Smart Grid intrusion detection system based on AI techniques

Mounir Mohammad Abou-Elasaad[1,*], Samir G. Sayed[2], Mohamed M. El-Dakroury[3]

[1]Department of Electronics & Communications, Faculty of Engineering Egypt, Helwan University, Egypt
[2]Professor, Department of Electronics and communication Engineering, Helwan University, Egypt
[3]Assistant Professor, Department of Electronics and Communications Engineering, Helwan, Egypt

Emails: Mounir_abouelkhair@h-eng.helwan.edu.eg; samir_abdelgawad@h-eng.helwan.edu.eg; mdakroury@h-eng.helwan.edu.eg

Abstract

Smart grids (SGs) are integral to modern utility systems, managing power generation, energy consumption, and communication networks. However, as these systems become increasingly interconnected, they are exposed to sophisticated cyber threats that can compromise their functionality and security. To address these challenges, this paper presents an AI-driven detection framework designed to significantly enhance cybersecurity in smart grids. The proposed system combining Recurrent Neural Networks (RNNs) with Support vector classifier to improve detection accuracy, recognition capabilities, and system robustness. The methodology comprises four main stages: (1) data preprocessing to ensure high-quality input for analysis, (2) traffic detection using RNNs to capture temporal patterns, (3) classification of traffic as normal or abnormal via support vector classifier (SVC), and (4) identification of specific attack types through another SVC for refined threat categorization. This integrated approach enables real-time detection of both known and emerging threats, focusing on minimizing false positives and maximizing detection precision. The system was evaluated on three comprehensive benchmark datasets: UNSW_NB15 and BoT-IoT, achieving an average accuracy of 100%. These results underscore the superiority of this AI-based solution over traditional intrusion detection systems, providing a robust and scalable framework for securing smart grids and other critical infrastructures.

Keywords: Smart Grid; Cyber-Attacks; Vulnerabilities; Artificial Intelligence; Detection Method; Advanced Technologies

1. Introduction

The global shift towards more sustainable and efficient energy systems has accelerated the widespread adoption of smart grids (SGs), which represent a transformative advancement in energy management. By integrating advanced communication technologies, data analytics, and automated control systems, smart grids optimize energy generation, distribution, and consumption. Unlike traditional grids, SGs rely on the Internet of Things (IoT) devices, big data, and machine learning to facilitate real-time monitoring, self-diagnostics, and automated responses to energy demands and operational faults [1]. However, this heightened intelligence and interconnectivity have simultaneously exposed SGs to complex cybersecurity risks. AI-based methods have shown remarkable potential in detecting both known and novel threats, analyzing vast datasets to discern malicious patterns without misidentifying legitimate traffic as threats [12]. Additionally, the scalability of AI is well suited to the substantial data volumes generated by SGs, enabling faster threat detection and response times [13]. Unlike traditional detection methods with rigid patterns, AI-driven systems are adaptive and learn continuously, providing more robust defenses against dynamic cyber threats [14].

In recent years, cyber threats targeting smart grids have surged in both sophistication and frequency, encompassing Distributed Denial of Service (DDoS) attacks, malware infections, and data manipulation techniques [2]. These cyberattacks compromise the reliability of electricity supply, causing substantial economic damage and posing risks to national security by targeting critical infrastructure. Traditional security measures, such as encryption and firewalls, offer limited protection against these advanced threats, underscoring the need for more sophisticated detection mechanisms. Common attack vectors include phishing, where attackers deceive personnel into revealing

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

login credentials, enabling unauthorized access and operational disruptions. Malware is another prevalent threat, infiltrating systems to cause intermittent or permanent outages, data theft, or manipulation, often resulting in severe degradation of essential grid components. Such cyber incidents can lead to widespread blackouts, impact millions, and inflict financial losses while threatening critical services, including healthcare and emergency response [3]. The interconnected nature of SGs exacerbates this vulnerability; a security breach in one component can propagate throughout the network, magnifying the scale of an attack and its consequences [4].

To address the increasing cybersecurity challenges in smart grids (SGs), numerous detection methodologies have been developed, with signature-based and anomaly-based detection systems being the most prevalent approaches. Signature-based detection systems operate by identifying cyberattacks through pre-established patterns or known signatures. While effective for recognizing previously documented threats, these systems face significant limitations when dealing with zero-day vulnerabilities or novel attack vectors that do not match existing signatures. This inherent constraint makes them less effective in dynamic and evolving threat landscapes. In contrast, anomaly-based detection systems monitor operational behaviors to identify deviations from standard or expected patterns, flagging these deviations as potential intrusions. These systems are advantageous in detecting unknown threats and sophisticated attacks that evade signature-based mechanisms. However, a key challenge with anomaly-based systems lies in their high false-positive rates, which can lead to unnecessary alarms, resource misallocation, and delayed response times ultimately reducing their overall reliability in high-stakes environments like SGs. To overcome these challenges, AI-driven detection mechanisms have emerged as a transformative solution, leveraging advanced algorithms to enhance accuracy and responsiveness. For example, Muneeswari et al. proposed a multi-stage cyber intelligence framework designed specifically to mitigate false positives while improving detection accuracy in SG environments [1]. This approach integrates multiple AI techniques to provide a more robust and nuanced analysis of potential threats. Further, research conducted by Alsuwian and Koduru has demonstrated the effectiveness of AI-powered systems in detecting and responding to cyber threats. Their studies highlight AI's capability to recognize complex and evolving anomalous patterns within SG infrastructures, enabling rapid countermeasures that minimize system disruptions and ensure operational continuity [5,6].

In this context, this paper proposes a novel AI-driven detection system designed to address the limitations of current methods. The proposed system incorporates advanced artificial intelligence (AI) techniques, including RNNs, and SVC, to create a resilient framework capable of detect and recognize both known and unknown cyber threats in real time. This AI-based model aims to reduce false positives while enhancing overall detection accuracy and speed, providing a comprehensive approach to SG cybersecurity. Real-time capability is crucial in the proposed AI-driven cybersecurity framework for smart grids. It is a cornerstone of the effectiveness of the AI-based framework for enhancing cybersecurity in smart grids, enabling instant threat detection and response before significant damage occurs or operations are disrupted. This minimizes downtime, ensures grid stability, and protects critical infrastructure. By analyzing temporal patterns using Recurrent Neural Networks (RNNs), the system can identify evolving and dynamic threats, while Support Vector Classifiers (SVCs) provide precise threat classification, reducing false alarms and enabling targeted responses. The system also supports rapid decision-making by delivering immediate insights to operators and enhances collective security for interconnected grids through real-time information sharing. Furthermore, it dynamically adapts to changing conditions, such as rerouting energy or isolating compromised components, bolstering resilience against large-scale attacks. Thus, real-time capability serves as the foundation for ensuring the secure and efficient operation of smart grids in the face of escalating cyber threats.

