Electric Vehicles Charging and Discharging Strategies Based on Federated Learning

Zhanlian Li[1], Xiaoning Zhao[2], Jiehong Ye[2], Mengxuan Yan[1,3],Chun Sing Lai[1,4,*]

[1]School of Automation, Guangdong University of Technology, Guangzhou, China
[2]Guangdong Power Grid Corporation, China Southern Power Grid, Guangzhou, China
[3]Jiangmen Power Supply Bureau of Guangdong Power Grid Co Ltd., Jiangmen, China
[4]Brunel University of London, London, UK
E-MAIL: 1112104006@mail2.gdut.edu.cn, 18602068850m@sina.cn, jiehong_ye@163.com, 1122004002@mail2.gdut.edu.cn, chunsing.lai@brunel.ac.uk

Abstract:

Electric vehicles (EVs) are the key solution for achieving zero-carbon emissions in transportation and play an essential role in global climate change mitigation. However, large-scale EV integration introduces major challenges for power grid stability and energy management due to increasingly complex and decentralized charging and discharging behaviors. Traditional centralized optimization methods face difficulties in addressing privacy concerns and meeting the scalability needs of practical applications. To overcome these issues, this paper proposes a federated learning-based strategy that enables collaborative optimization among multiple EVs or charging stations without sharing raw data, thus protecting user privacy. By leveraging distributed data and computational resources, this approach adapts to diverse user behaviors and grid conditions and reduces communication overhead. Through a simulated case, the effectiveness of the model has been verified, which also demonstrates strong potential as a scalable, privacy-preserving framework for intelligent EV charging and discharging management in future zero-carbon energy systems.

Keywords:

Electric vehicles (EVs); Zero-carbon emissions; Federated learning; Power grid stability; Privacy-preserving optimization

1 Introduction

The rapid electrification of the transportation sector is a key strategy for reducing greenhouse gas emissions [1]. This transition is essential for achieving global carbon neutrality targets [2]. Electric vehicles (EVs) are being integrated into energy and mobility systems at an unprecedented rate [3]. This growth is driven by technological advances and supportive public policies. Many studies have shown that widespread EV adoption improves urban air quality [4]. It also accelerates progress toward decarbonization goals at both national and regional levels. Despite these benefits, large-scale EV deployment introduces significant challenges for power systems[5]. These challenges include increased peak loads, voltage fluctuations, and greater privacy concerns [6]. Therefore, it is crucial to optimize charging and discharging strategies for EVs. This optimization supports the sustainability, resilience, and efficiency of modern energy infrastructure.

Most traditional approaches to EV charging and discharging optimization use centralized architectures [7]. These methods rely on techniques such as genetic algorithms, particle swarm optimization[8], and reinforcement learning (RL) [9]. Classical optimization methods include linear programming, mixed-integer programming, and heuristic algorithms. Centralized schemes often face risks such as privacy leakage and single points of failure. Distributed optimization algorithms can help address these issues[10]. However, they still encounter difficulties with coordination efficiency, communication overhead, and scalability. These problems are especially pronounced in heterogeneous and dynamic environments. With the development of artificial intelligence, recent advances in intelligent algorithms have improved EV charging and discharging management. Reinforcement learning (RL) and deep reinforcement learning (DRL) are highly adaptable[11]. They perform well under dynamic pricing, uncertain demand, and complex system constraints. Policy gradient, Q-learning, and actor-critic algorithms have achieved strong results in real-time scheduling and decision-making[12]. However, these approaches usually require sharing large amounts of data and centralized training. This can compromise user privacy and system security.

Federated learning (FL) is a promising solution for privacy-preserving and scalable collaborative optimization in smart grids [13]. FL allows multiple agents to train models together without sharing raw data. This method addresses privacy and security concerns directly. When federated learning is combined with reinforcement learning, the result is called federated reinforcement learning (FedRL). FedRL increases adaptability and robustness in distributed optimization. It works well in dynamic and uncertain conditions. FedRL has shown potential in EV charging and discharging scenarios. It helps balance global system objectives with local user preferences. However, current solutions are still limited in managing large-scale and heterogeneous EV networks. The challenges become greater when user behavior changes rapidly or electricity prices fluctuate. It is difficult to achieve optimal performance, privacy protection, and computational efficiency at the same time. These frameworks should use the strengths of federated learning and advanced reinforcement learning algorithms for EV charging and discharging optimization.

