{RT}-\lambda^{DA})(\mathrm{T}\hat{v}_a^{RT} + \mathrm{BKa})\\
\text{s.t.} \quad & \mathrm{T}\Delta v_a^{RT} - \mathrm{BKa} = 0, && (4.3a)\\
& \|(\mathrm{I}-\mathrm{HK})(F^{RT}+\mathrm{a})\|_2 \le \varepsilon, && (4.3b)\\
& F_l^{\min} - v_a^{RT} \le 0, \quad \forall l \in \mathcal{L}, && (4.3c)\\
& v_a^{RT} - F_l^{\max} \le 0, \quad \forall l \in \mathcal{L}. && (4.3d)
\end{aligned}
$$

where $B_{ij} = -1/X_{ij}$ if i and j are connected and $i \neq j$; $B_{ii} = \sum_{j=i,j\neq i}^{N} 1/X_{ij}$ if i and j are connected; otherwise, $B_{ij} = 0$. $\mathrm{BKa} = \mathrm{B}\hat{\mathrm{x}}_{\mathrm{bad}} - \mathrm{B}\hat{\mathrm{x}} = \mathrm{B}\hat{\theta}_{\mathrm{bad}} - \mathrm{B}\hat{\theta}$ represents the deviation in active power injection estimates caused by FDIA. Based on the requirement of the market participant as discussed in 4.3.2.2, $v_a^{DA} = v_a^{RT}$. However, owing to the capacity of FDIA to deceive system operators, leading them to make biased estimations, there is a deviation $\Delta v_a^{RT} = v_a^{RT} - \hat{v}_a^{RT}$ between the settlement trading volume of a_{RT}^T the system operator and the real trading volume $\hat{v}_a^{RT}$ executed by the attacker. Note that the attack profit is settled by the system operator. The attacker must comply with market regulations when buying and selling in both markets according to Eqn. (4.1) with $\hat{v}_a^{RT} = \hat{v}_a^{DA}$. T is a incidence matrix where the element T_{il} represents the relationship between node i and line l. If line l starts at node i and ends at node j, then $\mathrm{T}_{il} = 1$ and $\mathrm{T}_{jl} = -1$. All other elements of T are 0.

4.3.2.1 Price Gap between DA and RT Market

In normal operational scenarios, both the DA LMP and RT LMP can be obtained from (2.7), as discussed in Section 2.2.3.2. Based on accurate load forecasts and stable DA operational conditions, the price gap tends to be sufficiently small to be ignored.

This is mainly due to the secure and efficient dispatch mechanisms that prevent line congestion. However, when attackers intervene, a significant price gap between the DA and RT markets may arise. Specifically, attackers can induce state estimation errors by injecting false data vectors, leading to artificial line congestion, which in turn results in a considerable price gap.

4.3.2.2 Market Requirements for Virtual Bidding Mechanisms

Virtual bidding mechanisms are implemented in market operators, such as ISO-New England, to enhance market competition and liquidity within the electricity markets. Generally, market participants are not necessarily required to have actual generation or electricity consumption. Consequently, if attackers purchase and sell a certain amount of virtual power at bus i in the DA market, they are obligated to sell and purchase an equivalent amount of virtual power at bus i in the RT market, to maintain the stable operation of the power system.

4.3.3 LMP Calculation with Carbon Considerations

In this proposed framework, the attack vector a is injected into RT measurements to compromise the accuracy of the economic dispatch process, thereby manipulating the LMPs. Notably, the objective of the optimal economic dispatch in this section incorporates carbon emission costs, with a comprehensive analysis ensuring alignment with low-carbon system operation principles.

4.3.3.1 Problem Formulation

The carbon-aware LMP calculation models are presented in the following sections.

$$\min_{\Delta P_{g_i}, \Delta E^{co}} \quad \mathcal{G}(\Delta P_{g_i}, \Delta E^{co}) \tag{4.4}$$

$$= \sum_{i=1}^{N} C_i(BKa + \Delta P_{g_i} + \hat{P}_{g_i}) + I_t(\Delta E^{co} + \hat{E}^{co}) \cdot (\Delta E^{co} + \hat{E}^{co})$$

$$\text{s.t.} \quad \sum_{i=1}^{N}(\Delta P_{g_i} + \hat{P}_{g_i}) - \sum_{i=1}^{N}(\hat{P}_{d_i}) = 0, \tag{4.4a}$$

$$\sum_{i=1}^{N}(e_i(\Delta P_{g_i} + \hat{P}_{g_i})) - \alpha \sum_{d=1}^{N}(\hat{P}_{d_d}) - (\Delta E^{co} + \hat{E}^{co}) = 0, \tag{4.4b}$$

$$\Delta P_{g_i}^{\min} - \Delta P_{g_i} \leq 0, \quad \forall i \in \mathcal{N}, \tag{4.4c}$$

$$\Delta P_{g_i} - \Delta P_{g_i}^{\max} \leq 0, \quad \forall i \in \mathcal{N}, \tag{4.4d}$$

$$\sigma F_l^{\min} - \Delta F_l \leq 0, \quad \forall l \in \mathcal{L}, \tag{4.4e}$$

$$\Delta F_l - \sigma F_l^{\max} \leq 0, \quad \forall l \in \mathcal{L}, \tag{4.4f}$$

$$\Delta E^{co} + \hat{E}^{co} - E_i^{max} \leq 0 \quad \forall i \in \mathcal{N}, \tag{4.4g}$$

where $\Delta P_{g_i}^{\max}$ and $\Delta P_{g_i}^{\min}$ are usually chosen to be 0.1 MWh and -2 MWh respectively [38]. $E_i^{max} = \mathrm{CSF}_i \hat{P}_{d_i}$ is the carbon emission quota. Typically, the optimization problem is reformulated as a Lagrangian function, with the resulting Lagrange multipliers used to compute the LMP. While this section employs a reinforcement learning approach rather than the Lagrangian method for problem-solving, the subsequent analysis of nodal economic risks requires the computation of carbon-aware LMP based on Lagrange multipliers. Therefore, this section introduces the Lagrangian function to enable a comparative analysis of carbon-aware LMP and traditional LMP according to Eqn. (2.9).

$$\begin{aligned}\max_{\gamma,\rho,\mu_i^-,\mu_i^+,\eta,\zeta,\iota_i} \quad & \mathcal{L}(\gamma,\rho,\mu_i^-,\mu_i^+,\eta,\zeta,\iota_i) \\ & = \textstyle\sum_{i=1}^{N} C_i(BKa + \Delta P_{g_i} + \hat{P}_{g_i}) \\ & + I_t(\Delta E^{co} + \hat{E}^{co}) \cdot (\Delta E^{co} + \hat{E}^{co}) \\ & + \gamma(\textstyle\sum_{i=1}^{N}(\hat{P}_{d_i}) - \sum_{i=1}^{N}(\Delta P_{g_i} + \hat{P}_{g_i})) \\ & + \rho(\textstyle\sum_{i=1}^{N}(e_i(\Delta P_{g_i} + \hat{P}_{g_i})) - \alpha \sum_{i=1}^{N}(\hat{P}_{d_i}) \\ & - (\Delta E^{co} + \hat{E}^{co})) + \textstyle\sum_{i=1}^{N} \mu_i^-(\Delta P_{g_i}^{\min} - \Delta P_{g_i}) \\ & + \textstyle\sum_{i=1}^{N} \mu_i^+(\Delta P_{g_i} - \Delta P_{g_i}^{\max}) \\ & + \textstyle\sum_{l=1}^{L} \zeta_l(\sigma F_l^{\min} - \Delta F_l) + \sum_{l=1}^{L} \eta_l(\Delta F_l - \sigma F_l^{\max}) \\ & + \textstyle\sum_{i=1}^{N} \iota_i(\Delta E^{co} + \hat{E}^{co} - \mathrm{CSF}_i \hat{P}_{d_i}).\end{aligned} \tag{4.5}$$

To satisfy the Karush-Kuhn-Tucker (KKT) conditions, it yields constraints as

$$\nabla_{\Delta P_{g_i}} \mathcal{L}(\Delta P_{g_i}, \Delta E^{co}, \gamma, \rho, \mu_i^-, \mu_i^+, \eta, \zeta) = 0, \tag{4.5a}$$

$$\nabla_{\Delta E^{co}} \mathcal{L}(\Delta P_{g_i}, \Delta E^{co}, \gamma, \rho, \mu_i^-, \mu_i^+, \eta, \zeta) = 0, \tag{4.5b}$$

$$\mu_i^-(\Delta P_{g_i}^{\min} - \Delta P_{g_i}) = 0, \mu_i^- \geq 0, \quad \forall i \in \mathcal{N}, \tag{4.5c}$$

$$\mu_i^+(\Delta P_{g_i} - \Delta P_{g_i}^{\max}) = 0, \mu_i^+ \geq 0, \quad \forall i \in \mathcal{N}, \tag{4.5d}$$

$$\zeta_l(\sigma F_l^{\min} - \Delta F_l) = 0, \zeta_l \geq 0, \quad \forall l(i,j) \in \mathcal{L}, \tag{4.5e}$$

$$\eta_l(\Delta F_l - \sigma F_l^{\max}) = 0, \eta_l \geq 0, \quad \forall l(i,j) \in \mathcal{L}, \tag{4.5f}$$

$$\iota_i(\Delta E^{co} + \hat{E}^{co} - \mathrm{CSF}_i \hat{P}_{d_i}) = 0, \iota_i \geq 0, \quad \forall i \in \mathcal{N}, \tag{4.5g}$$

where the constraints and can be derived from the Lagrangian function, which is as

$$\nabla_{\Delta P_{g_i}} \mathcal{L} = 2a\Delta P_{g_i} + b - \gamma + \rho \cdot e_i - \mu_i^- + \mu_i^+ + \textstyle\sum_l^M \mathrm{GSF}_{l-i} \left(\eta_l^{RT} - \zeta_l^{RT}\right). \tag{4.6}$$

$$\nabla_{\Delta E_i^{co}} \mathcal{L} = 2\kappa\Delta E^{co} + 2\kappa\hat{E}^{co} - \kappa E^{\min} + I^{\min} - \rho - \iota_i \mathrm{CSF}_i. \tag{4.7}$$

Based on the previous equations, the carbon-aware LMP is

$$\lambda_i^{RT} = \gamma^{RT} - \sum \mathrm{GSF}_{li} \eta_l^{RT} + \sum \mathrm{GSF}_{li} \zeta_l^{RT} - \rho \cdot \alpha - \iota_i \cdot \mathrm{CSF}_i, \tag{4.8}$$

where are $\eta^{RT}, \zeta^{RT}, \rho$ and α are derived from the lower-level economic dispatch problem in (4.5). Compared to the Eqn. (2.9), an additional term $-\rho \cdot \alpha - \iota_i \cdot \mathrm{CSF}_i$ is included, which introduces complexities of the price gap.

4.3.3.2 Price of Carbon Emissions Trading

In economic theory, the electricity price would increase with the increased demand of consumers with the constrained supply. Given the limited carbon emission rights, a rise in the carbon price, corresponding to the growing demand for carbon emission rights, can penalize high-emission users while promoting the users with low emissions [43]. Without loss of generalization, the model for carbon emission pricing can be given as

$$\begin{aligned} I_t &= \frac{I^{\max} - I^{\min}}{E^{\max} - E^{\min}}(\Delta E^{co} + \hat{E}^{co} - E^{\min}) + I^{\min} \\ &= \kappa(\Delta E^{co} + \hat{E}^{co} - E^{\min}) + I^{\min}, \end{aligned} \tag{4.9}$$

where κ denotes the coefficient, quantifying the relationship between changes in carbon emission rights and corresponding changes in the carbon price in the specific power system. When the carbon emission of the whole system does not exceed the allocated limitation, the surplus emission rights can be introduced into the carbon trading market. Consequently, the carbon emission right E is represented as a negative value.

4.4 Optimization of Stackelberg-based FDIA model based on Hierarchical Multi-agent deep deterministic policy gradient algorithm

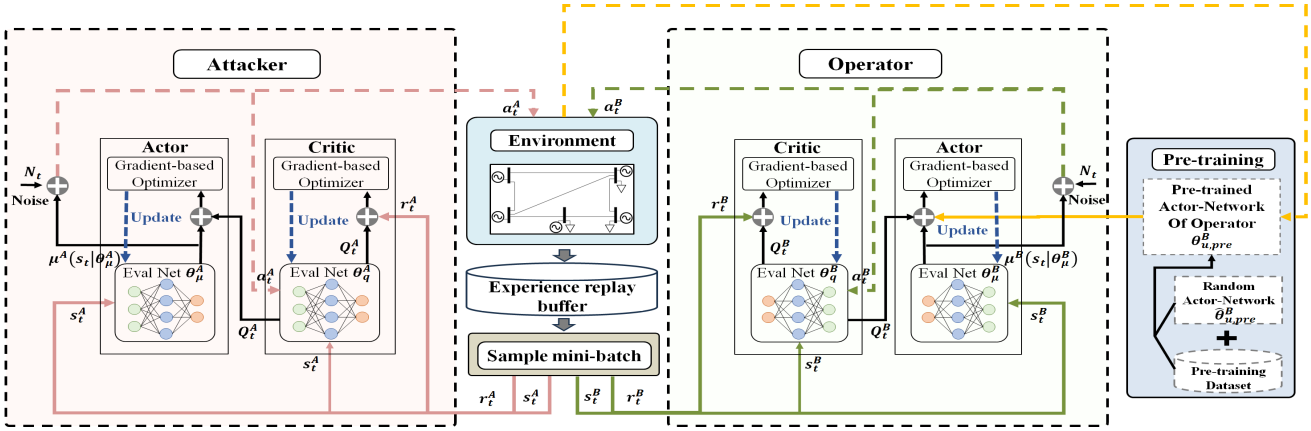


FIGURE 4.3: Hierarchical multi-agent deep deterministic policy gradient algorithm.

To solve the Stackelberg game, H-MADDPG enables hierarchical agents, such as leaders and followers, to focus on their roles while accounting for the strategies and actions of other levels.

4.4.1 Mapping Stackelberg game to Markov Decision Process

The optimization process of interactions between attackers and system operators follows MDP properties, mapping the Stackelberg game to an MDP framework with two types of agents: the attackers as the leader agents and the system operators as follower agents. The main elements of the agents are listed as follows.

• ***State:*** For the leader agent, the power states $s^{RT} = [F_l, \hat{P}_{d_i}, \hat{P}_{g_i}, \hat{E}^{co}]$ and the RT LMP are considered as the environment. The actual power states are observed by attackers, while the RL LMPs are calculated by ISO and provided as the feedback to attackers. Based on the above information, the attacker agent seeks the optimal action a^{A*}. For the follower agent, ISO, can observes the power states and the injected false data from the leader to determine the optimal action a^{B*}.

The states of the leader and the follower are denoted as $s^A = [s^{RT}, c^B]$ and $s^B = [s^{RT}, c^A]$, respectively. s^{RT} represents that environment state. $c^B = \lambda^{RT}$ represents the received message for attackers from operators and $c^A = [a, \hat{v}_a^{RT}, \Delta v_a^{RT}]$ represents the message to attackers.

• ***Action:*** The actions of the leader and follower are $a = [a^A, a^B]$, where $a^A \in \Omega^A$ and $a^B \in \Omega^B$.

• ***Reward:*** In RL, the action space is typically unconstrained and may include actions that fail to meet certain constraints. However, in real-world power system applications, the actions of attackers and operators must satisfy specific physical constraints. To address this challenge effectively, the algorithm restricts actions in two primary ways: by designing specific penalty functions and developing the actor-network architectures in section 4.4.3. This section details the design of two distinct reward functions for the agents. Each function incorporates a penalty function to discourage out-of-bound actions by reducing their values, thereby guiding agents toward high-value strategies within defined physical boundaries.

For a variable x, if the constraint $f(x)$ is bounded by $[x^{\min}, x^{\max}]$, the penalty function is defined as $\phi(\cdot) = \ln \frac{|x-x^{\max}|+|x-x^{\min}|}{(x^{\max}-x^{\min})}$. Similarly, when $f(x) = 0$ is required, the penalty function is defined as $\phi'(\cdot) = |f(x)|^2$.

The constraints in Eqn. (4.3a) and Eqn. (4.3b) are incorporated into the reward function r^A as penalty functions ϕ_1^A and ϕ_2^A to penalize violations of these constraints, respectively. The constraints in Eqn. (4.3c) and Eqn. (4.3d) are directly enforced by

applying restrictions to the action outputs of the leader's actor-network, ensuring that they satisfy these constraints, as detailed in Eqn. (4.17). The immediate reward function for attackers is

$$r_t^A = \mathcal{U}(s_t^A, a_t^A) - \omega_a^A \cdot \phi_1^A - \omega_b^A \cdot \phi_2^A. \tag{4.10}$$

The constraints in Eqn. (4.4c), (4.4d) and (4.4g) are implemented by applying restrictions to the action outputs of the actor-network of the follower, as detailed in Eqn. (4.22). Subsequently, the constraints in Eqn. (4.4a)-(4.4b) and (4.4e)-(4.4f) are formulated as penalty functions and incorporated into the reward function, respectively, which is defined as

$$r_t^B = \mathcal{G}(s_t^B, a_t^B) - \omega_a^B \phi_a^B - \omega_b^B \phi_b^B - \omega_e^B \phi_e^B - \omega_f^B \phi_f^B, \tag{4.11}$$

where ω^A and ω^B are the penalty coefficients. If all constraints are satisfied, the leader and follower agents can obtain actions that correspond to high-value rewards.

4.4.2 Supervised Pre-training for the Actor-Network of the Follower

Effective feature capture is crucial for RL algorithms. Pre-training, particularly using supervised learning techniques, can improve the quality of feature extraction. However, the lack of historical attack data limits the availability of labeled data for supervised pre-training of the leader model. In contrast, the follower model is based on the well-studied optimal economic dispatch problem in power systems. Existing simulators can generate optimal dispatch solutions and calculate LMPs. These outputs can be utilized to construct a pre-training database, enabling supervised learning-based pre-training for the follower model. This section details the pre-training process for the follower model, emphasizing its ability to leverage pre-training datasets to improve performance. This enhanced performance subsequently facilitates the training of the leader model through interactions.

4.4.2.1 Data Generation and Labeling

This subsection outlines the process of using the Matpower simulator to generate the pre-training dataset, thereby pre-training the actor-network of the follower. The pre-training dataset is constructed based on the output of the optimal power flow function in Matpower simulator, defined as $\mathrm{y} = u_{sim}(s^{RT}) = \mathrm{y} = [\Delta P_g, \Delta E^{\mathrm{co}}, \lambda^{RT}]$. Based on this, the input-output pairs for the dataset are defined as

$$\mathcal{D} = \{(s_t^{RT}, \mathrm{y}_t)\}_{n=1}^{N_{pre}}. \tag{4.12}$$

Afterward, a neural network $u^B_{pre}(:;\theta^B_{u,pre})$ based on Long Short-Term Memory (LSTM) is designed to predict the target vector $\hat{y}$, which is as

$$\hat{y}_t = u^B_{pre}(s^{RT}_t;\theta^B_{u,pre}). \tag{4.13}$$

4.4.2.2 Pre-training

The LSTM-based neural network $u^B_{pre}(:;\theta^B_{pre})$ is trained by minimizing the error between predicted $\hat{y}_i$ and actual output y_i using mean squared error (MSE) loss fucntion:

$$L_{pre}(\theta^B_{pre}) = \frac{1}{N_{pre}}\sum_{i=1}^{N_{pre}}(\hat{y}_i - y_i)^2. \tag{4.14}$$

To update the model parameters θ^{pre}, stochastic gradient descent (SGD) is applied as $\theta^B_{pre,t} = \theta^B_{pre,t} - \eta^{pre}\nabla_{\theta^B_{pre,t}} L_{pre}(\theta^B_{pre,t})$, where η^{pre} is the learning rate.

In multi-agent systems, where the actions of each agent are influenced by the decisions of others, the rapid learning of one agent can accelerate the training time to achieve the optimal policy. As the follower agent quickly finds an optimal action from the pre-training in the initial training as $a^B_{ini} = u^B_{pre}(:;\theta^B_{pre})$, the leader can adjust their actions by referencing their behavior, creating a feedback mechanism that promotes faster overall learning.

4.4.3 H-MADDPG with Constrained Action Search

H-MADDPG is employed to address the Stackelberg game problem proposed in Section 4.3.1 due to its two primary advantages: 1) its capability to optimize policies within continuous action spaces, and 2) its effectiveness in managing interactions between agents with limited information about each other. The fully distributed characteristics of MADDPG are particularly well-suited for scenarios involving attackers and operators, where these agents can independently optimize their policy function based on partial observations.

4.4.3.1 Attacker Architecture

The objective of attackers is to optimize the policy function μ^A_t to maximize the final profits:

$$\begin{aligned} J^A_t(\theta^A_{\mu,t}) &= \mathbb{E}_{s^A_t\sim\mu^A_t, a^A_t\sim\mu^A_t}[Q^A_t(s^A_t, a^A_t)] \\ &= [Q^A_t(s^{RT}_t, \lambda^{RT}_t, a_t, \hat{v}^{RT}_{a,t}, \Delta v^{RT}_{a,t})]. \end{aligned} \tag{4.15}$$

• *Training of Actor-Network:* A Transformer-based neural network parameterized by $\theta^A_{\mu,t}$ to capture the attackers' actions is defined as $a^A_t = u^A_t(s^A_t; \theta^A_{u,t}) = \text{Transformer}(s^A_t; \theta^A_{u,t})$, which is updated using the Monte-Carlo method as

$$\nabla_{\theta^A_{\mu,t}} J^A_t(\theta^A_{\mu,t}) \approx \frac{1}{N} \sum\nolimits_{n \in N} \{\nabla_{\theta^A_{\mu,t}} Q^A_t(s^A_t, a^A_t)\}. \tag{4.16}$$

where the parameter $\theta^A_{\mu,t}$ is updated by the Adam optimizer $\theta^A_{\mu,t} = \theta^A_{\mu,t} + \eta^A \nabla_{\theta^A_{\mu,t}} J^A_t(\mu^A_t)$.

