BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 74(1), 2026, Article number: e155897
DOI: 10.24425/bpasts.2025.155897

ARTIFICIAL AND COMPUTATIONAL INTELLIGENCE

Towards sustainable wireless rechargeable sensor networks: a federated multi-agent reinforcement learning approach for cooperative wireless charging

C.N. VANITHA and **P. ANUSUYA***

Department of Information Technology, Karpagam College of Engineering, Coimbatore, India

Abstract. Wireless rechargeable sensor networks (WRSNs) face persistent energy limitations due to the finite battery capacity of sensor nodes, which can compromise network reliability in remote or dynamic environments. To address these challenges, this paper proposes a novel federated multi-agent reinforcement learning (FedRL-MARL) framework for adaptive and cooperative energy replenishment using multiple mobile wireless chargers (MWCs). Unlike traditional centralized approaches, FedRL-MARL leverages decentralized policy learning, enabling each MWC to train on local observations while contributing to a globally aggregated model through federated updates. The problem is formulated as a Markov decision process (MDP), allowing agents to make intelligent charging and routing decisions in real time, even in the presence of obstacles and changing node demands. Simulation results demonstrate that the proposed method improves network lifetime by up to 16%, enhances energy efficiency by over 9%, and significantly reduces communication overhead when compared to state-of-the-art approaches. This research sets a strong direction for scalable, decentralized energy replenishment in next-generation sensor networks. It lays the groundwork for resilient, efficient power management across diverse applications such as smart cities, environmental sensing and autonomous IoT deployments.

Keywords: wireless sensor networks; wireless charging; federated reinforcement learning; multi-agent reinforcement learning; energy efficiency.

1. INTRODUCTION

Wireless sensor networks (WSNs) have become integral to a wide range of modern applications, such as industrial automation, smart agriculture, environmental monitoring, healthcare systems and smart city infrastructures [1]. These networks comprise spatially distributed sensor nodes that perform data collection, processing and transmission to enable real-time and intelligent decision-making. One of the foremost challenges in sustaining WSN performance is energy depletion at sensor nodes, which directly affects the network's operational lifespan and reliability [2].

Traditional energy management techniques such as energy harvesting, energy-aware routing and static wireless charging have seen varied levels of success [3]. Among them, wireless charging using mobile chargers has gained attention for its flexibility and potential to recharge nodes on demand to extend lifetime of networks [4]. In this approach, MCs move through the network and wirelessly replenish the energy of critical nodes. However, efficiently scheduling the movement and charging paths of MCs is no trivial task due to the chargers' limited battery capacity and the dynamic nature of energy consumption in the network

Wireless recharging, particularly via mobile chargers [5], introduces flexibility by enabling targeted power delivery to sensor nodes based on their energy levels and spatial distribution. These systems typically employ inductive or resonant coupling techniques to transfer energy without physical contact, minimizing the need for manual intervention. However, the effectiveness of this approach is bounded by multiple operational constraints [6]. Mobile chargers have finite energy resources and must be intelligently scheduled to maximize network coverage and efficiency while avoiding redundant charging tasks [7]. Moreover, environmental factors, such as node density, obstacle placement and fluctuating communication loads, introduce additional complexity. To ensure timely and efficient energy replenishment, advanced optimization techniques [8, 9] are required to dynamically plan charger routes and prioritize nodes based on real-time energy demands and network topology [10]. This underlines the need for intelligent [11], adaptive frameworks [12] that can handle the inherent uncertainties and scalability challenges of large-scale WSNs [13]. Centralized approaches [14, 15] for optimizing MC schedules often face scalability issues, particularly in large-scale WSN deployments. These limitations highlight the need for a more distributed and scalable learning approach [16]. To overcome these limitations, federated reinforcement learning (FedRL) is proposed as a decentralized solution [17]. Unlike centralized RL approaches, where all data are sent to a central server for training, FedRL allows each mobile charger to train a local reinforcement learning model using its own observations. These models are periodically synchronized by sharing only model updates such as Q-values instead of raw data. This approach not only reduces communication overhead but also addresses data privacy concerns [17–19], particularly in sensitive applications such as

*e-mail: anusuyamathan.ece@gmail.com

Manuscript submitted 2025-05-23, revised 2025-09-09, initially accepted for publication 2025-09-22, published in January 2026.

military and healthcare monitoring [20]. FedRL leverages the distributed nature of WSNs, making it more adaptable to real-world constraints [21].

However, FedRL alone is insufficient for coordinated behavior among multiple chargers taken as multiple agent systems [22]. When chargers act independently, issues such as redundant charging of the same sensor node and inefficient path planning may arise. To address this, we incorporate multi-agent reinforcement learning (MARL), allowing mobile chargers to learn collaboratively [23]. MARL introduces a shared learning environment where agents not only optimize their own actions but also adapt to the behaviors of others in real time [24]. This inter-agent coordination significantly improves system-wide efficiency in dynamic and large-scale WSN deployments [25]. In MARL, each mobile charger acts as an autonomous agent, capable of learning optimal strategies while considering the actions of other chargers [26]. This cooperative behavior ensures better task allocation, prevents overlaps in charging efforts, and allows the system to respond effectively to fluctuating energy demands in the network. By combining FedRL with MARL, we propose an intelligent, privacy-preserving and coordinated charging system referred to as FedRL-MARL. This hybrid framework ensures that mobile chargers can learn optimal paths and schedules in a decentralized environment while effectively collaborating with other agents to enhance charging performance across the network.

1.1. Key contributions

We propose a federated learning (FL) approach for decentralized wireless charging in WSNs, ensuring privacy and reduced communication overhead. A MARL-based cooperative strategy enables multiple MCs to coordinate, minimizing redundancy and optimizing energy use. The framework supports real-time adaptability to dynamic energy demands and is lightweight enough for deployment on edge devices (Raspberry Pi, ESP32, Jetson), eliminating reliance on cloud infrastructure. Comprehensive simulations show superior performance over RL, FL and heuristic methods in charging efficiency, network lifetime and scalability. Unlike prior RL-based charging schemes that are either centralized or assume isolated learners, our framework jointly leverages FedRL and MARL to achieve privacy-preserving, decentralized training with cooperative execution among multiple MWCs.

Concretely, we: (i) share model updates only (no raw data) to reduce communication and protect privacy; (ii) use a coordinated multi-agent policy to prevent redundant charging and cut path overlap; and (iii) validate scalability under increasing node densities and multiple MWCs with obstacles. This pairing of FedRL aggregation with MARL-based cooperative execution for on-demand WRSN charging is, to our knowledge, not reported in prior works and yields measurable gains in network lifetime and energy efficiency.

The rest of this paper is structured as follows: Section 2 provides a comprehensive literature review of existing works on WSN charging strategies, federated learning and multi-agent RL. Section 3 and 4 describes the system model and detailed methodology of our proposed FedRL-MARL framework. Section 5 presents the experimental setup, performance metrics, and simulation results. Section 6 concludes the paper with future research directions.

2. LITERATURE REVIEW

Recent contributions can be broadly classified into the four domains listed below.

2.1. Heuristic / optimization-based charging strategies

These methods focus on charging tour planning, scheduling and resource allocation through mathematical programming and heuristic approaches. Bai *et al.* [27] introduced a guided search twin-dueling-double deep Q-network (GS-TD3QN) to optimize UAV flight paths, charging strategies and data upload intervals in AAV-assisted data collection. Jiang *et al.* [28] studied the charging sequence scheduling problem (CSCE) and developed an improved deep Q-network (IDQN-CSCE) for efficient and cost-effective scheduling. Ri *et al.* [29] proposed an integrated fuzzy cognitive network process and Q-learning-based scheduling (iFQS) to balance trade-offs among multiple performance metrics. Vuong *et al.* [30] presented an adaptive charging scheme using graph neural networks (GNNs) that significantly reduced sensor node failures.

2.2. RL based methods

These apply deep reinforcement learning to optimize charging order, routes and resource allocation without federated coordination. Jiang *et al.* [31] proposed a deep reinforcement learning with hybrid action space (DRLH-JCSCT) method combining DQN and DDPG to jointly optimize charging sequences and times, extending network lifetime. Sun *et al.* [26] introduced a UAV-based dynamic charging strategy employing the g-MAPPO algorithm to optimize UAV flight paths and energy allocation. Singh *et al.* [32] developed an age-aware UAV-aided energy harvesting framework for wireless rechargeable mobile networks (WRMNs), employing DDPG and MDP modeling to extend system longevity.

