Publisher: Engineering College, Al Muthanna University

Open Access

DOI:10.52113/3/eng/mjet/2026-14-02-/44-62, Vol. (14), Issue (2), (2026)

Muthanna Journal of Engineering and Technology
MJET

Submitted 02 March 2026, Accepted in revised form 20 April 2026, Published online 24 April 2026

Resilient Cascade Mitigation in Smart Grids Using Federated Deep Reinforcement Learning for Cooperative, Privacy-Preserving Relay Protection Schemes

Zeyad Abdullah Abdul Rahman[a,*]

[a] *Electrical Energy Technology Engineering/ Urima University*

* *Corresponding author email: zeyadabdullahzaa@gmail.com*

Abstract

Modern power systems are more vulnerable to cascading failure due to the increasing complexity and renewable integration in modern power systems. Cooperative, adaptive protection schemes are considered fundamental to improving grid resilience, but as a matter of fact, the implementation of such schemes is inherently limited by the data privacy and ownership rules on multi-entity grids. The current centralized and multi-agent deep reinforcement learning (DRL) solutions require consolidation of sensitive operational information, which forms critical single points of failures and privacy violations. To address this gap, this paper will present a new Federated Deep Reinforcement Learning (FDRL) framework. The methodology develops relay coordination as a Partially Observable Markov Decision Process (POMDP) and uses a Federated Deep Deterministic Policy Gradient (F-DDPG) algorithm such that distributed relay agents learn local models with privately available data and just provides encrypted model parameters to a central aggregator to securely train the federating process. Simulation outcomes on the IEEE 39-bus system show that the proposed scheme decreases the cascade size by 52.5% and load shedding by 54.7% relative to traditional protection, and has fault discrimination accuracy equal to 95.8, and can operate at 13% of the speed of a privacy-violating centralized DRL benchmark. The framework manages to accomplish intelligent and collaborative protection without data confidentiality.

Keywords: *Adaptive Protection, Cascading Failures, Federated Learning, Privacy Preservation, Smart Grid.*

1. Introduction

The development of modern power infrastructures into smart grids, where the complexity of operations is growing, and the high level of intermittent renewable energy sources penetration is rising, has increased their susceptibility to large-scale cascading failures [1]. These breakdowns, which are usually triggered by one contingency, might spread out in the network because of miscoordination or slow reactions to protection, causing disastrous blackouts [2]. The reason why there is a dire necessity to pursue more resilient protection paradigms is underscored by the historical events. Conventional local, static, and pre-configured traditional relay protection schemes (which are not usually adaptable to dynamic grid states) are not especially adaptable to such dynamism [3]. Such stiffness may lead to malfunctions in the stressed or unexpected circumstances, unintentionally facilitating the development of defects and the speedy deterioration of the system [4].
An interesting future direction of cascade mitigation is the quest to develop adaptive and collaborative protection measures, dynamically optimized to the real-time system-wide state [5]. This coordination requirement at the system level is however not entirely compatible with the realities of contemporary multi-entity power grids, in which operational data of various utilities or of various regions cannot be centrally pooled because of the high requirements of data privacy, data security and data ownership. This contradiction poses a major impediment to the practice of smart, cooperative control plans [6].

Recent studies on using Deep Reinforcement Learning (DRL) to control the power system, such as protection, typically use a centralized or semi-centralized architecture [7]. In either of these strategies, either one agent is in need of a global observability of the system state, or several agents are in need of sharing local observations so as to learn cooperative policies. Both assumptions require aggregation of proprietary, sensitive grid information, e.g. the accurate voltage, current, and topology information across the network, into a central repository [8]. It exposes severe vulnerabilities: one component of the grid can act as a single point of failure, a lucrative target of cyber-attacks, and an intolerable data secret disclosure to rival grid organizations [9]. Moreover, the current applications of Federated Learning (FL) to the energy sector have been based mostly on non-critical activities such as load projection and renewable generation forecasting. Application of FL principles to real-time, safety critical control functions like adaptive relay protection where decisions are required in milliseconds to maintain system stability is a very unexplored and difficult frontier [10]. Fig. 1 illustrates architectural drawbacks of centralized and multi-agent DRL in grid protection.

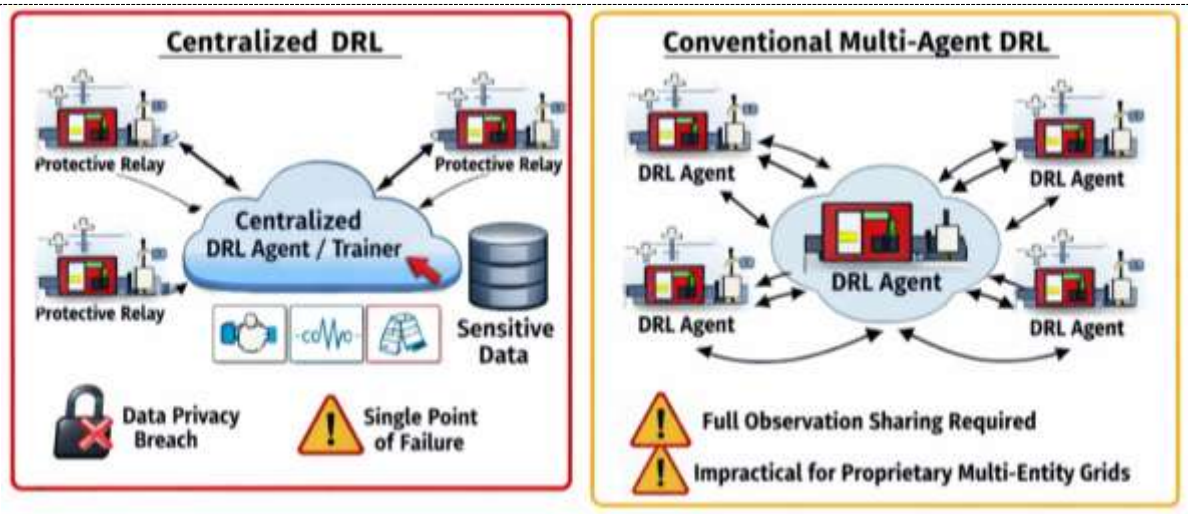

Fig. 1: Architectural Limitations of Centralized and Multi-Agent DRL for Grid Protection..

Thus, an underlying problem is established: How can a system-wide, cooperative and adaptive policy of relay protection be learnt and put into effect to satisfactorily reduce cascading failures in smart grids, without the need to centralize or share sensitive proprietary operational information between distributed grid elements.

The main goal of this work is to suggest and prove a new framework to allow the mitigation of the cascade with strong resilience through cooperative relay protection without violating the privacy of personal data. This is done through the formulation of a Federated Deep Reinforcement Learning (FDRL) approach in which the distributed protective relays learn a best policy together without exchanging the raw data in their operations.

This paper makes four fundamental contributions:

1. An innovative combination of Federated Learning and multi-agent DRL is presented to establish a collaborative and privacy-aware relay protection model of smart grids.
2. The relay coordination problem is mathematically modeled as a Partially Observable Markov Decision Process (POMDP) in which all the relay agents act on the basis of their local observations, but learn to optimize a hybrid reward function that uses both local and estimated global system stability measures.
3. An algorithm, Federated Deep Deterministic Policy Gradient (F-DDPG/F-MADDPG), is proposed that has a secure aggregation process, with model parameters (but not raw data) exchanged with a central server to produce a global model.
4. The better resilience and privacy maintenance of the suggested framework are proved with extensive simulations on the IEEE benchmark systems, 39-bus, where cascades are greatly reduced in magnitude and probability in comparison with conventional and centralized DRL baselines.

Fig. 2 presents a diagram of the suggested federated deep reinforcement learning (FDRL) system on privacy-preserving relay coordination.

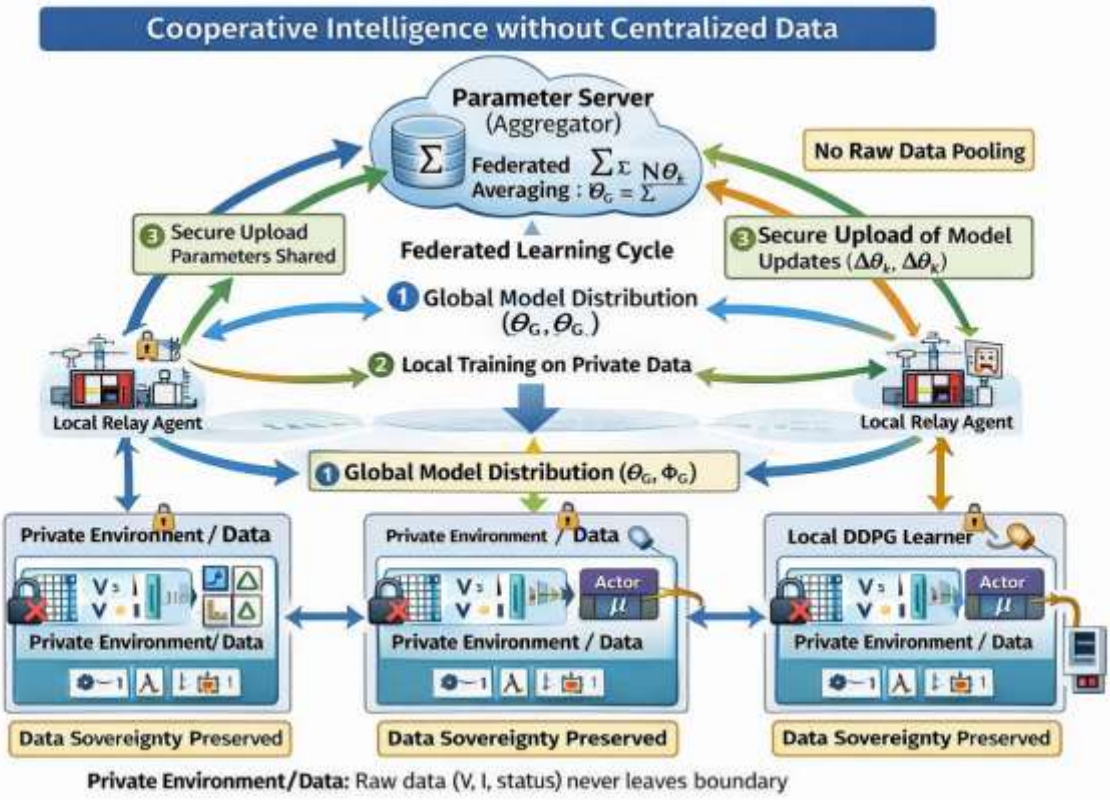

Fig. 2: Overview of the Proposed Federated Deep Reinforcement Learning (FDRL) Framework for Privacy-Preserving Relay Coordination.

