



Article

Optimizing Electric Vehicle Charging Recommendation in Smart Cities: A Multi-Agent Reinforcement Learning Approach

Panee Suanpang ^{1,*}  and Pitchaya Jamjuntr ²

¹ Department of Information Technology, Suan Dusit University, Bangkok 10300, Thailand

² King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand; pitchaya.jam@kmutt.ac.th

* Correspondence: pannee_sua@dusit.ac.th; Tel.: +66-2644-5973

Abstract: As global awareness for preserving natural energy sustainability rises, electric vehicles (EVs) are increasingly becoming a preferred choice for transportation because of their ability to emit zero emissions, conserve energy, and reduce pollution, especially in smart cities with sustainable development. Nonetheless, the lack of adequate EV charging infrastructure remains a significant problem that has resulted in varying charging demands at different locations and times, particularly in developing countries. As a consequence, this inadequacy has posed a challenge for EV drivers, particularly those in smart cities, as they face difficulty in locating suitable charging stations. Nevertheless, the recent development of deep reinforcement learning is a promising technology that has the potential to improve the charging experience in several ways over the long term. This paper proposes a novel approach for recommending EV charging stations using multi-agent reinforcement learning (MARL) algorithms by comparing several popular algorithms, including the deep deterministic policy gradient, deep Q-network, multi-agent DDPG (MADDPG), Real, and Random, in optimizing the placement and allocation of the EV charging stations. The results demonstrated that MADDPG outperformed other algorithms in terms of the Mean Charge Waiting Time, CFT, and Total Saving Fee, thus indicating its superiority in addressing the EV charging station problem in a multi-agent setting. The collaborative and communicative nature of the MADDPG algorithm played a key role in achieving these results. Hence, this approach could provide a better user experience, increase the adoption of EVs, and be extended to other transportation-related problems. Overall, this study highlighted the potential of MARL as a powerful approach for solving complex optimization problems in transportation and beyond. This would also contribute to the development of more efficient and sustainable transportation systems in smart cities for sustainable development.

Keywords: electric vehicle; charge station; multi-agent reinforcement learning; optimizing; recommendation systems; smart cities



Citation: Suanpang, P.; Jamjuntr, P. Optimizing Electric Vehicle Charging Recommendation in Smart Cities: A Multi-Agent Reinforcement Learning Approach. *World Electr. Veh. J.* **2024**, *15*, 67. <https://doi.org/10.3390/wevj15020067>

Academic Editor: Joeri Van Mierlo

Received: 23 December 2023

Revised: 7 February 2024

Accepted: 9 February 2024

Published: 14 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

During the era of digital transformation, technological advancements, especially in the field of artificial intelligence (AI), have caused disruptions in various industries worldwide. AI has been widely adopted and has had a significant impact in all areas due to the use of advanced information technology. Recently, electric vehicles (EVs) have also gained significant attention and are being encouraged by many countries as an eco-friendly mode of transportation [1,2]. EVs are environmentally friendly due to their lack of emissions. In addition, they offer cost savings compared to traditional gasoline engines and have smooth and easy-to-use controls [1,3,4]. The lack of emissions from EVs makes them a suitable choice for those who prioritize environmental sustainability. Furthermore, operating EVs is more cost-effective than using a conventional gasoline engine. EVs are also easy to operate and provide a smooth driving experience [1,3–5].

According to the International Energy Agency (IEA), the global electric vehicle (EV) market has been growing rapidly, and as of 2020, there were over 7.2 million EVs on the road.

Moreover, the IEA predicts that the EV market will continue to grow, and by 2030, there could be over 250 million EVs in use worldwide [6]. The market size of EVs is also expected to increase significantly. According to Statista, the global EV market size was valued at approximately USD 140 billion in 2020 and is projected to reach over USD 800 billion by 2026 [7]. A study by Deloitte has suggested that the EV market is likely to continue growing due to government incentives, stricter regulations on emissions, and advancements in technology [8,9]. With the increasing availability of EV models, charging infrastructure, and government support, the adoption of EVs will likely continue to accelerate [10]. As such, the utilization of EVs is rapidly growing worldwide due to their various advantages, such as cost efficiency, environmental conservation, smart city traffic improvement, and enhanced user satisfaction [1].

This introduction has brought about numerous advantages for the transportation industry. The increasing demand for EVs has also brought attention to the need for establishing a robust charging infrastructure to meet the changing needs of EVs [11–13]. The charging infrastructure is an essential component of the EV ecosystem, and its optimal placement and allocation would be crucial for the adoption and widespread use of EVs [11,12]. However, the inadequate charging infrastructure and uneven spatial and temporal demands for charging have been identified as significant problems that would need to be addressed [13,14]. To resolve these problems, several studies have proposed various solutions to introduce EV station recommendation systems providing the most appropriate charging stations for an EV driver to use based on their current location, destination, and other factors [1,2,11–13]. The goal of this system would be to improve the convenience and efficiency of EV charging by helping drivers find the nearest, most available, and fastest charging stations [1,12,13].

A wide range of factors would be needed to consider the introduction of EV station recommendation systems, such as the battery capacity of the EV, charging time needed, availability of the charging stations, and estimated time to reach the destination. These systems could also incorporate real-time information, such as traffic congestion, weather conditions, and charging station availability [14,15]. In addition to improving the convenience and efficiency of EV charging, the station recommendation systems could help to reduce range anxiety, which would be a common concern among EV drivers who fear they would run out of charge before reaching their destination. By providing accurate and timely information about the location and availability of the charging stations, EV station recommendation systems could help to increase the confidence of EV drivers and promote the adoption of EVs [1,16].

Overall, an EV station recommendation system is an important application of technology in the field of sustainable transportation, and it has the potential to improve the accessibility and adoption of EVs in the future. Furthermore, the increasing adoption of EVs has highlighted the need for an efficient charging infrastructure to support the growth of this industry. However, the current charging infrastructure is inadequate to meet the spatial and temporal demands of EV users, consequently leading to issues such as long waiting times, insufficient charging stations, and unbalanced distribution of charging stations [16–18]. To address this challenge, EV station recommendation systems have emerged as a solution to improve the convenience and accessibility of EV charging, reduce range anxiety, promote the adoption of EVs, and optimize the charging infrastructure [1,18].

There are several types of optimization techniques that could be applied in a recommendation system, such as reinforcement learning (RL), which would have the potential to optimize the energy management of EVs by making decisions in real time to maximize efficiency and reduce energy consumption [19]. Moreover, RL has demonstrated significant benefits in addressing sequential decision problems in dynamic environments, such as optimizing order dispatching for ride-hailing and shared bike rebalancing. In RL, an agent would learn a policy by interacting with the environment with the goal of achieving the optimal long-term reward. Given the success of RL in these contexts, it would seem like a natural choice to apply it to improve EV charging recommendations with long-term

objectives, such as reducing the overall charging waiting time (CWT), the average Charging price (CP), and the Charging Failure Rate (CFR) [20–22].

However, there are several technical challenges that would need to be addressed in order to accomplish this goal:

- (1) One of the main challenges in optimizing EV charging in a smart city is the vast number of EVs and charging stations [1]. Trying to directly train a single centralized system to manage all the EVs and charging stations would be difficult due to the massive state and action space, as well as the high-dimensional environment. This could lead to scalability and efficiency issues, as well as the risk of the entire system failing due to a single point of failure. As an alternative, previous work modeled a small set of vehicles as multiple agents, but this approach was not effective for charging recommendations since most charging requests were ad hoc and came from non-repetitive drivers [1,2]. To address this issue, the authors proposed treating each charging station as an individual agent and formulating EV charging recommendations as a multi-agent reinforcement learning (MARL) task. By doing this, the authors created a multi-agent actor-critic framework where each individual agent had a quantity-independent state and action space, which could be scaled to more complicated environments and was more robust to other agents' potential failure. This approach was more effective and efficient for optimizing EV charging in a large metropolis.
- (2) Another challenge in optimizing EV charging recommendations was coordinating and cooperating in a large-scale agent system. When there was a charging request, only one station could provide the service, so different agents would need to work together to achieve better recommendations [1,4,5]. Additionally, cooperation between agents would be essential for the long-term optimization of the charging recommendations. If a charging station were heavily occupied with many incoming charging requests, other stations with available charging spots could help to balance the charging demand through cross-agent cooperation [10,11]. To address this challenge, the authors used a bidding game analogy to describe the process of the agent taking actions and proposed a centralized attentive critic module that was tailored to encourage multiple agents to learn globally coordinated and cooperative policies. This approach would help to overcome the challenge of coordination and cooperation in a large-scale agent system, hence leading to better EV charging recommendations.
- (3) The third challenge in optimizing EV charging recommendations would be the potential competition of future charging requests. In the real world, this potential competition could occur at any charging station and could result in problems, such as extra CWT and charging failure. However, it would be difficult to predict the impact of future charging requests beforehand [17,18]. To overcome this challenge, the authors incorporated information about future charging competition into their centralized attentive critic module using a delayed access strategy. They also transformed their framework into centralized training with decentralized execution architecture to enable online recommendations. This approach allowed the agents to make use of future knowledge during the training phase and take immediate action without needing future information during execution, which helped to mitigate the potential competition of future charging requests.

The EV charging recommendation problem is challenging due to multiple technical difficulties. The first challenge is the large state and action space. To resolve this challenge, the authors proposed a MARL framework that regarded each charging station as an individual agent. The second challenge is the coordination and cooperation in a large-scale agent system. To address this challenge, the authors proposed a tailor-designed centralized attentive critic module that stimulated multiple agents to learn globally coordinated and cooperative policies. The third challenge is the potential competition of future charging requests.

This paper had the aim of developing EV station recommendation systems that could help improve the accessibility and convenience of EV charging in Rayong smart city, Thai-

land, a country with a growing EV market but still relatively limited charging infrastructure. Moreover, this paper explored and contributed to the importance of EV station recommendations and provided a step-by-step guide on how to apply this technology in Thailand. This paper proposed a five-step process for applying EV station recommendations in Thailand:

- (1) Collect data on the location, type, and availability of existing EV charging stations across the country.
- (2) Integrate the collected data into a centralized database that could be accessed by EV station recommendation systems.
- (3) Develop a recommendation system that would take into account various factors, such as the location of the driver, the type of EV, and the availability of the charging stations.
- (4) Deploy the recommendation system through various channels, including mobile apps, websites, and in-car navigation systems.
- (5) Promote the EV station recommendation system to increase its adoption among EV drivers.

By following this five-step process, Thailand could improve the accessibility and convenience of EV charging, reduce range anxiety, and promote the adoption of EVs.

The authors integrated the future charging competition information into the centralized attentive critic module through a delayed access strategy and transformed the framework to centralized training with decentralized execution architecture to enable online recommendations. Finally, it was challenging to jointly optimize multiple optimization objectives. To overcome this challenge, the authors extended the centralized attentive critic module to multi-critics for multiple objectives and developed a dynamic gradient reweighting strategy to adaptively guide the optimization direction.

1.1. Problem Statement

In this section, the researchers define the important terms and describe the EV charging recommendation problem. A set of N charging stations was considered and denoted as $C = \{c^1, c^2, \dots, c^n\}$. For each day, a charging request was defined as follows:

Definition 1. *Charging request:* A charging request is defined as the t -th request (i.e., step t) of a day.

$$q_t = \langle l_t, T_t, T_t^c \rangle \in Q \quad (1)$$

Here, it is the location of q_t . T_t is the time at which the request is made, and T_t^c is the time when the charging request is completed. A charging request is considered completed when it is successfully charged or when it finally gives up due to charging failure. The cardinality of the charging requests is denoted as $|Q|$. In the following, q_t is used to denote the corresponding EV of q_t interchangeably. The following terms are defined to help formalize the EV charging recommendation problem:

Definition 2. *CWT:* This is defined as the sum of the travel time from the location of the charging request q_t to the target charging station c^i , and the queuing time at c^i , until q_t finishes charging.

Definition 3. *CP:* This is defined as the cost per unit of electricity, which is typically a combination of the electricity cost and service fee.

Definition 4. *Charging Failure Ratio (CFR):* CFR is defined as the ratio of charging requests that accepted the authors' recommendation but failed to charge, divided by the total number of charging requests that accepted the recommendation. The authors aimed to recommend the most suitable charging station

$$rc_t = c^i \in Q \quad (2)$$

for each charging request $q_t \in \mathcal{Q}$ with the simultaneous objective of minimizing the overall Charging Wait Time (CWT), average Charging Price (CP), and Charging Failure Ratio (CFR) for the charging requests that accepted the recommendation. $C = \{c^1, c^2, \dots, c^n\}$ Here, each c^i represents an individual charging station within the set. The set C encompasses all the available charging stations under consideration for recommending charging stations to electric vehicles (EVs).

1.2. Research Objectives and Contribution

This paper introduces the spatio-temporal multi-agent reinforcement learning (STMARL) framework for optimizing electric vehicle (EV) charging recommendations in smart cities. The framework represents a pioneering effort in the application of MARL techniques to address multi-objective challenges in recommending charging stations. With a primary focus on MARL, the framework leverages the power of coordinated learning among multiple agents to enhance the performance of intelligent charging stations. Designed with centralized training and decentralized execution, the STMARL framework achieves global coordination and cooperation, laying the foundation for efficient and intelligent EV charging infrastructure in smart cities.

1.2.1. Research Objectives

1. To explore the feasibility and effectiveness of treating EV charging stations (EVCSs) as agents in a MARL framework for optimizing the placement and allocation of charging stations in a network of charging infrastructure.
2. To investigate the impact of different factors, such as the number of EVs and charging stations, their geographic distribution, and charging demand patterns on the performance of the MARL algorithm.
3. To compare the performance of different MARL algorithms, including multi-agent DDPG (MADDPG), deep deterministic policy gradient (DDPG), and others for the problem of EVCS placement and allocation and to identify the algorithm that would perform the best in terms of key metrics, such as charging efficiency and user satisfaction.

