Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids

A SEMINAR REPORT

Submitted by

SHADMA SUBAIR C P

(MES20CS102)

to

The APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of the Degree

of

Bachelor of Technology

in

Computer Science and Engineering

Department of Computer Science and Engineering
[B.Tech. Programme accredited by NBA and NAAC]
MES College of Engineering Kuttippuram
Thrikkanapuram P.O., Malappuram Dt., Kerala, India 679582

2023-2024

DECLARATION

I, hereby declare that the seminar report ”Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids”, submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done under the supervision of Mr. Harikrishnan G R, Assistant Professor, Computer Science and Engineering. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University

Place: Kuttippuram
Date: 01-12-2023

Shadma Subair C P

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MES COLLEGE OF ENGINEERING, KUTTIPPURAM

CERTIFICATE

This is to certify that the report entitled **"Electricity Theft Detection Using Deep Reinforcement Learning in Smart Power Grids"** submitted by **SHADMA SUBAIR C P**, to the APJ Abdul Kalam University in partial requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering is a bonafide record of the seminar work carried out under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Mr. Harikrishnan G R
Assistant Professor
Dept. of Computer Science and Engineering .
MES College of Engineering

Dr. Anil K Jacob
Professor & Head
Dept.of Computer Science and Engg.
MES College of Engineering

ACKNOWLEDGEMENT

First of all I wish to thank God Almighty for blessing that made this work a success. I am grateful to ***Dr. Rahumathunza I****, Principal, MES College of Engineering, Kuttippuram,* for providing the right ambiance to do this seminar. I would like to extend my sincere gratitude to ***Dr. Anil K Jacob,*** *Head of the Department, CSE, MES College of Engineering, Kuttippuram.*

I am deeply indebted to the seminar coordinator ***Ms. Aswathy Babu C A****, Assistant Professor, Department of Computer Science and Engineering* for her continued support.

It is with great pleasure that I express deep sense of gratitude to my seminar guide ***Mr. Harikrishnan G R,*** *Assistant Professor, Department of Computer Science Engineering,* for his guidance, supervision, encouragement and valuable advice in each and every phases.

I would like to thank all other faculty members and fellow students of MES College of Engineering, Kuttippuram for their warm friendship, support and help.

SHADMA SUBAIR C P

ABSTRACT

In the realm of smart power grids, the deployment of Smart Meters (SMs) at the customer end facilitates the periodic reporting of fine-grained power consumption readings to utility providers for effective energy management and load monitoring. Despite their benefits, these systems are susceptible to cyber-attacks, specifically electricity theft, wherein fraudulent customers compromise their SMs to submit false readings and reduce their electricity costs. While supervised Machine Learning methods have been explored for attack detection, this article introduces a novel approach employing Deep Reinforcement Learning (DRL) to address the electricity theft problem. DRL, utilizing exploration and exploitation mechanisms for adaptive learning, presents a promising solution due to its capacity to efficiently handle dynamic cyber-attack scenarios and changing consumption patterns. The proposed DRL approach is presented across four scenarios, incorporating a global detection model using both Deep Q Network (DQN) and Double Deep Q Network (DDQN) architectures, customization for new customers to enhance detection accuracy and prevent zero-day attacks, consideration of changing consumption patterns for existing customers, and addressing challenges in defending against newly launched cyber-attacks. The experimental results demonstrate the efficacy of the DRL approach in boosting detection accuracy for electricity theft cyber-attacks, as well as its ability to adeptly learn new consumption patterns, changes in existing patterns, and counter newly launched attacks.

CONTENTS

LIST OF FIGURES

ABBREVIATIONS

AI	Artificial Intelligence
ML	Machine Learning
CNN	Convolutional Neural Network
FFNN	Feed Forward Neural Netwok
GRU	Gated Recurrent Unit
RNN	Recurrent Neural Network
AMI	Advanced Metering Infrastructure
SM	Smart Meters
RL	Reinforcement Learning
DL	Deep Learning
SO	System Operator
DQN	Deep Q Network
DDQN	Double Deep Q Network

CHAPTER 1

INTRODUCTION

The Smart Grid (SG) is a modernized power grid designed to enhance the efficiency and reliability of electricity delivery. It includes components like electricity production stations, Advanced Metering Infrastructure (AMI) network, System Operator (SO), and transmission and distribution systems. The AMI facilitates bidirectional communication between Smart Meters (SMs) at customers' homes . Unlike traditional monthly billing, SG records fine-grained electricity consumption readings every few minutes, aiding in demand management, load monitoring, and dynamic pricing. However, cyber-attacks, particularly electricity theft, pose a serious threat to SG, leading to financial losses and grid performance degradation.

Traditional power grids face physical tampering for electricity theft, while SGs are vulnerable to cyber-attacks where fraudulent customers hack their SMs to manipulate consumption readings. The financial impact of electricity theft globally is estimated at 89.3 billion dollars, affecting both developed and developing countries. To address this issue, Artificial Intelligence (AI) and Machine Learning (ML) have been applied, but existing approaches, especially Deep Learning (DL), have limitations in adapting to changing consumption patterns and new cyber-attacks.

Reinforcement Learning (RL), a type of ML, emerges as a promising solution to overcome DL limitations. RL, in to human learning, adapts to the environment through exploration and exploitation, making optimal decisions autonomously. The paper explores the use of RL to detect electricity theft cyber-attacks across four scenarios. It proposes global detection models using Deep Q Network (DQN) and

Double Deep Q Network (DDQN) with various DL architectures. The RL detector's ability to adapt to new consumption patterns, detect changes in consumption, and address newly launched cyber-attacks is investigated. Experimental results show that the proposed RL approach improves detection accuracy and adapts efficiently to evolving scenarios, outperforming existing DL techniques.

In conclusion, this method introduces RL as a novel approach for electricity theft detection in Smart Grids, addressing the limitations of traditional DL methods. The proposed RL models demonstrate superior adaptability and performance, offering a promising solution to enhance the security of Smart Grids against cyber-attacks.

CHAPTER 2

LITERATURE SURVEY

In the realm of Advanced Metering Infrastructure (AMI), where Smart Meters (SMs) transmit fine-grained power consumption data for load monitoring and billing, the escalating concern of electricity theft necessitates innovative approaches. This method underscores the need for efficient and privacy-preserving solutions, considering the trade-offs inherent in existing models. The proposed scheme aims to mitigate communication and computation overheads significantly compared to previous approaches, offering a promising avenue for enhancing privacy-preserving electricity theft detection.