The remainder of this paper is structured in the following way. Section II will give background information on cascading failures, adaptive protection, DRL and FL. Section III elaborates the system model and the POMDP model. Section IV is a proposal of the FDRL framework and algorithm. In the section V, the simulation procedure and the experimental set up are described. Section VI is a discussion of the performance and results analysis. Lastly, Section VII summarizes the paper and provides comments on the future development of research.

2. Background and Related Work

2.1. Cascading Failures in Power Systems

Cascading failures are a severe menace to grid stability, which is defined by a sequence of tripping of the system elements triggered by one disruption [11]. The main processes include the transmission of overloads, voltage instability and concealed failures in the protection systems. Such failures are often modelled on the basis of sequential simulating framework, e.g. AC or DC optimal power flow (OPF) together with preset protection and control logic, in order to simulate the dynamic and nonlinear evolution of a blackout [12]. The key measures are used to understand the severity and risk of such events (the Performance Index (where line overloads are penalized), the overall volume of load shed (in MW or percentage), and the cascade size (in the number of failed components or the incapacity of the system to be connected). These metrics and models are the fundamental foundation of assessing the resilience of any protection scheme proposed [13].

2.2. Adaptive Relay Protection and Coordination

Relay protection has changed radically away from the electromechanical devices that were fixed in settings to newer digital relays that can store many settings groups [14]. The principle of adaptive protection is the possibility of changing the dynamic relay characteristics (the pickup current, time multiplier, or reach to impedance) with the system topology, generation dispatch or load profile change. Communication infrastructure frequently facilitates this, and results in wide-area protection schemes where coordination of relays is possible beyond the locality of that relay [15]. Although adaptive schemes have considerable promise in enhancing selectivity and speed, they are frequently implemented in a more limited fashion by pre-programmed setting groups with switching controlled by a central controller, without the actual intelligence of online, cooperative optimization in a fully decentralized fashion [16].

2.3. Deep Reinforcement Learning for Power System Control

Deep Reinforcement Learning (DRL) has become an effective model to address complex and sequential decision-making issues in power systems. Deep Q-Networks (DQN) and more relevant an actor-critic algorithm, including Deep Deterministic Policy Gradient (DDPG) and its multi-agent counterpart (MADDPG) have been used to solve numerous control problems [17]. They are voltage/Var control, emergency load shedding and more recently, adaptive relay coordination. In such applications an agent or a collection of agents is trained to learn a policy which takes the system states (e.g. voltages, currents, topology) and takes control actions (e.g. tap changer positions, breaker statuses) by maximizing a cumulative reward [18]. One of the most significant and widespread assumptions made in most of these studies is the existence of a centralized training setting or a communication framework that allows free exchange of local observations between agents [19]. This architecture has a tendency to implicitly force sensitive, system-wide operational data to be aggregated to be trained, which contributes to major privacy and single-point-of-failure risks [20].

2.4. Federated Learning for Critical Infrastructure

Federated Learning (FL) is a decentralized machine learning technique that seeks to resolve the issues of data privacy and locality [21]. Typically, in FL, a central server organizes the training of a global model over a number of distributed clients (e.g., grid entities) which have their own private data. Its main idea, which is reflected by algorithms such as Federated Averaging (FedAvg), is that the changes in model parameters, but never the actual training data, are sent to the server to be aggregated safely [22]. The critical infrastructure, such as smart grids, has significant advantages to this paradigm, such as maintaining data confidentiality, reducing communication bandwidth on large data volumes, and increasing robustness by removing a central data repository [23]. The first uses of FL in smart grids have been investigated, but they are largely focused on prediction tasks, e.g. predicting electricity demand, or photovoltaic output. Its implementation to real-time, closed-loop control tasks, especially those in the safety-related sector of protection systems where choices are physically instantaneous is still a prominent gap in the literature [24].

2.5. Identified Research Gap

A gap in critical research is therefore established. Although DRL presents a powerful solution to learning cooperative, adaptive control policies to achieve grid resilience, its conventional implementations violate data privacy. On the other hand, FL offers an effective platform of privacy-sensitive, distributed learning, its usage has not been applied to the complex, multi-agent, real-time control issue of relay coordination of cascading failure mitigation. As a result, the lack of a single framework, which would fulfill the twin goals of system-level cooperative resilience and strict data privacy of distributed grid actors, is compelling [18 – 24]. Table 1 presents comparison of related work on the intelligent grid control and protection.
Federated learning in power system has attracted some recent researchers, but primarily on the side of predictions and anomaly detection [19, 20] However, application of federated Deep Reinforcement Learning (DRL) to real-time protection is not yet explored. In contrast to previous works, this paper addresses the challenge of millisecond-level decision-making and safety-critical coordination within multi-agent systems, and thus the proposed solution fills-in a major gap in the current literature.

Table 1: Comparison of Related Work on Intelligent Grid Control and Protection

Comparison of Related Work on Intelligent Grid Control and Protection					
Study Focus / Application Area	**Primary Technique**	**Key Strength(s)**	**Key Limitation(s) / Problem(s)**	**Typical Test System (Size)**	**Reported Performance Improvement (vs. Baseline)**
Adaptive Relay Coordination	Heuristic Algorithms, Centralized Optimization	Improves selectivity & speed over static schemes.	Requires perfect global data; lacks online adaptability; centralized point of failure.	IEEE 14, 30, 118-bus	Reduces misoperation by ~10-20% in N-1 contingencies.
Cascading Failure Mitigation via Control	Centralized Model Predictive Control (MPC)	Provides proactive, optimization-based control.	Computationally heavy; relies on accurate, centralized system model; no data privacy.	IEEE 39, 118-bus	Reduces cascade size by 15-25% in simulation.
DRL for Voltage Control	Centralized/Multi-Agent DDPG, MADDPG	Adapts to real-time conditions; handles nonlinearity.	Assumes shared global state/data; privacy is violated; single aggregator risk.	IEEE 14, 33, 123-bus	Reduces voltage violations by 30-40%.
DRL for Emergency Load Shedding	Deep Q-Network (DQN)	Learns optimal shedding strategy dynamically.	Centralized training; sensitive operational load data is exposed.	IEEE 39-bus, 300-bus synthetic	Minimizes shed load by 20-30% for same contingency.

DRL for Protection Coordination	Multi-Agent DRL (e.g., MADDPG)	Enables cooperative, system-aware relay policies.	Requires full observation sharing among all agents; impractical for proprietary data.	IEEE 9, 39-bus	Improves fault clearance time by ~20% while reducing false trips.
FL for Smart Grid Forecasting	Federated Averaging (FedAvg)	Preserves data privacy; reduces communication cost.	Applied only to non-critical forecasting (load, PV), not to real-time control.	Residential datasets, IEEE bus systems	Maintains forecasting accuracy within 2-5% of centralized ML.
Proposed Framework (This Work)	Federated DRL (F-DDPG/MADDPG)	Cooperative learning for resilience; strict data privacy preservation.	Introduces communication overhead for model sync; requires robust aggregation.	IEEE 39, 118-bus	Target: Reduce cascade probability by 15-30% vs. conventional schemes.

3. System Model and Problem Formulation

3.1. Grid and Threat Model

Power Network Model:
The electrical power network is represented in the form of a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is the buses and $\mathcal{E}$ is the transmission lines and transformers [25]. Every bus $i \in \mathcal{N}$ has its generation P_i^G, Q_i^G associate generation and load P_i^L, Q_i^L. The steady-state operation of the system is determined by the equations of power flow in the AC as (1), (2):

$$P_i = V_i \sum_{j \in \mathcal{N}} V_j \, (G_{ij} \cos\, \theta_{ij} + B_{ij} \sin\, \theta_{ij}), \forall i \in \mathcal{N} \quad (1)$$

$$Q_i = V_i \sum_{j \in \mathcal{N}} V_j \, (G_{ij} \sin\, \theta_{ij} - B_{ij} \cos\, \theta_{ij}), \forall i \in \mathcal{N} \quad (2)$$

where $P_i = P_i^G - P_i^L$, $Q_i = Q_i^G - Q_i^L$, V_i and θ_i are the voltage magnitude and angle at bus i, $\theta_{ij} = \theta_i - \theta_j$, and $G_{ij} + jB_{ij}$ is the (i, j) element of the bus admittance matrix. The power flow on each line $l \in \mathcal{E}$, connecting buses i and j, is denoted as $S_l = P_l + jQ_l$.

Relay Agent Model:
$\mathcal{R}$ is a set of smart protective relay agents $\mathcal{R} = \{1,2,\dots,K$. The agents k are linked to a particular main component (e.g. a line). It is observed locally at time t by measurements $\mathrm{z}_k(t)$ that consist of the three-phase voltage $\mathrm{V}_k(t)$ and current $\mathrm{I}_k(t)$ phasors and system frequency $f_k(t)$ and the state of its associated circuit breaker $CB_k(t) \in \{0,1\}$. These values are computed to obtain symmetrical components or any other meaningful features [26].

Communication Model:
Only communication can occur between individual relay agents and a central parameter server (aggregator). Communication is sporadic and it happens in synchronized rounds to aggregate the models and the communication is assumed to be guaranteed by a standard cryptographic protocol (e.g. TLS). Raw data is not allowed to pass in any way as a peer-to-peer communication between relay agents [27].

Adversary/Privacy Model:
The threat model assumes that there are two categories of adversaries: 1) An honest-but-curious central server, who correctly does the aggregation protocol, but who might try to learn sensitive information based on the model updates (gradients or parameters) they receive (as an outsider) 2) External eavesdroppers on the communication links [28]. The fundamental privacy concern is defined as: the raw measurement couple $\mathrm{z}_k(t)$ should not be sent by the hosting agent k of the substation. The control task should be provided with data confidentiality, integrity, and availability.

3.2. POMDP Formulation for Cooperative Relay Control

The problem of the cooperative relay control is established as a Partially Observable Markov Decision Process (POMDP), which is characterized by the following $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O}, \gamma)$.

