Effective Charging Planning Based on Deep Reinforcement Learning for Electric Vehicles

Cong Zhang, *Graduate Student Member, IEEE*, Yuanan Liu, *Member, IEEE*, Fan Wu, *Member, IEEE*, Bihua Tang, and Wenhao Fan, *Member, IEEE*

***Abstract*— Electric vehicles (EVs) are viewed as an attractive option to reduce carbon emission and fuel consumption, but the popularization of EVs has been hindered by the cruising range limitation and the inconvenient charging process. In public charging stations, EVs usually spend a lot of time on queuing especially during peak hours of charging. Therefore, building an effective charging planning system has become a crucial task to reduce the total charging time for EVs. In this paper, we first introduce EVs charging scheduling problem and prove the NP-hardness of the problem. Then, we formalize the scheduling problem of EV charging as a Markov Decision Process and propose deep reinforcement learning algorithms to address it. The objective of the proposed algorithms is to minimize the total charging time of EVs and maximal reduction in the origin-destination distance. Finally, we experiment on real-world data and compare with two baseline algorithms to demonstrate the effectiveness of our approach. It shows that the proposed algorithms can significantly reduce the charging time of EVs compared to EST and NNCR algorithms.**

***Index Terms*— Electric Vehicle, EVs charging scheduling system, deep reinforcement learning.**

I. Introduction

ELECTRIC vehicles as a significant part of the next-generation smart grid, have been drawing a lot of attention in recent years due to their energy-saving, carbon reduction, and environment protection. Compared with traditional combustion engine vehicles, EVs are more efficient and adopted by many customers [1]. Many countries around the world have set relevant policies to facilitate the development and popularization of EVs [2]. Between 2016 and 2017 alone, there are 3 million EVs sold worldwide which sales increased by 50%. It is expected that by 2030, more than 130 million EVs will enter the market [3]. By the end of September 2017, according to the official statistics, there are 141,094 EVs running in Beijing and 14,612 public charging piles have been deployed in the city [4]. However, since the cruising range is limited to the battery capacity, the EVs need to be charged frequently especially in the long-distance travel, making the charging scheduling system of EVs essential.

An increasing number of EVs driving on the road in the world, how to make the charging process of EVs more convenient has become a key point in researches. The literature [5] proposes that if public charging stations are not well deployed, the total time spent on charging would be dramatic. Afterwards, some methods to solve the charging pile planning problem are proposed. In [6], the authors utilize parking information of personal trips and use a regression model containing various variables to estimate parking demand. The authors in [7] use Monte Carlo simulations of EV behaviors to compare the service capacities and earnings of EV charging and battery swapping for both taxi and bus fleets. They discuss the possible reasons for today's less prevalence of battery swapping stations. The literature [8] assumes that the flow pattern of EVs follow the fossil-fuel vehicle traffic and uses them to estimate the charging demand. The authors in [9] propose an algorithm optimizing EV charging station to compute the optimal allocation of charging stations. The literature [10] uses a genetic-algorithm-based method to find sub-optimized locations to deploy charging stations.

However, these works merely focus on charging stations deployment while ignoring how to schedule EVs. In order to reduce the total amount of time spent on charging with a limited number of charging stations, it is more significant to design an optimal and efficient charging strategy for EVs.

In this paper, we propose a combination objective of optimal EVs charging scheduling scheme. Our works focus on reducing the total charging time of EVs and the origin-destination distance of EVs charging en-route. There are three challenges to conduct such an electric vehicle charging scheduling task:

- **Arbitrary Spatial Distribution.** Different from the EVs' charging at fixed times and locations, the location of each EV in our problem is totally arbitrary, which makes the charging scheduling routes vary.
- **Diverse Unbalanced Charging Rate.** In our problem EV charging modes can be classified into: slow charging, and fast charging [10]. In the slow charging mode, the charging power is low. It takes a long time for EVs to be fully charged. Since the charging process is slow, this charging mode requires a long parking time. If all EVs choose the fast charging pile, it would cause charging congestion when a large number of EVs arrive at one

Manuscript received December 17, 2019; revised May 21, 2020; accepted June 10, 2020. Date of publication June 24, 2020; date of current version December 24, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61821001 and Grant 61701041, in part by the YangFan Innovative and Entrepreneurial Research Team Project of Guangdong Province, and in part by the Beijing Key Laboratory of Work Safety Intelligent Monitoring (Beijing University of Posts and Telecommunications). The Associate Editor for this article was X. Chen. *(Corresponding authors: Yuanan Liu; Fan Wu.)*

The authors are with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100088, China (e-mail: cong1126@bupt.edu.cn; yuliu@bupt.edu.cn; wufanwww@bupt.edu.cn).

Digital Object Identifier 10.1109/TITS.2020.3002271

fast-charging pile at the same time. Such behaviors may lead to traffic congestion around fast charging stations as well as the power overload of these areas.
- **Limited Travel Distance.** The electricity left in the battery prevents an EV from arriving at some charging stations. Therefore, the travel distance of the scheduled EV is limited and the EVs cannot arrival at all the charging stations.

In this paper, we design an EVs charging route scheduling method for users. This system aims to schedule the EVs to appropriate charging piles for charging, so that the total time spent on charging could be minimized and the reduction in the distance of EVs' origin-destination(OD) would be maximized.

The contributions of the paper can be summarized as follows:
- We propose a novel EVs charging scheduling route problem, which takes the electricity consumption rate, charging rate constraint, and maximum travel distance into consideration.
- We propose a deep reinforcement learning algorithm to solve the EVs charging route scheduling problem.
- Experiments show the charging scheduling average finish charging time of the EVs recommended by DQL algorithm is 1.374-hour shorter than EST method and 2.824-hour shorter than NNCR method, reduced by 74.3% and 152.8% respectively.

The rest of the paper is organized as follows: Section II gives a brief overview of the related work. Section III describes the problem and the system overview. Electricty consumption model of EVs is discussed in Section IV. Section V gives the solution of EVs charging scheduling problem. Experiments are given in Section VI. Section VII concludes the paper.