This method addresses the persistent challenge of electricity theft in utility companies despite the implementation of Advanced Metering Infrastructure (AMI). Existing machine learning-based detectors often fall short in capturing the complex patterns and temporal correlations within the time-series profiles of energy consumption data. In response, this method introduces a novel solution, a deep recurrent vector embedding model, to effectively identify electricity theft cyber-attacks. Leveraging vector embedding to represent energy consumption profiles and incorporating gated recurrent units for capturing time-series nuances, the proposed model demonstrates significant improvements. Through a sequential grid-search hyperparameter optimization algorithm, detection performance is further enhanced. Testing against real datasets yields promising results, with a 95.8 percentage detection rate, 2.1 percentage false alarm rate, and 93.7 percentage highest difference. The method also emphasizes the broader context of electricity losses, emphasizing the economic impact, particularly in developing countries, and the vulnerabilities introduced by smart

meters in the Advanced Metering Infrastructure (AMI). The proposed model is positioned as a comprehensive and effective solution to address the complexities of electricity theft cyber-attacks and improve the overall stability of smart grids .

The evolving landscape of Advanced Metering Infrastructure (AMI) in smart grids introduces challenges, particularly concerning the privacy of consumers' power consumption readings. In this context, the Change And Transmit (CAT) approach efficiently collects readings, transmitting data only when a significant change in consumption occurs. However, this approach raises privacy concerns, particularly the risk of a Presence-Privacy Attack (PPA) where an attacker infers sensitive information about consumers' activities by analyzing transmission patterns. A prior countermeasure involved sending spoofing transmissions to obfuscate patterns, but it suffered from a high attacker success rate. The proposed defense model, trained on transmission patterns during consumer presence, significantly reduces the attacker's success rate and the number of transmitted readings compared to existing schemes. This work stands out as the first to explore PPA using a deep-learning approach, offering a more sophisticated attack model and a general defense model applicable to diverse consumers. The contributions address gaps in existing literature, providing a robust solution for preserving consumer privacy in AMI networks.

The literature on Advanced Metering Infrastructure (AMI) and smart grid technologies highlights the significance of addressing privacy concerns, particularly in the context of Presence-Privacy Attacks (PPA) arising from the Change and Transmit (CAT) approach for fine-grained power consumption readings. Previous studies focused on encryption as a countermeasure, but its limitations prompted exploration of alternative strategies. A notable contribution proposes a defense mechanism involving spoofing transmissions; however, its drawbacks, including high attacker suc-

cess rates and consumer-specific dependencies, indicate room for improvement. In response, this method introduces the "STID" scheme, leveraging an interactive deep-learning defense model to efficiently collect power consumption readings while significantly mitigating PPA, thus advancing the literature with a sophisticated approach that enhances both security and efficiency.

The literature on detecting false-reading attacks in Advanced Metering Infrastructure (AMI) networks primarily focuses on consumption metering and Feed-In Tariff (FIT) systems, with limited attention given to the unique challenges presented by the net-metering system. In the broader context of AMI security, existing works have proposed various solutions for detecting false readings, considering scenarios where malicious customers attempt to manipulate consumption or generation data. These approaches often rely on data analysis, anomaly detection, and machine learning techniques. However, the specific characteristics and complexities of the net-metering system have been overlooked until this study. In consumption metering, attackers aim to reduce readings for financial gains, while in FIT systems, they seek to increase generation readings. The net-metering system introduces a more intricate challenge, as false readings must account for both consumption and generation patterns simultaneously. This method contributes significantly by filling this gap and introducing a novel multi-data-source deep hybrid learning-based detector designed explicitly for the net-metering system. The proposed detector leverages correlations between net meter readings and external data sources, such as solar irradiance and temperature, demonstrating promising results in terms of accuracy and performance. Overall, this method underscores the importance of addressing the distinct challenges posed by the net-metering system, offering a valuable contribution to the evolving field of AMI security..

CHAPTER 3

PROPOSED METHODOLOGY

This paper is focused on the designed to provide a comprehensive understanding of the process undertaken for developing, training, and evaluating the Reinforcement Learning (RL) models for electricity theft detection. The following steps outline the methodology:

3.1 Data Collection

Efficient data preprocessing is paramount in guaranteeing the quality and reliability of machine learning models designed for electricity theft detection. This crucial phase encompasses a series of essential steps focused on converting raw smart grid data into a structured format suitable for model training and rigorous evaluation. The meticulous execution of these preprocessing steps lays the groundwork for robust and accurate model development, ensuring that the insights derived from the data are meaningful and contribute effectively to the overall success of electricity theft detection systems.

3.2 Data Collection

At the heart of this research lies the bedrock of a high-quality and diverse dataset. The historical data hails from smart grid systems and is sourced from the ISSDA and State Grid Corporation of China. This comprehensive dataset encapsulates rich information, including energy consumption patterns, billing records, and

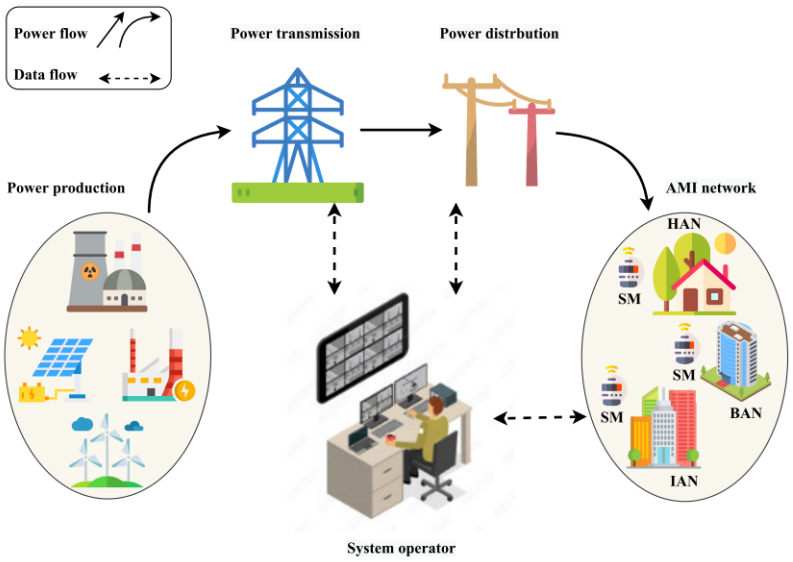


Figure 3.1: Smart grid model architcture

documented instances of confirmed electricity theft. The judicious utilization of such a robust dataset serves as a cornerstone, enabling a thorough exploration of patterns and anomalies essential for developing effective models in the domain of electricity theft detection.

