Local Demand Management of Charging Stations using Vehicle-to-Vehicle Service: A Welfare Maximization-Based Soft Actor-Critic Model

Akhtar Hussain[1, *], Van-Hai Bui[2], and Petr Musilek[1,3]

[1] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2G2, Canada.

[2] Department of Electrical Engineering at SUNY Maritime College, NY 10465, USA.

[3] Department of Applied Cybernetics, University of Hradec Králové, 500 03 Hradec Králové, Czech Republic.

Abstract- Transportation electrification has the potential to reduce carbon emissions from the transport sector. However, the increased penetration of electric vehicles (EVs) can potentially overload the distribution systems. This becomes prominent in locations with multiple EV chargers and charging stations with many EVs. Therefore, this study proposes a welfare maximization-based soft actor critic (SAC) model to mitigate transformer overload in distribution systems due to the high penetration of EVs. The demand of each charging station is managed locally to avoid network overload during peak load hours in two steps. First, a welfare maximization-based optimization model is developed to maximize the welfare of electric vehicle owners by performing vehicle-to-vehicle(V2V) service. In this step, the sensitivity of EV owners to different parameters (energy level, battery degradation, and incentives provided by fleet operators) is considered. Then, a deep reinforcement learning-based method (soft-actor critic) is trained by incorporating the welfare value (obtained from the welfare maximization model) in the reward function. The total power demand (at the transformer level) and transformer capacity are also included in the reward function. The agent (fleet operator) learns the optimal pricing strategy for local demand management of EVs by interacting with the environment. Each electric vehicle responds to the action (price) by deciding the amount of power they are willing to charge/discharge (V2V) during that interval. Training is performed offline, and the trained model can be used for real-time demand management of different types of charging stations. The simulation results have shown that the proposed method can successfully manage the demand of different charging stations, via V2V, without violating the transformer capacity limits.

Keywords- Electric vehicles, deep reinforcement learning, demand management, soft actor-critic, vehicle-to-vehicle (V2V), welfare maximization.

---

[*]Corresponding author: akhtar3@ualberta.ca (A. Hussain).

1. Introduction

The electrification of transportation is considered a viable option for reducing greenhouse gas emissions from the transport sector by reducing the utilization of fossil fuels. However, the increased penetration of electric vehicles (EVs) comes with several challenges for both the power and transport sectors. For example, at the power system level, more power plants will be required to supply the power to EVs and to serve as a reserve [1]. In addition, it may overload local equipment and cause several technical issues (voltage drop, network congestion, phase imbalances, etc.) at the distribution level [2]. Similarly, due to the use of power electronics in the charging infrastructure, a higher penetration of EVs can cause several power quality issues in power systems [3]. For example, harmonics can cause thermal overloading of transformers, and voltage deviations can cause network instability. Distribution systems are the immediate victims of the exacerbated loads due to the direct connection of EVs with them. Coordination between the power and transport sectors is required during the planning and operation phases to mitigate the above-mentioned issues [4]. Several studies have been conducted to reduce the congestion of the distribution system under the high penetration of EVs.

Existing studies on the congestion management of distribution systems can be broadly categorized into two groups. In the first group, system-level mechanisms are used to manage EV loads in response to market price signals and incentives, i.e., demand response. For example, day-ahead congestion signals are used in [5]–[7] to communicate tariffs to fleet operators (FO). In [5], the main objective is to minimize cost while reducing network congestion, and EVs are imposed a higher charging cost (penalty) in [6] if they do not switch to lower-cost slots. Finally, three main actors (EV owner, FO, and distribution system) are considered in [7]. Other studies use hierarchical methods: in [8], market- and price-based controls are implemented at each level, while in [9] the economic dispatch of the distribution system is carried out in the upper layer, while the lower layer is responsible for the aggregator energy management. In addition, various studies have considered vehicle-to-grid (V2G) services to relieve the grid during the congestion period. For example, reactive power injection is considered in [10] to mitigate voltage drops, and V2G is implemented considering the energy level of EVs in [11]. In [12], different indicators are defined and used to measure and mitigate the voltage and congestion impacts of EVs. Dynamic pricing for spatial load shifting is proposed in [13] by incentivizing EVs, and incentive-based charging control for EVs with sufficient energy is proposed in [14] to mitigate network congestion due to EVs.

Most of these approaches are based on different tariff structures, as noted in [15]. Different prices during different times of the day can reduce the EV charging price. It has been demonstrated in [16] that changing the price tariff based on carbon intensity can significantly reduce the EV charging price compared to a flat tariff. However, only time-of-use tariffs have been shown to be inefficient for the demand management of EVs, as they merely shift the peak loads [17]. In addition, it has been demonstrated that the average EV load is more significant at the local transformer level than at the network level [18]. Therefore, local demand management is required, and consequently, the second group (local demand management) of studies has been conducted, as explained below.

To reduce the peak load of the distribution transformers, demand-side management of EVs is proposed in [19] considering the weak buses (voltage stability) of the distribution system. In [20], a least laxity first approach is proposed to divide the available system capacity among the EVs requiring recharge during system congestion. In [21], a game model is proposed to minimize the charging cost of each EV while respecting transformer capacity constraints. Deep reinforcement learning (DRL)-based methods are also used by various researchers for the demand management of EVs. For example, the transformer level information is used in [22] to compute the network level information in a distributed way. Similarly, EV clustering is considered in [23] and real-time scheduling of EVs is proposed to avoid congestion during peak times. Soft actor-critic (SAC) models are used in [24] and [25] to manage the load of EVs in distribution systems. In [24], both a collective policy mode and an independent learner mode are proposed for each EV, while SAC is combined with a nodal multi-target model in [25] to reduce the dimensionality of the neural network. In [26], a bi-level scheduling problem is proposed, where a constrained DRL method is used in the upper layer to obtain the net EV demand, while an optimal descending order charging policy is used in the lower layer to promptly determine the behavior of EVs.

However, none of these studies have considered vehicle-to-vehicle (V2V) services to mitigate transformer overload. With the increase in useable battery sizes and mileage efficiencies of EVs, V2V services are becoming feasible [27]. EVs that have an excess of energy during system congestion can provide power to EVs that require recharge via V2V. In addition to congestion mitigation, it would be beneficial in several ways. For example, deferral of distribution system upgrade, additional revenues for EV owners participating in V2V, and the elimination of the requirement of network configuration information. In addition, the sensitivities of individual EV owners to the state-of-charge (SoC) level, battery degradation, and incentives offered by fleet managers could be different. The inclusion of these factors requires a huge amount of data, since these factors are subject to human behavior/preferences and the type of EV (useable battery size and energy efficiency). These factors are also not extensively studied in the existing literature. Finally, some studies have considered training neural networks for individual EVs, which results in 1) reusability issues due to diversified travel patterns, EV types, and preferences of the EV owner. 2) computational complexity due to the requirement to train one neural network for each EV, especially with higher EV penetration levels.

Meanwhile, to mitigate equipment overloading, a collective limit is imposed on the net load of all EVs registered with the same charging station due to the provision of power to all EVs by the same transformer. For example, a contracted capacity limit is imposed on the charging station in [28] and then charging/discharging priorities are determined to limit the energy consumption of EVs below the contracted capacity limit in that charging station. The power dispatch strategies are then determined to minimize the EV charging cost. In these types of problems, the charging/discharging decisions of EVs are influenced by the decision of the FO and vice versa. Similarly, the charging/discharging decision of one EV is also influenced by the charging/discharging decision of other EVs due to the imposition of the limit on the total charging station load. These types of problems can be readily solved using game-theoretic approaches. However, with game theory alone, the charging station demand management problem cannot be solved for real-time applications due to the requirement of parameter updating after the arrival/departure of each EV. In addition, exact sensitivity (SoC level and battery degradation) information of EV is required, which is not readily available. This problem can be solved by combining the welfare maximization and DRL models. A welfare maximization-based model will capture the interdependence in the actions of different actors (EVs, FO, and distribution system) in the network. The DRL model will eliminate the need for information exchange after arrival/departure of each EV. Therefore, the trained model can be used to manage the load of charging stations in real time.

To utilize the merits of both welfare maximization (core component a game model) and DRL, this article proposes a welfare maximization-based SAC algorithm to locally manage the load of EVs. The proposed model is developed and trained using the following steps. First, a welfare maximization-based optimization model is developed using the sensitivity of individual EVs to SoC and battery degradation along with the price determined by the FO. Then, a reward function is devised using the welfare values of the EVs and the parameters of the distribution system. These parameters include the capacity of the distribution transformer and the electric load of the distribution system (at the transformer level). Finally, the price signals are generated within the regulated bounds. The agent interacts with the environment during numerous episodes to learn an optimal pricing strategy for local charging demand management of EVs. During each episode, the EV parameters are generated within the specified bounds, and the welfare maximization-based optimization model is executed. The net welfare and prices used for each episode are stored and used to update the reward function. Due to the approximation ability of deep neural networks, the trained model can be used to locally manage the load of different types of charging stations (residential, commercial, industrial, mixed, etc.). Training is performed offline, and the trained model can determine the charging/discharging demand of EVs in real time. Simulation results have shown that the proposed method can manage the load of different types of charging station while respecting the capacity limits of distribution transformers.

