

Coordinated Optimization of Orderly Charging and Grid Interaction at Electric Vehicle Charging Stations Based on Multi-Agent Reinforcement Learning

Yixin Wang

School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, Henan, China

Corresponding author: Yixin Wang (e-mail: 16692696699@163.com)

ABSTRACT Uncoordinated charging of large-scale electric vehicles exacerbates peak-valley differences and voltage exceedance risks in the power grid, while existing scheduling methods still have limitations in distributed decision-making, dynamic pricing, and multiobjective balancing. These problems become more significant in charging station clusters where power-electronic converters, communication links, and complex electromagnetic operating environments jointly affect grid interaction stability. In this paper, a collaborative optimization framework based on multi-agent reinforcement learning is proposed for orderly charging at electric vehicle charging stations and coordinated interaction with the power grid. First, each charging station is modeled as an autonomous agent, and distributed environment modeling is realized based on local observation information and Markov decision processes. Second, a proximal policy optimization algorithm is used to generate a dynamic service fee multiplier in a continuous action space, which is combined with a demand elasticity module to form an adaptive pricing mechanism. Finally, a composite reward system integrating grid stability, operational revenue, and user satisfaction is developed, and multi-agent convergence training is achieved through parameter sharing and generalized advantage estimation. The results confirm the overall benefits of joint optimization in load shaping, economic performance, and robustness, providing a technical reference for intelligent charging coordination under grid interaction and electromagnetic compatibility constraints.

INDEX TERMS multi-agent reinforcement learning, electric vehicle charging scheduling, orderly charging, grid interaction, electromagnetic compatibility

I. INTRODUCTION

AS the global energy transition gains momentum, electric vehicles (EVs) have become the heart of transportation electrification, due to their low carbon emissions and large energy efficiency. Nevertheless, the increased charging load due to the considerable numbers of EVs entering the power system is a grave threat to the safe and cost-effective functioning of the power system. The normal daily load curves indicate that the peak charging period following get off work coincides considerably with the peak power consumption of household gadgets in the evening, causing numerous over-limit operations of distribution network transformers, a reduced node voltage qualification rate, and a dramatic rise in line losses. How to scientifically schedule the charging behavior of massive numbers of EVs and achieve multi-party collaborative optimization of “vehicle-station-grid” has become a key scientific problem that urgently needs to be solved in the field of smart grids. The electric vehicle charging scheduling

problem is simply a complex decision making problem that has several constraints, several stakeholders and several coupled goals. In space, geographically scattered charging stations have independent charging pile resources, cater to various user populations, and their feeder topologies are independent. Temporarily, vehicles arrival times, initial state of charge, as well as expected departure times are highly randomized. Profit-wise, charging station operators are after maximizing income, users are after low prices and quality of services, whereas the power grid is concerned with load relief and safety margins. Conventional centralized optimization algorithms (including mixed integer programming and alternating direction multiplier techniques) can solve optimal in small-scale applications, whereas with high concentrations of charging stations and as vehicle arrivals become more frequent, centralized architectures are limited by factors like exploding communication overhead, high single-point failure vulnerability, and limited real-time responsiveness. Static pricing mechanisms, such as time-of-

use (TOU) tariffs, fail to respond to intra-hour grid fluctuations. This lack of responsiveness significantly hinders the mobilization of demand-side flexibility among electric vehicle users. The latency inherent in time-of-use pricing and other static mechanisms prevents real-time adaptation to grid state fluctuations, thereby limiting the activation of user-side flexibility.

The current paper suggests a collaborative optimization framework of orderly charging and grid interaction in electric vehicle charging stations using multi-agent reinforcement learning. The main contributions include: abstracting each charging station as an independent agent, achieving decentralized decision-making based on local observation information and Markov decision processes, significantly reducing communication bandwidth requirements and centralized computing bottlenecks; employing a near-end policy optimization algorithm to generate dynamic service fee multipliers in a continuous action space, combined with a demand elasticity assessment module based on historical data, enabling charging stations to autonomously evolve optimal pricing strategies based on real-time grid conditions and user response characteristics; designing a three-dimensional reward function that integrates grid stability (load peak-valley difference, voltage over-limit penalty), operating revenue (electricity sales profit, equipment utilization rate), and user satisfaction (on-time completion rate, power smoothness), and ensuring grid safety is given the highest priority through weight configuration; and adopting a “centralized training, distributed execution” framework, combining parameter sharing and generalized advantage estimation methods to effectively alleviate the non-stationarity problem in multi-agent environments, improving training efficiency and policy generalization ability.

