IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 July 2023, accepted 12 August 2023, date of publication 17 August 2023, date of current version 23 August 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3305966

RESEARCH ARTICLE

Context Aware-Resource Optimality in Electric Vehicle Smart2Charge Application: A Deep Reinforcement Learning-Based Approach

MUDDSAIR SHARIF[1,2], GERO LÜCKEMEYER[2], AND HUSEYIN SEKER[1]
[1]Faculty of Computing, Engineering and the Built Environment, Birmingham City University, B5 5JU Birmingham, U.K.
[2]University of Applied Sciences, 70174 Stuttgart, Germany

Corresponding author: Muddsair Sharif (muddsair.sharif@mail.bcu.ac.uk)

ABSTRACT Electric vehicle (EV) adoption is expanding, posing new issues for grid operators, fleet operators, charging station operators, and EV owners. The challenge is to devise an efficient and cost-effective strategy for managing EV charging that takes into account the demands and objectives of all parties. This study offers a context-aware EV smart charging system based on deep reinforcement learning (DRL) that takes into account all participants' requirements and objectives. The DRL-based system adjusts to changing contexts such as time of day, location, and weather to optimize charging decisions within an instantaneous fashion by balancing the trade-offs among charging cost, grid strain reduction, fleet operator preferences, and energy efficiency of charging station maintainer while providing EV owners with a convenient and cost-effective charging experience for its ability to handle sequential decision-making, capture complex patterns in data, and adapt to changing contexts. The proposed system's performance has been evaluated using simulations and compared with existing solutions. The results demonstrate that the proposed system is capable of balancing the trade-offs between different objectives and providing an energy-efficient solution which is approximately 15% better than traditional approach, and about 10% more cost-effective charging experience for EV owners while reducing grid strain by 20% and CO2 emissions by 10% as a result of using a natural energy source. The proposed system has then resulted in achieving the needs for efficient and optimised resource scheduling of fleet operators and charging station maintainers.

INDEX TERMS Context-aware, deep reinforcement learning, electric vehicle, resource optimisation.

I. INTRODUCTION

The growing popularity of electric vehicles (EVs) unveils novel challenges for industry stakeholders which are the main drivers of the industry for example grid operators, fleet operators, charging station operators, and EV owners. One of the primary challenges involves figuring out trade-offs between several objectives, such as lowering the cost of charging for EV users, reducing demand on the power grid, optimizing fleet management, and improving energy efficiency at charging stations. The suggested approach should be a context-aware EV smart charging system that is capable of adapting to changing contexts and optimizing choices regarding charging in an immediate fashion to negotiate compromises between charging cost, grid strain reduction, fleet operator preference, and energy efficiency of charging station maintainers while providing EV owners with a convenient and cost-effective charging experience.

For the last few decades, researchers have been working on EV smart charging. Initially, the emphasis was to establish fundamental charging infrastructure and standards [1], [2], [3]. In recent years, the focus has shifted to designing advanced charging systems that are capable of optimal trade-offs between several objectives [4], [5], [6], [7], [8], [9], [10]. The increasing adoption of electric vehicles (EVs) is a major trend in the transportation

The associate editor coordinating the review of this manuscript and approving it for publication was M. Anwar Hossain.

sector, driven by concerns about energy security, climate change, and air pollution [12]. Though, the integration of EVs into the electricity grid raises additional challenges for grid operators, fleet operators, charging station operators, and EV owners. The key problem is balancing trade-offs between many objectives, such as lowering EV charging costs, reducing the load on the power grid, optimizing the management of fleets, and promoting energy efficiency at charging stations [13], [14]. To present such issues, researchers have proposed numerous solutions, which involve including time-of-use pricing schemes [15], [16], [17], [18], dynamic load management [19], [20], and smart charging algorithms such as A Stochastic Game Approach [21], Vehicle-to-Grid (V2G) Optimization [22], Pareto Optimal solution in Multi-Objective Optimization [23], Real-Time Energy Management Systems [24], Blockchain-based Charging Systems [25], [26], [27]. Ultimately, these aggregated approaches mainly lack the way to engage with changing parameters such as time of day, location, weather, and other factors that may have significant effects on EV charging patterns and electrical infrastructure obligations. Other than this, such outcomes are usually reliant on centralized decision-making, which can be rigid and incapable of satisfying the ever-changing needs of various stakeholders.

In line with the availability of data being generated by EV vehicles and its stakeholders, the application of complex algorithms and technologies, such as optimization, control, and machine learning, is at the forefront of EV smart charging [11]. Reinforcement learning (RL) is a suitable approach for this research study due to several reasons. Firstly, RL is a type of machine learning that focuses on making sequential decisions in an environment to maximize a cumulative reward. In the context of the EV charging infrastructure, the decision-making process involves dynamically assigning charging resources to meet the demands of EV end-users while considering the constraints and objectives of other stakeholders. Secondly, the application of deep reinforcement learning (DRL) allows for the integration of deep neural networks, enabling the model to capture complex patterns and representations from large-scale data. In the context of resource allocation in an EV charging infrastructure, DRL can learn from historical charging data, user preferences, grid conditions, and fleet operator requirements to make informed decisions on resource optimization. Furthermore, the context-awareness aspect of the research topic aligns well with the capabilities of reinforcement learning. RL algorithms can adapt to changing contexts, such as variations in user demand, charging station availability, and grid conditions. By continuously learning and updating its decision-making policies, RL can optimize resource allocation based on the current context and contribute to increased efficiency and user satisfaction in the charging process. Deep reinforcement learning (DRL) is currently gaining popularity as a solution to the problem of EV smart charging. By maximizing a reward signal, DRL is a strong machine-learning technique that can learn to make judgments in complicated, dynamic contexts. It has been revealed to be advantageous for a variety of applications, including gaming, robotics, and energy management. Particularly, DRL-based EV smart charging systems can adapt to changing settings and realistically optimize charging decisions to balance trade-offs between several objectives that have been presented. DRL-based EV smart charging systems, for example, have been proposed to optimize charging schedules to minimize charging costs and reduce grid strain while taking the context of the charging session into account [28], [29], [30]. Likewise, DRL-based EV smart charging systems have been proposed to optimize charging schedules across multiple EV fleets by considering the preferences and objectives of different fleet operators into account [31], [32]. Using this approach may promote an additional flexible and decentralized decision-making process that can more effectively meet the varying requirements of various stakeholders.

In a nutshell, identifying an efficient and cost-effective way to manage EV charging that takes into account the demands and objectives of all participants is a complicated and difficult process. According to a recent investigation, DRL-based EV smart charging systems can be an effective solution that can adapt to changing settings and optimize charging decisions realistically to balance trade-offs between multiple objectives. In the outcomes from above, more research in this area is needed to improve these technology platforms by integrating more context-aware and decision-making skills that take into account the demands of all users.

II. LITERATURE REVIEW

Despite the significant progress that has been made in EV smart charging research in recent years, there is still a significant gap in the literature when it comes to addressing the trade-offs between different objectives with respect to all stakeholders which are involved in this research. More specifically, still there is a lack of research that considers the needs and preferences of all stakeholders, such as EV owners, grid operators, fleet operators, and charging station maintainers [33], [34]. Presently, the research focus mainly on addressing the needs of a single stakeholder, such as minimizing charging costs for EV owners or reducing grid strain for grid operators, without fully considering the impact on other stakeholders and their concerns. So, this can lead us to sub-optimal solutions that do not fully optimize the trade-offs between different objectives.

Little research has been done on the integration of context-awareness in EV smart charging systems from the perspective of multiple stakeholders. For example, a survey on the current state of the art of context-aware EV charging systems focused mainly on one stakeholder perspective, such as the EV end-users [35]. There are also limited studies on the use of deep reinforcement learning for EV smart charging, particularly in a multi-stakeholder context. A multi-agent reinforcement learning framework has been proposed to coordinate EV charging, but does not take into account the context-awareness or multiple objectives of different stakeholders [36]. It is reported that recent EV smart charging systems don't seem to have been designed to adapt to changing contexts, which limits their effectiveness in real-world situations. For example, in "A review of electric vehicle charging control strategies: From off-line to real-time approaches" [37], the authors have reviewed several EV charging control strategies, but they do not seem to discuss how to adapt the charging system to a changing context. There also seems to be a lack of research on how to balance the different objectives of the various stakeholders involved in EV smart charging, such as EV end-users, grid operators, fleet operators, and charging station maintainers. For example, in "Optimizing Electric Vehicle Charging Schedules: A Review" [38], authors have reviewed several EV charging optimization methods, but they do not discuss how to balance the different objectives of different stakeholders. Additionally, the literature review on similar models in other domain is also considered to evaluate its contribution to the advancement of various fields, including X-ray detection technology [39], predictive maintenance planning [40], and battery prognostics [41]. They demonstrate the effectiveness of data-driven approaches and highlight the potential for improved accuracy and cost reduction in these areas of research.

