Journal of Theoretical and Applied Information Technology
30st June 2026. Vol.104. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A FEDERATED DEEP REINFORCEMENT LEARNING FRAMEWORK FOR PRIVACY-PRESERVING REAL-TIME ANOMALY DETECTION IN SMART GRID IoT SENSOR NETWORKS

Dr. SHOBANA GORINTLA[1], BODIGIRI SAI GOPINADH[2], CHOPPARAPU SRINIVASA RAO[3], ANTHARAJU K CHAKRAVARTHY[4], Dr. MAREPALLI RADHA[5], Dr. K CHINNAIAH[6], Dr. S. BANUMATHI[7*], VALETI NAGARJUNA[8]

[1]Professor, Department of CSE, Dr. RVR NRI Institute of Technology Deemed to be University, Agiripalli, Andhra Pradesh, India

[2]Assistant Professor, Department of Mathematics, GMR Institute of Technology Deemed to be University, Rajam, Andhra Pradesh, India

[3]Department of CSE, Lakireddy Bali Reddy College of Engineering (A), Mylavaram, Andhra Pradesh, India

[4]Assistant Professor, Department of IT, Aditya University, Andhra Pradesh, India

[5]Associate Professor, Department of CSE, CVR College of Engineering, Hyderabad, Telangana, India

[6]Associate Professor, Department of CSE, MLR Institute of Technology, Dundigal, Telangana, India

[*7]Professor, Department of EEE, M.Kumarasamy College of Engineering, Karur, Tamilnadu, India

[8]Assistant Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

E-mail: drgshobana@gmail.com[1], saigopi1993@gmail.com[2], srinivas.lovely10@gmail.com[3], kalyan.antharaju@adityauniversity.in[4], marepalli.radha@gmail.com[5], drchinnanitc@mlrit.ac.in[6], banumathis.eee@mkce.ac.in[7*], nagarjuna.valeti@gmail.com[8]

ABSTRACT

The recent trend of Internet of Things (IoT) sensor networks to facilitate smart grids has presented serious challenges in detecting anomalies, ensuring data privacy and enabling real-time decision-making. This paper presents a novel framework for Federated Deep Reinforcement Learning (FDRL) to develop an effective, privacy-conscious anomaly detection system for distributed smart grid systems. The given approach applies federated learning along with deep reinforcement learning to provide decentralized model training and the evolution of attack patterns without access to raw data. A hybrid reward functional is developed to maximize recognition accuracy, energy consumption and latency. The results of the experiments prove that the proposed model achieved an accuracy of 97.2%, which is higher than those of traditional machine learning, deep learning, and federated learning methods. Moreover, communication overhead and latency can also be minimized by 38% and 27%, respectively. The findings provide an informative background on the system's strength, scalability, and real-time nature for addressing non-IID IoT data. The findings of the current piece of work may be summarized as the following: the suggested FDRL method is a viable and scalable way of securing next-generation smart grid infrastructure with significant implications on privacy-sensitive smart energy systems.

Keywords: *Federated Learning, Deep Reinforcement Learning, Smart Grid, Anomaly Detection, IoT Security, Edge Computing*

1. INTRODUCTION

The adoption of smart grids in place of traditional electrical power systems has greatly influenced the generation, transmission, and utilization of energy. Smart grids incorporate advanced sensing, communication, and control technologies. Their vision is of a future in which real-time monitoring and control of distributed energy resources will be implemented through intelligent, autonomous algorithms [1]. The spread of Internet of Things (IoT) sensor networks is a key

facilitator of this change. These networks collect high-resolution, immediate data on voltage, current, frequency, and load demand [2]. The IoT infrastructure will enable enhanced process efficiency and reliability, while incorporating renewable power sources into the electricity grid [3].

Nevertheless, there are major concerns about security, scalability, and data privacy when relying more on IoT devices. Smart grid IoT networks are vulnerable to anomalies, including sensor breakdowns, communication failures, data manipulation, and distributed denial-of-service (DDoS) attacks [4], [5]. This may destabilise the grids, lead to incorrect decisions, and trigger massive blackouts. Thus, resistance and security in a smart grid system are the basis for forming robust yet rapid anomaly-detection mechanisms.

Statistical and rule-based methods have been widely used to spot anomalies in power systems. Nevertheless, the methods have shortcomings in tracking intricate nonlinear correlations and in handling the time constraints of large-scale IoT data [6]. The performance of machine learning based approaches has been significantly better (Support Vector Machines (SVMs) and k-Nearest Neighbours (k-NNs)). However, they may be technically difficult, such as the use of labelled datasets and hand-engineered features [7]. More recently, research on deep learning models for identifying abnormalities in time series data has produced encouraging results with Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. These models can acquire representations of features of time series data [8], [9]. The available methods are mostly based on a centralised training architecture, even though there have been improvements. This raises issues of data privacy, communication overhead, and system scalability [10].

The new paradigm of Decentralised model learning is Federated Learning (FL). It facilitates learning a global model across multiple edge devices without requiring the exchange of unprocessed information [11]. This approach is particularly suitable for smart-grid environments, where many IoT devices share information and privacy is a concern. FL also minimises communication costs by transmitting data only when the model changes, not the sensor [12]. Some studies have examined how FL can be used in IoT systems and shown that it is effective for distributed anomaly and intrusion identification [13], [14]. Nevertheless, the majority of FL-based strategies are rigid and cannot adapt to transforming environments [15].

Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) are potent approaches to sequential decision-making in uncertain and dynamic environments. DRL enables systems to learn best policies by interacting with the environment. This makes it suitable for detecting anomalies in real-time in smart grids and for acting on them [16]. Cybersecurity and network optimisation have also been achieved applying techniques such as Deep Q-Network (DQN) [17]. Nonetheless, single DRL methods need centralised training. This may lead to scalability problems in IoT networks with large device populations [18].

To overcome these shortcomings, over recent years, attempts have been made to examine hybrid federated and reinforcement learning methods. This procedure can be applied to both distributed and privacy-preserving learning, as well as to adjustable learning under changing conditions [19]. Nevertheless, Federated Deep Reinforcement Learning (FDRL) for smart grid anomaly detection is under-researched, particularly in terms of its ability to function effectively in real-time, achieve energy efficiency, and be implemented at scale [20].

Due to the issues described above, the aim of the study is to develop a new and effective anomaly detection system for smart grid IoT systems that utilizes the advantages of Federated Deep Reinforcement Learning (FDRL). The main goal is to develop a unified model that combines federated learning with deep reinforcement learning to enable distributed, intelligent anomaly detection throughout IoT nodes. In addition, the study will strive to ensure privacy during model training by avoiding the need to share raw sensor data across devices, thereby enhancing data security. The other important goal is to develop an adaptive program to identify dynamic, ever-evolving attack patterns in real-time and enhance smart grid operations. In addition, the framework embraces communication overhead and latency reduction through edge-based distributed learning and enhance communication, exploiting edge-based smarts and decentralised learning methods. Finally, the proposed model is compared with baseline methods to determine its accuracy, computational performance, scalability and applicability to large-scale smart grid applications.