2.3. FL based methods

These methods emphasize privacy preservation and communication efficiency by enabling local training with aggregated model updates. Huang *et al.* [33] presented federated deep reinforcement learning (FDRL)-based joint energy replenishment and data gathering (FERG), improving energy efficiency and reducing latency. Li *et al.* [34] proposed a deep reinforcement learning-based dynamic charging-recycling scheme (DCRS) using DDQN to jointly optimize charging and recycling schedules, minimizing node failures and charging delays.

2.4. MARL based methods

MARL-based methods extend reinforcement learning into multi-agent coordination for distributed scheduling and cooperative charging. Liang *et al.* [35] introduced asynchronous and scalable multi-agent hybrid PPO (ASM-HPPO) for online scheduling with partial charging in multi mobile wireless chargers (MWC) scenarios, which dynamically adapts charging durations to enhance efficiency. Cao *et al.* [36] proposed the

MARLCS framework for underwater WRSNs, demonstrating improved survival rates and reduced charger energy consumption through cooperative strategies.

Table 1 classifies the existing works based on the strategies used and methods proposed in wireless recharging. Despite the significant advancements in WRSN charging strategies, many methods assume synchronized charging without considering the asynchronous nature of mobile chargers, leading to inefficient scheduling. Moreover, privacy concerns and high communication overhead are evident in centralized learning approaches. Some works neglect multi-agent collaboration, resulting in redundant charging or inefficient movement strategies. Additionally, traditional RL based models often fail to adapt dynamically to real-time sensor energy fluctuations and environmental changes. The proposed FedRL-MARL framework addresses these limitations by integrating privacy-preserving federated learning, decentralized cooperative charging, and adaptive scheduling strategies. Unlike the already existing methods, our approach ensures real-time decision-making, enhances multi-agent coordination and optimizes charging efficiency while reducing energy waste and communication costs. By leveraging FL, our model preserves data privacy while maintaining global learning capabilities, ensuring scalability and robustness in large-scale WRSN deployments [37, 38].

Table 1
Comparison between existing methods

Ref.	Optimization based	Learning based	Strategy	Contribution
[27]	×	✓	Periodic	UAV-assisted data collection
[31]	✓	×	Demand	Hybrid action space for charging
[26]	✓	×	Periodic	UAV path optimization
[28]	×	✓	Demand	Cost-efficient scheduling
[29]	✓	×	Demand	Fuzzy decision-making
[30]	✓	×	Demand	Adaptive charging
[35]	✓	×	Demand	Asynchronous MARL
[36]	×	✓	Periodic	Underwater charging
[33]	×	✓	Demand	FedRL for energy & data gathering
[34]	×	✓	Demand	Dynamic charging & recycling
[32]	×	✓	Periodic	UAV energy harvesting

3. NETWORK MODEL AND SYSTEM DESCRIPTION

The WSN in this study consists of N sensor nodes and M MWCs operating in a 2D plane with obstacles.

3.1. WSN initialization and parameters

Each sensor node has an initial energy level and consumes energy due to sensing, communication and processing tasks [1]. MWCs are introduced with limited energy reserves, requiring intelligent movement strategies to maximize efficiency. The deployment follows the following parameters:

Sensor node deployment: $S = \{S_1, S_2, \ldots, S_N\}$ where each sensor node s_i is randomly deployed within the environment.

Energy consumption model: each sensor node consumes energy due to communication energy C_{comm} and sensing energy C_{sensing}, residual energy at any given time t is: $E_i(t) = E_i(t) - C_{\text{comm}} - C_{\text{sensing}}$. If $E_i(t)$ drops below threshold E_{thresh}, the node is marked as critical and prioritized for charging.

Obstacle constraints: obstacles are represented as a set, $O = \{O_1, O_2, \ldots, O_L\}$, MWCs must navigate around obstacles using an optimized path-planning strategy.

Figure 1 represents a WSN with MWCs, where sensors have different energy levels (full, medium, low) and communicate with a central server to request charging. MWCs move along optimized paths to charge critical nodes while avoiding obstacles and maintaining communication with the depot. The dotted circles indicate the communication range of sensors.

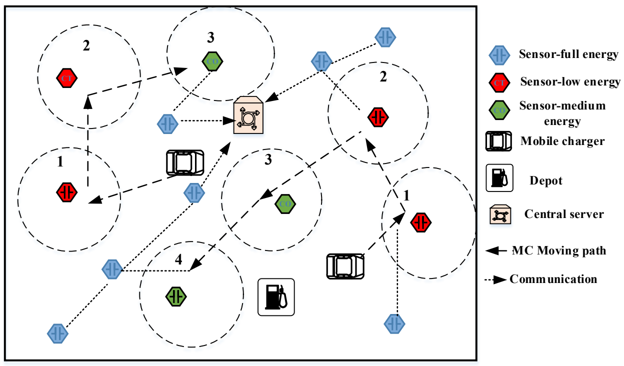


Fig. 1. WSN environment with wireless charging architecture

3.2. Reinforcement learning and federated multi-agent setup

To formulate the charging decision problem in WRSNs, each MWC is modeled as an autonomous RL agent. The process is expressed as MDP consisting of state space, action space and reward function, combined with federated learning for distributed coordination.

State space (S): (agent j at time t)

$$S_t^j = \Big\{p_{c_j}(t),\, E_j^{\text{MWC}}(t),\, \big\{p_{s_i}, E_i(t) \,|\, s_i \in N_j(t)\big\},\, Q_j(t), H_j(t)\Big\}. \quad (1)$$

In (1) above, $p_{c_j}(t)$ is 2D position of charger c_j at t, $E_i(t)$, $E_i^{\text{MWC}}(t)$ is residual energy of s_i and c_j, fixed position of sensor s_i; $N_j(t)$ set of sensor nodes within charging radius R_c of charger c_j, $Q_j(t)$ is task queue of pending tasks and $H_j(t)$ is a short history of task completions/failures.

Action space (A):

$$A = \Big\{\text{Move}_N, \text{Move}_S, \text{Move}_E, \text{Move}_W, \text{Charge}, \text{Idle}\Big\}. \quad (2)$$

As per (2), each MWC can move, charge or remain idle based on its learned policy.

Reward function: The per-step reward balances successful charging, energy consumption and task failures, as defined in (3).

$$R_t^j = \alpha . r_{c,t}^j - \beta . e_{t,j}^{\text{cost}} - \gamma . f_{\text{task,j}}^j , \tag{3}$$

where, $r_{c,t}^j$ is the number of successful charges attributable to c_j, $e_{t,j}^{\text{cost}}$ is energy consumed by c_j , $f_{\text{task,j}}^j$ denotes penalties for failed or uncompleted tasks.

Reward components:

To improve interpretability, the compact reward in (3) is decomposed into its components.

Charging reward as of (4) prioritizes charging of critically low-energy nodes.

Movement penalty as of (5) penalizes unnecessary travel based on distance.

Collision penalty as of (6) penalizes collisions with obstacles or other MWCs.

$$R_{t,j,i}^{\text{charge}} = \alpha_c (E_{\max} - E_i(t)), \tag{4}$$

$$R_{t,j}^{\text{move}} = -\beta_m d(p_{c_j}(t), p_{s_i}), \tag{5}$$

$$R_{t,j}^{\text{coll}} = -\gamma_c . \tag{6}$$

The total shaped reward, combining (4)–(6), is expressed in (7).

$$R_t^j = \sum_{i \in N_j(t)} R_{t,j,i}^{\text{charge}} + R_{t,j}^{\text{move}} + R_{t,j}^{\text{coll}} . \tag{7}$$

Policy optimization:

Each MWC aims to learn an optimal policy that maximizes the cumulative discounted reward, as shown in (8).

$$\pi_j^* = \arg\max_{\pi_j} E_{\pi_j} \left[\sum_{t=0}^{T} \gamma^t R_t^j \right]. \tag{8}$$

Local learning updates:

During training, MWCs iteratively update their models. This can be represented as a gradient update (see: (9a)) or equivalently as a Q-learning update (see: (9b)).

$$\theta_j \leftarrow \theta_j + \eta \nabla_{\theta_j} L_j(\theta_j), \tag{9a}$$

$$Q(s,a;\theta_j) \leftarrow Q(s,a;\theta_j) + \alpha_Q \left[r + \gamma \max_{a} Q\left(s', a'; \theta_j^-\right) - Q(s,a;\theta_j) \right]. \tag{9b}$$

Federated model aggregation:

To preserve privacy and reduce communication cost, MWCs share only their model parameters θ_j, which are aggregated using the FedAvg rule in (10).

$$Q_{\text{global}} = \frac{1}{M} \sum_{j=1}^{M} Q_j , \tag{10}$$

where M is the number of MWCs. The aggregated model is redistributed to MWCs to refine policies collaboratively.