Earlier studies have mainly focused on one objective or one stakeholder and lack a comprehensive approach to quantify the different combinations of the factors and the trade-offs between various objectives for the involvement of multiple stakeholders [44], [45], [46], [47], [48], [49]. As seen from the literature, the EV smart charging problem is a complex and dynamic one, involving multiple agents (EVs, grid, charging stations) with different objectives and constraints as seen in Tables 1, 2, 3, and 4. Therefore, there is a need for more research on how to effectively coordinate and optimize the decisions of multiple stakeholders in a decentralized manner. To effectively coordinate and optimize the decisions of these agents, it is necessary to develop methods that can handle the uncertainty and non-stationary of the problem. Moreover, to enable a better decision-making process, there is a necessity for more research that brings outcomes with further enhancements using combination, communication, and coordination among different stakeholders. The tables 1, 2, 3, and 4 include the description of several metrics that can be used to evaluate the resource optimality of different solutions for EV charging. The metrics are described along with their perspective of each stakeholder. By considering the perspectives of all stakeholders, a more comprehensive evaluation of the solutions can be made, taking into account the specific requirements and constraints of the EV charging environment.

Given the importance of context-aware systems and multiple-factors, the theoretical framework for this research is built upon three main components, namely, context-aware systems, resource optimality, and deep reinforcement learning. The main aim of this research is therefore to develop a practical and smart context-aware EV smart charging system that can optimize the trade-offs between different objectives for all stakeholders involved. The system should be able to adapt to the changing context and make decisions accordingly. The concept of resource optimality is used to guide the research in finding the optimal charging schedules

TABLE 1. EV end user metrics.

Metric	Description	EV end user
Total cost of charging	The overall cost incurred for charging EVs	Wants to minimize the cost of charging
Total time taken to reach the charging station	The time taken for the EV to reach the charging station and complete the charging process	Wants to minimize the time taken to charge the EV
Total CO2 emissions	The total emissions produced from the charging process	Wants to minimize emissions for environmental and personal reasons
Total energy consumption	The overall energy consumed for charging the EVs	Wants to minimize energy consumption

TABLE 2. Grid operator metrics.

Metric	Description	Grid operator
Total cost of charging	The overall cost incurred for charging EVs	Concerned with the cost of electricity, network upgrades, and energy management systems
Total time taken to reach the charging station	The time taken for the EV to reach the charging station and complete the charging process	Concerned with network upgrades and grid stability
Total CO2 emissions	The total emissions produced from the charging process	Concerned with reducing emissions and meeting regulatory requirements
Total energy consumption	The overall energy consumed for charging the EVs	Concerned with energy demand and grid stability

TABLE 3. Fleet operator metrics.

Metric	Description	Fleet operator
Total cost of charging	The overall cost incurred for charging EVs	Wants to minimize the cost of charging for their fleet of EVs
Total time taken to reach the charging station	The time taken for the EV to reach the charging station and complete the charging process	Wants to minimize the time taken to charge their fleet of EVs
Total CO2 emissions	The total emissions produced from the charging process	Wants to minimize emissions for environmental and regulatory reasons
Total energy consumption	The overall energy consumed for charging the EVs	Wants to minimize energy consumption and costs

TABLE 4. Charging station maintainer objectives.

Metric	Description	Charging station metrics
Total cost of charging	The overall cost incurred for charging EVs	Wants to maximize revenue from charging services
Total time taken to reach the charging station	The time taken for the EV to reach the charging station and complete the charging process	Wants to minimize downtime and increase usage of charging stations
Total CO2 emissions	The total emissions produced from the charging process	Wants to minimize emissions for environmental and regulatory reasons
Total energy consumption	The overall energy consumed for charging the EVs	Wants to minimize energy consumption and costs

for EVs that take into account the different objectives and constraints of the stakeholders as seen in Table 4. This concept is based on the idea that resources should be used in the most optimised way possible. For this purpose, Deep reinforcement learning using Q-networks is used as the main method to optimize the charging schedules.

In summary, the three main components are considered to guide the research in developing a context-aware EV smart charging system that can optimize the trade-offs between different objectives for all stakeholders involved as described below.

III. THE STATE-OF-THE-ART FRAMEWORK

The overall objective of this research is to present a framework that facilitates non-gasoline vendors such as EV end-user, grid-operator, fleet operators, charging-station to receive optimum outcomes with the help of provided key factors i.e. fleet-booking, charging station, charging station maintainer, charging point demand, location, charging station availability, using machine learning technique. Additionally, in this framework, a deep reinforcement learning (DRL) technique which works as State, Reward, and Environment pattern is used to optimise the process of decision-making for the most optimum EV charging model.

Artificial intelligence is a broad category of computer science that aids in the development of smart processes and computers capable of doing activities that would typically need human intelligence. Furthermore, these machines are capable of resolving issues, enhancing decisions, as well as performing tasks that were previously undertaken by human beings [28], [29]. *Machine Learning*(ML) is a subpart of Artificial Intelligence which is a data-driven and trained using the data. Additionally, ML-based models are capable of discovering their own findings from data being provided to machine without any explicit human involvement. Moreover, ML uses numerous algorithmic procedures that consume data to figure-out how to progress, make predictions, and explain data. These models can be trained via *supervised,*

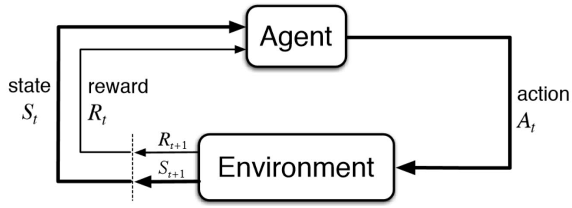

FIGURE 1. Reinforcement learning model.

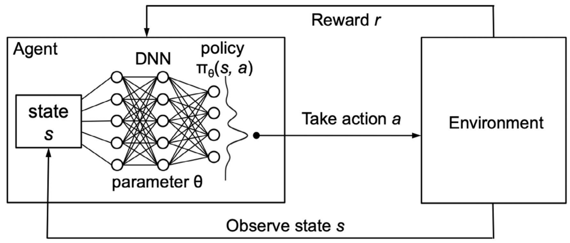

FIGURE 2. Reinforcement learning with policy presented via DNN source: DeepRM-HotNets.

unsupervised, semi-supervised and reinforcement learning based strategy. Among all these approaches in ML, *Deep Learning* is the subsection of ML that consists of a set of algorithms, based on the concept of artificial neural networks that have a self-learning capability over the multi-layer neural networks with the help of data that allows them to accomplish tasks (e.g., image recognition, speech recognition) [30].

Reinforcement Learning (RL) is the science of decision making which also lies in the machine learning section where computer program act as an intelligent agent that interacts with the environment and learns to act within that. For example, how a Robotic Agent learns the movement of its foot to play football is an example of reinforcement learning [31]. The basic RL model involves an agent interacting with an environment to learn an optimal policy for taking actions in different states. At each time step t, the agent observes the current state St of the environment and selects an action Ai based on its policy. The environment transitions to a new state St+1, and the agent receives a reward Rt for the action taken in state St. The goal of the agent is to learn an optimal policy that maximizes the expected cumulative reward over time. The value of a state-action pair (St, Ai) is defined as the expected cumulative reward starting from state St and taking action Ai, and following the optimal policy thereafter. This value is denoted as Q(St, Ai).

In *figure* 2, the agent is the learner and the decision-maker. The environment is the point of interaction between the agent and its objective. The environment reacts to the agent's actions by presenting it with fresh scenarios. Rewards are another result of the environment; their unique numerical values are what the agent attempts to maximize over time by selecting activities. Formalizing the agent's purpose or aim in terms of a unique signal termed the reward that travels from the environment to the agent. The reward is a straightforward number $R_t \epsilon$ R at each time step. The agent's informal objective is to maximize the overall

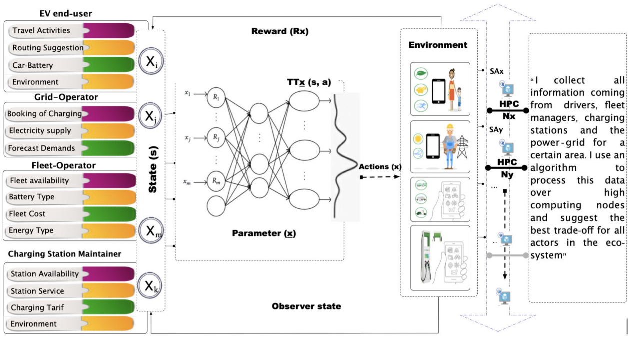

FIGURE 3. The state of the art architecture.

reward it accumulates. This involves optimizing long-term cumulative reward rather than immediate reward. The return is a function of the agent's desire to maximize future benefits (in expected value). It has numerous distinct definitions based on the nature of the activity and whether or not delayed reward is desired. The unaccounted formulation is suitable for continuous tasks in which the operative such as agent-environment interaction certainly breakdowns into episodes; the discounted formulation is suitable for continuous tasks in which the interaction does not naturally breaks into episodes but lasts indefinitely. We attempt to describe the returns for the two types of jobs in such a way that a single set of equations can be applied to both episodic and ongoing scenarios. By solving the Bellman optimal equations, which are certain consistency conditions that the optimal value functions must satisfy, it is fairly straightforward to build an optimum policy from the optimal value functions.