II. Related Workes

As the widespread application of EVs in urban transportation, related researches have made corresponding progresses. They can be classified into two main fields: 1) price-based EVs charging schedule approach, and 2) charging time-based EVs charging schedule approach.

A. Price-Based EVs Charging Schedule Approach

In [11], the authors present a survey on recent EV charging control strategies, and propose the price-based EVs charging coordination approach that has been studied by many research works and implemented by many experiments [12], [13]. The literature [14] presents a Q-learning algorithm based dynamic charging scheduling scheme which intent to optimize the operation benefit for EVs. In [15], the authors have developed smart charging strategy according to time-of-use price from day-ahead predictions. Due to uncoordinated EV charging behaviors will impact the power grid and degrade power quality, the literatures [16]–[18] propose some methods to relieve the stress on the grid. Because the existence of randomness in traffic conditions, users' commuting behaviors, and the pricing process of the utility, the literatures [19] proposes a model-free approach to determine an optimal charging strategy for EVs charging.

B. Charging Time-Based EVs Charging Schedule Approach

The literature [20] uses optimal control to optimally allocate EV charging time and energy, but the algorithm requires users to provide charging schedule, which may degrade customers' satisfaction due to acquring inputs frequently. The authors in [21] formulate the charging of EVs in a single charging station as a mixed integer programming (MIP) problem and propose offline and online scheduling algorithms for this problem. The literature [22] proposes a distributed scheme that schedules charging activities temporally and spatially to minimize the waiting time of EVs. The authors in [20] develop a decentralized algorithm to schedule EV charging. In [23], a simplified charge-control (SCC) algorithm is presented. In the deterministic case, it can simplify the charging control decisions within an SCC set. In the stochastic case, an online state recursion algorithm is designed, which can provide a charging navigation utilizing online information. In [24], the authors propose a global optimal scheduling scheme and a local optimal scheduling scheme for EV charging and discharging. The literature [25] determines the charging order by a linear rank function, which is based on the estimated arrival time, the waiting time bound, and the amount of demanded electricity. For EVs charging scheduling, optimal charging scheduling problems have been considered, where optimal objective might be charging delays [26], energy-efficient route [27], [28], or a combination objective of travel time, charging time, and energy consumption [29]. However, these charging scheduling optimal objective of EVs ignore some realistic constraints and requirements faced (e.g. charging queuing time management which likes active queue management in the internet [30], [31] and demand of users' driving destination, etc). Different from those works, our work analyzes the EVs charging route scheduling problem under the real word conditions, and we focus on minimizing the charging time of EVs and maximizing the reduction of OD driving distances caused by charging.

III. Overview

In this section, we define the EVs charging-scheduling problem, and outline our solution framework.

A. Preliminaries

In this subsection, we will introduce some notations and the definition of EVs charging-scheduling to better describe the problem. A brief summarization of some notations used in our task definition is given in Table I.

In a city, its urban public EVs charging-scheduling problem includes public charging stations and EVs. We define the EVs charging-scheduling graph as follows.

Definition 1 (Electric Vehicles Charging-Scheduling Graph): In a city, there are $N = \{n_1, n_2, \ldots .n_n\}$ charging stations and y EVs $EV = \left\{ev_1, ev_2, \ldots .ev_y\right\}$. Charging station n_i has m_i charging piles. The number of fast charging piles in charging station n_i is m_i^f, and the number of slow charging is m_i^s. The EVs, charging stations, and the roads form a graph $G = < \upsilon, \xi >$. In graph G, the vertex set $\upsilon = N \cup EV$ stands for the set of charging stations and EVs,

TABLE I
NOTATIONS AND DESCRIPTION IN TASK-LEVEL

Notation	Description
N	Charging stations set
EV	Electric vehicles set
ev_i	The ith element in EVs set
n_j	The jth element charging stations set
$d_{i,j}$	Shortest distance between ev_i and n_j
m_i^f	The number of fast charging piles in n_i
m_i^s	The number of slow charging piles in n_i
$G =< v, \xi >$	Electric Vehicles & Charging Stations Graph
E_{higher}	Higher bound of EV's battery
E_{lower}	Lower bound of EV's battery
$\triangle e_{fast}$	Charged electricity in unit time at the fast pile
$\triangle e_{slow}$	Charged electricity in unit time at the slow pile
$E_{i,j}$	Left electricity of ev_i, when it starts charging in n_i
$T_{i,j}^C$	Charging time of ev_i, which charging in n_i
$\mathscr{F}$	Electricity consuming function of EV
E_i^{init}	Initially remained electricity of ev_i
ς_i	Maximum travel distance of ev_i
$T_{i,j}^T$	Travel time (Arrival time) of ev_i travel to n_j
v_i	Velocity of ev_i
Ω_j^f	EVs set which choose the fast charging pile at n_j
Ω_j^s	EVs set which choose the slow charging pile at n_j
u_j^f	Number of fast charging piles currently used at n_j
u_j^s	Number of slow charging piles currently used at n_j
$ev_{\hat{l},j}^f$	The $\hat{l}$th EV of Ω_j^f
$\beta_{\hat{l},j}^f$	Finish charging time of $ev_{\hat{l},j}^f$
$T_{\hat{l},j}^Q$	Queue time of $ev_{\hat{l},j}^f$ of Ω_j^f
$\delta_{i,j}$	Whether ev_i can reach n_j
$\wp_{i,j}^f$	Whether ev_i choose the fast charging pile at n_j
$\wp_{i,j}^s$	Whether ev_i choose the slow charging pile at n_j
$x_{i,j}^{f,\hat{l}}$	Whether ev_i is the $\hat{l}th$ in fast charging pile of n_j
$x_{i,j}^{s,\hat{l}}$	Whether ev_i is the $\hat{l}th$ in slow charging pile of n_j
OD_i	Distance from the current location of ev_i to its destination
$OD_{i,j}$	Distance from n_j to the destination of ev_i

and the edge set ξ stands for the set of roads that connect EVs and charging stations. We use the shortest distance $d_{i,j}$ between EV ev_i and charging stations n_j to represent the weight of each edge, which uses the shortest path algorithm to calculate from the graph G. Figure G is used to describe the connection relationship between the EV and the charging station. The number of edges of the charging stations connected to the EV indicates the candidate set that the current EV can choose to charge. The number of EVs connected to the charging station indicates the number of EVs that the charging station currently has potential for charging. The edge weights indicate the distance traveled by the EV during the charging process.