3.2.1 Normalization

Normalization is a crucial preprocessing technique employed in data analysis and machine learning to standardize the scales of features within a dataset. The primary objective is to mitigate issues stemming from varying magnitudes and units across different features. By transforming numerical values into a standardized range, often between 0 and 1, normalization ensures that no single feature disproportionately influences the learning process. This fosters fair and unbiased contributions from all variables during model training, preventing the dominance of high-magnitude features.

In practical terms, normalization enhances the stability and convergence

of machine learning algorithms, allowing them to effectively generalize across diverse datasets. Common normalization methods include Min-Max scaling, Z-score normalization, and robust scaling, each catering to specific characteristics of the data and the requirements of the modeling task. Through normalization, practitioners can create a level playing field for features, promoting a more accurate and efficient learning process while facilitating meaningful comparisons and interpretations of the model's performance.

3.2.2 Anomaly Removal

Anomaly removal is a critical facet in the data preprocessing pipeline, essential for enhancing the reliability and performance of machine learning models. This process involves the identification and elimination of outliers or irregularities within a dataset that may distort the learning patterns of the model. By systematically detecting and appropriately handling anomalies, the overall integrity of the dataset is preserved, leading to more accurate and robust model training. Anomalous data points, which could arise from errors, noise, or genuine but rare occurrences, are carefully addressed to prevent their undue influence on the learning process. The objective is to create a refined dataset that captures the underlying patterns of normal behavior, allowing the machine learning model to discern subtle deviations indicative of potential electricity theft. An effective anomaly removal strategy contributes significantly to the model's generalization capacity, enabling it to better adapt to unseen data and ultimately improving the accuracy and reliability of electricity theft detection systems.

3.3 Model Construction

The model construction phase is a pivotal component of this research, involving the development of advanced machine learning architectures tailored for the task of electricity theft detection in smart grid systems. Two main categories of models are explored: Reinforcement Learning (RL) models and traditional Deep Learning (DL) models.

3.3.1 Reinforcement Learning Models

3.3.1.1 Deep Q-Network (DQN)

The study relies on the Deep Q-Network (DQN), a fundamental Reinforcement Learning (RL) model. This model harnesses the power of a neural network to approximate the Q-function, which assesses the desirability of taking a particular action in a given state. By aiming to learn a policy that maximizes cumulative rewards over time, the DQN equips itself to make informed decisions pertaining to energy consumption patterns.The training scheme for DQN method is illustrated in the Figure 3.2. This foundational RL approach forms the cornerstone of the study's methodology, facilitating a deeper understanding of complex decision-making processes in the context of smart power grids.

3.3.1.2 Double Deep Q-Network (DDQN)

Building upon the powerful foundation of the Deep Q-Network (DQN), the Double Deep Q-Network (DDQN) emerges as a sophisticated enhancement designed to tackle the challenge of potential overestimation biases in predicting anomalous energy consumption patterns within the context of smart grids. In the pursuit of accurate

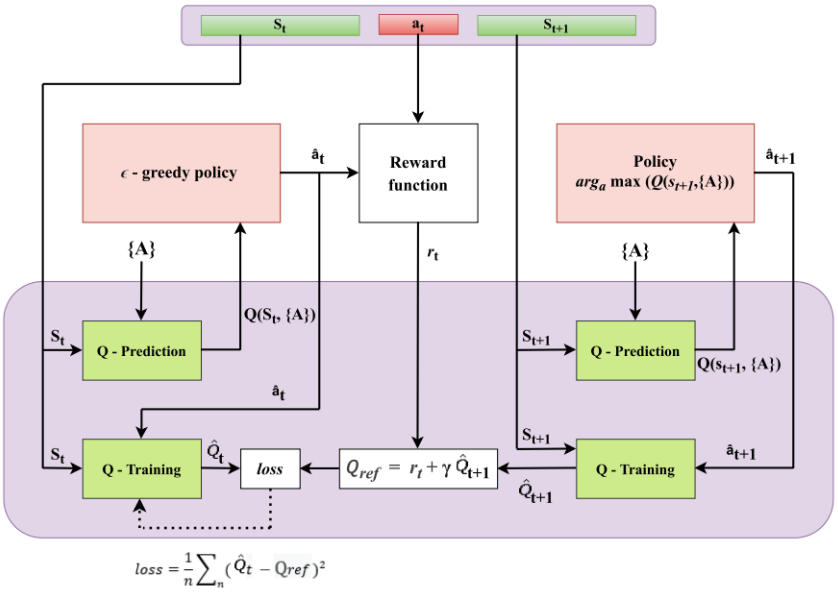


Figure 3.2: DQN training scheme

anomaly detection, the DDQN introduces a nuanced approach by decoupling action selection and evaluation. The DQN, while robust, is susceptible to overestimating Q-values, leading to suboptimal decision-making, especially when errors accumulate over time. The action selection network focuses on choosing the most promising action based on the current state, while the Q-value evaluation network provides a more reliable assessment, mitigating the potential for overestimation biases. The training scheme for DDQN method is illustrated in the Figure 3.3. This architectural innovation refines the learning process, enhancing the model's ability to discern subtle deviations in energy consumption patterns and identify anomalies such as electricity theft. The DDQN represents a strategic evolution, offering improved accuracy and effective decision-making in the dynamic and complex smart grid environment, thus contributing to the continuous refinement of machine learning architectures for real-world applications.