This paper presents FedRL-EV, which stands for Federated Reinforcement Learning for Electric Vehicles. FedRL-EV is a distributed optimization framework for EV charging and discharging. It is based on federated reinforcement learning. The FedRL-EV framework integrates federated learning with policy gradient reinforcement learning. This combination enables efficient, privacy-preserving, and scalable scheduling across multiple charging stations. The main contributions of this work are as follows. First, a FedRL-EV framework supporting collaborative learning among multiple charging stations is designed to protect user privacy. Second, a policy gradient-based charging and discharging scheduling algorithm is developed, tailored for large-scale and dynamic scenarios. Third, comprehensive simulation studies are conducted to validate the effectiveness and scalability of the proposed approach. The rest of this paper is structured as follows. Section 2 describes the problem formulation and methodology. Section 3 presents the experimental results and analysis. Section 4 concludes the paper and discusses future research directions.

2 Problem Formulation

This section presents the collaborative charging and discharging optimization framework for multiple electric vehicles (EVs) based on federated learning. The proposed approach emphasizes distributed, privacy-preserving, and scalable multi-agent coordination under time-of-use electricity pricing, aiming to maximize overall system economic benefits while accounting for battery aging and operational constraints.

2.1 Research Strategy

A federated multi-agent learning paradigm is adopted, where each EV independently trains its local policy and only uploads model parameters to a central aggregator. The methodology consists of the following key components:

2.1.1 Federated Multi-Agent Modeling

The multi-EV charging/discharging problem is formulated as a federated multi-agent decision process. Each agent i observes its local state $s_{i,t}$ and independently updates its policy π_{θ_i}. Agents periodically upload their model parameters to a central server, which aggregates them to form a new global model, subsequently distributed to all agents for further local training. The aggregation rule is as follows:

$$\theta_{k+1}^{(g)} = \mathcal{A}\left(\{\theta_{k+1}^{(i)}\}_{i=1}^{N}\right) \tag{1}$$

where $\theta_{k+1}^{(g)}$ denotes the global model parameters after the $(k+1)$-th aggregation, $\theta_{k+1}^{(i)}$ is the locally trained model parameters of agent i, N is the total number of agents, and $\mathcal{A}(\cdot)$ represents the aggregation operator (e.g., weighted averaging).

2.1.2 Privacy-Preserving Collaborative Optimization

To ensure data privacy, all raw states, actions, and reward trajectories remain local, in which only model weight differences or gradients are exchanged. The federated optimization objective minimizes the average loss across all agents:

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}_{s_{i,t},a_{i,t}} \left[\mathcal{L}_i \left(\pi_\theta(a_{i,t}|s_{i,t})\right)\right] \tag{2}$$

where θ denotes the globally shared policy parameters, $\mathcal{L}_i(\cdot)$ is the local loss function for agent i, $s_{i,t}$ and $a_{i,t}$ are the local state and action at time t, $\pi_\theta(a_{i,t}|s_{i,t})$ is the probability of taking action $a_{i,t}$ in state $s_{i,t}$ under policy parameter θ, and $\mathbb{E}$ denotes the expectation operator.

2.1.3 Simulation and Validation

A distributed simulation environment is constructed, with each EV agent interacting with its own local environment. The system supports asynchronous participation, heterogeneous data distributions, and dynamic fleet sizes. The performance of the federated approach is benchmarked against centralized and isolated (non-collaborative) optimization schemes.