Suppose the fast charging piles and slow charging piles are identical in all charging stations, we define the charging time of EVs at different charging pile as follows.

Definition 2 (Electric Vehicles Charging Time): In order to avoid the destroying of the EV's battery, we denote EV's batteries have a upper bound E_{higher} and lower bound E_{lower}, where E_{lower} denotes the stop discharging energy to protect the battery cycle life. Before reaching E_{higher}, the rapid charger keeps the charging current constant [32]. In this charging process the charging curve is approximately linear and the charging process stops when the energy in battery rises to E_{higher}. We denote $\triangle e_{fast}$ as the charged energy in unit time at the fast pile and $\triangle e_{slow}$ as the charged energy in unit time at the slow pile. When an ev_i starting to charge at charging station n_j has the arrival energy $E_{i,j}$. If charging in a fast charging pile at charging station n_j, the charging time $T_{i,j}^C$ can be expressed as

$$T_{i,j}^C = \frac{E_{higher} - E_{i,j}}{\Delta e_{fast}}. \quad (1)$$

Otherwise in a slow charging pile at charging station j, the charging time $T_{i,j}^C$ is calculated by the following equation

$$T_{i,j}^C = \frac{E_{higher} - E_{i,j}}{\Delta e_{slow}}. \quad (2)$$

Definition 3 (Travel Electricity Consumption): When the EV drives to a charging station, it will consume the corresponding battery power. The initial left electricity that an EV starts to charge at the charging station varies with the arrived travel distance to a different charging station. As a result, the EVs' travel distance among different charging stations varies significantly, for which we train an EVs' travel electricity consumption model using real-world data to better characterize the EVs charging scheduling events. The electricity consuming function $\mathscr{F}$ is detailed in Section IV

Definition 4 (Maximum Travel Distance Limit): As EVs' residuel electricity is limited, they cannot arrive at some charging stations since the distance is too long. The maximum distance of EV ev_i is the travel distance before its electricity reaches the lower bound E_{lower}. The maximum travel distance is related to the initially remained electricity. If we use E_i^{init}, and ς_i to denote the initially remained electricity and the maximum travel distance of EV ev_i, respectively, ς_i can be calculated as the equation 3

$$\varsigma_i = \mathscr{F}^{-1}(E_i^{init} - E_{lower}). \quad (3)$$

where function $\mathscr{F}^{-1}$ is the inverse of $\mathscr{F}$.

Definition 5 (Electric Vehicles Travel Time): Since different EVs have various distances from the same charging station, the travel time of different EVs to the same charging station varies. This causes different queuing sequences for the charging station. If we schedule EV ev_i to charging station n_j, its travel time $T_{i,j}^T$ can be expressed the following equation:

$$T_{i,j}^T = \frac{d_{i,j}}{v_i}. \quad (4)$$

where v_i is the velocity of EV ev_i and $d_{i,j}$ indicates the travel distance between EV ev_i and charging station n_j.

Definition 6 (Electric Vehicles Queuing Time): The queuing time of the EVs is determined by the charging time of other EVs in front of these EVs at the charging station queue.

Here, we assume that the order of EVs arriving at the charging station is determined by their arrival time, the number of remaining charging piles, and the number of fast and slow charging in the current charging station. The arrival time of EVs refers to the travel time of the EVs from the current location to the selected charging station location. Let $\boldsymbol{\Omega}_j^f$ ($N_j^f = \left|\boldsymbol{\Omega}_j^f\right|$) and $\boldsymbol{\Omega}_j^s$ ($N_j^s = \left|\boldsymbol{\Omega}_j^s\right|$) denote the ordered queue of EVs, which choose the fast charging pile and the slow charging pile at charging station n_j for charging. The order of the queues $\boldsymbol{\Omega}_j^f$ and $\boldsymbol{\Omega}_j^s$ are determined by the arrival time of the EVs. Assume the number of fast charging piles and slow charging piles currently used at charging station n_j is u_j^f and u_j^s. The $\hat{l}$th EV of $\boldsymbol{\Omega}_j^f$ is denoted as $ev_{\hat{l},j}^f$ and $\hat{\beta}_{\hat{l},j}^f$ denote the finish charging time (total charging time) of the EV $ev_{\hat{l},j}^f$ at charging station n_j using fast charging pile, where $\hat{l} \in \left\{1, 2, \cdots, N_j^f\right\}$. The charging time of an EV includes the time when the EV arrives at the charging station (EVs Travel Time), the time that the EV queues at the charging station (EVs Queuing Time), and the time required for the battery of the EV to charge from the moment of charging (EVs Charging Time). Then the queue time of $ev_{\hat{l},j}^f$ of $\boldsymbol{\Omega}_j^f$ can be calculated as the following equations:

$$T_{\hat{l},j}^Q = \hat{\beta}_{\hat{l},j}^f - T_{\hat{l},j}^T - T_{\hat{l},j}^C. \tag{5}$$

If $u_j^f + \hat{l} \leq m_j^f$:

$$\hat{\beta}_{\hat{l},j}^f = T_{\hat{l},j}^T + T_{\hat{l},j}^C. \tag{6}$$

If $\hat{l} = 1 \quad and \quad u_j^f + \hat{l} > m_j^f$:

$$\hat{\beta}_{\hat{l},j}^f = max(\min_{\hat{l}+u_j^f-m_j^f}(\tau_j^f), T_{\hat{l},j}^T) + T_{\hat{l},j}^C. \tag{7}$$

If $\hat{l} > 1 \quad and \quad u_j^f + \hat{l} > m_j^f$:

$$\hat{\beta}_{\hat{l},j}^f = max(\min_{\hat{l}+u_j^f-m_j^f}(\tau_j^f \cup (\hat{\beta}_{1,j}^f, \cdots, \hat{\beta}_{\hat{l}-1,j}^f)), T_{\hat{l},j}^T) + T_{\hat{l},j}^C. \tag{8}$$

where τ_j^f represents the set of EVs' finish charging time which choose the fast pile charging at charging station n_j and function $\min\limits_{\hat{l}}(\tau_j^f)$ is to select the $\hat{l}th$ smallest element from set τ_j^f.

