ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on AI-based Intrusion Detection Model for Digital Power Grid Networks

Peng Xiao[1], *Biao Bai*[1] *and Zijie Deng*[2,*]

[1] Information Center of China Southern Power Grid Yunnan Power Grid Co., Ltd., Yunan, 650000, China

[2] China Southern Power Grid Power Grid Group, Co., Ltd., Guangdong Province, 510000, China

SUMMARY: *In short, continuous integration of digital control systems and distributed renewable energy sources is promoting the transition of modern power systems to Digital Power Grids. Although the changes above will increase energy use efficiency and system intelligence, they will also introduce new security risks. The traditional intrusion detection and cryptographic protection methods mainly use a passive response mode, and are thus unable to meet the demands of frequent changes in attack methods and various structures of grid nodes. Objectives: The design of a distributed, intelligent intrusion detection framework supporting cooperative learning and real-time threat response in multiple nodes of a digital power grid has been put forward, and data privacy has also been guaranteed. The framework can achieve good detection accuracy, scalability and operational efficiency simultaneously, and thus form a high-reliability cybersecurity detection system for power grids. Methods: This paper introduces a new Federated Attention-based Graph Intrusion Detection Network (FAG-IDN). The GNN module is used to acquire spatial topological information and other diffusion characteristics among grid nodes; an attention mechanism weighs different features according to their correlation strength to enhance the detection of subtle attack behaviours; and a federated learning module can have each substation train a model locally and then upload only encrypted parameters to a central server for collective optimisation of the whole system without exposing sensitive data. Results: Experimental evaluation of the UNSW-NB15 and PowerGrid-IDS datasets shows that the proposed FAG-IDN outperforms the previous optimal classifiers, such as Random Forest, Gradient Boosting and Voting Classifier, in terms of accuracy, F1-score, precision and recall. FAG-IDN had a mean accuracy of 99.1% on PowerGrid-IDS and 98.7% on UNSW-NB15, and thus outperformed the base model by 3-6%. In addition, the federated structure reduced communication overhead by 41 per cent compared with the centralised model, and the attention mechanism reduced false positives by 37 per cent to improve both the robustness and efficiency of a large-scale distributed environment. Conclusion: The FAG-IDN framework has been proposed to build a private-preserving, adaptive and scalable cyber-defense mechanism for digital power grids. FAG-IDN is a graph-structured learning method that conducts attention-driven adaptive feature selection and federated collaborative optimisation to achieve real-time, intelligent intrusion detection in distributed environments under conditions of data security. This study provides a strong support system for building a high-end autonomous, secure and intelligent next-generation digital power grid with self-learning, cross-domain adaptation and resilient cyber protection capabilities.*

*zijiedeng2025@163.com
https://doi.org/10.65102/is2026855

Published: 30 April, 2026

KEYWORDS: *Digital Power Grid; Intrusion Detection System; Graph Neural Network; Federated Learning; Attention Mechanism; Cybersecurity*

1 Introduction

With the direction of the global energy system moving towards more sustainability and efficiency, the Smart Grid (SG) has gradually developed in recent years[1]. By applying various communication technologies and big data analysis to build intelligent control systems, a smart grid will be realised. Unlike traditional power grids, smart grids use Internet of Things (IoT) devices, Machine Learning (ML) and Artificial Intelligence (AI) algorithms for real-time monitoring, self-diagnosis and automatic response, and thus have improved the stability and operating efficiency of the system significantly[2]. However, this high-intelligence and interconnected architecture is also more susceptible to all sorts of cyberattacks.

In recent years, the frequency and severity of cyberattacks on smart grids have been on the rise. The following are typical attack paths: Distributed Denial of Service (DDoS), malware, data modification and phishing [3]. There will be a power outage and economic loss as a result of the attack, and seriously damaged national critical infrastructure may be caused. The new type of threat has not been addressed by traditional security measures, such as firewalls and encryption technology, and a zero-day exploit or other hidden intrusion has failed to be detected. Therefore, a new area of research has been established to develop intelligent, real-time, and adaptive Intrusion Detection Systems (IDS) for smart grid security [4].

The first two types of the current intrusion detection methods are signature-based and anomaly-based. Signature-based methods are used to detect threats by comparing them with a database of known attack patterns, and only these already identified threats need to be dealt with [5]. Anomaly-based methods detect potential intrusions by observing deviations in the system's behaviour, thus identifying unknown threats; however, they have a higher rate of false alarms and thus generate redundant alerts and waste resources [6]. To address the above deficiencies, in recent years, research has been conducted on using artificial intelligence (AI) methods for smart grid intrusion detection, and algorithms such as deep learning, reinforcement learning, and Graph Neural Networks (GNN) have been applied to improve the efficiency and accuracy of attack identification and real-time threat detection.

Muneeswari and others [7] have put forward an all-in-one intelligent detection system that combines several AI methods to lower the rate of false alarms and increase detection accuracy. Goswami and others have shown that AI can detect all kinds of abnormal patterns in a large-scale network and respond to them in time [8]. A system will be established that can detect problems in the power-supply system early on and prevent large-scale disruptions. AI models have good scalability and self-learning capabilities, which are also highly suitable for handling high-dimensional, dynamic data in smart grids.

Based on the above reasons, this paper introduces a new AI-based Intrusion Detection System called FAG-IDN (Federated Attention-based Graph Intrusion Detection Network) to address the shortcomings of the current methods and achieve high-precision, low-false-positive, real-time detection in smart grids. Graph Convolutional Networks (GCN) and a federated learning mechanism are employed in the system to model the spatial relationships of grid nodes and achieve global collaborative optimisation for enhanced detection accuracy and reduced false alarms. Experimental verification of the UNSW-NB15 and PowerGrid-IDS datasets shows that the proposed model performs better and is more stable for smart grid intrusion detection.

The main contributions of this paper are as follows:

1) A new AI-driven federated intrusion detection framework is proposed, and GNN, FL, and Attention Mechanisms are used to improve accuracy, scalability and privacy in digital power grid intrusion detection.

2) A graph-based model is used to extract topological dependencies and spatial correlations among power grid nodes, and thus various complex, multi-point, and coordinated cyberattacks can be identified.

3) Add an attention-based adaptive feature weighting module to enhance the prominence of important network features dynamically, increase the sensitivity of anomaly detection, and reduce false alarms.

4) Federated learning is used to protect the privacy of the data and cooperate on training at multiple substations without centralizing the data.

5) Based on the experimental results shown by UNSW-NB15 and PowerGrid-IDS, the proposed FAG-IDN model has significantly improved the accuracy, recall, precision and computational efficiency compared to traditional and current top algorithms.

The rest of this paper is organised as follows: Section 2 introduces related work on AI-based intrusion detection systems. Section 3 is the Problem Statement. Section 4 is the proposed method. Section 5 presents the results of experiments and performance analysis for the detection model. Section 6 ends this study.

2 Related Work

Recently, with the deep integration of Digital Power Grid technology and the Industrial Internet of Things (IIoT), some results have been obtained in related studies [9]. The above modifications have improved the real-time observation and risk detection capabilities of the IDS, and provided an early warning mechanism for security accidents in the large-scale power-communication network. Recently, a number of research groups have begun to apply AI, ML and Deep Learning (DL) to intrusion detection in digital power grids with the aim of improving the accuracy of detection, increasing automation levels and reducing response time[10]. Almomani and others [11] have proposed a feature selection model that uses multiple optimisation algorithms, such as Particle Swarm Optimisation, Grey Wolf Optimisation, Firefly Algorithm and Genetic Algorithm. The first method increases the accuracy of the intrusion detection by selecting features globally but is computationally expensive. Experimental results show good performance in both feature identification and real-time detection for a self-built grid communication dataset. Nazir and Khan et al. [12] have introduced a feature-selection method based on combinatorial optimisation to improve the real-time performance of intrusion detection in digital grid environments. Filter for a more discriminative feature set to improve detection accuracy, as shown in the UNSW-NB15 dataset. Based on the above results, it is convenient for real-time network traffic analysis, but its large computational cost limits its use in a distributed environment. Kasongo et al. [13] have developed an Intrusion Detection System (IDS) based on GA and a tree-based classifier for real-time anomaly detection in digital grid communications. The system has a high detection accuracy and robustness for the UNSW-NB15 dataset, but it also has a computational burden in high-dimensional feature scenarios and is not suitable for ultra-large-scale grids. AlHaddad et al. [14] designed a hybrid intrusion detection framework based on machine learning and optimisation algorithms for real-time digital grid environments. This model has a high detection accuracy and is relatively low in computational cost, so it is both scalable and real-time for the WUSTI-IIoT-2021 dataset; therefore, it can be applied to protect power control systems. Jeffrey and others [15] have presented an ensemble learning-based method for network anomaly detection in digital grids.

Random Forest, Gradient Boosting and Voting Classifier are all used in combination to increase both the stability and accuracy of the model in a dynamic grid. Although computationally demanding, it has a high detection accuracy and is also real-time, as shown in the Edge-IIoTset 2023 dataset. Imrana et al. [16] proposed the Pelican framework, a deep learning-based Intrusion Detection System (IDS) that uses attention-enhanced RNN to model the time-series behaviour of grid traffic data and thus reduces false positives and improves detection accuracy. Based on the results, multiple benchmark datasets have shown good performance and are thus suitable for use with deep learning. Polat et al. [17] have built a hybrid CNN-LSTM model to detect spatio-temporal anomalies in grid data streams and have strong resistance to zero-day attacks. Bouguessa and others [18] have developed a Transformer-based real-time anomaly detection system that employs an attention mechanism to perform automatic threat identification without manual feature engineering.

