Stochastic charging scheduling of a parking lot with wind power: a state-aggregation method based on Markov decision processes

Fangzhu Ming | Feng Gao | Kun Liu | Yuzhou Zhou

Faculty of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an 710049, China

Correspondence
Kun Liu, Faculty of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an 710049, China.
Email: kliu@sei.xjtu.edu.cn

Abstract
Optimal charging scheduling of electric vehicles with renewable energy could greatly save the electricity cost in parking lots, and also helps for environment protection. However, uncertainties of renewable energy and electric vehicles charging demand bring great challenges to obtaining the optimal charging policy. To this end, in this paper, a state aggregation based dynamic programming method is presented for such a stochastic charging scheduling problem. Specifically, first, a new Markov decision process based formulation is established for the multi-stage stochastic programming problem. Second, a novel state-aggregation method is proposed to relieve the dimension curse due to large state-action space without sacrificing the optimality of the charging policy. Besides, the consistency of the problem before and after state aggregation also could be verified by proposed theorems both on feasibility and optimality. Moreover, economic efficiency and computational efficiency are both improved in numerical testing, which shows the effectiveness of the proposed method.

1 INTRODUCTION

To alleviate the environmental pollution problem, electric vehicles (EVs) have been developed vigorously due to their zero-carbon emission characters [1–3]. Meanwhile, the integration of renewable energy also helps for environmental protection. To this end, optimal scheduling of EVs with renewable energy has attracted more and more attention and becomes an urgent problem.

Charging EVs with renewable energy is significant in both environmental protection and energy saving. In recent years, the studies of charging EVs with renewable energy are mainly from two main aspects. One is the safe operation of the power grid, which focuses on the centralized schedule of EV groups. They usually schedule the charging time and quantity of EVs to transfer the charging load according to the amount of renewable energy [4–9]. The other one is the economic dispatch of EVs which decides the charging behaviour of EVs to minimize the charging cost [10, 11]. In this paper, we focus on the economical schedule of multi-EVs coordinating and propose an effective method to solve the problem considering the uncertainties of both EVs charging demand and renewable energy.

For the charging problem, most of the existing works focus on EVs. However, these charging policies based on EVs require the high cooperation of EV users, it is hard to obey for all users sometimes. Thus, we try to schedule the charging problem from the point of charging piles and conduct the research based on the charging piles in the parking lot. Scheduling the charging behaviours of charging piles in the parking lot is meaningful. With the growth of EVs, parking lots are the most common charging place in the city [12–14]. And it is more feasible to manage the charging piles than to control the EV users for the manager. In this paper, we take the charging scheduling in an enterprise private parking lot as the application scenario, considering the willingness of EV users to cooperate and the regularity of EV journey in such parking lots.

To solve the charging scheduling problem in the parking lot, there are three main challenges as follows. First, there exist large uncertainties on both the charging demand of EVs and renewable energy, which generates a large number of charging states. Second, the problem is time-coupled and multi-processes. Each EV has a charging deadline which means the current charging decision would affect future decision-making. And wind power breaks the independence of charging piles, which turns...
the charging process of a single charging pile into co-operation charging processes of multi charging piles. Third, a large number of uncertainties make it hard and time-consuming to solve the stochastic problem.

To address the challenges above, we model the multistage stochastic optimization problem as a Markov decision process (MDP). It is very clear to represent the stochastic in the problem and illustrate the relation between the adjacent stages. Many literatures also model the charging problem as MDP, but they decide the total charging power for an EV aggregation, which cannot guarantee the satisfaction of the charging constraint of each EV. In [15], it models the profit-maximizing joint charging problem for a public charging station as an MDP, and decides the total charging rate for all EVs. In [16], the joint optimization problem of EVs is modelled as MDP, which aggregates the load of all EVs and decides the total charging power at each stage. However, the MDP model in our paper describes a joint-charging problem of EVs, whose action is a high-dimension vector composed of the charging action of each EV, and its state is the combination of the states of all EVs. Thus, the MDP directly decides the action of each EV and chooses the optimal action in a feasible action zone calculated by the action constraint to ensure the feasibility. There are also a few literature building the MDP model considering the state of each EV for the joint charging scheduling. They usually take the parking time and charging demand of each parked EV as states [17, 18]. We add a new state for each EV which called the occupied state indicating the parking state of each EV. It can better describe the randomness of the arrival of EVs. However, solving MDP is another difficulty due to the dimension curse when the state space and action space are very large. Dynamic programming (DP) is the traditional method to solve MDP problems which would get the global optimal solution [19]. But it is time-consuming to solve the problems with a large number of states. In [20], the stochastic dual DP (SDDP) method is used to solve the charging scheduling problem considering the uncertainty of future charging demand. It takes several hours to solve the problem. Approximately dynamic programming (ADP) is another kind of method to solve the MDP problems [21], which reduces the complexity of computation by sacrificing part of the solution accuracy. The ADP methods to solve the charging problems can be divided into three categories. The first one is to define a deterministic charging rule which has little complexity to make a decision. In [22], an energy scheduling problem of EVs in the residential distribution network is modelled as MDP, and it proposes a rule-based charging policy according to the contribution of EVs to guide the charging action. The second one is called the lookahead policies, which simulates the state transition and action selection in the future for a period of time to evaluate the value function such as Monte Carlo simulation (MC) and rollout algorithm. In [17, 18], a simulation-based policy improvement method is proposed to solve the charging scheduling problem, which samples a large number of scenarios to approximate the value function. In [23], a two-stage approximately dynamic programming algorithm is proposed to obtain the optimal charging policy for EVs in a commercial building parking lot. It reduces the computation complexity by generating a large number of deterministic samples containing the information of stochastic vehicle arrival and departure as well as dynamic electricity price. The third one is the method of function estimation, which fits the value function directly by constructing a parameterized function [24]. With the development of neural network (NN), it usually combines the traditional methods with NN to reduce the difficulty of feature construction. In [25], it models the charging problem of EVs as two-level MDP and the proximal policy optimization (PPO) is taken to solve both the lower and upper level MDP problem. In [26], a safe deep reinforcement learning (SDRL) is proposed to solve the constraint charging scheduling problem to minimize the charging cost. However, the rule-based method cannot deal with the dynamics of the future stages, which causes a solution error in stochastic problems. The accuracy of lookahead policies is related to the number of samples, and it is difficult to balance the accuracy and calculation time. The estimation of the value function takes a long time to train the function and is easy to fall into a local optimal solution. While our methods can get the global optimal solution at an acceptable cost of time.