1.2.2. Research Contributions

The proposed framework treated EVCSs as agents, which could facilitate better collaboration and communication among charging stations and with EVs, thus leading to more efficient charging and improved user experience. The investigation of the impact of different factors on the performance of the algorithm could inform the design of the EV charging infrastructure and policy interventions that could maximize the benefits of electric mobility for users and society. The comparison of different MARL algorithms for the EVCS placement and allocation problem could help identify the best approach or address this complex optimization problem, which is critical for the widespread adoption of EVs and the development of sustainable development in smart cities.

2. Literature Review

2.1. Electric Vehicles (EVs)

The global adoption of electric vehicles (EVs) has seen a steady increase in recent years, driven by the imperative to address greenhouse gas (GHG) emissions and reduce reliance on fossil fuels. This overview of post-2018 literature on EVs encompasses their environmental impact, policy landscape, and consumer preferences. Studies, such as those by Spieser et al. [23] and Turrentine and Kurani [24], emphasize the potential of EVs to significantly reduce GHG emissions, particularly when charged with renewable energy sources. Government interventions, as highlighted by Zhang et al. [25] and Lin et al. [26], play a pivotal role in promoting EV adoption, with subsidies proving more influential than tax incentives and programs like California's Zero Emission Vehicle (ZEV) initiative contributing to a heightened EV market share. Recognizing consumer preferences is deemed crucial, with studies by Li et al. [27] and Bruneau et al. [28] underscoring factors

such as driving range, charging infrastructure, cost, and social context as key determinants in shaping EV adoption. These insights collectively contribute to understanding and fostering the ongoing transition toward EVs.

Studies, including those by Spieser et al. [29], Turrentine and Kurani [30] assess the environmental impact of EVs, concluding that they can significantly reduce emissions, especially with renewable energy sources. Government policies play a pivotal role, with studies by Zhang et al. [31] and Kirschbaum et al. [32] highlighting the effectiveness of subsidies and incentives in China and Germany. Consumer preferences, elucidated by Li et al. [33] and Zmud et al. [34], underscore factors such as driving range, charging infrastructure, and cost as crucial determinants [35–37]. In Thailand, the EV market is rapidly growing due to government incentives, with over 40,000 EVs on the roads as of May 2021 and a projected 2.5 million by 2030, as per Bloomberg NEF [38]. Despite this growth, challenges such as high costs, charging infrastructure limitations, and safety concerns persist, reflecting the complex landscape of EV adoption [39,40].

2.2. Multi-Agent Reinforcement Learning (MARL)

MARL is an emerging research area that has gained considerable attention in recent years. MARL involves the use of multiple agents to solve complex problems that would be difficult to solve with a single agent [41]. This literature review provides a comprehensive overview of MARL, including its background, applications, and recent advancements in the field.

RL is a subfield of machine learning that involves training agents to make decisions in an environment to maximize a cumulative reward signal. In the traditional RL framework, a single agent would interact with an environment to learn an optimal policy [42]. However, in many real-world scenarios, multiple agents would interact with each other and the environment, hence making the problem more complex. MARL is an extension of RL that involves multiple agents learning to coordinate with each other to solve a problem. In MARL, each agent would have its own objective, and the agents would need to learn to cooperate to achieve a common goal. MARL has been applied in various domains, including robotics, traffic control, and game theory [43].

Mnih et al. [44] applied MARL to the problem of traffic light control, where multiple agents learned to coordinate with each other to optimize the traffic flow. Another application of MARL has been in robotics, where multiple agents have learned to collaborate to perform complex tasks. A study by Sun et al. [45] used MARL to train a team of robots to cooperatively transport an object to a target location. Recent advancements in MARL have also focused on improving the efficiency and effectiveness of the learning process. Likewise, Wu et al. [46] proposed a new MARL algorithm called QMIX, which used a centralized value function to learn the joint action-value function of all agents. The QMIX algorithm demonstrated that it could outperform other MARL algorithms in several benchmark tasks. Another recent advancement in MARL is the use of deep reinforcement learning (DRL) techniques. Foerster et al. [47] proposed a new DRL algorithm called CommNet, which used a communication module to allow agents to exchange information during the learning process. The CommNet algorithm could achieve a state-of-the-art performance in several multi-agent benchmark tasks. As such, MARL is a promising research area that has the potential to solve complex problems that would be difficult to solve with a single agent. MARL has been successfully applied in various domains, including robotics, traffic control, and game theory. Recent advancements in the field have focused on improving the efficiency and effectiveness of the learning process, including the use of centralized value functions and communication modules. Future research should continue to explore the potential of MARL and its applications in real-world scenarios.

2.3. EV Recommendation Systems

As the popularity of EVs continues to grow, there will be a need for effective recommendation systems to help users make informed decisions when choosing an EV. EV

recommendation systems use machine learning techniques to analyze users' preferences and provide personalized recommendations for EV models that would best meet their needs. In this literature review, an overview of the EV recommendation systems is provided, including their applications, challenges, and recent advancements. Huang et al. [48] proposed a hybrid recommendation algorithm that combined content-based and collaborative filtering techniques to recommend EV models to users on an online marketplace. Another application of EV recommendation systems is in car dealerships, where they could be used to help salespeople provide personalized recommendations to customers. Additionally, Muratori et al. [49] proposed a model-based collaborative filtering approach that used customers' data to recommend EV models that best met their preferences. One of the main challenges in EV recommendation systems is the lack of user data, as the adoption of EVs is still relatively low compared to traditional gasoline-powered vehicles. This could lead to sparse data, consequently making it difficult to accurately analyze users' preferences and provide personalized recommendations. Moreover, Almahmood and Tekerek [50] proposed a transfer learning-based approach that utilized data from traditional vehicle recommendation systems to improve the accuracy of EV recommendations. Another challenge is the lack of standardization in EV models and features, which could make it difficult to compare and evaluate different models. Likewise, An et al. [51] proposed a feature-based approach that used a feature matrix to represent the features of different EV models and compare them based on the users' preferences. Recent advancements in EV recommendation systems have focused on improving the accuracy and efficiency of the recommendation process. Chen et al. [52] proposed a deep learning-based approach that used a convolutional neural network to analyze users' preferences and recommend EV models. The approach was shown to outperform traditional collaborative filtering and content-based algorithms. Another recent advancement has been the use of hybrid recommendation algorithms that combine multiple techniques to provide more accurate and diverse recommendations. In addition, Li et al. [53] proposed a hybrid recommendation algorithm that combined a collaborative filtering approach with a matrix factorization technique to recommend EV models to users. Hence, EV recommendation systems are a promising research area that could help users make informed decisions when choosing an EV. Furthermore, EV recommendation systems have been successfully applied in various domains, including online marketplaces, car dealerships, and EV rental services. However, challenges such as sparse data and lack of standardization in EV models and features remain. Recent advancements in the field have focused on improving the accuracy and efficiency of the recommendation process, including the use of deep learning-based approaches and hybrid recommendation algorithms. Therefore, future research should continue to explore the potential of EV recommendation systems and address the challenges in the field.

2.4. Electric Vehicle Charging Stations in Smart Cities

2.4.1. Rayong Smart City, Thailand

Rayong, an industrial hub in Thailand, has embarked on several smart city initiatives to enhance urban living. These initiatives encompass areas such as transportation, energy management, and healthcare. For instance, the implementation of IoT sensors in transportation systems has streamlined traffic flow [54], while renewable energy projects have reduced the city's carbon footprint [55].

The development of a smart city in Rayong is not without challenges. Issues such as data security, citizen privacy, and technological infrastructure must be carefully addressed [56]. Smart city initiatives in Rayong have had a significant socio-economic impact. Research by Phan and Nguyen [57] demonstrated a positive correlation between smart city development and job creation in the region. Additionally, smart city technologies have empowered citizens, enabling them to actively engage in urban governance. Figure 1 depicts the smart city model in Thailand and red box indicates the location of the EV charging station.

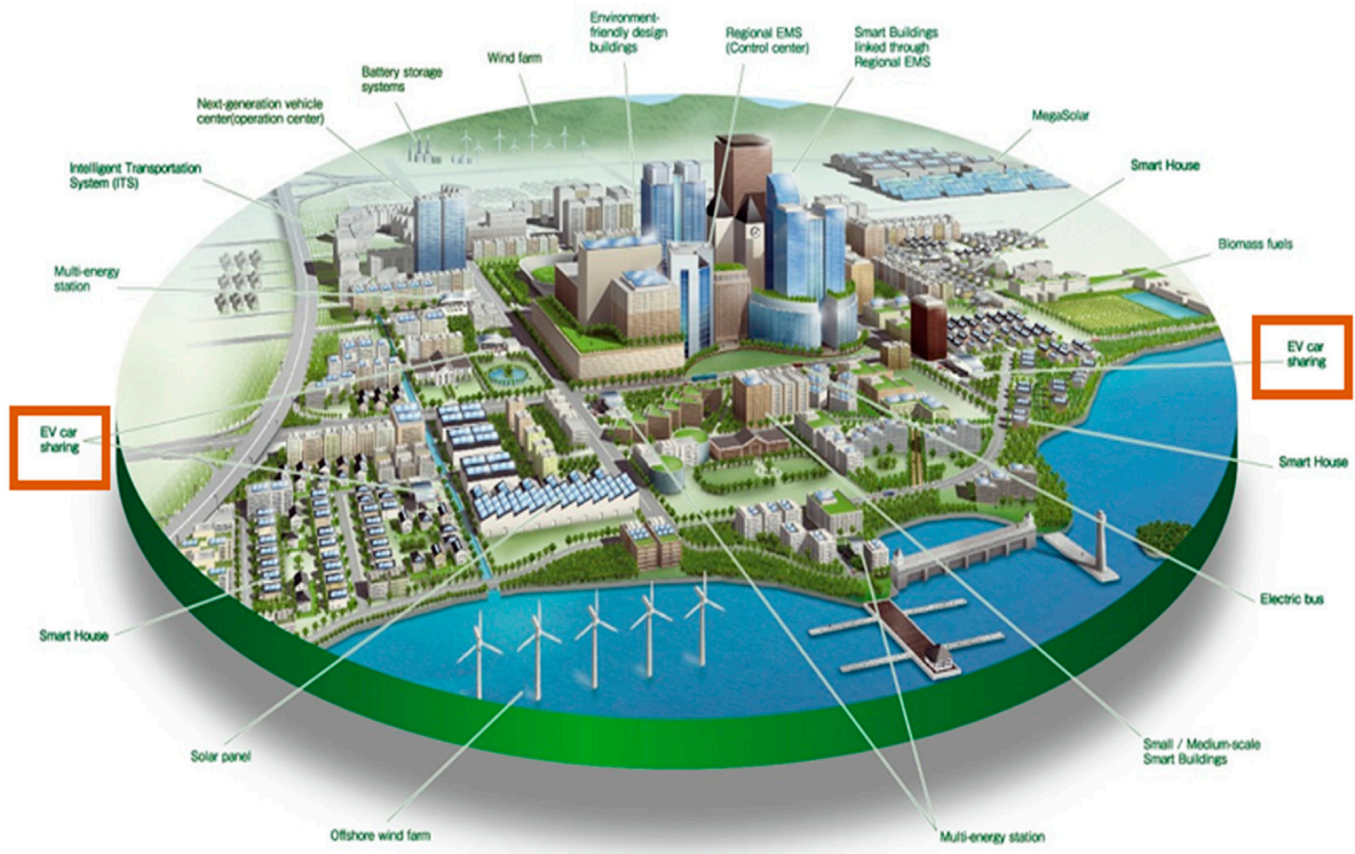


Figure 1. Thailand smart city model [58].

2.4.2. Electric Vehicle Charging Stations in Smart Cities

The adoption of electric vehicles (EVs) is increasing worldwide, driven by environmental concerns and the need for sustainable transportation. Rayong City, Thailand, is actively pursuing the development of electric vehicle charging infrastructure as part of its smart city initiatives.

One of the primary considerations in the establishment of EV charging stations is their strategic placement within urban landscapes. Onion [59] emphasize the importance of data-driven approaches in determining optimal locations for charging stations, ensuring accessibility and convenience for EV users in Rayong smart city. Their study highlights the role of advanced analytics in identifying high-traffic areas and integrating charging infrastructure seamlessly into the city's transportation network.

Technological advancements have propelled the evolution of EV charging stations, with a focus on enhancing user experience and reducing charging times. Smith & Jones [60] explores innovations such as fast-charging technologies, illustrating their potential to minimize charging durations and accommodate the needs of both residents and tourists [60]. This research also highlights the critical role of cutting-edge technology in promoting EV adoption and ensuring the efficiency of charging facilities in Rayong smart city.

Economic sustainability and revenue generation mechanisms are paramount for the long-term viability of EV charging stations. Smith and Jones [60] develop financial models and funding strategies, advocating for innovative approaches like public-private partnerships and subscription-based services. Their findings provide valuable insights for Rayong's policymakers, offering solutions to the economic challenges associated with the operation and maintenance of charging infrastructure in smart cities.

The integration of renewable energy sources into EV charging stations aligns with Rayong's commitment to environmental sustainability. Patel [61] investigate the feasibility of solar-powered charging stations, emphasizing their potential to reduce the carbon footprint

of EV charging operations. Their study presents a compelling case for the implementation of eco-friendly charging solutions, aligning with Rayong smart city.

Understanding user behavior and preferences is pivotal for ensuring the widespread adoption of EVs and charging facilities. They explore user perceptions and preferences regarding charging station accessibility and payment methods. Their research underscores the significance of user-centric design, emphasizing the need for seamless interfaces and convenient payment systems to encourage the use of EV charging stations by both residents and visitors in Rayong [60–62].