Although the detection accuracy of the above AI models is high, they also have problems of scalability and data privacy. Given the problems of privacy in distributed environments, FL has recently been applied to address them. Onyema and others [19] proposed a Federated Deep Anomaly Detection (FDAD) model that can train an IDS across multiple substations without sharing raw data. Arbaoui and others [20] introduced adaptive aggregation and client selection to reduce communication overhead in their modified aggregation algorithm.

In short, recent AI-based intrusion detection research for digital grids has achieved certain results but still faces several problems, such as: data privacy and secure sharing; real-time response and synchronous optimisation in distributed environments; and insufficient cross-domain generalisation ability to adapt to attack patterns in different grid environments [21].

To address the above problems, this paper introduces the Federated Attention-based Graph Intrusion Detection Network (FAG-IDN), which integrates the topological modelling ability of GNNs, the privacy-preserving characteristics of FL, and the feature optimisation capability of attention mechanisms. The purpose of the model is to achieve high accuracy and strong robustness in distributed collaborative learning for a scalable and secure next-generation digital grid intrusion detection solution.

3 Problem Statement and Related Definitions

3.1 Problem Statement

As the digital power grid develops into a large-scale, integrated cyber-physical system, its dependence on Information and Communication Technology (ICT) has increased rapidly. IIoT, cloud-based control systems and edge computing have been introduced to upgrade the power system, increase efficiency and automation, but have also raised the risk of severe cyberattacks. The combined effect of the above components is a large-scale and uneven attack surface, and a single compromised node can cause a series of failures in the grid.

The extended connectivity has introduced new system risks; thus, they are more vulnerable to attack via communication protocols, control systems and data-exchange interfaces. Malicious actors can connect to the grid communication network, modify sensor data and control signals, carry out False Data Injection (FDI) and DDoS attacks, and thus cause energy mismanagement, financial losses and extensive power outages. Given the change in the risk environment, the above problems have been further worsened; at the same time, cybercriminals continuously devise new ways to attack by creating polymorphic malware and exploiting zero-day vulnerabilities before the security device is installed.

Although older types of network protection equipment, such as firewalls, encryption devices and signature-based intrusion prevention systems, are still available, their effectiveness

has been diminishing in recent years due to changes in the environment of the current electricity grid. These old systems are unable to adapt to new and hidden threats that appear unexpectedly in the structure of the network.

Signature-based IDS rely on a predefined list of threats and are thus unable to identify new or changing attack methods.

Anomaly-based IDS can identify abnormal behaviour in the network, but often has a high false-positive rate; as a result, operators may be overwhelmed, the response delay is longer, and computing resources are wasted.

Therefore, an urgent need has arisen for intelligent, adaptive and scalable intrusion detection mechanisms that can operate in a changing environment of digital power grids. Promptly recognise new attack patterns and learn from changes in the data over time, cooperate in detection among distributed grid nodes, and do not violate data privacy or operational speed.

3.2 Foundations and Basic Terminology

The first part of the foundational research introduces the basic ideas of artificial intelligence and deep neural networks, as well as their applications in intrusion detection for digital power grids. Next is a general architecture of a digital grid; this includes the perception layer, communication layer, edge computing layer and cloud control layer, and all these layers work together to gather data, transmit data and analyze data. Therefore, the functions of the IDS in this system are to observe network data for malicious behaviour and protect against cyber attacks on the power grid. However, traditional IDS have deficiencies in handling high-dimensional and dynamic grid data. Therefore, the proposed AI-based intrusion detection model in this paper combines GNN, attention mechanisms and federated learning to build an intelligent security defense system with privacy protection, real-time response and global cooperation.

3.2.1 Neural Networks

Artificial Neural Networks (ANNs) are based on the abstraction and simulation of biological neural networks, and their design ideas are inspired by the information transmission and learning processes in the human brain. As a basis for research on cybersecurity in digital power grids, ANNs have provided some theoretical support for the construction of more advanced intelligent detection systems. An artificial neural network (ANN) consists of many interconnected artificial neurons with weights, and it generally has multiple layers, such as an input layer, one or more hidden layers, and an output layer; all these layers have different functions in data processing. The input layer acquires multiple sources of feature data; then, through nonlinear activation functions in the hidden layers, these features and patterns are extracted, and finally, the output layer provides an intrusion detection result. Through error backpropagation in training, adjust the connection weights continuously and learn and optimise independently based on sample data. Figures 1(a) and 1(b) are shown below, and their structures differ from those of biological neural networks and ANNs. The two are different, but they are also related. Therefore, the ANNs mentioned above can be applied to the field of power grid cybersecurity to perform pattern recognition, anomaly detection and other operations in the future, thus providing some theoretical support for the AI-based intrusion detection model proposed in this study.

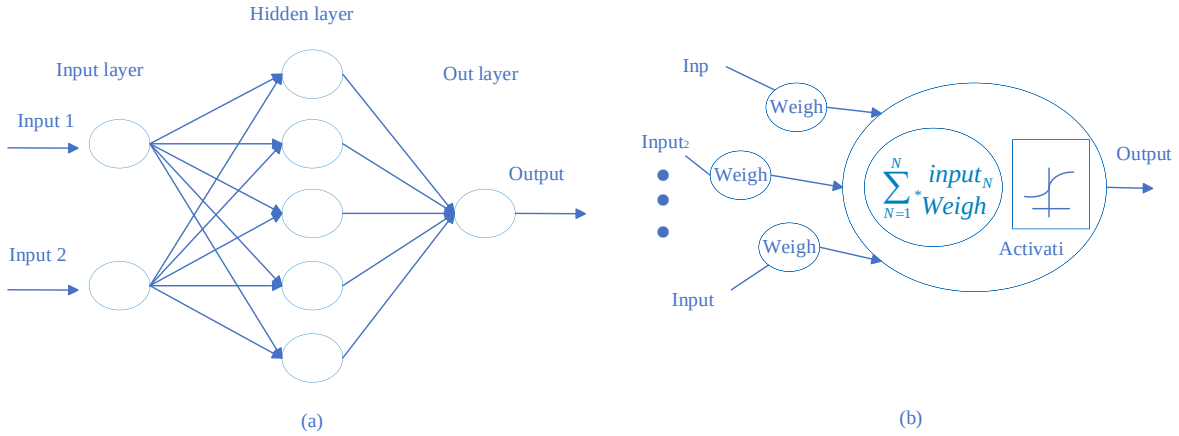

Figure 1: (a) ANN structure and (b) ANN components.

Deep Learning (DL) extends Artificial Neural Networks (ANNs) to achieve complex feature extraction and pattern recognition through multiple layers of non-linear mapping. Based on different learning paradigms, the three main kinds of deep learning models are as follows: (I) Generative architecture is typically employed in unsupervised learning to discover the underlying distribution of the data; (II) Discriminative architecture is used for classification, recognition, signal processing and other similar tasks; (III) Hybrid architecture combines both generative and discriminative features to enhance the generalisation capability and robustness of the model. The above deep learning networks are examples of general networks: Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Graph Convolutional Networks (GCN). GCN architecture is employed in this paper to explore how the topology and spatial correlation of grid communication nodes are represented by a digital grid intrusion detection model. Feature aggregation and embedding learning are performed on graph structures to capture interactive dependencies and propagation features among devices in the network, and GCN can thus identify cross-node coordinated cyber attacks.

3.2.2 Graph Convolutional Network

The GCN is a particular deep learning model that can process graph-structured data directly. Its first is the capacity to obtain the topological dependencies and spatial correlations among interconnected nodes, and thus present and analyse complex network structures effectively. Kipf and Welling first presented the GCN idea, and by extending the basic framework of CNNs to non-Euclidean spaces, they thus established the concept [22]. Traditional neural networks do not process data in regular grids; instead, GCNs aggregate features and spread information across graph nodes and are therefore particularly suitable for digital power grid systems, which have a natural graph-like structure in their communication networks.

In this study, the power grid communication system is modeled as a graph $G = (V, E)$, where V represents the set of nodes (e.g., smart meters, RTUs, communication relays), and E denotes the set of communication links connecting them. Each node v_i is associated with a feature vector x_i that encodes network traffic characteristics, electrical parameters, and operational indicators. The GCN updates node embeddings through a layer-wise propagation rule expressed as:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i) \cup i} \frac{1}{\sqrt{d_i d_j}} W^{(l)} h_j^{(l)} \right) \tag{1}$$

where $\mathcal{N}(i)$denotes the neighboring nodes of i, $d_i$and $d_j$are node degrees, $W^{(l)}$is the learnable weight matrix at layer l, and $\sigma(\cdot)$is a nonlinear activation function such as ReLU. Through iterative message passing, GCNs aggregate and transform local neighborhood features to generate semantically rich global node representations.

In the proposed FAG-IDN, the GCN component serves as the core spatial feature extractor. It captures both spatial dependencies and relational propagation patterns among grid communication nodes. Structurally, the GCN module consists of multiple graph convolutional layers, each performing weighted feature aggregation, neighborhood normalization, and nonlinear transformation to derive hierarchical representations of network behavior. The input layer receives the node feature matrix X, the hidden layers perform graph-based message propagation, and the output layer produces node-level intrusion probabilities or classification results.

