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

Design of AI-based Defense Mechanism for Digital Power Grid Attack Chains in Large-scale Renewable Energy Integration

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

[1] Information Center of China Southern Power Grid Yunnan Power Grid Co., Ltd., 650000,Yunnan, China
[2] China Southern Power Grid Power Grid Group Co., Ltd., Guangdong Province, 510000,China

SUMMARY: *In short, while the combined deployment of distributed renewable energy sources and intelligent control in smart power grids has improved their efficiency, risks in cyberspace have also increased. To address the deficiencies in reactive intrusion detection and traditional protection methods, this paper proposes a distributed intelligent defense framework that integrates GNN, reinforcement learning and federated transfer learning. Multimodal states are formed by topology, communication actions and logic semantic meanings, and PPO is used by people to generate adaptive defence policies. Knowledge distillation also transfers the cloud-side capacities to a lightweight edge agent. Experiments conducted on a simulated cyber-physical power grid show that the framework can achieve 94% operational stability and is therefore 3-5% better than the baseline method. It has also reduced the number of model parameters by approximately 68%, maintained inference delay time under 20ms, and decreased communication cost expenditure by 42%. The above results show that this framework can provide support for active, expandable and privacy-preserving network protection in intelligent power grids.*

KEYWORDS: *Smart grid cybersecurity, Graph Neural Network, Reinforcement Learning, Proximal Policy Optimization, Federated Transfer Learning, Knowledge Distillation*

1 Introduction

Electric power systems are developing into cyber-physical energy systems through communication networks, intelligent control and automation [1]. The increase in flexibility and situation awareness is enabled by the convergence of IT and OT at SCADA, IEDs and substation controllers, expanding the attack surface [2, 3]. In 2015 and 2016, Ukraine's grid attacks were coordinated; as a result, control commands were manipulated and sensing data corrupted, leading to a large-scale outage [4, 5]. Therefore, the foundation of an intelligent, distributed and data-driven power system is now cyber resilience.

Existing smart-grid defences include machine-learning anomaly detection, intrusion prevention and blockchain-secured communication [6]. Three boundaries remain. Many models are only activated after the appearance of an attack [7]. Static strategies are ineffective once the system's behaviour has been analysed by attackers [8]. Cryptographic and blockchain schemes strengthen integrity but are often too computationally expensive for real-time, large-scale grid operation [9]. These constraints need to have flexible, distributed and resource-aware

*zijiedeng2025@163.com
https://doi.org/10.65102/is2026853
Published: 30 April, 2026

defences.

A Hybrid GNN-RL-FTL Framework for Intelligent Power Grid Cyber Defence is Proposed in this Paper. GNNs can learn the connections among physical devices and communication nodes to model how an attack spreads in a coupled layer. A PPO-based agent is used to learn a defence policy from network traffic, system logs and topological embeddings, and balances stability, response speed and computational cost. Knowledge distillation and federated transfer learning are used to connect cloud-based teacher models with edge-based student models, and privacy-preserving knowledge sharing among substations and regional control centres is achieved. The local operating data are stored at the site and periodically collected jointly for defence.

Dynamically adjust the policy's entropy and exploration strength based on the degree of defence failure, response delay and stability of the current policy. Compared with previous DRL studies on game theory and graphs for grid defense [10-12], the proposed design incorporates structural modeling, adaptive policy learning and distributed coordination in a single defense loop. The main achievements are topology-aware cyber-physical modelling, PPO-based adaptive defence, cloud-edge KD-FTL deployment that reduces model size by 68% and keeps inference latency below 20ms, and closed-loop policy adjustment for stable cooperative defence of distributed power-grid infrastructure.

2 Related Work

Recently, in smart-grid cybersecurity research, deep learning, graph neural networks, reinforcement learning, federated learning and knowledge distillation have been applied to achieve lightweight deployment [13]. Although the above studies have provided a foundation for intelligent and distributed defence, limitations remain in generalisation, real-time response and privacy-preserving coordination. Tirulo and his colleagues have presented some methods for the detection of FDIA and DoS using CNN-, RNN- and autoencoder-based approaches; however, these methods also require large-scale labelled datasets and demonstrate poor generalization capabilities in non-stationary grid environments [14]. Wang and others put forward FedGraph-KD to compress knowledge from high-capacity teacher models into light-weight local student models, reducing both communication costs and privacy risks compared with FedAvg [15]. It is still mainly classification in scope, has limited temporal modelling, and lacks a reinforcement-learning-based defence policy.

Federated and Graph-based Studies Further Expand Distributed Monitoring. Bondok and others have put forward a distillation-based adversarial defence with perturbation-smoothing constraints to increase the resistance of malicious updates, but asynchronous communication can still cause parameter drift and slower convergence [16]. Ma and others used federated learning and graph neural networks to conduct distributed intrusion detection in a 6G-IoT environment, confirmed the benefit of structural modelling for heterogeneous networks, but this method still only addresses static detection rather than real-time response [17]. Kim et al. have built a GNN-based asynchronous federated transfer-learning model for digital-twin-driven distributed systems that enhances privacy-preserving cooperation and convergence stability, but they have not yet added reinforcement-learning policy optimisation for adaptive defence.

RL-based defence has addressed the problem of dynamics more directly. Suresh designed a PPO-based adaptive intrusion-response system that can choose defensive measures at different times during an attack [18]. This way shows that reinforcement learning can be applied to sequential defence, but policy oscillation and slow convergence have not been solved yet. Wang and others have also put forward a multi-agent reinforcement learning scheme for regional-grid

cooperation to enable policy sharing and game-theoretic reasoning under attack propagation [19, 20]. Its centralised training and communication coordination have reduced the scalability of a large-scale grid.

In short, the three current research deficiencies are: First, most of the methods are still detection-focused and offer little support for active adaptive defence. Second, graph-based models have enhanced structural awareness but are often unable to capture the time-varying nature of multi-stage attack propagation. Thirdly, although the above federated and distillation-based schemes support distributed learning, asynchronous updates and non-IID data may still result in model drift and a decline in performance. To address the above deficiencies, this paper constructs a distributed defence framework that integrates GNN, PPO, KD and FTL to support cross-region and multi-level protection and continuous learning in smart grids.

3 The Proposed Method

This chapter introduces the general Design of an AI-based defence against cyber attack chains in digital power grids. The goals of the proposed framework above are to obtain information about the system structure using GNN, design strategies with DRL, and build a high-speed response function that determines the required action via a loop defense mechanism.

3.1 Architecture Overview

Figure 1 shows a three-layer AI-driven defence architecture for cyber-physical power grids to deal with stealthy, fast-moving and multi-layered attack chains. A closed-loop architecture for perception-decision making-execution has been designed to achieve dynamic grid modelling, adaptive defence selection and real-time response.