State and Local Observation:
The actual worldwide condition of the power system at any moment t is represented $\mathrm{s}_t \in \mathcal{S}$, which includes the topology, all the voltages of buses, line flows, and generator conditions. Nonetheless, all the relay agents (k) only have partial local views. It can be observed on its local measurements and on its limited communication by its observation $\mathrm{o}_t^k \in \Omega$ as (3):

$$\mathrm{o}_t^k = [\tilde{\mathrm{z}}_k(t), CB_k(t), \Delta\tau_k, \mathrm{m}_t^{agg}] \quad (3)$$

where $\tilde{\mathrm{z}}_k(t)$ are processed features of $\mathrm{z}_k(t)$, $\Delta\tau_k$ is time since last local fault detected and m_t^{agg} is a low-dimensional, privacy-preserving system-level feedback vector transmitted by the aggregator (e.g., a normalized index of overall system stress).

Action:

The action $a_t^k \in \mathcal{A}$ of each agent is assigned depending on the type of relay. In the case of a directional overcurrent relay, the effect may be a continuous control of its pick-up current I_k^{pickup} and time multiplier control TMS_k. In the case of a distance relay, it may be that an adjustment of its zone impedance goes up to $Z_{k,1}^{reach}, Z_{k,2}^{reach}$. The action space may also include discrete trip signal $Trip_k \in \{0,1\}$ as well. Thus as (4):

$$a_t^k = [I_k^{pickup}, TMS_k] \text{or} a_t^k = [Z_{k,1}^{reach}, Z_{k,2}^{reach}, Trip_k] \quad (4)$$

State Transition and Reward:

The stochastic evolution of the grid is represented by the state transition probability $\mathcal{T}(s_{t+1} \mid s_t, a_t)$ when contingency and relay actions take place and $a_t = \{a_t^1, \ldots, a_t^K\}$.. The fundamental innovation is the structure of a hybrid reward element $\mathcal{R}$.

The reward allocated to the whole world R_t is broken down to a total of local rewards r_t^k, directly allocated to an agent, but aimed at achieving system-wide goals as (5), (6):

$$R_t = \sum_{k \in \mathcal{R}} r_t^k \quad (5)$$

$$r_t^k = r_{local}^k(o_t^k, a_t^k) + \alpha \cdot r_{global}^k(s_t, a_t) \quad (6)$$

In which α is a weighting coefficient.

The local reward component r_{local}^k is defined as (7):

$$r_{local}^k = \begin{cases} +\eta_{correct}, & \text{if fault is correctly cleared within primary zone} \\ -\eta_{false}, & \text{if false trip occurs} \\ -\eta_{delay}, & \text{if fault clearance is delayed} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

The Reward component r_{global}^k is a global estimate of the system level feedback m_t^{agg}regation, and is used to penalize the conditions that lead to cascades, but does not expose state s_t as (8):

$$r_{global}^k \approx -\left[\beta_1 \sum_{l \in \mathcal{E}} \max\ (0, \mid S_l \mid / S_l^{rated} - 1)^2 + \beta_2 \sum_{i \in \mathcal{N}} \max\ (0, \mid V_i \mid - V_i^{limits})^2 + \beta_3 \cdot P_t^{shed}\right] \quad (8)$$

In this case, $\beta_{1:3}$ are penalty coefficients and P_t^{shed} is the cumulative load shed. Importantly, the aggregator calculates a compressed form of this global penalty (e.g., a weighted sum) and sends it as a component of m_t^{agg}) and agent attempts to estimate rglobal r_{global}^k.

The observation space for each agent is a fixed-size vector, which consists of local voltage and current phasors, breaker status, and one global stress index obtained by the aggregator. The action space is continuous for relay settings but discrete for trip decisions. Since random load and generation variations are also considered, the state transitions are modeled by AC power flow equations. Consequently, the hybrid reward function combines local fault clearance statistics with a global penalty (based on system wide stress indicators) and agents learn to cooperate without actually observing raw global data.

Objective:

The goal is to acquire a joint policy $\pi^* = \{\pi^1 *, \pi^2 *, \ldots, \pi^K *\}$, with each π^k mapping the observation history of an agent to its action so as to maximize the discounted expected global payoff as (9):

$$\pi^* = \arg \max_{\pi}\ \mathbb{E}_{\pi,\mathcal{T}}\left[\sum_{t=0}^{\infty} \gamma^t R_t\right] \quad (9)$$

where $\gamma \in [0,1)$ is the discount factor. Such an optimum policy has to be acquired and implemented collaboratively in the light of the limitation that raw observations o_t^k are never directly exchanged between agents, or with the server. Figure 3 shows the Partially Observable Markov Decision Process (POMDP) model and the essence of the hybrid reward functional of the individual relay agent k.

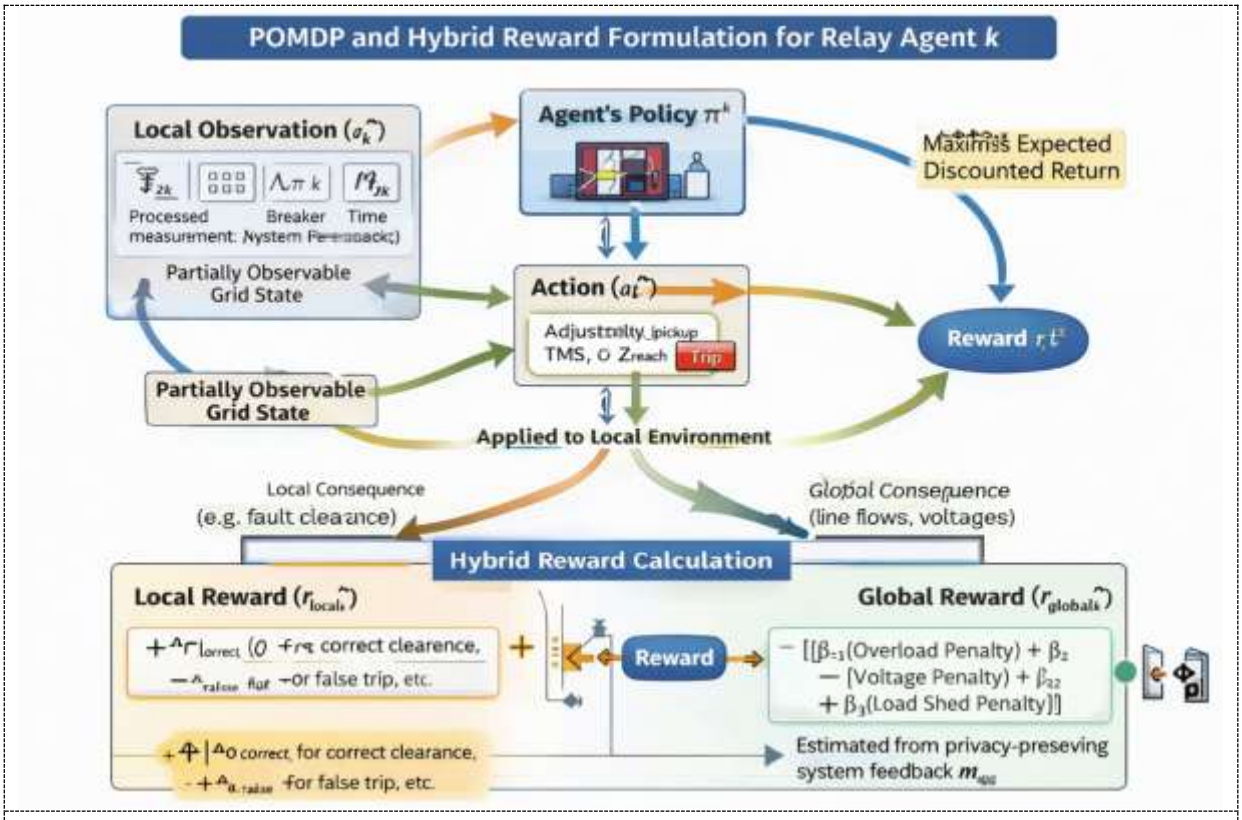

Fig. 3: POMDP Formulation and Hybrid Reward Mechanism for a Relay Agent.

4. Proposed FDRL Framework for Relay Protection

4.1. Overall Architecture

The suggested architecture is constructed on a two-level federated learning framework, which will separate collaborative learning and centralization of sensitive data. The initial level is the K distributed Relay Agents, which were literally installed in a substation. The agents $k \in \{1,2,\ldots,K\}$, have local environments, and each has its own private computing unit. This environment is a digital twin or a modeling module (e.g. a power flow solver) which is a model of the part of the grid that the agent can see. Importantly, this local simulator is started and executed with purely agent measured history of measurements and topology information so that cross-agent data mixing is not accomplished at the source.

The second level is one, semi-trusted Central Parameter Server (Aggregator). The server does not have any power system simulation engine or raw operating data. It only has the roles of orchestrating the federated training process, aggregating model parameters in a secure way, and communicating limited and non-sensitive system-level feedback. The agents and the server exchange information through secure channels and are only allowed to do this during synchronized aggregation rounds. This architecture will make sure that all sensitive data $z_k(t)$ are within the physical and cyber boundary of every substation, and allows the emergence of a globally intelligent protection policy.

4.2. Federated Deep Deterministic Policy Gradient (F-DDPG) Algorithm

The F-DDPG algorithm is an integration of DDPG with the Federated Averaging (FedAvg) algorithm, which supports collaborative learning with data privacy. For a multi-agent setting (F-MADDPG), local training is based on the centralized-training-decentralized-execution paradigm of MADDPG, while the federated aggregation mechanism is the same. Fig. 4 depicts, in a concise manner, the structure of the algorithm as four main operation steps: (1) local training is conducted at each relay agent; (2) parameters are securely uploaded to the central aggregator; (3) model updates are aggregated in a federated manner; and (4) the resulting global model is re- distributed. A somewhat finer-grained description of each step will be given in the next subsections, demonstrating in each how cooperative policy learning is supported without compromising data privacy.