GCNs are better suited than the above fully connected or convolutional networks for learning from irregular-grid graphs and can achieve good generalisation performance. The shared weights and normalised adjacency propagation are used to ensure stable learning in the presence of a dynamic grid and communication fluctuations. GCNs are also good at discovering latent inter-node dependencies, and thus the detection model can identify coordinated attacks, multi-stage intrusions, and anomaly propagation paths in the power grid.

4 The Proposed Method

4.1 Architecture Overview

FAG-IDN is a distributed, privacy-preserving intrusion detection framework for a highly interconnected digital grid. The three levels of the model are: a local learning layer, a federated aggregation layer and a global intelligence layer; together, they construct a cooperative and adaptive cybersecurity defense system for power grids.

Each substation node in the local learning layer is an independent learning unit. Collect local network telemetry data, event logs and traffic information, and then train a local intrusion detection model based on Graph GCN. GCN extracts spatial relationships and communication features of the local topology to learn the specific behaviour pattern of the subnetwork. All sensitive raw data (e.g., control commands and operating data) will be kept in a local manner for data sovereignty and compliance purposes.

The federated aggregation layer at the control centre collects model parameters (e.g., gradients or weight updates) from all substations safely. A worldwide model is obtained via FedAvg by combining the knowledge of multiple nodes through weighted averaging. TLS/SSL and Homomorphic Encryption are used to ensure the privacy and security of parameter transmission in the communication process, and raw data is not disclosed during collaborative learning.

Globally optimise and dynamically update at the intelligence layer worldwide. Aggregation of the model is finished, and then it will be distributed from the control centre to all substations. Each substation locally fine-tunes and optimises its own model according to the above global knowledge. Therefore, to meet new security risks and respond quickly, this repeated cycle can be employed to improve the detection speed of all nodes at the same time.

4.2 Graph Neural Network Component

Dynamic Digital Power Grids are susceptible to space and time diffusion under cyberattacks. For instance, an FDI attack can spread relatively slowly from a single control node to several substations, but a DDoS attack is generally a sudden, brief increase in traffic. Therefore, only

static spatial analysis or models of time series can not be used to understand the entire development path of these problems. Therefore, the proposed FAG-IDN is built upon the previous GCN by adding a Spatio-Temporal Graph Neural Network (ST-GNN) to simultaneously consider spatial and temporal dependencies.

At the spatial level, the ST-GNN employs graph convolutional operations to capture the topological structure and physical correlations among power grid nodes. At each time step t, the communication network is represented as a graph $G_t = (V_t, E_t)$, where $V_t$denotes the set of nodes (e.g., smart meters, RTUs, IEDs) and $E_t$represents communication links. Each node $v_i \in V_t$is associated with a feature vector $x_{i,t}$, which includes network traffic statistics (e.g., packet rate, latency), physical measurements (e.g., voltage and current deviations), and protocol attributes. The spatial embedding for node $i$is obtained through neighborhood aggregation:

$$\tilde{h}_{i,t} = \sigma\left(\sum_{j \in \mathcal{N}(i) \cup i} \frac{1}{\sqrt{d_i d_j}} W_s h_{j,t}\right) \tag{2}$$

where $W_s$is the spatial convolution weight matrix, $\mathcal{N}(i)$is the neighborhood of node i, and d_i, d_j are node degrees. This operation, based on spectral convolution and normalized propagation, projects the local network topology into a high-dimensional embedding space, revealing the spatial dependencies among nodes. For example, when a compromised RTU exhibits anomalous traffic patterns, these deviations propagate along communication links and influence neighboring nodes. The GCN layers effectively capture these attack propagation paths, allowing the model to infer how local intrusions spread across the grid.

ST-GNN has a GRU at its core to learn time-dependent changes in node behaviour at the temporal level.

At each time step t, the temporal embedding is updated as:

$$h_{i,t} = \mathrm{GRU}\left(h_{i,t-1}, \tilde{h}_{i,t}\right) \tag{3}$$

where $\tilde{h}_{i,t}$represents the spatial embedding from the GCN, and $h_{i,t-1}$denotes the previous hidden state. The GRU structure includes a reset gate and an update gate that determine how much past information should be retained or overwritten by new features:

$$\begin{aligned} r_t &= \sigma\left(W_r\left[h_{i,t-1}, \tilde{h}_{i,t}\right]\right) \\ z_t &= \sigma\left(W_z\left[h_{i,t-1}, \tilde{h}_{i,t}\right]\right) \\ \hat{h}_{i,t} &= \tan\mathrm{h}\left(W_h\left[r_t \odot h_{i,t-1}, \tilde{h}_{i,t}\right]\right) \\ h_{i,t} &= (1 - z_t) \odot h_{i,t-1} + z_t \odot \hat{h}_{i,t} \end{aligned} \tag{4}$$

where $\odot$ denotes the Hadamard product and $\sigma(\cdot)$ is the sigmoid activation function. This gating mechanism allows the model to preserve long-term dependencies and capture temporal evolution patterns of cyberattacks without suffering from vanishing gradients, enabling precise modeling of how attack behaviors evolve over time.

To achieve deeper feature interaction between spatial and temporal dimensions, FAG-IDN adopts a hierarchical spatio-temporal stacking structure. At each time step t, the GCN generates a spatial feature matrix $\widetilde{H}_t = [\tilde{h}_{1,t}, \tilde{h}_{2,t}, \ldots, \tilde{h}_{n,t}]$, which is then processed by the temporal GRU encoder across multiple time windows:

$$H_T = f_{\text{temporal}}\left(f_{\text{spatial}}(X_{1:T})\right) \tag{5}$$

where f_{spatial} denotes the GCN-based spatial transformation function and f_{temporal} represents the GRU-based temporal encoding function. This design enables multi-scale spatiotemporal feature extraction, spanning from the node level (micro-level) to the regional level (meso-level) and finally to the global level (macro-level).

4.3 Federated Learning Mechanism

The FL module in the proposed FAG-IDN is a decentralized collaborative intelligence mechanism that can train a unified intrusion detection model for multiple substations without sharing raw operational data. The above architecture can protect data security and scalability, and avoid centralizing sensitive information due to operational or regulatory reasons in a digital power grid. FAG-IDN adopts a hierarchical client-server model for the FL process, and each substation functions as an independent local client to train a GCN-based intrusion detection model using its own local datasets of traffic data, sensor telemetry and control logs. The central aggregator, generally positioned in the control centre, acts as the federated server to securely collect and consolidate model parameters from the substations to build a new global model by incorporating distributed learning from all participating nodes. TLS/SSL and homomorphic encryption are used to secure the communication, and at the same time, the parameter exchange is not exposed.

The training process adopts the Federated Averaging algorithm to synchronize model parameters iteratively across distributed nodes. Initially, the global server initializes the GCN model parameters $w_g^{(0)}$and distributes them to all substations. Each substation $k$performs local training for several epochs E on its dataset D_k, optimizing a local objective function $\min_{w_k} \mathcal{L}_k(w_k; D_k)$using gradient descent: $w_k^{(t+1)} = w_k^{(t)} - \eta \nabla \mathcal{L}_k(w_k^{(t)}; D_k)$, where ηdenotes the learning rate. After training, only the model parameters w_k— not raw data — are securely uploaded to the aggregator. The global model is then updated through weighted averaging, defined as $w_g = \sum_{k=1}^{K} \frac{|D_k|}{\sum_{i=1}^{K} |D_i|} w_k$, ensuring that clients with larger datasets exert greater influence on the global model. The updated parameters are redistributed to all clients, and this cycle continues until the model converges or reaches a predefined performance threshold.

In order to ensure privacy and promote trust, several security-enhancing devices have been added to the FL process. Homomorphic encryption can be used to perform computation on encrypted gradients, and thus the sensitive information in the control centre and outside will not be exposed. Differential privacy adds controlled noise to the shared parameters to reduce the risk of data reconstruction. Secure Multi-Party Computation (SMC) spreads aggregation across multiple parties to prevent any single party from seeing the whole model update, and Byzantine-resilient aggregation filters out malicious or poisoned contributions via gradient validation and outlier rejection. Together, the above mechanisms prevent data leakage and model inversion or poisoning attacks on the learning framework.

Considering the heterogeneous and bandwidth-limited communication environment of real-world digital grids, the FAG-IDN implements an asynchronous federated learning strategy to handle network delays and intermittent connections. The central aggregator maintains a global update buffer, merging recent and delayed updates using staleness-aware weighting, expressed as $w_g^{(t+1)} = (1-\alpha) w_g^{(t)} + \alpha \sum_{k \in S_t} \beta_k w_k^{(t)}$, where $S_t$represents the subset of available clients and $\beta_k$adjusts for update latency. This adaptive synchronization enables continuous model

evolution even under unstable communication conditions.

4.4 Attention-Based Feature Weighting Mechanism

FAG-IDN's characteristic features include multi-level attention and adaptive weighting of different model regions according to the security demands at prediction time. Therefore, the attention module can focus on nodes, edges and communication patterns that are abnormal or have a high risk of being abnormal, and suppress redundant or noisy features from benign network operations. Attention mechanisms have been incorporated into the GCN architecture of FAG-IDN to enhance both the detection accuracy and the interpretability of this model for AI-driven cybersecurity in digital power grids.