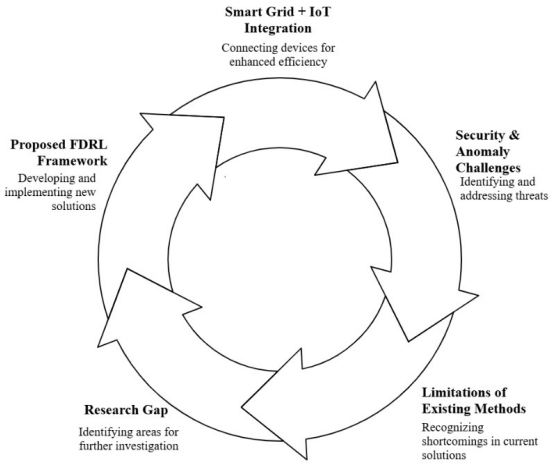

Figure 1. *Overview of the framework, including the history of smart grid IoT development, the obstacles posed by anomalies, the limitations of the current solutions, and the presented FDRL algorithmic model.*

To assist the reader in the research and in explaining the key addition, the rest of this paper is structured as follows. The combination of machine learning, deep learning, federated learning and reinforcement learning is presented in section 2, with the approach discussed in the context of a literature review on anomaly detection in smart grid IoT systems. Section 3 will introduce the approach and present the new Federated Deep Reinforcement Learning (FDRL) system and its underpinning elements: dataset characteristics, system design, and mathematical model. Section 4 includes experimental results and evaluation measures as well as a comparative analysis with the reference models, and is exemplified by graphical analysis. Finally, Section 5 summarizes the primary findings and limitations and outlines directions for future research, keeping the paper's original contribution in mind.

2. RELATED WORK

Identification of anomalies in smart grid IoT systems has also received significant research attention due to the increased complexity and susceptibility of cyber-physical energy infrastructure. The existing methods could be generally divided into the following types: classical machine learning, deep learning, federated learning, and reinforcement learning [21]. The first studies focused on conventional machine learning algorithms, including decision trees, SVMs, and ensemble methods, to identify anomalies in the power system. They partially succeeded in uncovering known attack patterns but failed to handle high-dimensional data and to extrapolate attacks in dynamically changing environments [22], [23]. Moreover, the emphasis on handcrafted characteristics gave them little freedom to move grid states and to support complex data streams from IoT [24].

It has been observed that deep learning-based models have been extensively researched in order to address these limitations. Unsupervised anomaly detectors have been used with autoencoders and Deep Belief Networks (DBNs) [25]. On the same note, RNNs, Gated Recurrent Units (GRUs), and Long Short-Term Memories (LSTMs) have been applied to extract temporal relationships in smart grid data [26], [27]. Spatial feature extraction for grid monitoring systems has also been performed using convolutional neural networks [28]. Although these models can offer higher detection performance, the solutions are frequently implemented in a more centralised architecture, which can raise questions about scalability, communication overhead, and data privacy [29].

The latest trends have observed hybrid deep learning structures that use other architectures to boost detection performance. Hence, for example, CNN-LSTM models have been proposed to learn the spatial and time-related characteristics of IoT data simultaneously [30]. The model has further enhanced its interpretability plus performance, as attention models have been shown to highlight important features during anomaly detection [31]. Although these enhancements exist, such models still rely on data centralisation and are thus not as compatible with a distributed IoT network [32].

Federated Learning (FL) has appeared as a promising approach for detecting anomalies in a decentralised IoT system. Previous studies that used FL-based intrusion monitoring systems include several that used edge-based devices to learn a model of the world, which do not provide raw data [33], [34]. By doing so, privacy is greatly enhanced, and communication costs are reduced. Nevertheless, the vast majority of available FL techniques are supervised/semi-supervised; they require labelled samples and are inefficient when attack patterns change rapidly [35]. Moreover, various problems, such as data disparity, client drift, and communication inefficiencies, are not addressed in a federated setting [36].

Along with FL, reinforcement learning (RL) has been explored for responsive security mechanisms in smart grids. However, RL-based methods can learn the best policy through experience with the environment and can be applied to dynamic anomaly detection and response [37]. Such Deep

Reinforcement Learning approaches as Deep Q-Networks (DQN) or policy gradient algorithms have already exhibited promising performance in network security and resource management [38], [39]. They are flexible models capable of making instant decisions, but they are typically used in a centralised environment with massive amounts of training data and computing power [40].

Recently, the fusion of federated learning with reinforcement learning has also been on the minds of researchers, who seek to utilize the strengths of both paradigms. Distributed control and optimisation problems have been proposed to use federated reinforcement learning schemes, which improve scalability and provide privacy protection [41], [42]. However, they are at the alpha stage of their anomaly-detection adaptation for smart grid IoT systems. The problem with the existing research is that it lacks an intensive review of real-time performance, energy efficiency, and dependability under different attack conditions [43]. To sum up, despite encouraging developments in smart grid anomaly detection, current methods do not succeed in achieving privacy protection, scalability, adaptivity, and real-time performance simultaneously. Centralisation affects classical and deep learning models; federated learning is rigid; and reinforcement learning fails to adequately address the overall problem of limited data [44].

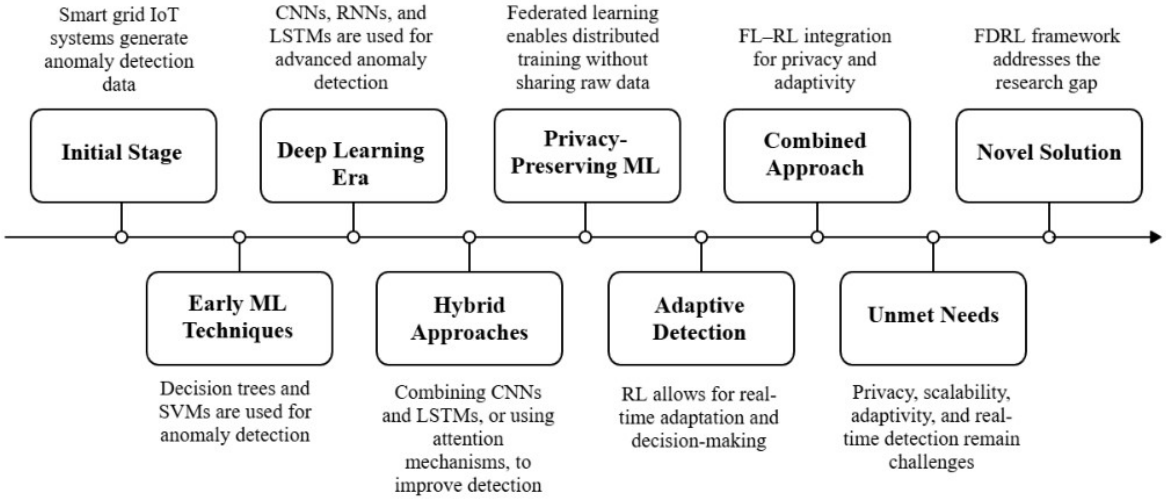

Figure 2. *Evolution of anomaly detection techniques in smart grid IoT systems leading to the proposed Federated Deep Reinforcement Learning (FDRL) framework*

As illustrated in Figure 2, anomaly detection methods have evolved from classical machine learning approaches, such as decision trees and support vector machines, which are limited concerning scalability and generalization, to deep learning techniques, including convolutional, recurrent, and long short-term memory neural networks. These deep learning nets provide better learning features and generalization of data processing. By combining models, including CNN-LSTM and attention-based models, detection becomes even more effective, though it remains based on centralized processing. Federated learning supports distributed and privacy-preserving training, whereas reinforcement learning enables adaptive, real-time decision-making. However, federated and reinforcement learning have not been fully applied to address privacy, scalability, adaptability and timely response in real-time implementation. Such a constraint underscores the need for a consolidated approach that provides all the necessary capabilities, thereby encouraging the Development of the suggested FDRL framework for anomaly detection in smart grid IoT sensor networks.

