Optimized AI Detection Methods for Countering Adversarial Attacks Against Vehicle-to-Microgrid Services

by

Ahmed Mohamed Elsayed Omara

Thesis Supervisor : Prof. Burak Kantarci

A thesis
presented to the University of Ottawa
in fulfillment of the
thesis requirement for the degree of
Ph.D

School of Electrical and Computer Engineering
Faculty of Engineering
University of Ottawa

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In the past decade, communication technologies such as cellular networks, Wi-Fi, and optical communication have significantly advanced, impacting daily life and enhancing urban preparedness for power outages. Smart grids, unlike traditional utility grids, enable bi-directional flows of electricity and information, improving efficiency and minimizing power losses by exchanging grid status and customer requirements. However, these advancements have also increased the attack surface, introducing new cyber vulnerabilities that adversaries can exploit, posing threats to smart grids. Our work addresses the security challenges arising from integrating Electric Vehicles (EVs), smart microgrids, and Artificial Intelligence (AI) in Vehicle-to-Microgrid (V2M) applications. First, the research investigates the growing attack surface resulting from the integration of EVs and smart grids, particularly focusing on data integrity attacks that pose a significant threat to V2M applications. A scheme leveraging unsupervised ML techniques is proposed to model and detect these attacks. Extensive simulations demonstrate the scheme's effectiveness in reducing the impact of data integrity attacks by up to 76.5%. Next, we explore Adversarial Machine Learning (AML) attacks targeting V2M services. These attacks exploit vulnerabilities in the ML classifiers, enabling adversaries to deceive the system and disrupt microgrid operations. To anticipate and counteract these threats, we conduct an anticipatory analysis of a multi-stage attack. By simulating adversary behavior, we evaluate the robustness of the ML classifier and develop effective countermeasures. Our findings reveal that the multi-stage gray-box attack achieves an Evasion Increase Rate (EIR) of up to 73.2%, using 40% less data than traditional white-box attacks. To enhance the security of AI-based microgrid control systems in V2M services, we propose a comprehensive defense framework integrating a Generative Adversarial Network (GAN) model and a robust ML classifier. The GAN model generates realistic adversarial samples, enabling the ML classifier to learn and adapt to novel attack patterns. Additionally, the ML classifier is trained on a diverse dataset comprising both legitimate and adversarial samples, improving its ability to distinguish between normal and malicious activities. Simulations validate the effectiveness of the proposed defense mechanism, achieving an Adversarial Detection Rate (ADR) of 90.2%. To address the limited computational power and memory in V2M edge settings, we examine different model optimization techniques, such as projection, pruning, and quantization, to optimize the model's size without compromising detection performance. The proposed method integrates model design and compression, resulting in an optimized detection model that remains robust against adversarial attacks. This approach ensures that the model remains compact and maintains high accuracy. For instance, the Convolutional Neural Network (CNN) model's detection rate against Fast Gradient Sign Method (FGSM) attacks is 92.5% and 91% before and after compression, respectively.

Acknowledgements

I would like to express my deepest appreciation to all those who provided me the possibility to make this thesis. A special gratitude I give to my thesis supervisor and mentor, Prof. Burak Kantarci, whose contribution of stimulating suggestions, encouragement, consistent care, and advice helped me to coordinate my thesis. His patience and unwavering support during challenging times provided me with the resilience and confidence needed to persevere in my studies.

I would like also to thank my wife, Sarah, for her strong support and belief in me. I would like to express my deepest gratitude to my parents for their unconditional love and support throughout my academic journey. My father has always been a source of inspiration and encouragement for me, and I owe him a lot for his guidance and wisdom. My mother, who passed away during my studies, was a pillar of strength and resilience for me, and I dedicate this thesis to her memory.

Table of Contents

List of Tables

List of Figures

List of Symbols

A_{im}^{nr}	Binary variable defines the multiplication of O_m^{nr} by O_m^{nr+1}
C	Waiting cost per hour ($/h)
D_m	Energy demand of microgrid m (kWh)
E_n	Initial battery level of EV n (kWh)
H	Electrical vehicle charging power (kW)
K_{im}	Distance between microgrid i and microgrid m (km)
M	Number of microgrids
N	Number of participating electrical vehicles (EVs) in the power supply request
O_m^{nr}	Binary variable is one if microgrid m is served by EV n in order r
P_n	Electrical energy price offered by EV n ($/kWh)
S_m^n	Amount of energy that EV n provides to microgrid m (kWh)
T_m	Distance from the initial location of EV n to first microgrid m (km)
U_n	Revenue of EV n ($)
α_n	Percentage value determine the contribution value of the EV n
v	Constant value defines the cost of the service request made by the microgrid ($/request)
z	Average energy consumption per km (kWh/km)

List of Abbreviations

$ADSL$	Asymmetric Digital Subscriber Line
AON	Active Optical Network
AMI	Advanced Metering Infrastructure
AML	Adversarial Machine Learning
ADR	Adversarial Detection Rate
BAN	Building Area Network
$BB-PLC$	Broadband PLC
BS	Base Station
BFS	Backward Feature Selection
bps	bit per second
CNR	Carrier-to-Noise Ratio
CR	Cognitive Radio
DER	Distributed Energy Resources
DMS	Distribution Management System
DG	Distributed Generator
DSL	Wireless Local Area Network
DSM	Dynamic Spectrum Management
$DSSS$	Direct Sequence Spread Spectrum
$DSRC$	Dedicated Short Range Communication
DTV	Digital TV

DT	Decision Tree
DWT	Discrete Wavelet Transform
EV	Electric Vehicle
EPC	Evolve Packet Core
EIR	Evasion Increase Ratio
FD	Frequency Domain
FL	Fuzzy Logic
FFS	Forward Feature Selection
GPR	Gaussian Process Regression
GSM	Global System for Mobile Communications
$HetNet$	Heterogeneous Network
$HSPA+$	Evolved High-Speed Packet Access
IoT	Internet of Things
LEO	Low Earth Orbits
LTE	Long Term Evolution
$LP-WAN$	Lower Power Wide Area Networks
$mmWave$	millimetre Wave
$MIMO$	Multi-Input, Multi-Output
$MILP$	Mixed Integer Linear Programming
$NB-PLC$	Narrowband PLC
$NREL$	National Renewable Energy Laboratory
$OFDMA$	Orthogonal Frequency Division Multiple Access
$OFDM$	Orthogonal Frequency Division Multiplexing
PCA	Principle Component Analysis
$PDSCH$	Physical Down-link Shared Channel
PON	Passive Optical Network
PLC	Power Line Communication

PLR	Packet Loss Rate
PMU	Phasor Measurement Unit
QoS	Quality of Service
RAN	Radio Access Network
RSN	Roadside Unit
$RMSE$	Root Mean Square Error
SAE	Stacked Auto-Encoder
$SCADA$	Supervisory Control And Data Acquisition
SGD	Stochastic Gradient Descent
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
TCP	Transmission Control Protocol
$UMTS$	Universal Mobile Telecommunications System
UE	User Equipment
$V2G$	Vehicle-to-Grid
$VDSL$	Very-high bit rate Digital Subscriber Line
$V2G$	Vehicle-to-Grid
$V2I$	Vehicle-to-Infrastructure
$V2M$	Vehicle-to-Microgrid
$V2V$	Vehicle-to-Vehicle
$WLAN$	Wireless Local Area Network
$WWAN$	Wireless Wide Area Network
$WiMAX$	Worldwide Interoperability for Microwave access
$WPAN$	Wireless Personal Area Network

Candidate's Publications During his Studies

- **Ahmed Omara**, and B. Kantarci, "AI Model Optimization for Adversarial Attacks Detection on Edge Devices in V2M Services," under submission
- **Ahmed Omara**, and B. Kantarci, "Generative Adversarial Networks to Secure Vehicle-to-Microgrid Services," IEEE Virtual Communications Conference (VCC), Nov, 2023.
- **Ahmed Omara**, and B. Kantarci, "An AI-Driven Solution to Prevent Adversarial Attacks on Mobile Vehicle-to-Microgrid Services," in Elsevier Journal of Simulation Modelling Practice and Theory, Sept, 2024.
- **Ahmed Omara**, and B. Kantarci, " Adversarial Machine Learning-Based Anticipation of Threats Against Vehicle-to-Microgrid Services," 2022 IEEE Global Communications Conference (GlobeCom), Rio, Brazil, 2022
- **Ahmed Omara**, and B. Kantarci, "On the Impact of Data Integrity Attacks on Vehicle-to-Microgrid Services," 2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), 2021
- M. Simsek, **Ahmed Omara**, and B. Kantarci, "Cost-aware Data Aggregation and Energy Decentralization with Electrical Vehicles in Microgrids through LTE Links," 2020 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Odessa, Ukraine, 2020

Chapter 1

Introduction

Reliable, sustainable, and resilient electric power systems are essential for modern societies. These goals require the distribution and diversification of power sources, which could be facilitated by smart grids [1]. A smart grid is an electric power system comprised of sensors, communication technologies, and control units to provide customers with better power services [2]. It enables bi-directional communication between the control units and the end-loads, contrary to the traditional utility grid that uses uni-directional communication (generation to consumers) [3]. The integration of Internet-of-Things (IoT) networks allows smart grids to monitor, control and manage the grid [4]. Power system reliability, sustainability, and resiliency as major concerns in smart grids call for proactive emergency preparedness and self-recovery solutions [5].

Transactive energy represents an innovative approach to achieving a balanced electricity supply and demand within an electric power system, leveraging communication networks, market-based mechanisms, and information technology. By optimizing the utilization of available energy resources, transactive energy systems aim to enhance grid reliability, stability, and cost-effectiveness [6]. These systems rely on the integration of smart grids, Advanced Meter Infrastructure (AMI), and communication networks to facilitate real-time information and electricity exchange among prosumers. The utilization of communication networks, such as cellular networks and Wi-Fi, plays a crucial role in enabling transactive energy systems. Smart grids leverage these communication technologies to monitor energy demand and supply in real-time [7]. Additionally, AMI utilizes cellular networks and Wi-Fi to communicate with smart grids, providing valuable insights into energy consumption patterns [8]. This empowers prosumers to monitor their energy usage and make adjustments based on market conditions. Furthermore, communication networks foster the emergence of new business models like peer-to-peer energy trading. This enables indi-

viduals and organizations to trade energy directly with one another, bypassing traditional utility companies. The integration of communication networks also facilitates the seamless integration of renewable energy sources, such as solar and wind power, into the power grid, contributing to the reduction of carbon emissions and the promotion of a sustainable energy system [9].

Within the realm of transactive energy systems, V2M systems stand out as a prominent application. These systems leverage EVs to supply excess electricity to localized smart microgrids [10]. Similar to other transactive energy systems, V2M systems heavily rely on communication technologies to enable real-time information exchange between EVs and smart microgrids. Notably, communication protocols like the Open Automated Demand Response (OpenADR) protocol facilitate seamless communication between microgrids and EVs through a central hub. Additionally, V2M systems make use of wireless communication technologies, such as Wi-Fi or cellular networks, to enable remote monitoring, control, and data exchange between EVs and other components of the smart grid ecosystem, including renewable energy systems and energy storage devices. By harnessing the power of communication technologies, V2M systems offer a multitude of benefits for both EV owners and microgrid operators. Enabled by communication capabilities, V2M services provide an efficient and reliable approach to managing energy resources, allowing microgrids to dynamically balance electricity supply and demand in real-time based on the availability and charging status of EVs [11]. This optimization leads to reduced energy costs, enhanced grid stability, and optimal utilization of available energy resources.

To optimize the operation of V2M systems, ML algorithms play a critical role. These algorithms utilize data-driven insights to enhance the efficiency, reliability, and cost-effectiveness of V2M operations [12]. ML algorithms can effectively predict the power demand of microgrids and optimize the charging and discharging schedules of EVs accordingly. By ensuring a continuous and sufficient power supply, ML-powered V2M systems prevent power outages and other disruptions, improving overall system performance [13].

1.1 Motivation

The possibility of using EV batteries to provide ancillary services to support smart grids has been studied for more than one decade [14]. EVs can be connected to a smart microgrid using bi-directional charging technology. This technology allows the EVs not only to draw energy from the grid to charge their batteries but also to supply energy back to the grid [15]. Such connections are facilitated through smart charging stations that are integrated into the microgrid's management system, enabling precise control over both

charging and discharging processes to optimize grid stability and reliability during various conditions, including power outages [16]. Furthermore, commercial EVs such as Kia EV6 and Nissan Leaf are equipped with bi-directional technologies and battery capacity to power a household for three consecutive days, as in F-150 Lightning [17]. Moreover, businesses such as Fermata Energy is another example of V2M feasibility and growing interest.

During power outages where traditional grid infrastructure is compromised, the literature suggests innovative approaches for direct energy transactions between EVs and smart grids, such as Vehicle-to-Grid (V2G) [18]. This concept, outlined in various studies, emphasizes the role of EVs as decentralized energy resources capable of supporting the stability and reliability of smart grids through direct connections. This approach bypasses traditional grid dependencies, offering a resilient alternative for energy distribution and management. V2G has been intensively studied in the literature with different objectives such as incentive-participation schemes [19] [20], data privacy and security [21] and energy management [22]. The authors in [23] provided a systematic literature review of the V2G technological and methodological frameworks enabling such interactions, including the importance of smart charging infrastructure and the potential for EVs to contribute to grid reliability and distributed generation. In the context of incentive-participation schemes, there are many research works that focus on defining the social and economic benefits for EV owners to sell their battery energy back to the grid [24] [25]. Other examples from the industry include the collaboration between Nissan and Enel power company to launch a V2G trial project in the UK [26]. This initiative involved setting up 100 V2G chargers for use by Nissan Leaf drivers, offering them the possibility to sell excess energy back to the grid for a profit. Given that most vehicles remain parked 95% of the time, their batteries have the potential to supply electricity back to the grid, potentially benefiting utility companies by up to $4000 annually per vehicle [27]. Recently, the V2G concept has evolved to V2M to provide additional layers of flexibility and robustness to smart grids [28]. This allows EV owners to directly support smart microgrids with their electrical needs.

Significant strides have been made in communication technologies, such as cellular networks, Wi-Fi, and optical communication, in recent years. These advancements have not only transformed our daily lives but have also empowered cities to enhance their resilience in the face of power outages. The collection and exchange of data enabled by these technologies have facilitated real-time monitoring of transmission and distribution lines, bolstering preparedness for disruptions. Smart grids, in contrast to conventional utility grids, have emerged as dynamic systems capable of facilitating the bidirectional flow of electricity and information among various entities within the grid.

The advent of smart grids has brought numerous benefits, including reduced power losses and improved efficiency in electricity generation and distribution. This is achieved through

the seamless exchange of information between subsystems, enabling enhanced grid status monitoring and better alignment with customer requirements. However, alongside these technological advancements, a pressing concern has emerged—smart grids have expanded the attack surface and introduced new cyber vulnerabilities. These vulnerabilities have created opportunities for adversaries to exploit and unleash devastating cyberattacks targeting smart grids.

As we delve into the domain of V2M systems, which harness the potential of transactive energy and employ advanced communication technologies, the need to comprehensively understand and address these cyber vulnerabilities becomes paramount. Adversaries can manipulate these vulnerabilities to compromise the integrity and functionality of V2M systems. Consequently, there is a strong motivation to investigate these emerging security challenges within the context of V2M systems. In addition, it is essential to design compact AI solutions that can work in a computationally constrained edge environment. We also identified some gaps in the literature. For instance, there is a lack of understanding on how to effectively map the increasing complexities in the attack vectors that come with the integration of EVs, smart microgrids, and AI systems in V2M infrastructures. Furthermore, the current research has not sufficiently proposed and evaluated preventive and detective measures tailored to the unique cybersecurity challenges faced by V2M services against adversarial machine learning attacks. In addition, there is a lack of research work on studying the impact of Mobile Edge Computing (MEC) computational resources on the performance of AI-based detection methods under V2M settings.

1.2 Objectives

By studying the vulnerabilities and potential attack vectors introduced by these technological advancements, this research aims to develop proactive defensive mechanisms to fortify the security of V2M systems. In addition, this research strives to enhance the resilience and robustness of V2M systems against cyber attacks. The ultimate objective is to ensure the uninterrupted provision of V2M services and safeguard the stability and reliability of the power grid against adversarial disruptions. Through this thesis, we address the following research questions and objectives:

1. What are the potential adversarial threats and vulnerabilities in V2M systems, and what are the impacts of data integrity attacks and AML threats on the reliability and security of these systems if an adversary has limited knowledge of the V2M system? The objective is to conduct a comprehensive analysis of data integrity attacks and AML threats specific to V2M systems. This allows us to understand the implications

of these attacks on system's reliability and security, leading to the development of effective countermeasures. We also take into consideration the adversary's knowledge as a factor in our study and analysis.

2. What are the effective defense mechanisms that can be developed to detect and prevent adversarial attacks in V2M systems?
   The objective is to design and implement novel defense mechanisms that leverage AI-based detectors to detect and prevent adversarial attacks. The aim is to enhance the security and resilience of V2M services by mitigating the impact of data integrity attacks and minimizing the success rates of AML threats.

3. How can we address the trade-off between the detection model's complexity (i.e., MEC resources) and the detection rate performance of the detection model?
   We focus on investigating how the optimization of MEC and edge resources affects the performance of Deep Learning (DL) models used in the proposed defense framework. The objective is to evaluate the detection performance under different compression methods and different evasion attacks.

1.3 Contributions

Our contributions in this thesis can be summarized as follows:

- Defense Scheme for Data Integrity Attacks: We have identified the potential attack vectors of data integrity attacks; and developed a novel defense scheme that employs an unsupervised ML technique to mitigate the impact of those attacks. Through extensive simulations, we have demonstrated a significant reduction in the impact of these data integrity attacks.

- Effective Countermeasures for Adversarial ML Attacks: Our research has identified potential vulnerabilities in V2M services targeted by Adversarial ML attacks. We have conducted an anticipatory analysis of a multi-stage gray-box attack and white-box attack. Through our anticipatory study, we showed the severe impacts of Adversarial ML attacks against V2M services.

- Comprehensive Defense Framework: To fortify the security of AI-based microgrid control systems, we have proposed a defense framework that integrates a GAN model and a robust ML classifier. This framework enables the ML classifier to learn and adapt to novel attack patterns by generating realistic adversarial samples. Through

simulations, we have demonstrated the effectiveness of this defense mechanism result in higher detection rate of adversarial attacks.

- Optimized edge environment resources: To address the limited computational power and memory in V2M edge settings, we studied different model optimization (i.e., model design and compression) to optimize the model's size without compromising the detection performance. Several AI model compression techniques, such as projection, pruning, and quantization, were discussed. Our proposed method integrates model design and compression, resulting in an optimized detection model that remains robust against adversarial attacks.

1.4 Outline of the thesis

The rest of the thesis is organized as follows:

In Chapter 2, we study the literature works of the following: In Section 2.1 we lay the foundation with an in-depth exploration of adversarial machine learning attacks, including various techniques and their applications. More specifically, we cover inference attacks, evasion attacks, adversaries' knowledge, and adversarial ML attacks against V2M services in Sub-Sections 2.1.1 to 2.1.2. In Section 2.2, we review the communication technologies used in smart grids; Sub-Section 2.2.1 focuses on the wired communication technologies; Sub-Section 2.2.2 discusses the role of wireless communications in power systems management by breaking it into three groups (based on the coverage distance): WPAN, WLAN and WWAN.

In Chapter 3, we address adversarial attacks from the vehicle's side as part of the V2M system. We try to understand the feasibility of launching such attacks in a V2M context where the considered attack surface is the vehicle. We focus on data integrity attacks as a case study for adversarial attacks within V2M systems. We begin with an overview of the problem in section 3.1, followed by a detailed system model for V2M services in Sub-Sections 3.2.1 and 3.2.2. This chapter also includes modeling of data integrity attacks against V2M applications, exploring data integrity attacks and anomaly detection in Sub-Sections 3.3.1to 3.3.2, and concluding with a performance evaluation and summary in sections 3.4 and 3.5, respectively.

In Chapter 4, as we realized that it is significantly challenging for adversaries to launch adversarial attacks against vehicles, we shift the focus of the attack surface to the microgrid aspect of V2M. We introduce the threat model in 4.2, which will be used in the rest of the dissertation. We anticipate adversarial ML-based attacks against V2M services,

focusing on inference and evasion attacks and Conditional Generative Adversarial Networks in sections 4.3. A performance evaluation follows, including experiment setup, evaluation metrics, and numerical results in Sub-Sections 4.4.1 to 4.4.3, and the chapter concludes with a summary in section 4.5.

In Chapter 5, we address the detection of adversarial machine learning attacks against V2M services. We introduce the topic in section 5.1, discuss the methodology, including AML-based attacks against V2M services and a GAN-based detection technique in Sub-Sections 5.2.1 and 5.2.2, followed by performance evaluation with experiment setup and results in Sub-Sections 5.3.1 and 5.3.2, and end with a summary of the chapter in section 5.4.

In Chapter 6, we begin by establishing the significance and necessity of AI model compression in a V2M environment in section 6.1, emphasizing the constraints of edge devices on computational and memory resources. In section 6.2, we provide an overview of various AI model compression techniques, leading up to our proposed method in Sub-Section 6.2.1. We then evaluate the performance of our method against different types of evasion attacks in Sub-Sections 6.3. Finally, we conclude the chapter with final remarks and key findings in section 6.4.

Finally, Chapter 7 concludes the thesis with final remarks, open issues and future work.

Chapter 2

Background and Literature study

In this chapter, we will explore adversarial machine learning, beginning with the general pillars of threat models and focusing on the relevant attack models in a machine learning context. We will then dive into the background of adversarial attacks against V2M services. Subsequently, we will discuss the role of communication technologies and edge computing in smart microgrids. The subsection will conclude with a discussion on the potential vulnerabilities associated with integrating these technologies into smart microgrids. Finally, we will end the chapter with an overview of machine learning model compression.

2.1 Adversarial Machine Learning Attacks

Adversarial Machine Learning (AML) is a subcategory of adversarial attacks that focuses on targeting ML models' vulnerabilities [29, 30]. The life-cycle of a ML system typically involves three phases: training, deployment, and inference [31]. Based on these stages, AML attacks can be classified into three attack paradigms that occur during each stage: training-time attacks [32], deployment-time attacks [33] and inference-time attacks [34]. In this work, we focus on AML attacks against smart grids that occur during inference time, specifically inference and evasion attacks. Additionally, we will provide an overview of other prevalent adversarial attacks. Before examining AML attacks against smart grids, it is important to provide an overview of the main pillars of the threat model, including the adversary's resources, access, goals and strategy.

2.1.1 Attack Taxonomy

Adversary's Resources

The adversary's resources reflect the capabilities needed to launch a successful attack. This includes knowledge of the system under attack, the tools and equipment necessary for the attack, and the time required to penetrate the system. In the context of AML attacks, these resources can be redefined with specific examples. For instance, adversary's equipment refers to the computational power needed to execute the attack strategy. Additionally, adversary's knowledge in AML settings pertains to how much the adversary knows about the victim's machine learning model. Consequently, this knowledge can be categorized into three types of attacks: white-box, black-box and gray-box attacks [35]. In a white-box attack, the victim's ML model is known to the adversary and can be easily replicated using the model's hyper-parameters and the training dataset. This allows the adversary to create an exact copy of the victim's ML model, which they can use to infer its statistics and launch other types of attacks [36]. White-box attacks are considered to be the most powerful type of inference attack because the adversary has complete knowledge of the victim's ML model [37].

In a black-box attack, the adversary cannot access the victim's ML model's hyper-parameters or the training dataset. This makes it more difficult for them to create a surrogate model, but it is not impossible. For example, the adversary might use techniques such as model inversion or membership inference attacks to infer information about the victim's ML model [38]. Black-box attacks are considered to be less powerful than white-box attacks because the adversary has limited knowledge of the victim's ML model [39].

In a gray-box attack, the adversary has access to some of the victim's ML model's hyper-parameters and the training dataset [40]. This allows the adversary to create a surrogate model, but the performance of the surrogate model will depend on the quality and quantity of the collected observations. For example, the adversary might query the victim's ML model to build a dataset, but they may not be able to capture enough observations to train a high-quality surrogate model. Gray-box attacks are considered to be intermediate in terms of their power, as the adversary has partial knowledge of the victim's ML model.

Adversary's Access

Adversary's access defines the type of entry point to the network. For an adversary to mount an attack, either physical or cyber access to the network is required. Physical access can be direct, such as through a malicious insider with privileges in the network, or indirect, involving an outsider attempting to gain privileges [41]. It is more common for attacks to be initiated by outsiders who escalate their privileges within the network, especially in V2M services as we will discuss in the communication vulnerabilities part of

this chapter. Cyber-access, on the other hand, allows the adversary to connect remotely to the network through insecure communication channels.

Adversary's Goals

The attacker's goal typically focuses on compromising three key aspects: confidentiality, integrity, and availability, collectively known as the 'CIA triad'. A confidentiality breach occurs when sensitive information from applications or ML models is exposed, potentially leading to user privacy violations and unauthorized commercial exploitation through surrogate model development [42]. Various types of AML attacks fall under this category, including model extraction (where the model's parameters are extracted), membership inference attacks (which determine if a particular sample was used in training), and model inversion (which deduces information about the input by analyzing the output). Integrity violations arise when ML models are deceived into generating incorrect results for both malicious and benign inputs. This could happen in either training-time or inference-time. This often involves altering input data, such as adding perturbations, to manipulate output classifications (e.g., changing a predicted label from malicious to benign) or values (e.g., altering a sensor reading to a specific target) [43]. Evasion attack is one example of the AML attacks that target ML model's integrity. Availability breaches occur when the ML model is unable to perform its intended services, either due to functional impairments that prevent it from processing legitimate inputs, or because its accuracy has degraded below acceptable levels, resulting in excessive false positives and negatives [44].

Adversary's Strategy

Adversary strategies can be categorized into two main types: gradient-based and non-gradient-based. Gradient-based adversarial example generation methods utilize the gradient information of the victim model to craft adversarial examples, with this gradient data playing a crucial role in the computations for generating adversarial examples. Notable attack algorithms identified in the literature include the Fast Gradient Sign Method (FGSM) [45], Basic Iterative Method (BIM) [46], Projected Gradient Descent (PGD) [47], Jacobian-based Saliency Map Attack (JSMA) [48], and Carlini & Wagner (C&W) [49]. The non-gradient-based approach involves identifying the most effective features and perturbations for creating adversarial examples using a variety of techniques. This approach can be further split into two groups: score-based and decision-based attacks. Zeroth Order Optimization (ZOO) [50], Particle Swarm Optimization [51] and Genetic algorithm [52] are some of the most common techniques under the score-base method. In addition, decision-based method includes include Reinforcement Learning (RL) [53] and brute-force [54].

Attack Types

AML attacks can occur at different stages of the ML model life-cycle. Inference attacks and evasion attacks typically take place during the inference-time stage, while poisoning attacks occur during the training-time stage. Inference and evasion attacks can occur either together or separately, depending on the adversary's knowledge of the victim's ML model. For example, in a white-box attack scenario, the adversary, having full knowledge of the model, can launch an evasion attack directly without needing an inference attack. However, in a black-box scenario, where the adversary lacks direct knowledge of the model and the training dataset, an inference attack is necessary to gather sufficient information before launching an evasion attack.

Inference attacks are a type of AML attack that aim to infer information about the victim's machine learning model [55]. The goal of these attacks is to create a surrogate or "shadow" model of the victim's ML algorithm that replicates its statistical distributions and functionalities. This surrogate model can then be used by the adversary to infer the statistics of the victim's ML model. Once the adversary has this information, they can use it to launch other types of attacks. For example, they might use the surrogate model to launch an integrity attack, which aims to misclassify inputs or reduce the victim's model confidence. This can have significant impacts on the performance of the victim's ML model and the systems that rely on it. The adversary's goal of this attack type is to violate the confidentiality and privacy of the ML model.

Moreover, evasion attacks are designed to deceive a machine learning model by feeding carefully perturbed samples, known as adversarial instances [56]. The goal of these attacks is to significantly reduce the integrity of the victim's ML model by causing it to make wrong decisions [57]. To increase the success rate of the attacks, the adversary first studies the victim's ML model by performing an inference attack. This allows them to infer information about the victim's ML model, such as its statistical distributions and functionalities [58]. With this information, the adversary can design adversarial instances that are more likely to be misclassified by the victim's ML model. The adversary chooses the adversarial instances based on the output labels' distances to the decision boundaries. Output labels that are adjacent to the decision boundaries tend to increase the likelihood of mis-classification in the victim's ML model. The adversary's goal of this attack type is to violate the integrity of the ML model.

On the other hand, attacks targeting the training phase aim to compromise the model's training dataset. For example, data poisoning attacks occur during the training phase and involve either the addition of new data or the modification of existing training data. A specific variant, known as backdoor attacks, occurs when hidden triggers are embedded in the model during the training phase. These triggers are subsequently activated during the inference phase through specially crafted inputs, causing the model to generate incorrect outputs. Depending on the attacker's objective, this type of attack can disrupt either the

availability or the integrity of the ML model.

2.1.2 Adversarial ML Attacks Against V2M Services

In the context of V2M services, Adversarial ML attacks can have significant impacts. For example, an adversary might pursue an inference attack to manipulate the output of a machine learning model that is used to predict the demand for electricity in a microgrid. This could cause the microgrid to operate inefficiently and potentially lead to power outages. Similarly, an evasion attack could cause a machine learning model used for scheduling vehicle charging to make incorrect decisions, leading to poor performance and reduced reliability of the V2M system.

There are research efforts to investigate Adversarial ML attacks against smart grids. For instance, [59] [60] [61] discussed the impact of Adversarial ML attacks on smart grids without proposing defence strategies to prevent the attacks. In addition, the authors in [62] studied the detection of adversarial attacks on smart grids by using adversarial training. However, in a recent study [63], the authors reported that adversarial training failed to detect novel adversarial attacks. In Table. 2.1, we provide an overview of the main differences between our work and the other recent studies in AML attacks against smart grids. Our work focuses on V2M services, a crucial aspect of smart grid operations, which has not been the primary focus in the mentioned references nor in the literature. Unlike the other studies that either assume the adversary's knowledge of the victim's ML model or focus solely on full-knowledge scenarios of the training dataset, our work encompasses no-knowledge, partial-knowledge and full-knowledge scenarios (i.e., black-box, gray-box and white-box attacks), offering a more realistic and challenging perspective. Moreover, we propose a detection method using GAN models to safeguard the victim's model against evasion attacks, a strategy not explored in [59] [60] [61]. Moreover, we study both inference and evasion attacks unlike the other studies that focus on one attack type only. Furthermore, we tailor our detection model to work in the constrained nature of edge environments where computational resources and memory are limited. Hence, we designed a model compression pipeline to produce an optimal lightweight detection model against adversarial attacks in V2M settings. Additionally, we provide a detailed threat model(4.1) (introduced in chapter 4), ensuring that our defense mechanism stems from an accurate understanding of potential attack vectors. Furthermore, most of the works in the field of AML attack consider full knowledge of the victim's training dataset (i.e., white-box attack), however, we consider all different levels of adversary's resources. That is, white-box, gray-box and black-box attacks, which correspond to, full knowledge, partial knowledge and no knowledge of the victim's dataset, respectively.

It is worth noting that developing a unified defense strategy against Adversarial ML attacks within smart grid applications faces several substantial challenges. Firstly, the complexity and diversity of smart grid services cover a broad spectrum of functionalities, each requiring different data types, and operational methods [64]. This diversity makes a one-size-fits-all defense mechanism impractical, as techniques effective in one application may not transfer well to others [65]. For instance, employing GANs for adversarial ML attack detection in V2M services might not suit other smart grid areas due to differing operational characteristics, threat models and data types. Secondly, the efficacy of ML models for attack detection depends crucially on the data's quality and specificity, which varies greatly across smart grid applications [66]. The data involved in V2M services (e.g., power consumption and generation patterns) significantly differs from those in applications such as power quality monitoring (e.g., phasor measurements, current and voltage signals). Thirdly, adversaries' knowledge and capabilities can vary extensively, from having limited system data access (black-box scenarios) to possessing in-depth system knowledge (white-box scenarios). This makes the design of a defense method that can address the entire potential adversarial spectrum significantly challenging. Lastly, the evolving attack surface, with adversaries constantly developing new attack techniques and exploiting novel vulnerabilities [67]. This means that defense methods must continuously adapt and improve, further complicating the development of a generalized defense approach. These factors highlight the need for defense mechanisms to be specifically tailored to unique requirements and threat models of each smart grid. Other existing research works studied Adversarial ML attacks in wireless communication settings. For instance, [77] and [78] studied the impact of evasion attacks on modulation recognition. In addition, inference attacks have been studied in the context of spectrum sensing [79] [80] and network traffic [81]. Other Adversarial ML attacks against wireless communication settings include the data fusion process in IoT [82] and signal authentication [83]. In addition to the wireless communication field, Adversarial ML attacks have been studied in other domains such as computer vision and NLP [84]. However, the proposed solutions in those domains are unable to address the special challenges presented by the wireless communication of V2M systems [82].

2.2 Enabling Communication Infrastructures and Technologies for Smart Microgrids

Smart grids employ different communication technologies to provide various services. Based on the coverage area, communication technologies in smart grids can be grouped into three main categories: (i) Wide Area Network (WAN), (ii) Local Area Network (LAN), and

Table 2.1: Summary of common AML attacks in smart grids

Attack type	**Attack method**	**Task/Service**	**ML problem**	**Metric**	**Ref.**
Evasion	TFGSM	FDIA detection	Classification	Accuracy	[68]
Poisoning	Simulated Annealing	Load forecasting	Regression	MAPE	[69]
Poisoning	-	Electricity theft detection	Classification	F1-score	[70]
Evasion	FGSM	Power quality recognition	Classification	MAPE	[71]
Causative	BIM/MIM	FDIA detection	Classification	ASR	[61]
Inference	CGNA/FGSM	V2M services	Classification	ADR	[72]
Evasion	DeepFool/FGSM	FDIA detection	Classification	False Negative	[59]
Poisoning	-	AC state estimation	Regression	MAE	[73]
Poisoning	FGSM	Event diagnosis	Classification	Accuracy	[74]
Evasion	PGA	Load monitoring	Classification	ASR	[75]
Evasion	FGSM/BIM/C&W	Electricity theft	Classification	Accuracy/F1-score	[76]
Inference and evasion	CGAN-based	V2M	Classification	ADR and EIR	This work

(iii) Personal Area Network (PAN). WAN is the largest computer network that covers a big geographical area. It is a private network used to interconnect multiple distributive LANs [85]. There are many examples of smart grids communication technologies in WAN such as WiMAX technology, cellular networks, and satellite communication. A LAN, on the other hand, connects different devices together within a small geographical area such as school, office building and residence. The most common two examples for LAN in smart grids communication are: Ethernet and WiFi. In general, routers and gateways are used to link a LAN with a WAN. A smaller area can be covered by a PAN which is used mainly for low power and short-distance networks. ZigBee, Bluetooth, and Z-Wave are examples of smart grids communication technologies in PAN.

Smart microgrid applications require a communication component to provide their services as desired. It is also an essential criterion when assuring the QoS, especially for

mission critical applications such safe and security applications [86]. Smart grids can exchange information between its entities using wire or wireless communication systems. In wire-based systems, the data can be transmitted using various methods such as telephone networks, fiber-optic networks and power line communication (PLC) [87].

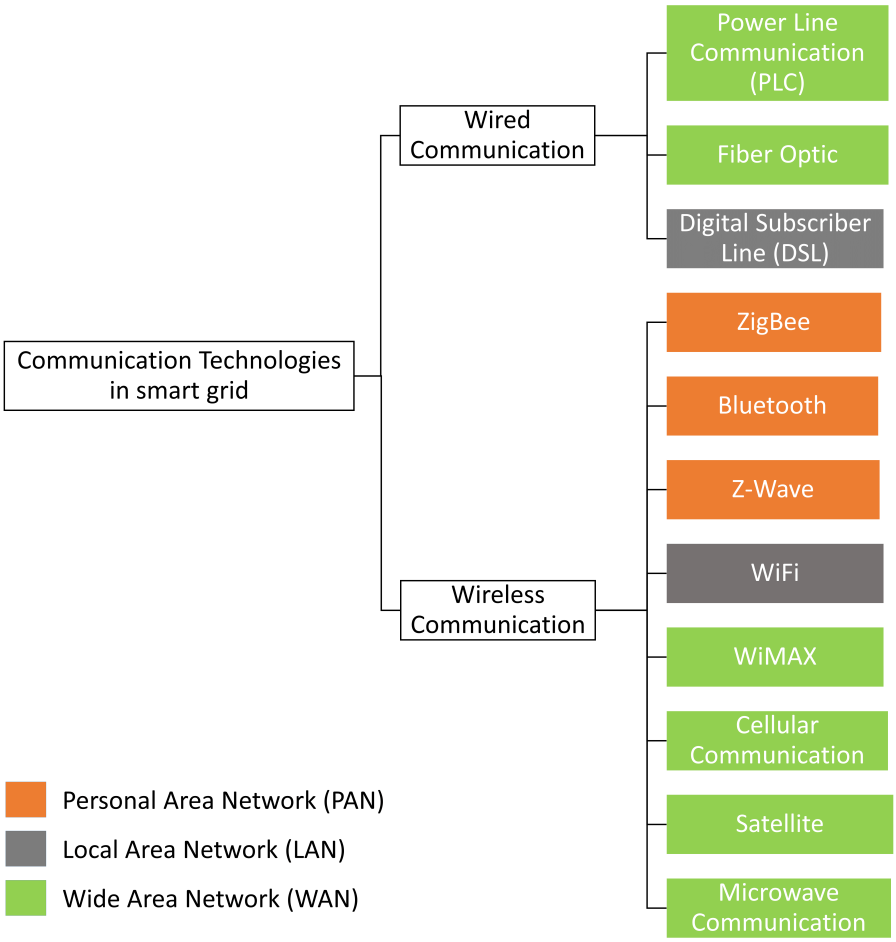


Figure 2.1: Communication classification in smart grids

Table 2.2: Comparison of different technologies used in smart grids communication

Technology	**Spectrum**	**Data Rate**	**Transmission range**	**Latency**
BB-PLC and NB-PLC	1-30MHz	1-200Mbps	Up to 3Km	5-7ms
ZigBee (IEEE 802.15.4)	868-915 MHz, 2.4GHz	0.25Mbps	Up to 100m	15ms
6LoWPAN (IETF REC 4944)	2.4GHz	0.25Mbps	Up to 100m	-
NB-IoT (3GPP in LTE)	2.4GHz	Up to 0.25Mbps	Up to 35Km	<10s
Bluetooth (IEEE 802.15.1)	2.4GHz	Up to 1Mbps	Up to 100m	<40ms
Microwave communication	2-40GHz	155Mbps	60Km	<2ms
Wi-MAX (IEEE 802.16d/e/j/m)	2.5, 3.5, 5.8GHz	Up to 75Mbps	Up to 1Km	10-50ms
Wi-Fi (IEEE 802.11b/g/n)	2.4, 5GHz	2-600Mbps	Up to 1Km	3.2-17ms

On the other hand, in wireless communication systems, the data is transferred wirelessly using different technologies such as Wireless Personal Area Network (WPAN), Wireless Local Area Network (WLAN), and Wireless Wide Area Network (WWAN). In this section, we will explain the wired and wireless communication technologies used in the smart grid. The communication technologies used in smart grid communication are shown in Figure 2.1 [88] and Table. 2.2 [89] [90]. Table. 2.3 summarizes the research on adversarial machine learning attacks in wireless communications. The examined works studied evasion impacts on modulation recognition, inference attacks on spectrum sensing

and network traffic, and adversarial threats to IoT data fusion and signal authentication.

Attack type	**Attack method**	**Task/Service**	**ML problem**	**Ref.**
Evasion	FGM	Modulation	Classification	[91]
Exploratory	LEB attack	Cooperative spectrum sensing	Classification	[92]
Exploratory	Membership inference attack (MIA)	Wireless signal	Classification	[93]
Evasion	Adversarial Mutation Network (AMN)	Modulation	Classification	[94]
Exploratory	FGSM	Spectrum sensing	Classification	[95]
Evasion	BIM / MIM	Modulation	Classification	[96]
Evasion	Reinforcement learning-based attack	Signal authentication	Classification	[97]
Evasion	FGSM	Modulation	Classification	[98]

Table 2.3: Summary of attack types and methods in wireless communication

2.2.1 Wired Communication Technologies

Service suppliers in smart grids prefer wired communication technologies to exchange data using wired networks as the name implies [99] [100]. The wired communication has many advantages, mainly connection reliability and security. Fiber optic and digital subscriber line (DSL) are two examples of the wired communication used in the smart grid. However, the PLC technology is the most wired communication that is widely used in the smart grid [101]. DSL and fiber optic can support data transmission up to 10 Gbps and 150 Gbps, respectively.

Power Line Communication

Electromagnetic environments (e.g. transformers) cause interference and disruptive effects to the PLC, resulting in abrupt propagation behaviour [102]. To overcome that technical

challenge, the PLC can operate in two different bandwidths, namely BroadBand PLC (BB-PLC) and NarrowBand PLC (NB-PLC) [103]. The BB-PLC has a maximum data rate of 200 Mbps (on very short distance); it operates at frequency between 2 MHz and 30 MHz, covering a range up to 1.5 km. On the other hand, the NB-PLC is used for low data rate applications, such as smart metering, and it operates at 500 MHz frequency. It can reach a coverage area up to 150 km with two voltage line modes: low voltage line and high voltage line [104]. In addition to the disruptive effects, the PLC suffers from low data rate transmission for long distances as the signals experience more channel interference and higher losses [105]. Another drawback is bandwidth limitations that makes PLC technique inadequate for high bandwidth applications [106]. On the other hand, the PLC technology has low operation and maintenance cost [100].

Fiber Optic Communication

Another wired communication system used in smart grid is fiber optic. It provides higher data rate comparing to the PLC. In addition, fiber optic technology can cover several kilometers area with a data rate in range of gigabits [107]. There are various types of fiber optic standards such as AON (IEEE 802.3ah), BPON (ITU-T G.983), and EPON (IEEE 802.3ah) [108]. The Active Optical Network (AON) can transmit data up to 100 Mbps on both up and down streams with a maximum distance of 100 km. Furthermore, the two standards of the Passive Optical Networks (PON) have different data rate and distance range. For instance, the Broadband PON (BPON) can achieve a data rate of 622 Mbps and covering an area range up to 60 km, whereas the Ethernet PON (EPON) can reach a maximum data rate of 1 Gbps with a distance of 20 km [109].
Fiber optic brings many advantages to the smart grid applications and services. For example, it can deliver data in diverse transmission rates from different distances ranges while preserving the signals strength and robustness. However, the fiber optic suffers from high implementation cost [110].