The primary objectives of this research are as follows:

1) Identify key cybersecurity vulnerabilities in SGs by analyzing prominent security gaps that make SGs susceptible to cyber-attacks.
2) Develop a novel AI-based detection system capable of identifying both traditional and emerging cyber threats in real-time.
3) Evaluate the effectiveness of the proposed system using various datasets to assess its accuracy, precision, recall, and overall performance compared to existing intrusion detection systems (IDS).
4) Expand the applicability of the system by applying the AI-driven detection model to other critical infrastructures and industrial IoT systems.

The findings of this paper mark a significant advancement in smart grid (SG) security and the application of AI for detecting cyber-attacks, with a strong emphasis on the importance of real-time capabilities. The research introduces a real-time, AI-driven model that overcomes the limitations of traditional methods by effectively identifying both known and emerging cyber threats, ensuring immediate response to potential risks. By integrating deep learning techniques, such as Recurrent Neural Networks (RNNs), with traditional machine learning methods like Support Vector Classifiers (SVCs), the study enhances SG protection systems in terms of detection accuracy, scalability, and adaptability to dynamic cyber landscapes [15]. The real-time aspect is particularly crucial, as it

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

enables anomaly detection across various SG architecture layers, ensuring comprehensive and resilient security. This is vital given the interconnected nature of these layers, where vulnerabilities can propagate rapidly without prompt intervention [16]. Moreover, the study underscores the value of using publicly available datasets, such as UNSW_NB15 and BoT-IoT, for model training, establishing a foundation for standardized testing and benchmarking. By leveraging real-time capabilities, this research provides a robust framework for proactive and scalable SG cybersecurity, setting a new standard for future advancements in the field.

The remainder of this paper is structured as follows: Section 2 shows the problem statement of this paper. Section 3 provides Foundations and Basic Terminology. Section 4 describes the related works of the proposed AI-based system. Section 5 presents the methodology. Section 6 describes the used data sets. Section 7 shows experimental results and performance analysis of the detection model. Section 8 discusses the implications of the finding and give the research conclusion.

2. Problem Statement

As smart grids become increasingly reliant on interconnected Information and Communication Technologies (ICT), their susceptibility to cyberattacks has grown significantly. The integration of IoT, Industrial IoT (IIoT), and cloud computing into the smart grid infrastructure has notably expanded the attack surface, offering malicious actors new opportunities to exploit vulnerabilities and gain unauthorized access to critical components [17]. This interconnectedness introduces systemic risks, as any compromise in one segment of the infrastructure can cascade into widespread disruptions. The rapid evolution of cyber threats exacerbates these vulnerabilities, enabling attackers to swiftly adapt their strategies and exploit newly uncovered weaknesses. Traditional security mechanisms, such as encryption, firewalls, and even intrusion prevention systems, often fall short in addressing the multifaceted and adaptive nature of these threats. Although these measures provide a foundational layer of security, they are insufficient for detecting and mitigating emerging threats like zero-day vulnerabilities or sophisticated, persistent attacks. Signature-based detection systems, which rely on predefined patterns to identify threats, are ineffective against novel and evolving cyber threats. Similarly, anomaly-based detection systems—while adept at identifying deviations from normal behavior—suffer from high false-positive rates, which can result in delayed responses and strained resources. These limitations highlight the urgent need for more advanced cybersecurity solutions tailored to the unique challenges of smart grids. AI-driven detection systems offer a promising alternative by leveraging machine learning and deep learning to enhance detection capabilities. Unlike traditional methods, AI-based approaches can adapt to evolving threat landscapes, identify subtle and complex patterns indicative of potential breaches, and enable real-time responses. For instance, these systems can differentiate between benign anomalies and actual threats, thereby reducing false positives and ensuring timely intervention. The importance of these advancements is underscored by the rising frequency and sophistication of cyberattacks targeting smart grids. Cyber intrusions in this domain can lead to catastrophic consequences, including the compromise of critical infrastructure, financial losses, and large-scale disruptions to energy distribution. This underscores the necessity of developing proactive, robust, and scalable detection frameworks that can not only identify and mitigate risks but also fortify the overall resilience of the smart grid infrastructure.
By implementing AI-driven Intrusion Detection Systems (IDS), smart grids can achieve a higher level of security and reliability. These systems are capable of continuously monitoring network activities, predicting potential threats, and facilitating swift countermeasures. As global reliance on secure and stable energy systems grows, addressing smart grid vulnerabilities with real-time AI-based techniques has become indispensable for safeguarding these critical infrastructures against the increasingly complex cyber threat landscape [18].

3. Foundations and Basic Terminology

This section establishes foundational knowledge crucial to understanding the research context. It begins with an overview of neural networks, emphasizing their architecture and operational principles as the basis for advanced models such as Recurrent Neural Networks (RNNs). RNNs are highlighted for their ability to process sequential data, making them particularly effective for applications requiring temporal pattern recognition. The section also examines the standard architecture of Industrial Internet of Things (IIoT) systems, which typically comprises multiple interconnected layers, including sensing, network, edge, and cloud layers. Each layer plays a vital role in data collection, transmission, and processing, forming a comprehensive framework for industrial operations. Within this IIoT architecture, the role of Intrusion Detection Systems (IDS) is critically analyzed. IDS are integral to the security framework, tasked with monitoring network traffic, identifying potential threats, and mitigating cyber risks. This discussion lays the groundwork for exploring how RNN-based and other AI-driven IDS can address the unique cybersecurity challenges posed by IIoT environments.

A. Neural Networks

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

Biological neural networks refer to the complex networks of neurons found in living organisms, serving as an inspiration for the development of their artificial counterparts. On the other hand, artificial neural networks (ANNs) are computational models designed to mimic the functionality of biological neural networks. ANNs consist of fundamental building blocks called neurons, which are interconnected and work collaboratively to process information and solve specific problems. These neurons are organized into layers, including the input layer, hidden layers, and the output layer, each performing distinct roles in data processing. The neurons within a layer are connected to those in adjacent layers, forming a network architecture that enables the model to learn and adapt through processes such as weight adjustments and activation functions. Figures 1(a) and 1(b) provide a visual comparison, illustrating the structural differences between biological neural networks and ANNs, respectively. This comparison underscores the conceptual inspiration behind ANNs and their evolution into powerful tools for solving complex problems in domains such as pattern recognition, anomaly detection, and predictive analytics within IIoT environments.