We reviewed a number of research initiatives completed by various organizations that work independently adequately but lead to the wastage of resource due to a lack of collaboration in-between. In this context, let's consider an example in *figure* 3 where the first stakeholder, EV end-user is interested to find an appropriate charging point on their way from 'location X' towards 'location Y' with minimal time to charge the vehicle as well as lower charging cost. The second stakeholder, the grid operator, is responsible for generating electricity which is required to meet the requirements of inhabitants without knowing their required demands in a certain region. The charging station maintainer, as the third stakeholder, plays a crucial role in ensuring that the charging point is always ready for use before anyone reserves it. However, it's worth noting that this situation rarely happens without prior knowledge of how the charging station operates. On the other hand, the fleet operator stakeholder takes on the responsibility of determining the optimal number of vehicles available for booking. Their goal is to meet the recommended requirements and preferences of the customers, ensuring a smooth and satisfactory experience. Moreover, all of these stakeholder work independently without knowing the on-demand requirements from other vendors realistically. However, the proposed state-of-the-art methodology in this paper overcomes these issues and incorporates realistically with preferred demands from different vendors, and uses the resources more efficiently.

The following part explains thoroughly the manner in which the suggested architecture works. We demonstrate how the algorithm makes use of contextual data to determine the win-win requirement for each stakeholder. We define four different sets of stakeholders in the efficient transportation eco-system including **EV end-users**, **Grid-Operator**, **Fleet-Operator**, and **Charging Stations Maintainer**.

1) **EV end-users:** EV end-user should share their travel activity such as their movement plan (e.g. start-location and end-location of their journey). Moreover, the EV end-user will get the routing suggestions from which needs choosing one of the paths. The technical specifications of the vehicle (e.g. a battery type) are also decided by the EV end-user. The algorithm produces the best potential route options based on these inputs and the provided key performance indicators criteria, including pricing and the accessibility of charging stations. The EV end-user might select the routing option that outfits according to his environment as well as the best from the algorithm recommendations.
2) **Grid-Operator:** it can provide information on feeder and transformer loads, such as charging and electric supply booking, and so contribute to and have an impact on the charging station's grid-friendly use. Usually, the grid-operator can approximate the feeder and transformer loading for up to the upcoming twenty-four hours utilizing cutting-edge distribution network modeling technologies.
3) **Fleet-Operator:** The primary role of the fleet operator is to keep eye on available fleet for booking and what type of energy type is used in it (e.g. hydrogen, gas, benzine, electric). Furthermore, battery usage data, such as discharge rate, will be available to the fleet manager, who may use it to troubleshoot issues and organize repairs. In addition to this, what type of fleet is requested and its cost as well as meeting the required criteria of load according to customer requirements.
4) **Charging Stations Maintainer:** It will keep the charging station operational in order to meet end-user needs and deliver suitable services in the event of an unintentional failure. When the cost of renewable energy falls, the proprietor of the charging station may send alerts to customers to get the charging done with a minimal cost. Furthermore, before visiting the charging station, the end-user can book a charging station for their particular fleet.

The information received from individual stakeholders is represented in (X_i, X_j, X_k, ...,X_m) with associated initial rewards respectively. These parameters represent the state(s) that are further given to the model as inputs, as shown in the bottom-left-hand side of Figure 3. These data sets, on the other hand, are examined using a cutting-edge approach based on deep reinforcement learning method where the computer learns weights of DRL parameters from sets of

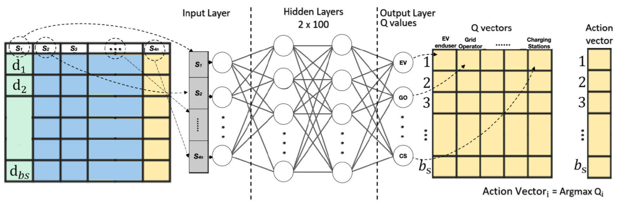

FIGURE 4. **DQN model prediction using states and deep neural networks, the outputs are Q-values, and actions are computed based on Argmax Q_i for the current State.**

inputs, recommended domains, and their restrictions. When the anticipated barrier is attained, the precise result is produced in Figure 3 on the right, where different output is displayed according to a specific stakeholder such as ''EV end-user: will get the best schedule and routing option based on the personal needs of the car's battery and the environment'' is displayed, Grid-Operator: will obtain expected power demands for a specific region based on charging station reservations, as well as eliminate electric fluctuation, and so on.'' Otherwise, the system constantly learns from its surroundings by interpolating weights, and so on.

$$\sum_{t=0}^{t=\infty} \gamma^{t} r(x(t), a(t)) \tag{1}$$

There is a demonstration of an objective function for reinforcement learning, which describes its aim. Here we express our computation reward function r applies-over t for time step calculations. Furthermore, we may acquire all of the rewards by running with the aid of this goal function. At a particular time step, state is represented by x, and action taken at that state is represented by a. A reward r represents the computed outcome over the state x and the action a. In addition, each job aims to maximize a discounted sum of its rewards γ for particular time-steps [32].

As illustrated in Fig.4, the DQN agent receives input states from four different stakeholders: EV end users, Grid Operators, Charging Station Maintainers, and Fleet Operators. These input states consist of 16 features represented in X_1 to X_{16} in Figure 3. The DQN agent uses a batch size of d_1 until d_{bs} for each input-feature state, which is represented in S_1, S_2 until S_n in Figure 4. For each state, the DQN agent retrieves a batch of records from memory, which can vary in size from 50 to 200, and feeds them into a batch table. The DNN used by the DQN agent has sixteen input features and two hidden layers with hundreds of interconnected nodes. The DNN has four output states, represented by a linear number starting from 1 until b_s number, corresponding to the number of participants.

The four output states represent the Q-values of each action for each participant, which are used to determine the optimal action for each participant in the given state. The action vector is also represented in a similar manner in figure 5. An action is assumed to be the choice made by the agent after processing the environment during a specified time window. The network agent provides a list of actions as an action vector by combining the neural network input with its features. The resulting Q-values are subsequently employed to predict if the required information was effectively gathered. It feeds the current DQN the state vector with the batch size. The agent then evaluates the current DQN output based on threshold rates as Q-values to calculate the Q-threshold value for classifying stakeholders. Overall, the DQN agent uses the input states from stakeholders to learn the optimal policy for coordinating the charging of electric vehicles in a distributed manner, which is presented in detail in the forthcoming methodology section and explained with an example.

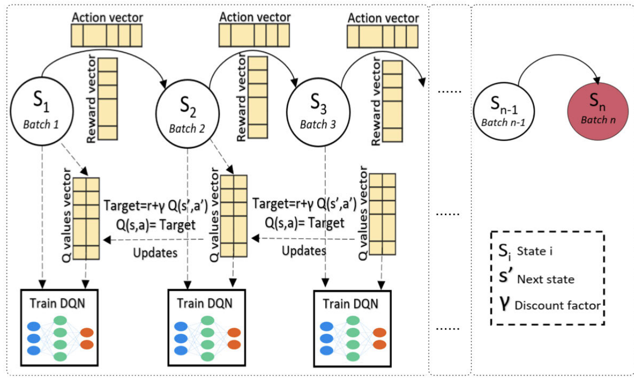

FIGURE 5. **Based on training and prediction of the current and subsequent states, the DQN agent state transition Markov diagram illustrates the learning process.**

The essential functionality has been developed as a software package that enables customers across various sectors to interact using our platform. This interaction is made simpler by middleware developed as a service component. With the help of this modification, we can demonstrate the model's utility at the urban scale level under high computing demand, high-dimensional data, and model scalability. As an example, suppose *"Stakeholder 'X' want to collect all data originating from electric-vehicle end-users, fleet management, charging stations, and the power grid for a certain area. In that case, we can utilize an algorithm to analyze this data across several computer nodes and offer the best trade-off for all potential participants in EV charging ecosystem."*

A. TEMPLATE BASE AUTOMATION USING CI/CD

Continuous Integration/Continuous Deployment (CI/CD) is a software development practice that involves continuously integrating and testing code changes, and then deploying those changes automatically to the production of its software product. CI/CD can be particularly helpful in running AI solutions on High Performance Computing (HPC) systems for several reasons. For example, it can enable faster development and deployment of AI solutions [42]. By automating the process of integrating and testing code changes, CI/CD can help reduce the time it takes to develop and deploy AI solutions, especially when those solutions involve complex algorithms and large amounts of data. Another example is that it can improve the scalability, reliability and stability of AI solutions. By continuously testing code changes and automatically deploying only those that pass all the tests, CI/CD can help ensure that AI solutions are robust and work as intended. This can be especially important when running AI solutions on HPC systems, which may be used to solve complex and mission-critical problems. Overall, CI/CD can be a valuable tool for enabling the development and deployment of AI solutions on HPC systems, helping to make the process faster, more reliable, and more efficient.

