Multi-Agent Reinforcement Learning for EV energy management and trading using the Lightning Network

Thomas K. Dasaklis
School of Social Sciences
Hellenic Open University
Patras, Greece
0000-0003-1240-4822

Panagiotis G. Giannopoulos
School of Social Sciences
Hellenic Open University
Patras, Greece
0000-0001-7530-2458

Vangelis Malamas
Department of Informatics
University of Piraeus
Piraeus, Greece
0000-0002-9154-8340

Georgios Tantis
Department of Public Administration
Panteion University of Social and Political Sciences
Athens, Greece
0009-0002-0394-0447

Constantinos Patsakis
Department of Informatics
University of Piraeus
Piraeus, Greece
0000-0002-4460-9331

***Abstract*—Micropayments, involving low-value transactions (e.g., fractions of a euro/dollar), are critical for unlocking granular digital services. In this paper we present a novel high-level architecture integrating artificial intelligence (AI) agents with the Bitcoin Lightning Network (LN) to enable efficient micropayments for electric vehicle (EV) charging and peer-to-peer energy trading. The proposed architecture leverages the ultra fast and low-cost nature of the LN to enforce trustless payments upon verified energy delivery. AI agents embedded in EVs and charging stations autonomously negotiate dynamic pricing and energy allocation using reinforcement learning (RL) approaches, optimizing grid load balancing and enhancing profitability compared to on-chain methods. Based on a comprehensive use case involving EV owners, operators and energy providers, we demonstrate the system's viability, supported by a prototype implementation on the LN Testnet. Results show a 98.2% success rate for micropayments during simulated charging sessions, with AI agents reducing latency by prioritizing high-liquidity payment channels.**

***Index Terms*—Micropayments, Lightning network, Artificial Intelligence, Agents, Electric Vehicles**

I. Introduction

Electric vehicles (EVs) are rapidly transforming global transportation and energy ecosystems. During the last years EVs have evolved into mobile energy storage units, thus they can participate in a wide array of granular digital services, from dynamic toll payments to peer-to-peer (P2P) energy tradings [1]. Notwithstanding significant progress towards facilitating multiple trading schemes, an urgent requirement of enabling high-frequency, low-value transactions (micropayments), still exists. Enabling such micropayments is driven by three converging trends: (i) Decentralized Energy Markets, where vehicle-to-grid (V2G) networks rely on automated compensation mechanisms for distributed energy contributions; (ii) Micropayment Scalability, as on-chain transactions for per-minute parking or per-kWh energy trades are prohibitively slow and expensive; and (iii) Autonomy, since increasingly intelligent EVs minimize the feasibility of manual oversight for routine transactions.

The Lightning Network (LN) is a Layer-2 protocol built on top of the Bitcoin blockchain, providing a promising solution for instant, extremely cheap micropayments. By handling most transactions off-chain, the LN avoids the congestion and high transaction costs associated with on-chain settlements (Layer-1) on the Bitcoin blockchain [2]. This capability is critical for distributed energy scenarios, where EVs frequently exchange energy with charging stations or other vehicles. Coupled with EVs' expanding battery capacities, the LN's ability to handle sub-cent payments has the potential to exemplify advanced services, such as real-time billing and micro-compensation for energy grid support, by providing a feasible and cost-effective solution.

Despite the potential of EVs to enhance grid stability while creating new revenue streams, real-time coordination of micropayments and energy trades presents significant challenges. As EV fleets expand, the complexity of dynamic negotiations over energy pricing, route optimization, and grid constraints surpasses the capabilities of manual approaches. Moreover, both on-chain micropayment solutions and traditional centralized systems often impose prohibitively high transaction fees, with centralized platforms charging near-dollar or even higher fees; on-chain solutions incurring even greater costs under congestion, rendering them economically unviable for high-frequency trading.