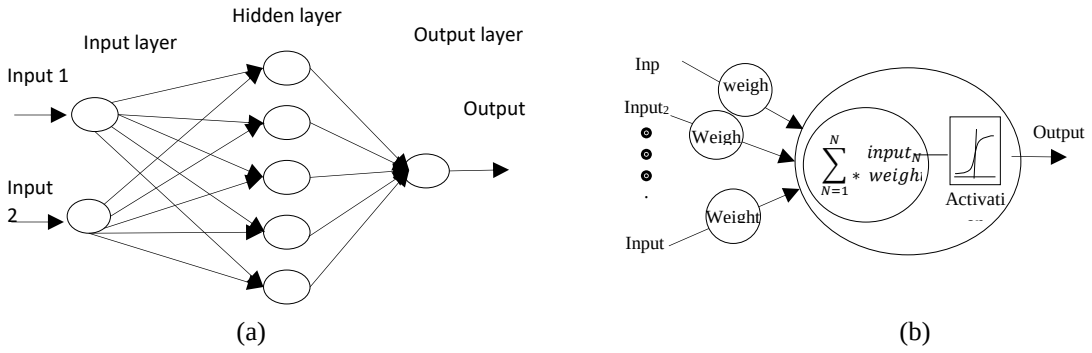

Figure 1. (a) ANN structure, and (b) ANN components

Deep learning (DL) refers to the use of artificial neural networks (ANNs) with multiple hidden layers to perform complex learning tasks. It is classified into three major classes (I) Generative architecture uses an unsupervised deep learning model. Commonly, it grew when there is limited data to train for the difficult network. (II) Discriminative architecture is widely used in information and signal processing, for example, language processing, recognition applications, and (III) Hybrid architectures comprise both generative and discriminative processes [35]. Among the most well-known varieties of deep learning networks are Deep Belief Networks (DBNs), Long Short-Term Memory networks (LSTMs), Recurrent Neural Networks (RNNs), Restricted Boltzmann Machines (RBMs), and Convolutional Neural Networks (CNNs) [7]. In this paper, we focused on RNNs for traffic detection in smart grid environments. RNNs were chosen due to their ability to model sequential patterns and dependencies in network traffic data. The results achieved using this approach were highly promising, demonstrating the potential of RNNs in enhancing traffic analysis and intrusion detection accuracy.

B. Recurrent neural network (RNN)

Recurrent Neural Networks (RNNs) are a specialized type of generative neural network characterized by their ability to use prior outputs as inputs, enabling them to process sequential and time-series data effectively. The foundational concept for RNNs was introduced by David Rumelhart in 1986, while the formal development was credited to Williams and Zipser in the late 1980s [36]. Additionally, John Hopfield's discovery of Hopfield networks in 1982 laid the groundwork for understanding a subset of RNN architectures, highlighting their capacity for associative memory. RNNs are uniquely designed to handle input data with varying lengths, such as sequential or time-series data. Their architecture is structured as a chain of repeating modules, often referred to as RNN cells, which sequentially process data, as illustrated in Figure 2. In this configuration: x represents the input at a given time step. h denotes the hidden state, which stores intermediate computations and incorporates information from previous steps. y is the output generated at each time step. These RNN cells function based on the distribution of hidden states, which dynamically update their weights using non-linear operations. This dynamic nature allows RNNs to learn temporal dependencies and make predictions based on sequential patterns in the input data. One of the defining features of RNNs is their ability to store information about past inputs efficiently, enabling the network to utilize historical context when generating outputs. This property is particularly advantageous in tasks such as natural language processing, speech recognition, and anomaly detection, where understanding sequential dependencies is critical. Moreover, RNNs can generate outputs not only based on current input data but also on predicted or inferred inputs, making them highly effective in predictive modeling and generative tasks.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

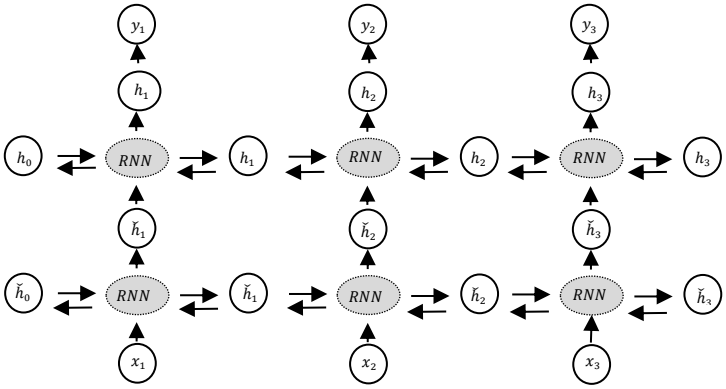

Figure 2. Repeating modules of recurrent neural network (RNN)

Recurrent Neural Networks (RNNs) are particularly effective for tackling challenges that involve sequence prediction, such as language modeling, text generation, video tagging, and music composition [36]. Their unique ability to process sequential data and capture temporal dependencies makes them suitable for tasks requiring the prediction of patterns over time. RNN-based sequence prediction problems can be categorized into three main types based on the relationship between the input and output sequences. (I) One-to-Many: In this sort of issue, an observation is mapped as an input to a sequence with numerous steps as an output. (II) Many-to-One: In this scenario, a series of numerous steps serving as the input are mapped to a class or quantity forecast. (III) Many-to-Many: A series of numerous steps serving as input is translated into a series of multiple steps serving as output. These categorizations highlight the versatility of RNNs in addressing diverse sequence prediction challenges across various domains, from natural language processing to video analysis. By tailoring their architectures to specific input-output relationships, RNNs can effectively model and predict patterns within sequential data. In the proposed system, we leverage the sequential processing capabilities of Recurrent Neural Networks (RNNs) to handle data in a structured, time-dependent manner. The sequential nature of RNNs enables the model to capture temporal dependencies and extract meaningful patterns from data streams that unfold over time. This approach is particularly advantageous for the problem at hand, as it aligns with the inherent structure of the input data, which is typically presented as time-series or sequential datasets. By processing the data systematic and maintaining a hidden state that encapsulates information from prior steps, the system ensures that contextual relationships within the sequence are preserved and utilized to make informed predictions. Such a design is especially critical in traffic detection, where the ability to understand and predict patterns based on historical data significantly enhances detection accuracy and system reliability. The use of RNNs in this sequential framework ensures that the proposed system is well suited for identifying dynamic trends and responding to them in real time.