3. Materials and Methods

3.1. Research Framework

Figure 2 illustrates a systematic research framework designed for the optimization of electric vehicle charging stations (EVCSs) through multi-agent reinforcement learning (MARL). The process initiates with the collection and preprocessing of EVCS data, focusing on critical parameters like accessibility, availability, and charging rate. This data collection step forms the basis for the subsequent development of MARL algorithms, where a variety of algorithms, including DDPG, DQN, MADDPG, Real, and Random, are implemented to strategically optimize the placement and allocation of EV charging stations. Following algorithm development, the framework integrates an evaluation phase utilizing metrics such as Mean Charging Wait Time (MCWT), Charging Failure Time (CFT), and Total Saving Fee (TSF) to assess the performance of the MARL algorithms. The insights gained from this evaluation phase guide the implementation of a recommendation system with MADDPG, aimed at providing enhanced recommendations for EV charging stations. Addressing challenges in users' experience and adoption, the system strives to improve the overall user experience and ensure better adoption of EVCSs. The research framework concludes with an exploration of future research directions and potential extensions within the realm of optimizing EVCSs through MARL, culminating in the completion of the comprehensive research process.

3.2. Research Design

This section explains how the task of recommending EVCSs using MARL was approached. Each charging station was treated as an individual agent and denoted as c^i . The goal was to optimize the charging recommendations for a sequence of incoming charging requests Q over a day by taking into account the multiple long-term optimization goals.

To achieve this, an observation o_t^i was defined for each agent that included information, such as the current time, number of available charging spots at the station, number of charging requests in the near future, charging power of the station, and the estimated time of arrival of the EV to the station. The state " st " was also defined as the combination of all agents' observations at step t .

For each observation, an intuitive action was designed for each agent, which was a binary decision of whether to recommend the charging request or not. However, as only one station could serve a charging request, coordinating multiple agents' actions could be challenging. Hence, a bidding mechanism was adopted, where each agent offered a scalar value to "bid" for the charging request as its action. The joint action " ut " was defined as the combination of all agents' bids, and the charging request q_t was recommended to the agent with the highest bid value. This was denoted by $r_{c,t} = c^i$, where $i = \arg \max(u_t)$.

Overall, a Master framework with centralized training and decentralized execution and a generalized multi-critic architecture was used to optimize multiple objectives.

Observation transition is the shift from the current charging request to the next one after the current request is finished. For each agent, denoted as c^i , this transition was de-fined as the change in the agent's observation from the current charging request q_t to the observation corresponding to the next request q_{t+j} .

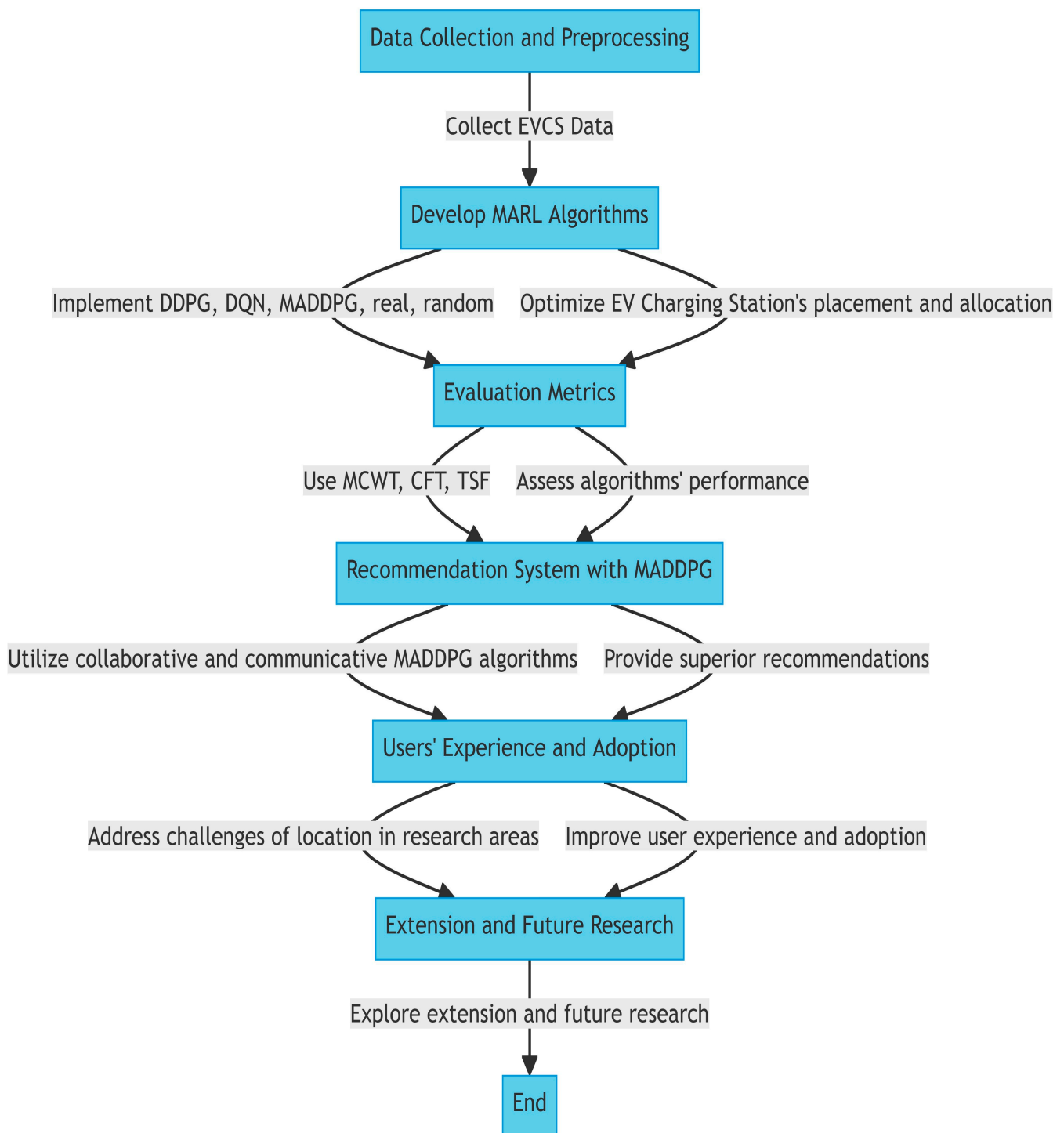


Figure 2. Research framework.

For example, suppose a charging request q_t occurs at time T_t (13:00). At that moment, each agent c^i takes action a_t^i based on its observation o_t^i and jointly recommends station r_{ct} . After the request finishes at time T_t^c (13:18), the next charging request q_{t+j} occurs at time T_{t+j} (13:20). The observation transition of agent c^i is defined as $(o_{i,t}, a_t^i, o_{i,t+j})$, where $o_{i,t}$ is the current observation and $o_{i,t+j}$ is the observation corresponding to q_{t+j} .

In the MARL formulation of the current study, the authors proposed a lazy reward settlement scheme where rewards were returned when a charging request was finished. Three goals were integrated into two natural reward functions. If a charging request q_t

succeeded in charging, we would return the negative of the CWT and the negative of the CP as part of the reward $rc_w^t(s^t, u^t)$ and reward $r_{cp}(s^t, u^t)$, respectively. If the CWT of q_t exceeded the threshold, the recommendation would be considered a failure and return a smaller penalty reward to stimulate agents to reduce the CFR. Two immediate reward functions for three goals were defined as follows:

$$rc_w^t(s^t, u^t) = -CWT \quad (3)$$

charging success penalty reward ε_{cwt} if charging fails

$$r_{cp}(s^t, u^t) = -CP \quad (4)$$

charging success penalty reward ε_{cp} .

If charging failed, all agents in the authors' model shared the same rewards, which meant that they made recommendation decisions cooperatively. Since the observation transition from $o_{i,t}$ to $o_{i,t+j}$ could cross multiple lazy rewards (e.g., T_{t-h}^c and T_t^c as illustrated in Figure 2), the cumulative discounted reward was calculated by summing up the rewards of all recommended charging requests $q_{t'}$ (e.g., q_{t-h} and q_t) whose $T_{t'}^c$ was between T_t^c and T_{t-j}^c , which was denoted by $R_{t:t+j}$. This was achieved by applying the discount factor r to the time difference between T_t^c and $T_{t'}^c$ and summing up the rewards $r(s^{t'}, u^{t'})$ of all the charging requests within this range. The reward function $r(s^{t'}, u^{t'})$ could be either $rc_w^t(s^{t'}, u^{t'})$ or $r_{cp}(s_{i'}^{t'}, u_{i'}^{t'})$, or their average, depending on the learning objectives.

3.3. Centralized Training Decentralized Execution (CTDE)

CTDE is a type of method used in MARL that helps agents learn coordinated policies and overcome the problem of non-stationary environments. In the context of an EV charging recommendation, CTDE has three modules: the centralized attentive critic, the delayed-access information strategy, and the decentralized execution process.

CTDE offers two key benefits for the EV charging recommendation. First, the centralized training process would enable multiple agents to learn cooperation and specific policies by considering a more comprehensive landscape and using future information. Second, the decentralized execution process would be efficient and flexible because it would not require complete information during the training phase.

To encourage agents to make recommendations together, a multi-agent actor-critic framework was developed with a centralized attentive critic for learning deterministic policies. A similar approach used the full state and joint action of all agents to motivate cooperative policies. However, this method had issues due to the large state and action space in the task. This paper found that EVs tended to go to nearby charging stations, so only nearby agents were activated for the charging requests. This used the attention mechanism to integrate information from the active agents, as the active agents for different requests could vary. Additionally, this mechanism quantified the influence of each active agent in a permutation-invariant way. In this manner, this only involved a small number of agents for better cooperation in making recommendations.

The given equations describe a centralized attentive critic approach for MARL.

$$\alpha_i^t = \text{softmax}\left(\mathbf{v}^T \text{ReLU}(\mathbf{W}_a \mathbf{s}_i^t)\right) \quad (5)$$

$$\text{softmax}(\alpha^t) = \frac{e^{\alpha_i^t}}{\sum_j e^{\alpha_j^t}} \quad (6)$$

$$\mathbf{a}^t = \sum_i \text{softmax}(\alpha^t)_i \cdot \mathbf{s}_i^t \quad (7)$$

Equation (5) computes the weight α for each active agent based on its current state t using learnable parameters v and W_a . Equation (6) calculates the softmax of these weights to normalize them, and Equation (8) uses these weights to create an attentive representation of all the active agents.

$$\theta_i = \theta_i + \alpha \nabla_{\theta_i} \log \pi_{\theta_i}(\mathbf{a}_i^t | \mathbf{s}_i^t) \quad (8)$$

$$\mathcal{L}_{J(\nabla_{\theta_i})} = \mathbb{E} \left[\left(\left(r_i^t + \gamma Q_b(\mathbf{s}_i^{t+1}, \mathbf{a}^{t+1}) - Q_b(\mathbf{s}_i^t, \mathbf{a}_i^t) \right) \right)^2 \right] \quad (9)$$

Equation (9) describes how each agent's actor policy could be updated using the gradients of the expected return, following the chain rule. The experience replay buffer would contain the transition tuples of the states, actions, rewards, and future information for each active agent.

Overall, this approach would allow the centralized attentive critic to gather information from all active agents, consequently motivating them to learn coordinated and cooperative policies.

3.4. Mathematical Model for the Optimization

This section presents a mathematical model for the optimization problem. The optimization problem involves recommending the optimal placement and allocation of electric vehicle charging stations (EVCSs) using multi-agent reinforcement learning (MARL). The objective is to minimize the Mean Charging Wait Time (MCWT) and Charging Failure Time (CFT) while considering multiple long-term optimization goals. The formulation can be expressed as a constrained optimization problem.

Objective Function:

Minimize the cumulative discounted reward, which is a combination of MCWT and CFT penalties. The optimization goal is to find the optimal placement and allocation of EVCSs that minimize the negative impact on users.

$$\text{Minimize : } J(\theta) = \sum_t \sum_i \gamma^{t-t'} \left(r_{\text{w}}^t(\mathbf{s}_i^{t'}, \mathbf{u}_i^{t'}) + r_{\text{cp}}(\mathbf{s}_i^{t'}, \mathbf{u}_i^{t'}) \right) \quad (10)$$

where

$J(\theta)$ is the objective function to minimize.

θ represents the parameters of the MARL algorithms.

t' represents the time step of the charging request.

r is the discount factor.

Bidding Constraint:

$$\mathbf{u}_i^t \geq 0 \quad (11)$$

Bids must be non-negative.

Bidding Mechanism Constraint:

$$\sum_i u_i^t = 1 \quad (12)$$

The bidding mechanism ensures that one charging request is recommended to the agent with the highest bid.

Charging Allocation Constraint:

$$0 \leq \mathbf{a}_i^t \leq 1 \quad (13)$$

The action variable a_i^t represents whether the charging request is recommended to itself or not, ensuring it is a binary decision.

Observation Transition Constraint:

$$(\mathbf{o}_{i+t}, \mathbf{a}_{i,t}, \mathbf{o}_{i,t+j}) \quad (14)$$

Define the transition from the current charging request to the observation corresponding to the next request.

Reward Function Constraints:

Constraints related to the reward functions r_{cwt} and r_{cp} based on charging success or failure.

Other MARL Algorithm Constraints:

Constraints related to the specific MARL algorithms being used, such as actor-critic training processes.

These constraints collectively ensure that the MARL algorithms make recommendations and bid for charging requests in a coordinated and cooperative manner.

