Article

Collaborative Optimization of Electric Vehicles Based on MultiAgent Variant Roth–Erev Algorithm

Jianwei Gao ^{1,2}, Yu Yang ^{1,2,*} , Fangjie Gao ^{1,2} and Haoyu Wu ^{1,2}

¹ School of Economics and Management, North China Electric Power University, Changping, Beijing 102206, China; gaojianwei111@sina.cn (J.G.); gfg930818@163.com (F.G.); vorlewu@163.com (H.W.)

² Beijing Key Laboratory of New Energy and Low-Carbon Development, North China Electric Power University, Beijing 102206, China

* Correspondence: yang_yang_yu_yu@163.com

Abstract: With the implementation of the carbon neutral policy, the number of electric vehicles (EVs) is increasing. Thus, it is urgently needed to manage the charging and discharging behavior of EVs scientifically. In this paper, EVs are regarded as agents, and a multiagent cooperative optimization scheduling model based on Roth–Erev (RE) algorithm is proposed. The charging and discharging behaviors of EVs will influence each other. The charging and discharging strategy of one EV owner will affect the choice of others. Therefore, the RE algorithm is selected to obtain the optimal charging and discharging strategy of the EV group, with the utility function of the prospect theory proposed to describe EV owners' different risk preferences. The utility function of the prospect theory has superior effectiveness in describing consumers' utility. Finally, in the case of residential electricity, the effectiveness of the proposed method is verified. Compared with that of random charging, this method reduces the total EV group cost of EVs by 52.4%, with the load variance reduced by 26.4%.

Keywords: optimized dispatching; prospect theory; electric vehicles (EVs); RE algorithm; multiagent games

Citation: Gao, J.; Yang, Y.; Gao, F.; Wu, H. Collaborative Optimization of Electric Vehicles Based on MultiAgent Variant Roth–Erev Algorithm. *Energies* **2022**, *15*, 125. <https://doi.org/10.3390/en15010125>

Academic Editor: Calin Iclodean

Received: 29 November 2021

Accepted: 22 December 2021

Published: 24 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, China vigorously promoted the development and application of electric vehicles (EVs). EVs are an alternative to fossil fuels, and their application helps to reduce greenhouse gas emissions [1–3]. With the increase in EVs, their charging and discharging behaviors threatened the safe operation of the distribution network [4].

Electric vehicles can be utilized as an energy storage system to alleviate the shortage during the electricity consumption peak, decreasing power prices [5–7]. For example, batteries are enabling technologies for Vehicle to Grid (V2G) because they act as storage devices that provide electricity during grid demand peak. As a decarbonization technology, batteries that coordinate EVs and grid are receiving increasing attention between EVs, and the grid is receiving increasing attention from the providers of auxiliary services. During the application of EVs, the aggregator can play a coordination role [8].

The notion of V2G was first introduced by Kempton William and Amory Lovins [9]. They tried to explain the future wide application of EVs, which can equip the grid with valuable storage facilities [9–11]. The basic idea underlying V2G is that when demand rises, electricity in the batteries of parked EVs will be pumped into the grid; when demand decreases, batteries of EVs will be charged [12].

The disorderly charging behavior of EVs not controlled by aggregators was studied in the literature [13–18]. Large-scale uncontrolled EV charging threatens the safe operation of the power grid. To alleviate the impact of the grid-connected EVs on the distribution network, protect the grid's stability, and stabilize the power quality, the centralized and coordinated control of the charging and discharging of EVs was researched in the literature [19–22].

These studies showed that controlling the charging and discharging behavior of EVs can effectively alleviate its impact on the distribution network, but they did not take EV

owners into consideration. The willingness of EV owners with various risk preferences to participate in charge and discharge control is different.

In this paper, the reinforcement learning (RL) algorithm is used to learn the optimization scheduling behavior of the EV fleet. Reinforcement learning is challenged by continuous and very large states and action spaces [23]. Shi and Wong [24] proposed an RL method to provide V2G services through EV fleet. The Markov decision process (MDP) used is considered from the perspective of an EV, that is, without coupling constraints (constraints involving the behavior of multiple EVs). Dawei Qiu and Yujian Ye [25] proposed a new method of deep reinforcement learning (DRL). The deep deterministic policy gradient (DDPG) principle is combined with the priority experience replay (PER) strategy, and the problem is set in a multidimensional continuous state and action space to solve the studied EV pricing problem. How to make a daily charge consumption plan for a group of EVs and make them follow the plan during operation, control EV charging through heuristic schemes during operation, and use batch mode RL to learn the EV charging behavior were discussed [26]. In the present paper, Variant Roth–Erev (VRE) algorithm [27] in RL is proposed to optimize the charging and discharging scheduling of EVs. The traditional Roth–Erev (RE) algorithm can only deal with positive feedback value, but the improved VRE algorithm can deal with both positive and negative feedback value and zero feedback value.