To address these challenges, we propose a three-layer decentralized framework that seamlessly integrates bidirectional charging infrastructure, autonomous Artificial Intelligence (AI) agents and LN-empowered micropayments. Bidirectional chargers, compliant with ISO 15118 and CHAdeMO standards, facilitate energy exchange between EVs, buildings and the grid. AI agents, embedded in both EVs and charging stations (CSs), are trained using reinforcement learning (RL) to develop adaptive strategies for autonomous price negotiation, energy allocation, and risk management (e.g., battery degradation, market volatility). This dual objective supports both grid stabilization and economic viability for the trading ecosystems. The LN, settled over Hashed Time-Locked Contracts (HTLCs), serves as the payment mechanism within the framework, enabling trustless, low-latency transactions and making sub-$0.01/kWh trades feasible. By leveraging the synergies between AI and LN, our work lays the groundwork for decentralized energy ecosystems that capitalize on the advanced capabilities of modern EVs. Beyond demonstrating the high-level architecture, our work also provides empirical evidence showcasing multi-agents ability to stabilize grid demand while converging toward a profitable operational scheme for both CSs and EVs.

II. Literature review

A substantial body of the current literature has focused on the integration of multi-agent RL frameworks to model the complex interactions among CSs, EV users and other energy market participants. For instance, in [3] the authors introduce a game-theoretic multi-agent system in which fast charging stations and EV users interact to determine personalized electricity prices based on real-time factors. Similar multi-agent paradigms are evident in auction-based systems where charging stations, mobile chargers, and EVs participate in distributed energy trading. [4] presents a hierarchical auction mechanism that uses second-price and double auction models, while [5] and [6] model competitive pricing as non-cooperative games and multi-leader–common-follower games, respectively. Other novelties modeling the interfaces between participants in such ecosystems, include exploring P2P energy trading using decentralized architectures where EVs with energy storage and generation capabilities (prosumers) engage in market-driven transactions [1], [7]. On the RL-centric component implementation, recent novelties include the integration of cluster-based grid management systems, by modeling prosumers' behavior through the use of double deep Q-networks (DDQN) to coordinate trading activities while reducing reliance on the grid [8].

Except for the decentralized architectures, other implementation venues are oriented towards developing centralized learning - decentralized execution paradigms, as those proposed in [9] which models a multi-agent demand response system where plug-in EV (PEV) stations participate in V2G transactions, dynamically adjusting power distribution based on grid conditions. Other works extend agent-like frameworks to model wireless EV charging systems, where vehicles charge while in motion or parked, requiring real-time adjustment of charging and discharging prices based on supply–demand fluctuations [10]. In residential and multi-home energy management systems, autonomous agents (e.g. smart homes equipped solar photovoltaic systems, EVs) interact with a central aggregator to optimize real-time pricing, energy allocation, and battery degradation [11], [12]. Furthermore, integrated systems that couple EV charging with building HVAC operations have been proposed to reduce grid dependency and enhance energy efficiency [13].

Arguably, the incorporation of RL techniques presents a promising avenue. The choice of RL algorithm varies across studies, reflecting the diversity of system requirements and complexity. In particular, several works have employed classical model-free approaches such as Q-learning [3], [14] and SARSA [14], while others leverage deep RL methods such as Deep Q-Networks (DQN) [6], [12] and Double Deep Q-Networks (DDQN) [15]. Policy-gradient methods, including Proximal Policy Optimization (PPO) [13], [16], [17] and Deep Deterministic Policy Gradient (DDPG) [11], [16], have also been adopted to address continuous action spaces inherent in dynamic pricing. In addition, advanced frameworks in the literature incorporate graph-based RL techniques to capture spatial dependencies and inter-agent influences in routing and on-route charging guidance [18]. Blockchain and federated learning have also emerged as promising avenues for ensuring data privacy and security while enabling distributed learning. For example, the authors in [19] propose a blockchain-enabled V2X trading system that uses Proof of State of Charge (PoSOC) to incentivize strategic energy transactions. In [20] the authors employ a federated RL approach to optimize bidirectional EV charging while preserving user privacy.