A Graph Neural Network (GNN) is used in the perception layer to model the physical and communication layers of the grid as a directed weighted graph G(V,E). Each node vi is a grid device or control entity with electrical and operational attributes, and each edge weight wij indicates the reliability, latency and availability of communication or energy flow. A graph convolution and an attention mechanism have been added to learn the changes in topology and identify structurally weak links for real-time prediction of an attack path in this model. The Decision Layer is a DRL-based Actor-Critic model. The Actor outputs a probability distribution over defence actions, and the Critic values the total reward. A multi-objective reward function is used to address the problems of detection accuracy, response delay, system stability and defence costs simultaneously, and thus modify the behaviour of the agent according to changes in the threat characteristics. DCA agents in the execution layer carry out the chosen actions across the grid network to isolate, reroute, roll back, and configure reset within a few milliseconds. Execution delay is also passed back to update the model; on the other hand, in response to new conditions in the environment, defensive means will be adjusted.

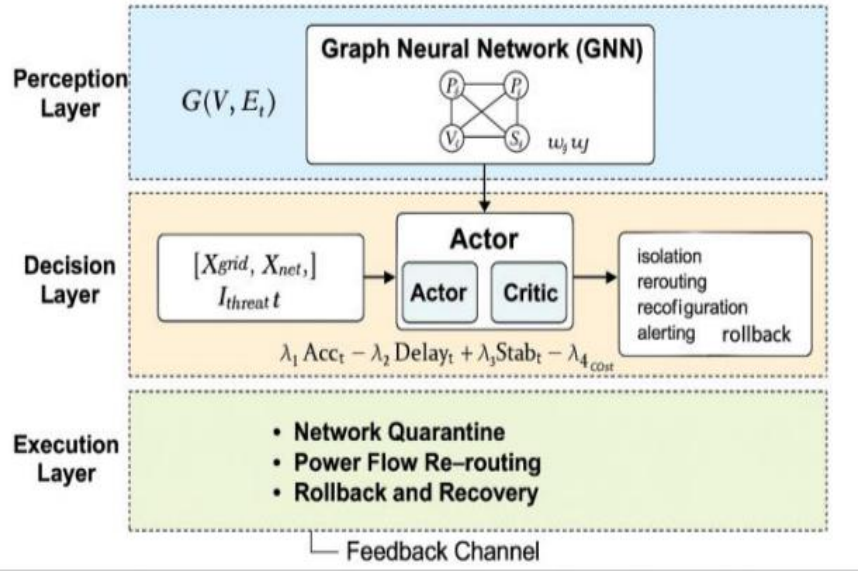

Figure 1: Schematic Diagram of the AI-Powered Defense Architecture

3.2 Attack Chain Modeling

Given the development of the digital electric network, network attacks have continuously changed and evolved in recent years, spreading through a combination of cyber and physical layers. In order to express this kind of temporal and structural dynamics accurately, we model the attack chain as a time-evolving directed weighted graph G(v, Et), where the vertex set v represents the key entities of the system, and the edge set Et records the possible propagation paths of the attack that change with time t. The model can form the joint expression of physical-cyber dependencies and time connections between all attack steps. Figure 2 shows the frame picture of the cyber kill chain building model.

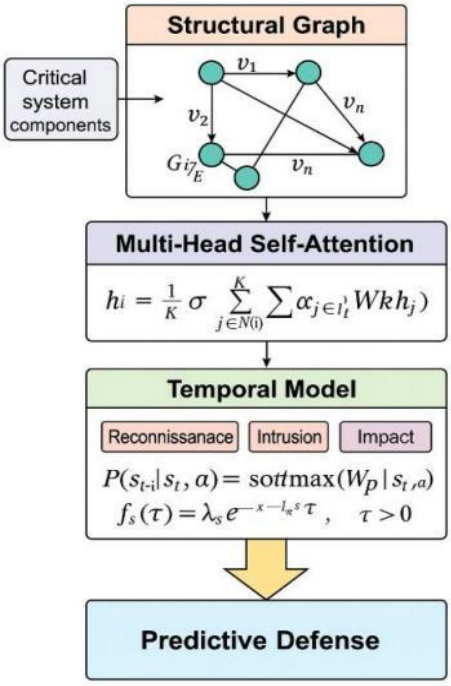

Figure 2: Attack Chain Modelling Framework

3.2.1 Graph-Based Structural Definition

The intelligent electricity network is considered by the public to be a dynamic attack graph that contains essential control, communication and interface nodes. The possible connection relationships among nodes are represented as possible invasion paths, and their strength indicates the degree of weakness in connecting, trust situations, etc. Since these connections change during the attack, the topology is considered to be in a state of change over time rather than fixed. Provide a foundation for the representation of multi-step attack spreading in cyber-physical electric power systems.

3.2.2 Multi-Head Self-Attention Enhanced Graph Encoding

A multi-head self-attention mechanism is used to extract high-dimensional dependency features of the dynamic graph structure, and multi-path attack signals can be effectively combined. The first is:

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

where: h_i^{I} is the updated representation for node v_i, N(i) is the set of neighboring nodes of v_i, α_{ij}^{k} is the attention weight assigned to neighbor node v_j by the k-th attention head, w_k is the linear transformation matrix for the k-th attention head, and $\sigma(\cdot)$ is the non-linear activation function (LeakyReLU is used here). The attention weight α_i is calculated according to the similarity of node features:

$$\alpha_{ij}^{k} = \frac{\exp\left(\text{LeakyReLU}\left(a^{T}\left[W^{k}h_i \parallel W^{k}h_j\right]\right)\right)}{\sum_{m \in N(i)} \exp\left(\text{LeakyReLU}\left(a^{T}\left[W^{k}h_i \parallel W^{k}h_m\right]\right)\right)} \tag{2}$$

The two are as follows: the two models focus on different types of attack channels at the level of attention heads.

Head-1: Anomalies in the communication layer (e.g., sniffing, traffic redirection).

• Head-2: Manipulations in the control layer (e.g., false instruction injection, state tampering),

Head-3: Physical Layer Impacts (e.g., voltage disturbances and frequency attacks).

Head-4: Anomalies in the data chain (e.g., timing pollution, fake sampling).

Concatenating the outputs of all attention heads is done.

The final graph representation HI = [h, h, ..., h] is generated.

Attack features at all levels to provide a more stable input for the following decisions.

3.2.3 Attack Stage Transition Modeling

The network invasion of intelligent power grids is a multi-stage process that, as a general rule, progresses from information gathering to intrusion, horizontal expansion, and finally, control. To address the time-related dependence, the attacking procedure is now being modelled as a one-by-one state-change mechanism. This expression provides a basis for showing the organised development of attack steps in cyber-physical electric power systems.

$$P(s_{t+1}|s_t, a_t) = \text{softmax}(w_p [s_t, a_t]) \tag{3}$$

where at is the defensive action taken (e.g., isolation, rerouting, rollback), and wp is the stage

transition weight matrix. Softmax normalizes the probabilities of all the stages to produce a distribution for the next step in the attack chain.