2. The Related Work

In this paper, each EV owner is regarded as an agent of the proposed multiagent EV collaborative optimization scheduling based on the prospect theory value function and optimized charging and discharging of the EV. The different risk preferences of electric vehicle owners make them have different willingness to participate in the response of charge and discharge demand. Thus, the participation enthusiasm of EV owners will be increased, and the impact of EV groups on power distribution will be stabilized.

The multiagent RE method is used to establish a collaborative optimization model for EVs in this research. EVs managed by the same aggregator are regarded as multiple agents that influence each other and affect the overall load fluctuation and cost. By Roth–Erev (RE) algorithm, the multiagent can reach the optimal equilibrium of the game.

This article's contributions can be summarized as follows.

- (1) A variant parameter value function of prospect theory is proposed to describe the diversity of multiple agents' risk attitudes.
- (2) The endogenous retrovirus (ERV) intelligent algorithm is proposed to optimize the charging and discharging strategy of EVs in the dispatching area.
- (3) The case analysis showed that the ERV intelligent algorithm Boltzmann cooling parameter value is suitable for the charging and discharging optimization of EVs.

The rest of the paper is laid out as follows. The optimized scheduling model framework and variant parameter value function of prospect theory are described in Section 3. The price-based demand response optimization model and the strategy suggested in this paper are presented in Section 4. Section 5 describes the application of the Variant Roth–Erev algorithm in detail. A case study is presented in Section 6 to demonstrate the efficiency of the proposed strategy. Section 7 presents the conclusion.

3. The Related Theory

3.1. The Optimized Scheduling Model Framework

The EV ADR optimization model based on Multiagent RE Algorithm is depicted in Figure 1. The distribution network provides electricity to the demand side through the aggregator that optimizes the electricity consumption strategy of the demand side through the control operation center. In this paper, the electricity for other demand sides, except EVs, is taken as the base load, and the optimal scheduling of EVs' charging and discharging were considered separately.

Figure 1. Multiagent RE algorithm model based on EV ADR optimization.

This paper focuses on a residential area with a centralized control operation center where an aggregator manages the electricity demand and obtains electricity from the distribution network. This paper assumes that residents' other electricity demands, except that for EVs, are all base load demands. Electric vehicles and aggregators signed unified deployment agreements. Aggregator commerce must ensure that the regular driving pattern of EV owners will not be affected, and on this basis, optimization scheduling is carried out.

3.2. Variant Parameter Value Function of Prospect Theory

Unlike the traditional utility function that assumes the consumers are rational, the value function of the prospect theory in this paper regards EV owners participating in demand response as bounded rational consumers. The prospect theory's value function presents the S shape, better describing the behavior decision of multiple agents involved in demand response. The prospect theory value function is described in detail below.

The value function was developed by Kahneman and Tversky to describe psychological effect [28]:

$$v(x) = \begin{cases} (x - b)^{\alpha_0}, & x \geq b, \\ -\theta_0(b - x)^{\beta_0}, & x < b, \end{cases} \quad (1)$$

where $0 < \alpha_0/\beta_0/\theta_0 < 1$; the loss aversion coefficient $\theta_0 > 1$; b represents the reference point. x can be regarded as a gain or a loss depending on b . When the value of b increases/decreases, the whole function moves horizontally to the right/left. In other words, if we want to obtain the same utility value as that before b is not moved, the value of x will increase/decrease accordingly. When x exceeds the reference point, it represents a relative gain, otherwise, a relative loss.

The b value can be a constant for one consumer, but it is not applicable in the case of multiple agents. Therefore, this paper proposes an improved value function of prospect theory where the value b is set as a variable $b(l)$. As different individuals have different

values of $b(l)$, the improved value function of prospect theory can describe multiagent situations where risk preferences are diverse. The EV owners have varying risk preferences, and even owners of the same EV type may have a different appetite for risk. Risk preference is controlled by α_0, β_0 , and its basic reference point is determined by $b(l)$ for different EV purchase time or different models. In Equation (3) $q_i (i = 1, \dots, m)$ represents different groups, and the parameters of $b(l)$ function are selected according to $l \in q_i (i = 1, \dots, m)$. Through such classification, the willingness of EV owners to participate in demand response can be better described in practical application. EV owners have different loss aversion coefficients. Therefore, a variant parameter value function of prospect theory is proposed.