II. RELATED WORK

With the large-scale access of electric vehicles to the power grid, the charging scheduling problem has evolved from the energy allocation of a single charging station to a complex decision problem involving multiple constraints, multiple agents, and multiple objectives. Saner et al. [1] proposed an electric vehicle charging scheduling framework based on a cooperative hierarchical multi-agent system. Shao et al. [2] have advanced cooperative structures, the field also relies heavily on non-reinforcement-learning optimization methods. Specifically, Model Predictive Control (MPC) and Genetic Algorithms (GA) have been widely adopted as robust baselines for solving multi-objective scheduling problems due to their ability to handle explicit constraints. However, compared to these traditional meta-heuristic and control-theoretic approaches, MARL offers superior adaptability in decentralized environments where communication overhead is a bottleneck. Qian et al. [3] established a maximum charging power function dependent on SOC and embedded this function into a typical charging scheduling optimization framework. Traditional electric vehicle charging pricing mechanisms use a single rate or simple time-of-use pricing,

which makes it difficult to simultaneously consider the revenue of charging station operators and the charging costs of users. Adetunji et al. [4] gave different discounts/premiums to users whose charging start time is earlier than a certain threshold and later than another threshold, guiding users to actively adjust their charging time. Qureshi et al. [5] modeled the problem of how to simultaneously plan the path of a mobile charging station and the charging scheduling problem under uncertain travel time as a coupling of the vehicle path problem with random travel time and the charging scheduling problem.

Qureshi et al. [6] modeled the charging scheduling problem as a stochastic convex programming. Arikumar et al. [7] designed an SDN-assisted charging scheduling and management architecture, collected all charging piles, electric vehicles, grid loads and road network traffic information, and formed a unified global data view. Kiani et al. [8] proposed a hierarchical single-loop framework based on the alternating direction multiplier method (ADMM). Sone et al. [9] constructed a worst-case expected cost minimization model, describing the uncertainty set of the distribution through the Wasserstein sphere. Affolabi et al. [10] proposed a hierarchical interactive energy scheduling framework. Existing pricing mechanisms mostly adopt preset rules or simple feedback, failing to incorporate the economic game process between charging stations and users into the dynamic decision-making closed loop, resulting in limited adaptability of the strategy in the real market environment. This paper proposes a collaborative optimization framework based on multi-agent reinforcement learning, which aims to achieve a better collaborative balance between grid security, operational economy and user satisfaction through endogenous dynamic pricing game, heterogeneous multi-objective composite reward function and stable convergence architecture combining centralized training and distributed execution.

III. METHODS

A. DISTRIBUTED ENVIRONMENT MODELING BASED ON MULTI-AGENT COLLABORATION

In the development of the multi-agent reinforcement learning model, the system represents each geographically distributed charging station as an autonomous agent capable of independent decision-making. These agents are aware of the immediate condition of power grid operation by constantly communicating with the external environment, such as the remaining capacity of transformers, the signal of time-of-use price of electricity in the market at that time, the trend of fluctuation of the basic electricity load in the area, and the real-time occupancy rate and health of equipment of each charging pile. All agents need to do is process the local information and do not need to transfer all the data up to the central server, which is very important in terms of bandwidth communication needs or the risk of having only one point of failure. Simultaneously, parallel decision-making enhances the ability of the system to respond in real-time as well

[11,12]. The model uses a Markov decision process to model state evolution to take care of the randomness of the way electric vehicles enter charging stations. The assumption in this process is that the state at the following moment of time is only dependent on the present state and present action which is quite similar to real physical process of charging scheduling. In order to properly describe the uncertainty of vehicle arrival, a Poisson arrival model taking place on the background of historical operational data is established in a simulation environment, and the intensity of arrival at various times is adjusted according to the real data of the charging stations. Not only the physical constraints of a power grid are part of the state space of this agent, but also the actual state of charge of every vehicle waiting to be charged, the predetermined expected departure time and the projected target charge of each vehicle are all monitored in real time. Individual charging station agents have a dynamic charging task table, where the remaining charging time, power and departure deadline of every vehicle undergoing charging or queuing is recorded. It is all this information that contributes to the decision-making of the agent.