Global (joint) optimization objective:

The collective objective of all MWCs is to maximize the total expected reward across the network, as defined in (11)

$$\max_{(\pi_j)_{j=1}^M} \sum_{j=1}^{M} E_{\pi_j} \left[\sum_{t=0}^{T} \gamma^t R_t^j \right]. \tag{11}$$

MARL coordination:

Although each MWC is trained independently, the framework is embedded within a MARL paradigm. MARL ensures cooperative behavior by applying the centralized training and decentralized execution (CTDE) principle: global updates are aggregated centrally using (10), while individual agents execute their policies locally as per (8). Conflict avoidance and fairness are enforced through the reward components calculated in (4)–(7), preventing redundant charging and ensuring balanced workload distribution across MWCs. Figure 2 represents a FedRL framework, where MWCs independently train on local environments and share only Q-value updates with a central server, ensuring data privacy and efficient global policy updates.

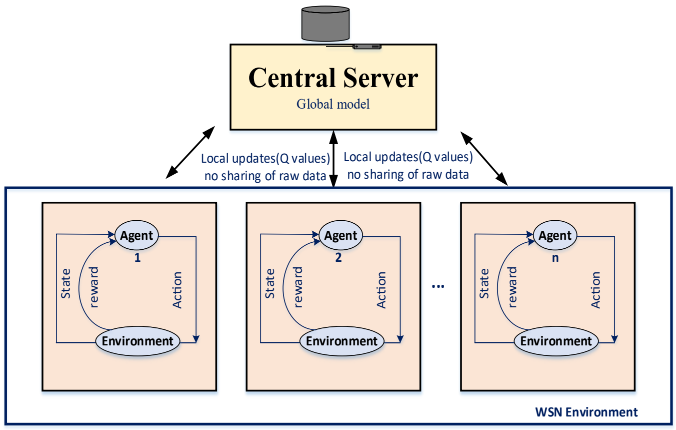


Fig. 2. Federated learning framework

4. PROPOSED METHODOLOGY

The proposed framework integrates FedRL with MARL to develop an efficient and adaptive wireless charging strategy for WSNs. Unlike traditional centralized charging methods that rely on a single decision-maker and require raw data transmission, this approach enables multiple MWCs to independently learn and make optimal charging decisions while sharing only model updates to enhance network longevity. It also reduces communication overhead and improves energy efficiency. The methodology is structured into three primary phases: (i) FedRL-based learning for global knowledge sharing, (ii) MARL-based cooperative charging for local decision-making, and (iii) an integration phase ensuring continuous learning and adaptation. Figure 3 illustrates the integration of FedRL and MARL, where MWCs collaboratively learn optimal charging strategies.

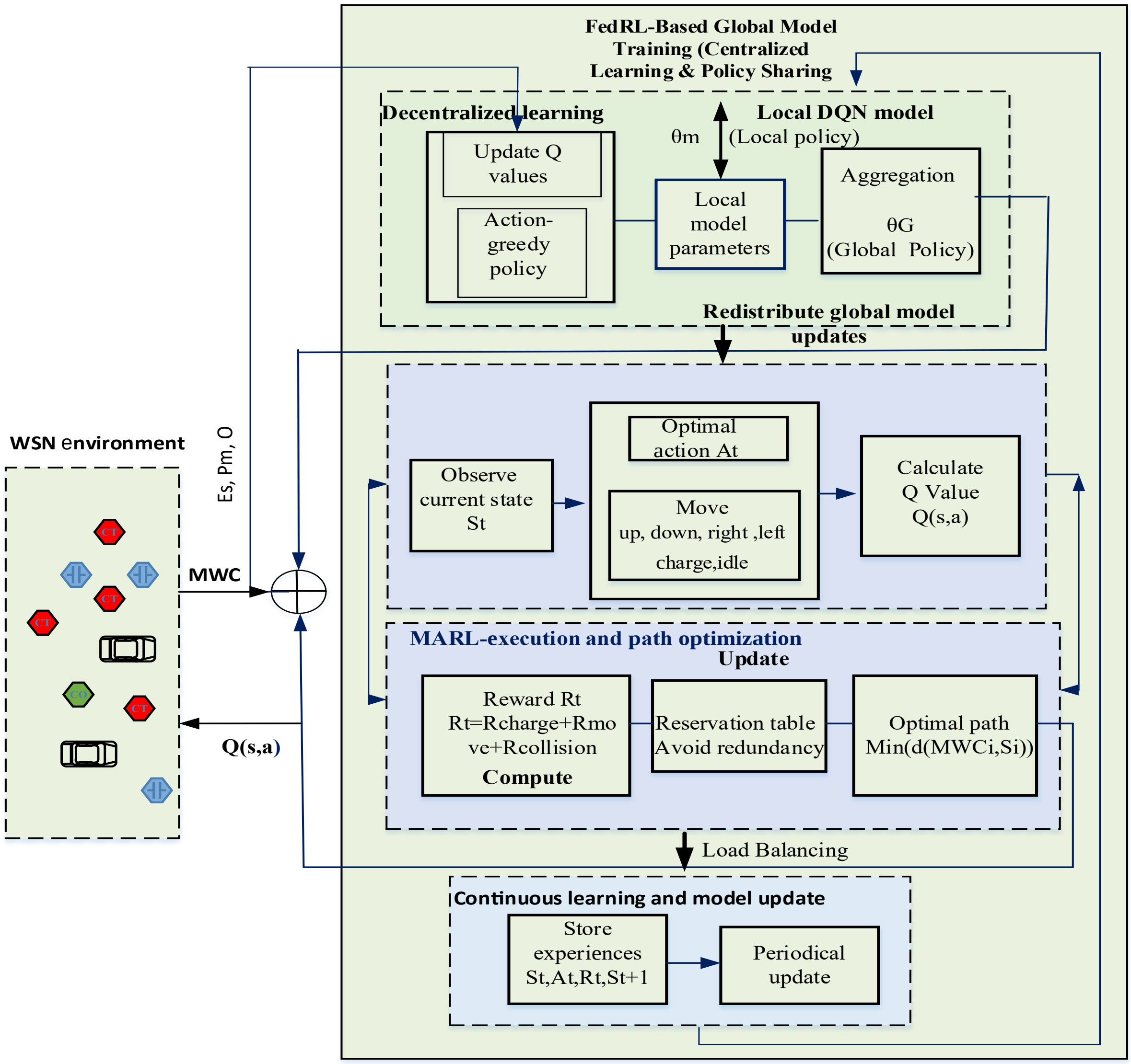


Fig. 3. Proposed architecture of FedRL-MARL framework

4.1. FedRL based Learning

In this phase, FedRL is employed to train a global model that enables MWCs to learn an optimal charging strategy without centralized data sharing [39]. Traditional RL requires a large amount of training data, often collected in a centralized manner, which is impractical in distributed WSNs due to communication constraints and privacy concerns.

FedRL overcomes these limitations by allowing MWCs to train models locally on their respective data and share only model updates with a global aggregator. Each MWC acts as an independent RL agent and follows the MDP formulation described in (1)–(7).

Once local training is completed, each MWC updates its local parameters θ_j using the learning rules in (9a)–(9b). These parameters are then sent to a global server, where the aggregation function FedAvg from (10) combines all received models into a single global model θ_G. This aggregated model is redistributed to all MWCs, improving their decision-making capability without requiring raw data exchange. This decentralized approach enhances data privacy, scalability, and adaptability to varying network conditions. By integrating local learning with federated aggregation, FedRL effectively balances computational efficiency with network-wide coordination.