To improve the model's sense of time, a stage residence time distribution $f_s(\tau)$ is introduced to show the duration characteristics of an attack at stage s:

$$f_S(\tau) = \lambda_S e^{-\lambda s \tau}, \tau > 0 \quad (4)$$

λs is the rate of change at the stage. Rapid calculation of the expected length of the attack and an effective response by means of the model.

3.2.4 Attack Chain Prediction and Intervention

By repeatedly updating the values of stage-transition probability, the model we have constructed can be used for prediction of attack evolution. When the risk of it being promoted to the impact stage is high and exceeds the intervention threshold, in advance, measures such as node isolation and traffic arrangement will be taken. Thus, the framework will have a proactive defence capability to cut off an attack before it occurs rather than responding after an incident has taken place.

3.3 Reinforcement Learning Strategy Design

To have adaptable and changeable defenses, DRL will be added to the cyber-defense system of the digital power grid. By constantly observing changes in its environment, it builds a "sense of situation" to drive the learning of a defense agent. PPO is not restricted to discrete action spaces as in conventional Q-learning/DQN, and it can learn continuous and stable policies for the non-stationary, high-dimensional cyber-physical grid environment.

3.3.1 Algorithm Framework

PPO (Proximal Policy Optimisation) is a method that extends Policy Gradient (PG) by introducing a constraint on the step size of policy updates to prevent large deviations in the policy during optimisation. PPO aims to maximise the advantage function.

$$L_{PPO}(\theta) = E_t [\min(r_t(\theta)\hat{A}^t, clip(r_t(\theta), 1 - E, 1 + E)\hat{A}^t)] \quad (5)$$

where: $r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$ represents the ratio of the new

and the old policy probabilities, t = Rt - V(st), are the estimated advantage functions, V(st) is the value function that approximates the cumulative expected reward, E is the clipping factor, usually set between 0.1 and 0.3.

The Actor network is used to generate the policy $\pi_\theta(s_t)$ at time t during training, and the Critic network evaluates the value function $V(s_t)$ of the current state. Both networks are updated synchronously by gradient descent to improve the policy and maintain stability constraints:

$$\theta^I \leftarrow \theta + \eta \nabla_\theta L_{PPO}(\theta) \quad (6)$$

Where η is the learning rate. To improve the stability of the policy further, Generalized Advantage Estimation (GAE) is introduced in this paper.

$$\hat{A}^t = \sum_{l=0}^{T-t} (\gamma\lambda)^l \delta_{t+l}, \delta_t = r_t + \gamma V(s_{t+1}) - V(s_t) \quad (7)$$

GAE is used to reduce both bias and variance by adding a smoothing coefficient λ, and thus enhances the stability and convergence of learning in dynamic grids.

3.3.2 State Representation Optimization

The state expression of a cyber-physical power grid contains heterogeneous information such as topological structure, communication network and operating behaviour, and therefore exhibits strong high-dimensional characteristics, dynamic changes and cross-domain connections. In order to model the above characteristics, a multi-modal state embedding network has been proposed to jointly encode structural, network and behaviour features in a unified representation space.

$$s_t = [H_t^{GNN} \times F_t^{NetFlow} \times M_t^{syslog}] \quad (8)$$

where: H is the topological embedding extracted by the

Graph Neural Network (GNN) for structural dependencies among grid components; FtNetFlow to extract temporal features of communication dynamics from network flow sequences; Mtsyslog to extract behavioral semantics from host-level system logs; Concatenation operation for multimodal feature fusion.

All of the above are relatively independent.

contextual information: H captures inter-node

FtNetFlow is about communication changes, and Mtsyslog is for system command meanings. Together, they make up the complete state vector that can describe both the structure and operation of the cyber-physical grid.

The physical and communication structures of the grid are shown as a directed graph G(V, E), where V = {v1, v2, ..., vn} is the set of electrical or control nodes, and E is the flow of energy or data among these nodes.

A GNN aggregates local neighbourhood information through a message-passing mechanism to update the hidden representations of all nodes as follows:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{c_{ij}} W^{(l)} h_j^{(l)} + b^{(l)}\right) \quad (9)$$

where: hi(l) is the node representation at layer l; N(i) denotes the neighbor set of node vi; cij is a normalization constant (usually degree-based); w(l) and b(l) are learnable parameters; σ(.) is the activation function, and LeakyReLU is used to strengthen negative gradient propagation.

One GNN based on two-way attention is used to capture two-way information flow and the interdependence of complex nodes. The above Design can be used to collect both the signals entering and leaving a node; thus, it is more stable in the face of a dynamic topology change, such as node disconnection or link failure. Based on the above, the GNN module will construct a model for the dynamic dependency graph of the system and enhance structural awareness in recognising coordinated or cascading attacks. At the level of communication, characteristics such as sudden increases in network flow, packet retransmission, and time anomalies can be used to describe sequence rules, and thus Temporal Convolutional Networks are employed to extract these rules.

$$FtNetFlow = TCN(xtflow ; k, d) \tag{10}$$

where: xtflow: Input flow time series (packet rate, directional entropy, latency, etc.);k : Convolution kernel size;

d: Dilation rate, controlling the temporal receptive field.

The TCN uses ReLU activation, Batch Normalisation and a Dropout layer (dropout rate p=0.3) to reduce overfitting. Dilated convolution design can address the problem of long-term temporal dependency without increasing computational cost, and thus is suitable for detecting low- and slow attacks that have extended stealthy behaviour.

Mtsyslog is a module in the system that provides high-level semantic information about the operations of hosts and controls, such as command sequences, access paths, and changes in privileges. A Transformer-based log encoder is used to encode the above sequences:

$$M_t^{syslog} = TransformerEncoder\ (E_t^{log}) \tag{11}$$

where Etlog refers to the embedded word representations of log entries.

The number of heads in the encoder's multi-head self-attention is 8, and each head has a dimension of 64. Sinusoidal function-based positional encodings are added to retain the sequence information, and then mean pooling is used to generate sentence-level embeddings.

This module can also identify fine-grained dependencies in the sequence of operation commands, such as the quick execution of a command after an unauthorised login or abnormal power promotion. By this way, the syslog coding device provides semantic abstraction for command-level behaviour and can help discover anomalies related to control-layer intrusion. A single multi-layer fusion module has been proposed for the three-modality integration. First of all, all kinds of features are made to reside in a common representation space; then, attention and gating mechanisms are used to regulate how much each modality contributes, and finally, normalisation and residual connections are added to improve the stability and convergence of training. The resulting combined state representation can perform adaptive cross-domain feature fusion and is thus situation-aware, even in the event of partial data loss, such as missing sensor readings or damaged logs. By jointly aligning and fusing multiple embedded vectors, a single stable-state representation for decision-making in dynamic confrontation-oriented grid environments has been built.