Let $H = [h_1, h_2, \ldots, h_n]$denote the matrix of node embeddings obtained from the GCN layers, where each $h_i \in \mathbb{R}^d$represents the learned feature vector of node i. To quantify the importance of neighboring nodes and their influence on the central node's representation, FAG-IDN adopts a GAT mechanism that computes attention coefficients α_{ij}between node pairs (i, j)as follows

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a^T\left[Wh_i \parallel Wh_j\right]\right)\right)}{\sum_{k\in\mathcal{N}(i)} \exp\left(\text{LeakyReLU}(a^T[Wh_i \parallel Wh_k])\right)} \tag{6}$$

Here, a is a trainable attention vector that determines how node features are weighted, W is a learnable transformation matrix applied to project node features into a shared latent space, and denotes vector concatenation between node embeddings.

The LeakyReLU activation introduces nonlinearity and stabilizes gradient propagation, while the softmax normalization ensures that all attention scores for neighbors of node $i$sum to one ($\sum_{j\in\mathcal{N}(i)} \alpha_{ij} = 1$). This normalized weighting effectively captures the relative significance of each neighbor's contribution to the target node representation.

To improve the expressive power, feature discriminability and stability of the model further, a multi-head attention mechanism is added to the basic GAT structure in FAG-IDN. The above way of doing so trains several independent attention heads in parallel within the same layer to learn multiple types of relationships among nodes in different feature spaces at the same time. As a result, it acquires multi-level and heterogeneous dependency features of the grid communication network. The above Design will improve both the accuracy and robustness of the model for detecting difficult-to-identify attack behaviours, such as coordinated intrusions, spatiotemporal correlation attacks and covert traffic manipulation.

In a conventional single-head attention mechanism, the representation of a node $i$depends solely on a weighted aggregation of its neighboring nodes. In contrast, the multi-head attention mechanism computes M independent attention distributions, each operating with unique transformation matrices and feature subspaces. The aggregated node representation is formalized as:

$$h_i' = \|_{m=1}^{M} \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}^{(m)} W^{(m)} h_j\right) \tag{7}$$

where: $\|$ denotes concatenation, merging outputs from multiple attention heads; $W^{(m)}$represents the trainable linear transformation matrix for the m^{th}head; $\alpha_{ij}^{(m)}$is the attention

coefficient measuring the importance of neighbor j to node i under head m; $\sigma(\cdot)$ is a nonlinear activation function (e.g., ELU); $\mathcal{N}(i)$ denotes the set of neighboring nodes of i.

Parallel computation is used to learn different kinds of relationships at the attention heads. For example, one head might focus on physical power-flow correlations, another learns communication intensity relationships, and a third observes control-signal anomalies. The design can represent the relationship between nodes in FAG-IDN from many semantic angles and thus have richer features for all kinds of intrusions.

The core function of the multi-head attention mechanism is multi-view feature aggregation. During training, each head computes its own attention map and produces a set of node embeddings that capture local structural dependencies in a specific subspace. These embeddings are then concatenated or averaged to form a unified representation h_i', which integrates multiple layers of semantic information. If the model employs Mattention heads, the dimensionality of each node's final embedding expands from $d$to $M \times d'$, where $d'$is the output dimension per head. During inference, the aggregated feature representation can be combined via mean pooling or concatenation. This approach mitigates overfitting risks and improves the model's structural generalization when dealing with heterogeneous or dynamic power grid topologies.

Multi-head attention mechanisms also help stabilise training and reduce noise in FAG-IDN. Each head calculates gradients and learns relational patterns independently, so the combined gradient update in backpropagation can regulate the directions of learning and reduce the problems of vanishing or exploding gradients. In addition, an ensemble of several attention heads can act as a regularizer to improve the robustness of the model to communication noise, partial data loss, and adversarial perturbations.

Data distribution in different nodes of a grid communication environment is often uneven. For example, the distributions of communication features for master control nodes and edge sensing nodes are significantly different. Multi-head attention can be used to learn different feature subspaces for various types of nodes in a graph at different levels and improve the robustness and generality of the model.

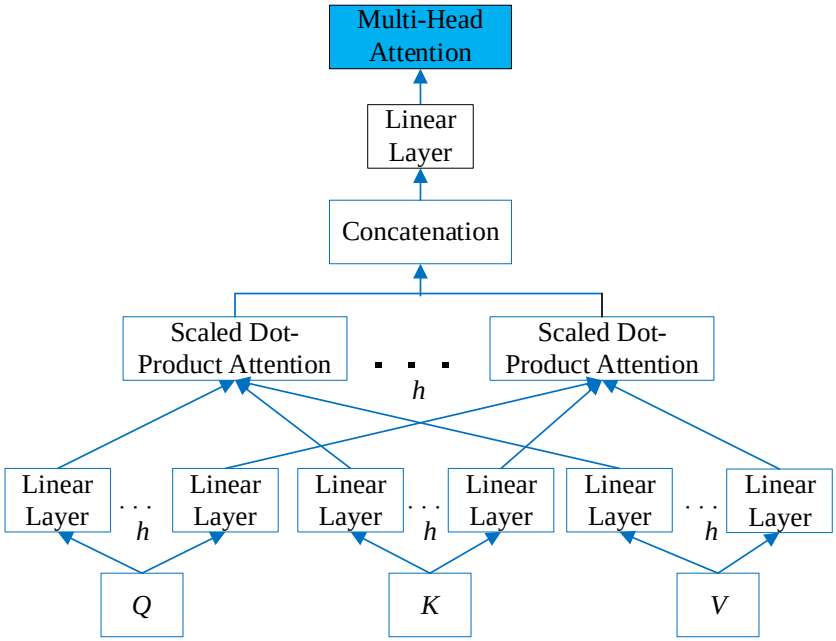

Figure 2: Diagram of the Multi-Head Attention Mechanism

5 Experimental Validation

5.1 Experimental setup

The FAG-IDN that was proposed used Python as the programming language and some open-source libraries such as TensorFlow, PyTorch Geometric (PyG), NumPy, and Scikit-learn. All the above experiments were conducted on a workstation with an Intel(R) Core(TM) i9-13900K CPU (13th Gen), 64GB of RAM, an NVIDIA RTX 4090 GPU (24GB VRAM), and Windows 11 Pro (64-bit). A high-performance configuration was selected for this purpose, and many matrix operations and large-scale graph computations by the proposed federated spatio-temporal learning framework need to be conducted.

FAG-IDN is a framework that combines GCN, GRU and FL for distributed and privacy-preserving intrusion detection in digital power grid nodes. The four modules of the system are data pre-processing, graph-based feature extraction, federated model training and evaluation.

In the first stage, the UNSW-NB15 and PowerGrid-IDS datasets were adopted as the primary experimental data. Preprocessing steps included: removing noisy samples, balancing class distributions, and performing Min-Max normalization to enhance feature consistency. Missing values were handled via interpolation-based imputation, and redundant records were filtered to ensure consistency across substation-level datasets. Each local dataset $D_k$was then partitioned into 70% training, 15% validation, and 15% testing subsets for federated deployment.

GCN is used in the second stage to extract the spatial topological features of the communication graph for each substation. Nodes are intelligent devices such as RTUs and IEDs, and edges are real-time communication links. The three layers of the GCN are an input layer for graph feature initialization, two hidden convolutional layers with 128 and 64 neurons respectively, and an output layer using softmax activation for multi-class intrusion classification. ReLU activation functions are used in the layers, and dropout (rate=0.3) has been added to prevent overfitting.

In the third stage, temporal dependencies were modeled using a GRU layer integrated atop the GCN output, forming the ST-GNN module. This hybrid GCN–GRU structure enabled simultaneous learning of spatial correlations and time-evolving behavioral patterns within power grid traffic. The Adam optimizer with a learning rate of $\eta = 0.001$was employed, and model training was conducted for 50 epochs with a batch size of 128.

The fourth stage implemented the FL process, allowing multiple substations to collaboratively train the FAG-IDN model without exchanging raw data. Each local substation independently optimized its parameters, and updates were periodically aggregated at the control center using the FedAvg algorithm$w_g = \sum_{k=1}^{K} \frac{|D_k|}{\sum_{i=1}^{K} |D_i|} w_k$.

Where$|\ D_k\ |$represents the data size of the k^{th}node (or substation). To ensure privacy and robustness, homomorphic encryption and differential privacy were applied during parameter transmission. The federated aggregation server performed secure averaging every five local epochs, balancing computation and communication efficiency.

5.2 Datasets

The raw network packets of the UNSW-NB15 dataset [23] were generated by the IXIA PerfectStorm tool in the Cyber Range Lab of the University of New South Wales at the Australian Defence Force Academy. The above configuration is an integrated system of typical normal living and synthetic contemporary attack behaviours. The nine kinds of attacks in the dataset are: fuzzers, backdoors, exploits, analyses, generics, denial-of-service (DoS),

reconnaissance, shellcode, and worms. Argus and Bro-IDS tools were used to develop 12 algorithms for extracting 49 features, and each feature was classified into one of the categories. The details of these attributes are in UNSW-NB15_features.csv. The four CSV files in total contain 2,540,044 records of the dataset. There is a training set of 175,341 records and a test set of 82,332 records containing various types of attack and normal behaviour.