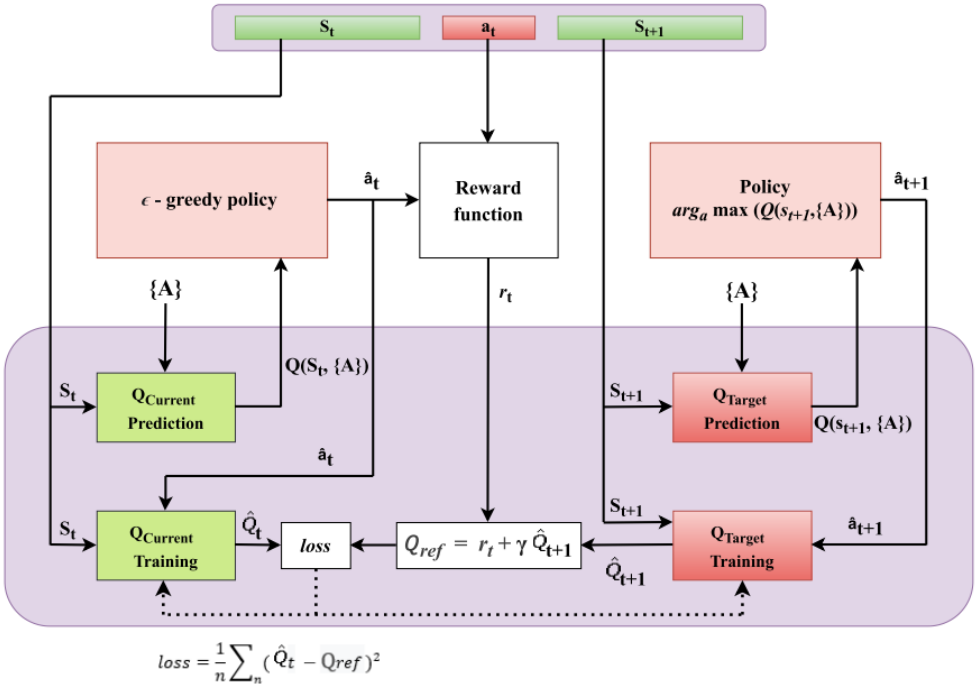


Figure 3.3: DDQN training scheme

3.3.2 Comparison Deep Learning Models

To benchmark the RL models, a suite of traditional Deep Learning (DL) models is constructed, each tailored to capture specific aspects of the dataset:

3.3.2.1 Feedforward Neural Network (FFNN)

The utilization of a Feedforward Neural Network (FFNN) proves instrumental in unraveling intricate relationships embedded within structured data, specifically tailored for the discerning task of detecting anomalies in the dynamic realm of energy consumption and billing records within smart grids. Renowned for its capacity to navigate and discern complex dependencies, the FFNN serves as a formidable tool for capturing the nuanced patterns and correlations inherent in the diverse input features. Through its layered architecture as shown in Figure 3.4, the FFNN excels in

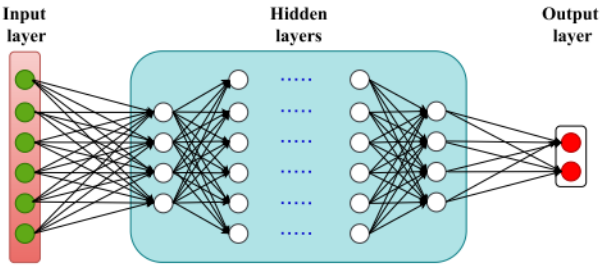


Figure 3.4: FFNN Architecture

processing information, making it particularly adept at recognizing subtle anomalies indicative of irregular energy consumption patterns. This model stands as a robust solution, leveraging its ability to unveil the intricacies of structured data, ultimately contributing to the overarching goal of enhancing anomaly detection mechanisms in the context of smart grid applications.

3.3.2.2 Convolutional Neural Network (CNN)

Recognizing the spatial patterns inherent in multidimensional data, a Convolutional Neural Network (CNN) is crafted. This model excels at feature extraction, enabling it to identify spatial anomalies that may indicate electricity theft.

A Convolutional Neural Network (CNN) is a specialized type of artificial intelligence model, often used for tasks like image recognition. What makes it stand out is its ability to automatically learn and extract important features from data, making it particularly effective in identifying patterns and anomalies within spatial information. CNNs have proven to be valuable in various fields, from image analysis to electricity theft detection, thanks to their knack for understanding complex relationships in multidimensional data. The typical architecture of CNN is shown in Figure 3.5.

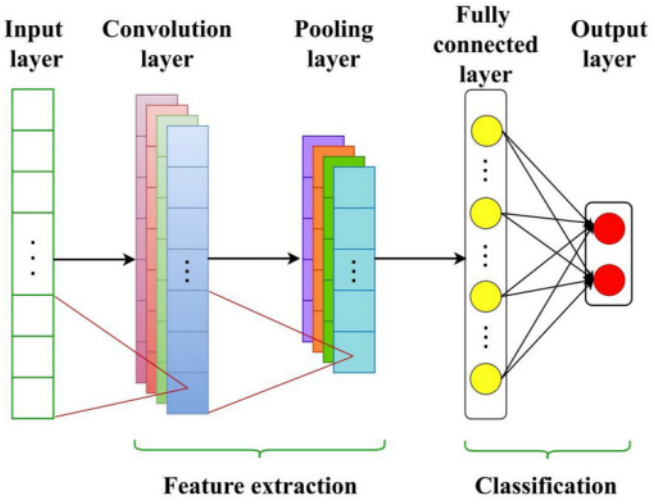


Figure 3.5: CNN Architecture

3.3.2.3 Recurrent Neural Network (RNN)

A Recurrent Neural Network (RNN) is a type of neural network designed to process sequential data, making it adept at capturing patterns over time as shown in Figure 3.6. RNN plays a crucial role in analyzing temporal aspects of the data, especially in understanding changes in electricity consumption patterns. By leveraging the sequential nature of smart grid data, RNN enhances the model's ability to recognize subtle variations and trends that may indicate electricity theft. Its application in this context allows for a more comprehensive analysis, complementing the Convolutional Neural Network (CNN) in providing a holistic understanding of both spatial and temporal patterns for effective detection.

3.3.2.4 Gated Recurrent Unit (GRU)

Acknowledging the sequential nature inherent in energy consumption data, the integration of a Gated Recurrent Unit (GRU) emerges as a pivotal augmentation. As a specialized variant within the Recurrent Neural Networks (RNN) framework,

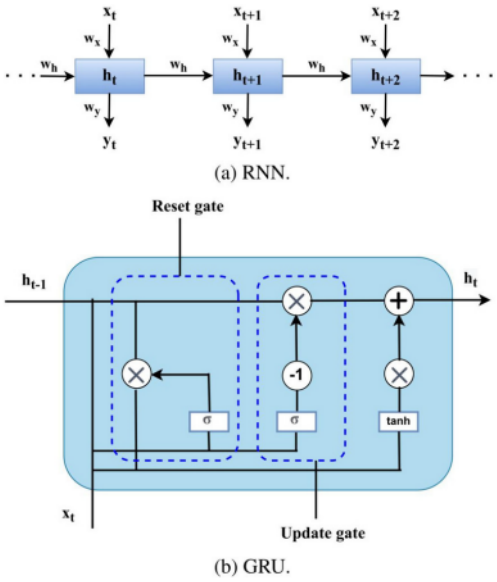


Figure 3.6: GRU and RNN Architecture

the GRU exhibits exceptional proficiency in capturing temporal dependencies.The typical architecture of RNN is shown in Figure 3.6. By seamlessly incorporating the GRU, the model achieves an elevated capacity to comprehensively analyze the temporal dimensions, contributing significantly to the efficacy of anomaly detection, especially in the nuanced context of electricity theft.