Definition 1. A variant parameter value function of prospect theory is proposed to describe the diversity of multiple agents' risk attitudes, and the formula is shown below:

$$v(l, x) = \begin{cases} (x - b(l))^{\alpha_0}, & x \geq b, \\ -\theta(b(l) - x)^{\beta_0}, & x < b, \end{cases} \quad (2)$$

$$b(l) = \begin{cases} b_1, & l \in q_1 \\ b_2, & l \in q_2 \\ \dots \\ b_m, & l \in q_m \end{cases} \quad (3)$$

$q_i (i = 1, \dots, m)$

where $0 < \alpha_0, \beta_0 < 1$. θ is the loss aversion coefficient and satisfies $\theta > 1$. $b(l)$ is the agent's value function zero-point and varies in different agent groups. $l \in q_i (i = 1, 2, \dots, m)$ means agent l belongs to risk attitude type q_i . There are M types of risk attitudes: $q_1, q_2, \dots, q_m$.

4. Modeling

EV owners are encouraged to participate actively in aggregator optimization scheduling through price optimization. On the demand side, the objective function minimizes the consumption cost of EV owners. The cost of battery loss is considered while seeking the lowest charging price and the highest discharging price. Furthermore, this strategy can cover charging needs without requiring EV owners to change their regular driving patterns.

The strategy based on price optimization will prompt most EVs to choose the initial low price. As the electricity price depends on the total load, when most EVs charge simultaneously and the load increases, the electricity price will rise. Then, the optimal charge/discharge strategy based on the initial price will be deficient. With considering EVs as agents, the application of multiagent RL is proposed in this paper to find the optimal charging/discharging strategy for EVs.

4.1. Model Component

4.1.1. The Real-Time Aggregator Pricing (The Aggregator's Pricing Action)

Electricity price is a rising convex function of total load, according to current power market practices. Equation (4) can model the electricity price that is affected by the total load consumption of price-making agents [29–31]:

$$\rho(k) = \rho(L_{u^*}(k)) \quad (4)$$

where $\rho(*)$ is an increasing convex function. $L_{u^*}(k)$ means the total load on the demand side at the time k based on the optimization strategy u^* . The charge and discharge price of EVs varies according to the total load on the demand side. When the load climbs, the price increases, and vice versa. For EV owners willing to engage in the aggregator organization's optimized schedule, such a setting would be advantageous.

4.1.2. Network Information Transmission

In this paper, a parking and charging station managed by the aggregator is studied. When the electric cars drive into the station, they need to upload their driving information to the aggregator. The following are expected to be the vehicle's important parameters.

$$V_l = \{T_{in,l}, T_{out,l}, S_{O,l}, S_{E,l}, Q_l, P_{c,l}, P_{d,l}\} \quad (5)$$

where $T_{in,l}$, $T_{out,l}$ respectively represent EV entrance and departure time; $S_{O,l}$, $S_{E,l}$ represents initial SOC and expected SOC, respectively; Q_l represents the trolley capacity; $P_{c,l}$, $P_{d,l}$ represents the charging and discharging power, respectively. When entering the parking station, the EV is connected to the grid immediately, and its information is transmitted to the aggregator in the operation center. According to the information of the EV, the aggregator optimizes the scheduling of the vehicle charging and discharging based on the intelligent algorithm.

4.2. Objective Function

Each EV's objective function that defines an optimization model based on pricing demand response is shown below:

$$\min_{u_l} C_l = \sum_{k \in T_{m,l}} \{ [P_{c,l} + P_{d,l}] \cdot \omega_{m,l}(k) I_0 \cdot \rho(k) \Delta t + |[P_{c,l} + P_{d,l}] \cdot \omega_{m,l}(k) I_0| \cdot \varepsilon \} \quad (6)$$

$$\omega_{m,l}(k) = \begin{cases} 1, & k \in T_{m,l} \\ 0, & \text{other wise} \end{cases} \quad k \in T_{m,l} \quad (7)$$

where I_0 is the basic charging/discharging state variable; $I_0 = 1$ means the EV is charging, $I_0 = -1$ indicates the EV is discharging, and 0 means EV is neither charging or discharging; $\omega_{m,l}(k)$ is the EV battery's maneuverability; $\rho(k)$ is real-time electricity price at the time k ; ε is the battery loss factor. The objective function considers the interests of EV owners from the demand side and pursues the minimum cost for EV owners. When the total load on the demand side is high, the electricity price is high, so the owners of EVs choose to discharge to earn profits. On the contrary, when the total load on the demand side is low, the electricity price is low; thus, the EV owners choose to charge to reduce the charging cost. However, given that the charging and discharging of EVs will deplete the EVs' batteries, the cost of battery depletion is also included.