In this paper, we aim at getting an optimal charging policy to minimize the cost of purchasing electricity from the power grid. And a novelty state-aggregation method is proposed to decrease the computation complexity and the state aggregation based dynamic programming method (SABDP) method is proposed to obtain the optimal charging policy. The main contributions in this paper are as follows.

- MDP is taken to model the joint-charging scheduling problem of multi-EVs to describe the uncertainties of each EV and the time-coupled relation clearer.
- A novelty state aggregation method is proposed which aggregates the states by merging the states with the same constituents elements but a different arrangement.
- SABDP is proposed to solve the MDP problem, and numerical experiments show that it gets the global optimal value as well as the optimal charging policy with less calculation time and space cost compared with other classic methods.

The rest of the paper is organized as follows. Section 2 formulates the proposed optimal charging model, and the state aggregation method and solution method are introduced in Section 3. The numerical testing is implemented in Section 4 and Section 5 concludes the paper.

2 | MDP MODEL FOR MULTISTAGE STOCHASTIC SCHEDULING PROBLEM

2.1 | System description and assumption

In this paper, we focus on a stochastic charging scheduling problem in the parking lot. There are $N$ fast charging piles, AC/DC converters and small wind turbines in the parking lot, shown in Figure 1. The charging energy can be obtained directly from both the wind generators and purchased from the power grid. The cost of wind power is regarded as zero. Each EV
submits the charging time and charging demand to the parking lot after parking. Then an intelligent scheduler in the parking lot decides the charging actions of EVs, considering the charging information of EVs and power resources. A basic consensus should be obeyed that the charging behaviour of EVs during the parking period can be completely controlled by the scheduler and the charging station must meet the charging demand of each EV with a low charging cost before the EV leaves. Moreover, the parking lot can obtain the distribution of charging demand and the wind power output by statistical historical data to predict the future uncertainty on both the EVs and the wind power.

The model is established on the following assumptions.

1) The output power of the charging pile is constant.
2) The charging process of EVs can be interrupted.

Assumption (1) simplifies the charging decisions to only decide which charging pile to work while not decides the charging power [17]. Assumption (2) means that the smart scheduler allows a discontinuous charging process and flexible control to cope with the fluctuation of renewable energy and electricity price for better economic performance in the parking lot. This kind of charging mode is mainly used in research on the economics of EV charging currently [14, 15, 25].

2.2 Markov decision processes model

MDP is used to model the stochastic dynamic programming problem. An MDP model contains 5 factors < S, A, P, R, γ > , which are decision time, state space, action space, state transition matrix, reward function and discounted factor [27]. The established MDP based model for a general charging scheduling problem in the parking lot is described as follows.

2.2.1 State space

State space S contains all the possible states of the problem. Each state for the joint-scheduling problem is a high-dimension state, consisting of the state of the charging piles and renewable energy, denoted as $S = [s_1^i, s_2^i, ..., s_N^i, v_i]$. The state in our paper is defined as $S = [S_{CP}, S_{W}^i]$, $S_j \in S_j$ for simplicity. $S_{CP}$ is formulated as $S_{CP} = [s_1^j, s_2^j, ..., s_N^j]$, where $t = 1, 2, ..., T$. The state of each charging pile is denoted as $s_i^j = [a_i^j, l_i^j, \ell_i^j]$. The specific definition of state $s_i^j$ is as follows.

$$d_i = \begin{cases} 1, & \text{if occupied} \\ 0, & \text{if not charging} \end{cases}$$

(1)

$$\ell_i^j = \begin{cases} (\ell_i^j > 0, \ell_i^j \geq 0), & \text{if } d_i = 1 \\ (\ell_i^j = 0, \ell_i^j = 0), & \text{if } d_i = 0 \end{cases}$$

(2)

$$S_{W}^i = \begin{cases} W_{cap}, & \text{if } v_{rated} < v_i \leq v_{out} \\ W_{cap} \left( \frac{v_i}{v_{rated}} \right) ^{3}, & \text{if } v_{low} < v_i \leq v_{out} \\ 0, & \text{otherwise} \end{cases}$$

(3)

For the convenience of discretization, $S_{W}^i$ is mainly considered to be affected by the velocity of wind. The calculation formula is shown in Equation (3).

2.2.2 Action space

The joint action of the problem is to decide which charging pile to work at stage $t$, defined as $A_j = [d_1^j, d_2^j, ..., d_N^j], A_j \in A_j$.

$$d_i = \begin{cases} 1, & \text{if charging} \\ 0, & \text{if not charging} \end{cases}$$

(4)

if $d_i^j = 1$, the $i$th charging pile is chosen to work.

2.2.3 State transition matrix

$P_i$ describes the transition probability between states of the adjacent decision stage. Before giving the state transition matrix, state transition functions are obtained according to the physical characteristics of states.

$$\ell_{i+1}^j = \begin{cases} 1, & \text{if } \ell_i^j \geq 0, \ell_i^j + 1 > 0 \\ 0, & \text{if } \ell_i^j \geq 0, \ell_i^j + 1 = 0 \end{cases}$$

(5)

$$\alpha_{i+1} = \begin{cases} \ell_i^j - 1, & \text{if } d_i = 1 \\ \alpha_i^j, & \text{if } d_i = 0, d_{i+1} = 1 \\ 0, & \text{if } d_i = 0, d_{i+1} = 0 \end{cases}$$

(6)

$$\phi_{i+1} = \begin{cases} \ell_i^j - d_i^j P_i \Delta t, & \text{if } d_i = 1 \\ P_i^j, & \text{if } d_i = 0, d_{i+1} = 1 \\ 0, & \text{if } d_i = 0, d_{i+1} = 0 \end{cases}$$

(7)

Equation (5) describes the transition of the occupancy state, where $d_i = 1$ represents the charging pile is occupied. $\Delta t$ is the decision interval. Equations (6) and (7) are the state transition...
funtion of $l'_i$ and $\ell'_i$. The first lines in Equations (6) and (7) show that the transition process is deterministic for the $i$th charging pile, if it has been parked at time $t$. The second lines indicate that the $l'_i$ and $\ell'_i$ of the $i$th charging pile in the future is uncertain if the charging pile is not occupied at time $t$. $\alpha'_{t+1}$ and $\beta'_{t+1}$ are two random values following a certain distribution, representing the possible parking time and charging demand.

