

## Research Article

# SV2G-ET: A Secure Vehicle-to-Grid Energy Trading Scheme Using Deep Reinforcement Learning

Aparna Kumari <sup>1</sup>, Mihir Trivedi <sup>2</sup>, Sudeep Tanwar <sup>3</sup>, Gulshan Sharma <sup>4</sup>,  
and Ravi Sharma<sup>5</sup>

<sup>1</sup>Computer Science and Engineering Department, Institute of Computer Technology, Ganpat University, Mehsana, Gujarat, India

<sup>2</sup>Department of Electrical Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

<sup>3</sup>Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

<sup>4</sup>Department of Electrical Engineering Technology, University of Johannesburg, Johannesburg, South Africa

<sup>5</sup>Centre for Inter-Disciplinary Research and Innovation, University of Petroleum and Energy Studies, Dehradun, India

Correspondence should be addressed to Sudeep Tanwar; [sudeep.tanwar@nirmauni.ac.in](mailto:sudeep.tanwar@nirmauni.ac.in) and Gulshan Sharma; [gulshanmail2005@gmail.com](mailto:gulshanmail2005@gmail.com)

Received 23 January 2022; Revised 14 March 2022; Accepted 8 April 2022; Published 29 April 2022

Academic Editor: Sitharthan R

Copyright © 2022 Aparna Kumari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, advancements in electric vehicle (EV) technology and rising petrol prices have increased the demand for EVs and also made them important for the Smart Grid (SG) economy. During the high energy demand, Vehicle to Grid (V2G) comprises a notable feature that returns the stored energy back to the grid. However, due to dynamic nature of energy prices and EVs availability, determining the best charging and discharging strategy is quite difficult. The existing approaches need a model to predict the uncertainty and optimize the scheduling problem. Further, other issues like security, scalability, and real-time data accessibility of EVs energy trading (ET) data at low cost also exist. Though many solutions exist, they are not adequate to handle the aforementioned issues. This paper proposes a Secure V2G-Energy Trading (SV2G-ET) scheme using deep Reinforcement Learning (RL) and Ethereum Blockchain Technology (EBT). The proposed SV2G-ET scheme employs a deep Q-network for EVs scheduling for charging/discharging. SV2G-ET scheme uses InterPlanetary File System (IPFS) and smart contract (SC) for secure access of EV's ET data in real time. The experimental results prove the efficacy of the proposed SV2G-ET scheme that leads to improved scalability, saving the EVs charging cost, low ET data storage cost, and increased EV owner's profit.

## 1. Introduction

Electric vehicles (EVs) have become popular due to several environmental and economical advantages compared to fossil fuel cars such as reducing pollution, greenhouse gas emission, and operation cost [1]. In modern transportation systems, EVs are the preferred choice due to their energy and environmental sustainability, although they often fail to find proper charging stations due to the limited infrastructure of Smart Grid (SG) and unbalance demand of charging. SG offers bidirectional communications with distributed computers to increase the grid's quality-of-service, efficiency, reliability, and stability. In SG, the availability of public charging infrastructure is critical to accommodate more renewable energy, lower carbon emissions, and reduce peak charging demands. So, the development of

smart networks enables the mechanism of Vehicle-to-Grid (V2G), where energy flows from EVs to the grid, rather than Grid-to-Vehicle (G2V) mechanism.

The International Energy Agency (IEA) reports that, by 2030, the number of EVs will reach 145 million, which will account for 7% of the total global vehicle fleet. Therefore, the efficient management of the EVs charging/discharging to reduce cost has become challenging due to random EV's arrival and departure times, randomness of the EVs' utility, and dynamic price of energy. In V2G, EVs can be used for excess energy storage from the SG to balance energy demand, smooth demand peak, and eliminate distribution losses [1].

Various research work has been done in past years to reduce the charging operation cost of EVs [2]. Yao et al. [3] optimally solved the charging scheduling problem. Then, Wang et al. [4]

presented an online charging algorithm to maximize the profit of charging stations with Demand Response (DR) mechanisms. Here, a scheduling mechanism has been proposed in order to efficiently manage day-ahead charging and discharging of EVs. Ortega-Vazquez et al. [5] proposed a vigorous optimization mechanism for EV charging integrated price uncertainty and battery degradation. In [6], a two-stage stochastic approach has been proposed by the authors for EV charging scheduling. Energy demand of EVs and energy prices are dynamic in nature; hence, real-time scheduling techniques have grabbed the attention using various Artificial Intelligence (AI) mechanisms, for instance, Machine Learning (ML) and Deep Learning (DL). Das et al. [7] designed an ML-based algorithm to learn the EV's status and inform when the EVs should perform G2V/V2G using support vector machines (SVM). Donovan et al. [8] presented a prediction mechanism for the EV's availability to perform energy trading (ET) and used the Gurobi optimization to minimize the consumer's energy bill.

Several benefits of V2G include efficiency, peak load shaving, clean fuel, quiet operation, grid infrastructure reduction, etc.; at the same time, there are disadvantages like adverse battery effects, new infrastructure, new trading mechanism, and security from cyberattacks, replay attacks, data manipulation attacks, etc. For example, Bartłomiej et al. [1] presented the effectiveness of the V2G process using predictive model [9, 10]. Moreover, the various ML- and DL-based approaches exist, but they have various challenges; for example, ML/DL models are trained on the training data and then tested over the test dataset. Whenever unseen data pass to the ML/DL model, it does not work as per the need. To handle the aforesaid challenges, the emergent model-free mechanism, i.e., Reinforcement Learning (RL), is a viable solution [11].

Recently, the model-free approach has proven its effectiveness in various decision-making applications such as e-healthcare and intelligent transportation systems (ITS). RL learns a good control policy, takes appropriate actions to achieve desired goals, and does not require prior knowledge of the system [12–14]. RL approaches like deep Q-network (DQN) and Q-learning have been adopted widely in the SG system [15–17]. Dang et al. [18] implemented the charging scheduling problem using the Q-learning mechanism. Chis et al. [19] presented an RL-based algorithm to schedule the individual charging of EVs at home that reduces energy costs of EV owners.

Some issues are identified in the existing approaches, for instance, no guarantee for optimality, financial losses, impracticality to implement, and scalability issue. Other issues like single-point-failure, reliability, trust, data security, QoS, and transparency also exist in the available approaches [20–23]. Further, high deployment and usage of EVs for energy exchange raise various security and privacy concerns like unauthorized access to EVs data, denial-of-service, jamming, data injection attack, eavesdropping, and hijacking [24, 25]. These security attacks either harm the EVs or stop them from performing their committed task. This motivates us to develop and design a solution that copes up with the aforesaid security issues.