Problem Definition:

Setting EVs' driving speed υ, status of charging station N, EVs' initial electricity, EVs' electricity consumption model, EVs' charging rate, the road network RN and EVs charging-scheduling graph. *The objective of the EV charging scheduling problem* aims to schedule the EVs to appropriate charging piles for charging, so that the total time spent on charging is minimized and the reduction in the distance of EVs' origin-destination(OD) is maximized. The total time of EVs spent on charging includes the travel time to the charging station, the queuing time, and the actual charging time which can be obtained using equations (6) (7) and (8). Obviously, the solution to this problem can be reduced as finding the nearest fast charging pile, with ignoring the EVs queuing time. If we only consider the EVs queuing time, the strategy is to choose a charging station where few EVs are waiting for charging. However, in the real world, the EV charging scheduling problem has several constraints: (1) Each EV can only choose one charging pile for charging until its battery reaches higher bound E_{higher}; (2) The electricity left in the battery prevents an EV from arriving at some charging stations; and (3) The number and charging rate of fast charging pile m_i^f and slow charging m_i^s at a charging station is limited, which leads to a long time for actually recharging a battery. Given all the above constraints, we define the following 0-1 variables to represent these constraints:

$$\delta_{i,j} = \begin{cases} 1 & \text{if } d_{i,j} \leq \varsigma_i \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

$$\wp_{i,j}^f = \begin{cases} 1 & \text{if } ev_i \ choose\ the\ fast\ charging\ pile\ at\ n_j \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

$$\wp_{i,j}^s = \begin{cases} 1 & \text{if } ev_i \ choose\ the\ slow\ charging\ pile\ at\ n_j \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

$$x_{i,j}^{f,\hat{l}} = \begin{cases} 1 & \text{if } ev_i \ is\ the\ \hat{l}th\ in\ fast\ charging\ pile\ of\ n_j \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

$$x_{i,j}^{s,\hat{l}} = \begin{cases} 1 & \text{if } ev_i \ is\ the\ \hat{l}th\ in\ slow\ charging\ pile\ of\ n_j \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Our EVs charging scheduling problem can be formulated as follows:

$$min: \sum_{j\in N}(\sum_{\hat{l}\in\left\{1,\cdots,N_j^f\right\}} \hat{\beta}_{\hat{l},j}^f \cdot x_{i,j}^{f,\hat{l}} + \sum_{\hat{l}\in\left\{1,\cdots,N_j^s\right\}} \hat{\beta}_{\hat{l},j}^s \cdot x_{i,j}^{s,\hat{l}}). \tag{14}$$

$$max: \sum_{j\in N}\sum_{i\in EV}\delta_{i,j}(OD_i - OD_{i,j}). \tag{15}$$

$$str. \quad \sum_{j\in N}(\delta_{i,j}\cdot\wp_{i,j}^f + \delta_{i,j}\cdot\wp_{i,j}^s) = 1. \tag{16}$$

$$\sum_{\hat{l}\in\left\{1,\cdots,N_j^f\right\}}\sum_{j\in N}(x_{i,j}^{f,\hat{l}} + x_{i,j}^{s,\hat{l}}) = 1. \tag{17}$$

where OD_i represents the distance from the current location of EV ev_i to its destination and $OD_{i,j}$ denotes the distance from charging station n_j to the destination of EV ev_i. Equations (14) and (15) are the objective function. Equations (16) and (17) are subjection. Equation (16) ensures every EV can arrive at the chosen charging station and be scheduled just once. Equation (17) ensures that any EV is charged only at the Lth of the charging queue of one charging station.

Such a problem of scheduling resource-constrained EVs to appropriate charging piles for charging with a minimized total