The remainder of the paper is organized as follows. The introduction section is followed by the local charging demand management section (section 2), where the charging station demand management problem is formulated as a welfare maximization model. In section 3, the SAC model is developed, and the training process is explained. In section 4, the performance of the SAC model is compared with other deep reinforcement learning methods. In addition, the performance of the proposed method is analyzed for different days and charging stations. The network level impact analysis (impact on transformer overloading) and comparison with game-model is carried out in section 5 which is followed by conclusions.

2. Local Charging Demand Management

2.1 Local Charging Demand Management System

The local demand management system proposed in this study comprises two parts, as shown in Fig. 1. The first part is the power distribution system and the second part is the charging station with the EV fleet. The FO is responsible for managing the charging demand of the charging station. The FO requires information about the transformer rating and the electric load of the distribution system. This information is used along with the EV fleet information to manage the charging station load locally. The FO contains the trained SAC model developed in this study (discussed in the next section). It also has information about EVs registered with the charging station in the database. The FO needs to communicate with the EVs and with the distribution system operators; therefore, it includes a communication module.

The upstream grid operators (distribution system operators) only interact with the FO. The FO is responsible for managing the load of the charging station by adjusting the EV charging/discharging price for each interval, considering transformer loading conditions and net EV load. Then, each EV reacts to the price set by the FO during different intervals by deciding the amount of power they are willing to charge/discharge during that interval. Therefore, the FO has a price regulation module with information about maximum and minimum price bounds, which are used to determine the charging/discharging price for different time intervals of the day. The EVs that have decided to charge will pay the per unit price, and

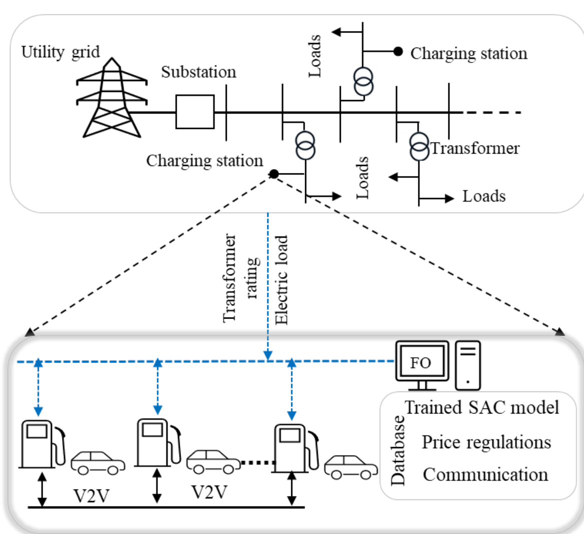

Fig. 1 Configuration of the proposed local charging demand management system.

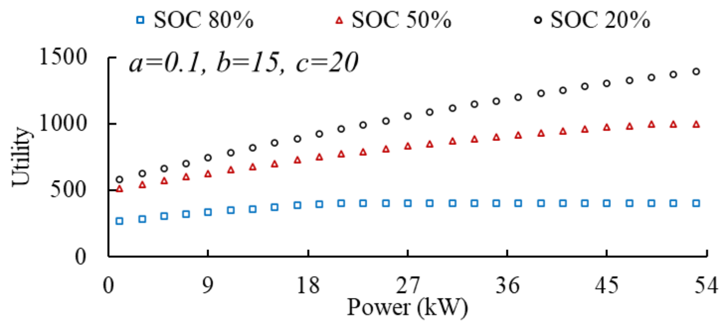

Fig. 2 Overview of the proposed quadratic utility function under different SoC levels.

the EVs that have decided to discharge will receive the per-unit price as an incentive for each kWh of energy they discharge. This is because the EVs in the charging station are capable of providing V2V services. Participation of EVs in V2V (EV discharging their power) will increase their revenues and minimize the load on the power system during system congestion intervals.

2.2 Welfare Maximization and Charging Demand Management

To manage the charging demand of EVs locally, coordination is required between FO and EV owners. The role of FO is to devise a charging mechanism considering the net charging demand of EVs and the capacity of the transformer. Similarly, the role of EV owners is to decide the amount of power they are willing to charge or discharge based on the price set by the FO. However, the decision of FO is influenced by the decisions of the EV owners and vice versa. Similarly, the decision (charging/discharging) of one EV is impacted by the decision of other EVs [27]. Therefore, in this study, a welfare maximization-based model is used to generate EV-related data (net load, total utility and total welfare), which is then used to train the neural network (discussed in the subsequent section).

2.3 Utility and Welfare Modeling

A utility/welfare function that encompasses all influencing factors is considered the core of welfare maximization problems. The utility function represents the level of satisfaction against the amount of power consumed by EVs. The utility function needs to fulfill certain properties, as outlined in [29]:

1. Non-decreasing monotonicity: The utility function should increase with an increase in power consumption until it reaches saturation.
2. Concavity: The utility function should be concave to ensure that the marginal benefit of consumers is a non-increasing function, and the utility gradually reaches saturation.

These properties are fulfilled by several utility functions such as logarithmic, exponential, and quadratic functions. However, quadratic functions are convex and trackable [30], which makes them suitable for solving using commercial optimization software. Therefore, a quadratic utility function is developed in this study. An overview of the proposed utility function for different SoC levels of an EV is presented in Fig. 2. It can be observed that the utility of EVs with different levels of SoCs increases with an increase in power consumption. Similarly, the utility of an EV with 80% SoC saturates around 20kW, and that of an EV with 50% SoC saturates around 50kW. Due to these desired characteristics, the following quadratic utility function ($U_{v,t}$) is designed for each EV (indexed by *v*) during time interval *t*

$$U_{v,t} = -a \cdot (e_{v,t} - \Delta t \cdot p_{v,t})^2 + b \cdot (e_{v,t} - \Delta t \cdot p_{v,t}) + c \cdot \Delta t \cdot p_{v,t} \cdot \Delta t \cdot p_{v,t-1}, \tag{1}$$

$$\text{where}\, e_{v,t} = B_v^{\max} \cdot (1 - SoC_{v,t});\ a, b, c > 0\ .$$

This function represents the utility of the v^{th} EV for charging (positive) or discharging (negative) a power amount $p_{v,t}$ for a duration of Δt hours. It should be noted that Δt is time duration in hours while *t* is time step and it could have any time units (hours, minutes, seconds, etc.). The first two terms consider the sensitivity to the SoC level by considering the available energy factor ($e_{v,t}$), where *a* and *b* represent the sensitivity of the v^{th} EV. The last term represents the sensitivity of the EV to cyclic degradation considering the amount of power charged/discharged during the previous ($p_{v,t-1}$) and current ($p_{v,t}$) intervals and *c* depict the sensitivity level. It discourages frequent charging/discharging cycles as it is reported to cause battery degradation [31]. From the SoC factor, it is evident that the utility of EVs is proportional to $(1 - SoC_{v,t})$. This means that for the same amount of power, EVs with a higher SoC will get lower utility compared to EVs with a lower SoC (Fig. 2). Finally, $B_v^{\max}$ represents the usable battery size of the v^{th} EV in kWh. Parameters *a*, *b*, and *c* are nonnegative, and their units are utils/kWh^2, utils/kWh, and utils/kWh^2, respectively. The parameter value ranges in the utility function are defined considering different factors such as the EV battery size, the required saturation level, and the depth of discharge. The ranges for parameters *a, b,* and *c* are determined to be [0.05, 0.5], [5, 15], and [15, 30], respectively. The utility function of EVs can be used along with the price of FO (PR_t) to determine the welfare ($W_{v,t}$) of the EVs, as follows

$$W_{v,t} = U_{v,t} - PR_t \cdot p_{v,t} \cdot \Delta t \cdot \rho\ . \tag{2}$$

It should be noted that PR_t is the per-unit electricity price (¢/kWh) set by the FO for charging EVs during interval *t*. This price is determined by the FO considering the electricity price of the upstream grid, remaining capacity of the transformer, and the total EV load. The parameter ρ is the utility factor for the energy consumption and its units are utils/cent. Therefore, the unit of welfare will be utils (same as those of the utility). The price of the FO is regulated between an upper ($PR^{\max}$) and a lower ($PR^{\min}$) bound, i.e.

$$PR_t \in \left[PR^{\max}, PR^{\min} \right]. \tag{3}$$

2.4 Welfare Maximization Model

The objective of the model is to maximize the welfare of the EV fleet, as given by the following equation

$$\max_{P_{v,t}} \sum_{v\in V_t} W_{v,t} \ . \tag{4}$$

Several constraints need to be considered for the welfare maximization problem. For example, EVs cannot receive more power than the required amount ($e_{v,t}^{req}$), as below

$$p_{v,t} \cdot \Delta t \le e_{v,t}^{req} \ . \tag{5}$$

Generally, overcharging and deep discharging are prohibited to elongate the life span of batteries. Therefore, personalized SoC limits are imposed on each EV. The upper bound for charging ($SoC_v^{\max}$), set by the EV owner, as given by

$$p_{v,t} \cdot \Delta t \le \left(SoC_v^{\max} - SoC_{v,t}\right) \cdot B_v^{\max} \ , \tag{6}$$

where $SoC_{v,t}$ is the current SoC level of v^{th} EV. Conversely, EVs cannot be discharged below the lower SoC bound ($SoC_v^{\min}$) set by the EV owner, as given by the following expression

$$p_{v,t} \cdot \Delta t \ge -\left(SoC_{v,t} - SoC_v^{\min}\right) \cdot B_v^{\max} \ . \tag{7}$$

Finally, the net EV load should be lower than or equal to the available capacity (Tx_t^{ac}) of the transformer

$$\sum_{v\in V} p_{v,t} \le Tx_t^{ac} \ . \tag{8}$$

The available capacity refers to the remaining capacity of the transformer after serving other electric loads (excluding the EV load).