Digital Subscriber Line

DSL is a digital data transmission technology that uses telephone lines to exchange information over the smart grid. Hence, it can be cheaper to implement [111]. However, it suffers from low data rate and signal interference [100]. Table. 2.4 summarizes some of the common used standards in the smart grid.

Table 2.4: Comparison of different DSL standards used in smart grids

Name	**Standard**	**Data rate in (Mbps)**	**Distance in (km)**
Asymmetric DSL (ADSL)	ITU G.992.1	8 downstream / 1.3 upstream	up to 5
ADSL2	ITU G.992.2	12 downstream / 3.5 upstream	up to 7
ADSL2+	ITU G.992.5	24 downstream / 3.3 upstream	up to 7
Very high speed DSL (VDSL)	ITU G.993.1	[52 - 85] downstream / [16 - 85] upstream	up to 1.2
VDSL2	ITU G.993.2	200 downstream/upstream	[0.3 - 1.5]

2.2.2 Wireless Communication Technologies

Wireless communication technologies can be categorized based on their transmission ranges as mentioned previously. In the following, we will explain the most important wireless communication technologies which are suitable to smart microgrids, ordered by the coverage range from the smallest to the largest.

Wireless Personal Area Network

The WPAN is based on IEEE 802.15 family and it includes several standards such as ZigBee (IEEE 802.15.4) and Bluetooth (IEEE 802.15.1). The ZigBee technology is used for low data rate transmission in smart grids, specifically 256 kbps [112]. It is considered as a power efficient technology with a low range communication coverage [113]. The ZigBee technology can cover an area up to 100 m and it operates at frequency band of 2.4 GHz, except for Europe and the USA, where ZigBee operates at 864 MHz and 915 MHz, respectively [114]. The drawbacks of the ZigBee technology, such as slow processing, limited memory storage,

and short battery life, makes it more suitable for home automation systems [115]; its various applications include smart meter reading, load control, and security systems [107].

Bluetooth is also a power efficient communication technology for short range coverage. It can support a data rate up to 721 kbps within a 100 m range. Similar to the 802.15.4 technology, the IEEE 802.15.1 standard (Bluetooth) operates in the 2.4 GHz frequency band, which makes it more susceptible to interference from other communication technologies that operate at the same frequency band, such as WiFi [116]. One of the Bluetooth application areas in a smart grid is online monitoring of local substations [117].

Another technology used in the WPAN is Z-wave communication which was developed by Z-wave Alliance. Since it operates on a lower frequency band (900 MHZ), the Z-wave technology can penetrate obstacles while preserving its signal strength [114]. Although it has a low data rate transmission (100 kbps), the Z-wave is the leading technology in home automation applications, compared to Bluetooth and ZigBee technologies. The reason is the low cost and the simple implementation of the Z-wave technology [118].

Wireless Local Area Network

WLAN is based on the IEEE 802.11 family that enables mobile devices to connect wirelessly. The market name of the 802.11 standard is known as WiFi. Although the 802.11 standard can operate on a wide range of frequencies, the 2.4 GHz and the 5.0 GHz frequency bands are the most popular operating frequencies that are used in the smart grid [119]. There are different versions of the 802.11 standard that vary in terms of two main components: the maximum data rate, the coverage area, and the bandwidth [120]. In the following, we discuss some of the commonly used standards under the IEEE 802.11 family [121] [122] [123]. Some of these standards are used in smart grids communication according to [100], [87], [124], [125], and [126].

• 802.11a: this standard was released in 1999 and it operates at a frequency of 5 GHz with a maximum data rate transmission of 54 Mbps. It uses the orthogonal frequency division multiplexing (OFDM) as a modulation scheme, and it can operate indoors and outdoors with range up to 35 m and 120 m, respectively. The main disadvantage of this standard is the small coverage area which makes it more suitable for indoor areas. However, the presence of indoor obstacles may cause difficulties in the data transmission.

• 802.11b: similar to the 802.11a standard, the 802.11b was ratified in 1999. It operates at frequency of 2.4 GHz with maximum data rate of 11 Mbps. It can cover outdoor areas up to 140 m and indoor range of 38 m. The 802.11b standard uses a modulation technique known as Direct Sequence Spread Spectrum (DSSS). As many of the IEEE 802.11

standards operate at 2.4GHz frequency, the 802.11b standard is more likely to experience high interference. However, the low operating frequency compared to the 802.11a makes it signal more immutable to the obstacles.

• 802.11g: this standard was released in 2003. Similar to the 802.11b standard, it operates at a frequency of 2.4 GHz. However, the data transmission rate is greater than the 802.11b as it can reach a data rate up to 54 Mbps. It can use two modulation schemes, either the OFDM modulation or the DSSS modulation, and it can cover a similar range as in the 802.11b (i.e. 38 and 140 for indoor and outdoor respectively). This standard outperforms the 802.11a and 802.11b since it combines the advantages of the maximum data rate and the obstacle penetration.

• 802.11n: this standard was released in 2009 and it operates at two different frequency bands, 2.4 GHz and 5 GHz. It has a higher maximum data rate of 600 Mbps, when compared to the 802.11a/b/g standards. It uses the OFDM modulation scheme, and Multi-Input, Multi-Output (MIMO) antenna. It reaches a wide range up to 70 m and 250 m for indoor and outdoor areas, respectively. This standard has the highest coverage range among the others mentioned standard.

• 802.11ac: the 802.11ac only uses one frequency band of 5 GHz, and it was released in 2013. It can achieve a data rate up to 6900 Mbps, but it can operate indoor only with a range up to 35 m. It uses OFDM modulation with multi-user MIMO (MU-MIMO) technique in the down-link [127].

• 802.11ax: this standard was released in September 2019 and it operates on the both frequency bands, 2.4 GHz and 5.0 GHz. It can reach a transmission data rate up to 10000 Mbps, which makes it the highest achievable data rate when compared to the 802.11a/b/g/n/ac standards. Contrary to the 802.11ac, the 802.11ax standard can cover indoor and outdoor areas with ranges up to 35 m and 120 m, respectively. It uses the orthogonal frequency division multiple access (OFDMA) modulation with multi-user MIMO (MU-MIMO) technique in both up-link and down-link [128]. This standard is a promising solution for IoT and smart grid applications [129].

The WiFi protocol standards (802.11a/b/g/n/ac/ax) are used to enable a two-way communication for the smart grid applications and services. However, since wireless communication suffers from lack of security, many changes and improvements were adopted in the smart grid. For instance, the IEEE 802.11i standard was ratified in 2004 to provide security improvements [130]. In addition, IEEE 802.11w and IEEE 802.11e provided protected frame management [131], and enhanced QoS (e.g, traffic prioritization, scheduling and admission control [132]) for wireless LAN applications, respectively. The 802.11 standards discussed previously are summarized in Table 2.5 [100] [87] [124] [125] [126].

Table 2.5: Comparison of different versions of IEEE 802.11 standard used in smart grids communication

IEEE standard	**Frequency in (GHz)**	**Data Rate in (Mbps)**	**Outdoor Transmission range (m)**
802.11a	5.0	54	120
802.11b	2.4	11	140
802.11g	2.4	54	140
802.11n	2.4/5.0	600	250
802.11ac	5.0	6900	NA

Wireless Wide Area Network

WWAN includes various technologies to provide wireless communication to smart microgrids. It covers the largest geographical area among our list. In the following, we will discuss four examples of WWAN technologies, namely WiMAX (IEEE 802.16), cellular communication, satellite communication, and microwave communication.

• WiMAX: IEEE 802.16 standard was first released in 2001 under a commercial name known as Worldwide Interoperability for Microwave Access (WiMAX). The 802.16 standard is considered a good solution for wireless networks as it offers low cost implementation, a high data rate, and long distance coverage [133]. Typically, the 802.16 standard can cover an area range up to 10 km with a maximum data rate of 128 Mbps and 28 Mbps for down-link and up-link, respectively [134]. The smart grid distribution domain is one of the application areas that suits the WiMAX technology when compared to the GSM solution [135].

To improve the 802.16 standard, different versions were released over the past decade. For instance, the 802.16j standard brings many features such as supporting OFDMA modulation along with other types of coding and modulation schemes, better handover techniques, and multi-cast and broad-cast services [136]. Moreover, 802.16j defines different

types to mutli-hop relaying techniques [137]. Another version of the WiMAX is 802.16m, which can support high mobility speed (350 km/h) while providing at least a data rate of 100 Mbps and 1 Gbps for lower speeds (i.e. fixed users) [138]. It can provide a coverage area of [30 - 100] km with reduced performance. An acceptable performance can be reached with a distance between 5 km and 30 km, whereas the optimum performance can be archived within 5 km range.

• Cellular communication: includes a wide range of network technologies such as 2G, 2.5G, 2.75G, 3G, 4G (LTE and LTE-A), and 5G; however, smart grids use 3G and beyond for wireless communication. The Universal Mobile Telecommunications System (UMTS), which is the most commonly known standard of 3G, provides a maximum data rate of 168 Mbps and 22 Mbps, offered by the Evolved High-Speed Packet Access standard (HSPA+), in the down-link and the up-link, respectively [139]. The 3rd Generation Partnership Project (3GPP) developed another cellular technology know as Long Term Evolution Advanced (LTE-A) standard which is an enhanced version of the LTE. The implementation of cellular technologies in a smart grid is expensive and may suffer from network congestion (due to the spectrum sharing [140]); however, the LTE-A brings various advantages to cellular technology such as flexible bandwidth, better handover between the cells, and more support for heterogeneous Network (HetNets) [141].

Heterogeneous Networks are composed of multiple tiers of wireless network cells. The main advantage of HetNets is to support wireless networks with coverage in different environments (e.g., offices, tunnels, and city centers). Macro-cells, pico-cells, and femto-cells are some examples of HetNets. From an architectural point of view, HetNets are built on top of the functionalities offered by conventional macro Radio Access Network (RAN), RAN transport capability, small cells, and WiFi.

The role of (HetNets) in microgrid control and management has been considered in the last few years by several researchers. In [142], the authors presented the first work that introduces multi-agent coordination via a HetNet infrastructure in order to address the trade-off between two cost components: communication and power generation. In [143], the authors tackle the automation problem in microgrids; they propose a heterogeneous and converged fiber-wireless network infrastructure as the communication medium for the addition and/or deployment of renewables as well as storage systems into the power grid. To address the interoperability issues, the proposed system utilizes the IEC 61850 standard for the power grid end; whereas, off-the-shelf automation protocols, such as PROFINET and Modbus TCP, are utilized for the interoperability within building automation.

The authors in [144], proposed a framework to process smart microgrid data in HetNet using unsupervised machine learning algorithm. A multi-class queuing system in the pico-cell tier was introduced to prioritize the data based on time-sensitivity. A real-time dataset

was collected from 443 microgrids for one day. The authors used the k-means algorithm to cluster the data into three categories: delay-tolerant, delay-medium, and delay-sensitive. The results showed that the proposed approach significantly reduces the delivery delay of messages carrying time sensitive events from the microgrid. In addition, the results demonstrated that the framework is able to process the time sensitive events faster and more accurately in comparison to a single-tier network infrastructure where the messages are severed on First In, First Out (FIFO) basis.

The authors extended their work to include different HetNet tiers such as pico-cell,micro-cell, and macro-cell [86]. The work tested the proposed technique under different number of users with two communication scenarios. TCP and UDP protocols were employed with different time sensitive data to improve the delay while keeping the packet loss rate at low level. Through simulations, the authors showed that the proposed technique reduced the queuing delay by 93% for the packets of delay-sensitive (urgent) messages and the packet loss rate by 7% when compared to the benchmark where no aggregation mechanism exists prior to the small cell base stations.

The massive number of wireless devices produces large amounts of data that need to be processed and transmitted over the network [145]. That said, communication networks need to provide higher QoS performance in terms of communication delay, reliability, and security. To this end, 5G cellular networks are introduced to meet the mentioned QoS requirements and as well as other requirements. The features implemented in 5G include massive MIMO, optimal utilization of network resources, and the usage of higher frequency (i.e. millimetre-Wave (mmWave) frequency band), which will increase the maximum data rate, reduce the network delay and provide better throughput performance. For instance, the 5G wireless HetNet challenges in using massive MIMO and mmWave technologies were discussed and analyzed [146]. Furthermore, there are different issues in HetNet that need to be resolved by the network operators, e.g., cells interference, implementation cost, and data flow management between the different cells. Hence, this will help the HetNet technology to grow faster with a reliable and robust services while providing a better performance in terms of delay and throughput as discussed in [147].

• Satellite: it can be expensive and difficult to provide rural areas with communication services through traditional systems. Hence, smart grid applications such as the SCADA system uses satellite communication technology to provide services for remote areas and substations [148]. Satellite communication systems offer various communication performance in terms of latency and bandwidth based on the orbit altitude of the satellite [149]. There are three main orbit altitudes: (i) Low Earth Orbits (LEO) (ii) Medium Earth Orbit (MEO) (iii) Geostationary Earth Orbit (GEO) [150]. It can be relatively cheap to implement a smaller satellite stations in a high altitude orbit [151], however, this might degrade

the communication performance by increasing the delay.

• Microwave: microwave technology is widely used for line of sight communication (i.e., point to point). More than 50% of the base-stations in the world communicate through microwave technology [152]. It can provide a long distance coverage up to 60 km with a maximum data rate of 155 Mbps. The configuration of the line of sight has to be precise and accurate to avoid transmission loss and/or service interruptions. In addition, microwave technology suffers from channel fading due multipath interference and precipitation [153]. Transfer trips between smart grid units is one application of microwave technology [154].

2.2.3 Edge Computing and V2G Security

The implementation of V2G is categorized into two distinct approaches: centralized dispatching and decentralized dispatching, as identified in the mainstreams of V2G [155]. In the centralized approach, a unified strategy is employed where an energy coordinator (EC) optimizes EV charging and discharging in accordance with the grid status and specific EV charging requirements. This method, however, has its drawbacks. It necessitates the uploading of sensitive EV charging data to the EC, raising concerns about data privacy and security [155]. Additionally, as the number of EVs increases, the complexity and computational requirements grow exponentially, posing a significant challenge in scalability [156]. The decentralized method, in contrast, employs an incentive-based strategy where EV charging behaviors are indirectly coordinated through electricity pricing mechanisms. This approach requires understanding the feedback of EV users to electricity pricing, an interdisciplinary challenge that involves complex economic and behavioral considerations. The readiness for electricity market liberalization varies across countries and regions, making the decentralized approach more challenging in certain areas [157].

From a technological perspective, V2G involves both hardware and software aspects. High-performance sensing devices are crucial for ensuring smooth communication between EV users and ECs [158]. V2G systems can be classified as either bidirectional or unidirectional based on the capacity of the charging points (CPs) to support bidirectional power flows [155]. Furthermore, V2G dispatching frameworks can be designed as either single-layer or hierarchical. The hierarchical framework, though more complex in terms of communication networks, offers the advantage of dividing a large problem into smaller, more manageable segments, thereby improving efficiency and reducing solving time [159][160]. A noteworthy example is the double-layer V2G model developed by the authors in [161], which significantly outperforms the single-layer model in terms of computational speed.

On the software front, V2G control flow is divided into day-ahead and real-time dispatching. The day-ahead approach necessitates accounting for the uncertainties in RES

output and charging behavior [162]. Real-time V2G, on the other hand, demands prompt decision-making, irrespective of the EV population size. The increase in the number of scheduled EVs poses a challenge for traditional mathematical optimization methods, often leading to a situation known as the "curse of dimensionality" [156]. While heuristic algorithms provide a more efficient alternative, they carry the risk of resulting in suboptimal solutions [163]. Customized algorithms have been developed for specific scenarios, but their adaptability across different scenarios remains a subject of discussion [162][164]. The introduction of parallel computing techniques is a recent advancement aimed at enhancing the efficiency of V2G dispatching [165].

Demand response is another V2G service that has recently attracted considerable attention, particularly in the context of edge computing. This focus is seen in several innovative approaches designed to integrate demand response mechanisms within edge computing frameworks. The authors in [166] introduced an innovative online auction system that encourages edge networks to engage in Energy Demand Response (EDR). This mechanism is designed to provide incentives to participants, thereby promoting active involvement in EDR. Similarly, the authors in [167] developed a sophisticated online task scheduling algorithm. This algorithm is strategically designed to select specific clusters for workload dispatch, aiming to achieve energy reduction targets efficiently. In another development, the authors presented a reverse auction model that involves local generators. This model is specifically geared towards ensuring the achievement of targeted EDR power reduction.

The authors in [168] have taken a step further by designing an online auction mechanism that simultaneously addresses both power EDR and computing EDR in the realm of edge computing. This dual-focused approach signifies an advancement in integrating energy management with computing resource allocation. Moreover, the authors in [169] contributed to this field by developing a two-phase game-theoretical algorithm. This algorithm addresses the challenges posed by the mobile edge computing EDR problem, providing a strategic solution that balances multiple factors in a complex computing environment. Additionally, the authors in [170] created efficient online auctions specifically for scheduling cloud computing jobs. These auctions take into account completion deadlines, offering a time-sensitive approach to job scheduling in cloud computing environments.

However, despite these varied and significant contributions, none of these studies have explored the potential of using EVs as energy sources in their demand response models. Moreover, the consideration of bid deadlines in designing auction mechanisms for EDR has not been addressed in these works. This gap indicates a potential area for future research, where the integration of EVs as a dynamic and mobile energy source could offer new dimensions and efficiencies in demand response strategies within edge computing and cloud computing environments.

In V2G, edge computing is utilized by local controllers who primarily rely on local data. This decentralized approach facilitates the coordination among a large number of local controllers, substantially enhancing the privacy of users and considerably reducing the complexity associated with V2G computational problems [171][157]. For example, the authors in [162] explored a combined routing and V2G challenge and proposed a distributed algorithm based on dual decomposition, specifically designed to protect the privacy and autonomy of EV users. Similarly, the authors in [172] approached a related challenge using an approximate distributed algorithm characterized by its lower computational demands. Another significant contribution in this field is from the authors in [173], who introduced a distributed control algorithm that caters to a variety of V2G objectives. This algorithm is particularly effective in enhancing the stability of the power system and in reducing the total cost associated with EV charging. Recent advancements in distributed V2G systems are exemplified by the development of the Internet of Smart Charging Points (ISCP) [158][174]. In this system, each CP is equipped with a computational unit, termed a Smart Charging Point (SCP). These SCPs collaborate among themselves, thereby boosting computational performance. The communication network employed in this system utilizes the small-world network model, which is known for its efficiency in transmission and cost-effectiveness in terms of wiring. A key feature of this system is its focus on maintaining the privacy of EV users; it achieves this by processing and storing EV charging data locally at the SCPs and implementing a stringent protocol for handling sensitive information post-dispatch, known as "burn after dispatching."

The authors in [175] proposed a unique auction framework tailored for EVs. This framework utilizes a smoothed analysis mechanism, which aims to motivate EVs to participate in EDR. The design of this mechanism takes into account the variability and unpredictability of EV participation, offering a more flexible and adaptive approach to integrating EVs into EDR systems. Moreover, the authors in [176] developed a novel approach that combines a polynomial-time online algorithm with an auction mechanism. This dual approach is designed to incentivize EVs that have surplus energy to sell their excess energy. The goal is to meet the charging demands of other EVs, thereby creating a more efficient and cooperative energy distribution system among EV users.

The authors in [177] introduced a mechanism that focuses on stimulating energy interactions between EVs and the power grid, leveraging V2G technologies. This mechanism is significant as it opens up possibilities for bi-directional energy flows, allowing EVs to not just consume energy but also to provide it back to the grid, creating a more dynamic and interactive energy network. Additionally, the authors in [178] devised an online mechanism that addresses the EV charging scheduling problem. This mechanism considers both the charging costs and the potential dissatisfaction of EV users, indicating a balance between

economic efficiency and user satisfaction in EV charging management. Moreover, the authors in [179] presented an online linear programming model to manage the variability of charging rates in each control period. This model addresses the challenge of fluctuating charging rates, ensuring more stable and predictable charging processes for EVs.

While there are significant advancements in the field of EVs and V2G/V2M technologies, particularly in the realm of edge computing, these developments also face notable challenges and limitations. A prominent focus of current studies is on EV charging or their interaction with the grid. However, these studies often do not directly align with edge computing contexts. An essential aspect missing in these works is the consideration of the constrained aspect of the edge environment such as computational resources and memory. The absence of this consideration in existing mechanisms opens up a potential area for further research and development. Integrating time-sensitive strategies into these frameworks could significantly enhance the efficiency and applicability of these online mechanisms for EVs across various energy and computing scenarios. On the other hand, the application of distributed edge computing in V2G systems comes with its own set of challenges. The strategies outlined in [158] and [174], for instance, rely heavily on accurately predicting future power grid data—a complex and largely unresolved issue. Accurately forecasting grid load consumption and renewable energy source generation remains a difficult task. Therefore, while distributed edge computing presents promising opportunities for enhancing V2G systems, addressing these challenges is crucial to ensure their effective, safe, and robust implementation in a variety of scenarios.

In addition, the technological advances in V2G/V2M also bring to light new concerns, particularly in cybersecurity. The expanded attack surface of smart grids introduces new vulnerabilities, creating opportunities for adversaries to launch cyber attacks. This scenario is particularly pertinent in the domain of V2M systems, which utilize transactive energy and advanced communication technologies along with AI-based technologies. Understanding and addressing these emerging cyber vulnerabilities in V2M systems is essential, as adversaries can exploit these vulnerabilities to compromise the integrity and functionality of these systems. Thus, the exploration of these security challenges within V2M systems becomes a crucial aspect of advancing these technologies safely and securely. It is evident in the literature that data integrity attacks (e.g., false data injection attacks) are real threats that can take place against smart meters and AMI systems. Adversaries can launch such attacks by exploiting the vulnerabilities in those systems. For instance, the authors in [180] launched an FDI attack against their designed SCADA testbed system by using the vulnerabilities of Modbus protocol. The authors were able to manipulate the smart meter data despite password protection in ICS and SCADA system. Other works such as [181] considered the privacy and security threats related to smart meters. The authors

studied the smart meter's attack surface and the possible attack vectors. In [182] and in their previous work [183], the authors proposed a statistical anomaly detection method to prevent false data attacks that exploit AMI vulnerabilities. Other studies that focus on electricity theft built their threat model based on the assumption of existing vulnerabilities of AMI and smart meters [184] [185] [186] [187] [188].

The authors in [189] presented a threat model and developed attack strategies enabling adversaries to manipulate a large number of charging stations by exploiting existing and newly discovered vulnerabilities in the V2G protocols ISO 15118 and OCPP. A test setup was developed to evaluate several open-source tools to validate the success of their V2G attack strategies. In addition, it was demonstrated that ISO 15118 messages sent over PLC could be intercepted in plaintext, as shown in [190] and [191]. Specifically, the testbed in [191] permitted attackers to access and alter the entire stack by injecting perturbed V2G messages.
The authors in [192] proposed a relay attach on EV charging system in which adversaries can steal energy during a charging session. The authors developed their attack strategy based on the vulnerability of ISO 15118 standard. The authors proposed an extension to the standard to prevent such attacks. Other works such as [193] also provided enhancement to the current functionalities of ISO 15118 standard.
The authors in [194] exploited wireless communication vulnerabilities between the EV and charging station through the ISO 15118-8 and IEEE 802.11 standards. The authors were able to launch a DoS attack risking the availability of the V2G service. The authors in [195] analyzed the cyber threats against ISO 15118 standard where the charging service availability and integrity can be compromised. The authors showed that adversaries can exploit non-binding certificate authorities to perform DoS attacks on charging stations. Another study [196] on ISO 15118 demonstrated the impersonate attack by copying transactions in the vehicle's RFID chip and disguises itself as another vehicle. Moreover, adversaries can exploit the standard vulnerabilities to report false State of Charge (SoC) data to the charging infrastructure [197]. ISO 15118 is also susceptible to other attacks that fabricate metering data and battery level to give malicious EVs smaller bills and higher charging priority. A malicious EV could send a charging request indicating a lower SoC to secure a higher charging priority. For instance, the authors in [198] developed reinforcement learning framework to generate intelligent and stealthy attacks to falsify the SoC by exploiting ISO 15118 vulnerabilities. This tactic could be expanded to mobilize a large number of compromised EVs that then obtain higher priority, blocking regular users from charging and effectively leading to a denial-of-charge attack. Despite all of the mentioned research works on the susceptibility of V2G to cyberattacks, we believe, evident by this thesis, that it is more difficult to launch a cyberattack on a vehicle when compared to a microgrid's smart meter.

Chapter 3

The Impact of Data Integrity Attacks on V2M Systems

In this chapter, our focus is on addressing data integrity attacks against V2M services. It is worth noting that we tackle data integrity attacks from the vehicle's side, whereas the next chapters are mainly focused on adversarial attacks from the microgrid's perspective. We try to understand the feasibility of launching such attacks in V2M context where the attack surface under study is the vehicle. We analyze the V2M system interaction and identify the possible vulnerabilities to gain a deeper understanding of the attacks characteristics and potential impacts on the system. We develop a detector that leverages ML methods to effectively counter data integrity attacks. The utilization of unsupervised ML methods allows us to detect anomalous patterns and deviations in the data, enabling the detector to proactively respond to emerging attack vectors. Our proposed method aims at enhancing the security and resilience of V2M systems by accurately identifying and mitigating potential threats posed by data integrity attacks.

3.1 Introduction

Smart grids enable sustainable and resilient electricity services by allowing consumers to act as producers (a.k.a prosumers) via energy trading. Energy generated from the Distributed Generators (DGs), such as solar panels and wind turbines, can be shared among other entities connected to the grid, forming Community Resilience Microgrids (CRMs) [199]. Recent research aims at maximizing the V2G efficiency of the delivered power while reducing the cost using various approaches. For instance, the study in [10] has proposed

a V2G cost-objective optimization model that aims at finding the closest EVs to a microgrid considering the communication aspects. In addition to optimization models, machine learning techniques such as Reinforcement Learning (RL) are used in power management for grid-tied microgrid problems where V2G service is considered as an alternative power source [12]. Another study models the interactions between the EVs and microgrids where the suppliers (i.e. EVs) specify the plug-in length, arrival times and the amount to supply/sell [200].

As demand on the power grid continues to rise, EVs are increasingly utilized as mobile energy storage units to facilitate energy trading and prevent power shortages. This integration of EVs with smart grids has resulted in an expanded attack surface, enabling adversaries to launch sophisticated attacks on the system. Consequently, data integrity attacks in modern smart grids, particularly in Vehicle-to-Grid (V2G) and V2M applications, are anticipated to escalate. To address this issue, we propose a novel approach for modeling data integrity attacks in V2M applications. By harnessing the power of unsupervised machine learning, we develop an intelligent detector capable of identifying and countering these data integrity attacks.

3.2 System Model for V2M Services

The threat model builds on the optimization model presented in [10] so it is worth revisiting the implemented optimization model before proceeding with the the threat model.

3.2.1 Optimization Model Revisited

The power management framework uses real time information of the microgrid power demand to find the optimal set of EVs to participate in the process, considering reliable communication between the cellular base-station and EVs. The optimization model in [10] is framed as follows: a set of smart microgrids M, that are predicted to suffer from power outage, send highly-time sensitive service requests to the cell's base-station [8]. To ensure reliability, the base-station uses the Transmission Control Protocol (TCP) to acknowledge the reception of the requests [13]. The requests contain information of the anticipated energy demand D_m(kWh) and the microgrid locations. The base-station broadcasts the request to N EVs within the coverage range. Each EV responds to the request with three pieces of information: (i) selling price P_n(\$/kWh), (ii) contribution percentage α_n of the EV's battery E_n, and (iii) current location. Then, the base-station computes the distances between the microgrids K_{im}(km) and chooses an optimal set of EVs to serve the microgrids'

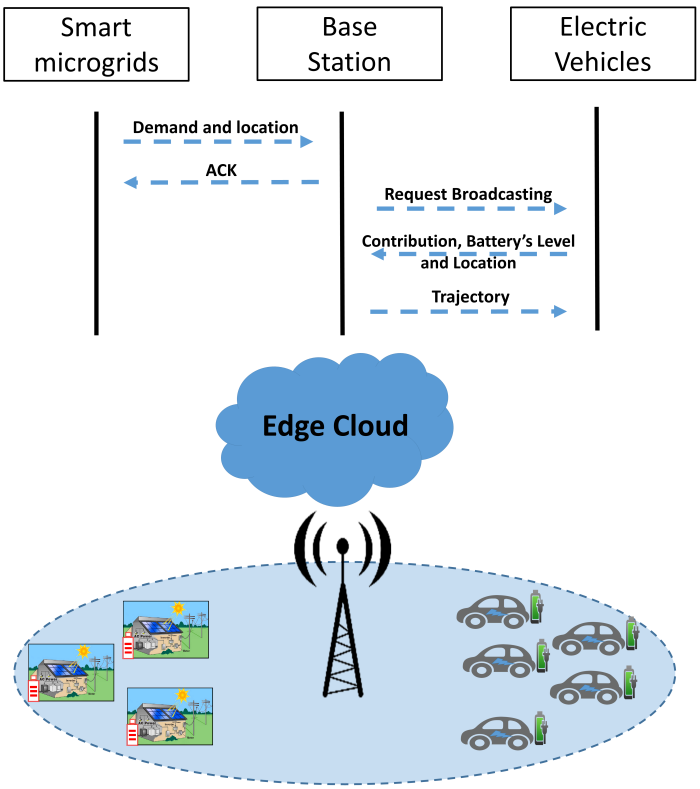


Figure 3.1: Messages flow in our framework

requests. Lastly, the base-station sends the trajectories to the EVs, starting with the first microgrid T_m to be serviced. The message flow for the V2M service is depicted in Figure 3.1.

The V2M model builds on a MILP formulation to find the optimal set of EVs to supply the affected microgrids with the required power until the main grid is restored. The objective of this optimization problem is to find the charging order of the microgrids to minimize operational costs, specifically focusing on reducing the travel distances of the EVs and the waiting times for charging. Practically, the deployment of this optimization model should run in the computational component (i.e., edge computing) of the gNodeB. Details of this model are presented in our previous work [10]. The objective function in (3.1) can be solved by the set of constraints where (3.2) and (3.6) presents the key subset of these constraints.

$$Minimize \sum_{m=1}^{M} cost_m \tag{3.1}$$

subject to

$$cost_m - (\sum_{n=1}^{N} S_m^n + \sum_{i=1}^{M}\sum_{n=1}^{N}\sum_{r=1}^{M} z \times K_{im} \cdot A_{im}^{nr} + \sum_{n=1}^{N} O_m^{n1} \cdot T_m \times z) \times P_n - (\sum_{n=1}^{N} S_m^n/H) \times C - \sum_{n=1}^{N}\sum_{r=1}^{M} O_m^{nr} \times v = 0, \forall m \in M \tag{3.2}$$

$$A_{im}^{nr} - O_m^{nr} \leq 0, \forall i \in M, \forall m \in M, \forall n \in N, \forall r \in M \tag{3.3}$$

$$A_{im}^{nr} - O_i^{nr+1} \leq 0, \forall i \in M, \forall m \in M, \forall n \in N, \forall r \in M \tag{3.4}$$

$$O_i^{nr} + O_m^{nr+1} - A_{im}^{nr} \leq 1, \forall i \in M, \forall m \in M, \forall n \in N, \forall r \in M \tag{3.5}$$

The optimization model requires a formulation of the conditional variable indicating whether microgrid i and m are served by EV n at the r-th and $(r+1)$-th order, respectively. This is represented by the product of two variables, i.e., $O_i^{nr} \times O_m^{nr+1}$. The non-linearity of this product is linearized through the constraints specified in (3.3)-(3.5). The first two constraints ensure that A_{im}^{nr} will be zero if either O_i^{nr} or O_m^{nr+1} is zero. The final constraint guarantees that A_{im}^{nr} will be set to one if both binary variables are one, thus linearizing A_{im}^{nr} within the model without affecting the other constraints. The constraint in (3.6) guarantees that the total energy transferred from EV n plus the energy consumed during travel does not exceed the proportion α of the initial energy level of the EV's battery E_n.

$$\sum_{m=1}^{M} S_n^m + \sum_{i=1}^{M}\sum_{m=1}^{M}\sum_{r=1}^{M} z \times K_{im} \times A_{im}^{nr} + \sum_{m=1}^{M} O_m^{n1} \cdot T_m \times z \leq E_n \times \alpha_n, \forall n \in N \tag{3.6}$$

$$\sum_{m=1}^{M}\sum_{r=1}^{M} O_m^{nr} - \sum_{i=1}^{M}\sum_{m=1}^{M}\sum_{r=1}^{M} A_{im}^{nr} = 1, \forall n \in N \tag{3.7}$$

$$\sum_{m=1}^{M} O_m^{nr} \leq 1, \forall n \in N, \forall r \in M \tag{3.8}$$

$$\sum_{n=1}^{N} \sum_{r=1}^{M} O_m^{nr} \geq 1, \forall m \in M \tag{3.9}$$

The constraint (3.7) specifies the number of trips made by EV n between two microgrids. The constraint in (3.8) ensures that for a given EV n, each microgrid m can be assigned order r at most once. Assuming that EVs are the sole source of power, the constraint in (3.9) stipulates that each microgrid m must be served by at least one EV to meet partially/fully its power needs. It is worth noting that the attack will have an indirect impact of the total cost (i.e., the optimization model's objective). The manipulation of the contribution factor (α) will result in changing the amount of the total cost as observed from equation (3.2) and (3.6). The alteration of this factor can lead to changes in how resources are allocated and priced within the microgrid system. If (α) is increased or decreased through malicious tampering, it can cause the system to overestimate or underestimate the actual contribution of an EV.

It is important to mention that the sole purpose of the aforementioned optimization model is to match sellers (i.e. EVs) with buyers (i.e. smart microgrids) for a V2M application while finding the optimal microgrids charging order.

3.2.2 Threat Models in V2M Services

To understand the threats and possible attacks on V2M application, one should comprehend the involved entities that form the threat model. Figure 3.2 presents a high level abstraction of the threat model entities including adversaries and vulnerabilities. We conceptually follow the threat model in [201] that is presented for IoT security in smart grids.

1) The adversary represents the first entity of the threat model. The threat level of an adversary is defined by three main elements, namely the required access, the adversary's resources and their motivation.

(i) *Required access:* defines the type of entry point to the network. In order for the adversary to mount an attack, a physical or cyber access to the network is required. Physical-access can be direct through a malicious agent (i.e. an insider) who has privileges in the network, or indirect by an outsider who seeks to gain privileges. It is more likely for attacks to be launched by outsiders who escalate their privileges in the network rather than

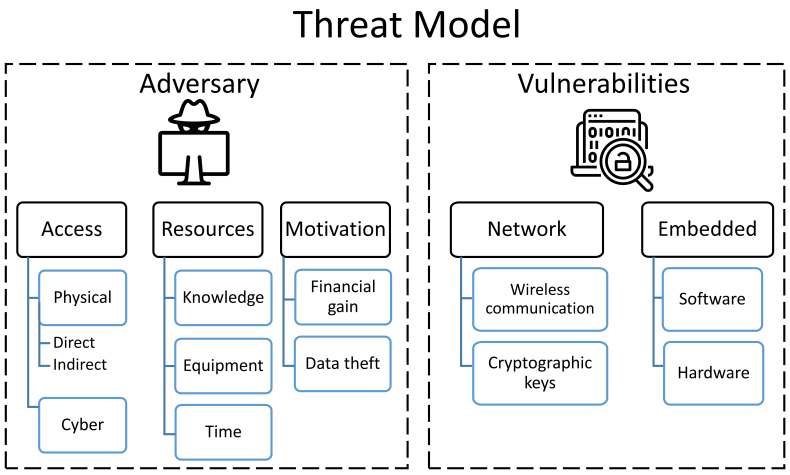


Figure 3.2: High-level abstraction of the threat model for V2M

malicious users, especially in a V2M service, as we will discuss in the performance evaluation section. On the other hand, cyber-access allows the adversary to connect remotely to the network via insecure communication channels. For instance, the adversary can exploit network-based vulnerabilities to gain a full control over the participated EVs during the V2M operation.

(ii) *Required resources:* represent the needed capabilities of the adversary to launch a successful attack. The adversary's resources are defined by the adversary's expertise of the system, the required equipment and time. If the attacker is an expert and aware of the system interactions, it is called a white-box attack. On the contrary, if the attacker is a novice and lacks the necessary knowledge to initiate an attack on the system, it is called black-box attack. In addition, the adversary might need different types of equipment to attack V2M network. For instance, to access an insecure communication channel, the attacker needs to be within the communication range of the traveling EVs; thus, using a vehicle is necessary to launch an attack [202]. However, other attacks might need simpler equipment, such as a desktop computer. Another vital resource for the adversary in a V2M application is the amount of time required to perform the attack. The V2M application poses dynamic changes of the network, hence, a dynamic attack surface. Therefore, the adversary has to perform the attack within a limited time window.

(iii) *Motivation:* can be viewed as another definition of the attacker's utility (i.e. the possible gains of performing a successful attack considering the associated efforts and risks to the attacker). Energy and data theft, financial gain, and service disruption are typical examples that drive the adversary [201].

2) The second entity in Figure 3.2 is the vulnerabilities in V2M services. There are two parties that form the attack surface in a V2M service, that is, vehicles and smart microgrids. Both parties share network-based and embedded vulnerabilities.

(i) *Network vulnerabilities:* wireless communication and key management expose the network to various types of vulnerabilities. It is worth noting that Autonomous Vehicles (AVs) can be considered as EVs if they use battery units. Thus, network-based vulnerabilities of AVs hold for EVs as well. One of the most commonly used wireless technology for in-vehicle network is Bluetooth. Its ability to hop fast, resist noisy environments and support multi-channel communication through Frequency Hopping Spread Spectrum (FHSS) technology made it a favourable technology for Controller Area Network (CAN) communication [203]. However, Bluetooth technology suffers from various vulnerabilities, as reported in [202].

(ii) *Embedded system vulnerabilities* are another way for adversaries to break into the EVs. Firmware and operating system vulnerabilities are two main issues for the software-layer of embedded systems. For instance, firmware vulnerabilities allow the attacker to gain full access to the system, which can provide the attacker with the control to read and change messages [204] [205]. Moreover, the attacker can pursue and achieve privilege escalation by exploiting operating system vulnerabilities. Hence, the attacker performs the attack seamlessly. Other software vulnerabilities, such as in Engine Control Unit (ECU), are possible in [206] whereas embedded hardware in EVs present other exploitable vulnerabilities. The lack of proper hardware implementations and cryptographic algorithms ease the path for adversaries to launch attacks on EVs. Side-channel attacks such as power analysis attacks allow the attacker to extract the encryption keys [207].

3.3 Modelling of Data Integrity Attack Against V2M Application

We apply the discussed threat model to analyze the impact of the possible security flaws on the V2M operation (i.e. optimization model). We assume an outside adversary with cyber-access to the system is exploiting in-vehicle communication network using cryptographic key vulnerabilities. The adversary is assumed to be able to alter the messages exchanged between the EVs and the base-station. Hence, we implement a data integrity attack, assuming a white-box attack with an expert adversary of the underlying interactions within a V2M service. However, the base-station / edge node is equipped with an intelligent system to detect data integrity attack attempts.

3.3.1 Data Integrity Attack

The aim of study is primarily to quantify the impact of cyber-security on energy trading process between EVs and smart microgrids under the presence of an anomaly detector. After analyzing the risks of different parameters involved in the process, we have empirically chosen the contribution value (α_n) to be the best target for an adversary to manipulate. The contribution value is responsible for the amount of traded energy during a V2M service. The adversary's motivation is to weaken the resiliency of the smart microgrid community by reducing the amount of energy provided to microgrids (a.k.a buyers). The adversary aims to change the contribution value of the EVs, as they are more susceptible to cyberattacks. The false contribution values are received at the base-station and used as inputs to the optimization model. However, different false contribution values affect the V2M service differently. For instance, some of the data integrity attacks have zero impact on the service. Thus, we had to assess not only the success of the data integrity attacks but also whether the attacks impacted the service. We define the Impactful Attacks (IA) in formula 3.10 as the attacks that can bypass DBSCAN and affect the V2M service.

$$IA = \frac{\text{Number of undetected attacks of non-zero impact}}{\text{Total number of attacks}} \tag{3.10}$$

3.3.2 Anomaly Detection

We use DBSCAN to detect the data integrity attack on the V2M application. However, some of the attacks can bypass DBSCAN and impact the V2M services. DBSCAN algorithm is controlled by two parameters, ϵ and $min-points$ where ϵ is responsible for the neighborhood search radius and the $min-points$ parameter controls the minimum number of points needed to establish a cluster. The parameters were empirically chosen to make it harder on the adversary to bypass the detector. DBSCAN algorithm works as follows: 1) An initial point is selected and marked as visited. 2) The points within the search radius of epsilon are counted and added to a set. 3) The initial point is considered as a new cluster if the number of points exceeds the predefined min-point value. This process is continued for all points in the neighbourhood. 4) If the number of points is less than the min-point, the point is defined as noise. 5) These steps are repeated until all points are clustered.

DBSCAN has a time complexity of $O(n^2)$ but this can be reduced to $O(nlogn)$ with parameter optimization [208]. Unlike K-means, DBSCAN does not require pre-specification of the number of clusters, which makes it a good fit for anomaly detection problems.

3.4 Performance Evaluation

In this section, we present a detailed description of the simulation settings employed in our study. We discuss the various parameters and configurations used to evaluate the performance of our proposed method. Furthermore, we analyze the obtained results, providing comprehensive insights into the effectiveness and limitations of our proposed method.

3.4.1 Simulation Settings

To assess the associated risks of the proposed data integrity attack on the contribution value of the optimization model, Optimization and Simscape Electrical toolboxes are used. All simulations are performed using Intel Core i5-7500 CPU with 16GB of RAM running on a Windows 10 system. The optimization toolbox solves the cost-based MILP model, whereas the Simscape Electrical Toolbox simulates microgrids and EVs to provide synthetic power data [209]. Table 3.1 presents the used simulation parameters.