Dynamic pricing mechanisms often incorporate auction-based models and game-theoretic principles to reflect market competition. In an online continuous progressive second price (OCPSP) auction mechanism, EV owners act as bidding agents and the charging station allocates resources based on bid values, therefore promoting truthful bidding and Nash equilibrium [21]. Similarly, Stackelberg game formulations have been used to model the interaction between electricity markets and EV charging operators, where the leader–follower dynamics guide pricing decisions [22]. These approaches illustrate how economic theories can be integrated with AI-driven optimization to balance supply–demand and ensure equitable energy distribution.

A recurring challenge across the literature is the design of reward functions that encapsulate diverse objectives such as cost reduction, grid stability, user comfort, and infrastructure profitability. In many models, the reward functions are multi-objective, converting various operational metrics—including fixed infrastructure costs, SOC-based charging needs, grid load fluctuations, and even battery degradation—into a single utility function that drives the learning process [3], [20], [23]. This multi-objective optimization is critical for ensuring that the resulting pricing policies are both economically and operationally sustainable.

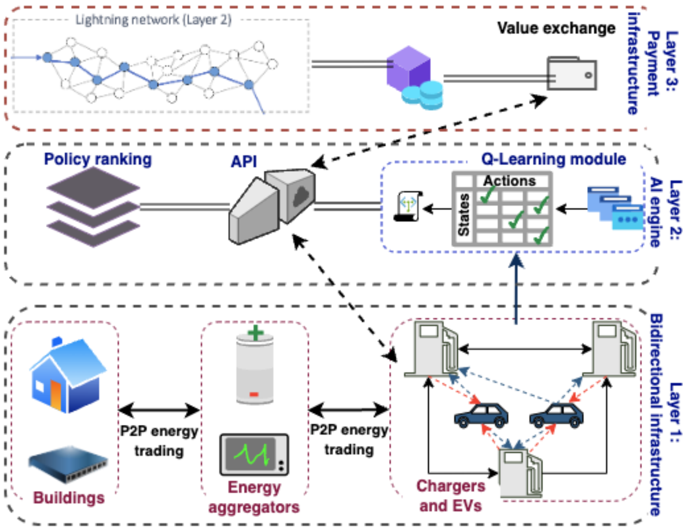

Fig. 1. A high level architecture leveraging integration of AI agents and LN

III. High-level architecture

To facilitate the rising need of activating the prosuming role of EVs within the grid management layers, we propose an integrated, LN-infused architecture, capable of addressing challenges related to the on-chain and centralized payment solutions. Leveraging the capabilities of LN, our architecture enables high-frequency, low-volume payments, allowing the incorporated agents to negotiate fairly and maximize their revenue in a trustless environment. As depicted in Figure 1, the overall architecture consists of three layers: (i) the bidirectional infrastructure, (ii) the AI engine and (iii) the LN, which acts on top of the decisions made by the underlying agents, serving the scope of value exchanges.

A. Bidirectional Infrastructure

The foundation of the proposed framework is a bidirectional charging infrastructure that supports two-way energy flow between EVs, buildings, and the power grid. Unlike conventional setups where EVs only consume electricity, bidirectional chargers compliant with standards such as ISO 15118 and CHAdeMO enable EVs to discharge stored energy back into the grid or share it with other vehicles. This capability transforms EVs into mobile energy reservoirs, capable of providing on-demand power during peak load or selling excess energy to neighbors. Beyond offering new revenue streams for EV owners, bidirectional charging alleviates grid stress by dynamically adjusting supply and demand. It also reduces reliance on centralized power plants, fostering a more resilient and sustainable energy ecosystem.

B. AI Decision Engine

An AI engine is incorporated in the proposed architecture to orchestrate energy trading and resource allocation across distributed EV networks. This engine is built upon the capabilities of multi-agent RL approaches, incorporating value-based methods to model and train two types of agents (i.e. EVs and CSs) to learn a synergistic policy that maximizes the profit of the CSs while minimizing grid overload conditions, to optimize the networks' resilience. AI engine acts in real-time operations and continuously updates bidding strategies, pricing and routing decisions. In this manner, EVs autonomously negotiate energy trades, respond to grid signals and manage their charging or discharging schedules. The result is an adaptive market mechanism that not only reduces transaction and coordination costs but also optimizes battery health, cost savings and the overall grid stability.