3.3.3 Dynamic Policy Adaptation

Given that the characteristics of cyberattacks on digital power grids are non-stationary and unpredictable, we put forward an Adaptive Policy Regulation Mechanism (APRM) to adjust the defence policy dynamically within a continuous action space. It continuously observes changes in the environment and variations in behaviour, and thus achieves an optimum trade-off between exploration and exploitation to adapt to all stages of an attack.

The decision-making policy of the defense agent is a continuous action distribution, and thus it can explore the cyber-physical environment stochastically but under control. At each time step t, the policy output is given by:

$$a_t = \pi_\theta\ (s_t) + E,\ E \sim N(0, \sigma_t^2) \tag{12}$$

where $\pi_\theta(s_t)$ is the mean action predicted by the policy network, and E is zero-mean Gaussian exploration noise with variance σ_t^2. The variance term σt dynamically adjusts to the uncertainty of the environment according to:

$$\sigma_t = \sigma_0 (1 + \beta_t), \beta_t = \tan h(\alpha u_t), u_t = Var(\hat{A}^t) \quad (13)$$

σ_0 is the base noise scale, β_t is the dynamic exploration factor, u_t is the variance of the advantage function at time t, and α_i is a scaling coefficient (usually 0.1-0.3).

The above design will automatically adjust the level of exploration in response to changes in the environment: When uncertainty is relatively high, such as when a new type of attack appears or the threat suddenly escalates, the system will increase the scope of exploration to obtain more data; conversely, if uncertainty is low, this exploration will be reduced to ensure the stability of decision-making.

In addition to other aims, an entropy-regularisation term has also been introduced to promote exploration and avoid premature convergence of the policy in the objective function:

$$a_t = \pi_\theta (s_t) + E, E \sim N(0, \sigma_t^2) \quad (12)$$

λH is the entropy weight coefficient for balancing exploration and convergence. To suit the different stages of training, λ_H was reduced exponentially with time as shown below:

$$\lambda_H (t + 1) = \lambda_H (t) \times e^{-rt} \quad (13)$$

The entropy annealing scheme ensures that the policy is relatively random in the early stages of learning; thus, it is encouraged to explore broadly; and then, as we approach the end of training, it begins to settle down somewhat. The decrease in entropy can be smooth and stable; thus, a relatively stable and reliable defense plan can be learned that is less susceptible to transient noise or brief fluctuations. Good policy optimisation can be achieved in practice by a favourable exploration-exploitation trade-off throughout the learning process.

Given that it is an unstable environment, a feedback-based adjustment method for the policy will be adopted. The three real-time performance indicators in the system are:

1. Policy Stability (sstab): Cosine similarity of consecutive policy vectors is used to evaluate behavioural smoothness.

2. Response Efficiency (Ref), the mean of defence success rate per unit of time;

3. Computation Cost (cload) is the mean decision latency of the defence agent.

When the system finds that the stability is too low (small sstab) or the defense efficiency has dropped sharply (small Reff), it will automatically adjust both the value of the dynamic exploration factor β_t and the entropy weight λ_H. The closed-loop feedback mechanism restores balance by increasing randomness and searching upon detecting instability, and then reduces this increase as the policy becomes more consistent.

Therefore, in the face of new cyber threats, it will continuously strengthen adaptive self-regulation. By adding uncertainty-aware exploration, entropy-driven diversity control and feedback-based adjustment, a good control loop has been constructed to make the policy adaptable in the presence of uncertainty and stable when the environment stabilizes. Thus, the above dynamic regulation framework can help the system balance the demands of learning adaptability, behaviour stability and computational efficiency to achieve good performance in highly dynamic and adversarial grid conditions.

3.3.4 Self-Learning and Transfer Mechanism

In order to improve the performance of making it larger, to cooperate better with many small components, and to reduce the space at the boundary too much, we have developed a mixture of two methods named KD and FTL that allows computers to continue learning and safely share

their knowledge even when far apart. In the cloud setting, a large-capacity teacher model $\pi_T(a \mid s)$ is trained using extensive data and computational resources to achieve a good policy expression ability and decision-making accuracy. To perform real-time inference at the network edge, a lightweight student model $\pi_S(a \mid S)$ is placed on the edge equipment, such as substation protection terminals. The model of how students learn from the teacher's model is that, through practice, the Kullback-Leibler (KL) divergence between their policy distributions decreases.

$$\mathrm{LKD} = \mathrm{DKL}\,(\pi\mathrm{T}(\mathrm{a}|\mathrm{S}) \,||\, \pi\mathrm{S}(\mathrm{a}|\mathrm{S})) \tag{16}$$

To make it larger, perform better with various small parts and have little space at the edges, we used a combination of two things called KD and FTL that allow computers to continuously learn and share their ideas safely even when they are far apart from each other.

Based on KD, a Federated Transfer Learning (FTL) mechanism has been proposed to facilitate cooperative learning and private knowledge sharing in distributed grid areas. Each regional node locally trains and updates its parameters θ_i, and periodically uploads gradients or weights to a central aggregation server for global synchronization:

$$\theta_{global} = \sum_i \frac{n_i}{N} \theta_i \tag{17}$$

FTL is not an intensive training mode, and it does not share the original attack data; only parameter updates occur, thus protecting data privacy and enabling distributed nodes to participate in a general-purpose defense model. The above measures can promote generalisation by combining attack knowledge from different areas, and are thus still compatible with various hardware and network environments.

Knowledge distillation can be used to build a continuous learning structure for FTL that supports partial parameter updates, federated data consistency and fast adaptation to new attack types without full retraining. Therefore, this frame promotes cross-area adaptive ability, collective self-promotion, and privacy-protecting cooperation under the condition of data-separation restriction. At the core of the mixed KD–FTL scheme is a scattered cyber-protection space that can be continuously observed, locally changed, and developed together with intelligent power grid basic facilities.

4 Experimental Validation and Case Study

4.1 Case Study Setup

In order to verify the validity of the proposed AI-based cyber-physical protection scheme for digital networks, a case study has been conducted on a medium-sized city distribution network that has been regarded as an intelligent microgrid including PV generation. The testbed is a real-world CPS that has both fluctuations in renewable energy and coordinated cyberattacks simultaneously, allowing us to test the multi-layered defence mechanisms we have proposed against such threats.