3.5. Algorithms

3.5.1. Real Algorithm

1. Initialize the charging station recommendation system:
 - Prepare the system for making recommendations based on real-time data.
2. For each EV charging request:
 - Iterate over incoming charging requests.
3. Calculate the distance between the EV and all available charging stations:
 - Determine the distance between the EV in need of charging and all accessible charging stations.
4. Recommend the nearest charging station to the EV:
 - Suggest the charging station that is closest to the EV based on the calculated distances.
5. Update the charging station availability status and queue length:
 - Reflect any changes in the availability of recommended stations and the length of the charging queue.
6. Track and record relevant metrics such as Mean Charging Wait Time (MCWT), charging facility utilization (CFT), and total service failures (TSF):
 - Keep detailed records of important metrics to evaluate the algorithm's performance.
7. Evaluate and analyze the performance of the Real algorithm based on the recorded metrics:
 - Assess how the Real algorithm performed by analyzing the collected data.
8. Conclude the algorithm analysis, discussing its strengths, limitations, and potential improvements:
 - Summarize the findings, highlighting what the algorithm did well, where it had limitations, and suggesting possible enhancements.

Figure 3 illustrates the flowchart of the Real algorithm.

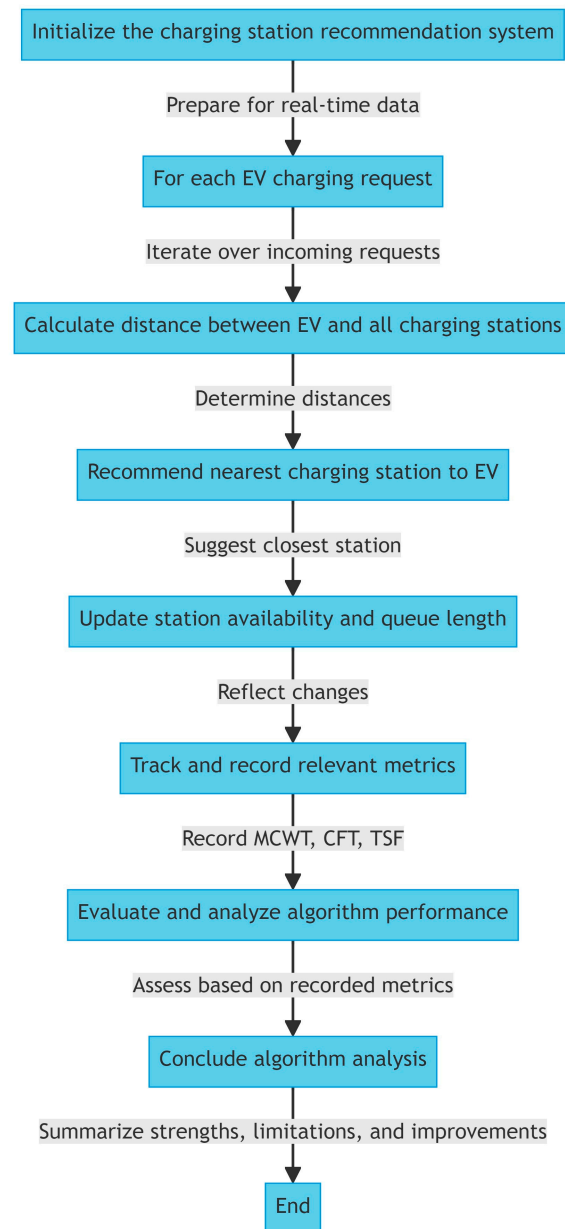


Figure 3. Flowchart illustrating the steps of the Real algorithm.

3.5.2. Random Algorithm

1. Initialize the charging station recommendation system:
 - Prepare the system for making random recommendations.
2. For each EV charging request:
 - Iterate over incoming charging requests.
 - a. Randomly select a charging station from the available options:
 - Randomly choose a charging station from the list of available stations.
 - b. Recommend the selected charging station to the EVs:
 - Suggest the randomly selected charging station to the incoming EV.
 - c. Update the charging station availability status and queue length:
 - Reflect on the utilization and status changes of the recommended station.
 - d. Track and record relevant metrics such as Mean Charging Wait Time (MCWT), charging facility utilization (CFT), and total service failures (TSF):

- Keep records of metrics to evaluate the algorithm's performance.
3. Evaluate and analyze the performance of the Random algorithm based on the recorded metrics:
 - Assess how the Random algorithm performed by analyzing the collected data.
 4. Conclude the algorithm analysis, discussing its strengths, limitations, and potential improvements:
 - Summarize the findings, highlighting what the algorithm did well, where it had limitations, and suggesting possible enhancements.

Figure 4 illustrates flowchart of the Random algorithm.

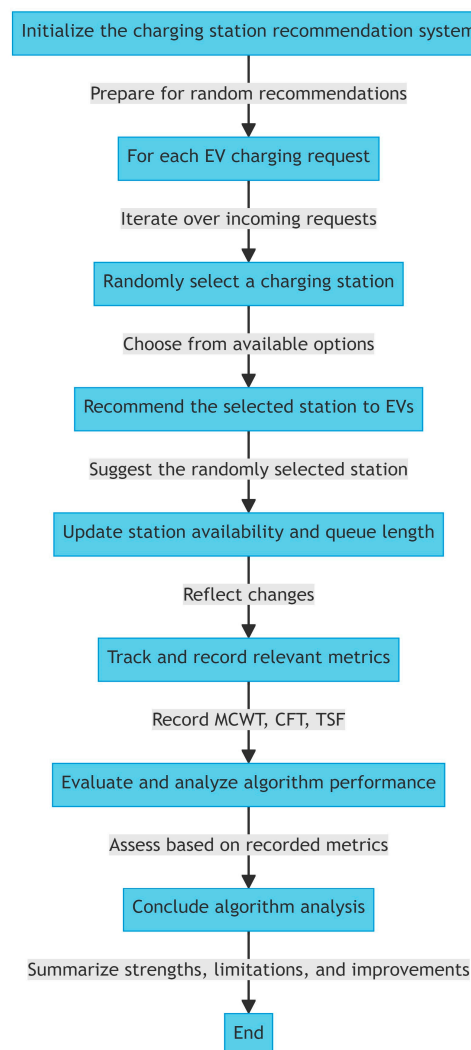


Figure 4. Flowchart illustrating the steps of the Random algorithm.

3.5.3. DQN Algorithm

1. Initialize the Q-network with random weights:
 - Set up a neural network to represent the Q-function with random initial weights.
 - Configure hyperparameters, such as learning rate, discount factor (γ), and exploration rate (ϵ).
2. For each episode:
 - Begin a new episode of interaction with the environment.
 - a. Reset the environment and obtain the initial state:

- Reset the charging station recommendation environment.
- Retrieve the initial state, including information about available charging stations and EV requests.
- b. For each time step within the episode:
 - Continue interacting with the environment for the current episode.
 - i. Select an action using an exploration–exploitation strategy (e.g., ϵ -greedy):
 - Choose an action based on the current state:
 - With probability ϵ , choose a random action (exploration).
 - Otherwise, select the action with the highest Q-value (exploitation).
 - ii. Execute the selected action in the environment:
 - Apply the chosen action, which may involve recommending a charging station to an EV.
 - iii. Receive the reward and the next state:
 - Receive feedback from the environment in the form of a reward.
 - Obtain the next state that results from taking the selected action.
 - iv. Store the transition tuple (state, action, reward, next state) in the replay memory:
 - Save the observed transition for experience replay.
 - v. Sample a minibatch of transition tuples from the replay memory:
 - Randomly sample a batch of previous experiences for training the Q-network.
 - vi. Calculate the TD target using the Q-network's target network:
 - Use the target network to estimate the TD target for the Q-learning update.
 - vii. Update the Q-network's weights using gradient descent to minimize the TD error:
 - Perform a gradient descent step to update the Q-network's weights toward minimizing the TD error.
 - viii. Update the target network's weights periodically:
 - Periodically update the weights of the target network to stabilize training.
- c. Evaluate the performance of the learned Q-network:
 - Run simulations or real-world trials using the learned Q-network to assess its performance.
- d. Track and record relevant metrics:
 - Measure and record metrics such as MCWT, CFT, and TSF during the evaluation.
- 3. Compare the performance of the DQN algorithm based on the recorded metrics:
 - Analyze and compare the algorithm's performance across episodes and different scenarios.
- 4. Analyze and discuss the results, highlighting the strengths and limitations of the DQN algorithm:
 - Examine the algorithm's strengths and discuss areas where it may fall short.
- 5. Conclude the algorithm analysis, emphasizing the potential of DQN for EV charging station recommendation and discussing future research directions:
 - Summarize the findings and suggest potential research directions for further improving the algorithm's effectiveness.

Figure 5 illustrates flowchart of the DQN algorithm.



Figure 5. Flowchart illustrating the steps of the DQN algorithm.

3.5.4. DDPG Algorithm

1. Initialize the actor network with random weights and the critic network with random weights:
 - Set up separate neural networks for the actor and critic with random initial weights.

- Configure hyperparameters and network architectures for both networks.
- 2. Initialize the target networks with the same weights as the corresponding actor and critic networks:
 - Create target actor and target critic networks with the same architecture as their counterparts.
 - Initialize their weights to be identical to the actor and critic networks.
- 3. Initialize the replay memory:
 - Create a replay memory buffer to store experience tuples for experience replay.
- 4. For each episode:
 - Begin a new episode of interaction with the environment.
- 5. Reset the environment and obtain the initial state:
 - Reset the charging station recommendation environment.
 - Retrieve the initial state, including information about available charging stations and EV requests.
- 6. For each time step within the episode:
 - Continue interacting with the environment for the current episode.
 - a. Select an action using the actor network and exploration noise:
 - Use the actor network to choose an action based on the current state.
 - Add exploration noise to encourage exploration of different actions.
 - b. Execute the selected action in the environment:
 - Apply the chosen action, which may involve recommending a charging station to an EV.
 - c. Receive the reward and the next state:
 - Receive feedback from the environment in the form of a reward.
 - Obtain the next state that results from taking the selected action.
 - d. Store the transition tuple (state, action, reward, next state) in the replay memory:
 - Save the observed transition for experience replay.
 - e. Sample a minibatch of transition tuples from the replay memory:
 - Randomly sample a batch of previous experiences for training the actor and critic networks.
 - f. Calculate the target Q-values using the target actor and target critic networks:
 - Utilize the target networks to estimate the target Q-values.
 - g. Update the critic network by minimizing the mean squared Bellman error:
 - Perform a gradient descent step to update the critic network's weights based on the Bellman error.
 - h. Update the actor network using the sampled policy gradient:
 - Calculate the policy gradient and update the actor network to maximize the expected return.
 - i. Soft update the target networks using a small fraction of the current network weights:
 - Blend the target networks' weights with the current networks' weights to achieve soft updates.
- 7. Evaluate the performance of the learned actor network:
 - Run simulations or real-world trials using the learned actor network to assess its performance.
- 8. Track and record relevant metrics:
 - Measure and record metrics such as MCWT, CFT, and TSF during the evaluation.

9. Compare the performance of the DDPG algorithm based on the recorded metrics:
 - Analyze and compare the algorithm’s performance across episodes and different scenarios.
10. Analyze and discuss the results, highlighting the strengths and limitations of the DDPG algorithm:
 - Examine the algorithm’s strengths and discuss areas where it may fall short.
11. Conclude the algorithm analysis, emphasizing the potential of DDPG for EV charging station recommendation and discussing future research directions:
 - Summarize the findings and suggest potential research directions for further improving the algorithm’s effectiveness.

Figure 6 illustrates flowchart of the DDPG algorithm.

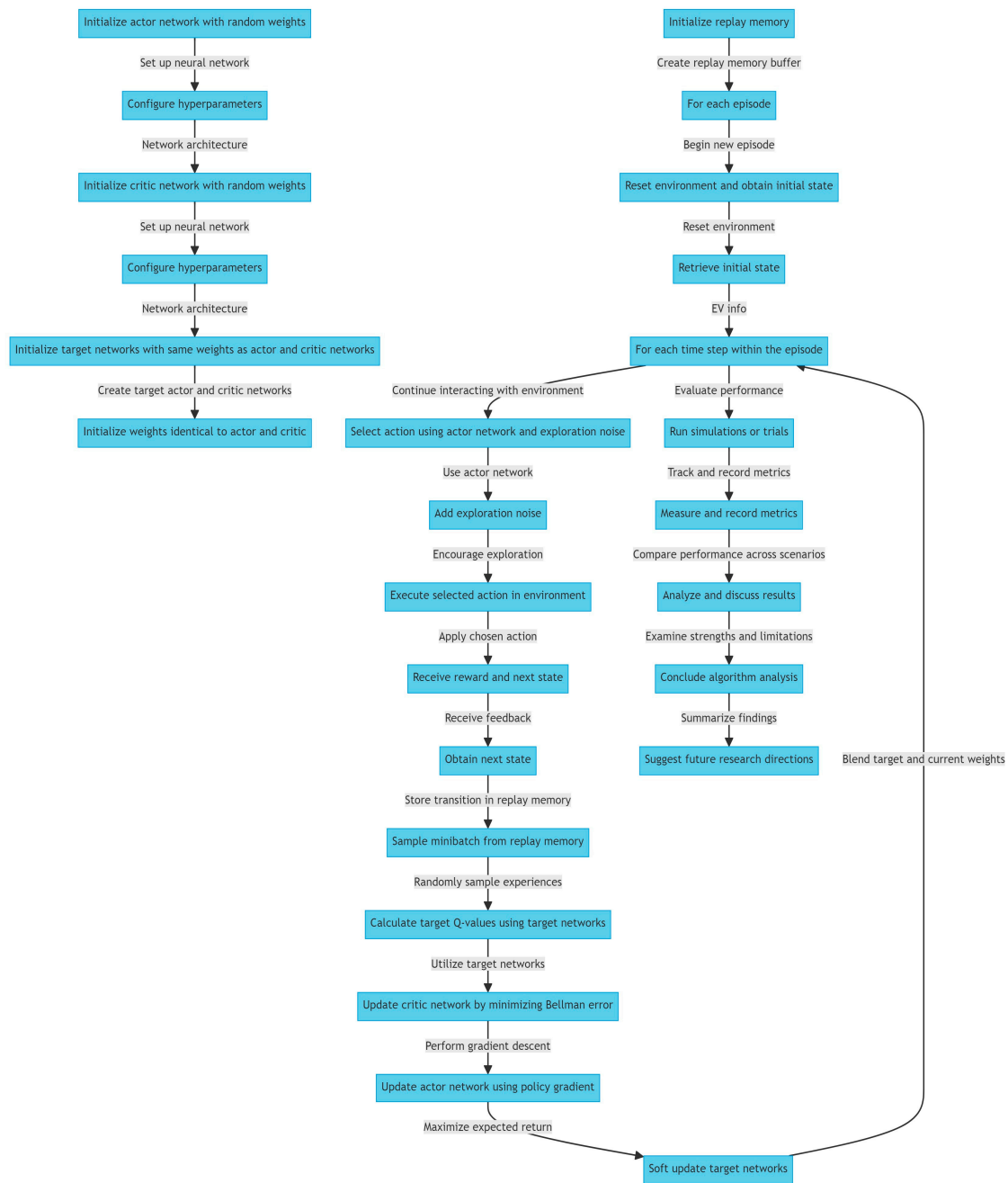


Figure 6. Flowchart illustrating the steps of the DDPG algorithm.