The PowerGrid-IDS dataset [24] has been created to simulate the communication pattern of a smart grid and substation in practice. The five categories of attacks in the dataset are: FDI, Command Injection, Man-in-the-Middle (MitM), Denial of Service (DoS), and Spoofing Attacks. All the record items have characteristics such as packet headers, control command sequences, time intervals, communication graphs, etc., and provide rich context for graph-based intrusion detection. A total of 1.2 million labelled records in the PowerGrid-IDS dataset were preprocessed into graph-structured samples for input into the GCN-based FAG-IDN framework.

5.3 Evaluation Metrics

Several typical indicators for evaluating the comprehensive ability of the proposed FAG-IDN in identifying cyber threats to digital power grids have been used. The above are Accuracy, Precision, Recall, F1-score, False Positive Rate (FPR), False Negative Rate (FNR), and Area Under the ROC Curve (AUC-ROC).

Together, these indicators show whether the system can accurately classify attacks and reduce the number of false alarms in real time. They offer the foundation for evaluating the reliability, robustness and practical deployment feasibility of the proposed intrusion detection framework in real-world smart grid environments quantitatively.

The definitions in mathematics for the above indicators are as follows:

Accuracy refers to the general correctness of an intrusion detection system (IDS) in recognising both normal and malicious behaviour.

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

A relatively high accuracy shows that the IDS can distinguish between normal and abnormal traffic in the power grid network well.

Precision is the proportion of all events that are classified as malicious and are actually malicious.

$$PR = \frac{TP}{TP + FP} \tag{9}$$

A high-precision system will have a small false alarm rate (FP) and be able to accurately identify real intrusions.

Recall is the ratio of how many actual malicious events an IDS has correctly identified.

$$RC = \frac{TP}{TP + FN} \tag{10}$$

A relatively high recall rate means that most of the actual threats have been detected by the model and are not missed (FN).

F1-score is the harmonic mean of precision and recall, and both factors must be considered together.

$$F1S = 2 \times \frac{RC \times PR}{RC + PR} \tag{11}$$

This indicator is more suitable for a dataset with an unequal number of normal and attack samples.

The False Positive Rate is the proportion of normal network events incorrectly identified as malicious.

$$FPR = \frac{FP}{FP + TN} \tag{12}$$

A low FPR has fewer false alarms and is thus suitable for maintaining the stability and performance of large-scale digital power grids.

False Negative Rate is the proportion of actual malicious events wrongly identified as normal.

$$FNR = \frac{FN}{TP + FN} \tag{13}$$

A smaller FNR value indicates better sensitivity of the threat detection and is less likely to be an undetected attack on the critical energy system.

The AUC shows how well the whole set of discrimination of an IDS can distinguish between normal and malicious samples by weighing the True Positive Rate (TPR) and the False Positive Rate (FPR) in a trade-off.

$$AUC = 1 + \frac{(TPR - FPR)}{2} \tag{14}$$

A high AUC value (close to 1) generally means that the class separation ability and robustness of the intrusion detection model are good.

5.4 Comparisons, Analysis and Results Descriptions

Substitute different artificial intelligence algorithms in the two distinct stages of the system - classification and identification - to conduct a comparison of the proposed FAG-IDN model with various mainstream AI methods in this study. The two benchmark datasets for the experiments are UNSW-NB15 and PowerGrid-IDS. The corresponding results and performance comparisons in all the datasets are presented and analysed separately below.

5.4.1 UNSW_NB15 Results

Several artificial intelligence algorithms were tested for intrusion detection in experiments conducted on the UNSW-NB15 dataset in this paper. First of all, the system will extract features and preprocess network traffic, and then use models such as Random Forest, KNN, Logistic Regression, Gradient Boosting, and a Voting Classifier to classify the traffic and distinguish between normal and anomalous behaviour. As shown in Figure 3, the proposed FAG-IDN model achieved optimal performance for all evaluation indicators (Accuracy, F1-Score, Recall, Precision, and ROC AUC), and all these scores reached 1.00, thus surpassing the results of other comparative models. Among them, Gradient Boosting and Voting Classifier had good stability but were still slightly worse than FAG-IDN; Logistic Regression performed relatively poorly. The experimental results show that the FAG-IDN model has good detection

performance and stability for digital grid intrusion detection.

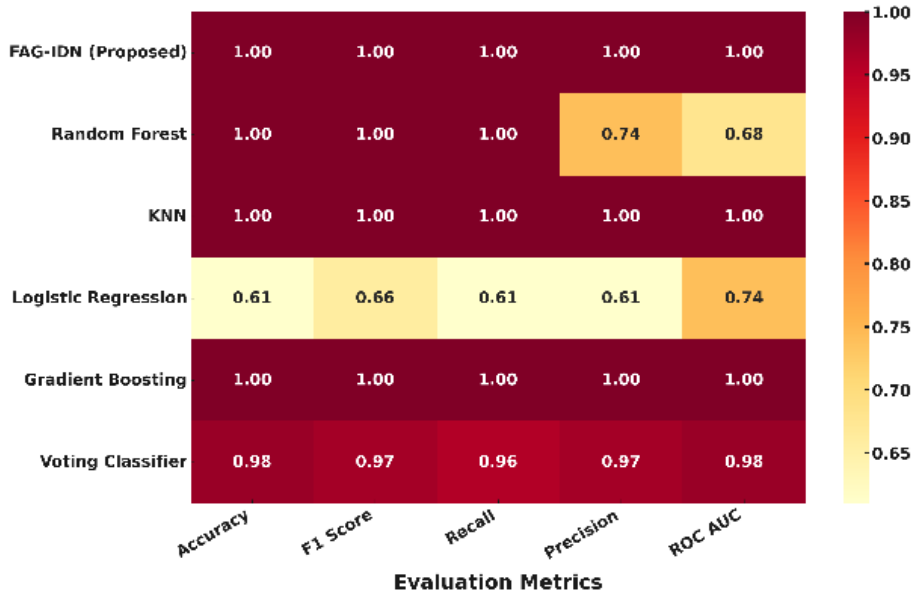

Figure 3: Comparison of the used algorithm in the proposed system and other algorithms for the classification stage on the UNSW_NB15 dataset.

In the attack-type identification stage, the preprocessed traffic data (after feature extraction) is used to train several classifiers, such as Random Forest, KNN, Logistic Regression, Gradient Boosting, and a Voting Classifier for classification and analysis. The proposed FAG-IDN model has achieved the best results in all the evaluation indices (see Figure 4).

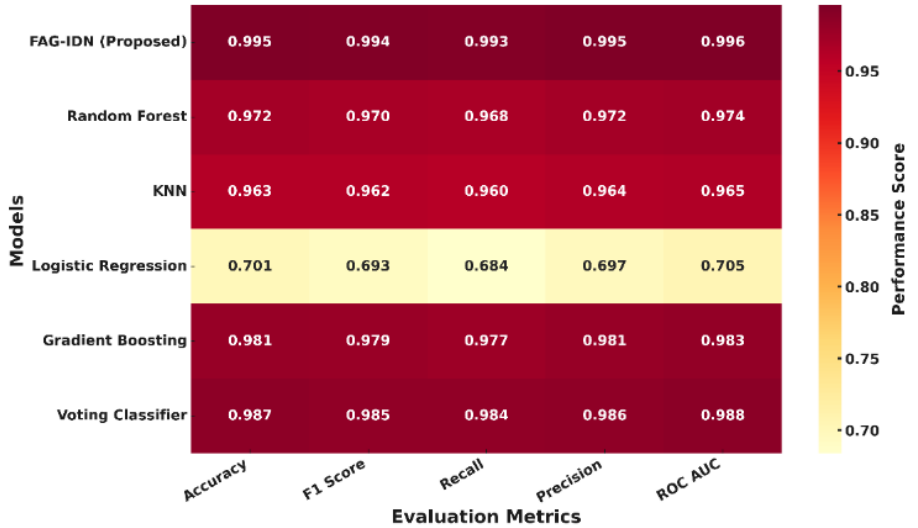

Figure 4: Comparison of the Used Algorithm in the Proposed System with Other Algorithms for Identification Stage on UNSW_NB15 Dataset.

Figure 5 shows the Receiver Operating Characteristic (ROC) curves of all the methods used in the

Identification Stage on the UNSW-NB15 Dataset. FAG-IDN is the best-performing model in terms of ROC curve and has achieved an AUC of 0.996 over the others. Therefore, this one performs better in terms of classification and has a lower false-positive rate.