3. Deep Reinforcement Learning and Local Demand Management

The welfare maximization model developed in the previous section can be used to solve the local demand management problem in a centralized manner [32]. However, the computational burden would increase with an increase in the EV fleet size, making the determination of the optimal charging/discharging demand of EVs in real-time challenging. Therefore, decentralized solutions have been proposed for similar problems [27], [33]. However, these solutions are iterative, and their convergence time increases significantly with an increase in the number of EVs. In addition, the information exchange process must be repeated every time a new EV arrives or departs. These constraints make such approaches difficult to apply for real-time or near-real-time demand management of EVs.

To overcome these limitations, a DRL-based model is combined with welfare maximization model in this study. The presented model emulates the behavior of EVs by randomly generating sensitivity parameters in a predetermined range for a large number of EVs. The welfare maximization model is incorporated into the reward function of the neural network and trained based on the generated data. Neural networks are known to be good approximators and thus can be used for different sizes of EV fleets and different types of charging stations (residential, commercial/industrial, and mixed). Training is an offline process, and the trained model can be used to manage the load of EVs in real-time.

3.1 Demand Management Modeling using Soft Actor-Critic

SAC has gained popularity over other DRL methods due to its ability to learn quickly, its immunity to trapping in local optima, and its stable operation [34]. These traits are achieved by adding an entropy term to the objective function of the SAC, which measures the predictability of the random variable [35]. In addition, a policy network is used to restrict the agent from choosing the same action repeatedly. The local demand management of charging stations also involves several uncertainties such as the arrival/departure time of EVs, the SoC of EVs, the sensitivity of EV owner to battery degradation and SoC, and the upstream electricity price. Therefore, in this study, SAC is used to solve the EV local demand management problem.

The learning agent learns two Q-networks ($Q_{\Phi 1}, Q_{\Phi 2}$) and a policy network (π_0). The use of only one Q-network results in a positive bias in the policy improvement step. Different studies have shown that this bias degrades the performance of value-based methods [36]. Therefore, in this study, two Q-networks are employed. Additionally, the use of two Q-networks significantly reduces training time, particularly for larger tasks [37]. Both Q-networks are trained independently, and the Q-function with a lower value is utilized for the stochastic gradient and the policy gradient. The soft copy method is employed to update the two target networks at each training step, thereby stabilizing the learning process. Finally, the learning agent tries to maximize both the expected future rewards and the future entropy. This way, local trapping can be avoided while the learning process is accelerated.

A comprehensive reward function is formulated to train the SAC model to maximize the utility of EVs while respecting the capacity constraints of the transformers. The reward function is computed for each episode. It should be noted that an episode refers to one complete cycle over the scheduling horizon (24-time intervals in this study). In addition, the number of EVs present in the charging station during different time intervals could be different. Therefore, the total number of EVs is a function of time (V_t). The reward function for episode k (R_k) contains two terms, as shown below

$$R_k = \sum_{t \in T} \left(\sum_{v \in V_t} U_{v,t,k} - \max\left\{ \left(P_{t,k} - Tx_{t,k}^{ac} \right) \cdot C, 0 \right\} \right), \tag{9}$$

$$\text{where } P_{t,k} = \sum_{v \in V_t} p_{v,t,k} .$$

The first term ($\sum_{v \in V_t} U_{v,k}$) shows the total utility gained by all EVs during episode k. A positive sign with this term indicates that the objective is to maximize the utility of all EVs. The second term is introduced to penalize any positive deviations ($P_{t,k} > Tx_{t,k}^{ac}$). The parameter C represents the penalty cost, and its units are utils/kW. The objective of this formulation is to ensure that the net EV load is not greater than the available transformer capacity. However, it could be lower than the available capacity. A negative sign with the last term indicates that the objective is to minimize the value of this term.

3.2 Training and Testing of Soft Actor-Critic Model

During each training episode, several steps are involved. First, EV parameters are randomly generated within predefined ranges (discussed in Section 2). Next, the welfare maximization model is executed using these EV parameters to obtain EV values. These EV data values are then used as input by the SAC model. Based on these values, the model takes an action and transitions to the next state. This process is repeated until the end of all episodes. The training model consists of the following four major steps:

a. **Initialization of input parameters**: These parameters include the total number of episodes, the maximum number of EVs, and the remaining transformer capacity.
b. **Random generation of EV parameters**: These parameters include the sensitivity to SoC level (*a* and *b*), degradation sensitivity (*c*), and the SoC level itself.
c. **Obtaining EV values**: This step involves calculating the utility, welfare, and net EV load. The welfare optimization algorithm is executed during this step.
d. **SAC execution**: This step includes taking a step in the environment using the EV values, computing a reward based on the action, and updating the new state.

Steps b to d are repeated until the end of the total number of episodes defined in step a.

During the testing phase, the trained SAC model receives information on the total number of EVs at time *t* and the remaining capacity of the transformer. The model generates EV parameters considering the upper and lower bounds for each parameter stored in the FO database. The SAC model then provides information about the amount of power each EV is willing to charge/discharge during the time interval *t*.

The step-by-step process for the training and testing of the proposed SAC-based model for local demand management of EVs is shown in Fig. 3. First, the number of episodes, interval number, and maximum number of EVs are initialized. In addition, the maximum and minimum number of EVs (V_t^{max}, V_t^{min}), the available capacity of transformers ($Tx^{\max}$, $Tx^{\min}$), and the EV load (p^{max}, p^{min}) are defined. These values are used to normalize the states in the environment. Then, during each episode, the EV parameters are set using Algorithm I.

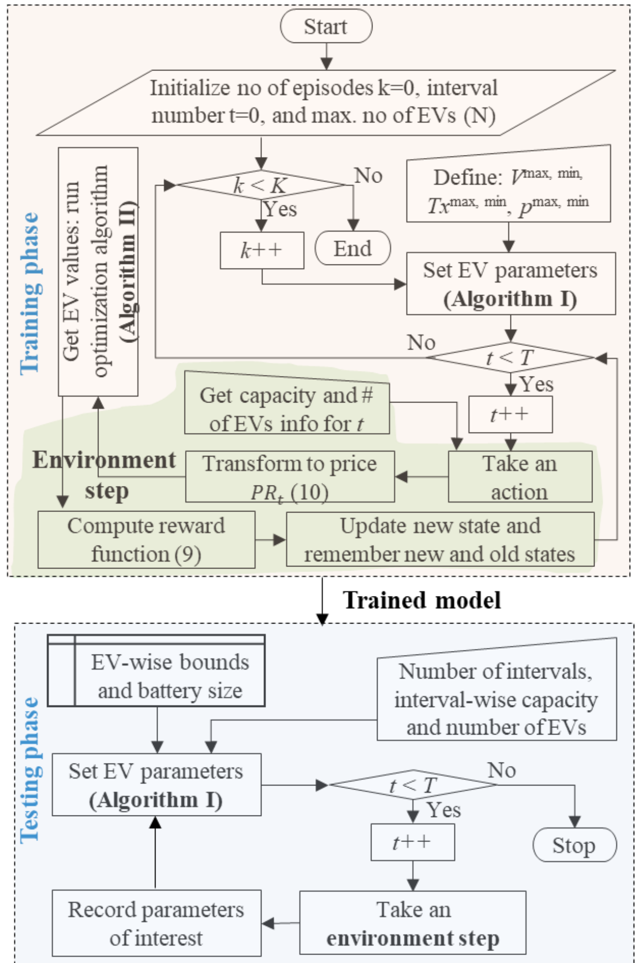

Fig. 3 Training and testing of the proposed SAC-based model for local demand management of EVs.

Algorithm I shows that during each episode k, the useable battery size information is defined for all available EV models. Then, for each time interval t, a random number is generated between the maximum and minimum EV values. An EV fleet is formulated by choosing an EV from the available EVs and allocating an SoC level to it. Then, for each EV v, daily mileage is calculated considering a log-normal distribution with a mean of μ km and a standard deviation of σ km. The amount of energy consumed by each EV is computed using the energy efficiency of that EV (η_v) and the traveled distance. Then, the required energy for each EV is determined. Finally, the sensitivity parameters (as formulated in the utility function) are randomly generated between the predefined bounds. All this information is returned to the main function.