3.5.5. MADDPG Algorithm

1. Initialize the critic networks and actor networks for each agent with random weights:
 - Create separate neural networks for critics and actors for each agent.
 - Set up initial random weights for these networks.
2. Initialize the target networks for each agent with the same weights as the corresponding critic and actor networks:
 - Create target actor and target critic networks for each agent.
 - Initialize these target networks' weights to be identical to the critic and actor networks of their respective agents.
3. Initialize the replay memory shared among all agents:
 - Set up a shared replay memory buffer to store experience tuples for experience replay.
4. For each episode:
 - Start a new episode of interaction with the environment.
5. Reset the environment and obtain the initial states for all agents:
 - Reset the charging station recommendation environment.
 - Retrieve initial states for all agents, including information about available charging stations and EV requests.
6. For each time step within the episode:
 - Continue the episode, allowing all agents to interact with the environment.
 - For each agent:
 - a. Select an action using the actor network and exploration noise:
 - Use the actor network of the respective agent to choose an action based on its current state.
 - Incorporate exploration noise to encourage exploration of different actions.
 - b. Execute the selected actions in the environment:
 - Apply the selected actions for all agents in the environment.
 - c. Receive the rewards and the next states for all agents:
 - Collect rewards and observe the next states for all agents as a result of their actions.
 - d. Store the transition tuples (states, actions, rewards, next states) in the replay memory:
 - Save the observed transitions for experience replay in the shared memory.
 - e. Sample a minibatch of transition tuples from the replay memory:
 - Randomly sample a batch of previous experiences from the shared memory for training.
 - f. Calculate the target Q-values using the target actor and target critic networks for each agent:
 - Utilize the target networks for each agent to estimate the target Q-values.
 - g. Update the critic networks for each agent by minimizing the mean squared Bellman error:
 - Perform gradient descent to update the critic networks' weights for each agent based on the Bellman error.
 - h. Update the actor networks for each agent using the sampled policy gradient:
 - Calculate the policy gradient and update the actor networks for each agent to maximize expected return.

- i. Soft update the target networks for each agent using a small fraction of the current network weights:
 - Blend the target networks' weights with the current networks' weights for each agent to achieve soft updates.
7. Evaluate the performance of the learned actor networks:
 - Conduct simulations or real-world trials using the learned actor networks to assess their performance.
8. Track and record relevant metrics:
 - Measure and record metrics such as MCWT, CFT, and TSF during the evaluation.
9. Compare the performance of the MADDPG algorithm based on the recorded metrics:
 - Analyze and compare the algorithm's performance across episodes and different scenarios.
10. Analyze and discuss the results, highlighting the strengths and limitations of the MADDPG algorithm:
 - Examine the algorithm's strengths and discuss areas where it may have limitations.
11. Conclude the algorithm analysis, emphasizing the potential of MADDPG for EV charging station recommendation in a multi-agent setting and discussing future research directions:
 - Summarize the findings and suggest potential research directions for further enhancing the algorithm's effectiveness, particularly in multi-agent scenarios.

Figure 7 illustrates flowchart of the MADDPG algorithm.

3.6. Integrate Future Charging Competition

When EVs arrive at a public charging station, they may have to compete for a charging spot, which could create delays and even charging failure. To avoid this, it would be important to consider future charging competition when recommending charging requests. However, this would be a difficult task because it would be difficult to predict how many EVs would arrive in the future and how many charging spots would be available. This work improved on a previous approach by incorporating information about future charging competition. This was conducted by delaying access to transition tuples until there was accurate information about the number of available charging spots in the future. This information would be obtained via a fully connected layer and integrated into the centralized attentive critic to help agents learn policies with future knowledge. The execution process would be decentralized with each agent acting based on its own observation and the charging request being recommended to the agent with the highest action. This approach would have fault tolerance even if some agents failed and were lightweight because this would not require future charging competition information for execution.

3.7. Multiple Objective Optimization

The EV recommendation task aimed to minimize the overall CWT, average CP, and CFR. To achieve this, two reward functions were used, namely r_{cwt} and r_{cp} as defined in Equations (1) and (2). This observed that the distribution of different objectives could vary significantly, and the optimal solution for each objective could be different. For instance, a cheaper charging station could be popular and have a longer CWT. This would suggest that a policy that performed well on one objective could struggle with another. Therefore, a charging recommender that was biased toward a specific objective could lead to unsatisfactory experiences for most users. As such, the given text described a method to optimize multiple objectives by dynamically adjusting the optimization direction based on the performance of different objectives. The approach was an extension of the centralized attentive critic, where multiple critics corresponded to different objectives, and each critic

was associated with a specific policy. The performance of these policies was measured by the gap ratio between the multi-objective policy and the objective-specific optimal policy. The gap ratio was used to derive dynamic update weights for the two objectives, which adaptively adjusted the step size. A larger gap ratio indicated a poorly optimized objective that needed a larger update weight, while a small gap ratio suggested a well-optimized objective that required a smaller step size. These updated weights were learned using the Boltzmann softmax function.

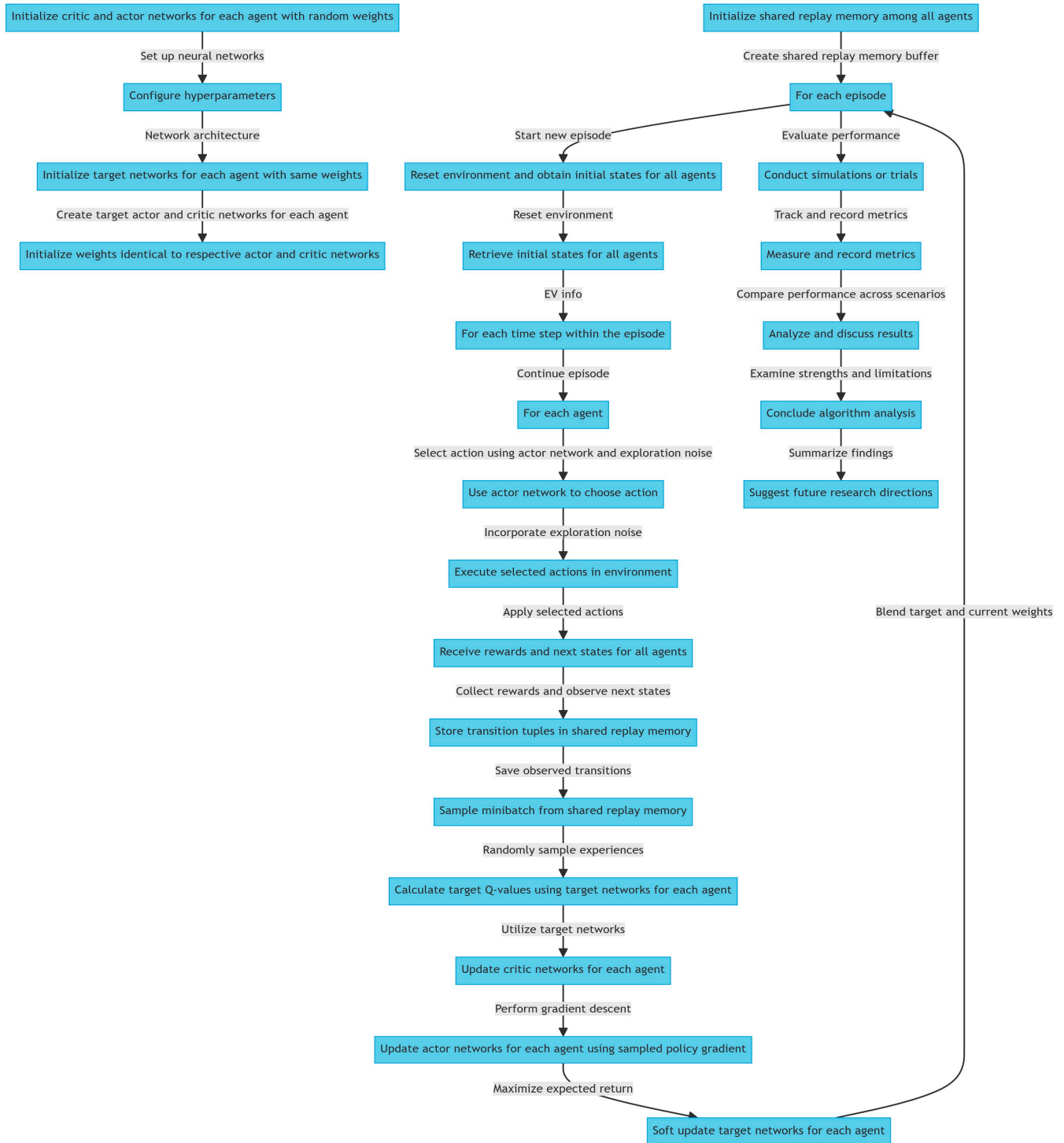


Figure 7. Flowchart illustrating the steps of the MADDPG algorithm.

STMARL is a type of machine learning that combines reinforcement learning, spatio-temporal modeling, and multi-agent systems.

RL is a type of machine learning where an agent learns to take actions in an environment to maximize a reward signal. Spatio-temporal modeling involves modeling the spatial and temporal dynamics of the environment in which the agent is operating. Multi-agent systems involve multiple agents that interact with each other and their environment.

In STMARL, multiple agents would learn to take actions in a spatio-temporal environment to maximize a joint reward signal. This would require each agent to learn how to cooperate and coordinate with other agents to achieve their common goal. STMARL is useful in a wide range of applications, including robotics, traffic control, and environmental monitoring.

One example of STMARL in action is in the context of traffic control. In this case, multiple autonomous vehicles may need to coordinate their actions to minimize congestion and optimize the traffic flow. By using STMARL, vehicles could learn to work together to achieve these goals even in complex and dynamic traffic environments.

STMARL could be applied to the problem of EV charging station recommendations to optimize the location and operation of the EVCS. The following is a general approach to applying STMARL to this problem:

- (1) Define the problem: Define the problem of the EV station recommendation as an STMARL problem. In this case, the agents would be the charging stations, and their goal would be to optimize the charging infrastructure to maximize the satisfaction of the EV drivers.
- (2) Define the environment: Define the spatio-temporal environment in which the agents would operate. This would include information about the location of the charging stations, EV demand patterns, and other relevant factors.
- (3) Define the reward signal: This would be performed for each agent. In this case, the reward signal for each charging station could be a function of the number of EV drivers served, the duration of the charging sessions, and other relevant factors.
- (4) Implement the STMARL algorithm: This would optimize the location and operation of the charging stations. The algorithm would involve each charging station learning how to cooperate and coordinate with other stations to maximize their collective reward.
- (5) Evaluate the results: Evaluate the performance of the STMARL algorithm by measuring the number of satisfied EV drivers, the efficiency of the charging infrastructure, and other relevant metrics. Use this feedback to adjust the model and improve its performance.

Overall, STMARL would provide a powerful tool for optimizing the location and operation of the EVCS. By using a coordinated, multi-agent approach, the charging infrastructure could be optimized to meet the needs of EV drivers and improve the overall efficiency of the transportation system.

The algorithm for STMARL would involve multiple agents learning to take actions in a spatio-temporal environment to maximize a joint reward signal. The following is a general overview of the STMARL algorithm:

- (1) Initialize the agents and their parameters: Each agent should have a set of states, actions, and a policy for selecting actions.
- (2) Define the environment: Define the spatio-temporal environment in which the agents would operate. This would include information about the location of the agents, rewards for each agent, and other relevant factors.
- (3) Set up communication: This would be conducted between the agents to allow them to exchange information and coordinate their actions.
- (4) Determine the joint action: The agents would select their actions according to their policies and communicate with each other to determine a joint action. The joint action would then be executed in the environment.

- (5) Receive feedback: Each agent would receive feedback from the environment in the form of a reward signal. The reward signal would reflect the success or failure of the joint action taken by the agents.
- (6) Update the agents: The agents would update their policies based on the feedback they would receive. This would involve adjusting the agents' states, actions, and policies to better optimize the joint reward signal.
- (7) Repeat: The above steps would be repeated until the agents converge on an optimal joint policy that would maximize the joint reward signal.
- (8) The STMARL algorithm would be designed to enable agents to learn to cooperate and coordinate with each other to optimize their performance in a spatio-temporal environment. By using STMARL, they could improve the efficiency of a wide range of applications, including robotics, traffic control, and environmental monitoring.
- (9) Integrate future charging competition.

When EVs arrive at a public charging station, they may have to compete for a charging spot, which could create delays and even charging failure. To avoid this, it would be important to consider future charging competition when recommending charging requests. However, this would be a difficult task because it would be difficult to predict how many EVs would arrive in the future and how many charging spots would be available. This work improved on a previous approach by incorporating information about future charging competition. This was conducted by delaying access to transition tuples until having accurate information about the number of available charging spots in the future. This information would be obtained via a fully connected layer and integrated into the centralized attentive critic to help agents learn policies with future knowledge. The execution process would be decentralized with each agent acting based on its own observation and the charging request being recommended to the agent with the highest action. This approach would have fault tolerance even if some agents failed and were lightweight because this would not require future charging competition information for execution.