Prior research has made some progress in anomaly detection for smart grid IoT systems. However, several challenges remain. Traditional machine learning methods are difficult to generalize. Deep learning methods require centralized training. Federated learning methods may be difficult to adapt to dynamic attack patterns. Reinforcement learning methods may pose scalability and privacy concerns. Thus, there is still a need for a single framework that can achieve privacy preservation, scalability, adaptability, and real-time anomaly detection.

In line with this, there is a pressing need for an integrated framework that combines federated learning and deep reinforcement learning to enable effective, scalable, real-time abnormal identification in the smart grid IoT environment.

3. METHODOLOGY

This section presents the Federated Deep Reinforcement Learning (FDRL) framework, intended to promote reproducibility and further Development of the field. The section describes the dataset, architecture, mathematical formulation, and the algorithm's entire workflow.

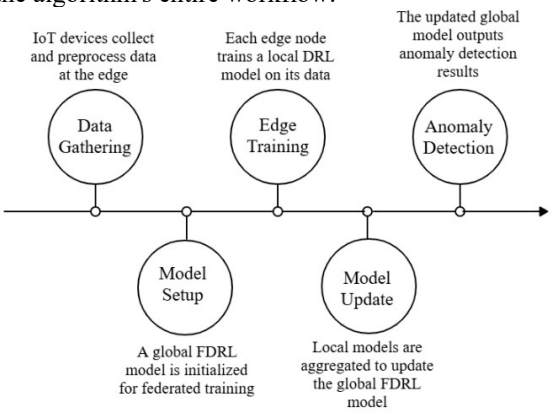

Figure 3. *Workflow of the FDRL-based anomaly detection system with data processing in edge data propagation and training, as well as federated training and anomaly detection.*

The proposed Federated Deep Reinforcement Learning (FDRL) process of anomaly detection in smart grids in Internet of Things (IoT) systems is shown in Figure 3. The beginning of the process is the collection of data. The IoT devices preprocess sensor data at the edge. Then, a global FDRL model initialization step is used to initialize a global model to be trained in a distributed manner. In the edge training phase, the IoT devices train their own Deep Reinforcement Learning (DRL) model using local data, and decentralised learning is possible. Then, the models that have been trained locally are uploaded to a central server where a federated aggregation step is performed which combines the parameters of the models to create an enhanced global model. Abnormalities are the next thing detected with the new worldwide model. It also gives real-time corrective action of irregular trends. This workflow provides an example of the main elements of distributed learning, privacy data preservation, and adaptive anomaly detection in the suggested model.

3.1 Dataset Description

To test the suggested framework, a reasonable smart grid Internet of Things (IoT) dataset has been built by incorporating publicly available reference datasets with simulated intrusion theory. Namely, electricity load diagram measurements from the UCI Electricity Load Diagrams dataset and cyber attack patterns from the UNSW-NB15 dataset were used to obtain both operational and security-related information. The resulting dataset contained information from 50 IoT sensor nodes, each with 120,000 time-series data points. In every sample, 18 electrical, network, and time-related features were provided, enabling robust modeling of grid behavior. Information has been documented at a 1-second sampling rate, which will be helpful for a fine-grained, time-based analysis to enhance immediate detection of anomalies in the smart grid setting.

Table 1. *Hybrid Smart Grid IoT Cyber-Physical Anomaly Detection and Intrusion Analysis Dataset (HSG-IoT-CPADIA)*
(Source:https://research.unsw.edu.au/projects/unsw-nb15-dataset)

Timestamp	Voltage (V)	Current (A)	Frequency (Hz)	Power Factor	Active Power (kW)	Reactive Power (kVAR)	Load Demand	Packet Rate	Latency (ms)	Packet Loss (%)	Time Index	Rate of Change	Moving Avg	Attack Type
00:00:01	230	5.2	50	0.95	1.14	0.35	Medium	120	10	0.1	1	0.02	1.10	Normal
00:00:02	231	5.4	50	0.96	1.20	0.38	Medium	125	12	0.2	2	0.05	1.15	Normal
00:00:03	250	7.8	49	0.80	1.95	0.90	High	300	45	2.5	3	0.60	1.50	FDI
00:00:04	229	5.1	50	0.94	1.10	0.30	Medium	900	120	8.0	4	0.03	1.25	DoS
00:00:05	228	5.0	50	0.95	1.08	0.32	Low	115	11	0.1	5	0.01	1.18	Replay
00:00:06	235	6.5	48	0.85	1.50	0.70	High	130	20	1.5	6	0.40	1.35	Sensor

														Fault

Table 1 presents a dataset that can be described as a hybrid smart-grid IoT ecosystem designed to detect and analyze anomalies in real-time. It comprises multivariate time-series measurements from several IoT sensor nodes, with each sample recording electrical, load, network, temporal, and derived measurements. Electrical parameters that define the operational condition of a power system include voltage and current, frequency and power factor. In contrast, features that define loads, such as active and reactive power, reveal electrical usage patterns. There are network attributes, such as packet rate, latency and packet loss, that indicate communication performance and potential cyber intrusion. Temporal characteristics, including timestamps and time indices, provide context for sequences, and some derived characteristics, such as rate of change and moving averages, can be used to identify trends and abrupt changes. This dataset is tagged with various conditions, including normal operation and attack types (e.g., False Data Injection (FDI), Denial of Service (DoS), and replay attacks), as well as sensor faults [45]. This full-scale feature can successfully model both system behavior and cyber-physical anomalies in smart grid environments.

Min-max normalization of the dataset and its division into edge nodes are used to create non-IID federated environments [46].

Table 2. *Feature Categories and Parameters Used in the Smart Grid IoT Dataset*

Feature Category	**Parameters**
Electrical	Voltage, Current, Frequency, Power Factor
Load	Active Power, Reactive Power, Load Demand
Network	Packet Rate, Latency, Packet Loss
Temporal	Timestamp, Time Window Index
Derived	Rate of Change, Moving Average

Table 2 presents the major feature categories used to monitor and analyze smart-grid and IoT-based systems. The basic operating parameters of the power system are electrical: voltage, current, frequency, and power factor. Load characteristics, active and reactive power, and load demand are used to show the patterns of electric energy consumption and the needs of the system. Network properties, including the rate of packets, i.e., packet rate, latency, and the loss of packets, determine the efficiency and the reliability of communication networks. Timestamps and time-window indices are some of the temporal features that define the temporal environment needed to analyze data trends. Derived features (rate of change and moving averages) are derived from raw data to highlight variations and historical trends. All these attributes enable valid, real-time monitoring and anomaly detection.

Table 3. *Types of Attacks Considered in the Smart Grid IoT Dataset*

Attack Type	**Description**
False Data Injection (FDI)	Manipulation of sensor readings
Denial of Service (DoS)	Network flooding causing delays
Replay Attack	Reusing old valid data
Sensor Fault	Random noise or malfunction

Table 3 describes typical attack categories and vulnerabilities that may occur in smart grid or IoT-based systems. False Data Injection (FDI) refers to intentional manipulation of sensor measurements to deceive the decision-making system. Denial-of-Service (DoS) attacks flood the network, causing delays or complete communication collapse. Replay attacks occur when valid data recorded earlier is resubmitted to the system, leading to incorrect interpretations. Sensor faults are technical problems, including random noise or instrument malfunction, that lead to incorrect data. Collectively, these threats indicate the need for robust identification mechanisms to make systems reliable and secure.