C. IIoT System Architecture

The architecture of the Industrial Internet of Things (IIoT) typically comprises several layers, each fulfilling distinct functions that collectively enhance the overall efficiency and reliability of industrial processes [10]. The primary layers include (see Figure 3):

1. **Perception Layer**: This is a foundational layer consists of various sensors and devices that gather data from the physical environment. Key parameters such as temperature, pressure, humidity, and other relevant metrics are collected, serving as crucial inputs for monitoring industrial processes. As the cornerstone of the IIoT architecture, the perception layer provides the raw data essential for subsequent processing and analysis [19].
2. **Network Layer**: It is responsible for data communication and connectivity; the network layer ensures that data collected by the perception layer is transmitted to the appropriate destinations. This layer employs a range of wired and wireless communication technologies, such as Ethernet, Wi-Fi, and 5G, to facilitate reliable and secure data transmission [19].
3. **Application Layer**: The application layer is tasked with processing and analyzing the transmitted data to generate actionable insights that support decision-making in industrial contexts. By leveraging advanced analytics, machine learning, and artificial intelligence, this layer transforms raw data into valuable information. Its role is crucial for optimizing industrial operations, enhancing efficiency, and reducing costs [20]. Additionally, IIoT architecture often incorporates cloud computing, which allows for data processing closer to the source, thereby reducing latency and bandwidth usage [21].

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

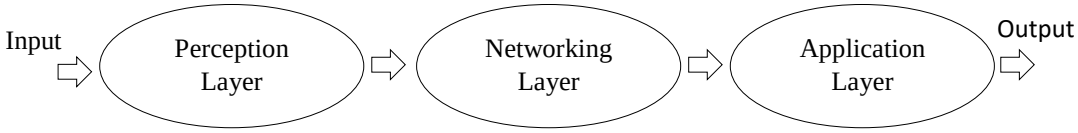

Figure 3. The architecture of IIoT system.

D. Intrusion Detection Systems

Intrusion Detection Systems (IDS) play a vital role in securing Industrial Internet of Things (IIoT) environments. These systems can be broadly categorized into two main types: Network-based IDS (NIDS) and Host-based IDS (HIDS) [22]. 1) Network-based IDS (NIDS): This type of IDS continuously monitors network traffic for suspicious activities. It analyzes traffic patterns to identify potential intrusions and provides real-time alerts regarding anomalous behaviors within the network [23]. 2) Host-based IDS (HIDS): In contrast, HIDS focuses on individual devices or hosts, monitoring them for signs of compromise. This system analyzes various host-based activities, including system logs; file integrity, and user behavior, to detect potential security breaches [22]. It is worth mentioning, the proposed system is a Network-based IDS, designed to enhance the security of IIoT environments by leveraging real-time traffic analysis and anomaly detection.

IDS employs a variety of detection methods to identify potential threats:

- **Signature-based Detection**: This method relies on predefined patterns or signatures of known threats. While it is effective in identifying established attacks, it is unable to detect new or unknown threats [24].
- **Anomaly-based Detection**: This approach establishes a baseline of normal behavior for the monitored environment and identifies deviations from this baseline. While it can effectively uncover unknown threats, it may also result in higher false positive rates [25].
- **Hybrid Detection**: Combining both signature-based and anomaly-based detection methods, hybrid detection systems offer a balanced approach. They leverage the strengths of each method to detect known threats while also identifying new and emerging threats [24].

It is worth mentioning, the proposed method utilizes a hybrid detection approach using deep learning. This strategy not only improves the robustness of the system but also yields superior results in experimental evaluations.

4. Related work

The field of Intrusion Detection Systems (IDS) for the Industrial Internet of Things (IIoT) has seen significant advancements, especially with the integration of smart grid technologies. These innovations have notably enhanced IDS capabilities, enabling the more effective identification and mitigation of potential threats in real-time environments. Several studies have focused on applying artificial intelligence (AI) methodologies, including machine learning, deep learning, and neural networks, to develop robust and real-time IDS solutions for IIoT environments. Almomani [25] developed a feature selection model using optimization algorithms such as Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), Firefly Algorithm (FFA), and Genetic Algorithm (GA). These techniques improve detection accuracy by optimizing feature selection, although they come with increased computational complexity. The model demonstrated improved feature identification and real-time detection rates when tested on a proprietary dataset. Nazir and Khan [26] introduced a novel feature selection approach based on combinatorial optimization that significantly enhances the performance of network intrusion detection in real-time. The method refines the selection of relevant features for intrusion detection, leading to higher detection accuracy, although it requires substantial computational resources. The approach was validated using the UNSW-NB15 dataset, displaying its applicability for real-time network monitoring. Kasongo [27] proposed IDS for IIoT environments that integrates Genetic Algorithms (GA) with tree-based algorithms, improving both accuracy and adaptability in detecting complex intrusions within dynamic and real-time IIoT environments. The system achieved high detection rates when validated using the UNSW-NB15 dataset, though it did involve considerable computational overhead. Gaber et al. [28] presented a hybrid method combining machine learning and optimization techniques for real-time IIoT intrusion detection. This approach maintained high accuracy and efficiency, enabling real-time detection while minimizing computational costs. Their method, validated using the WUSTL-IIOT-2021 dataset, demonstrated both scalability and real-time effectiveness in IIoT security. Jeffrey et al. [29] explored ensemble-learning techniques for anomaly detection in cyber-physical systems. By combining multiple models, their system enhanced the robustness of IDS in real-time scenarios. Despite higher computational requirements, their approach improved detection accuracy, validated on the Edge-IIoTset 2023 dataset, and proved effective in dynamic environments. Wu et al. [30] developed Pelican, a deep

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

residual network tailored for network intrusion detection in IIoT environments. This model enhances the scalability and accuracy of IDS, making it well suited for large-scale, real-time deployments. While the model significantly improved detection performance, it required substantial computational resources to operate effectively in real-time contexts. These studies underline the increasing importance of real-time IDS in IIoT, with a focus on improving detection accuracy and efficiency while managing the challenges of computational overhead. The integration of AI and deep learning techniques has shown substantial promise in enhancing real-time intrusion detection, making systems more adaptive, accurate.

5. Methodology

A. Implementation and Setups.

Our system is implemented using Python programming language with some libraries, such as Tensorflow, Keras, and open-CV. We ran the network using Keras and Tensorflow on a laptop with an Intel(R) Core(TM) i7-1165G7CPU; generation 11; 64 GB of RAM; Windows 11; a 64-bit operating system with an x64-based processor.