According to the reference [32], the conditions of their vehicles influence the risk attitudes of EV owners. If the vehicle is new (purchased within two years), the owner tends to be risk-averse, and if the vehicle is old (purchased more than two years ago), the owner tends to be risk-seeking. In the present study, EV owners are the average consumers in prospect theory, with similar value sensitivity to the benefits. The zero-point of the value functions varies with the condition of the vehicle. In the case of low benefits, the psychological value of owners of new vehicles is low, and they prefer not to participate in charge and discharge scheduling. In contrast, the owners of old vehicles prefer to participate actively in the scheduling to fully utilize the EVs.

Agents' willingness to participate in ADR can be determined in this fashion based on the probability of their previous utility value. As a result, based on $p_j(d)$, the agent's participation factor can be simulated.

Because the risk attitude of EV owners influences their participation desire, we apply the value function to express that attitude towards relative value. For ADR event d , the utility value of the vehicle l can be determined as follows:

$$v(l, x)_d = \begin{cases} (\pi_{l,d} - b(l))^{\alpha_0}, & \pi_{l,d} \geq b(l), \\ -\theta(b(l) - \pi_{l,d})^{\beta_0}, & \pi_{l,d} < b(l), \end{cases} \quad (8)$$

where $\pi_{l,d}$ represents the benefit of l for d . $\pi_{l,d}$ in Equation (17) consists of two cost components and

$$\pi_{l,d} = C_{0,l} - C_l(d) \quad (9)$$

where each net benefit component of $\pi_{l,d}$ is the difference between the profit and cost derived from the participation of the agent in DR.

4.3. Constraints

Overcharging and discharging shorten the life of the lithium battery. Thus, the charge state of the battery must be limited, as illustrated in Equations (10)–(12) below. Equations (10) and (11) indicate that the SOC of EVs at each moment should be neither less than the minimum value nor greater than the maximum one. Equation (12) ensures that an EV cannot be charged and discharged simultaneously.

$$S_l(k) = S_l(k-1) + [P_{c,l}\xi_c + P_{d,l}/\xi_d]\omega_{m,l}(k)I_0\Delta t/Q_l \quad (10)$$

$$P_{c,l} \cdot P_{d,l} = 0 \quad (11)$$

$$S_{\min} \leq S_l(k) \leq S_{\max} \quad (12)$$

$$S_{O,l} + \sum_{k \in T_{m,l}} \Delta t \frac{[P_{c,l}\xi_c + P_{d,l}/\xi_d]\omega_{m,l}(k)I_0}{Q_l} \geq S_{E,l} \quad (13)$$

$$0 \leq S_{O,l} \leq 1 \quad (14)$$

$$0 \leq S_{E,l} \leq 1 \quad (15)$$

$$T_{m,l} > T_{c,l}, \quad l = 1, 2, \dots, n \quad (16)$$

$$T_{c,l} = (S_{E,l} - S_{O,l})Q_l/P_{c,l}\xi_c \quad (17)$$

Equations (13)–(17) constrain the minimum charging time of EVs and guarantee that the participation of EVs in the optimal scheduling of aggregator organizations will not affect the regular driving pattern of EV owners. Constraint (13) means that when the EV leaves the parking station, the charge state of the battery should meet the expectation. Constraints (16) and (17) indicate that EVs can carry out repeated charging and discharging, but the total charging time cannot be less than the minimum charging time.

5. Variant Roth–Erev Algorithm

The EV agent decides for the EV owner when considering the owner's attitude towards risk. The price-based optimal charging/discharging strategy leads to the concentration of EV charging, changes the original load curve shape, and affects the real-time electricity prices, making the initial price-based optimal strategy flawed. Therefore, we used the reinforcement learning algorithm to solve the problem of the multiagent game.

We defined $q_j(d)$ as the propensity of the agent l to perform the action a_j ($a_j \in S_l, u_{1,l}(d) = a_j$) on ADR event d . Before the learning phase, we specified the initial propensity and initial selected action. We assumed that $q_j(d)$ takes an initial value $q_j(0)$, and for any l , the initial action is $a_j = u_{1,l}(1)$ (all agents take price-based optimization strategy). It was assumed that $a_l = u_{1,l}(d)$ was acted on ADR event d . After the ADR scheduling, $v_{l,d}$ was fed into the learning algorithm, denoting the reward l for taking action $a_l = u_{1,l}(d)$ in the event d . The action propensity in the ADR event $d+1$ can be calculated by:

$$\begin{cases} q_j(d+1) = (1-r)q_j(d) + (1-Ex)v_{l,d}, & \text{if } a_l = a_j \\ q_j(d+1) = (1-r)q_j(d) + \frac{q_j(d)}{(N-1)}Ex, & \text{if } a_l \neq a_j \end{cases} \quad (18)$$

The experimental parameter controlling the degree of exploration is denoted by the letter Ex . The propensity of not choosing the action of ADR event d was raised by in-

creasing the experimentation parameter value. The retrospective parameter r allows for a “forgetting” effect, enabling the agent to disregard rewards from previous acts since they may become suboptimal over time. The forgetting effect is amplified as r increases.