Regarding action space design, the intelligent agent has a fine-grained control of charging power by changing the output current coefficient of the charging piles. Each charging pile can be regulated to constant output power of zero to 100 percent of its rated power, as opposed to simply on/off control. This constant adjustment ability enables the smart agent to gradually add or reduce the load thus preventing the voltage flicker that occurs when power is suddenly changed. In a case, when it is observed that the rate of loading the transformer is approaching the safety limit, the intelligent agent adjusts the current coefficient of all charging vehicles proportionally, instead of simply disconnecting a number of charging piles to keep the grid safe, thus avoiding the consequences of user experience. This distributed modeling method is efficient in reducing the computation bottleneck of centralized computing in situations where huge amounts of charging requests may occur, enhancing the real-time response rate of the system.

B. GAME PRICING AND STRATEGY EVOLUTION OF DEEP REINFORCEMENT LEARNING ALGORITHMS

The financial relationship between the charging stations and the users is the fundamental connection that directs the charging pattern. This approach uses a deep reinforcement learning algorithm that uses continuous action space and in particular, generates dynamic pricing strategy. The pricing rationale ceases to be based on the predefined fixed rules but rather dynamically adjusted based on the balance signal of the supply and demand of the power grid in real time. The proposed model utilizes a Proximal Policy Optimization (PPO) algorithm. The Actor and Critic networks both feature a 3-layer MLP architecture (64-128-64 neurons) using ReLU activation and a Tanh output for continuous action scaling. Training hyperparameters include a clip range of 0.2, an entropy coefficient of 0.01, and a learning rate of

3×10^{-4} with Adam optimization. Convergence is defined by the stability of the mean episode reward over 100 consecutive epochs [13].

Under light load conditions with renewable energy production, the intelligent agent will be more appealing to vehicles to be connected to charging as immediate service fee is minimized, creating optimal shaving and valley filling. It is at this point that the intelligent agent is informed by the grid side with a consumption signal, which synthesizes renewable energy generation capacity, system frequency deviation and load margin. On the other hand, the intelligent agent can cause the non-urgent charging demand to switch to future times when the grid load is at a safe level using price levers. The smart agent does not only raise the price charged in the present period but also at the same time releases previews of the future low-price periods, which will motivate the users to book delayed charging via the application. This is an expectation-based dynamic pricing approach that is more adaptable than the old-fashioned peak-valley pricing and can adapt to variations in grid circumstances within minutes.

The model incorporates a demand elasticity module utilizing a log-linear functional form: $\ln(Q) = \alpha + \beta \ln(P)$, where β represents the elasticity coefficient estimated at -1.45 via historical regression ($R^2 = 0.92$). The system employs a Multinomial Logit Model to determine the probability distribution of users choosing to “charge now,” “postpone,” or “relocate” based on the price signal. This module captures the probability distribution of users opting to charge now, postpone charging or go to other charging points at various price levels. The system is capable of fitting a group demand elasticity curve to users in the service area of individual charging stations through a large amount of interaction history. The agent also uses user queue length and average waiting time currently being experienced when coming up with pricing strategies to prevent user churn-off as a result of overpricing. The agent can develop over time a dynamically optimal pricing model that guarantees basic profits at charging stations, but also the best coordination of the demand with grid regulation, by repeatedly trying various pricing ranges and monitoring user acceptance and patterns of charging loads shifting.