3.3.2.5 Hybrid CNN-GRU Model

To synergize spatial and temporal feature extraction, a hybrid model combining CNN and GRU architectures is implemented. This amalgamation aims to provide a holistic understanding of the dataset, considering both spatial and temporal intricacies.

3.4 Benchmarking

DL models, including Feedforward Neural Network (FFNN), Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), and a combination of CNN and GRU (CNN+GRU), are strategically employed as benchmarks in the comprehensive evaluation of the performance of Reinforcement Learning (RL) detection models. The utilization of these diverse DL benchmarks allows for a nuanced analysis, enabling a thorough understanding of the relative strengths and weaknesses of RL detection models across a spectrum of key performance metrics. This comparative framework contributes valuable insights into the potential advancements and distinctive attributes that RL models bring to the forefront in the domain of detection methodologies.

3.5 Dataset Preparation

In this section, we detail the preparation of our dataset for training and evaluating the proposed electricity theft detection model. The Irish smart energy trails dataset, released by Electric Ireland and the Sustainable Energy Authority, is used, focusing on readings from 130 randomly selected customers, resulting in 69,680 benign samples.The main characteristics of this dataset are presented in Table 3.1. To supplement the dataset with malicious samples for model training, they employ cyber-attacks from prior work. Six distinct attacks manipulate true electricity consumption readings, simulating instances of theft comprehensive dataset, featuring both benign and synthetically generated malicious samples, forms the basis for training and evaluating our model, aiming to enhance its robustness in detecting real-world electricity

Table 3.1: Irish Dataset Charecteristics

Description of data	Value
Data consumption time frame	536 days
No. of customers	3600+
Fine-grained interval	30 Min
No. of readings per day (R)	48
No. of the considered customers	130

Table 3.2: Cyber attack functions

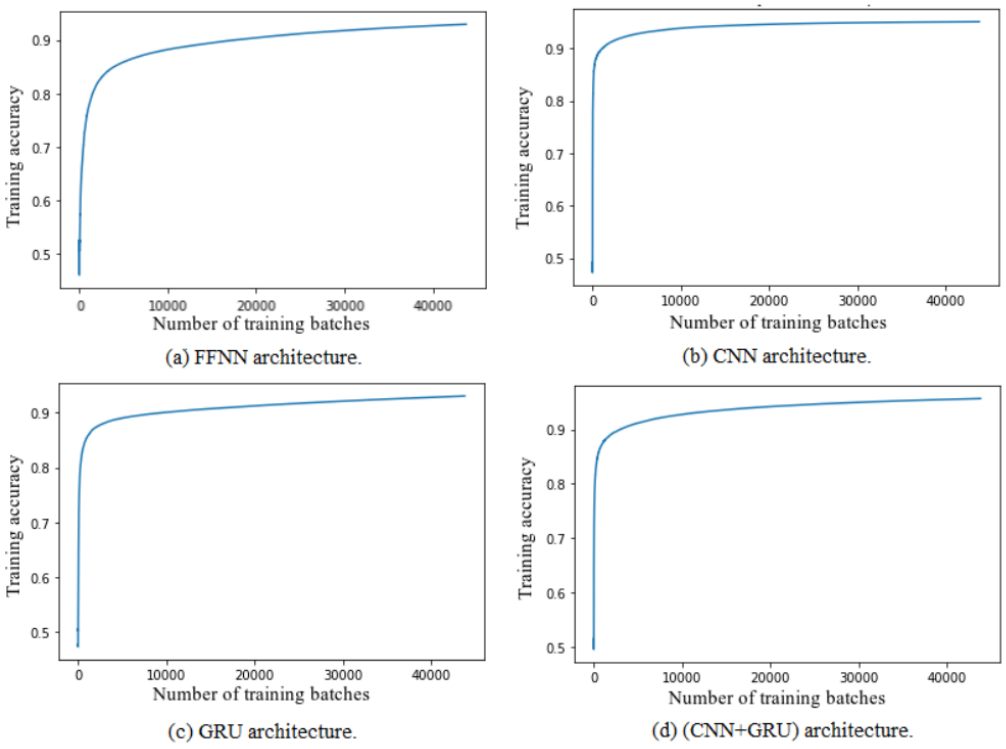
Attack No	Attack function
1^{st}	$f_1(x_i(t)) = \beta x_i(t)$
2^{nd}	$f_2(x_i(t)) = \beta_t x_i(t)$
3^{rd}	$f_3(x_i(t)) = mean(x_i)$
4^{th}	$f_4(x_i(t)) = \beta_t mean(x_i(t))$
5^{th}	$f_5(x_i(t)) = \begin{cases} 0 & t \in [t_s, t_e] \\ x_i(t) & t \notin [t_s, t_e] \end{cases}$
6^{th}	$f_6(x_i(t)) = x_i(R - t)$

theft scenarios. The six different attack functions are illustrated in the Table 3.2.

3.6 Training Accuracy

The evaluation of training accuracy serves as a critical element in appraising the effectiveness of diverse architectures within Double Deep Q-Network (DDQN)-based global models, as depicted in Figure 3.7. This visual representation offers a transparent understanding of how training accuracy dynamically evolves across various model structures. A notable observation unveils a discernible upward trend in training accuracy as the training batches expand. This trend suggests a positive correlation between batch size and the efficiency of model learning. The nuanced scrutiny

Figure 3.7: Training accuracy of different architectures of DDQN-based global model



of these patterns provides valuable insights into the iterative refinement and convergence of DDQN-based global models, shedding light on the profound impact of training dynamics on their overall accuracy.