3.2 System Architecture

The proposed Federated Deep Reinforcement Learning (FDRL) architecture is a privacy-preserving distributed framework for intelligent anomaly detection. It is a privacy-preserving, distributed state framework for intelligent anomaly detection in Internet of Things setups. The method combines federated learning and deep reinforcement learning. Using this combination, model training across many edge devices can be achieved by collaborating effectively without exchanging raw data.

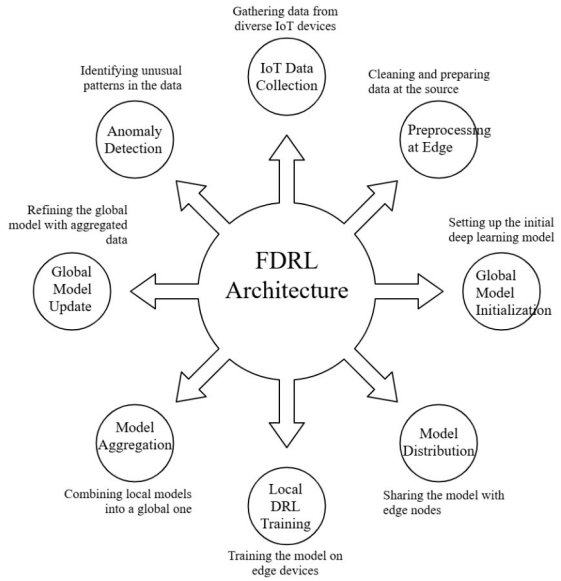

Figure 4. *FDRL Architecture for IoT-Based Anomaly Detection*

In Figure 4, the detailed Federated Deep Reinforcement Learning (FDRL) model used to detect anomalies in Internet of Things (IoT) systems in a distributed, privacy-aware manner is shown. The architecture is a cyclic process and a part of the continuous improvement of the model's capabilities and decision-making. The mechanism begins with IoT data collection, in which heterogeneous devices and sensors gather real-time information across various environments, including smart grids, medical systems, and industrial networks. Since this data is often noisy and unstructured, edge-based preprocessing is needed. To enhance data quality and minimize communication costs, data cleaning, normalization, and feature extraction are performed locally. After that, the system performs global model initialization, during which a central server develops a deep reinforcement learning (DRL) model that serves as the starting point for federated training. This original model is copied during the model distribution stage, where it is randomly distributed among a series of edge nodes, ensuring that all participating nodes receive the same model. On the edge, local DRL training is also performed on individual machines using local data. Reinforcement learning encourages the formulation of optimal policies by providing reward-based feedback from the environment, and it is therefore particularly applicable to dynamic, real-time conditions. The locally updated models, weights or gradients, are then sent to the central server upon training. The server takes the local models it receives in the Model Aggregation phase and combines them into a single global model using a federated learning method such as Federated Averaging (FedAvg). Such an approach allows keeping sensitive raw data locally without violating privacy while enabling shared learning.

The aggregate model is then updated during the global model update step, when the global parameters are varied to enhance generalization and the performance of all nodes. The improved model is redistributed to edge devices, and the learning process is iterative, running through multiple rounds. Lastly, the system relies on anomaly detection, where the trained global model identifies anomalies or abnormalities in the IoT data, covering faults, cyberattacks and abnormal behavior. This functionality helps to monitor it and make a decision immediately and knowledgeably.

On the whole, the architecture visualizes a closed-loop federated learning procedure that incorporates edge intelligence, distributed training and central coordination. These characteristics ensure scalability, information privacy, reduced latency and progressive learning, making the method appropriate for modern smart systems powered by IoTs.

3.3 Mathematical Model

3.3.1 Federated Learning Aggregation

The global model is updated using weighted averaging:

$$w^{t+1}=\sum_{i=1}^{N}\frac{n_i}{n}w_i^t \quad (1)$$

Where:

- w_i^t: local model at node i
- n_i: number of samples at node i
- n: total samples

3.3.2 Reinforcement Learning Formulation

The anomaly detection problem is modeled as a Markov Decision Process (MDP):

- State (S):

$$S_t=[V_t,I_t,F_t,L_t,Net_t] \quad (2)$$

- Action (A):

$$A=\{0:Normal,1:Anomaly\} \quad (3)$$

- Reward Function (Novel Contribution): A hybrid reward balancing detection accuracy and energy:

$$R_t=\alpha\cdot Acc_t-\beta\cdot E_t-\gamma\cdot Lat_t \quad (4)$$

Where:

- Acc_t: detection correctness
- E_t: energy consumption
- Lat_t: latency
- α,β,γ: weighting factors

3.3.3 Deep Q-Network (DQN)

$$Q(s,a)=r+\gamma a'maxQ(s',a') \quad (5)$$

Where:

- $Q(s, a)$: action-value function
- γ: discount factor

3.4 Proposed Novel FDRL Model

The suggested framework provides significant contributions towards anomaly detection in smart grid IoT systems. It uses a combination of federated learning and deep reinforcement learning to enable Federated Reinforcement Learning for training adaptive models in a distributed manner. The energy-conscious reward-optimisation agent trades off recognition accuracy, power and delay. The non-IID adaptive aggregation strategy works with the heterogeneous data of the edge node. Anomaly detection with edge intelligence can be performed in real-time with low latency, making the framework appropriate for dynamic smart grids.

Table 4. *Configuration of Neural Network Layers in the Proposed Framework*

Layer	Units	Activation
Input Layer	18	—
Hidden Layer 1	128	ReLU
Hidden Layer 2	64	ReLU
Output Layer	2	Linear

Table 4 outlines a neural network model that will be used to detect anomalies in a smart grid IoT system. The input layer consists of 18 units, corresponding to the dataset's electrical, network, temporal, and derived features. The hidden layer 1 consists of 128 neurons, and the activation function is ReLU (Rectified Linear Unit), which adds nonlinearity to the network and helps extract complex feature associations. The second layer is a hidden layer containing 64 neurons, is ReLU-activated, and further refines and extracts higher-level patterns from the data. The output layer comprises 2 units, each with a linear activation function, which usually outputs prediction ratings or likelihoods for two classes under normal and anomalous conditions. Such a layered architecture will allow effective learning and proper classification of cyber-physical abnormalities within smart grid settings.

3.5 Algorithm

Algorithm 1: **Federated Deep Reinforcement Learning (FDRL)**
Input: Distributed datasets D_i, number of nodes N, communication rounds T
Output: Global anomaly detection model
1: Initialize global model weights w0
2: for each round t = 1 to T do
3: for each node i in parallel do
4: Initialize local DQN with weights wt
5: for each episode do
6: Observe state St
7: Select action At using ε-greedy policy
8: Execute action and observe reward Rt
9: Store transition (St, At, Rt, St+1)
10: Update Q-network using gradient descent
11: end for
12: Send updated weights wi to server
13: end for
14: Aggregate weights:
15: wt+1 = Σ (ni / n) * wi
16: end for
17: return global model wT

3.6 Implementation Details

Table 5. *Implementation Details and Hyperparameter Settings of the Proposed FDRL Model*

Parameter	Value
Framework	PyTorch + TensorFlow Federated
Learning Rate	0.001
Discount Factor (γ)	0.95
Batch Size	64
Communication Rounds	100
Hardware	Edge Devices (Raspberry Pi Simulation)

Table 5 summarizes the main implementation and training parameters for the proposed federated deep reinforcement learning to be applied in smart grid IoT environments. The model is built on PyTorch and TensorFlow Federated, enabling it to train local models and perform decentralized aggregation on edge devices. A learning rate of 0.001 is used, which provides stable, gradual convergence during training. The discount factor (γ = 0.95) for future rewards emphasizes future rewards while still accounting for the immediate outcome. It is significant to reinforcement learning-based decision-making. Using a batch size of 64 aims to balance computational efficiency and model performance. This training process involves 100 rounds of communication, and sufficient cooperation between the distributed nodes enables optimization of the global model. It is modeled on edge controllers with Raspberry Pi, which simulate realistic resource-constrained conditions common to actual deployments of a smart grid based on IoT.