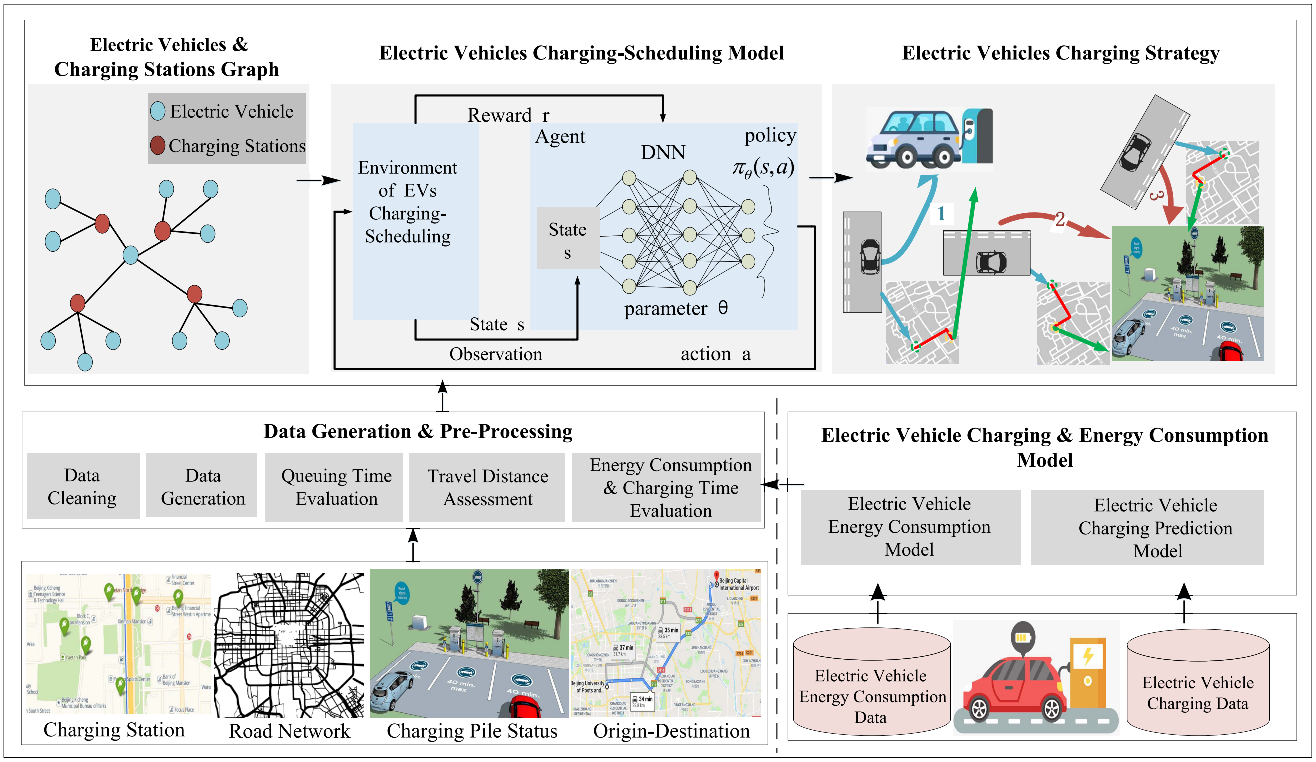


Fig. 1. System overview.

charging time and a maximal reduction in the OD distance is NP-hard as proven in Lemma 1 below.

Lemma 1 (NP-difficulty): When the number of charging piles in the charging station, the charging rate of the charging pile, and EVs' maximum travel distance is constrained, scheduling EVs to appropriate charging piles for charging with a minimized total charging time and a maximal reduction in the OD distance is NP-hard.

Proof 1: The EVs charging scheduling problem is scheduling the EVs to appropriate charging piles for charging. As a consequence, we can reduce the problem of scheduling EVs with a minimized total charging time from a kind of Parallel Machine Scheduling (PMS) problems, when EVs' maximum travel distance and capacity of charging station are constrained. Every single EVs' charging task could be regarded as a job while each charging stations could be considered as a machine. The constraint relationship is that each job has to be processed on just one of unrelated machines. Every piece of job has a particular processing time, while the processing speed of the machines are not exactly identical (i.e., the charging rates of fast piles and slow piles are totally different) and the size of the job could be various (i.e., every EV starting to charge with various arrival energy). The arrival time of every EV to charging stations is viewed as the release time of one job.

Furthermore, not all charging station $n_j \in N$ can be reached by every EV in the problem. The charging precedence of the EVs are various for different charging stations, which are determined by the arrival time, the number of remaining charging piles, and the number of fast and slow charging in the current charging station. For each machine in PMS, the release time of each job is deterministic and identical. But in our problem, the start charging time of any EV is determined by both its arrival time to charging station and the queuing time in charging station. These are determined by the scheduling decisions. Our goal is to minimize the total time spending on EVs' charging and maximize the reduction in the distance of EVs' origin-destination (OD) which means the selected charging station to the destination of the EVs is the closest. Thus, our problem is more complicated than PMS, and if we simplify some of the constraints of the problem, it can be boiled down to a PMS, which is known to be NP-complete [33], [34]. As a result, completing the proof of our problem is NP-difficulty.

We develop a deep reinforcement learning algorithm in Section V to tackle this NP-hard issue.

B. System Overview

Figure 1 presents an overview of our system, which consists of four main components: (1) *Electric Vehicle Charging & Energy Consumption Model*, which calculates travel electricity consumption for each EVs. In this mode, we use real-world data of EVs to simulate the consumption of battery electricity in an EV to a charging station. It takes the EVs' parameters, e.g., the EVs initial battery electricity, distance from EVs to charging station, and outputs the EVs' electricity consumption (detailed in Section IV); (2) *Data Generation & Pre-Processing and EVs & Charging Stations Graph* component takes the results of the energy consumption model, the road network data, and the charging station data as input. In this part, it cleans the raw data and uses the map-matching to map each generated EV's GPS point onto the corresponding

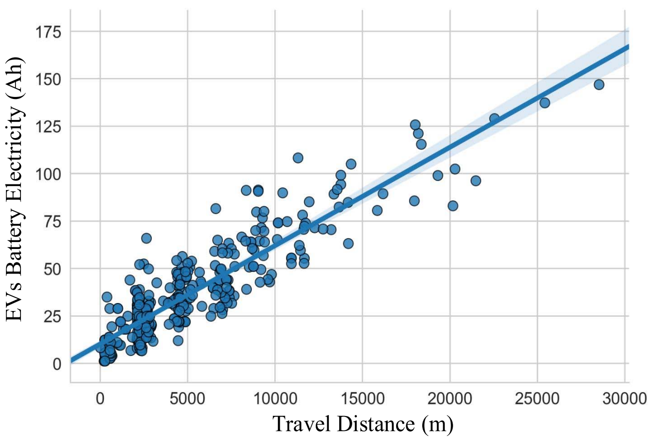


Fig. 2. Correlation between travel distance of EVs and electricity consumption of EVs batteries.

road segment (detailed in Section VI-B), and establishes the distribution graph of the EVs and charging station (detailed in definition 1). These will be used to build a simulation environment for an EVs charging-scheduling system. (3) *Electric Vehicles Charging-Scheduling Model*, which is a deep reinforcement learning model based on Deep Q network and recommend the optimal scheduling decisions for all EVs which need to be charged (detailed in Section V).