Algorithm I Setting EV parameters for episode *k*.

```
Define useable battery size (B_v^max) for available EV models
for all t ∈ T
  Get the total number of EVs for t: V_t = randint(V_t^max, V_t^min)
  for all v ∈ V_t do % EV fleet formation
    Choose available EV: B_{v,t}^max = B_v^max(randint(len(B_v^max)))
    Initialize current SoC: SoC_{v,t} = rand(SoC_v^max, SoC_v^min)
    Compute daily mileage: d_{v,t} = lognrnd(μ, σ)
    Compute the amount of energy consumed: e_{v,t}^con = d_{v,t} * η_v
    Compute the amount of energy required: e_{v,t}^req = SoC_{v,t} − e_{v,t}^con / B_{v,t}^max
    Define sensitivity to SoC level: a_v = rand(a_v^max, a_v^min) and
    b_v = rand(b_v^max, b_v^min)
    Define sensitivity to degradation: c_v = rand(c_v^max, c_v^min)
  end for
end for
return vectors: V_t, SoC_{v,t}, e_{v,t}^req, B_v^max, a_v, b_v, c_v
```

Based on the information on the available capacity of the transformer and the total number of EVs, the agent takes an action. The action is then transformed into the price, based on the following equation

$$PR_t = \left(PR^u + PR^{\min}\right) + act \cdot PR^u \, . \tag{10}$$

where $act \in [-1,1]$, and the upper price limit (PR^u) fulfills the following condition

$$PR^u = \frac{PR^{\max} - PR^{\min}}{2} \, . \tag{11}$$

The parameters $PR^{\max}, PR^{\min}$ are the regulated bounds of the price, as defined in (3). After obtaining the price value, the optimization algorithm (4)-(8) is executed. The detailed process is shown in Algorithm II.

After getting the EV parameters from Algorithm I, the net EV demand is initialized to zero. Then, the total EV demand is computed by accumulating the demands of individual vehicles. After running the optimization algorithm, the net EV demand (the amount of power the EV is willing to charge/discharge) of each EV is extracted and accumulated. Then, using the value of the welfare function (objective function), the total utility of the fleet is computed based on the amount of power charged/discharged and the price settled for that interval *t*. This information is used to compute the reward function (9). Then, the state is updated, and this process is repeated until the end of the total number of episodes (*K*). The state s_t contains information on the number of EVs, the transformer capacity, the total energy demand of EVs, and the current interval

$$s_t = \left[V_t,\ Tx_t^{ac},\ e_t^{net},\ t\right]. \tag{12}$$

Algorithm II Getting EV values during interval *t*.

Get EV parameters returned from Algorithm I.
Initialize total energy required by EV fleet: $e_t^{req} = 0$
for all $t \in T$ **do**
  for all $v \in V$ **do**
    Compute total required energy: $e_t^{req} += e_{v,t}^{req}$
  end for
  Initialize net required energy: $e_t^{net} = 0$
  Formulate and run optimization problem: (4)-(8)
  for all $v \in V$ **do**
    Extract the net energy demand of each EV ($e_{v,t}^{net}$)
    Compute net energy for interval *t*: $e_t^{net} += e_{v,t}^{net}$
  end for
  Extract total welfare: objective function value (W_t^{tot})
  Compute total utility: $U_t^{tot} = W_t^{tot} + e_t^{net} * PR_t$
end for
return: $W_t^{tot}, e_t^{net}, e_t^{req}, U_t^{tot}$

The state parameters are then normalized for stability and fast convergence of the learning processing

$$s_t[k] = \frac{s_t[k] - s^{\min}[k]}{s^{\max}[k] - s^{\min}[k]}, \tag{13}$$

where $s^{min}[k]$ and $s^{max}[k]$ are, respectively, the minimum and maximum value of the k^{th} state parameter. This way, the model is trained, and the trained model can be used to manage the load of any charging station that has a finite number of EVs.

To test the performance of the trained model, it is loaded first. Then, the interval-wise number of EVs and the available capacity of the transformer is defined. The model sets the parameters for EVs based on Algorithm I and takes a step in the environment (takes and action, gets reward, and moves to next state). After each interval, parameters of interest, such as price, amount of power charged/discharged by each EV, etc., can be recorded. It can be seen that the proposed trained model only requires information about the total number of EVs and the available capacity of the transformer. It can generate EV parameters based on the preset bounds of each EV (FO database). The trained model can produce results in real time with any finite number of EVs for different types of the charging station. These charging stations could be residential, commercial/industrial, or mixed residential and commercial/industrial.

4. Numerical Simulations

4.1 Test Network

To analyze the performance of the proposed method, yearly data are used from a real distribution system [38] located in the Midwest of the US. This distribution system belongs to a municipal utility and is fully observable, i.e., smart meters are installed at all customer sites. This publicly available data set describes

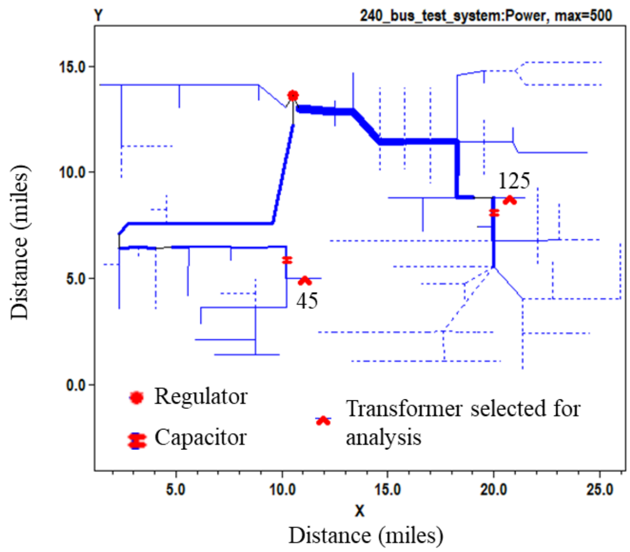

Fig. 4 Test network used for evaluation of the proposed SAC-based local demand management method.

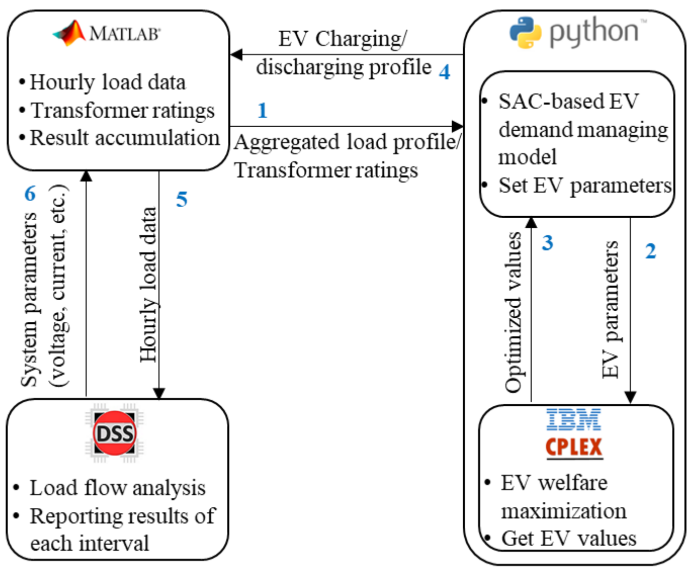

Fig. 5 Simulation framework for testing the network-level impact of the proposed method.

three feeders supplied with a 69-kV substation. The feeders contain residential and commercial loads and also single-phase and three-phase transformers. An overview of the network is shown in Fig. 4. As indicated in the figure, two transformers are selected to evaluate the performance of the proposed method, one being predominantly residential (bus 45) and the other being predominantly commercial/industrial (bus 125).

The simulation framework used for testing the network-level impacts is shown in Fig. 5. First, the hourly network data is exported to MATLAB. It includes transformer-level active and reactive powers. This information, along with the transformer rating, is sent to Python, where the SAC-based model is trained

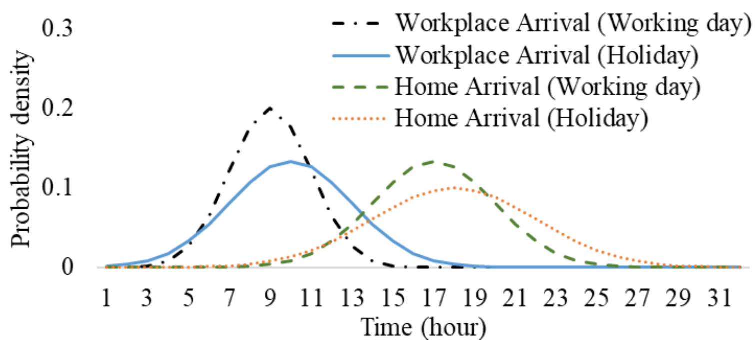

Fig. 6 Daily arrival and departure probabilities of vehicles at/from home and workplace.

using the provided data. The DRL model is trained using data of hourly resolutions for a horizon of 24-hours (T=24). The welfare maximization problem is also simulated in Python using the CPLEX optimization tool [39], as outlined in Section III-B.

The optimized values (charging/discharging decisions of EVs and total welfare) of the model are sent back to MATLAB for quasi-static time-series analysis. An OpenDSS [40] model is developed using the system parameters, such as the load along with the transformer, lines, capacitors, and regulator ratings. The model is called from MATLAB for each hour and the load flow analysis is performed. Parameters of interest, such as line currents and transformer power, are extracted and accumulated for the entire year for further analysis. The extracted parameters are then used to compute the transformer overload, as discussed in the subsequent sections. The combined use of Matlab and Python in this study allows to leverage the advantages of both software platforms. However, either of these programming languages can also be used individually to solve the problem, yielding the same results.