The abovementioned issues can be self-addressed with the adoption of blockchain technology in the EVs trading

system [26], which is distributed in nature and integrated trust, immutability, and security [27]. The trust among two parties in the EVs trading system can be achieved using self-executable and self-verifiable smart contracts (SCs) [28]. Several blockchain-based V2G energy trading systems have a number of issues like limited data storage, EVs data storage cost, etc.

To handle the aforementioned issues, this paper proposed a Secure V2G Energy Trading (*SV2G-ET*) scheme using deep RL and Ethereum Blockchain Technology (EBT). The proposed *SV2G-ET* scheme benefits the EVs owner by maximizing their profits and securing energy trading data using blockchain (through SCs). Here, the proposed scheme uses EBT; the reasoning behind this is EBT was designed to be open, flexible, low-cost, and suited for multiple parties for trading cooperation in real time. EBT architecture has unique layers that strengthen energy trading systems and create new possibilities [29, 30]. *SV2G-ET* uses EBT for V2G trading to handle the interoperability among various entities like EVs owners, utility providers, and grid operators via wireless (like 5G-enabled tactile internet and 6G) or wired communication channel [31]. Table 1 shows a relative comparison of the proposed *SV2G-ET* scheme with existing approaches considering their merits and demerits.

*1.1. Research Gaps.* Following are the research gaps identified in the existing approaches:

- (i) Some issues are identified like no guarantee for optimality, financial losses, and impracticality to implement [20, 22, 23].
- (ii) Then, issues like single-point-failure, reliability, trust, data security, QoS, and transparency also exist in the available approaches [29, 30].
- (iii) Next, blockchain-based existing approaches suffer from limited data storage, data storage cost, and scalability issues [31].

To handle the aforementioned issues, this paper proposes *SV2G-ET* scheme based on deep RL and EBT. Here is the list of motivations for this research.

*1.2. Motivation.*

- (i) The importance of V2G energy trading is vital due to the high deployment and usage of EVs which raises various security concerns like unauthorized data access, denial-of-service, jamming, data injection attack, and hijacking [24, 25].
- (ii) Most of the existing approaches emphasize on centralized solutions for EVs energy trading that has a single-point-failure, security, and privacy issues.
- (iii) Security attacks either harm the EVs or stop them from performing their committed task. This motivates us to develop and design a solution that copes up with the aforesaid issues.
- (iv) Motivated from the above facts, this paper proposed a *SV2G-ET* scheme using deep RL and EBT. Here,

TABLE 1: A comparative analysis of the proposed *SV2G-ET* scheme with existing approaches.

year	Proposed approaches	Short descriptions	Merits	Demerits
2018	Wan et al. [32]	A real-time EV charging scheduling mechanism using deep RL	Realistic variability of the EV traveling patterns and wholesale prices are considered	Scalability issues along with no data accessibility for user
2019	Wang et al. [22]	Maximizing the objective of a charging scheduling and real-time pricing problem using an RL algorithm hyperopia state-action-reward-state-action (SARSA)	138.5% higher charging-station profit compared to benchmark algorithms	Need to be included data security
2019	Dang et al. [18]	Q-learning algorithm-based dynamic charging scheduling scheme	Multidimensional Q-learning is applied for optimal dispatch	Compute-intensive in case the size of Q-table is huge
2020	Mhaisen et al. [33]	To control the charging and discharging operation of EVs using Q-learning	Model-free and off policy mechanism is employed	State elements shift is not considered
2020	Mohanty et al. [7]	ML-based support vector machine (SVM) algorithm is applied for scheduling of EVs	Various parameters such as energy consumption, trip time, and charging state for every interval are considered	Advanced ML algorithms like deep neural network (DNN), etc. need to be explored depending upon the length of the data
2020	Qiu et al. [12]	A deep RL method for pricing EVs using discrete charging levels	This method achieves 31% higher profit compared to Q-learning-based approach	Realistic variability of the EV traveling patterns and wholesale prices are not considered
2021	Lin et al. [34]	Deep RL for EVs routing problem along with time windows	Solve EVs routing problem instances of large sizes	Scalability and data security issues
2021	Ding et al. [35]	Secure policy for energy trading in smart microgrids for peer-to-peer trading	Optimal pricing method is used on real dataset	Scalability and data storage cost issue on blockchain
2022	The proposed <i>SV2G-ET</i> scheme	Secure V2G energy trading using deep RL and EBT	Handling data storage cost issue on blockchain, increasing EV's owner profit, improving scalability, saving EVs charging cost	—

deep RL is employed for charging/discharging schedule of EVs.

- (v) The reason behind employing EBT is that EBT was designed to be open, flexible, low-cost, and suited for multiple parties for trading cooperation in real time. Next, the deep RL works well on unseen data due to its model-free mechanism [29, 30].

*1.3. Research Contributions.* Following are the research contributions of this paper.

- (i) A model-free deep RL-based scheme, i.e., *SV2G-ET* is proposed for the charging/discharging schedule of EVs.
- (ii) Secure SCs are designed for decentralized V2G and also secure EVs data using the InterPlanetary File System (IPFS) mechanism.
- (iii) The proposed scheme facilitates the EVs owner to perform V2G/G2V with a nearby charging station to either buy/sell energy in a secure way.
- (iv) The effectiveness of the proposed scheme is evaluated based on various parameters like profit maximization of EVs owner and high system throughput by low data storage cost, data security, and scalability.

*1.4. Organization of the Paper.* The remaining sections of the paper are organized as follows. Section 2 narrates the system

model of the proposed *SV2G-ET* scheme and includes problem formulation of it. Then, Section 3 specifies the workflow of the proposed *SV2G-ET* scheme. Section 4 presents the experimental results of *SV2G-ET* scheme and lastly, the paper is concluded in Section 5.

## 2. System Model and Problem Formulation

This section presents *SV2G-ET*'s system model and problem formulation of the proposed scheme.

*2.1. System Model.* Figure 1 shows the system model of the proposed *SV2G-ET* scheme, where the EV charging station's operation is seen over the horizon of time  $T$ . Here, arrival of EVs to an EV charging station is a random process.

In *SV2G-ET* scheme, the energy is getting generated at the grid and supplied to charging station via a utility supplier. Then, EVs are arriving at  $t$  time slot,  $N_t$  indicates the volume of EV, and energy is consumed by end-consumer, i.e., EVs owner through EVs. An  $i^{th}$  EV gets charged (in timeslot  $t$ ) at charging station based on the following parameters:  $\{p_t, t_{arrival}, d_t\}$ . Here,  $p_t$  represents the energy price at time  $t$ ,  $t_{arrival}$  is arrival time, and  $d_t$  is the charging demand. All these parameters are included in one vector  $K_t$ , which is represented as follows:  $K_t = (p_t, t_{arrival}, d_t)$ .