IV. Electricity Consumption Model of EVs

Since the initial left electricity of EVs' battery will affect its charging time and available distance to travel, we use the distance traveled by the EVs to predict the corresponding electricity consumption. As shown in the Figure 2, we analyzed the relationship between the travel distance of EVs and their battery electricity consumption on the real-word data, which shows they have an approximate linear relationship. In this paper, the electricity consumption model of EVs uses linear regression (LR) [35] model to predict EVs electricity consumption values.

$$\mathscr{F}(x) = wx + b. \tag{18}$$

where x is the travel distance of an EV, w is regression coefficient, and b is the constant.

Here, we get the optimal w and b by minimizing the predicted value of the EVs' battery electricity consumption $\mathscr{F}(x)$ and the mean square error of the true battery electricity consumption value y in the data. Then regression coefficient w and constant b can be calculated as the following equations:

$$\begin{aligned}(w^*, b^*) &= \underset{(w,b)}{arg\ min} \sum_{i=1}^{m} (\mathscr{F}(x_i) - y_i)^2 \\ &= \underset{(w,b)}{arg\ min} \sum_{i=1}^{m} (wx_i + b - y_i)^2. \end{aligned} \tag{19}$$

Finally, we use EV energy consumption data (detailed in Section VI-A) to train the model parameters and can get the closed-form of the optimal solution of w and b.

$$w = \frac{\sum_{i=1}^{m} y_i(x_i - \bar{x})}{\sum_{i=1}^{m} x_i^2 - \frac{1}{m}(\sum_{i=1}^{m} x_i)^2} \tag{20}$$

$$b = \frac{1}{m}\sum_{i=1}^{m}(y_i - wx_i). \tag{21}$$

where $\bar{x}$ is the mean of x. It can be denoted as $\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$.

V. Electric Vehicle Charging Route Planning

The EVs charging-scheduling problem can be seen as a sequential decision problem. The aim of this issue is to minimize the total time spending on charging of all EVs and reduce the OD distance of EVs charging en-route. Such a sequential decision making processes can be naturally modeled as Markov decision processes (MDPs). We will introduce the preliminaries of MDPs, and explain how to model the EVs charging route choice process as a MDP.

A. EVs Charging Scheduling Markov Decision Process (MDP)

Sequential decision problem evolves probabilistically based on a finite and discrete set of states. MDP [36] provides a mathematical method for modeling sequential decision processes. At each time steps of MDP, firstly the agent observes the state s_n of the process. Next it selects and executes an action a which is optional at state s_n. Then the agent receives a reward $R(s, a)$ according to the action a. At the next time, the process moves into a new state s_{n+1} and the probability of process from a state s_n to a new state s_{n+1} is influenced by the chosen action. We use a tuple $< S, A, p, R, \gamma >$ to represent the MDP, where S is a finite set of states and A is a finite set of possible actions. $p : S \times A \times S \rightarrow [0, 1]$ is the transition probability from state s_n to state s_{n+1} when an action at state s_n is taken. $R : S \times S \rightarrow R$ is the reward function, i.e. $R(s, a)$ is the reward received by the agent when taking action a in state s. $\gamma \in [0, 1)$ is the discount factor. The sequential decision problem in an MDP aims to find an optimal policy $\pi : S \rightarrow A$ maximizing the long-term cumulative reward.

B. Modeling EVs Charging Route Choices With MDP

The EVs charging schedule system can be defined as an "agent", which completes the charging of all EVs by making a sequence of decisions on the selection of charging stations and charging paths for all the EVs that need to be charged. We model the EV charging system scheduling processes using MDP model. Next, we introduce each component in an MDP.

1) Enviroment of EVs Charging Scheduling: In EVs charging scheduling tasks the agent requires interacts with an environment in which an agent can take actions and observe the results. Therefore, we use the OSM road network data, charging pile data, EV status data, EV electricity consumption model and EV charging model to build an emulator for EV charging scheduling environment. In the emulator, we can get the travel time to the charging stations, the queuing time, and

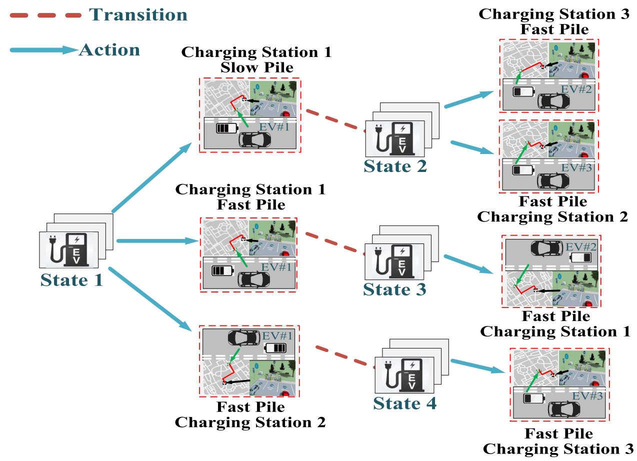


Fig. 3. Illustration: transit choices as an MDP.

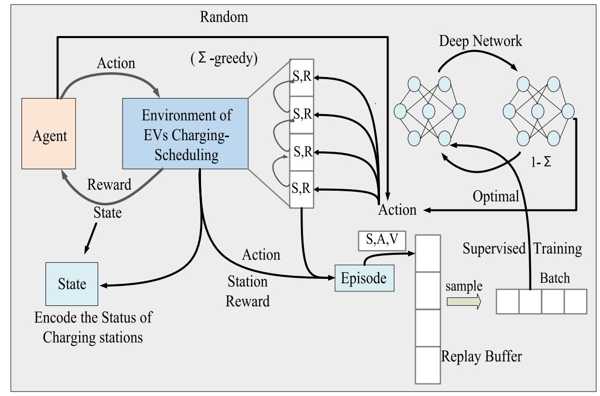


Fig. 4. The structure diagram of DQL-based EVs' charging scheduling algorithm.

the actual charging time of the EV in the charging scheduling process (detailed in Section VI-B). The agent selects an action from the set of EVs valid actions, at each time-step. The action is passed to the emulator, then emulator gives the next state and feedback in the form of a reward to agent.