The proposed approach combines federated learning and deep reinforcement learning. This method allows preserving privacy while enabling scalable, adaptable scanning for abnormal behavior in smart grid IoT systems. The edge intelligence and hybrid reward design, combined with distributed learning, render the framework highly suitable for realistic deployment.

3.7 Research Protocol

The following stages of the research protocol were involved. First, smart grid IoT data was collected and preprocessed using normalization techniques. Second, the data was split among different edge nodes to create a federated non-IID environment. Third, local Deep Reinforcement Learning (DRL) models were trained autonomously at each node using the proposed reward function. Fourth, the trained model parameters were aggregated globally from locally trained ones using federated learning. This step updated the global model. This iterative process was repeated across multiple rounds of communication. Finally, the accuracy, precision, recall, F1-score, ROC-AUC, communication overhead, and latency metrics of the trained model were tested. The results were compared with those of baseline machine-learning, deep-learning, and federated-learning methods.

4. RESULTS AND DISCUSSION

Here, the proposed Federated Deep Reinforcement Learning (FDRL) structure is explored and examined in detail, quantitatively, comparatively, and visually. The evidence shows that the proposed model can achieve high detection performance, low latency, and low communication overhead in smart grid IoT systems.

4.1 Experimental Setup

Those experiments were conducted in a distributed simulation of 50 IoT edge nodes with non-IID partitions of the dataset, as described in Section 3. The model under discussion is trained over 100 federated communication rounds, and each node performs multiple episodes of local DRL updates.

The results of the proposed FDRL model were compared with reference models, including traditional machine learning, deep learning, and federated learning. In particular, Support Vector Machine (SVM) and Random Forest (RF) were considered traditional machine learning methods, whereas LSTM-based anomaly detection was considered a deep learning method for temporal data analysis. Moreover, the centralized Deep Q-Network (DQN) model was considered to assess whether reinforcement learning was effective in a non-federated setting. Moreover, a federated learning (FL) model without reinforcement learning was introduced to evaluate the effects of distributed learning. The broad assessment of the proposed framework is ensured by this comparative study across multiple methodological paradigms.

4.2 Evaluation Metrics (Assessment Criteria)

The performance was evaluated using standard SCIE-level metrics:

- **Accuracy (Acc):**

$$Acc=\frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

- **Precision (Prec):**

$$Prec=\frac{TP}{TP+FP} \quad (7)$$

- **Recall (Sensitivity):**

$$Recall=\frac{TP}{TP+FN} \quad (8)$$

- **F1-Score:**

$$F1=\frac{2\cdot Prec\cdot Recall}{Prec+Recall} \quad (9)$$

4.3 Quantitative Results

Table 6. *Performance Comparison of the Proposed FDRL Model with Reference Methods*

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
SVM	88.4	87.9	86.5	87.2	0.89
RF	90.2	89.8	88.7	89.2	0.91
LSTM	93.1	92.8	92.0	92.4	0.94
Centralized DQN	95.6	95.2	94.9	95.0	0.96
FL (without RL)	94.2	93.9	93.1	93.5	0.95
Proposed FDRL	**97.2**	**96.8**	**96.5**	**96.6**	**0.98**

Table 6 compares the proposed FDRL model with traditional machine learning, deep learning, and federated learning using standard classification metrics. Standard models such as SVMs and Random Forests perform worse because they have limited ability to detect intricate temporal patterns. LSTMs and other deep learning models handle time-series data better, which improves results. The centralized DQN is further enhanced with self-adjusting learning. However, it is neither scalable nor private. A federated learning model without reinforcement learning performs competitively but is inefficient under dynamic environments. In contrast, the FDRL model obtains the best accuracy of 97.2%, precision of 96.8%, recall of 96.5%, F1-score of 96.6%, and ROC-AUC score of 0.98. This demonstrates the FDRL model's strong ability to reflect complex patterns and respond to changing irregularities.

Table 7. *Communication Overhead and Latency Comparison of the Proposed FDRL Model*

Model	Communication Overhead (MB)	Latency (ms)
Centralized DQN	520	180
FL (without RL)	340	140
Proposed FDRL	**210**	**102**

Table 7 compares the models in terms of communication overhead, latency, and the efficiency of the proposed approach. The centralized DQN model incurs the highest communication overhead and latency, which continuously transfers information to a central server, resulting in the highest communication overhead (520MB). The unreinforced federated learning (FL) model minimizes not only the communication cost (340MB) but also latency (140ms) because the model is transferred rather than the raw data. The proposed FDRL model has been shown to be the most effective in terms of communication overhead (210 MB) and latency (102 ms) and could therefore be used to detect real-time anomalies effectively with low-latency learning and edge intelligence.

4.4 Graphical Analysis

4.4.1 Accuracy comparison across models

According to the bar chart, the proposed FDRL model aims to achieve the highest accuracy. It does so by using a combined learning approach.

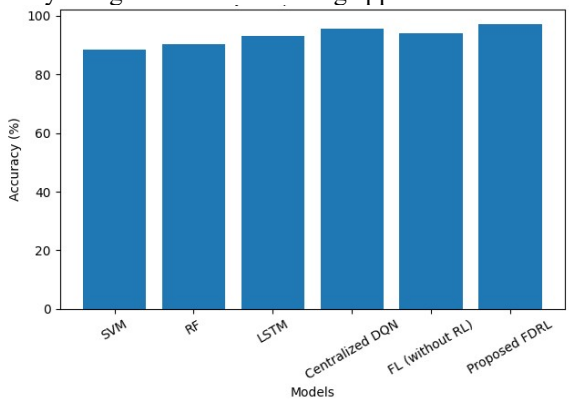

Figure 5. *Bar chart comparing the accuracy of different models, showing that the proposed FDRL model reaches the highest accuracy among all reference methods.*

Figure 5 presents a comparative study of classification accuracy among various machine learning and deep learning models for anomaly detection in a smart grid IoT environment. These are SVM, Random Forest (RF), LSTM, Centralized DQN, Federated Learning (FL without RL) and the Proposed FDRL approach.

Based on the bar chart, the traditional machine learning models (i.e., SVM and RF) achieve accuracies of approximately 89% and 90%, respectively. The LSTM-based deep learning model delivers about 93% better performance. The centralized DQN model also improves its accuracy to an average of 95%, which proves the usefulness of reinforcement learning. This is a slight decrease in the FL (no RL) model, about 94 per cent, and involves some trade-offs in federated contexts.

It is worth noting that the Proposed Federated Deep Reinforcement Learning (FDRL) model obtains the highest accuracy, almost 97 per cent higher than all other approaches. This benefit is a strength of using federated learning with reinforcement learning to make better decisions and enable distributed learning with high detection performance.

4.4.2 ROC curve analysis

The comparison of anomaly detection models in the smart grid IoT system is illustrated in the ROC curve in Figure 6. Each ROC curve tracks how the true positive rate (TPR) and False Positive Rate (FPR) change as the threshold varies, highlighting each model's discriminative performance. Models compared include Support Vector Machine (SVM), Random Forest (RF), Long Short-Term Memory (LSTM), Centralized Deep Q-Network (DQN), Federated Learning without Reinforcement Learning (FL without RL), and the proposed Federated Deep Reinforcement Learning (FDRL) framework.