3.7 Performance Evaluation

In the expansive domain of performance evaluation, a sophisticated arsenal of key metrics, notably comprising Accuracy (ACC), Precision, Recall, False Alarm Rate (FA), False Negative Rate (FNR), Hamming Distance (HD), and F1 score, stands as the linchpin for meticulously scrutinizing the effectiveness of detectors. This comprehensive framework, akin to a prism, allows for a nuanced dissection of detector performance, shedding light on intricate aspects often obscured in simpler analyses. Venturing into the heart of this evaluative endeavor, the focal point lies in

an exhaustive comparative exploration of diverse Deep Learning (DL) and Reinforcement Learning (RL) detection models, unfurling the tapestry of their unique attributes and vulnerabilities within the tapestry of real-world scenarios.

The discerning examination of these cardinal metrics not only lays the groundwork for a profound understanding of the intricacies characterizing DL and RL detection models but also serves as a beacon for informed decision-making, guiding their application in specific operational contexts. Embedded within this evaluative fabric, Equations 1 through 7 emerge as formidable tools, intricately calculating the system's Accuracy, Precision, Recall, False Alarm Rate, False Negative Rate, Hamming Distance, and F1 score. Together, they constitute a quantitative compass, navigating the labyrinth of detector performance with an unparalleled level of granularity and sophistication.

Equations 1, 2, 3, 4, 5, 6 and 7 are used to determine the system's Accuracy, precision, recall, False Alarm Rate, False Negative Rate, Highest Distance and F1 score respectively.

$$\text{Accuracy (ACC)} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} \tag{1}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}} \tag{3}$$

$$\text{False Alarm Rate (FA)} = \frac{\text{False Positives}}{\text{False Positives + True Negatives}} \tag{4}$$

$$\text{False Negative Rate (FNR)} = \frac{\text{False Negatives}}{\text{False Negatives + True Positives}} \tag{5}$$

$$\text{Highest Distance (HD)} = \frac{\text{Unequal Positions}}{\text{Total Positions}} \tag{6}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision + Recall}} \tag{7}$$

CHAPTER 4

EXPERIMENTS AND RESULTS

4.1 Experimental Results of Scenario 1

In scenario one, a comprehensive evaluation of global RL detection models is conducted, incorporating Double Deep Q-Network (DDQN) architectures with various Deep Learning (DL) structures, namely Feedforward Neural Network (FFNN), Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), and a hybrid architecture (CNN+GRU). Utilizing these DL models as benchmarks, training is performed on the dataset outlined in Section IV, with specific parameters detailed in Tables 4.1-4.7. Figure 3.7 illustrates the training accuracy's correlation with the number of training batches, showcasing superior and faster convergence in both CNN and hybrid (CNN+GRU) architectures compared to FFNN and GRU.

The performance evaluation in scenario one, quantified in Table 4.5, demonstrates the efficacy of DL-based, DQN-RL-based, and DDQN-RL-based global detectors. Notably, CNN and GRU detectors outperform FFNN detectors due to their distinctive characteristics. The hybrid architecture of CNN and GRU excels, leveraging the effective combination of their unique features. All detectors exhibit superior performance when utilizing RL compared to DL, attributed to RL's ability to optimize actions using reward concepts during training. Specifically, the DDQN-RL-based hybrid (CNN+GRU) detector surpasses all others, emphasizing its effectiveness in discerning anomalies related to electricity theft. The experimental results reveal notable improvements in key metrics such as Highest Distance (HD), False

Table 4.1: The parameters of the FFNN detection model

Architecture	Parameters		
	Layer	Number of units	AF
FFNN	Input	48	Linear
	Dense	512	Relu
	Dense	700	Relu
	Dense	850	Relu
	Dense	1024	Relu
	Dense	512	Relu
	Dense	256	Relu
	Dense	200	Relu
	Dense	50	Relu
	Output	2	Softmax

Table 4.2: The parameters of the CNN detection model

Architecture	Parameters		
	Layer	Number of units	AF
CNN	Input	48	Linear
	Conv1D	32	Relu
	Conv1D	64	Relu
	Conv1D	128	Relu
	Dense	64	Relu
	Dense	128	Relu
	Dense	256	Relu
	Dense	256	Relu
	Dense	512	Relu
	Dense	2	Softmax

Alarm (FA) reduction, and F1 score enhancement, emphasizing the superior performance of DDQN-RL-based detectors over DL-based counterparts and highlighting the hybrid (CNN+GRU) structure as the optimal choice for subsequent scenarios.

4.2 Experimental Results of Scenario 2

Scenario two unfolds with a strategic focus on customization, where a dedicated Double Deep Q-Network (DDQN)-based detection model is meticulously crafted for a new customer joining the smart grid. The approach involves leveraging the readings of the new customer to retrain a personalized iteration of the global detec-

Table 4.3: The parameters of the GRU detection model

Architecture	Parameters		
	Layer	Number of units	AF
GRU	Input	48	Linear
	GRU	64	Sigmoid
	GRU	64	Tanh
	Dense	64	Relu
	Dense	128	Relu
	Dense	2	Softmax

Table 4.4: The parameters of the hybrid (CNN+GRU) detection model

Architecture	Parameters		
	Layer	Number of units	AF
CNN+GRU	Input	48	Linear
	Conv1D	64	Relu
	Conv1D	64	Relu
	Conv1D	128	Relu
	GRU	64	Tanh
	GRU	64	Tanh
	GRU	64	Tanh
	Dense	64	Relu
	Dense	128	Relu
	Dense	2	Softmax

Table 4.5: Comparison between the performance of DL, DQN-RL, and DDQN-RL global models

Metrics	DL				DQN-RL				DDQN-RL			
	FFNN	CNN	GRU	CNN+GRU	FFNN	CNN	GRU	CNN+GRU	FFNN	CNN	GRU	CNN+GRU
ACC (%)	92.42	93.14	91.10	94.71	95.02	95.84	95.84	96.84	94.63	95.22	95.82	97.33
Precision (%)	92.41	92.78	91.67	93.68	95.10	95.93	95.90	96.89	94.86	95.50	95.94	97.38
Recall (%)	92.40	93.52	90.38	95.84	95.02	95.84	95.84	96.84	94.63	95.22	95.82	97.33
FA (%)	7.56	7.23	8.17	6.42	4.47	2.99	3.75	2.68	2.43	1.37	2.48	2.06
FNR (%)	7.59	6.47	9.62	4.15	5.48	5.32	4.57	3.64	8.30	8.16	5.86	3.27
HD (%)	84.85	86.30	82.21	89.43	90.55	92.86	92.09	94.16	92.20	93.86	93.35	95.27
F1 (%)	92.40	93.15	91.02	94.75	95.06	95.89	95.87	96.86	94.74	95.36	95.88	97.35