The calculated result of Equation (18) is the propensity of each action. During the actual calculation, it is necessary to further calculate the probability of each action strategy. In this way, the capture can be performed in the next iteration according to the probability. $p_j(d)$ was calculated using a Gibbs-Boltzmann distribution:

$$p_j(d) = \frac{e^{\frac{q_j(d)}{cp}}}{\sum_{i=1}^N e^{\frac{q_i(d)}{cp}}} \quad (19)$$

The Boltzmann cooling parameter is denoted by CP. The propensity value Q for each strategy is determined in Equation (18). When Q is converted to P , the CP value is used to control the conversion rate, as shown in Equation (19). The total number of activities is denoted by N . Different cases have different values of Boltzmann cooling parameters. When the small $q_j(d)$ value is coupled with a large CP value, the value e of the molecule in Equation (19) is 1. In this case, the changes of $q_j(d)$ value cannot cause variations in probability. Therefore, it is worthwhile to investigate the selection of appropriate parameter values of the Boltzmann cooling according to the actual situation. The participation factor can be determined by variant Roth–Erev algorithm, as shown in Figure 2, the virtual runtime environment (VRE), and the value function.

Figure 2. Variant Roth–Erev Algorithm flow chart.

6. Case Analysis

6.1. Case Parameter

To verify the effectiveness of the study method, a living area with 100 EVs was selected as the sample for case analysis in this paper. Other loads in the living area are regarded as base loads, such as household electricity consumption, air conditioning, and washing machine. In this study, base loads are temporarily excluded from participating in the demand response. Workstations are deployed by an aggregator to direct the charging and discharging of EVs to maximize profit. Since the electricity price fluctuates according to the load, the price of electricity is high when the load is high, and EVs tend to discharge to earn profits. Given the low electricity prices at low loads, EVs tend to store electricity at the lowest possible price for charging. In this way, EVs can participate in demand response, reducing peak-to-valley differences on the demand side and enhancing the safety of the power grid.

According to the National Household Travel Survey (NHTS) of the US Department of Transportation in 2001 [29,30]. As a general living area, these data are of the reference value, so the simulation is carried out based on that. The starting charging time and daily mileage of the general vehicle are generated.

According to reference [31], the EVs in living areas are simplified into three types, as shown in Table 1. These three types represent EVs commonly used by residents, including small, medium, and large by their capacities. Among those EVs, the number of small and medium-sized EVs is high, and the number of large EVs is low, consistent with the actual situation. Despite the growing popularity of EVs, people still choose gas-powered vehicles or gas-electric hybrid vehicles when traveling away. EVs are chosen for short-distance travels, so fewer people choose EVs with large battery capacities.

Table 1. Information about EVs.

Types	Proportion (%)	Battery Capacity (kw•h)
small	40	24
medium	40	30
large	20	53

According to reference [32], the risk attitudes of EV owners can be anticipated, as indicated. The data in the reference illustrates the risk attitude of EV users in Beijing, China. Based on the risk attitudes, the EV owners are divided into two groups in this study. Table 2 shows how parameters of the utility function are set based on various risk preferences.

Table 2. Classification of EV owners.

Classification	51.8% (A)	18.2% (B)	9% (C)	21% (D)
$b(l)$	1	0	1	0
$\alpha_0\beta_0$	0.88z	0.88	1.21	1.21

6.2. Analysis Results

6.2.1. Cost and Profit

During the learning process, different values of Boltzmann cooling parameters result in different learning convergence rates. Six different parameter settings are proposed to find the optimal Boltzmann cooling parameter value applicable to the case in this study.

In this study, Boltzmann cooling parameter values were set to 1, 0.5, 0.25, 0.1, 0.01 and 0.008, with each value iterated for 100,000 times, 200,000 times and 500,000 times. The total charging cost for 100 electric cars and the total profit for 100 electric cars relative to the disordered charging case was calculated. In addition, the case of Boltzmann cooling parameter value taking other parameters was examined. The results showed that too large

parameter values led to the too slow learning convergence, which could not be applied to the actual situation. Too small parameter values tend to hinder the promotion of learning.

As shown in Figure 3, with the decrease of Boltzmann cooling parameter value and the increase of iteration order, the total charging cost of EVs decreases and gradually stabilizes. In Figure 4, as the Boltzmann cooling parameter value decreases and the iteration order increases, the total profit of EV charging increases and gradually stabilizes.

Figure 3. EV cost.

Figure 4. EV profit.

In Figures 3 and 4, when the Boltzmann cooling parameter value is 0.01 and 0.008, the total cost and profit are relatively stable. However, when the Boltzmann cooling parameter value is 0.1 and the number of iterations is 500,000, the total charging cost of EVs is the minimum, and the total profit is the maximum. Therefore, the Boltzmann cooling parameter value of 0.1 was used for the subsequent study.