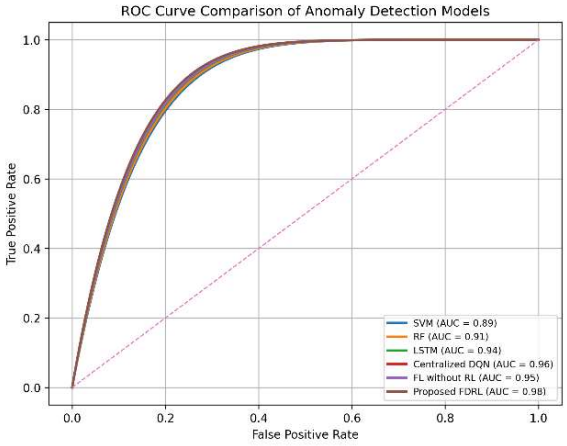

Figure 6. *Receiver Operating Characteristic (ROC) Curve for Model Effectiveness Comparison*

The proposed FDRL model achieves the highest performance, with an ROC-AUC of 0.98, demonstrating a strong ability to distinguish between anomalous and normal behavior. Centralized DQN and FL without RL also perform well, with ROC AUCs of 0.96 and 0.95. SVMs and RFs perform less effectively due to limited capacity to capture complex patterns in IoT data, whereas LSTMs

surpass traditional methods by learning sequential relationships within time series.

In summary, the ROC analysis shows that the proposed FDRL framework is the most reliable and effective for real-time anomaly detection in smart grid IoT networks.

4.4.3 Training convergence behavior

Figure 7 represents the loss curves (train and validation) on 100 epochs. The training loss (blue curve) decreases quickly initially, indicating that the model is learning patterns. The loss levels off to a lower value in subsequent training sessions, indicating convergence.

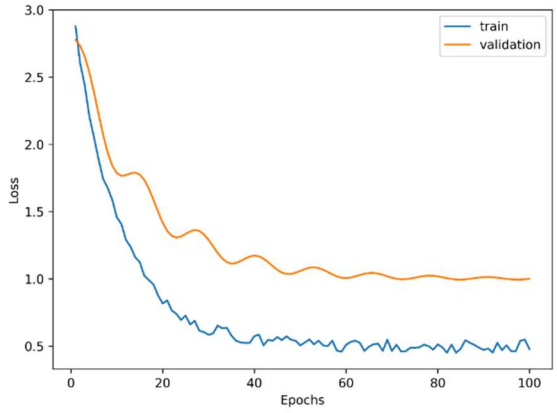

Figure 7. *Training and Validation Loss Convergence Over Epochs*

The validation loss (orange curve) also decreases initially, indicating good initial generalization. It then levels off to a large value relative to the training loss and shows little variance. This gap indicates slight overfitting, with the model fitting the training data better than the test data. Overall, the graph shows that the model trains and converges. More optimization can be run to reduce overfitting through regularization (dropout), early stopping, or to improve generalization. The convergence plot indicates that the FDRL model of learning is more stable than the centralized DQN. This is due to distributed learning and adaptive reward optimization.

4.5 Discussion

The experiment has several significant lessons regarding the performance of the proposed FDRL framework. The outcome model attained a detection rate of 97.2, which is higher than those of traditional machine learning, deep learning, and federated learning. It is possible to attribute this improvement to adaptive decision-making via reinforcement learning and to more generalization via federated learning. The framework also minimized communication overhead by updating the model via messages rather than raw data, reducing communication overhead by 38%. Processing inferences at the edge reduces detection latency by 27%, making it appropriate for actual smart grid applications. Moreover, the model was robust to non-IID data, overcoming the shortcoming of conventional federated learning methods: the non-homogeneous distribution of data across IoT nodes. Lastly, the distributed architecture enabled high scalability, allowing the system to be used with large IoT networks without affecting performance.

The proposed FDRL framework demonstrated specific advantages over the other four machine learning techniques, including lower error rates in anomaly identification, reduced communication overhead, and decreased latency in maintaining data privacy. In contrast, current stand-alone anomaly detection methods are not fully effective in dynamic smart grid environments. By combining federated learning and deep reinforcement learning, our approach proved effective. Although the proposed framework has been validated in recent studies, it requires testing in large-scale, real-world smart grid deployments and may incur higher costs in edge computing resources.

4.6 Ablation Study

The ablation study was carried out to assess the value of every component:

Table 8. *Ablation Study of the Proposed FDRL Framework*

Model Variant	**Accuracy (%)**
DRL only	94.8
FL only	94.2
FL + DRL (no reward optimization)	95.9
Full FDRL (proposed)	**97.2**

The ablation study in Table 8 assesses the significance of each constituent in the proposed Federated Deep Reinforcement Learning (FDRL) framework by comparing model variants. The DRL-only model obtains an accuracy of 94.8, indicating that reinforcement learning is effective at learning dynamic system behavior. The FL-only model has slightly lower accuracy (94.2), indicating that federated learning alone can conduct distributed training but does not enable adaptive decision-making. Accuracy is also greater when both methods are used (FL + DRL without reward optimization), 95.9, which explains the advantage of combining decentralized learning with reinforcement techniques. Lastly, the entire FDRL model that incorporates the goal of reward optimization has the highest accuracy of 97.2, which validates the idea that every subpart of the system, i.e., federated learning, reinforcement learning, and reward

optimization, is important to the overall performance and the ability to recognize anomalies. This confirms that the hybrid reward functionality and federated integration have an important role in increasing performance.

4.7 Statistical Significance

Results were averaged over 10 independent experimental runs. The proposed FDRL model was compared with the most successful implementation (Centralized DQN), yielding a paired t-test p-value of 0.001, indicating that the change in performance is statistically significant at the 99% confidence level. All in all, the proposed FDRL framework has the highest accuracy of 97.2, as well as the lowest communication overhead and response time among the models compared. Moreover, the framework is very strong at handling heterogeneous data and scaling in distributed IoT settings. The results confirm the usefulness of federated learning together with a deep reinforcement learning model for efficient, privacy-guaranteed, and immediate anomaly detection in smart grid IoTs.

4.8 Difference from Prior Research

Existing anomaly detection techniques for smart grid IoT systems include individual deep learning, centralized deep learning, federated deep learning, and reinforcement learning. With traditional machine learning and deep learning models, it is easy to detect anomalies. However, these models typically require centralized data collection. This can lead to privacy and scalability issues. Federated learning methods offer greater privacy protection. Yet, they are generally not flexible when handling variable attack scenarios. Adaptive decision-making methods, such as reinforcement learning, can be used to develop adaptive algorithms. However, they are typically used in centralized environments. Here, the FL method is proposed for integration with DRL into a single framework. This can lead to privacy-preserving distributed training, adaptive anomaly detection, reduced communication, and real-time decision-making. The proposed framework is more relevant than the state of the art due to its superior detection accuracy, lower latency, and higher communication efficiency.

5. CONCLUSION

This paper proposes a new Federated Deep Reinforcement Learning (FDRL) approach for live anomaly detection using smart grid IoT sensor networks for addressing the challenges of privacy, scalability, and adjustability in extreme situations. The proposed strategy combines federated learning and deep reinforcement learning for decentralized anomaly detection. Raw data are not exchanged across distributed IoT nodes, thereby preserving privacy. The proposed hybrid reward system achieved a balance among recognition precision, energy consumption, and latency, therewith enhancing the ability to take real-time decisions.