4.2. MARL-based cooperative charging

After receiving the updated global model, MWCs enter the MARL-based cooperative charging phase. Here, multiple MWCs interact in real time to optimize charging schedules while avoiding redundancy. MARL ensures cooperation through the CTDE paradigm, where policies are trained with shared global knowledge as per (10) but executed locally according to the optimal policy, to be found in (8). In this phase, MWCs make charging decisions based on real-time observations, such as:

- energy demand of sensor nodes: low-energy nodes are prioritized (reward component in (4)).

- path optimization: MWCs minimize travel distance using the movement penalty as of (5).
- collision avoidance: MWCs avoid conflicts with obstacles or other chargers as of (6).

To illustrate this behavior, Table 2 provides example reward values for different actions. These numerical cases are consistent with the shaped reward design of (7).

Table 2
Reward function calculation for MWC action

Action (Ai)	C_i	E_i	Reward function R_t	Computed reward
Moves to a critical node & charges	5	10	$R_t = C_i - E_j$	−5
Moves to a critical node but fails to charge	0	8	$R_t = -E_j$	−8
Moves unnecessarily (no critical node nearby)	0	7	$R_t = -E_j$	−7
Stays idle when needed	0	0	$R_t = 0$	0
Charges a non-critical node	2	6	$R_t = C_i - E_j$	−4
Avoids obstacle while moving	–	2	$R_t = 2 - E_j$	0
Moves optimally & charges multiple nodes	8	12	$R_t = C_i - E_j$	−4

4.3. Decentralized FedRL-MARL integration

The integration of FedRL and MARL ensures that MWCs can effectively learn global charging strategies while executing decentralized, cooperative decision-making in real time.

FedRL phase: MWCs learn locally using state/action/reward definitions to be found in (1)–(7), and share only model updates, aggregated via (10).

MARL phase: MWCs execute charging decisions cooperatively, guided by the optimal policy in (8), while avoiding redundant charging through conflict-aware strategies.

Continuous adaptation: experiences collected during execution are fed back into FedRL, ensuring the global model evolves dynamically to reflect changing WSN conditions.

This iterative loop ensures scalability, robustness and efficiency in dynamic WSN environments, significantly improving network lifetime as compared to traditional approaches.

Algorithm 1. Federated RL training

Input: WSN environment, number of sensor nodes N, number of MWCs M, energy levels $E_s(i)$, learning rates l_{rQ}, l_{rn}, discount factor γ, batch size B_c, exploration rate ϵ, training episodes N_e.
Output: Optimized charging policy $\pi*(s,a)$ for MWCs.
Initialization:
Deploy N sensor nodes and M MWCs.
Initialize energy levels $E_s(i)$ for each sensor node S_i.
Set movement constraints and obstacles O.
Initialize local DQN models for MWCs with random weights θ_Q.
Initialize federated aggregation model with global weights θ_{global}.
Initialize experience replay buffer with capacity B_c.
Training Phase:
for episode = 1 to N_e **do**
Observe initial state $S(0)$ for all MWCs.
for step = 1 to K **do**
Update sensor node energy:
$E_s(i) = E_s(i) - C_{\text{comm}} - C_{\text{sensing}}$
if $E_s(i) \le E_{\text{thresh}}$, mark S_i as critical.
Compute state S_t (sensor energy, location, obstacles).
Compute action A_t (move, charge, idle) using ϵ-greedy policy.
Compute reward: $R_t = w_1 C - w_2 E$
Store (S_t, A_t, R_t, S_{t+1}) in replay buffer B_c.
Federated model aggregation:
if local training iteration is complete then
MWCs send Q-values to central server.
Aggregate global Q-values:
$Q_{\text{global}} = \frac{1}{M} \sum_{j=1}^{M} Q_j$
Redistribute updated model parameters to MWCs.
end for
Check convergence criteria; if met, proceed to Phase 2.

Figure 4 illustrates the workflow of the proposed FedRL–MARL charging strategy. After deploying sensor nodes and

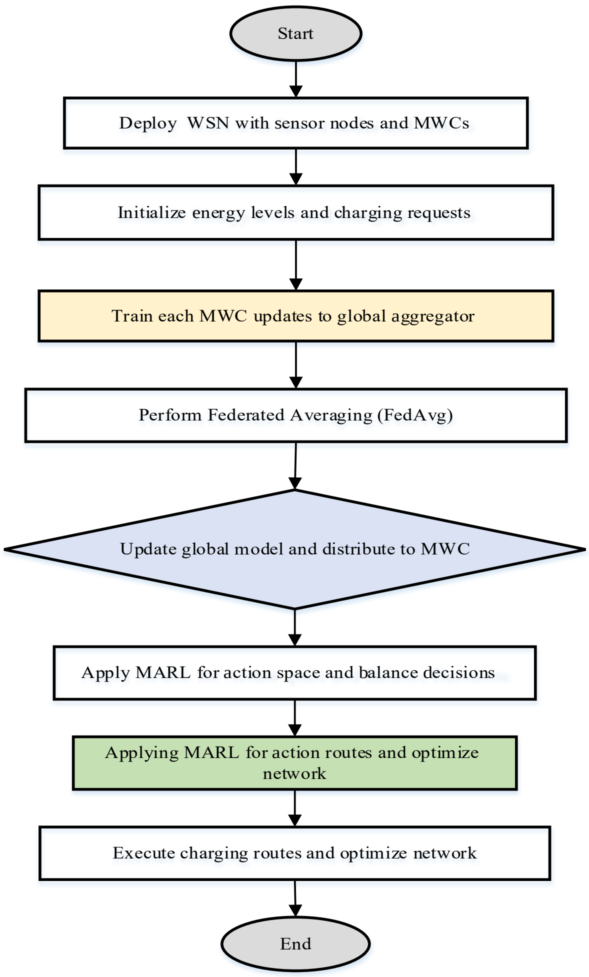


Fig. 4. Flowchart for FedRL-MARL execution

MWCs, the system monitors node energy levels and identifies charging needs. Each MWC trains a local RL model, which is aggregated using FedAvg as per (10) into a global model and redistributed for collective learning. Using this shared knowledge, MARL coordinates MWCs to balance charging loads and avoid conflicts. The updated policies are then executed to determine optimal charging routes, improving efficiency and extending WSN lifetime.

Algorithm 2. MARL-based execution & optimization

Execution & path planning:
for each MWC in deployment do
  Execute real-time actions based on learned policy.
  Compute optimal path to critical nodes:
    Optimal Path = $\min \sum_{j=1}^{M} d(C_j, S_i))$
Use reservation tables to avoid conflicts and redundant charging.
Policy refinement & learning update:
  Compute new reward function: $R_t = w_1 C - w_2 E$
  Update Q-learning policy:
    $Q(S,a) \leftarrow Q(S,a) + \alpha[R + \gamma \max Q(S',a') - Q(S,a)$
Check for convergence:
  if policy converges then terminate execution.
  else repeat MARL execution.
end for
Output: Optimized policy for MWCs ensuring efficient sensor node charging.

5. SIMULATION ANALYSIS AND DISCUSSION

Simulation of the FedRL-MARL framework is conducted in a WSN environment with varying sensor node densities and multiple MWCs. The environment consists of randomly deployed nodes that consume energy for sensing and communication, while MWCs navigate the network to recharge critical nodes avoiding obstacles. The simulation framework was developed in Python 3.10, utilizing OpenAI Gym for single-agent environment design and PettingZoo for MARL interactions. Reinforcement learning algorithms, including DQN, were implemented using Stable-Baselines3, which is fully based on PyTorch 1.13. Fed aggregation was carried out using a PyTorch/NumPy-based implementation of FedAvg, ensuring efficient synchronization of local and global models. The evaluation is conducted under three different scenarios, each characterized by increasing sensor node density and the number of MWCs. The simulation parameters are summarized in Table 3.

Table 4 presents the key hyperparameters for FedRL-MARL training. A learning rate ($\alpha = 0.0005$) and discount factor ($\gamma = 0.95$) ensure stable convergence, while the exploration rate ($\epsilon = 1.0$–0.05) is to balance exploration and exploitation. Training was performed with a batch size of 64, a replay buffer of 10 000, and 500 episodes per run. Federated aggregation was conducted every 50 episodes to synchronize local and global models. These values were selected through preliminary tuning and adapted from previous WRSN studies.