In the subsequent study, the parameter value of Boltzmann cooling was used 0.1 for further learning, with the algorithm iterated 1,000,000 times. As shown in Figure 5, when the number of iterations is 1,000,000, the total cost increases, and the total profit decreases, indicating that the learning results are unstable when the Boltzmann cooling

parameter value is 0.1. In addition, 1,000,000-time iteration requires higher computation speed and longer computation time of the workstation, which is inconsistent with practical applications.

Figure 5. Costs and profits with Boltzmann cooling parameter value of 0.1.

According to the total cost and profit data, the Boltzmann cooling parameter values of 0.01 and 0.008 represent stable data performance.

6.2.2. Load

In the case analysis of load, the calculation results of 500,000 iterations with different Boltzmann cooling parameter values were chosen for comparison with the data of disordered charging of EVs. Figure 6 shows that the load fluctuations after optimal scheduling are much smaller than those after random charging of EVs, regardless of the parameter values. Therefore, the effectiveness of the proposed method can be verified.

Figure 6. Load fluctuations.

To more accurately show the advantages and disadvantages of the learning results of different Boltzmann cooling parameter values, the load variance is proposed as an index to select the optimal Boltzmann cooling parameter value.

As shown in Table 3, when the Boltzmann cooling parameter value is 0.01, the load variance on the demand side is the smallest, and the charging cost of EVs is the second lowest. Although the current result shows the lowest cost when the parameter is 0.1, the result is unstable with a large load variance, indicating that this choice needs to be excluded.

Table 3. Data results with different parameters.

Classify	1	0.5	0.25	0.1	0.01	0.008	Disorder
Cost	169.74	169.74	163.01	126.66	131.78	133.23	276.98
Load variance	423	423	403	364	354	358	481

According to the above analysis, the best data results can be obtained when the Boltzmann cooling parameter value is 0.01. According to Table 3, as the Boltzmann cooling parameter decreases, the variance of the overall load decreases within 48 h. This result is because the Boltzmann cooling parameter decreases, and the previous learning result has a greater impact on later learning. Within a limited time range, a smaller value of the Boltzmann cooling parameter allows faster searching of the algorithm for better results. However, when the Boltzmann cooling parameter value is too small, the learning algorithm stops with a local optimal result, failing to find the global optimal result. Therefore, it is critical to find the appropriate value of the Boltzmann cooling parameter. Since this paper is focused on day-ahead optimization scheduling, the learning time of the learning algorithm needs to be strictly defined. In consideration of the learning speed and results, 0.01 is selected as the final value of the Boltzmann cooling parameter in this paper. However, the result is better than that of disordered charging regardless of the parameter values, demonstrating the effectiveness and superiority of the algorithm proposed in this study.

When 500,000 iterations are performed, the calculated total charge and discharge cost of EVs is 52.4%, less than the total cost of random charge, with the load variance reduced by 26.4%.

Compared with that of the disordered charging of EV groups, the impact of EV charging on the power grid is subdued through the optimized scheduling method proposed in this paper, as shown in Figure 7. The yellow curve in the figure represents the load after optimized scheduling, and the load peak is lower than that of disordered charging. In addition, the off-peak load is higher than that of disordered charging, proving the effectiveness of the proposed method.

Figure 7. Optimization results were compared with disordered charging.

7. Conclusions

Based on reinforcement learning and prospect theory, a multiagent scheduling learning model is proposed to optimize the charging and discharging strategy of EVs. According to different risk preferences of EV owners, the charging and discharging strategies of EVs are optimized to increase the income of EV owners and alleviate the impact of EV cluster grid connection on the distribution network.

Compared with that of the traditional RE algorithm, the Variant-Re algorithm applied in this paper is more suitable for the optimal scheduling of EVs. The traditional RE algorithm can only deal with the case of positive income. Since EV owners have positive and negative income, the Variant-Re algorithm is more suitable as it can deal with the whole real number field.

Since other reinforcement learning algorithms are not considered in this study, the optimization algorithms will be added in future studies by comparing them with other reinforcement learning algorithms. In addition, the algorithm in this study has a long computation time, which affects the effectiveness of optimization scheduling. In future studies, the improvement of the learning speed and innovation of the algorithm will be considered.

The optimization scheduling model proposed in this paper is applicable to residential areas where EV charging piles are installed. Through the overall management of aggregators, EV charging port switch devices are controlled by intelligent instruments to optimize the charging and discharging scheduling of EV groups in residential areas. In this paper, each EV is regarded as an agent, and charging and discharging activities of multiple EVs are carried out under the management of the same polymer: a multiagent optimization model. The effectiveness of the method is verified by an example. The total charge and discharge cost calculated by the optimization method is 52.4% lower than the total cost of random charging, and the load variance is reduced by 26.4%.