4.2 Input Data

In this study, Edmonton's daily travel data is used. Traffic data [41] are used to estimate the daily arrival times of vehicles at home and work. Similarly, different density functions are estimated for home and workplace departure times, as shown in Fig. 6. As expected, noticeable peaks are observed during weekdays. Similarly, Edmonton's household travel survey data [42] are used to estimate the daily mileage of vehicles. The mean and standard deviation for the Edmonton region are 40.9km and 1.5km, respectively. The parameters of the EV models available in Canada (as of March 2023), used in this study, are listed in Table I.

Table I: Parameters of EV models used in this study.

Model name	Battery size (kWh)	Efficiency (Wh/km)
BMW i3	37.9	165
Chevrolet BOLT	66	254
Ford Mustang Mach-E	68	197

Hyundai IONIQ Electric	38.3	153
Hyundai KONA Electric	64	162
Jaguar I-PACE	84.7	223
Kia Niro	64	173
Kia Soul Electric	39.2	170
MINI Cooper SE	28.9	156
Nissan LEAF	56	172
Smart Fortwo Electric	16.7	176
Tesla Model 3	51	146
Tesla Model S	90	162
Tesla Model X	90	198
Tesla Model Y	76	177
Tesla Cybertruck(pre-order)	100	256
Volkswagen e-Golf	35.8	214

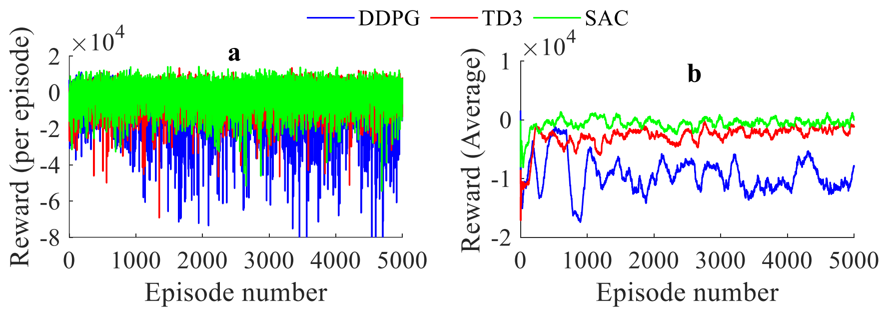

Fig. 7 Convergence analysis of different DRL methods: a) per-episode reward, b) running average reward of 100 episodes.

4.3 Comparative Analysis of Different DRL Methods

In this section, the performance of the SAC model is compared with two other state-of-the-art DRL methods, i.e., the deep deterministic policy gradient (DDPG) and twin delayed DDPG (TD3). All models are trained using the same data (yearly network data with the same EV fleet). The per-episode reward and the moving average reward (100 episodes) of all three models are shown in Fig. 7. It can be observed that the performance of SAC is more stable (smaller fluctuations) compared to the other two methods; DDPG has the worst performance. The fluctuations in episode reward in all methods are due to the random selection of EVs and the random generation of EV parameters during each episode. For example, SoC, sensitivity parameters (*a, b,* and *c*), daily mileage, and battery capacity and mileage efficiency. In addition, the convergence speed of SAC (converged around 3000 episodes) is also faster compared to TD3 and DDPG. Finally, SAC has reached the highest reward level, which signifies that it has the ability to better explore the environment and find better actions. These results are in alignment with several similar studies [37], [43], [44] conducted on performance comparison of DRL methods. Due to these desirable traits of SAC, it is adopted for simulation in this study and only the results of SAC will be discussed in the remainder of the paper.

4.4 Local Charging Demand Management Results

The SAC model is trained using the yearly data of the distribution network and the generated EV load profiles, as discussed in Section III. The test results are then obtained by defining the number of EVs, transformer capacity, and the price range of the charging stations. Similarly, the EV parameters are generated using the specified bounds of each EV. The upper and lower bounds of the price are chosen as 55¢/kWh and 15¢/kWh.

In this section, the performance of the proposed method is analyzed for different hours of a selected day. The day is selected based on the criterion that it has a transformer overload for both selected transformers (45 and 125). The load of selected buses, before the integration of EVs, is shown in Fig. 8. The total load (after integrating EVs) and capacity of the transformers along with the price determined for each interval

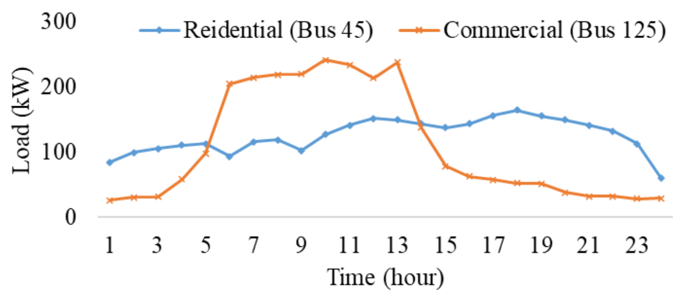

Fig. 8 Electric load (without including EV load) of selected buses for analysis.

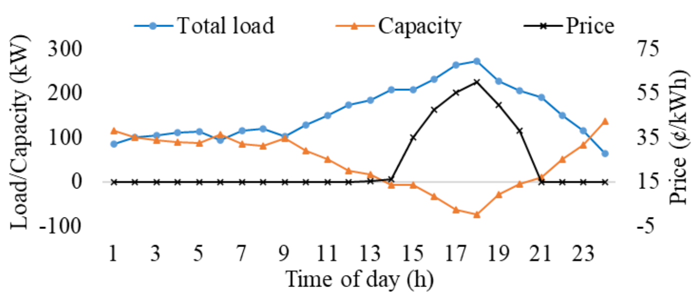

Fig. 9 Load, remaining capacity, and price results of the residential transformer (bus 45) for the selected day.

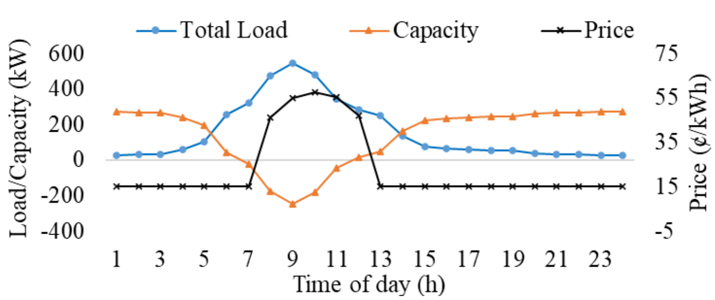

Fig. 10 Load, remaining capacity, and price results of the commercial transformer (bus 125) for the selected day.

are shown in Figs. 9 and 10. It should be noted that, in these figures, capacity refers to the remaining capacity of the transformers. It is obtained by subtracting the original capacity of the transformer from the total load. Therefore, a positive value of capacity indicates that the transformer is not overloaded and still has the capacity to serve loads. Conversely, a negative capacity value indicates that the transformer is overloaded.

It can be observed from Fig. 9 that the price signal, determined by the SAC algorithm for that interval, follows the capacity and net load of the system. For example, when the capacity is positive and the total load is lower (intervals 1-13), it chooses the lowest price. However, it increases the price with an increase in load and negative capacity (intervals 14-20). Remarkably, it reaches the highest price value (55¢/kWh) during interval 18. Being a residential transformer, it has the highest residential load and the EV load in the evening hours. The same is true for bus 125. The model correctly traces the peak load hours and the available capacity of the transformer. Being a commercial transformer, the peak load occurs in the morning hours, and the price increases during those intervals, as shown in Fig. 10. The price is set to the lowest value (15¢/kWh) during the evening and afternoon hours (intervals 13-24). This analysis shows that the trained model can be used for different types of charging stations (commercial or residential).

5. Discussion and Analysis

5.1 Flexibility and Scalability Analysis

In this section, the flexibility of the proposed method is analyzed through a detailed analysis of EV charging/discharging decisions in different types of charging stations, such as residential and commercial/industrial. Similarly, a varying number of EVs are selected for each charging station to assess the scalability of the proposed method. An overview of the number of EVs present in each charging station during different hours of the day is presented in Figure 11.

5.1.1 Residential Charging Station

In this section, interval 19 of the residential transformer (bus 45) is selected for detailed analysis. It can be observed that there are 9 EVs at the residential charging station during this interval (interval 19) and the

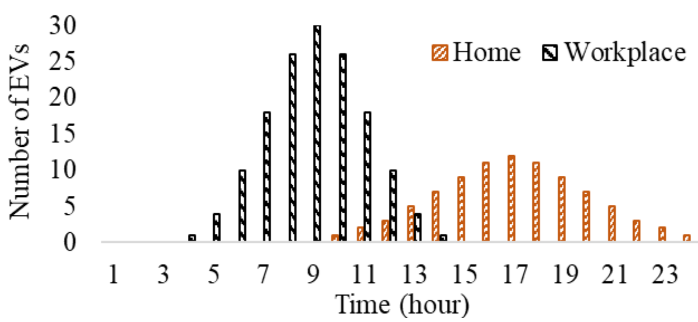

Fig. 11 Number of EVs present in the charging station during different intervals of the day.