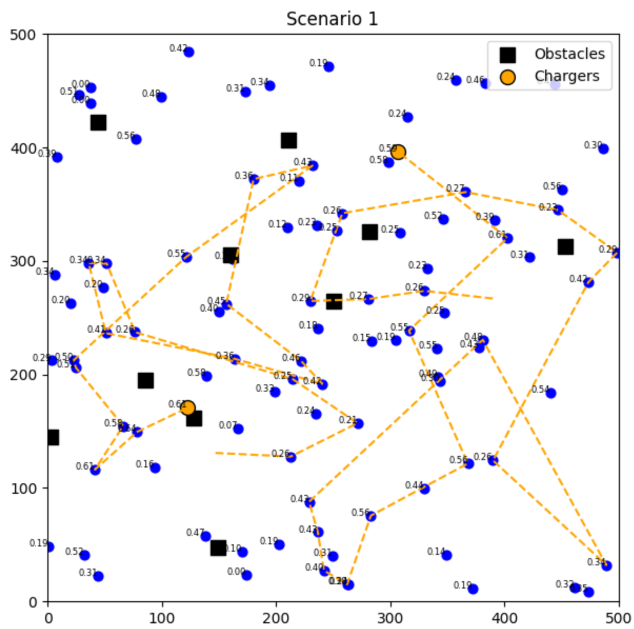
Figures 5a–5c show the MWC path trajectories in three WSN scenarios with obstacles. In Scenario 1 (100 nodes, 2 MWCs),

Table 3
Network simulation parameters

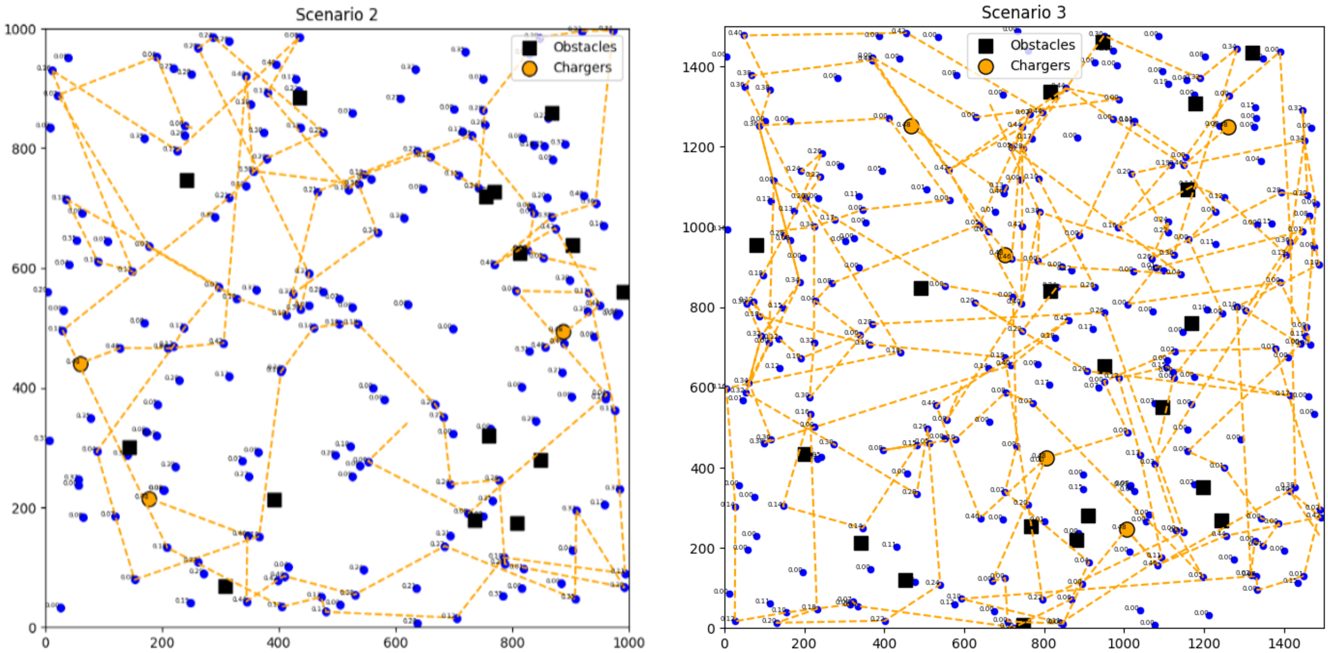
Parameter	Scenario 1	Scenario 2	Scenario 3
Simulation area (m^2)	500 × 500	1000 × 1000	1500 × 1500
Number of sensor nodes (N)	100	200	300
Number of MWCs (M)	2	3	5
Initial node energy J (Joules)	1.5 J	1.5 J	1.5 J
Charging efficiency (%)	88%	92%	94%
Communication overhead (MB)	5.2 MB	7.8 MB	9.5 MB
Obstacle density (%)	10%	15%	20%
Movement speed (m/s)	2 m/s	2.5 m/s	3 m/s
Threshold energy (J)	0.3 J	0.3 J	0.3 J

Table 4
Hyperparameters for training

Hyperparameter	Value / Setting
Learning rate (α)	0.0005
Discount factor (γ)	0.95
Exploration rate (ϵ)	1.0 → 0.05 (decay)
Batch size	64
Replay buffer size	10 000
Episodes per run	500
Federated aggregation rounds	Every 50 episodes



(a)

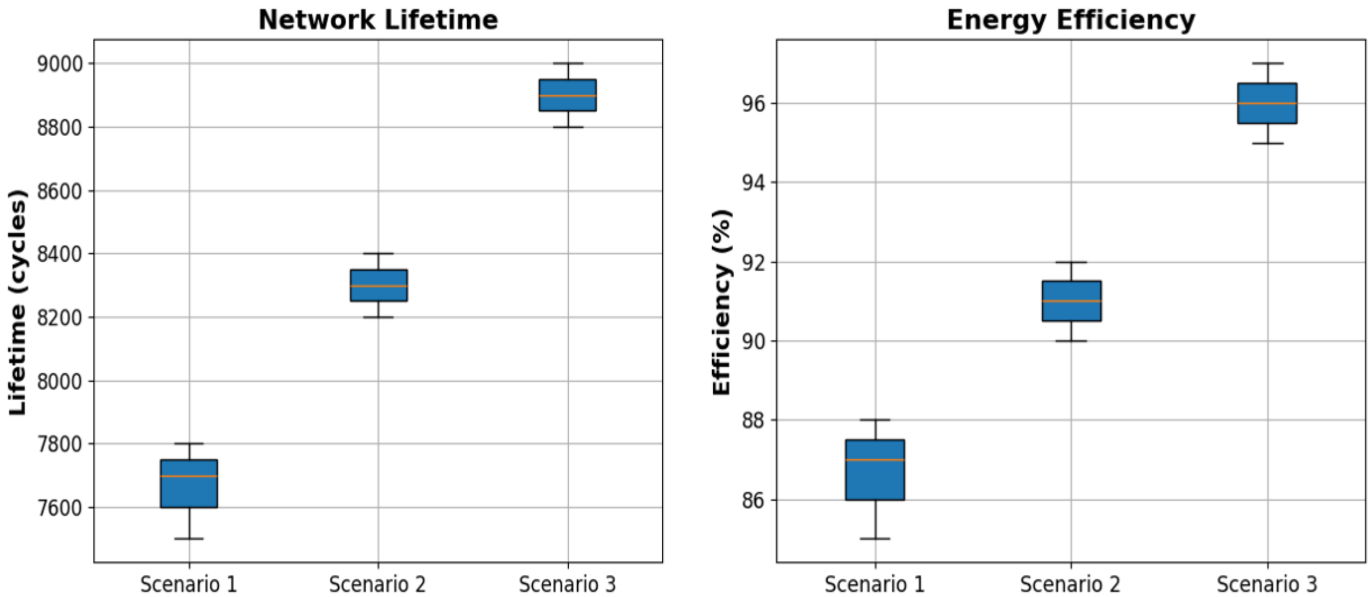


(b) (c)

Fig. 5. MWC path trajectories (a) Scenario 1, (b) Scenario 2, (c) Scenario 3

routes are short and efficiently cover the area. Scenario 2 (200 nodes, 3 MWCs) increases path complexity as MWCs coordinate to balance coverage. Scenario 3 (300 nodes, 5 MWCs) depicts dense trajectories in a constrained space, where MWCs adapt to obstacles and minimize redundant movements. Dotted paths highlight collaborative navigation for efficient charging.