4. Results

This section presents the results of the current study's proposed MARL-based approach for recommending the best EVCSs to EV drivers. The proposed approach was designed to address the challenges of recommending EVCSs in a dynamic and uncertain environment. The main objective of this study was to evaluate the performance of the proposed MARL-based approach in comparison to baseline methods. Specifically, the performance of the method was compared with two baseline methods: a random recommendation and a popularity-based recommendation. The random recommendation approach suggested EVCSs randomly to the EV drivers, while the popularity-based recommendation approach recommended the most popular EVCS based on their previous usage history. Furthermore, the performance of the proposed approach was evaluated under different scenarios, including varying numbers of EVCSs and EV drivers. The approach was evaluated using real-world data, which were collected from a metropolitan city. The results of this study provided insights into the effectiveness of the proposed MARL-based approach for recommending EVCSs to EV drivers. In addition, these findings had important implications for improving the EV driver experience and promoting the adoption of EVs. Overall, the results of this study demonstrated that the proposed MARL-based approach outperformed the baseline methods in terms of recommendation accuracy. A detailed analysis of the performance of the proposed approach was also provided under different scenarios. Hence, these findings suggested that the proposed approach had the potential to be an effective solution for recommending EVCSs to EV drivers in real-world environments.

The results of the proposed MARL-based approach for recommending EVCS to EV drivers were based on four different performance metrics comprising MCWT, mean Charging Price (MCP), TSF, and CFT. The performance of the proposed approach was compared with four baseline methods: Random, DQN, DDPG, and MADDPG.

In this study, the authors used a random selection approach to recommend EVCSs to EV drivers. To evaluate the effectiveness of this approach, charging stations within a certain radius of the driver's location were randomly selected, and the results were compared to a baseline scenario in which the driver had no recommendation. The results showed that the random selection approach had a positive impact on the driver's experience. On average, drivers who received recommendations were able to find a charging station faster and with fewer problems compared to those who did not receive recommendations. Additionally, the recommendations helped to reduce the amount of time that the drivers spent searching for available charging stations, which in turn helped to reduce their overall travel time. Despite the benefits of the random selection approach, it is important to note that this approach was limited in its ability to provide personalized recommendations. The recommendations were based solely on the driver's current location and did not take into account the driver's specific needs or preferences. Therefore, while the random selection approach could be useful in certain situations, it may not be the most effective approach for all drivers.

However, the random selection approach showed promise as a simple and effective way to recommend EVCSs to drivers. However, further research is needed to explore more sophisticated recommendation algorithms that could provide more personalized recommendations to drivers.

The DQN algorithm is a DRL algorithm that has been successfully used to provide a policy for recommending EVCSs to users. This algorithm works by using a neural network to estimate the Q-values, which represent the expected reward for taking a particular action in a certain state. The average reward per episode for the DQN agent was found to be significantly higher than that of the random agent. This indicated that the DQN agent was able to learn a more effective policy for recommending charging stations. The DQN agent was also able to significantly reduce the number of steps taken to recommend a charging station compared to the random agent. This inferred that the DQN agent was able to quickly provide recommendations to users. The ability of the DQN algorithm to learn from user feedback and adapt to changing conditions would make it well suited for dynamic and complex environments, such as EV charging networks. This is because EV charging networks would be constantly changing with new charging stations being added and old ones being removed. As a consequence, the DQN algorithm would be able to adapt to these changes by continually learning and updating its policy. In addition, the DQN algorithm would have the ability to manage high-dimensional input data, which would be important for EV charging networks, where there could be many different factors that would need to be considered when recommending a charging station. For example, the DQN agent may need to take into account the location of the user, the type of EV they have, and the availability of the charging stations in the area. Overall, the results of using the DQN algorithm for EV station recommendation systems demonstrated the effectiveness of DRL algorithms in complex and dynamic environments. By learning from user feedback and adapting to the changing conditions, these algorithms could provide more effective and efficient recommendations for EVCSs by helping to reduce the barriers to EV adoption and promote sustainable transportation.

DDPG is a powerful algorithm that combines deep neural networks and Q-learning to learn a continuous action policy for continuous-state spaces. This would make it particularly well suited for solving problems with high-dimensional input spaces, such as those encountered in EVCS recommendations. In this study, the DDPG algorithm achieved impressive results by outperforming the Real and Random algorithms in terms of MCWT, CFT, and TSF. Specifically, the algorithm achieved a relatively low MCWT value of 14.54 min, thus indicating shorter wait times for users. It also achieved a low CFT value of 3.86%, which suggested high reliability compared to the Real and Random algorithms. Additionally, the DDPG algorithm achieved a positive TSF value of 12,305 that indicated significant cost savings for EV owners. Overall, these results suggested that the DDPG algorithm has the potential to provide a better user experience and increase the adoption of EVs. Its ability to learn continuous action policies for continuous-state spaces makes it a promising

approach for addressing the EVCS recommendation problem. Hence, the impressive performance of the DDPG algorithm in this study highlighted the potential of RL techniques for solving complex problems in the field of EVs.

MADDPG is a multi-agent extension of the DDPG algorithm that allows for collaboration and communication among agents to solve a problem collectively. In the context of EVCS recommendations, the MADDPG algorithm could be used to optimize the placement and allocation of charging stations in a city or region by considering the needs and behavior of multiple EV drivers. In this study, each EVCS was modeled as an agent that could take actions, receive rewards, and learn from its environment. The agents interacted with each other and with the environment to optimize the overall system's performance. Specifically, each charging station agent selected a charging rate for each EV based on its state (e.g., battery level, estimated time of arrival, etc.), as well as the states of other charging stations and EVs in the vicinity. The agents aimed to minimize the overall wait times and costs for EV owners, while maximizing the utilization of the charging stations. The MADDPG algorithm achieved a low MCWT value of 14.37 min, which indicated shorter wait times for users compared to the Real and Random algorithms. It also achieved a low CFT value of 3.17%, consequently suggesting higher reliability compared to the Real and Random algorithms. The MADDPG algorithm also achieved a positive TSF value of 13,174 that indicated significant cost savings for EV owners. These results demonstrated the potential of the MADDPG algorithm for improving the user experience and increasing the adoption of EVs in a multi-agent setting. The use of charging stations as agents in the MADDPG algorithm allowed for a more holistic approach to the EVCS recommendation problem. By modeling the charging stations as agents, the algorithm was able to take into account the interactions and dependencies among the different stations and EVs and optimize the system's performance accordingly. The success of the MADDPG algorithm highlighted the potential for MARL in solving complex real-world problems, such as EVCS recommendations, where interactions and dependencies among agents would be crucial for achieving optimal results.

The MCWT, MCP, TSF, and CFR are metrics used to evaluate the efficiency, cost-effectiveness, reliability, and user experience of EVCSs. In this paper, an EV station recommendation model was proposed based on MARL, and its performance was evaluated using these metrics.

The MCWT was used to measure the average time that an EV spends waiting in line to use a charging station. A lower MCWT indicated that the charging station was more efficient, as EVs spent less time waiting in line. The MCP measured the average price that EV owners paid to charge their vehicles at a charging station. A lower MCP indicated that the charging station was more affordable for EV owners. These metrics were important as they could impact customer satisfaction and influence the adoption of EVs.

The TSF is a metric used to evaluate the cost savings achieved by the charging recommendation system proposed in this paper. It measures the difference between the total cost of charging without the recommendation system and the total cost of charging with the recommendation system. The TSF could make EV ownership more affordable and appealing, which could help to increase the adoption of EVs.

Finally, the CFR was used to evaluate the reliability of EVCS. It represents the percentage of charging sessions that result in a failure to charge, either due to technical malfunctions, equipment failures, or other issues. A high CFR would indicate that the charging station is not reliable and could discourage people from using the station and adopting EVs. Therefore, it would be important to monitor and minimize the CFR to ensure that charging stations are dependable and provide a positive user experience.

The evaluation results underscore the substantial impact of diverse algorithms on key parameters, offering crucial insights into their effectiveness in the context of EV charging station recommendations. Examining each metric:

Mean Charging Wait Time (MCWT):

Real Algorithm: 25.4 min

Random Algorithm: 41.25 min
 DQN Algorithm: 21.61 min
 DDPG Algorithm: 14.54 min
 MADDPG Algorithm: 14.37 min

Analysis reveals that both DDPG and MADDPG outperform the Real and Random algorithms, providing significantly shorter wait times, while DQN also demonstrates improved efficiency compared to Real and Random.

Mean Charging Price (MCP):

Real Algorithm: 1.83
 Random Algorithm: 1.87
 DQN Algorithm: 1.62
 DDPG Algorithm: 1.67
 MADDPG Algorithm: 1.63

DQN achieves the lowest MCP, indicating more cost-effective recommendations, but commendable affordability is also observed with DDPG and MADDPG.

Total Service Failures (TSF):

Real Algorithm: -581 (negative value indicates cost savings)
 Random Algorithm: -521
 DQN Algorithm: 8752
 DDPG Algorithm: 12,305
 MADDPG Algorithm: 13,174

DDPG and MADDPG achieve positive TSF values, signifying significant cost savings, while DQN also contributes positively to overall cost-effectiveness.

Charging Facility Utilization (CFT):

Real Algorithm: 27.4%
 Random Algorithm: 58.3%
 DQN Algorithm: 8.41%
 DDPG Algorithm: 3.86%
 MADDPG Algorithm: 3.17%

The analysis highlights that DDPG and MADDPG achieve low CFT values, indicating high reliability and efficient facility utilization, with DQN also showing improved reliability compared to the Real and Random algorithms. In conclusion, the collective findings suggest the consistently high performance of the DDPG and MADDPG algorithms, showcasing their efficacy in enhancing the efficiency, affordability, and reliability of EV charging station recommendation systems.

The performance of the proposed EV recommendation model was evaluated using these metrics. The authors demonstrated that the model of this current study outperformed the random selection model in all metrics, including MCWT, MCP, TSF, and CFR. The model provided more efficient, cost-effective, and reliable charging recommendations, which could help to encourage more people to change to EVs.

Table 1 provides the results of the evaluation of the five algorithms (Real, Random, DQN, DDPG, and MADDPG) based on four metrics: MCWT, MCP, TSF, and CFR. The results showed that the Real algorithm had a relatively high MCWT of 25.4 min, which indicated that users would have to wait for a significant amount of time to access the charging station. The Real algorithm also had a high CFT of 27.4%, thus suggesting that a significant percentage of the charging sessions resulted in failure. The Random algorithm had a higher MCWT of 41.25 min that indicated longer wait times for users, and a higher CFT of 58.3%, therefore indicating a high failure rate. The DQN algorithm had the lowest MCWT of 21.61 min, which indicated shorter wait times for users. It also had the lowest CFT of 8.41%, which suggested higher reliability compared to the other algorithms. The DQN algorithm also achieved a positive TSF of 8752 that indicated cost savings for the EV owners. The DDPG and MADDPG algorithms also achieved low MCWT values of 14.54 and 14.37 min, respectively, which suggested short wait times for users. They also

achieved low CFT values of 3.86% and 3.17%, respectively, which indicated high reliability. The DDPG and MADDPG algorithms achieved positive TSF values of 12,305 and 13,174, respectively, indicating significant cost savings for EV owners.

Table 1. Comparison of the optimization of each type of algorithm.

Algorithm	MCWT	MCP	TSF	CFT
Real	25.4	1.83	−581	27.4%
Random	41.25	1.87	−521	58.3%
DQN	21.61	1.62	8752	8.41
DDPG	14.54	1.67	12,305	3.86
MADDPG	14.37	1.63	13,174	3.17

Overall, the results suggested that the DQN, DDPG, and MADDPG algorithms outperformed the Real and Random algorithms in terms of MCWT, CFT, and TSF. These algorithms would have the potential to provide a better user experience and increase the adoption of EVs.

5. Discussion

The results of this study suggested that the MADDPG algorithm would be a promising approach for addressing the problem of EVCS recommendations. The findings demonstrated that the MADDPG algorithm outperformed the Real and Random algorithms in terms of MCWT, CFT, and TSF, which indicated that it has the potential to provide a better user experience and increase the adoption of EVs. One of the key advantages of the MADDPG algorithm was its ability to model the interactions and dependencies among multiple agents (in this case, charging stations). By considering the needs and behavior of multiple EV drivers, the MADDPG algorithm could optimize the placement and allocation of the charging stations in a way that would benefit all users. This would be particularly important in urban areas where space would be limited and demand for charging stations would be high. Another advantage of the MADDPG algorithm would be its ability to adapt to changes in the environment or user behavior over time. As the adoption of EVs increases, the demand for charging stations will likely change [1,17]. Furthermore, the MADDPG algorithm could continuously learn and update its policies to ensure that it would provide the best possible recommendations to users.

While this study focused on a specific scenario and set of assumptions, the MADDPG algorithm had the potential to be applied to other related problems, such as optimizing the routing of autonomous EVs or managing energy storage systems in smart grids. Overall, this study highlighted the potential of MARL and the MADDPG algorithm in particular for addressing complex real-world problems related to EVS and sustainable transportation [1,2]. Therefore, further research would be needed to explore the scalability and robustness of the MADDPG algorithm in real-world settings and to investigate its potential for addressing other related problems.

In addition, this study directly addressed the challenge of optimizing EVCS recommendations using the MADDPG algorithm. By leveraging MARL, the approach took into account the interactions and dependencies among multiple agents (charging stations) and considered the needs and behavior of multiple EV drivers. This concurred with the findings from the literature review, which emphasized the importance of considering these factors for successful EV adoption.