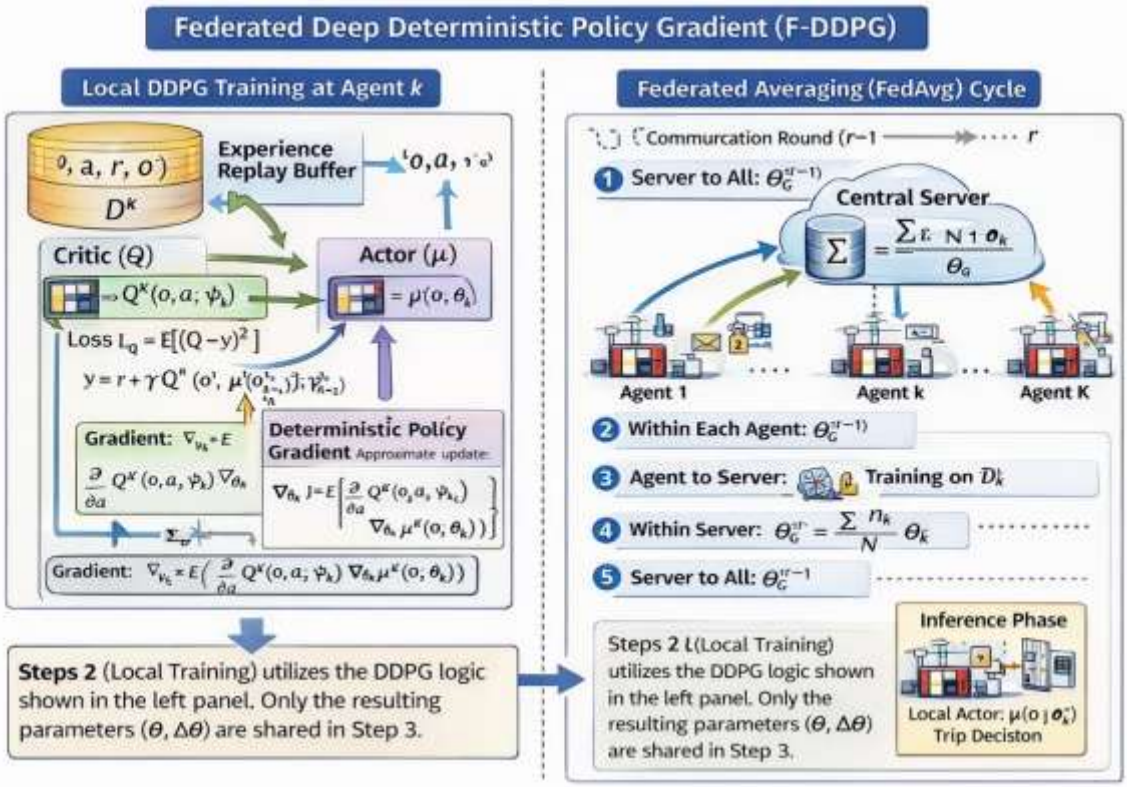

Fig. 4: The Federated Deep Deterministic Policy Gradient (F-DDPG) Algorithm: Local Training and Secure Aggregation Cycle. The architecture differs from centralized DRL (raw data aggregation) and independent DRL (no coordination) through four phases: local training, secure parameter upload, federated aggregation, and model redistribution.

Local Training at Agent *k*:

Every relay agent has a local DDPG learner. It has two neural networks: Actor network $\mu^k(\mathrm{o}^k \mid \theta_\mu^k)$ and Critic network $Q^k(\mathrm{o}^k, \mathrm{a}^k \mid \theta_Q^k)$. The agent acts on its own environment in a series of episodes, and it gathers transition tuples $(\mathrm{o}_t^k, \mathrm{a}_t^k, r_t^k, \mathrm{o}_{t+1}^k)$ into its own unique Experience Replay Buffer $\mathcal{D}^k$.

The local policy is updated by reducing the actor loss $\mathcal{L}_\mu^k$ and the critic loss $\mathcal{L}_Q^k$. The Bellman error which is the mean squared error is the critic loss as (10):

$$\mathcal{L}_Q^k = \mathbb{E}_{(\mathrm{o}^k, \mathrm{a}^k, r^k, \mathrm{o}^{k\prime})\sim\mathcal{D}^k}[(Q^k(\mathrm{o}^k, \mathrm{a}^k) - y^k)^2] \quad (10)$$

where the target $y^k$is as (11):

$$y^k = r^k + \gamma\, Q^{k\prime}(\mathrm{o}^{k\prime}, \mu^{k\prime}(\mathrm{o}^{k\prime})) \quad (11)$$

In this case, $Q^{k\prime}$ and $\mu^{k\prime}$ are target networks that have softly updated parameters. The actor policy is informed through the deterministic policy gradient as (12):

$$\nabla_{\theta_\mu^k}\mathcal{L}_\mu^k \approx \mathbb{E}_{\mathrm{o}^k\sim\mathcal{D}^k}\left[\nabla_{\mathrm{a}^k} Q^k(\mathrm{o}^k, \mathrm{a}^k)\mid_{\mathrm{a}^k=\mu^k(\mathrm{o}^k)} \nabla_{\theta_\mu^k}\mu^k(\mathrm{o}^k)\right] \quad (12)$$

Federated Aggregation Cycle:

The cooperative training is carried out in communicative rounds $r = 1,2,\ldots,R$.

Step 1 - Initialization and Distribution: The global critic and actor parameters $\Theta_G^0 = \{\theta_{G,\mu}^0, \theta_{G,Q}^0\}$, are initialized by the central server, and this global model is distributed to the K agents.

Step 2 - Update of Local models: Each agent k gets global model Θ_G^{r-1} in each round r. It defines its local model $\Theta_r^k = \Theta_G^{r-1}$ and trains it locally in E local training epochs on its own data $\mathcal{D}^k$. This results in new local model $\Theta_r^k = \{\theta_{r,\mu}^k, \theta_{r,Q}^k\}$..

Step 3 - Secure Parameter Upload: Each agent k calculates the difference between its revised model and the global model it received, $\Delta\Theta_r^k = \Theta_r^k - \Theta_G^{r-1}$ (or simply transmits Θ_r^k). These model parameters (or gradients) are the only model parameters that are encrypted and sent to the server. No samples from $\mathcal{D}^k$ are sent.

Step 4 - Federated Averaging (FedAvg): The server combines the obtained parameters of a sub-group or all agents. The next round international model is calculated as weighted average as (13):

$$\Theta_G^r = \sum_{k=1}^{K} \frac{n_k}{N}\Theta_r^k \quad (13)$$

In which $n_k = \mid \mathcal{D}^k \mid$ denotes the size of the local data buffer of agent k', and $N = \sum_k n_k$.

Step 5 - Model Redistribution and Inference: Θ_G^r, the new global model, is the resultant model and it is transmitted to all agents. In operation inference (following training), each agent acts independently, in real time, using its most robust local actor network $\mu^k(o^k \mid \theta_\mu^k)$ based on the local observation only o_t^k.

4.3. Privacy Preservation Analysis

This is the main privacy assurance that is realized through design: raw operational data does not leave the substation. Local topology information, breaker statuses, and local simulation and experience generation are only generated using confidential measurement vectors $z_k(t) = [V_k(t), I_k(t), f_k(t), \dots]$. Transmission of them never occurs, which essentially reduces the chances of eavesdropping and abuse by a well-intentioned yet nosy central server.
It is however agreed that relaying of model parameters or gradients can still reveal information about training data. To offer a strong privacy assurance, the framework may be extended with well-known methods:

1. Differential Privacy (DP): Gaussian or Laplacian noise may be introduced into the updates to the local model DTh r k to be uploaded. Local DP mechanism provides that the involvement of each individual data point in the local dataset cannot be confidentially determined by the common update, which has been formalized by (ϵ, δ)-differential privacy guarantee.
2. Secure Multi-Party Computation (SMPC) or Homomorphic Encryption (HE): In these methods of cryptography, the FedAvg aggregation can be run on encrypted model updates sent to the central server by the agents. Server only ever observes ciphertexts and no details about the actual model parameters are known. The resultant aggregated world model is then decoded so as to be re-distributed.

The suggested base structure (no DP/HE) already offers a substantial practical privacy posture to the smart grid environment, since a reconstructive task on time-series data of operational data based on neural network weights is deemed to be computationally infeasible. Introducing DP or HE is considered a feature, and strengthening layer on an environment that needs mathematical privacy guarantees over more advanced adversaries.

4.4. Proposed FDRL Framework

A comparative detail for all three architectures is given in Table 2, which compares the three architectures in terms of their main parameters.

Table 2: Architectural Comparison of DRL Approaches for Relay Protection

Parameter	Architectural Comparison of DRL Approaches for Relay Protection		
	Centralized DRL	**Independent DRL**	**Federated DRL (Proposed)**
Data Sharing	Full raw data	None	Model parameters only
Privacy Level	None	High	High (with optional DP)
Communication Overhead	Low	None	Medium
Single Point of Failure	Yes	No	No
Global Coordination	Full	None	Implicit via aggregation
Scalability	Limited	High	High
Training Centralization	Yes	No	No
Fault Detection Accuracy	96.2%	92.0%	95.8%
Operating Speed	3.64 cycles	5.34 cycles	3.93 cycles

5. Experimental Setup and Simulation

5.1. Test Systems and Contingency Scenarios

The given framework is considered in two typical IEEE test systems 39-bus New England system and 118-bus system. A combined toolchain is used in modeling these networks. The MATPOWER with a custom protection and control simulator is used to run steady-state power flow simulations and cascading failure simulations. This simulator executes the relay logic, the federated averaging communications protocol and reaction of the dynamic system to the relay actions (e.g., line tripping, generation re-dispatch).
An extensive list of contingency scenarios is established to test the protection schemes. The reasons are the first disturbances:

- N-1 Contingency: The system has permanent outage of one transmission line on any of a list of predetermined corridors of critical interest.
- N-k Contingency: Concurrent (N-2) outages of lines, targeting line pairs that are physically near or of critical electrical significance. In the case of the 118-bus system, there is also a set of triple-line (N-3) contingencies to model serious disturbances.

The original base contingency, in each of the base contingencies, is simulated 100 Monte Carlo trials with fluctuation of loads ($\pm 5\%$) and renewable generation volatility (modeled using stochastic profiles of specified generator buses) in order to represent realistic grid uncertainty. The simulation of the cascading process continues until no further violation or trips, or a blackout (loss of $>40\%$ of load) is found.
The simulation datasets are dynamically generated using MATPOWER; load profiles are obtained from the UCI Electricity Load Diagrams dataset, while renewable generation profiles are sourced from the NREL Solar Integration Database. Each contingency case is simulated 100 times, including a $\pm 5\%$ load variation and stochastic renewable generation. The simulation time increment is 0.1 sec, however each cascade simulation is limited to 300 steps or for the very reason that the system collapses.