Targeted Adjustment of Action Range: The intermediate output of actor-network is $\mathbf{h}^A \in \mathbb{R}^{B \times d_K \times 1}$, where $d_K = 3L$ and each element is denoted as $\mathrm{h}^A[b,k,0]$. According to Eqn. (4.3d) and (4.3c), $\mathrm{h}^A[:,k,0]$ is contained within $[F^{min}_k, F^{max}_k]$, where $k \in (L+1:2L)$. To implement this constraint, the layer normalization $\text{Norm}(\cdot)$ is used to normalize the $\hat{\mathrm{h}}^A[b,k,0] = \text{Norm}(\mathrm{h}^A[b,:,0])_{k=L+1:2L}$. Next, the tanh activation function is used to constrain the output $\hat{\mathrm{h}}^A[:,k,0]$ within $[-1,1]$. Subsequently, a linear transformation is applied to map the $\hat{\mathrm{h}}^A[b,k,0]$ to $[F^{min}_i, F^{max}_i]$ by the following function:

$$a^A[b,k,0] = \frac{(F^{max}_k - F^{min}_k)}{2}(\tanh(\hat{\mathrm{h}}^A[b,k,0]) + 1) + F^{min}_k, \quad \forall k \in \mathcal{L}. \tag{4.17}$$

• *Training of Critic-Network:* The critic network is trained by minimizing MSE in estimating the Q-function Q^A_t:

$$L^A(\theta^A_{q,t}) = \frac{1}{N} \sum\nolimits_{n \in N} (r^A_t - Q^A_t(s^A_t, a^A_t; \theta^A_{q,t}))^2, \tag{4.18}$$

where $Q^A_t(s^A_t, a^A_t; \theta^A_{q,t}) = \text{LSTM}(s^A_t \oplus a^A_t; \theta^A_{q,t})$. r^A_t is updated from the Eqn. (4.10). and θ^A_q is updated as

$$\theta^A_q \leftarrow \theta^A_q - \alpha \nabla_{\theta^A_q} L^A(\theta^A_q). \tag{4.19}$$

4.4.3.2 Independent System Operator Architecture

The objective of system operators is to optimize the policy function μ^B_t to minimize the optimal economic dispatching in Eqn. (4.4):

$$\begin{aligned} J^B_t(\theta^B_{\mu,t}) &= \mathbb{E}_{\mathcal{R}^B_t \sim \mu^B_t, a^B_t \sim \mu^B_t}[Q^B_t(s^B_t, a^B_t) - \left\| a^B_t - u^B_{pre}(s^B_t; \theta^B_{pre}) \right\|^2] \\ &= [Q^B_t(s^{RT}_t, a_t, \hat{v}^{RT}_{a,t}, \Delta v^{RT}_{a,t}, \Delta P_{g_t}, \Delta E^{co}_t, \lambda^{RT}_t) - \left\| a^B_t - u^B_{pre}(s^B_t; \theta^B_{pre}) \right\|^2]. \end{aligned} \tag{4.20}$$

• *Training of Actor-Network:* The actor-network is modeled as a Transformer parameterized by $\theta^B_{\mu,t}$ to capture the operators' actions as

$a_t^B = u_t^B(s_t^B; \theta_{u,t}^B) = \text{Transformer}(s_t^B; \theta_{u,t}^B)$, which is updated as:

$$\nabla_{\theta_{\mu,t}^B} J_t^B(\theta_{\mu,t}^B) \approx \frac{1}{N} \sum_{n \in N} \{\nabla_{\theta_{\mu,t}^B} Q_t^B(s_t^B, a_t^B)\}, \tag{4.21}$$

where the parameter $\theta_{\mu,t}^B$ is updated by $\theta_{\mu,t}^B = \theta_{\mu,t}^B + \eta^B \nabla_{\theta_{\mu,t}^B} J_t^B(\mu_t^B)$.

Targeted Adjustment of Action Range: $\mathbf{h}^B \in \mathbb{R}^{B \times d'_K \times 1}$, where $d'_K = 2N + 1$ and each element is denoted as $\mathrm{h}^B[b, k, 0]$. According to Eqn. (4.4c) and (4.4d), $\mathrm{h}^B[b, k, 0]$ is contained within $[\Delta P_{g_k}^{\min}, \Delta P_{g_k}^{\max}]$, where $k \in (1 : N)$. Then, $\mathrm{h}^B[b, N+1, 0]$ is constrained within $[0, E_i^{\max} - \hat{E}^{co}]$ where $i \in \mathcal{N}$ according to the constraint in Eqn. (4.4g). $\mathrm{Norm}(\cdot)$ is used to normalize the $\hat{\mathrm{h}}^B[b, k, 0] = \mathrm{Norm}(\mathrm{h}^B[b, :, 0])_{k=1:N+1}$. Next, a linear transformation is applied to map the $\hat{\mathrm{h}}^B[b, k, 0]$ to satisfy these constraints using the formula:

$$\begin{aligned} a^B[b, k, 0] &= \frac{(\Delta P_{g_k}^{\max} - \Delta P_{g_k}^{\min})}{2} (\tanh(\hat{\mathrm{h}}^B[b, k, 0]) + 1) + \Delta P_{g_k}^{\min}, \quad \forall k \in \mathcal{L}. \\ a^B[b, N+1, 0] &= \frac{E_{co} - E_{\max,i}}{2} (\tanh(\hat{\mathrm{h}}^B[b, N+1, 0]) + 1), \quad \forall i \in \mathcal{N}. \end{aligned} \tag{4.22}$$

• *Training of Critic-Network:* The critic network is trained by minimizing MSE in estimating the Q-function Q_t^B:

$$L^B(\theta_t^{q^B}) = \frac{1}{N} \sum_{n \in N} (r_t^B - Q_t^B(s_t^B, a_t^B))^2, \tag{4.23}$$

where $Q_t^B(s_t^B, a_t^B; \theta_{q,t}^B) = \text{LSTM}(s_t^B \oplus a_t^B; \theta_{q,t}^B)$. r_t^B is updated from the Eqn. (4.11). and θ_q^B is updated as

$$\theta_q^B \leftarrow \theta_q^B - \beta \nabla_{\theta_q^B} L^B(\theta_q^B). \tag{4.24}$$

4.5 Economic Vulnerability Assessment of Transmission Line

Traditional FDIA requires certain conditions to be stealthy for BDD. Moreover, to induce errors in state estimation that could lead to incorrect operational commands, such as load shedding, more stringent conditions are necessary, such as a sufficiently large magnitude of the attack. However, considering scenarios where attackers derive profit-oriented FDIAs, even minimal-intensity FDIAs that sufficiently relax congestion in transmission lines can be profitable. Consequently, profit-oriented FDIA poses a greater risk. Particularly after the introduction of carbon emission constraints, the congestion caused by these constraints could further expand the manipulation scope for attackers, thereby increasing the feasibility and potential benefits.

This section discusses the economic risk of each transmission line l while attackers engage in RL-market trading between nodes i and j. The risk analysis introduces, for

the first time, the risks caused by carbon emission constraints, which are more applicable to a future low-carbon-oriented societal framework.

4.5.1 Vulnerability Analysis under Attack Profit Model

A model for market traders to profit from the DA market and the RT market on the trading between bus i and j is formulated as

$$\begin{aligned}\text{Payoff}_l &= \text{Payoff}_l^{RT} - \text{Payoff}_l^{DA} \\ &= (\lambda_i^{RT} - \lambda_j^{RT}) - (\lambda_i^{DA} - \lambda_j^{DA}) \cdot v_a^{RT} \\ &= \textstyle\sum_{l\in\mathcal{L}} \eta_l^+ (\text{GSF}_{l,j} - \text{GSF}_{l,i}) + \sum_{l\in\mathcal{L}} \zeta_l^- (\text{GSF}_{l,i} - \text{GSF}_{l,j}) \\ &- (\iota_i \text{CSF}_i - \iota_j \text{CSF}_j) - (\lambda_i^{DA} - \lambda_j^{DA}) \cdot v_a^{RT}.\end{aligned} \tag{4.25}$$

Given that the subsequent discussion focuses on the sign, v_a^{RT} is set to 1 for simplification. Herein, the payoff is always positive if the subsequent three conditions are satisfied [120]:

$$\lambda_i^{DA} - \lambda_j^{DA} < 0, \tag{4.25a}$$

$$F_l < {F_l}^{\max}, \forall l \in \left\{\text{GSF}_{l,j} - \text{GSF}_{l,i} < 0\right\}, \tag{4.25b}$$

$$F_l > {F_l}^{\min}, \forall l \in \left\{\text{GSF}_{l,i} - \text{GSF}_{l,j} < 0\right\}, \tag{4.25c}$$

$$\hat{E}^{co} < \text{CSF}_i - \Delta E^{co}, \forall i \in \mathcal{N}. \tag{4.25d}$$

Specifically, when $\text{GSF}_{l,j} - \text{GSF}_{l,i} < 0$, if $F_{l,t} > F_l^{\min}$, the constraint Eqn. (4.5e) is slack. Therefore, η_l^+ is set to zero, and $\sum_{l\in\mathcal{L}} \eta_l^+ (\text{GSF}_{l,j} - \text{GSF}_{l,i}) \geq 0$. Similarly, $\sum_{l\in\mathcal{L}} \zeta_l^- \left(\text{GSF}_{l,i} - \text{GSF}_{l,j}\right) \geq 0$. In addition, $\iota_i = 0$ while $\hat{E}^{co} < \text{CSF}_i - \Delta E^{co}$. Hence, the traders earn a positive profit from markets.

4.5.2 Vulnerability Analysis under Attack Profit Model with FDIAs

In addition to meeting the conditions specified in Eqn. (4.25a)-(4.25c) to conduct profits via virtual bidding in markets, attackers can also gain more profits than ordinary traders by injecting false data. This manipulation leads to a differential in payoff, which is expressed as

$$\begin{aligned}\text{Payoff}_l &= \text{Payoff}_{l,att}^{RT} - \text{Payoff}_l^{RT} \\ &= (\lambda_{i,att}^{RT} - \lambda_{j,att}^{RT}) - (\lambda_i^{RT} - \lambda_j^{RT}) \\ &= \Sigma_{l\in\mathcal{L}}(\eta_{l,att}^{+} - \eta_l^{+})(\text{GSF}_{l,j} - \text{GSF}_{l,i}) \\ &+ \Sigma_{l\in\mathcal{L}}(\zeta_{l,att}^{-} - \zeta_l^{-})(\text{GSF}_{l,i} - \text{GSF}_{l,j}) \\ &+ (-\iota_{i,att} + \iota_i)\text{CSF}_i + (\iota_{j,att} - \iota_j)\text{CSF}_j.\end{aligned} \tag{4.26}$$

Here, the discussion is restricted to the case where $\text{GSF}_{l,j} - \text{GSF}_{l,i} < 0$, under which $\eta_{l,att}^{+} = 0$ and $\iota_{i,att} = 0$. To ensure Eqn. (4.26) is always positive, there are three cases: 1) $\zeta_{l,att}^{-} = \zeta_l^{-} = 0$ and $\eta_l^{+} > 0$; 2) $\zeta_{l,att}^{-} > \zeta_l^{-} > 0$; 3) $\zeta_{l,att}^{-} > 0$ and $\zeta_l^{-} = 0$.

Case 2 requires that the constraint in Eqn. (4.5e) is tight both with FDIAs and without FDIAs. Furthermore, these constraints must be even tighter when under attack. However, effectively implementing such measures in real-time scenarios proves challenging. Case 3 requires that the originally tight constraint in Eqn. (4.5e) is relaxed subject to FDIAs. Hence, if $\text{GSF}_{l,j} - \text{GSF}_{l,i} < 0$, the FDIA necessitates relaxing Eqn. (4.5e) to ensure a positive payoff. Similarly, when $\text{GSF}_{l,j} - \text{GSF}_{l,i} > 0$, the FDIA requires the relaxation of Eqn. (4.5f) to ensure a positive payoff.

In addition, if the following conditions are satisfied, the constraints caused by carbon emissions are always positive: 1) $\iota_{j,att} = \iota_j = 0 \; and \; \iota_i > 0$; 2) $\iota_{j,att} > \iota_j > 0$; 3) $\iota_{j,att} > 0 \; and \; \iota_j = 0$. In practice, both case 1 and case 2 necessitate the constraint in Eqn. (4.5g) for j to be tight and relaxed, respectively, both with and without the FDIAs. However, the profits from case 1 are minimal, and case 2 requires a stronger tightness post-attack, which is challenging to achieve. Therefore, case 3 is the preferred condition for attackers.

To ensure that attackers gain benefits, the following constraints, based on Eqns. (4.25a)–(4.25d), need to be satisfied:

$$\begin{cases} (4.25a), (4.25d) \\ F_l > F_l^{\min} \; and \; F_l = F_l^{\min}, \forall l \in \left\{\text{GSF}_{l,j} - \text{GSF}_{l,i} < 0\right\}, \\ F_l < F_l^{\max} \; and \; F_l = F_l^{\max}, \forall l \in \left\{\text{GSF}_{l,j} - \text{GSF}_{l,i} > 0\right\}, \\ F_{l,att} < F_l^{\max}, \\ F_{l,att} > F_l^{\min}. \end{cases} \tag{4.27}$$

4.6 Simulation Results and Analysis

The proposed multi-transaction profit-oriented FDIA is tested on a modified version of PJM 5-bus system. All simulations are conducted using MATLAB 2022 and Python on a PC with an Intel Core i7-10750H CPU (2.60 GHz) and 32 GB of RAM.

The attack profits of various strategies are analyzed and compared under two scenarios: with and without consideration of carbon emissions. In addition, the unit attack profit for each node is analysed to guide the vulnerability assessment for defenders.

4.6.1 System Setup

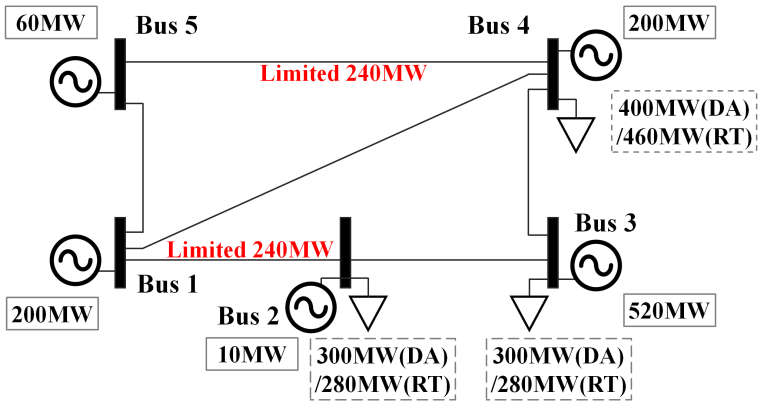


FIGURE 4.4: The modified version of PJM 5-bus test system.

The system parameters and topology of the PJM 5-bus system are detailed in [121]. To facilitate an intuitive comparison of the carbon allocation constraints for each bus, one generator was moved from bus 2 to bus 3. In addition, the load capacities for buses 2, 3, and 4 were configured in both the DA and RT markets in Table 4.1, and the limits of the generation capacity were specified for each plant.

TABLE 4.1: Load values for RT and DA markets at different times.

Time (t)	Market	Bus 2 (MW)	Bus 3 (MW)	Bus 4 (MW)
$t = 5h$	RT	2.2308	2.2308	3.0308
	DA	2.1808	2.1808	2.9808
$t = 10h$	RT	3.0516	3.0516	3.8516
	DA	3.0016	3.0016	3.8016
$t = 16h$	RT	3.0070	3.0070	3.8070
	DA	2.9570	2.9570	3.7570

4.6.2 Validation of the H-MADDPG algorithm with Supervised pre-training

The proposed algorithm is investigated on the Modified version of PJM test system with the actual load profile of NYISO [122] for January 2023 with 15-minute time intervals. As shown in Fig. 4.5, the proposed algorithm significantly improves

performance in two aspects: the quality of the rewards and the efficiency of training time.

Hyperparameters Configurations: The proposed Actor-Critic model utilizes a Transformer-based Actor-network and an LSTM-based Critic-network as detailed in Table 4.2. The model is trained using mini-batches of 64 instances. The Actor-network employs 6 Transformer encoder layers with 2 attention heads, while the Critic-network uses an LSTM architecture with a hidden dimension of 128. The model is trained using mini-batches of 64 instances and conducted over 60 epochs, with each epoch comprising 40960 instances.

TABLE 4.2: Hyperparameters Configurations.

Hyperparameters	**Value**	**Hyperparameters**	**Value**
Batch size	64	No. of epochs	60
Learning rate	1e-4	Instances per epoch	40,960
Transformer-based Actor Network			
No. of encoder layers	6	No. of attention heads	2
Optimizer	Adam	Hidden dimension	128
LSTM-based Critic Network			
Hidden dimension	128	Optimizer	Adam

In the multi-agent system modeled as a Stackelberg games, the independent objective functions of the leader and follower introduce challenges in identifying equilibrium points. The iterative updating of actions by both agents depends on the responses of the counterpart, adding complexity to the interaction and exploration of the solution space. Within this framework, the follower model integrates both discrete and continuous feature learning, which can pose significant challenges when performance is not ideal. These limitations in the follower model significantly impede the exploration capabilities of the attacker with the interactions. Pre-training the follower model effectively enhances its initial performance, thereby improving the leader's ability to navigate the solution space, as shown in Fig. 4.5. This enhancement enables the leader to communicate precise, real-time messages to the follower model, significantly accelerating the convergence rate of the entire system model.

4.6.3 Comparison of Our Proposed Attack Strategy with Other Strategies in Carbon and Non-Carbon Scenarios

This section presents a comparative analysis of the proposed multi-transaction and carbon-aware FDIA strategy against other attack strategies, including the multi-transaction arbitrage, the single-transaction FDIA and the single-transaction

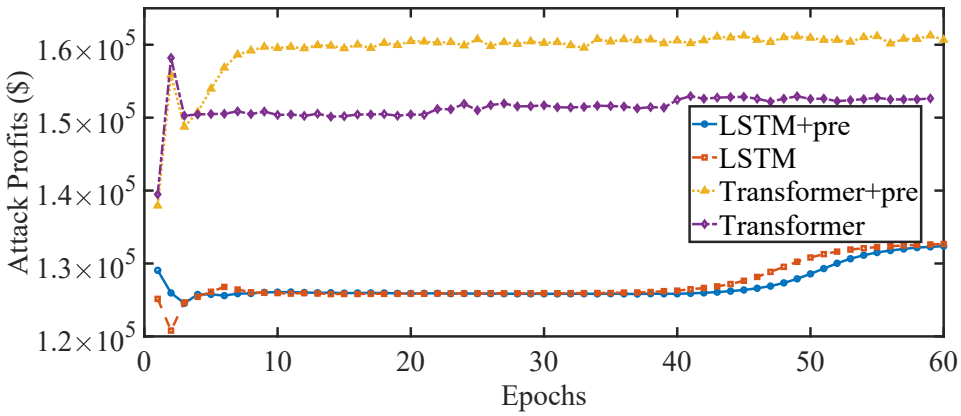


FIGURE 4.5: Validation curves.

arbitrage under two scenarios: Scenario I (with carbon emission considerations) and Scenario II (without carbon emission considerations). The features of each attack strategy are detailed in Table 4.3.

TABLE 4.3: Features of various attack strategies.

Attack Strategy	Multi-Transaction	FDIA
Proposed Strategy (A)	+	+
Multi-Transaction Arbitrage (B)	+	-
Single-Transaction FDIA (C)	-	+
Single-Transaction Arbitrage (D)	-	-

In both scenarios, Strategy A consistently outperforms Strategies B, C, and D, especially as load increases. This is analyzed from three aspects.

4.6.3.1 Impact of Transaction Types (Multi-transaction or Single-transaction)

Multi-transaction strategies (A and B) allow attackers to engage in multiple transactions, fully exploiting price gaps across the market. For example, Strategy A outperforms Strategy C by 2.06×10^4\$ with carbon costs and 1.13×10^4\$ without carbon costs at time step 6h, indicating that participating in multiple transactions significantly increases profit potential.

4.6.3.2 Impact of FDIA

The introduction of FDIA allows attackers to further exploit market vulnerabilities by manipulating power states, which leads to more drastic price fluctuations and greater arbitrage opportunities. As a result, Strategy A exceeds Strategy B by 4.78×10^4\$ and

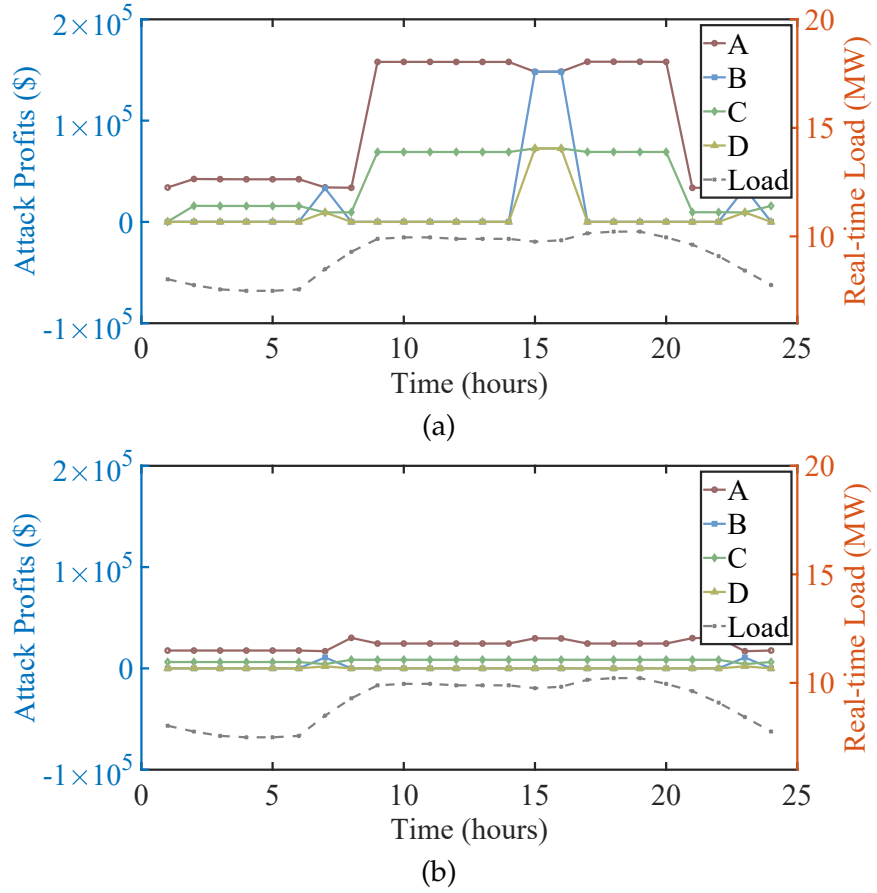


FIGURE 4.6: Comparison of various strategies in scenario I (a) and II (b).