In the proposed *SV2G-ET* scheme, at each time slot  $t$ , the utility supplier supplied hourly energy prices to the charging station to perform energy trading between grid and EVs, i.e.,

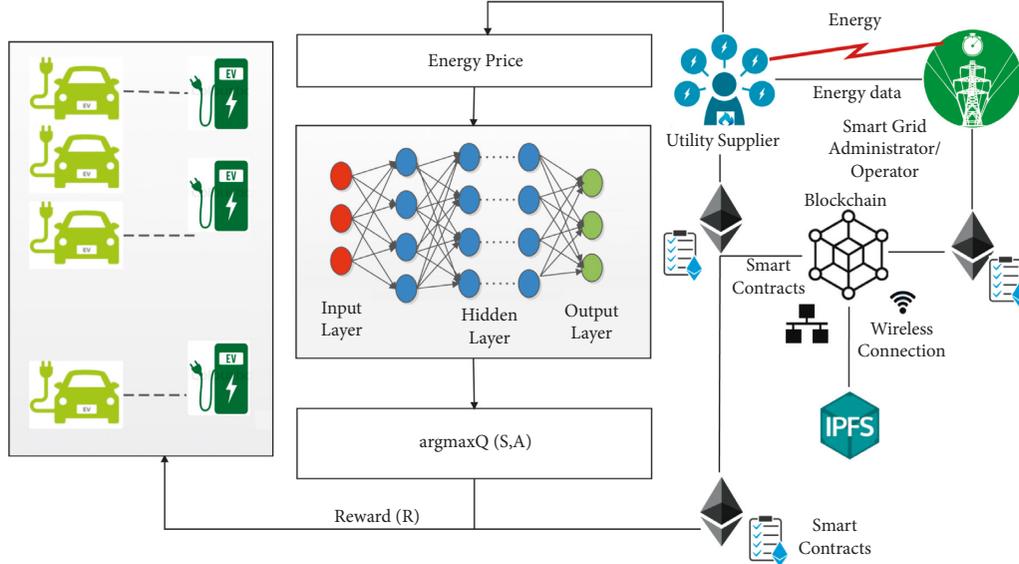


FIGURE 1: SV2G-ET system model.

V2G/G2V. The proposed scheme employed deep Q-network (DQN) based on the dynamic energy prices to perform V2G/G2V operation. Here, DQN uses the RL method, as well as Convolutional Neural Network (CNN) to approximate the Q-function, i.e., action-value function. The DQN agent determines the rates of energy and in return decides whether the action is G2V or V2G.

Once G2V/V2G operation is performed, then entire details of the energy trading are shared securely in real time among all entities. Every entity  $E$ , i.e., utility supplier, grid administrator, and EVs are connected through EBT. The EBT is quite costly to store data, so the proposed scheme incorporates the IPFS mechanism to store data in off-chain mode. Then, only the hash-key of data is stored over EBT, which reduces energy data storage cost and improves system scalability. Moreover, IPFS stores EV's owner details, energy records, etc. These EV's ET data can be accessed using the hash-key over EBT. All the transactions are completed with full encryption and safety using unique SCs, designed over EBT.

**2.2. Problem Formulation.** In the proposed SV2G-ET scheme, the aim is to design the feature that will make a decision to perform V2G/G2V operation based upon the present energy price and the history of previous energy prices. Based on the EV's ET obtained on an hourly basis, the amount of profit gets calculated. The problem is formulated by applying Markov Decision Process (MDP) in discrete time steps.

At every time step  $t$  (i.e., discrete in nature), the agent observes the energy price and availability of EV and then based on the agent action EVs either consume energy or distribute energy. Here, mathematical architecture is supported by MDP to perform decision-making (outcomes are partially controlled by a decision-maker and partly random). MDP comprises five-tuple  $(s, a, p, r, \delta)$ . Here,  $s$  represent

system states,  $a$  shows finite actions,  $p$  is identified as probability for state transition,  $\delta$  is the discount factor, and  $\delta$  is reward.

The details of MDP formulation for the proposed scheme are shown as follows.

- (ii) Action: in the given state  $\mathcal{S}_t$ , the action  $\mathcal{A}_t$  represents the charging and discharging of energy. The action of charging or G2V is selected when the energy price is *low* and action of discharging or V2G is selected when the energy price is *high*. The charging/discharging constraints are defined as follows:

$$-E_{\text{dis}} \leq a_t \leq E_{\text{ch}}. \quad (1)$$

- (iii) Reward: the reward is calculated using current action ( $\mathcal{A}_t$ ) and current state ( $\mathcal{S}_t$ ). The reward function at  $t$  time step is as follows:

$$r_t = \begin{cases} ep_t \cdot a_t, & l_t \neq 0, \\ ep_t \cdot a_t - |e_{\text{max}} - e_t|, & l_t = 0. \end{cases} \quad (2)$$

The reward is represented as the energy price at which the action has occurred and multiplied by the action performed. At the time of departure, the reward is reduced by the price difference of the selling and buying action of the agent.

- (iv) Probability for state transition:

$$p = (a_t, s_t, s_{(t+1)}). \quad (3)$$

Here, the above equation represents transition probability  $p$  from state  $s_t$  to  $s_{(t+1)}$  with respect to the action  $a_t$ . In the proposed scheme, there is no model hence the scheme needs to learn by sampling.

- (v) Action-value function: it predicts sum of future rewards (i.e., for  $k$  time steps window), used to

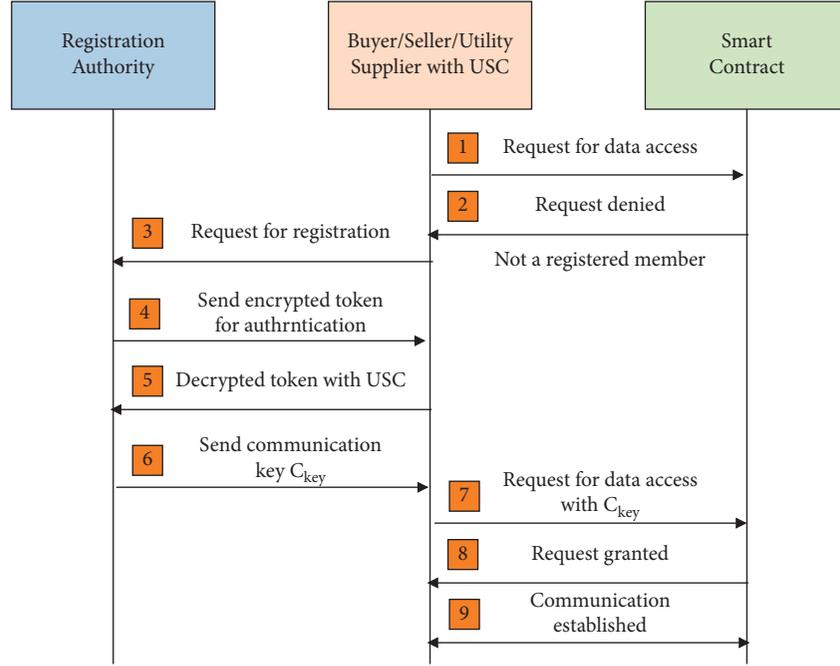


FIGURE 2: Registration of all stakeholders with trusted registration authority for real-time data access in the blockchain.