We present two sets of microgrids M=$\{4, 8\}$, where a microgrid's demand (D_m) is picked randomly between [5-30] kWh based on the synthetic power data. Similarly, we consider two scenarios for the number of EVs in the V2M service operation N=$\{6, 12\}$, with different battery levels (E_n) between [10-40] kWh. The adversary targets the contribution value of the EVs. However, the number of exploitable EVs can vary for diverse reasons. Therefore, another objective of this study is to anticipate the number of EVs to be attacked, that would have the heaviest impact on the microgrids' resiliency. Thus, we present three different percentage of EVs that could be exploitable: 33%, 66% and 100%. That is, for N=6, we study the impact of having 2, 4 and 6 exploitable EVs; and for N=12, the number of exploitable EVs is set to 4, 8 and 12. The adversary aims at changing the EVs' contribution values with 100% reduction. However, DBSCAN will prevent that from occurring. After DBSCAN's fine-tuning, we select the ϵ and $min - points$ parameters as 1.5 and 5, respectively.

3.4.2 Numerical Results

In this section we analyze the impact of the data integrity attack on the contribution value (α_n) with different number of exploitable EVs. The optimization model has three outputs, outstanding demand (kWh), total cost ($) and average vehicle's revenue ($). The outstanding demand defines the amount of the microgrid's energy that could not be supplied by the EVs. The total cost represents the microgrid's cost of exchange for the

Table 3.1: Simulation parameters

Notations	**Value**
Number of EVs	$\{6, 12\}$
Number of microgrids (M)	$\{4, 8\}$
Selling price (P_n)	0.201 \$/kWh
Average energy consumption per km (z)	0.18 kWh/km
EV charger's power (H)	20 kW
Waiting time price (C)	10 \$/h
Service request price (v)	1 \$/request
Distance between microgrids (K_{im})	[0.2-3] km
Distance from EV n initial location to first microgrid m (T_m)	[0.2-3] km
Microgrid's demand (D_m)	[5-30] kWh
EV's initial battery level (E_n)	[10-40] kWh
EV's original contribution percentage (α_n)	{20,40,90}%
Reduction percentages of the original contribution	{30,50,70,90}%
ϵ	1.5
$min - points$	5
Number of exploitable EVs for N=6 under 33%, 66% and 100%	2, 4 and 6, respectively
Number of exploitable EVs for N=12 under 33%, 66% and 100%	4, 8 and 12, respectively

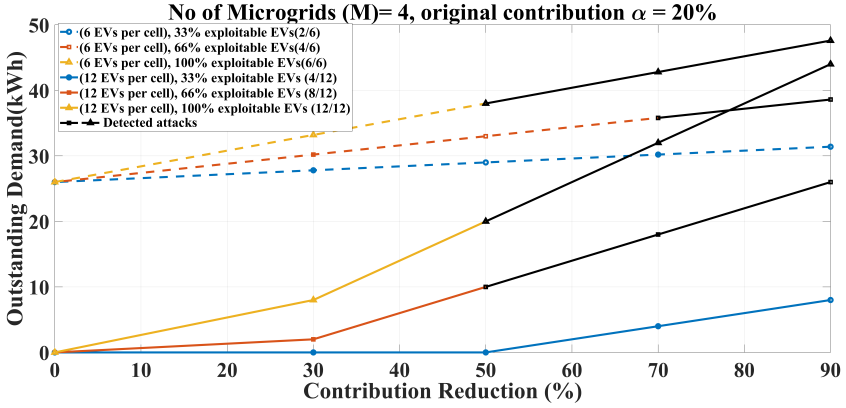
EVs' energy; and each EV makes a revenue by participating in the request defined by the vehicle's average revenue. For the purpose of this chapter, we limit our focus to analyzing the impact of data integrity attacks on the outstanding demand.

Analyzing data integrity attacks under four microgrids

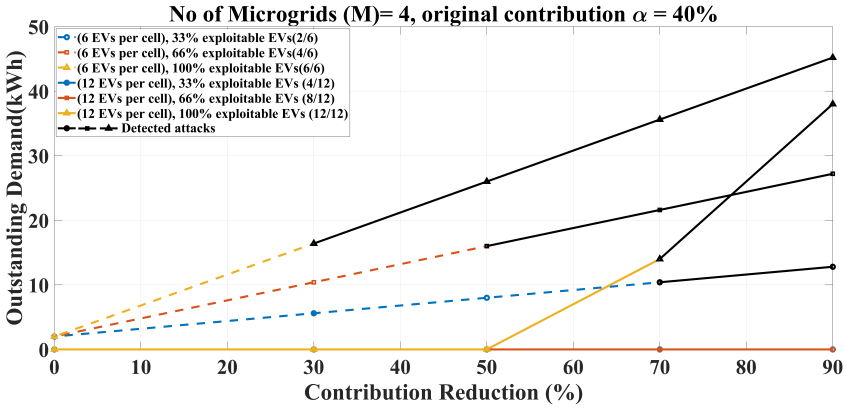
The outstanding demand of the 4-microgrid energy request is depicted in Figure 3.3a. The EVs are willing to contribute to the request with 20% of their batteries. However, the adversary aims at reducing the original contribution with different percentages as shown on the x-axis. It is important to note that 0% reduction means that the original contribution value has not been changed. In other words, the 0% reduction denotes a "no-attack" case on the V2M operation. We present the no-attack case to show the benign scenario of the model. The dotted-lines represent scenario-1, with six EVs, with different number of

exploitable EVs. The solid-lines represent scenario-2, with twelve EVs. The black lines denote the detected attacks by DBSCAN. Under the no-attack case (i.e., 0% reduction), the outstanding demand of scenario-1 is greater than scenario-2. That is because there are more EVs in scenario-2 that can contribute to the request when compared to the number of EVs in scenario-1. Beyond the no-attack point, the effect of the data integrity attacks on the contribution value starts to emerge. The number of exploitable EVs has different impacts of the outstanding demand. For instance, under the same scenario, the number of attacked EVs is directly proportional to the outstanding demand. The 33% attacked EVs has the least impact on the outstanding demand, whereas the 100% attacked EVs has the highest impact. However, when we compare the two scenarios against each other, one interesting observation from Figure 3.3a is that the data integrity attacks on 33% of scenario-1 is more impactful than 100% attacked EVs of scenario-2, for a contribution reduction of 30%. That means attacking two EVs in scenario-1 is more critical and risky to the V2M operation than attacking twelve EVs in scenario-2. This occurs because under the no-attack case, the outstanding demand of scenario-1 is higher than scenario-2. Thus, the six EVs of scenario-1 cannot cover all of the microgrid demands when compared to the twelve EVs scenario, which would have surplus energy even after covering all the demands.

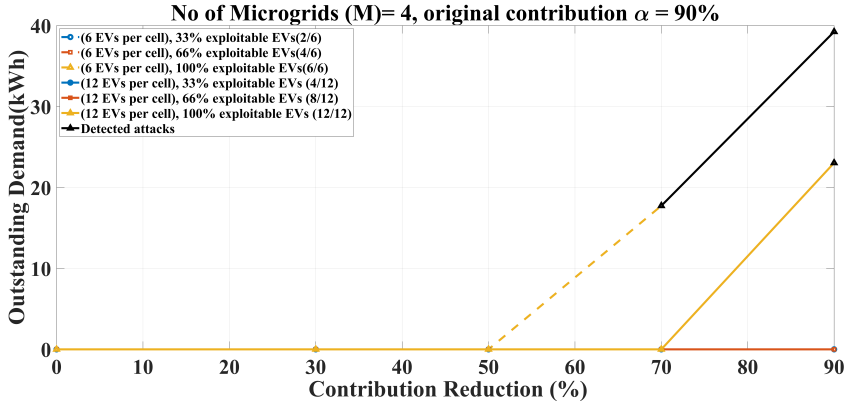
Consequently, even with a 30% reduction in the contribution of the twelve EVs in scenario-2, the EVs can still meet the microgrid demands sufficiently. Furthermore, the 30% contribution reduction attack translates into more EV requests to compensate the energy loss caused by the attack. This results in an excessive communication request which will overload the communication infrastructure, especially if those requests are clustered as highly-time sensitive [13]. Moreover, in scenario-2 under 33% attacked EVs, the data integrity attacks have an absolute zero impact on the outstanding demand until a reduction value of 50%. The adversary's utility is negative since the attack has no gains to the adversary. Starting at a reduction of 50% and onward, DBSCAN detects the adversarial attempts on attacking 66% and 100% of the EVs for both cases. That exhibits a promising performance of the applied anomaly detector. However, DBSCAN fails to detect the 33% attacked EVs, for both scenarios, for all reduction percentages. The adversary succeeds to impact the outstanding demand, hence the microgrids' resiliency, by reducing the original contribution value by 70% with attacking 33% EVs of scenario-1.



(a)



(b)



(c)

Figure 3.3: Impact of the data integrity attack on the outstanding demand for M=4. (a) original contribution = 20%, (b) original contribution = 40%, (c) original contribution = 90%

In Figure 3.3b, the EVs' contribution percentage doubles (i.e. 40%) with same number of microgrids M=4. Outstanding demand of the no-attack scenario has dropped when compared to the 20% contribution case. The adversary is successfully able to reduce the original contribution by 30% without being detected for scenario-2. Similarly, for scenario-1, all the data integrity attacks deceived DBSCAN detector except for the 100% exploitable EVs case. At 50% reduction, for scenario-1, the data integrity attack on the 33% exploitable EVs remains undetected, whereas for the other two cases (66% and 100% exploitable EVs), DBSCAN detects the attacks. Beyond the 50% reduction, all the data integrity attacks on the scenario-1 are detected. On the other hand, the data integrity attacks under scenario-2 remain undetected for the 33% and 66% cases until 90% reduction. However, even though the attacks successfully deceived DBSCAN, the attacks have almost zero impact on the outstanding demand. Hence, not all successful attacks can impact the V2M operation. Furthermore, that can be seen, as the original contribution increases to 90% as depicted in Figure 3.3c. It is worth noting that none of the undetected attacks have any impact on the outstanding demand. In addition, the 100% exploitable EVs case for scenario-1 and scenario-2 are detected successfully. Hence, the performed attacks on an original contribution of 90% have no impact on the V2M operation for both scenarios. That is because the EVs for both scenarios have enough energy to meet the microgrid demands even under different contribution reduction attacks. However, as mentioned earlier, the attacks on the contribution values will result in heavier communication load. Lastly, increasing the original contribution from 20% to 40% results in doubling the detected data integrity attacks of scenario-1.

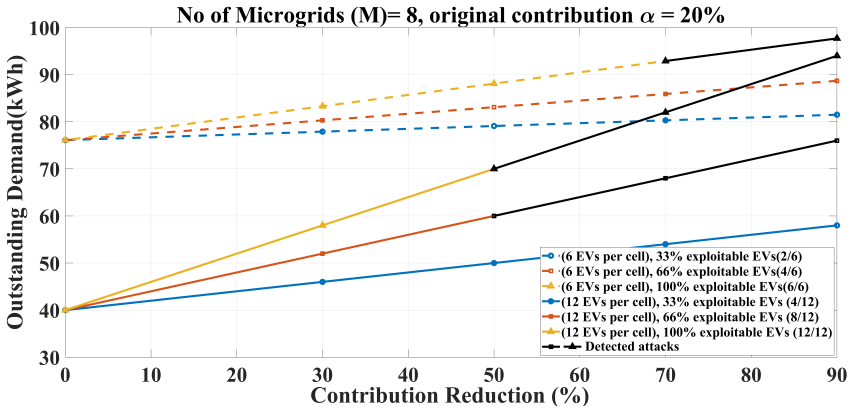
Analyzing data integrity attacks under eight microgrids

We also double the number of microgrids (i.e. M= 8) to further investigate its relation to the outstanding demand. In Figure 3.4a under 20% original contribution and 8 microgrids, the outstanding demand increases by around 40kWh for both scenarios under the no-attack case (i.e. 0% contribution reduction) when compared to the 4-microgrids case. That means neither the six EVs scenario nor the twelve EVs scenario can sufficiently meet the microgrid demands. Contrary to the 4-microgrids case, the data integrity attack against the 33% exploitable EVs of scenario-2 has impacted the outstanding demand under all different contribution reduction attacks. Hence, it is shown that increasing the number of microgrids has a linear impact on the outstanding demand under 20% original contribution. On one hand, doubling the original contribution to 40% lowers impact of the integrity attack on the outstanding demand for both scenarios as shown in Figure 3.4b. On the other hand, doubling the number of microgrids to 8 results in an increase in the outstanding demand. Hence, the positive impact of increasing the original contribution from 20%

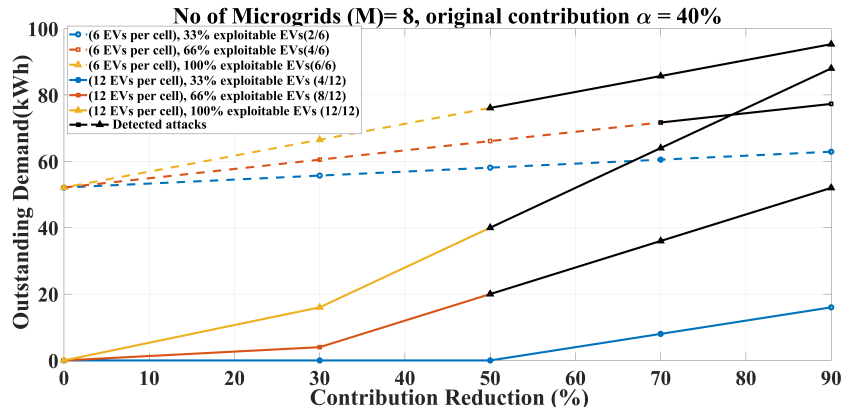
to 40% cancels the negative impact of increasing the number of microgrids from 4 to 8. Thus, Figure 3.4b could represent 4 microgrids with 20% original contribution (i.e. an exact replica of Figure 3.3a). As we further increase the original contribution to 90%, the data integrity attacks lead to a lower impact on the outstanding demand as depicted in Figure 3.4c. However, when compared to Figure 3.3c with M=4 and 90% original contribution, the impact of different exploitable cases under scenario-1 becomes noticeable. Since not all the data integrity attacks affect the outstanding demand of the V2M operation, we evaluate the impact of the performed attacks on 4 and 8 microgrids as presented in Figure 3.5. It is shown that the impactful attacks percentage is affected by four factors: (i) the original contribution (ii) the number of EVs (N) (iii) exploitable EVs percentage (iv) the number of microgrids (M). As the original contribution percentage increases, the impactful attacks either decrease or remain constant. For instance, in 33% exploitable EVs for N=6, the impactful attacks rate drops from 100% to 50% as the original contribution increases from 20% to 40%. Similarly, the impactful attacks rate drops from 100% to 50% as the N grows from 6 to 12 under the same contribution value of 20% and 33% exploitable EVs. Furthermore, increasing the exploitable EVs percentage results in a decrease of the impactful attacks rate when the other factors remain the same.

3.5 Summary

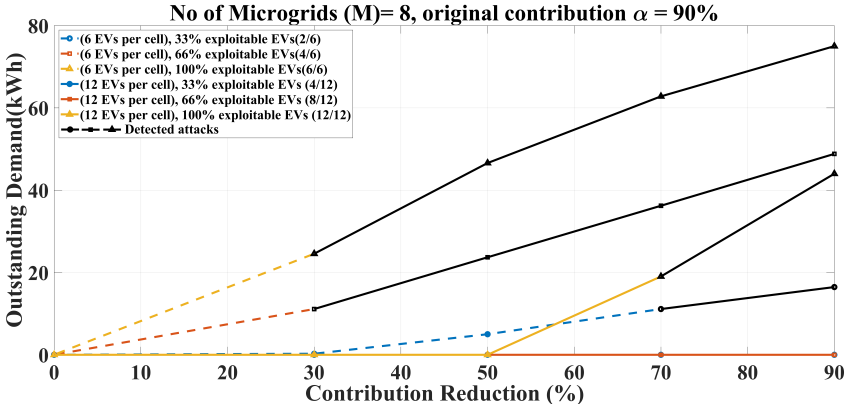
In this chapter, we studied data integrity attacks as potential cyber-threats on Vehicle-to-Microgrid (V2M) service operation, in which the adversary alters the original contribution of the EVs. We have modeled and performed an in-depth impact analysis for these threats considering the microgrid demands that could not be covered by the EVs (i.e. the outstanding demand), in a V2M operation. Numerical results have shown that the risk of data integrity attacks on the outstanding demand as well as the impactful attacks rate can drop by increasing the original contribution and the number of EVs; the risk rises when the number of exploitable EVs and the number of microgrids increase. For instance, increasing the original contribution from 20% to 90% resulted in dropping the outstanding demand by up to 100% and 93% for the 4 and 8 microgrids cases, respectively. Similarly, rising the number of EVs from 6 to 12 resulted in dropping the outstanding demand by up to 76.5% and 30% for the 4 and 8 microgrids cases, respectively.
It is worth noting that this chapter served as a guideline on whether impactful adversarial attacks can be launched on vehicles. We concluded that while it is possible for such attacks to target vehicles, the complexity of the attack surface makes unlikely to happen. Hence, we switch our focus in the next chapters to adversarial attacks on smart meters (i.e., from microgrid's side).



(a)

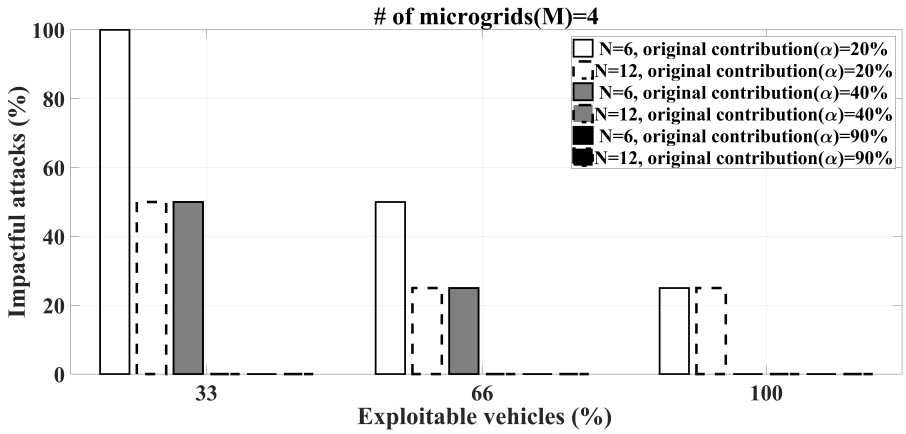


(b)

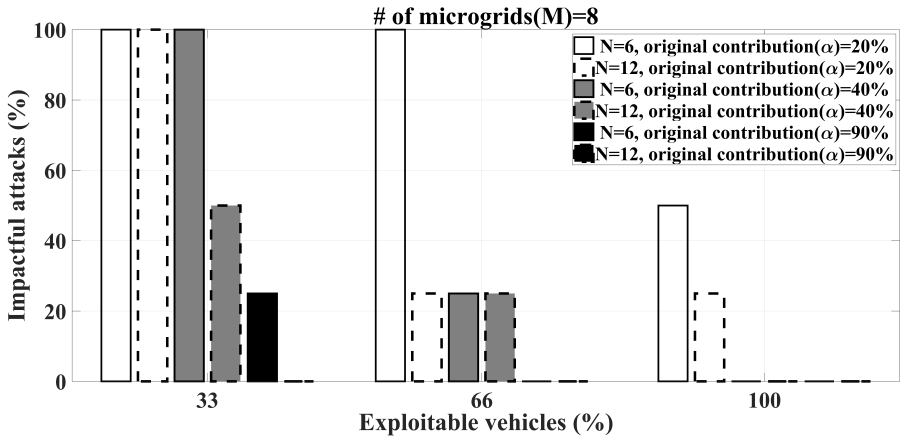


(c)

Figure 3.4: Impact of the data integrity attack on the outstanding demand for M=8. (a) original contribution = 20%, (b) original contribution = 40%, (c) original contribution = 90%



(a)



(b)

Figure 3.5: Impactful attacks on the V2M operation for different exploitable EVs (a) M=4 (b) M=8

In addition, the work of this chapter has shown the complexity and difficulty launching impactful and successful adversarial attacks against V2M services, from a vehicular aspect. Moreover, adversaries require advanced technical experience to execute their attacks. Hence, for the next chapters, we will study the adversarial attacks from the microgrids perspective. We will not use α as the adversary's targeting point. However, the adversary will aim at changing the power data reported by smart meter units.

Chapter 4

The Impact of Adversarial Machine Learning Attacks on Vehicle-to-Microgrid Services

As we realized that it is significantly challenging for adversaries to launch adversarial attacks against vehicles, we shifted the focus of the attack surface to the microgrid aspect of V2M. In this chapter, we focus on two main types of adversarial machine learning attacks that occur during inference (test) time, namely inference attacks and evasion attacks. We study the impact of both attacks as they are potential venues of an adversary against V2M services. Hence, throughout this chapter we seek answers to the following research question: what is the impact of Adversarial ML-based attacks against V2M services if an adversary has limited knowledge of the V2M service? The motivation is to have better understanding of the vulnerabilities in a V2M system that are likely to be exploited by an adversary. Knowing these vulnerabilities will help defend V2M services against Adversarial ML-based attacks by creating powerful mitigation systems as well as early-detection techniques.

4.1 Introduction

Smart microgrids rely on wireless communication devices to remotely connect with the existing power grid [210]. Wireless devices use data interfaces to communicate with other devices on a local power grid, such as substations and transformers. These interfaces make them susceptible to being hacked by sophisticated technology trying to modify the behavior

of the device or cause physical damage to it. Over the past few years, we have already seen the rise of malicious actors targeting smart grids and EVs' infrastructure [211].

This chapter studies the impact of potential attacks against V2M services built upon adversarial machine learning. It is important to emphasize that we will not use α as the adversary's targeting point. However, we study the attack from the microgrid's side (i.e., smart meters). This attack consists of multiple stages, where the adversary first launches an inference attack to train a surrogate model to perform an evasion attack. The inference attack uses a deep neural network that generates synthetic requests which are indistinguishable from real microgrid requests. The adversary uses a unique type of generative model, namely a Conditional Generative Adversarial Network (CGAN) [212], that is trained to learn to synthesize data samples that are statistically similar to real data samples. To the best of our knowledge, for the first time, this chapter studies a gray-box evasion attacks in V2M settings. It is important noting synthetic data generation using simpler methods such as Gaussian Mixture Model (GMM) cannot produce high-quality synthetic data for adversarial training purposes. These methods fail to capture complex data distributions. Hence, we decided to implement a powerful AI model such as CGAN model for data generation.

4.2 Threat Model

To carry out an adversarial attack, one must leverage pre-existing vulnerabilities within the targeted system. A vulnerability is identified as a defect within the system that can be exploited by outside parties [213]. Within the context of this system, the total energy usage of each consumer is measured by smart meters, which then transmit this data to electric power utility companies (EPUs). EPUs may store these data either locally or on cloud-based platforms. The literature often assumes vulnerabilities such as manipulating the data from a single or multiple devices like smart meters [214], [215], attacking the communication networks [216], [217], or directly hacking into EPUs' control/data centers to alter the stored data [218], [219]. By taking advantage of any of these vulnerabilities, adversaries can initiate adversarial attacks. The goal of the adversary is to deliberately create stealthy perturbations that will be introduced to smart meter readings through these vulnerabilities to compromise ML models [220], [221]. The depicted threat model in Figure 4.1 demonstrates the transmission of smart meter data from its origin (e.g., a consumer's home) to the electric power utility (EPU), including various potential attack points. Adversaries, operating remotely, are capable of exploiting vulnerabilities in both communication protocols and software infrastructures. Non-Intrusive Load Monitoring (NILM) is commonly used to analyze smart meter data to facilitate decision-making.

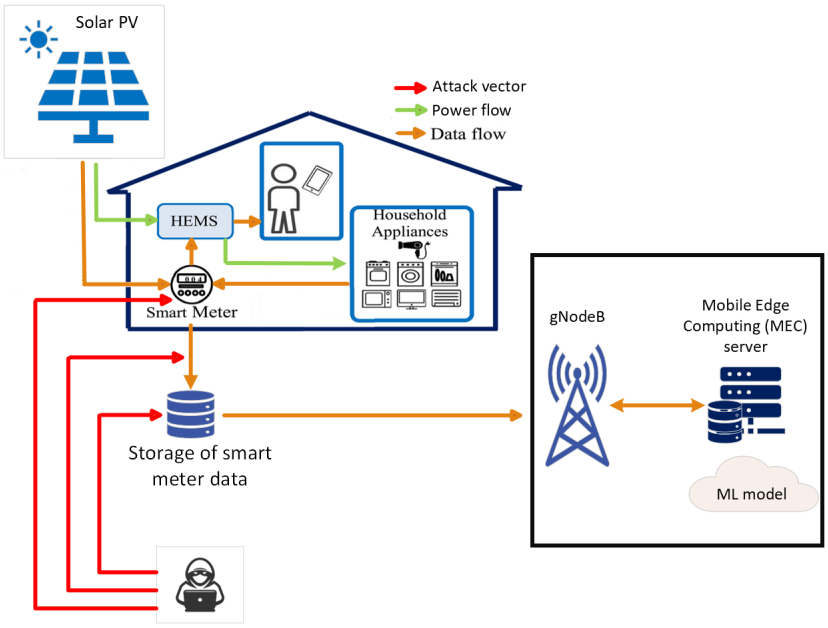


Figure 4.1: The threat model under study in V2M system

However, adversaries aim to manipulate NILM outputs, resulting in inaccurate decisions within the power system. Hence, the attack under study has two distinct impacts. Firstly, it negatively affects demand response initiatives. NILM plays a crucial role in assessing consumer flexibility for these programs [222]. EPUs send management signals for demand to consumers and control flexible devices identified through NILM [223][224]. Demand response aims to lower peak system demands through load reduction and the rescheduling of appliances. However, inaccurate data about appliance status leads to EPUs wrongly estimating the grid's load flexibility and issuing incorrect control signals, causing supply-demand imbalances and potentially overloading distribution networks, which might lead to cascading failures. Secondly, the attack impacts Home Energy Management Systems (HEMS) that rely on NILM for scheduling appliance use in homes to minimize energy consumption [225] during peak times, causing these systems to function poorly. In the context of V2M, this causes NILM to inaccurately report the smart microgrids electrical needs, resulting in shifting the priorities of the dispatched EVs to the affected smart microgrids. In this chapter, we consider the second impact as it is directly related to the studied V2M scenarios. In addition, we consider the second impact as the adversary's goal in the threat model.

To launch successful adversarial attacks, we assume that the adversary can gain partial or complete possession of the victim's ML model's training dataset. Adversaries can obtain

training datasets through various methods. For example, data breaches, notably prevalent in the energy sector, offer one avenue for adversaries to access these datasets [226]. For instance, 35% of attacks that targeted the energy industry were attempted data theft and leaks in 2021 [227]. In addition, the energy sector faced significant challenges in 2023, marked by a 90% rise in data breaches [228].

4.3 Anticipation of Adversarial ML-based Attacks Against V2M Services

In this section, we present an anticipatory framework to assess the impact of Adversarial ML-based attacks that can be potentially launched against V2M services. In consideration of the involved components, the following two assumptions are made:

(i) An adversary in a V2M setting can directly obtain the output (i.e., classification label) of the targeted machine learning classifier at the mobile edge.

(ii) An adversary can manipulate the input data (i.e., observations/features) of the targeted machine learning classifier, and perform a data integrity attack, as demonstrated in our previous work [41]. In addition, the authors in [182] showed that data integrity attacks are likely to occur by exploiting vulnerabilities of Advance Meter Infrastructure (AMI). The authors presented a threat model with four different falsification modes where an adversary can change and manipulate AMI's inputs. We explained the feasibility of the assumption in 2.

Before discussing the potential impact of AML-based attacks, it is worth covering the interaction among the three main entities in a V2M setting, namely EVs, smart microgrids and gNodeB. Smart microgrids, that are predicted to experience a power outage, send service requests to the gNodeB. The machine learning model at the gNodeB classifies the incoming requests into high-priority, medium-priority or low-priority requests. The three priority groups represent the level of energy necessity for microgrids, taking into consideration power generation and power consumption factors. The requests are labeled based on the input features sent by the smart microgrids [10]. The gNodeB broadcasts the request to the EVs within its coverage range. The EVs respond to the request with the following: (i) selling price, (ii) contribution percentage of the EV's battery, and (iii) current location. Then, the gNodeB calculates distances between the microgrids and the EVs, and chooses an optimal set of EVs to serve each microgrid's request. Lastly, the gNodeB sends the trajectories to the EVs [8].

In the following subsections, we discuss Adversarial ML-based attacks, and their use of CGAN to generate synthetic data in the context of V2M settings.

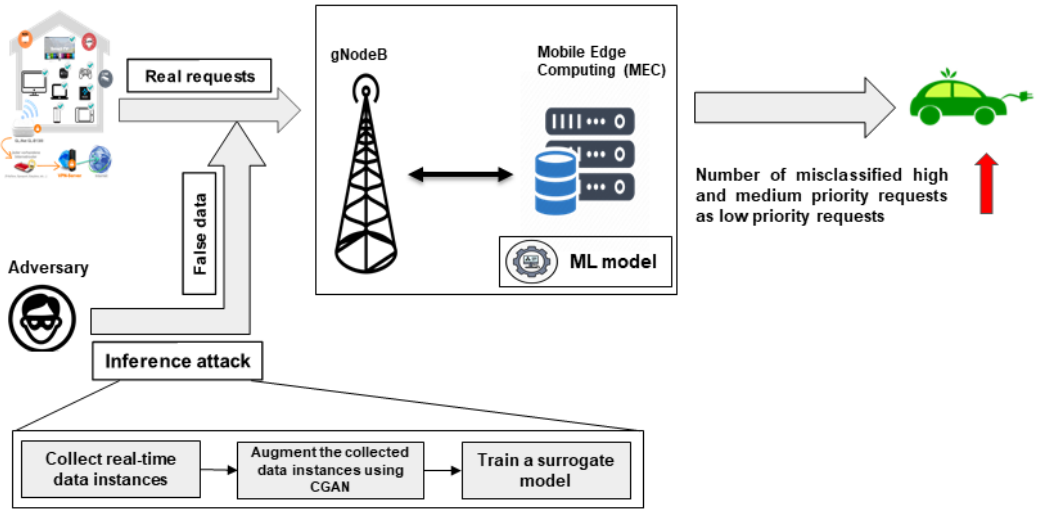


Figure 4.2: An overview of anticipated attacks on V2M services

4.3.1 Inference and Evasion Attacks Against V2M Services

The anticipated inference attack launched against the discussed V2M framework consists of three steps as depicted in Figure 4.2. First, an adversary collects sufficient real-time data instances (i.e., observations) and output data (i.e., priority labels). It is shown that microgrids' communication systems are susceptible to cybersecurity attacks which can lead to data leaks [229]. Collecting enough observations is a difficult task for the adversary. Hence, the second step is using the CGAN model to generate synthetic data which, with addition to the observations, builds a complete dataset. Third, the adversary trains different surrogate models on the complete dataset, and the model with the best performance, specifically accuracy and F1-score, is selected as the optimal surrogate model. With the help of the surrogate model, the adversary launches an evasion attack where they inject adversarial samples into the microgrid's real requests. As a result, the operating ML classifier at the mobile edge classifies the high and medium priority requests as low priority requests.

4.3.2 Conditional Generative Adversarial Network

To augment the collected observations (i.e., real-time data instances), the adversary uses a CGAN to generate synthetic data. However, with too few collected observations, the CGAN cannot accurately model the distribution of the real data. Hence, we investigate the impact of different amounts of collected observations on the CGAN performance. This paves the way for unveiling the potential vulnerabilities of a protected V2M system that can still be exploited, and leads to solid guidelines for the design of resilient techniques to mitigate the anticipated damage with possible prevention of adversarial attacks.

Generative Adversarial Networks (GANs) can be used to produce synthetic data that resembles the real data input to the networks. CGANs can use data labels during the training process to generate data belonging to specific categories. Additional information that is correlated with the input data, such as class labels, can be used to improve GAN performance. For instance, synthetic data can be generated in a better quality while maintaining more stability or faster training. CGANs are trained in such a way that both the generator and the discriminator models are conditioned on the class label, so that when the trained generator is used as a standalone model to generate samples in the domain, samples of a given type can be generated. CGANs consist of two networks that train together as adversaries:

Generator network - Given a label and random array as input, this network generates data with the same structure as the training data observations corresponding to the same label. The objective of the generator is to generate synthetic-labeled data that the discriminator classifies as "real."

Discriminator network - Given batches of labeled data containing observations from both training data and generated data from the generator. The discriminator tries to classify the instances as "real" or "synthetic". The objective of the discriminator is to not be deceived by the generator when given batches of both real and synthetic (i.e., generated) labeled data. The entire process corresponds to a minimax game played between D and G as stated in Eq. (4.1) [230]:

$$\min_{G}\max_{D}\ \mathbb{E}_{x,y\sim p_{data}(x,y)}log[D(x,y)] + \mathbb{E}_{y\sim p_Y(y),z\sim p_Z(z)}log[1-D(G(z,y),y)] \tag{4.1}$$

,where y represents the extra information (i.e., labels) drawn from a distribution $p_Y(y)$.

4.4 Performance Evaluation

In this section, we present a detailed description of the dataset and simulation settings employed in our study. We discuss the various parameters and attacks' configurations used to evaluate the performance of our approach. Furthermore, we analyze the obtained results, providing comprehensive insights into the effectiveness and limitations of our proposed method.

4.4.1 Experiment Setup

To assess the impact of inference and evasion attacks on V2M services, MATLAB Deep Learning and Simscape Electrical toolboxes are used. All simulations are performed using Intel Core i7-10700F CPU with 32GB of RAM and NVIDIA GeForce GTX 1660 SUPER GPU running on a Windows 10 system. The Deep Learning Toolbox is used to run the CGAN model, whereas the Simscape Electrical Toolbox simulates the EVs [209]. The results of this chapter onwards are the average of 10 runs with a confidence interval of 95% as shown in 4.2, where $\bar{x}$ is the mean, SD denotes the standard deviation, and N denotes the number of samples.

$$error = \bar{x} + \frac{t * SD}{\sqrt{N}} \tag{4.2}$$

Dataset preparation and mobile edge-based classifier

We use the iHomeLab RAPT dataset to run our experiments [231]. The dataset consists of electrical power consumption data as well as power generation data for five residential households in Switzerland. The power consumption data can be found in two resolutions: appliance-tier data and aggregated household-tier data. The power generation data comes from Photovoltaic panels (PVs). The data was collected in a period of 1.5 to 3.5 years, with different sampling frequency for each household. In this chapter, we chose one residential household from the iHomeLab RAPT dataset with a sampling frequency of 10 minutes. In addition, we supplement the iHomeLab RAPT dataset with a previously used dataset in [41]. The supplement dataset comprises of a power generation module. The power generation module, which uses a wind farm and back-up batteries as alternative resources of power, contains power generation data such as wind farm's generation profile and the capacity of the back-up batteries.

We use K-means, an unsupervised machine learning algorithm, to cluster the aggregated dataset (D), which consists of iHomeLab RAPT and our own dataset, into three priority groups, namely low-priority, medium-priority and high-priority. Each one of the three priority groups represents the level of energy requirement for the microgrids; the priority groups take energy generation and power consumption factors into consideration. Hence, K-means clusters the dataset based on the power consumption and the energy generation profiles, as in [13]. After clustering and labeling the aggregated dataset, we use the dataset (D) to train a K-NN model to operate as an ML classifier at the network edge. K-NN was selected as it outperforms other ML classifiers that were trained on the labeled dataset, such as Support Vector Machine (SVM) and Logistic Regression (LR).

Inference attack's settings

As previously mentioned, the inference attack starts with an adversary collecting real-time data instances to build a training dataset. To provide a comprehensive analysis, we compare five cases, where an adversary has access to different amounts of data. For the white-box attack, the adversary has full access to the labeled (training) dataset (D) of the operating ML classifier at the edge. For the gray-box attack, the adversary can collect limited real-time data instances from the communication between the microgrids and gNodeB, namely 1,600, 1,200, 800 and 400, which correspond to 80%, 60%, 40% and 20% of the training dataset's size.
In the gray-box attack, after obtaining the data instances (i.e., data observations), the adversary complements the few collected observations using CGAN to generate a complete dataset. To perform a fair comparison, we assume that CGAN generates enough data to complement the missing percentage of the dataset, such that the adversary will always have the same dataset's size in all cases. For instance, if the adversary collects 1,800 data instances (i.e., 80% of the training dataset's size), CGAN is used to generate the remaining 200 data instances (i.e., 20% dataset's size), as illustrated in Table 4.1. For the white-box attack, where the adversary has 100% access to the training set, the data augmentation step is skipped. The CGAN's architecture is depicted in Figure 4.3 where the generator and the discriminator are explained in detail. A brief description of the generator operations can be summarized in the following four points: projecting and reshaping the random noise, converting the categorical labels to embedding vectors, concatenating the outputs of the random noise and labels, and upsampling the concatenated arrays using a series of transposed convolution, batch normalization and Rectified Linear Unit (ReLU) layers. The discriminator operates similarly except for the last point, where the discriminator downsamples the concatenated arrays using a series of convolution layers with leaky ReLU layers.

The last part of the inference attack is to choose a surrogate model that captures the underlying functionalities of the targeted classifier at the mobile network edge. Six different models are trained on the complete dataset (i.e., real and synthetic data). The six models that we consider are Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), K-NN, SVM and Naive Bayes (NB). The models are trained and tested on the 80% and 20% partition of the complete dataset, respectively. The model with the best accuracy and F1-score performance is selected as the surrogate model. It is worth mentioning that the test set (i.e., 20%) consisted of the real-time data instances collected by the adversary, not data generated by the CGAN. Hence, this allows us to assess the capability of a CGAN model to capture the distribution of real-time data.

4.4.2 Evaluation Metrics

In this section, we introduce different metrics to assess the potential damage of the inference attack and the evasion attack on V2M services. The impact of the inference attack is assessed based on the surrogate's model performance, particularly on the testing data. high performance of the surrogate model demonstrates the effectiveness of the CGAN's model to generating high-quality data that mimics the statistical distribution of real-time data. We evaluate the six models in terms of accuracy $((TP + TN)/(TP + FP + TN + FN))$ and F1-score $(2 \cdot (Precision \cdot Recall)/(Precision + Recall))$ where $Precision = \frac{TP}{TP+FP}$ and $Recall = \frac{TP}{TP+FN}$. TP, TN, FP and FN stand for True Positive, True Negative, False Positive and False Negative, respectively.

The anticipated impact of the evasion attack is determined based on the adversary's ability to deceive the K-NN classifier at the gNodeB such as the number of high-priority and medium-priority requests that are falsely classified as low-priority requests. Two metrics are considered to evaluate the evasion attack: Evasion Increase Rate (EIR) and Adversarial Detection Rate (ADR) (i.e., True Negative Rate (TNR)). The ADR measures how well the system identifies adversarial examples correctly. EIR quantifies the increase in evasion capability of adversarial inputs compared to normal inputs, highlighting the degradation in a system's ability to detect adversarial examples.

$$ADR = \frac{TN}{TN + FP}$$

$$EIR = 1 - \frac{TNR_{adversarial}}{TNR_{original}}$$

The $TNR_{adversarial}$ defines the adversarial detection rate, whereas $TNR_{original}$ defines the original detection rate of the classifier before the attack. It is worth noting that traditional GAN metrics such as Inception Score is designed to evaluate the general quality and diversity of the generated data, which does not translate to effectiveness in adversarial contexts. In addition, Maximum Mean Discrepancy (MMD) and Perceptual Path Length focus more on statistical or perceptual similarity and continuity, respectively. They do not provide information about the adversarial strength or the security aspects of the generated data, which are crucial in adversarial contexts. While the quality of the generated data can be high according to the traditional GAN metrics, it is not guaranteed that the generated data would bypass the detection model. However, ADR and EIR metrics are guaranteed to measure the impact of the launched adversarial attacks.

4.4.3 Numerical Results

We first start with the numerical results of the inference attack. Case-A is the white-box attack, where the adversary has access to 100% of the training dataset of the K-NN classifier at the mobile network edge. Cases B, C, D and E correspond to 80%, 60%, 40% and 20% of real-time data instances, respectively.
As the adversary has access to the complete training dataset in case-A, the synthetic data generation using CGAN is skipped, as shown in Table 4.1. In all other cases, the CGAN is built using the available data instances collected by the adversary. After the CGAN converges, the CGAN's generator is used to generate data, which is used to complement the missing percentage of the collected dataset, creating a final combined dataset, consisting of collected and generated data. We ensure that all cases have a combined dataset of the same size. The combined dataset is used to train and test the surrogate models. To guarantee fair comparison, all surrogate models are trained and tested under 80% and 20% of the combined dataset, respectively. In addition, we use a balanced dataset, and the results of the surrogate models represent the average of 10 different runs.