1.97×10^4\$ with carbon costs and without carbon costs at time step 6h, respectively. It shows that FDIA expands the space for price manipulation, thereby increasing profits.

4.6.4 Arbitrage Opportunities induced by Carbon Emission Costs

As illustrated in Fig. 4.6, attackers can achieve higher profits with carbon emission costs compared to without such carbon costs. This section provides a detailed analysis of the arbitrage opportunities arising from carbon emission costs under two scenarios: with and without carbon emissions considerations.

4.6.4.1 Arbitrage Opportunities in Scenario I

The LMPs for each bus in the DA and RT market with carbon cost considerations are presented in Fig. 4.7 (a), along with real-time LMP calculation parameters at three-time points (t = 5 10 and 16) in Table 4.4. To analyze the DA and RT price gap induced by attackers under different loads, the LMPs at three time points (t = 5, 10, and 16) were compared, which are illustrated in Fig. 4.9.

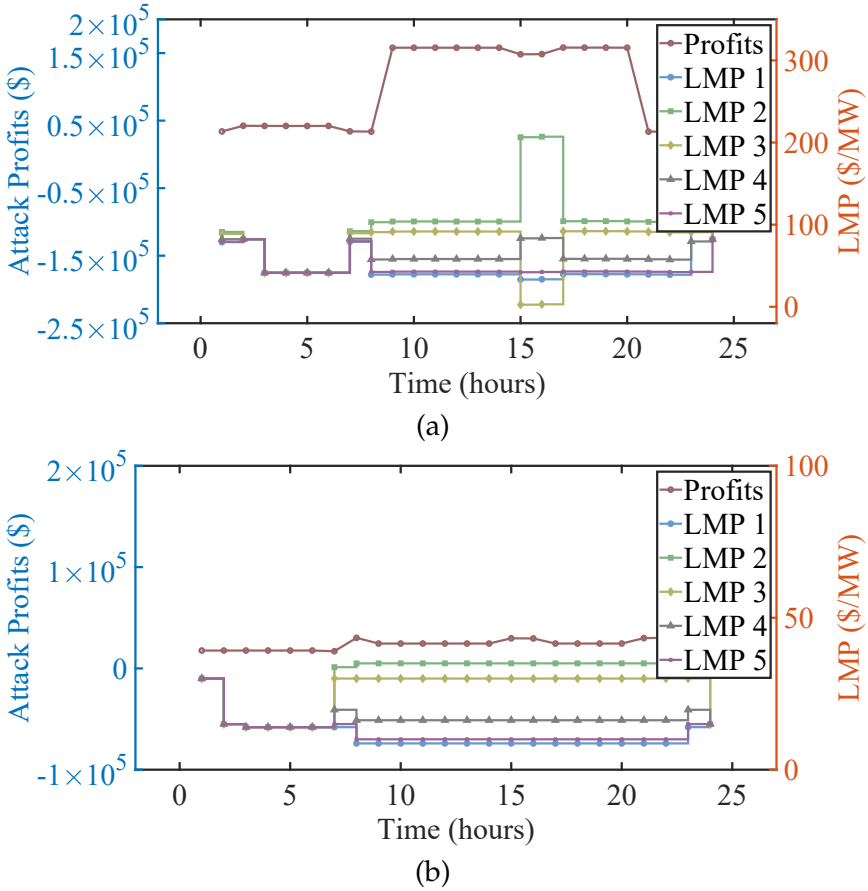


FIGURE 4.7: Comparison of LMPs in scenario I (a) and II (b).

In the DA market at $t = 5$, congestion on transmission line 1 results in a significant increase in the LMP at Bus 2. Conversely, in the RT market, the relaxation of this constraint, with only the carbon constraints remaining strict, which results in consistent LMPs across all buses. At $t = 10$, congestions on both transmission line 1 and the carbon constraint at Bus 3 raise LMPs in the DA market, while the relaxation of these constraints in the RT market reduces the LMP at Bus 2 and increases it at Bus 3, thereby creating arbitrage opportunities. By $t = 16$, congestion on transmission lines 1 and 6, along with the carbon constraint at Bus 3, elevates LMPs in the DA market, whereas the relaxation of line 6 in the RT market induces an LMP increase at Bus 2.

4.6.4.2 Arbitrage Opportunities in Scenario II

The LMPs for each bus in both DA and RT markets under the proposed attack strategy are shown in Fig. 4.7 (b), along with real-time LMP calculation parameters at two-time points ($t = 5$ and 10) in Table 4.5. At $t = 5$h, in the DA market, transmission lines 1 and 6 are congested. The attacker exploits this by injecting false data, misleading operators into perceiving these congested lines as relaxed, thereby creating arbitrage opportunities. At $t = 10$h, the attacker relaxes the previously congested line

TABLE 4.4: LMP comparisons between DA and RT markets in scenario I.

T	Bus	LMP_{RT}^{ATT}	γ_{RT}	$-\sum GSF\eta$	$\sum GSF\zeta$	$-\rho\alpha$	$-\iota CSF$	LMP_{DA}	γ_{RT}	$-\sum GSF\eta$	$\sum GSF\zeta$	$-\rho\alpha$	$-\iota CSF$
5h	1	41.0877	+	-	-	+	-	36.4367	+	+	-	+	-
	2	41.0877	+	-	-	+	-	96.9367	+	+	-	+	-
	3	41.0877	+	-	-	+	-	85.4741	+	+	-	+	-
	4	41.0877	+	-	-	+	-	53.9520	+	+	-	+	-
	5	41.0877	+	-	-	+	-	39.5419	+	+	-	+	-
10h	1	39.5058	+	+	-	+	-	33.2406	+	++	-	+	-
	2	103.778	+	+	-	+	-	206.5691	+	++	-	+	-
	3	91.6008	+	+	-	+	-	2.6325	+	++	-	+	+
	4	58.1133	+	+	-	+	-	83.4208	+	++	-	+	-
	5	42.8046	+	+	-	+	-	42.1368	+	++	-	+	-
16h	1	39.4063	+	+	-	+	-	47.9635	+	+	+	+	-
	2	103.556	+	+	-	+	-	92.6081	+	+	+	+	-
	3	91.4021	+	+	-	+	-	90.1134	+	+	+	+	-
	4	57.9783	+	+	-	+	-	83.2528	+	+	+	+	-
	5	42.6988	+	+	-	+	-	42.0125	+	+	+	+	-

TABLE 4.5: LMP comparisons between DA and RT markets in scenario II.

T	Bus	LMP_{RT}^{ATT}	γ_{RT}	$-\sum GSF\eta$	$\sum GSF\zeta$	LMP_{DA}	γ_{RT}	$-\sum GSF\eta$	$\sum GSF\zeta$
5h	1	14	+	-	-	14.0000	+	+	+
	2	14	+	-	-	29.4609	+	+	+
	3	14	+	-	-	30.0000	+	+	+
	4	14	+	-	-	31.4825	+	+	+
	5	14	+	-	-	10.0000	+	+	+
10h	1	8.6479	+	+	-	16.9774	+	-	+
	2	34.9911	+	+	-	26.3845	+	-	+
	3	30.0000	+	+	-	30.0000	+	-	+
	4	16.2745	+	+	-	39.9427	+	-	+
	5	10.0000	+	+	-	10.0000	+	-	+

6, altering the LMP calculation structure and enabling arbitrage opportunities between the DA and RT markets.

Consequently, without carbon constraints, the attacker can only manipulate power states to obscure the transition of transmission line constraints from congested to relaxed or mask relaxed constraints as tightened. However, when carbon emission costs are considered, the attacker can additionally obscure carbon constraints, thereby expanding the potential scope for LMP manipulation.

4.6.5 Vulnerability Analysis based on Carbon Emission Costs

Without an attack strategy, fluctuations between the DA and RT markets are minimal, resulting in negligible differences in system constraints and a relatively small price gap at t = 10h and t = 16h, both with and without carbon considerations, as shown in Fig. 4.8 (a) and Fig. 4.9 (a).

Under attack strategy A, attackers can alter measurements and then mislead operators to perceive the constraints on transmission line 1 as relaxed instead of constrained. This manipulation results in price gaps, as depicted in Fig. 4.8 (a) and Fig. 4.9 (a).

In scenarios where carbon costs are not considered, at t = 10 and t = 16, attackers can easily alter measurements, misleading operators to perceive the constraints on transmission line 1 as relaxed instead of constrained. This manipulation results in price gaps, as shown in Fig. 4.9. With minor load fluctuations, and transitioning from the DA market to the RT market, it is difficult for attackers to induce price gaps without launching the FDIA.

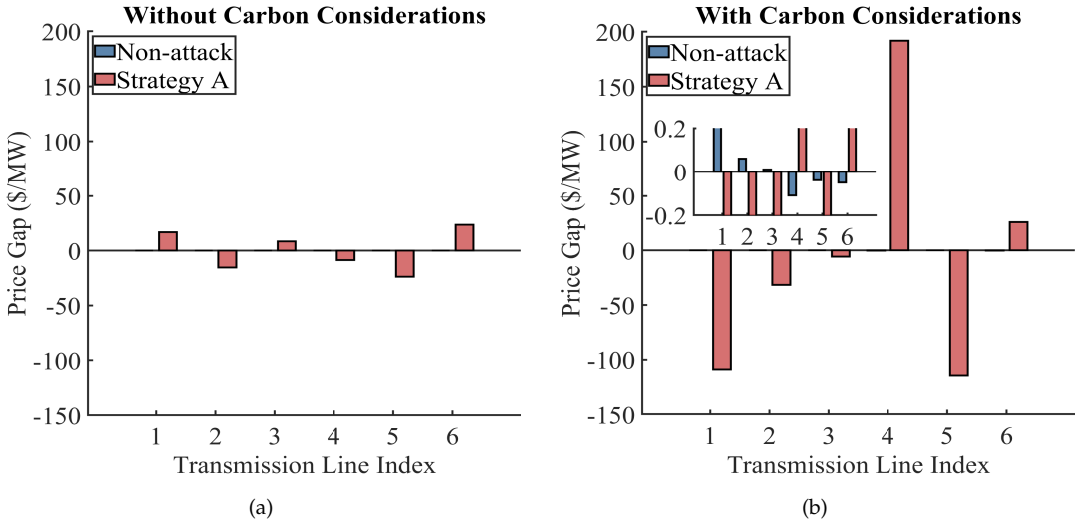


FIGURE 4.8: Vulnerability analysis for each transmission line at t = 10h without (a) and with (b) carbon emission consideration.

As depicted in Fig. 4.8 (b), when there is no attack, the carbon emission-related LMP constraints and power transmission constraints in both the DA and RT markets remain unchanged due to minimal load fluctuations. Consequently, the price gap remains relatively consistent under no-attack conditions. In contrast, when strategy A is implemented, it causes relaxation of carbon emission-related LMP constraints at bus 3, significantly increasing the vulnerability on this bus. This results in a large price gap in transactions involving transmission line 4 between bus 3 and bus 2.

At t = 16h, in both attack and non-attack scenarios, the previously tightened constraints in the DA market change to relax in the RT market due to load

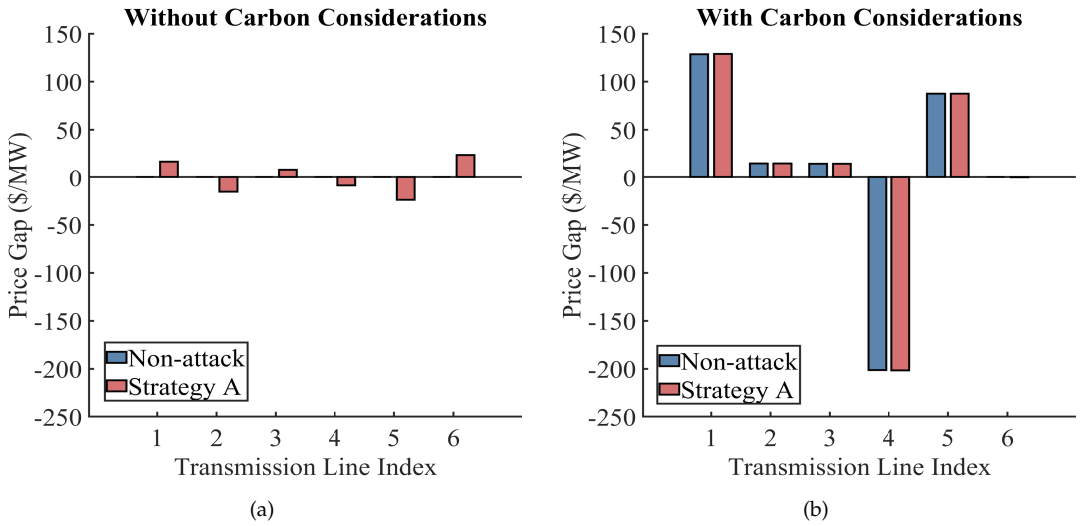


FIGURE 4.9: Vulnerability analysis for each transmission line at t = 16h without (a) and with (b) carbon emission consideration.

fluctuations. This results in a large price gap in transactions involving transmission line 4 between bus 3 and bus 2. Consequently, the space of constraints that attackers can manipulate (i.e., originally tightened constraints misrepresented as relaxed) is nearly non-existent. Consequently, under both attack and no-attack scenarios, the price gap remains relatively unchanged.

The analysis concludes that when carbon emissions are not considered, transmission lines 5 and 6 are vulnerable due to the tightened constraints on transmission lines 1 and 6 in the DA market, which attackers can manipulate to appear relaxed from the operators' perspective in the RT market. In contrast, when carbon emissions are considered, the vulnerability of transmission line 4 increases, necessitating prioritized protection due to the non-critical facilities at bus 3 and the reduced allocation of carbon emissions allowances. Notably, this vulnerability provides opportunities for pure arbitrage attackers, even without FDIAs.

4.7 Summary

In conclusion, this chapter represents the first attempt to broaden the scope of exploring profit-oriented FDIAs, extending beyond the traditional LMP calculation that only focuses on electricity costs to include carbon emission costs. A Stackelberg-Game-based profit-oriented FDIA model is proposed, which can analyze the attack behavior and the interaction with ISO operators through LMP calculations. To optimize attack strategies, an H-MADDPG algorithm is developed, employing pre-training to improve convergence rates and computational efficiency. Building upon the most threatening attack strategy identified, a novel vulnerability assessment

framework is proposed to evaluate economic risks associated with carbon emissions. Simulation results reveal that this framework can identify increases in economic losses of up to $201.61/MWh on certain transmission lines, highlighting the heightened vulnerabilities introduced by integrating carbon considerations into electricity markets. This research underscores the significant economic risks posed by carbon emission considerations and highlights the need to protect the integrity of integrated C&E markets while advancing low-carbon energy goals.

Chapter 5

Micro-segmentation to mitigate False Data Injection in Cyber-Physical Power Systems

5.1 Introduction

As mobile devices, IoT devices, and Cloud computing platforms are increasingly integrated into CPPS, traditional perimeter-based security approaches are becoming less effective in isolating suspicious components that are already inside the network. Moreover, the interaction between the cyber and power layers within a CPPS exposes the latter to a wider attack surface, thereby facilitating the lateral spreading of attackers across the whole network. Indeed, once an adversary breaches the CPPS perimeter, it can easily spread laterally and compromise several components to launch sophisticated attacks. Therefore, it is critical to secure each network component in a finer-grained manner, rather than focusing solely on perimeter protection.

5.1.1 Overview

Among the various types of cyberattacks targeting CPPS, FDIA has been recognized as a highly threatening attack [123]. FDIA aims to mislead the PSSE [32] by manipulating power measurements in a correlated way to bypass the residual-based BDD, which relies on measuring residues to identify bad data. To be successful, an FDIA requires two preconditions: (i) sufficient knowledge of the structural network and parameters for constructing a stealthy attack vector and (ii) a sufficient number of measurements for collaborative manipulation.

Several approaches have been proposed to mitigate FDIA by targeting those preconditions [44]. One of these involves providing measurement redundancy in SE by deploying PMUs [14], which can identify the measurement manipulations the adversary needs to meet precondition (i). Another attractive security approach is the MTD [124, 125], which utilizes distributed flexible AC transmission system (D-FACTS) devices. MTD can proactively alter impedance perturbations and hide the measurement matrix information, thereby hindering precondition (ii). Although these two approaches can effectively increase the accuracy of the BDD and mitigate FDIA, their implementation in power systems is impractical due to the high cost and limited scalability of deploying PMUs and D-FACTS. For example, protecting all the buses in the IEEE-14 system requires almost 61.9% of the transmission lines to be deployed with D-FACTS [46], which is infeasible due to the high costs involved.

Existing techniques against FDIA assume that the adversary has already compromised the required measuring devices. The previous penetration stages of the attack are therefore not considered at all, neglecting further opportunities for detection and prevention. Indeed, before executing a stealthy FDIA, the attacker needs to get access to the sensor network and spread laterally across it to infect other devices. Hampering the penetration stage would prevent the attacker from taking control of enough devices. This can be achieved using a ZTA, where authentication and authorization are required for each access request of every network component [22], regardless of whether it is located within or outside the perimeter. In this way, the lateral movement of an attacker within the network can be prevented, thereby reducing the number of compromised devices and making it more difficult to satisfy precondition (ii) for executing an FDIA. Furthermore, this solution would not incur the costs and scalability issues of deploying PMUs and D-FACTS.

5.1.2 Contribution

In prior research [126], a cyber-physical ZTA designed for CPS was proposed, and the challenges of implementing ZTA in CPS were analyzed. Although this Cyber-Physical ZTA shows its potential against FDIAs, assessment and refinement of its effectiveness are essential due to limited security resources and evolving threats in real-world scenarios. Consequently, this chapter proposes a novel micro-segmentation technique to secure CPPS from FDIAs by enhancing its residual-based BDD detection capability. To assess the effectiveness of the micro-segmentation strategy, a new combined cyber-physical metric is devised, and a combinatorial optimization problem is formulated to optimize it. However, this problem is challenging to solve in polynomial time. Given the critical nature of power systems, a very short response time is required (e.g., load-shedding decisions for operators within milliseconds) to effectively mitigate potential cross-layer cascading risks. Therefore, a novel heuristic

optimization algorithm based on GAT+RL is designed to search for a near-optimal micro-segmentation strategy within the inference time. The main contributions of this chapter are summarized below.

1. A novel micro-segmentation technique based on the concept of ZTA; it restricts lateral attack propagation, reduces the stealthiness of FDIAs, and is proven effective against FDIAs under the DC model.
2. An optimization of the proposed micro-segmentation technique; it leverages a cyber-physical metric and a GAT+RL algorithm to enhance its effectiveness.
3. A GAT-based extraction algorithm; it captures cross-layer features of both cyber and power components.
4. Simulation-based evaluation showing that deploying the proposed micro-segmentation technique significantly improves the detection rate of residual-based BDD against FDIAs, increasing from 5.23% to 94.02%, and highlights the effectiveness of the GAT+RL optimization algorithm, which considerably outperforms state-of-the-art algorithms in computing time while maintaining solution quality.

5.1.3 Structure of Chapter

The chapter is organized as follows: Section 5.2 illustrates the PSSE and the stealthiness of FDIAs. Section 5.3 introduces the implementation of the proposed micro-segmentation technique in CPPS and proves its effectiveness against FDIAs. Section 5.4 formulates the optimal micro-segmentation as a combinatorial optimization problem and proposes a cyber-physical-BDD-enhancement-metric that simultaneously considers the impact of the micro-segmentation strategy on lateral spreading capability and BDD detection probability. Section 5.5 presents a GAT+RL algorithm designed to solve the proposed combinatorial optimization problem. Section 5.6 provides simulation results and analysis, and Section 5.7 concludes the chapter.

5.2 System Model and Attack Model

In the background section 2.2.3, the stealthiness of FDIA and its impact on PSSE was briefly discussed. This section briefly introduces the stealthiness of FDIA and its properties.

5.2.1 System Model

As illustrated in Fig. 5.1, the SCADA system collects power measurements via measuring devices (i.e., sensors) and transmits them to the EMS for PSSE. Residual-based BDD ensures data reliability by detecting abnormal measurements. The refined states are used to determine control signals for other EMS functions, such as optimal economic dispatch.

5.2.1.1 Power system state estimation (PSSE)

The PSSE aims to estimate state variables with the measurements [127]. Assume that N' power system state variables $\mathbf{x} = (x_1, x_2, \cdots, x_{N'})^T$ are evaluated M ($N' \ll M$) based on measurements $\mathbf{z} = (z_1, z_2, \cdots, z_M)^T$. This yields $\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{e}$, where $\mathbf{H} \in \mathbb{R}^{M \times N'}$ is the measurement matrix and $\mathbf{e} \in \mathbb{R}^M$ is the independent noise. When $\mathbf{e}$ is normally distributed with zero means, the estimated state is $\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{W}\mathbf{h})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z} \triangleq \mathbf{K}\mathbf{z}$, where $\mathbf{W}$ is a diagonal matrix.