The transition probability of $l'_i$ is calculated by Equation (8), which is the joint probability of $\ell'_i$,

$$p(\ell'_{t+1} \mid l'_{t}, a'_t) = p(l'_{t+1} \mid \ell'_{t}) \times p(l'_{t+1} \mid l'_{t}) \times p(\ell'_{t+1} \mid \ell'_{t})$$ (8)

The transition probability of $\ell'_i$, $l'_i$, $\epsilon'_i$ is defined as follows.

$$p(\ell'_{t+1} \mid \ell'_{t}) = \begin{cases} 1, & \text{if } l'_t > 0 \\ \xi'_{t+1}, & \text{if } l'_t = 0 \end{cases}$$ (9)

$$p(l'_{t+1} \mid l'_{t}) = \begin{cases} l', & \text{if } \ell'_t = 1 \\ t'_{t+1}, & \text{if } \ell'_t = 0 \end{cases}$$ (10)

$$\epsilon'_{t+1} = \begin{cases} 1, & \text{if } \ell'_t = 1 \\ \gamma'_{t+1}, & \text{if } \ell'_t = 0 \end{cases}$$ (11)

Where, $t'_{t+1}$ obeys the normal distribution $N(\mu_{t+1}, \sigma_{t+1})$, $\gamma'_{t+1}$ is regarded as the Chi-square distribution $\chi^2(k)$. The transition of $S_{W,t}$ is regarded as a Markov process, which means the wind power is transferred to a random value according to a probability distribution, and the transition probability is denoted as $p(S_{W,t+1} \mid S_{W,t})$.

Therefore, the total transition probability that combines the transition probability of $S_{W,t}, S_{E} T , t$ is shown in Equation (12).

$$p(S_{t+1} \mid S_{t}, A_{t}) = p(S_{W,t+1} \mid S_{W,t}) \prod_{i=1}^{N} p(\epsilon'_{t+1} \mid \epsilon'_{t}) \times p(l'_{t+1} \mid l'_{t}, a'_t)$$ (12)

2.2.4 Reward function

The reward function calculates the immediate reward after taking an action and reflects the worth of the action. In this paper, the reward is the expected purchasing cost of electricity.

$$R(S_{t+1} \mid S_{t}, A_{t}) = \sum_{S_{t+1}} p(S_{t+1} \mid S_{t}, A_{t}) \times P(r(t))$$ (13)

$$\times \max(P_i \sum_{i=1}^{N} d'_i - S_{W,t}, 0)$$

2.2.5 Objective function

The goal is to get the optimal charging policy to minimize the total expected cost of purchasing electricity, shown in Equation (14).

$$\min J(S_0) = E_{\pi} \sum_{t=1}^{T} [P(t) \max(P_i \sum_{i=1}^{N} d'_i - S_{W,t}, 0)]$$ (14)

Where $E(\cdot)$ is the expected purchasing cost under the strategy $\pi$. The first constraint restricts the total charging power at stage $t$ to be less than the bus power $P_{\text{max}}$. The second one is the action constraint which represents the charging demand must be satisfied before the EV leaves. $\Delta t$ is the maximum redundancy time, which is the maximum time difference between the occupancy time and the actual charging time.

Bellman function is taken to iteratively calculate the optimal value in (15), where $V^* = V^*(S_t)$ is the optimal value. Its the basis of getting the optimal results in MDP.

$$V^*(S_t) = \min_{\Delta t \in A} R(S_t, A_t) + \sum_{S_{t+1} \in S} P(S_{t+1} \mid S_t, A_t) \times V^*(S_{t+1})$$ (15)

The established MDP based model is shown as Equations (1)–(14). However, the explosion problem of the combination of states makes it hard to get the optimal results of the MDP problem. Thus, more efficient methods need to be proposed.

3 SOLUTION METHODOLOGY

Solving the MDP problem may be trapped in the curse of dimension especially when the problem scale is large. In order to reduce the complexity of calculation caused by the large scale of states, we propose an optimality-consistency state aggregation method (OCSAM) to reduce the number of states in the problem while retaining the optimality of the solution. Moreover, we propose an improved dynamic programming method based on the OCSAM to improve the solution efficiency of the problem.

3.1 The optimality-consistency state aggregation method

State aggregation is quite important for reducing the computational burden of large-scale MDP problems. Different from existing literature [28, 29], a novel state aggregation method is proposed in this subsection, without sacrificing the accuracy of
the optimal results, and the consistency of results precision also could be guaranteed in the proposed method.

The main difficulty in solving the charging problem comes from the state combination explosion of multiple charging piles. Each $S_{C_P}$ is with three elements and each element has several discrete values. The number of the state of a charging pile is shown in Equation (16).

$$N_i^j = N_i^0 \times N_i^j \times N_i^f$$

(16)

$$N_0^j = N_0^0 \times \prod_{i=1}^{N} N_i^j$$

(17)

It is shown that the state number is exponential growth with the number of charging piles increasing in Equation (17). If traversing through all the states, it is time-consuming and high space-consuming. In this paper, we reduce the states by making use of the symmetry of charging piles, that is the states are generated by state combination instead of state permutation. We propose two theorems to prove the feasibility of the method and the consistency of the results before and after aggregation. The proof process is in the Appendix.