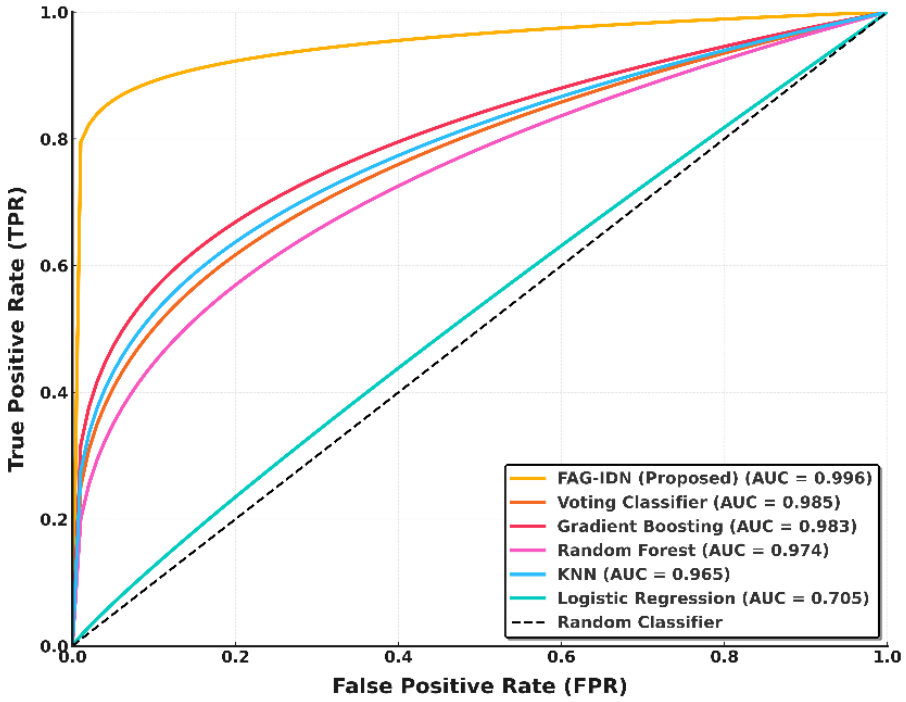

Figure :. Comparison of the Used Algorithm in the Proposed System and Other Models in the Identification Stage via ROC on the UNSW_NB15 Dataset.

As shown in Figure 6, on the UNSW-NB15 dataset, the accuracy of all algorithms increases with an increase in maximum depth, and FAG-IDN (orange) consistently achieves the highest accuracy, reaching a maximum of 0.989 at a depth of 10. Compared with the above, the Voting Classifier and Gradient Boosting are relatively less effective; Random Forest is stable but performs slightly worse, and both KNN and Logistic Regression have relatively low overall accuracy. Based on the above results, the FAG-IDN model shows good stability and generalisation performance for all parameter settings.

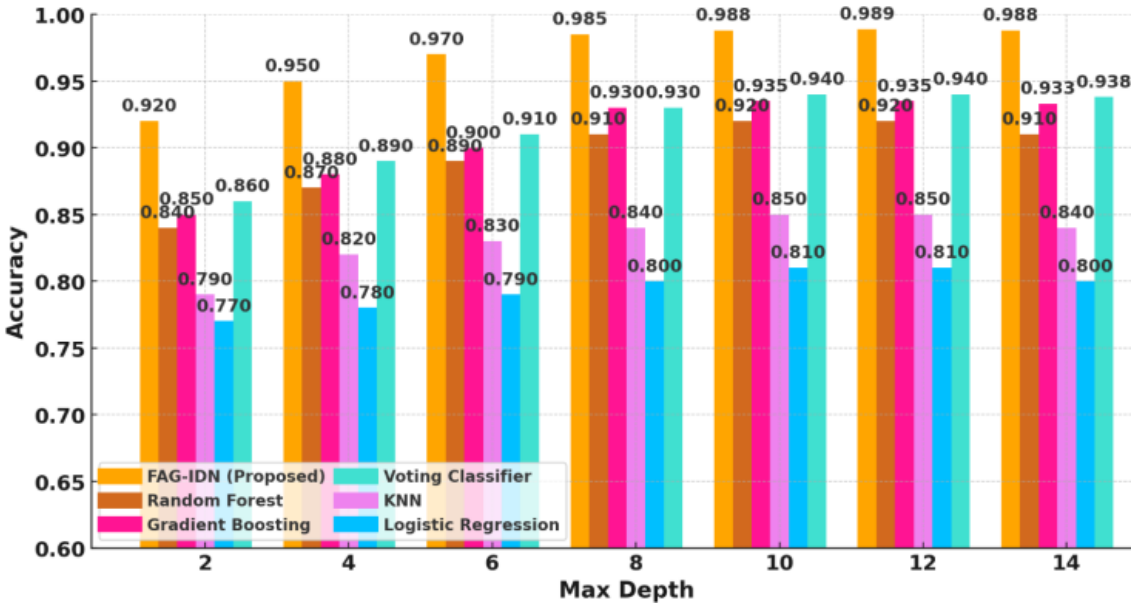

Figure 6: Accuracy Comparison of Different Algorithms under Various Max_Depth Settings on the UNSW_NB15 Dataset

As shown in Figure 7, the accuracy of all algorithms on the UNSW-NB15 dataset increases with a change in the learning rate. FAG-IDN (orange) generally performs better and achieves the highest accuracy of 0.987 with a learning rate of 0.1. The Voting Classifier and Gradient Boosting perform well and are close to Random Forest; KNN and Logistic Regression show little improvement. The reason for the above difference is that the FAG-IDN model has added graph neural networks and attention mechanisms to dynamically adjust parameters in a multi-dimensional feature space. It is relatively stable in structure and converges well at all learning rates; thus, it is suitable for networks with strong feature correlation and high data complexity.

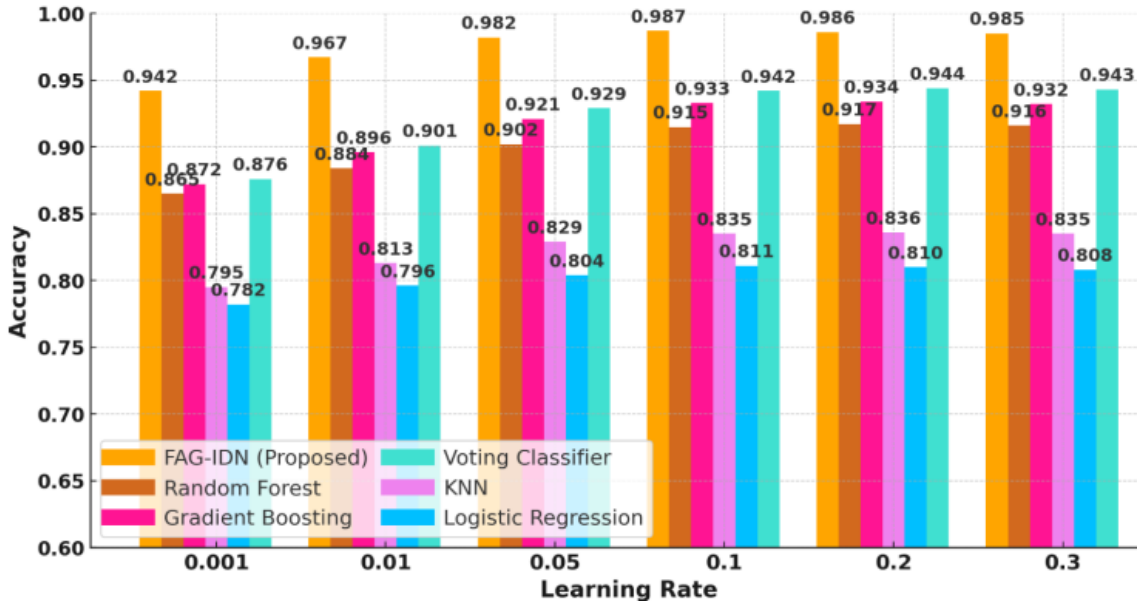

Figure 7: Accuracy Comparison of Different Algorithms Under Different Learning Rate Settings on the UNSW-NB15 Dataset

5.4.2 PowerGrid-IDS Results

As shown in Figure 8, on the PowerGrid-IDS dataset, the proposed FAG-IDN model has an accuracy of 0.992, and the Voting Classifier and Gradient Boosting models are at 0.982 and 0.972, respectively. Logit regression has a much smaller value of about 0.73 by comparison. This performance gap is due to the FAG-IDN model's integration of GCN and federated learning mechanisms; it can fully utilize the spatial dependency and global topological features among grid nodes, and at the same time preserve local feature information of each substation during training to improve the generalization ability and detection accuracy of the model. The Attention mechanism will give more weight to key feature areas and thus help the model spot irregularities in complicated grids more accurately and sensitively.

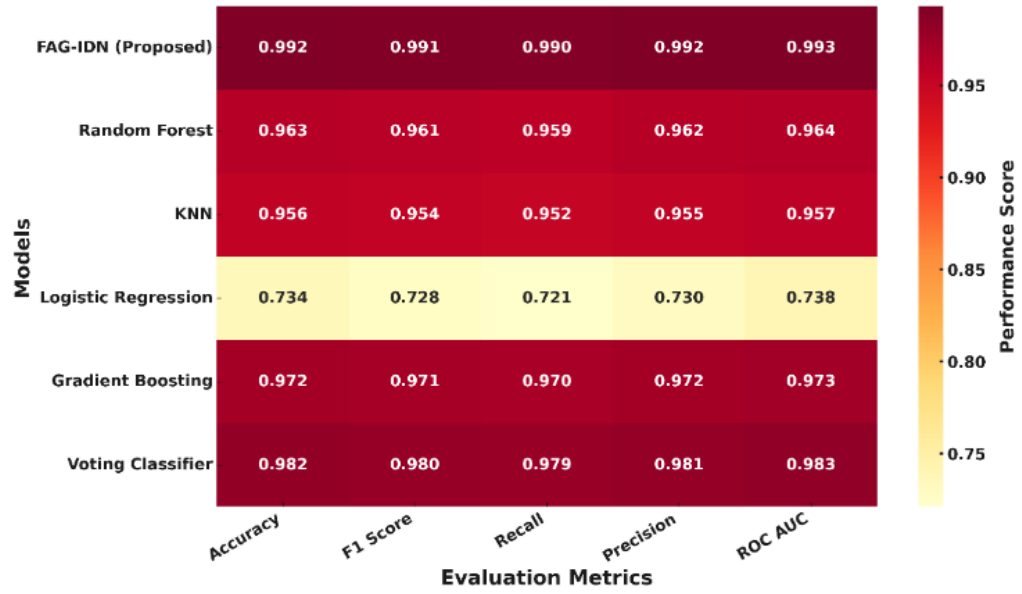

Figure 8: Comparison of the Used Algorithm in the Proposed System and Other Algorithms for Classification on PowerGrid-IDS Dataset

As shown in Figure 9, in the identification stage of the PowerGrid-IDS dataset, the proposed FAG-IDN model has an accuracy of 0.991 and is much higher than that of other algorithms. The Voting Classifier and Gradient Boosting model perform relatively well, but Logistic Regression has a lower overall accuracy of around 0.72. The two performers are different because the FAG-IDN model uses GCNs to model the topology of grid communication nodes and applies an attention mechanism for adaptive feature weighting. Thus, more accurately capture the spatial correlations and propagation features of attack behaviour. In addition, the federated learning framework can address the problem of data silos and coordinate the training of a global model by the substations, thereby improving both the accuracy and generalisation ability of identification in a privacy-preserving manner.