5.2.1.2 Bad data detector

In PSSE, the residual-based BDD is widely utilized to localize and detect abnormal measurements. Specifically, the residual $\mathbf{r} = (\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{z}$ is used to detect bad data, where $\mathbf{I}$ is the identity matrix. If $\|\mathbf{r}\|_2 = \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\|_2 > \tau_r$, bad data are successfully detected.

5.2.2 Attack Model: Mechanisms and Properties

We consider the stealthy FDIA targeting PSSE proposed in [32], which can bypass the BDD by constructing a completely stealthy attack vector, rather than random bad data.

5.2.2.1 Stealthy FDIAs

The BDD-bypassing attack vector is defined as $\mathbf{a} = \mathbf{H}\mathbf{c}$, where $\mathbf{c} \in \mathbb{R}^{N'}$ is an arbitrary vector. Without considering measurement noise $\mathbf{e}$, this attack vector should satisfy the following condition [32] to bypass the BDD:

$$\begin{aligned} \|\mathbf{r}\| &= \|\mathbf{z}_\mathbf{a} - \mathbf{H}\hat{\mathbf{x}}_{\text{bad}}\| \\ &= \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}} + (\mathbf{H}\mathbf{c} - \mathbf{H}\mathbf{c})\| = \|\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\| \leq \tau_r, \end{aligned} \tag{5.1}$$

where $\mathbf{z}_\mathbf{a} = \mathbf{z} + \mathbf{a}$ is the measurement after FDIA and $\hat{\mathbf{x}}_{\text{bad}}$ is the estimated results calculated from $\mathbf{z}_\mathbf{a}$.

Lemma 5.1. *If the attack vector can satisfy* $\mathbf{a} \in \mathrm{col}(\mathbf{H})$*, then it can bypass the BDD [32].*

5.2.2.2 Attack Properties

Assume that attackers have the following properties:

- *Accessibility*: Attackers can inject false data via compromised sensors, disrupting PSSE accuracy. Following the initial compromise, the attack can spread laterally, disrupting adjacent devices.
- *Attack Target*: When inaccurate state estimates $\hat{\mathbf{x}}_{\text{bad}}$ caused by FDIAs are transmitted to the EMS, they result in suboptimal control signals, leading to operational issues such as voltage instability, frequency fluctuations, blackouts, and financial losses [128].
- *Attack Knowledge*: The knowledge of attackers is analyzed in two scenarios. In Scenario I, attackers possess sufficient knowledge, including the power network topology, parameters, and the indicator matrix $\mathbf{B}_{\text{i}}$ for the targeted i-th security segment. In Scenario II, attackers do not have access to $\mathbf{B}_{\text{i}}$.

5.3 ZTA in CPPS: Leveraging Micro-Segmentation

The ZTA discussed in this chapter is a security architecture rather than a standalone detector. When integrated into the SCADA system, it enhances the detection capability of the existing residual-based BDD against stealthy FDIAs.

5.3.1 Implementation of ZTA within CPPS Architecture

Traditional perimeter-based security architectures only analyse the traffic flowing between the external (untrusty) and the internal networks (trusty) [129], therefore cannot prevent lateral movement within their perimeters. Once attackers breach the perimeter, e.g., by exploiting a vulnerability, they can access any trusted components. To address this issue, ZTA has been proposed, shifting the focus from perimeter-based to internal, resource-centric defense strategies.

As shown in a NIST report [22], ZTA addresses the aforementioned limitations by prioritizing the protection of resources, assets, and components, instead of the perimeter [82]. When one component (the *subject*) requests access to another component (the *resource*), this architecture determines whether the request can be served. PAPs define and manage access control policies used by PDPs to evaluate and

decide on access requests. PEPs enforce these policies, acting as gateways between components, either at security boundaries or before specific components. To support PAPs, supplementary modules provide contextual data, including risk assessments, traffic logs, and threat intelligence.

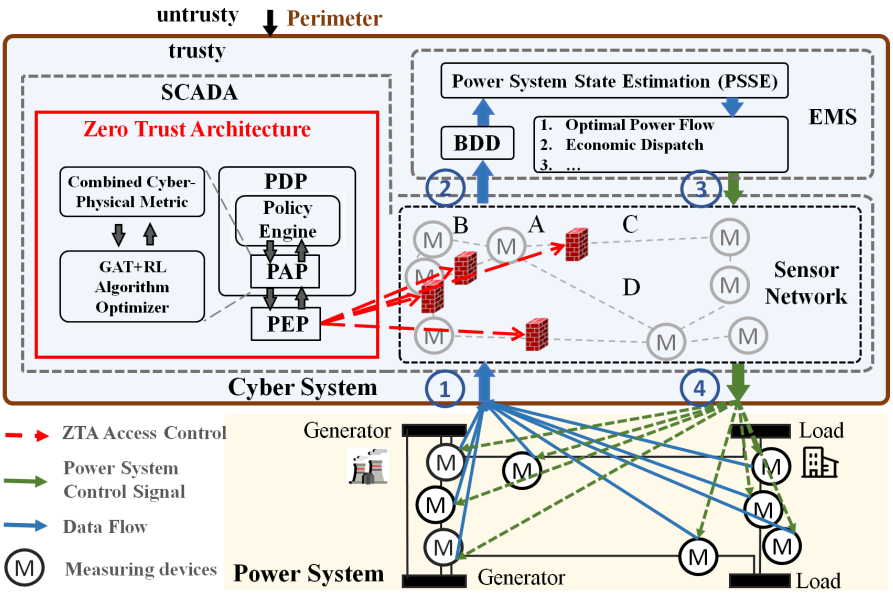


FIGURE 5.1: Architectural framework for implementing ZTA in CPPS.

In CPPS, SCADA systems are pivotal in connecting sensors, controllers, and actuators. Traditional security architectures cannot mitigate lateral spreading within SCADA systems, posing risks like the spread of corrupted sensor data to control signals. With reference to Fig. 5.1, integrating ZTA into SCADA systems can restrict attack spread between internal components, thereby enhancing the overall security of CPPS. The next section presents the proposed micro-segmentation technique to enable a ZTA-based mitigation of FDIAs.

5.3.2 Micro-Segmentation technique against FDIA

Fig. 5.2 shows an example to highligh the differences between a network without ZTA (above in Fig. 5.2) and one with ZTA (below in Fig. 5.2). The sensor network with ZTA is segmented into several security segments, and the access requests between devices in different segments are monitored by the policy engine. In scenarios where an attacker has breached breach device A, the micro-segmentation technique restricts their ability to extend this breach to devices C, D, E, and F via device A. This hinders FDIA precondition (ii) for cooperative measurement manipulation, thereby enhancing the BDD detection capability for identifying attack behavior.

We define the sensor network as a weighted graph $G = G(E, V)$ with node set V, $|V| = M$, edge set E and communication adjacent matrix $\mathbf{A}_{M*M} = \left\{a_{ij}^0\right\}_{i,j=1}^{M}$. A

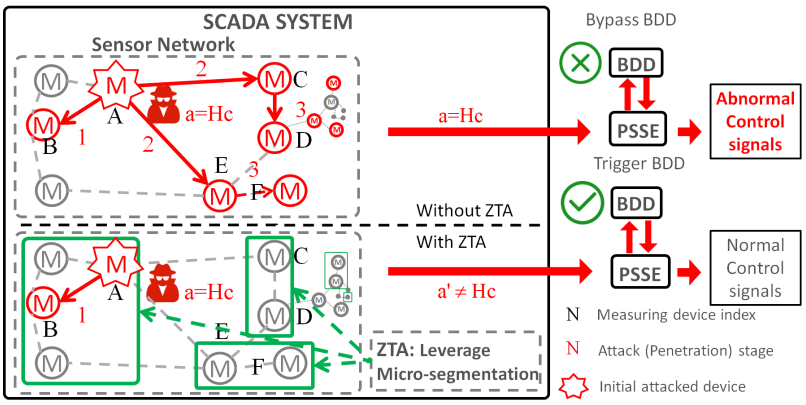


FIGURE 5.2: Illustrative example of micro-segmentation mechanism.

micro-segmentation strategy segments the network into K segments, represented by an indicator matrix $\mathbf{B}_{M*K} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_K]^T$. Each indicator vector $\mathbf{h}_k = [h_{1,k}, h_{2,k}, \cdots, h_{M,k}]^T$ specifies the assignment of nodes to segments, where $h_{ij} = 1$ if the i-th node is in the j-th segment; otherwise, $h_{ij} = 0$. Suppose the micro-segmentation strategy with the indicator matrix $\mathbf{B}_{M,K} = [\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_K]^T$ is implemented in the sensor network, where M measuring devices are logically segmented into $\mathcal{K} = \{\mathcal{N}_1, \cdots, \mathcal{N}_K\}$ security segments. During each attack period, the attack vector $\mathbf{a} = \mathbf{Hc}$ will be reconstructed as $\mathbf{a}' = \mathbf{B}_i\mathbf{Hc} = \mathbf{H}'_i\mathbf{c}$, where $\mathbf{B}_i = diag(\mathbf{B}_i)$ and $i = 1, 2, \cdots, K$. Note that $\mathbf{H} = \mathbf{H}'_1 + \mathbf{H}'_2 + \cdots + \mathbf{H}'_k$. Given a matrix $\mathbf{H}'_i$, let $\mathbf{H}'^{-0}_i$ be the matrix obtained by deleting all-zero rows from $\mathbf{H}'_i$. It follows that $rank(\mathbf{H}'^{-0}_i) = rank(\mathbf{H}'_i)$.

The effectiveness of the micro-segmentation technique under the DC model is demonstrated in two scenarios that differ in the attacker's knowledge.

5.3.2.1 Scenario I Sufficient knowledge

Despite possessing sufficient knowledge, including power network topology, parameters, and the indicator matrix $\mathbf{B}_i$ for the i-th security segment, attackers can design a stealthy attack vector only if $\mathbf{B}_i$ satisfies a specific condition, which is analyzed subsequently. If this condition is unsatisfied, attackers cannot design the stealthy $\mathbf{a}$ to bypass the BDD, resulting in an increased detection rate.

Lemma 5.2. *If and only if the security segment matrix* $\mathbf{B}$ *satisfies the* $rank(\mathbf{H} - \mathbf{H}'_i) < N'$ *for any* i, *attackers can construct a completely stealthy attack vector to bypass the BDD, i.e., there exist solutions for* $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_i\mathbf{c}_i$ *and* $\mathbf{a}' \neq \mathbf{0}$.

Proof. (Sufficiency) The condition that there exists solution for $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_i\mathbf{c}_i \neq \mathbf{0}$ can be transformed into the equal condition that there exists $\mathbf{c}'' = [\mathbf{c}\ \mathbf{c}_i]^T$ satisfying

$[\ \mathbf{H}\ \mathbf{H}'_{\mathrm{i}}\ ][\ \begin{matrix}\mathbf{c}\\ \mathbf{c}_{\mathrm{i}}\end{matrix}\ ] = \mathbf{0}$. Define $\mathbf{H} - \mathbf{H}'_{\mathrm{i}}$ as $\mathbf{H}'_{-\mathbf{i}}$. Equally, the original condition is further reconstructed as $[\ \begin{matrix}\mathbf{H}'^{-0}_{\mathrm{i}} & \mathbf{H}'^{-0}_{\mathrm{i}}\\ \mathbf{H}'^{-0}_{-\mathbf{i}} & 0\end{matrix}\ ][\ \begin{matrix}\mathbf{c}\\ \mathbf{c}_{\mathrm{i}}\end{matrix}\ ] = 0$, $\mathbf{a}' \neq \mathbf{0}$ and $\mathbf{c} \neq \mathbf{0}$.

Here, if the $rank(\mathbf{H} - \mathbf{H}'_{\mathrm{i}}) = rank(\mathbf{H}'^{-0}_{-\mathbf{i}}) < N'$, there exists non-zero solution for $\mathbf{H}'^{-0}_{-\mathrm{i}}\mathbf{c} = \mathbf{0}$. Thus, given any exact non-zero solution $\mathbf{c}' \neq \mathbf{0}$ and $\mathbf{c}'_{\mathrm{i}} = -\mathbf{c}' \neq \mathbf{0}$, it yields that $\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c}' + \mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c}'_{\mathrm{i}} = \mathbf{0}$. Next, we need to prove that $\mathbf{a}' \neq \mathbf{0}$. Referring to $\mathbf{H}'^{-0}_{-\mathrm{i}}\mathbf{c} = 0$, $\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c}'_{\mathrm{i}} \neq \mathbf{0}$ should be proved consequently. Here, we resort to a contradiction. Assuming that there exists an non-zero element $\mathbf{c}$ satisfying $\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c} = -\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c}_{\mathrm{i}} = \mathbf{0}$. It yields there exists non-zero solution for $\mathbf{Hc} = \mathbf{0}$. However, since $rank(\mathbf{H}) = N$, there exists no non-zero solutions for $\mathbf{Hc} = \mathbf{0}$. There is a contradiction. Hence, $\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c} \neq \mathbf{0}$, i.e. , $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_{\mathrm{i}}\mathbf{c}_{\mathrm{i}} \neq \mathbf{0}$. In this case, after launching an attack in security segment i, attackers can construct a completely stealthy attacker vector to bypass the BDD.

(Necessity) Suppose there exists solution for $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_{\mathrm{i}}\mathbf{c}_{\mathrm{i}}$ and $\mathbf{a}' \neq \mathbf{0}$. That is, there exist non-zero solutions $\mathbf{c} = [\mathbf{c}\ \mathbf{c}_{\mathrm{i}}]^T$ satisfying $\left[\begin{matrix}\mathbf{H}'^{-0}_{\mathrm{i}} & \mathbf{H}'^{-0}_{\mathrm{i}}\\ \mathbf{H}'^{-0}_{-\mathrm{i}} & 0\end{matrix}\right]\left[\begin{matrix}\mathbf{c}\\ \mathbf{c}_{\mathrm{i}}\end{matrix}\right] = 0$. Given one exact solution $\mathbf{c}' = [\mathbf{c}'\ \mathbf{c}'_{\mathrm{i}}]^T$, it yields that $\mathbf{H}'^{-0}_{\mathrm{i}}\mathbf{c}'_{\mathrm{i}} \neq \mathbf{0}$ and $\mathbf{c}' = -\mathbf{c}'_{\mathrm{i}} \neq \mathbf{0}$. Hence, the non-zero solution $\mathbf{c}'$ can satisfy $\mathbf{H}'^{-0}_{-\mathrm{i}}\mathbf{c}' = \mathbf{0}$. Thus, it yields that $rank(\mathbf{B}^{-0}_{-\mathrm{i}}) < N$‘. □

Remark: Under the deployment of micro-segmentation, the penetration capability of attackers is restricted. The attack vector $\mathbf{a}'$ injected into measurements differs from the originally attack-constructed vectors $\mathbf{a}$. The corresponding expected residual $\mathrm{E}(\|\mathbf{r}\|)$ induced by $\mathbf{a}'$ cannot satisfy the Lemma 5.1. Consequently, $\mathbf{a}'$ cannot eliminate the increase of $\mathrm{E}(\|\mathbf{r}\|)$, leading to a higher detection rate of the BDD.

5.3.2.2 Scenario II Insufficient knowledge

When attackers lack knowledge of the deployed micro-segmentation strategy $\mathbf{B}_{\mathrm{i}}$, they cannot design and execute a completely stealthy attack vector $\mathbf{a}$. A less effective attack vector $\mathbf{a}'$ is injected, resulting in an increased detection rate.

Proposition 5.3. *After implementing the micro-segmentation technique, the practical attack vector $\mathbf{a}'$ cannot bypass the BDD in a completely stealthy fashion. In other words, there exists no solution for $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_{\mathrm{i}}\mathbf{c}_{\mathbf{i}}$ where $\mathbf{c} \in \mathbb{R}^{N'}$, $\mathbf{c}_{\mathbf{i}} \in \mathbb{R}^{N'}$ and $\mathbf{a}' \neq \mathbf{0}$. Without a stealthy attack vector, the attack actions are easily detected by the BDD, resulting in an expected increase in the detection rate.*

Proof. Assuming that the network is divided into two segments with matrix $\mathbf{H}'_1$ and $\mathbf{H}'_2$, there exists $\mathbf{c}_1$ and $\mathbf{c}_2$ satisfying $\mathbf{a}' = \mathbf{H}'_1\mathbf{c}_1 = \mathbf{H}'_2\mathbf{c}_2$, $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^{N'}$. It equals that there

exists $\mathbf{c}' = [\mathbf{c}_1\mathbf{c}_2]^T$ satisfying $\begin{bmatrix} \mathbf{H}'_1 \ \mathbf{H}'_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = 0$, $\mathbf{a}' \neq 0$ and $\mathbf{c}' \neq \mathbf{0}$. Referring to $\mathbf{H}'_i = \mathbf{B}_i\mathbf{H}$ and $\mathbf{H} = \mathbf{H}'_1 + \mathbf{H}'_2$, the assumption is transferred into $\begin{bmatrix} \mathbf{H}'^{-0}_1 & 0 \\ 0 & \mathbf{H}'^{-0}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = 0$. Note that exchanging the rows of $\mathbf{H}'_i$ has no impact on the solutions of this equation, and thus it yields that $\mathbf{H}'^{-0}_2\mathbf{c}_2 = 0$ and $\mathbf{H}'^{-0}_1\mathbf{c}_1 = 0$. Thus, $\mathbf{H}'_2\mathbf{c}_2 = 0$ and $\mathbf{H}'_1\mathbf{c}_1 = 0$. There exists a contradiction with the assumption. Hence, the proposition has been proved. That is, there exists no solution for $\mathbf{c_i}$ can satisfy $\mathbf{H}'_i\mathbf{c_i} \in \text{col}(\mathbf{H})$ besides $\mathbf{c_i} = \mathbf{0}$. Thus, attackers cannot construct such attack vector $\mathbf{a}' = \mathbf{Hc} = \mathbf{H}'_i\mathbf{c}_i$ which can bypass the BDD under the micro-segmentation. In this case, under the micro-segmentation technique, FDIA vectors cannot bypass the BDD in a stealthy fashion, and the practical attack vector can be detected by chance. □

5.4 Optimization of Micro-Segmentation Strategies with Cyber-Physical-BDD-Enhancement Metrics

The previous discussion shows that micro-segmentation $\mathbf{B}$ results in a practical attack vector $\mathbf{a}'$, diverging from the attacker-constructed vector $\mathbf{a}$. This deviation enhances the detection capability of the BDD, thereby reducing lateral spread in the sensor network. Specifically, the micro-segmentation $\mathbf{B}$ enhances the detection probability of BDD by raising the expected residue $\mathbb{E}(\|\mathbf{r}\|)$, as stated in Eqn. (5.1).

Consequently, defenders aim to maximize $\mathbb{E}(\|\mathbf{r}\|)$ by deploying $\mathbf{B}$, treating this problem as a combinatorial optimization problem. Given the sequential nature of attack propagation in the sensor network and its cross-layer impacts on the power system, a combined cyber-physical metric is proposed as the optimization objective, which conjointly integrates impacts on both cyber and power systems. This metric is used to assess the effectiveness of $\mathbf{B}$, compared with two standard metrics (i.e., physical metric and cyber metric), as illustrated in Section 5.4.2 and 5.4.3. This section elaborates on the problem definition and formulation of the combined cyber-physical metric, followed by an in-depth analysis of the physical and cyber metrics, respectively.

5.4.1 Objective Definition: Cyber-Physical-BDD-Enhancement Metric

A residual-based metric is discussed in Section 5.4.2 to evaluate the increase in $\mathbb{E}(\|\mathbf{r}\|)$ after deploying strategy $\mathbf{B}$. However, this metric has limitations as it assumes that attackers can only manipulate measurements on devices they have already compromised, leaving those from uncompromised devices unaltered [130]. In practice, the process of attacker penetration is sequential and dynamic, rather than

fixed, making such assumptions impractical. Therefore, the $\mathbb{E}(\|\mathbf{r}\|)$ is not only affected by the physical measuring matrix $\mathbf{H}$, but also by the lateral spreading capability of attackers prior to the cooperative measurement manipulation, as depicted in Fig. 5.1. This capability can be quantified using the infection probability of devices, which depends on their locations, neighboring devices, and the entire network topology, as discussed in Section 5.4.3. To solve this issue, a combined cyber-physical metric, denoted as L_{cp}, is formulated which integrates the infection probability into the standard residue-based metric. This is formulated as

$$
\begin{aligned}
L_{cp} &= \max_{\mathbf{B}} \ \mathbb{E}\left(\left\|(\mathbf{I}-\mathbf{HK})\left(\mathbf{z}+\frac{\mathbf{a}''}{\|\mathbf{a}''\|}\right)\right\|\right) \qquad (5.2)\\
&= \mathbb{E}\left(\left\|(\mathbf{I}-\mathbf{HK})\left(\mathbf{z}+\frac{\mathrm{diag}(\mathcal{P}')\mathbf{a}}{\|\mathrm{diag}(\mathcal{P}')\mathbf{a}\|}\right)\right\|\right)\\
&= \mathbb{E}\left(\left\|\mathbf{Q}(\mathbf{z}+\frac{\mathrm{diag}\left(\rho^{*}-\mathbf{T}(\mathbf{G}^{0}-\mathbf{G}^{0}\circ(\mathbf{B}\mathbf{B}^{\top}-\mathrm{E}))\rho^{*}\right)\mathbf{a}}{\left\|\mathrm{diag}\left(\rho^{*}-\mathbf{T}(\mathbf{G}^{0}-\mathbf{G}^{0}\circ(\mathbf{B}\mathbf{B}^{\top}-\mathrm{E}))\rho^{*}\right)\mathbf{a}\right\|}\right\|\right)\\
\text{s.t.}\quad & \sum_{j=1}^{K} h_{ij}=1, \forall\, i\in V \qquad (5.2a)\\
& \sum_{i=1}^{M} h_{ij}=1, \forall\, j\in \mathcal{K}. \qquad (5.2b)
\end{aligned}
$$

This metric can consider the impact of $\mathbf{B}$ on the lateral spreading capability and the BDD detection probability simultaneously. It can be used to guide defenders to obtain an effective $\mathbf{B}$ against FDIA. Specifically, it models the trade-off between mitigating the lateral spreading capability at the cyber layer and reducing the magnitude of injected false data at the physical layer. The explanations for each symbol can be found in the following Sections 5.4.2 and 5.4.3. In addition, the solution to this optimization problem is discussed in Section 5.5.