Below, the visualization in Fig. 6 compares network lifetime and energy efficiency across three different scenarios. The first plot illustrates the variation in network lifetime, showing how sensor networks perform under different numbers of nodes and chargers. The second plot highlights the energy efficiency trends, indicating improvements in power utilization with optimized charging strategies. The results in Table 5 highlight the average charging distance, standard deviation and energy efficiency of each method across the three scenarios.



(a) (b)

Fig. 6. Performance comparison over 3 scenarios

Table 5
Metric evaluation of mean distance travelled by MWCs, variability in MWC travel distance and overall efficiency in recharging sensor nodes

Methods	Scenario 1 (100 nodes, 2 MWCs)			Scenario 2 (200 nodes, 3 MWCs)			Scenario 3 (300 nodes, 5 MWCs)		
Metrics	Avg. distance (m)	Std dev	Energy efficiency (%)	Avg. distance (m)	Std dev	Energy efficiency (%)	Avg. distance (m)	Std dev	Energy efficiency (%)
FedRL-MARL	24.75	0.72	88.32	39.12	1.20	89.45	50.89	2.55	90.72
DRLH-JCSCT	31.62	0.93	72.60	47.23	1.28	74.20	61.23	2.14	76.35
FDRL-JERDG	30.81	0.99	74.89	46.08	1.74	76.30	59.08	2.33	78.42
GS-TD3QN	28.06	0.87	80.97	42.02	1.43	82.12	55.19	2.83	83.95
UAV-gMAPPO	27.14	0.78	82.57	40.62	1.26	84.30	52.62	2.69	85.90
DCRS	25.91	0.74	84.47	38.63	1.22	86.69	48.63	2.62	88.58

5.1. Performance comparison over evaluation metrics

The comparative analysis of six different charging strategies including FedRL-MARL (proposed), DRLH-JCSCT [31], FDRL-JERDG [33], GS-TD3QN [27], UAV-gMAPPO [26], and DCRS [34], demonstrates superiority of the FedRL-MARL method. The performance of the proposed FedRL-MARL strategy is evaluated using key metrics, each quantified using the following formulas:

Dead node ratio: This metric indicates the proportion of sensor nodes that have exhausted their energy and are no longer functional. Here, N_{dead} is the number of sensor nodes with depleted energy, and N_{total} is the total number of deployed sensor nodes.

$$\text{DNR (\%)} = \frac{N_{\text{dead}}}{N_{\text{total}}} * 100. \tag{12}$$

Charging latency: This refers to the average time delay between a sensor node's request for charging and the actual initiation of charging. In the equation, R is the total number of charging requests, $t^i_{\text{Charge start}}$ is the time at which the i-th request was made, and t^i_{request} is the time at which charging started for that request.

$$\text{CL} = \frac{1}{R}\sum_{i=1}^{R}\left(t^i_{\text{Charge start}} - t^i_{\text{request}}\right). \tag{13}$$

Communication overhead: This measures the total data transferred in the network in megabytes (MB), accounting for packet size and count. Here, P_i is the number of packets sent by node i, and S_i is the average size of each packet in bytes.

$$\text{CO (MB)} = \sum_{i=1}^{N}\frac{P_i * S_i}{10^6}. \tag{14}$$

Network lifetime: This metric calculates the duration the WRSN remains operational from the initial deployment to a defined failure threshold. Here, T_{initial} is the deployment time and T_{final} is the time at which the network reaches failure.

$$\text{NL} = T_{\text{final}} - T_{\text{initial}}. \tag{15}$$

Energy efficiency: This denotes how effectively the MWCs utilize their energy to deliver usable energy to sensor nodes. $E_{\text{delivered}}$ is the total energy received by the sensor nodes, and E_{supplied} is the total energy output by the MWCs.

$$\text{EE (\%)} = \frac{E_{\text{delivered}}}{E_{\text{supplied}}} * 100. \tag{16}$$

Equations (12) to (16) define the dead node ratio, charging latency, communication overhead, network lifetime and energy efficiency, respectively. These formulas ensure consistent and accurate benchmarking of the proposed FedRL-MARL framework against baseline approaches. Hence, this method consistently outperforms existing charging strategies. Figures 7 to 12 highlight FedRL-MARL's superior performance across all evaluation metrics. It achieves the lowest dead node ratio, minimal charging latency and communication overhead, highest network lifetime and best energy efficiency with faster convergence as compared to existing methods.

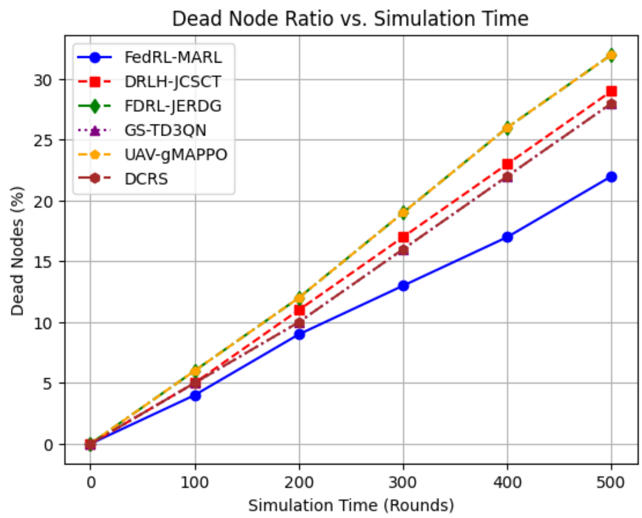


Fig. 7. Dead node ratio over time for various algorithms

Federated aggregation propagates effective policies globally without raw-data exchange, allowing MWCs trained in distinct local conditions to benefit from shared policy priors. The most pronounced gains appear when node density and obstacle complexity increase: (i) lifetime rises due to less idle travel and better

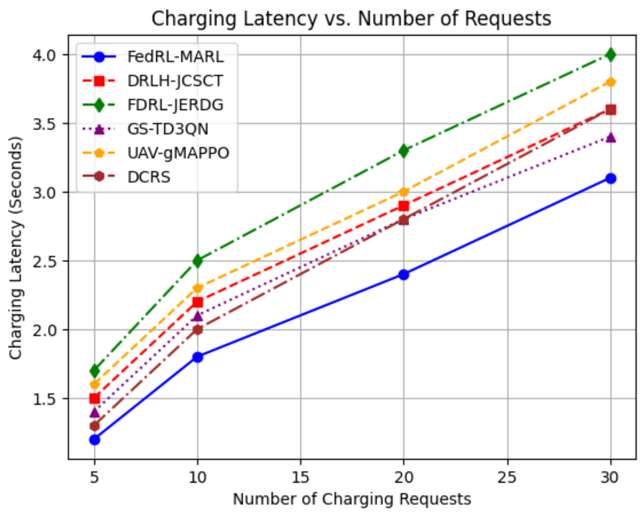


Fig. 8. Impact of request load on charging latency

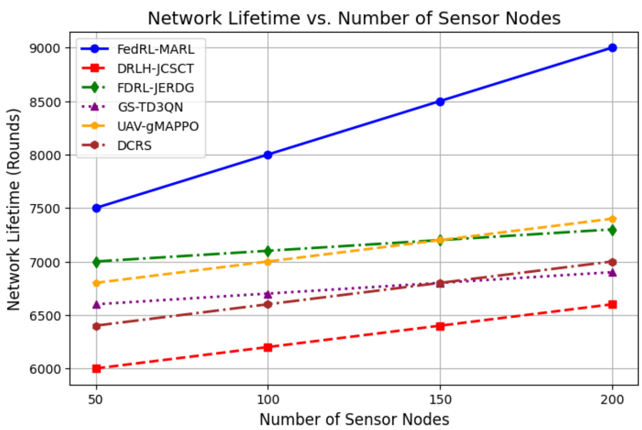


Fig. 9. Lifetime performance of charging strategies

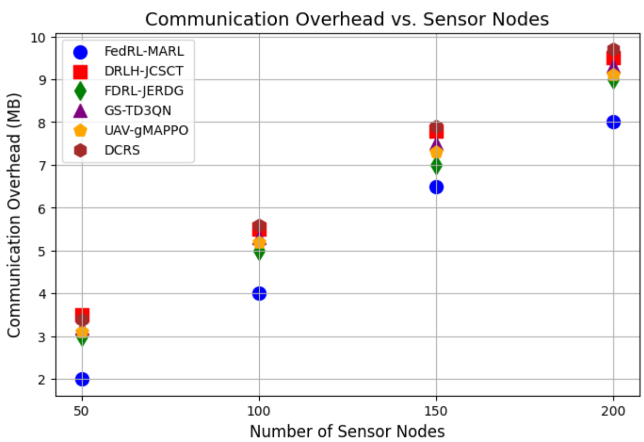


Fig. 10. Impact of node count on communication overhead

prioritization of critical nodes (ii) latency drops via conflict-aware routing, and (iii) energy efficiency improves as fewer "wasted" movements occur before a successful charge.