Table II: EV parameters of the residential charging station.

EV ID	SoC	*a*	*b*	*c*
1	0.69	0.3	8.5	16
2	0.85	0.46	11.5	18
3	0.4	0.1	14.5	23
4	0.41	0.07	9	16
5	0.36	0.2	5.5	22
6	0.73	0.26	14.5	26
7	0.32	0.25	8.5	22
8	0.41	0.47	13.5	26
9	0.35	0.22	6	22

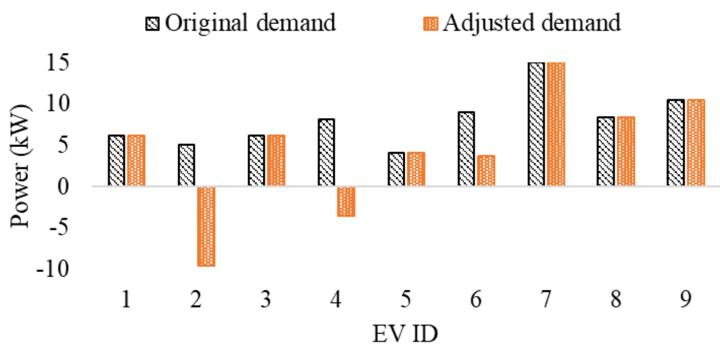

Fig. 12 EV-wise charging/discharging decisions for the selected interval.

remaining transformer capacity (excluding the EV load) is 50kW. The original net demand for the 9 EVs is 72.7kW. Therefore, the demand must be reduced by at least 22.7kW to avoid transformer overload. The SoC level and sensitivity parameters generated for this interval for each EV are tabulated in Table II. Parameters *a*, *b*, and *c* are the sensitivity parameters used in the utility function (1). It should be noted that the higher value of *a* and the lower value of *b* correspond to the higher sensitivity of any EV owner to the SoC level and vice versa. Similarly, a higher value of *c* corresponds to higher sensitivity to battery degradation and vice versa.

The original and adjusted demands of all EVs are shown in Fig. 12. Adjusted demand refers to the amount of power an EV has decided to charge/discharge based on the price derived for this interval. It can be observed that EV2 has decided to discharge since it has a higher SoC (85%) compared to other EVs and a lower value of *c* (less sensitive to battery degradation). Similarly, EV 4 has also decided to discharge due to the lowest sensitivity to SoC level (lowest *a*) and battery degradation (lower *c*). EV6 has reduced its demand from 9.01kW to 3.77kW due to a higher SoC. Note that it has not decided to discharge, despite having a lower sensitivity to SoC (smaller *a* and higher *b*), due to a higher sensitivity to battery degradation (highest *c*). The remaining EVs have decided to charge their full demands mainly due to lower SoCs. This analysis shows that the proposed methods determine the charging/discharging power of EVs in a charging station in the desired manner based on their sensitivity to different parameters.

5.1.2 Industrial/Commercial Charging Station

In this section, the interval with the maximum number of EVs (interval 9) at the commercial charging station (bus 125) is selected for a detailed analysis. It can be observed that there are 30 EVs present during this interval. The remaining capacity of the transformer is 58kW, and the net EV demand (original demand) is 239.14kW. Therefore, the net demand needs to be reduced by at least 181.14kW to avoid overloading the transformer. The state of charge (SoC) level, sensitivity parameters (*a*, *b*, and *c*), and the adjusted and original demands are presented in Table III.

Table III: EV parameters and demand of the commercial charging station.

EV	SOC	*a*	*b*	*c*	Original demand	Adjusted demand
1	0.63	0.11	11.5	27	12.70	4.64
2	0.31	0.05	9.5	29	12.90	5.27
3	0.33	0.21	9.5	15	14.20	10.20
4	0.6	0.45	12.5	27	8.30	8.30
5	0.35	0.15	8	18	12.10	-4.33
6	0.4	0.08	10.5	16	12.10	-5.78
7	0.54	0.38	8	27	10.00	10.00
8	0.6	0.35	11	29	3.70	3.31
9	0.42	0.49	7	28	8.02	8.02
10	0.48	0.35	11.5	22	5.70	5.70
11	0.61	0.11	9	20	4.84	-6.85
12	0.61	0.06	6	19	14.20	-36.90
13	0.66	0.43	14.5	27	2.40	2.40
14	0.55	0.32	10	29	6.70	6.70
15	0.61	0.42	8	20	13.44	13.44
16	0.58	0.19	8	17	5.34	-6.35
17	0.75	0.14	13.5	18	4.34	-5.76
18	0.42	0.12	5.5	20	11.00	7.44
19	0.63	0.22	9	19	8.90	0.74
20	0.5	0.38	13.5	25	3.30	3.30
21	0.36	0.47	8	21	2.50	2.50
22	0.48	0.4	11.5	28	3.00	3.00
23	0.55	0.07	10	19	12.60	-10.65
24	0.54	0.46	9	20	3.10	3.10
25	0.64	0.42	11	20	8.16	8.16
26	0.32	0.32	7	25	9.00	9.00
27	0.5	0.12	13.5	17	6.90	-1.55
28	0.37	0.37	13	28	5.10	5.10
29	0.44	0.32	7.5	29	9.00	9.00
30	0.49	0.36	8	29	5.60	5.60

It can be observed from Table III that the proposed method has successfully reduced the load of EVs. The adjusted load has been reduced to 56.75kW, which is less than the available transformer capacity (58kW). EVs 5, 6, 11, 12, 16, 17, 23, and 27 have decided to discharge primarily due to lower values of the parameter *c* (sensitivity to battery degradation) and lower values of the parameter *a* (sensitivity to battery SoC). Similarly, EVs 1, 2, 3, 8, 18, and 19 have decided to reduce their demands due to lower values of the parameter *a* and lower SoCs. It is worth noting that these EVs have not decided to discharge mainly due to higher values of the parameter *c*. All other EVs have decided not to change their demands due to either higher sensitivity to SoC (higher values of *a* and lower values of *b*) and battery degradation (higher values of *c*) or lower SoC.

This analysis has shown that the proposed method can successfully manage the load of both residential and commercial/industrial charging stations. Furthermore, it validates that the proposed method is scalable and can be applied to charging stations with either a small or large number of EVs.

5.2 Performance Comparison

In this section, the performance of the proposed method is compared to a game-theory-based allocation method [27] for load management in charging stations, considering both residential and commercial scenarios. The EV parameters for the residential charging station are the same as shown in Table II. The results of both the proposed method and the game theory-based method are presented in Figure 13. It can be observed that both methods exhibit a similar trend. For instance, EVs 2 and 4 have decided to discharge, EV 6 has reduced its demand, and all other EVs have chosen to maintain their original demands. This demonstrates that the proposed method performs comparably to the game-theory-based method, despite its ability to handle different types of charging stations.

Similarly, the performance of the proposed method and the game-theory-based method is compared for the commercial charging station. The EV parameters provided in Table III are used for this analysis, and the results are presented in Table IV. It can be observed that both methods exhibit a similar trend in this case as well, mirroring the observations made for the residential charging station. For example, the same EVs have decided to discharge (EVs 5, 6, 11, 12, 16, 17, and 23), reduce their demand (EVs 1, 2, 3, 8, 18, and 19), or maintain their original demand (EVs 4, 7, 9, 10, 13-15, and 20-30). This indicates that the proposed method, using the same model, is capable of managing commercial charging stations with a higher number of EVs.

It has been validated that the proposed method is capable of effectively managing the load of various charging station types, achieving desired outcomes similar to a game-theory model, but without the need for multiple iterations in real time. For instance, the game-theory method required 21 and 78 iterations to converge for residential and commercial charging station cases, respectively. It is important to note that the game-theory model needs to repeat the entire process whenever an EV joins or leaves the charging station, which is not the case with proposed method (retraining is not required). Instead, the proposed model demonstrates its ability to handle uncertainties by incorporating various EV sensitivity factors during off-line training. The model incorporates approximated parameters for different combinations of EV factors, thereby enhancing its adaptability to handle unpredictable situations. The key difference lies in the amount of power discharged or reduced by different EVs. This disparity is primarily attributed to the game-theory model's inclination to precisely match the adjusted load with the remaining transformer capacity, which is usually not desirable.

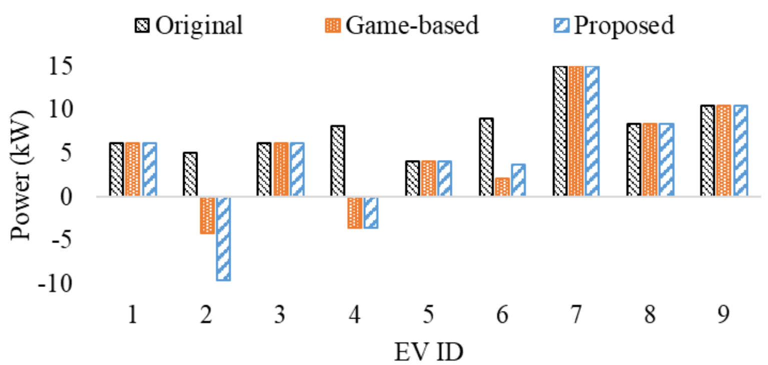

Fig. 13 Result comparison for a residential charging station.