C. CONVERGENCE MECHANISM DESIGN FOR HETEROGENEOUS MULTI-OBJECTIVE REWARD FUNCTIONS

Multi-agent collaborative optimization is often characterized by conflicting interests of the parties involved so that a scientific reward function is needed to balance the interests. The paper forms a composite rewarding function that encompasses three aspects: grid stability, operating revenue, and user satisfaction. The grid-side reward aims at reducing the standard deviation of the load curve, reducing variations by rewarding actions that increase the so-called secondary peaks. The system determines the sliding variance of the overall feeder load each fifteen minutes; when a variance

is greater than it was in the last period, negative rewards are given to all the agents in question. Once the voltage at any node is over the limit, all agents that contributed to the overage are heavily penalized immediately, with a weight much greater than that of traditional rewards of revenue, hence giving grid safety the highest priority. Grid safety is prioritized through strict weight configuration: the safety penalty weight (ω_s) is set to 10.0, the operating revenue weight (ω_r) is 2.5, and the user satisfaction weight (ω_u) is 3.5. These fixed coefficients ensure that the grid stability penalty is four times more significant than revenue gains. The real-time electricity price difference and the charging volume is directly related to operational rewards whereby the agent is motivated to enhance the turnover efficiency. After each step of the decision, the agent computes the multiplication of actual electricity sold and time-of-use electricity price, less the cost of buying electricity, and receives the gross profit reward. At the same time, to encourage better use of charging piles, two-cycle idle charging piles will give the agent a small negative reward, and the agent will adjust to attract vehicles by reducing prices or making announcements. The weight of profit rewards will, however, be less than one-third the weight of safety rewards, which will prioritize safety.

The on-time rate of charging task accomplishment is the main factor of reward system in the user-side. The agent will experience considerable negative feedback in case the battery level of the vehicle when leaving the station is not sufficient to meet the user-established threshold. The agent is rewarded positively to the vehicles that finish the task ahead of schedule and surpass the minimum requirement but marginal rewards decrease with the battery level to avoid overcharging. Another proxy measure of user satisfaction introduced by the system is that when the number of times a vehicle has been interrupted during its charging exceeds a predefined threshold, the user will earn a negative reward, which implicitly obliges the agent to ensure that the output power change is smooth.

This framework uses a mixture of parameter sharing and dominance function in order to make the algorithm converge in a multi-agent competitive environment. The parameters of the policy and value networks of all agents have the same feature extraction layers, so each agent can learn to rely on the experience of other agents and speed up the overall learning process. At the same time, the output layer of each agent has its own policy head and promotes the search of alternative pricing and power distribution schemes with the help of an anti-entropy regularization term. The dominance function is based on a generalized dominance estimation approach to trade-off bias and variance to a more stable action evaluation. In training, a central coordinator periodically consolidates gradient information across all agents, updates the shared network and redistributes it to all agents. While this “centralized training, distributed execution” (CTDE) framework enhances convergence speed through parameter sharing, it inherently assumes a degree

of homogeneity among agents. We acknowledge that in scenarios involving highly heterogeneous charging stations—characterized by vastly different transformer capacities, distinct user demographics, or localized grid constraints—this approach may require further partitioning of agent groups or the introduction of agent-specific layers to maintain scalability and representational accuracy. This is a centralized training, distributed execution architecture that uses global information to enhance the efficiency of training, and to provide low communication overhead in execution. Table 1 outlines the important parameter configuration of the system at the training and run time.

The parameter settings described above combine the physical operating principles of actual power distribution networks with the general standards of market-oriented operation. The scale of the experience playback pool ensures that the intelligent agent can extract patterns from long-term historical fluctuations. The transformer rated load rate serves as a hard constraint to force intervention actions. The user’s acceptance probability of delayed charging is based on research into typical urban vehicle usage scenarios, ensuring the feasibility of the strategy in practical applications.

D. SIMULATION SETUP AND ENVIRONMENT

The proposed framework is validated using a customized Python-based power system simulator integrated with OpenAI Gym. The EV arrival process is modeled as a non-homogeneous Poisson process, with daily arrivals averaging 150 vehicles per station. Battery capacities are sampled from a Gaussian distribution $N(60, 15)$ kWh, and charging power levels are constrained between 7 kW (AC) and 120 kW (DC). User behavior, including departure deadlines and target SOC, is synthesized from the 2023 National Household Travel Survey (NHTS) dataset. The ‘typical daily load data’ for the regional distribution network (used in Table 2) is derived from the open-source IEEE 33-node test system, scaled to reflect the peak demand of an urban residential district.