There are 33 distributed PV units on the microgrid platform, each with a rating of 50-200kW, and the total capacity is 3.5MW. Every 5 minutes, based on CAISO irradiance and load data, a demand of 1.8-3.0 MW is dispatched in real time. To reproduce a multi-stage cyberattack, false data will be injected into 35% of the PV inverter command channels and a DoS interruption applied to 20% of the communication nodes in the environment. The four links of the attack are

reconnaissance, intrusion, control-of-the-system, and impact. A digital-twin deception framework of 100 virtual PV instances that refreshes every 250ms simulates cyber-physical interaction during an attack. The detection dataset has added 2.5 million labelled samples of real-world incident records and synthetic adversarial examples. NSGA-III is used to optimise the defence policies over a rolling 48-hour horizon, and MATLAB/Simulink, Pyomo, CPLEX and TensorFlow are deployed on a Xeon-A100 high-performance computing platform.

As shown in Figure 3, the combined attack significantly reduces the PV output. Approximately 12 units have been damaged, their outputs are between 50 and 160 kW, and the unaffected units are still relatively stable at around 80-200 kW. Lower median output, larger variance and multiple outliers in the compromised units all point to inverter-control disruption and reduced dispatch stability. Based on the above results, a united cyberattack will directly affect the stability of production and cause widespread disorder across the whole distribution network; thus, adaptable AI-based defences are needed.

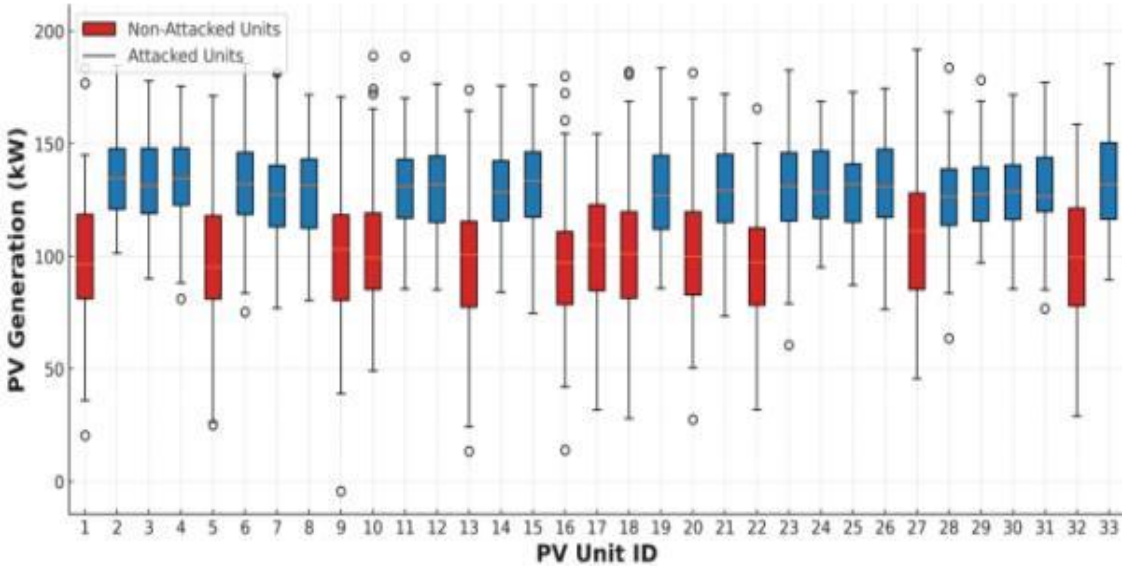

Figure 3: PV Generation Profile and Cyberattack Status

4.2 Comparison Methods

The four baseline sets for validation are:

Graph Neural Network Intrusion Detection (GNN-ID): A topology-aware detection method based on Graph Convolutional Networks for grid communication dependency modeling.

Reinforcement Learning-based Anomaly Defense (RL-AD): A deep Q-learning-driven defense policy that learns by repeatedly interacting with simulated attack environments.

DT-DD: A deception-based cyber-defense framework that uses digital twin replicas of PV and control nodes to mislead attackers.

Federated Transfer Learning Defense (FTL-DD): A cooperative learning model that shares defense knowledge among multiple grid nodes via gradient aggregation instead of sharing raw data.

4.3 Efficiency and analysis

4.3.1 Adaptive Evolution of Attack Detection Performance

Figure 4 shows how the detection of attacks changes in different conditions of cyber defence and PV grid response. In an environment of low complexity, protection based on rules can only

maintain a detection accuracy of 60%–75%, and thus some stealthy attacks will go unnoticed. After the introduction of reinforcement learning-based adaptive detection, the accuracy in medium- and high-complexity cases has exceeded 90%, indicating that the recognition of multi-stage attacks is also relatively strong. Using an RL strategy and a digital twin cheat method, a detection rate of 98% has been achieved, and thus the danger can be identified before causing a disruption to the power grid's normal operation. Although the transition zone has a small oscillation of 85% and 90%, the whole range is not random, and thus a stronger defense design will achieve better detection results. This result provides support for the effect of the AI-driven hierarchical cyber defence structure that we have proposed.

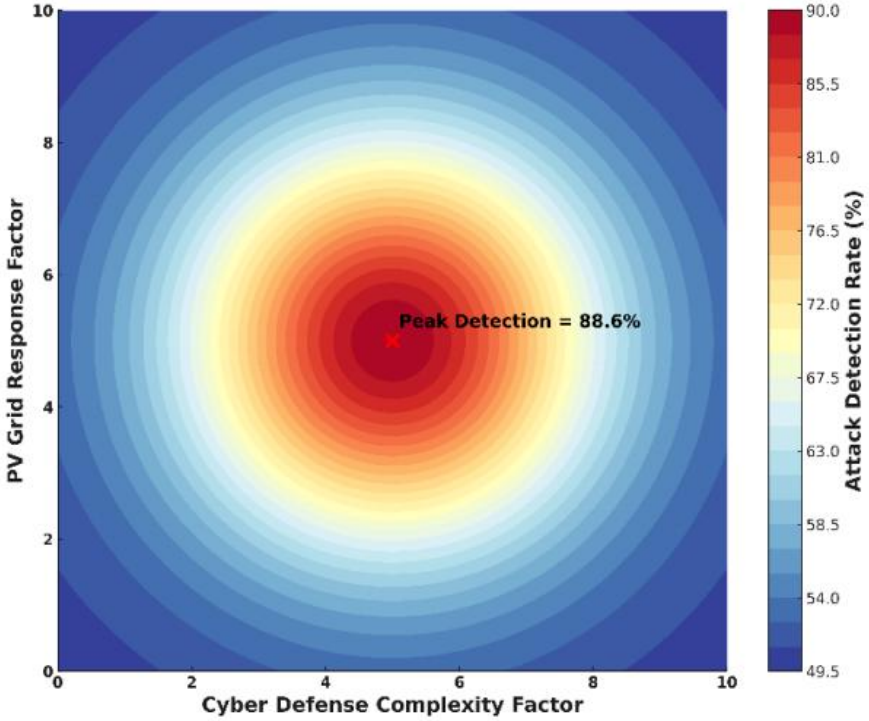

Figure 4: Cyberattack Detection Rate Contour

4.3.2 The Dilemma between Security and Efficiency

As shown in Figure 5, the complexity of cybersecurity has reduced the speed at which the digital grid with PV integration can be operated in real time. Under the condition of protection from light, the efficiency of the system is still over 85% and can provide continuous energy supply. As security technology has developed increasingly sophisticated, so have the particular demands for blockchain verification; thus, digital twins and reinforcement learning-based defences have emerged, but their overall efficiency is currently around 50% to 60%. The above results show that strong protection and high operating speed have been balanced. Therefore, this figure shows that the design of cyber defence must also meet the demand for real-time performance in cyber-physical power systems.