5.4.2 Increasing the detection capability of BDD: Physical Metric

According to Proposition 5.3, it proves that after deploying micro-segmentation, attackers cannot execute the attack vector in a stealthy fashion as they intended. However, it is worth mentioning that the BDD has a tolerance level for the predetermined threshold $\tau > 0$. In certain scenarios, a practical attack vector can still bypass the BDD by chance if it satisfies $\|\mathbf{r}\| \leq \tau$ in Eqn. (5.1). This situation poses a challenge in terms of localizing the security group where the malware is launched and identifying the precise attack vector, which leads to a malicious threat to the CPPS.

To solve this problem, micro-segmentation designed in this section aims to maximize the expected residue $\mathbf{E}(\|\mathbf{r}\|)$. In this way, a minor attack vector can result in a large

residue after deploying the strategy, which improves the detection capability of the BDD. The relationship between residue $\mathbf{r}$ and attack vector $\mathbf{a}$ referring to Eqn. (5.1) is given as:

$$\mathbf{r} = (\mathbf{I} - \mathbf{HK})(\mathbf{z} + \mathbf{a}). \tag{5.3}$$

By segmenting the measuring devices with strongly correlated measurements into different segments, these measurements can be detected with a high detection probability. In other words, when measurement $\mathbf{z}_i$ in the first group is injected with the attack vector $\mathbf{a}_i$, the BDD can detect this attack with a high probability, relying on the correct measurements (correlative with $\mathbf{z}_i$) in other groups referring to ***lemma 2*** in [46]. Here, the residue after micro-segmentation for security segment i can be defined as:

$$\mathbf{r}_i = (\mathbf{I} - \mathbf{HK}) \left(\mathbf{z} + \frac{\text{diag}(\mathbf{B}_{:,i})\mathbf{a}}{\|\text{diag}(\mathbf{B}_{:,i})\mathbf{a}\|} \right), \tag{5.4}$$

where $\text{diag}(\mathbf{B}_{:,i})\mathbf{a}$ represents the practical attack vector injected into measurements after the micro-segmentation scheme. It is worth noting that with the increase of attack strength (i.e., the number of attack vectors injected into measurements), the probability of success of detection increases [131]. To effectively compare the effectiveness of various micro-segmentation strategies, it is crucial to maintain consistent attack strengths. This ensures that any differences in security performance are attributed to the micro-segmentation strategies, rather than to variations in attack strength. In this metric, the attack vector $diag(\mathbf{B}_{:,i})\mathbf{a}$ is set with the unit attack strength $\frac{\text{diag}(\mathbf{B}_{:,i})\mathbf{a}}{\|\text{diag}(\mathbf{B}_{:,i})\mathbf{a}\|} = 1$. Here, the physical metric is modeled as:

$$\begin{aligned} \mathrm{L}_p &= \max_{\mathbf{B}} \min_{i \in \mathcal{K}} \ \mathbb{E}(\|\mathbf{r}_i\|) \\ &= \mathbb{E}\left(\left\| (\mathbf{I} - \mathbf{HK}) \left(\mathbf{z} + \frac{\text{diag}\left(\mathbf{B}_{:,i}\right)\mathbf{a}}{\|\,\text{diag}\left(\mathbf{B}_{:,i}\right)\mathbf{a}\|} \right) \right\| \right). \end{aligned} \tag{5.5}$$

Incorporating this physical metric into micro-segmentation increases the system's sensitivity to FDIA vectors. Specifically, even smaller FDIA vectors can cause a significant residue to be detected with $|\mathbf{r}| > \tau$, thereby narrowing the attack space of FDIA vectors. Hence, the level of security architecture in CPPS increases.

5.4.3 Limiting the lateral spreading capability: Cyber Metric

The effectiveness of strategies on the BDD detection capability has been analyzed in Sections 5.3.2 and 5.4.2. However, the attack vector $\mathbf{a}' = \frac{\text{diag}(\mathbf{B}_{:,i})\mathbf{a}}{\|\,\text{diag}(\mathbf{B}_{:,i})\mathbf{a}\|}$ after deploying the strategy is constructed in an ideal scenario, assuming that all measuring devices within the compromised security segment are accessible to attackers, while devices in other segments are inaccessible. In fact, the PEP deployed on the security segment is incapable of isolating the lateral spreading with a 100% probability. Specifically, there

is a negligible probability that malware targeting one security segment might laterally spread to other segments. In contrast, within a compromised security group, certain measuring devices might remain uninfected since the malware propagation follows a Susceptible-Infected-Susceptible (SIS) dynamics model [132].

Consequently, the infection probability is used to represent the infection state of each measuring device, rather than relying on their locations within or outside security segments. In a fully connected cyber network, the sum of infection probabilities for all devices is a general metric for evaluating the spreading capability of a network, which is mainly determined by its topology. By deploying the micro-segmentation strategy, the links between different security segments are inaccessible to attackers, resulting in an updated topology as expected by the attackers.

The original network topology and updated topology after the strategies are defined as $\mathbf{G}^0 = \left\{g_{ij}^0\right\}_{i,j=1}^{M}$ and $\mathbf{G}'$, respectively. The change of topology is $\Delta\mathbf{G} = \mathbf{G}^0 - \mathbf{G}'$, where $\mathbf{G}' = \mathbf{G}^0 \circ (\mathbf{B}\mathbf{B}^\top - \mathrm{E})$. Given that malware spreading relies on the correlation between neighboring nodes, the SIS model of a correlated complex network is discussed to analyze the malware spreading dynamics, which is given as [132]:

$$\frac{d\rho_i(t)}{dt} = -\rho_i(t) + \beta\left[1 - \rho_i(t)\right]\sum_{j=1}^{M} g_{ij}^0 \rho_j(t), \tag{5.6}$$

where ρ_i is the infection probability of node i and β is the infection rate. In the steady state, at $t \to \infty$, the infection probability $\rho_i^* \equiv \rho_i(\infty)$ with topology $\mathbf{G}^0$ can be approximated as:

$$\rho_i^* = \beta\left(1 - \rho_i^*\right)\sum_{i=1}^{M} g_{ij}^0 \rho_j^*. \tag{5.7}$$

Eqn. (5.7) can be rewritten as:

$$\rho^* = \beta\left(\mathbf{1} - \rho^*\right) \circ \left(\mathbf{G}^0 \rho^*\right), \tag{5.8}$$

where $\rho^* = [\rho_1^*, \rho_2^*, \ldots, \rho_M^*]^T$, $\circ$ is the Hadamard product, and $\mathbf{1} = [1, 1, \ldots, 1]^T$. Deploying the micro-segmentation strategy, the infection probability distribution is updated as $\mathcal{P}' = [\rho_1', \rho_2', \ldots, \rho_M']^T$, which is formulated as:

$$\begin{aligned}
\mathcal{P}' &= \rho^* - \Delta\rho^* = \rho^* - (\mathrm{E} - \beta\mathrm{X})^{-1}\mathrm{Y} \\
&= \rho^* - (\mathrm{E} - \beta\mathrm{X})^{-1}\beta\,\mathrm{diag}\left(1 - \rho^*\right)\Delta\mathbf{G}\rho^* \\
&= \rho^* - \mathbf{T}(\mathbf{G}^0 - \mathbf{G}^0 \circ (\mathbf{B}\mathbf{B}^\top - \mathrm{E}))\rho^*,
\end{aligned} \tag{5.9}$$

where $\Delta\rho^* = 1^{\mathrm{T}}(\mathrm{E} - \beta\mathrm{X})^{-1}\mathrm{Y}$ referring to ***lemma 1*** in [133].

$\mathrm{X} = \mathrm{diag}\left(\mathbf{1} - \rho^*\right)\mathbf{G}^0 - \mathrm{diag}(\mathbf{G}^0\rho^*)$, $\mathrm{Y} = \beta\,\mathrm{diag}\left(1 - \rho^*\right)\Delta\mathbf{G}\rho^*$.

After deploying the micro-segmentation strategy, the lateral spreading capability in the sensor network becomes restricted, leading to a decrease in the infection probability of each node. Note that the decrease in infection probability can be used as a cyber metric when only evaluating the impact of the strategy on lateral spreading capability within the sensor network, which is defined as:

$$\mathrm{L}_c = \mathbf{1}^{\mathrm{T}} \Delta \rho^*. \tag{5.10}$$

The metric is discussed in the comparative simulations in the following sections.

5.5 GAT+RL-based Algorithm for Optimizing Micro-Segmentation Strategies

The process of searching for the optimal micro-segmentation strategy $\mathbf{B}^*$ is modeled as an NP-hard combinatorial optimization problem, as stated in Eqn. (5.2). To simplify the discussion, we refer to this problem as the MSC problem in the following discussion. It can be solved by exact methods, approximate methods or heuristic methods [134]. In real-world scenarios, it is essential to select the algorithms that can balance the quality of solutions and the computation time. In terms of the smart grid security events, an effective optimization algorithm is expected to respond promptly, minimizing the malicious consequences, and preventing a cascade of cross-layer impacts. Among the three types of methods, since heuristic methods can usually get a feasible solution with a faster speed, they are favorable for optimizing micro-segmentation strategies.

In this section, an optimization algorithm that combines a GAT-based [135] encoder-decoder model and the reinforcement learning (RL) algorithm is proposed to produce good timeliness and generalization. Within the algorithm, the encoder-decoder model is utilized to extract features of the MSC problem while the RL algorithm is utilized to maximize the objective in Eqn. (5.2) by searching for the

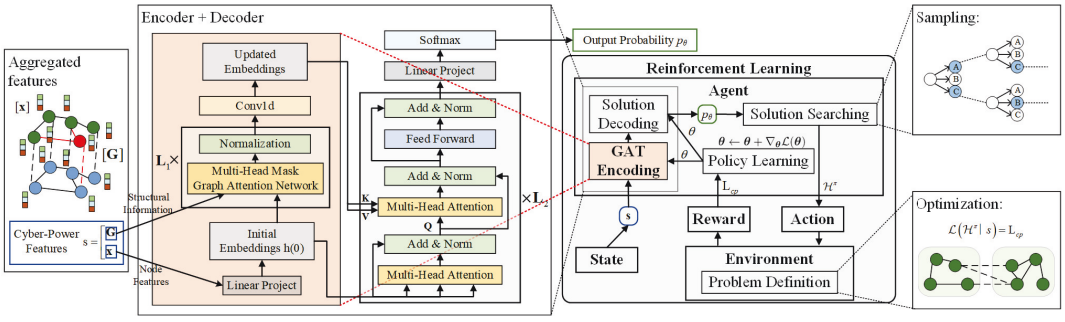


FIGURE 5.3: Structure of GAT+RL optimization algorithm for the MSC problem.

optimal solution $\boldsymbol{\mathcal{B}}^{\pi*} = \{\pi_1^*, \pi_2^*, \ldots, \pi_M^*\}$, which is defined as:

$$\mathcal{L}(\boldsymbol{\mathcal{B}}^{\pi*} \mid s) = \mathrm{L}_{cp}, \tag{5.11}$$

where the state s is defined as all the features of measuring devices, including power measurements, the topology of the power electric grid, cyber features (i.e., infection probability), and the topology of the sensor network. A solution $\boldsymbol{\mathcal{B}}^{\pi} = \{\pi_1, \pi_2, \ldots, \pi_M\}$ is a permutation with repetition of the security group index $\{1, 2, \ldots, K\}$, and each measuring device should be divided into one security group. Note that $\boldsymbol{\mathcal{B}}^{\pi}$ has the same meaning as $\mathbf{B}$, albeit represented with a different symbol. Specifically, π_{i} represents the column index of the element 1 in row i in $\mathbf{B}$.

Based on the formulation of the MSC problem, an encoder-decoder model with parameters $\boldsymbol{\theta}$ is designed, which takes the state s as input, and outputs $\boldsymbol{\mathcal{B}}^{\pi}$. Concretely, given s, this model defines a stochastic policy $p_{\boldsymbol{\theta}}\left(\boldsymbol{\mathcal{B}}^{\pi} \mid s\right)$ for searching for $\boldsymbol{\mathcal{B}}^{\pi}$:

$$p_{\boldsymbol{\theta}}(\boldsymbol{\mathcal{B}}^{\pi} \mid s) = \prod_{\mathrm{i}=1}^{M} p_{\boldsymbol{\theta}}\left(\pi_{\mathrm{i}} \mid s, \boldsymbol{\mathcal{B}}^{\pi}{}_{1\sim\mathrm{i}-1}\right). \tag{5.12}$$

In detail, the encoder takes the original state s as input and outputs the encoder embeddings. Afterward, decoding happens sequentially from i = 1 and stops until all the measuring devices have been divided into the security group, the output of which is a probability distribution. Afterward, the exact solution $\boldsymbol{\mathcal{B}}^{\pi}$ is selected according to this distribution. In this way, the procedure of searching for the optimal solution for the MSC problem is translated to seeking an optimal set of $\boldsymbol{\theta}^*$, with which the model can produce the optimal $\boldsymbol{\mathcal{B}}^{\pi*}$.

5.5.1 Encoder-Decoder Model

For an MSC problem, two types of features are significantly informative, namely sequential information and relation information between neighbors. The encoder maps a sequence of measuring device representations to a sequence of security group representations $\mathbf{v} = (v_1, \ldots, v_K)$, where the above two types of features are included. Given $\mathbf{v}$, the decoder outputs a sequence $\mathbf{p} = (p_1, \ldots, p_M)$, where each element represents the probability of segmenting each of the M measuring devices to the target security group (see Eqn. (5.20)).

The original state $\mathbf{s} = \{\mathbf{G}, \mathbf{x}\}$ consists of both the node features and the structural information. Generally, the electric grid network and the sensor network cannot be modeled in the Euclidean space effectively due to the graph-based topology [136]. A graph-type architecture is required to capture the spatial correlations in the electric

grid network and the sensor network, respectively. Hence, a GAT-based encoder is designed in this subsection to extract structural information.

5.5.1.1 Encoder

As shown in the first column of Fig. 5.3, the encoder of the proposed model is composed of a linear transformation layer, L_1 encoder layers, and a convolutional layer.

For the input of the encoder, the features of each node are mapped to the d_h-dimensional node embeddings:

$$\mathbf{H}^{(0)} = W^{(0)}\mathbf{x} + \mathbf{b}, \tag{5.13}$$

where $\mathbf{x} = [\mathbf{x}^{\mathrm{c}}, \mathbf{x}^{\mathrm{p}}]$ is the input vector. $W^{(0)}$ and $\mathbf{b}$ are learnable parameters of the fully connected layer. $\mathbf{b}^{(0)}$ is the initial node embeddings.

Afterward, $\mathbf{H}^{(0)}$ is processed through L_1 encoder layers and updated as the output embeddings. For each encoder layer, the multi-head graph attention network [135] is introduced to compute d_h-dimensional representations for both sequential information learning and structural information refining. Here, the multi-head attention (MHA) is used to stabilize the learning capability of self-attention [137], where the self-attention with an identical structure is performed N_K times. The attention coefficients can be expressed as:

$$\alpha_{\mathrm{ij}} = \frac{\exp\left(\mathrm{ReLU}\left(W'^{T}\left[Wh_{\mathrm{i}}^{(0)} \| Wh_{\mathrm{j}}^{(0)}\right]\right)\right)}{\sum_{\mathrm{k}\in\mathcal{G}_{\mathrm{i}}}\exp\left(\mathrm{ReLU}\left(W'^{T}\left[Wh_{\mathrm{i}}^{(0)} \| Wh_{\mathrm{k}}^{(0)}\right]\right)\right)}, \tag{5.14}$$

where $[\cdot\|\cdot]$ denotes the concatenation operation. $\mathcal{G}_{\mathrm{i}}$ denotes the set of first-order neighbors of i, which is obtained from the inputted graph features $\mathbf{G}$. The mechanism can acquire structural information by performing masked attention during the learning process. LeakyReLU is used as the activation operation. The computed N_K embeddings are concatenated as the output embeddings $\mathbf{H}_{enc}^{\ell}$ of the ℓ_{th} encoder layer, which can be defined as:

$$h_{\mathrm{i}}^{\ell} = \mathrm{Norm}^{\ell}\left(\sigma\left(\frac{1}{N_K}\sum_{n_k=1}^{N_K}\sum_{\mathrm{j}\in\mathcal{G}_{\mathrm{i}}}\alpha_{\mathrm{ij}}^{n_k}W^{n_k}h_{\mathrm{j}}^{(\ell-1)}\right)\right). \tag{5.15}$$

The final embeddings of the encoder $\mathbf{H}_{enc}^{(out)} \in \mathbb{R}^{H\times M\times d_h}$ is derived from the L_1-th layer embedding $\mathbf{H}^{L_1} \in \mathbb{R}^{H\times M\times d_k}$ via a 1D convolutional neural network. $\mathbf{H}_{enc}^{(out)}$ is the representation for each security group k.

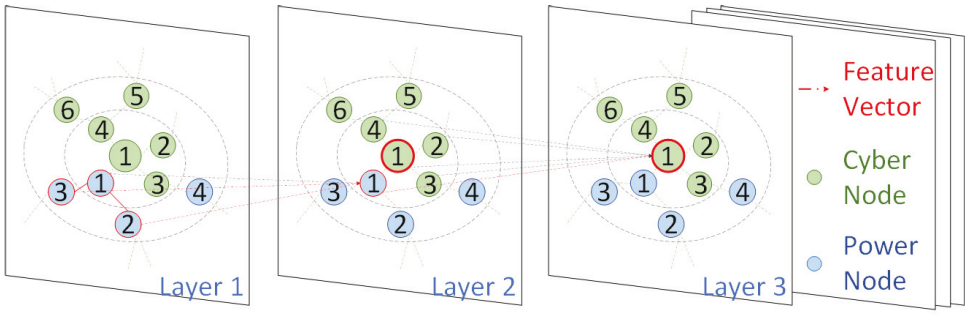


FIGURE 5.4: Cross-layer feature embedding via graph attention network.

To effectively analyze the security of each cyber node in CPPS, it is crucial to capture not only its node features and structural information at the cyber layer but also its cross-layer relationship with power nodes. Graph-structured Neural Networks provide a promising solution by allowing the stacking of multiple layers to learn more complex representations of the graph structure. This approach enables the representation of each cyber node to capture its cross-layer relationship with power nodes, even if they are not immediately apparent in the local neighborhood of this node. As shown in Fig. 5.4, by stacking two layers, the representation v_1^c of cyber node 1 in the third layer can capture its cross-layer relationship with power nodes 1, 2, and 3. Specifically, the representation v_1^p of power node 1 in the second layer captures the features of power nodes 2 and 3 in the first layer. By leveraging v_1^p, v_1^c in the third layer can also capture the features of power nodes 2 and 3 in the first layer. As a result, the representation v_1^c of cyber node 1 captures the cross-layer relationships of power nodes 1, 2, and 3, thereby leading to a more effective security analysis for cyber node 1.

5.5.1.2 Decoder

As shown in the second column of Fig. 5.3, similar to the decoder in the Transformer, the decoder of the proposed model is composed of a stack of decoder layers (L_2 layers) and a fully connected layer.

In each decoder layer, two MHAs are used for both sequential information learning and relation information refining, three "Add & Norm" layers, and an FF layer are also used.

Specifically, the first MHA in the first decoder layer takes the $\mathbf{H}^{(0)}$ as input, while the first MHA in the subsequent L_2 decoder layers takes the output from the former decoder layer as input. The output embeddings can be defined as:

$$\mathbf{H}^{tmp} = \text{Norm}^{\ell}\left(\mathbf{H}'^{(\ell-1)} + \text{MHA}_1^{\ell}\left(\mathbf{H}'^{(\ell-1)}\right)\right), \tag{5.16}$$

$$\mathbf{H}^{tmp1} = \text{Norm}^{\ell}\left(\mathbf{H}^{tmp} + \text{FF}^{\ell}\left(\mathbf{H}^{tmp}\right)\right). \tag{5.17}$$

The output $\mathbf{H}^{tmp1}$ is the representation of each measuring device i. Different from the first MHA in each decoder layer, the second MHA in each decoder layer takes inputs from different modules. Specifically, it takes $\mathbf{H}^{tmp1}$ as the *query* and takes $\mathbf{H}_{enc}^{(out)}$ as the *key* and *value*. In this module, more attention is obtained by a *value* if its *key* is more compatible with the query. That is, each measuring device i is sequentially segmented into one security group k. For each device i, among the features of all security groups (*key*s $\{\mathrm{key_k}\}_{\mathrm{k}=1}^{K}$), the one that matches the features of measuring device i (*Query*s $\{\mathrm{query_i}\}_{\mathrm{i}=1}^{M}$) most is selected as the segmentation group for device i. The output embeddings of this MHA can be defined in the following manner:

$$\mathbf{H}'^{\ell} = \mathrm{MHA}_2^{\ell}\left(\mathbf{Q} = \mathbf{H}_{enc}^{(out)}, \mathbf{K} = \mathbf{H}^{tmp1}, \mathbf{V} = \mathbf{H}^{tmp1}\right). \tag{5.18}$$

After L_2 decoder layers, the output of the fully connected layer in the last of the decoder, i.e., $\mathbf{H}_{dec}^{(out)}$, is defined as:

$$\mathbf{H}_{dec}^{(out)} = W'_{dec}\mathbf{H}'^{(L_2)} + \mathbf{b}', \tag{5.19}$$

where W'_{dec} and $\mathbf{B}'$ are learnable parameters of the layer.