**Theorem 1.** $\forall S_i^j, S_j^t \in S_t$, if $S_i^j \neq S_j^t$, $\text{Sort}(S_i^j) = \text{Sort}(S_j^t)$, then $V(S_i^j) = V(S_j^t)$, where $\text{Sort}(\cdot)$ is as follows.

$$\text{Sort}([s_i^j, s_j^t]) = \begin{cases} [s_i^*, s_j^*], & \text{if } t^* \leq l^*_i, v^*_j < e^*_j \\ [e^*_i, l^*_j], & \text{otherwise} \end{cases}$$

(18)

$\text{Sort}(\cdot)$ is a sorting function which arranges the state through ascending order of $l$ and $e$. Theorem 1 proves that any state with the same composition $[s_i^j, s_j^t, \ldots, s_n^t]$ has the same $V$ value whatever the arrangement order of $s_i^j$ is.

**Theorem 2.** Supposing $S_t$ and $S_i^t$ are the state sets and the aggregation state sets, where $S_t = \{S_i^j, S_j^t, \ldots, S_n^t\}$, $S_i^t = \{S_i^j, S_j^t, \ldots, S_n^t\}$, $t = 1, 2, \ldots, T$. If $\text{Sort}(S_i^j) = \text{Sort}(S_j^t)$, then:

$$\begin{align*}
E^*(V(S_i^j)) &= E^*(V(S_j^t)) \\
P(S_{i+1}^j|S_i^j) &= N_0^{G_{0i}} P(S_{i+1}^j|S_i^j)
\end{align*}$$

(19)

Where $S_i^j$ is the substitution state for a set of states $G_{0i}$ that have the same composition with $S_i^j$ in $S_t$. It shows the state aggregation method does not influence the optimality of results.

We take a small scale case to illustrate the aggregation method. Suppose there are three charging piles, the aggregation rules are written as follows, where $a, b, c$ represent three different states for a charging pile.

$$R = \begin{cases} r_1 = [a, b, c], [a, c, b], [b, a, c], [b, c, a], [c, a, b], [c, b, a] \\ r_2 = [a, a, b], [a, b, a], [b, a, a] \\ r_3 = [a, a, a] \end{cases}$$

(22)

**Algorithm 1** The optimality-consistency state aggregation method

1: Calculate all the aggregation rules, denoted as $R$. Each rule shows a set of state combination forms, recorded as $r$.
2: Aggregate the states into $M$ aggregations according to the rule $r_i, r_j \in R$ and all the states in an aggregation $G_{0i}^m$ have the same constitutes.
3: Replace all the states in $G_{0i}^m$ by any state $S_i^j \in G_{0i}^m$, and the state transition probability is rewritten as below.
4: Repeat step (3) until calculating all aggregations.
5: The total number of states at stage $t$ after reduction is calculated by the following formula, where $x$ is the number of different kinds of $s_i^j$ in $S_i^j$.

**Algorithm 2** State-aggregation based dynamic programming

1: Given the state space $S_t$, transition matrix $P_t$ of each charging pile and the initial state $S_0$.
2: Generate all the possible states of each charging pile from $t$ to $T$.
3: if $t = T$, generate the aggregated state space $S_T'$ and initial the value of all states as $V(S_T') = V_0$, $S_T' \in S_T$.
4: if $t > 0$, set $i = t-1$, aggregate the states and update state space $S_i$ and transition matrix $P_i$. Then calculate the optimal value of each state in $S_i$ by the Bellman function Equation (15).
5: if $t = 0$, end the iteration.

In Equation (22), $r_1$ shows all the combinations with three different kinds of state and $r_2, r_3$ are the combination rule of two and one kind of states. The states satisfying the same rule $r_i \in R$ are aggregated in a group and replaced by one state in the group.

### 3.2 State-aggregation based dynamic programming method

SABDP combines the OCSAM with DP, which reduces the computational complexity while maintaining the optimality of results. The method is presented in Algorithm 2.

### 4 NUMERICAL RESULTS

In this section, we conduct the simulation experiments of charging scheduling in the parking lot of an enterprise and analyse the experiment results from three aspects to show the advantages of the SABDP method. First, we get the optimal expected charging cost under different scenarios, which shows that our method is flexible and universal. Second, we compare the time and space complexity with those of DP in solving the charging problem to verify that our method could reduce the scale of state space greatly and improve computational efficiency. Third, we compare the SABDP method with two experienced charging policies and two ADP methods to show the optimality of decisions in our charging policy.
TABLE 1 Parameters of normal distribution for occupation time

| Time | μ  | σ  | Time | μ  | σ  | Time | μ  | σ  |
|------|----|----|------|----|----|------|----|----|
| 1    | 8.84 | 4.589 | 9    | 5.00 | 3.81 | 17   | 8.19 | 7.24 |
| 2    | 7.31 | 3.29  | 10   | 3.94 | 4.09 | 18   | 8.42 | 7.01 |
| 3    | 5.52 | 3.44  | 11   | 4.32 | 5.19 | 19   | 10.50| 5.34 |
| 4    | 5.08 | 4.04  | 12   | 4.58 | 5.47 | 20   | 10.15| 5.03 |
| 5    | 4.22 | 3.88  | 13   | 4.81 | 5.75 | 21   | 9.98 | 4.79 |
| 6    | 6.43 | 4.19  | 14   | 5.60 | 6.78 | 22   | 10.13| 6.14 |
| 7    | 6.56 | 3.62  | 15   | 6.03 | 6.72 | 23   | 9.98 | 4.79 |
| 8    | 5.85 | 3.52  | 16   | 7.40 | 7.21 | 24   | 9.68 | 4.92 |

4.1 Parameters

The parking lot equips with fast charging piles. Considering the cost of charging piles, the number of fast charging piles in a parking lot is usually less than eight [30]. For a fast charging pile, it takes about 30 min to 2 h to fully charge an EV [17]. We take 2 h as full charge time according to the parameters of BYD e6 and the decision interval is 20 min [31].