Author Contributions: In this research activity, all authors were involved in the data collection and preprocessing phase, model constructing, empirical research, case analysis, discussion, and manuscript preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Natural Science Foundation of Beijing Municipality under Grant No. 9202017.

Acknowledgments: This paper was completed with the help of many teachers and classmates. We would like to express our gratitude to them for their help and guidance.

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

Nomenclature

J	The number of time phases for the study
Δt	The duration of each time phase
T	The entire amount of time spent on study scheduling
V_l	The vehicle's information
$T_{in,l}/T_{out,l}$	The moment when EV enters/exits a parking lot.
$SO_{c,l}$	The initial state of charge (SOC)
$SE_{e,l}$	The EV leaves expecting SOC
Q_l	The vehicle's battery capacity
$P_{c,l}/P_{d,l}$	The charging/discharging rate of the vehicle l
$\rho(k)$	Real-time electricity price at the time k
$L_{u^*}(k)$	Based on the optimization strategy u^* , the total load on the demand side at the time k
$T_{m,l}$	The period of vehicle l enters the parking station connected to the grid
C_l	The cost of the vehicle l
ε	The battery loss factor

I_0	The charge/discharge indicator variable
$\omega_{m,l}(k)$	The EV battery's maneuverability
$S_l(k)$	The vehicle's SOC that can be scheduled throughout the time period
ξ_c/ξ_d	The charging/discharging efficiency of the battery
$S_{\min}/S_{\max}$	The minimum/maximum of the allowable SOC
$T_{c,l}$	The time it takes to charge the EV in the shortest possible time
n	The number of vehicles
$\alpha_0/\beta_0/\theta_0$	The parameter of the prospect theory utility function
$\pi_{l,d}$	The benefit of EV l for event d
$C_{0,l}$	The disorderly charging cost of EV l