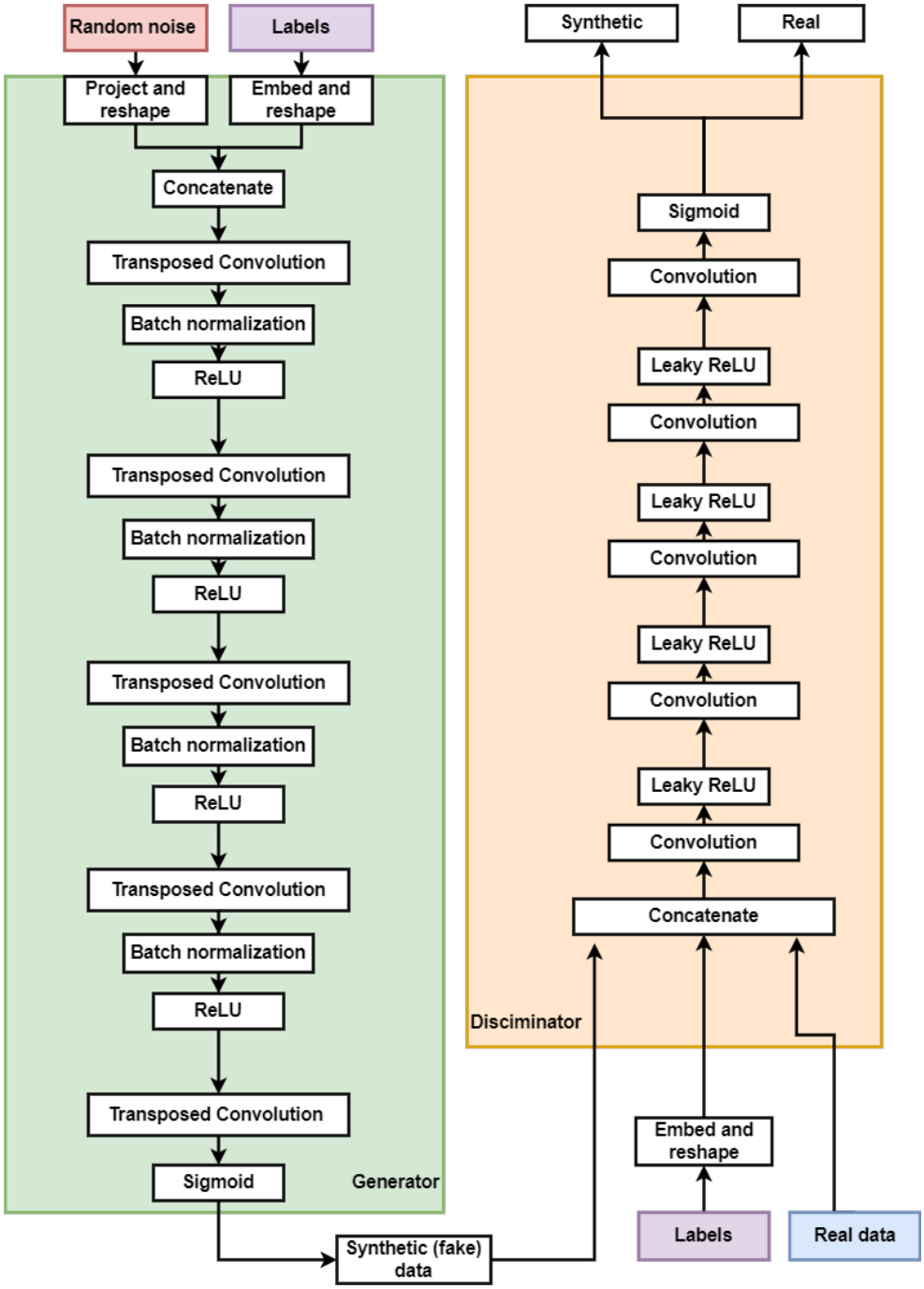


Figure 4.3: Generator and discriminator architectures of CGAN

Table 4.1: Description of the combined dataset

Case		Combined dataset	
		Adversary's contribution to the dataset (real-time collected data)	CGAN's contribution to the dataset (generated data)
White-box	A	100%	0%
Gray-box	B	80%	20%
	C	60%	40%
	D	40%	60%
	E	20%	80%

Table 4.2 presents the anticipated performance of the surrogate model as a part of an inference attack against V2M services. In case-A, case-B and case-D, the adversary is able to choose a surrogate model similar to the target edge classifier (i.e., K-NN). For case-A, based on 98.35% accuracy and 97.5% F1-score, K-NN is selected as the surrogate model. Similarly, in case-B, the adversary was again able to select the same operating classifier at the edge. Even with decreasing the available real-time data instances from 100% to 80%, K-NN can still achieve 95.46% accuracy and 94.9% F1-score.
In case-C, with 60% access to training data, SVM outperforms the other models and serves as the surrogate model. It is important to note that despite SVM getting selected as the surrogate model for its accuracy of 89.88% and an F1-score of 90%, K-NN performed slightly lower than SVM with an accuracy of 89.65% and F1-score 89.5%. In case-D, although the adversary only accesses 40% of the training dataset, the performance in terms of accuracy and F1-score are acceptable, yielding 81.4% and 82.3%, respectively. The K-NN is selected as the surrogate model which matches the targeted classifier. In case-E, with 20% collectable data instances, the performance drops significantly, leading to 70.15% accuracy and 69% F1-score under LR as the surrogate model. It is noted that in case-E, the number of collected real-time data instances are insufficient for the CGAN model to capture the underlying statistics of the data. Thus, when the LR classifier is tested on real-time data instances, it demonstrates low performance as compared to the other cases. In addition, K-NN is not selected as the surrogate model because of the insufficient number of data instances. It is observed that as the adversary has access to a smaller partition of the real-time dataset, the performance of the surrogate model decreases. It is worth noting that the accuracy and F1-score drop is logarithmic, not linear. For instance, between case-A and case-B, the accuracy decreases by 3% and the F1-score decreases by 4%; whereas, between case-D and case-E, the accuracy decreases by 11% and the F1-score decreases by

Table 4.2: Anticipated accuracy and F1-score of the surrogate models as part of inference attack (when using CGAN in cases B-E)

Case		Selected surrogate model	Performance	
			Accuracy	F1-score
White-box	A	K-NN	98.35%	97.5%
Gray-box	B	K-NN	95.46%	94.9%
	C	SVM	89.88%	89.43%
	D	K-NN	81.4%	80.9%
	E	LR	70.15%	69.2%

12%.

Results of the evasion attack are presented in Table 4.3. Performance of case-A is significantly high with an ADR of 2.4% and an EIR of 97.5%. The reason behind that high impact of damage is because of the adversary's complete access (i.e., white-box) to the training dataset of the K-NN classifier at the edge. The performance starts dropping for the gray-box attacks as the adversary collects limited real-time data instances. For instance, in case-B, although only 80% data instances were collected, we notice that the adversary is able to deceive the gNode's K-NN classifier with ADR of 4.8% and EIR of 95.1%. However, the anticipated damage of the attack decreases significantly for case-C with an ADR of 20.6% and an EIR of 79.3%. This significant decrease is due to the mismatching between the selected surrogate model (i.e., SVM) and the operating classifier at the network edge (i.e., K-NN). In addition, the boundary regions of the SVM for the low-priority requests are different than the K-NN, which might misguide the adversary about the feature values that lead to incremental false positives. On the contrary, in case-D, as the selected surrogate model matches the operating classifier, we observe a less significant decrease in the ADR and EIR results with 26.5% and 73.2%, respectively. Thus, the adversary is able to perform an impactful evasion attack by collecting only 40% of the real-time data instances and generating 60% synthetic data using CGAN. However, a significant decrease occurs under case-E due to two reasons. First, with only 20% of collected real-time data instances, the CGAN is unable to generate high quality synthetic data as it cannot capture the statistical distribution of the real-time data. Second, since the decision boundaries of the selected surrogate model (i.e., LR) and the operating K-NN classifier are different, the adversary cannot design adversarial perturbed samples that can successfully deceive the classifier. In Table 4.4, we show the potential impact of the evasion attack without using CGAN. Thus, without using CGAN, the impact of the ADR and EIR is reduced by at least 7% when compared to not using CGAN. In addition, we notice that the EIR of case-D with CGAN (73.2%) is higher than the EIR of case-C without CGAN

Table 4.3: Anticipated ADR and EIR of evasion attack against V2M services (when using CGAN in cases B-E)

Case		Selected surrogate model	Performance	
			ADR	EIR
White-box	A	K-NN	3.4%	96.5%
Gray-box	B	K-NN	4.8%	95.1%
	C	SVM	20.6%	79.3%
	D	K-NN	26.5%	73.2%
	E	LR	43.5%	56.4%

(69.9%). In other words, augmenting 40% of real-time data instances (case-D) with CGAN results in an EIR of 73.2%, whereas 60% of real-time data instances (case-C) results in an EIR of 69.9% when not using CGAN. Hence, an adversary can potentially pose more damage to V2M services when collecting fewer real-time data instances and augmenting them with CGAN.

Table 4.4: Anticipated ADR and EIR of evasion attack against V2M services (without using CGAN)

Gray-box Case	Performance	
	ADR	EIR
B	11.7%	88.1%
C	29.9%	69.9%
D	40.7%	59.2%
E	59.3%	40.5%

4.5 Summary

In this chapter, we presented an anticipatory study of a multi-stage adversarial machine learning attack in Vehicle-to-Microgrid (V2M) settings, consisting of an inference attack and an evasion attack. Knowledge of the adversary has been taken into consideration as we quantify the impact of the collectable real-time data instances. We have compared five different cases of combined dataset, one case under white-box setup and four cases under gray-box setup. In the white-box case, the adversary has 100% access to the training dataset of the K-NN classifier; whereas, in the gray-box cases, we have assumed that the adversary collects different amounts of real-time data instances, and the rest of the dataset

is complemented using a CGAN model. Through simulations, we have presented the selected surrogate model for each case based on the accuracy and F1-score performance. As the number of collected data decreases, the surrogate model's performance drops logarithmically. In addition, we have shown that an adversary is able to deceive the K-NN classifier at the edge, achieving high ADR and EIR when the selected surrogate model is K-NN, matching the operating classifier. Moreover, we have anticipated that with 40% collected data, the potential damage of evasion attack results in 26.5% ADR and 73.2% EIR. We have concluded that, with the lack of powerful mitigation techniques, Adversarial ML-based attacks on V2M services can significantly damage the system.

Chapter 5

Detection of Adversarial Machine Learning Attacks Against Vehicle-to-Microgrid services

As the use of AI becomes more prevalent in the electrical power sector, there has been a rise in the number of adversarial attacks targeting ML models. Therefore, it has become imperative to safeguard ML models against such attacks. Our analysis shows that adversaries aim to deceive the victim's ML classifier at the network edge to misclassify the incoming energy requests from microgrid users. In this chapter, we introduce an AI-powered framework to detect new instances of AML attacks against V2M systems. The proposed framework uses Generative Adversarial Network (GAN) model and ML classifiers to accurately detect novel AML instances.

5.1 Introduction

Transactive energy is a concept that involves the use of communication networks, market-based mechanisms, and information technology to balance electricity supply and demand in an electric power system. This innovative approach aims to optimize the use of available energy resources, reduce energy costs, and improve the reliability and stability of the power grid [6]. Through the use of smart grids, Advanced Meter Infrastructure (AMI) and communication networks, transactive energy systems allow for real-time exchange of information and electricity among prosumers. For instance, smart grids use communication technologies such as cellular networks and Wi-Fi to monitor the energy demand and

supply in real-time [8]. The use of communication networks in transactive energy systems also enables new business models to emerge, such as peer-to-peer energy trading. Communication networks also facilitate the integration of renewable energy sources, such as solar and wind power, into the power grid. This is critical for reducing carbon emissions and achieving a more sustainable energy system [9]. However, The rising need of using transactive energy networks has paved the way for new attack vectors to emerge.

The goal of this chapter is to better understand the vulnerabilities in V2M systems that are likely to be exploited by an adversary, with the aim of developing effective mitigation systems and early-detection techniques to defend against adversarial attacks. In our recent work [232], we have demonstrated how susceptible V2M systems are to adversarial attacks, especially when adversaries use Conditional Generative Adversarial Network (CGAN) model to increase the impact and success rate of their attacks. In this chapter, we implement the detection phase by investigating the effectiveness of DBSCAN as an AI-based countermeasure and proposing a GAN-based solution as a new strategy for defending against adversarial attacks.

5.2 Methodology

5.2.1 AML-based attacks against V2M services

In this section we present a brief overview of the AML-based attack scheme that was initially studied in our previous work [232]. The adversary's aim is to misclassify the incoming requests to the ML model at the mobile edge. To this end, the adversary launches an inference attack followed by an evasion attack. The inference attack aims at creating a surrogate model of the victim's ML model (i.e., the ML model resides at the mobile edge). The inference attack consists of three steps. First, the adversary builds a dataset to train the surrogate model. In this chapter, we assume that the adversary can either have partial access or full access to the dataset of the victim's ML model. Second, in case of the partial access to the dataset, the adversary uses a CGAN model to generate synthetic data instances to augment the partial dataset. Hence, the adversary creates a complete dataset that is a combination of synthetic data and real data. Notations used in this section are summarized in Table 5.1 while the terms and their corresponding definitions are presented in Table 5.2.

Table 5.1: Notation and definition used in this work.

Notation	**Definition**
O_d	The training dataset of ML classifier operating at the mobile edge
α	A hyper-parameter used in the GAN model to determine the learning rate (i.e., step size) of the SGD model
G_{params}, D_{params}	Define the weights and biases parameters of the generator and discriminator networks, respectively
L_G, L_D	Define the loss function of the generator and discriminator networks, respectively
∇	A gradient operator
x	A data instance
z	A random noise input to G()
G(z)	The output of the generator network
D(x)	The output of the discriminator network

Table 5.2: Terms and definition used in this work.

Terms	**Definition**
Legitimate requests	Requests sent by microgrids
Illegitimate requests	Requests sent by an adversary
GAN	A deep learning model used as a detection method in the detection phase
DBSCAN	An unsupervised ML algorithm used as a detection method in the detection phase
CGAN	A deep learning model used by the adversary for data augmentation in the attack phase
Classifier-1	A ML classifier used at the mobile edge as part of the detection phase and trained on the GAN output to detect AML attacks (i.e., distinguish between legitimate and illegitimate requests)
Classifier-2	A ML classifier operates at the mobile edge and targeted by adversaries with their AML attacks

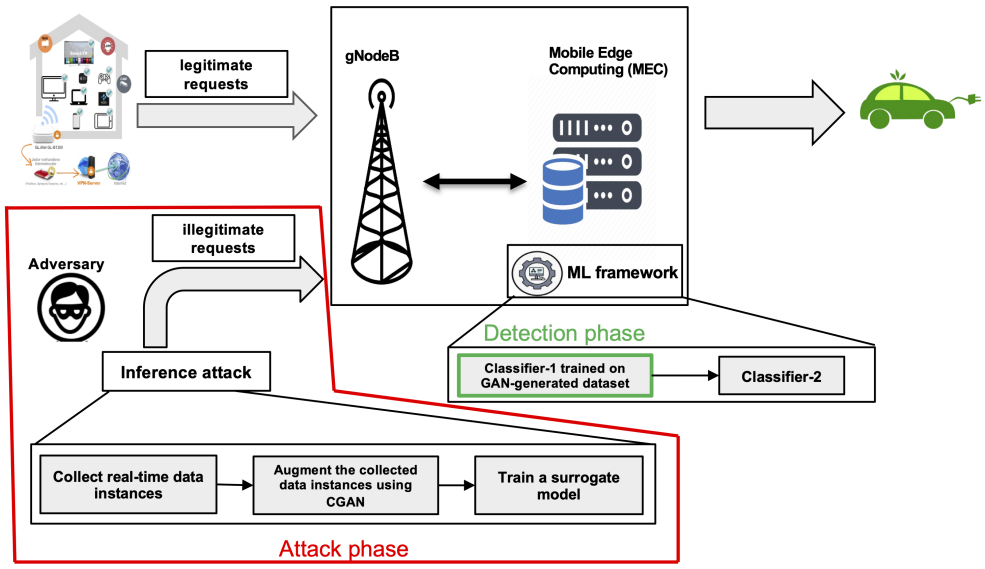


Figure 5.1: An overview of the anticipated attacks and the proposed detection system

5.2.2 Proposed solution: GAN-based detection technique

A GAN is made up of two parts: a generator network and a discriminator network, as shown in Figure 5.2. The generator network is responsible for generating new data instances whereas the discriminator network is responsible for determining whether a given data instance is real or synthetic (i.e., fake). The two models are trained simultaneously, with the generator network trying to produce data that is indistinguishable from real data, and the discriminator network trying to identify synthetic data. The competition between the two models ultimately leads to the generation of high-quality, synthetic data [230]. The training process can be summarized by the following steps:

1. The generator network takes in a random noise vector as input and generates a synthetic data sample.

2. The discriminator network receives both the fake data sample and a real data sample from the training dataset and attempts to classify them as either real or fake.

3. The generator and discriminator networks are trained using backpropagation and gradient descent, with the objective of minimizing a loss function. The loss function for the generator network is defined as the cross-entropy loss between the fake data samples and the labels "real", while the loss function for the discriminator network

is defined as the cross-entropy loss between the real and fake data samples and the labels "real" and "fake", respectively.

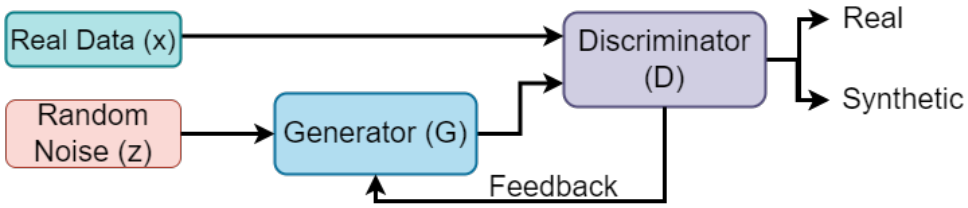


Figure 5.2: An overview of the GAN model

Generator network - the generator network in a GAN is typically a neural network that learns to map from a latent space to the data space. This means that it takes a random noise vector as input, z, and produces the output as a synthetic data instance, $G(z)$, that resembles the real data. The objective of the generator network is to generate data that is indistinguishable from real data, so the loss function for this model is typically defined as the cross-entropy between the generated data and the real data. This loss function can be expressed mathematically as in Eq. 5.1:

$$L_G = -E[log(D(G(z)))] \tag{5.1}$$

where $D(x)$ is the output of the discriminator network for a given data instance x, $G(z)$ is the output of the generator network for a given latent variable z, and $E[.]$ is the expectation operator. The parameters of the generator network, denoted by G_{params}, can then be updated using stochastic gradient descent according to the following rule in Eq. 5.2 [233]:

$$G_{params} = G_{params} - \alpha * \nabla(L_G, G_{params}) \tag{5.2}$$

where α is the learning rate which determines the step size for updating the network parameters during training. G_{params} represents the weights and bias parameters of the generator network.

Discriminator network - the discriminator network in a GAN is also a neural network that learns to distinguish between real and fake data. It takes as input a data instance, x, and produces as output a value, D(x), that indicates the probability that the input is a real data instance (i.e., D(x) = 1) or a fake data instance (i.e., D(x) = 0). The loss function for the discriminator network is typically defined as the binary cross-entropy between the predicted labels (real or fake) and the true labels. This loss function can be expressed mathematically as shown in Eq. 5.3.

$$L_D = -E[log(D(x))] - E[log(1 - D(G(z)))] \tag{5.3}$$

where $D(x)$ is the output of the discriminator network for a given data instance x, $G(z)$ is the output of the generator network for a given latent variable z, and $E[.]$ is the expectation operator. The parameters (i.e., weights and biases) of the discriminator network, denoted by D_{params}, can then be updated using stochastic gradient descent according to the following rule in Eq. 5.4.

$$D_{params} = D_{params} - \alpha * \nabla(L_D, D_{params}) \tag{5.4}$$

In this chapter, we propose using a GAN model to defend V2M services against AML attacks particularly, when an adversary uses CGAN model to strengthen their attacks, ash shown in Figure 5.1. We use a GAN model at the mobile edge to create a high-quality synthetic dataset to enhance the original dataset (O_d) which is used to train classifier-2. Consequently, we use the enhanced dataset to train classifier-1. Ultimately, classifier-1 should be able to distinguish between legitimate requests, that are sent by microgrids, and illegitimate requests, that are sent by an adversary. Upon the completion of the classification process at classifier-1, requests that are identified as legitimate will only pass to classifier-2 to be classified as low, medium or high priority requests. It is worth noting that a GAN model operates similar to a CGAN model with one main difference. That is, a CGAN model generates synthetic data that is conditional to a specific class. Hence, we use a GAN model at the mobile edge to defend all classifier-2 classes as we cannot anticipate the class the adversary may target with their attacks.
In Algorithm 1, we explain our detection technique. The algorithm provided outlines a robust adversarial training strategy that uses a Generative Adversarial Network (GAN) to augment a training dataset for a binary classifier. The GAN is first trained on a set of clean (legitimate) data samples until the generator within the GAN is proficient in creating new data samples that closely mimic the original clean data. These synthetic samples are of high quality, meaning they are realistic enough to be considered similar to actual clean data by the discriminator component of the GAN. Once the GAN is adequately trained, it produces synthetic clean samples which are then combined with the original training data to create an expanded training dataset. This new dataset has the benefit of greater diversity and volume, potentially improving the robustness of the classifier being trained. The classifier, denoted as f_θ, with θ representing its learnable parameters, is then trained on this enriched dataset. The goal of the classifier is to correctly identify legitimate samples during its operation, effectively filtering out any non-legitimate data. To achieve this, the classifier undergoes several training iterations, with each iteration refining the model's parameters θ. This iterative process, determined by N_{iter}, is designed to incrementally optimize the classifier's ability to distinguish between legitimate and non-legitimate samples; if the sample is benign, the label is 0; otherwise, the label is 1. The result of this iterative training and optimization process is a single binary classifier model,

f_θ , that is expected to be robust against adversarial examples or anomalies in the data. By leveraging the GAN-generated samples for training, the model aims to have a stronger generalization capability and an enhanced resistance to adversarial attacks when deployed, allowing only clean, legitimate data to pass through.

Algorithm 1 GAN-based Detection Procedure

Input O_d: A training dataset,
N_{iter}: The number of iterations used for training
Output: One robust classifier f_θ,
where θ is a set of learnable parameters of the model f

1: Train a GAN model on the dataset until the generator converges
2: Generate a set of high-quality synthetic samples $X^{`}$ using the generator model
3: Combine the samples from O_d and generated samples from $X^{`}$
4: Train a binary classifier f_θ using the combined dataset. At inference time, the model will only pass legitimate data.
5: **for** 1, 2, 3, ... N_{iter} **do**
6: Optimize θ by training f_θ
7: **end for**
8: **Return** One binary classifier model f_θ allowing only legitimate data to pass.

5.3 Performance Evaluation

In this section, we present a detailed description of the experiment setup in our study. We discuss the various parameters and attacks' scenarios used to evaluate the performance of our proposed detection framework [234]. Furthermore, we analyze the obtained results, providing comprehensive insights into the effectiveness and limitations of our proposed method.

5.3.1 Experiment Setup

Inference and evasion attacks and the proposed defence strategy are simulated using MATLAB Deep Learning and Simscape Electrical toolboxes. All simulations are performed using Intel Core i7-10700F CPU with 32GB of RAM and NVIDIA GeForce GTX 1660 SUPER GPU running on a Windows 10 system. The Deep Learning Toolbox is used to

run CGAN and GAN models; whereas, the Simscape Electrical Toolbox simulates the EVs. Terminology used in this section is summarized in Table 5.3.

Table 5.3: Terminology used in this section.

Terminology	**Definition**
Scenario-1	The adversary targets the low-priority class to misclassifying the medium and high priority requests under low-priority requests
Scenario-2	The adversary aims at misclassifying the high-priority requests under low and medium priority requests
Scenario-3	Equal coexistence of both scenarios: scenario-1 and scenario-2.
Case-A	A white-box attack case in which an adversary can access 100% of O_d
Case-B, Case-C, Case-E, Case-D	Gray-box attack cases in which an adversary can access 80%, 60%, 40% or 20% of O_d, respectively
Attack phase - category 1	CGAN is disabled in the attack phase
Attack phase - category 2	CGAN is enabled in the attack phase
Detection phase - mode 0	No detection method is implemented in the detection phase
Detection phase - mode 1	DBSCAN-based detection is implemented in the detection phase
Detection phase - mode 2	GAN-based detection is implemented in the detection phase
True Positive (TP)	Legitimate requests predicted as legitimate
True Negative (TN)	Illegitimate requests predicted as illegitimate
False Positive (FP)	Illegitimate requests predicted as legitimate
False Negative (FN)	Legitimate requests predicted as illegitimate

Dataset description

The iHomeLab RAPT dataset contains data on the electrical power consumption and generation for five Swiss households [235]. The power consumption data is available at two levels of detail: per appliance and for the entire household. The power generation data comes from photovoltaic panels. The data was collected over a period ranging from 1.5 to 3.5 years, with different sampling frequencies for each household. For the purpose of this chapter, we selected one household from the iHomeLab RAPT dataset with a sampling frequency of 10 minutes. Additionally, we included a supplementary dataset previously used in [236] that includes power generation data from a wind farm and backup batteries. This supplementary dataset includes information on the wind farm's power generation profile and the capacity of the backup batteries. We apply K-means, an unsupervised machine learning algorithm, to divide the combined dataset (O_d) - consisting of the iHomeLab RAPT dataset and the supplementary dataset - into three energy requirement categories: low-priority, medium-priority, and high-priority. These categories, which consider energy generation and power consumption, reflect the energy needs of microgrids. Therefore, K-means clusters the dataset based on power consumption and energy generation patterns, similar to the approach described in [9]. After clustering and labeling the aggregated dataset, we use the dataset (O_d) to train a K-NN model to operate as an ML classifier at the network edge. K-NN was selected as it outperforms other ML classifiers that were trained on the labelled dataset such as Support Vector Machine (SVM) and Logistic Regression (LR).

Attack phase setup

We start the attach phase setup with discussion on the different scenarios and cases used in this work. Subsequently, we describe three additional attack techniques, demonstrating the effectiveness of CGAN-based adversarial attacks.

To thoroughly evaluate the results, we evaluate our proposed solution under three scenarios and five cases. In scenario-1, the adversary tries to compromise the integrity of the low-priority class by wrongly categorizing medium and high priority requests as low-priority requests. In scenario-2, the adversary tries to wrongly categorize high-priority requests as either low or medium priority. In scenario-3, both scenario-1 and scenario-2 are present. In addition, we compare five cases where the adversary has access to varying amounts of data observations. Case-A is the white-box attack, where the adversary has access to 100% of the training dataset of the K-NN classifier at the MEC. Cases B, C, D and E correspond to 80%, 60%, 40% and 20% access to the training dataset, respectively.

In the white-box attack scenario, the adversary has complete access to the labeled (i.e.,

training) dataset (O_d) used by the operating ML classifier at the mobile edge. For the gray-box attack, the adversary is able to gather a limited number of the data instances of the O_d dataset, specifically 4000, 3000, 2000, and 1000 instances, which correspond to 80%, 60%, 40%, and 20% of the size of the training dataset, respectively.
In the gray-box attack, after collecting a certain number of data instances, the adversary uses a CGAN to generate additional data in order to create a complete dataset. To ensure a fair comparison, we assume that the CGAN generates enough data to complement the missing portion of the dataset in all cases, so that the adversary always has the same size dataset. For example, if the adversary has access to 4,000 data instances (80% of the training dataset), the CGAN is used to generate the remaining 1,000 data instances (20% of the training dataset). In the white-box attack, where the adversary has full access to the training set, the data augmentation step is not necessary.

To evaluate the impact of CGAN adversarial attacks, we show the ADR performance under three additional evasion techniques, namely FGSM, BIM and C&W.

1. FGSM technique: This attack works by using the gradients of the neural network to create an adversarial example. For an input (x), the method add perturbation to (x) in the direction of the sign of the gradient of the cost function (J) with respect to the input to create a new (x') that maximizes the loss. The parameter ϵ controls the magnitude of the perturbation. This process is expressed as in Eq. 5.5.

$$x' = x + \epsilon \cdot \text{sign}(\nabla_x J(\theta, x, y)) \tag{5.5}$$

2. BIM technique: This attack is an extension of FGSM applied multiple times with small steps α. After each step, values of intermediate results are clipped to ensure that they are within an ϵ-neighborhood of the original input (x). This method generally produces more effective adversarial examples compared to FGSM. This process is expressed as in Eq. 5.6.

$$x'_{n+1} = \text{Clip}_{x,\epsilon}\{x'_n + \alpha \cdot \text{sign}(\nabla_x J(\theta, x'_n, y))\} \tag{5.6}$$

3. C&W technique: The C&W attack is a more complex and powerful attack compared to FGSM and BIM. It involves an optimization process that finds the smallest perturbation δ that can cause misclassification. The constant (c) is a regularization parameter that is determined through a binary search. The objective function f is designed to have its minimum when the classifier output is on the decision boundary. This process is given as in Eq. 5.7, where f is the objective function such that $x + \delta$ is misclassified.

$$\text{Minimize } \|\delta\|_2 + c \cdot f(x + \delta) \tag{5.7}$$

Detection phase setup

We start with defining our baseline solution and continue with the details of the proposed solution (i.e., GAN-based model). We use DBSCAN algorithm as a potential intelligent solution to define the baseline. DBSCAN is a clustering algorithm that is used to identify clusters, or groups, of data points in a dataset. It is a density-based algorithm, meaning that it looks for clusters of data points that are densely packed together and separates them from other data points that are less densely packed. DBSCAN works by first defining a distance (called Eps) between data points. It then looks for data points that are within this distance of each other, and considers these points to be part of the same cluster. If a data point has fewer than a specified number of other points within the Eps distance (called $MinPts$), it is considered to be an outlier and is not included in any cluster. DBSCAN then iteratively expands the clusters by looking for other data points that are within the Eps distance of the points already in the cluster. If a point is within the Eps distance of a cluster, it is added to the cluster. This process continues until all points have been either included in a cluster or deemed to be an outlier. One of the advantages of DBSCAN is that it is able to identify clusters of different shapes and sizes, and it can also handle data points that are noise (i.e., points that do not belong to any cluster). It is also relatively efficient, as it does not require the user to specify the number of clusters in advance. However, it is sensitive to the choice of Eps and $MinPts$, and these parameters must be carefully chosen in order to obtain meaningful results.
However, DBSCAN fails to detect intelligently designed AML attacks, particularly when an adversary uses CGAN for dataset augmentation. Therefore, we use a GAN model as an alternative to the DBSCAN algorithm. The architecture of the GAN model consists of a generator and a discriminator. Figure 5.3 depicts the GAN training performance as produced from MATLAB simulation. The GAN architecture is shown in Figure 5.4 and includes a generator and a discriminator. The generator takes in random noise and categorical labels and produces new data instances through a series of operations including projecting and reshaping the noise, converting the labels to embedding vectors, concatenating the noise and labels, and upsampling the concatenated arrays using transposed convolution, batch normalization, and Rectified Linear Unit (ReLU) layers. The discriminator operates in a similar way except that it downsamples the concatenated arrays using convolution layers with leaky ReLU layers. It is worth noting that a GAN model operates similar to a CGAN model with one main difference. That is, a CGAN model generates synthetic data that is conditional to a specific class.

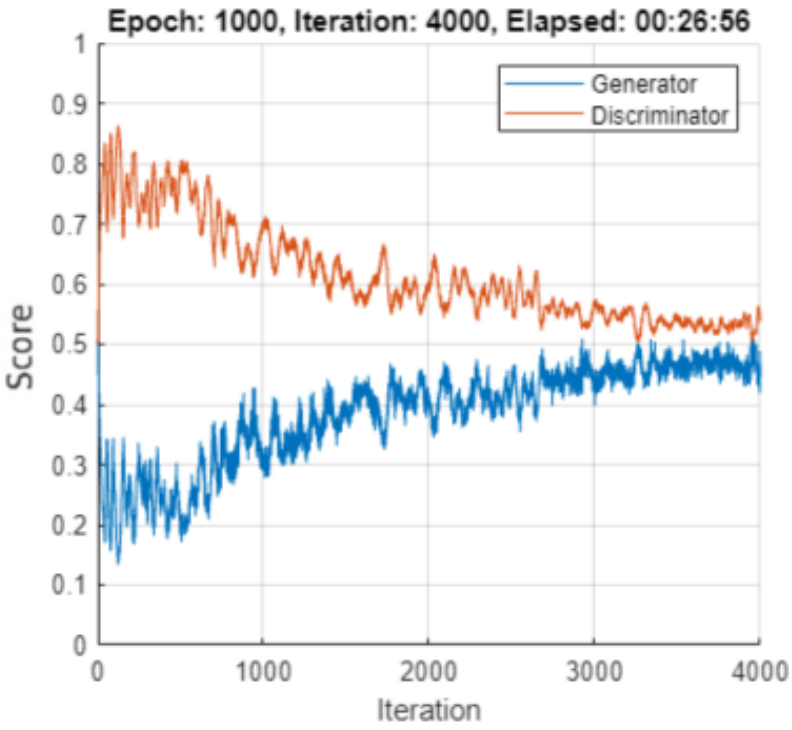


Figure 5.3: GAN training performance

In addition to DBSCAN and GAN model, we present the results of three types of detection models, namely 2-class SVM, LSTM and AAE. We provide an overview of each detection model.

1. (2-Class) SVM: This machine learning model excels in binary classification by constructing an optimal hyperplane to differentiate between authentic and adversarial examples. It's particularly effective when adversarial examples are close to the decision boundary, as it focuses on maximizing the margin between classes.

2. LSTM: LSTMs process input data in a sequential manner, making them effective in identifying non-obvious temporal anomalies indicative of adversarial interference, thus providing robustness in applications where data points are interdependent over time.

3. AAE: The Attentive Autoencoder uses its attention mechanism not just to detect the presence of adversarial noise, but to learn which data features are most vulnerable, improving its ability to guard against future adversarial manipulations.

5.3.2 Results

In this section, we demonstrate the ADR performance under two categories. The first category studies the AML attacks and countermeasures (i.e., DBSCAN) in the absence of CGAN in the attack phase. The second category studies AML attacks and countermeasures (i.e., DBSCAN or GAN) in the presence of CGAN for data augmentation in the attack phase.

Attack phase - category 1 (without CGAN)
Table. 5.4 presents the ADR performance under the three attack scenarios, the five access cases and three detection modes. Mode-0 is when the detection is off. That is, no detection method is used in the detection phase. Mode-1 is when we use DBSCAN as a detection method. All the results are obtained when the adversary does not use a CGAN model to strengthen their AML attacks.

Table 5.4: Assessing the ADR performance with and without implementing DBSCAN at the mobile edge. All the cases are presented in the absence of CGAN at the attack phase.

Access (%) to O_d		ADR (%) of the detection phase modes					
		Mode-0			Mode-1 (DBSCAN)		
		Scn-1	Scn-2	Scn-3	Scn-1	Scn-2	Scn-3
White-box	100% (case-A)	3.4%	3%	3.2%	3.8%	3.4%	3.6%
Gray-box	80% (case-B)	11.7%	9.2%	10.3%	13.5%	10.3%	13.3%
	60% (case-C)	29.9%	25.8%	27.7%	34.4%	29.3%	31.0%
	40% (case-D)	40.7%	35.5%	37.6%	48.1%	39.8%	42.1%
	20% (case-E)	59.3%	51.9%	54.8%	71.2%	58.1%	61.4%

In all scenarios, it is noticed that as the adversary's access to O_d becomes limited (i.e., case-A to case-E), the ADR percentage increases proportionally. This means an AML attack can be detected when the adversary has limited access to the training dataset, regardless of the implementation of a detection method. Furthermore, we observe the following remarks:

- Scenario-1 leads to the highest ADR performance in all access cases when compared to scenario-2 and scenario-3 in with/out DBSCAN by [0.2-13]%. That means it is easier to detect an AML attack that aims at misclassifying high- and medium-priority requests into low-priority than detecting an AML attack that aims at misclassifying low and medium priority requests into high-priority or the coexistence of both scenarios (i.e, scenario-3).

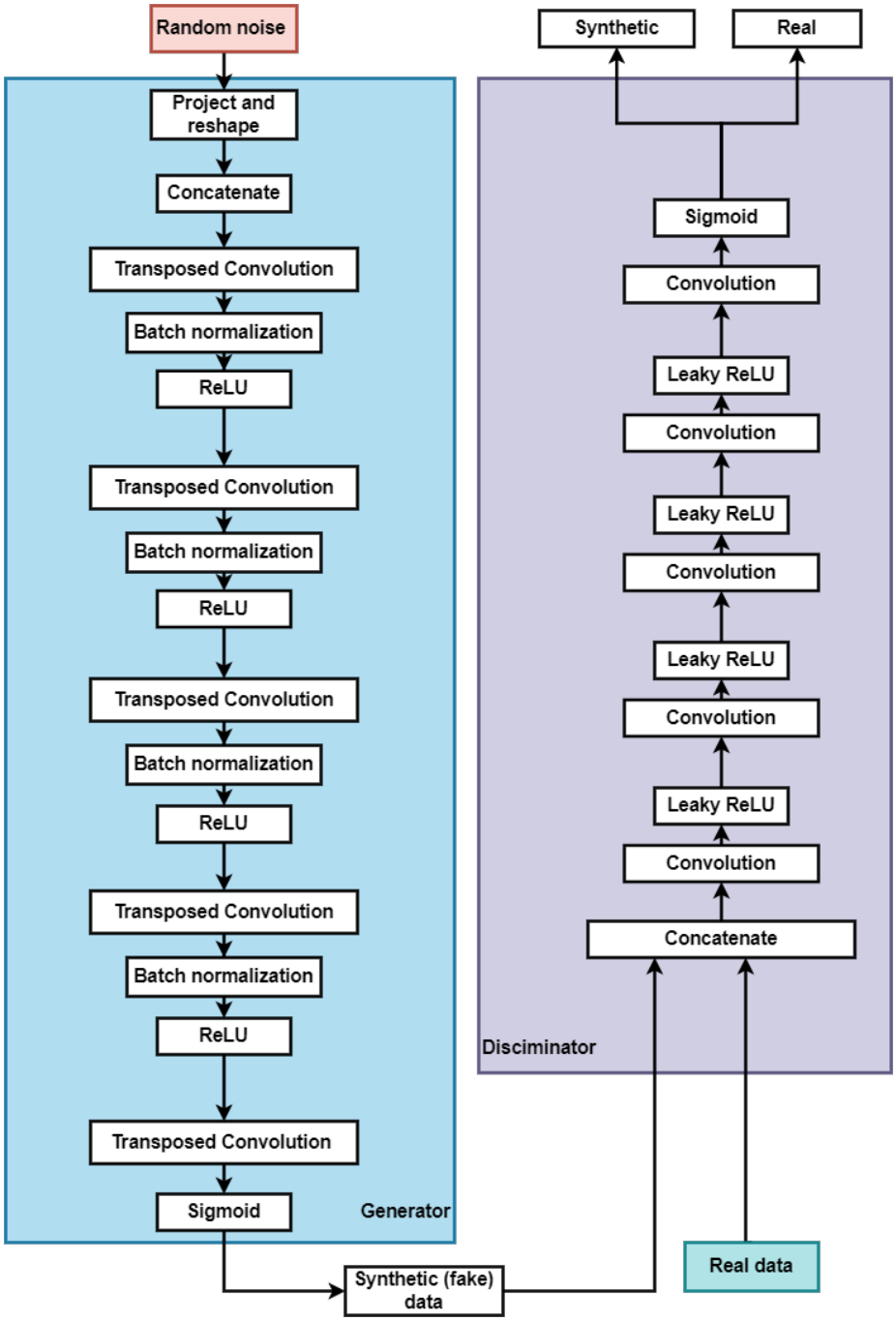


Figure 5.4: Generator and discriminator architectures of GAN

- We notice that by implementing a DBSCAN algorithm, the ADR performance improves between [0.4-12]% when compared to the same case and scenario where DBSCAN is not implemented. For instance, when using DBSCAN, the ADR of Case-E scenario-1 increases by 12% in comparison to the ADR of Case-E scenario-1 in the absence of DBSCAN. Hence, implementing DBSCAN algorithm can help detect AML attacks against V2M services as DBSCAN improves the adversarial detection rate.

Attack phase - category 2 (with CGAN)

In this category, we examine the ADR performance under the five access cases and three detection modes. It is important to note that DBSCAN and GAN are used in detection mode 1 and mode 2, respectively. As depicted in Figure 5.5, DBSCAN fails to detect AML attacks, particularly when the adversary uses CGAN in the attack phase to generate synthetic data instances to augment the accessed portion of the dataset O_d. We make the following observations:

- Similar to Table. 5.4, the ADR percentages improve as the adversary's access to O_d shrinks. The ADR performance in all three scenarios and five cases drop significantly when compared to category 1 (i.e., attack phase without CGAN). For instance, Case-E (20% access to O_d) under detection mode 0 of scenario-1, we notice that the ADR drops by 15.8%, from 59.3% in category 1 to 43.5% in category 2.

- It is clear that under the detection mode 1 (i.e., when using DBSCAN), ADR performance degrades by [0.1 - 23.2]% when compared to its performance under the same mode in category 1.

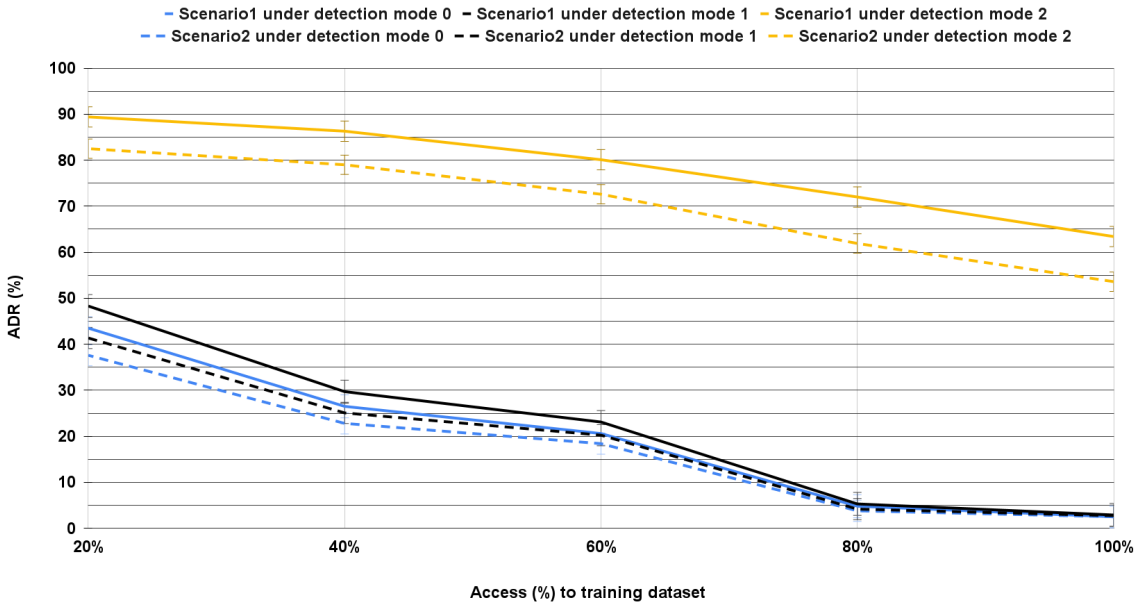


Figure 5.5: ADR performance of scenario-1 and scenario-2

- Under detection mode 2 (i.e., when using GAN), the GAN-based detection method outperforms the DBSCAN-based detection method significantly for all scenarios and cases, especially in the gray-box attack cases (i.e., cases B-E). For instance, in case-E of scenario-2, the GAN solution achieves 83.2% whereas the DBSCAN solution achieves 41.4% only. This means an adversarial ML attack is about 2 times more likely to be detected when the GAN solution is implemented over the DBSCAN solution. Moreover, the ADR performance improves even further under scenario-1 for the same case, as it attains 88.4% and 48.3% for the GAN and DBSCAN solutions, respectively. Even when the adversary's access to O_d increases to 80% (case-B), the GAN-based detection method maintains a high ADR performance as it achieves 72.2% and 62.1% for scenario-1 and scenario-2, respectively, when compared to 5.3% and 4.2% for the same scenarios under the DBSCAN-based detection method. However, we notice that all the detection methods fail to perform well when access to O_d becomes 100%. The rationale is that an adversary can create an exact replica of classifier-2 (i.e, the operating classifier at the mobile edge) during an attack phase of category 2, which makes it significantly challenging to detect.