IV. METHODOLOGY

The following section's primary purpose is to provide an in-depth explanation of the research approach used to build and assess the suggested deep reinforcement learning algorithm for the optimality of the electric vehicle smart2charge application. This comprises data collection, preliminary processing and purification, data normalization, and the incorporation of critical information from all stakeholders involved in the electric vehicle (EV) charging process.

A. DATA COLLECTION

Sources: The data was collected from various sources such as real-world electric vehicle (EV) charging data, power grid load data and other relevant datasets from the stakeholders involved in the EV charging process, such as the EV end-user, grid operator, fleet operator, and charging station operator. Moreover, relevant part of the data is selected and anonymised. *Data Pre-processing*: To ensure the quality and consistency of the data, various pre-processing steps were taken, such as removing irrelevant or duplicate data, normalizing the data to a consistent format, and integrating the data from different sources [43].

Cleaning: The suggested gathering of data and facts obtained from various sources was extensively cleansed to ensure that it was precise and trustworthy for use in training the deep reinforcement learning algorithm. This implies omitting any missing or inconsistent values and assuring that the data was in the proper format for use in training the algorithm. *Data Normalization*: The data need to be normalized to a consistent format so that they can utilized throughout the training and evaluation operations. This process helps transform the information into a standard format, such as altering the facts into values that are numerical, standardizing a range of values, and converting the data into a form consistent with the sophisticated methodology. *Locations*: Latitude and longitude points have been added as an additional column named ''locations'' to the dataset, which consists of the route direction's geographical coordinates. This data is utilized to combine the charging station dataset to compute the distance from the present position toward the station for charging. *Energy source*: The dataset now includes another parameter called ''energy source'' that specifies the type of energy used by each charging station operator during the charging of the vehicle. All of the procedures mentioned above have been finished in order to ensure that the input information is in the best possible condition for training and evaluating the deep reinforcement learning algorithm.

```
Algorithm 1: Deep Q-learning agent training based on Smart2ChargeApp environment
Data:  Smart2Charge Dataset                                /* Environment */
Data:  DQL Parameters                                 /* Agent parameters */

begin
                          /* Pre-processing and Parameters Initialization */
    Normalize Smart2ChargeDS
    Initialize parameters as mention in  figure 3
    Batch_Size <— 50 ( varies between 50 to 200 )
    State <—  fetch ( Smart2ChargeDS, Batch_size )
    Create_model (States, layers_hidden, ReLU, layers_output )
      /* DQL agent learning episodes and iterations */
    foreach epoch ∈ num-episode do
        Reset(states)
        Create(q_val_List[size=bs,Action_size])
        foreach T ∈ num_iteration do
            Initialize parameters
             /* With probability of ϵ:  */
            AV_i ← Create_random(Action_space) ∀i ∈ bs
            ϵ ← ϵ ∗ decoy_rate
             /* With probability of 1 − ϵ:  */
            QV_i ← model.predict(current-state)
            AV_i = Argmax(QV_i) ∀i ∈ bs
             /* Compute rewards */
            RV_i ← Compute_reward(AV,labels) ∀i ∈ bs
             /* Agent's learning improvement */
            Q'← model.predict(state')
            QT_i←RV_i+γ ∗ Q[state′, action′]
            Model.train(state,QT_i)
            Compute_loss(QV_i,QT_i)
            State←State′
        end
    end
end
```

FIGURE 6. Algorithm 1: Deep Q-learning agent training based on Smart2ChargeApp environment.

B. ALGORITHMS IMPLEMENTATION

This section demonstrates the high-level blueprint of how the strategy has been implemented using deep reinforcement learning. The algorithm is a Deep Q-learning (DQL) agent training algorithm for the Smart2ChargeApp environment. It starts by taking Smart2ChargeDS data as an input, pre-processing it, and initializing the DQL parameters. The DQL agent's neural network model is then created with hidden layers, ReLU activation function, and output layers. The algorithm trains the DQL agent with a number of episodes and iterations. For each episode, the states are reset, and the algorithm loops over the iterations. The states can represent the current state of the EV battery level, the location of the EV, the cost of charging at the current location, the distance to the nearest charging station, etc. For each iteration, the action values are set randomly with probability epsilon and obtained by predicting the actual state with probability 1-epsilon. The actions can represent the decisions made by the EV end-user, such as choosing to charge at the current location or to drive to a different location. The reward is computed for each action taken by the agent, and the Q-value for the next state is predicted by the model.

The rewards can represent the cost of charging the EV, the time taken to reach the next charging station. The rewards should be designed for the agent to make decisions that result in lower charging costs and shorter charging times. The target Q-value is then computed, and the model is trained on the current state and target Q-value. The loss is computed, and

FIGURE 7. Algorithm 1: Deep Q-learning agent computational assessment on Smart2ChargeApp environment.

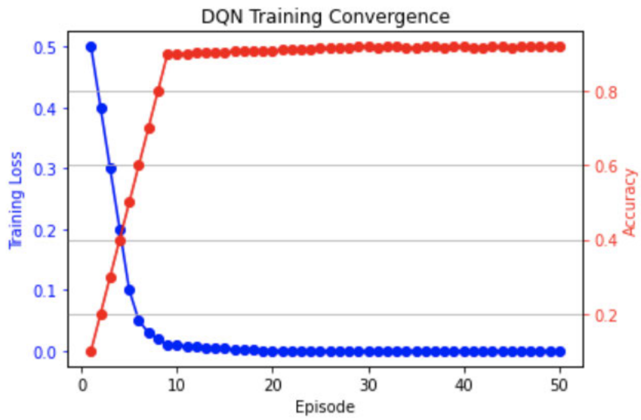

FIGURE 8. Algorithm 1: DQN Accuracy and Convergence.

the state is updated to the next state until the iteration is complete. The algorithm repeats this process for each episode until the training is complete. The agent's computational performance is then evaluated by comparing it with the desired results and by monitoring the performance metrics such as loss/reward, Discount factor and computational time as depicted in figure 7.

The computational graph showcases the relationship between discount factors (γ), loss and reward values, and computational time in the context of the DQN learning process. The loss and reward values illustrate the performance of the DQN model across different discount factors. As the discount factor increases, the loss decreases, indicating improved convergence and learning. Similarly, the reward increases with higher discount factors, suggesting more successful and rewarding agent behavior. The computational time graph demonstrates the time required for the DQN learning process as a function of the number of episodes. The computational time appears to be relatively consistent across different discount factors and increases gradually with a higher number of episodes. This indicates that the DQN model's computational complexity is primarily influenced by the number of episodes rather than the discount factor. At the end, the choice of discount factor significantly affects the learning process's effectiveness, as reflected in the loss and reward values. However, the computational time remains relatively stable across different discount factors, with the number of episodes playing a more significant role. These insights can inform decision-making when configuring and optimizing the DQN learning process based on desired trade-offs between learning performance and computational efficiency.

Rev3.5 In this figure 8, the training_loss values represent the training loss for each episode during the DQN training process, and the accuracy values inside the graph represent the accuracy achieved for each episode. The plot displays two y-axes, one for training loss represented with the color blue and the other color representing accuracy red. The training loss is plotted as a blue line with markers, and the accuracy is plotted as a red line with markers.

C. SIMULATION SETUP

In this simulation environment, we need to consider electric vehicle charging with an associated route from Stuttgart to Germany with an approximate distance 129 km.

1) Parameters: The simulation environment includes the following parameters:
   a) Number of EVs: 3 sample EVs are considered for the simulation.
   b) Charging stations: The dataset includes information about the charging stations along the route.
   c) Charging rate of the EVs: The rate of charging the EVs is considered as an input parameter.
   d) Cost of electricity: The cost of electricity at each charging station is considered as an input parameter.
   e) Route direction: The route direction from Stuttgart to Heidelberg is considered as an input parameter.
   f) Environmental factors: the factors such as weather conditions and wind direction/speed are considered as input parameters for the simulation.
   g) Energy source: this parameter gives information about the source of the energy such as coals, gas, solar, and wind.
2) Constraints: The simulation need to be considered the following constraints of the EV charging scenario;
   a) The number of EVs and charging stations simulated should not exceed the actual number of EVs and charging stations present in the scenario.
   b) The charging rate of the EVs should not exceed the maximum charging rate of the EVs.
   c) The basic price calculated at each of the charging stations must lie within the limit of the sum of the actual cost of all charging stations.
   d) The simulation should consider the impact of other environmental factors such as weather, wind on the EV charging process.