evaluate EV's charging/discharging scheduling for  $\{action, state\}$  pair, i.e.,  $\{a, s\}$ .

$$Q_\rho(s, a) = \mathbb{E}_\rho \left\{ \sum_{k=0}^K \delta^k \cdot r_{t+k} | s_t, a_t \right\}. \quad (4)$$

Here,  $Q_\rho(s, a)$  is the action-value function.  $\rho$  is the EV charging/dischARGE policy that translates a system condition to a charging/dischARGE schedule and  $\delta$  is the discount factor. Here,  $\delta$  weighs the significance of present and future benefits [32]. It improves the profit  $\mathbb{P}$  of EVs' owners through V2G/G2V operation.

Once EV's ET is done, then the complete details of the transaction are shared among all stakeholders using EBT. The EBT is incorporated with IPFS mechanism to improve ET data security, data storage cost ( $DS C$ ), and scalability ( $Scal$ ) of the scheme. In view of the above discussion,  $SV2G-ET$ 's objective function is formed as follows.

$$\mathbb{O} = \text{MAX} \left\{ \mathbb{P}, \text{Scal}, \sum_{i=0}^2 \text{Security}\{ET_i\} \right\} + \text{MIN}(DS C) \quad (5)$$

$\forall i \in \{0 \rightarrow \text{utility\_supplier}, 1 \rightarrow \text{EV\_Owner}, 2 \rightarrow \text{SG}\}$ ,

where security  $\rightarrow \{ET_i\}$  denotes that the proposed  $SV2G-ET$  scheme ensures ET data security to be maintained over EBT. Then,  $\mathbb{P}$  denotes profit to EV\_Owner,  $Scal$  denotes scalability, and  $DS C$  denotes the data storage cost over EBT. Here, limits of charging and discharging of EVs are system-dependent parameters.

### 3. The Proposed Approach

Initially, all stakeholders are registered with the EBT-based attack model through Trusted Registration Authority (TRA) to participate in the energy trading. The grid administrator generates a Unique Secrecy Code (USC) to all stakeholders such as buyer, seller, and utility supplier. Stakeholders can use USC to prove their identity and authenticity over authentication protocol of EBT using SCs. The TRA has all the information of USC of all stakeholders, which is required for registration process. The TRA verifies token for registration and provides the communication key ( $C_{key}$ ) for interaction over EBT using SCs. Figure 2 shows the complete registration procedure for buyer, seller, and utility providers along with SG administrators. In case of security attack, the attacker first requests the data access to EBT, as attacker is not registered with the system (as registration required USC) so data accessibility will be denied by the smart contracts (SCs). Here, defense layer is created by the SCs to handle the security issues in the proposed  $SV2G-ET$  scheme.

Figure 3 illustrates the workflow of the  $SV2G-ET$  scheme incorporated with deep RL (i.e., DQN) and EBT. Here, hourly energy prices are supplied to the DQN model and accordingly reward is getting calculated using action-value function and random action selection. In the proposed scheme, an action-value function is getting updated based on Bellman equation, which is as follows:

$$Q_{(j+1)}(s, a) = \mathbb{E} \left\{ r_t + \delta \max_{a_{(t+1)}} Q_j(s_{(t+1)}, a_{(t+1)}) | (s_t, a_t) \right\}, \quad (6)$$

From the above equation, we get expected return value of current state by calculating sum of reward at time  $t$  and expected reward at time  $t + 1$  by taking action  $a$  with discount factor  $\delta$ . Moreover, after converging to the optimal

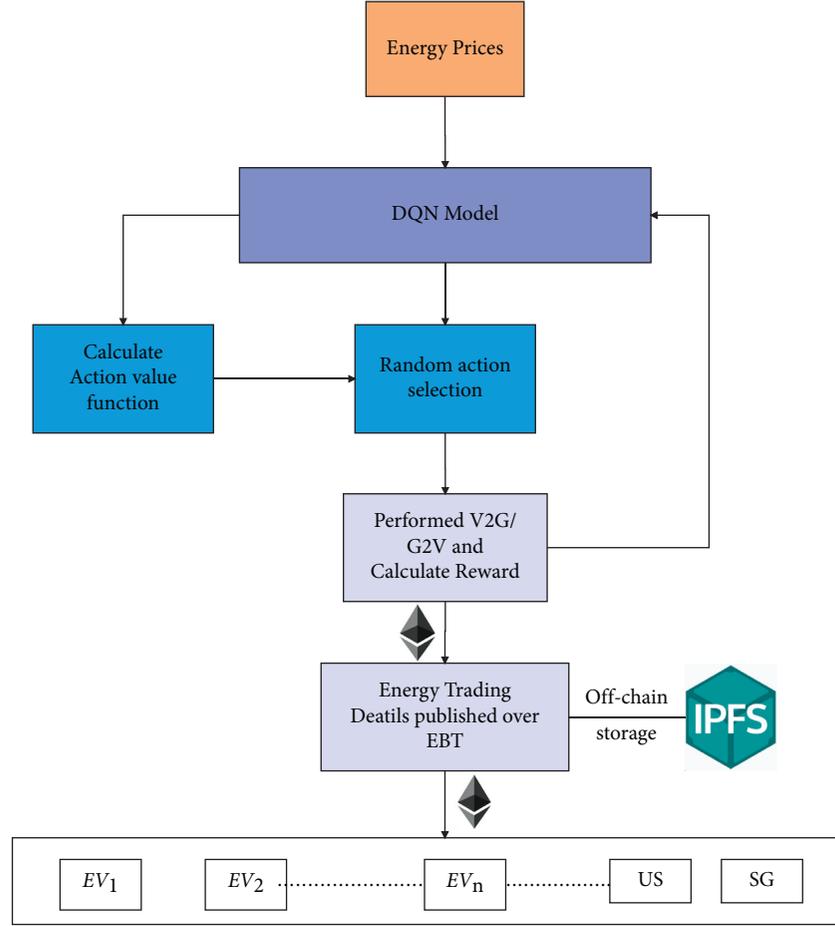


FIGURE 3: The workflow of the proposed SV2G-ET scheme.

value function, schedules are calculated using greedy approach, which is defined as follows:

$$a^* = \arg \max Q^*(s, a). \quad (7)$$

DQN employs the Q-learning, which is one of the most frequently used RL methods, and also includes CNN in order to calculate the Q-function, i.e., action-value function. Here, DQN uses the method of experience replay and with probability  $\epsilon$  random action  $a$  is getting selected, as shown in

$$a(s, a) = \max_{\rho} Q_{\rho}(s, a), \quad (8)$$

and then the loss is calculated by the loss function and update the episodes.

$$L_i(\phi_i) = E \left[ \left( r + \delta^* \max_{a'} \tilde{Q}(s', a', \tilde{\phi}_i) - Q(s, a, \phi_i) \right)^2 \right], \quad (9)$$

where  $\delta$  is the discount factor, i.e., decay rate,  $r$  shows the reward,  $\phi_i$  are network parameters of episodes  $i$  for Q-network, and it is used to compute the target. Here, we calculate loss by adding reward with our target reward and subtracting the predicted values.