Table IV: Result comparison for a commercial charging station.

EV ID	Original	Game-based	Proposed	EV ID	Original	Game-based	Proposed
1	12.7	6.1	4.6	16	5.3	-6.4	-6.4
2	12.9	10.3	5.3	17	4.3	-15.9	-5.8
3	14.2	10.1	10.2	18	11.0	9.7	7.4
4	8.3	8.3	8.3	19	8.9	6.2	0.7
5	12.1	-4.3	-4.3	20	3.3	3.3	3.3
6	12.1	-5.8	-5.8	21	2.5	2.5	2.5
7	10.0	10.0	10.0	22	3.0	3.0	3.0
8	3.7	2.4	3.3	23	12.6	-13.4	-10.7
9	8.0	8.0	8.0	24	3.1	3.1	3.1
10	5.7	5.7	5.7	25	8.2	8.2	8.2
11	4.8	-6.9	-6.9	26	9.0	9.0	9.0
12	14.2	-29.0	-36.9	27	6.9	-8.7	-1.6
13	2.4	2.4	2.4	28	5.1	5.1	5.1
14	6.7	6.7	6.7	29	9.0	9.0	9.0
15	13.4	13.4	13.4	30	5.6	5.6	5.6

5.3 Network Level Impact Analysis

In this section, the impact of EVs on the distribution system and the performance of the proposed method in mitigating transformer overload are analyzed. The proposed model is tested for different charging stations (located at different transformers) having more than 5 EVs. For the sake of visualization, two cases are particularly analyzed in this section. In the first case, a residential transformer (bus 45) is analyzed, and a commercial/industrial transformer (bus 125) is analyzed in the second case. In each case, the utilization ($U_{t,tx}$) of the transformers is computed using

$$U_{t,tx} = \left(\frac{S_{t,tx}}{S_{tx}^{rat}} \right) \cdot 100 \, . \tag{14}$$

where $S_{t,tx}$ is the power flowing through transformer tx during time t. Similarly, S_{tx}^{rat} is the rated kVA of the transformer.

The utilization of the selected transformers is analyzed for the entire year. It can be observed from Fig. 14 that transformer utilization is below the 100% threshold without the integration of EVs, i.e., the transformers are not overloaded. However, a significant overload can be observed in the case "With EVs". Being a residential transformer, overloading is prominent during evening hours, when most EV owners return home. They use their home appliances, and most of them tend to charge their EVs. Overloading has been mitigated by the proposed method, using the V2V service, as shown in the rightmost figure. Similarly, for the commercial transformer, overloading can be observed when EVs are integrated, especially during

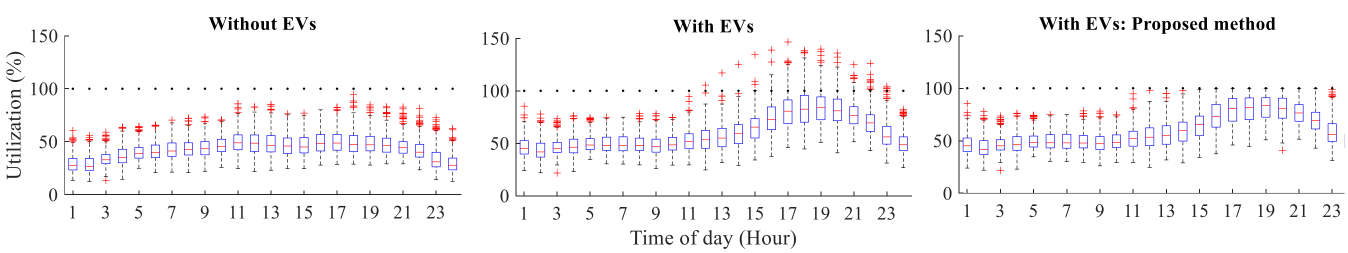

Fig. 14 Transformer utilization under different scenarios for residential transformer (bus 45).

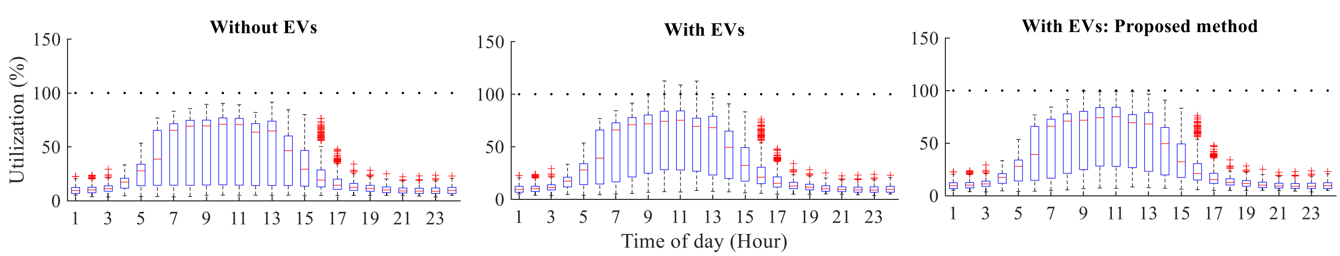

Fig. 15 Transformer utilization under different scenarios for commercial transformer (bus 125).

office hours. The proposed method has also successfully mitigated the overloading for this case, as shown in the rightmost part of Fig. 15. It can be concluded from this analysis that the proposed method can mitigate transformer overloading in different types of transformers by managing the load locally through V2V services.

6. Conclusions

In this study, a soft actor-critic model is trained to manage the load of charging stations locally to relieve the distribution transformers from overloading. Comparative analysis has shown that the soft actor-critic method outperforms other state-of-the-art deep reinforcement learning methods in terms of convergence speed and stability. The vehicle-to-vehicle service is realized as a welfare maximization model with the objective of maximizing the welfare of electric vehicle owners. Simulations have shown that by using the proposed method, price signals can be generated to maximize the welfare of electric vehicles while respecting the capacity limits of distribution transformers. It has been demonstrated that the proposed method can be used to manage the load of different types of charging stations, such as residential, commercial/industrial, and mixed residential and commercial/industrial. The performance of the proposed model is similar to game-theory based models for these diverse charging stations. This indicates that the proposed model effectively addresses the challenges and requirements of managing different types of charging stations, using a single model. In addition, the proposed model can solve the problem in a single run in contrast to the iterative game models. Analysis of individual electric vehicles has shown that electric vehicles can optimally determine their charging and discharging amounts, based on their sensitivities to different factors (battery degradation and energy level) and price determined by the fleet operator, using the proposed method. Yearly data analysis has shown that transformer overload can be mitigated for different types of days by managing the charging station load locally. In addition, training is an offline process and the trained model can be used to manage the load of charging stations in real-time.

Acknowledgments

This research has been supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada grant number ALLRP 549804-19, and by the Alberta Electric System Operator (AESO), AltaLink, ATCO Electric, ENMAX, EPCOR Distribution and Transmission Inc., and FortisAlberta.

References

[1] W. Sun, F. Neumann, and G. P. Harrison, “Robust Scheduling of Electric Vehicle Charging in LV Distribution Networks under Uncertainty,” IEEE Trans. Ind. Appl., vol. 56, no. 5, pp. 5785–5795, 2020, doi: 10.1109/TIA.2020.2983906.

[2] A. Hussain and P. Musilek, “Resilience Enhancement Strategies for and Through Electric Vehicles,” Sustain. Cities Soc., vol. 80, 2022, pp. 103788.

[3] 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, 2019, pp. 100006.

[4] T. Unterluggauer, J. Rich, P. B. Andersen, and S. Hashemi. "Electric vehicle charging infrastructure planning for integrated transportation and power distribution networks: A review." ETransportation, vol. 12, 2022, pp. 100163.

[5] N. O’Connell, Q. Wu, J. Østergaard, A. H. Nielsen, S. T. Cha, and Y. Ding, “Day-ahead tariffs for the alleviation of distribution grid congestion from electric vehicles,” Electr. Power Syst. Res., vol. 92, pp. 106–114, 2012, doi: 10.1016/j.epsr.2012.05.018.

[6] R. Rana and S. Mishra, “Day-Ahead Scheduling of Electric Vehicles for Overloading Management in Active Distribution System via Web-Based Application,” IEEE Syst. J., vol. 13, no. 3, pp. 3422–3432, 2019, doi: 10.1109/JSYST.2018.2851618.

[7] A. H. Einaddin and A. S. Yazdankhah. "A novel approach for multi-objective optimal scheduling of large-scale EV fleets in a smart distribution grid considering realistic and stochastic modeling framework." Int. J. Electri. Pow. Energy Sys., vol. 117, 2020, pp. 105617.

[8] J. Hu, C. Si, M. Lind, and R. Yu, “Preventing distribution grid congestion by integrating indirect control in a hierarchical electric vehicles’ management system,” IEEE Trans. Transp. Electrif., vol. 2, no. 3, pp. 290–299, 2016, doi: 10.1109/TTE.2016.2554469.