The proposed system employs a multi-faceted approach to bolster cyber threat detection within smart grid environments. By integrating recurrent neural networks (RNNs), and support vector classifier (SVC), this framework offers a robust, interpretable, and effective solution for identifying and mitigating cyber vulnerabilities. The data processing pipeline within the system comprises four key stages: Initially, data undergoes meticulous preprocessing to ensure the dataset is clean, normalized, and optimized for efficient learning. This step enhances the data quality, making it suitable for the subsequent stages of analysis and detection. An RNN-based detection model designed to identify abnormal patterns in smart grid operations processes the preprocessed data. RNNs are particularly suited for extracting complex features from large datasets, enabling the identification of anomalies in network traffic and irregularities in operational commands. The employed RNN architecture comprises four layers: an input layer, two hidden layers, and an output layer, with neurons distributed as follows: 784 in the input layer, 128 and 64 in the hidden layers, and 10 in the output layer. This configuration allows the model to effectively learn from complex data landscapes, enhancing its overall detection capabilities. In the third stage, a SVC algorithm is applied to classify traffic as either normal or suspicious. The SVC is fine-tuned with parameters such as max_depth=25, n_estimators=500, and random_state=45 to achieve high accuracy and reliability in distinguishing between benign and potentially harmful traffic. Finally, another SVC is employed to recognize the type of detected attack. The random forest model uses parameters such as max_depth=25, n_estimators=200, and random_state=45 to accurately classify attacks, providing a more detailed and actionable analysis of the detected threats. This integrated approach significantly enhances the cybersecurity resilience of smart grids, providing a dynamic, multi-layered defense against evolving cyber threats. Figure 4 shows the proposed system framework.

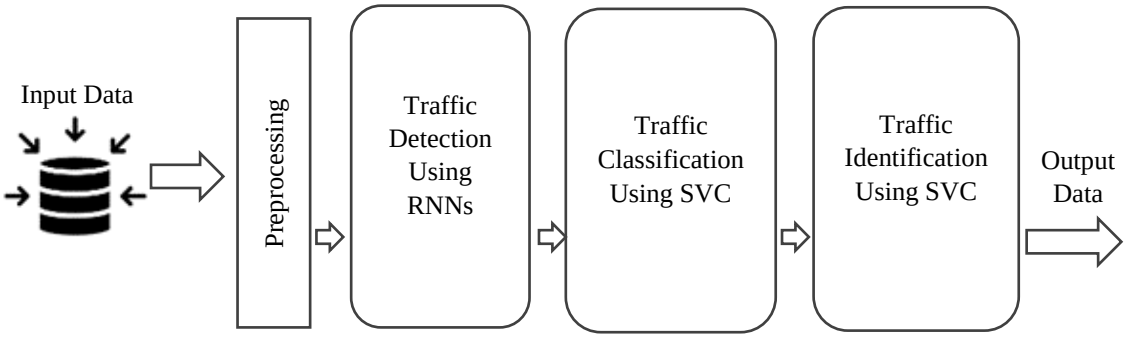

Figure 4. The Proposed system framework.

B. Preprocessing

The preprocessing phase involved several critical steps aimed at preparing the datasets for effective model training. Initially, data cleaning procedures were implemented to eliminate duplicates and manage null values, ensuring the integrity and reliability of the data. This step is essential for mitigating any potential biases that could affect the model's performance. Normalization techniques were then applied to scale features uniformly; facilitating learning that is more efficient during the training process. This was particularly important given the diverse range of attributes present in the datasets.
Figure 5 shows the framework of the data preprocessing in the proposed system.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

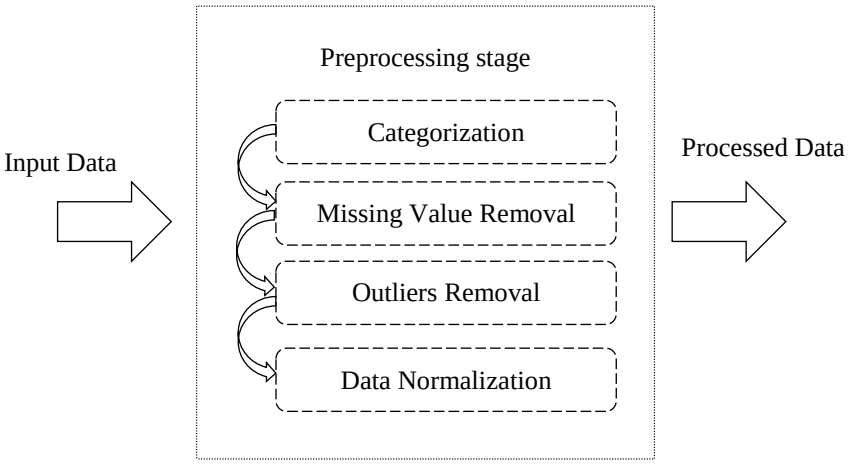

Figure 5. The Preprocessing stage in the proposed system

Feature engineering played a pivotal role in identifying and constructing key attributes that contribute significantly to the model's decision-making process. Attributes such as packet size, transmission timing, and traffic patterns were meticulously examined to enhance the predictive power of the model.

To optimize model efficiency, we employed feature selection techniques, notably Recursive Feature Elimination (RFE) [31]. This method enabled us to systematically reduce the dataset to the most impactful features, thereby streamlining the model and improving its interpretability.

For optimal model performance, hyperparameters including learning rate, batch size, and the number of hidden layers—were meticulously fine-tuned. Both grid search and random search methods were employed to achieve a balance between bias and variance [32]. This rigorous tuning process was crucial for enhancing the model's accuracy and overall effectiveness in real-time detection scenarios.

C. Evaluation Metrics

To provide a comprehensive assessment of the model's performance in detecting cyberattacks, several performance metrics were utilized, including accuracy, precision, recall, F1 score, and AUC-ROC. These metrics are critical for understanding the model's effectiveness in identifying threats, minimizing false positives, and reducing the likelihood of false negatives [40]. The use of these metrics allows for a thorough examination of the model's capabilities and its potential for deployment in real-world smart grid environments. Following are the definitions and mathematical equations for each evaluation metric used.

Accuracy measures the overall correctness of the IDS in identifying both normal and malicious activities

$$AC = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}) \quad (1)$$

Precision quantifies the proportion of correctly identified malicious activities among all detected activities.

$$PR = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \quad (2)$$

Recall (Sensitivity) measures the proportion of actual malicious activities correctly identified by the IDS.

$$RC = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \quad (3)$$

F1 Score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance

$$F1S = 2\ (\mathrm{RC}.\,\mathrm{PR})/(\mathrm{RC} + \mathrm{PR}) \quad (4)$$

False Positive Rate (FPR) indicates the rate of normal activities incorrectly identified as malicious, representing false alarms.

$$FPR = \mathrm{FP}/(\mathrm{FP} + \mathrm{FN}) \quad (5)$$

False Negative Rate (FNR) quantifies the rate of malicious activities incorrectly identified as normal, indicating missed detections.

$$FNR = \mathrm{FN}/(\mathrm{FP} + \mathrm{FN}) \quad (6)$$

Area Under the Curve (AUC) assesses the IDS's ability to distinguish between classes (malicious and normal) by plotting the ROC curve.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

$$AUC = 1 + (\text{TPR} - \text{FPR})/2 \quad (7)$$

Moreover, the confusion matrix provides a detailed breakdown of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) classifications, offering deeper insights into the IDS's performance.