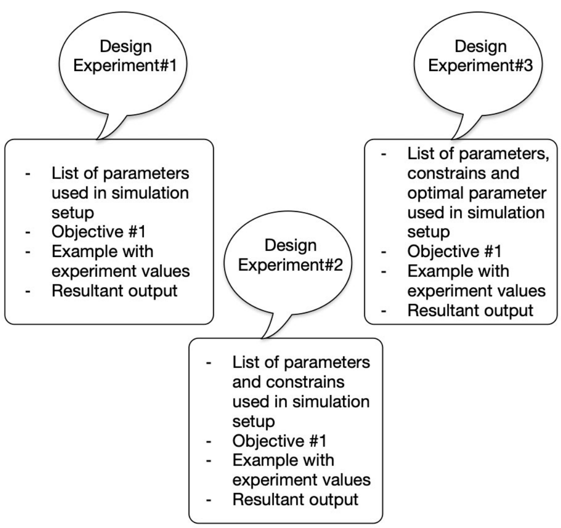

FIGURE 9. Experiment design using simulation setup.

3) Optimal Parameters: The simulation need to be also considered the following optimal parameters for energy sources:
   a) The optimal charging rate of the EVs to ensure maximum efficiency and minimize the cost of electricity.
   b) The optimal route direction to reach the charging station with the minimum cost of electricity.
   c) The optimal choice of charging station based on the cost of electricity, distance to the charging station, and availability of renewable energy sources.
   d) The simulation should also consider the impact of environmental factors such as weather and wind on the optimal parameters for energy sources.

D. EXPERIMENT DESIGN

The basic purpose of such experiments is to develop strategies for optimizing the use of electric car resources, such as reducing charging time and cost by selecting the nearest and most cost-effective charging station, increasing the use of renewable energy-based sources by selecting charging stations powered by renewable energy such as photovoltaic (PV) or wind rather than coal or oil. Moreover, the usage of these energy sources produces a direct environmental impact by reducing CO2 emissions and encourages the use of eco-friendly energy sources by EV users. The proposed experiment design is broken down into three main steps: Experiment Design One, Experiment Design Two, and Experiment Design Three, as shown in Figure 9.

1) Objective:
   a) To minimize the cost of charging for the EV end-user by selecting the nearest and most cost-effective charging station.

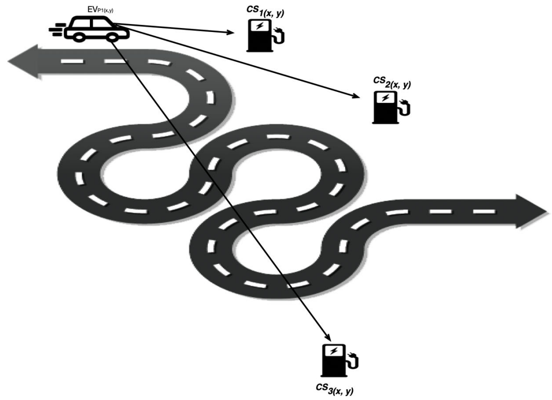

FIGURE 10. Simulation of EV without constraints and optional parameters.

   b) To maximize the use of renewable energy sources by selecting charging stations powered by renewable energy.
   c) To minimize the time taken to reach the charging station and minimize the impact of traffic congestion, weather conditions, and wind direction on the charging process.
   d) To minimize the environmental impact by reducing CO2 emissions.

E. EVALUATION METRICS

The basic idea behind the evaluation metrics is to evaluate the developed strategy to get the assurance of the used resources in electric vehicle charging that meets the proposed objectives as proposed by all of the participants. There are several common evaluation metrics that are used in this context such as energy efficiency, charging time, charging cost, battery life, grid impact, environment impact. Meanwhile, the main experiments to be presented in this paper will be built upon the evaluation of the cost of charging for EV owners.

1) Experiment Design 1: Suppose that there are three charging stations available to the EV end-user, A, B, and C. Station A is powered by renewable energy and charges $0.15 per kilowatt-hour, station B is powered by conventional energy and charges $0.20 per kilowatt-hour, and station C is also powered by conventional energy but charges $0.10 per kilowatt-hour.
   If the EV has a range of 100 miles and requires 20 kilowatt-hours of energy to fully charge, the cost of charging at each station can be calculated as follows:
   - Station A: 20 kilowatt-hours * $0.15/kilowatt-hour = $3.00
   - Station B: 20 kilowatt-hours * $0.20/kilowatt-hour = $4.00
   - Station C: 20 kilowatt-hours * $0.10/kilowatt-hour = $2.00

   According to the aforementioned inputs, the above calculation determines that charging station C offers the lowest calculated rates per kilowatt-hour. Accordingly, it will be the most suitable option for the EV end-user

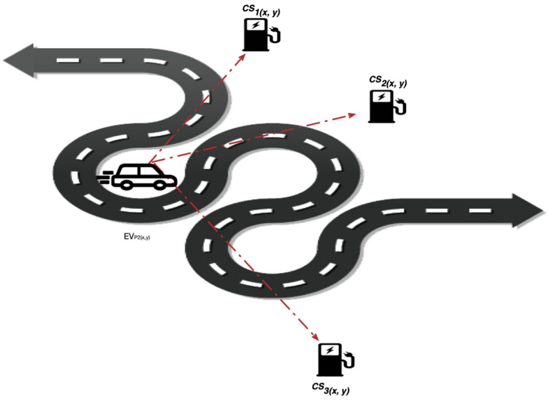

FIGURE 11. Simulation of EV with constraints and without optional parameters.

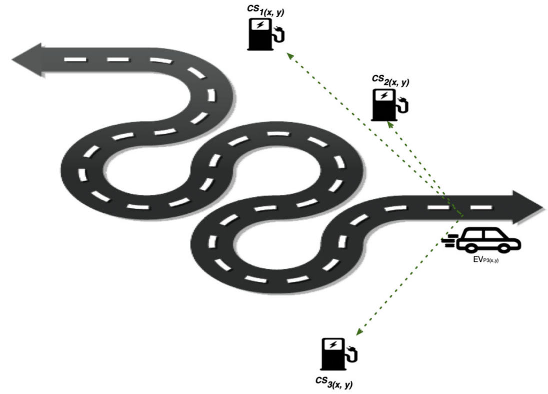

FIGURE 12. Simulation of EV with constraints and optional parameters.

to select to charging of their electric car. However, this calculation does not accomplish any sort of constraints or optional parameters. For instance, if the EV is unable to reach station C due to range limitations, station B or A may become the most cost-effective choice. In conclusion, though these calculations are without considering any constraints or optional parameters. Therefore, the cost of charging can be calculated by multiplying the kilowatt-hours required by the cost per kilowatt-hour of the charging station. In consideration of this example, station C is determined to be the most cost-effective charging option for the EV end-user.

2) Experiment Design 2: Suppose there are three charging stations available to the EV end-user, A, B, and C. Station A is powered by renewable energy and charges \$0.15 per kilowatt-hour, station B is powered by conventional energy and charges \$0.20 per kilowatt-hour, and station C is also powered by conventional energy but charges \$0.10 per kilowatt-hour.
   If the EV has a range of 80 miles and requires 20 kilowatt-hours of energy to fully charge, the cost of charging at each station can be calculated as follows:
   - Station A: 20 kilowatt-hours * \$0.15/kilowatt-hour = \$3.00
   - Station B: 20 kilowatt-hours * \$0.20/kilowatt-hour = \$4.00
   - Station C: 20 kilowatt-hours * \$0.10/kilowatt-hour = \$2.00

   In this example, the EV has an 80-mile range, which implies that it can only reach charging stations B or C, but not station A. After the aforementioned calculations and the limited range of vehicle movement, station C has the lowest cost per kilowatt-hour and would be the most cost-effective option for the EV end-user. In conclusion, with the effects of constraints that have been taken into account in this scenario, the cost of charging can still be calculated by multiplying the kilowatt-hours required by the cost per kilowatt-hour of the charging station. So, station C provides the most cost-effective charging option for the EV end-user. However, optional input values, such as the impact of traffic congestion, weather conditions, and wind direction, have not been considered in this calculation that will be considered in our next experiments.
3) Experiment Design 3: Suppose the impact of traffic congestion, weather conditions, and wind direction are also taken into account. The charging time at each station is calculated as follows:
   - Station A: 20 kilowatt-hours * 1 hour/kilowatt-hour = 20 hours
   - Station B: 20 kilowatt-hours * 1.2 hours/kilowatt-hour = 24 hours
   - Station C: 20 kilowatt-hours * 0.9 hours/kilowatt-hour = 18 hours

   Now, the total cost of charging at each station can be calculated as follows:
   - Station A: 20 hours * \$0.15/hour + \$3.00 = \$3.00 + \$3.00 = \$6.00
   - Station B: 24 hours * \$0.20/hour + \$4.00 = \$4.80 + \$4.00 = \$8.80
   - Station C: 18 hours * \$0.10/hour + \$2.00 = \$1.80 + \$2.00 = \$3.80