C. Lightning Network Layer

All financial transactions in the proposed framework rely upon the LN. Through payment channels, EVs and CSs can transmit micropayments, often amounting to fractions of a cent, almost instantaneously and at negligible cost. Smart contracts known as HTLCs ensure that funds are only released upon verifiable delivery of energy, effectively mitigating counter-party risk. This trustless and distributed settlement process is essential for high-frequency, small-value transactions, such as paying per kWh of energy consumed or compensating EV owners for grid stabilization services. By integrating LN micropayments into the AI-driven trading engine and bidirectional infrastructure, the proposed architecture delivers a holistic solution that is both economically and technically feasible and scalable.

IV. Implementation details

This section outlines the technical implementation of the architecture's two core components: AI agents for autonomous decision-making and LN for trustless micropayments. The implementation scheme is tailored to addressing the main scenarios analyzed in Table I.

To model the ecosystems we develop two types of agents: CSs and EVs, both operating under a Q-learning framework to adapt their decisions dynamically. EV agents make decisions based on their SOC, which is discretized into ten bins and can choose to either remain idle, charge by selecting a fraction of their available capacity or discharge through V2G. The charging decision depends on the time-of-day pricing and the availability of energy at different CSs, with EVs selecting the cheapest available station, which may apply discounts when its buffer is above a certain threshold. In contrast, V2G actions allow EVs to sell energy back to a selected CS at a fixed rate.

CS agents, on the other hand, update their pricing dynamically based on grid demand and the average price across all stations is used in computing the cost of charging. By dynamically buying from the grid when needed, selling to EVs at adaptive prices and accepting V2G energy, CS agents manage energy demands and continuously update their cost basis. The overall implementation process is built upon the centralized learning - decentralized execution (CL-DE) paradigm, to streamlining multiple classes of agents in a manner that advances the overall performance of the system. In this paradigm, learning is centralized, meaning that all agents

TABLE I
SUMMARY OF AI-DRIVEN ENERGY TRADING AND MANAGEMENT SCENARIOS IN EV ECOSYSTEMS

Scenario	Objective	AI decision cactors	Constraints
Residential P2P	Neighbors trade excess solar energy via EVs, maximizing revenue while ensuring affordable community pricing.	Surplus energy, local grid rates, time-of-day demand.	Avoid deep discharge (SOC $\geq$ 20%), prioritize community pricing ($\leq$ grid rate).
Workplace V2G	EVs sell energy to the grid during peak hours while reserving sufficient charge for commuting.	Grid demand signals, battery reserves, degradation costs.	Reserve energy for commute (SOC $\geq$ 30%), avoid excessive battery wear.
Fleet Optimization	Delivery vans optimize energy arbitrage across locations while ensuring route feasibility.	Route energy needs, price differentials, battery cycles.	Meet next-day route energy needs (SOC $\geq$ 40%).
Public Charging Auctions	EVs participate in real-time bidding for charging slots at high-demand hubs, aiming to minimize costs.	Urgency, competing bids, charger availability.	Bid $\leq$ 150% of grid rate, avoid overpaying.
Emergency Backup	EVs provide power to critical infrastructure (e.g., hospitals) during outages, overriding all other priorities.	Energy triage, regulatory frameworks, battery health.	Allow deep discharge (SOC $\geq$ 10%).

are trained within a shared environment incorporating system-wide metrics, while execution remains decentralized, as each EV and CS agent makes autonomous decisions based on its individual state and learned behavior or policy. Over time, this reinforcement mechanism enables optimized decision-making, ensuring scalability, adaptability and economic efficiency within the energy trading system.