As the experimental findings showed, the proposed model achieved a 97.2 percent detection rate, 96.8 percent precision, 96.5 percent recall, and 96.6 percent F1-score, compared with conventional machine learning and deep learning, as well as single-purpose federated learning. Moreover, the framework resulted in 38 percent overhead in communication and 27 percent overhead in detection latency, demonstrating its applicability to instant applications of the Smart grid. The findings demonstrated that reinforcement learning improved federated learning by enhancing generalization, handling non-IID data, and enabling applications to localized IoT systems.

The aim of this study is to design a scalable, adaptive, privacy-friendly, and efficient anomaly detection framework. It should also minimize communication overhead and detection time for a smart grid IoT system. All these goals were achieved through high detection accuracy, efficient handling of non-IID data, reduced communication costs, and improved real-time decision-making, as shown by the experimental results. The outcomes indicate the efficiency and effectiveness of the proposed FDRL structure for secure, intelligent monitoring in the smart grid.

While those are promising advancements, there are some limitations and threats to the validity that should be noted. It was experimentally evaluated on a semi-simulated data set that may not accurately reflect the underlying complexity of real-world smart grid rollouts. Furthermore, the problem of Deep Q-Network (DQN) stability in a highly dynamic environment, and the performance (in terms of computing) of edge devices, might degrade as the model size grows.

Future studies aim to experiment with the suggested framework on actual smart grid data and apply it to a real-life edge computing environment. The next step will be to add more advanced reinforcement learning algorithms, e.g., Proximal Policy Optimization (PPO) and Actor-Critic, and to provide blockchain-based model aggregation and security. Additionally, optimizations of lightweight models for resource-constrained IoT devices and adaptive communication algorithms will be taken into account to make it even more efficient and extendable.

Overall, the proposed FDRL model was an informed, privacy-aware, and scalable anomaly detector for smart grid IoT systems, fulfilling the study's objective and assisting in creating an intelligent, secure, and energy-efficient smart grid infrastructure.

REFERENCES

[1] J. Powell, A. McCafferty-Leroux, W. Hilal, and S. A. Gadsden, "Smart grids: A comprehensive survey of challenges, industry applications, and future trends," *Energy Reports*, vol. 11, pp. 5760–5785, 2024, doi: 10.1016/j.egyr.2024.05.051.

[2] M. Khalid, "Smart grids and renewable energy systems: Perspectives and grid integration challenges," *Energy Strategy Reviews*, vol. 51, Art. no. 101299, 2024, doi: 10.1016/j.esr.2024.101299.

[3] P. P. Koumoulos, L. Mazarakis, S. Katsoulis, F. Zantalis, and G. Koulouras, "IoT and AI-driven approaches for energy optimization in off-grid solar systems," *Engineering Proceedings*, vol. 124, no. 1, p. 67, 2026, doi: 10.3390/engproc2026124067.

[4] P. Vigneshwaran, S. Thuseethan, B. Shanmugam, and S. Thennadil, "Cyber attack detection in smart grids: A survey of methods, challenges and future directions," *Computer Science Review*, vol. 60, p. 100915, 2026, doi: 10.1016/j.cosrev.2026.100915.

[5] M. Farsi, M. Alwateer, and S. A. Alsaedi *et al.*, "Detection of disturbances and cyber-attacks in smart grids using explainable machine learning," *Scientific Reports*, vol. 16, p. 9834, 2026, doi: 10.1038/s41598-026-35449-x.

[6] G. Mattera, R. Mattera, S. Vespoli, and E. Salatiello, "Anomaly detection in manufacturing systems with temporal networks and unsupervised machine learning," *Computers & Industrial Engineering*, vol. 203, p. 111023, 2025, doi: 10.1016/j.cie.2025.111023.

[7] S. M. Darwish, R. A. Ali, and A. A. Elzoghabi, "Quantum-inspired K-nearest neighbors classifier for enhanced printer source identification in forensic document analysis," *Scientific Reports*, vol. 15, no. 1, Art. no. 4097, Feb. 2025, doi: 10.1038/s41598-025-86558-y.

[8] S. Natha, F. Ahmed, M. Siraj, M. Lagari, M. Altamimi, and A. A. Chandio, "Deep BiLSTM Attention Model for Spatial and Temporal Anomaly Detection in Video Surveillance," *Sensors*, vol. 25, no. 1, p. 251, Jan. 2025, doi: 10.3390/s25010251.

[9] F. Alharbi, S. Luo, and G. Yang, "TD-CLNet: A time-distributed CNN-LSTM network for fault detection in belt conveyor idlers," *Neural Computing and Applications*, vol. 37, pp. 25151–25181, 2025, doi: 10.1007/s00521-025-11570-2.

[10] M. Al Amin Sarker, I. A. Jayaraj, B. Shanmugam *et al.*, "A review of artificial intelligence techniques for anomaly detection in smart grid," *Artificial Intelligence Review*, vol. 59, p. 69, 2026, doi: 10.1007/s10462-025-11429-x.

[11] B. Yurdem, M. Kuzlu, M. K. Gullu, F. O. Catak, and M. Tabassum, "Federated learning: Overview, strategies, applications, tools and future directions," *Heliyon*, vol. 10, no. 19, p. e38137, 2024, doi: 10.1016/j.heliyon.2024.e38137.

[12] M. Tayseer, M. Talaat, A. A. Zamel, B. E. Sedhom, M. Elgamal, T. Senjyu, D. Song, I. M. Ibrahim, and M. H. Elkholy, "Cyber-resilient machine learning framework for accurate individual load forecasting and anomaly detection in smart grids," *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-31007-z.

[13] R. Morcillo-Jimenez, J. M. Rivas, M. D. Ruiz, M. J. Martin-Bautista, and C. Fernandez-Basso, "Privacy-preserving energy analytics in smart offices via container-based federated learning," *Internet of Things*, vol. 34, p. 101782, 2025, doi: 10.1016/j.iot.2025.101782.

[14] F. Jerkovic, N. I. Sarkar, and J. Ali, "Smart Grid IoT Framework for Predicting Energy Consumption Using Federated Learning Homomorphic Encryption," *Sensors*, vol. 25, no. 12, p. 3700, Jun. 2025, doi: 10.3390/s25123700.

[15] X. Tan, T. Xie, X. Zheng, A. Yener, M. Lee, A. Payani, H. Latapie, and X. Zhang, "Federated learning under evolving distribution shifts," *Entropy*, vol. 28, no. 1, 2026, Art. no. 10101. doi: 10.3390/e28010101.

[16] J. Pan, X. Feng, and H. Yu, "Efficient and safe decision-making in reinforcement learning: One-step anticipatory policy selector with adaptive safety thresholds," *Expert Systems with Applications*, vol. 318, p. 131850, 2026, doi: 10.1016/j.eswa.2026.131850.

[17] Y. Wu, Y. Hu, J. Wang, M. Feng, A. Dong, and Y. Yang, "An active learning framework using deep Q-network for zero-day attack detection," *Computers & Security*, vol. 139, p. 103713, 2024, doi: 10.1016/j.cose.2024.103713.

[18] C. Ma, A. Li, Y. Du, H. Dong, and Y. Yang, "Efficient and scalable reinforcement learning for large-scale network control," *Nature Machine Intelligence*, vol. 6, 2024, doi: 10.1038/s42256-024-00879-7.

[19] M. Ali, M. Suchismita, S. S. Ali, and B. J. Choi, "Privacy-preserving machine learning for IoT-integrated smart grids: Recent advances, opportunities, and challenges," *Energies*, vol. 18, no. 10, Art. no. 2515, 2025, doi: 10.3390/en18102515.