2.2 Model Preprocessing

2.2.1 Local State Representation

The local state of each agent at time t is encoded as:

$$s_{i,t} = [\phi_1(t), \phi_2(SOC_{i,t}), \phi_3(p_t), \phi_4(\Delta E_{i,t}), \phi_5(\eta_{i,t})] \quad (3)$$

where $s_{i,t}$ is the local state vector of agent i at time t, $\phi_1(t)$ is a time feature extraction function, $\phi_2(SOC_{i,t})$ is the normalized state-of-charge (SOC) feature, $\phi_3(p_t)$ is the time-of-use price feature, $\phi_4(\Delta E_{i,t})$ is the energy increment feature, and $\phi_5(\eta_{i,t})$ is the battery health feature.

2.2.2 Action Space

The action space for each agent is defined as:

$$a_{i,t} \in \mathcal{A} = \{\text{charge}, \ \text{idle}, \ \text{discharge}\} \quad (4)$$

where $a_{i,t}$ denotes the action chosen by agent i at time t, and $\mathcal{A}$ is the set of possible actions: charging, idling, and discharging. Each action corresponds to a fixed power level, such as P_{charge} and $P_{discharge}$ for charging and discharging power (e.g., 50 kW).

2.2.3 Constraint Handling

All agent actions must satisfy the following local and global constraints:

$$SOC_{min} \le SOC_{i,t+1} \le SOC_{max} \quad (5)$$

$$\left| \sum_{i=1}^{N} P_{i,t} \right| \le P_{grid}^{\max} \quad (6)$$

$$SOC_{i,t_{commute}} \ge SOC_{commute} \quad (7)$$

where SOC_{min} and SOC_{max} are the minimum and maximum allowable state-of-charge, $SOC_{i,t+1}$ is the SOC of agent i at time $t+1$, $P_{i,t}$ is the charging/discharging power of agent i at time t, $P_{grid}^{\max}$ is the maximum grid power limit, $SOC_{i,t_{commute}}$ is the SOC of agent i at the required departure time, and $SOC_{commute}$ is the minimum SOC required for commuting.

2.2.4 Federated Experience Replay

Each agent locally maintains experience tuples $(s_{i,t}, a_{i,t}, r_{i,t}, s_{i,t+1})$ without sharing any raw data with other agents or the aggregator.

2.3 Federated Policy Update Mechanism

2.3.1 Local Policy Optimization

Each agent optimizes the following surrogate objective using its local trajectory data:

$$\mathcal{J}_i(\theta) = \mathbb{E}_{\tau_i} \left[\sum_{t=0}^{T-1} \gamma^t r_{i,t} \right] \quad (8)$$

where $\mathcal{J}_i(\theta)$ is the local policy objective for agent i, τ_i is the complete trajectory of agent i, T is the decision horizon, γ is the discount factor ($0 < \gamma \le 1$), and $r_{i,t}$ is the immediate reward at time t.

2.3.2 Federated Aggregation Protocol

After E rounds of local updates, agents upload their parameters $\theta^{(i)}$ to the server, which aggregates them using the FedAvg protocol:

$$\theta^{(g)} \leftarrow \sum_{i=1}^{N} \frac{n_i}{\sum_j n_j} \theta^{(i)} \quad (9)$$

where $\theta^{(g)}$ denotes the aggregated global model parameters, n_i is the number of training samples used by agent i in this round, $\theta^{(i)}$ is the locally updated model parameters of agent i, and $\sum_j n_j$ is the total number of samples across all agents. Here, the left arrow "$\leftarrow$" indicates an assignment or update operation, i.e., the right-hand side aggregation result is assigned to $\theta^{(g)}$ as the new global parameter, rather than representing a mathematical equality. This notation is widely adopted in algorithms and federated learning literature to distinguish variable updates from equalities.

2.4 Battery Aging and Reward Modeling

2.4.1 Battery Aging Penalty

Battery aging is penalized via a differentiable cost function:

$$\mathcal{C}_{age,i,t} = \kappa \cdot \psi(P_{i,t}, T_{i,t}, SOC_{i,t}) \quad (10)$$

where $\mathcal{C}_{age,i,t}$ is the battery aging cost for agent i at time t, κ is the cost scaling coefficient, $\psi(\cdot)$ is a nonlinear function reflecting the impact of power, temperature, and SOC on aging, and $T_{i,t}$ is the battery temperature.