References

- Madura, J.M.; Lieu, D.K. Analysis of an Electromechanical Flywheel for Use as a Dedicated High-Power Device in a Hybrid Electric Vehicle. *Int. J. Electr. Electron. Eng. Telecommun.* **2019**, *8*, 127–132. [CrossRef]
- Krawiec, K. A Concept of Conventional or Mixed Bus Fleet Conversion with Electric Vehicles: A Planning Process. *Int. J. Electr. Electron. Eng. Telecommun.* **2020**, *9*, 8–12. [CrossRef]
- Zhang, Y.; Zhang, X.; Lan, L. Robust optimization-based dynamic power generation mix evolution under the carbon-neutral target. *Resour. Conserv. Recycl.* **2021**, *178*, 106103. [CrossRef]
- Liu, C.; Chau, K.; Wu, D.; Gao, S. Opportunities and challenges of vehicle-to-home, vehicle-to-vehicle, and vehicle-to-grid technologies. *Proc. IEEE* **2013**, *101*, 2409–2427. [CrossRef]
- Kempton, W.; Steven, L. Electric vehicle as a new power source for electric utility. *Transp. Res. D* **1997**, *2*, 157–175. [CrossRef]
- Kempton, W.; Tomic, J.; Letendre, S.; Brooks, A.; Lipman, T. *Vehicle-to-Grid Power: Battery, Hybrid, and Fuel Cell Vehicles as Resources for Distributed Electric Power in California*; Contract UCD-ITS-RR-01-03; Institute of Transportation Studies: Davis, Ca, USA, 2001.
- Nicolo, D.; Aruna, S.; Polak, J.W. Electric vehicle charging choices: Modelling and implications for smart charging services. *Transp. Res. Part C* **2017**, *81*, 36–56.
- Solanke, T.U.; Ramachandaramurthy, V.K.; Yong, J.Y.; Pasupuleti, J.; Kasinathan, P.; Rajagopalan, A. A review of strategic charging–discharging control of grid-connected electric vehicles. *J. Energy Storage* **2020**, *28*, 101193. [CrossRef]
- Kempton, W.; Tomić, J. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *J. Power Sources* **2005**, *144*, 268–279. [CrossRef]
- Vyas, A.; Santini, D. Use of National Surveys for Estimating ‘Full’ PHEV Potential for Oil Use Reduction [EB/OL]. 21 July 2008. Available online: <http://www.transportation.anl.gov/pdfs/HV/525.pdf> (accessed on 25 November 2021).
- Guille, C.; Gross, G. A conceptual framework for the vehicle-to-grid (V2G) implementation. *Energy Policy* **2009**, *37*, 4379–4390. [CrossRef]
- Sovacool, B.K.; Noel, L.; Axsen, J.; Kempton, W. The neglected social dimensions to a vehicle-to-grid (V2G) transition: A critical and systematic review. *Environ. Res. Lett.* **2017**, *13*, 013001. [CrossRef]
- Kong, P.Y.; Karagiannidis, G.K. Charging schemes for plug-in hybrid electric vehicles in smart grid: A survey. *IEEE Access* **2016**, *4*, 2169–3536. [CrossRef]
- John, J.; Antunes, S.; Eduardo, M.; Dos, S.; Ana, P.; Carboni, M.; Bernardon, D.P. Control Strategies for Smart Charging and Discharging of Plug-In Electric Vehicles. *Smart Cities Technol.* **2016**, *1*, 121–141.
- Beck, L.J. *V2G-101: A Text About Vehicle-to-Grid, the Technology Which Enables a Future of Clean and Efficient Electric-Powered Transportation* 332; Leonard Beck: Newark, DE, USA, 2009.
- Tamor, M.A.; Milačić, M. Electric vehicles in multi-vehicle households. *Transp. Res. Part C* **2015**, *56*, 52–60. [CrossRef]
- Chellaswamy, C.; Ramesh, R. Future renewable energy option for recharging full electric vehicles. *Renew. Sustain. Energy Rev.* **2017**, *76*, 824–838. [CrossRef]
- Salisbury, M.; Toor, W. How and why leading utilities are embracing electric vehicles. *Electr. J.* **2016**, *29*, 22–27. [CrossRef]
- Wenzel, G.; Negrete-Pincetic, M.; Olivares, D.E.; MacDonald, J.; Callaway, D.S. Real-time charging strategies for an electric. *IEEE Trans. Smart Grid* **2018**, *9*, 5141–5151. [CrossRef]
- Seddig, K.; Jochem, P.; Fichtner, W. Integrating renewable energy sources by electric vehicle fleets under uncertainty. *Energy* **2017**, *141*, 2145–2153. [CrossRef]
- Mousavi, S.M.; Flynn, D. Controlled charging of electric vehicles to minimize energy losses in distribution systems. *IFAC-Pap. Line* **2016**, *49*, 324–329. [CrossRef]
- Galus, M.D.; Andersson, G. Demand management of grid connected plug-in hybrid electric vehicles (PHEV). In Proceedings of the IEEE Energy 2030 Conference, Atlanta, GA, USA, 17–18 November 2008.
- Powell, W.B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality (Probability and Statistics)*; Wiley: Hoboken, NJ, USA, 2007.
- Shi, W.; Wong, V.W. Real-time vehicle-to-grid control algorithm under price uncertainty. In Proceedings of the International Conference on Smart Grid Communications (SmartGridComm), Brussels, Belgium, 1–20 October 2011; pp. 261–266.
- Qiu, D.; Ye, Y.; Papadaskalopoulos, D.; Strbac, G. A Deep Reinforcement Learning Method for Pricing Electric Vehicles with Discrete Charging Levels. *IEEE Trans. Ind. Appl.* **2020**, *56*, 5901–5912. [CrossRef]

26. Vandael, S.; Claessens, B.; Ernst, D.; Holvoet, T.; Deconinck, G. Reinforcement Learning of Heuristic EV Fleet Charging in a Day-Ahead Electricity Market. *IEEE Trans. Smart Grid* **2015**, *6*, 1795–1805. [[CrossRef](#)]
27. Lu, R.; Hong, S.H.; Zhang, X. A Dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach. *Appl. Energy* **2018**, *220*, 220–230. [[CrossRef](#)]
28. Bahrani, S.; Parniani, M.; Vafaeimehr, A. A modified approach for residential load scheduling using smart meters. In Proceedings of the IEEE Conference Innovative Smart Grid Technologies Europe (ISGT Europe), Berlin, Germany, 14–17 October 2012.
29. Bessembinder, H.; Lemmon, M.L. Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. *J. Finance* **2002**, *57*, 1347–1382. [[CrossRef](#)]
30. Zeng, B.; Wu, G.; Wang, J.; Zhang, J.; Zeng, M. Impact of behavior-driven demand response on generation adequacy in smart distribution systems. *Appl. Energy* **2017**, *202*, 125–137. [[CrossRef](#)]
31. Xiao, L.; Mandayam, N.B.; Poor, H.V. Prospect Theoretic Analysis of Energy Exchange Among Microgrids. *IEEE Trans. Smart Grid* **2014**, *6*, 63–72. [[CrossRef](#)]
32. Pan, L.; Yao, E.; MacKenzie, D. Modeling EV charging choice considering risk attitudes and attribute non-attendance. *Transp. Res. Part C Emerg. Technol.* **2019**, *102*, 60–72. [[CrossRef](#)]