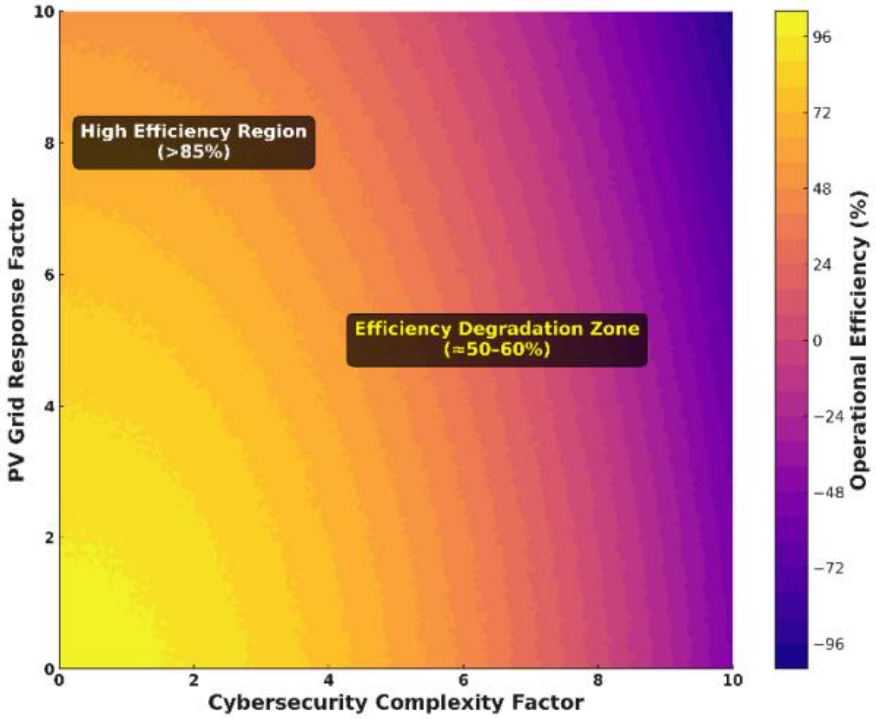

Figure 5: Computational overhead Impact Contour

4.3.3 Comparative Evaluation of Cyber-Defense Scenarios

Table 1 shows the comparison of the five cybersecurity configurations for a PV-integrated digital grid under coordinated cyberattacks. PV stability index is a measure of how far the generation output has moved from the normal state; a larger difference indicates a more severe disturbance caused by an attack. The correctness of the attack check indicates the proportion of harmful events that have been correctly identified, and this rate is significantly improved after introducing digital-twin cheating and reinforcement-learning-dependent self-adapting protection. The capacity of the system to distribute power during a power failure is called dispatch efficiency, and with the continuous development of advanced AI-driven protection systems, it has also been rising. Computational cost adds to the extra load of each defence layer; although it is larger with stronger protection, it remains within a range suitable for the real-time operation of a power grid. The resilience score is a general index of cyber-physical robustness that indicates how quickly a system can return to its original state after a cyber attack under the joint support of blockchain verification, digital twin model construction, and reinforcement learning-based decision intelligence.

Table 1: *Performance Indicators in Various Scenarios.*

Scenario	PV Stability (kW dev.)	Detection (%)	Dispatch (%)	Comp.Overhead (ms)	Resilience Score (0-100)
Baseline (No Attack)	5.2	0.0	97.8	5	85.2
Under Attack (No Def.)	38.7	45.2	68.4	0	45.3
Digital Twin Defense	12.4	78.6	85.1	42	70.1
RL-Based Cyber Defense	9.1	92.3	91.6	78	88.5
Full Cyber-Resilient Optimization	6.8	98.5	96.2	120	94.8

4.4 Comparative Analysis

4.4.1 Detection Rate vs. Network Complexity

Figure 6 shows the heatmap contrast of attack check rates for five defence methods at different degrees of network complexity: GNN-ID, RL-AD, DT-DD, FTL-DD, and the framework we proposed. The way we have proposed so far has consistently achieved a relatively high detection accuracy, reaching 97%-98% even in very complex network environments; therefore, it is considered to be highly generalizable and adaptable to new kinds of security threats. FTL-DD and DT-DD are relatively good enhancers, with check rates in the range of 85%-90%, while GNN-ID and RL-AD are still below 82% and show poor stability at higher complexity levels. Although all the methods are better with an increase in network complexity, only the framework we have put forward can maintain this advantage and be very robust. Therefore, the above results indicate that a hierarchical AI-driven defence structure is relatively superior in handling the problem of complex cyberattacks on digital electric power grids with high detection rates.

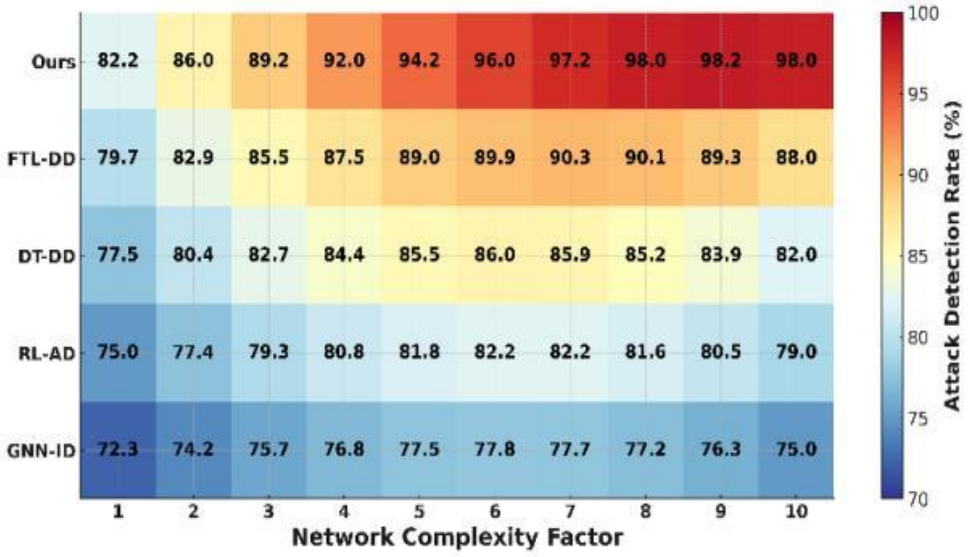

Figure 6: Comparison of Attack Detection Rates Under Different Network Complexity Levels