[9] L. Ni, F. Wen, W. Liu, J. Meng, G. Lin, and S. Dang, “Congestion management with demand response considering uncertainties of distributed generation outputs and market prices,” J. Mod. Power Syst. Clean Energy, vol. 5, no. 1, pp. 66–78, 2017, doi: 10.1007/s40565-016-0257-9.

[10] Haider, R. e. Z. Rizvi, J. Walewski, and P. Schegner, “Investigating peer-to-peer power transactions for reducing EV induced network congestion,” Energy, vol. 254, p. 124317, 2022, doi: 10.1016/j.energy.2022.124317.

[11] S. Deb et al., “Charging Coordination of Plug-In Electric Vehicle for Congestion Management in Distribution System Integrated with Renewable Energy Sources,” IEEE Trans. Ind. Appl., vol. 56, no. 5, pp. 5452–5462, 2020, doi: 10.1109/TIA.2020.3010897.

[12] R. C. Leou, C. L. Su, and C. N. Lu, “Stochastic analyses of electric vehicle charging impacts on distribution network,” IEEE Trans. Power Syst., vol. 29, no. 3, pp. 1055–1063, 2014, doi: 10.1109/TPWRS.2013.2291556.

[13] S. Lai, J. Qiu, Y. Tao, J. Zhao, and S. Member, “Pricing for Electric Vehicle Charging Stations based on the Responsiveness of Demand,” IEEE Trans. Smart Grid, vol. PP, p. 1, 2021, doi: 10.1109/TSG.2022.3188832.

[14] J. Soares, J. Almeida, L. Gomes, B. Canizes, Z. Vale, and E. Neto. "Electric vehicles local flexibility strategies for congestion relief on distribution networks." Energy Reports, vol. 8, 2022, pp. 62-69.

[15] G. R. Pleat, “Allowing for Off-Peak Price Incentives to Reduce Electric Vehicle Distribution Grid Congestion,” Electr. J., vol. 25, no. 6, pp. 89–97, 2012, doi: 10.1016/j.tej.2012.06.009.

[16] J. Dixon, W. Bukhsh, K. Bell, and C. Brand. "Vehicle to grid: driver plug-in patterns, their impact on the cost and carbon of charging, and implications for system flexibility." ETransportation, vol. 13, 2022, pp. 100180.

[17] F. L. Da Silva, C. E. H. Nishida, D. M. Roijers, and A. H. R. Costa, “Coordination of Electric Vehicle Charging through Multiagent Reinforcement Learning,” IEEE Trans. Smart Grid, vol. 11, no. 3, pp. 2347–2356, 2020, doi: 10.1109/TSG.2019.2952331.

[18] FleetCarma, “Charge the North - Results from the world’s largest electric vehicle charging study.,” pp. 0–7, 2018.

[19] A. Khamis, M.H. Aiman, W.M. Faizal, C.Y. Khor, "Charging strategy in electric vehicle chargers by utilizing demand side management scheme", Electric Power Systems Research, vol. 220, 2023, pp. 109240.

[20] M. Zeballos, A. Ferragut, and F. Paganini, “Proportional Fairness for EV Charging in Overload,” IEEE Trans. Smart Grid, vol. 10, no. 6, pp. 6792–6801, Nov. 2019, doi: 10.1109/TSG.2019.2911231.

[21] J. Li, C. Li, Y. Xu, Z. Y. Dong, K. P. Wong, and T. Huang, “Noncooperative Game-Based Distributed Charging Control for Plug-In Electric Vehicles in Distribution Networks,” IEEE Trans. Ind. Informatics, vol. 14, no. 1, pp. 301–310, 2018, doi: 10.1109/TII.2016.2632761.

[22] Z. Zhang, Y. Wan, J. Qin, W. Fu, and Y. Kang, “A Deep RL-Based Algorithm for Coordinated Charging of Electric Vehicles,” IEEE Trans. Intell. Transp. Syst., vol. PP, pp. 1–11, 2022, doi: 10.1109/TITS.2022.3170000.

[23] A. Bertolini, M. S. E. Martins, S. M. Vieira, and J. M.C. Sousa. "Power output optimization of electric vehicles smart charging hubs using deep reinforcement learning." Expert Systems with Applications, vol. 201, 2022, pp. 116995.

[24] L. Yan, X. Chen, Y. Chen, and J. Wen. "A cooperative charging control strategy for electric vehicles based on multiagent deep reinforcement learning." IEEE Transactions on Industrial Informatics, vol. 18, no. 12 (2022): 8765-8775.

[25] J. Jin, and Y. Xu. "Optimal policy characterization enhanced actor-critic approach for electric vehicle charging scheduling in a power distribution network." IEEE Transactions on Smart Grid, vol. 12, no. 2 (2020): 1416-1428.

[26] F. Ming, F. Gao, K. Liu, and X. Li. "A constrained DRL-based bi-level coordinated method for large-scale EVs charging." Applied Energy, vol. 331 (2023): 120381.

[27] A. Hussain and P. Musilek, “A Game-Theoretic Approach for Charging Demand Management of Electric Vehicles during System Overload,” in 2021 IEEE Electrical Power and Energy Conference, EPEC 2021, Dec. 2021, pp. 353–358, doi: 10.1109/EPEC52095.2021.9621384.

[28] D. Li, A. Zouma, J.-T. Liao, and H.-T. Yang. "An energy management strategy with renewable energy and energy storage system for a large electric vehicle charging station." Etransportation, vol. 6, 2020, pp. 100076.

[29] P. Samadi, A.-H. Mohsenian-Rad, R. Schober, V. W. S. Wong, and J. Jatskevich, “Optimal Real-Time Pricing Algorithm Based on Utility Maximization for Smart Grid,” pp. 415–420, Nov. 2010, doi: 10.1109/SMARTGRID.2010.5622077.

[30] B. H. Lim and H. S. Lee. "Portfolio decision with a quadratic utility and inflation risk." Advances in Difference Equations, vol. 2018, no. 1 (2018): 1-16.

[31] A. Tomaszewska, Z. Chu, X. Feng, S. O'kane, X. Liu, J. Chen, C. Ji et al. "Lithium-ion battery fast charging: A review." ETransportation, vol. 1, 2019, pp. 100011.

[32] N. Liu, X. Yu, C. Wang, and J. Wang, “Energy Sharing Management for Microgrids with PV Prosumers: A Stackelberg Game Approach,” IEEE Trans. Ind. Informatics, vol. 13, no. 3, pp. 1088–1098, Jun. 2017, doi: 10.1109/TII.2017.2654302.

[33] X. Luo, W. Shi, Y. Jiang, Y. Liu, and J. Xia. "Distributed peer-to-peer energy trading based on game theory in a community microgrid considering ownership complexity of distributed energy resources." J. Cleaner Production, vol. 351, 2022, pp. 131573.

[34] V. H. Bui, A. Hussain, and W. Su, “A Dynamic Internal Trading Price Strategy for Networked Microgrids: A Deep Reinforcement Learning Based Game-Theoretic Approach,” IEEE Trans. Smart Grid, 2022, doi: 10.1109/TSG.2022.3168856.

[35] T. Haarnoja et al., “Soft Actor-Critic Algorithms and Applications,” Dec. 2018, doi: 10.48550/arxiv.1812.05905.

[36] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor.” PMLR, pp. 1861–1870, Jul. 03, 2018, Accessed: Sep. 04, 2022. [Online]. Available: https://proceedings.mlr.press/v80/haarnoja18b.html.

[37] S. Fujimoto, H. Hoof, and D. Meger. "Addressing function approximation error in actor-critic methods." In International conference on machine learning, pp. 1587-1596. PMLR, 2018.

[38] F. Bu, Y. Yuan, Z. Wang, K. Dehghanpour, and A. Kimber, “A Time-Series Distribution Test System Based on Real Utility Data,” 51st North Am. Power Symp. NAPS 2019, Jun. 2019, doi: 10.1109/NAPS46351.2019.8999982.

[39] “CPLEX Optimizer | IBM.” https://www.ibm.com/analytics/cplex-optimizer (accessed Sep. 04, 2022).

[40] “OpenDSS.” https://www.epri.com/pages/sa/opendss (accessed Sep. 04, 2022).

[41] “Toronto traffic report | TomTom Traffic Index,” 2021. https://www.tomtom.com/en_gb/traffic-index/edmonton-traffic/ (accessed Sep. 04, 2022).

[42] R. A. M. & A. Ltd., “City of Edmonton, 2018, "2015 Edmonton and Region Household Travel Survey - Google Search.” .

[43] A. Hussain, V.-H. Bui, and H.-M. Kim, “Deep reinforcement learning-based operation of fast charging stations coupled with energy storage system,” Electr. Power Syst. Res., vol. 210, p. 108087, 2022, doi: 10.1016/j.epsr.2022.108087.

[44] V. H. Bui and W. Su, “Real-time operation of distribution network: A deep reinforcement learning-based reconfiguration approach,” Sustain. Energy Technol. Assessments, vol. 50, p. 101841, Mar. 2022, doi: 10.1016/J.SETA.2021.101841.