A. QL-based agents

The implementation of the proposed CL-DE approach is based on the established Q-learning technique. This approach applies Q-learning with explicit Q-tables, where each agent updates its Q-values using the Bellman equation 1 to iteratively refine its decision-making process.

$$Q(s,a) \leftarrow Q(s,a) + \alpha(r + \gamma \max Q(s',a') - Q(s,a)) \quad (1)$$

In this context, distinct reward mechanisms are designed for the two agent classes to effectively incentivize decision-making based on their respective states. This differentiation ensures that EV and CS agents align with their specific objectives while maintaining overall system coordination. EV agents receive rewards ($r_{\text{EV}}^{(t)}$) based on energy purchase costs ($C_{\text{EV}}^{(t)}$) and battery degradation from discharging ($D_{\text{EV}}^{(t)}$), as these represent their two primary decision-making actions. Remaining idle yields a reward of zero. Conversely, CS agents receive rewards ($r_{\text{CS}}^{(t)}$) based on their profit from energy sales ($R_{\text{sales},i}^{(t)}$), while accounting for storage costs ($C_{\text{buffer},i}^{(t)}$) and operational costs ($C_{\text{oper},i}^{(t)}$). By structuring these distinct reward functions (see Eq. 2) within a decentralized execution paradigm, the framework enables each agent to independently optimize its behavior while interacting within a shared environment. This confirms that agents learn to act optimally across different states rather than following a predefined policy, reinforcing the value-based learning approach, where decisions are guided by learned Q-values instead of a centralized policy enforcement.

$$r^{(t)} = \begin{cases} r_{\text{EV}}^{(t)} = \begin{cases} -\left(C_{\text{EV}}^{(t)} + D_{\text{EV}}^{(t)}\right), & \text{if an action taken} \\ 0, & \text{if idle} \end{cases} \\ r_{\text{CS}}^{(t)} = \sum_{i=1}^{N_S} \left(R_{\text{sales},i}^{(t)} - C_{\text{buffer},i}^{(t)} - C_{\text{oper},i}^{(t)}\right) \end{cases}, \quad (2)$$

B. Micropayments

To validate the proposed micropayment framework we leverage the LN Testnet, a publicly available sandbox blockchain that uses valueless Testnet Bitcoin (TBTC), thus enabling developers to experiment with channel openings, routing and payment flows without risking real funds.

Testnet Setup: The first step is configuring a LN node—such as LND or Core Lightning—to connect specifically to the Testnet blockchain. Appropriate node settings (e.g., –bitcoin.testnet) designate which blockchain the node will interact with, ensuring that any funds committed or transactions performed remain purely for test purposes. After the node is running, it must be funded with TBTC to open payment channels. Testnet coins can be obtained from a community-managed faucet—a service that provides small amounts of TBTC at no cost. After creating a funded node, then the node is connected to other Testnet nodes, effectively joining the wider Testnet LN. By establishing these connections, the node gains visibility into potential payment routes and can route transactions to or from other participants (e.g., other EVs, CSs).

Channel Management on Testnet: Once the node is set up and funded, channel management is the next step. A Lightning channel is created by committing TBTC on-chain, effectively locking in a portion of funds for off-chain transactions. When opening a channel to a peer (EVs, CSs, etc.) the local amount of TBTC determines the capacity available for sending payments through that channel. To maintain operational efficiency, channels may need rebalancing over time. For instance, if an EV node frequently receives payments but rarely sends them, it may accumulate the majority of liquidity on its side, limiting its capacity to route outgoing transactions.

Simulating EV Energy Transactions: With channels open and sufficient liquidity, the next step involves simulating transactions that replicate real-world EV energy trades. For instance, a bidirectional CS acting as a power seller may issue an invoice for a specified amount of TBTC, corresponding to a defined energy quantity (e.g., "5 kWh @ 300 sat/kWh"). The buyer, whether another CS or an EV, settles this invoice off-chain through existing Lightning Network channels. If a direct channel between the buyer and seller is unavailable, the network automatically routes the payment via intermediate nodes. These off-chain transactions are both fast and cost-efficient, making them well-suited for the micropayments required in EV trading scenarios. The trustless nature of these transactions is ensured through HTLCs, which ensure that the power seller receives payment only upon delivering the agreed-upon energy, while the buyer's funds remain secure until proof of service is verified.