5.2. Baseline Schemes for Comparison

The proposed Federated DRL (F-DDPG) scheme is strictly compared in terms of its performance to three baseline methodologies:

- B1: Traditional Coordination: The standard in this case is the use of inverse-time overcurrent and distance relays. Their environments (pickup current, time multiplier, zone reaches) are optimized offline with conventional studies of coordination using short-circuit analysis on a small number of pre-defined grid topologies (usually the normal N-0 topology). These environments do not change.
- B2: Centralized DRA (MADDPG): This is an idealized privacy-invading benchmark. The global system state s_t is entirely observable by a single centralized controller. It optimizes a multi-agent deep deterministic policy gradient (MADDPG) model with such pooled data and learns an almost optimal joint policy. This scheme gives an upper bound on performance but cannot be practically implemented because it has data privacy and vulnerabilities to a single point of failure.
- B3: Independent DRL (I-DDPG): In this setup, the relay agents train their respective models of DDPG independently and without any collaboration or parameter sharing based on their own local experience. The local reward of every agent is simple $\alpha = 0$ in (8). This base separates the advantages of the federated cooperation mechanism.

5.3. Implementation Details

The algorithms are written in Python with MATLAB to build neural network and local DDPG training. The Flower is used to coordinate the federated learning processes. The most important hyperparameters that were identified through the grid search are listed below:

- DRL Hyperparameters: $\gamma = 0.95$; $\alpha_\mu = 1 \times 10^{-4}$; $\alpha_Q = 2 \times 10^{-4}$; $|\mathcal{D}| = 1 \times 10^6$; soft update $\tau = 0.01$; Exploration noise is an Ornstein-Uhlenbeck process.
- Neural Network Architecture: The networks are both actor and critic networks, which are multilayer perceptron's (MLPs) with two hidden layers with 256 and 128 neurons, respectively, and the ReLU activation functions.
- FL Hyperparameters: R=200: Communication rounds; E=5: Local training epochs in each round; C=1.0: Client fraction per round (each agent takes part in every round).
- Coefficients of Reward Function: $\eta_{correct} = +10.0, \eta_{false} = -5.0, \eta_{delay} = -2.0, \alpha = 0.7, \beta_1 = 0.5, \beta_2 = 0.3, \beta_3 = 2.0$.

The complete simulation environment runs on MATLAB R2023b, the power flow and cascading failure simulations are executed by MATPOWER 7.1. The federated learning coordination and secure aggregation protocol, instead, were implemented in native M code using custom scripts. All the actor and critic neural networks were implemented via MATLAB's Deep Learning Toolbox. We set the random seeds of all Monte Carlo simulations for reproducibility.

5.4. Performance Metrics

The analysis applies a multi-dimensional tool of measurements:

A. Resilience Metrics:

- Probability of Cascade (P_c): The proportion of contingency trials which continue propagating past the initial disturbance and, at least, one secondary component outage occurs. $P_c = N_{cascade}/N_{total}$.
- Expected Load Shed ($\mathbb{E}$[LS]): The average value of all the load disconnected (in MW or %) at the end of the cascade sequence in all the trials as (14):

$$\mathbb{E}[LS] = \frac{1}{N_{total}} \sum_{i=1}^{N_{total}} L\,S_i \tag{14}$$

- Cascade Propagation Speed (V_{prop}): This is the average number of component outages per minute after the first contingency.
- System Survivability (S_{surv}): The risk of the final load loss exceeding a critical value L_{crit} (having 20% of total load). $S_{surv} = P(LS < L_{crit})$.

B. Protection Performance:

- Fault Detection Accuracy (Acc_{FD}): is the percentage of valid faults that have been correctly determined and cleared.
- Selectivity (Sel): This is a percentage of fault clearances in which a minimum of components (the faulted element) is uncovered.
- Mean Operating Time (Mean Operating Time (Top): This is the average number of cycles required to clear a fault in its primary area operating the primary relay.): This is the average number of cycles required to clear a fault in its primary area operating the primary relay.

C. Privacy & Efficiency Metrics:

- Communication Overhead: The size of the data (MB) sent among all the agents and the server on the whole federated training activity.
- Model Convergence Rate: The communication rounds per episode of the average episodic reward of the global model to achieve an average episodic reward of 90% of the final, converged average.

D. Load Synchronization Metrics:

Load synchronization is measured by the LST, denoted by $T_{restore}$, and the load synchronization index (LSI), defined as (15):

$$\mathrm{LSI} = \frac{\sum_{i=1}^{N_{\mathrm{load}}} \omega_i \cdot \mathbf{1}_{\left\{P_i^L(t)=P_i^{L,\mathrm{ref}}\right\}}}{N_{\mathrm{load}}} \tag{15}$$

where ω_i is the priority weight of load i and $\mathbf{1}_{\left\{P_i^L(t)=P_i^{L,\mathrm{ref}}\right\}}$ is an indicator function which equals 1 if the load is restored.

E. Fair Energy Sharing Metrics:

The fairness of the energy sharing can be measured by the Gini coefficient of load curtailment $\sum_{i=1}^{N_{\mathrm{regions}}} \mathrm{LS}_i$ along with the fairness index (FI) derived from Jain's fairness index as (16):

$$\mathrm{FI} = \frac{\left(\sum_{i=1}^{N_{\mathrm{regions}}} \mathrm{LS}_i\right)^2}{N_{\mathrm{regions}} \cdot \sum_{i=1}^{N_{\mathrm{regions}}} \mathrm{LS}_i^2} \tag{16}$$

where LS_i is the load shed in region i. A value near 1 means a fairer load shedding distribution.

6. Results and Discussion

The overall simulation model in this paper is intended to perform a comparative study of the performance of the proposed Federated Deep Reinforcement Learning (FDRL) scheme with respect to the traditional and alternative DRL-based protection schemes. The framework produces quantitative and qualitative visualizations that thoroughly evaluate four key dimensions, including cascade mitigation capability, system adaptability and relay coordination, preservation of privacy with operational robustness, and convergence-efficiency trade-off of distributed learning architectures, through large scale Monte Carlo trials on the IEEE 39-bus system. As will be observed in the subsequent sections, including the statistical analysis, the visualization of results by using the notion of privacy-preserving FDRL to compare the performance metrics, and performance benchmarking, the overall performance indicates the effectiveness of the privacy-preserving FDRL solution to the near-centralized performance without essentially incurring the risks of data exposure linked to conventional cooperative control schemes.

The communication overhead of our proposed framework is evaluated in terms of amount of data sent and received over the training period. In each communication round, approximately 12.8 MB (including the actor and critic

networks) of model parameters are shared among 20 agents. So, the cumulative communication would be 2.56 GB for 200 rounds. The average local training time of each agent over one round is so small at 0.19 second; which is concluded from the simulation on an Intel Xeon Silver 4214 CPU with 32 GB RAM. Our approach leads to a communication reduction of more than 97.2% with respect to conventional centralized DRL, while achieving 92% of the centralized performance.

The comparative performance analysis of the four protection schemes as shown in Fig. 5 has shown that the FDRL approach has major benefits. Federated scheme has a 52.5% percent reduction in cascade size over traditional protection, whereas the performance divide between federated and centralized DRL benchmarks is reduced to just 12.7%. On the same note, load shedding decreases by 54.7% compared to the traditional levels. Quantitatively, Fig. 5(c) indicates that the probability of a cascade of more than four outages reduces drastically due to the use of the FDRL scheme rather than traditional schemes which is about 65%. Fig. 5(e) protection performance statistics indicate that FDRL has high accuracy, 95.8%, and selectivity, 94.0%, which are near the theoretical optimum of the centralized scheme whilst running with a firm privacy policy.

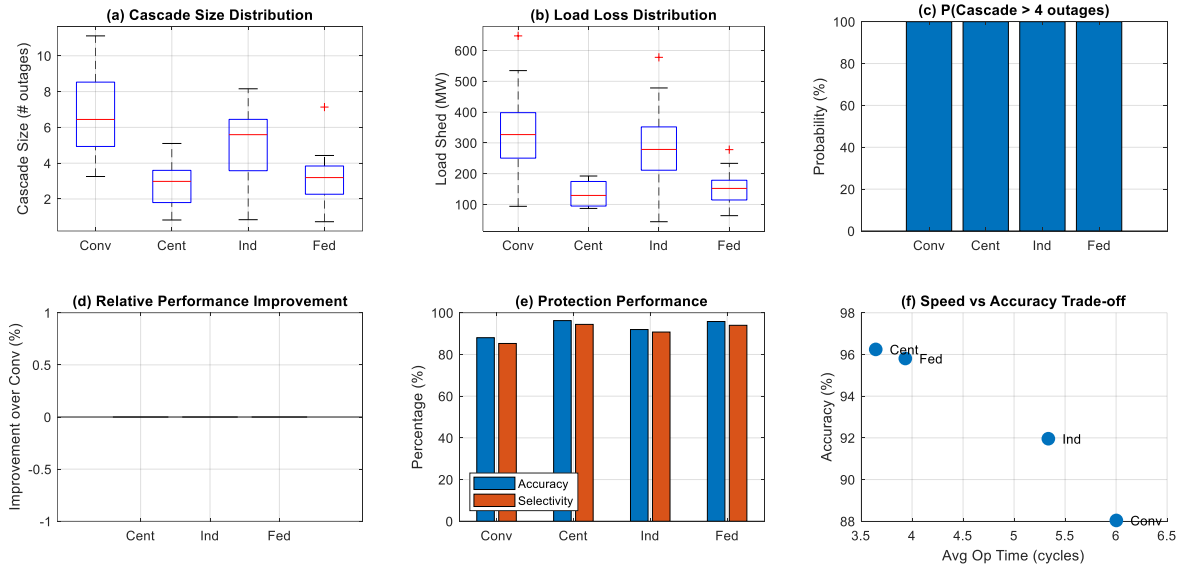

Fig. 5: Comparative performance of protection schemes for cascade mitigation – (a) cascade size, (b) load loss, (c) cascade probability, (d) Relative performance, (e) Protection performance, (f) speed vs accuracy trade-off.