Many works have proved that wind power is mainly related to wind speed. The wind speed is regarded as obeying Weibull distribution. The shape and ratio parameters of Weibull distribution are 1.309 and 7.0576 got from the data of the National Renewable Energy Laboratory of the United States fitted in [17]. For simplicity, we divide the output of wind power into six levels and the wind power is calculated by Equation (23). At each time, the wind power may choose any value in the six wind power values with different probability, and it means that the number of states after transferring is at least six.

\[ S_{W,i} = i \cdot P_{W}, i = 1, 2, ..., 6 \]  
(23)

The occupancy time obeys normal distribution. Table 1 shows the means μ and variance σ at different hours. The initial charging demand is correlated with \( d_t \), which obeys the Chi-squared distribution. The variance \( \nu \) is 8.488 [17]. The relation between charging demand and mileage is shown in Equation (24).

\[ e_t = d_t \cdot e_c \]  
(24)

We take the time-of-use (TOU) electricity price of Xian [32]. Some important parameters are shown in Table 2.

4.2 Experiment results and analysis

4.2.1 Optimal charging cost and charging policy

We conduct the experiments under different wind power, different decision stages and different scales of charging piles and get the optimal charging cost respectively.

| PW | Cost (RMB) | AW (kWh) |
|----|------------|----------|
| 0  | 105.1      | 70.7     |
| 5  | 259.1      | 144.3    |
| 12 | 342.0      | 234.0    |
| 15 | 484.6      | 306.5    |

TABLE 2 Parameters table

| Parameter | Value | Parameter | Value |
|-----------|-------|-----------|-------|
| \( \lambda \) | 0.5483 | \( \lambda \) | 0.2983 |
| \( RMB/kWh \) | (8:00–22:00) | \( RMB/kWh \) | (23:00–7:00) |
| \( W^{max} \) | 60 kWh | \( v_{total} \) | 15 m/s |
| \( P_i \) | 30 kW/h | \( v_{iu} \) | 3.5 m/s |
| \( d_t \) | 0.2 kWh/km | \( v_{out} \) | 3.5 m/s |
| \( P_{W} \) | 30 kW/km | \( \Delta t \) | 20 min |

FIGURE 2 The evolution of charging process

In Figure 2, it shows the evolutionary process from the initial state to the terminal states. The blue circles at stage \( t \) are the \( V \) values of states that \( S_{t-1} \) can transfer to and the red solid point is the expected value of all states at stage \( t \). The line in black is one of the deterministic charging scenarios sampling from the charging process. We analyse the expected charging cost based on the whole charging process and analyse the optimal charging decisions based on the sampled scenarios.

Table 3 shows the optimal expected cost and the abandoned wind power (AW) under different charging scenarios calculated by SABDP. It also shows the relation about the optimal charging cost and the wind power. From the table, we get the following results: (1) The optimal cost decreases greatly with the increase of wind power. Taking the seven charging piles and 72 decision stages as an example, compared with the charging cost

TABLE 3 The optimal cost and wind abandon rate of different wind power

| PW | Cost (RMB) | AW (kWh) |
|----|------------|----------|
| 0  | 105.1      | 70.7     |
| 5  | 259.1      | 144.3    |
| 12 | 342.0      | 234.0    |
| 15 | 484.6      | 306.5    |

FIGURE 2 The evolution of charging process

In Figure 2, it shows the evolutionary process from the initial state to the terminal states. The blue circles at stage \( t \) are the \( V \) values of states that \( S_{t-1} \) can transfer to and the red solid point is the expected value of all states at stage \( t \). The line in black is one of the deterministic charging scenarios sampling from the charging process. We analyse the expected charging cost based on the whole charging process and analyse the optimal charging decisions based on the sampled scenarios.

Table 3 shows the optimal expected cost and the abandoned wind power (AW) under different charging scenarios calculated by SABDP. It also shows the relation about the optimal charging cost and the wind power. From the table, we get the following results: (1) The optimal cost decreases greatly with the increase of wind power. Taking the seven charging piles and 72 decision stages as an example, compared with the charging cost...
with no wind, the optimal cost is reduced by 62.28%, 93.16% and 96.36% as the wind power changes from 5 kW to 12 kW.

(2) Appropriate allocation of wind power is significant to reduce the charging cost and improve energy utilization. If the wind power is too large, the charging demand can be totally met by wind power and it leads to waste of wind power. If the wind power is too small, it buys the electricity to meet the remaining demand, which causes a large charging cost. However, when the wind power is close to the maximum charging demand of EVs, there is a good balance between the charging cost and the utilization of wind power, which makes the optimization of EVs more significant.

In Figure 3, we contrast the charging policies under different wind power to demonstrate the impact of wind power. The charging scenarios start from the same initial state and evolve according to their own charging policies. At each stage, the green line and red line are the maximum and minimum charging energy the charging piles could choose, which form the feasible zone of action. The blue line is the actual charging energy and the light blue line is the wind power. The optimal charging decision is mainly affected by wind power, electricity price and action constraint. When the wind power is far less than the charging demand, as shown in Figure 3(a), it has to buy extra electricity after fully using the wind power almost all the time, which means the wind has little with charging actions. The charging policy is mainly affected by the electricity price so that it tends to buy as much electricity as possible at a low price before stage 22. When the wind power is large enough as shown in Figure 3(c), the charging decision is almost the same as the myopic policy, which chooses the charging action making the current charging cost the lowest. It considers little about the fluctuation of wind power in the future. The results show that the charging policy obtained by SABDP is very close to the experienced charging policies when the wind power is too small or too large. There seems unnecessary to take SABDP to get the global optimal solution since the experienced charging policies can get a near-optimal solution with less complexity. However, the charging policy considers both the information of wind power and the electricity price when the wind power is appropriate shown in Figure 3(b), which reduces the charging cost obviously. For example, comparing Figure 3(a,b) at decision stage three, the maximum charging demand is 60 kWh, which means all charging piles have the charging demand. The charging policy of Figure 3(a) chooses to charge all charging piles because there is no need to delay the charging demand to the next stage since the wind power is inadequate all the time. However, in Figure 3(b), it charges less EVs at stage three when the wind power is small. And it delays some of the charging demand to stage four which has larger wind power to avoid purchasing electricity at stage three.