The superior performance of the FedRL-MARL framework

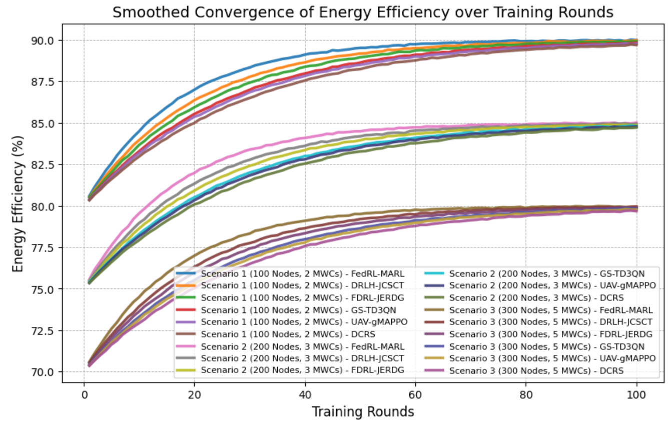


Fig. 11. Energy efficiency convergence across scenarios and charging methods

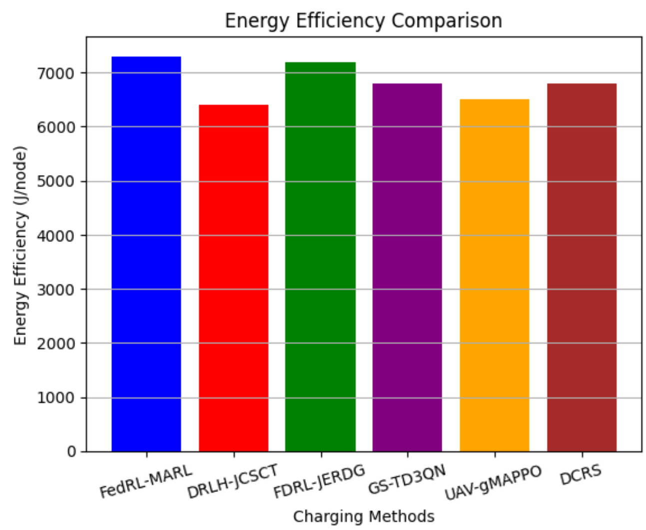


Fig. 12. Performance evaluation: energy efficiency of charging technique

arises from two key factors: (i) federated aggregation, which enables each MWC to benefit from the collective experiences of all agents without sharing raw data, thereby reducing redundancy and improving policy generalization; and (ii) multi-agent cooperation, which minimizes overlapping charging actions and ensures balanced energy distribution across nodes. In contrast, baseline methods such as DQN [40] and MARL-only approaches either suffer from overfitting to local experiences or from inefficient charger coordination. These results demonstrate that the integration of FL and MARL leads to a synergistic improvement, particularly in dense deployments with high obstacle density.

6. CONCLUSIONS

This paper proposed a FedRL-MARL framework for efficient wireless charging in WRSNs. By integrating federated aggregation with cooperative MARL decision-making, the framework enables distributed learning, privacy preservation, and effective coordination among multiple MWCs. Simulation results across three deployment scenarios demonstrated that the proposed approach improves network lifetime by up to 16% and energy efficiency by 9% as compared to baseline RL and heuristic methods. These improvements highlight the effectiveness of combining FL with MARL for real-world WRSNs. This study also establishes a scalable foundation for future intelligent energy management systems in dynamic WRSN environments. By integrating federated learning into real-time mobile charger coordination, our work opens new avenues for decentralized energy replenishment strategies, relevant for smart infrastructure, remote monitoring and autonomous IoT systems. Future work will focus on extending this framework to larger-scale networks, incorporating UAV-based charging strategies, and enhancing fault tolerance mechanisms to address charger or node failures. Such extensions will further improve the scalability, robustness and applicability of the proposed method in practical IoT and industrial WSN environments.

REFERENCES

[1] K. Gulati, R.S.K. Boddu, D. Kapila, S.L. Bangare, N. Chandnani, and G. Saravanan, "A review paper on wireless sensor network techniques in Internet of Things (IoT)," *Materi. Today-Proc.*, vol. 51, pp. 161–165, 2022, doi: 10.1016/j.matpr.2021.05.067.

[2] P. Anusuya, C. Vanitha, J. Cho, and S.V. Easwaramoorthy, "A comprehensive review of sensor node deployment strategies for maximized coverage and energy efficiency in wireless sensor networks," *Peer J. Comput. Sci.*, vol. 10, p. e2407, 2024, doi: 10.7717/peerj-cs.2407.

[3] J. Xu, X. Yuan, Z. Wei, J. Han, L. Shi, and Z. Lyu, "A wireless sensor network recharging strategy by balancing lifespan of sensor nodes," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–6.

[4] O. Banimelhem and S. Bani Hamad "A Proactive Charging Approach for Extending the Lifetime of Sensor Nodes in Wireless Rechargeable Sensor Networks," *J. Sens. Actuat. Netw.*, vol. 14, no. 2, p. 26, 2025, doi: 10.3390/jsan14020026.

[5] W. Yang *et al.*, "Robust wireless rechargeable sensor networks," *IEEE/ACM Trans. Netw.*, vol. 31, no. 3, pp. 949–964, 2022, doi: 10.1109/TNET.2022.3199389.

[6] M. Erol-Kantarci and H.T. Mouftah, "Suresense: sustainable wireless rechargeable sensor networks for the smart grid," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 30–36, 2012, doi: 10.1109/MWC.2012.6231157.

[7] X. Rao, P. Yang, Y. Yan, H. Zhou, and X. Wu, "Optimal recharging with practical considerations in wireless rechargeable sensor network," *IEEE Access*, vol. 5, pp. 4401–4409, 2017, doi: 10.1109/ACCESS.2017.2665471.

[8] A. Kaswan, P.K. Jana, and S.K. Das, "A survey on mobile charging techniques in wireless rechargeable sensor networks," *IEEE Commun. Surv. Tutorials*, vol. 24, no. 3, pp. 1750–1779, 2022, doi: 10.1109/COMST.2022.3189387.

[9] L. Fu, P. Cheng, Y. Gu, J. Chen, and T. He, "Optimal charging in wireless rechargeable sensor networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 278–291, 2015, doi: 10.1109/TVT.2015.2391119.

[10] N. Gharaei, Y.D. Al-Otaibi, S. Rahim, H.J. Alyamani, N.A. Khani, and S. Malebary, "Broker-based nodes recharging scheme for surveillance wireless rechargeable sensor networks," *IEEE Sens. J.*, vol. 21, no. 7, pp. 9242–9249, 2021, doi: 10.1109/JSEN.2021.3053203.

[11] H. Wang, J. Li, and W. Xiao, "Reinforcement learning-based charging cluster determination algorithm for optimal charger placement in wireless rechargeable sensor networks," *Ad Hoc Netw.*, vol. 164, p. 103605, 2024, doi: 10.1016/j.adhoc.2024.103605.

[12] J. Chen, X. Li, Y. Ding, B. Cai, J. He, and M. Zhao, "Charging Efficiency Optimization Based on Swarm Reinforcement Learning under Dynamic Energy Consumption for WRSN," *IEEE Sens. J.*, vol. 24, no. 20, pp. 33427–33441, 2024, doi: 10.1109/JSEN.2024.3407748.

[13] B. Qureshi *et al.*, "A state-of-the-art survey on wireless rechargeable sensor networks: Perspectives and challenges," *Wireless Netw.*, vol. 28, no. 7, pp. 3019–3043, 2022, doi: 10.1007/s11276-022-03004-x.

[14] B.H. Liu, N.T. Nguyen, V.T. Pham, and Y.X. Lin, "Novel methods for energy charging and data collection in wireless rechargeable sensor networks," *Int. J. Commun. Syst.*, vol. 30, no. 5, p. e3050, 2017, doi: 10.1002/dac.3050.