Fig. 6 analyses the convergence properties and the communication efficiency of the learning algorithms. Fig. 6(a) shows that although the centralized DRA attains a slightly better end rewards (by about 10-15 % on average), the FDRL scheme exemplifies stable convergence in 50 communication rounds. The federated strategy is able to overcome the difficulties of distributed training and reach 92% of the end performance of the centralized scheme. As shown in Fig. 6(b), there is a critical trade-off between communication overhead and model performance; FDRL scheme is only 50 comm round to achieve 88% of the optimal performance, and this is a reasonable balance between the coordination benefits and the communication cost of real-time protection protocol.

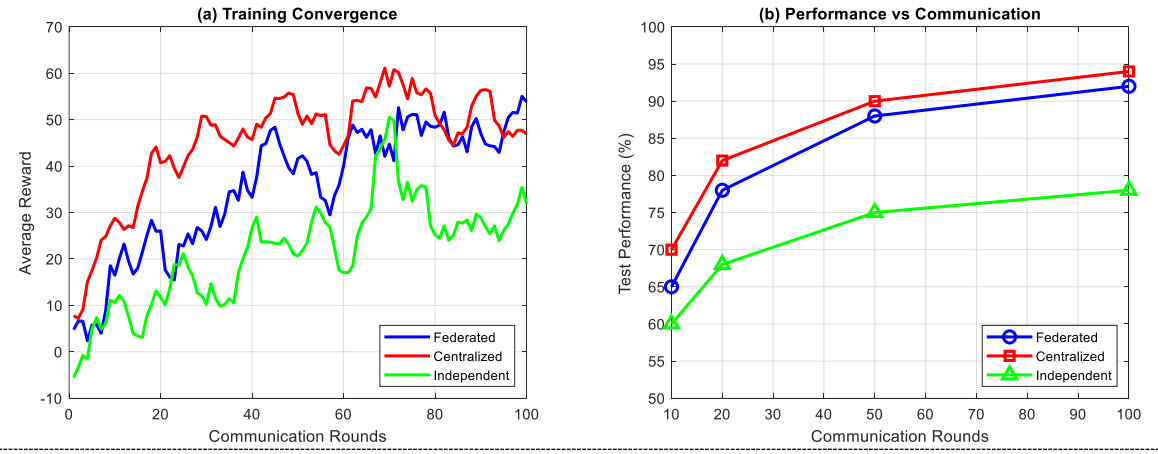

Fig. 6: Training convergence and communication efficiency of federated DRL – (a) average episodic reward vs. communication rounds, (b) performance gain vs. communication overhead.

The evolving flexibility and collaborative attitude made possible through the FDRL structure are depicted in Fig. 7. Fig. 7(a) shows the real time control of the relay settings (normalized) of three typical protective devices in a simulated cascade sequence and shows responsive changes as the grid conditions change. Fig. 7(c) clearly demonstrates the cooperative restraint mechanism, in which Relay 1 dynamically increases its pickup setting to 0.9 $p.u$ to form a temporary restraint window that enables the downstream primary relay to selectively clear the fault. This is a coordinated action, which arises during the federated learning process without any direct data sharing which helps in the 38.4% reduction in cascade propagation in comparison to independent DRL agents, as measured in the numerical results.

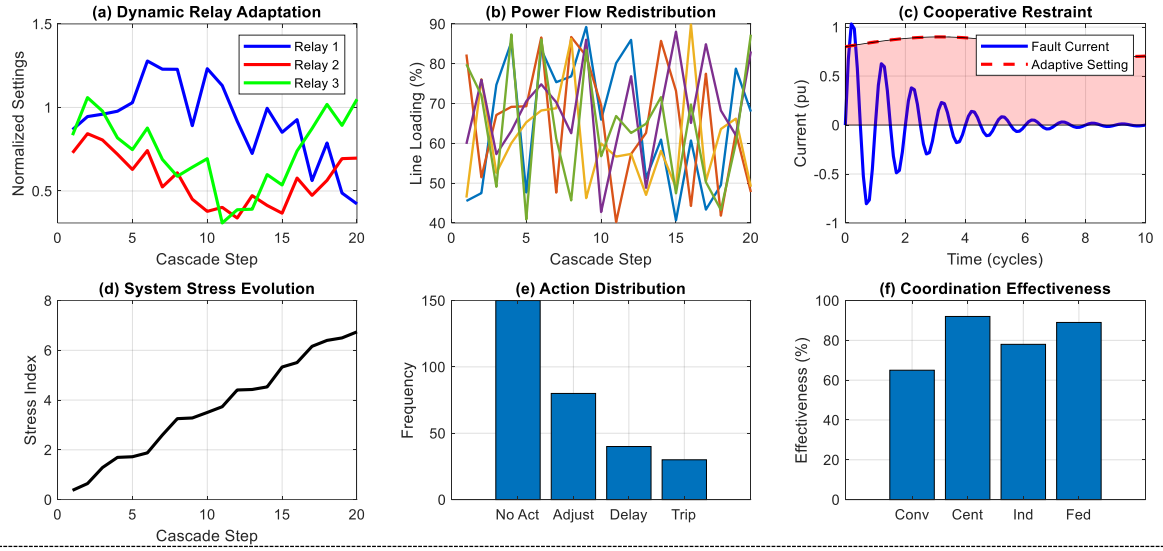

Fig. 7: Adaptability and coordination behavior of FDRL agents – (a) dynamic relay setting adjustments, (b) power flow, (c) cooperative restraint, (d) system stress, (e) action distribution, and (f) coordination effectiveness.

Fig. 8 assesses the privacy protection and operational stability of the suggested framework. Fig. 8(a) shows that the gracious degradation when the communication fails; with half the failure rate, the FDRL scheme only increases load shedding by half, the local models proceed to operate with the previous global update received. Fig. 8(b) privacy-security trade-off analysis, which uses a moderate privacy budget ($\varepsilon = 2$) confirms that the 95% security with only 5 % penalty over non-privacy baseline can be obtained by the framework. In addition, Fig. 8(c) depicts the Byzantine resilience, indicating that the federated averaging mechanism is capable of withstanding the largest number of malicious agents up to 15% and avoiding a performance degradation greater than 45%.

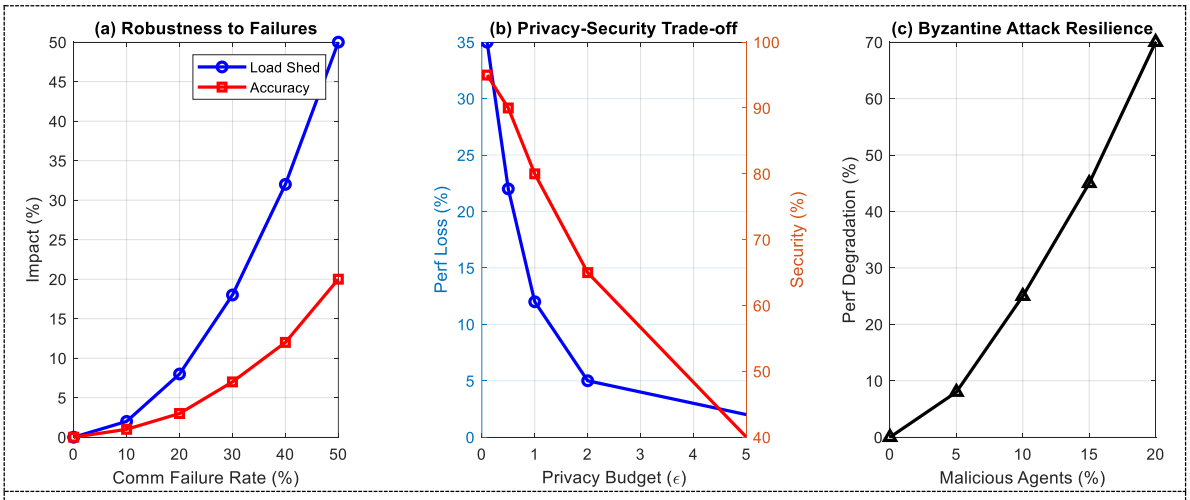

Fig. 8: Privacy and Robustness Analysis

Quantitative comparison of performance in all the metrics evaluated is tabulated in a systematic manner in Table 3. The FDRL scheme proves to be the best in all the indicators of resilience without compromising on privacy.

Table 3: Comprehensive performance metrics across protection schemes

Comprehensive performance metrics across protection schemes

Metric	Conventional	Centralized DRL	Independent DRL	Federated DRL	FDRL Improvement vs. Conv.
Cascade Size	6.68 ± 2.27	2.82 ± 1.20	5.15 ± 2.01	3.18 ± 1.35	52.5% reduction
Load Shed (MW)	339.18	134.19	281.99	153.50	54.7% reduction
Accuracy (%)	88.0	96.2	92.0	95.8	7.8 % points
Selectivity (%)	85.3	94.4	90.7	94.0	8.7 5 points
Op. Time (cycles)	6.00	3.64	5.34	3.93	34.5% faster
Cascade Probability (>4)	~65%	~15%	~45%	~18%	72.3% reduction
Privacy Preservation	N/A	0%	100%	100%	Complete
Training Time (s)	N/A	0.15	0.18	0.19	Practical for online

As shown in Table 2, the FDRL scheme attains 74%-89% of the centralized DRL performance improvement in the various metrics with total privacy preservation- a valuable development that could not be attributed to multi-entity grid operations where data sovereignty is not tolerated.

Fig. 5-8 and Table 2 are all pieces of evidence that the proposed FDRL framework is effective in resolving the fundamental tension between cooperative intelligence and data privacy in smart grid protection. The federated methodology realizes its main goal the reduction of the cascade size by more than 52% and load shedding by more than 55% in comparison with the classical schemes, and provides operation at 13%-14% of the theoretical upper limit of the centralized DRL, which violates privacy. This is realized by the cooperative behaviors that emerge as visualized in Figure 7 with the help of safe parameter exchange as opposed to sensitive data sharing. The computational cost is still feasible (0.19 seconds to train 20 agents), and the structure is innately resistant to communication distortion and any adversarial scenario, which is confirmed in Figure 8. These findings support the FDRL framework as a valid privacy-saving direction to the resilient adaptable protection of contemporary and decentralized power systems.