4.2.2 Time and space complexity

State aggregation reduces the scale of state-space greatly. As the number of states decreases, the computation complexity is also reduced. We contrast DP and SABDP on both the space-time complexity and solution accuracy. The results indicate that the SABDP method effectively reduces time and space complexity without affecting the accuracy of the results.

In Table 4, we contrast the number of states before and after state aggregation. The SABDP method cuts down the hidden danger of dimension disaster to some degree because it saves the storage space greatly. Especially, as the number of charging piles increases, the reduction ratios of states reach 99.99%.

The charging cost and computation time are shown in Table 5. With the increase in the number of charging piles, the computation time of SABDP is significantly lower than that of DP. And SABDP expands the problem scale that DP can solve. Moreover, the optimal charging cost of DP and SABDP...
TABLE 4  The number of states before and after reduction

| N  | EV number | Original states | Optimized states | Reduction ratio |
|----|-----------|-----------------|------------------|-----------------|
| 1  | 21        | 18              | 18               | 0%              |
| 2  | 42        | 324             | 171              | 47.22%          |
| 3  | 63        | 5832            | 1140             | 80.45%          |
| 4  | 84        | 104976          | 5985             | 94.30%          |
| 5  | 105       | 1889568         | 26334            | 98.61%          |
| 6  | 126       | 34012224        | 100947           | 99.70%          |
| 8  | 168       | 1.1020e10       | 1081575          | 99.99%          |
| 9  | 189       | 1.9386e11       | 3124550          | 99.99%          |
| 10 | 210       | 3.57e12         | 8436285          | 99.99%          |

TABLE 5  The computation time of DP and SABDP

| N  | T   | Cost (s) | Computation time(s) |
|----|-----|----------|----------------------|
| 3  | 24  | 0.95     | DP 0.95 SABDP 1.9   |
| 48 | 2.20| 2.20     | DP 1.9 SABDP 3.3   |
| 72 | 2.87| 2.87     | DP 2.8 SABDP 4.6   |
| 24 | 24  | 8.91     | DP 4.9 SABDP 11.2  |
| 48 | 8.91| 8.91     | DP 8.9 SABDP 16.8  |
| 72 | 14.99| 14.99   | DP 14.9 SABDP 23.9 |
| 7  | 24  | 28.63    | DP 28.6 SABDP 34.2 |
| 48 | 46.23| 46.23   | DP 46.2 SABDP 62.5 |
| 72 | 56.09| 56.09   | DP 56.1 SABDP 62.5 |

are exactly the same. It proves that SABDP could get optimal results while reducing the calculation complexity.

4.2.3  Comparison of charging policies

To clearly show that the SABDP method could get the optimal charging policy to get the minimum charging cost more clear, we compare the charging policy and charging cost of SABDP with those of the other two experience charging policies and two ADP methods. For a more concise description, we define the methods as follows.

PCP: Plug and charge policy, which means that if an EV plugs in, charge it immediately.

IRP: The minimum immediate reward policy chooses the action to make sure the one-step reward minimum.

MC: Monte Carlo method, a simulation-based method that approximates the charging cost by sampling a large number of deterministic scenarios.

DQN: Deep Q-learning network, a deep reinforcement learning method, which gets the charging cost and the charging strategy by estimating the value function of states.

Table 6 shows the optimal expected cost under three charging policies with the wind power $P_w = 12 kW$. S/P means the comparison of the SABDP and PCP. And S/I is the comparison of SABDP and IRP. The results show that the cost of SABDP cuts down by 65.51% and 5.92% on average compared with PCP and IRP. PCP always charges with the maximum charging demand and does not consider the current wind power and electricity price. Thus it may buy electricity with a high electricity price when wind power is small and abandons wind power when the charging demand is far less than the wind power. IRP guarantees the immediate optimal reward, but it cannot ensure the global optimal value for it does not consider the effect of future wind power and electricity price.

Table 7 shows the charging cost and calculation time of SABDP, MC and DQN. Because the DQN is difficult to converge with the increase of state space and action space, we only take the scenario of five charging piles in 24 time periods and the wind power is $P_w = 8 kW$ as an example to analyse the results. The results show that the charging cost of SABDP is lower than that of MC and higher than that of DQN. MC has a large error compared with SABDP by sampling part of scenarios to estimate the charging cost. However, the optimal charging cost estimated by DQN is unreasonable, which is even better than that of SABDP because of the overestimation. Therefore, SABDP can achieve a more accurate charging cost than that of the other two methods in an acceptable time.

Figure 4 shows the charging process of three policies under the same initial state and wind power output curve. PCP gets the largest charging cost among the three policies because it charges EVs with the maximum charging demand all the time no matter how much the wind power and electricity price are. For example, from decision stage 1 to 6, PCP buys 92 kWh electricity to meet the maximum charging demand and abandons 36 kWh electricity generated from the wind power. However, the charging policy of SABDP only buys 60 kWh electricity.
and abandons 0 kWh wind power to meet the same charging demand. IRP gets near-optimal charging decisions that are closest to SABDP, but its charging cost is still higher than that of SABDP. The main reason is that IRP does not consider the future wind power and electricity price. For example, IRP and SABDP have the same charging demand from decision stage 1 to 3. At stage 1, the wind power is 36 kWh which is enough to charge three EVs but not enough to charge four EVs. IRP chooses to charge three EVs in order not to buy extra electricity and delays some of the charging demand to the next stage. But it has to buy 12 kWh electricity at stage 2 to satisfy both the original and delay charging demand due to the insufficient wind power. While SABDP does not delay the charging demand and it just buys 8 kWh to satisfy the same charging demand. SABDP makes the charging decision considering the wind power and electricity price comprehensively. If the electricity price is unchanged for a period of time, the charging decision is mainly influenced by the future wind power. The price influences the charging decision especially at the time of price change. If the electricity price is low, the policy would buy electricity to satisfy the charging demand as more as possible rather than delaying part of charging demand to the next stage, even though the wind power is insufficient. For example, compared with that of IRP, SABDP charges the EVs at the maximum charging demand while the IRP charges at the minimum charging demand at stage 21, when the wind power is not large enough to satisfy the maximum charging demand. The main reason is that the electricity price changes to a high price at stage 22, which would cause a high cost to buy extra electricity if delaying charging demand to stage 22.
In Figure 5, we compare the charging actions of different methods under the same scenario to show the optimality in action selection. The results show that the charging policy of MC is almost the same as that of SABDP under the same scenario, but the precision is related to the number of samples. The charging policy obtained from DQN is not sensitive to wind power. The charging demand of EVs it chooses may be far larger than the wind power such as stages 1 and 2. This charging policy would buy more electricity than SABDP. However, SABDP chooses the optimal charging decision considering both the wind power and the electricity price, which can get a lower charging cost and higher wind power utilization.