The extraction of distinguish characteristics from raw data is a critical step in improving the action-value function approximation. The scheduling method can reduce charging

costs by utilizing these characteristics. Here, CNN is used with one or more layers of units between the input and output layers to feed forward. The output units represent a hyperplane in the space of the input patterns. For  $M$  number of layers, the weights are denoted as  $W^{m-1}$ . Then,  $\theta_i^m$  represents the bias,  $o_i^m$  shows the output, and  $\phi_i^m$  denotes the activation function. Then, the output of neural network is then fed into the Q-network for the estimation of approximate optimum action-value function.

In order to reduce the unwillingness of participation of the EV owner, the opinion of the owner should be known in prior. When the owner decides to not participate in the trading, the vehicle will not be able to gain or release energy; hence there will be no change in the charging state of the vehicle. At the time of peak load condition, the bidding strategy will depend on the number of different responses given by EVs owner. When the owner is unwilling to participate, those participating vehicles should not be included in the trading. As there will be maximum participation when there are maximum gains for EVs owners using the buy and sell price in the proposed scheme, such scenario has been considered in the proposed scheme. Further, to reduce the participation unwillingness in EVs energy trading (by the EVs owner), it is handled by providing incentives to the EVs owner from the grid, which is an extension of this research work.

Input: Energy price, vehicle volume, Reward  $r$   
Output: Deep Q-Network parameters.

- (1) DNN parameters  $\phi$  is initialized randomly.
- (2)  $\tilde{\phi}$ , a target parameter is initialized.
- (3) for episode = 1 to  $X$  do.
- (4) Get State;  $S_t$ .
- (5) for timestep  $t = 1$  to  $t$  do.
- (6) With  $\epsilon$  probability select action  $a$  randomly.
- (7) Select action  $a$  and reward ( $r$ ) is calculated to proceed to the next state, i.e.,  $s_t$ .
- (8) State transition details ( $s, a, s, s_t$ ) are stored.
- (9) Calculate loss function  $L_i(\phi_i) = E[(r + \delta^* \max_{\tilde{\phi}_i} \hat{Q}(s_t, a_t, \tilde{\phi}_i) - Q(s, a, \phi_i))^2]$ .
- (10) Update the parameter  $\phi$  using Gradient Descent algorithm
- (11) For each  $N$ , copy weights from  $\phi$  to  $\tilde{\phi}$
- (12) end for
- (13) end for

ALGORITHM 1: Deep Q-Network.

Algorithm 1 shows the steps for DQN parameters and Algorithm 2 denotes the steps for reward calculation for V2G/G2V operation. Once V2G/G2V operation is performed, then ET details are published over EBT and stored in an off-chain storage system using IPFS. Further, private and public keys are generated and shared among all stakeholders for real-time secure access to ET data.

In the existing EBT-based approaches, the EVs ET details are stored into the EBT, which is costly operation, i.e.,  $\approx$  around USD 76,000 per 1 GB of storage. So, a secure, distributed, and immutable IPFS mechanism could handle this ET data storage issue using its Merkle direct-acyclic graph structure. This speeds up the ET data storage process by using low bandwidth and also handles data duplication, i.e., generating a unique hash value for each ET data. The integration of IPFS mechanism with EBT improves the scalability and transaction latency of the proposed scheme.

The proposed SV2G-ET scheme comprises IPFS-based EBT, which involves only storage cost of hash-key in place of complete EVs energy trading data storage cost. Then, IPFS is a distributed storage system (i.e., open-source) that is free and immutable. IPFS receives the energy trading data tuple, i.e.,  $\{\text{EV\_Owner}_{\text{ID}}, \text{Energy}_{\text{Cost}}, t, G\_Station_{\text{id}}, \mathbb{P}, E_{\text{amt}}\}$ , from utility\_supplier and generates the hash-key of this tuple. Here,  $G\_Station_{\text{id}}$  represents the charging/discharging station of grid. IPFS splits the energy trading file into small chunks and encrypts it with a random encryption key and satisfies predefined conditions of SCs. Next, these encrypted chunks are stored over EBT. The hash-key size uses 256 bits for SHA-256 on EBT and a single tuple involves only 1-word size in SV2G-ET scheme.

## 4. Performance Evaluation

This section highlights the performance evaluation of the proposed SV2G-ET scheme with respect to the scalability, ET data storage cost, and profit to the EVs owner. The detailed explanation is as follows.

**4.1. Dataset Description.** The energy price dataset is referred to from *Nord Pool* [36]. In the proposed SV2G-ET scheme,

the hourly prices are divided into training and testing part. The first 20 days of every month consists of training datasets and the remaining 10 days of the month contains test dataset. Furthermore, the EVs data is referred to from International Energy Agency (IEA) [37], which comprises energy capacity of total 30 EVs belonging to two different categories: battery electric vehicle (BEV) and plug-in hybrid electric vehicles (PHEVs). To evaluate the performance of the proposed SV2G-ET scheme, we integrate the energy price data and EVs data. The obtained results from this evaluation may vary due to different parameters such as block size, block creation time, endorsement policy, etc.

Figure 4 illustrates the opening (marked in green color) and closing (marked in blue color) energy prices dated from the month of January to September 2021. Then, Figure 5 shows the hourly price variation throughout the day, i.e., highest energy price (marked in blue color) and lowest energy price (marked in orange color).

**4.2. Experimental Results.** In the beginning, all simulation parameters are set. Table 2 shows these parameters. It is worth emphasizing that the proposed learning-based scheme makes no hypotheses about the distributions of random variables like arrival/departure time and energy consumption while creating the dataset. The battery capacity is considered to be 30kWh for each EV. The charging actions are (0,1) for the states representing discharging and charging of the EV.

The discount factor  $\delta$  is set to 0.9 for the agent to evaluate each of its actions. The price history for the last 24 hours is set in the representation network. The output is then fed into the Q-network for the calculation of the action-value function. Here, parameters are randomly updated using the gradient descent algorithm.