IV. RESULTS AND DISCUSSION

A. ANALYSIS OF THE SHAPING EFFECT OF COLLABORATIVE OPTIMIZATION ON POWER GRID LOAD CURVE

By simulating typical daily load data of a regional power distribution network, we compared the load evolution characteristics before and after the collaborative optimization strategy. Without orderly guidance, peak charging times for electric vehicles and peak residential electricity consumption highly overlap, causing distribution network transformers to operate on the verge of overload for extended periods in the evening. This not only shortens transformer lifespan but also significantly increases line resistance losses. After introducing multi-agent collaborative optimization, the system can accurately identify off-peak periods and, through price incentives and power limiting, distribute charging demand that was originally concentrated during peak hours to the

TABLE 1. Key Parameter Settings

Parameter classification	Parameter name	Physical meaning and setting logic	Benchmark reference value	Dynamic adjustment range
Environmental perception	Node voltage deviation coefficient	Measuring the safety margin on the distribution network side	0.05 pu	0.02 - 0.10 pu
Grid constraints	Transformer rated load rate	Determine the hard upper limit of charging load	0.85	0.70 - 0.95
Learning mechanism	Experience replay pool size	Memory length for storing historical interaction data	500,000	100,000 - 1M
Pricing Logic	Marginal profit rate of return	Minimum revenue ratio for maintaining the operation of a charging station	12.5%	8% - 20%
User behavior	Delayed charging acceptance probability	Probability distribution of user response to price guidance	0.65	0.40 - 0.90

night or early morning. This smoothing of the spatiotemporal dimensions greatly reduces the peak-to-valley difference in the load curve.

Table 2 below quantifies the power grid operating parameters under different scheduling strategies. The data fully demonstrates the outstanding performance of collaborative optimization in improving the quality of power grid operation.

Experimental data shows that the peak load in the multi-agent collaborative mode is only 1.68 MW, a 34.9% reduction compared to the disordered charging mode, while the peak-valley difference reduction rate reaches 48.9%. More importantly, the node voltage qualification rate increases from 84.6% to 99.4%, virtually eliminating all voltage over-limit risks. The reduction in network losses not only means improved energy utilization but also directly reduces the economic cost of grid operation. This improvement mainly stems from the agent's ability to perceive the grid topology, enabling it to proactively avoid high-loss operating points while meeting charging demands.

B. EVALUATION OF CHARGING STATION OPERATION EFFICIENCY AND RESOURCE ALLOCATION EFFICIENCY

Besides contributing to power grid, the multi-agent cooperative algorithm also exhibited high efficiency in enhancing the performance of charging stations. The experiment was conducted with long-term tracking survey of the rate of utilization of charging piles, capital recovery cycle, and the average user waiting time. The findings indicated that, during the busy schedule of charging tasks, with the help of a reinforcement learning algorithm to dynamically schedule the charging tasks, the idle time of charging piles was minimized, particularly during the daytime scattered charging periods. In fine-tuning the prices, the agent was able to attract some vehicles whose demand is flexible. In addition to this, the agent demonstrated a high level of logical thinking in the way he organized the arrangement of heterogeneous charging piles (a combination of fast and slow charging). The operational efficiency of the charging

station at each stage of operation is shown in Table 3.

The analysis of the data indicates that the average utilization rate of the equipment did rise as compared to the initial stage (32.4% to the stable operation period of 61.2%). The drastic rise in the revenue per charging pile each day illustrates the economic feasibility of the approach. The mean battery level at the point of leaving the station was 94.8, which means that the basic requirements of users were completely ensured in the course of the optimization of the charging path. The important increase in the speed of reaction to abnormal shutdowns shows the innate benefits of the multi-agent structure in fault tolerance and fast reconfiguration plans, which allow timely scheduling of neighboring charging piles to replace fault work.