5 CONCLUSION

In this paper, we consider a joint-charging scheduling problem in the parking lot with wind power. The charging problem is modelled as MDPs to describe the uncertainties of wind power and the charging events. Then we propose a state-aggregation method to reduce the number of states and a novelty method based on aggregation to get the global optimal value more efficiently. In addition, several experiments are conducted to verify that our methods could get the optimal charging cost with less computational time compared with other methods. Moreover, it also shows that the SABDP method can be extended to different scales of charging problems. In future research, we will pay more attention on the charging scheduling problem with large scale and multi-scale.

ACKNOWLEDGMENT

This work was supported in part by National Key R&D Program of China (2016YFB0901900), and National Natural Science Foundation of China (61903293).

NOMENCLATURE

A. Indices

| t | Time step indices |
| i | The index of charging piles |
| m | The index of state aggregators |

B. Parameters

| St | State space at stage t |
| Ai | Action space at stage t |
| Pt | Transition probability matrix |
| St | The state space of a charging pile |
| P | Transition probability of a charging pile |
| T | One-step reward function |
| N | The number of charging piles |
| v | Wind speed at stage t |
| vF | Wind speed at stage t |
| vcut | Wind speed at stage t |
| αt+1 | Initial occupancy time of the i-th charging pile |
| βt+1 | Charging demand of the i-th charging pile |
| ζt+1 | Occupied probability |
| τt+1 | The probability distribution of αt+1 and βt+1 |
| 𝑔𝑡+1 | The price of electricity at t |
| N | The number of available state values for ζt+1 |

C. Variables

| sj | State at stage t |
| si | State of the i-th charging pile |
| st | State of the renewable energy |
| 𝑆CPC | Joint state of charging piles |
| 𝑆W | State of wind power |
| δt | Occupancy state of the i-th charging pile |
| ηt | Occupancy time of the i-th charging pile |
| a | Charging demand of the i-th charging pile |
| c | Charging decision of the i-th charging pile |
| A | Joint action of all charging piles |

REFERENCES

1. Zhang, X., Bai, X.: Incentive policies from 2006 to 2016 and new energy vehicle adoption in 2010–2020 in china. Renewable Sustainable Energy Rev. 70, 24–43 (2017)
2. Mukherjee, J.C., Gupta, A.: A review of charge scheduling of electric vehicles in smart grid. IEEE Syst. J. 9(4), 1541–1553 (2014)
3. International Energy Agency: Global ev outlook: Understanding the electric vehicle landscape to 2020. Paris, France (2013)
4. Clement-Nyns, K., Haesen, E., Driesen, J.: The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. IEEE Trans. Power Syst. 25(1), 371–380 (2010)
5. Deilami, S., et al.: Real-time coordination of plug-in electric vehicle charging in smart grids to minimize power losses and improve voltage profile. IEEE Trans. Smart Grid 2(3), 456–467 (2011)
6. Wang, J., et al.: Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power. Energy Policy 39(13), 2761–2769 (2011)
7. Zhang, H., et al.: Pev fast-charging station siting and sizing on coupled transportation and power networks. IEEE Trans. Smart Grid 9(4), 2595–2605 (2016)
8. Liu, H., et al.: Frequency droop control with scheduled charging of electric vehicles. IET Gener. Transm. Distrib. 11(3), 649–656 (2017)
9. Zhang, J., et al.: Optimal charging strategy for electric vehicles using symbolic–graphic combination principle. IET Gener. Transm. Distrib. 13(13), 2761–2769 (2019)
10. Shi, W., Wong, V.K.: Real-time vehicle-to-grid control algorithm under price uncertainty. In: 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm). Brussels, Belgium, pp. 261–266 (2011)
11. Zhang, T., et al.: Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price. IEEE Trans. Veh. Technol. 63(6), 2600–2612 (2013)
12. Zhang, L., Li, Y.: A game-theoretic approach to optimal scheduling of parking-lot electric vehicle charging. IEEE Trans. Veh. Technol. 65(6), 4068–4078 (2015)
13. You, P., et al.: Optimal cooperative charging strategy for a smart charging station of electric vehicles. IEEE Trans. Power Syst. 31(4), 2946–2956 (2016)
APPENDIX A: PROOF OF THEOREM 1

\[ \exists S_{CP,t}^{i} = \{s_{1}^{i}, s_{2}^{i}, ..., s_{N}^{i}\}, \quad S_{CP,t}^{j} = \{s_{1}^{j}, s_{2}^{j}, ..., s_{N}^{j}\}, \quad S_{CP,t}^{i} \subseteq S_{CP,t}, \quad t = 0, 1, ..., T, \text{ satisfying:} \]

\[ \text{Sort}(S_{CP,t}^{i}) = \text{Sort}(S_{CP,t}^{j}); \]

\[ \exists S_{CP,t}^{i} = \text{Sort}(S_{CP,t}^{j}); \]

\[ \forall s_{j}^{i} \in S_{CP,t}^{i}, \exists s_{j}^{m} \in S_{CP,t}^{j}, \text{ making } s_{j}^{m} = s_{j}^{i}; \]

\[ \text{Num}(S_{CP,t+1}^{i}) = \text{Num}(S_{CP,t+1}^{j}); \]

Where \( S_{CP,t}^{i} \) is the state space of \( S_{CP,t}^{j} \) and \( S_{CP,t+1}^{i} \) is the state set that \( S_{CP,t}^{j} \) can transfer to. \( \text{Num}(\cdot) \) is the function to calculate the number of states.