2) EVs State Space S: The agent fully observes the current charging situation which can realize the system state s_n representation at step t_n. State $s_n \in S$ is represented by a vector of $2n$ variables: $s_n = L_n$. The L_n indicates the number of charging piles that are not currently used in all charging stations at step t_n (the time of scheduled EVs arriving at the target charging station), denoted as n tuple $L_n = \left\{(l_n^{1,f}, l_n^{1,s}), (l_n^{2,f}, l_n^{2,s}), \cdots, (l_n^{n,f}, l_n^{n,s})\right\}$, where $l_n^{1,f}$ and $l_n^{1,s}$ show the number of fast charging piles and slow charging piles that are not used in charging stations. For example, in Fig 3, the charging scheduling agent takes different actions at station s_n, and the charging situation may be in different state. The arriving time of scheduling EVs to different charging stations is various, and state transition might take place when a new EVs arrives or an EVs has been fully charged. Thus, it causes the charging state transition of EVs under different scheduling actions.

3) EVs Action Space A: An action $a_n \in A$ is a scheduling choice decision. Action a_n means that the EV at state s_n selects a charge station which the EV can reach, e.g., the electricity left in the battery prevents an EV from reaching some charging piles (fast charging pile or slow charging pile). Action a_n is represented by a vector of three variables: $a_n = [\delta_n^{ev}, \delta_n^{CS}, \delta_n^{CM}]$. We use $\delta_n^{ev} \in \left\{ev_1, ev_2, \ldots .ev_y\right\}$ to denote the current scheduled EVs. The δ_n^{CS} indicates the chosen charging station and δ_n^{CM} is a $0-1$ valued binary variable which we use to denote the outcome of EVs' charging model selection, where 1 stands for choosing fast pile, and 0 stands for choosing slow pile. A_n is the set of valid actions at t_n. If the state is s_n, the valid action is $a_n \in An$. The valid action in A_n represents that the EVs not do repeatedly schedule, and do not select the charging station which the EVs cannot reach and the charging mode which is not supported by the current charging station. For example, in Fig 3, the actions of EVs#1 can reach two different charging stations. We denote a_n as a scheduling choice decision at step t_n.

4) EVs Scheduling Reward R: Reward R reflects the desirablity of a particular state transition that is observed by performing action a starting in the state s_n and resulting in next state s_{n+1}. Fig 4 shows state-action transition in the EVs charging scheduling task. Its goal is to minimize the total time spent on charging of all EVs and maximize reduction in the distance between the EV's origin and destination, while meeting certain constraints. This specific objective can be fed to RL agent by means of a reward function. Thus, the reward function serves as a function similar to an objective function formulation in EVs charging scheduling task. The reward function is calculated by the following equations:

$$R(s_n, a_n) = \begin{cases} \dfrac{\alpha}{\hat{\beta}_{(s_n,a_n)}} + \dfrac{\beta \times \mathbb{RED}(OD_{(s_n,a_n)})}{OD_{(s_n,a_n)}} & \text{if } a_n \in = A_n \\ -\infty & \text{if } a_n \notin = A_n. \end{cases} \tag{22}$$

where $\hat{\beta}_{(s_n,a_n)}$ represents the finish charging time of the current EV when the a_n action is performed in the state s_n. $\mathbb{RED}(OD_{(s_n,a_n)})$ and $OD_{(s_n,a_n)}$ denote the value of the EVs' OD distance decrease and the initial EVs' OD distance when selected action is a_n in the state s_n, respectively. α and β are tuning parameters to set the preference on reducing the total charging time and OD distance of EVs. A larger α and smaller β mean that the shorter total charging time of the EVs' scheduling option is preferred to select and will get a greater reward, while a smaller α and larger β indicate that more attention is paid to the reduction of EVs' OD distance.

The optimal scheduling of EVs charging task requires the determination of a stationary policy π, defining which scheduling action a_n should be applied at step t_n. As such, a EVs charging scheduling policy π induces a stationary mass distribution over the realizations of the stochastic process s_n, a_n. A sequence of functions is $\pi = (a_1, a_2, a_3, \cdots, a_n)$, in which a_n is an admissible policy if $a_n \in An$. The goal of EV scheduling task is learned to find an optimal policy π^* which could maximize the total EVs scheduling rewards. When the EVs scheduling system is following the scheduling policy π, the action at t_n is $a_n = \pi(s_n)$. Therefore, the optimal policy π^* is calculated by the following equations:

$$\pi^* = \arg\max_{\pi} \sum_{s_n \in S} Q^{\pi}(s_n, a_n). \tag{23}$$

and

$$\begin{aligned} Q^{\pi}(s_n, a_n) &= E_{\tau}[\sum_{k=0} \gamma^k R(s_{n+k}, a_{n+k})] \\ &= \sum_{(s_n, a_n, \cdots) \sim \tau} \pi(a_n|s_n) \\ &\quad \times \rho(s_{n+1}|s_n, an) \cdots \sum_{k=0} \gamma^k R(s_{n+k}, a_{n+k}) \\ &= \sum_{a_n} \pi(a_n|s_n)\rho(s_{n+1}|s_n, an) \\ &\quad \times [R(s_n, a_n) + Q^{\pi}(s_{n+1}, a_{n+1})]. \end{aligned} \tag{24}$$

where τ is a sequence obtained by sampling according to policy and state transitions. $Q^{\pi}(s_n, a_n)$ is the long-term return expectation generated by the action of a certain scheduling policy in the current state s_n. And the $\gamma \in (0, 1]$ is a discount factor.