Testnet-Specific Tools: To streamline development and testing, a variety of Testnet-specific tools are available, designed to replicate mainnet counterparts and refine workflows. These include: (i) Lightning Explorers, which provide graphical dashboards displaying channel status, route availability, and node information on the Testnet network; (ii) Faucets and Channel-Opening Services, which offer free TBTC or pre-funded channels to simplify the initial setup process; and (iii) Hosted Node Platforms, such as Voltage, which enable seamless deployment of Testnet LN nodes with managed liquidity and user-friendly dashboards.

V. Experimental evaluation

To provide some evidence regarding the ability of the LN-enabled AI agents to stabilize the grid distribution network while achieving profitable trade-offs, we first analyze a baseline scenario. In particular, we compare two infrastructure models for AI-driven energy trading. First, a high-fee model is set to reflect costs incurred in traditional centralized or congested cryptocurrency infrastructures, specifically by setting an average 2.3% rate of transaction's value. Second, an LN-based infrastructure applies a minimal 0.0029% fee per transaction [1]. .

The number of EV agents and CS agents is set to 1,000 EVs and 60 CSs, respectively, considering that specific proportions of EVs serve distinct objectives, as outlined in Table I. The EVs are categorized into P2P residential (20%), V2G (70%), and fleet (10%) groups, reflecting diverse usage patterns. Also, different baseline prices for selling electric power are set. These prices reflect a point based on which the agents negotiate prices in different time slots. In our case, we consider that variations in baseline prices exist for three different slots within a day, while further refining of pricing strategies is allowed through offering specific discounts when an agent has excess of energy. Ultimately, to simulate an urgency-response scenario, our framework also includes a 5% hourly blackout probability, requiring 50 kWh of critical load coverage. Regarding the implementation of QL we have used the following settings: $\alpha = 0.1, \gamma = 0.9$, to facilitate agents adapt their charging and discharging behavior to maximize accepted transactions, minimize costs, and stabilize the grid. For the LN-based payments, after developing the full network form we record a 98.2% success rate for completed transactions, consistent with previous studies [24], which is used as fixed rate in the implementation of the AI agents.

After conducting simulations for both payment infrastructures over a one-year period, using detailed 24-hour time steps, we obtained the results presented in Figure 2. Notably, within approximately two months of operation, both classes of agents reach an optimal operating state. This is evidenced by the stabilization of transaction volumes (Figure 2a) and the consistency in the amount of energy traded per episode (Figure 2b). The convergence observed within this time-frame highlights the rapid adaptability of the proposed framework. Another significant finding concerns the profitability of each type of agent. As expected, the CS agents stabilize their profitability at positive net values. Meanwhile, the EV agents, within the same number of episodes, establish an operational scheme that enhances their role as prosumers. This results in near-zero profit values, indicating that while they continuously incur costs for energy acquisition and degradation, the volume of transactions ultimately ensures profitability.

VI. Discussion and conclusions

In this paper we have introduced a framework for the implementation of intelligent multi-agent systems aimed at maximizing the impact of EVs as prosumers. The framework optimizes profitability and reduces fluctuations in the energy demand from the grid. Given the great popularity of EVs, with an estimated 300 million EVs projected globally by 2030, the system's scalability, facilitated by the LN's high transaction throughput (exceeding potentially 1 million transactions per second) and the fast-convergence ability of AI agents, positions our framework as a foundational concept for next-generation smart grids.

Findings at a proof-of-concept level confirm the effectiveness of value-driven multi-agent approaches in simulating a wide range of trading EV-CS ecosystems. A key observation is the rapid convergence of energy traded within the agent ecosystems after a relatively short number of episodes, which indicates that energy demand from local grids could be stabilized over time. Furthermore, the integration of the LN within these ecosystems results in a significantly higher volume of energy transactions compared to traditional systems. This implies that the overall energy demand from local grids would be substantially reduced, as more energy transactions are handled within the decentralized LN-enabled network, thereby alleviating pressure on centralized grid infrastructure.

Specific extensions could further enhance the framework's applicability. Future research should justify and discuss the

[1]Cointelegraph, "Bitcoin Lightning Network is 1,000x cheaper than Visa and Mastercard: Data," available at https://cointelegraph.com/news/bitcoin-lightning-network-is-1-000x-cheaper-than-visa-and-mastercard-data.