Table 4.6: Comparison between the performance of the global and customized models for different customers

Metrics	Global model	Customized models		
		Customer 14	Customer 30	Customer 48
ACC (%)	97.332	99.439	98.102	99.502
Precision (%)	97.388	99.441	98.119	99.507
Recall (%)	97.332	99.439	98.102	99.502
FA (%)	2.06	0.25	0.93	0.14
FNR (%)	3.27	0.86	2.88	1.00
HD (%)	95.27	99.19	97.17	99.36
F1 (%)	97.355	99.440	98.110	99.504

tion model. This process, conducted sample by sample, results in the emergence of a finely tuned DDQN-based customized detection model. The training accuracy progression for the DDQN-based hybrid (CNN+GRU) customized model for customer 14 is vividly depicted in Figure 4.1, showcasing a commendable increase with the augmentation of training batches

Table 4.6 provides a comprehensive comparison between the performance of the global detection model and the customized detection models for three randomly selected new customers. To ensure equitable comparison, all DDQN-based global and customized detection models share the same hybrid (CNN+GRU) structure. The results outlined in Table 9 underscore the superiority of the customized detection models, demonstrating enhanced performance for different customers. Across various metrics, including ACC, precision, recall, HD, and F1-score, the customized models consistently outperform the global detection model. Notably, false alarm (FA) and false negative rate (FR) are lower, emphasizing the efficacy of the tailored approach in achieving superior detection accuracy for individual customers.

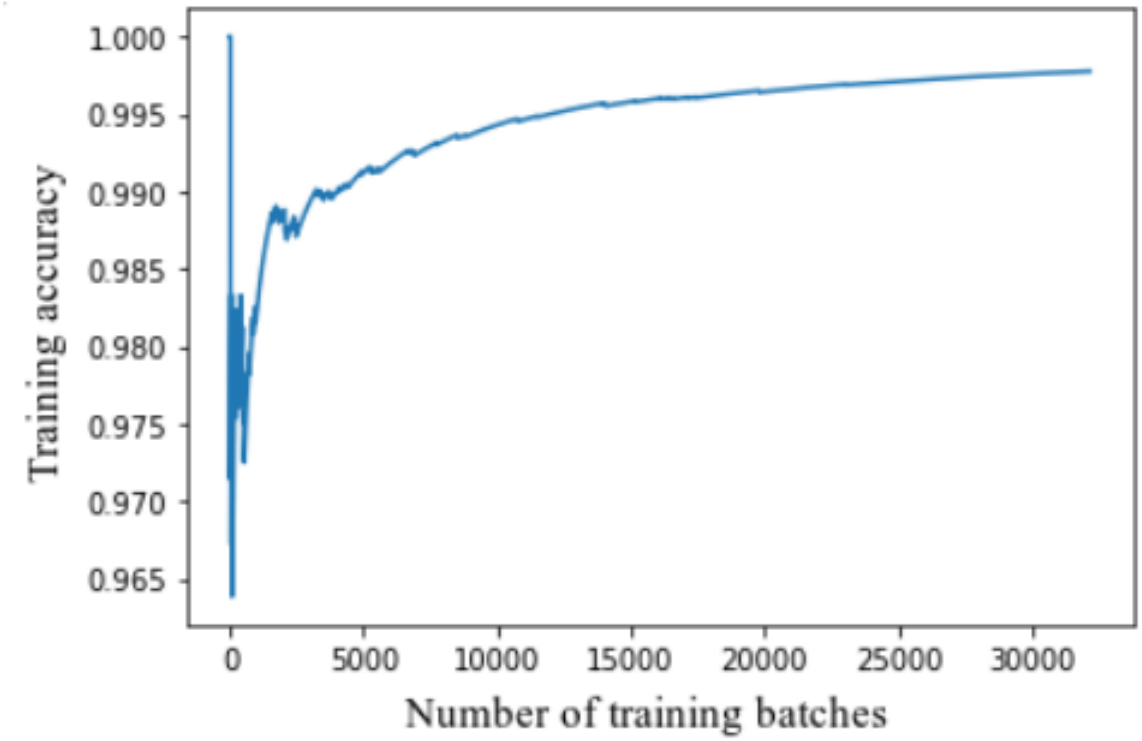


Figure 4.1: Training accuracy of DDQN-based hybrid(CNN+GRU) customized model for customer 14

4.3 Experimental Results of Scenario 3

In scenario three, the dynamic aspect of customer behavior is addressed by considering changes in the consumption pattern of an existing customer. Customer consumption patterns can shift due to various factors such as alterations in the number of occupants or the acquisition of new electric appliances. In response to these changes, the DDQN-based customized detection model for the affected customer undergoes a meticulous retraining process, incorporating the new consumption readings. This adaptive approach ensures the model remains attuned to evolving consumption behaviors. Table 4.7 provides a comprehensive comparison between the results of the initial and updated customized detection models for different customers, reflecting the impact of changes in their consumption patterns.

All DDQN-based customized detection models maintain a consistent hybrid (CNN+GRU) structure. The results in Table 10 demonstrate the superior performance of the updated customized detection models, post-retraining, compared to their pre-retraining counterparts. Notably, both accuracy (ACC) and Hamming Distance (HD) witness significant improvements, with increases of up to 20 percentage and 55 percentage, respectively. Moreover, false alarm (FA) and false negative rate (FNR) show marked reductions of up to 33 percentage and 26 percentage, respectively. Comparing the performance of the updated customized detection models to the original models in Table 4.6 reveals a close match, underlining the adaptability of the models to changes in consumption patterns for different customers.

Table 4.7: The performance of old and updated customized models for different customers due to changing their consumption behavior

Metrics	Old customized model			Updated customized models		
	Customer 14	Customer 30	Customer 48	Customer 14	Customer 30	Customer 48
ACC (%)	78.095	79.595	79.9	99.222	99.626	99.248
Precision (%)	79.865	79.717	80.519	99.233	99.234	99.230
Recall (%)	78.095	79.595	79.9	99.222	99.626	99.248
FA (%)	33.56	19.14	12.58	0.11	0.37	0.12
FNR (%)	10.0	15.06	27.84	1.57	0.38	1.45
HD (%)	44.54	60.455	67.32	99.11	99.26	99.10
F1 (%)	78.970	79.655	80.208	99.228	99.429	99.238

4.4 Experimental Results of Scenario 4

In scenario four, the investigation pivots towards assessing the model's resilience against recently discovered cyber-attacks. Initial training of a global detection model involves specific attacks outlined in Table 3.2 (1st, 2nd, and 4th attacks). Subsequently, customers deploy novel techniques in the form of the 3rd, 5th, and 6th attacks. To counter these evolving threats, malicious samples from these new attacks are strategically employed for the meticulous retraining of the global detection model. This retraining process unfolds in a granular manner, with each sample contributing to the refinement of a Double Deep Q-Network (DDQN)-based customized detection model.