After the training for 10000 episodes, the optimal scheduling for V2G/G2V for EV charging is done. In each episode, the reward is calculated and learning starts after 1000 episodes. Figure 6 shows the cumulative rewards, which are being smooth over per 100 episodes. It is observed from graph that rewards start to increase as soon as agent

Input: State  $S_t$ , energy prices, vehicle volume  
Output: EV G2V/V2G schedules

- (1) Implement the DNN parameters  $\phi$  trained by Algorithm 1.
- (2) for Timestep  $t$ : 1 to  $t$  do
- (3) Obtain past 2 days energy prices.
- (4) Network extracts the features from the data of energy prices.
- (5) Calculation of action-value  $Q(s_t, a, \phi)$  by Q network
- (6) The action  $a$  is selected randomly, otherwise  $a = \arg \max_a Q(s, a)$
- (7) Output G2V/V2G schedule.
- (8) Execute ET\_Transaction\_Details() for secure access
- (9) Publish ET details over EBT and stored in off-chain mode IPFS.
- (10) Generate private\_key and public\_key.
- (11) Distribute private\_key and public\_key to all stakeholders i.e., EV\_Owner, utility\_supplier, SG administrator.
- (12) end for return  $a_{1:t}$

ALGORITHM 2: V2G or G2V for EV.

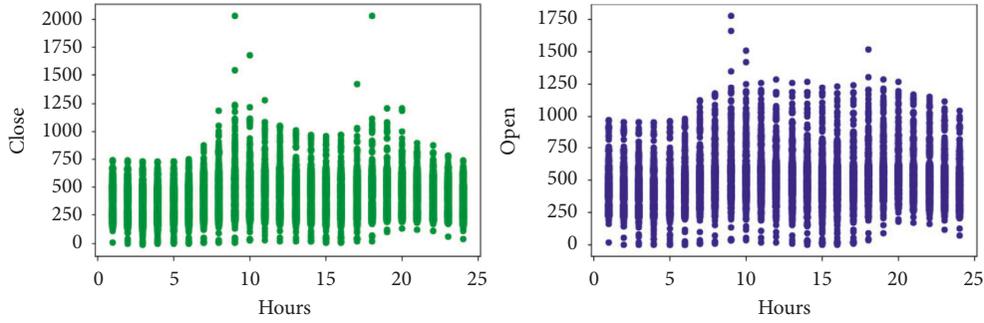


FIGURE 4: Opening and closing energy prices.

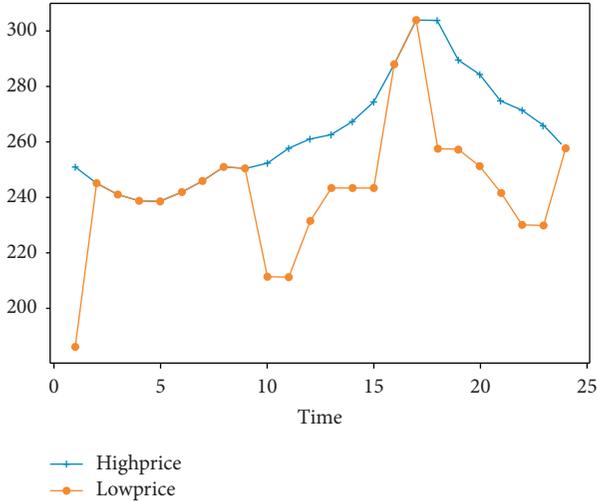


FIGURE 5: Price variation for a specific day.

starts to learn. Then, at episode 8000, the rewards gradually increase due to exploration and exploitation mechanism.

To investigate the performance of the proposed SV2G-ET scheme, ET occurs by EVs over a period of the days. Figure 7 illustrates the ET, where vertical axis shows the prices that are considered to be the mean price of the hour. Then, x-axis shows the time at which the ET takes place. Further, green and red circles show the discharge/V2G and

TABLE 2: System parameters.

Arrival time	$\mathcal{N}(17, 1^2)$
Departure time	$\mathcal{N}(8, 1^2)$
Energy consumed	$\mathcal{N}(0.5, 0.1^2)$
Learning episodes	$10^5$
Discount factor $\delta$	0.9
Beginning-exploration rate $\epsilon$	1
Exploration fraction $\epsilon$	0.1
End-exploration rate $\epsilon$	0.02
Batch size, Batch <sub>size</sub>	32
Number of transactions	1000

charging/G2V, respectively. The graph depicts that the proposed scheme is able to detect high prices based on the previous action when it makes the V2G transaction. When the energy price is low, it performs G2V transactions in the scenario. The agent gets data in real time and acts on it using a smart pattern that allows EV to get discharge at peak (near about) prices and to get charge at low prices. The activity is mainly inactive when it comes to average pricing.

In the proposed SV2G-ET scheme, the accumulative charging cost of EVs throughout 10 days (test dataset) is calculated. Here, profit  $\mathbb{P}$  indicates the number of units of currency starting with 1.0 unit.

Figure 8 shows the result for ET, where profit is on y-axis and x-axis contains days. Then, the baseline-1 approach (Mhaisen et al. [33]) uses the heuristic strategy which

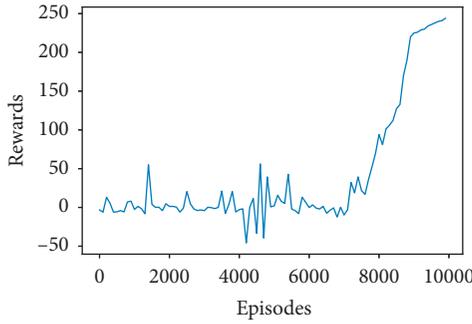


FIGURE 6: Cumulative reward.

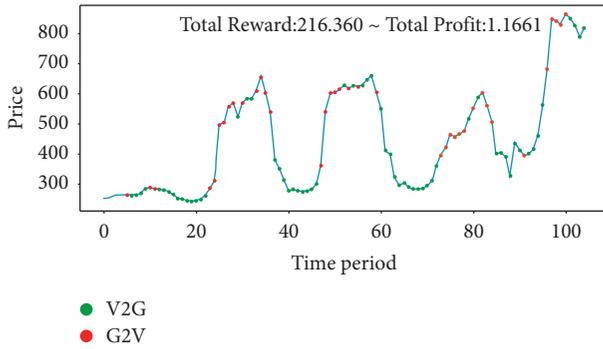


FIGURE 7: G2V/V2G operation.

represents the common strategy, where the EV gets charged fully and further its discharge until it reaches a certain threshold limit of the stored energy. In *SV2G-ET* scheme, the threshold limit is 80%. Profit indicates the number of units of currency, which is starting with 1.0 unit. For baseline-2, the optimal case is taken and considered as the future energy prices are known for maximum profit. The profit generated by the proposed *SV2G-ET* scheme results in a promising outcome compared to baseline-1. More, the optimal method always generates profit because it knows the future prices, which is not possible in a real-life scenario.