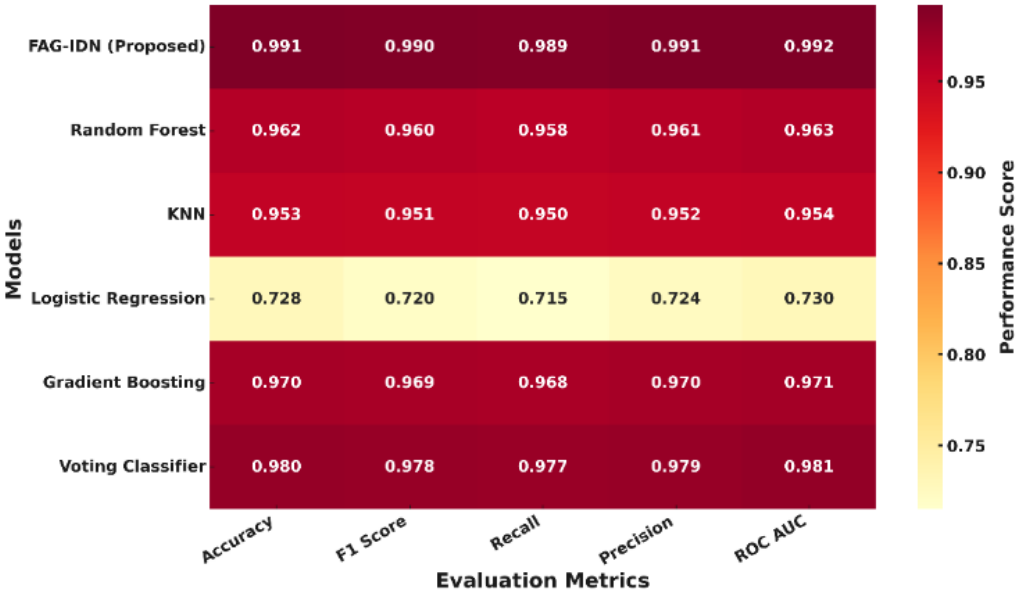

Figure 9: Comparison of the Used Algorithm in the Proposed System and Other Algorithms for Identification Stage on PowerGrid-IDS Dataset

As shown in Figure 10, the proposed FAG-IDN model had the best performance among the compared algorithms in the identification stage on the PowerGrid-IDS dataset and achieved an AUC of 0.992 on the ROC curve. Therefore, the model has a better discriminatory effect and a lower false-positive rate. The Voting Classifier (AUC=0.981) and Gradient Boosting (AUC=0.978) are not very good, and Logistic Regression is significantly worse (AUC=0.730). The first is that the FAG-IDN model is a graph neural network and employs attention mechanisms to extract rich, high-level relational features among nodes. In conjunction with federated learning for cross-node collaborative optimisation, significantly increase the robustness and detection accuracy of the model during the identification stage.

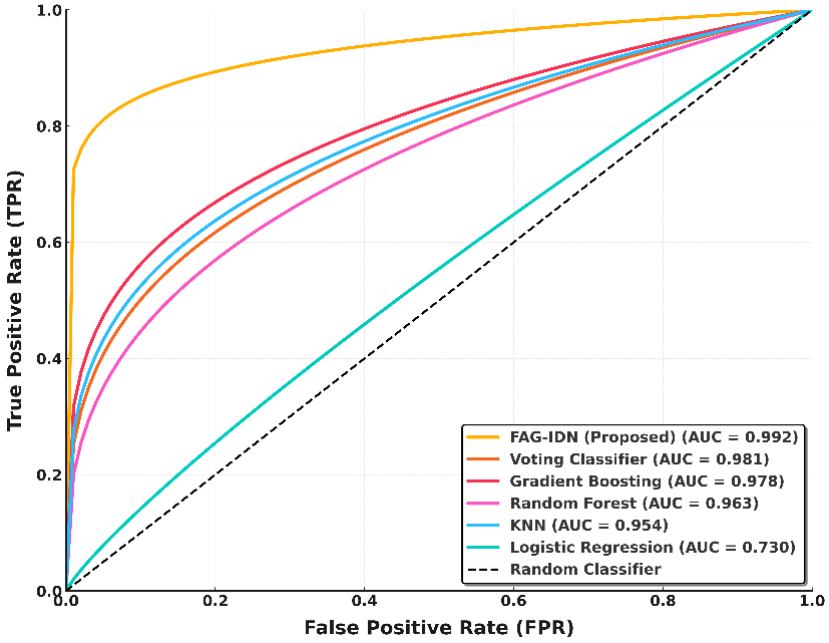

Figure 10: Comparison of the Used Algorithm in the Proposed System with Other Models in the Identification Stage via ROC Curve on PowerGrid-IDS Dataset

As shown in Figure 11, the accuracy of all algorithms on the PowerGrid-IDS dataset generally increases with an increase in maximum depth. Among them, the FAG-IDN model (orange) performs better at all depth levels and has reached a maximum accuracy of 0.992 in the depth range of 10-12. The Voting Classifier and Gradient Boosting models have relatively stable performance and maintain an accuracy of about 0.94-0.95; Random Forest is slightly lower than the other two, and KNN and Logistic Regression have performed worse overall. The performance difference is attributed to the combination of Graph Convolutional Networks (GCN) and federated learning in the FAG-IDN model; therefore, spatial structure information of a multi-node grid topology is effectively obtained, and deep feature weights can be adaptively adjusted for more precise classification. Increase the depth of the model, add an attention mechanism, improve the learning capability of FAG-IDN for complex intrusion patterns and reduce overfitting. Therefore, it has good stability and generalisation performance with various parameter settings.

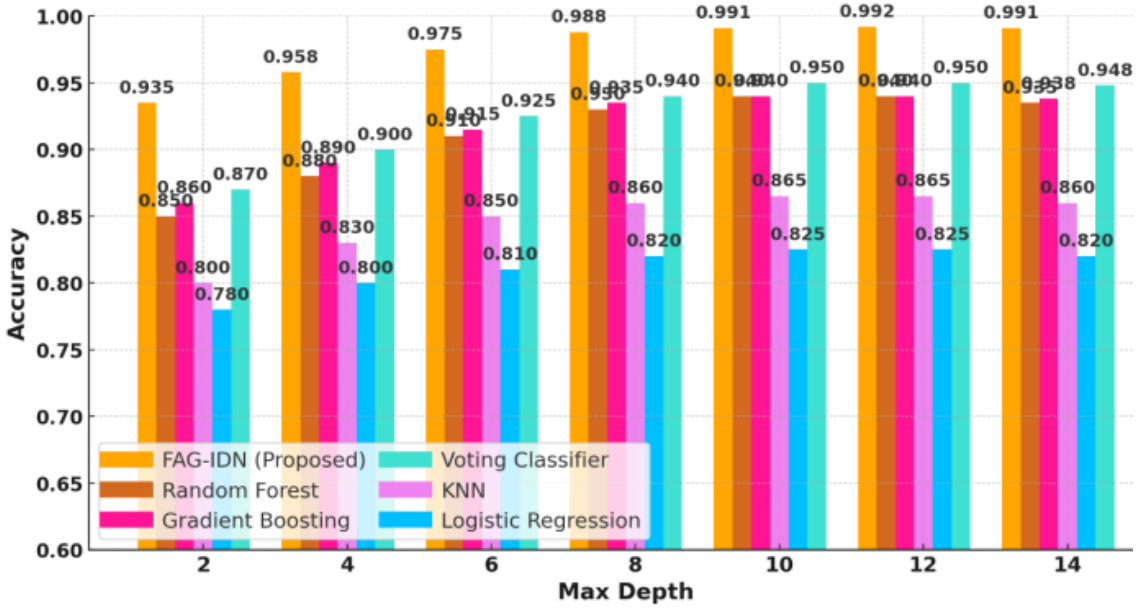

Figure 11: Accuracy Comparison of Different Algorithms under Various Max_Depth Settings on PowerGrid-IDS Dataset

As shown in Figure 12, the accuracy of all algorithms on the PowerGrid-IDS dataset increases with an increase in the learning rate before stabilizing. The FAG-IDN model (orange) is best in terms of performance for all learning rates.