[15] S.M.A. Rahaman, M. Azharuddin, and P. Kuila, "Efficient scheduling of charger-uav in wireless rechargeable sensor networks: Social group optimization based approach," *J. Netw. Syst. Manage.*, vol. 32, no. 3, p. 55, 2024, doi: 10.1007/s10922-024-09833-9.

[16] R. Anusuya and D. Karthika Renuka, "FedAssess: analysis for efficient communication and security algorithms over various federated learning frameworks and mitigation of label-flipping attack," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 72, no. 3, 2024, doi: 10.24425/bpasts.2024.148944.

[17] S. Ali, Q. Li, and A. Yousafzai, "Blockchain and federated learning-based intrusion detection approaches for edge-enabled industrial IoT networks: A survey," *Ad Hoc Netw.*, vol. 152, p. 103320, 2024, doi: 10.1016/j.adhoc.2023.103320.

[18] J. Chen, H. Yan, Z. Liu, M. Zhang, H. Xiong, and S. Yu, "When federated learning meets privacy-preserving computation," *ACM Computing Surveys*, vol. 56, no. 12, pp. 1–36, 2024, doi: 10.1145/3679013.

[19] J. Pei, W. Liu, J. Li, L. Wang, and C. Liu, "A review of federated learning methods in heterogeneous scenarios," *IEEE Trans. Consum. Electron.*, vol. 70, no. 3, pp. 5983–5999, 2024, doi: 10.1109/TCE.2024.3385440.

[20] J. Wang, M.T. Quasim, and B. Yi, "Privacy-Preserving Heterogeneous Multi-Modal Sensor Data Fusion via Federated Learning for Smart Healthcare," *Inf. Fusion*, vol. 120, p. 103084, 2025, doi: 10.1016/j.inffus.2025.103084.

[21] S. Moreno-Álvarez, M.E. Paoletti, A.J. Sanchez-Fernandez, J.A. Rico-Gallego, L. Han, and J. Haut, "Federated learning meets remote sensing," *Expert Syst. Appl.*, vol. 255, p. 124583, 2024, doi: 10.1016/j.eswa.2024.124583.

[22] N.E.H. Dehimi, Z. Tolba, M. Medkour, A. Hadjadj, and S. Galland, "Improving testing of multi-agent systems: An innovative deep learning strategy for automatic, scalable, and dynamic error detection and optimisation," *Bull. Pol. Acad. Sci. Tech. Sci.*, vol. 73, no. 3, p. e154062, 2025, doi: 10.24425/bpasts.2025.154062.

[23] H. Moudoud, Z. Abou El Houda, and B. Brik, "Advancing security and trust in wsns: A federated multi-agent deep reinforcement

learning approach," *IEEE Trans. Consum. Electron.*, vol. 70, no. 4, pp. 6909–6918, 2024, doi: 10.1109/TCE.2024.3440178.

[24] M. Zhao, J. Li, X. Chai, and J. Li, "An efficient dynamic energy replenishment and data gathering strategy based on deep reinforcement learning in wireless rechargeable sensor networks," *Alexandria Eng. J.*, vol. 123, pp. 170–179, 2025, doi: 10.1016/j.aej.2025.03.069.

[25] Z. Ning, and L. Xie, "A survey on multi-agent reinforcement learning and its application," *J. Automat.*, vol. 3, no. 2, pp. 73–91, 2024, doi: 10.1016/j.jai.2024.02.003.

[26] H. Sun, J. Fei, G. Zhang, and M. Xie, "A multi-agent-based dynamic charging strategy for UAV-assisted wireless rechargeable sensor networks," *Comput. Netw*., vol. 259, p. 111081, 2025.

[27] S. Bai, X. Wang, M.C. Gursoy, G. Jiang, and S. Xu, "Deep Reinforcement Learning for Rechargeable UAV-Assisted Data Collection from Dense Mobile Sensor Nodes," *IEEE Access*, vol. 13, pp. 28398–28407, 2025, doi: 10.1109/ACCESS.2025.3539888.

[28] C. Jiang, W. Chen, J. Wang, Z. Wang, and W. Xiao, "An improved deep Q-network approach for charging sequence scheduling with optimal mobile charging cost and charging efficiency in wireless rechargeable sensor networks," *Ad Hoc Netw.*, vol. 157, p. 103458, 2024, doi: 10.1016/j.adhoc.2024.103458.

[29] M.G. Ri, C.H. Kim, S.H. Pak, and C.M. Pong, "iFQS: An Integrated FCNP-Q-Learning-Based Scheduling Algorithm for On-Demand Charging in Wireless Rechargeable Sensor Networks," *Int. J. Distrib. Sens. Netw.*, vol. 2024, no. 1, p. 4418058, 2024, doi: 10.1155/2024/4418058.

[30] A.D. Vuong *et al.*, "An adaptive charging scheme for large-scale wireless rechargeable sensor networks inspired by deep Q-network," *Neural Comput. Appl.*, vol. 36, no. 17, pp. 10015–10030, 2024, doi: 10.1007/s00521-024-09658-2.

[31] C. Jiang, W. Chen, X. Chen, S. Zhang, and W. Xiao, "Deep reinforcement learning approach with hybrid action space for mobile charging in wireless rechargeable sensor networks," *Expert Syst. Appl.*, vol. 249, p. 123752, 2024, doi: 10.1016/j.eswa.2024.123752.

[32] A. Singh and R.M. Hegde, "Age-Aware UAV-Aided Energy Harvesting for the Design of Wireless Rechargeable Mobile Networks," *IEEE Trans. Artif. Intell.*, vol. 6, no. 7, pp. 1797–1807, 2025, doi: 10.1109/TAI.2025.3528377.

[33] H. Huang, J. Zhang, B. Wang, W. Miao, and G. Min, "Joint Mobile Energy Replenishment and Data Gathering in Wireless Sensor Networks via Federated Deep Reinforcement Learning," *IEEE Trans. Mob. Comput.*, vol. 24, no. 7, pp. 6460–6473, 2025, doi: 10.1109/TMC.2025.3543009.

[34] L. Li, Y. Feng, N. Liu, Y. Li, and J. Zhang, "Deep Reinforcement Learning Based Dynamic Charging-Recycling Scheme for Wireless Rechargeable Sensor Networks," *IEEE Sens. J.*, vol. 24, no. 9, pp. 15457–15471, 2024, doi: 10.1109/JSEN.2024.3380592.

[35] Y. Liang, H. Wu, and H. Wang, "Asynchronous multi-agent reinforcement learning for collaborative partial charging in wireless rechargeable sensor networks," *IEEE Trans. Mob. Comput.*, vol. 23, no. 5, pp. 4806–4818, 2023, doi: 10.1109/TMC.2023.3299238.

[36] J. Cao, J. Liu, J. Dou, C. Hu, J. Cheng, and S. Wang, "Multi-agent reinforcement learning charging scheme for underwater rechargeable sensor networks," *IEEE Commun, Lett.*, vol. 28, no.'3, pp. 508–512, 2023, doi: 10.1109/LCOMM.2023.3345362.

[37] S.K. Das, R. Mudi, M.S. Rahman, K.M. Rabie, and X. Li, "Federated Reinforcement Learning for Wireless Networks: Fundamentals, Challenges and Future Research Trends," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 1400–1440, 2024, doi: 10.1109/OJVT.2024.3466858.

[38] P. Wang*et al.*, "Decentralized navigation with heterogeneous federated reinforcement learning for UAV-enabled mobile edge computing," *IEEE Trans. Mob. Comput.*, vol. 23, no. 12, pp. 13621–13638, 2024, doi: 10.1109/TMC.2024.3439696.

[39] Y. Gao, M. Liu, X. Yuan, Y. Hu, P. Sun, and A. Schmeink, "Federated deep reinforcement learning based trajectory design for UAV-assisted networks with mobile ground devices," *Sci. Rep.*, vol. 14, no. 1, p. 22753, 2024, doi: 10.1038/s41598-024-72654-y.

[40] S.P.R. Banoth, P.K. Donta, "Dynamic mobile charger scheduling with partial charging strategy for WSNs using deep-Q-networks," *Neural Comput. Appl.*, vol. 33, no. 22, pp. 15267–15279, 2021, doi: 10.1007/s00521-021-06146-9.