Intuitively, for a given measuring device i, a probability distribution $\{p'_1, p'_2, \cdots, p'_K\}$ is obtained, where $p'_{\mathrm{k}} = \frac{e^{\mathbf{H}_{\mathrm{dec,ik}}^{(out)}}}{\sum_{\mathrm{j}\leq K} e^{\mathbf{H}_{\mathrm{decc,ij}}^{(out)}}}$. This distribution represents the probability of segmenting measuring device i into each of the K security groups. Assuming a greedy strategy is leveraged, the k-th security group with the largest probability p'_{i} is selected as the target group that device i is segmented into, denoted as $\mathcal{B}_{\mathrm{i}}^{\pi} = \mathrm{k}$. The final probability p_{i} of the proposed model is obtained via a softmax function, which is defined as:

$$p_{\mathrm{i}} = p_{\boldsymbol{\theta}}\left(\mathcal{B}_{\mathrm{i}}^{\pi} = \mathrm{k} \mid s, \mathcal{B}_{1\sim \mathrm{i}-1}^{\pi}\right) = \frac{e^{\mathbf{H}_{dec,\mathrm{ik}}^{(out)}}}{\sum_{\mathrm{j}\leq K} e^{\mathbf{H}_{dec,\mathrm{ij}}^{(out)}}}. \tag{5.20}$$

5.5.2 Training with Reinforcement Learning

Given an MSC problem with state s, the proposed encode-decoder model with parameters $\boldsymbol{\theta}$ can provide the probability distribution $p_{\boldsymbol{\theta}}(\mathcal{B}^{\pi} \mid s)$ for each measuring device in each segment according to Eqn. (5.12). Afterward, by sampling from the probability, a set of solutions $\mathcal{B}^{\pi}$ at state s can be obtained. Following Eqn. (5.11), the objectives $\mathcal{L}$ for each solution can be obtained.

Since the objective is to minimize $\mathcal{L}$, the rewards for reinforcement learning can be computed via the Monte Carlo rollout. Then, the agent learns to improve itself toward

the solution with minimal rewards by a state-action-reward tuple. The parameters $\boldsymbol{\theta}$ can be updated using the policy gradient methodology:

$$\begin{aligned} \nabla_\theta \mathcal{L}(\theta) &= \mathbf{E}_{p_{\boldsymbol{\theta}}(\mathcal{B}^\pi|s)} \left[\nabla \log p_\theta(\mathcal{B}^\pi \mid s) L(\mathcal{B}^\pi)\right] \\ \theta &\leftarrow \theta + \nabla_\theta \mathcal{L}(\boldsymbol{\theta}). \end{aligned} \tag{5.21}$$

The rewards used to update the parameters $\boldsymbol{\theta}$ during each training step correspond to the total objectives of a set of solutions. Note that the micro-segmentation strategy for each measuring device in these solutions is sampled from the probability function. This sampling approach effectively mitigates the occurrence of high variance that may arise due to the uncertainty of the sampling process. This high variance can slow down the rate of convergence by giving the model contradictory descent directions to learn. It is necessary to introduce the baseline $b(s)$ to reformulate the policy gradient by Eqn. (5.21):

$$\nabla_\theta \mathcal{L}(\theta) = \mathbf{E}_{p_{\boldsymbol{\theta}}(\mathcal{B}^\pi|s)} \left[\nabla \log p_\theta(\mathcal{B}^\pi \mid s)(L(\mathcal{B}^\pi) - b(s))\right]. \tag{5.22}$$

Generally, $b(s)$ can be computed by many methods, such as the critic network, the greedy rollout of policy, and so on. Different methods are appropriate to different scenarios. The baseline functions $b(s)$ in the proposed problem are computed from the average value of all solutions selected greedily in the first epoch.

Scenario-Specific Reward Functions: The reward function, aimed at optimizing the metric specified in Eqn. (5.2), incorporates penalty terms to enforce the constraints detailed in Section 5.3.2 from Scenarios I and II, which make some solutions infeasible. These penalties guide the agent towards feasible solutions with scenario-specific micro-segmentation strategies.

In Scenario I, constraints in Eqn. (5.2a) and (5.2b) must be satisfied as detailed in Eqn. (5.2). For constraint Eqn. (5.2a), the decoder outputs a probability distribution vector for each device. The solution search module then selects the one with the highest probability. This selection process ensures that each device is allocated into a single security group and, thus, satisfies constraint Eqn. (5.2a). Concerning Eqn. (5.2b), it is equivalent to $rank(\mathbf{B}^\top \mathbf{B}) = K$. To guide the RL exploration towards satisfying this constraint, a penalty function, denoted as $\mathrm{L}_1 = f(rank(\mathbf{B}^\top \mathbf{B}) < K)$, is designed. Herein, f functions as an indicator, taking a value of 1 when the inequality holds true and 0 otherwise. Consequently, the reward function in Scenario I is reformulated to $\mathcal{L}_1\left(\mathcal{B}^{\pi*} \mid s\right) = \mathrm{L}_{cp} - \lambda_1 \mathrm{L}_1$, employing the penalty coefficient λ_1 to modulate the intensity of L_1.

In Scenario II, according to Lemma 5.1, the number of security groups i for which the $rank(\mathbf{H} - \mathbf{B}_i\mathbf{H}) = N$ represents the number of security groups that there are no

stealthy FDIA vectors. Based on this, a penalty function denoted as $L_2 = \sum_i^k [f(rank\,(\mathbf{H} - \mathbf{B}_i\mathbf{H}) < N)]$ is integrated into rewards to guide the solutions towards satisfying this constraint. Simultaneously, constraints Eqn. (5.2a) and (5.2b) in Scenario I should also be satisfied. Consequently, the reward function in Scenario I is redesigned as $\mathcal{L}_2\left(\mathcal{B}^{\pi*} \mid s\right) = L_{cp} - \lambda_{21}L_1 - \lambda_{22}L_2$, where λ_{21} and λ_{22} are penalty coefficients.

5.6 Simulation Analysis and Results

In this section, the effectiveness of the designed micro-segmentation technique is tested on a CPPS network as shown in Fig. 5.5. This network consists of an IEEE 30-Bus system, with parameters cited in MATPOWER[138], and a 39-node sensor network. In addition, the performance of the GAT+RL algorithm is demonstrated and compared with other optimization algorithms.

5.6.1 FDIA Vector Generation

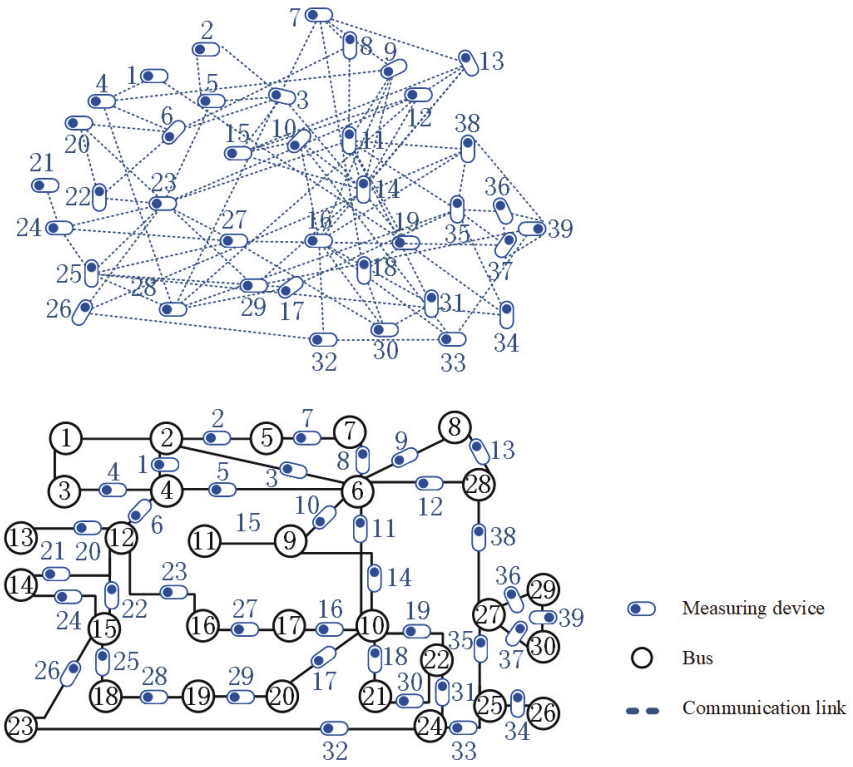


FIGURE 5.5: Measuring devices deployment of the IEEE-30 bus system.

To demonstrate the effectiveness of the proposed micro-segmentation technique, the FDIA attacking pool [139] is generated with 10000 attack vectors. Specifically, the bias **c** utilized to generate **a** is randomly sampled referring to a uniform distribution

U(-0.1, 0.1). Assume that the variance of the measurement noise is σ^2 = 0.01, and the predetermined threshold is $\sigma\sqrt{\chi^2_{m-n,\alpha}}$. Note that with the increase of attack strength (i.e., the number of attack vectors injected into measurements), the probability of success of detection increases [131]. Thus, this section does not compare the effectiveness of micro-segmentation strategies with different attack strengths. In this case, a coefficient ρ is set to limit the $\|\mathbf{a}' = \rho \cdot \mathbf{a}\|$ to a unit of value. In addition, the actual load profile of NYISO [122] for January 2023 with 15-minute time intervals is used to generate a realistic power-state dataset using MATPOWER.

The detection probability is defined as the ratio of the number of unsuccessful attack vectors to 10000. This rate serves as an indicator of BDD detection capability, thereby evaluating the effectiveness of strategies in this section.

5.6.2 Comparisons of the Proposed Cyber-Physical-BDD-Enhancement Metric with other Metrics in Two Scenarios

Under the FDI attacks discussed above, the performance of the proposed cyber-physical metric L_{cp} (CP-Metric) is demonstrated by comparing its results with standard residue-based physical metric L_p (P-Metric) and cyber metric L_c (C-Metric) in Scenario I and Scenario II, respectively. It is unrealistic to obtain the exact optimal solution for each metric due to the complexity of MSC the problem. Hence, we select the near-optimal solutions found by the four compared heuristic algorithms, namely SA, SD, GA, and PSO as the optimal solutions discussed in this section. To reduce the impact of uncertainties, each algorithm has been independently run 30 times and during each iteration, 10000 is set as the maximum number of generations. In addition, the impact of several setting parameters of micro-segmentation strategies on the detection probability is discussed. These parameters specifically include the number of security groups and the missing alarm rate.

5.6.2.1 Number of the Security Groups

In this subsection, the near-optimal segmentation scheme is tested with the increasing number of security groups, K in both scenarios I and II. As shown in Fig. 5.6, the detection rate without micro-segmentation is approximately 5.23%, which increases to 82.97%, 82.06%, and 42.18% for the CP-Metric, C-Metric, and P-Metric in Scenario I, and from 10.28% to 94.02%, 93.03%, and 55.18% in Scenario II, respectively. Notably, the CP-Metric outperforms the C-Metric and P-Metric in both scenarios. The detection rates of the C-Metric and CP-Metric initially increase with K and then fluctuate, with the CP-Metric reaching over 82.4% and 94.02% at $K = 6$ in Scenario I and II, respectively. In contrast, the P-Metric continues to increase slightly.

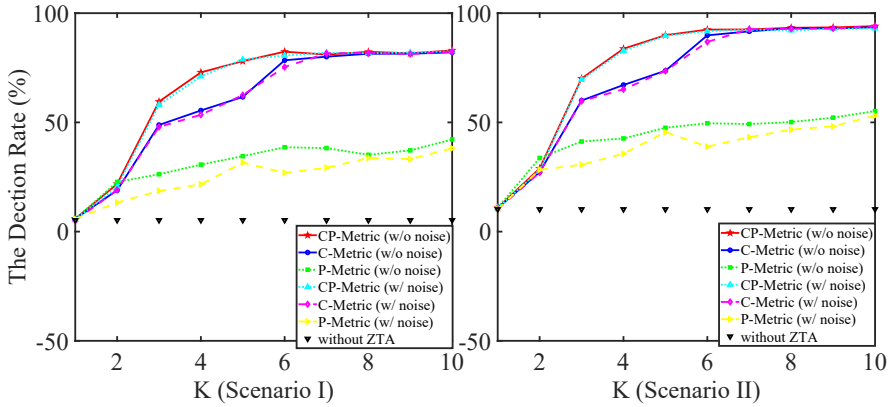


FIGURE 5.6: The detection probability with the number of security groups.

This can be attributed to the increasing degree of network isolation as K grows. Initially, when K is small, the network is weakly segmented, and the degree of network isolation is relatively low. In this case, each security segment contains more nodes, allowing greater potential for attack propagation within each group. Under such conditions, metrics that incorporate network structural information, such as the CP-Metric and C-Metric, demonstrate stronger detection performance due to their sensitivity to internal connectivity and spread dynamics. As K gradually increases, the intensity of segmentation grows, resulting in higher levels of network isolation. However, this finer segmentation might separate critical power-correlated nodes into different security groups, thereby reducing the capability of CP-Metric and C-Metric to effectively assess segmentation strategies. In contrast, the P-Metric, which focuses solely on power correlations while neglecting the global network structure, shows limited effectiveness in evaluating the segmentation strategy.

Overall, the increasing number of security groups can increase the effectiveness of micro-segmentation against FDIAs in both scenarios. Among the three metrics, the CP-Metric consistently achieves superior results by balancing network isolation and critical node correlations, making it a robust metric for segmentation assessment. In contrast, the P-Metric shows limited sensitivity to network structure, resulting in relatively weaker performance improvements.

5.6.2.2 Missing Alarm Rate of the PEPs

In practice, the missing alarm rate (MAR) of a PEP is defined as the ratio of actual malicious access requests that are undetected by the PEP to the total number of malicious requests monitored. In this subsection, the near-optimal micro-segmentation strategy is tested using three metrics with 1e-4 and 1e-5 MAR in two scenarios, respectively. In this subsection, the near-optimal micro-segmentation strategy is tested using three metrics with $1e-4$ and $1e-5$ MAR in two scenarios,

respectively. As shown in Fig. , the detection rate without micro-segmentation is close to 0% in both scenarios. With the increase in the number of security groups K, the detection performance improves significantly across all metrics and MAR values. Specifically, in Scenario I, when K=6, the detection rates for the CP-Metric are 80.81%, 79.6%, and 75.1% under MAR = 0, $1e-5$, and $1e-4$, respectively. The C-Metric achieves 78.7%, 76.10%, and 76.3%, while the P-Metric yields 37.39%, 37.17%, and 36.97% under the same conditions. In Scenario II, the detection rates at K=6 reach 91.68%, 90.24%, and 88.89% for CP-Metric, 90.82%, 90.97%, and 89.47% for C-Metric, and 52.57%, 51.88%, and 52.03% for P-Metric, respectively. The detection rate decreases as MAR increases, particularly for CP-Metric and C-Metric. In addition, the CP-Metric consistently outperforms the other metrics across all MAR levels..

This is primarily due to the inherent enforcement limitations of the PEP, which are often constrained by hardware factors such as gateway configurations, resulting in a non-negligible probability of attack propagation between security segments. As a result, metrics that incorporate the degree of network isolation, such as CP-Metric and C-Metric, are more susceptible to the impact of MAR. In contrast, the P-Metric, which focuses on power correlations, remains relatively unaffected.

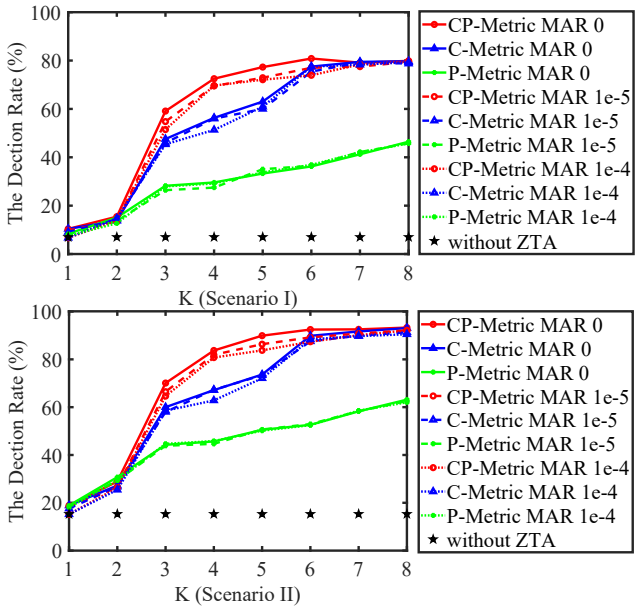


FIGURE 5.7: The detection probability with MAR 0, $1e-5$ and $1e-4$.

5.6.3 Validation of the GAT-RL Algorithm

The presented algorithm is investigated on the IEEE-30 system to demonstrate the performance in terms of time complexity and learning ability.

5.6.3.1 Experiment Setup

Implementations: All simulations are conducted using Python on a PC with an Intel Core i7-10750H CPU (2.60 GHz) and 32 GB of RAM.

Hyperparameters Configurations: The proposed model employed mini-batches of 64, as outlined in TABLE 5.1. The architecture consists of a GAT encoder, configured with 3 layers and 8 heads, and a similar setup for the solution decoder. Both components are optimized using SGD. The model, with 128 hidden dimensions, was trained for 50 epochs, with each epoch comprising 10,240 examples.

TABLE 5.1: Hyperparameters configurations.

HyperParameters	Value	HyperParameters	Value
Batch size	64	No.of epoch	100
Learning rate	1e-4	No. of instances per epoch	10240
GAT-based Encoder			
No. of GAT layers	3	No. of heads	8
Optimizer	SGD	Hidden dimension	128
Multi-Head Attention Decoder			
No. of Decoder layers	3	No. of heads	8
Optimizer	SGD	Hidden dimension	128

5.6.3.2 Comparisons against Other Algorithms on Execution Time and Gap

The section models the micro-segmentation strategy optimization problem as a combinatorial optimization problem. In this optimization problem, the contributions of decision variables and their impact on the objective are not additive, making it challenging to apply existing specialized heuristic algorithms. Hence, some general non-learned heuristic algorithms are included as competitors, especially for the MHA+RL algorithm referring to [134].

In addition, to demonstrate the quality of the solved solutions, the presented approach is also compared with the state-of-the-art exact optimization solvers. TABLE 5.2 depicts the performance of the GAT+RL and the compared algorithms in terms of both running time and optimality gap.

TABLE 5.2: GAT+RL vs state-of-the-art heuristic algorithms.

Method		Obj.↑		Gap ↓		Time ↓	
Exact Algorithm	Cplex	-		0.00%		>1h	
	OR-tools (CP-SAT)	-		0.00%		>1h	
	Gurobi	-		0.00%		>1h	
Scenario		I	II	I	II	I	II
Heuristic Algorithm	GAT+RL	0.2543	**0.1932**	9.38%	**9.72**%	1.91s	2.23s
	MHA+RL	0.2540	0.1903	9.42%	11.07%	**1.56s**	**1.84s**
	SA	0.2532	0.1729	9.70%	19.20%	16.18min	20.44min
	GA	**0.2545**	0.1882	**9.24%**	12.06%	12.58min	18.42min
	PSO	0.2542	0.1835	9.34%	14.25%	15.50min	19.60min
	SD	0.2531	0.1792	9.74%	16.26%	12.91min	15.88min

Performances over the running time: Concerning the running time, our proposed algorithm significantly outperforms other compared exact algorithms and non-learned heuristic algorithms in terms of running time. The exact algorithms are inappropriate for the MCS problem as they take over one hour to search for the exact solution in the test system. In practice, the security strategy must be provided as fast as possible to prevent malicious cascading damage in the power system during attack emergencies. In addition, without compromising the quality of the results, our proposed algorithm significantly outperforms non-learned heuristic algorithms, reducing running time from minutes to seconds.

Performances over the optimality gap: It is obvious that the GAT+RL algorithm can reach significantly closer optimality (except for GA) than other classical non-learned heuristic algorithms within 10000 iterations. Although there is still an optimality gap with the exact algorithm (no more than 10%), the GAT+RL algorithm shows significant advantages in running time as aforementioned.

It is worth mentioning that the GAT+RL algorithm is trained on a dataset consisting of different trained MSC problem instances without supervision. The results are derived by generalizing the model with trained parameters to an MSC problem with different states $\mathbf{s}$. Hence, it is reasonable that there is a slight optimality gap between the GAT+RL algorithm and compared heuristic algorithms. Another concern is that our objective is not to design a specialized, properly designed heuristic algorithm to outperform others in terms of both the running time and the quality of solutions. Instead, our objective is to investigate an appropriate algorithm across all the solvers (exact algorithm, classical non-learned, heuristic algorithm, and so on) to resolve the proposed problem under urgent circumstances. Obviously, the reinforcement learning

algorithm demonstrates its superior performance in terms of the inference time in the MSC problem.

5.6.3.3 Comparisons Against Deep Learning Baselines on Training

The point network (PN) model [140] is the first deep learning model capable of solving combinatorial optimization issues effectively, which is a common baseline method. The learning ability of our proposed GAT+RL algorithm in the training process is evaluated compared with the state-of-the-art algorithm in [134] across Scenario I and Scenario II, where MHA replaces RNN. The performance comparisons of the proposed method and compared baseline method are shown in Fig. 5.8. It is observed that our proposed method definitely outperforms the state-of-the-art baseline in terms of convergence speed. Although MHA can effectively capture sequential information, the GAT mechanism used in the proposed method is more effective at capturing the structural information of measuring devices in the MSC. For the large-scale power grid in the real world, the proposed method shows perspectives on running time.