Table 4.8 serves as a comprehensive repository, presenting a comparative analysis between the outcomes of the global models (pre-retraining) and the customized detection models tailored to address new cyber-attacks across diverse customers. With a unified adherence to a hybrid (CNN+GRU) structure, the results accentuate the adaptive prowess of the customized detection model. Across varying customers, discernible enhancements in accuracy (ACC) and Hamming Distance (HD) stand out, showcasing increases of up to 21 percentage and 30 percentage, re-

Table 4.8: Comparison between the performance of the global and customized models for newly launched attacks for different customers.

Metrics	Global model				Newly launched attacks customized models			
	Customer 5	Customer 20	Customer 25	Customer 35	Customer 5	Customer 20	Customer 25	Customer 35
ACC (%)	77.753	81.704	77.069	79.060	99.782	98.444	98.630	99.751
Precision (%)	80.271	83.648	81.416	85.170	99.387	98.481	98.648	99.752
Recall (%)	77.753	81.704	77.069	79.060	99.782	98.444	98.630	99.751
FA (%)	8.04	6.20	4.45	3.01	0.65	0.19	0.57	0.42
FNR (%)	36.21	30.48	41.14	42.68	0.44	2.90	2.16	0.49
HD (%)	69.72	75.5	72.62	82.16	99.13	98.26	98.07	99.33
F1 (%)	78.992	82.665	79.182	82.001	99.584	98.462	98.639	99.751

spectively. Simultaneously, there are substantial reductions in false alarms (FA) and false negative rates (FNR), reaching up to 7.4 percentage and 42 percentage, respectively. These compelling findings affirm the superior capability of the customized detection model in effectively identifying instances of electricity theft amidst emerging and dynamic cyber threats

4.5 A Comparative Analysis of Four Scenarios

The evaluation of the proposed electricity theft detection model unfolds through four distinct scenarios, each offering unique insights into the model's adaptability and effectiveness. In Scenario One, where a global detection model is crafted using various Double Deep Q-Network (DDQN)-based architectures, the hybrid (CNN+GRU) structure emerges as the standout performer, showcasing superior convergence and performance compared to other architectures.

Scenario Two delves into the customization realm, constructing DDQN-based personalized detection models for new customers. The results affirm the effectiveness of this tailored approach, consistently outperforming the global detection model across different metrics.

In Scenario Three, the model's robustness is tested against changes in con-

sumption patterns for existing customers. The adaptive retraining process yields significant improvements in accuracy and Hamming Distance, highlighting the model's resilience in accommodating dynamic shifts in customer behavior.

The investigation extends to Scenario Four, exploring the model's response to newly discovered cyber-attacks. The customized detection model, refined through sample-by-sample retraining, demonstrates notable enhancements in accuracy and Hamming Distance, coupled with substantial reductions in False Alarms and False Negative Rates.

The experimental results emphasize the DDQN-based electricity theft detection model's versatility and efficacy. The hybrid (CNN+GRU) architecture consistently proves adaptable and robust across diverse smart grid scenarios.

CHAPTER 5

CONCLUSION

In conclusion, this seminar paper explores the application of Reinforcement Learning (RL) in identifying electricity theft cyber-attacks within smart power grids. Through a systematic investigation, a set of RL-based detectors is proposed, utilizing malicious reading samples generated from real power consumption data subjected to cyber-attacks. The findings reveal that RL-based detectors, particularly the global detectors of RL-based Deep Q Network (DQN) and Double Deep Q Network (DDQN), outperform their Deep Learning (DL) counterparts, showcasing lower False Alarm (FA) rates and higher Hamming Distance (HD).

Notably, the hybrid architecture of Convolutional Neural Network and Gated Recurrent Unit (CNN+GRU) emerges as the most effective, capitalizing on the distinctive characteristics of both components. Customized RL-based detectors for new customers demonstrate superior performance compared to global detection models, emphasizing the adaptability of RL in personalized scenarios. Furthermore, the study explores changes in consumption patterns and underscores the resilience of RL-based detection models, maintaining comparable performance even after such alterations. The fourth scenario delves into the capability of the model to adapt to new cyber-attacks, showcasing the model's capacity to learn and achieve high accuracy, recall, and HD, while concurrently reducing FA. Overall, this research underscores the efficacy of RL-based detectors in addressing dynamic challenges within smart power grids, offering promising avenues for enhanced cybersecurity measures in the detection of electricity theft cyber-attacks.

REFERENCES

[1] M. I. Ibrahem, M. Nabil, M. M. Fouda, M. M. E. A. Mahmoud, W. Alasmary, and F. Alsolami, "Efficient privacy-preserving electricity theft detection with dynamic billing and load monitoring for AMI networks", *IEEE Internet Things J., vol. 8, no. 2, pp. 1243–1258, Jan. 2021.*

[2] A. Takiddin, M. Ismail, M. Nabil, M. M. E. A. Mahmoud, and E. Serpedin "Detecting electricity theft cyber-attacks in AMI networks using deep vector embeddings", *IEEE Syst. J., vol. 15, no. 3, pp. 4189–4198, Sep. 2021.*

[3] M. I. Ibrahem, M. M. Badr, M. M. Fouda, M. Mahmoud,W. Alasmary, and Z. Md. Fadlullah "PMBFE: Efficient and privacy-preserving monitoring and billing using functional encryption for AMI networks", *in Proc. Int.* Symp. Netw., Comput. Commun. (ISNCC), Oct. 2020, pp. 1–7.

[4] M. I. Ibrahem, M. M. Badr, M. Mahmoud, M. M. Fouda, andW. Alasmary. "Countering presence privacy attack in efficient AMI networks using interactive deep-learning", *in Proc. Int. Symp. Netw., Comput. Commun*.

[5] M. M. Badr, M. I. Ibrahem, M. Mahmoud, M. M. Fouda, F. Alsolami, and W. Alasmary, "Detection of false-reading attacks in smart grid netmetering system", *IEEE Internet Things J., vol. 9, no. 2, pp. 1386–1401, Jan. 2022*.