4.4.2 Operational Efficiency vs. Security Layers

Figure 7 shows the quantity of safety stratification and the operational degree of the digital electric power grid under various methods of resistance to malicious actors on the Internet. When the number of safety layers is between 1 and 5, the efficiency of all methods decreases gradually; therefore, a larger amount of work is required by the computer and its communication with other parts of the system becomes relatively cumbersome. But our Ours framework still has the highest efficiency at the top, around 92%, and is still above 82% even with strong security. Thus it has been shown that it is more suitable for achieving a good balance between the security of cyberspace and the real-time operational requirements of power grids. On the other hand, GNN-ID and RL-AD, these traditional methods, have much poorer results, and their effectiveness falls below 75%; thus, they are unable to perform expansion in the face of complex multi-layer defences. FTL-DD and DT-DD are moderate in performance and also drop significantly when using more complex defences. From the above figure, we can see that the AI-driven architecture proposed here has reached an excellent equilibrium point between security depth and continuous operation development; thus, strong safety can be guaranteed without sacrificing system response speed.

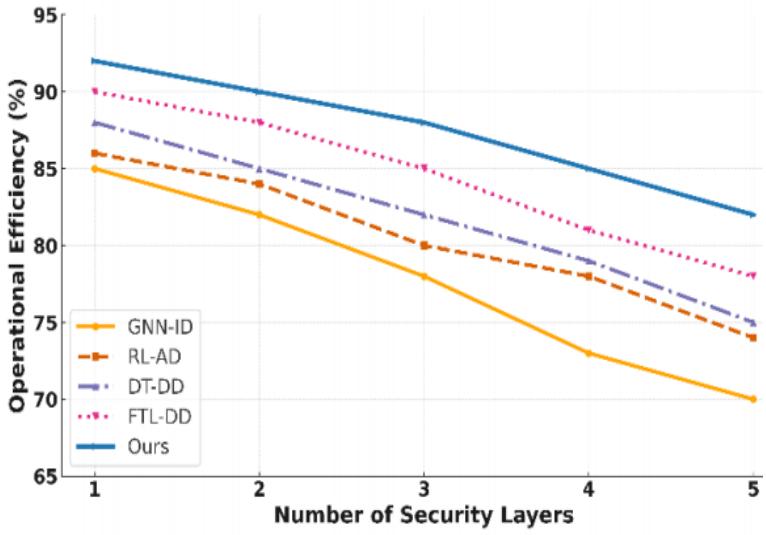

Figure 7: Changes in Real-Time Operational Efficiency at Different Depths of the Security Layer.

4.4.3 Response Time vs. Defense Adaptability

As shown in Figure 8, the mean response time of the different cybersecurity measures for handling a specific level of defence adaptability varies. As shown in the figure above, all the methods have a downward trend in response time with an increase in adaptability; therefore, those with higher adaptability are better suited for real-time decision-making. GNN-ID and RL-AD have a slow response time at low adaptability (over 150ms) because there is little room for policy updates and fixed defence thresholds. DT-DD and FTL-DD have moderate improvements, and the response time is around 110-130ms under the adaptive reinforcement mechanism. Our proposed framework always achieves the lowest average response time (less than 100ms) at all levels of adaptability and shows good dynamic coordination between reinforcement learning agents and digital-twin deception modules. It shows that resources are being used efficiently and policies are converging quickly; both are necessary for real-time cyber-physical defence in PV-integrated digital grids. Based on the above results, it can be concluded that the addition of adaptive intelligence generally improves both the speed of response and flexibility to changes in the attack.

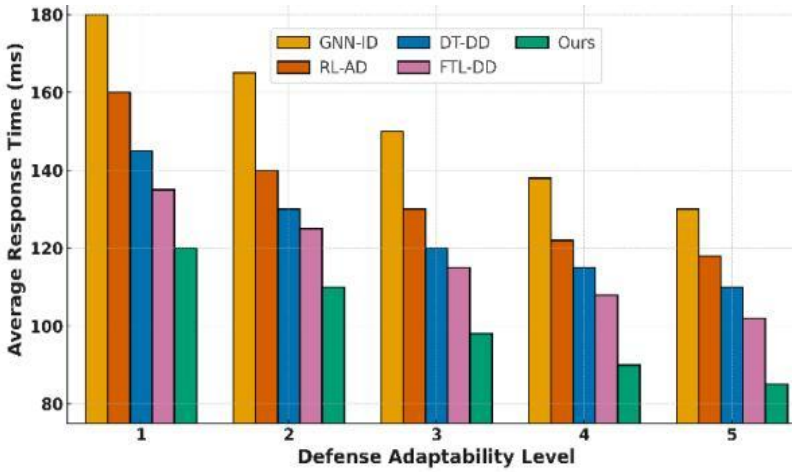

Figure 8: Average Response Time under Different Levels of Defense Adaptability

4.4.4 Operational Stability Comparison Across Defense Methods

Fig. 9 is a violin plot showing the distribution of operational stability for the five cyber-defenses under dynamic grid conditions. Each violin plot shows the probability density and variation of system stability under different defence mode choices. As shown in the above results, our method has the highest mean stability (94%) and a narrow distribution; thus, it is more robust and less sensitive to cyber-physical disturbances. FTL-DD and DT-DD have relatively stable values (90-92%) and are slightly more distributed; therefore, they are somewhat more sensitive to changes in the system. GNN-ID and RL-AD have a wider and lower-centered distribution (80-85%), so they are more unstable and inconsistent in the presence of adaptive attacks. In short, the above system can keep working well in the face of continuously developing cyber threats and maintain the stability of the power grid.