6. Datasets

The data collection phase was meticulously designed to compile a comprehensive set of datasets that accurately represent both real world and simulated smart grid environments. This approach facilitated the exploration of a wide range of cyber-attack scenarios, which is essential for building a resilient AI-driven detection model. For this research, we specifically selected two well-established benchmark datasets: UNSW_NB15 [37] and BoT-IoT [33]. It is worth noting that 70% of each dataset was used for training the proposed system, while the remaining 30% was reserved for testing. The inclusion of these datasets is crucial, as they provide a solid foundation for the model's ability to detect both known and novel cyber threats. By leveraging the extensive samples of normal and abnormal smart grid traffic from these datasets, the AI model can learn to distinguish subtle differences in traffic patterns, thereby improving its accuracy and reliability in real-time detection. The next subsections give a brief overview of these datasets.

A. UNSW_NB15 Dataset

The raw network packets of the UNSW-NB15 dataset [37] were generated using the IXIA Perfect Storm tool in the Cyber Range Lab at UNSW Canberra to simulate a combination of real modern normal activities and synthetic contemporary attack behaviors. This dataset includes nine types of attacks: Fuzzers, Backdoors, Exploits, Analysis, Generic, DoS, Reconnaissance, Shellcode, and Worms. The Argus and Bro-IDS tools were employed, and 12 algorithms were developed to extract 49 features, each labeled with its class. These features are detailed in the UNSW-NB15_features.csv file. The dataset consists of 2,540,044 records, stored across four CSV files. It was divided into training and testing sets, with 175,341 records in the training set and 82,332 records in the testing set, representing various attack and normal activity types.

B. BoT-IoT Dataset

The BoT-IoT dataset [33] was created by designing a realistic network environment in also the Cyber Range Lab at UNSW Canberra, combining both normal and botnet traffic. The dataset's source files are available in various formats. The dataset consists of over 72,000,000 records includes attacks such as DDoS, DoS, OS and Service Scans, Keylogging, and Data Exfiltration, with DDoS and DoS attacks further categorized by the protocol used.

7. Comparisons, Analysis and Results Descriptions

The proposed system is compared with several AI algorithms by altering the algorithms used in the two stages of the system: the classification stage and the identification stage. Following are the result description on every dataset

A. UNSW_NB15 Results

During the experiment using the UNSW_NB15 dataset, after the preprocessing and the traffic detection, which used RNNs, the detected traffic was classified as normal or abnormal using random forest, KNN, SVC, logistic regression, gradient boosting, and a voting classifier. As shown in Figure 6, SVC, which used in the proposed system provided the best results across all evaluation metrics and achieved the smallest processing time, which was 0.49 seconds.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

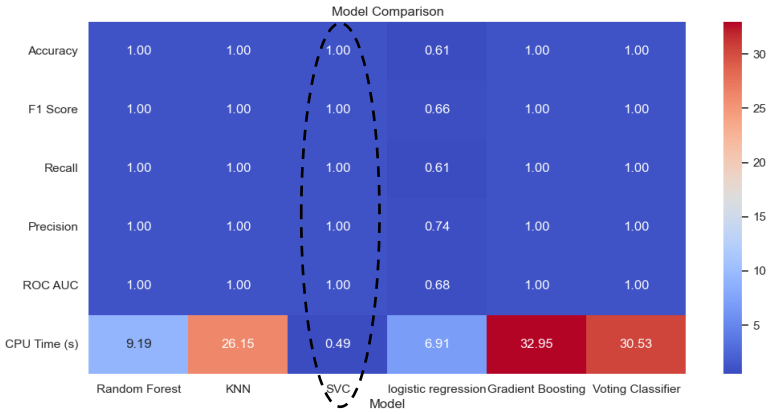

Figure 6. Comparison between the used algorithm in the proposed system and other algorithms in the classification stage on UNSW_NB15 dataset.

For traffic identification or for identifying the type of attack the classified data passes from the SVC to random forest, KNN, SVC, logistic regression, gradient boosting, and voting classifier the SVC also perform an optimum results in all the evaluation matrices and provided the smallest processing time equal 0.6 second (see figure 7).

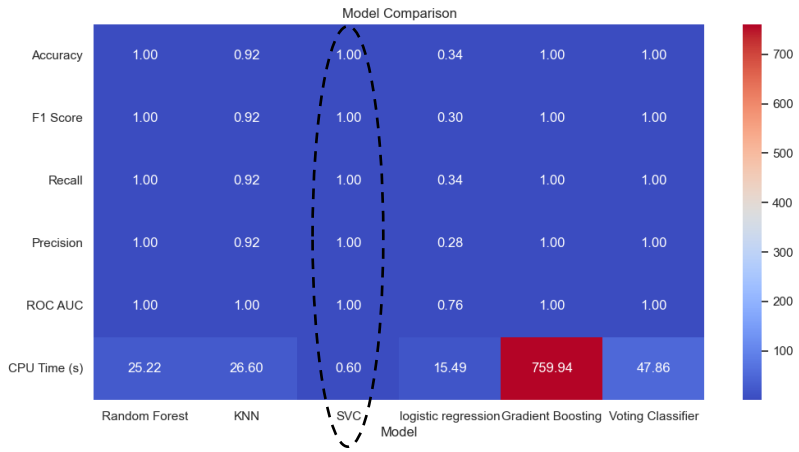

Figure 7. Comparison between the used algorithm in the proposed system and other algorithms in the identification stage on UNSW_NB15 dataset.

Figure 8 illustrates the Receiver Operating Characteristic (ROC) curves for all the methods employed during the identification stage on the UNSW-NB15 dataset.

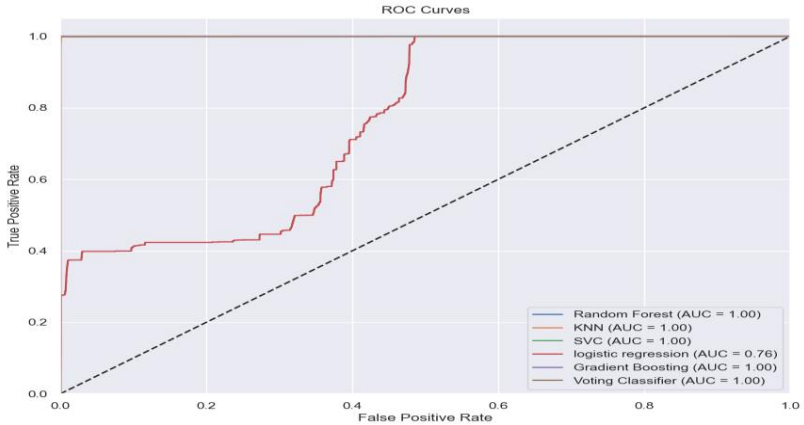

Figure 8. Comparison between the used algorithm in the proposed system and other models in the identification stage through the ROC on UNSW_NB15 dataset.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

B. BoT-IoT Results

During the experiment with the BoT-IoT dataset, detected traffic was classified as normal or abnormal using random forest, KNN, SVC, logistic regression, and gradient boosting. As shown in Figure 9, SVC achieved the best results across all evaluation metrics and the shortest processing time at 10.12 seconds.