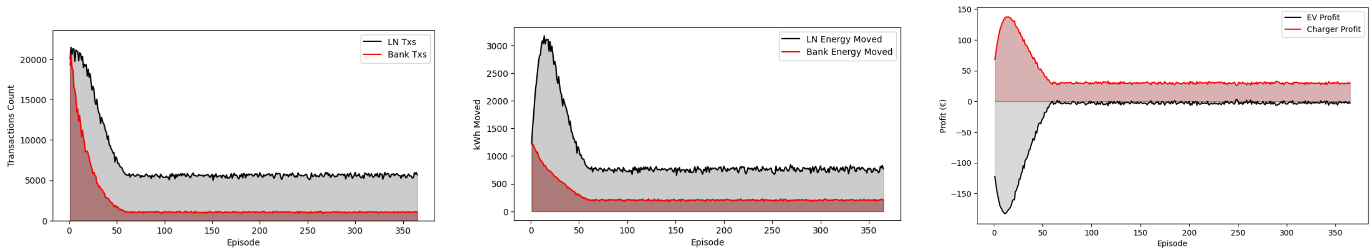

Fig. 2. Contribution of the LN-enabled agents in network's stability and system's profitability.

simulation assumptions, explore the carbon footprint of BTC mining on system sustainability, validate battery degradation models across EV types with real data and develop integration and transition strategies for existing charging infrastructures lacking bidirectional standard compatibility. Finally, future research should assess the AI component's convergence under varying numbers of both EV and CS agents.

ACKNOWLEDGMENT

This research has been funded by the Hellenic Open University under the project "Multi-agent systems and Generative Artificial Intelligence in Management science: Innovative Applications in Human Resources and Operations management-PELOPAS".

REFERENCES

[1] A. Alferidi, M. Alsolami, B. Lami, and S. B. Slama, "AI-Powered Microgrid Networks: Multi-Agent Deep Reinforcement Learning for Optimized Energy Trading in Interconnected Systems," *Arabian Journal for Science and Engineering*, 2024.

[2] T. K. Dasaklis and V. Malamas, "A review of the lightning network's evolution: Unraveling its present state and the emergence of disruptive digital business models," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 18, no. 3, pp. 1338–1364, 2023.

[3] S. Bae, B. Kulcsár, and S. Gros, "Personalized dynamic pricing policy for electric vehicles: Reinforcement learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 161, 2024.

[4] Y. Han, J. Meng, and Z. Luo, "Multi-Agent Deep Reinforcement Learning for Blockchain-Based Energy Trading in Decentralized Electric Vehicle Charger-Sharing Networks," *Electronics (Switzerland)*, vol. 13, no. 21, 2024.

[5] X. Yang, T. Cui, H. Wang, and Y. Ye, "Multiagent Deep Reinforcement Learning for Electric Vehicle Fast Charging Station Pricing Game in Electricity-Transportation Nexus," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 4, pp. 6345 – 6355, 2024.

[6] T. Qian, C. Shao, X. Li, X. Wang, Z. Chen, and M. Shahidehpour, "Multi-Agent Deep Reinforcement Learning Method for EV Charging Station Game," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 1682 – 1694, 2022.

[7] N. A. Alghanmi and H. Alkhudhayr, "EnergyShare AI: Transforming P2P energy trading through advanced deep learning," *Heliyon*, vol. 10, no. 17, 2024.

[8] M. Yavuz and O. C. Kivanc, "Optimization of a Cluster-Based Energy Management System Using Deep Reinforcement Learning Without Affecting Prosumer Comfort: V2X Technologies and Peer-to-Peer Energy Trading," *IEEE Access*, vol. 12, pp. 31 551 – 31 575, 2024.

[9] D. Choudhary, R. N. Mahanty, and N. Kumar, "Demand management of plug-in electric vehicle charging station considering bidirectional power flow using deep reinforcement learning," *Engineering Applications of Artificial Intelligence*, vol. 139, 2025.

[10] M. B. Hossain, S. R. Pokhrel, and H. L. Vu, "Efficient and Private Scheduling of Wireless Electric Vehicles Charging Using Reinforcement Learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4089 – 4102, 2023.