C. SYSTEM RESPONSE AND GAME EQUILIBRIUM ANALYSIS UNDER COMPLEX MULTI-CONDITIONS

In real-world operating environments, charging stations face far more complex conditions than laboratory simulations. This section delves into the dynamic response characteristics of multi-agent systems under various complex conditions, including fluctuations in distributed renewable energy, extreme load shocks, and sudden changes in user behavior. By introducing a set of random operating conditions generated by Monte Carlo simulation, we tested the decision-making path of the agents in response to extreme power demand. Experiments revealed that when a sudden surge in grid load occurs, the agents do not adopt a simple "complete power outage" strategy. Instead, based on game theory equilibrium, they prioritize all charging vehicles within the station within milliseconds. This dynamic allocation mechanism, based on a "demand-value" trade-off, allows the system to maximize the charging continuity for high-value users while protecting grid security.

Table 4 shows the ability of the collaborative optimization strategy to maintain the critical state of the system under different external disturbance intensities.

Analysis of the data from the above scenarios reveals that even under extreme weather conditions with a disturbance intensity coefficient as high as 0.92, the system's response

TABLE 2. Power Grid Operation Quality Performance

Scheduling mode	Assess peak load (MW)	Valley load (MW)	Peak-to-valley difference reduction rate (%)	Node voltage qualification rate (%)	Daily average network loss reduction
Traditional disorder mode	2.58	0.75	-	84.6%	-
Time-of-use electricity pricing guidelines	2.15	1.10	25.4%	91.2%	12.4 kWh
Centralized Optimization (CO)	1.62	-	51.2%	99.6%	37.1 kWh
Single-Agent RL (SARL)	1.84	-	40.5%	96.8%	28.9 kWh
Multi-agent collaboration	1.68	1.25	48.9%	99.4%	35.8 kWh

TABLE 3. Operational efficiency of charging stations at different stages of operation

Key operational metrics	Initial running period	Strategy optimization period	Stable income period	Year-on-year growth/improvement
Equipment utilization rate (%)	32.4%	48.9%	61.2%	+88.8%
Average daily revenue per pile (RMB)	125.60	184.50	242.30	+92.9%
Average off-site electricity consumption (%)	74.2%	86.5%	94.8%	+27.8%
Abnormal shutdown response time (s)	12.5	5.4	1.2	-90.4%

TABLE 4. System dynamic response performance under different external disturbance scenarios

External disturbance scenario	Disturbance intensity coefficient	Response latency (ms)	System load recovery time (min)	Loss rate of return (%)	Strategy success rate (%)
Distributed photovoltaic power generation has decreased sharply.	0.45	12.0	4.2	2.15	99.8
Regional power grid overfrequency	0.60	8.5	3.1	1.80	100.0
User concentrating access impact	0.85	15.6	7.4	3.42	98.6
Extreme weather interference	0.92	22.4	12.8	5.60	96.5

latency remains within 22.4 milliseconds, far exceeding the reaction limits of human dispatchers. The system load recovery time is also kept within a reasonable range, and due to the precision of the strategy, the revenue loss rate is kept low. More representatively, in the scenario of a sudden reduction in distributed photovoltaic power, the intelligent agent, by initiating V2G (vehicle-to-grid) reverse discharge compensation, not only maintained the stability of the local grid frequency but also offset some of the losses from charging operations through electricity sales revenue. This collaborative game among multiple intelligent agents is not only reflected in power regulation but also in the cross-temporal arbitrage of energy asset value, thereby constructing a stable dissipative structure with self-healing capabilities in an uncertain environment.

V. CONCLUSION

The problems that this paper tackles are the widening of peak-valley disparities and voltage surpassing thresholds in the power grid due to unordered charging of large-scale electric vehicles. It suggests a joint optimization model of both the organization of charging at charging

stations and communication with the power grid, which is grounded on multi-agent reinforcement learning. The framework balance the requirements of power grid security, operational economics and user satisfaction through the joint design of distributed environment modeling, adaptive game-theoretic pricing mechanisms and heterogeneous multi-objective composite reward functions. Nevertheless, there are still some limitations in this research: the existing model presupposes the absolute reliability of information exchange between charging stations and does not take into account the effect of actual communication delays or information loss; moreover, the model of user demand elasticity is founded on the past static data which does not allow changing it instantly when there is a shift in the behavior of the group. The future studies will focus on two directions, including first, presenting the strong multi-agent training mechanism in the case of dynamic changes in communication topology; and second, integrating generative models of online prediction and adaptive calibration of the real-time user behavior to further enhance the generalization and robustness of the strategy in real-world deployment settings.