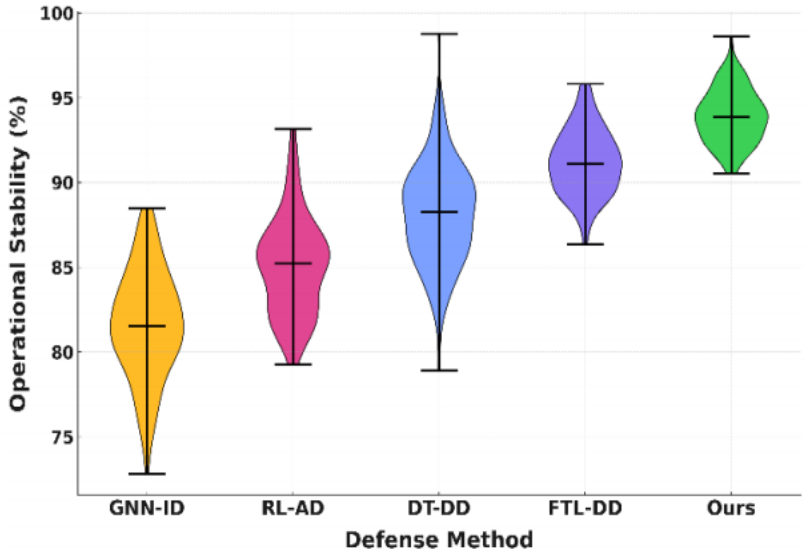

Figure 9: Violin Plot of Operational Stability Distribution under Different Defense Strategies

5 Conclusion

Construction of a mixed network defense framework for intelligent power grids based on the combination of GNN, PPO, federated transfer learning and knowledge distillation in this study. The above three are jointly used to form a combined state expression, and thus can be flexibly adapted to deal with changes in the order of attacks. Experimental results show that compared with traditional static projects, the improved system has better working stability, response delay less than 20 ms, an attack recognition rate of 98.5 per cent, and steady performance in the presence of topology changes and partial data loss. The main contribution is that it brings together structural model construction, reinforcement decisions, and distributed continuous studies in a single cyber-defence theory system. Thus, the intelligent power grid has achieved scalability in protection and maintained privacy. In the future, the work we carry out will expand this framework to multi-energy systems and larger power grid environments.

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 earned the title of M.S.E. from the South China University of Technology in Guangzhou, China, in 2018. He is an engineer at China Southern Power Grid Power Grid Group Co., Ltd. in Guangdong Province, China. The first field of study is Digital Technology and Cybersecurity.

Biao Bai is the general manager of the Information Centre at China Southern Power Grid Yunnan Power Grid Co., Ltd. in Yunan, China. The first is research in digital technology and cyber security.

References

[1] Inderwildi, O., Zhang, C., Wang, X., et al. (2020). The impact of intelligent cyber-physical systems on the decarbonization of energy. Energy & Environmental Science, 13(3), 744-771.

[2] Ajayi, O. O., Alozie, C. E., & Abieba, O. A. (2025). Enhancing cybersecurity in energy infrastructure: Strategies for safeguarding critical systems in the digital age. Trends in Renewable Energy, 11(2), 201-212.

[3] Aslam, M. M., Tufail, A., Apong, R. A. A. H. M., et al. (2024). Scrutinizing security in industrial control systems: An architectural vulnerabilities and communication network perspective. IEEE Access, 12, 67537-67573.

[4] Tirulo, A., Chauhan, S., & Dutta, K. (2024). Machine learning and deep learning techniques for detecting and mitigating cyber threats in IoT-enabled smart grids: A comprehensive review. International Journal of Information and Computer Security, 24(3-4), 284-321.

[5] Tatipatri, N., & Arun, S. L. (2024). A comprehensive review on cyber-attacks in power systems: Impact analysis, detection, and cyber security. IEEE Access, 12, 18147-18167.

[6] Aljumah, A., Ahanger, T. A., Ullah, I., et al. (2025). Smart methodology for defence asset management in blockchain environment. Cluster Computing, 28(11), 690.

[7] Abubakar, R., Aldegheishem, A., Majeed, M. F., et al. (2020). An effective mechanism to mitigate real-time DDoS attack. IEEE Access, 8, 126215-126227.

[8] Afianian, A., Niksefat, S., Sadeghiyan, B., et al. (2019). Malware dynamic analysis evasion techniques: A survey. ACM Computing Surveys, 52(6), 1-28.

[9] Musa, H. S., Krichen, M., Altun, A. A., et al. (2023). Survey on blockchain-based data storage security for android mobile applications. Sensors, 23(21), 8749.

[10] Fang, D., Guan, X., Lin, L., et al. (2020). Edge intelligence based economic dispatch for virtual power plant in 5G internet of energy. Computer Communications, 151, 42-50.

[11] Zhu, M., Anwar, A. H., Wan, Z., et al. (2021). A survey of defensive deception:

Approaches using game theory and machine learning. IEEE Communications Surveys & Tutorials, 23(4), 2460-2493.

[12] Pazho, A. D., Noghre, G. A., Purkayastha, A. A., et al. (2023). A survey of graph-based deep learning for anomaly detection in distributed systems. IEEE Transactions on Knowledge and Data Engineering, 36(1), 1-20.

[13] Waikhom, L., & Patgiri, R. (2023). A survey of graph neural networks in various learning paradigms: Methods, applications, and challenges. Artificial Intelligence Review, 56(7), 6295-6364.

[14] Tirulo, A., Chauhan, S., & Dutta, K. (2024). Machine learning and deep learning techniques for detecting and mitigating cyber threats in IoT-enabled smart grids: A comprehensive review. International Journal of Information and Computer Security, 24(3-4), 284-321.

[15] Wang, S., Xie, J., Lu, M., et al. (2023). FedGraph-KD: An effective federated graph learning scheme based on knowledge distillation. In 2023 IEEE 9th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) (pp. 130-134). IEEE.

[16] Bondok, A., Badr, M. M., Mahmoud, M., et al. (2025). Securing smart grid federated learning against advanced evasion attacks using ensemble-based adversarial training. IEEE Internet of Things Journal.

[17] Ma, X., Hu, J., Liang, S., et al. (2025). Federated learning and resource-aware graph neural network for intrusion detection in 6G-IoT driven healthcare system. IEEE Internet of Things Journal.

[18] Suresh, A., & Cyril Jose, A. (2025). Adaptive network intrusion detection using reinforcement learning with proximal policy optimization. ACM Transactions on Privacy and Security, 28(4), 1-24.

[19] Wang, Y., Liu, X., & Yu, X. (2025). Research on joint game-theoretic modeling of network attack and defense under incomplete information. Entropy, 27(9), 892.

[20] Kim, Y. J., Kim, H., Ha, B., et al. (2025). Federated digital twins: A scheduling approach based on temporal graph neural network and deep reinforcement learning. IEEE Access.

[21] Sinneh, I. S., & Yanxia, S. (2025). Federated deep MPC-enabled digital twin and multiagent learning framework for secure and scalable smart nano grid energy management. Renewable Energy Focus, 100762.

[22] Jiang, L., Li, Q., Che, X., et al. (2025). A knowledge distillation enhanced semi-supervised federated learning framework for intrusion detection in EV charging networks. IEEE Internet of Things Journal.