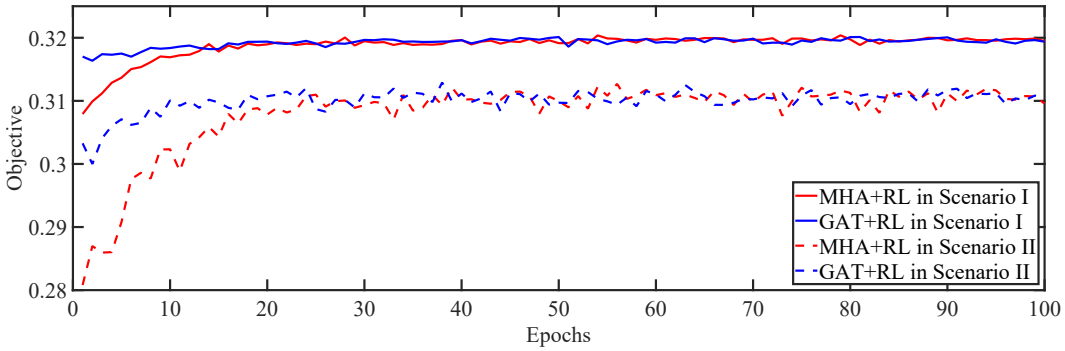


FIGURE 5.8: Validation curves of GAT-RL.

5.7 Summary

In this chapter, a novel micro-segmentation technique against the FDIA is designed, which can be deployed in SCADA and enhance the detection capability of the traditional BDD. This technique segments measuring devices into multiple security groups, effectively restricting lateral attack propagation and significantly improving the accuracy of residual-based BDD mechanisms. Theoretical analysis demonstrates that attackers cannot construct completely stealthy attack vectors to bypass BDD, even with sufficient system knowledge, due to the stringent conditions imposed by the deployment of micro-segmentation configuration.

To obtain a highly effective micro-segmentation strategy in polynomial time, a GAT+RL optimization algorithm is designed to seek the micro-segmentation strategy to maximize the combined cyber-physical metric. Simulations demonstrate the effectiveness of the proposed micro-segmentation technique, with a significant improvement in BDD detection rates against FDIAs, increasing from 5.23% to 94.02%. In addition, it also demonstrates the computational efficiency of the GAT+RL algorithm, which outperforms state-of-the-art methods while maintaining solution quality. These findings underscore the effectiveness of the proposed micro-segmentation technique in enhancing the security and robustness of modern CPPS.

Chapter 6

Conclusions and Future Works

This research investigates the problem of analyzing and defending two representative cross-layer threats in CPPS, including degraded QoS and FDIAs. Several defensive approaches are designed and enhanced to evaluate and mitigate the cross-later impact of QoS and FDIAs. In this chapter, the findings and limitations identified in the previous chapters are first summarized, and then potential future directions are discussed.

6.1 Conclusions

- **Conclusions to Research Objective 1:**

  Chapter 3 proposes a joint optimization framework to optimize the D2D communication resource allocation strategy, with the objectives of minimizing QoS disruptions and mitigating their cross-layer impacts on microgrid stability. To enhance the effectiveness of this optimization, the AW-MRAS algorithm is introduced, which efficiently identifies effective D2D allocation strategies.

- **Conclusions to Research Objective 2:**

  Chapter 4 proposes an economic vulnerability assessment to assess the economic risks induced by carbon emission costs, which can effectively identify the vulnerability power nodes while considering carbon cost. To assess node vulnerability under the most threatening scenarios, an attack strategy optimization framework is proposed, enabling the prediction of the most threatening attack behaviors while considering interactions with ISO operators.

- **Conclusions to Research Objective 3:**

  Chapter 5 proposes a micro-segmentation technique to strategically restrict lateral attack spreading among measuring devices, thereby disrupting the

stealthiness conditions of FDIA and improving the accuracy of residual-based BDD. To enhance the effectiveness of this technique, a cyber-physical-BDD-enhancement metric and a GAN+RT algorithm are developed. These developments facilitate the design of segmentation strategies that achieve polynomial-time efficiency and real-time responsiveness.

6.2 Future Work

In the summary sections of **Chapters 3, 4 and 5**, this research briefly discusses the limitations and future work. In addition to the future work mentioned in these chapters, additional limitations and directions for further exploration are highlighted as follows.

In **Chapter 3**, the effectiveness of optimal D2D communication allocation strategy cannot be ensured under extreme bandwidth limitations conditions, where time delays might exceed the tolerance of control algorithms, posing significant challenges to the proposed solutions. It is essential to fundamentally resolve bandwidth limitations, such as repairing damaged base stations to mitigate the challenges posed by such extreme scenarios.

In **Chapter 4**, the integration of market mechanisms, such as carbon emission credits, expands the attack surface and increases the dynamic characteristics of CPPS. This challenges both the predictive accuracy of attacker behavior and the effectiveness of the vulnerability assessment framework. In such scenarios, in addition to the vulnerability assessment, it is also necessary to strengthen market regulatory mechanisms to constrain attackers' abnormal arbitrage behavior.

In **Chapter 5**, the proposed micro-segmentation technique primarily targets cyber devices within sensor networks, simplifying the analysis by considering only transmission and measurement functions. However, in practice, the communication nodes are often heterogeneous, featuring diverse functions and network topologies, which leads to significant challenges for modeling and analysis of micro-segmentation techniques. This simplification limits the applicability of the proposed model in practical scenarios, which remains an open problem arising from the inherent interdependence of CPPS.

In **Chapters 3, 4 and 5** propose optimization framework incorporates both discrete and continuous features, presenting challenges for efficient resolution within polynomial time. As a result, heuristic algorithms are often preferred in practical applications, as they typically yield feasible solutions more rapidly than approximate or exact algorithms [67]. While the proposed heuristic algorithm facilitates prompt defensive decision-making, it has limitations in balancing solution quality with

computational efficiency during both the solving and training processes. Enhancing the solution quality of heuristic algorithms remains an open area of research.

In summary, CPPSs are complex systems with inherent interdependencies and cascading failures. Nowadays, no universal solution exists, as each specific cross-layer threat requires tailored defensive measures. Future research should focus on addressing these complexities by developing defensive approaches that comprehensively analyze multi-layer node interactions and collaboratively leverage strategies from all layers to design rapid response mechanisms that enhance scalability and adaptability in CPPSs. These unresolved challenges offer significant opportunities to advance the resilience and security of CPPS.

Appendix A

A.1 Supporting Materials for Chapter 3

A.1.1 Generation of Reconstructing sampling set S″

The generation of reconstructing the new sampling set S'' from the original set S is given as follows.

- The unique elements $\{x'^i_k\}_{i=1}^{N'}$ are selected from the sampling set $S = \{x^i_k\}_{i=1}^{N}$ in step 2 in ***Algorithm 1***.
- These elements $\{x'^i_k\}_{i=1}^{N'}$ are in a new set S' with $|S'| = N'$.
- For each element x'^i_k in set S', it is sampled n^i times. n^i is proportional to the value of $S(\mathcal{J}(x'^i_k))$.
- After repetitively sampling, these elements $\{y^i_k\}_{i=1}^{N''}$ are used to generate a new sampling set S''. $|S''|$ is defined as $N'' = n^1 + n^2 + \cdots + n^{N'}$.

A.1.2 Proof of Lemma 3.3

Referring to [113], the parameter θ_{k+1} is updated by

$$\begin{aligned}
\theta_{k+1} &= \arg\max_{\theta\in\Theta} E_{\theta_k}[\frac{[S(\mathcal{J}(X))]^k}{f(X,\theta_k)} I_{\{\mathcal{J}(X)\geq\tilde{\gamma}_{k+1}\}} \ln f(X,\theta)], \\
&= \arg\max_{\theta\in\Theta} \int_{\mathcal{X}} [\frac{[S(\mathcal{J}(X))]^k}{f(x,\theta_k)} I_{\{\mathcal{J}(x)\geq\tilde{\gamma}_{k+1}\}} \ln f(x,\theta)\cdot f(x,\theta_k)]\nu(dx), \\
&= \arg\max_{\theta\in\Theta} \int_{\mathcal{X}} [S(\mathcal{J}(x))]^k I_{\{\mathcal{J}(x)\geq\tilde{\gamma}_{k+1}\}} \ln f(x,\theta)\nu(dx),
\end{aligned}$$

The equal problem of estimating the parameters θ in $f(x,\theta)$ from sampling set $S'' = \{y_k^i\}_{i=1}^{N''}$ can be formulated as

$$\begin{aligned}
\theta_{k+1} &= \arg\max_{\theta\in\Theta} \prod_{i=1}^{N''} f(y_k^i,\theta) \\
&= \arg\max_{\theta\in\Theta} [\ln \prod_{i=1}^{N''} f\left(x_k^i,\theta\right)] = \arg\max_{\theta\in\Theta} \ln\{[\prod_{n^1} f(Y = x'^1_k,\theta) \\
&\cdot \prod_{n^2} f(Y = x'^2_k,\theta) \cdots \prod_{n^{N'}} f(Y = x'^{N'}_k,\theta)]\} \\
&= \arg\max_{\theta\in\Theta} \ln[f(Y = x'^1_k,\theta)^{S(\mathcal{J}(x'^1_k))]^k I\left\{\mathcal{J}(x'^1_k)\geq\bar{\gamma}_k\right\}} \\
&\cdot f(Y = x'^2_k,\theta)^{S(\mathcal{J}(x'^2_k))]^k I\left\{\mathcal{J}(x'^2_k)\geq\bar{\gamma}_k\right\}} \\
&\cdots f(Y = x'^{N'}_k,\theta)^{S(\mathcal{J}(x'^{N'}_k))]^k I\left\{\mathcal{J}(x'^{N'}_k)\geq\bar{\gamma}_k\right\}}] \\
&= \arg\max_{\theta\in\Theta} [S(\mathcal{J}(x'^1_k))]^k I\{\mathcal{J}(x'^1_k)\geq\bar{\gamma}_k\} \ln f(Y = x'^1_k,\theta) \\
&+ S(\mathcal{J}(x'^2_k))]^k I\{\mathcal{J}(x'^2_k)\geq\bar{\gamma}_k\} \ln f(Y = x'^2_k,\theta) \\
&\cdots S(\mathcal{J}(x'^{N'}_k))]^k I\{\mathcal{J}(x'^{N'}_k)\geq\bar{\gamma}_k\} \ln f(Y = x'^{N'}_k,\theta)] \\
&= \arg\max_{\theta\in\Theta} \int_{\mathcal{X}} [S(\mathcal{J}(x))]^k I_{\{\mathcal{J}(x)\geq\bar{\gamma}_k\}} \ln f(x,\theta)\nu(dx).
\end{aligned}$$

Hence, the optimization problem of Eqn. (3.19) can be converted to the problem of estimating the probability distribution function $f(x,\theta)$ from sampling set S'' via Eqn. (3.20).

A.1.3 Analysis of the Optimal Transmit Power

The asymptotic analysis is conducted to explore the relationship between energy efficiency and transmit power. Specifically, it investigates the energy efficiency in scenarios where the transmit power is either very small or very large. This analysis reveals that energy efficiency initially increases and then decreases as transmit power p^d rises, indicating the presence of an optimal transmit power.

Specifically, the energy efficiency is defined as $f_1^c(\mathbf{P}^d,\xi) = \frac{\mathcal{R}^c(\mathbf{P}^d,\xi)}{\mathcal{P}^c(\mathbf{P}^d,\xi)}$, where the total throughput is $\mathcal{R}^c(\mathbf{P}^d,\xi) = \sum_{i=1}^{\mathbf{I}} \mathcal{R}_i^c + \sum_{i=1}^{\mathbf{I}}\sum_{l=1}^{\mathbf{L}} \xi_{l,i}\mathcal{R}_l$ and the total power consumption is $\mathcal{P}^c(\mathbf{P}^d,\xi) = \frac{1}{2}(\sum_{i=1}^{\mathbf{I}} p_i^c + p_{BS}) + \sum_{i=1}^{\mathbf{I}}\sum_{l=1}^{\mathbf{L}} \xi_{l,i}\epsilon_l p_{l,i}^d$.

To simplify the description and facilitate the discussion, some constants are consolidated for ease of representation. Furthermore, once a specific power allocation matrix ξ is given, this energy efficiency in D2D linked l can be expressed as $f_1^c = \frac{\mathcal{R}^c(p^d)}{\mathcal{P}^c(p^d)}$, where $\mathcal{R}^c = \mathcal{R}^{c_i^{sum}} + \mathcal{R} = \mathcal{R}^{c_i^{sum}} + B^c \log_2(1+\gamma)$ and $\mathcal{P}^c = \mathcal{P}^{BS+c} + \epsilon \cdot p^d$. $\gamma_l = \frac{\sum_{i=1}^{\mathcal{I}} \xi_{l,i} p_{l,i}^d(k) g_l^d}{\sigma_0^2 + \sum_{i=1}^{\mathcal{I}} \xi_{l,i}\left(p_i^c g_{i,l}^c + \sum_{l'=1,l'\neq l}^{\mathcal{L}} \xi_{l',i} p_{l',i}^d(k) g_{l',l}^d\right)}$ is simplified as $\gamma = \frac{p^d g^d}{(\sigma_0^2 + I^c + I')}$.

(1) Low Transmit Power p^d Case: When p^d is very small, the SINR is also very small. Using the linear approximation for small SINR, it yields $\mathcal{R} = B^c \log_2\left(1 + \frac{p^d g^d}{\sigma_0^2 + I^c + I'}\right) \approx B^c \frac{p^d g^d}{\sigma_0^2 + I^c + I'} \cdot \log_2 e$. Then, the energy efficiency is $f_1^c \approx \frac{B^c \cdot \log_2 e \cdot \frac{p^d g^d}{\sigma_0^2 + I_c + I'} + \mathcal{R}^{c_i^{sum}}}{p^d + \mathcal{P}^{BS+c}}$. When p^d is very small and much smaller than $\mathcal{P}^{BS+c}$, the energy efficiency approximates to: $f_1^c \approx \frac{B \cdot \log_2 e \cdot \frac{p^d g^d}{\sigma_0^2 + I_c + I'} + \mathcal{R}^{c_i^{sum}}}{\mathcal{P}^{BS+c}} \approx \frac{B^c \cdot p^d \cdot g^d \cdot \log_2 e + \mathcal{R}^{c_i^{sum}} \cdot \left(\sigma_0^2 + I_c + I'\right)}{\left(\sigma_0^2 + I_c + I'\right) \cdot \mathcal{P}^{BS+c}}$. Therefore, the energy efficiency increases linearly with a small p^d.

(2) High Transmit Power p^d Case: When p^d is large, the SINR γ is also large. Using the logarithmic approximation for large SINR, it yields $\mathcal{R} \approx B^c \log_2\left(\frac{p^d g^d}{\sigma_0^2 + I_c + I'}\right)$. Then, the energy efficiency is $f_1^c \approx \frac{B^c \log_2\left(\frac{p^d g^d}{\sigma_0^2 + I_c + I'}\right) + \mathcal{R}^{c_i^{sum}}}{\mathcal{P}^{BS+c} + \epsilon \cdot p^d}$. Since the logarithmic function grows more slowly while the total power consumption ($\mathcal{P}^{BS+c} + \epsilon \cdot p^d$) increases linearly, the energy efficiency f_1^c will start to decrease when p^d is large. Based on the analysis, when p^d is relatively small, the initial value of energy efficiency is relatively low, and it increases as p^d rises. Conversely, when is p^d becomes substantially large, the initial energy efficiency is higher, but it begins to decrease with further increases. Without loss of generality, the complex relationship between energy efficiency and transmit power can be represented by the curve trend shown in Fig. 3.

References

[1] Hui Hou, Jianzhong Zhou, Yongchuan Zhang, and Xiongkai He. A brief analysis on differences of risk assessment between smart grid and traditional power grid. In *2011 Fourth International Symposium on Knowledge Acquisition and Modeling*, pages 188–191, 2011.

[2] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid — the new and improved power grid: A survey. *IEEE Communications Surveys & Tutorials*, 14(4):944–980, 2012.

[3] Mohammad Kamrul Hasan, AKM Ahasan Habib, Zarina Shukur, Fazil Ibrahim, Shayla Islam, and Md Abdur Razzaque. Review on cyber-physical and cyber-security system in smart grid: Standards, protocols, constraints, and recommendations. *Journal of Network and Computer Applications*, 209:103540, 2023.

[4] 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.

[5] Tullio Facchinetti and Marco L. Della Vedova. Real-time modeling for direct load control in cyber-physical power systems. *IEEE Transactions on Industrial Informatics*, 7(4):689–698, 2011.

[6] Huan Pan, Honghui Lian, Chunning Na, and Xiao Li. Modeling and vulnerability analysis of cyber-physical power systems based on community theory. *IEEE Systems Journal*, 14(3):3938–3948, 2020.

[7] Shahab Karamdel, Xiaodong Liang, Sherif O Faried, and Massimo Mitolo. Optimization models in cyber-physical power systems: A review. *IEEE Access*, 10:130469–130486, 2022.

[8] Baozhong Ti, Gengyin Li, Ming Zhou, and Jianxiao Wang. Resilience assessment and improvement for cyber-physical power systems under typhoon disasters. *IEEE Transactions on Smart Grid*, 13(1):783–794, 2022.

[9] 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.

[10] 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, 2017.

[11] Md. Mahmud Hasan and Hussein T. Mouftah. Optimal trust system placement in smart grid scada networks. *IEEE Access*, 4:2907–2919, 2016.

[12] Guoqing Zhang, Wengen Gao, Yunfei Li, Xinxin Guo, Pengfei Hu, and Jiaming Zhu. Detection of false data injection attacks in a smart grid based on wls and an adaptive interpolation extended kalman filter. *Energies*, 16(20), 2023.

[13] Qiuye Sun, Bin Wang, Xiaomeng Feng, and Shiyan Hu. Small-signal stability and robustness analysis for microgrids under time-constrained dos attacks and a mitigation adaptive secondary control method. *Science China Information Sciences*, 65, 2022.

[14] Morteza Sarailoo and N. Eva Wu. Cost-effective upgrade of pmu networks for fault-tolerant sensing. *IEEE Transactions on Power Systems*, 33(3):3052–3063, 2018.

[15] Subhash Lakshminarayana, E. Veronica Belmega, and H. Vincent Poor. Moving-target defense against cyber-physical attacks in power grids via game theory. *IEEE Transactions on Smart Grid*, 12(6):5244–5257, 2021.

[16] Chiara Bordin, Anne Håkansson, and Sambeet Mishra. Smart energy and power systems modelling: an iot and cyber-physical systems perspective, in the context of energy informatics. *Procedia Computer Science*, 176:2254–2263, 2020. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020.

[17] Ahmed Yousuf Saber and Ganesh Kumar Venayagamoorthy. Efficient utilization of renewable energy sources by gridable vehicles in cyber-physical energy systems. *IEEE Systems Journal*, 4(3):285–294, 2010.

[18] Yi Tang, Qian Chen, Mengya Li, Qi Wang, Ming Ni, and XiangYun Fu. Challenge and evolution of cyber attacks in cyber physical power system. In *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pages 857–862, 2016.

[19] Yunqi Wang, Jing Qiu, Yuechuan Tao, and Junhua Zhao. Carbon-oriented operational planning in coupled electricity and emission trading markets. *IEEE Transactions on Power Systems*, 35(4):3145–3157, 2020.

[20] Ruilong Deng, Peng Zhuang, and Hao Liang. Ccpa: Coordinated cyber-physical attacks and countermeasures in smart grid. *IEEE Transactions on Smart Grid*, 8(5):2420–2430, 2017.

[21] Eklas Hossain, Imtiaj Khan, Fuad Un-Noor, Sarder Shazali Sikander, and Md. Samiul Haque Sunny. Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access*, 7:13960–13988, 2019.

[22] Scott Rose, Oliver Borchert, Stuart Mitchell, and Sean Connelly. Zero trust architecture, 2020-08-10 04:08:00 2020.

[23] Kaiyisah Hanis Mohd Azmi, Nurul Asyikin Mohamed Radzi, Nayli Adriana Azhar, Faris Syahmi Samidi, Izzati Thaqifah Zulkifli, and Alisadikin Muhammad Zainal. Active electric distribution network: Applications, challenges, and opportunities. *IEEE Access*, 10:134655–134689, 2022.

[24] Andreas Ulbig, Tobias Rinke, Spyros Chatzivasileiadis, and Göran Andersson. Predictive control for real-time frequency regulation and rotational inertia provision in power systems. In *52nd IEEE Conference on Decision and Control*, pages 2946–2953, 2013.

[25] Shichao Liu, Xiaoyu Wang, and Peter Xiaoping Liu. Impact of communication delays on secondary frequency control in an islanded microgrid. *IEEE Transactions on Industrial Electronics*, 62(4):2021–2031, 2015.

[26] Claudio De Persis and Pietro Tesi. Input-to-state stabilizing control under denial-of-service. *IEEE Transactions on Automatic Control*, 60(11):2930–2944, 2015.

[27] Luo Xu, Qinglai Guo, Tianyu Yang, and Hongbin Sun. Robust routing optimization for smart grids considering cyber-physical interdependence. *IEEE Transactions on Smart Grid*, 10(5):5620–5629, 2019.

[28] A. Termehchi and M. Rasti. Joint sampling time and resource allocation for power efficiency in industrial cyber–physical systems. *IEEE Transactions on Industrial Informatics*, 17(4):2600–2610, 2021.

[29] Peng Yong Kong. Radio resource allocation scheme for reliable demand response management using d2d communications in smart grid. *IEEE Transactions on Smart Grid*, 11(3):2417–2426, 2020.

[30] Pudong Ge, Fei Teng, Charalambos Konstantinou, and Shiyan Hu. A resilience-oriented centralised-to-decentralised framework for networked microgrids management. *Applied Energy*, 308:118234, 2022.

[31] Mauro De Falco. Stuxnet facts report: A technical and strategic analysis, 2012.

[32] 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.

[33] Jingwen Liang, Oliver Kosut, and Lalitha Sankar. Cyber attacks on ac state estimation: Unobservability and physical consequences. In *2014 IEEE PES General Meeting — Conference & Exposition*, pages 1–5, 2014.

[34] György Dán and Henrik Sandberg. Stealth attacks and protection schemes for state estimators in power systems. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 214–219, 2010.

[35] Henrik Sandberg, André Teixeira, and Karl H Johansson. On security indices for state estimators in power networks. In *First workshop on secure control systems (SCS), Stockholm*, 2010.

[36] Le Xie, Yilin Mo, and Bruno Sinopoli. False data injection attacks in electricity markets. In *2010 First IEEE International Conference on Smart Grid Communications*, pages 226–231, 2010.