AUTHOR CONTRIBUTIONS

Yixin Wang designed, collected and analyzed the data, and drafted the manuscript. Yixin Wang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Yixin Wang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] C. B. Saner, A. Trivedi, and D. Srinivasan, "A cooperative hierarchical multi-agent system for EV charging scheduling in presence of multiple charging stations," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 2218-2233, 2022, doi: 10.1109/TSG.2022.3140927.
- [2] S. Shao, H. Sartipizadeh, and A. Gupta, "Scheduling EV charging having demand with different reliability constraints," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 11018-11029, 2023, doi: 10.1109/TITS.2023.3279070.
- [3] K. Qian, R. Fachrizal, J. Munkhammar, et al., "The impact of considering state-of-charge-dependent maximum charging powers on the optimal electric vehicle charging scheduling," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 3, pp. 4517-4530, 2023, doi: 10.1109/TTE.2023.3245332.
- [4] K. E. Adetunji, I. W. Hofsajer, A. M. Abu-Mahfouz, et al., "A two-tailed pricing scheme for optimal EV charging scheduling using multiobjective reinforcement learning," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3361-3370, 2023, doi: 10.1109/TII.2023.3305682.
- [5] U. Qureshi, A. Ghosh, and B. K. Panigrahi, "Scheduling and routing of mobile charging stations with stochastic travel times to service heterogeneous spatiotemporal electric vehicle charging requests with time windows," *IEEE Transactions on Industry Applications*, vol. 58, no. 5, pp. 6546-6556, 2022, doi: 10.1109/TIA.2022.3182323.
- [6] U. Qureshi, A. Ghosh, and B. K. Panigrahi, "Multiobjective pareto-optimal intelligent electric vehicle charging schedule in a commercial charging station: A Stochastic convex optimization approach," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 11, pp. 12620-12632, 2024, doi: 10.1109/TII.2024.3423373.
- [7] K. S. Arikumar, S. B. Prathiba, R. S. Moorthy, et al., "Software defined networking assisted electric vehicle charging: Towards smart charge scheduling and management," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 163-173, 2023, doi: 10.1109/TNSE.2023.3293053.
- [8] S. Kiani, K. Sheshyekani, and H. Dagdougui, "ADMM-based hierarchical single-loop framework for EV charging scheduling considering power flow constraints," *IEEE Transactions on Transportation Electrification*, vol. 10, no. 1, pp. 1089-1100, 2023, doi: 10.1109/TTE.2023.3269050.
- [9] S. P. Sone, J. J. Lehtomaki, Z. Khan, et al., "Robust EV scheduling in charging stations under uncertain demands and deadlines," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 12, pp. 21484-21499, 2024, doi: 10.1109/TITS.2024.3466514.
- [10] L. Affolabi, M. Shahidehpour, F. Rahimi, et al., "Hierarchical transactive energy scheduling of electric vehicle charging stations in constrained power distribution and transportation networks," *IEEE Transactions on Transportation Electrification*, vol. 9, no. 2, pp. 3398-3409, 2022, doi: 10.1109/TTE.2022.3219721.
- [11] A. S. Bouhouras, D. Kothona, P. A. Gkaidatzis, et al., "Distribution network energy loss reduction under EV charging schedule," *International journal of energy research*, vol. 46, no. 6, pp. 8256-8270, 2022, doi: 10.1002/er.7727.
- [12] M. Farhoumandi, S. Bahramirad, M. Shahidehpour, et al., "Blockchain for Peer-to-Peer Energy Trading in Electric Vehicle Charging Stations With Constrained Power Distribution and Urban Transportation Networks," *Energy Internet*, vol. 2, no. 1, 2025, doi: 10.1049/ein2.12026.
- [13] S. S. Shuvo and Y. Yilmaz, "Demand-side and utility-side management techniques for increasing EV charging load," *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 3889-3898, 2023, doi: 10.1109/TSG.2023.3235903.