Figure 9 depicts the comparison of *SV2G-ET*'s scalability with the existing approaches based on the transaction time of EVs trading and blocks mined. In the proposed scheme, ET data is stored in an off-chain system, i.e., IPFS, and merely the hash-key of ET is sent to EBT. Here, size of hash-key is 160 bits, which is lower than the original EV's ET transaction size, which is in bytes. It facilitates adding more transactions to EBT. Therefore, *SV2G-ET* scheme offers more transactions and services to more EV owners at the same quantum of time, so improving the scalability of the system.

Figure 10 illustrates a comparative analysis of ET data storage cost in existing EBT-based approaches and the proposed *SV2G-ET* scheme, which uses IPFS mechanism for off-chain storage of data. In the existing approaches, the ET details are stored on the EBT itself, which is a costly operation. Another side of the coin is that the proposed *SV2G-ET* scheme uses a distributed and off-chain mechanism, i.e., IPFS, to store ET data that is a relatively low-cost storage system compared to the existing approaches.

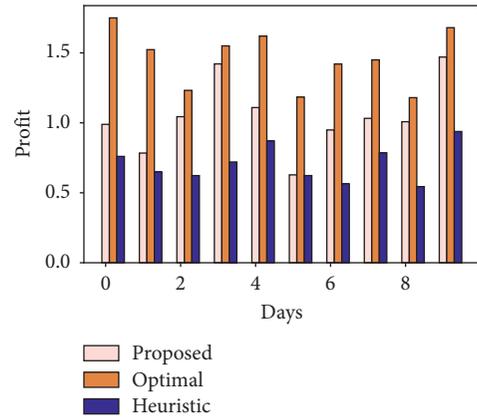


FIGURE 8: Profit generated.

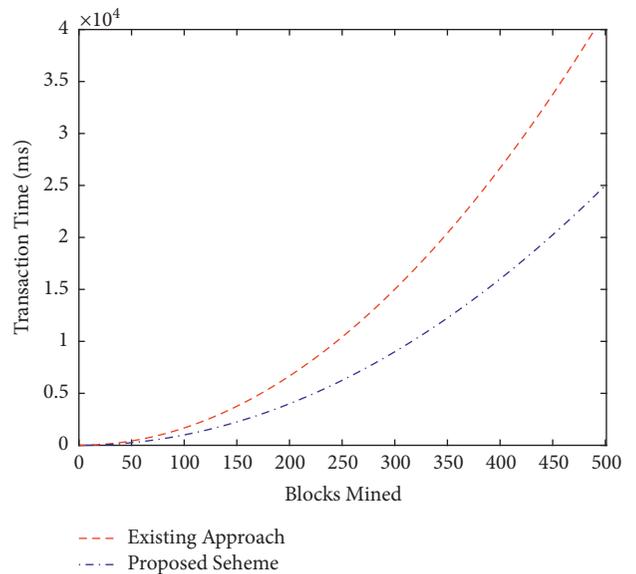


FIGURE 9: Scalability comparison.

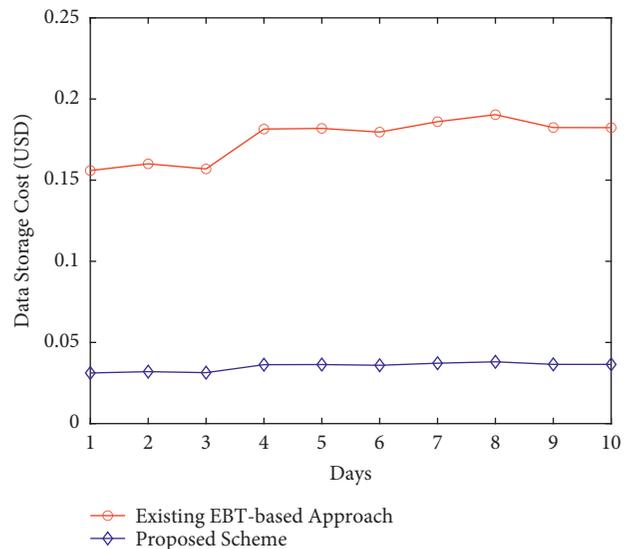


FIGURE 10: ET data storage cost comparison.

EVs are limited by range and speed, and most of these EVs range from 50 to 100 miles and are required to be recharged again [38]. So, EVs cannot be used for long journeys with single charging, which results in that V2G energy trading is required to happen within this range only. RL agent makes charging/discharging decisions mostly as per the battery level of an EV (one of the major challenges due to mobility of vehicles). Therefore, RL agent can generate feasible solutions but it will be difficult to find optimal solutions, which is an interesting direction for future research. Further, to handle the participation unwillingness in EVs energy trading (by the EVs owner) also needs attention and that is handled by providing incentives to the EVs owner from the grid, which is an extension of this research. There are various other parameters that require attention and will be an extension of this study; these parameters are communication latency, communication cost, and EV owner's privacy.

## 5. Conclusion

In this paper, EV charging/discharging is formulated using model-free RL methodology, which securely supports V2G/G2V incorporation with EBT. The randomness of energy prices and EVs commutes are considered for problem formulation to obtain an optimal mechanism for scheduling problems. This paper proposed a secure model-free scheme, i.e., SV2G-ET, which does not require environment uncertainty information. Here, a deep Q-network is employed to estimate the action-value function and securely provide accessibility on ET data among all stakeholders like SG, EV owners, and utility suppliers using EBT. Then, the performance of the proposed SV2G-ET scheme is evaluated by displaying the charging/discharging trend versus real-time pricing. The promising result shows the effectiveness of the SV2G-ET scheme, which is able to learn useful charging and discharging patterns of EVs during off-peak/peak hours. Further, it calculated the profit generated by the EVs owner over 10 days and showed a better result than the optimal scenario and heuristic approach.

In the future, the communication cost, EV owner's privacy, and latency will be an extension of this research work while accessing the EVs' ET data in real time using blockchain technology.