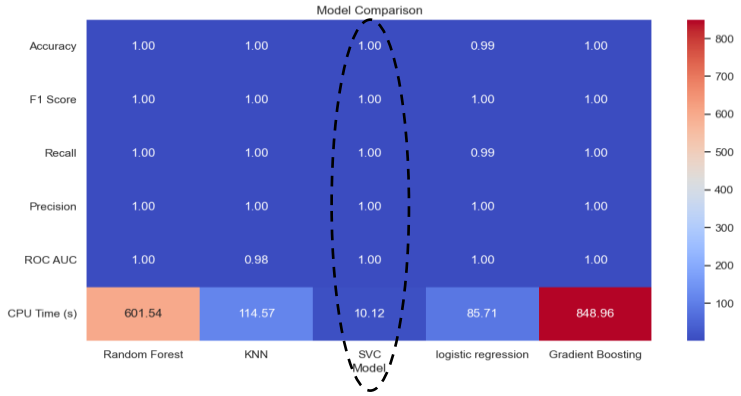

Figure 9. Comparison between the used algorithm in the proposed system and other models in the classification stage on BoT-IoT dataset

For attack type identification, the classified data is passed from SVC to random forest, KNN, SVC, logistic regression, gradient boosting, and a voting classifier. As shown in Figure 10, SVC achieved the best results across all evaluation metrics and the shortest processing time at 11.35 seconds.

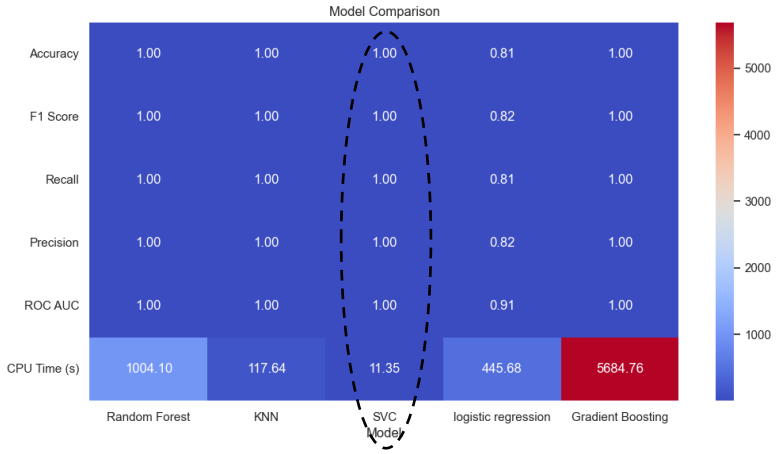

Figure 10. Comparison between the used algorithm in the proposed system and other models in the identification stage on BoT-IoT dataset.

Figure 11 presents the ROC curves for all methods used in the identification stage on the BoT-IoT dataset.

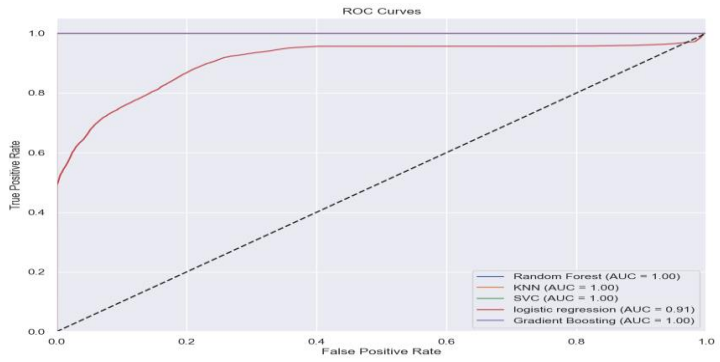

Figure 11. Comparison between the used algorithm in the proposed system and other models in the identification stage through the ROC on UNSW_NB15 dataset.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

6. Conclusion and Discussion

The findings of this study highlight significant advancements for the future of smart grid cybersecurity. The high detection rates and low false positive results demonstrate that AI-based detection methods provide robust security for critical infrastructures like smart grids. Leveraging deep learning and hybrid AI models enables operators to identify known and emerging threats with greater speed and reliability than traditional methods. Integrating real-time data from diverse sources, supported by state-of-the-art machine learning, is essential for these detection systems, as it allows them to adapt to new attack patterns as they arise. This adaptability is particularly crucial in smart grids, where threats continue to evolve and adapt within their respective environments. Furthermore, the scalability of the proposed method allows for seamless implementation across various smart grids, from micro grids to national-scale systems. The success of this AI-based approach underscores the potential of AI in advancing cybersecurity methods for smart grids. As smart grid infrastructure becomes increasingly complex, the need for efficient detection systems capable of processing large data volumes and identifying diverse attack types will continue to grow. This research was designed to address some of the vulnerabilities of smart grids to cyberattacks through a new AI-enhanced detection mechanism. The study reveals that combining deep learning with SVM significantly enhances the performance of proposed models compared to standard detection systems like signature-based and anomaly-based systems. This approach achieved a high accuracy of 100% in identifying various forms of cyber-attacks with minimal false alarms. These results were made possible by technical improvements, such as real-time big data analytics, refined feature selection, and advanced training techniques. Additionally, the versatility of the proposed method, especially in adapting to dynamic threats, makes it a viable solution for safeguarding smart grid assets.

References

[1] Muneeswari, G., Rose, R. M., Balaganesh, S., Prasath, G. J., & Chellam, S., Mitigation of attack detection via multi-stage cyber intelligence technique in smart grid. Measurement: Sensors, 33, 101077, 2024.

[2] Bouramdane, A. A., Cyberattacks in smart grids: Challenges and solving the multi-criteria decision-making for cybersecurity options, including ones that incorporate artificial intelligence, using an analytical hierarchy process. Journal of Cybersecurity and Privacy, 3(4), 662-705, 2023.

[3] Beg, O. A., Khan, A. A., Rehman, W. U., & Hassan, A. A review of AI-based cyber-attack detection and mitigation in microgrids. Energies, 16(22), 7644, 2023.

[4] Mazhar, T., Irfan, H. M., Khan, S., Haq, I., Ullah, I., Iqbal, M., & Hamam, H. Analysis of cyber securi ty attacks and its solutions for the smart grid using machine learning and blockchain methods. Future I nternet, 15(2), 83, 2023.