2.4.2 Composite Reward Function

The immediate reward for agent i at time t is defined as:

$$r_{i,t} = \lambda_1 \cdot \Pi_{i,t} - \lambda_2 \cdot \mathcal{C}_{age,i,t} - \lambda_3 \cdot \mathbb{I}_{\mathcal{C}}(a_{i,t}) \tag{11}$$

where $r_{i,t}$ is the immediate reward, $\lambda_1, \lambda_2, \lambda_3$ are weighting coefficients, $\Pi_{i,t}$ is the net economic benefit, $\mathcal{C}_{age,i,t}$ is the battery aging cost, and $\mathbb{I}_{\mathcal{C}}(a_{i,t})$ is an indicator function that equals 1 if action $a_{i,t}$ violates constraints, and 0 otherwise.

2.5 Evaluation Metrics

The following metrics are used to evaluate the performance of the federated learning framework:

Total Net Profit

$$\mathcal{P}_{FL} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\Pi_{i,t} \tag{12}$$

where $\mathcal{P}_{FL}$ is the average total net profit under federated learning, N is the number of agents, T is the simulation horizon, and $\Pi_{i,t}$ is the net profit of agent i at time t.

Total Battery Aging Cost

$$\mathcal{A}_{FL} = \sum_{i=1}^{N}\sum_{t=1}^{T}\mathcal{C}_{age,i,t} \tag{13}$$

where $\mathcal{A}_{FL}$ is the total battery aging cost across all agents, and $\mathcal{C}_{age,i,t}$ is the battery aging cost for agent i at time t.

Constraint Violation Rate

$$\mathcal{V}_{FL} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{\mathcal{C}}(a_{i,t}) \tag{14}$$

where $\mathcal{V}_{FL}$ is the constraint violation rate, and $\mathbb{I}_{\mathcal{C}}(a_{i,t})$ indicates whether the action of agent i at time t violates constraints.

3 Case Studies

To validate the effectiveness of the proposed federated reinforcement learning (FLRL) based multi-EV charging strategy, comprehensive case studies were conducted using simulated data. The federated framework enables distributed learning and decision-making across multiple EVs, preserving data privacy while achieving coordinated optimization. The test environment setup and simulation parameters are summarized in Table 1, with federated learning specific parameters highlighted.

The case study implemented a scalable policy gradient framework within a federated learning architecture. Each EV agent performs local policy updates using its own data, and periodically synchronizes model parameters with a central aggregator via the FedAvg algorithm. The simulation environment is implemented in Python with TensorFlow 2.10.1 and GPU acceleration. The policy network adopts a three-layer architecture (128-64-3) with experience replay, ensuring stable convergence and efficient training. The federated approach demonstrated significant economic benefits and privacy preservation, with each EV achieving an average daily profit of 60.08 yuan by strategically charging during off-peak (23:00-07:00) and discharging during peak (10:00-14:00, 17:00-21:00) periods. The charging control system operates at a 15-minute resolution, with 50 kW power and 20%-90% SOC constraints, optimizing charging for 10 & 20 Tesla EVs (78.4 kWh battery) while considering battery degradation via an integrated aging model.

3.1 Electricity Pricing

The time-of-use (TOU) electricity pricing scheme is characterized by distinct peak and off-peak periods. The federated strategy leverages this structure by coordinating EV charging and discharging actions among clients. Specifically, EVs are scheduled to charge during off-peak hours (23:00–07:00) and discharge during peak periods (10:00–14:00, 17:00–21:00). This coordination is achieved without exchanging raw data between EVs, thus safeguarding user privacy. At the same time, such a strategy contributes to grid stability by smoothing the overall load profile. The adopted TOU pricing profile is depicted in Fig. 1(b).