## Data Availability

No data are associated with this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] B. Mroczek and A. Kołodyńska, "The v2g process with the predictive model," *IEEE Access*, vol. 8, pp. 86947–86956, 2020.
- [2] W. Tang, S. Bi, and Y. J. Zhang, "Online coordinated charging decision algorithm for electric vehicles without future information," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2810–2824, 2014.
- [3] L. Yao, W. H. Lim, and T. S. Tsai, "A real-time charging scheme for demand response in electric vehicle parking station," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 52–62, 2017.
- [4] S. Wang, S. Bi, Y.-J. A. Zhang, and J. Huang, "Electrical vehicle charging station profit maximization: admission, pricing, and online scheduling," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1722–1731, 2018.
- [5] M. A. Ortega-Vazquez, "Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty," *IET Generation Transmission & Distribution*, vol. 8, pp. 1007–1016, 2014.
- [6] D. Wu, H. Zeng, C. Lu, and B. Boulet, "Two-stage energy management for office buildings with workplace ev charging and renewable energy," *IEEE Transactions on Transportation Electrification*, vol. 3, no. 1, pp. 225–237, 2017.
- [7] P. K. Mohanty, P. Jena, and N. P. Padhy, "Home electric vehicle charge scheduling using machine learning technique," in *Proceedings of the 2020 IEEE International Conference on Power Systems Technology (POWERCON)*, Bangalore, India, September 2020.
- [8] D. Aguilar-Dominguez, J. Ejeh, A. D. F. Dunbar, and S. F. Brown, "Machine learning approach for electric vehicle availability forecast to provide vehicle-to-home services," *Energy Reports*, vol. 7, pp. 71–80, 2021.
- [9] R. Shipman, R. Roberts, J. Waldron et al., "We got the power: predicting available capacity for vehicle-to-grid services using a deep recurrent neural network," *Energy*, vol. 221, Article ID 119813, 2021.
- [10] S. Das, P. Acharjee, and A. Bhattacharya, "Charging scheduling of electric vehicle incorporating grid-to-vehicle (g2v) and vehicle-to-grid (v2g) technology in smart-grid," in *Proceedings of the 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*, Cochin, India, January 2020.
- [11] Y. Tao, J. Qiu, S. Lai, X. Zhang, Y. Wang, and G. Wang, "A human-machine reinforcement learning method for cooperative energy management," *IEEE Transactions on Industrial Informatics*, vol. 18, 2021.
- [12] D. Qiu, Y. Ye, D. Papadaskalopoulos, and G. Strbac, "A deep reinforcement learning method for pricing electric vehicles with discrete charging levels," *IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5901–5912, 2020.
- [13] E. Mocanu, D. C. Mocanu, P. H. Nguyen et al., "On-line building energy optimization using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3698–3708, 2019.
- [14] A. Kumari and S. Tanwar, "Reinforcement learning for multiagent-based residential energy management system," in *Proceedings of the 2021 IEEE Globecom Workshops (GC Wkshps)*, Madrid, Spain, December 2021.
- [15] S. Najafi, M. Shafie-khah, P. Siano, W. Wei, and J. P. Catalão, "Reinforcement learning method for plug-in electric vehicle bidding," *IET Smart Grid*, vol. 2, no. 4, pp. 529–536, 2019.
- [16] X. Sun and J. Qiu, "A customized voltage control strategy for electric vehicles in distribution networks with reinforcement learning method," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 6852–6863, 2021.
- [17] H. Wei, Y. Zhang, Y. Wang, W. Hua, R. Jing, and Y. Zhou, "Planning integrated energy systems coupling v2g as a flexible storage," *Energy*, vol. 239, Article ID 122215, 2022.

- [18] Q. Dang, D. Wu, and B. Boulet, "A q-learning based charging scheduling scheme for electric vehicles," in *Proceedings of the 2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, Detroit, MI, USA, June 2019.
- [19] A. Chiş, J. Lundén, and V. Koivunen, "Reinforcement learning-based plug-in electric vehicle charging with forecasted price," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 3674–3684, 2017.
- [20] H. Ko, S. Pack, and V. C. M. Leung, "Mobility-aware vehicle-to-grid control algorithm in microgrids," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 7, pp. 2165–2174, 2018.
- [21] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, "Verification and validation techniques for streaming big data analytics in internet of things environment," *IET Networks*, vol. 8, no. 3, pp. 155–163, 2019.
- [22] S. Wang, S. Bi, and Y. A. Zhang, "Reinforcement learning for real-time pricing and scheduling control in ev charging stations," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 849–859, 2021.
- [23] A. Perera and P. Kamalaruban, "Applications of reinforcement learning in energy systems," *Renewable and Sustainable Energy Reviews*, vol. 137, Article ID 110618, 2021.
- [24] R. Khalid, N. Javaid, S. Javaid, M. Imran, and N. Naseer, "A blockchain-based decentralized energy management in a p2p trading system," in *Proceedings of the ICC 2020—2020 IEEE International Conference on Communications (ICC)*, Dublin, Ireland, June 2020.
- [25] M. Baza, A. Sherif, M. M. E. A. Mahmoud et al., "Privacy-preserving blockchain-based energy trading schemes for electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 9, pp. 9369–9384, 2021.
- [26] S. Tanwar, Q. Bhatia, P. Patel, A. Kumari, P. K. Singh, and W.-C. Hong, "Machine learning adoption in blockchain-based smart applications: the challenges, and a way forward," *IEEE Access*, vol. 8, pp. 474–488, 2020.
- [27] T. AlSkaif, J. L. Crespo-Vazquez, M. Sekuloski, G. van Leeuwen, and J. P. S. Catalão, "Blockchain-based fully peer-to-peer energy trading strategies for residential energy systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 231–241, 2022.
- [28] A. Kumari, R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "When blockchain meets smart grid: secure energy trading in demand response management," *IEEE Network*, vol. 34, no. 5, pp. 299–305, 2020.
- [29] A. Kumari and S. Tanwar, "A reinforcement-learning-based secure demand response scheme for smart grid system," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 2180–2191, 2022.
- [30] A. Muzumdar, C. Modi, G. M. Madhu, and C. Vyjayanthi, "A trustworthy and incentivized smart grid energy trading framework using distributed ledger and smart contracts," *Journal of Network and Computer Applications*, vol. 183–184, Article ID 103074, 2021.
- [31] A. Kumari, A. Shukla, R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "Et-Deal: A p2p smart contract-based secure energy trading scheme for smart grid systems," in *Proceedings of the IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Toronto, ON, Canada, July 2020.
- [32] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time ev charging scheduling based on deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5246–5257, 2019.
- [33] N. Mhaisen, N. Fetais, and A. Massoud, "Real-time scheduling for electric vehicles charging/discharging using reinforcement learning," in *Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, February 2020.
- [34] B. Lin, B. Ghaddar, and J. Nathwani, "Deep reinforcement learning for the electric vehicle routing problem with time windows," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2021.
- [35] S. Ding, J. Zeng, Z. Hu, and Y. Yang, "A peer-2-peer management and secure policy of the energy internet in smart microgrids," *IEEE Transactions on Industrial Informatics*, p. 1, 2021.
- [36] "Nordpool, historical market data," 2021, <https://www.nordpoolgroup.com/historical-market-data/>.
- [37] "Electric vehicles," 2021, <https://www.iea.org/articles/global-ev-data-explorer>.
- [38] "Conserve energy future," 2022, <https://www.conserve-energy-future.com/advantages-and-disadvantages-of-electric-cars.php>.