The above model is better because, through the combination of federated learning and a graph neural network architecture, it can perform adaptive optimisation of the feature space. FAG-IDN can capture the topology and data dependencies among nodes efficiently at any learning rate. The attention mechanism will learn to weigh the important parts more heavily and thus avoid oscillations or overfitting problems caused by a large learning rate. Therefore, the model is relatively stable and accurate compared with the others.

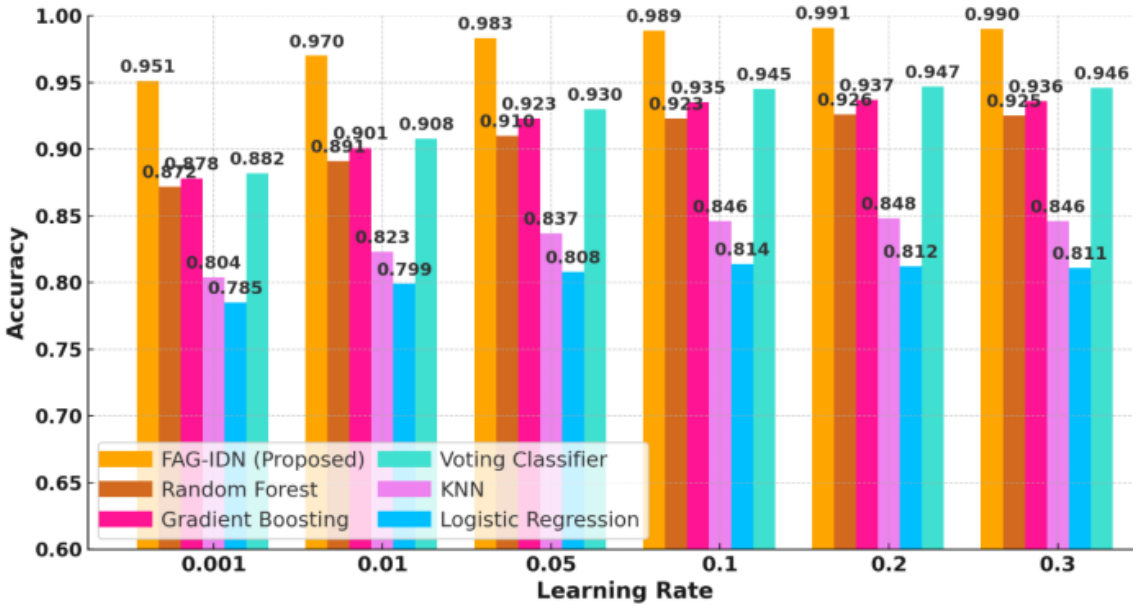

Figure 12: Accuracy Comparison of Different Algorithms under Various Learning Rate Settings on PowerGrid-IDS Dataset

6 Conclusion

The complex security risks in digital power grids under intelligent and networked conditions have increased; therefore, this paper proposes an intelligent intrusion detection model that combines Graph Neural Networks (GNN), Federated Learning (FL), and an Attention Mechanism. A model that supports multi-node collaborative training and global detection optimisation without exposing raw data. Build graph structures to acquire the spatial topology of grid nodes, then use an attention mechanism to dynamically weigh features and enhance the accuracy and stability of anomaly detection. According to the experimental results, FAG-IDN has performed well on both the UNSW-NB15 and PowerGrid-IDS datasets, and is thus efficient, private and scalable.

However, there are some deficiencies in the proposed way. First, in cases of highly imbalanced cross-regional data distribution, federated learning may lead to reduced generalisation performance of the global model. Second, the computational cost of the attention mechanism is relatively high, and it may reduce the speed of real-time detection as the number of nodes increases. The present model mainly addresses static structures with known communication topologies, and therefore has not been modified to deal with dynamic topology changes or node failures in power grids.

The following research directions will be explored in the future: (1) Incorporate Reinforcement Learning (RL) to achieve adaptive cooperation and strategy optimisation among federated nodes; (2) Introduce quantum-safe encryption and differential privacy mechanisms to enhance model security in the post-quantum era; (3) Design Dynamic Graph Neural Networks (Dynamic GNN) to improve system robustness and real-time performance under the condition of topological changes; and (4) Investigate cross-domain Federated Transfer Learning for knowledge sharing and security collaboration among different regional power grids.

About the Author

Peng Xiao, obtained a Bachelor's degree in Software Engineering from Dianchi College, Yunnan University.Currently serving as the Deputy Manager of the Network Security Management Center of the Information Center of Southern Power Grid Yunnan Power Grid Co., Ltd., a Level 3 leading professional technical expert, with a main research focus on information security assessment technology, including network attack and defense technology, network security management, and enterprise security system construction.

Zijie Deng obtained a M.S.E. degree from the South China University of Technology, Guangzhou, China in 2018. He is an engineer in China Southern Power Grid Power Grid Group, Co., Ltd., Guangdong Province, China. His main research direction is digital technology and Cyber Security.

Biao Bai is a general manager in Information Center of China Southern Power Grid Yunnan Power Grid Co., Ltd., Yunan, China. His main research direction is digital technology and Cyber Security.

References

[1] Al-Shetwi, A. Q., Hannan, M. A., Al-Masri, H. M. K., et al. (2024). Latest advancements in smart grid technologies and their transformative role in shaping the power systems of tomorrow: An overview. Progress in Energy.

[2] Egbuna, I. K., Salihu, F. B., Okara, C. C., et al. (2025). Advances in AI-powered energy management systems for renewable-integrated smart grids. World Journal of Advanced Engineering Technology and Sciences, 15(2), 2300-2325.

[3] Hasan, M. K., Habib, A. K. M. A., Islam, S., et al. (2023). DDoS: Distributed denial of service attack in communication standard vulnerabilities in smart grid applications and cyber security with recent developments. Energy Reports, 9, 1318-1326.

[4] Sahani, N., Zhu, R., Cho, J. H., et al. (2023). Machine learning-based intrusion detection for smart grid computing: A survey. ACM Transactions on Cyber-Physical Systems, 7(2), 1-31.

[5] Kaur, R., & Singh, M. (2014). A survey on zero-day polymorphic worm detection techniques. IEEE Communications Surveys & Tutorials, 16(3), 1520-1549.

[6] Aghazadeh Ardebili, A., Hasidi, O., Bendaouia, A., et al. (2024). Enhancing resilience in complex energy systems through real-time anomaly detection: A systematic literature review. Energy Informatics, 7(1), 96.

[7] Muneeswari, G., Rose, R. A. M., Balaganesh, S., et al. (2024). Mitigation of attack detection via multi-stage cyber intelligence technique in smart grid. Measurement: Sensors, 33, 101077.

[8] Goswami, M. J. (2024). AI-based anomaly detection for real-time cybersecurity. International Journal of Research and Review Techniques, 3(1), 45-53.

[9] Ahmed, S. F., Alam, M. S. B., Hoque, M., et al. (2023). Industrial Internet of Things enabled technologies, challenges, and future directions. Computers and Electrical Engineering, 110, 108847.

[10] Sarker, I. H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. SN Computer Science, 2(6), 1-20.

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

[12] Nazir, A., & Khan, R. A. (2021). A novel combinatorial optimization based feature selection method for network intrusion detection. Computers & Security, 102, 102164.

[13] Kasongo, S. M. (2021). An advanced intrusion detection system for IIoT based on GA and tree based algorithms. IEEE Access, 9, 113199-113212.

[14] AlHaddad, U., Basuhail, A., Khemakhem, M., et al. (2023). Ensemble model based on hybrid deep learning for intrusion detection in smart grid networks. Sensors, 23(17), 7464.

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

[16] Imrana, Y., Xiang, Y., Ali, L., et al. (2024). CNN-GRU-FF: A double-layer

feature fusion-based network intrusion detection system using convolutional neural network and gated recurrent units. Complex & Intelligent Systems, 10(3), 3353-3370.

[17] Polat, O., Ahmad, A. A., Oyucu, S., et al. (2025). Temporal-spatial feature extraction in IoT-based SCADA system security: Hybrid CNN-LSTM and attention-based architectures for malware classification and attack detection. IEEE Access.

[18] Bouguessa, A., Mostefaoui, S. A. M., Daoud, M. A., et al. (2025). TBAC-IDS: Enhancing intrusion detection with transformer-based alerts correlation. Cluster Computing, 28(16), 1012.

[19] Onyema, E. M., Dalal, S., Romero, C. A. T., et al. (2022). Design of intrusion detection system based on cyborg intelligence for security of cloud network traffic of smart cities. Journal of Cloud Computing, 11(1), 26.

[20] Arbaoui, M., Brahmia, M. A., Rahmoun, A., et al. (2024). Federated learning survey: A multi-level taxonomy of aggregation techniques, experimental insights, and future frontiers. ACM Transactions on Intelligent Systems and Technology, 15(6), 1-69.

[21] Hassan, A., Nizam-Uddin, N., Quddus, A., et al. (2024). Navigating IoT security: Insights into architecture, key security features, attacks, current challenges and AI-driven solutions shaping the future of connectivity. Computers, Materials & Continua, 81(3).

[22] Wu, K., Zhang, C., Hao, F., et al. (2025). Fc-gcn: A formal concept-enhanced graph convolution network model. Soft Computing, 29(6), 2715-2725.

[23] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In 2015 Military Communications and Information Systems Conference (MilCIS) (pp. 1-6). IEEE.