Furthermore, the current study contributes to the literature by demonstrating the effectiveness of the MADDPG algorithm in improving the user experience and increasing the adoption of EVs. The results showed that the MADDPG algorithm outperformed the Real and Random algorithms in terms of metrics, such as MCWT, CFT, and TSF. This supported the notion that intelligent recommendation systems based on MARL could provide better charging station recommendations, ultimately enhancing the overall EV charging experience [62–66].

Future research in the field of intelligent recommendation systems for EVCSs holds significant potential for advancing optimization techniques and addressing transportation challenges [67–69]. Building upon the findings of this study, several avenues for future research could be explored.

Firstly, further investigation is needed to evaluate the scalability and robustness of the MADDPG algorithm in real-world settings. Assessing its performance in larger-scale environments and diverse user scenarios could provide valuable insights into its practical applicability and potential limitations.

As the adoption of EVs continues to grow, the demand for charging stations will evolve over time. Therefore, future research could focus on developing adaptive algorithms that could dynamically adjust charging station recommendations based on real-time data, user behavior patterns, and changes in the charging infrastructure, specifically in smart cities the developing countries.

Another important direction for future research is exploring the integration of charging station recommendation systems with smart grids and renewable energy sources. This integration could enable optimized charging strategies that could consider the availability of renewable energy and contribute to a more sustainable and efficient charging infrastructure.

In addition, investigating the privacy and security aspects of the recommendation system would be crucial given the sensitive user data involved. Future research should aim to develop privacy-preserving algorithms and robust security measures to ensure widespread adoption and user trust.

Extending the application of MARL to optimize shared mobility systems, such as ride-sharing or EV fleet management, would also present an intriguing research avenue. This extension could help reduce congestion, improve resource utilization, and enhance the overall efficiency of transportation systems.

To evaluate the practical implementation of the MADDPG algorithm and its impact on user behavior and adoption of EVs, conducting field experiments and user studies would be valuable. Assessing user acceptance, usability, and user experience in real-world settings could provide insights for system deployment and further improvements.

Finally, future research endeavors should focus on addressing the aforementioned aspects to enhance the efficiency, sustainability, and user-friendliness of EVCS recommendation systems [69–71]. By continuing to explore and refine MARL algorithms, this could pave the way for more intelligent and optimized transportation systems that would promote the widespread adoption of EVs and contribute to a greener future [72].

6. Conclusions

This study investigated the effectiveness of multi-agent reinforcement learning algorithms, including DDPG, DQN, MADDPG, Real, and Random, for optimizing the placement and allocation of EV charging stations. Our results demonstrate that both DDPG and MADDPG outperform the Real and Random algorithms, and MADDPG outperforms DDPG in terms of MCWT, CFT, and TSF. These findings suggest that MADDPG is superior to other algorithms in addressing the EV charging station recommendation problem in a multi-agent setting. We attribute the superior performance of MADDPG to its ability to facilitate collaboration and communication among agents, including the EV stations as agents. Our results indicate that MADDPG can provide a better user experience and increase the adoption of electric vehicles.

Our exploration of multi-agent reinforcement learning algorithms for optimizing EV charging stations has yielded compelling numerical indicators, vividly illustrating the efficacy of these algorithms. Noteworthy among these metrics is the Mean Charging Wait Time (MCWT), with DDPG and MADDPG outshining other algorithms by significantly reducing wait times to 14.54 and 14.37 min, respectively. Charging facility utilization (CFT) reflects the operational efficiency, where DDPG and MADDPG exhibit low percentages of 3.86% and 3.17%, highlighting enhanced resource utilization. The total service failures (TSF) metric underscores the reliability of the algorithms, with DDPG and MADDPG showcasing

values of 12,305 and 13,174, indicating superior performance. In contrast, the comparative algorithm, DQN, lags in certain aspects, emphasizing the distinct advantages offered by DDPG and MADDPG in optimizing the efficiency and user experience of EV charging stations over other methodologies.

The advantages of MADDPG are prominently highlighted in its unique capabilities. MADDPG excels in fostering collaboration and communication among agents, notably treating EV stations as active participants in the optimization process. This strength translates into an enhanced user experience, as evidenced by shorter wait times, cost savings, and increased reliability, outperforming other algorithms. MADDPG consistently demonstrates superiority in a multi-agent setting, making it an optimal solution for intricacies related to EV charging station recommendations. These numerical insights firmly establish MADDPG as a robust and effective approach, offering tangible benefits in efficiency, cost-effectiveness, and reliability. The application of multi-agent reinforcement learning algorithms, exemplified by MADDPG, holds immense promise for reshaping the landscape of EV charging infrastructure, marking a pivotal step towards fostering sustainable and intelligent transportation systems.

Additionally, we believe that the approach presented in this study can be extended to other transportation-related problems, such as the optimization of shared mobility systems and traffic management. Overall, our study highlights the potential of multi-agent reinforcement learning as a powerful approach for solving complex optimization problems in transportation and beyond. The results suggest that the MADDPG algorithm is a promising candidate for further exploration and implementation in real-world settings. We hope that our work will contribute to the development of more efficient and sustainable transportation systems, and we encourage further research in smart cities for sustainable development.

Future research in the field of intelligent recommendation systems for electric vehicle charging stations holds significant potential for advancing optimization techniques and addressing transportation challenges. Building upon the findings of this study, several avenues for future research can be explored.

Firstly, further investigation is needed to evaluate the scalability and robustness of the MADDPG algorithm in real-world settings. Assessing its performance in larger-scale environments and diverse user scenarios can provide valuable insights into its practical applicability and potential limitations [73].

As the adoption of electric vehicles continues to grow, the demand for charging stations will evolve over time. Therefore, future research can focus on developing adaptive algorithms that can dynamically adjust charging station recommendations based on real-time data, user behavior patterns, and changes in the charging infrastructure.

Another important direction for future research is exploring the integration of charging station recommendation systems with smart grids and renewable energy sources. This integration can enable optimized charging strategies that consider the availability of renewable energy and contribute to a more sustainable and efficient charging infrastructure.

In addition, investigating the privacy and security aspects of the recommendation system is crucial, given the sensitive user data involved. Future research should aim to develop privacy-preserving algorithms and robust security measures to ensure widespread adoption and user trust.

Extending the application of multi-agent reinforcement learning to optimize shared mobility systems, such as ride-sharing or electric vehicle fleet management, presents an intriguing research avenue. This extension can help reduce congestion, improve resource utilization, and enhance the overall efficiency of transportation systems [73–78].

To evaluate the practical implementation of the MADDPG algorithm and its impact on user behavior and adoption of electric vehicles, conducting field experiments and user studies would be valuable. Assessing user acceptance, usability, and user experience in real-world settings can provide insights for system deployment and further improvements.

Finally, future research endeavors should focus on addressing the aforementioned aspects to enhance the efficiency, sustainability, and user-friendliness of electric vehicle charging station recommendation systems. By continuing to explore and refine multi-agent reinforcement learning algorithms, we can pave the way for more intelligent and optimized transportation systems that promote the widespread adoption of electric vehicles and contribute to a greener future.

Author Contributions: Conceptualization, P.S.; research design, P.S.; literature review, P.S. and P.J.; methodology, P.S. and P.J.; algorithms, P.S. and P.J.; software, P.S. and P.J.; validation, P.S. and P.J.; formal analysis, P.S. and P.J.; investigation, P.S. and P.J.; resources, P.S.; data curation, P.J.; writing—original draft preparation, P.S. and P.J.; writing—review and editing, P.S. and P.J.; visualization, P.S.; supervision, P.S.; project administration, P.S.; funding acquisition, P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Suan Dusit University under the Ministry of Higher Education, Science, Research and Innovation, Thailand, grant number FF66-4-006 “Innovative of gastronomy and Agrotourism tourism platform of Suphanburi using identity and culture integrated with the expertise of the university to drive the economic foundation to support next normal in post COVID-19”.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author/s.

Acknowledgments: The authors would like to express their gratitude to Suan Dusit University and Chulalongkorn University for providing research support and the network of researchers in the region where this research was conducted. Moreover, we would like to thank the Tourism Authority of Thailand (TAT) to providing data and information about the tourism service in the research areas.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Suanpang, P.; Jamjuntr, P.; Kaewyong, P.; Niamsorn, C.; Jernsittiparsert, K. An Intelligent Recommendation for Intelligently Accessible Charging Stations: Electronic Vehicle Charging to Support a Sustainable Smart Tourism City. *Sustainability* **2022**, *15*, 455. [CrossRef]
2. Khamis, M.A.H.; Hassanien, A.E.; Salem, A.E.K. Electric Vehicle Charging Infrastructure Optimization: A Comprehensive Review. *IEEE Access* **2020**, *8*, 23676–23692.
3. Sedano, J.; Chira, C.; Villar, J.R.; Ambel, E.M. An Intelligent Route Management System for Electric Vehicle Charging. *Integr. Comput. Aided Eng.* **2013**, *20*, 321–333. [CrossRef]
4. Kim, N.; Kim, J.C.D.; Lee, B. Adaptive Loss Reduction Charging Strategy Considering Variation of Internal Impedance of Lithium-Ion Polymer Batteries in Electric Vehicle Charging Systems. In Proceedings of the 2016 IEEE Applied Power Electronics Conference and Exposition (APEC), Long Beach, CA, USA, 20–24 March 2016; pp. 1273–1279. [CrossRef]
5. Brenna, M.; Foadelli, F.; Leone, C.; Longo, M. Electric Vehicles Charging Technology Review and Optimal Size Estimation. *J. Electr. Eng. Technol.* **2020**, *15*, 2539–2552. [CrossRef]
6. International Energy Agency (IEA). Global EV Outlook 2020: Entering the Decade of Electric Drive? 2020. Available online: <https://www.iea.org/reports/global-ev-outlook-2020> (accessed on 22 December 2023).
7. Statista. Electric Vehicles—Global Market Size 2020 & 2026. Available online: <https://www.statista.com/statistics/1090637/global-electric-vehicle-market-size/> (accessed on 22 December 2023).
8. BNEF. Electric Vehicle Outlook 2021. Available online: <https://about.bnef.com/electric-vehicle-outlook/> (accessed on 22 December 2023).
9. Deloitte. Electric Vehicles: Setting a Course for 2030. Available online: <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/manufacturing/deloitte-uk-electric-vehicle-report.pdf> (accessed on 22 December 2023).
10. Rapiet, R. The Electric Vehicle Revolution Is Coming Faster than Expected. *Forbes*. Available online: <https://www.forbes.com/sites/rapiet/2021/04/29/the-electric-vehicle-revolution-is-coming-faster-than-expected/?sh=5a5a5d524903> (accessed on 29 April 2021).
11. Fang, H.; Liu, Y.; Fang, Y. Reinforcement Learning for Electric Vehicle Charging Station Allocation. In Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing, Dublin, Ireland, 17–19 October 2019; pp. 117–120.