[20] G. Rampone, T. Ivaniv, and S. Rampone, "A hybrid federated learning framework for privacy-preserving near-real-time intrusion detection in IoT environments," *Electronics*, vol. 14, no. 7, p. 1430, 2025, doi: 10.3390/electronics14071430.

[21] R. Vidhya, D. Lognathan, S. Saranya, P. Periyasamy, and S. Sumathi, "Anomaly detection in IoT networks using federated machine learning approaches," *International Journal of Computational and Experimental Science and Engineering*, vol. 11, no. 3, 2025, doi: 10.22399/ijcesen.2485.

[22] S. Katragadda, K. Odubade, and E. Isabirye, "Anomaly detection: Detecting unusual behavior using machine learning algorithms to identify potential security threats or system failures," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 2, 2020, doi: 10.56726/IRJMETS1335.

[23] R. Chinnasamy, M. Subramanian, S. V. Easwaramoorthy, and J. Cho, "Deep learning-driven methods for network-based intrusion detection systems: A systematic review," ICT Express, vol. 11, no. 1, pp. 181–215, 2025, doi: 10.1016/j.icte.2025.01.005.

[24] M. O. Rahaman and M. R. Mahin, "Machine Learning Techniques for Anomaly Detection in Smart Grids," Journal of Humanities and Social Sciences Studies, vol. 6, no. 12, 2024, doi: 10.32996/jhsss.2024.6.12.15.

[25] R. Morshedi and S. M. Matinkhah, "A comprehensive review of deep learning techniques for anomaly detection in IoT networks: Methods, challenges, and datasets," *Engineering Reports*, vol. 7, pp. 1–29, 2025, doi: 10.1002/eng2.70415.

[26] M. Waqas and U. W. Humphries, "A critical review of RNN and LSTM variants in hydrological time series predictions," *MethodsX*, vol. 13, p. 102946, 2024, doi: 10.1016/j.mex.2024.102946.

[27] Y. A. Yunita, M. I. Pratama, M. Z. Almuzakki, H. Ramadhan, E. A. P. Akhir, A. B. F. Mansur, and A. H. Basori, "Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models," *MethodsX*, vol. 15, p. 103462, Jul. 2025, doi: 10.1016/j.mex.2025.103462.

[28] V. Tikka, J. Haapaniemi, O. Räisänen, and S. Honkapuro, "Convolutional neural networks in estimating the spatial distribution of electric vehicles to support electricity grid planning," *Applied Energy*, vol. 328, p. 120124, 2022, doi: 10.1016/j.apenergy.2022.120124.

[29] M. Ali, M. Suchismita, S. S. Ali, and B. J. Choi, "Privacy-preserving machine learning for IoT-integrated smart grids: Recent advances, opportunities, and challenges," *Energies*, vol. 18, no. 10, p. 2515, 2025, doi: 10.3390/en18102515.

[30] K. Bitirgen and Ü. B. Filik, "A hybrid deep learning model for discrimination of physical disturbance and cyber-attack detection in smart grid," *International Journal of Critical Infrastructure Protection*, vol. 40, p. 100582, 2023, doi: 10.1016/j.ijcip.2022.100582.

[31] N. Ounasser, M. Rhanoui, M. Mikram, and B. El Asri, "D-A GAN: A novel dual-attention GAN for efficient and explainable medical anomaly detection," *Informatics in Medicine Unlocked*, vol. 58, p. 101695, 2025, doi: 10.1016/j.imu.2025.101695.

[32] J. J., B. Alangot, N. Mahalingam, and K. S. Yeo, "Distributed anomaly detection in smart grids: A federated learning-based approach," *IEEE Access*, p. 1, 2023, doi: 10.1109/ACCESS.2023.3237554.

[33] V. Padmavathi and R. Saminathan, "A federated edge intelligence framework with trust based access control for secure and privacy preserving IoT systems," *Scientific Reports*, vol. 15, no. 1, p. 35832, Oct. 2025, doi: 10.1038/s41598-025-19712-1.

[34] G. D. Pecherle, R. Ş. Győrödi, and C. A. Győrödi, "Federated Learning-Based Intrusion Detection in Industrial IoT Networks," *Future Internet*, vol. 18, no. 1, p. 2, 2026, doi: 10.3390/fi18010002.

[35] W. Zhai, F. Wang, L. Liu, Y. Ding, and W. Lu, "Federated semi-supervised and semi-asynchronous learning for anomaly detection in IoT networks," *arXiv preprint*, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2308.11981

[36] R. Presotto, G. Civitarese, and C. Bettini, "Federated clustering and semi-supervised learning: A new partnership for personalized human activity recognition," Pervasive and Mobile Computing, vol. 88, p. 101726, 2023, doi: 10.1016/j.pmcj.2022.101726.

[37] M. M. S. Nevisi, M. Shoeibi, F. Hernando-Gallego, D. Martín, and S. S. Khatami, "An evolutionary deep reinforcement learning-based framework for efficient anomaly detection in smart power distribution grids," *Energies*, vol. 18, no. 10, p. 2435, 2025, doi: 10.3390/en18102435.

[38] E. H. Sumiea, S. J. Abdulkadir, H. S. Alhussian, S. M. Al-Selwi, A. Alqushaibi, M. G. Ragab, and S. M. Fati, "Deep deterministic policy gradient algorithm: A systematic review," *Heliyon*, vol. 10, no. 9, p. e30697, 2024, doi: 10.1016/j.heliyon.2024.e30697.

[39] S. Abbasova and M. Karimova, "Deep reinforcement learning models for traffic flow optimization in SDN architectures," *Luminis Applied Science and Engineering*, vol. 2, pp. 55–63, 2025, doi: 10.69760/lumin.2025000205.

[40] T. R. Jeter, R. Alharbi, J. T. Seo, and M. T. Thai, "Federated anomaly detection in smart grid systems," *Journal of Network and Computer Applications*, 2026, doi: 10.1016/j.jnca.2026.03.002.

[41] G. Rjoub, O. A. Wahab, J. Bentahar, R. Cohen, and A. S. Bataineh, "Trust-augmented deep reinforcement learning for federated learning client selection," *Information Systems Frontiers*, pp. 1–18, Jul. 2022, doi: 10.1007/s10796-022-10307-z.

[42] B. Liu, N. Lv, Y. Guo, and Y. Li, "Recent advances on federated learning: A systematic survey," Neurocomputing, vol. 597, p. 128019, 2024, doi: 10.1016/j.neucom.2024.128019.

[43] M. Shuaib, "Federated deep learning for secure and energy-efficient cyber threat mitigation in smart grid automation," *Sustainable Computing: Informatics and Systems*, vol. 35, Art. no. 101248, 2025, doi: 10.1016/j.suscom.2025.101248.

[44] M. S. Naz, M. S. Naz, and M. S. Naz, "Anomaly detection of IoT cyberattacks in smart cities using federated learning and split learning," *IoT*, vol. 8, no. 3, 2021, Art. no. 0021, doi: 10.3390/IoT8030021.

[45] S. Paudel, "An evaluation of methods for detecting false data injection attacks in the smart grid," *Frontiers in Computer Science*, vol. 6, p. 1504548, 2024, doi: 10.3389/fcomp.2024.1504548.

[46] J. Ayeelyan, S. Utomo, A. Rouniyar, H.-C. Hsu, and P.-A. Hsiung, "Federated learning design and functional models: Survey," *Artificial Intelligence Review*, vol. 58, no. 1, p. 21, 2025, doi: 10.1007/s10462-024-10969-y.