TABLE 1. Simulation Parameters

Parameter	Value
Number of EVs (N)	1, 10, 20
Battery capacity (E_{cap})	78.4 kWh
Charging power (P_{rated})	50 kW (fixed per action)
Charging efficiency (η)	0.92
Initial SOC range	[20%, 80%]
Minimum SOC (SOC_{min})	20%
Maximum SOC (SOC_{max})	90%
Time step per episode (T)	96 (15 min per step)
Trip energy consumption	3.6 kWh (per trip)
Federated aggregation interval (FL)	Every 2 local epochs
Local training epochs (FL)	5
Policy network architecture	128-64-3 (hidden layers)
Learning rate (α)	0.001
Discount factor (γ)	0.99
Training episodes	20, 500
Max grid power ($P_{grid}^{\max}$)	500 kW
Commuting SOC requirement(SOC_{com})	40%
Reward weights (λ_1, λ_2, λ_3)	1, 0.2, 10
Battery aging coefficient (κ)	0.05

3.2 Results and Analysis

Based on the data presented in Fig. 1 and Table 2, the performance of FedRL and traditional RL in terms of Net Profit varies across different EV fleet sizes and training episodes. In the single-vehicle and medium-scale (10 EVs) scenarios, FedRL demonstrates improvements of 18.1% and 8.6%, respectively, over RL after 20 training episodes. This suggests that distributed information integration and collaborative learning can enhance overall economic returns in the early stages of training. As the number of training episodes increases to 500, the Net Profit gap narrows, with FedRL maintaining only a slight advantage (1.9% and 3.4% for 1 and 10 EVs, respectively). This indicates that, under sufficient data and prolonged training, the impact of privacy protection on policy optimization becomes limited, and both individual and global strategies tend to converge.

In the large-scale scenario (20 EVs), FedRL initially outperforms RL by 6.1% after 20 episodes. However, as training continues to 500 episodes, RL surpasses FedRL, with Net Profit for FedRL dropping to 729.27 yuan, compared to 771.70 yuan for RL. This reversal highlights that, as the fleet size increases and training extends, the privacy-preserving mechanisms of federated learning restrict the depth of information integration necessary for achieving global optimality. Model-level parameter aggregation can-

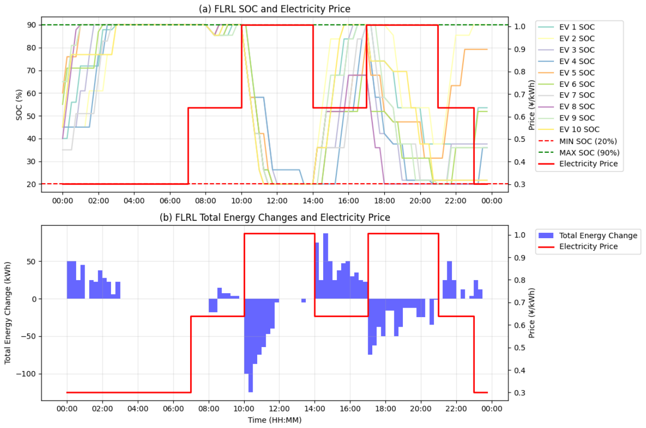

FIGURE 1. Optimal EV Charging/Discharging Strategy under Federated Reinforcement Learning

TABLE 2. Comparison of RL and Federated RL Performance Metrics

Episode	EV Number	Net Profit (¥)		Total Cost (¥)		Total Revenue (¥)	
		RL	FedRL	RL	FedRL	RL	FedRL
20	1 EV	48.63	57.46	39.50	43.26	88.13	100.72
20	10 EVs	353.33	383.82	565.74	434.53	919.07	818.35
20	20 EVs	660.89	701.41	812.14	708.67	1473.03	1410.08
500	1 EV	60.08	61.22	39.50	39.50	100.72	100.72
500	10 EVs	434.48	449.27	509.77	469.80	944.25	919.07
500	20 EVs	771.70	729.27	776.87	995.56	1548.57	1724.83

not fully capture each agent's optimal behavior, and the rising costs of system coordination and information barriers further limit the economic benefits of FedRL in large-scale, long-term scenarios.