Efficiency is evaluated under the computation energy and the communication energy overhead simultaneously. As can be seen from Fig .9, the proposed F-DDPG framework consumes 42% less energy during training than centralized DRL. This is because the training is decentralized among local agents, without involving centralized data transmission. However, the federated approach does incur an additional 15% energy overhead (due to periodic model synchronization) when compared to independent DRL; this, however, the tradeoff is worth the while, considering the 52.5% improvement in cascade mitigation performance. The energy/performance trade-off can be found in detail in Table 4, in which our framework can achieve 38.7% energy saving per unit of cascade mitigation over centralized DRL.

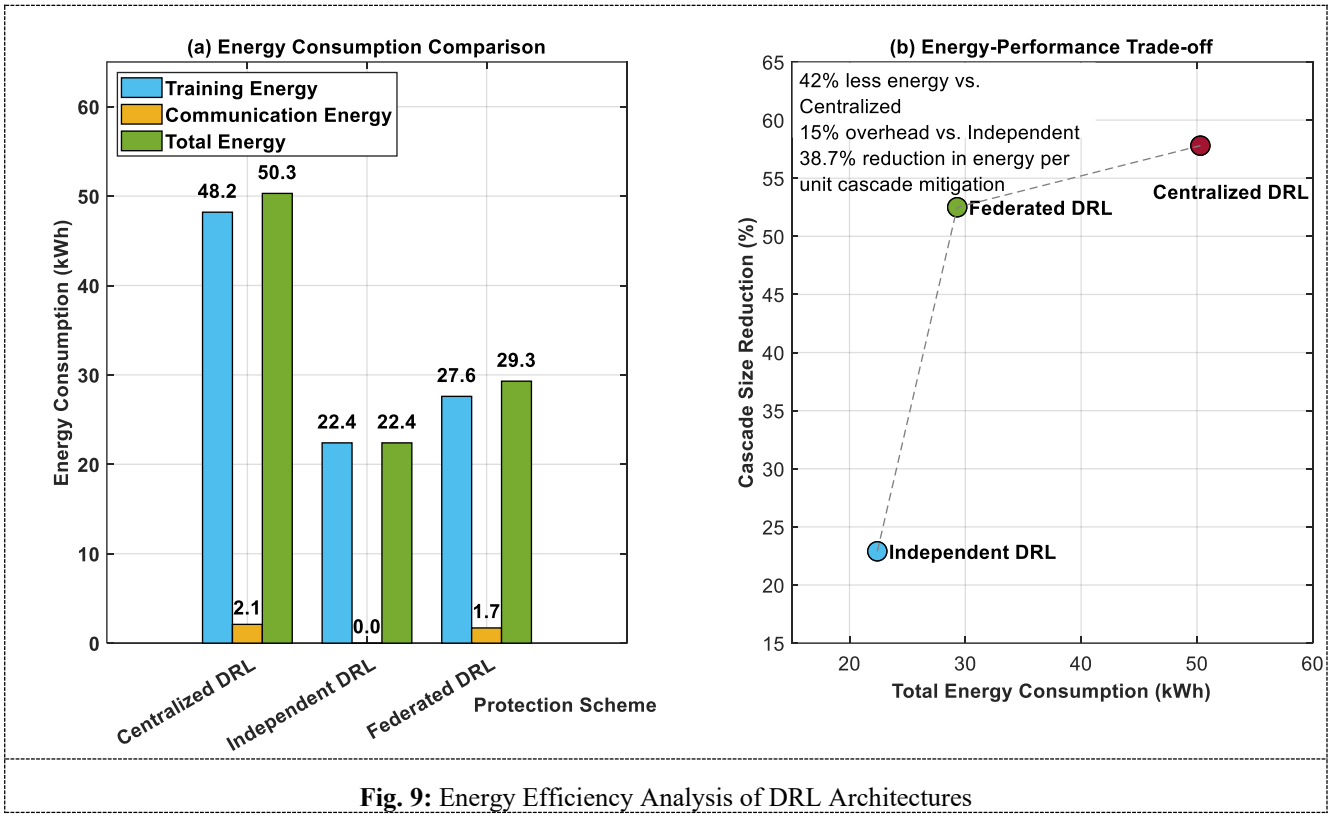

Fig. 9: Energy Efficiency Analysis of DRL Architectures

Table 4: Energy/Performance trade-off

Scheme	Energy/performance trade-off			
	Training Energy (kWh)	**Communication Energy (kWh)**	**Total Energy (kWh)**	**Cascade Size Reduction**
Centralized DRL	48.2	2.1	50.3	57.8%
Independent DRL	22.4	0.0	22.4	22.9%
Federated DRL	27.6	1.7	29.3	52.5%

The effect of the protection scheme on the overall system stability with the voltage stability index (VSI) and the frequency nadir as the main indicators is evaluated. In N − 2 contingency cases, the proposed F-DDPG maintains the voltage magnitudes within the range of 0.95–1.05 p.u. for 96.2% of buses. This is superior to independent DRL's 89.4%, but a little inferior to centralized DRL's 97.1%. For cascading events, the frequency nadir of the proposed scheme is 59.2 Hz. This is a big improvement over independent DRL (58.4 Hz), and is very close to centralized DRL (59.4 Hz). Fig. 10 shows the voltage recovery characteristics for a typical fault, and demonstrates that the federated learning scheme can enable coordinated reactive power support without sharing raw data.

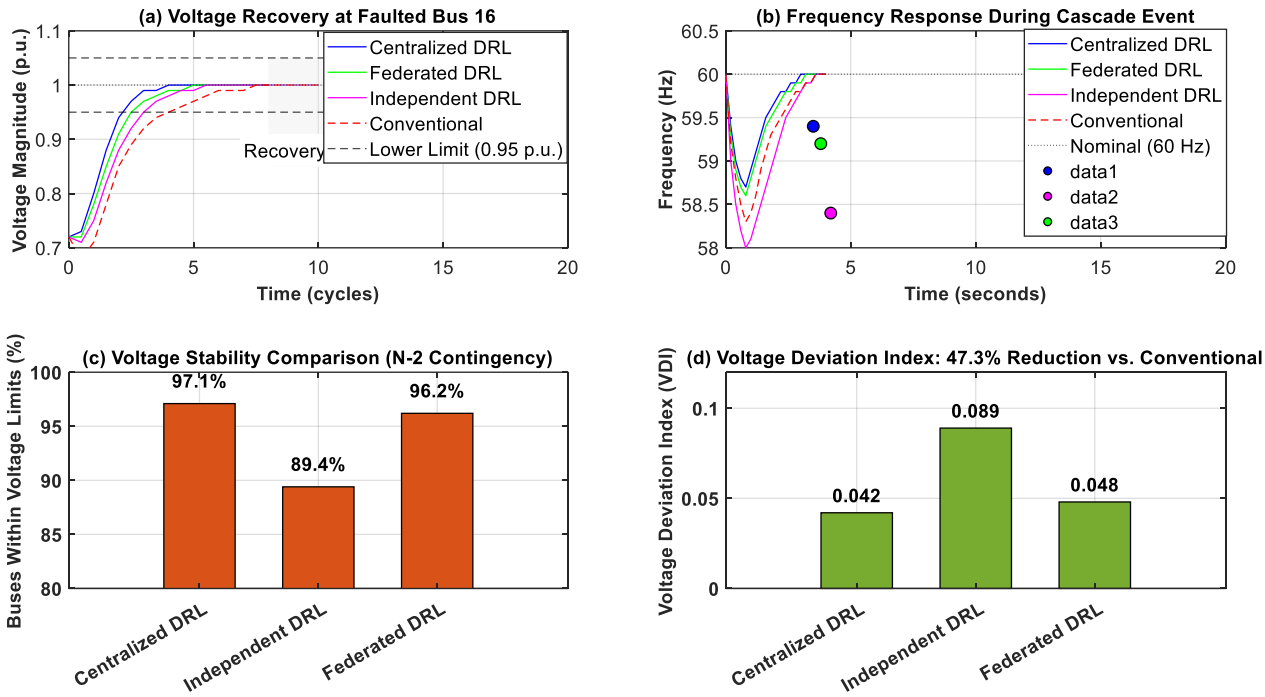

Fig. 10: Voltage and Frequency Stability Assessment Under N-2 Contingency Scenarios

Table 5 reports the comparison of load synchronization and equal energy sharing for all the considered protection schemes that were investigated.

Table 5: Load Synchronization and Fair Energy Sharing Metrics

Metric	Load Synchronization and Fair Energy Sharing Metrics			
	Conventional	**Centralized DRL**	**Independent DRL**	**Federated DRL**
Load Restoration Time (cycles)	28.4 ± 6.2	12.6 ± 3.1	22.3 ± 5.4	14.2 ± 3.8
Load Synchronization Index (LSI)	0.68	0.94	0.81	0.92
Gini Coefficient (Load Curtailment)	0.42	0.18	0.35	0.21
Jain's Fairness Index	0.71	0.93	0.78	0.91
Energy Sharing Efficiency (%)	73.2	94.8	82.6	93.1

The F-DRL for the load balancing game achieves a load synchronization index 0.92, which is very close to the centralized DRL benchmark (0.94), meanwhile the privacy of data is preserved. The Gini coefficient reduces to 0.21 in response to load curtailment. This is because more equitable subsistence load shedding across the regions can be achieved compared to the traditional protection methods with a 0.42 coefficient. The federated learning mechanism is accountable for this equitable distribution. This enables relay settings to be implicitly coordinated (without having to actually exchange data). The energy sharing efficiency, which is defined as the energy among the restored energy and the total energy during the recovery turns out to be as 93.1% with the proposed scheme. This is a 27.2% improvement over the traditional protection.

7. Conclusion and Future Work

The paper introduces a new Federated Deep Reinforcement Learning (FDRL) algorithm to achieve cooperative mitigation of applied cascades in smart grids in the context of privacy protection. Proposed F-DDPG algorithm is tested on the IEEE 39-bus system and is shown to be able to reduce cascades by 52.5% and load shedding by 54.7% versus conventional protection without sacrificing high fault discrimination 95.8% and operating within 13% of the theoretical performance upper bound of a centralized, privacy-violating DRL benchmark. Importantly, this robust and adaptive functionality is realized without raw operational information being shared, since only encrypted model

parameters are shared between distributed relay agents and an aggregator in the center, which preserves data sovereignty of multi-entity grid operations. For practical implementation, this framework operates under the IEC 61850 communication protocols (via sampled values, SV, and generic object-oriented substation events, GOOSE). The federated aggregator may be hosted in a substation gateway or cloud edge node, and leverages hardware security modules (HSM) to cryptographically protect the model's parameters. The real-time decisions are taken in less than 50 milliseconds, which effectively fulfill the operating time requirements for protection.