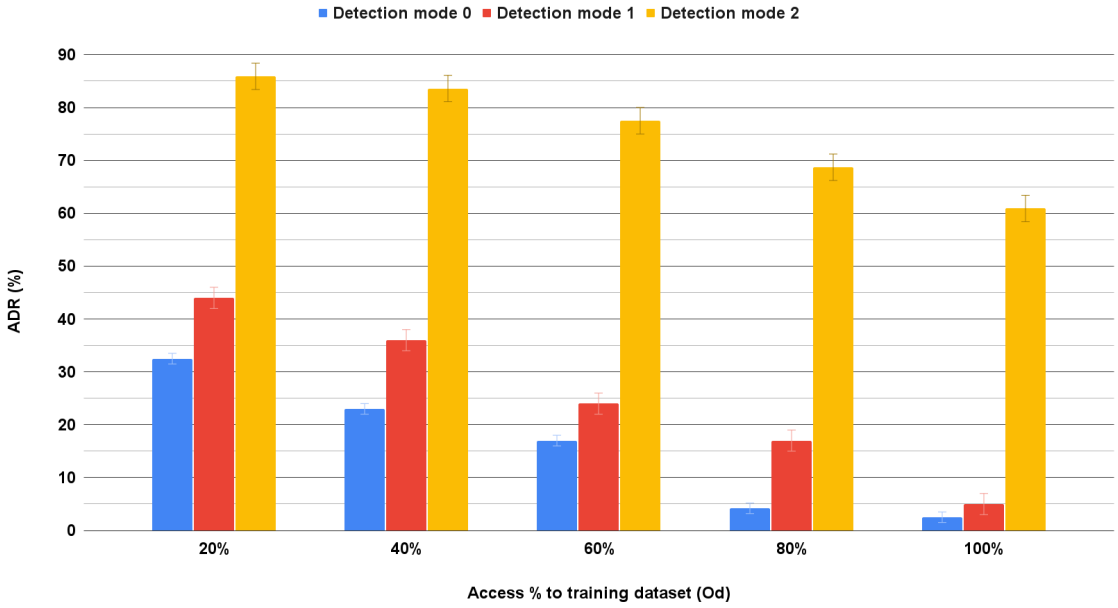


Figure 5.6: ADR performance of scenario-3 only

In Figure 5.6, we present the ADR results for scenario-3. Similar to the previous point, the GAN-based detection method outperforms the DBSCAN-based detection method in all scenarios and access cases. Additionally, the results demonstrate that the GAN-based solution can maintain high detection rate even with the existence of the two scenarios. This means that our proposed method is resilient and robust even under severe attack conditions.
Furthermore, we study the ADR and EIR performances under three additional adversarial detectors. The analysis of the dataset uncovers insights into the landscape of adversarial attack detection, illustrating general trends and specific detector vulnerabilities. It is evident that as adversaries gain more comprehensive access to the training dataset, the efficacy of detection systems uniformly declines across all models and types of attacks. In addition, the models become easily evasive.

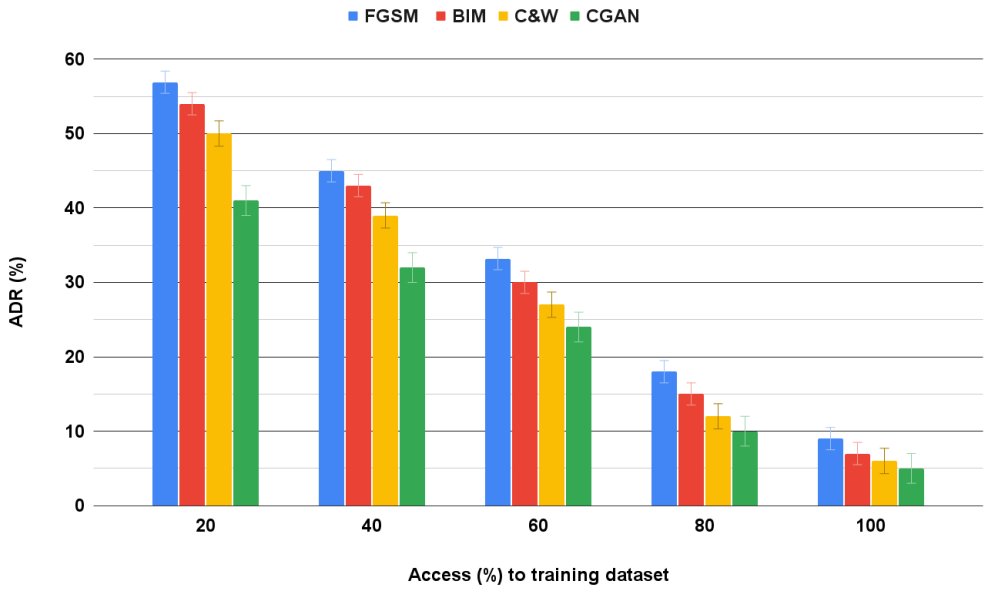


Figure 5.7: DBSCAN detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - ADR (%)

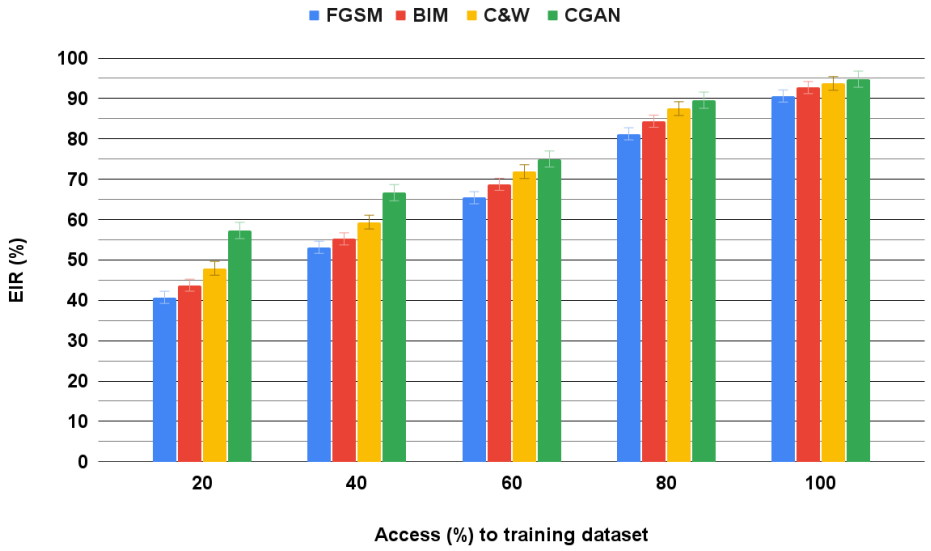


Figure 5.8: DBSCAN detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - EIR (%)

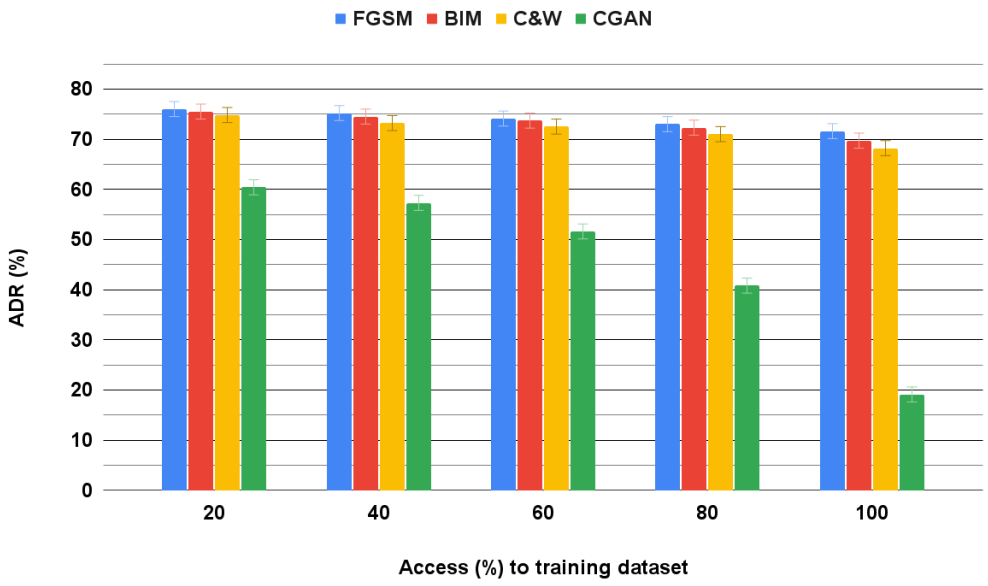


Figure 5.9: SVM detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - ADR (%)

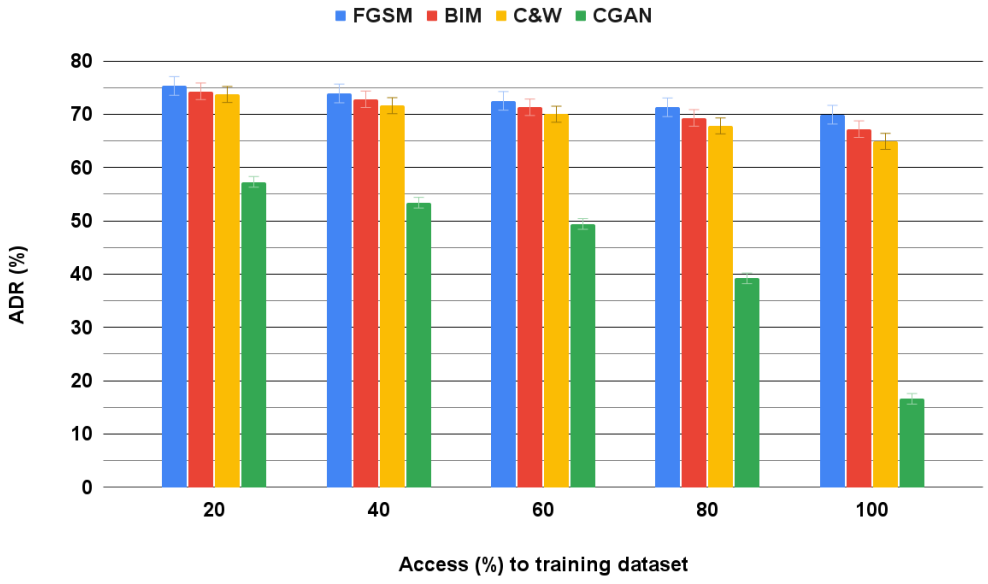


Figure 5.10: LSTM detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - ADR (%)

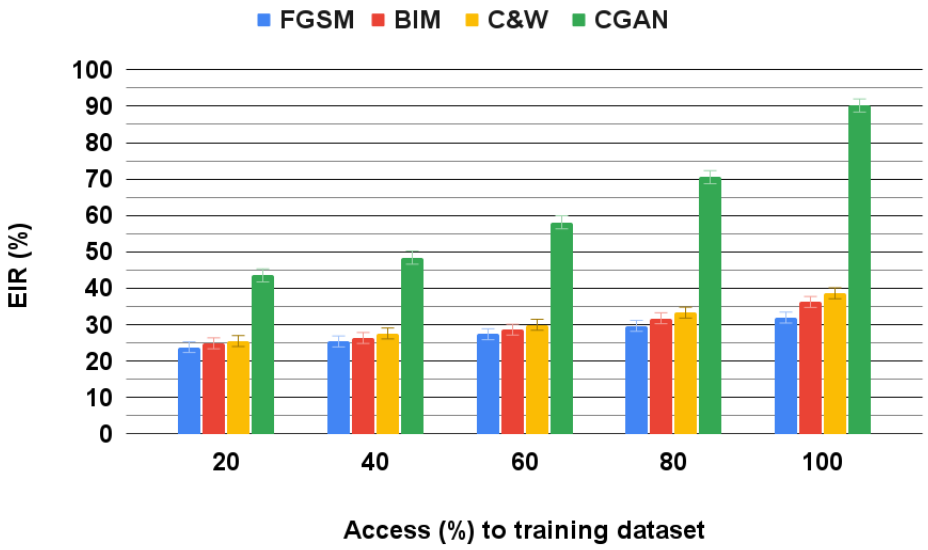


Figure 5.11: SVM detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - EIR (%)

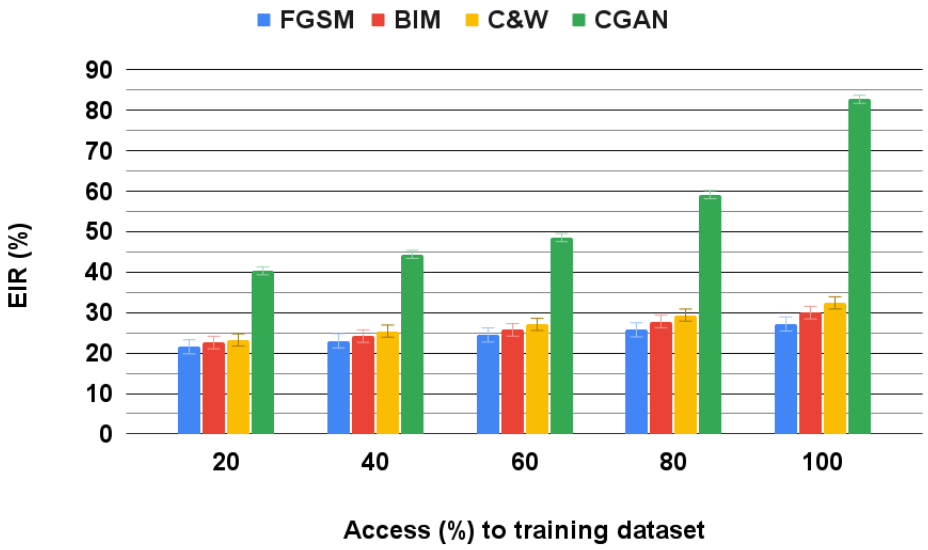


Figure 5.12: LSTM detection performance of scenario-3 under FGSM, BIM, C&W and CGAN attacks - EIR (%)

We also discuss the performance of individual detectors against the backdrop of CGAN's effectiveness. We illustrate the ADR performance of DBSCAN, as shown in Figure 5.7. Starting from a detection rate of 56.9% against FGSM attacks at 20% access, the performance drops dramatically to just 9% at 100% access. This trend is consistent across all attack types, including BIM, C&W, and CGAN, indicating that the DBSCAN detector's fails significantly to identify adversarial examples, with CGAN showing the lowest detection rates starting at 41% and decreasing to 5%. Conversely, the EIR performance in Figure 5.8 reveals an increasing trend, where evasion rates climb as adversaries gain more access to the training dataset. For FGSM attacks, the evasion increase rate jumps from 40.73% at 20% access to 90.63% at 100% access. In addition, the SVM detector reveals a particular weakness with its detection capability against CGAN attacks. We observe a sharp decrease from 54.2% at minimal data access to a mere 9.4% when adversaries have complete access, as shown in Figure 5.9. This reduction is a clear demonstration of SVM's struggles against the crafted adversarial examples generated by CGAN, a pattern shown in the LSTM detector's decreasing detection rate from 57.3% to 16.6% across similar access levels as depicted in Figure 5.10. We also observe the evasion rates for both detectors is high, indicating that the models fail to prevent the inference and evasion attacks, as shown in Figure 5.11 and Figure 5.12.

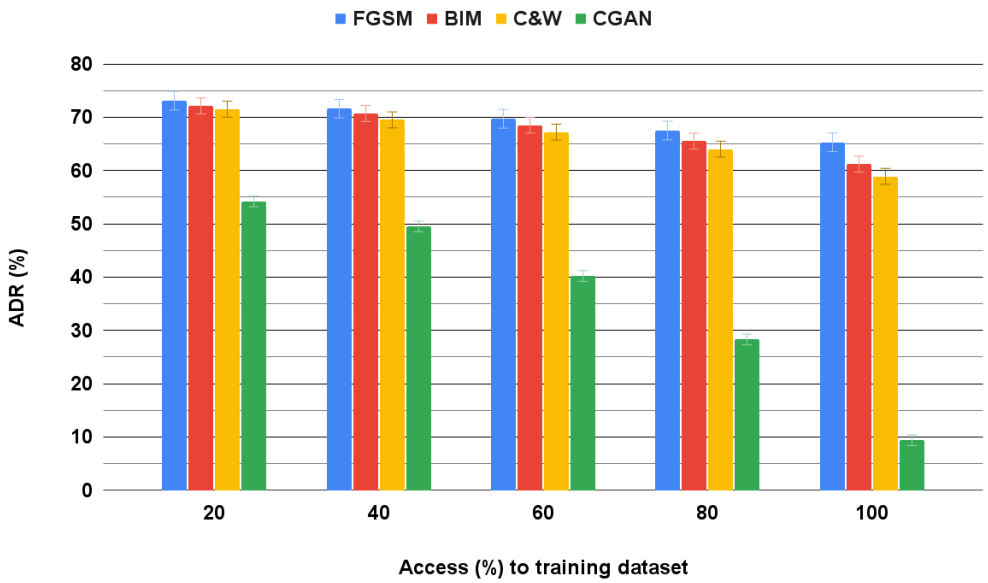


Figure 5.13: AAE detection of scenario-3 under FGSM, BIM, C&W and CGAN attack - ADR (%)

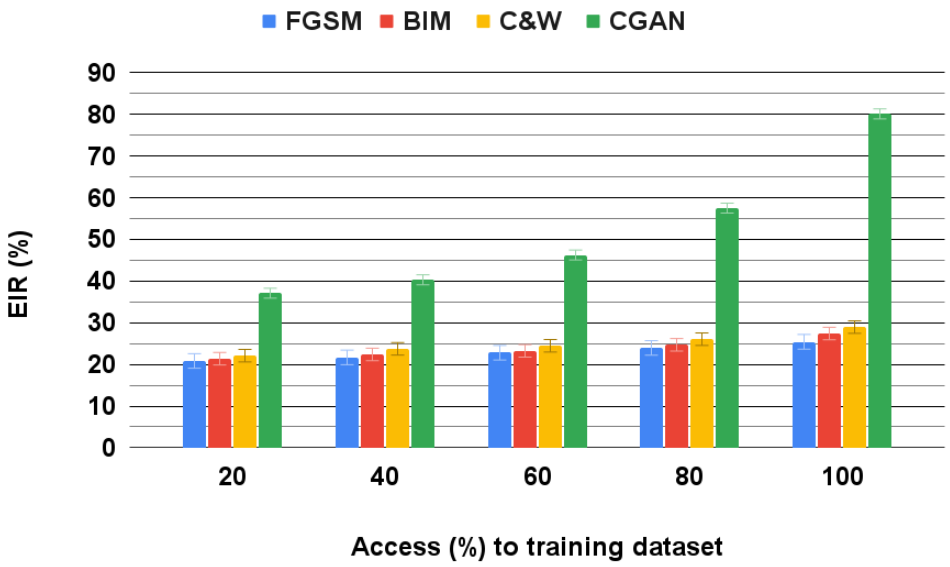


Figure 5.14: AAE detection of scenario-3 under FGSM, BIM, C&W and CGAN attack - EIR (%)

Moreover, the AAE detector in Figure 5.13, despite initially displaying a more robust defense with a 60.4% detection rate at 20% access, faces a steep decline to 19.1% at full access, highlighting the adaptive prowess of CGAN attacks even against advanced detection mechanisms. Regarding EIR performance, the AAE detector shows the better resilience to prevent inference and evasion attacks, as depicted in Figure 5.14. In contrast, the GAN-based detection system stands out for its resilience, maintaining relatively high detection rates even at full dataset access as shown in Figure 5.15. Although it registers a drop from 89% to 63%, its performance is notably superior to other models, emphasizing its potential as a more effective mechanism in the adversarial detection arms race. In addition, the GAN-based detector's performance against FGSM attacks dips from 92.5% to 87% as dataset access grows from 20% to 100%. Similar trends are observed with BIM and C&W attacks across all detectors, for instance, the SVM detector's detection rate for BIM attacks declines from 72.1% at 20% access to 61.2% at 100% access, and the AAE detector sees a reduction in C&W detection rates from 74.8% to 68.2% over the same access range. Conversely, the EIR performance demonstrated in Figure 5.16 illustrates that evasion rates escalate with increased adversarial knowledge. This starts at 3.65% for FGSM with 20% access and reaching to 9.38% at full access. Similar observations are noticed with BIM going from from 4.17% to 10.42%, C&W from 4.69% to 11.46%.

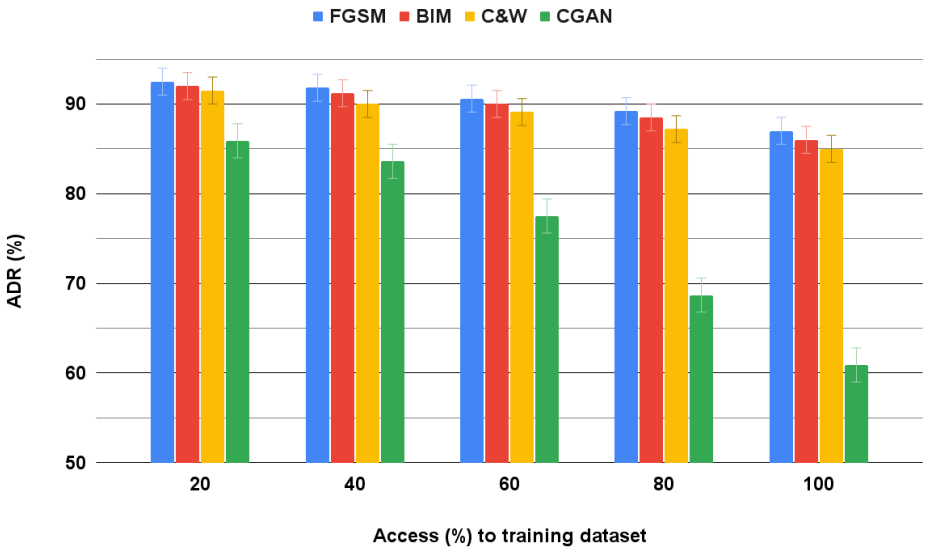


Figure 5.15: GAN detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - ADR (%)

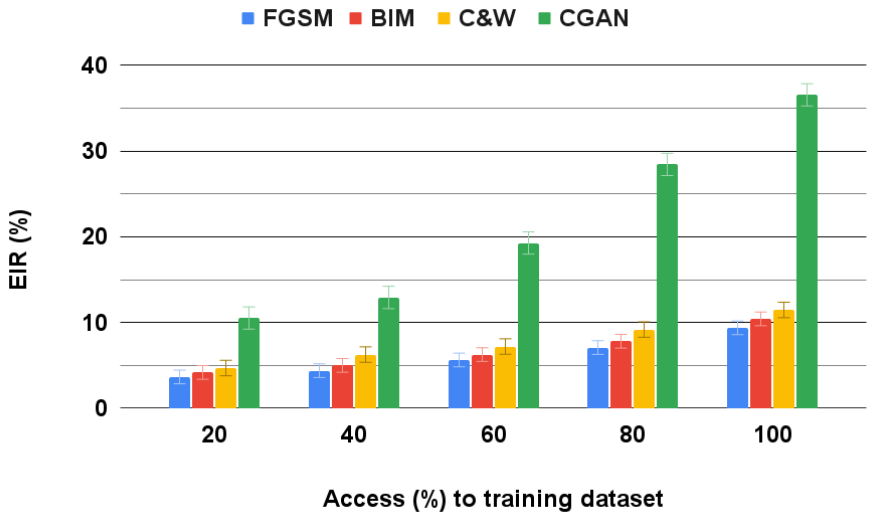


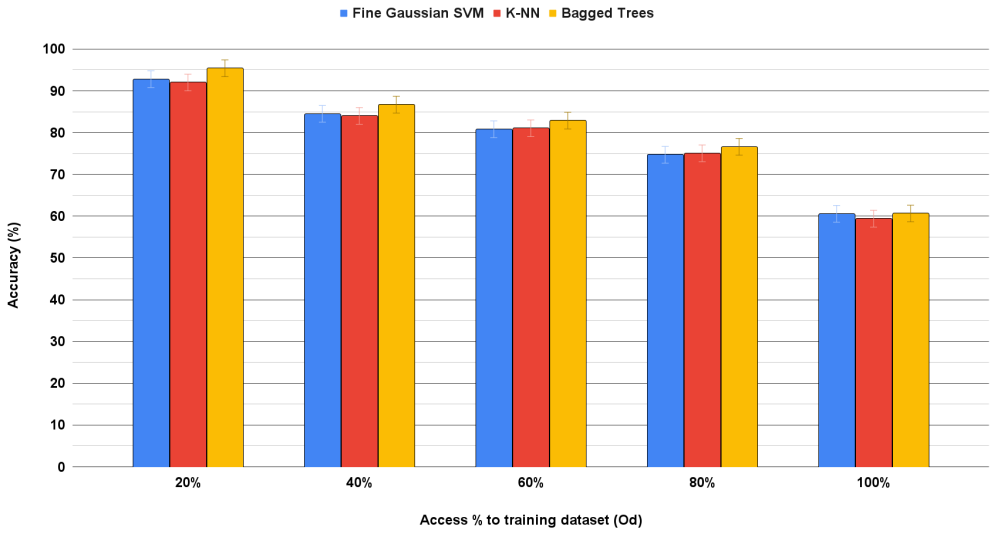
Figure 5.16: GAN detection of scenario-3 under FGSM, BIM, C&W and CGAN attacks - EIR (%)

We investigate the ADR values further by analyzing the confusion matrix of classifier-1 and classifier-2. In classifier-1, we compare three different classifiers, namely SVM, K-NN

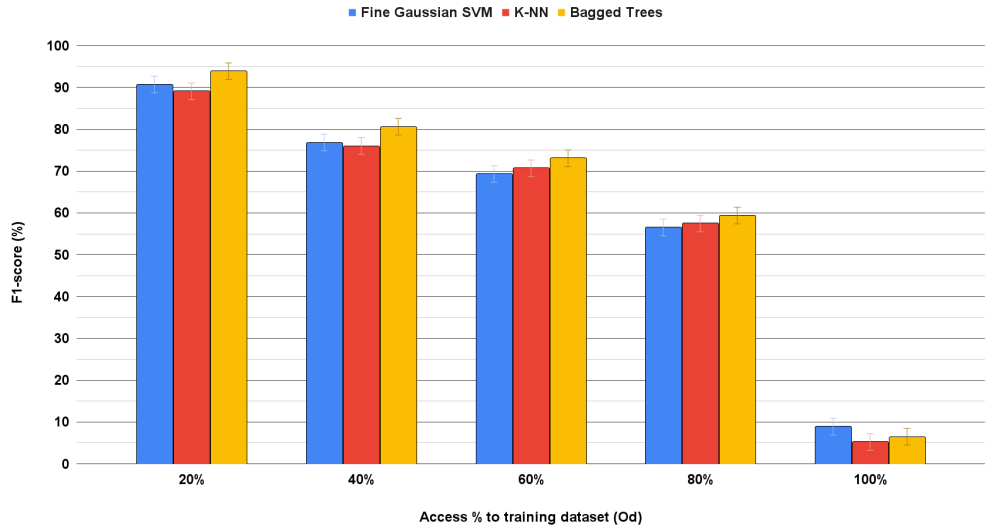
and Bagged Trees in terms of accuracy and F1-score as depicted in Figure 5.17. Classifier-1 is trained using the generated samples of the GAN model. Classifier-1 serves as a binary classifier, where it decides whether a request passes to classifier-2 or not. The results show that Bagged Trees classifier slightly outperforms the other two classifiers in terms of accuracy and F1-score under all five cases. Hence, we deploy Bagged Trees as classifier-1. It is worth noting that the results are taken over the average of 10 runs with a confidence interval of 95%.
Besides, the presented accuracy and F1-score results belong to the illegitimate class of scenario-2. We study scenario-2 in particular because as it has the lowest ADR performance of all the three scenarios.
In addition, we observe a repeated pattern as the accuracy and F1-score performances drop with the increment of the adversary's access to the training dataset (O_d). Furthermore, with access of 100% to O_d, the accuracy performance drops to an acceptable percentage. However, we notice a significant drop in the F1-score performance of all three classifiers. The is mainly because the Bagged Trees classifier is unable to distinguish between legitimate and illegitimate requests.

As shown in Figure 5.18, the number of TP and FN instances remain approximately the same across all five cases. On the contrary, the number of FP and TN instances changes according the different access cases to (O_d). As the adversary's access to (O_d) increases, the number of FP instances increases, whereas the number of TN instances decrease. For instance, with 20% access case, number of FP and TN instances are 45 and 455, respectively. On the contrary, under 100% access case, the number of FP and TN reach their maximum with 483 instances and lowest with 17, respectively. This means there are 483 FP instances that can bypass classifier-1 since they are predicted as legitimate requests. This explains the reason behind the low ADR performance under 100% access cases in all three scenarios as aforementioned. It is worth noting that requests predicted as legitimate (i.e., only TP and FP instances) can only pass to classifier-2. For example, under 20% access case, the number of TP and FP requests passing to classifier-2 are 738 and 45, respectively. Under 100% access case, the number of TP and FP requests passing to classifier-2 are 741 and 483, respectively. On the other hand, predicted requests as illegitimate (i.e., only TN and FN instances) will be blocked.

The requests passed to classifier-2 are split into three priority classes, namely low, medium and high priority. To distinguish between requests sent by microgrid's users and the adversary, we add "legitimate" and "illegitimate" as suffixes to the priority classes. Hence, "low legitimate" is a request with a low priority that is sent by microgrid users. Whereas "low illegitimate" is a low priority request that is sent by the adversary as shown in Figure 5.19. For 20% access case, 738 legitimate requests are classified into the different



(a)



(b)

Figure 5.17: Accuracy and F1-score performance of the illegitimate class under scenario-2 (a) Accuracy performance under different access cases to O_d, (b) F1-score performance under different access cases to O_d

priority classes, 246 requests are classified as low legitimate, 243 requests are classified as medium legitimate and 249 requests are classified as high legitimate as shown in Figure 5.19.a. On the other hand, 45 illegitimate requests are classified into either medium illegitimate (19 requests) or high illegitimate (26 requests). The low illegitimate class has zero requests as it is the adversary's intention under scenario-1 to misclassify the medium and high priority requests as low-priority requests. Hence, the adversary is not interested to send low-priority class requests.

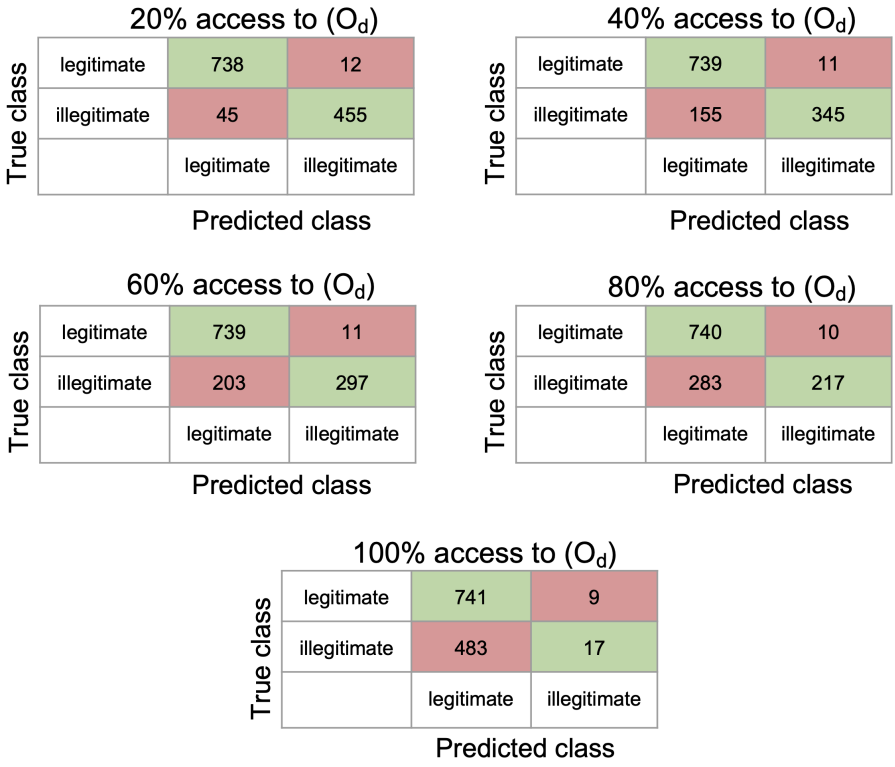
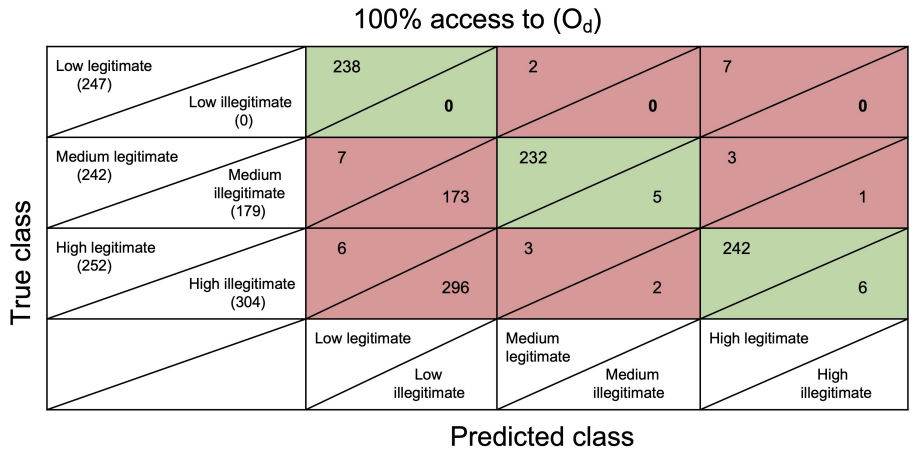
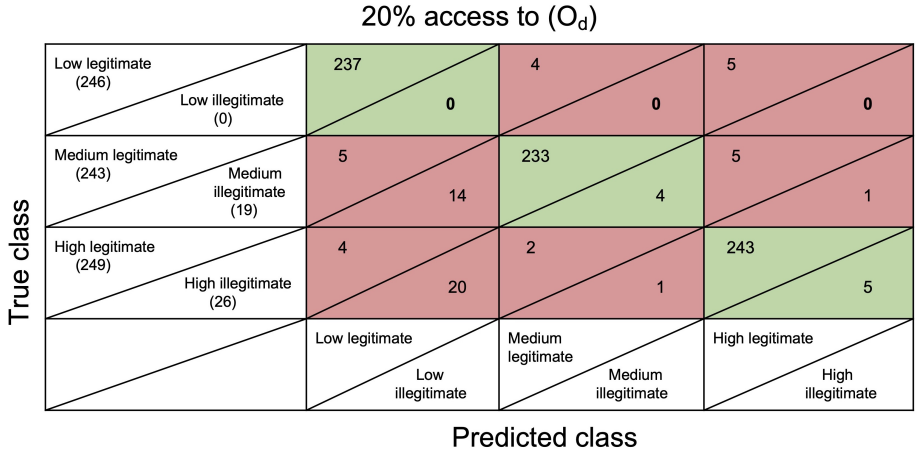


Figure 5.18: Confusion matrix of classifier-1 (Bagged Trees) under scenario-1 with the different access cases to (O_d)

For 100% access case, 741 legitimate requests are classified into the different priority classes, 247 requests are classified as low legitimate, 242 requests are classified as medium legitimate and 252 requests are classified as high legitimate as shown in Figure 5.19.b. On the other hand, 483 illegitimate requests are classified into either medium illegitimate (179 requests) or high illegitimate (304 requests).



(a)



(b)

Figure 5.19: Confusion matrix of classifier-2 under scenario-1 with two access cases. (a) case-E (20% access to (O_d)), (b) case-A (100% access to (O_d))

5.4 Summary

In summary, we presented an anticipatory study of two AML attacks in V2M settings under three scenarios. Knowledge of the adversary has been taken into consideration as we compared five different cases of access to the training dataset, one case under white-box setup and four cases under gray-box setup. In the white-box case, the adversary has 100% access to the training dataset of the K-NN classifier; whereas, in the gray-box cases,

the adversary can access [20-80]% of the training dataset. In addition, we provided two attack phase categories where each category indicates whether the adversary used CGAN during the attack or not. Moreover, we investigated three modes of the detection phase, mode 0 indicates the absence of the a detection method; mode 1 and mode 2 present a detection method using DBSCAN and GAN, respectively. Through simulations, we have presented the ADR performance under all three scenarios and five access cases for both attach and detection phases. Furthermore, we provided the comparison of three different classifiers in terms of the accuracy and F1-score and chose Bagged Trees classifier to be trained on the GAN output and serves as classifier-1. Moreover, we analyzed the confusion matrix of Bagged Trees classifier under scenario-1 for all five access cases. We also provided the confusion matrix of classifier-2 under two access cases (case-A and case-E). We show that DBSCAN fails to detect AML attacks, particularly when the attacks are intelligently augmented, obtaining an ADR of up to more than 50%. On the other hand, our proposed framework outperforms DBSCAN and achieves an ADR of 90% and above under bagged trees classifier. It is clear that under detection phase mode 1 (DBSCAN), ADR performance degrades by [0.1 - 23.2]% when compared to its performance under the same mode in category 1.

For the future work, we consider two main directions. The first direction focuses on reducing the number of FN instances since they are as important as blocking the illegitimate requests. We consider implementing a probabilistic approach to filter out the incoming requests. The second direction aims at increasing the ADR performance, particularly for the 100% access case. One way to address this issue is creating a binary classifier for each priority class. This will reduce the adversary chances to mount an impactful evasion attack against V2M systems. However, the computational cost of the system will need to be studied.

Chapter 6

AI Model Optimization for Adversarial Attacks Detection on Edge Devices in V2M Services

In this chapter, we focus on adapting the adversarial detection model to fit the limited capabilities of edge devices in V2M settings. Given that edge devices are typically enabled with low computational power and very small memory capacities, it is crucial to customize the detection model to operate effectively within these constraints. Our goal is to reduce the detection model size without compromising the model's detection performance. We begin the chapter by establishing the significance and necessity of AI model compression in a V2M environment, emphasizing how the constrained computational and memory resources of edge devices affect the model's inference time and throughput. Subsequently, we provide an overview of various AI model compression techniques, culminating in the introduction of our proposed method towards the end of the section. We then evaluate the performance of our method against different types of evasion attacks, comparing its effectiveness with other adversarial detection approaches. We conclude the chapter with final remarks, summarizing key findings and implications of our research.

6.1 Introduction

IoT sensors in electrical grids collect real-time data that improves grid stability, efficiency, and security. The traditional method of sending this vast amount of data to centralized cloud servers is becoming less feasible due to latency and bandwidth issues, necessitating

real-time, localized processing. This need has led to a shift towards Edge AI, where AI algorithms are deployed directly on edge devices like IoT sensors and mobile phones, enabling data to be processed locally [237]. This local processing reduces latency, enhances privacy and security, and decreases network bandwidth use. However, deploying AI on edge devices poses challenges due to their limited processing power, memory, and network bandwidth, along with constraints on battery life and power consumption [238]. Addressing these challenges requires model compression techniques to adapt AI models to the limited capabilities of edge devices. By optimizing AI models for edge deployment, smart meters in smart grids can process data more efficiently, enhancing their functionality and sustainability [239].
This chapter highlights the importance of AI model compression within a V2M environment and introduces an integrated compression mechanism that combines model design and compression into a single process. This approach results in an optimized detection model that functions efficiently in V2M edge environments, maintaining robust performance against adversarial attacks without sacrificing detection capabilities.

6.2 Model Optimization

In this section, we explore model optimization techniques focusing on model design and compression, employing state-of-the-art (SoTA) methods implemented using MATLAB. We begin by discussing various model compression strategies, such as projection, pruning and quantization. These methods are essential for enhancing model efficiency without significantly impacting accuracy. We then shift to a comprehensive model optimization framework that integrates initial model design with subsequent compression techniques. This integrated approach is designed to balance performance and efficiency, making models both compact and capable of maintaining high accuracy in under adversarial attacks against V2M services.

6.2.1 Model Compression

We implement a projection-based method that uses Principal Component Analysis (PCA) for compression [240]. PCA is a vital technique employed for compressing convolutional neural networks (CNNs). The method capitalizes on PCA to efficiently reduce the dimensionality of the layers' responses (i.e., output), which are inherently high-dimensional and computationally expensive to process. By applying PCA, the algorithm identifies the principal components that capture the maximum variance within the data. This allows

the network to retain the most informative components of the responses while discarding the less significant ones. This reduces the layer's complexity without severely impacting its functionality. The compression process involves representing the responses at a convolutional layer in a lower-dimensional subspace using a projection matrix M derived from PCA. Specifically, the method minimizes the reconstruction error between the original responses and their projections onto a low-rank subspace, formulated as:

$$\min_{M} \sum_{i} \|(y_i - \bar{y}) - M(y_i - \bar{y})\|_2^2 \quad \text{s.t.} \quad \text{rank}(M) \leq d'$$

This objective is crucial as it directly impacts the network's ability to approximate the original layer's responses accurately while significantly reducing the number of parameters. Determining M allows the transformation of the dense weight matrix W into two simpler matrices. This transformation reduces the computational complexity from processing the full dimensions of W to handling the lower-dimensional matrices that approximate the original responses. This allows the compression not to randomly discard information but instead to focus on the least significant variance directions.

In addition to the projection-based compression, we implement a pruning-based method that uses Taylor scores for compression. The pruning method works by estimating the impact of removing individual neurons or filters on the final loss. The method is based on the first-order Taylor expansion to approximate the importance of neurons. The importance $I(1)_m$ of a parameter w_m is calculated using the square of the product of the gradient of the loss with respect to the weight and the weight itself:

$$I(1)_m = (g_m w_m)^2$$

Here, g_m represents the gradient of the loss with respect to the weight w_m. This formulation allows for a computationally efficient estimation of importance, as it utilizes gradients that are readily available from the standard training process. The calculated importance scores are then used to iteratively prune the least important neurons or filters. This pruning is integrated within the training and fine-tuning cycles, allowing the network to dynamically adjust to changes in its architecture, thereby minimizing the impact on performance. It is proven that the pruning method can significantly reduce the computational resources required for deploying CNN model without compromising the performance.