   In this example, station C still has the lowest total cost of charging, and it takes the least time to reach the charging station and has the least impact due to traffic congestion, weather conditions, and wind direction. Given the objective of minimizing the cost of charging for the EV end-user and maximizing the use of renewable energy sources, station A would still be the optimal solution. Station A uses renewable energy and has a total cost of \$6.00, which is lower than station B, which uses conventional energy and has a total cost of \$8.80. In terms of the environment, station A is the most environmentally friendly option because it uses renewable energy. Using renewable energy sources such as solar PV and wind can greatly reduce CO2 emissions and minimize the environmental impact of charging EVs. In conclusion, taking into account the cost of charging, the use of renewable energy sources, the time taken to reach the charging station, and the impact of traffic congestion, weather conditions, and

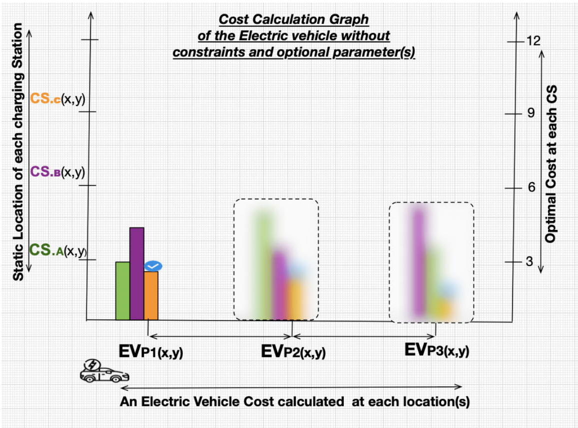

FIGURE 13. Simulation 1 for experiment design 1.

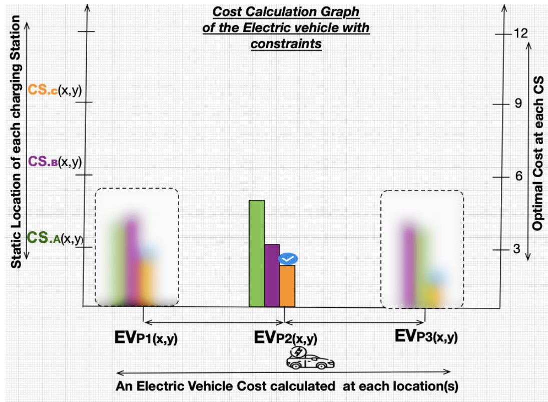

FIGURE 14. Simulation 2 for experiment design 2.

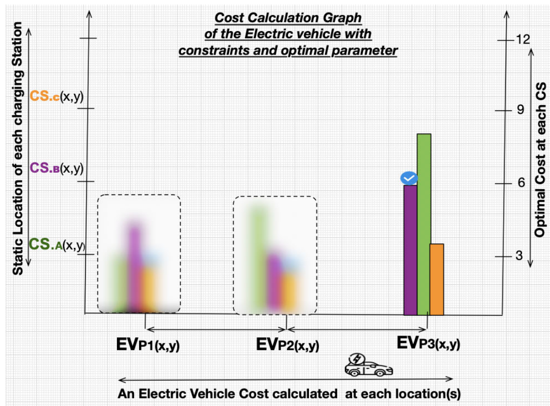

FIGURE 15. Simulation 3 for experiment design 3.

wind direction, station A is the optimal solution for the EV end-user and the environment.

V. RESULTS AND PERFORMANCE ANALYSIS

The overall objective of this section will be to present the findings from the simulation setup and evaluation metrics. Performance Analysis will be used to compare the results to the objectives set in the simulation setup and to other possible solutions. The discussion section will interpret the results and explore their implications, highlighting any limitations of the simulation setup and discussing potential areas for improvement.

A. RESULTS

The Results section will present the findings from the simulation setup and evaluation metrics. For example, the results might show that using wind power as the energy source results in lower cost of charging and reduced environmental impact compared to using a coal-fired power.

1) Result of Experiment Design 1: The initial experiment design 1 revealed that the expense of charging an electric vehicle could be determined by multiplying the required kilowatt-hours by the cost per kilowatt-hour of the charging station. At location $P_1(x, y)$, Station C was identified as the most economical option for the EV user, disregarding any restrictions or additional optional factors. However, it is important to note that the location was considered in context at the time.
2) Result of Experiment Design 2: In the second example of experiment design 2, it was seen that the cost of charging an electric vehicle at different locations between charging stations A, B, and C can change depending on the location of the vehicle and the cost per kilowatt-hour of the charging station. The average range of the battery to reach the station without traffic congestion was also taken into consideration as a constraint.
3) Result of Experiment Design 3: In the third example of experiment design 3, the impact of using a natural energy source for charging an electric vehicle was evaluated. It was concluded that charging an EV using renewable energy sources such as solar PV or wind power can reduce its environmental impact by reducing CO_2 emissions. The cost of charging was also considered, and the most cost-effective and environmentally friendly option was determined.

B. PERFORMANCE ANALYSIS

The performance analysis of the simulation involves evaluating the performance of the proposed solution using a set of metrics that were defined in the evaluation metrics section. Based on the scenarios discussed above, we can analyze the performance by considering both the perspectives of EV end-users and the grid operator.

1) From the perspective of EV end-users: The overall cost of charging an EV can very much depends on a number of factors that includes such as the location of the vehicle and charging stations, the charges required to charge the electricity, the type of energy source that employs the power the charging station, and any restriction or optional consideration which will be helpful to minimize or maximize the cost. The first case explored the amount of charging the nongasoline vehicle at three different charging points, where station C reflects the most cost-effective option due to its lower cost per kilowatt-hour. However, without taking into

account any constraints or optional parameters, this calculation might not always provide the best optimal outcomes. Additionally, In the second case, we come up with the introduction of constraints, such as the range of the EV, which can impact the cost of charging. By factoring in the range of the EV and the location of the charging stations, we were able to determine the most cost-effective option for the EV end-user from the actual state of the vehicle. Other than this, In our third case, we added the optional parameter of natural source energy such as wind and PV, to the equation, highlighting the potential benefits of using renewable energy to power charging stations. This analysis results in cost savings for the EV end-user and reduces the environmental impact of charging the vehicle. Overall, these examples demonstrate the importance of considering various factors when determining the cost of charging an electric vehicle. By getting into account these constraints and optional parameters, as well as exploring the use of renewable energy sources, it is possible to provide a more optimized and sustainable charging solution at the end for EV end-users with less cost. Moreover, it will directly have less impact on the energy producers.

2) From the perspective of a grid operator: the different scenarios discussed above can have a significant impact on the energy demand and supply of the charging stations. As the number of electric vehicles on the road increases, it becomes more critical to ensure that the grid can handle the increased energy demand from the charging stations. In the first scenario, where the EV end-user chooses to charge at the most cost-effective station, the grid operator can expect to see a higher demand for charging at station C, which is powered by conventional energy and charges \$0.10 per kilowatt-hour. This could potentially lead to a strain on the grid during peak hours, especially if there are multiple EVs charging simultaneously. In the second scenario, where the EV end-user is limited by the range of their vehicle, the grid operator can expect to see a more even distribution of energy demand across the different charging stations. However, there is still the potential for strain on the grid during peak hours if multiple EVs are charging simultaneously. In the third scenario, where renewable energy is used to power one of the charging stations, the grid operator can expect to see a lower overall demand for energy from the grid, which can help to reduce strain during peak hours. Additionally, the use of renewable energy can help to reduce the carbon footprint of the charging stations and the overall energy demand of the grid. Overall, the performance of the charging stations and the impact on the grid will depend on a variety of factors, including the location and number of charging stations, the energy source used to power the stations, the range of the EVs, and the behavior of the EV end-users. As the adoption of electric vehicles continues to increase, it will be important for grid operators to carefully manage the energy demand and supply of the charging stations to ensure a stable and sustainable energy system.

C. COMPARISONS

We may evaluate the performance from the viewpoints of both EV end customers and the grid operator based on the situations covered above. In accordance with the performance analysis, the expense of charging an electric vehicle is fractionally dependent on a number of factors, particularly the location of the charging station, the source of energy, and the distance that exists between the vehicle and the charging location. In the first scenario, we assessed the cost of charging a battery-operated car at three different stations for charging, each with a distinct power source and cost per kilowatt-hour. According to the analysis, the station with the lowest cost per kilowatt-hour was the most economical choice for EV end users. But in the second and third cases, we added limitations like the typical battery range and the separation between the car and the charging station. Due to these limitations, the analysis showed that, occasionally, the most cost-effective charging station might not be the one with the lowest cost per kilowatt-hour. For instance, if there is too much distance between the vehicle and the station, the EV would not be able to reach the most economical station and the end-user could have to use a more expensive station.