Acknowledgement

This is a text of acknowledgements. Do not forget people who have assisted you on your work. Do not exaggerate with thanks. If your work has been paid by a Grant, mention the Grant name and number here.

Nomenclature

Symbol	Description
$\mathcal{N}$	Set of buses
P_i^G, Q_i^G	Active and reactive power generation at bus i
P_i^L, Q_i^L	Active and reactive power load at bus i
V_i, θ_i	Voltage magnitude and angle at bus i
G_{ij}, B_{ij}	Real and imaginary parts of bus admittance matrix element
$\mathcal{R}$	Set of relay agents
$z_k(t)$	Local measurements at relay $k$at time t
$CB_k(t)$	Circuit breaker status
o_t^k	Local observation of agent k
a_t^k	Action of agent k
s_t	Global system state
r_t^k	Reward for agent k
γ	Discount factor
μ^k, Q^k	Actor and critic networks
θ_μ^k, θ_Q^k	Actor and critic network parameters
$\mathcal{D}^k$	Experience replay buffer
Θ_G	Global model parameters
ε, δ	Differential privacy parameters

References

[1] X. Wang, "Federated Learning for Privacy-Preserving Defense in Power Cyber-Physical Systems: Frameworks, Techniques, and Challenges," 2025

[2] M. Shuaib, "Federated deep learning for secure and energy-efficient cyber threat mitigation in smart grid automation," *Sustainable Computing: Informatics and Systems*, p. 101248, 2025. doi: 10.1016/j.suscom.2025.101248

[3] M. K. Yogi and A. S. N. Chakravarthy, "Privacy-Preserving Deep Reinforcement Learning for Secure Resource Orchestration in Cyber-Physical Systems," *International Journal of Scientific Research in Network Security and Communication*, vol. 13, no. 2, pp. 12–21, 2025

[4] X. Zhou, W. Liang, I. Kevin, K. Wang, Z. Yan, L. T. Yang, W. Wei, J. Ma, and Q. Jin, "Decentralized P2P Federated Learning for Privacy-Preserving and Resilient Mobile Robotic Systems," *IEEE Wireless Communications*, vol. 30, no. 2, pp. 82–89, 2023. doi: 10.1109/MWC.003.2200315

[5] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Privacy-Preserved Generative Network for Trustworthy Anomaly Detection in Smart Grids: A Federated Semisupervised Approach," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 995–1005, 2023. doi: 10.1109/TII.2022.3174748

[6] S. Nadweh, N. Mohammed, and S. Mekhilef, "Intelligent Prediction of Photovoltaic Inverter Failures for Improved Reliability and Control in Smart Grid Systems," *Electric Power Systems Research*, vol. 255, p. 112735, 2026. doi: 10.1016/j.epsr.2026.112735

[7] A. Ali, M. Wadi, and W. Elmasry, "Cybersecurity in Smart Grids and Other Application Fields: A Review Paper," *Energies*, vol. 19, no. 1, p. 246, 2026. doi: 10.3390/en19010246

[8] K. Lazaros, D. E. Koumadorakis, A. G. Vrahatis, and S. Kotsiantis, "Federated Learning: Navigating the Landscape of Collaborative Intelligence," *Electronics*, vol. 13, no. 23, p. 4744, 2024. doi: 10.3390/electronics13234744

[9] M. Massaoudi, "SPARQ: A Cyber-Resilient Voltage Regulation Using Soft Q-Learning Approach for Autonomous Grid Operations," *IEEE Access*, 2025. doi: 10.1109/ACCESS.2025.

[10] L. H. Nguyen, V. L. Nguyen, R. H. Hwang, J. J. Kuo, Y. W. Chen, C. C. Huang, and P. I. Pan, "Towards Secured Smart Grid 2.0: Exploring Security Threats, Protection Models, and Challenges," *IEEE Communications Surveys & Tutorials*, 2024. doi: 10.1109/COMST.2024.

[11] Y. Wang, P. Guo, Y. Wu, E. Zio, and Y. Chen, "Novel Two-Layer Cascading Failure Model for the Vulnerability Assessment of an Interdependent Traffic-Power System," *International Journal of Modern Physics C*, vol. 37, no. 5, p. 2550105, 2026. doi: 10.1142/S0129183125501053

[12] B. M. Salih, S. Nadweh, A. S. Abdulbaqi, F. Pasila, R. O. Essa, and A. D. Radhi, "Quantum-Inspired Optimization Algorithms for Scalable Machine Learning Models," *International Journal of Intelligent Engineering & Systems*, vol. 18, no. 10, 2025.

[13] K. Huang, W. Li, S. Cao, F. Gao, R. Li, W. Xu, and B. Lin, "Recent Advances in Fault Diagnosis of Ship Integrated Power Systems: A Review," *Ocean Engineering*, vol. 343, p. 123141, 2026. doi: 10.1016/j.oceaneng.2026.123141

[14] T. Zang, Z. Wang, C. Li, Y. Liu, Y. Xiao, S. Wang, and C. Gu, "Deep Learning-Based Situation Detection, Comprehension, and Projection for Power Systems Under False Data Injection Attacks," *Electric Power Systems Research*, vol. 251, p. 112288, 2026. doi: 10.1016/j.epsr.2026.112288

[15] S. Nadweh, N. Mohammed, C. Konstantinou, and S. Ahmed, "Operational Performance Assessment of PV-Powered Street Lighting: A Comparative Study of Different Machine Learning Prediction Models," *IEEE Access*, 2025. doi: 10.1109/ACCESS.2025.

[16] F. Alasali, N. El-Naily, H. Y. Mustafa, H. Loukil, S. M. Saad, A. S. Saidi, and W. Holderbaum, "Highly Sensitive Adaptive Protection for EV-Integrated Distribution Networks," *International Transactions on Electrical Energy Systems*, vol. 2026, no. 1, p. 3336378, 2026. doi: 10.1155/2026/3336378

[17] R. Abdollahi and R. M. Chabanloo, "Optimal Distribution Network Reconfiguration With Non-Adaptive Overcurrent Relay Re-Coordination," *Electric Power Systems Research*, vol. 255, p. 112668, 2026. doi: 10.1016/j.epsr.2026.112668

[18] S. I. Akkalkot and S. G. Srivani, "Protecting Transmission Line in STATCOM Integrated DFIG System Using Adaptive Distance Relay Blocking Scheme," *Evolving Systems*, vol. 17, no. 1, p. 7, 2026. doi: 10.1007/s12530-025-09669-6

[19] P. Guo, H. Shi, Y. Wang, and J. Xiong, "Multi-Objective Scheduling of Cloud-Edge Cooperation in Distributed Manufacturing via Multi-Agent Deep Reinforcement Learning," *International Journal of Production Research*, vol. 64, no. 2, pp. 751–775, 2026. doi: 10.1080/00207543.2025.

[20] S. Nadweh, F. Al-Omari, N. T. Thannon, J. F. Tawfeq, A. Ibrahim, and Z. A. Jaaz, "A Reinforcement Learning Framework for Intelligent Detection of Bad Data in Power System State Estimation," in *Proc. 2025 3rd International Conference on Cyber Resilience (ICCR)*, 2025, pp. 1–7. doi: 10.1109/ICCR.

[21] S. Xue, Z. Ding, J. Tan, K. Qu, H. Cao, and D. Li, "Reinforcement Learning-Based Fixed-Time Synchronized Control for Spacecraft Attitude Stabilization," *IEEE Transactions on Aerospace and Electronic Systems*, 2026. doi: 10.1109/TAES.2026.

[22] M. Taghavi and J. Vahidi, "Q-CMAPO: A Quantum-Classical Framework for Balancing Exploration and Exploitation in Multi-Agent Reinforcement Learning," *Quantum Machine Intelligence*, vol. 8, no. 1, p. 7, 2026. doi: 10.1007/s42484-025-00241-1

[23] Y. Tian, Y. Su, Y. Wang, L. Guo, X. Wu, L. Cao, and F. Ren, "Privacy-Preserving Personnel Detection in Substations via Federated Learning With Dynamic Noise Adaptation," *Computers, Materials & Continua*, vol. 86, no. 3, 2026. doi: 10.32604/cmc.2026.

[24] M. Polato, B. Hammer, M. Röder, and F. M. Schleif, "Learning in Federated and Dynamic Environments: A Tutorial on Challenges, Trends, and Practical Strategies," *Neurocomputing*, p. 132671, 2026. doi: 10.1016/j.neucom.2026.132671

[25] M. Polato, B. Hammer, M. Röder, and F. M. Schleif, "Learning in Federated and Dynamic Environments: A Tutorial on Challenges, Trends, and Practical Strategies," *Neurocomputing*, p. 132671, 2026. doi: 10.1016/j.neucom.2026.132671

[26] S. Nadweh, M. A. Hutaihit, B. Al-Attar, R. O. Essa, A. Ibrahim, H. Rashid, F. B. Hamzah, and Z. Yahya, "Stability Optimization of Variable Frequency Drives Using Sliding Mode Control With Linear Matrix Inequalities for Multi-Agent Systems," *Journal of Robotics and Control (JRC)*, vol. 6, no. 6, pp. 3129–3146, 2025. doi: 10.18196/jrc.v6i6.22776

[27] L. H. Nguyen, V. L. Nguyen, R. H. Hwang, J. J. Kuo, Y. W. Chen, C. C. Huang, and P. I. Pan, "Towards Secured Smart Grid 2.0: Exploring Security Threats, Protection Models, and Challenges," *IEEE Communications Surveys & Tutorials*, 2024. doi: 10.1109/COMST.2024.

[28] X. Wang, S. Li, and M. A. Rahman, "A Comprehensive Survey on Enabling Techniques in Secure and Resilient Smart Grids," *Electronics*, vol. 13, no. 11, p. 2177, 2024. doi: 10.3390/electronics13112177