[5] Alsuwian, T., Butt, A. S., & Amin, A. A., Smart grid cyber security enhancement: Challenges and solutions—A review. Sustainability, 14(21), 14226, 2022.

[6] Koduru, S., Machina, V. S. P., & Madichetty, S., Cyber-attacks in cyber-physical microgrid systems: A comprehensive review. Energies, 16(12), 4573, 2023.

[7] Berman, D. S., et al. A survey of deep learning methods for cyber security. Information, 10(4), 122, 20 19.

[8] Li, J. H. Cyber security meets artificial intelligence: a survey. Frontiers of Information Technology & Electronic Engineering, 19(12), 1462-1474, 2018.

[9] Stein, G., Chen, B., Wu, A. S., & Hua, K. Decision tree classifier for network intrusion detection with GA-based feature selection. In Proceedings of the 43rd annual Southeast regional conference—Volume 2 (pp. 136-141), 2005.

[10] Sanghavi, P., Solanki, R., Parmar, V., & Shah, K. Comprehensive study of cyber security in AI-based smart grid. In International Conference on Advances in Computing and Data Sciences, Cham: S pringer Nature Switzerland, pp. 189-202, Apr. 2023.

[11] Salem, A. H., Azzam, S. M., Emam, O. E., & Abohany, A. A. Advancing cybersecurity: A comprehen sive review of AI-driven detection techniques. Journal of Big Data, 11(1), 105, 2024.

[12] Dangi, A. K., Pant, K., Alanya-Beltran, J., Chakraborty, N., Akram, S. V., & Balakrishna, K. A review of use of artificial intelligence on cyber security and the fifth-generation cyber-attacks and its analysis. In 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), pp. 553-557, 2023.

[13] Nishat, F. Artificial intelligence-enabled anomaly IDS for IoT network: Trends, solutions, and challenges. In Artificial Intelligence for Intelligent Systems, CRC Press, pp. 190-202, 2025.

[14] Achaal, B., Adda, M., Berger, M., Ibrahim, H., & Awde, A. Study of smart grid cyber security, examining architectures, communication networks, cyber-attacks, countermeasure techniques, and challenges. Cybersecurity, 7(1), 10, 2024.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024

[15] Guato Burgos, M. F., Morato, J., & Vizcaino Imacaña, F. P. A review of smart grid anomaly detection approaches pertaining to artificial intelligence. Appl. Sci., 14, 1194, 2024.

[16] Ajala, O. A., Okoye, C. C., Ofodile, O. C., Arinze, C. A., & Daraojimba, O. D. Review of AI and mac hine learning applications to predict and thwart cyber-attacks in real-time. Magna Sci. Adv. Res. Rev., 10(1), 312-320, 2024.

[17] Guato Burgos, M. F., Morato, J., & Vizcaino Imacaña, F. P. A review of smart grid anomaly detection approaches pertaining to artificial intelligence. Appl. Sci., 14(3), 1194, 2024

[18] Saxena, K., Jeyakarthika, K., Dhaaraani, R., Goshwami, S., Raj, K. B., Rani, K. S., & Vyas, V. Enhancing cybersecurity in smart grids through machine learning-based intrusion detection systems. J. Electr. Syst., 20(7s), 2524-2533, 2024.

[19] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. Internet of things: A survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor, 17(4), 2347–2376, 2015.

[20] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. Edge computing: Vision and challenges. IEEE Internet Things J., 3(5), 637–646, 2016.

[21] Xu, L. D., He, W., & Li, S. Internet of things in industries: A survey. IEEE Trans. Ind. Inform., 10(4), 2233–2243, 2018.

[22] Scarfone, K., & Mell, P. Guide to intrusion detection and prevention systems (IDPS). Natl. Inst. Stand. Technol., NIST Spec. Publ. 800-94, 2007.

[23] Buczak, A. L., & Guven, E. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Commun. Surv. Tutor, 18(2), 1153–1176, 2015.

[24] Patcha, A., & Park, J.-M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. Comput. Netw. 51 (12), 3448–3470, 2007.

[25] Almomani, O. A feature selection model for network intrusion detection system based on PSO, GWO, FFA, and GA algorithms. Symmetry, 12(6), 1046, June 2020.

[26] Nazir, A., & Khan, R. A. A novel combinatorial optimization-based feature selection method for network intrusion detection. Comput. Secur. 102, 102164, March 2021.

[27] Kasongo, S. An advanced intrusion detection system for IIoT based on GA and tree-based algorithms. IEEE Access, 9, 113199–212, August 2021.

[28] Gaber, T., Awotunde, J., Folorunso, S., Ajagbe, S., & Eldesouky, E. Industrial internet of things intrusion detection method using machine learning and optimization techniques. Wireless Commun. Mobile Comput, 2023(1), 3939895, 2023.

[29] Jeffrey, N., Tan, Q., & Villar, J. Using ensemble learning for anomaly detection in cyber–physical systems. Electron, 13(7), 1391, April 2024.

[30] Wu, P., Guo, H., & Moustafa, N. PELICAN: A deep residual network for network intrusion detection. In Proc. 50th Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw. Workshops (DSN-W), June 2020, pp. 55–62.

[31] Krichen, M. Strengthening the security of smart contracts through the power of artificial intelligence. Comput, 12(5), 107, 2023.

[32] Shahin, M., Maghanaki, M., Hosseinzadeh, A., & Chen, F. F. Advancing network security in industrial IoT: A deep dive into AI-enabled intrusion detection systems. Adv. Eng. Inform, 62, 102685, 2024.

[33] Peterson, J. M., Leevy, J. L., & Khoshgoftaar, T. M. A review and analysis of the bot-IoT dataset. In 2021 IEEE International Conference on Service-Oriented System Engineering (SOSE) (pp. 20-27), August 2021.

[34] Moustafa, N. A new distributed architecture for evaluating AI-based security systems at the edge: network ToN_IoT datasets. Sustainable Cities and Society, 72, 202 1. Doi: 10.1016/j.scs.2021.102994.

[35] Smys, S., Chen, J. I. Z., & Shakya, S. Survey on neural network architectures with deep learning. J. So ft Comput. Paradigm, 2(3), 186–194, 2020.

[36] Wang, F., & Tax, D. M. Survey on the attention based RNN model and its applications in computer vis ion. ArXiv preprint arXiv: 1601.06823, 2016.

[37] Moustafa, N., & Slay, J. UNSWNB15: A comprehensive data set for network intrusion detection syste ms (UNSW-NB15 network data set). In Proc. IEEE Military Commun. Inf. Syst. Conf. (MilCIS), pp. 1–6, 2015.

DOI: https://doi.org/10.54216/JCIM.150215
Received: May 17, 2024 Revised: July 20, 2024 Accepted: November 04, 2024