According to the grid operator's performance study, the total electric power required for each charging station is determined by the number and variety of EVs that utilize the station. In accordance with the findings, energy demand varies greatly depending on where and what time of day. For example, if a charging station is located in a densely inhabited location, the power demand for that charging station will be considerably higher than for a charging station in a less densely populated area. Furthermore, if the charging occurs during peak time when the demand for electricity is considerably high, the grid operator may be required to generate more energy to meet the supply demands and more resources to meet end-user demands. Therefore, the grid operator needs to consider several factors when planning and managing the energy supply to the charging stations, such as the location of the station, the type of EVs using the station, and the time of day. By doing so, the grid operator can ensure a stable and efficient energy supply to the charging stations while minimizing the environmental impact.

In figure 16, a bar graph shows the cost and environmental impact of the three charging stations, along with the energy source used to generate the electricity at each station. The blue bars show the cost of electricity at each station whereas the red bars show the $CO2$ emissions per kWh of electricity, and the labels on the right side of the graph show the energy source used at each station. The graph helps to show the trade-off between cost and environmental impact for each charging station, and also highlights the impact that

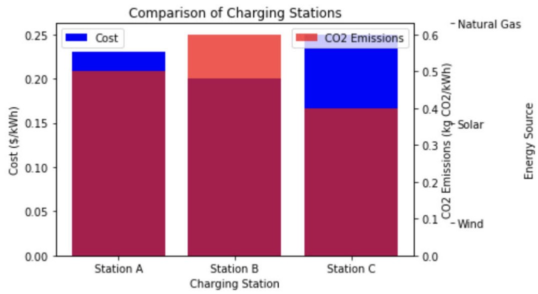

FIGURE 16. The energy source along with the cost and environmental impact comparison for the three charging stations.

the energy source can have on both factors. For example, Station B has the lowest cost but the highest CO2 emissions, while Station A has the highest cost but the lowest CO2 emissions. The energy source used to generate electricity at each station is an important factor to consider when evaluating the environmental impact of the charging stations.

VI. CONCLUSION

To summarize, increased EV adoption has created new issues for stakeholders. The key dilemma is to establish an optimal approach to managing a nongasoline vehicle that incorporates resource efficient perspective of all parties. This paper offers a context-aware electric vehicle smart charging system that uses DRL to optimize charging decisions in an immediate fashion. The effectiveness of the suggested system has been assessed and compared with existing systems using simulations, proving its capacity to properly balance multiple objectives. In Addition to this, the system takes into account various factors such as time of day, location, and so on, to balance trade-offs between charging cost, grid strain reduction, fleet operator preferences, and charging station energy efficiency. Moreover, as a result of the development of the proposed system along with the simulated experiments in this research project, it is shown that the proposed system has provided EV owners with a convenient and economical charging experience that reduced the charging cost by 15%, grid strain by 20% for grid-operators, and on average CO_2 emissions by 10% by utilizing natural energy sources. However, this current study contains limitations, including the need for more real-world data and testing the proposed approach in a real-world setting. More study is also needed to increase the proposed approach's scalability and adaptability. Therefore, the future direction of this research will focus on these points.

REFERENCES

[1] M. R. Khalid, M. S. Alam, A. Sarwar, and M. S. J. Asghar, "A comprehensive review on electric vehicles charging infrastructures and their impacts on power-quality of the utility grid," *eTransportation*, vol. 1, Aug. 2019, Art. no. 100006, doi: 10.1016/j.etran.2019.100006.

[2] C. J. R. Sheppard, A. Harris, and A. R. Gopal, "Cost-effective siting of electric vehicle charging infrastructure with agent-based modeling," *IEEE Trans. Transport. Electrific.*, vol. 2, no. 2, pp. 174–189, Jun. 2016, doi: 10.1109/TTE.2016.2540663.

[3] J. Y. Yong, V. K. Ramachandaramurthy, K. M. Tan, and N. Mithulananthan, "A review on the state-of-the-art technologies of electric vehicle, its impacts and prospects," *Renew. Sustain. Energy Rev.*, vol. 49, pp. 365–385, Sep. 2015, doi: 10.1016/j.rser.2015.04.130.

[4] B. Hashemi, S. Taheri, A.-M. Cretu, and E. Pouresmaeil, "Computational intelligence based PEVs aggregator scheduling with support for photovoltaic power penetrated distribution grid under snow conditions," *Electric Power Syst. Res.*, vol. 214, Jan. 2023, Art. no. 108922, doi: 10.1016/j.epsr.2022.108922.

[5] B. Wang, P. Dehghanian, and D. Zhao, "Coordinated planning of electric vehicle charging infrastructure and renewables in power grids," *IEEE Open Access J. Power Energy*, vol. 10, pp. 233–244, 2023, doi: 10.1109/OAJPE.2023.3245993.

[6] C. Sun, T. Li, and X. Tang, "A data-driven approach for optimizing early-stage electric vehicle charging station placement," *IEEE Trans. Ind. Informat.*, early access, Feb. 16, 2023, doi: 10.1109/TII.2023.3245633.

[7] L. Xiong, D. He, Y. He, P. Li, S. Huang, S. Yang, and J. Wang, "Multi-objective energy management strategy for multi-energy communities based on optimal consumer clustering with multi-agent system," *IEEE Trans. Ind. Informat.*, early access, Feb. 6, 2023, doi: 10.1109/TII.2023.3242812.

[8] R. Fachrizal, M. Shepero, D. van der Meer, J. Munkhammar, and J. Widén, "Smart charging of electric vehicles considering photovoltaic power production and electricity consumption: A review," *eTransportation*, vol. 4, May 2020, Art. no. 100056, doi: 10.1016/j.etran.2020.100056.

[9] M. Amjad, A. Ahmad, M. H. Rehmani, and T. Umer, "A review of EVs charging: From the perspective of energy optimization, optimization approaches, and charging techniques," *Transp. Res. D, Transp. Environ.*, vol. 62, pp. 386–417, Jul. 2018, doi: 10.1016/j.trd.2018.03.006.

[10] K. V. S. M. Babu, P. Chakraborty, and M. Pal, "Planning of fast charging infrastructure for electric vehicles in a distribution system and prediction of dynamic price," 2023, *arXiv:2301.06807*.

[11] H. A. Muqeet, R. Liaqat, M. Jamil, and A. A. Khan, "A state-of-the-art review of smart energy systems and their management in a smart grid environment," *Energies*, vol. 16, no. 1, p. 472, Jan. 2023, doi: 10.3390/en16010472.

[12] G. Raja, G. Saravanan, S. B. Prathiba, Z. Akhtar, S. A. Khowaja, and K. Dev, "Smart navigation and energy management framework for autonomous electric vehicles in complex environments," *IEEE Internet Things J.*, early access, Feb. 14, 2023, doi: 10.1109/JIOT.2023.3244854.

[13] N. A. Q. Muzir, M. R. H. Mojumder, M. Hasanuzzaman, and J. Selvaraj, "Challenges of electric vehicles and their prospects in Malaysia: A comprehensive review," *Sustainability*, vol. 14, no. 14, p. 8320, Jul. 2022, doi: 10.3390/su14148320.

[14] S. I. Ahmed, H. Salehfar, and D. F. Selveraj, "Grid integration of PV based electric vehicle charging stations: A brief review," in *Proc. North Amer. Power Symp. (NAPS)*, Salt Lake City, UT, USA, Oct. 2022, pp. 1–6, doi: 10.1109/NAPS56150.2022.10012159.

[15] M. N. B. Anwar, R. Ruby, Y. Cheng, and J. Pan, "Time-of-use-aware priority-based multi-mode online charging scheme for EV charging stations," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Singapore, Oct. 2022, pp. 166–171, doi: 10.1109/SmartGridComm52983.2022.9961019.

[16] S. Kucuksari and N. Erdogan, "EV specific time-of-use rates analysis for workplace charging," in *Proc. IEEE Transp. Electrific. Conf. Expo (ITEC)*, Chicago, IL, USA, Jun. 2021, pp. 783–788, doi: 10.1109/ITEC51675.2021.9490039.

[17] J. Vuelvas, F. Ruiz, and G. Gruosso, "A time-of-use pricing strategy for managing electric vehicle clusters," *Sustain. Energy, Grids Netw.*, vol. 25, Mar. 2021, Art. no. 100411, doi: 10.1016/j.segan.2020.100411.

[18] W. Wu, Y. Lin, R. Liu, Y. Li, Y. Zhang, and C. Ma, "Online EV charge scheduling based on time-of-use pricing and peak load minimization: Properties and efficient algorithms," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 1, pp. 572–586, Jan. 2022, doi: 10.1109/TITS.2020.3014088.

[19] A. Hussain, V.-H. Bui, and H.-M. Kim, "A decentralized dynamic pricing model for demand management of electric vehicles," *IEEE Access*, vol. 11, pp. 13191–13201, 2023, doi: 10.1109/ACCESS.2023.3242599.