We also study and implement quantization as one of the compression method that can significantly reduce a model's size. Fundamentally, quantization refers to reducing the precision of weights, parameters, biases, and activations so that they occupy less memory and the size of the model is reduced [241]. Quantization from 32-bit to 8-bit precision serves

as a critical method for optimizing neural networks to better suit environments where resources are constrained, such as in edge devices. In traditional setups, each weight in a neural network is typically stored as a 32-bit floating-point number. Given that advanced neural networks can contain millions of such parameters, the cumulative memory requirement becomes substantial. This high demand for storage not only impacts the amount of memory needed but also affects the speed and power efficiency of computations, which are critical factors in edge systems.
When we apply quantization, these 32-bit floating-point weights are converted into 8-bit integers. This transformation dramatically reduces the amount of memory required—by a factor of four, since 8-bit integers occupy significantly less space than 32-bit floating-point numbers. Beyond just reducing storage requirements, this also enhances computational efficiency. Processors can handle operations involving 8-bit integers faster than those involving 32-bit floating points due to the reduced data width, which accelerates the overall processing time of the neural network [242].

6.2.2 Proposed Framework

In Figure 6.1, we view the end-to-end pipeline for detection model deployment. As part of this process, model optimization is crucial for creating a compact and robust model in edge environments. A critical aspect of model optimization is model design, which is an exhaustive stage that requires extensive domain knowledge to select the right model architecture and hyperparameters. For this part, we use Neural Architecture Search (NAS) to find the best model architectures that deliver the best performance in terms of accuracy and F1-score. We then implement model compression techniques to reduce the model's size while maintaining optimal performance.

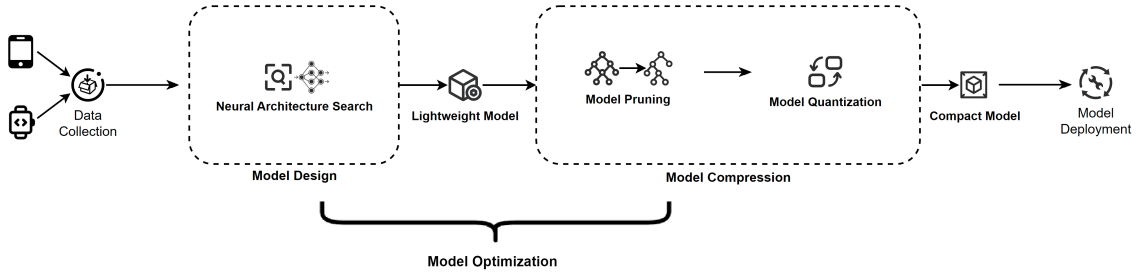


Figure 6.1: AI deployment pipeline

We explain our methodology here in details. In Algorithm 2, we implement NAS method using Bayesian Optimization [243], which will identify best performing CNN ar-

Algorithm 2 NAS with Bayesian Optimization Under Specific Performance Constraints

Input: Training data D_{train}, validation data D_{val}, accuracy threshold C, F1-score threshold F, number of top architectures n
Output: Set of top n best performance CNN architectures $\{M_i\}$
function NAS_Bayesian_Optimization($D_{\text{train}}, D_{\text{val}}, C, F$)
    Define search space S
    Initialize Bayesian model B
    Define acquisition function α
    while not stopping_criterion **do**
        $A \leftarrow$ select_next_architecture(B, α)
        perf $\leftarrow$ train_and_evaluate($A, D_{\text{train}}, D_{\text{val}}$)
        update_Bayesian_model($B, A,$ perf)
    $\text{Top}_n \leftarrow$ select_top_n_architectures(B, n, C, F)
    return Top_n
$\text{Top}_n \leftarrow$ NAS_Bayesian_Optimization($D_{\text{train}}, D_{\text{val}}, C, F$)

chitectures regardless of their size since the model compression will be part of the next algorithm. This process begins by defining the search space S, which encompasses various possible configurations of CNN architectures, including different types and arrangements of layers, activation functions, and other hyperparameters. The core of the NAS process involves a Bayesian optimization framework, initialized by setting up a Bayesian model B. This model is key in predicting the performance of various architectural configurations based on historical evaluation data. The algorithm employs an acquisition function α, which guides the search. It helps in balancing the exploration of new, untested architectural configurations against the exploitation of configurations known to perform well, thus ensuring an efficient search process.

During the NAS execution, the algorithm iteratively selects the next architecture to evaluate by utilizing the acquisition function. Each selected architecture is then trained and assessed on the given training and validation datasets D_{train} and D_{val}, respectively. The performance of each architecture is measured, and the results are used to update the Bayesian model, enhancing its prediction accuracy for future iterations. This loop continues until the predefined stopping criteria C and F are met. The stopping criteria serve different purposes. For instance, it acts as a guide for the Bayesian optimization process to explore the search space efficiently. It does not only facilitate a more structured search but also aid in evaluating the success of the search strategy itself. Without such criteria, the optimization process will not have metrics to assess the performance of CNN architectures. This will lead to continuously explore the search space and without exploiting. Although

the performance metrics can be set to (1.0) (as we did in this work), in practical scenarios, especially in constrained environment, achieving accuracy and F1-score of (1.0) may not be feasible due to limitations in computational resources and memory. Moreover, depending on the complexity of the application and dataset, achieving accuracy of (1.0) could be a sign of overfitting [244]. Hence, users can set lower performance criteria to avoid such issues. Besides, different applications may have varying requirements for accuracy and F1-score based on their specific needs and the consequences of misclassifications. Once the search is complete, the algorithm selects the top n architectures that not only meet but potentially exceed the defined accuracy and F1-score thresholds. This selection is based purely on performance metrics, without considering the computational or memory efficiency of the architectures. These top-performing models form the output of the NAS process and are poised for subsequent optimization steps, such as pruning and quantization, to enhance their efficiency for deployment in edge environments.

Algorithm 3 Select the Best Optimized and Compressed Architecture

Input: Set of top n optimized CNN architectures $\{M_i\}$, Training data D_{train}, Validation data D_{val}
Output: Optimized and compressed CNN model M_{best}
function PRUNE_AND_QUANTIZE_MODELS($\{M_i\}, D_{\text{train}}, D_{\text{val}}$)
    Initialize BestScore $\leftarrow -\infty$
    Initialize M_{best}
    for M in $\{M_i\}$ **do**
        $M_{\text{trained}} \leftarrow$ train_model(M, D_{train})
        $M_{\text{pruned}} \leftarrow$ iterative_pruning($M_{\text{trained}}, D_{\text{train}}$)
        $M_{\text{quant}} \leftarrow$ quantize_model($M_{\text{pruned}}, D_{\text{train}}, D_{\text{val}}$)
        $M_{\text{quant}} \leftarrow$ fine_tune($M_{\text{quant}}, D_{\text{train}}, D_{\text{val}}$)
        Accuracy $\leftarrow$ evaluate_accuracy($M_{\text{quant}}, D_{\text{val}}$)
        F1Score $\leftarrow$ evaluate_f1score($M_{\text{quant}}, D_{\text{val}}$)
        ModelSize $\leftarrow$ get_model_size(M_{quant})
        CompositeScore $\leftarrow$ calculate_composite_score(Accuracy, F1Score, ModelSize)
        if CompositeScore $>$ BestScore **then**
            BestScore $\leftarrow$ CompositeScore
            $M_{\text{best}} \leftarrow M_{\text{quant}}$
    return M_{best}
$M_{\text{best}} \leftarrow$ PRUNE_AND_QUANTIZE_MODELS($\{M_i\}, D_{\text{train}}, D_{\text{val}}$)

Following the selection of top-performing architectures algorithm.1, the next phase involves refining these models through pruning and quantization to enhance their com-

putational efficiency. Algorithm 3 begins by taking each architecture from the set of top n optimized CNN architectures. Initially, every model is trained to establish a strong baseline before any modifications that could potentially degrade the model's effectiveness. Once the models are fully trained, the next step is iterative pruning. During this stage, the algorithm systematically removes the least important connections within each model. Taylor-based pruning is used in this step. This method effectively reduces the model's complexity and size, which in turn decreases memory usage and computational demands without significantly sacrificing accuracy.

Following pruning, the models are subjected to quantization. This process converts the model's floating-point weights to a lower-precision format, such as 8-bit integers. Quantization significantly reduces the memory footprint of the model and is often accompanied by speed improvements during inference. However, quantization can sometimes lead to a loss in accuracy, which is why the next crucial step is fine-tuning. Fine-tuning the quantized models using both training and validation data helps recover any accuracy lost during the quantization process. It adjusts the model parameters within the constraints of their new precision levels to ensure the final model continues to perform robustly.

Subsequently, the algorithm evaluates each model on three parameters: accuracy, F1-score, and the size of the model. Accuracy and F1-score are direct indicators of the model's performance, assessing how well the model predicts correct outcomes and balances precision and recall, respectively. The size of the model is also evaluated to ensure the final model is not only high-performing but also compact enough for efficient deployment. To synthesize these diverse metrics into a singular decision metric, the algorithm computes a composite score for each model. This composite score is calculated through a weighted sum function as described in equation 6.1. For simplicity, we use equal weights for all three metrics. However, in real scenarios, different applications may have varying requirements for accuracy, F1-score and model's size based on their specific performance requirements and the available computational resources and memory.

$$CompositeScore = w_{acc} \cdot Accuracy + w_{F1} \cdot \text{F1-score} + w_{size} \cdot \left(1 - \frac{\text{Compressed Size}}{\text{Original Size}}\right) \quad (6.1)$$

6.3 Performance Evaluation

In this section, we present a detailed description of the experimental setup in our study. We discuss the various parameters and attacks' scenarios used to evaluate the performance of our proposed compression framework. Furthermore, we analyze the obtained results, providing comprehensive insights into the effectiveness and limitations of our proposed method. We present a thorough analysis of the anticipated AML attacks against V2M

services and the potential countermeasures by studying the V2M system under three different scenarios and five different cases. The scenarios represent the different AML attacks against the V2M services, whereas the cases represent the adversary's access to the training dataset of the victim's ML classifier.

6.3.1 Simulation settings

The adversary goal is to compromise the integrity of the low-priority class by wrongly categorizing medium- and high-priority requests as low-priority requests. In addition, we compare five cases where the adversary has access to varying amounts of the victim's ML training dataset. In white-box attacks, the adversary has access to 100% to the training dataset. Gray-box attacks, the adversary has 80%, 60%, 40% and 20% access to the training dataset. In the gray-box attack, after collecting a certain number of data instances, the adversary can use a CGAN (i.e., under the CGAN-based evasion attack) to generate additional data in order to create a complete dataset. It is important to note that GAN training performance is influenced by batch size and the number of epochs. Through experimentation, we found that fine-tuning these parameters is essential for optimal GAN performance in our V2M context. For instance, larger batch sizes (e.g., 128 in our case) allow the GAN to see a more representative sample of the data distribution in each training step, helping to stabilize training and capture complex energy patterns. A higher epoch count (1000 in our simulations) provides more opportunities for the GAN to refine its understanding of the data distribution which is crucial for detecting hidden anomalies that could indicate potential attacks. Our results showed optimal GAN performance with a batch size of 128 and 1000 epochs, striking a balance between computational efficiency and model effectiveness. However, these values may vary depending on specific applications and computational resources. For different V2M configurations or related applications, we recommend conducting a thorough hyperparameter search, potentially using techniques like grid search or Bayesian optimization.
Furthermore, we implement a black-box attack along with gray-box and white-box attacks. Black-box scenarios are more practical than the white-box model for ML applications, especially in smart grid contexts. In this scenario, the inputs and outputs of the deployed ML models are accessible to third parties through public querying, typically via cloud-based systems or APIs [245]. One common method for executing these queries is through the HTTP request-response protocol [246]. For instance, a public user can send a query to an ML model hosted on the Google AI platform using an HTTP request with necessary credentials and a JSON formatted input. The model then responds with a JSON payload containing the predicted output. Adversaries aim to gather enough data from limited queries to train a surrogate model, effectively turning a black-box scenario into a white-box

scenario. This surrogate model is then used to generate subtle perturbations to deceive the victim's model. Moreover, we use the CNN architecture depicted in Figure 6.2. The core of the network comprises multiple convolutional layers. The first layer uses 128 filters of size 3 with a stride of 1, employing the ReLU activation function. This is followed by two convolutional layers with 256 filters of the same size and stride to enhance feature extraction capabilities. The MaxPooling layer with a pool size of 2 is applied next. The final convolutional layer employs 512 filters to capture more complex patterns before the data is flattened into a one-dimensional array to facilitate dense layer processing. The dense segment of the network features a large, fully connected layer with 1024 units using ReLU activation, designed to synthesize the features extracted in previous layers. A dropout rate of 50% is implemented to prevent overfitting. The output layer is configured for a binary classification task which uses a sigmoid activation function to produce a probability indicating the likelihood of the input being legitimate or illegitimate task. In addition, we test our CNN detector under traditional attack scenarios that were initially defined in [184].It is worth noting that a CNN model trained on GAN-generated data offers better adversarial attack detection in V2M systems compared to using the GAN directly. This approach leverages the CNN's ability to learn from diverse synthetic adversarial examples, enhancing its generalization across various attack patterns. Unlike a GAN's discriminator, which primarily distinguishes real from fake data, the CNN can be specifically optimized to identify adversarial inputs, focusing on hidden features of manipulation. CNN flexible architecture allows for task-specific design, balanced training on normal and adversarial examples, and avoids GAN training instabilities. Moreover, CNNs offer computational efficiency for real-time detection on resource-constrained devices and adaptability through retraining as new attack types emerge. We also compare the CNN detector to other adversarial detectors such as ARIMA, 1-class SVM, 2-class SVM, LSTM and AAE. We use the EMSx dataset [247], provided by Schneider Electric, which is a comprehensive collection of data tailored for the analysis and management of electric microgrids. It encompasses historical observations and forecasts related to photovoltaic generation and energy demand across 70 industrial sites. This dataset is particularly designed for developing and testing control algorithms for microgrids that include photovoltaic units and energy storage systems. Each site in the dataset is well-documented with specific parameters related to battery operation and time-series data on energy usage and production. The photovoltaic data within the dataset is uniquely processed: a standard photovoltaic profile from a site in the South Central United States is adjusted and re-scaled for each site, compensating for the lack of detailed meteorological data which hampers accurate photovoltaic forecasting. This methodical approach allows the dataset to offer realistic scenarios for energy management system testing, providing both historical data and predictive forecasts to facilitate comprehensive studies on microgrid control and optimization.

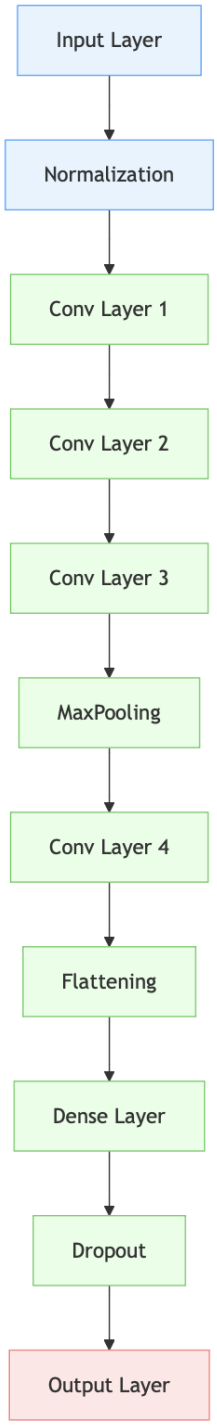


Figure 6.2: CNN architecture

To simulate adversarial activities, we constructed a malicious dataset using several functions designed to emulate various realistic malicious behaviors, as in [184]. These include partial reductions in reported power (f1(·) and f2(·)), selective bypassing of reporting (f3(·)), and manipulations based on time-of-use (ToU) pricing (f4(·), f5(·), and f6(·)). Specifically, these functions are defined as follows:

- f1(·): decreases the power reported by a smart meter by a constant fraction.
- f2(·): dynamically reduces the power consumption reported.
- f3(·): sets the power consumption to zero during specific ToU periods.
- f4(·): maintains a constant power consumption report throughout the day.
- f5(·): varies the power consumption report throughout the day.
- f6(·): alters the reported power consumption in accordance with ToU pricing.

6.3.2 Results

Before we discuss the model optimization results, it is important to validate the significance of the model optimization by understanding the impact edge resources on the CNN detection model performance, such as inference time and throughput. In Table. 6.1, we

Table 6.1: Impact of GPU utilization on the detection model performance

GPU Utilization (%)	Inference Time (s)	Throughput	Task Delay (s)
5	3.3	62	0.3
10	2	83	0.7
25	2	83	1.9
50	2	83	7.5
75	2.2	78	35

observe a clear trend of increasing task delay as the GPU utilization percentage rises. This suggests that higher utilization leads to longer wait times due to the increased processing load on the computing resources. Despite varying utilization levels, the inference time and throughput remain relatively stable. At the lowest utilization of 5%, the model demonstrates an inference time of 3.3 seconds with a throughput of 62, marking the lowest throughput observed. However, the task delay is minimal, indicating that the system is efficiently processing with ample available resources. As utilization increases to 10%, the

inference time decreases to 2 seconds, and the throughput peaks at 83, which it maintains through to 50% utilization. During this range, the task delay gradually increases from 0.7 seconds to 7.5 seconds, showing that even as the system handles a higher load, it begins to experience more significant delays. At a higher utilization of 75%, there is a slight increase in inference time to 2.2 seconds and a small drop in throughput to 78. More notably, the task delay dramatically increases to 35 seconds, indicating that at high levels of utilization, the processing efficiency starts to decrease, significantly impacting response times.
These observations suggest that the CNN model operates most efficiently under 10% utilization. Beyond this point, the increase in task delay becomes more pronounced, which could be detrimental in real-time applications where timely processing is critical. Maintaining a balance in GPU utilization is crucial; too low utilization underutilizes resources, potentially increasing operational costs without corresponding benefits in performance, while too high utilization leads to diminishing returns in terms of increased delays. Hence, lightweight deep learning models can help optimize performance, especially in environments where multiple tasks or models are run concurrently.

Table 6.2: Impact of RAM utilization (%) on the detection model performance

RAM Utilization (%)	**Inference Time (s)**	**Throughput**	**Task Delay (s)**
5	2.3	58	0.3
10	2	83	0.7
25	2	83	3.1
50	2	83	11.4
75	2.05	80	49

Table. 6.2 provides a detailed look at how RAM utilization affects the performance metrics of a CNN model, highlighting trends in inference time, throughput, and task delay as memory usage increases. One consistent trend across all levels of RAM utilization is the progressive increase in task delay. This suggests that higher memory use leads to longer processing times for tasks, due to the increased memory management overhead. Such activities can significantly impact the speed at which tasks are completed, which is crucial in performance-sensitive applications. At lower RAM utilization levels, specifically 5% and 10%, the model exhibits minimal inference times of 2.3 and 2 seconds, respectively. The throughput peaks at 83 at 10% utilization and is maintained up to 50% utilization, indicating that the system has enough RAM to handle tasks efficiently without significant memory management overhead. Task delays at these levels are also minimal, reinforcing the model's efficiency under these conditions. However, as RAM utilization increases to 25% and 50%, the trend changes slightly. Inference time remains stable, and throughput continues at its peak, but task delay begins to climb, reaching 3.1 seconds at 25% and

Table 6.3: Model Performance Comparison

Model Type	**Inference Time (s)**	**Memory Space (MB)**	**GPU Utilization (%)**
Original Model	3.2	20	5%
Quantized Model	2.74	6.5	3.75%
Pruned Model	2.4	13	4.62%
Projected Model	2.1	9.5	4.30%
Proposed Method	**0.9**	**1.35**	**2.68%**

jumping to 11.4 seconds at 50% utilization. These increases suggest that as more RAM is used, the system begins to experience difficulties in managing the available memory resources efficiently.

The most significant changes are observed at a high RAM utilization of 75%. At this level, there is a slight increase in inference time and a minor drop in throughput, but the most noticeable change is the dramatic increase in task delay to 49 seconds. This severe delay indicates significant constraints on the system's ability to manage memory efficiently and can be highly detrimental in environments that require even near-real-time processing. It is important to note that the high RAM utilization necessitates more time spent on memory management to decide which data to keep in RAM. This management overhead causes more delays to other threads and tasks. In addition, as RAM fills up, the operating system starts using disk space as virtual memory to manage overflow, which is much slower than RAM speed. On the GPU front, high utilization can cause thermal throttling, which happens when processors reduce speed to manage heat. Additionally, high GPU utilization lead to queuing and scheduling delays for new tasks due to the resource saturation. These combined effects create a bottleneck at 75% utilization.

Lightweight models are designed to require fewer computational resources, significantly lowering GPU utilization. This reduction in computational demand helps maintain faster inference times and minimizes the likelihood of processing bottlenecks, especially in environments with limited GPU capabilities. Additionally, lightweight CNNs typically have a smaller memory footprint, which reduces RAM utilization. This can minimize memory-related delays, enhancing the model's responsiveness and reliability, particularly in edge computing scenarios where memory resources are often limited.

Table 6.3 provides a detailed comparison of various models in terms of inference time, memory space, and GPU utilization. Notably, the proposed method significantly outperforms the other models across all metrics. With an inference time of just 0.9 seconds, the proposed method is markedly faster than the original model, which takes 3.2 seconds. This substantial improvement in inference time reduces the processing time by over 70%, hence,

enhancing the efficiency of real-time applications. In terms of memory space, the proposed method also excels, requiring only 1.35 MB, when compared to the original model's 20 MB. This substantial decrease in memory usage underscores the method's efficiency and suitability for environments with limited resources. Furthermore, the GPU utilization of the proposed method is only 2.68%, significantly lower than the original model's 5%. This reduction in utilization not only conserves computational resources but also reduces power consumption, making the proposed method more environmentally friendly and cost-effective. Additionally, the proposed method shows improvements over the quantized, pruned, and projected models in all aspects, demonstrating its robustness and superior design. These results indicate that the proposed method not only enhances performance but also optimizes resource usage, making it a compelling choice for future applications.

The performance metrics for the CNN detector under various evasion methods and differing levels of access to the training dataset reveal significant trends both before and after the application of a compression method. Before applying the compression method, the detection performance generally shows a decline as the adversary's access to the training dataset increases, as depicted in Figure 6.3. For FGSM, the detection rate starts at 92.5% when the adversary has 20% access and drops to 87% at 100% access. A similar trend is observed for BIM, starting at 92% and decreasing to 86%. The C&W attack sees a slight drop from 91.5% to 85%. The most significant drop is seen with CGAN, where detection performance jumps from 85.9% to 60.9%. Conversely, the EIR performance demonstrated in Figure 6.4 illustrates that evasion rates escalate with increased adversarial knowledge. This starts at 3.65% for FGSM with 20% access and reaching to 9.38% at full access.

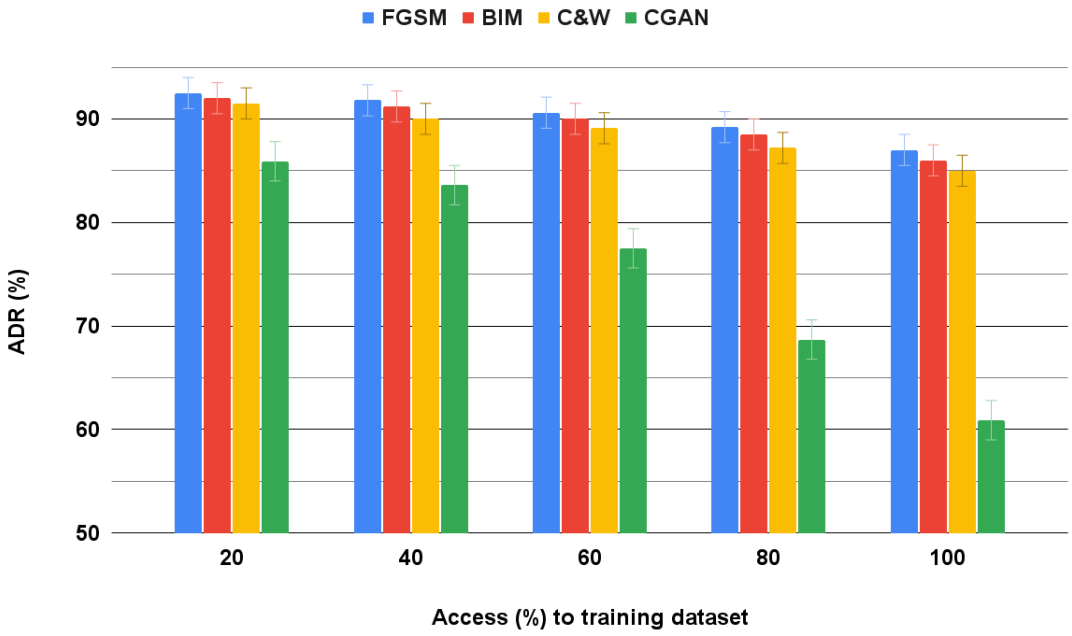


Figure 6.3: ADR (%) of CNN detection model before compression

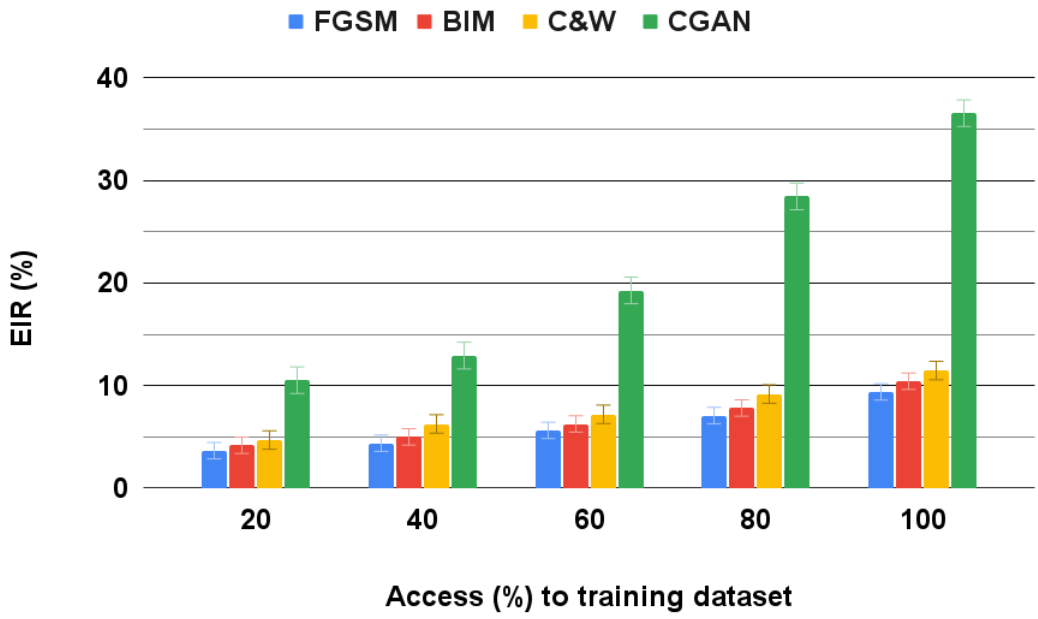


Figure 6.4: EIR (%) of CNN detection model before compression

Similar observations are noticed with BIM going from from 4.17% to 10.42%, C&W from 4.69% to 11.46%. After applying the compression method, there is a minimal reduction in detection performance across all attack methods, as depicted in Figure 6.5. For FGSM, the detection rate decreases from 91% at 20% access to 84% at 100% access, showing a similar trend to the pre-compression results. However, the model's compression strategy has minimal reduction. For instance, FGSM ADR at 20% access is 92.5% and 91% for before and after compression, respectively. We observe similar pattern under the rest of evasion methods. In general, the decrease in the ADR performance is about [1.5-2]%. The EIR performance increase with the same percentage as shown in Figure 6.6.

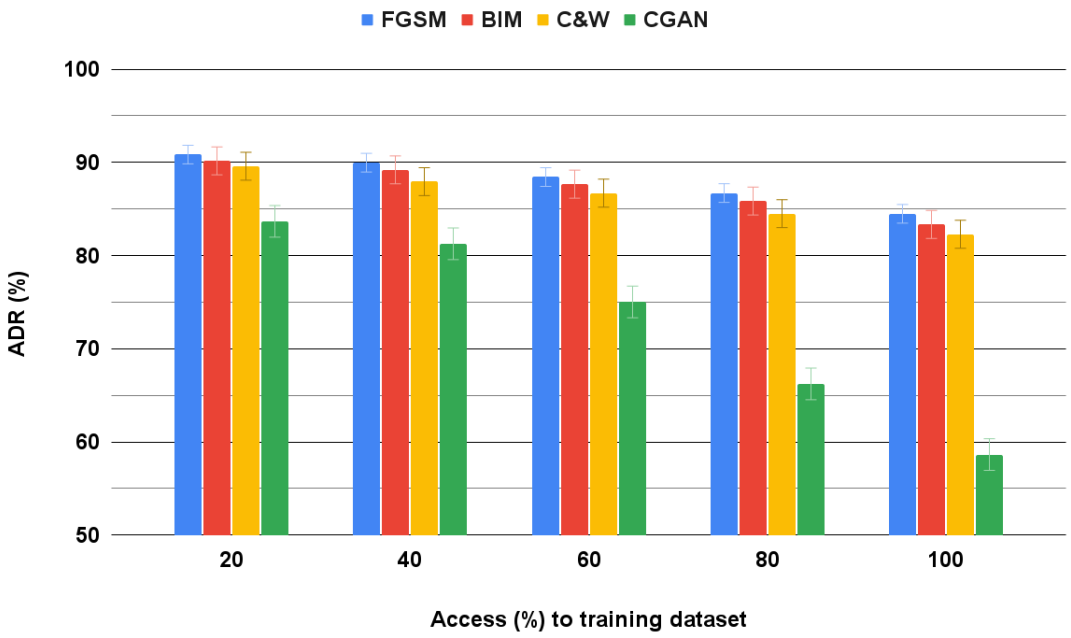


Figure 6.5: ADR (%) of CNN detection model after compression

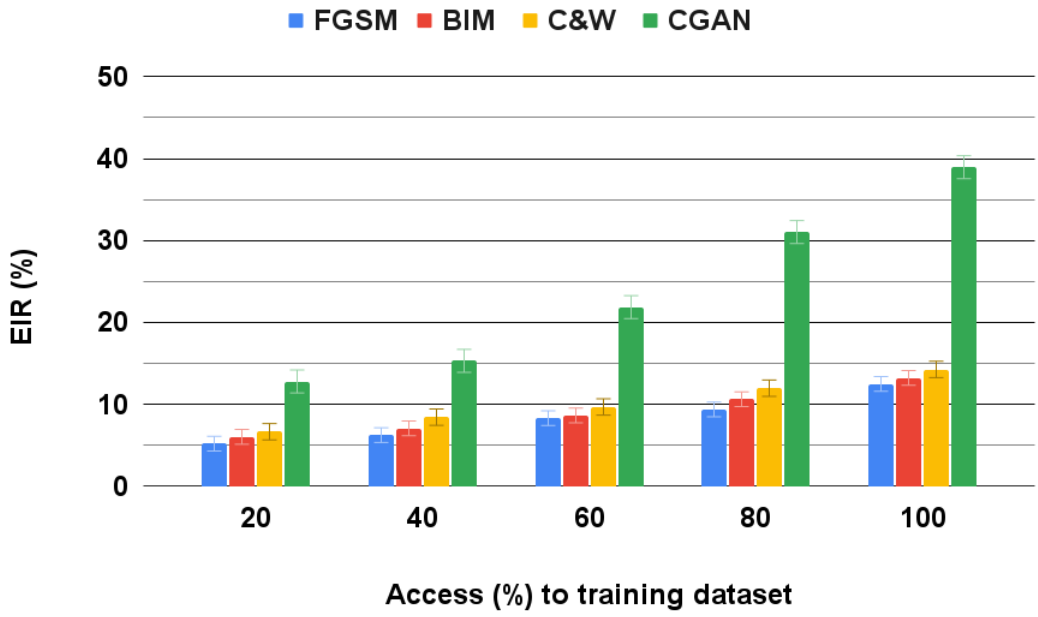


Figure 6.6: EIR (%) of CNN detection model after compression

It is evident that the compression method impacts the detector's overall performance, leading to slightly lower detection rates across all methods and access levels. The decline in detection performance after compression suggests a trade-off between model efficiency and robustness. While compression may improve model efficiency and resource utilization on edge devices, it appears to slightly compromise the detector's ability to accurately identify adversarial attacks.

Table 6.4: Comparison of ADR(%) and EIR(%) before and after compression in a black-box scenario under matching and mis-matching cases.

	Matching case		Mis-matching case	
	Before Compression	After Compression	Before Compression	After Compression
ADR(%)	91.9	89.3	94.8	92.4
EIR(%)	4.27	6.98	1.25	3.75

Table. 6.4 outlines the detection rates (ADR) and the evasion ratios (EIR) of the black-box attack scenario under both matching and mis-matching cases, comparing performances

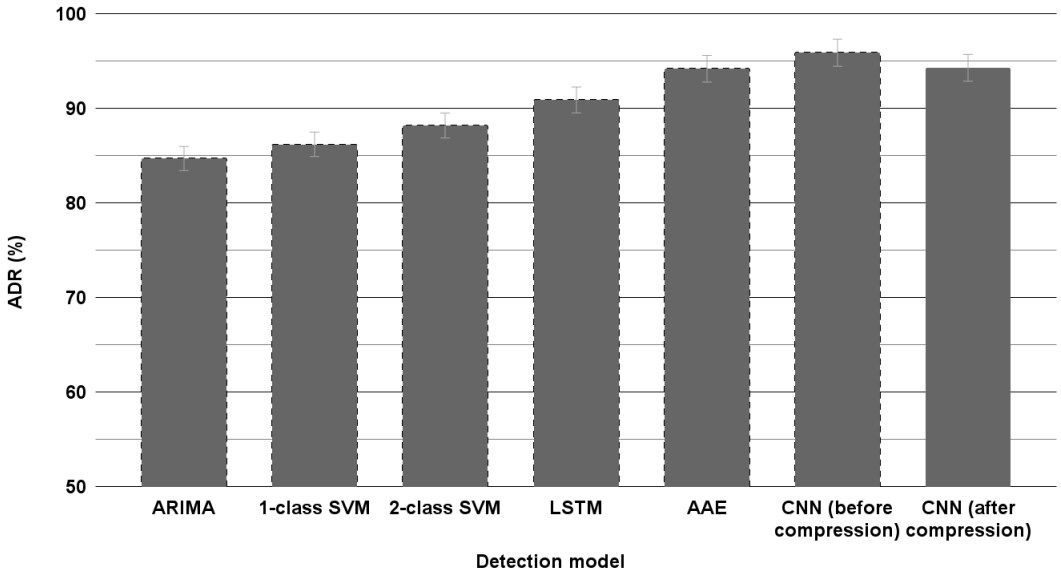


Figure 6.7: Different detection methods against traditional attacks - ADR (%) performance

before and after the application of a compression technique to a CNN model. In this scenario, the adversary does not know the victim's model or have access to the training dataset. In the matching case, where the adversary uses a surrogate model similar to the victim's model, the ADR before compression is 91.9%, and it slightly decreases to 89.3% after compression. This indicates a modest reduction in detection capability when the model is compressed. The slight decrease in performance could be attributed to the loss of some nuanced features or slight overfitting to the compressed model's characteristics, making it slightly less effective at detecting very closely aligned adversarial examples.
For the mis-matching case, where the adversary's surrogate model is different from the victim's model, the ADR before compression is higher at 94.8%, but it shows a decrease to 92.4% after compression. This suggests that the model, while still robust, loses some of its effectiveness against less similar adversarial models when compressed. The higher initial detection rate in the mis-matching case is due to the model being better at recognizing deviations from its learned patterns, which are more pronounced when the adversarial model differs more significantly. The results show that our CNN detector is robust in detecting adversarial attacks under black-box scenarios, even after compression.

In addition, we test our CNN detector under traditional attack scenarios that were defined in the simulation settings. Figure 6.7 presents the ADR for various models when sub-

jected to traditional attacks, where adversaries alter input values deterministically rather than using sophisticated algorithmic approaches such as FGSM and BIM. The results highlight the varying effectiveness of different detection models, from statistical methods like ARIMA to more complex deep learning approaches like CNN model. The ARIMA model has an ADR of 84.7%, the lowest among the detection models, suggesting limited effectiveness. As we move to machine learning models such as the 1-class SVM and 2-class SVM, there is a gradual improvement in detection rates, with ADRs of 86.2% and 88.2%, respectively. This increment indicates better handling and identification of adversarial inputs as model complexity increases. The LSTM model shows a more significant improvement, resulting in an ADR of 90.9%. This suggests better detection capabilities for attacks that may have temporal dependencies. Similarly, the AAE model achieves an ADR of 94.2%, indicating its strength in detecting sophisticated adversarial samples. The CNN model showcases the highest detection rate of 95.9% before compression. Although, the detection rate slightly decreases to 94.3% after compression, the model remains highly effective. This shows that our CNN detector is robust and effective against traditional attacks, even with compression. In Figure 6.8, we show the EIR performance for the detection models under traditional attacks. We noticed that our CNN detection model has the least evasion rate. This demonstrates the robustness of our detection model.

6.4 Summary

In summary, we focused on the challenge of adapting adversarial detection models for edge devices in V2M settings, where computational power and memory are limited. The primary goal is to reduce the size of the detection model without compromising its performance. Various AI model compression techniques, including projection, pruning, and quantization, were discussed. Our proposed compression method integrates model design and compression into a single process, resulting in an optimized detection model that maintains robust performance against adversarial attacks. This integrated approach ensures that the model remains compact and capable of maintaining high accuracy in V2M edge environments. For instance, before compression, the CNN model's detection rate against FGSM attacks is 92.5% when the adversary has 20% access to the training dataset, dropping to 87% at 100% access. After compression, there is only a modest reduction in performance, demonstrating the effectiveness of our approach.
We also analyze the performance metrics of a CNN model under different levels of GPU and RAM utilization. At 10% GPU utilization, the model demonstrates an inference time of 2 seconds and a throughput of 83. However, as utilization increases to 75%, inference

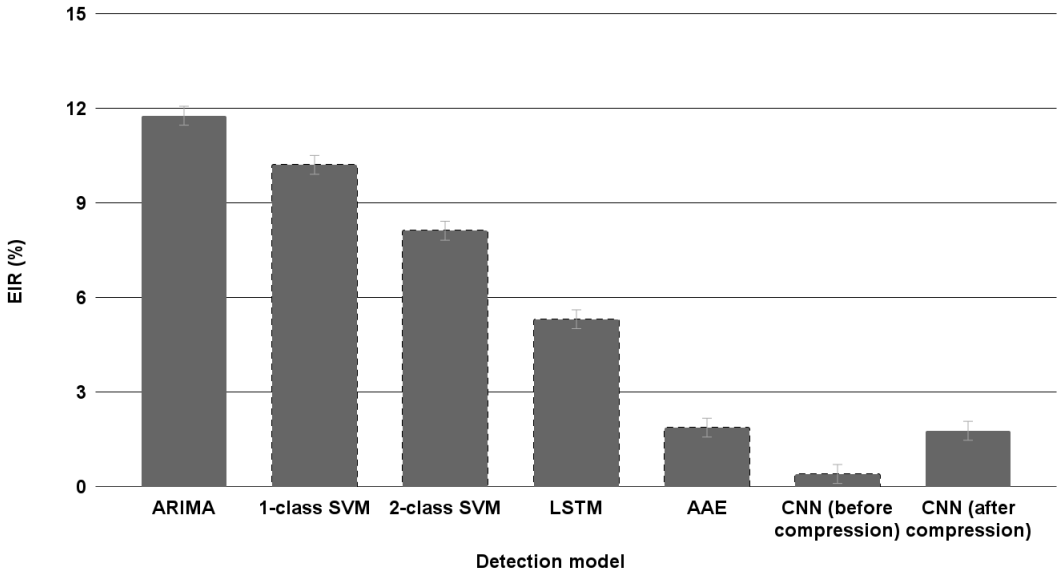


Figure 6.8: Different detection methods against traditional attacks - EIR (%) performance

time slightly increases to 2.2 seconds, and throughput drops to 78, with task delay dramatically increasing to 35 seconds. Similarly, at 10% RAM utilization, the inference time is stable at 2 seconds with a throughput of 83, but at 75% RAM utilization, task delay spikes to 49 seconds, indicating significant performance degradation.
Finally, we evaluate the CNN detector's performance against various evasion and traditional attack scenarios before and after compression. For example, before compression, the detection rate in a black-box attack scenario is 91.9% for matching cases and 94.8% for mis-matching cases. After compression, these rates slightly decrease to 89.3% and 92.4%, respectively. In traditional attack scenarios, the CNN model achieves the highest detection rate of 95.9% before compression, which slightly decreases to 94.3% after compression.

Chapter 7

Conclusion and Future Remarks

The integration of Electric Vehicles (EVs), smart grids, and Artificial Intelligence (AI) in Vehicle-to-Microgrid (V2M) applications presents security challenges that need to be addressed. This thesis proposal aimed to investigate and mitigate these challenges to enhance the resilience and reliability of V2M systems. The motivation behind this work stems from the remarkable advancements in communication technologies, including cellular networks, Wi-Fi, and optical communication. While these technologies have significantly impacted people's daily lives and improved cities' preparedness for power outages through real-time monitoring of transmission and distribution lines, they have also expanded the attack surface and introduced new cyber vulnerabilities within smart grids. Adversaries now have opportunities to exploit these vulnerabilities and launch devastating cyber attacks against smart grids. The objectives of this research are four-fold. First, it sought to investigate data integrity attacks in V2M systems, focusing on understanding the growing attack surface resulting from the integration of EVs, smart grids, and AI. Second, the research aimed to analyze Adversarial Machine Learning attacks specifically targeting V2M services, considering multi-stage gray-box attacks as potential threat scenarios. Third, a comprehensive defense framework was proposed to enhance the security of AI-based microgrid control systems in V2M applications. Lastly, a compact detection model that is well-suited for edge environment without compromising the detection performance.

The contributions of this work were significant. First, a novel scheme was developed to model and detect data integrity attacks in V2M systems, employing unsupervised machine learning techniques as an intelligent defense mechanism. Through simulations, the effectiveness of this scheme in mitigating the impact of data integrity attacks was demonstrated, achieving a reduction in their impact by up to 76.5%. Additionally, we conducted an anticipatory analysis of multi-stage gray-box attacks, exploring potential vulnerabilities

and devising effective countermeasures. Moreover, we studied the potential of a severe impact of a multi-stage gray-box attack, showing an Evasion Increase Rate (EIR) of up to 73.2%, using 40% less data than traditional white-box attacks.