[11] N. Kaewdornhan, C. Srithapon, R. Liemthong, and R. Chatthaworn, "Real-Time Multi-Home Energy Management with EV Charging Scheduling Using Multi-Agent Deep Reinforcement Learning Optimization," *Energies*, vol. 16, no. 5, 2023.

[12] T. Xu, T. Chen, and C. Gao, "Intelligent Energy Management Strategy with Internal Pricing Sensitivity for Residential Customers Based on Reinforcement Learning and Internet-of-Things," in *EI2 2022 - 6th IEEE Conference on Energy Internet and Energy System Integration*, 2022, pp. 1760 – 1765.

[13] A. Narayanan, S. Nagarathinam, P. Misra, and A. Vasan, "Multi-agent Reinforcement Learning for Joint Control of EV-HVAC System with Vehicle-to-Building Supply," in *ACM International Conference Proceeding Series*, 2024, pp. 332 – 341.

[14] S. Wang, S. Bi, and Y.-J. A. Zhang, "A Reinforcement Learning Approach for EV Charging Station Dynamic Pricing and Scheduling Control," in *IEEE Power and Energy Society General Meeting*, vol. 2018-August, 2018.

[15] S. B. Slama, "OptiPower AI: A deep reinforcement learning framework for intelligent cluster energy management and V2X optimization in industrial applications," *Computers and Industrial Engineering*, vol. 200, 2025.

[16] A. Ghabi, Z. Alatoom, M. Guizani, and H. Hamam, "Peer-to-Peer Energy Optimization in V2X Using Reinforcement Learning," in *20th International Wireless Communications and Mobile Computing Conference, IWCMC 2024*, 2024, pp. 1697 – 1704.

[17] P. Jain, A. Tandon, T. Soni, Y. Soni, V. Nirwan, and V. Mudgal, "Dynamic Tariff Optimization for EV Charging Stations Using Reinforcement Learning," in *Proceedings of the 3rd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems, ICPEICES 2024*, 2024, pp. 205 – 210.

[18] Y. Li, S. Su, M. Zhang, Q. Liu, X. Nie, M. Xia, and D. D. Micu, "Multi-Agent Graph Reinforcement Learning Method for Electric Vehicle on-Route Charging Guidance in Coupled Transportation Electrification," *IEEE Transactions on Sustainable Energy*, vol. 15, no. 2, pp. 1180 – 1193, 2024.

[19] M. Moniruzzaman, A. Yassine, and R. Benlamri, "Blockchain and Federated Reinforcement Learning for Vehicle-to-Everything Energy Trading in Smart Grids," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 839 – 853, 2024.

[20] Z. Zhang, Y. Jiang, Y. Shi, Y. Shi, and W. Chen, "Federated Reinforcement Learning for Real-Time Electric Vehicle Charging and Discharging Control," in *2022 IEEE GLOBECOM Workshops, GC Wkshps 2022 - Proceedings*, 2022, pp. 1717 – 1722.

[21] Y. Zhang, Z. Zhang, Q. Yang, D. An, D. Li, and C. Li, "EV charging bidding by multi-DQN reinforcement learning in electricity auction market," *Neurocomputing*, vol. 397, pp. 404 – 414, 2020.

[22] Y. Xiao, X. Lin, Y. Lei, Y. Gu, J. Tang, F. Zhang, and B. Qian, "Blockchain-Assisted Secure Energy Trading in Electricity Markets: A Tiny Deep Reinforcement Learning-Based Stackelberg Game Approach," *Electronics (Switzerland)*, vol. 13, no. 18, 2024.

[23] V. Moghaddam, A. Yazdani, H. Wang, D. Parlevliet, and F. Shahnia, "An Online Reinforcement Learning Approach for Dynamic Pricing of Electric Vehicle Charging Stations," *IEEE Access*, vol. 8, 2020.

[24] R. Financial, "The lightning network: Empirical performance analysis," *River Financial Report*, 2023. [Online]. Available: https://river.com/learn/files/river-lightning-report-2023.pdf