[37] 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, 2014.

[38] Le Xie, Yilin Mo, and Bruno Sinopoli. Integrity data attacks in power market operations. *IEEE Transactions on Smart Grid*, 2(4):659–666, 2011.

[39] Qiwei Zhang, Fangxing Li, and Xiaofei Wang. Cyber-impact analysis for iso revenue adequacy considering fdia in real-time market operations. *IEEE Transactions on Power Systems*, 38(5):4042–4053, 2023.

[40] Zhongjian Zhang, Shengrong Bu, Yanru Zhang, and Zhu Han. Market-level integrated detection against cyber attacks in real-time market operations by self-supervised learning. *IEEE Transactions on Smart Grid*, 15(4):4128–4142, 2024.

[41] Yujing Huang, Yudong Wang, and Nian Liu. Low-carbon economic dispatch and energy sharing method of multiple Integrated Energy Systems from the perspective of System of Systems. *Energy*, 244:122717, 2022.

[42] Qiang Fan, Dong Liu, and Jiaming Weng. Low-carbon operation of natural gas and electricity integrated energy systems considering carbon trading and demand response. In *2022 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5, 2022.

[43] Zhichao Yan, Chunyan Li, Yiming Yao, Weibin Lai, Jiyuan Tang, Changzheng Shao, and Qian Zhang. Bi-level carbon trading model on demand side for

integrated electricity-gas system. *IEEE Transactions on Smart Grid*, 14(4):2681–2696, 2023.

[44] Hang Zhang, Bo Liu, and Hongyu Wu. Smart grid cyber-physical attack and defense: A review. *IEEE Access*, 9:29641–29659, 2021.

[45] Kate L. Morrow, Erich Heine, Katherine M. Rogers, Rakesh B. Bobba, and Thomas J. Overbye. Topology perturbation for detecting malicious data injection. In *2012 45th Hawaii International Conference on System Sciences*, pages 2104–2113, 2012.

[46] Zhenyong Zhang, Ruilong Deng, David K. Y. Yau, and Peng Chen. Zero-parameter-information data integrity attacks and countermeasures in iot-based smart grid. *IEEE Internet of Things Journal*, 8(8):6608–6623, 2021.

[47] Edward A. Lee. The past, present and future of cyber-physical systems: A focus on models. *Sensors*, 15(3):4837–4869, 2015.

[48] Sizhe He, Yadong Zhou, Yujie Yang, Ting Liu, Yuxun Zhou, Jie Li, Tong Wu, and Xiaohong Guan. Cascading failure in cyber–physical systems: A review on failure modeling and vulnerability analysis. *IEEE Transactions on Cybernetics*, 54(12):7936–7954, 2024.

[49] Sayawu Yakubu Diaba, Miadrezah Shafie-khah, and Mohammed Elmusrati. Cyber-physical attack and the future energy systems: A review. *Energy Reports*, 12:2914–2932, 2024.

[50] Haicheng Tu, Yongxiang Xia, Jiajing Wu, and Xiang Zhou. Robustness assessment of cyber–physical systems with weak interdependency. *Physica A: Statistical Mechanics and its Applications*, 522:9–17, 2019.

[51] Seyed Mehrdad, Seyed Mousavian, Gholamreza Madraki, et al. Cyber-physical resilience of electrical power systems against malicious attacks: A review. *Current Sustainable Renewable Energy Reports*, 5(1):14–22, March 2018. Published on 25 January 2018.

[52] Kirti Gupta, Subham Sahoo, and Bijaya Ketan Panigrahi. A monolithic cybersecurity architecture for power electronic systems. *IEEE Transactions on Smart Grid*, 15(4):4217–4227, 2024.

[53] Abdullah Abusorrah, Ahmed Alabdulwahab, Zhiyi Li, and Mohammad Shahidehpour. Minimax-regret robust defensive strategy against false data injection attacks. *IEEE Transactions on Smart Grid*, 10(2):2068–2079, 2017.

[54] Kirti Gupta, Subham Sahoo, and Bijaya Ketan Panigrahi. Delay-aware semantic sampling in power electronic systems. *IEEE Transactions on Smart Grid*, 15(4):4038–4049, 2024.

[55] Kun Wang, Miao Du, Sabita Maharjan, and Yanfei Sun. Strategic honeypot game model for distributed denial of service attacks in the smart grid. *IEEE Transactions on Smart Grid*, 8(5):2474–2482, 2017.

[56] Liwei An and Guang-Hong Yang. Decentralized adaptive fuzzy secure control for nonlinear uncertain interconnected systems against intermittent dos attacks. *IEEE transactions on cybernetics*, 49(3):827–838, 2018.

[57] Hasan Ali and Dipankar Dasgupta. Effects of time delays in the electric power grid. In Jonathan Butts and Sujeet Shenoi, editors, *Critical Infrastructure Protection VI*, pages 139–154, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[58] Qiwei Zhang, Fangxing Li, Qingxin Shi, Kevin Tomsovic, Jinyuan Sun, and Lingyu Ren. Profit-oriented false data injection on electricity market: Reviews, analyses, and insights. *IEEE Transactions on Industrial Informatics*, 17(9):5876–5886, 2021.

[59] Fangxing Li, Yanli Wei, and Sarina Adhikari. Improving an unjustified common practice in ex post lmp calculation. *IEEE Transactions on Power Systems*, 25(2):1195–1197, 2010.

[60] Gabriela Hug and Joseph Andrew Giampapa. Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks. *IEEE Transactions on Smart Grid*, 3(3):1362–1370, 2012.

[61] Charalampos Kalalas, Linus Thrybom, and Jesus Alonso-Zarate. Cellular communications for smart grid neighborhood area networks: A survey. *IEEE Access*, 4:1469–1493, 2016.

[62] Kent Tsz Kan Cheung, Shaoshi Yang, and Lajos Hanzo. Achieving maximum energy-efficiency in multi-relay ofdma cellular networks: A fractional programming approach. *IEEE Transactions on Communications*, 61(7):2746–2757, 2013.

[63] Roya Arab Loodaricheh, Shankhanaad Mallick, and Vijay K. Bhargava. Energy-efficient resource allocation for ofdma cellular networks with user cooperation and qos provisioning. *IEEE Transactions on Wireless Communications*, 13(11):6132–6146, 2014.

[64] Josep M. Guerrero, Juan C. Vasquez, José Matas, Luis García de Vicuna, and Miguel Castilla. Hierarchical control of droop-controlled ac and dc microgrids—a general approach toward standardization. *IEEE Transactions on Industrial Electronics*, 58(1):158–172, 2011.

[65] John W. Simpson-Porco, Florian Dörfler, and Francesco Bullo. Synchronization and power sharing for droop-controlled inverters in islanded microgrids. *Automatica*, 49(9):2603–2611, 2013.

[66] Hadi Saadat. *Power System Analysis*. Wiley, 2002.

[67] Nina Mazyavkina, Sergey Sviridov, Sergei Ivanov, and Evgeny Burnaev. Reinforcement learning for combinatorial optimization: A survey. *Computers & Operations Research*, 134:105400, 2021.

[68] Chaitanya K. Joshi, Quentin Cappart, Louis-Martin Rousseau, and Thomas Laurent. Learning tsp requires rethinking generalization. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[69] Mengmeng Yu and Seung Ho Hong. A real-time demand-response algorithm for smart grids: a stackelberg game approach. *IEEE Transactions on Smart Grid*, 7(2):879–888, 2016.

[70] Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. *Advances in neural information processing systems*, 30, 2017.

[71] Mete Ozay, Iñaki Esnaola, Fatos T. Yarman Vural, Sanjeev R. Kulkarni, and H. Vincent Poor. Smarter security in the smart grid. In *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pages 312–317, 2012.

[72] Mete Ozay, Iñaki 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.

[73] Chao Ren and Yan Xu. A fully data-driven method based on generative adversarial networks for power system dynamic security assessment with missing data. *IEEE Transactions on Power Systems*, 34(6):5044–5052, 2019.

[74] Yuancheng Li, Yuanyuan Wang, and Shiyan Hu. Online generative adversary network based measurement recovery in false data injection attacks: A cyber-physical approach. *IEEE Transactions on Industrial Informatics*, 16(3):2031–2043, 2019.

[75] Bo Chen, Daniel W. C. Ho, Guoqiang Hu, and Li Yu. Secure fusion estimation for bandwidth constrained cyber-physical systems under replay attacks. *IEEE Transactions on Cybernetics*, 48(6):1862–1876, 2018.

[76] Jun Yan, Haibo He, Xiangnan Zhong, and Yufei Tang. Q-learning-based vulnerability analysis of smart grid against sequential topology attacks. *IEEE Transactions on Information Forensics and Security*, 12(1):200–210, 2016.

[77] Yingshuai Hao, Meng Wang, and Joe H Chow. Likelihood analysis of cyber data attacks to power systems with markov decision processes. *IEEE Transactions on Smart Grid*, 9(4):3191–3202, 2016.

[78] Chris Y. T. Ma, David K. Y. Yau, and Nageswara S. V. Rao. Scalable solutions of markov games for smart-grid infrastructure protection. *IEEE Transactions on Smart Grid*, 4(1):47–55, 2013.

[79] Longfei Wei, Arif I Sarwat, Walid Saad, and Saroj Biswas. Stochastic games for power grid protection against coordinated cyber-physical attacks. *IEEE Transactions on Smart Grid*, 9(2):684–694, 2016.

[80] Chris Y. T. Ma, David K. Y. Yau, Xin Lou, and Nageswara S. V. Rao. Markov game analysis for attack-defense of power networks under possible misinformation. *IEEE Transactions on Power Systems*, 28(2):1676–1686, 2013.

[81] Ruilong Deng, Gaoxi Xiao, and Rongxing Lu. Defending against false data injection attacks on power system state estimation. *IEEE Transactions on Industrial Informatics*, 13(1):198–207, 2015.

[82] Evan Gilman and Doug Barth. *Zero Trust Networks: Building Secure Systems in Untrusted Networks*. O'Reilly Media, Inc., 1st edition, 2017.

[83] Executive order on improving the nation's cybersecurity. Technical report, THE WHITE HOUSE, May 2021.

[84] Shalanda D. Young. Moving the u.s. government toward zero trust cybersecurity principles. https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf/, 2022. Accessed January 16, 2022.

[85] OMB. Office of management and budget releases federal strategy to move the u.s. government towards a zero trust architecture. https://www.whitehouse.gov/omb/briefing-room/2022/01/26/, 2022. Accessed January 16, 2022.

[86] Okta. The state of zero trust security 2022. https://www.okta.com/sites/default/files/2022-09/OKta_WhitePaper_ZeroTrust_H2_Campaign_.pdf/, 2022. Accessed January 16, 2022.

[87] C.L. Smith. Understanding concepts in the defence in depth strategy. In *Proceedings of the IEEE 37th Annual 2003 International Carnahan Conference on Security Technology*, pages 8–16, 2003.

[88] Zhang Xiaojian, Chen Liandong, Fan Jie, Wang Xiangqun, and Wang Qi. Power iot security protection architecture based on zero trust framework. In *2021 IEEE 5th International Conference on Cryptography, Security and Privacy (CSP)*, pages 166–170, 2021.

[89] Annamalai Alagappan, Sampath Kumar Venkatachary, and Leo John Baptist Andrews. Augmenting zero trust network architecture to enhance security in virtual power plants. *Energy Reports*, 8:1309–1320, 2022.

[90] Yahuza Bello, Ahmed Refaey Hussein, Mehmet Ulema, and Juanita Koilpillai. On sustained zero trust conceptualization security for mobile core networks in 5g and beyond. *IEEE Transactions on Network and Service Management*, 19(2):1876–1889, 2022.

[91] Luca Ferretti, Federico Magnanini, Mauro Andreolini, and Michele Colajanni. Survivable zero trust for cloud computing environments. *Computers & Security*, 110:102419, 2021.

[92] Sirshak Sarkar, Gaurav Choudhary, Shishir Kumar Shandilya, Azath Hussain, and Hwankuk Kim. Security of zero trust networks in cloud computing: A comparative review. *Sustainability*, 14(18), 2022.

[93] Baozhan Chen, Siyuan Qiao, Jie Zhao, Dongqing Liu, Xiaobing Shi, Minzhao Lyu, Haotian Chen, Huimin Lu, and Yunkai Zhai. A security awareness and protection system for 5g smart healthcare based on zero-trust architecture. *IEEE Internet of Things Journal*, 8(13):10248–10263, 2021.

[94] Sultana Maliha, Afrida Hossain, Laila Fabiha, Abu Taher Kazi, and Nazrul Islam Muhammad. Towards developing a secure medical image sharing system based on zero trust principles and blockchain technology. *BMC Medical Informatics and Decision Making*, 20(256):1472–6947, 2020.

[95] Rory Ward and Betsy Beyer. Beyondcorp: A new approach to enterprise security. https://research.google/pubs/beyondcorp-a-new-approach-to-enterprise-security/, 2014. Accessed: 2025-01-28.

[96] Zirak Zaheer, Hyunseok Chang, Sarit Mukherjee, and Jacobus Van der Merwe. Eztrust: Network-independent zero-trust perimeterization for microservices. In *Proceedings of the 2019 ACM Symposium on SDN Research*, SOSR '19, page 49–61, New York, NY, USA, 2019. Association for Computing Machinery.

[97] Netskope. Netskope announces general availability of zero trust secure access for hybrid it environments. https://www.netskope.com/press-releases/netskope-announces-general-availability-of-zero-trust-secure-access-for-hybrid-it-e 2025. Accessed: 2025-01-28.

[98] Zero trust maturity model. https://www.microsoft.com/en-us/security/blog/2020/04/02/announcing-microsoft-zero-trust-assessment-tool/, 2020.

[99] Zebing Feng, Peng Zhou, Qi Wang, and Weiqiang Qi. A dual-layer zero trust architecture for 5g industry mec applications access control. In *2022 IEEE 5th International Conference on Electronic Information and Communication Technology (ICEICT)*, pages 100–105, 2022.

[100] Mary K. Pratt. Zero-trust model case study: One ciso's experience. https://www.techtarget.com/searchsecurity/feature/Even-with-a-roadmap-zero-trust-model-an-ongoing-process, 2021. Accessed: 2025-01-28.

[101] Palo Alto Networks. Zero trust best practices. https://docs.paloaltonetworks.com/best-practices/zero-trust-best-practices, 2021. Accessed: 2025-01-28.

[102] Mathaios Panteli, Dimitris N. Trakas, Pierluigi Mancarella, and Nikos D. Hatziargyriou. Boosting the power grid resilience to extreme weather events using defensive islanding. *IEEE Transactions on Smart Grid*, 7(6):2913–2922, 2016.

[103] Chongyu Wang, Mingyu Yan, Kaiyuan Pang, Fushuan Wen, and Fei Teng. Cyber-physical interdependent restoration scheduling for active distribution network via ad hoc wireless communication. *IEEE Transactions on Smart Grid*, 14(5):3413–3426, 2023.

[104] Anushka M. Dissanayake and Nishantha C. Ekneligoda. Multiobjective optimization of droop-controlled distributed generators in dc microgrids. *IEEE Transactions on Industrial Informatics*, 16(4):2423–2435, 2020.

[105] Peng-Yong Kong. Multicell d2d communications for hierarchical control of microgrid system. *IEEE Systems Journal*, 15(2):1929–1938, 2021.

[106] Fanghong Guo, Changyun Wen, Jianfeng Mao, and Yong-Duan Song. Distributed secondary voltage and frequency restoration control of droop-controlled inverter-based microgrids. *IEEE Transactions on industrial Electronics*, 62(7):4355–4364, 2014.

[107] J.G. Proakis and M. Salehi. *Digital Communications, 5th edition*. McGraw-Hill Higher Education, 2008.

[108] Anushka M Dissanayake and Nishantha C Ekneligoda. Multiobjective optimization of droop-controlled distributed generators in dc microgrids. *IEEE Transactions on Industrial Informatics*, 16(4):2423–2435, 2019.

[109] Kaisa Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 2012.

[110] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[111] Xi Chen. *New model-based methods for non-differentiable optimization*. University of Illinois at Urbana-Champaign, 2015.

[112] Xiaoliang Ma, Yanan Yu, Xiaodong Li, Yutao Qi, and Zexuan Zhu. A survey of weight vector adjustment methods for decomposition-based multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 24(4):634–649, 2020.

[113] Jiaqiao Hu, Michael C. Fu, and Steven I. Marcus. A model reference adaptive search method for global optimization. *Operations Research*, 55(3):549–568, 2007.

[114] Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, 18(4):577–601, 2014.

[115] Pengcheng Fang, Jianjun Yang, Qingmiao Liao, Ray Y. Zhong, and Yuchen Jiang. Flexible worker allocation in aircraft final assembly line using multiobjective evolutionary algorithms. *IEEE Transactions on Industrial Informatics*, 17(11):7468–7478, 2021.

[116] Jun-Qing Li, Xiao-Long Chen, Pei-Yong Duan, and Jian-Hui Mou. Kmoea: A knowledge-based multiobjective algorithm for distributed hybrid flow shop in a prefabricated system. *IEEE Transactions on Industrial Informatics*, 18(8):5318–5329, 2022.

[117] Pathways to net-zero greenhouse gas emissions by 2050. `https://www.whitehouse.gov/wp-content/uploads/2021/10/US-Long-Term-Strategy.pdf`, 2021. Accessed November 18, 2023.

[118] Dae-Hyun Choi and Le Xie. Sensitivity analysis of real-time locational marginal price to scada sensor data corruption. *IEEE Transactions on Power Systems*, 29(3):1110–1120, 2014.

[119] Dajun Du, Minggao Zhu, Dakui Wu, Xue Li, Minrui Fei, Yukun Hu, and Kang Li. Distributed security state estimation-based carbon emissions and economic cost analysis for cyber–physical power systems under hybrid attacks. *Applied Energy*, 353:122001, 2024.

[120] Qiwei Zhang, Fangxing Li, Hantao Cui, Rui Bo, and Lingyu Ren. Market-level defense against fdia and a new lmp-disguising attack strategy in real-time market operations. *IEEE Transactions on Power Systems*, 36(2):1419–1431, 2021.

[121] Parth Pradhan, Kostas Hatalis, Shalinee Kishore, Rick S. Blum, and Alberto J. Lamadrid. Prospects of wave power grid integration. In *2014 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5, 2014.

[122] New york independent system operator (nyiso). https://www.nyiso.com/power-grid-data, 2023.

[123] Jiwei Tian, Chao Shen, Buhong Wang, Xiaofang Xia, Meng Zhang, Chenhao Lin, and Qian Li. Lesson: Multi-label adversarial false data injection attack for deep learning locational detection. *IEEE Transactions on Dependable and Secure Computing*, pages 1–15, 2024.

[124] Wangkun Xu, Martin Higgins, Jianhong Wang, Imad M. Jaimoukha, and Fei Teng. Blending data and physics against false data injection attack: An event-triggered moving target defence approach. *IEEE Transactions on Smart Grid*, 14(4):3176–3188, 2023.

[125] Wangkun Xu, Imad M. Jaimoukha, and Fei Teng. Robust moving target defence against false data injection attacks in power grids. *IEEE Transactions on Information Forensics and Security*, 18:29–40, 2023.

[126] Xiaomeng Feng and Shiyan Hu. Cyber-physical zero trust architecture for industrial cyber-physical systems. *IEEE Transactions on Industrial Cyber-Physical Systems*, 1:394–405, 2023.

[127] Mohammad Ashiqur Rahman and Amarjit Datta. Impact of stealthy attacks on optimal power flow: A simulink-driven formal analysis. *IEEE Transactions on Dependable and Secure Computing*, 17(3):451–464, 2020.

[128] Sara Mohammadi, Frank Eliassen, Yan Zhang, and Hans-Arno Jacobsen. Detecting false data injection attacks in peer to peer energy trading using machine learning. *IEEE Transactions on Dependable and Secure Computing*, 19(5):3417–3431, 2022.

[129] Christoph Buck, Christian Olenberger, André Schweizer, Fabiane Völter, and Torsten Eymann. Never trust, always verify: A multivocal literature review on current knowledge and research gaps of zero-trust. *Computers & Security*, 110:102436, 2021.

[130] Yonghe Guo, Chee-Wooi Ten, Shiyan Hu, and Wayne W. Weaver. Preventive maintenance for advanced metering infrastructure against malware propagation. *IEEE Transactions on Smart Grid*, 7(3):1314–1328, 2016.

[131] 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, 2021.

[132] Jichao Bi, Fengji Luo, Shibo He, Gaoqi Liang, Wenchao Meng, and other. False data injection-and propagation-aware game theoretical approach for microgrids. *IEEE Transactions on Smart Grid*, 13(5):3342–3353, 2022.

[133] Dong Wang, Michael Small, and Yi Zhao. Exploring the optimal network topology for spreading dynamics. *Physica A: Statistical Mechanics and its Applications*, 564, 2021.

[134] Kaiwen Li, Tao Zhang, Rui Wang, Yuheng Wang, Yi Han, and Ling Wang. Deep reinforcement learning for combinatorial optimization: Covering salesman problems. *IEEE Transactions on Cybernetics*, pages 1–14, 2021.

[135] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

[136] Yigu Liu, Haiwei Xie, Alfan Presekal, Alexandru Stefanov, and Peter Palensky. A gnn-based generative model for generating synthetic cyber-physical power system topology. *IEEE Transactions on Smart Grid*, 14(6):4968–4971, 2023.

[137] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, page 6000–6010, 2017.

[138] Matpower. `https://matpower.org/docs/MATPOWER-manual.pdf/`. Available September 13, 2022.

[139] Zhenyong Zhang, Ruilong Deng, Peng Cheng, and Mo-Yuen Chow. Strategic protection against fdi attacks with moving target defense in power grids. *IEEE Transactions on Control of Network Systems*, 2021.

[140] Oriol Vinyals, Meire Fortunato, and Navdeep Jaitly. Pointer networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, page 2692–2700, 2015.