To fortify the security of AI-based microgrid control systems in V2M services, we proposed a comprehensive defense framework, where we integrated a Generative Adversarial Network (GAN) model to generate realistic adversarial samples, enabling the ML classifier to learn and adapt to novel attack patterns. The ML classifier was trained on a diverse dataset, encompassing both legitimate and adversarial samples, to improve its ability to accurately distinguish between normal and malicious activities. Through simulations, the proposed defense mechanism achieved an impressive Adversarial Detection Rate (ADR) of 90.2%, significantly outperforming the baseline method (DBSCAN) with an ADR of up to 50.3%. Furthermore, this research presented a thorough analysis of the impact associated with the adversary's knowledge of the system. Five access cases to the victim's ML classifier training dataset were examined, including gray-box cases with different access percentages (20%, 40%, 60%, or 80%) and a white-box case with 100% access. We also assessed the effectiveness of DBSCAN method as a baseline solution in detecting adversarial attacks. We showed that DBSCAN method failed to block attacks when adversaries use a CGAN model, resulting in a decreased ADR of 48%. In contrast, the proposed GAN-based detector exceeded the limitations of DBSCAN and effectively detect adversarial attacks, achieving an ADR of 84% under scenario-1, and maintaining high ADR percentages of 79.6% and 81.6% under scenario-2 and scenario-3, respectively.

To address the limited computational power and memory in V2M edge settings, we studied different model optimization (i.e., model design and compression) to optimize the model's size without compromising the detection performance. Several AI model compression techniques, such as projection, pruning, and quantization were discussed. Our proposed method integrates model design and compression, resulting in an optimized detection model that remains robust against adversarial attacks. For example, the detection rate against FGSM attacks drops from 92.5% to 87% as adversary access increases from 20% to 100%, but the performance remains strong after compression. We analyze the performance of a CNN model under different GPU and RAM utilization levels. At 10% GPU utilization, the model shows an inference time of 2 seconds and a throughput of 83. At 75% utilization, inference time increases slightly to 2.2 seconds, throughput drops to 78, and task delay rises to 35 seconds. Similarly, at 10% RAM utilization, inference time is stable at 2 seconds with throughput at 83, but at 75% RAM utilization, task delay jumps to 49 seconds. The performance of the CNN detector against various attack scenarios is evaluated before and after compression. In black-box attack scenarios, the detection rate is 91.9% for matching cases and 94.8% for mis-matching cases before compression, slightly

decreasing to 89.3% and 92.4% after compression. In traditional attack scenarios, the CNN model achieves a detection rate of 95.9% before compression and 94.3% after compression.

This thesis establishes a guideline for future research directions in the area of securing V2M systems, with several opportunities for extending the current work. Throughout this work, we have identified and mitigated several less probable attack surfaces. Notably, the simulations presented in Chapter 3 highlight the complexities involved in launching successful attacks on the vehicle's contribution factor. Our findings suggest that even if adversaries possess the necessary skills for a successful attack, scaling such an operation to significantly impact V2M operations remains a challenge. Consequently, vehicles have been discounted as viable targets for disrupting V2M operations.
In contrast, Chapter 4 demonstrated that smart meters present a vulnerable attack surface, making them ideal targets for evasion attacks. A promising future direction would involve exploring the use of transformers for synthetic data generation, given their demonstrated efficacy in similar applications. This could potentially enhance the resilience of data-driven models against adversarial attacks.
Additionally, in Chapter 5, we explored the use of GAN models to detect adversarial attacks but relied on training a separate classifier for this purpose. Future research could explore the feasibility of integrating adversarial detection directly into the GAN model. For instance, the discriminator within a GAN could be repurposed as the primary classifier to identify adversarial examples, thereby eliminating the need for a separate classifier. This integration could reduce computational overhead and further optimize the system for resource-constrained edge devices, however, this would also necessitate robust performance evaluations to ensure detection accuracy is not compromised.
Finally, as Chapter 6 focused on optimizing AI models to run efficiently on edge devices, an essential next step would be to extend these optimizations by incorporating techniques such as federated learning to distribute the training load across multiple devices. Leveraging the distributed nature of edge devices can enhance the model performance without significantly increasing the computational burden on individual devices. Moreover, investigating the impact of various compression techniques, such as weight pruning or quantization-aware training, in combination with these distributed learning approaches, could lead to a highly efficient and robust adversarial detection framework suitable for real-time V2M systems.

References

[1] Rahat Hossain, Amanullah Maung, Aman Than Oo, and Shawkat Ali. *Smart Grid*, volume 132. 11 2013.

[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] SR Vijayan. Smart grid technologies: Distribution automation, microgrids, and cyber security. In *ISGW 2017: Compendium of Technical Papers*, pages 29–40. Springer, 2018.

[4] Larisa Dobriansky, Girish Ghatikar, and Dan Ton. Smart microgrids: Re-visioning smart grid and smart city development in india. In *ISGW 2017: Compendium of Technical Papers*, pages 273–288. Springer, 2018.

[5] Tie Qiu, Kaiyu Zheng, Houbing Song, Min Han, and Burak Kantarci. A local-optimization emergency scheduling scheme with self-recovery for a smart grid. *IEEE Transactions on Industrial Informatics*, 13(6):3195–3205, 2017.

[6] H. S. V. S. Kumar Nunna and Dipti Srinivasan. Multiagent-based transactive energy framework for distribution systems with smart microgrids. *IEEE Transactions on Industrial Informatics*, 13(5):2241–2250, 2017.

[7] Medhat Elsayed, Melike Erol-Kantarci, Burak Kantarci, Lei Wu, and Jie Li. Low-latency communications for community resilience microgrids: A reinforcement learning approach. *IEEE Transactions on Smart Grid*, 2019.

[8] Ahmed Mohamed Elsayed Omara. *Predictive Operational Strategies for Smart Microgrid Networks*. PhD thesis, University of Ottawa, 2020.

[9] Ahmed Omara, Wendong Yuan, Michele Nogueira, Burak Kantarci, and Lei Wu. Microgrid data aggregation and wireless transfer scheduling in the presence of time sensitive events. In *Proceedings of the 16th ACM International Symposium on Mobility Management and Wireless Access*, MobiWac'18, page 109–112, New York, NY, USA, 2018. Association for Computing Machinery.

[10] Murat Simsek, Ahmed Omara, and Burak Kantarci. Cost-aware data aggregation and energy decentralization with electrical vehicles in microgrids through lte links. In *IEEE Intl Black Sea Conference on Communications and Networking (BlackSeaCom)*, pages 1–6. IEEE, 2020.

[11] Mohammad Hossein Sarparandeh and M. Ehsan. Pricing of vehicle-to-grid services in a microgrid by nash bargaining theory. 2017.

[12] Arwa O Erick and Komla A Folly. Reinforcement learning approaches to power management in grid-tied microgrids: A review. In *2020 Clemson University Power Systems Conference (PSC)*, pages 1–6. IEEE, 2020.

[13] A. Omara, B. Kantarci, M. Nogueira, M. Erol-Kantarci, L. Wu, and J. Li. Delay sensitivity-aware aggregation of smart microgrid data over heterogeneous networks. In *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pages 1–7, 2019.

[14] Willett Kempton and Steven E Letendre. Electric vehicles as a new power source for electric utilities. *Transportation Research Part D: Transport and Environment*, 2(3):157–175, 1997.

[15] Metin Kesler, Mithat C Kisacikoglu, and Leon M Tolbert. Vehicle-to-grid reactive power operation using plug-in electric vehicle bidirectional offboard charger. *IEEE Transactions on Industrial Electronics*, 61(12):6778–6784, 2014.

[16] Mithat C Kisacikoglu, Metin Kesler, and Leon M Tolbert. Single-phase on-board bidirectional pev charger for v2g reactive power operation. *IEEE Transactions on smart grid*, 6(2):767–775, 2014.

[17] Ford Canada. FAQs: F-150 Lightning Intelligent Backup Power. Accessed March, 2024.

[18] Rashid Khan, Khawaja Khalid Mehmood, Syed Basit Ali Bukhari, Kashif Imran, Abdul Wadood, Sang Bong Rhee, and Sina Park. An optimization-based reliability enhancement scheme for active distribution systems utilizing electric vehicles. *IEEE Access*, 9:157247–157258, 2021.

[19] Qingyu Yang, Donghe Li, Dou An, Wei Yu, Xinwen Fu, Xinyu Yang, and Wei Zhao. Towards incentive for electrical vehicles demand response with location privacy guaranteeing in microgrids. *IEEE Transactions on Dependable and Secure Computing*, 19(1):131–148, 2020.

[20] Ke Zhang, Yuming Mao, Supeng Leng, Sabita Maharjan, Yan Zhang, Alexey Vinel, and Magnus Jonsson. Incentive-driven energy trading in the smart grid. *IEEE Access*, 4:1243–1257, 2016.

[21] Shubhani Aggarwal and Neeraj Kumar. A consortium blockchain-based energy trading for demand response management in vehicle-to-grid. *IEEE Transactions on Vehicular Technology*, 70(9):9480–9494, 2021.

[22] HSVS Kumar Nunna, Swathi Battula, Suryanarayana Doolla, and Dipti Srinivasan. Energy management in smart distribution systems with vehicle-to-grid integrated microgrids. *IEEE Transactions on Smart Grid*, 9(5):4004–4016, 2016.

[23] Vivian Sultan, Arun Aryal, Hao Chang, and Jiri Kral. Integration of evs into the smart grid: A systematic literature review. *Energy Informatics*, 5(1):65, 2022.

[24] Bo Zeng, Jiahuan Feng, Nian Liu, and Yixian Liu. Co-optimized parking lot placement and incentive design for promoting pev integration considering decision-dependent uncertainties. *IEEE Transactions on Industrial Informatics*, 17(3):1863–1872, 2020.

[25] Abdullah Al-obaidi and Hany EZ Farag. Optimal design of v2g incentives and v2g-capable electric vehicles parking lots considering cost-benefit financial analysis and user participation. *IEEE Transactions on Sustainable Energy*, 2023.

[26] Dai Wang, Jonathan Coignard, Teng Zeng, Cong Zhang, and Samveg Saxena. Quantifying electric vehicle battery degradation from driving vs. vehicle-to-grid services. *Journal of Power Sources*, 332:193–203, 2016.

[27] Huaqun Wang, Qihua Wang, Debiao He, Qi Li, and Zhe Liu. Bbars: Blockchain-based anonymous rewarding scheme for v2g networks. *IEEE Internet of Things Journal*, 6(2):3676–3687, 2019.

[28] Md Shamiur Rahman, Md Jahangir Hossain, Junwei Lu, Fida Hasan Md Rafi, and Sukumar Mishra. A vehicle-to-microgrid framework with optimization-incorporated distributed ev coordination for a commercial neighborhood. *IEEE Transactions on Industrial Informatics*, 16(3):1788–1798, 2019.

[29] Jinxin Liu, Michele Nogueira, Johan Fernandes, and Burak Kantarci. Adversarial machine learning: A multilayer review of the state-of-the-art and challenges for wireless and mobile systems. *IEEE Communications Surveys & Tutorials*, 24(1):123–159, 2021.

[30] Koosha Sadeghi, Ayan Banerjee, and Sandeep KS Gupta. A system-driven taxonomy of attacks and defenses in adversarial machine learning. *IEEE transactions on emerging topics in computational intelligence*, 4(4):450–467, 2020.

[31] Baoyuan Wu, Li Liu, Zihao Zhu, Qingshan Liu, Zhaofeng He, and Siwei Lyu. Adversarial machine learning: A systematic survey of backdoor attack, weight attack and adversarial example. *arXiv preprint arXiv:2302.09457*, 2023.

[32] Tianyu Gu, Kang Liu, Brendan Dolan-Gavitt, and Siddharth Garg. Badnets: Evaluating backdooring attacks on deep neural networks. *IEEE Access*, 7:47230–47244, 2019.

[33] Yingqi Liu, Shiqing Ma, Yousra Aafer, Wen-Chuan Lee, Juan Zhai, Weihang Wang, and Xiangyu Zhang. Trojaning attack on neural networks. In *25th Annual Network And Distributed System Security Symposium (NDSS 2018)*. Internet Soc, 2018.

[34] Yugeng Liu, Rui Wen, Xinlei He, Ahmed Salem, Zhikun Zhang, Michael Backes, Emiliano De Cristofaro, Mario Fritz, and Yang Zhang. {ML-Doctor}: Holistic risk assessment of inference attacks against machine learning models. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 4525–4542, 2022.

[35] Nicholas Carlini, Steve Chien, Milad Nasr, Shuang Song, Andreas Terzis, and Florian Tramer. Membership inference attacks from first principles. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 1897–1914. IEEE, 2022.

[36] Anish Athalye and Nicholas Carlini. On the robustness of the cvpr 2018 white-box adversarial example defenses. *arXiv preprint arXiv:1804.03286*, 2018.

[37] Yalin E Sagduyu, Tugba Erpek, and Yi Shi. Adversarial machine learning for 5g communications security. *Game Theory and Machine Learning for Cyber Security*, pages 270–288, 2021.

[38] Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z Berkay Celik, and Ananthram Swami. Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security*, pages 506–519, 2017.

[39] Wieland Brendel, Jonas Rauber, and Matthias Bethge. Decision-based adversarial attacks: Reliable attacks against black-box machine learning models. *arXiv preprint arXiv:1712.04248*, 2017.

[40] Ying Xu, Xu Zhong, Antonio Jimeno Yepes, and Jey Han Lau. Grey-box adversarial attack and defence for sentiment classification. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4078–4087, 2021.

[41] Ahmed Omara and Burak Kantarci. On the impact of data integrity attacks on vehicle-to-microgrid services. In *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1–7. IEEE, 2021.

[42] Heju Jiang, Jasvir Nagra, and Parvez Ahammad. Sok: Applying machine learning in security-a survey. *arXiv preprint arXiv:1611.03186*, 2016.

[43] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael P Wellman. Sok: Security and privacy in machine learning. In *2018 IEEE European symposium on security and privacy (EuroS&P)*, pages 399–414. IEEE, 2018.

[44] Felix O Olowononi, Danda B Rawat, and Chunmei Liu. Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for cps. *IEEE Communications Surveys & Tutorials*, 23(1):524–552, 2020.

[45] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.

[46] Alexey Kurakin, Ian J Goodfellow, and Samy Bengio. Adversarial examples in the physical world. In *Artificial intelligence safety and security*, pages 99–112. Chapman and Hall/CRC, 2018.

[47] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. *stat*, 1050(9), 2017.

[48] Nicolas Papernot, Patrick McDaniel, Somesh Jha, Matt Fredrikson, Z Berkay Celik, and Ananthram Swami. The limitations of deep learning in adversarial settings. In *2016 IEEE European symposium on security and privacy (EuroS&P)*, pages 372–387. IEEE, 2016.

[49] Nicholas Carlini and David Wagner. Towards evaluating the robustness of neural networks. In *2017 ieee symposium on security and privacy (sp)*, pages 39–57. Ieee, 2017.

[50] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[51] Russell Eberhart and James Kennedy. Particle swarm optimization. In *Proceedings of the IEEE international conference on neural networks*, volume 4, pages 1942–1948. Citeseer, 1995.

[52] David E Goldberg. Genetic and evolutionary algorithms come of age. *Communications of the ACM*, 37(3):113–120, 1994.

[53] Maryam M Najafabadi, Taghi M Khoshgoftaar, Clifford Kemp, Naeem Seliya, and Richard Zuech. Machine learning for detecting brute force attacks at the network level. In *2014 IEEE International Conference on Bioinformatics and Bioengineering*, pages 379–385. IEEE, 2014.

[54] Pin-Yu Chen, Huan Zhang, Yash Sharma, Jinfeng Yi, and Cho-Jui Hsieh. Zoo: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models. In *Proceedings of the 10th ACM workshop on artificial intelligence and security*, pages 15–26, 2017.

[55] Ishai Rosenberg, Asaf Shabtai, Yuval Elovici, and Lior Rokach. Defense methods against adversarial examples for recurrent neural networks. *arXiv preprint arXiv:1901.09963*, 2019.

[56] Yi Shi and Yalin E Sagduyu. Evasion and causative attacks with adversarial deep learning. In *MILCOM 2017-2017 IEEE Military Communications Conference (MILCOM)*, pages 243–248. IEEE, 2017.

[57] Zhiyan Chen and Burak Kantarci. Generative adversarial network-driven detection of adversarial tasks in mobile crowdsensing. In *ICC 2022 - IEEE International Conference on Communications*, pages 2780–2785, 2022.

[58] Md Ahsan Ayub, William A Johnson, Douglas A Talbert, and Ambareen Siraj. Model evasion attack on intrusion detection systems using adversarial machine learning. In *2020 54th annual conference on information sciences and systems (CISS)*, pages 1–6. IEEE, 2020.

[59] Amol Damare, Shouvik Roy, Scott A Smolka, and Scott D Stoller. A barrier certificate-based simplex architecture with application to microgrids. In *Runtime Verification: 22nd International Conference, RV 2022, Tbilisi, Georgia, September 28–30, 2022, Proceedings*, pages 105–123. Springer, 2022.

[60] Zhang Guihai and Biplab Sikdar. Adversarial machine learning against false data injection attack detection for smart grid demand response. In *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 352–357. IEEE, 2021.

[61] Guihai Zhang and Biplab Sikdar. Ensemble and transfer adversarial attack on smart grid demand-response mechanisms. In *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 53–58. IEEE, 2022.

[62] Jiwei Tian, Buhong Wang, Jing Li, and Zhen Wang. Adversarial attacks and defense for cnn based power quality recognition in smart grid. *IEEE Transactions on Network Science and Engineering*, 9(2):807–819, 2021.

[63] Sanglee Park and Jungmin So. On the effectiveness of adversarial training in defending against adversarial example attacks for image classification. *Applied Sciences*, 10(22):8079, 2020.

[64] Azwirman Gusrialdi and Zhihua Qu. Smart grid security: Attacks and defenses. *Smart Grid Control: Overview and Research Opportunities*, pages 199–223, 2019.

[65] Zubair A Baig and Abdul-Raoof Amoudi. An analysis of smart grid attacks and countermeasures. *J. Commun.*, 8(8):473–479, 2013.

[66] Eirini Anthi, Lowri Williams, Matilda Rhode, Pete Burnap, and Adam Wedgbury. Adversarial attacks on machine learning cybersecurity defences in industrial control systems. *Journal of Information Security and Applications*, 58:102717, 2021.

[67] Parya Haji Mirzaee, Mohammad Shojafar, Haitham Cruickshank, and Rahim Tafazolli. Smart grid security and privacy: From conventional to machine learning issues (threats and countermeasures). *IEEE access*, 10:52922–52954, 2022.

[68] Ali Sayghe, Junbo Zhao, and Charalambos Konstantinou. Evasion attacks with adversarial deep learning against power system state estimation. In *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2020.

[69] Yi Liang, Di He, and Deming Chen. Poisoning attack on load forecasting. In *2019 IEEE innovative smart grid technologies-Asia (ISGT Asia)*, pages 1230–1235. IEEE, 2019.

[70] Abdulrahman Takiddin, Muhammad Ismail, Usman Zafar, and Erchin Serpedin. Robust electricity theft detection against data poisoning attacks in smart grids. *IEEE Transactions on Smart Grid*, 12(3):2675–2684, 2020.

[71] Yize Chen, Yushi Tan, and Deepjyoti Deka. Is machine learning in power systems vulnerable? In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 1–6. IEEE, 2018.

[72] Ahmed Omara and Burak Kantarci. Adversarial machine learning-based anticipation of threats against vehicle-to-microgrid services. In *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pages 1844–1849, 2022.

[73] Tian Liu and Tao Shu. Adversarial false data injection attack against nonlinear ac state estimation with ann in smart grid. In *Security and Privacy in Communication Networks: 15th EAI International Conference, SecureComm 2019, Orlando, FL, USA, October 23–25, 2019, Proceedings, Part II 15*, pages 365–379. Springer, 2019.

[74] Milan Biswal, Satyajayant Misra, and Abu S Tayeen. Black box attack on machine learning assisted wide area monitoring and protection systems. In *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2020.

[75] Junfei Wang and Pirathayini Srikantha. Stealthy black-box attacks on deep learning non-intrusive load monitoring models. *IEEE Transactions on Smart Grid*, 12(4):3479–3492, 2021.

[76] Islam Elgarhy, Ahmed T El-Toukhy, Mahmoud M Badr, Mohamed Mahmoud, Mostafa M Fouda, Maazen Alsabaan, and Hisham A Kholidy. Secured cluster-based electricity theft detectors against blackbox evasion attacks. In *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, pages 333–338. IEEE, 2024.

[77] Yun Lin, Haojun Zhao, Ya Tu, Shiwen Mao, and Zheng Dou. Threats of adversarial attacks in dnn-based modulation recognition. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pages 2469–2478, 2020.

[78] Meysam Sadeghi and Erik G. Larsson. Adversarial attacks on deep-learning based radio signal classification. *IEEE Wireless Communications Letters*, 8(1):213–216, 2019.

[79] Yalin E. Sagduyu, Yi Shi, and Tugba Erpek. Iot network security from the perspective of adversarial deep learning. In *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 1–9, 2019.

[80] Yi Shi, Yalin E. Sagduyu, Tugba Erpek, Kemal Davaslioglu, Zhuo Lu, and Jason H. Li.

[81] Muhammad Usama, Adnan Qayyum, Junaid Qadir, and Ala Al-Fuqaha. Black-box adversarial machine learning attack on network traffic classification. In *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pages 84–89. IEEE, 2019.

[82] Damilola Adesina, Chung-Chu Hsieh, Yalin E. Sagduyu, and Lijun Qian. Adversarial machine learning in wireless communications using rf data: A review. *IEEE Communications Surveys Tutorials*, pages 1–1, 2022.

[83] Yi Shi, Kemal Davaslioglu, and Yalin E. Sagduyu. Generative adversarial network for wireless signal spoofing. In *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*, WiseML 2019, page 55–60, New York, NY, USA, 2019. Association for Computing Machinery.

[84] Adnan Qayyum, Muhammad Usama, Junaid Qadir, and Ala Al-Fuqaha. Securing connected autonomous vehicles: Challenges posed by adversarial machine learning and the way forward. *IEEE Communications Surveys Tutorials*, 22(2):998–1026, 2020.

[85] Murat Kuzlu, Manisa Pipattanasomporn, and Saifur Rahman. Communication network requirements for major smart grid applications in HAN, NAN and WAN. *Computer Networks*, 67:74–88, 2014.

[86] Ahmed Omara, Burak Kantarci, Michele Nogueira, Melike Erol-Kantarci, Lei Wu, and Jie Li. Delay sensitivity-aware aggregation of smart microgrid data over heterogeneous networks. In *IEEE International Conference on Communications (ICC)*, pages 1–7, 2019.

[87] Yasin Kabalci. A survey on smart metering and smart grid communication. *Renewable and Sustainable Energy Reviews*, 57:302–318, 2016.

[88] Gogulamudi Pradeep Reddy, Yellapragada Venkata Pavan Kumar, and Maddikera Kalyan Chakravarthi. Communication technologies for interoperable smart microgrids in urban energy community: A broad review of the state of the art, challenges, and research perspectives. *Sensors*, 22(15), 2022.

[89] Maedeh Ghorbanian, Sarineh Hacopian Dolatabadi, Maryam Masjedi, and Pierluigi Siano. Communication in smart grids: A comprehensive review on the existing and future communication and information infrastructures. *IEEE Systems Journal*, 13(4):4001–4014, 2019.

[90] Tania Gupta and Richa Bhatia. Communication technologies in smart grid at different network layers: An overview. In *2020 International Conference on Intelligent Engineering and Management (ICIEM)*, pages 177–182. IEEE, 2020.

[91] Brian Kim, Yalin E Sagduyu, Kemal Davaslioglu, Tugba Erpek, and Sennur Ulukus. Over-the-air adversarial attacks on deep learning based modulation classifier over wireless channels. In *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6. IEEE, 2020.

[92] Zhengping Luo, Shangqing Zhao, Zhuo Lu, Yalin E Sagduyu, and Jie Xu. Adversarial machine learning based partial-model attack in iot. In *Proceedings of the 2nd ACM workshop on wireless security and machine learning*, pages 13–18, 2020.

[93] Yi Shi, Kemal Davaslioglu, and Yalin E Sagduyu. Over-the-air membership inference attacks as privacy threats for deep learning-based wireless signal classifiers. In *Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning*, pages 61–66, 2020.

[94] Matthew DelVecchio, Vanessa Arndorfer, and William C Headley. Investigating a spectral deception loss metric for training machine learning-based evasion attacks. In *Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning*, pages 43–48, 2020.

[95] Yi Shi, Tugba Erpek, Yalin E Sagduyu, and Jason H Li. Spectrum data poisoning with adversarial deep learning. In *MILCOM 2018-2018 IEEE Military Communications Conference (MILCOM)*, pages 407–412. IEEE, 2018.

[96] Yun Lin, Haojun Zhao, Ya Tu, Shiwen Mao, and Zheng Dou. Threats of adversarial attacks in dnn-based modulation recognition. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pages 2469–2478. IEEE, 2020.

[97] Samurdhi Karunaratne, Enes Krijestorac, and Danijela Cabric. Penetrating rf fingerprinting-based authentication with a generative adversarial attack. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[98] Brian Kim, Yalin E Sagduyu, Kemal Davaslioglu, Tugba Erpek, and Sennur Ulukus. How to make 5g communications" invisible": Adversarial machine learning for wireless privacy. In *2020 54th Asilomar Conference on Signals, Systems, and Computers*, pages 763–767. IEEE, 2020.

[99] Wenye Wang, Yi Xu, and Mohit Khanna. A survey on the communication architectures in smart grid. *Computer Networks*, 55(15):3604–3629, 2011.

[100] Emilio Ancillotti, Raffaele Bruno, and Marco Conti. The role of communication systems in smart grids: Architectures, technical solutions and research challenges. *Computer Communications*, 36(17-18):1665–1697, 2013.

[101] Shichao Liu, Xiaoping P Liu, and Abdulmotaleb El Saddik. Modeling and distributed gain scheduling strategy for load frequency control in smart grids with communication topology changes. *ISA transactions*, 53(2):454–461, 2014.

[102] BN Lazarus. Smart grid enabled and enhanced by broadband powerline. In *Proceedings of ENERGY 2013, the Third International Conference on Smart Grids, Green Communications and IT Energy-Aware Technologies, Lisbon, Portugal*, pages 77–83, 2013.

[103] Nico Saputro, Kemal Akkaya, and Suleyman Uludag. A survey of routing protocols for smart grid communications. *Computer Networks*, 56(11):2742–2771, 2012.

[104] Stefano Galli, Anna Scaglione, and Zhifang Wang. For the grid and through the grid: The role of power line communications in the smart grid. *Proceedings of the IEEE*, 99(6):998–1027, 2011.

[105] Cristina Cano, Alberto Pittolo, David Malone, Lutz Lampe, Andrea M Tonello, and Anand G Dabak. State of the art in power line communications: From the applications to the medium. *IEEE Journal on Selected Areas in Communications*, 34(7):1935–1952, 2016.

[106] N Shaukat, SM Ali, CA Mehmood, B Khan, M Jawad, U Farid, Z Ullah, SM Anwar, and M Majid. A survey on consumers empowerment, communication technologies, and renewable generation penetration within smart grid. *Renewable and Sustainable Energy Reviews*, 81:1453–1475, 2018.

[107] Ekram Hossain, Zhu Han, and H Vincent Poor. *Smart grid communications and networking*. Cambridge University Press, 2012.

[108] Nima Zaker, Burak Kantarci, Melike Erol-Kantarci, and Hussein T Mouftah. Smart grid monitoring with service differentiation via epon and wireless sensor network convergence. *Optical Switching and Networking*, 14:53–68, 2014.

[109] Stuart Borlase. *Smart grids: infrastructure, technology, and solutions*. CRC press, 2017.

[110] Quang-Dung Ho, Yue Gao, Gowdemy Rajalingham, and Tho Le-Ngoc. *Wireless communications networks for the smart grid*, volume 2. Springer, 2014.

[111] Saida Elyengui, Riadh Bouhouchi, and Tahar Ezzedine. The enhancement of communication technologies and networks for smart grid applications. *arXiv preprint arXiv:1403.0530*, 2014.

[112] Chanmin Yoon and Hojung Cha. Experimental analysis of IEEE 802.15. 4a CSS ranging and its implications. *Computer Communications*, 34(11):1361–1374, 2011.

[113] Nicolas Fourty, Adrien Van Den Bossche, and Thierry Val. An advanced study of energy consumption in an IEEE 802.15. 4 based network: Everything but the truth on 802.15. 4 node lifetime. *Computer Communications*, 35(14):1759–1767, 2012.

[114] Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Yasir Mehmood, Abdullah Gani, Salimah Mokhtar, and Sghaier Guizani. Enabling communication technologies for smart cities. *IEEE Communications Magazine*, 55(1):112–120, 2017.

[115] IEEE Standard for Information technology—Telecommunications and information exchange between systems Local and metropolitan area networks—Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pages 1–3534, Dec 2016.

[116] N Phuangpornpitak and S Tia. Feasibility study of wind farms under the thai very small scale renewable energy power producer (vspp) program. *Energy Procedia*, 9:159–170, 2011.

[117] Ayindrila Roy, Jitendranath Bera, and Gautam Sarkar. Wireless sensing of substation parameters for remote monitoring and analysis. *Ain Shams Engineering Journal*, 6(1):95–106, 2015.

[118] Anzar Mahmood, Nadeem Javaid, and Sohail Razzaq. A review of wireless communications for smart grid. *Renewable and sustainable energy reviews*, 41:248–260, 2015.

[119] Prashant Pathak. A Comparison between WLAN (IEEE 802.11 a, b, g, n and ac) Standards. 2017.

[120] Michael Emmanuel and Ramesh Rayudu. Communication technologies for smart grid applications: A survey. *Journal of Network and Computer Applications*, 74:133–148, 2016.

[121] Bradley Mitchell. Wireless standards 802.11 a, 802.11 b/g/n, and 802.11 ac. *Verkkojulkaisu. Saatavissa: http://compnetworking. about. com/cs/wireless80211/a/aa80211standard. htm [viitattu 8.4. 2015]*, 2015.

[122] Rachana Khanduri and SS Rattan. Performance Comparison Analysis between IEEE 802.11 a/b/g/n Standards. *International Journal of Computer Applications*, 78(1):13–20, 2013.

[123] IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput. *IEEE Std 802.11n-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, and IEEE Std 802.11w-2009)*, pages 1–565, Oct 2009.

[124] Tarek Khalifa, Atef Abdrabou, Khaled Shaban, and Ahmed M Gaouda. Heterogeneous wireless networks for smart grid distribution systems: Advantages and limitations. *Sensors*, 18(5):1517, 2018.

[125] Farshad Koohifar, Nico Saputro, Ismail Guvenc, and Kemal Akkaya. Hybrid Wi-Fi/LTE aggregation architecture for smart meter communications. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 575–580, 2015.

[126] Stefano Rinaldi, Paolo Ferrari, Alessandra Flammini, Francesco Gringoli, Matteo Loda, and Nahla Ali. An application of IEEE 802.11 ac to Smart Grid automation based on IEC 61850. In *IECON 42nd Annual Conference of the IEEE Industrial Electronics Society*, pages 4645–4650, 2016.

[127] CISCO. IEEE 802.11ac: The Fifth Generation of Wi-Fi (Technical White Paper). Technical report, 2018.

[128] CISCO. IEEE 802.11ax: The Sixth Generation of Wi-Fi (Technical White Paper). Technical report, 2018.

[129] H. Yang, D. Deng, and K. Chen. On Energy Saving in IEEE 802.11ax. *IEEE Access*, 6:47546–47556, 2018.

[130] IEEE Standard for information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 6: Medium Access Control (MAC) Security Enhancements. *IEEE Std 802.11i-2004*, pages 1–190, July 2004.

[131] IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 4: Protected Management Frames. *IEEE Std 802.11w-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, and IEEE Std 802.11y-2008)*, pages 1–111, Sep. 2009.

[132] IEEE Standard for Information technology–Local and metropolitan area networks–Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements. *IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003)*, pages 1–212, Nov 2005.

[133] Honggang Wang, Yi Qian, and Hamid Sharif. Multimedia communications over cognitive radio networks for smart grid applications. *IEEE Wireless Communications*, 20(4):125–132, 2013.

[134] Carl Eklund, Roger B Marks, Kenneth L Stanwood, Stanley Wang, et al. IEEE standard 802.16: a technical overview of the WirelessMAN™ air interface for broadband wireless access. *IEEE Communications Magazine*, 40(6):98–107, 2002.

[135] Perumalraja Rengaraju, Chung-Horng Lung, and Anand Srinivasan. Communication requirements and analysis of distribution networks using WiMAX technology for smart grids. In *8th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 666–670. IEEE, 2012.

[136] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems. *IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004)*, pages 1–2080, May 2009.

[137] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 1: Multihop Relay Specification. *IEEE Std 802.16j-2009 (Amendment to IEEE Std 802.16-2009)*, pages 1–290, June 2009.

[138] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 3: Advanced Air Interface. *IEEE Std 802.16m-2011(Amendment to IEEE Std 802.16-2009)*, pages 1–1112, May 2011.

[139] Tarek Khalifa, Kshirasagar Naik, and Amiya Nayak. A survey of communication protocols for automatic meter reading applications. *IEEE Communications Surveys & Tutorials*, 13(2):168–182, 2010.

[140] Robert C Qiu, Zhe Chen, Nan Guo, Yu Song, Peng Zhang, Husheng Li, and Lifeng Lai. Towards a real-time cognitive radio network testbed: architecture, hardware platform, and application to smart grid. In *5th IEEE Workshop on Networking Technologies for Software Defined Radio Networks (SDR)*, pages 1–6, 2010.

[141] Liguang Zhang. Notice of Retraction Challenges of mobile learning in 4G era. In *IEEE International Conference on E-Business and E-Government (ICEE)*, pages 1–3, 2011.

[142] Hao Liang, Bong Jun Choi, Atef Abdrabou, Weihua Zhuang, and Xuemin Sherman Shen. Decentralized economic dispatch in microgrids via heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 30(6):1061–1074, 2012.

[143] Stefano Rinaldi, Paolo Ferrari, Nahla M Ali, and Francesco Gringoli. Iec 61850 for micro grid automation over heterogeneous network: Requirements and real case deployment. In *IEEE 13th International Conference on Industrial Informatics (INDIN)*, pages 923–930, 2015.

[144] Ahmed Omara, Wendong Yuan, Michele Nogueira, Burak Kantarci, and Lei Wu. Microgrid data aggregation and wireless transfer scheduling in the presence of time sensitive events. In *16th ACM Intl. Symp. on Mobility Management and Wireless Access*, MobiWac'18, pages 109–112, 2018.

[145] Kaile Zhou, Chao Fu, and Shanlin Yang. Big data driven smart energy management: From big data to big insights. *Renewable and Sustainable Energy Reviews*, 56:215–225, 2016.

[146] Tadilo Endeshaw Bogale and Long Bao Le. Massive mimo and mmwave for 5g wireless hetnet: Potential benefits and challenges. *IEEE Vehicular Technology Magazine*, 11(1):64–75, 2016.

[147] Ning Zhang, Nan Cheng, Amila Tharaperiya Gamage, Kuan Zhang, Jon W Mark, and Xuemin Shen. Cloud assisted hetnets toward 5g wireless networks. *IEEE communications magazine*, 53(6):59–65, 2015.

[148] Jinwen Zhu. Communication network for smart grid interoperability. In *IEEE International Conference on Communication Software and Networks (ICCSN)*, pages 260–265, 2015.

[149] Takuro Sato, Daniel M Kammen, Bin Duan, Martin Macuha, Zhenyu Zhou, Jun Wu, Muhammad Tariq, and Solomon Abebe Asfaw. *Smart grid standards: specifications, requirements, and technologies*. John Wiley & Sons, 2015.

[150] Yurong Hu and Victor OK Li. Satellite-based internet: a tutorial. *IEEE Communications Magazine*, 39(3):154–162, 2001.

[151] Ziming Zhu, Sangarapillai Lambotharan, Woon Hau Chin, and Zhong Fan. Overview of demand management in smart grid and enabling wireless communication technologies. *IEEE Wireless Communications*, 19(3):48–56, 2012.

[152] P Donegan. Ethernet backhaul: Mobile operator strategies & market opportunities. *Heavy Reading*, 5(8):4–65, 2007.

[153] Zabih Ghassemlooy, Wasiu Popoola, and Sujan Rajbhandari. *Optical wireless communications: system and channel modelling with Matlab®*. CRC press, 2019.

[154] SENETAS. Technical-paper – microwave links and security vulnerability. Technical report, 2015.

[155] Yanchong Zheng, Songyan Niu, Yitong Shang, Ziyun Shao, and Linni Jian. Integrating plug-in electric vehicles into power grids: A comprehensive review on power interaction mode, scheduling methodology and mathematical foundation. *Renewable and Sustainable Energy Reviews*, 112:424–439, 2019.

[156] Yitong Shang, Yanchong Zheng, Ziyun Shao, and Linni Jian. Computational performance analysis for centralized coordinated charging methods of plug-in electric vehicles: From the grid operator perspective. *International Transactions on Electrical Energy Systems*, 30(2):e12229, 2020.

[157] Yitong Shang, Man Liu, Ziyun Shao, and Linni Jian. A centralized vehicle-to-grid scheme with distributed computing capacity engaging internet of smart charging points: case study. *International Journal of Energy Research*, 45(1):841–863, 2021.

[158] Yitong Shang, Hang Yu, Songyan Niu, Ziyun Shao, and Linni Jian. Cyber-physical co-modeling and optimal energy dispatching within internet of smart charging points for vehicle-to-grid operation. *Applied Energy*, 303:117595, 2021.

[159] Xiangyu Chen, Ka-Cheong Leung, Albert YS Lam, and David J Hill. Online scheduling for hierarchical vehicle-to-grid system: Design, formulation, and algorithm. *IEEE Transactions on Vehicular Technology*, 68(2):1302–1317, 2018.

[160] Linni Jian, Xinyu Zhu, Ziyun Shao, Shuangxia Niu, and CC Chan. A scenario of vehicle-to-grid implementation and its double-layer optimal charging strategy for minimizing load variance within regional smart grids. *Energy conversion and management*, 78:508–517, 2014.

[161] Ubaid ur Rehman, Muhammad Riaz, and Muhammad Yaqoob Wani. A robust optimization method for optimizing day-ahead operation of the electric vehicles aggregator. *International Journal of Electrical Power & Energy Systems*, 132:107179, 2021.

[162] Linni Jian, Yanchong Zheng, and Ziyun Shao. High efficient valley-filling strategy for centralized coordinated charging of large-scale electric vehicles. *Applied Energy*, 186:46–55, 2017.

[163] Zhile Yang, Kang Li, Qun Niu, and Yusheng Xue. A novel parallel-series hybrid meta-heuristic method for solving a hybrid unit commitment problem. *Knowledge-Based Systems*, 134:13–30, 2017.

[164] Yanchong Zheng, Yitong Shang, Ziyun Shao, and Linni Jian. A novel real-time scheduling strategy with near-linear complexity for integrating large-scale electric vehicles into smart grid. *Applied Energy*, 217:1–13, 2018.

[165] Cristian Dimas, Gustavo Ramos, Luis Caro, and Adriana C Luna. Parallel computing and multicore platform to assess electric vehicle hosting capacity. *IEEE Transactions on Industry Applications*, 56(5):4709–4717, 2020.

[166] Shutong Chen, Lei Jiao, Lin Wang, and Fangming Liu. An online market mechanism for edge emergency demand response via cloudlet control. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 2566–2574. IEEE, 2019.

[167] Zhaoyan Song, Ruiting Zhou, Shihan Zhao, Shixin Qin, John CS Lui, and Zongpeng Li. Edge emergency demand response control via scheduling in cloudlet cluster. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 394–399. IEEE, 2020.

[168] Shutong Chen, Lei Jiao, Fangming Liu, and Lin Wang. Edgedr: An online mechanism design for demand response in edge clouds. *IEEE Transactions on Parallel and Distributed Systems*, 33(2):343–358, 2021.

[169] Guangming Cui, Qiang He, Xiaoyu Xia, Feifei Chen, Tao Gu, Hai Jin, and Yun Yang. Demand response in noma-based mobile edge computing: A two-phase game-theoretical approach. *IEEE Transactions on Mobile Computing*, 2021.

[170] Ruiting Zhou, Zongpeng Li, Chuan Wu, and Zhiyi Huang. An efficient cloud market mechanism for computing jobs with soft deadlines. *IEEE/ACM Transactions on networking*, 25(2):793–805, 2016.

[171] Li Lin, Xiaofei Liao, Hai Jin, and Peng Li. Computation offloading toward edge computing. *Proceedings of the IEEE*, 107(8):1584–1607, 2019.

[172] Xiaoying Tang, Suzhi Bi, and Ying-Jun Angela Zhang. Distributed routing and charging scheduling optimization for internet of electric vehicles. *IEEE Internet of Things Journal*, 6(1):136–148, 2018.

[173] Kenechukwu Ginigeme and Zhanle Wang. Distributed optimal vehicle-to-grid approaches with consideration of battery degradation cost under real-time pricing. *IEEE Access*, 8:5225–5235, 2020.

[174] Yitong Shang, Man Liu, Ziyun Shao, and Linni Jian. Internet of smart charging points with photovoltaic integration: A high-efficiency scheme enabling optimal dispatching between electric vehicles and power grids. *Applied Energy*, 278:115640, 2020.

[175] Ruiting Zhou, Zongpeng Li, Chuan Wu, and Minghua Chen. Demand response in smart grids: A randomized auction approach. *IEEE Journal on Selected Areas in Communications*, 33(12):2540–2553, 2015.

[176] Yulan Yuan, Lei Jiao, Konglin Zhu, and Lin Zhang. Scheduling online ev charging demand response via v2v auctions and local generation. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):11436–11452, 2021.

[177] Weifeng Zhong, Kan Xie, Yi Liu, Chao Yang, and Shengli Xie. Efficient auction mechanisms for two-layer vehicle-to-grid energy trading in smart grid. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2017.

[178] Hanling Yi, Qiulin Lin, and Minghua Chen. Balancing cost and dissatisfaction in online ev charging under real-time pricing. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 1801–1809. IEEE, 2019.

[179] Linqi Guo, Karl F Erliksson, and Steven H Low. Optimal online adaptive electric vehicle charging. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2017.

[180] Serkan Gönen, H Hüseyin Sayan, Ercan Nurcan Yılmaz, Furkan Üstünsoy, and Gökçe Karacayılmaz. False data injection attacks and the insider threat in smart systems. *Computers & Security*, 97:101955, 2020.

[181] Matthew Robinson, Pascal A Schirmer, and Iosif Mporas. Privacy and security threats from smart meters technology. In *2021 International Carnahan Conference on Security Technology (ICCST)*, pages 1–6. IEEE, 2021.