4 Conclusion

In summary, the proposed federated reinforcement learning framework enables efficient, privacy-preserving, and collaborative optimization of multi-EV charging and discharging strategies under time-of-use electricity pricing. The results demonstrate that FedRL enhances overall economic benefits and maintains user privacy, particularly in small to medium-scale scenarios. However, as fleet size and training duration increase, the advantages of FedRL diminish due to information barriers and coordination costs. These findings underscore both the potential and limitations of federated approaches for large-scale, long-term multi-agent energy management applications. Therefore, future work will further explore methods to enhance the performance of federated learning in this context.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (62206062).

References

[1] Y. G. Kim, H. Lim, and J. Lee, “Decarbonizing road transport in Korea: The role of electric vehicle transition policies,” Transportation Research Part D: Transport and Environment, vol. 128, p. 104084, 2024.

[2] S. Zhang, W. Cai, X. Zheng, X. Lv, K. An, Y. Cao, H. S. Cheng, J. Dai, X. Dong, S. Fan et al., “Global readiness for carbon neutrality: From targets to action,” Environmental Science and Ecotechnology, vol. 25, p.100546, 2025.

[3] F. Ahmad, B. K. Panigrahi, M. Longo, L. Al-Fagih, M. S. Alam, and H. A. Gaber, “Electric vehicle management in multi-energy systems,” p. 101608, 2024.

[4] Y. Wang, F. Yu, F. Lin, and K. Feng, “Air quality impacts of ev promotion: Evidence from local charging pile adoption in china,” Journal of Asian Economics, vol. 98, p. 101922, 2025.

[5] E. Çelik and M. A. Taş, “Large-Scale electric vehicle integration: electric vehicle power system challenges and solutions,” Electric Power Systems Research, vol. 241, p. 109147, 2024.

[6] M. Yilmaz and P. T. Krein, “Sustainable electric vehicle charging: strategies for resilient energy infrastructure,” IEEE Transactions on Transportation Electrification, vol. 10, no. 1, pp. 1231–1244, 2024.

[7] P. Alaee, J. Shao, J. Bemˇs, and J. M. Guerrero, “Coordinated routing optimization and charging scheduling in a multiple-charging station system: A strategic bilevel multi-objective programming,” Sustainable Energy, Grids and Networks, vol. 42, p. 101659, 2025.

[8] H. I. Shaheen, G. I. Rashed, B. Yang, and J. Yang, “Optimal electric vehicle charging and discharging scheduling using metaheuristic algorithms: V2G approach for cost reduction and grid support,” Journal of Energy Storage, vol. 90, p. 111816, 2024.

[9] P. Liu, Z. Liu, T. Fu, S. Garg, G. Kaddoum, and M. M. Hassan, “Optimization of multi-vehicle charging and discharging efficiency under time constraints based on reinforcement learning,” Alexandria Engineering Journal, vol. 105, pp. 724–735, 2024.

[10] X. Hu, R. Li, L. Wang, Y. Ning, and K. Ota, “A data sharing scheme based on federated learning in IoV,” IEEE Transactions on Vehicular Technology, vol. 72, no. 9, pp. 11 644–11 656, 2023.

[11] D. Chen, H. Li, C. S. Lai, and L. L. Lai, “Reinforcement learning-based profit maximization for battery energy storage systems with electric vehicles and photovoltaic systems,” in IET Conference Proceedings CP858, vol. 2023, no. 28. IET, pp. 35–40, 2023.

[12] Z. Li, M. Yan, C. S. Lai, and L. L. Lai, “Optimization of electric vehicle smart charging and discharging decisions based on policy gradient reinforcement learning,” in 2024 8th International Conference on Power and Energy Engineering (ICPEE). IEEE, pp. 426–430, 2024.

[13] A. Khalatbarisoltani, J. Han, M. Saeed, C.-z. Liu, and X. Hu, “Privacy-preserving integrated thermal and energy management of multi connected hybrid electric vehicles with federated reinforcement learning,” Applied Energy, vol. 385, p. 125386, 2025.