Because the charging piles share the same state space and action space, thus:

\[ \text{if } s_{j}^{i} = s_{j}^{m}, \text{ then } \]
\[ d_{j}^{i} = d_{j}^{m} \text{ and } s_{j}^{m} = s_{j}^{i}; \]
\[ \exists t \geq 1, \text{ making } s_{j}^{i} = s_{j}^{m}; \]

\[ \text{Num}(S_{CP,t+1}^{i}) = \text{Num}(S_{CP,t+1}^{j}); \]

\[ \exists \sum_{i=1}^{N} a_{j}^{m}; \]

\[ \text{V}(S_{t}^{i}) = R(S_{t}^{i}, A_{t}^{i}) + \sum_{S_{t+1}^{j}} P(S_{t+1}^{j} | S_{t}^{i}, A_{t}^{i}) \text{V}(S_{t+1}^{j}) \]
\[ = R(S_{t}^{i}, A_{t}^{i}) + \sum_{S_{t+1}^{j}} (p(S_{t+1}^{i} | S_{t}^{i}, A_{t}^{i}) \times \]
\[ \prod_{i=1}^{N} p(s_{j}^{m} | s_{j}^{i}, a_{j}^{m}) \times V(S_{t+1}^{j}) \]
\[ \text{V}(S_{t}^{i}) = R(S_{t}^{i}, A_{t}^{i}) + \sum_{S_{t+1}^{j}} (p(S_{t+1}^{i} | S_{t}^{i}, A_{t}^{i}) \times \]
\[ \prod_{i=1}^{N} p(s_{j}^{m} | s_{j}^{i}, a_{j}^{m}) \times V(S_{t+1}^{j}) \]
\[ R(S_{t}^{i}, A_{t}^{i}) = \max_{a_{j}^{m} \in A_{t}^{m}} (\sum_{i=1}^{N} (d_{j}^{i}) \times P_{t} - s_{j}^{i} \times 0) \times P(t), \text{ where } i \neq j \]

\[ L_{m} = \text{the number of states that } s_{j}^{i}, s_{j}^{m} \text{ transfer to.} \]

Calculate the optimal value at } t \text{ by reverse recursion.}
\[ V(S'_i) = R(S'_i, A'_i) + \sum_{i=1}^{\text{Num}(S'_i + 1)} P(S'_{i+1} | S'_i, A'_i) V(S'_{i+1}) \]

\[ = \Phi(S'_i, A'_i) + \sum_{i=1}^{\text{Num}(S'_i + 1)} P(S'_{i+1} | S'_i, A'_i) \times \ldots \times \sum_{i=1}^{\text{Num}(S'_i)} P(S'_{i+1} | S'_i, A'_i) V(S'_{i+1}) \]

\[ \Phi(S'_i, A'_i) = R(S'_i, A'_i) + \sum_{i=1}^{\text{Num}(S'_i + 1)} P(S'_{i+1} | S'_i, A'_i) \]

\[ R(S'_{i+1}, A'_{i+1}) + \ldots + \prod_{t=1}^{T} \sum_{i=1}^{\text{Num}(S'_i)} P(S'_{i+1} | S'_i, A'_i) \]

\[ P(S'_{ip} | S'_{ip-1}, A'_{jp-1}) \] is calculated similarly.

\[ V(S'_i) = \phi(S'_i, A'_i) + \prod_{t=1}^{T} \sum_{i=1}^{\text{Num}(S'_i)} P(S'_{ip} | S'_{ip-1}, A'_{jp-1}) V(S'_{ip}) \]

\[ \forall S'_i \in S'_i, 3 S'_i \in S'_i, \text{ satisfying:} \]

\[ \text{Sort}(S'_i) = \text{Sort}(S'_i), \]

\[ V(S'_i) = V(S'_i), \]

\[ : R(S'_i, A'_i) = R(S'_i, A'_i), \]

\[ P(S'_{i+1} | S'_i, A'_i) = P(S'_{i+1} | S'_i, A'_i), \]

\[ \Phi(S'_i, A'_i) = \phi(S'_i, A'_i), \]

\[ V(S'_i) = V(S'_i), t \in [t, T] \]

\section*{APPENDIX B: PROOF OF THEOREM 2}

Dividing \( S_i \) into \( M \) categories by the rules \( R \), then:

\[ S_i = \{ G'_1, \ldots, G'_M \}, G'_m = \{ S'_{i,m}, S'_{i,m}^2, \ldots, S'_{i,m}^{N'_{G,m}} \} \subseteq S_i, \]

\[ \sum_{m=1}^{M} N'_{G,m} = N_S; \]

\[ \forall S'_{i,m}, S'_{i,m} \in G'_m, \text{ Sort}(S'_{i,m}) = \text{Sort}(S'_{i,m}); \]

Thus choose \( S'_i = \{ S'_1, S'_2, \ldots, S'_M \}, S'_{i,m} \in G'_m, \)

From Theorem 1:

if \( \text{Sort}(S'_{i,m}) = \text{Sort}(S'_{i,m}), \) then:

\[ V(S'_{i,m}) = V(S'_{i,m}), \]

\[ P(S'_{i,m} | S'_{i-1,m}, A'_{i-1,m}) = P(S'_{i,m} | S'_{i-1,m}, A'_{i-1,m}) < 1, \]

\[ P(G'_m | S'_{i-1,m}, A'_{i-1,m}) = N'_{G,m} \cdot P(S'_{i,m} | S'_{i-1,m}, A'_{i-1,m}), \]

\[ E^*(V(S'_{i-1})) = \max(R(S'_{i-1}, A'_{i-1})) + \sum_{i=1}^{\text{Num}(S'_i)} P(S'_{ip} | S'_{ip-1}, A'_{jp-1}) V(S'_{ip}) \]

\[ = \max(R(S'_{i-1}, A'_{i-1})) + \sum_{m=1}^{M} N'_{G,m} P(S'_{i,m} | S'_{i-1,m}, A'_{i-1,m}) V(S'_{i,m}) \]

\[ \therefore E^*(V(S'_i)) = E^*(V(S'_i)), S'_i \in S'_i. \]