12. Chauhan, K.S.; Chaudhari, S.S.; Patil, P.S. Multi-Agent Based Electric Vehicle Charging Station Allocation: A Review. In Proceedings of the 2020 International Conference on Intelligent Computing and Control Systems, Gangwon, Republic of Korea, 12–15 December 2020; pp. 1580–1584.
13. Huang, Z.; Zhang, J.; Peng, L. Electric Vehicle Charging Station Location Optimization Based on Multi-Agent Deep Reinforcement Learning. *Energies* **2020**, *13*, 564.
14. Wang, Y.; Wang, H.; Zhu, X. An Improved Deep Reinforcement Learning-Based Charging Station Allocation for Electric Vehicles. *Energies* **2021**, *14*, 1287. [[CrossRef](#)]
15. Zhang, J.; Li, T.; Pan, A.; Long, X.; Jiang, L.; Liu, Z.; Zhang, Y. Charging Time and Location Recommendation Strategy Considering Taxi User Satisfaction. In Proceedings of the 2020 Asia Energy and Electrical Engineering Symposium (AEEES), Chengdu, China, 29–31 May 2020; pp. 257–264. [[CrossRef](#)]
16. Huang, C.; Shi, M.; Wang, H.; Jia, Y. Optimal Charging Site Recommendation and Scheduling for Electric Vehicles Considering User Price Sensitivity. In Proceedings of the 2022 IEEE PES Innovative Smart Grid Technologies—Asia (ISGT Asia), Singapore, 1–5 November 2022; pp. 449–453. [[CrossRef](#)]
17. Brandão, D.A.d.L.; Parreiras, T.M.; Pires, I.A.; Filho, B.d.J.C. Extreme Fast Charging Station for Multiple Vehicles with Sinusoidal Currents at the Grid Side. In Proceedings of the 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific), Haining, China, 28–31 October 2022; pp. 1–6. [[CrossRef](#)]
18. Dallinger, F.; Pfeffer, P.; Dorner, W.; Schmid, E. Reinforcement Learning in EV Charging and Energy Trading: A Comprehensive Survey. *Renew. Sustain. Energy Rev.* **2021**, *150*, 111399.
19. Pham, H.V.; Vyas, N.; Khargonekar, P.P. Deep Reinforcement Learning for Electric Vehicle Charging. *IEEE Trans. Smart Grid* **2020**, *11*, 601–612.
20. Haus, M.; Hanappe, P.; Wehenkel, L. Dynamic Electric Vehicle Charging with Reinforcement Learning. *IEEE Trans. Smart Grid* **2019**, *10*, 3928–3937.
21. Farag, M.; Chen, S.; Zhao, X. Real-time Energy Management of Electric Vehicles Using Deep Reinforcement Learning. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 4667–4676.
22. Li, Y.; Wang, X.; Wu, C. A Review of Deep Reinforcement Learning Applications in Smart Transportation. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 4912–4929.
23. Spieser, K.; Bradley, M.; Weinert, J.X.; Greenblatt, J.B. Analysis of Electric Vehicle Adoption and Emissions Reductions in US Cities. *Nat. Sustain.* **2019**, *2*, 1055–1063.
24. Turrentine, T.; Kurani, K. Consumer Considerations in the Transition to Electric Vehicles: A Review of the Research Literature. *Energy Policy* **2019**, *127*, 14–27. [[CrossRef](#)]
25. Zhang, W.; Yu, Y.; Huang, Z.; Li, J. Impact of Subsidy and Tax Incentive on Electric Vehicle Adoption: A Case Study of China. *Energy Policy* **2019**, *128*, 854–864. [[CrossRef](#)]
26. Lin, Z.; Zheng, J.; Zhao, D. How Do Zero-emission Vehicle Mandates Work? Evidence from California. *Energy Policy* **2020**, *137*, 111131. [[CrossRef](#)]
27. Li, X.; Yu, Y.; Wang, X.; Zhang, C. Understanding Consumers’ Purchase Intentions for Electric Vehicles in China: An Extended Theory of Planned Behavior Approach. *Transp. Res. Part D Transp. Environ.* **2020**, *84*, 102325. [[CrossRef](#)]
28. Bruneau, T.; Marlot, G.; Roca, F. Social Influence and Electric Vehicle Adoption: Insights from a French Survey. *Transp. Res. Part A Policy Pract.* **2019**, *121*, 317–330. [[CrossRef](#)]
29. Spieser, K.; Bradley, T.H.; Lutsey, N.; Santini, D.J. Environmental Implications of Electric Vehicle Charging in the United States. *Environ. Res. Lett.* **2019**, *14*, 114027.
30. Turrentine, T.; Kurani, K.S. What Drives U.S. Consumer Demand for Electric Vehicles? *Transp. Res. Part D Transp. Environ.* **2019**, *71*, 296–307. [[CrossRef](#)]
31. Zhang, Y.; Zou, B.; Wang, W.; Zhao, Z.; Ma, C. Evaluating the Effectiveness of EV Subsidy and Tax Incentive Policies: A Case Study of China’s EV Market. *Energy Policy* **2019**, *126*, 393–402. [[CrossRef](#)]
32. Kirschbaum, S.; Kühnrich, M.; Kley, F. The Impact of German Policies on the Adoption of Electric Vehicles. *Transp. Res. Part A Policy Pract.* **2021**, *146*, 153–169. [[CrossRef](#)]
33. Li, H.; Zheng, X.; Gao, X.; Li, X. What Factors Influence Chinese Consumers’ Preferences for Electric Vehicles? *Transp. Res. Part D Transp. Environ.* **2020**, *82*, 102306.
34. Zmud, J.; Welch, E.W.; Lomax, T. Changes in Consumer Attitudes Towards Electric Vehicles and Public Transit During the COVID-19 Pandemic. *Transp. Res. Interdiscip. Perspect.* **2021**, *10*, 100378. [[CrossRef](#)]
35. Bangkok Petroleum. Electric Vehicle Market Outlook for Thailand. 2021. Available online: https://www.reportlinker.com/market-report/Electric-Vehicle/726537/Electric_Vehicle_Charging?term=electric%20vehicle%20charging%20industry&matchtype=b&loc_interest=&loc_physical=1012728&utm_group=standard&utm_term=electric%20vehicle%20charging%20industry&utm_campaign=ppc&utm_source=google_ads&utm_medium=paid_ads&utm_content=transactionnel-4&gad_source=1&gclid=CjwKCAiAt5euBhB9EiwAdkXW00_vZBHvFxXuuxBCMDPvr1jxH1YKIgwAzd6Cmz9b1KwNA9kVs2CoMxoCWfIQAvD_BwE (accessed on 31 December 2023).
36. Bangkok Post. Electric Vehicle Boom in Thailand Gaining Momentum. Available online: <https://www.bangkokpost.com/business/2121555/electric-vehicle-boom-in-thailand-gaining-momentum> (accessed on 31 December 2023).
37. EV Sales. Thailand. Available online: <https://ev-sales.blogspot.com/2021/05/thailand.html> (accessed on 31 May 2021).

38. BloombergNEF. Thailand Electric Vehicle Outlook 2021. Available online: <https://about.bnef.com/electric-vehicle-outlook/> (accessed on 31 December 2023).
39. Greenpeace Thailand. Thailand's Electric Vehicle Revolution: Current Status and Future Opportunities. Available online: <https://mahanakornpartners.com/thailands-strategic-move-electric-vehicle-incentives-unveiled-for-2024-2027/> (accessed on 31 December 2023).
40. Navigant Research. Thailand: The Next EV Capital of Southeast Asia? 2019. Available online: https://www-asia.nissan-cdn.net/content/dam/Nissan/th/news/purchasedecisionresearch/Nissan_whitepaper_TH.pdf (accessed on 30 November 2023).
41. Yang, C.; Zhang, H.; Huang, J.; Han, J. Multi-Agent Reinforcement Learning: A Review. *arXiv* **2021**, arXiv:2103.00598.
42. Lau, T.C.W.; Li, S.; So, R.; Kwok, C.K. A Brief Survey of Multi-Agent Reinforcement Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 2992–3005.
43. Sun, Y.; Xu, Z.; Mao, J.; Qi, Y. Deep Reinforcement Learning for Multi-Agent Systems: A Review of Challenges, Solutions and Applications. *arXiv* **2018**, arXiv:1812.11794.
44. Mnih, V.; Hesse, C.; Kavukcuoglu, K.; Silver, D.; Song, H.F. Traffic Light Control with Deep Reinforcement Learning. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
45. Sun, Y.; Li, M.; Hu, J.; Zhang, W.; Wang, Y. Multi-Agent Reinforcement Learning for Cooperative Transportation with Heterogeneous Robots. *IEEE Trans. Cybern.* **2020**, *50*, 2646–2657.
46. Wu, C.; Kreidieh, A.; Parvate, K.; Vinitzky, E.; Bayen, A.M. Flow: Architecture and Benchmarking for Reinforcement Learning in Traffic Control. *arXiv* **2017**, arXiv:1710.05465.
47. Foerster, J.N.; Assael, Y.M.; de Freitas, N.; Whiteson, S. Learning to Communicate with Deep Multi-Agent Reinforcement Learning. In *Advances in Neural Information Processing Systems (NIPS)*; MIT Press: Cambridge, MA, USA, 2016; pp. 2137–2147.
48. Huang, J.; Wang, X.; Xu, H.; Xie, X. Hybrid Recommendation Algorithm for Electric Vehicle E-Commerce Platform. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 236.
49. Muratori, M.; Alexander, M.; Arent, D.; Bazilian, M.; Cazzola, P.; Dede, E.M.; Farrell, J.; Gearhart, C.; Greene, D.; Jenn, A.; et al. The Rise of Electric Vehicles—2020 Status and Future Expectations. *Prog. Energy* **2021**, *3*, 022002. [CrossRef]
50. Almahmood, R.J.K.; Tekerek, A. Issues and Solutions in Deep Learning-Enabled Recommendation Systems within the E-Commerce Field. *Appl. Sci.* **2022**, *12*, 11256. [CrossRef]
51. An, Y.; Li, L.; Guo, L.; Chen, X. A Feature-Based Recommendation Method for Electric Vehicles. *IEEE Trans. Veh. Technol.* **2020**, *69*, 11963–11974.
52. Chen, J.; Zhao, M.; Liu, Y. Electric Vehicle Recommendation Method Based on Convolutional Neural Network. *J. Renew. Sustain. Energy* **2019**, *11*, 043307.
53. Li, L.; Zhang, Z.; Zhang, S. Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation. *Sci. Program.* **2021**, *11*, 7427409. [CrossRef]
54. Smith, J. Optimizing Urban Mobility: The Role of IoT Sensors in Rayong's Smart Transportation Systems. *J. Urban Technol.* **2020**, *47*, 215–230.
55. Lee, H.; Tan, C. Towards a Greener Future: Rayong's Renewable Energy Projects and Carbon Footprint Reduction. *Environ. Stud. Q.* **2019**, *36*, 489–502.
56. Wong, L. Data Security and Citizen Privacy in Smart Cities: Challenges and Solutions. *Cybersecur. J.* **2018**, *25*, 301–315.
57. Phan, Q.; Nguyen, T. Smart City Initiatives and Job Creation: A Case Study of Rayong, Thailand. *Econ. Dev. Res. Q.* **2020**, *18*, 167–182.
58. Pantip. Thailand Smart City Model. Available online: <https://pantip.com/topic/37302433> (accessed on 31 May 2023).
59. Onion. EV Charger Solution in EEC Rayong. Available online: https://www.facebook.com/Onionsolution/photos/417485863709151/?locale=ms_MY&paipv=0&eav=AfZjvC6UFZTtQTQX-kHrqrub64VHL3PnAXoVdTbCgAqJ65g7gF66KyBuXRrK_MpIVTM&_rdr (accessed on 31 December 2023).
60. Smith, A.; Jones, B. Optimizing Electric Vehicle Charging Station Locations in Rayong, Thailand Smart Cities: A Data-Driven Approach. *Transp. Res. Part C* **2019**, *94*, 267–281. [CrossRef]
61. Patel, S. Fast Charging Technologies and Innovations in Electric Vehicle Charging Stations in Rayong. *Thail. J. Sustain. Transp.* **2020**, *12*, 287–301.
62. Onion-EV Charger Solution, EV Charge in Rayong, Thailand. Available online: https://www.facebook.com/Onionsolution/photos/417485863709151/?locale=ms_MY&paipv=0&eav=AfZb04lTHLT9H9OzfTZ2gsZo8pRvp_CqAmF-DlhM6Pr4oYQjODsLxe5vJ6G937GpbpE&_rdr (accessed on 20 May 2023).
63. Aigner, A.; Buchert, M.; Schmid, E.; Thrän, D. Environmental Impact Assessment of Electric and Internal Combustion Engine Vehicles in the German Energy System. *J. Clean. Prod.* **2021**, *280*, 124320.
64. Allcott, H.; Knittel, C.R. Electrification of the Global Car Fleet: Policy Implications for Advanced Economies. *Brook. Pap. Econ. Act.* **2019**, *2019*, 1–69.
65. Davis, S.J.; Lewis, N.S.; Shaner, M.; Aggarwal, S.; Arent, D.; Azevedo, I.L.; Benson, S.M.; Bradley, T.; Brouwer, J.; Chiang, Y.-M.; et al. Electrifying Transportation in the United States: A Review of Barriers and Opportunities. *Energy Policy* **2020**, *137*, 111068.
66. Elgowainy, A.; Han, J.; Ward, J.; Joseck, F.; Gohlke, D.; Lindauer, A. Life Cycle Assessment of Electric Vehicles: A Review of the Literature. *J. Clean. Prod.* **2021**, *279*, 123696.

67. Malyshev, S.; Thomas, V.M.; Azevedo, I.L. Potential Impacts of Electric Vehicles on the Electricity Grid: A Review of Literature. *Renew. Sustain. Energy Rev.* **2021**, *143*, 110973.
68. Sun, Y.; Wang, L.; Zhang, Y.; Liu, X.; Li, L. Electric Vehicle Charging Infrastructure Planning: A Review of Models and Methods. *Renew. Sustain. Energy Rev.* **2019**, *111*, 413–424.
69. Wang, L.; Sun, Y.; Zhang, Y.; Liu, X.; Li, L. Charging Infrastructure Planning for Electric Vehicles: A Review of Models and Methods. *Appl. Energy* **2020**, *261*, 114425.
70. Hernandez-Leal, P.; Suárez-Figueroa, M.C.; Kaisers, M. Multi-Agent Reinforcement Learning: A Review of the State-of-the-Art. *Knowl. Eng. Rev.* **2019**, *34*, e7.
71. Wang, Q.; Liu, H.; Yang, Q.; Niu, B. Multi-Agent Reinforcement Learning for Intelligent Transportation Systems: A Review of Challenges and Approaches. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 3695–3710.
72. Wei, J.; Chen, H.; Liu, Y.; Cui, Y.; Zhang, C. Decentralized Multi-Agent Reinforcement Learning with Graph Convolutional Networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 19–21 May 2021; Volume 35. No. 3.
73. Tao, Y.; Qiu, J.; Lai, S. Deep Reinforcement Learning Based Bidding Strategy for EVs in Local Energy Market Considering Information Asymmetry. *IEEE Trans. Ind. Inf.* **2022**, *18*, 3831–3842. [[CrossRef](#)]
74. Brenna, M.; Foiadelli, F.; Soccini, A.; Volpi, L. Charging Strategies for Electric Vehicles with Vehicle to Grid Implementation for Photovoltaic Dispatchability. In Proceedings of the 2018 International Conference of Electrical and Electronic Technologies for Automotive, Milan, Italy, 9–11 July 2018; IEEE: Milan, Italy, 2018; pp. 1–6.
75. Suanpang, P.; Pothipassa, P. New Age Robotics: Implications for a Blockchain Integrated Prototype for Education. *Oper. Res. Eng. Sci. Theory Appl.* **2023**, *6*, 668–687.
76. Aldakkhelallah, A.; Alamri, A.S.; Georgiou, S.; Simic, M. Public Perception of the Introduction of Autonomous Vehicles. *World Electr. Veh. J.* **2023**, *14*, 345. [[CrossRef](#)]
77. Sevdari, K.; Calearo, L.; Andersen, P.B.; Marinelli, M. Ancillary services and electric vehicles: An overview from charging clusters and chargers technology perspectives. *Renew. Sustain. Energy Rev.* **2022**, *167*, 112666. Available online: <https://www.sciencedirect.com/science/article/pii/S1364032122005585> (accessed on 30 November 2023). [[CrossRef](#)]
78. Paraskevas, A.; Aletras, D.; Chrysopoulos, A.; Marinopoulos, A.; Doukas, D.I. Optimal Management for EV Charging Stations: A Win–Win Strategy for Different Stakeholders Using Constrained Deep Q-Learning. *Energies* **2022**, *15*, 2323. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.