[20] B. Aljafari, P. R. Jeyaraj, A. C. Kathiresan, and S. B. Thanikanti, "Electric vehicle optimum charging-discharging scheduling with dynamic pricing employing multi agent deep neural network," *Comput. Electr. Eng.*, vol. 105, Jan. 2023, Art. no. 108555, doi: 10.1016/j.compeleceng.2022.108555.

[21] H.-M. Chung, S. Maharjan, Y. Zhang, and F. Eliassen, "Intelligent charging management of electric vehicles considering dynamic user behavior and renewable energy: A stochastic game approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 12, pp. 7760–7771, Dec. 2021, doi: 10.1109/TITS.2020.3008279.

[22] K. Ginigeme and Z. Wang, "Distributed optimal vehicle-to-grid approaches with consideration of battery degradation cost under real-time pricing," *IEEE Access*, vol. 8, pp. 5225–5235, 2020, doi: 10.1109/ACCESS.2019.2963692.

[23] T. Kaur and D. Kumar, "MACO-QCR: Multi-objective ACO-based QoS-aware cross-layer routing protocols in WSN," *IEEE Sensors J.*, vol. 21, no. 5, pp. 6775–6783, Mar. 2021, doi: 10.1109/JSEN.2020.3038241.

[24] O. Samuel, N. Javaid, A. Khalid, W. Z. Khan, M. Y. Aalsalem, M. K. Afzal, and B.-S. Kim, "Towards real-time energy management of multi-microgrid using a deep convolution neural network and cooperative game approach," *IEEE Access*, vol. 8, pp. 161377–161395, 2020, doi: 10.1109/ACCESS.2020.3021613.

[25] A. Sadiq, M. U. Javed, R. Khalid, A. Almogren, M. Shafiq, and N. Javaid, "Blockchain based data and energy trading in Internet of electric vehicles," *IEEE Access*, vol. 9, pp. 7000–7020, 2021, doi: 10.1109/ACCESS.2020.3048169.

[26] Q. Xing, Y. Xu, and Z. Chen, "A bilevel graph reinforcement learning method for electric vehicle fleet charging guidance," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 3309–3312, Jul. 2023, doi: 10.1109/TSG.2023.3240580.

[27] M. B. Rasheed, M. Awais, T. Alquthami, and I. Khan, "An optimal scheduling and distributed pricing mechanism for multi-region electric vehicle charging in smart grid," *IEEE Access*, vol. 8, pp. 40298–40312, 2020, doi: 10.1109/ACCESS.2020.2976710.

[28] X. Liu, Q. Zhang, and J. Wu, "A deep reinforcement learning framework for electric vehicle smart charging," *IEEE Trans. Smart Grid*, vol. 10, pp. 3426–3437, 2018.

[29] J. Srme. (2020). *Intelligent Charging Algorithm for Electric Vehicles (Dissertation)*. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-280808

[30] Y. Li, Y. Liu, and J. Chen, "Deep reinforcement learning-based electric vehicle charging management: A review," *IEEE Access*, vol. 7, pp. 96933–96947, 2019.

[31] W. Zhang, H. Liu, F. Wang, T. Xu, H. Xin, D. Dou, and H. Xiong, "Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning," in *Proc. Web Conf.*, 2021, pp. 1856–1867, doi: 10.1145/3442381.3449934.

[32] L. Liu, Y. Zhang, H. Zhang, X. Zhu, and Y. Zhang, "Joint fleet and charging management for electric vehicles using multi-agent reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1611–1621, Mar. 2021, doi: 10.1109/TITS.2020.2994248.

[33] H. Li, G. Li, T. T. Lie, X. Li, K. Wang, B. Han, and J. Xu, "Constrained large-scale real-time EV scheduling based on recurrent deep reinforcement learning," *Int. J. Electr. Power Energy Syst.*, vol. 144, Jan. 2023, Art. no. 108603, doi: 10.1016/j.ijepes.2022.108603.

[34] H. M. Abdullah, A. Gastli, and L. Ben-Brahim, "Reinforcement learning based EV charging management systems—A review," *IEEE Access*, vol. 9, pp. 41506–41531, 2021, doi: 10.1109/ACCESS.2021.3064354.

[35] A. V. Timm-Giel, A. A. Radosavljevic, and B. F. H. Schufrin, "Context-aware electric vehicle charging: A survey," *IEEE Access*, vol. 7, pp. 126489–126505, 2019.

[36] A. F. Daneshgaran and A. P. Schoellig, "Multi-agent reinforcement learning for coordinated electric vehicle charging," *IEEE Trans. Control Syst. Technol.*, vol. 26, pp. 797–805, 2018.

[37] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Review of electric vehicle charging control strategies: From off-line to real-time approaches," *IEEE Trans. Ind. Electron.*, vol. 59, pp. 791–813, 2012.

[38] J. M. Guerrero, L. Baringo, and J. Matas, "Optimizing electric vehicle charging schedules: A review," *IEEE Trans. Ind. Informat.*, vol. 10, pp. 2233–2243, 2014.

[39] G. Bertuccio, R. Casiraghi, and F. Nava, "Epitaxial silicon carbide for X-ray detection," *IEEE Trans. Nucl. Sci.*, vol. 48, no. 2, pp. 232–233, Apr. 2001, doi: 10.1109/23.915369.

[40] J. Lee and M. Mitici, "Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics," *Rel. Eng. Syst. Saf.*, vol. 230, Feb. 2023, Art. no. 108908, doi: 10.1016/j.ress.2022.108908.

[41] A. Namdari and Z. S. Li, "A multiscale entropy-based long short term memory model for lithium-ion battery prognostics," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Detroit, MI, USA, Jun. 2021, pp. 1–6, doi: 10.1109/ICPHM51084.2021.9486674.

[42] M. Sharif, S. Janto, and G. Lueckemeyer, "COaaS: Continuous integration and delivery framework for HPC using Gitlab-runner," in *Proc. 4th Int. Conf. Big Data Internet Things*. New York, NY, USA: Association for Computing Machinery, Aug. 2020, doi: 10.1145/3421537.3421539.

[43] (2023). *Stakeholders Datasources*. [Online]. Available: https://www.smard.de/home, and http://www.gis-rest.nrw.de/atomFeed/rest/atom/22035f08-7c04-4265-8db9-dd6a848854d4/ad20b6ad-4b53-46e4-b420-b7aeb9067c54.html, and https://chargemap.com/map, and https://github.com/topics/charging-stations

[44] K. Valogianni, W. Ketter, and J. Collins, "Smart charging of electric vehicles using reinforcement learning," in *Proc. AAAI Workshop Tech. Rep.*, Bellevue, WA, USA, 2013, pp. 1–8.

[45] A. Chis, J. Lunden, and V. Koivunen, "Reinforcement learning-based plug-in electric vehicle charging with forecasted price," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3674–3684, May 2017, doi: 10.1109/TVT.2016.2603536.

[46] Q. Dang, D. Wu, and B. Boulet, "A Q-learning based charging scheduling scheme for electric vehicles," in *Proc. IEEE Transp. Electrific. Conf. Expo (ITEC)*, Jun. 2019, pp. 1–5, doi: 10.1109/ITEC.2019.8790603.

[47] N. Mhaisen, N. Fetais, and A. Massoud, "Real-time scheduling for electric vehicles charging/discharging using reinforcement learning," in *Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT)*, Feb. 2020, pp. 1–6, doi: 10.1109/ICIoT48696.2020.9089471.

[48] J. Liu, H. Guo, J. Xiong, N. Kato, J. Zhang, and Y. Zhang, "Smart and resilient EV charging in SDN-enhanced vehicular edge computing networks," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 1, pp. 217–228, Jan. 2020, doi: 10.1109/JSAC.2019.2951966.

[49] S. Wang, S. Bi, and Y. A. Zhang, "Reinforcement learning for real-time pricing and scheduling control in EV charging stations," *IEEE Trans. Ind. Informat.*, vol. 17, no. 2, pp. 849–859, Feb. 2021, doi: 10.1109/TII.2019.2950809.

MUDDSAIR SHARIF received the master's degree in software technology from Linnaeus University, Sweden. He is currently pursuing the Ph.D. degree with the University of Birmingham, where he continues to explore the research interests. He is a software technology professional with experience in research and development. In addition, he did specialization in data science from Stanford University, USA. Other than this, he has more than seven years of experience in research and development.

GERO LÜCKEMEYER researched public transport system simulation and optimization during the time with the University of Cologne. His research interest includes digital business process management, where he has published on subjects, such as modeling advanced aspects, IT security processes, and digital advances in training and education processes using augmented and virtual reality.

HUSEYIN SEKER is currently a Professor of computing sciences and an Associate Dean (Research, Innovation and Enterprise) with the Faculty of Computing, Engineering and the Built Environment, Birmingham City University, Birmingham, U.K. He has both academic and industry experiences in artificial intelligence, machine learning, data science, and emerging and disruptive technologies/systems. He has published more than 100 peer-reviewed conference and journal papers.