[182] Shameek Bhattacharjee and Sajal K Das. Detection and forensics against stealthy data falsification in smart metering infrastructure. *IEEE Transactions on Dependable and Secure Computing*, 18(1):356–371, 2018.

[183] Shameek Bhattacharjee, Aditya Thakur, Simone Silvestri, and Sajal K Das. Statistical security incident forensics against data falsification in smart grid advanced metering infrastructure. In *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy*, pages 35–45, 2017.

[184] Paria Jokar, Nasim Arianpoo, and Victor CM Leung. Electricity theft detection in ami using customers' consumption patterns. *IEEE Transactions on Smart Grid*, 7(1):216–226, 2015.

[185] Mahmoud M Badr. *Security and privacy preservation for smart grid AMI using machine learning and cryptography*. PhD thesis, Tennessee Technological University, 2022.

[186] Abdulrahman Takiddin, Muhammad Ismail, and Erchin Serpedin. Robust data-driven detection of electricity theft adversarial evasion attacks in smart grids. *IEEE Transactions on Smart Grid*, 14(1):663–676, 2022.

[187] Zibin Zheng, Yatao Yang, Xiangdong Niu, Hong-Ning Dai, and Yuren Zhou. Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids. *IEEE Transactions on Industrial Informatics*, 14(4):1606–1615, 2017.

[188] Ejaz Ul Haq, Can Pei, Ruihong Zhang, Huang Jianjun, and Fiaz Ahmad. Electricity-theft detection for smart grid security using smart meter data: A deep-cnn based approach. *Energy Reports*, 9:634–643, 2023.

[189] Maria Zhdanova, Julian Urbansky, Anne Hagemeier, Daniel Zelle, Isabelle Herrmann, and Dorian Höffner. Local power grids at risk–an experimental and simulation-based analysis of attacks on vehicle-to-grid communication. In *Proceedings of the 38th Annual Computer Security Applications Conference*, pages 42–55, 2022.

[190] Richard Baker and Ivan Martinovic. Losing the car keys: Wireless {PHY-Layer} insecurity in {EV} charging. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 407–424, 2019.

[191] Sébastien Dudek, Jean-Christophe Delaunay, and Vincent Fargues. V2g injector: Whispering to cars and charging units through the power-line. In *Proceedings of the SSTIC (Symposium sur la sécurité des technologies de l'information et des communications), Rennes, France*, pages 5–7, 2019.

[192] Mauro Conti, Denis Donadel, Radha Poovendran, and Federico Turrin. Evexchange: A relay attack on electric vehicle charging system. In *European Symposium on Research in Computer Security*, pages 488–508. Springer, 2022.

[193] Andreas Fuchs, Dustin Kern, Christoph Krauß, and Maria Zhdanova. Trustev: trustworthy electric vehicle charging and billing. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 1706–1715, 2020.

[194] Sebastian Köhler, Simon Birnbach, Richard Baker, and Ivan Martinovic. On the security of the wireless electric vehicle charging communication. In *2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pages 393–398. IEEE, 2022.

[195] Kaibin Bao, Hristo Valev, Manuela Wagner, and Hartmut Schmeck. A threat analysis of the vehicle-to-grid charging protocol iso 15118. *Computer Science-Research and Development*, 33(1):3–12, 2018.

[196] Seokcheol Lee, Yongmin Park, Hyunwoo Lim, and Taeshik Shon. Study on analysis of security vulnerabilities and countermeasures in iso/iec 15118 based electric vehicle

charging technology. In *2014 International conference on IT convergence and security (ICITCS)*, pages 1–4. IEEE, 2014.

[197] Syed Rahman, Haneen Aburub, Yemeserach Mekonnen, and Arif I Sarwat. A study of ev bms cyber security based on neural network soc prediction. In *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, pages 1–5. IEEE, 2018.

[198] Mhd Ali Alomrani, Mosaddek Hossain Kamal Tushar, and Deepa Kundur. Detecting state of charge false reporting attacks via reinforcement learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[199] J. Zhang, J. Li, L. Wu, M. Erol-Kantarci, and B. Kantarci. Hierarchical optimal control of the resilient community microgrid in islanded mode. In *IEEE Power Energy Society General Meeting*, pages 1–5, 2019.

[200] Junghoon Lee and Gyung-Leen Park. A heuristic-based electricity trade coordination for microgrid-level v2g services. *International Journal of Vehicle Design*, 69(1-4):208–223, 2015.

[201] I. Stellios, P. Kotzanikolaou, M. Psarakis, C. Alcaraz, and J. Lopez. A survey of iot-enabled cyberattacks: Assessing attack paths to critical infrastructures and services. *IEEE Communications Surveys Tutorials*, 20(4):3453–3495, 2018.

[202] Kyounggon Kim, Jun Seok Kim, Seonghoon Jeong, Jo-Hee Park, and Huy Kang Kim. Cybersecurity for autonomous vehicles: Review of attacks and defense. *Computers & Security*, page 102150, 2021.

[203] Mark Bacchus, Alexander Coronado, and Maria A Gutierrez. The insights into car hacking, 2014.

[204] Zhiqiang Cai, Aohui Wang, Wenkai Zhang, M Gruffke, and H Schweppe. 0-days & mitigations: Roadways to exploit and secure connected bmw cars. *Black Hat USA*, 2019:39, 2019.

[205] Sen Nie, Ling Liu, and Yuefeng Du. Free-fall: Hacking tesla from wireless to can bus. *Briefing, Black Hat USA*, 25:1–16, 2017.

[206] Siddharth Shukla. Embedded security for vehicles: Ecu hacking, 2016.

[207] Eyal Ronen, Adi Shamir, Achi-Or Weingarten, and Colin O'Flynn. Iot goes nuclear: Creating a zigbee chain reaction. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 195–212. IEEE, 2017.

[208] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[209] Ahmed Omara, Wendong Yuan, Michele Nogueira, Burak Kantarci, and Lei Wu. Microgrid data aggregation and wireless transfer scheduling in the presence of time sensitive events. In *ACM Intl. Symp. on Mobility Management and Wireless Access*, pages 109–112, 2018.

[210] Sreedhar Madichetty and Sukumar Mishra. Cyber attack detection and correction mechanisms in a distributed dc microgrid. *IEEE Transactions on Power Electronics*, 37(2):1476–1485, 2021.

[211] Yin Haoyang, Liu Dong, and Weng Jiaming. Risk analysis of cyber physical distribution system considering cyber attacks on v2g system. 2021.

[212] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[213] Park Foreman. *Vulnerability management*. Auerbach Publications, 2019.

[214] Xingyu Zhou, Yi Li, Carlos A Barreto, Jiani Li, Peter Volgyesi, Himanshu Neema, and Xenofon Koutsoukos. Evaluating resilience of grid load predictions under stealthy adversarial attacks. In *2019 Resilience Week (RWS)*, volume 1, pages 206–212. IEEE, 2019.

[215] Tian Liu and Tao Shu. Adversarial false data injection attack against nonlinear ac state estimation with ann in smart grid. In *Security and Privacy in Communication Networks: 15th EAI International Conference, SecureComm 2019, Orlando, FL, USA, October 23–25, 2019, Proceedings, Part II 15*, pages 365–379. Springer, 2019.

[216] Tian Lan, Wenzong Wang, and Garng M Huang. False data injection attack in smart grid topology control: Vulnerability and countermeasure. In *2017 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2017.

[217] Yonghe Guo, Chee-Wooi Ten, Shiyan Hu, and Wayne W Weaver. Modeling distributed denial of service attack in advanced metering infrastructure. In *IEEE power & energy society innovative smart grid technologies conference (ISGT)*, pages 1–5, 2015.

[218] Zong-Han Yu and Wen-Long Chin. Blind false data injection attack using pca approximation method in smart grid. *IEEE Transactions on Smart Grid*, 6(3):1219–1226, 2015.

[219] Mohammad Esmalifalak, Huy Nguyen, Rong Zheng, and Zhu Han. Stealth false data injection using independent component analysis in smart grid. In *2011 IEEE international conference on smart grid communications (SmartGridComm)*, pages 244–248. IEEE, 2011.

[220] Yanan Sun, Lutz Lampe, and Vincent WS Wong. Smart meter privacy: Exploiting the potential of household energy storage units. *IEEE Internet of Things Journal*, 5(1):69–78, 2017.

[221] Günther Eibl and Dominik Engel. Differential privacy for real smart metering data. *Computer Science-Research and Development*, 32:173–182, 2017.

[222] Hu Yue, Kai Yan, Jianjun Zhao, Yougang Ren, Xihui Yan, and Hongshan Zhao. Estimating demand response flexibility of smart home appliances via nilm algorithm. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, volume 1, pages 394–398. IEEE, 2020.

[223] Jelena Ponoćko and Jovica V Milanović. Forecasting demand flexibility of aggregated residential load using smart meter data. *IEEE Transactions on Power Systems*, 33(5):5446–5455, 2018.

[224] Alexandre Lucas, Luca Jansen, Nikoleta Andreadou, Evangelos Kotsakis, and Marcelo Masera. Load flexibility forecast for dr using non-intrusive load monitoring in the residential sector. *Energies*, 12(14):2725, 2019.

[225] Antonio Ruano, Alvaro Hernandez, Jesus Ureña, Maria Ruano, and Juan Garcia. Nilm techniques for intelligent home energy management and ambient assisted living: A review. *Energies*, 12(11):2203, 2019.

[226] Phil Muncaster. Ninety percent of energy companies hit by cyber-attack in last year. https://www.infosecurity-magazine.com/news/ninety-percent-energy-companies/, December 2023.

[227] Justine Brown. Roundup: 2021 energy & utility data breaches and defenses in the news. https://securityintelligence.com/articles/energy-utility-data-breaches-2021/, October 2021.

[228] Josh Nadeau. Third-party breaches hit 90% of top global energy companies. https://securityintelligence.com/articles/third-party-breaches-top-global-energy-companies/, February 2024.

[229] Mohammad Ghiasi, Moslem Dehghani, Taher Niknam, Abdollah Kavousi-Fard, Pierluigi Siano, and Hassan Haes Alhelou. Cyber-attack detection and cyber-security enhancement in smart dc-microgrid based on blockchain technology and hilbert huang transform. *IEEE Access*, 9:29429–29440, 2021.

[230] Ender Ayanoglu, Kemal Davaslioglu, and Yalin E Sagduyu. Machine learning in nextg networks via generative adversarial networks. *IEEE Transactions on Cognitive Communications and Networking*, 2022.

[231] Patrick Huber, Melvin Ott, Martin Friedli, Andreas Rumsch, and Andrew Paice. Residential power traces for five houses: the ihomelab rapt dataset. *Data*, 5(1):17, 2020.

[232] Ahmed Omara and Burak Kantarci. Adversarial machine learning-based anticipation of threats against vehicle-to-microgrid services. In *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, pages 1844–1849, 2022.

[233] Yanchun Li, Qiuzhen Wang, Jie Zhang, Lingzhi Hu, and Wanli Ouyang. The theoretical research of generative adversarial networks: an overview. *Neurocomputing*, 435:26–41, 2021.

[234] Ahmed Omara and Burak Kantarci. An ai-driven solution to prevent adversarial attacks on mobile vehicle-to-microgrid services. *Simulation Modelling Practice and Theory*, 137:103016, 2024.

[235] Patrick Huber, Melvin Ott, Martin Friedli, Andreas Rumsch, and Andrew Paice. Residential Power Traces for Five Houses: the iHomeLab RAPT Dataset, December 2019.

[236] Ahmed Omara and Burak Kantarci. On the impact of data integrity attacks on vehicle-to-microgrid services. In *2021 IEEE 26th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1–7, 2021.

[237] Tuomo Sipola, Janne Alatalo, Tero Kokkonen, and Mika Rantonen. Artificial intelligence in the iot era: A review of edge ai hardware and software. In *2022 31st Conference of Open Innovations Association (FRUCT)*, pages 320–331. IEEE, 2022.

[238] Seungwoo Kum, Seungtaek Oh, Jeongcheol Yeom, and Jaewon Moon. Optimization of edge resources for deep learning application with batch and model management. *Sensors*, 22(17):6717, 2022.

[239] Zhuoqing Chang, Shubo Liu, Xingxing Xiong, Zhaohui Cai, and Guoqing Tu. A survey of recent advances in edge-computing-powered artificial intelligence of things. *IEEE Internet of Things Journal*, 8(18):13849–13875, 2021.

[240] Xiangyu Zhang, Jianhua Zou, Xiang Ming, Kaiming He, and Jian Sun. Efficient and accurate approximations of nonlinear convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and pattern Recognition*, pages 1984–1992, 2015.

[241] Chellammal Surianarayanan, John Jeyasekaran Lawrence, Pethuru Raj Chelliah, Edmond Prakash, and Chaminda Hewage. A survey on optimization techniques for edge artificial intelligence (ai). *Sensors*, 23(3):1279, 2023.

[242] Raghuraman Krishnamoorthi. Quantizing deep convolutional networks for efficient inference: A whitepaper. *arXiv preprint arXiv:1806.08342*, 2018.

[243] Colin White, Willie Neiswanger, and Yash Savani. Bananas: Bayesian optimization with neural architectures for neural architecture search. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 10293–10301, 2021.

[244] Michael Cogswell, Faruk Ahmed, Ross Girshick, Larry Zitnick, and Dhruv Batra. Reducing overfitting in deep networks by decorrelating representations. *arXiv preprint arXiv:1511.06068*, 2015.

[245] 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.

[246] Google. api - http body documentation. https://cloud.google.com/ai-platform/prediction/docs/reference/rest/v1/HttpBody. Accessed: 2024-05-05.

[247] Adrien Le Franc, Pierre Carpentier, Jean-Philippe Chancelier, and Michel De Lara. Emsx: a numerical benchmark for energy management systems. *Energy Systems*, 14(3):817–843, 2023.

[248] ME Ropp, K Aaker, J Haigh, and N Sabbah. Using power line carrier communications to prevent islanding [of PV power systems]. In *Conference Record of the Twenty-Eighth IEEE Photovoltaic Specialists Conference (Cat. No. 00CH37036)*, pages 1675–1678, 2000.

[249] Bikiran Guha, Rami J Haddad, and Youakim Kalaani. Anti-islanding techniques for inverter-based distributed generation systems-a survey. In *IEEE SoutheastCon*, pages 1–9, 2015.

[250] Suman Khichar and Mahendra Lalwani. An analytical survey of the islanding detection techniques of distributed generation systems. *Technology and Economics of Smart Grids and Sustainable Energy*, 3(1):10, 2018.

[251] M Ropp, D Larson, S Meendering, D McMahon, J Ginn, J Stevens, W Bower, S Gonzalez, K Fennell, and L Brusseau. Discussion of a power line carrier communications-based anti-islanding scheme using a commercial automatic meter reading system. In *IEEE 4th World Conference on Photovoltaic Energy Conference*, volume 2, pages 2351–2354, 2006.

[252] Wilsun Xu, Guibin Zhang, Chun Li, Wencong Wang, Guangzhu Wang, and Jacek Kliber. A power line signaling based technique for anti-islanding protection of distributed generators—part i: Scheme and analysis. *IEEE Transactions on Power Delivery*, 22(3):1758–1766, 2007.

[253] Irvin J Balaguer, Heung-Geun Kim, Fang Z Peng, and Eduardo I Ortiz. Survey of photovoltaic power systems islanding detection methods. In *34th Annual Conference of IEEE Industrial Electronics*, pages 2247–2252, 2008.

[254] SAM Javadian, R Tamizkar, and M-R Haghifam. A protection and reconfiguration scheme for distribution networks with DG. In *IEEE Bucharest PowerTech*, pages 1–8, 2009.

[255] Pukar Mahat, Zhe Chen, and Birgitte Bak-Jensen. Review of islanding detection methods for distributed generation. In *Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies*, pages 2743–2748. IEEE, 2008.

[256] Rohit S Kunte and Wenzhong Gao. Comparison and review of islanding detection techniques for distributed energy resources. In *40th North American Power Symposium*, pages 1–8. IEEE, 2008.

[257] Ahmad Yafaoui, Bin Wu, and Samir Kouro. Improved active frequency drift anti-islanding detection method for grid connected photovoltaic systems. *IEEE Transactions on Power Electronics*, 27(5):2367–2375, 2011.

[258] Aziah Khamis, Hussain Shareef, Erdal Bizkevelci, and Tamer Khatib. A review of islanding detection techniques for renewable distributed generation systems. *Renewable and Sustainable Energy Reviews*, 28:483–493, 2013.

[259] K Narayanan, Shahbaz Ahmed Siddiqui, and Manoj Fozdar. Hybrid islanding detection method and priority-based load shedding for distribution networks in the presence of DG units. *IET Generation, Transmission & Distribution*, 11(3):586–595, 2017.

[260] Farhan Noor, R Arumugam, and MY Vaziri. Unintentional islanding and comparison of prevention techniques. In *Proceedings of the 37th Annual North American Power Symposium*, pages 90–96. IEEE, 2005.

[261] CL Trujillo, D Velasco, E Figueres, and Gabriel Garcerá. Analysis of active islanding detection methods for grid-connected microinverters for renewable energy processing. *Applied Energy*, 87(11):3591–3605, 2010.

[262] IEEE Standards Board. *IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems: 1547-2003*. IEEE, 2003.

[263] SP Chowdhury, S Chowdhury, and PA Crossley. Islanding protection of active distribution networks with renewable distributed generators: A comprehensive survey. *Electric Power Systems Research*, 79(6):984–992, 2009.

[264] Math Bollen, Jin Zhong, Olof Samuelsson, and Johan Bjornstedt. Performance indicators for microgrids during grid-connected and island operation. In *IEEE Bucharest PowerTech*, pages 1–6, 2009.

[265] David Reigosa, Fernando Briz, Cristian Blanco Charro, Pablo García, and Juan Manuel Guerrero. Active islanding detection using high-frequency signal injection. *IEEE Transactions on Industry Applications*, 48(5):1588–1597, 2012.

[266] Hussein Al-Bataineh and Rajesh G Kavasseri. Islanding detection with data mining methods-a comparative study. In *Ninth Annual IEEE Green Technologies Conference (GreenTech)*, pages 104–109, 2017.

[267] Ward Bower and Michael Ropp. Evaluation of islanding detection methods for utility-interactive inverters in photovoltaic systems. *Sandia report SAND*, 3591:2002, 2002.

[268] Khalil El-Arroudi, Gza Joos, Innocent Kamwa, and Donald T McGillis. Intelligent-based approach to islanding detection in distributed generation. *IEEE transactions on power delivery*, 22(2):828–835, 2007.

[269] Mini S Thomas and Parveen Poon Terang. Islanding detection using decision tree approach. In *Joint International Conference on Power Electronics, Drives and Energy Systems & Power India*, pages 1–6. IEEE, 2010.

[270] NWA Lidula, N Perera, and AD Rajapakse. Investigation of a fast islanding detection methodology using transient signals. In *IEEE Power & Energy Society General Meeting*, pages 1–6, 2009.

[271] Mehrdad Heidari, Ghodratollah Seifossadat, and Morteza Razaz. Application of decision tree and discrete wavelet transform for an optimized intelligent-based islanding detection method in distributed systems with distributed generations. *Renewable and Sustainable Energy Reviews*, 27:525–532, 2013.

[272] Omar N Faqhruldin, Ehab F El-Saadany, and Hatem H Zeineldin. A universal islanding detection technique for distributed generation using pattern recognition. *IEEE Transactions on Smart Grid*, 5(4):1985–1992, 2014.

[273] SR Samantaray, Khalil El-Arroudi, Geza Joos, and Innocent Kamwa. A fuzzy rule-based approach for islanding detection in distributed generation. *IEEE Transactions on Power Delivery*, 25(3):1427–1433, 2010.

[274] Mohamed S ElNozahy, Ehab F El-Saadany, and Magdy MA Salama. A robust wavelet-ann based technique for islanding detection. In *IEEE Power and Energy Society General Meeting*, pages 1–8, 2011.

[275] Biljana Matic-Cuka and Mladen Kezunovic. Islanding detection for inverter-based distributed generation using support vector machine method. *IEEE Transactions on Smart Grid*, 5(6):2676–2686, 2014.

[276] Omar N Faqhruldin, EF El-Saadany, and HH Zeineldin. Naive bayesian islanding detection technique for distributed generation in modern distribution system. In *IEEE Electrical Power and Energy Conference*, pages 69–74, 2012.

[277] Manohar Mishra, Pravat Kumar Rout, Rituparna Sahu, Deepa Ray, and Swasti Swarup. Study the performance of s-transform based extreme learning machine for islanding detection in distributed generation. In *National Power Systems Conference (NPSC)*, pages 1–6. IEEE, 2016.

[278] Xiangrui Kong, Xiaoyuan Xu, Zheng Yan, Sijie Chen, Huoming Yang, and Dong Han. Deep learning hybrid method for islanding detection in distributed generation. *Applied Energy*, 210:776–785, 2018.

[279] Ahmad Darabi, Ali Moeini, and Mohsen Karimi. Distributed generation intelligent islanding detection using governor signal clustering. In *4th International Power Engineering and Optimization Conference (PEOCO)*, pages 345–351. IEEE, 2010.

[280] Victor Luiz Merlin, Ricardo Caneloi dos Santos, AP Grilo, JCM Vieira, Denis Vinicius Coury, and Mário Oleskovicz. A new artificial neural network based method for islanding detection of distributed generators. *International Journal of Electrical Power & Energy Systems*, 75:139–151, 2016.

[281] Mehrnoosh Vatani, Turaj Amraee, Ali Mohammad Ranjbar, and Babak Mozafari. Relay logic for islanding detection in active distribution systems. *IET Generation, Transmission & Distribution*, 9(12):1254–1263, 2015.

[282] Inamanamelluri Kumarswamy, Tara Kalyani Sandipamu, and Venkata Prasanth. Analysis of islanding detection in distributed generation using fuzzy logic technique. In *7th Asia Modelling Symposium*, pages 3–7. IEEE, 2013.

[283] C. R. Aguiar, R. F. Bastos, R. V. A. Neves, G. B. Reis, and R. Q. Machado. Fuzzy positive feedback for islanding mode detection in distributed generation. In *2013 IEEE Power Energy Society General Meeting*, pages 1–5, July 2013.

[284] Rubbens Boisguene, Sheng-Chia Tseng, Chih-Wei Huang, and Phone Lin. A survey on NB-IoT downlink scheduling: Issues and potential solutions. In *IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 547–551, 2017.

[285] Dizhi Zhou, Nicola Baldo, and Marco Miozzo. Implementation and validation of LTE downlink schedulers for ns-3. In *Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques*, pages 211–218. ICST (Institute for Computer Sciences, Social-Informatics and ..., 2013.

[286] Mattia Carpin, Andrea Zanella, Jawad Rasool, Kashif Mahmood, Ole Grøndalen, and Olav N Østerbø. A performance comparison of LTE downlink scheduling algorithms in time and frequency domains. In *IEEE International Conference on Communications (ICC)*, pages 3173–3179, 2015.

[287] Jawad Rasool, Vegard Hassel, Sébastien de la Kethulle de Ryhove, and Geir E Øien. Opportunistic scheduling policies for improved throughput guarantees in wireless networks. *EURASIP Journal on Wireless Communications and Networking*, 2011(1):43, 2011.

[288] Sem Borst and Phil Whiting. Dynamic channel-sensitive scheduling algorithms for wireless data throughput optimization. *IEEE Transactions on Vehicular Technology*, 52(3):569–586, 2003.

[289] Dizhi Zhou, Wei Song, Nicola Baldo, and Marco Miozzo. Evaluation of TCP performance with LTE downlink schedulers in a vehicular environment. In *9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1064–1069, 2013.

[290] Timotheos Kastrinogiannis and Symeon Papavassiliou. Probabilistic short-term delay and throughput requirements of multimedia services in high throughput wireless networks. In *IEEE Sarnoff Symposium*, pages 1–5, 2007.

[291] Narges Shojaedin, Majid Ghaderi, and Ashwin Sridharan. TCP-aware scheduling in LTE networks. In *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, pages 1–9, 2014.

[292] Xin Liu, Edwin KP Chong, and Ness B Shroff. A framework for opportunistic scheduling in wireless networks. *Computer networks*, 41(4):451–474, 2003.

[293] Arash Asadi and Vincenzo Mancuso. A survey on opportunistic scheduling in wireless communications. *IEEE Communications Surveys & Tutorials*, 15(4):1671–1688, 2013.

[294] Hongseok Kim and Gustavo De Veciana. Losing opportunism: Evaluating service integration in an opportunistic wireless system. In *IEEE International Conference on Computer Communications*, pages 982–990, 2007.

[295] Bilal Sadiq, Seung Jun Baek, and Gustavo De Veciana. Delay-optimal opportunistic scheduling and approximations: The log rule. *IEEE/ACM Transactions on Networking (TON)*, 19(2):405–418, 2011.

[296] Michael J Neely. Opportunistic scheduling with worst case delay guarantees in single and multi-hop networks. In *Proceedings IEEE INFOCOM*, pages 1728–1736, 2011.

[297] Kae Won Choi, Wha Sook Jeon, and Dong Geun Jeong. Resource allocation in ofdma wireless communications systems supporting multimedia services. *IEEE/ACM Transactions on Networking (TON)*, 17(3):926–935, 2009.

[298] Ramtin Kazemi Beidokhti, Mohammad Hossein Yaghmaee Moghaddam, and Jalil Chitizadeh. Adaptive qos scheduling in wireless cellular networks. *Wireless Networks*, 17(3):701–716, 2011.

[299] Peter Jacko. Value of information in optimal flow-level scheduling of users with markovian time-varying channels. *Performance Evaluation*, 68(11):1022–1036, 2011.

[300] The Mathworks, Inc., Natick, Massachusetts. *MATLAB 9.6.0.1072779 (R2019a)*, 2019.

[301] R.J. Campbell. Weather-related power outages and electric system resiliency. pages 103–118, 01 2013.

[302] Mladen Kezunovic, Ian Dobson, and Yimai Dong. Impact of extreme weather on power system blackouts and forced outages: New challenges. In *7th Balkan Power Conference*, pages 1–5. Citeseer, 2008.

[303] Mathaios Panteli and Pierluigi Mancarella. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. *Electric Power Systems Research*, 127:259 – 270, 2015.

[304] T. Liu, X. Tan, B. Sun, Y. Wu, X. Guan, and D. H. K. Tsang. Energy management of cooperative microgrids with p2p energy sharing in distribution networks. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 410–415, Nov 2015.

[305] Wayes Tushar, Tapan Kumar Saha, Chau Yuen, Thomas Morstyn, Malcolm D. McCulloch, H. Vincent Poor, and Kristin L. Wood. A motivational game-theoretic approach for peer-to-peer energy trading in the smart grid. *Applied Energy*, 243:10 – 20, 2019.

[306] L. Wu, J. Li, M. Erol-Kantarci, and B. Kantarci. An integrated reconfigurable control and self-organizing communication framework for community resilience microgrids. *The Electricity Journal*, 30(4):27 – 34, 2017. Special Issue: Contemporary Strategies for Microgrid Operation & Control.

[307] Mohammad Hossein Sarparandeh and Mehdi Ehsan. Pricing of vehicle-to-grid services in a microgrid by nash bargaining theory. *Mathematical Problems in Engineering*, 2017, 2017.

[308] M. Erol-Kantarci, B. Kantarci, and H. T. Mouftah. Reliable overlay topology design for the smart microgrid network. *IEEE Network*, 25(5):38–43, September 2011.

[309] Junghoon Lee and Gyung-Leen Park. A heuristic-based electricity trade coordination for microgrid-level V2G services. 2015.

[310] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification. Technical Specification (TS) 136.321, 3rd Generation Partnership Project (3GPP), 04 2015. Version 12.5.0.

[311] S. Shao, S. Guo, X. Qiu, L. Meng, and M. Lei. Traffic scheduling mechanism based on interference avoidance for meter data collection in wireless smart grid communication networks. *China Communications*, 12(7):142–153, July 2015.

[312] Sean Barker, Aditya Mishra, David Irwin, Emmanuel Cecchet, Prashant Shenoy, and Jeannie Albrecht. Smart*: An open data set and tools for enabling research in sustainable homes, 2012.

[313] R. H. Lasseter. Microgrids. *Proc. IEEE Power Eng. Soc. Winter Mtg.,2002*, pages 305–08.

[314] J. Huang, H. Wang, Y. Qian, and C. Wang. Priority-based traffic scheduling and utility optimization for cognitive radio communication infrastructure-based smart grid. *IEEE Transactions on Smart Grid*, 4(1):78–86, March 2013.

[315] Omnet++ discrete event simulator. https://omnetpp.org/. OMNeT 5.3 Released.

[316] Z. Feng, Q. Li, W. Li, T. A. Gulliver, and P. Zhang. Priority-based dynamic spectrum management in a smart grid network environment. *IEEE Journal on Selected Areas in Communications*, 33(5):933–945, May 2015.

[317] M. Rana. Architecture of the internet of energy network: An application to smart grid communication. *IEEE Access*, 5:4704–4710, April 2017.

[318] C. Hägerling, F. M. Kurtz, R. L. Olsen, and C. Wietfeld. Communication architecture for monitoring and control of power distribution grids over heterogeneous ict networks. In *2014 IEEE International Energy Conference (ENERGYCON)*, pages 838–845, May 2014.

[319] K. Wongwut and S. Nuchprayoon. Optimum hourly operation of a prosumer with battery energy storage system under time-of-use pricing. In *IEEE PES Asia-Pacific Power&Energy Eng. Conf.*, pages 1–6, Nov 2017.

[320] Melike Erol-Kantarci, Burak Kantarci, and Hussein T Mouftah. Reliable overlay topology design for the smart microgrid network. *IEEE Network*, 25(5), 2011.

[321] V. Theodorou, K. V. Katsaros, A. Roos, E. Sakic, and V. Kulkarni. Cross-domain network slicing for industrial applications. In *European Conf. on Networks and Communications*, pages 209–213, June 2018.

[322] Y. Yan, Y. Qian, H. Sharif, and D. Tipper. A survey on smart grid communication infrastructures: Motivations, requirements and challenges. *IEEE Comm. Surv. & Tutorials*, 15(1):5–20, First 2013.

[323] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke. Smart grid technologies: Communication technologies and standards. *IEEE Transactions on Industrial Informatics*, 7(4):529–539, Nov 2011.

[324] M. Samiullah, S. M. Abdullah, A. F. M. Imamul Hoq Bappi, and S. Anwar. Queue management based congestion control in wireless body sensor network. In *International Conference on Informatics, Electronics Vision (ICIEV)*, pages 493–496, May 2012.

[325] Junru Lin, Baohui Zhu, Peng Zeng, Wei Liang, Haibin Yu, and Yang Xiao. Monitoring power transmission lines using a wireless sensor network. *Wireless Communications and Mobile Computing*, 15(14):1799–1821.

[326] H. H. M. Tam, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor. Joint load balancing and interference management for small-cell heterogeneous networks with limited backhaul capacity. *IEEE Transactions on Wireless Communications*, 16(2):872–884, Feb 2017.

[327] Aron P Dobos. PVWatts version 5 manual. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014.

[328] Si Wu and Shun-Ichi Amari. Conformal transformation of kernel functions: A data-dependent way to improve support vector machine classifiers. *Neural Processing Letters*, 15(1):59–67, 2002.

[329] Simscape power systems, 2017. The MathWorks, Natick, MA, USA.

[330] MATLAB simscape power systems toolbox . https://www.mathworks.com/products/simpower.html. The MathWorks, Natick, MA, USA.

[331] Yang Zhang, Tao Huang, and Ettore Francesco Bompard. Big data analytics in smart grids: a review. *Energy Informatics*, 1(1):8, 2018.

[332] Yunxin (Jeff) Li. An overview of the dsrc/wave technology. In Xi Zhang and Daji Qiao, editors, *Quality, Reliability, Security and Robustness in Heterogeneous Networks*, pages 544–558, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[333] Daniel Jiang and Luca Delgrossi. IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. In *IEEE Vehicular Technology Conference (VTC)*, pages 2036–2040, 2008.

[334] Abdeldime MS Abdelgader and Wu Lenan. The physical layer of the IEEE 802.11 p WAVE communication standard: the specifications and challenges. In *Proceedings of the World Congress on Engineering and Computer Science*, volume 2, pages 22–24, 2014.

[335] IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments. *IEEE Std 802.11p-2010 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, IEEE Std 802.11n-2009, and IEEE Std 802.11w-2009)*, pages 1–51, July 2010.

[336] Xianan Huang, Ding Zhao, and Huei Peng. Empirical study of DSRC performance based on safety pilot model deployment data. *IEEE Transactions on Intelligent Transportation Systems*, 18(10):2619–2628, 2017.

[337] Maxime Guériau, Romain Billot, Nour-Eddin El Faouzi, Julien Monteil, Frédéric Armetta, and Salima Hassas. How to assess the benefits of connected vehicles? a simulation framework for the design of cooperative traffic management strategies. *Transportation research part C: emerging technologies*, 67:266–279, 2016.

[338] John Harding, Gregory Powell, Rebecca Yoon, Joshua Fikentscher, Charlene Doyle, Dana Sade, Mike Lukuc, Jim Simons, Jing Wang, et al. Vehicle-to-vehicle communications: readiness of V2V technology for application. Technical report, United States. National Highway Traffic Safety Administration, 2014.

[339] Hansong Xu, Jie Lin, and Wei Yu. Smart transportation systems: Architecture, enabling technologies, and open issues. In *Secure and Trustworthy Transportation Cyber-Physical Systems*, pages 23–49. Springer, 2017.

[340] Khadige Abboud, Hassan Aboubakr Omar, and Weihua Zhuang. Interworking of DSRC and cellular network technologies for V2X communications: A survey. *IEEE Transactions on Vehicular Technology*, 65(12):9457–9470, 2016.

[341] Min Wang, Martin Winbjork, Zhang Zhang, Ricardo Blasco, Hieu Do, Stefano Sorrentino, Marco Belleschi, and Yunpeng Zang. Comparison of LTE and DSRC-based connectivity for intelligent transportation systems. In *IEEE 85th Vehicular Technology Conference (VTC)*, pages 1–5, 2017.

[342] R Blasco, H Do, S Shalmashi, S Sorrentino, and Y Zang. 3GPP LTE enhancements for V2V and comparison to IEEE 802.11 p. In *Proc. ITS Eur. Congr.*, pages 1–10, 2016.

[343] Lei Wu, Jie Li, Melike Erol-Kantarci, and Burak Kantarci. An integrated reconfigurable control and self-organizing communication framework for community resilience microgrids. *The Electricity Journal*, 30(4):27–34, 2017.

[344] IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems - Amendment 1. *IEEE Std 1547a-2014 (Amendment to IEEE Std 1547-2003)*, pages 1–16, May 2014.

[345] Lei Wu, Tom Ortmeyer, and Jie Li. The community microgrid distribution system of the future. *The Electricity Journal*, 29(10):16–21, 2016.

[346] Thomas Ortmeyer, Lei Wu, and Jie Li. Planning and design goals for resilient microgrids. In *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5, 2016.

[347] A Fitah, A Badri, M Moughit, and A Sahel. Performance of DSRC and Wi-Fi for Intelligent Transport Systems in VANET. *Procedia Computer Science*, 127:360–368, 2018.

[348] Kenneth Sorle Nwizege, Mauro Bottero, Shedrack Mmeah, and Emmanuel D Nwiwure. Vehicles-to-Infrastructure Communication Safety Messaging in DSRC. *Procedia computer science*, 34:559–564, 2014.

[349] Caroline Carl. *Calculating solar photovoltaic potential on residential rooftops in Kailua Kona, Hawaii*. University of Southern California, 2014.

[350] M. Erol-Kantarci and H. T. Mouftah. Energy-efficient information and communication infrastructures in the smart grid: A survey on interactions and open issues. *IEEE Communications Surveys Tutorials*, 17(1):179–197, Firstquarter 2015.

[351] Melike Erol-Kantarci and Hussein T Mouftah. Energy-efficient information and communication infrastructures in the smart grid: A survey on interactions and open issues. *IEEE Communications Surveys & Tutorials*, 17(1):179–197, 2014.

[352] Kalpesh C Soni and FF Belim. Microgrid during grid-connected mode and islanded mode-a review. *International Journal of Advance Engineering and Research Development*, 2015.

[353] Adam Hirsch, Yael Parag, and Josep Guerrero. Microgrids: A review of technologies, key drivers, and outstanding issues. *Renewable and sustainable Energy reviews*, 90:402–411, 2018.

[354] Zeeshan Hameed Mir and Fethi Filali. LTE and IEEE 802.11 p for vehicular networking: a performance evaluation. *EURASIP Journal on Wireless Communications and Networking*, (1):89, 2014.

[355] Phebe Asantewaa Owusu and Samuel Asumadu-Sarkodie. A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering*, 3(1):1167990, 2016.

[356] Ashoke Kumar Basu, SP Chowdhury, S Chowdhury, and S Paul. Microgrids: Energy management by strategic deployment of ders—a comprehensive survey. *Renewable and Sustainable Energy Reviews*, 15(9):4348–4356, 2011.

[357] Tine L Vandoorn, Bart Meersman, Jeroen DM De Kooning, and Lieven Vandevelde. Transition from islanded to grid-connected mode of microgrids with voltage-based droop control. *IEEE Transactions on Power Systems*, 28(3):2545–2553, 2013.

[358] Shelby Brown. How to save money with uber and lyft. https://www.cnet.com/how-to/how-to-save-money-with-uber-and-lyft/. Accessed: 2019-06-23.

[359] Isidor Buchmann. BU-1003: Electric Vehicle (EV). https://batteryuniversity.com/learn/article/electric_vehicle_ev. Accessed: 2019-11-27.

[360] Ontario Hydro. Time-of-Use (TOU) Pricing. http://www.ontario-hydro.com/current-rates.

[361] Till Gnann, Simon Funke, Niklas Jakobsson, Patrick Plötz, Frances Sprei, and Anders Bennehag. Fast charging infrastructure for electric vehicles: Today's situation and future needs. *Transportation Research Part D: Transport and Environment*, 62:314–329, 2018.

[362] Willett Kempton, Yannick Perez, and Marc Petit. Public policy for electric vehicles and for vehicle to gridpower. *Revue d'économie industrielle*, (148):263–290, 2014.

[363] Benedetto Aluisio, Sergio Bruno, Luca De Bellis, Maria Dicorato, Giuseppe Forte, and Michele Trovato. Dc-microgrid operation planning for an electric vehicle supply infrastructure. *Applied Sciences*, 9(13):2687, Jul 2019.

[364] Mina Farmanbar, Kiyan Parham, Øystein Arild, and Chunming Rong. A widespread review of smart grids towards smart cities. *Energies*, 12(23):4484, Nov 2019.

[365] Nipendra Kayastha, Dusit Niyato, Ekram Hossain, and Zhu Han. Smart grid sensor data collection, communication, and networking: a tutorial. *Wireless Communications and Mobile Computing*, 14(11):1055–1087, 2014.

[366] Gowdemy Rajalingham, Quang-Dung Ho, and Tho Le-Ngoc. Evaluation of an efficient smart grid communication system at the neighbor area level. In *2014 IEEE 11th Consumer Communications and Networking Conference*, pages 426–431, Jan. 2014.

[367] Deepak Ronanki, Apoorva Kelkar, and Sheldon S. Williamson. Extreme fast charging technology—prospects to enhance sustainable electric transportation. *Energies*, 12(19):3721, Sep 2019.

[368] Mahmoud Saleh, Yusef Esa, Mohamed El Hariri, and Ahmed Mohamed. Impact of information and communication technology limitations on microgrid operation. *Energies*, 12(15):2926, Jul 2019.

[369] Roel Verdult, Flavio D Garcia, and Josep Balasch. Gone in 360 seconds: Hijacking with hitag2. In *21st {USENIX} Security Symposium ({USENIX} Security 12)*, pages 237–252, 2012.

[370] Flavio D Garcia, David Oswald, Timo Kasper, and Pierre Pavlidès. Lock it and still lose it—on the (in) security of automotive remote keyless entry systems. In *25th {USENIX} Security Symposium ({USENIX} Security 16)*, 2016.

[371] Solon Falas, Charalambos Konstantinou, and Maria K Michael. A modular end-to-end framework for secure firmware updates on embedded systems. *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, 18(1):1–19, 2021.

[372] 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.

[373] Martin Higgins, Fei Teng, and Thomas Parisini. Stealthy mtd against unsupervised learning-based blind fdi attacks in power systems. *IEEE Transactions on Information Forensics and Security*, 16:1275–1287, 2020.

[374] Xiao Zhaoxia, Li Hui, Zhu Tianli, and Li Huaimin. Day-ahead optimal scheduling strategy of microgrid with evs charging station. In *2019 IEEE 10th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, pages 774–780. IEEE, 2019.

[375] Roghieh A Biroon, Pierluigi Pisu, and Zoleikha Abdollahi. Real-time false data injection attack detection in connected vehicle systems with pde modeling. In *American Control Conference (ACC)*, pages 3267–3272, 2020.

[376] Mingshun Sun, Ming Li, and Ryan Gerdes. A data trust framework for vanets enabling false data detection and secure vehicle tracking. In *IEEE Conf. on Communications and Network Security (CNS)*, pages 1–9, 2017.

[377] Jinxin Liu, Michele Nogueira, Johan Fernandes, and Burak Kantarci. Adversarial machine learning: A multilayer review of the state-of-the-art and challenges for wireless and mobile systems. *IEEE Communications Surveys Tutorials*, 24(1):123–159, 2022.

[378] Zeyu Wang, Jianhui Wang, and Chen Chen. A three-phase microgrid restoration model considering unbalanced operation of distributed generation. *IEEE Transactions on Smart Grid*, 9(4):3594–3604, 2018.

[379] Yin Haoyang, Liu Dong, and Weng Jiaming. Risk analysis of cyber physical distribution system considering cyber attacks on v2g system. In *The 10th Renewable Power Generation Conference (RPG 2021)*, volume 2021, pages 841–846, 2021.

[380] Sohrab Nizami, Wayes Tushar, M.J. Hossain, Chau Yuen, Tapan Saha, and H. Vincent Poor. Transactive energy for low voltage residential networks: A review. *Applied Energy*, 323:119556, 2022.

[381] Zhaoyan Song, Ruiting Zhou, Shihan Zhao, Shixin Qin, John CS Lui, and Zongpeng Li. Emergency demand response in edge computing. *EURASIP Journal on Wireless Communications and Networking*, 2020:1–24, 2020.

[382] Jiaxiang Wu, Cong Leng, Yuhang Wang, Qinghao Hu, and Jian Cheng. Quantized convolutional neural networks for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4820–4828, 2016.