We define the optimal action-value function $Q^*(s_n, a_n)$ as the maximum expected return when the state of the agent is s_n and take optimal action a_n. $Q^*(s_n, a_n)$ is defined as equation 25:

$$Q^*(s_n, a_n) = E[R(s_n, a_n) + \gamma \max_{a_{n+1}}[Q^*(s_{n+1}, a_{n+1})]]. \tag{25}$$

C. Deep Q Networks Based EVs Scheduling

1) Main Idea: How to schedule the EVs to appropriate charging piles for charging to minimize the total time spent on charging of all EVs and maximize reduction in the distance between the EV's origin and destination. The main idea is that we consider task in which the agent interacts with the environment of EVs charging scheduling through a sequence of observations, actions and rewards, where it develops new strategies. The agent receives a reward and correlates actions a_n with rewards. The agent determines its actions by using a deep neural network and mixing the output of the deep neural network with random actions to sample its training set. In essence, the agent trains the deep neural network in such a way that it predicts the cumulative, weighted rewards for all actions. The relationship between EVs charging scheduling environment, agent and deep neural network is shown in Fig 4.

We use a Deep Q Network (DQN) [37] to model the EVs' charging scheduling. In DQN model, the approximate action-value function is represented as $Q(s_n, a_n; \theta)$, where $Q(s_n, a_n; \theta) \approx Q^*(s_n, a_n)$. In the learning process, the Q-network updates are based on a mini-batch of experiences $s_n, a_n, R(s_n, a_n, s_{n+1})$ and can be trained to learn the parameters θ of the action value function $Q(s_n, a_n; \theta)$ by minimizing a sequence of loss functions $L_i(\theta_i)$ that changes at each iteration i. The i_{th} loss function $L_i(\theta_i)$ is calculated by the equations 26:

$$\begin{aligned} L_i(\theta_i) = E[(&R(s_n, a_n) + \gamma \max_{a_{n+1}} Q(s_{n+1}, a_{n+1}; \theta_{i-1}) \\ &- Q(s_n, a_n; \theta_i))^2]. \end{aligned} \tag{26}$$

where θ_i is the neural network's parameters at the i_{th} update, and $R(s_n, a_n) + \gamma \max_{a_{n+1}} Q(s_{n+1}, a_{n+1}; \theta_{i-1})$ is the target for iteration i. The parameters from the previous iteration θ_{i-1} are held fixed when optimizing the loss function $L_i(\theta_i)$. Differentiating the loss function with respect to the neural network's parameters at iteration i, θ_i gives the following gradient:

$$\begin{aligned} \frac{\partial L_i(\theta_i)}{\partial \theta_i} = E[(&R(s_n, a_n) + \gamma \max_{a_{n+1}} Q(s_{n+1}, a_{n+1}; \theta_{i-1}) \\ &- Q(s_n, a_n; \theta_i))\frac{\partial Q(s_n, a_n; \theta_i)}{\partial x \theta_i}]. \end{aligned} \tag{27}$$

Our EVs scheduling model uses stochastic gradient descent [38] method to optimize the loss function. In order to balance the Exploration-Exploitation dilemma, our model learns the optimal action $a_n = \max_a Q(s_n, a_n; \theta)$. At the same time it still chooses random actions to ensure adequate exploration of the state space. Here, we use the $\epsilon - greedy$ strategy. In this strategy, the agent chooses the current optimal action with a probability $1 - \epsilon$, and chooses a random action with probability ϵ. The deep Q-learning algorithm of EVs charging scheduling is presented in Algorithm 1.

2) Algorithm Design: Algorithm 1 gives the pseudo-code of our deep reinforcement learning algorithm. In the *initialization stage*, firstly, the algorithm makes the replay buffer with initial capacity of N. the target network and the behavior network have exactly the same initialization parameters (Line 1- 3). In each iteration of the model training, as shown in Fig 4 the behavior $Q - network$ is responsible for interacting with the environment of EVs charging-scheduling to obtain interactive samples (Line 3- 9). The target value is calculated by target $Q - network$, then it is compared with the estimated value of behavior $Q - network$ to obtain the target value, and behavior $Q - network$ paramers θ is updated (Line 10- 15). Every time iteration C is completed, the model parameters of behavior $Q-network$ are synchronized to target $Q-network$ (Line 16). Finally, when the episode is equal to k, the model training terminates, and EVs charging scheduling route policy π^* are returned as the recommended EVs charging scheduling plan.

VI. Experiments

In this section, we conduct effective experiments to evaluate our system. Firstly we describe the dataset used in the paper. Then, we present the experiment results.

A. Datasets

1) Road Networks: We collect the road network data in Beijing, China from Open Street Map.[1]

2) EV Charging Piles Data: The EV charging stations data in Beijing, China are collected from ChargeBar.[2] Each EV charging station data contain charging station ID, charging station locations, longitude and latitude coordinates of charging station (GPS point), the number of fast charging piles and slow charging piles. As shown in Figure 5, there are 1525 charging stations in Beijing.

[1] https://www.openstreetmap.org/

[2] http://admin.bjev520.com/jsp/beiqi/pcmap/do/

Algorithm 1 DQL-Based EVs' Charging Scheduling Algorithm

Input: EVs charging-scheduling graph $G = < \upsilon, \xi >$, electricity consuming model $\mathscr{F}$, charging stations status, charging stations capacity $(m_1^f, m_1^s), \cdots, (m_n^f, m_n^s)$, charged electricity in unit time at the fast pile $\triangle e_{fast}$, charged electricity in unit time at the slow pile $\triangle e_{slow}$, discount factor γ, tuning parameter α, β and maximum number of iterations k.

Output: Optimal charging scheduling policy π^*.

//Stage 1: Initialization

Initialize replay memory $\mathbb{N}$ to capacity C
Initialize behavior $Q-network$ with random weights θ
Initialize target $Q-network$ with random weights $\theta^* = \theta$

//Stage 2: Deep Network Tranining

for $episode \leftarrow 1$ to k **do**
  Collect charging stations characteristic to realize state s_0
  for $i \leftarrow 0$ to T **do**
    Select $a_n \leftarrow \arg\max_a Q(s_n, a_n; \theta)$ with probability $1-\epsilon$ otherwise with probability ϵ select a random action a_n
    Execute a_n and observe $R(s_n, a_n)$ and s_{n+1}
    Store transition $(s_n, a_n, R(s_n, a_n), s_{n+1})$ in $\mathbb{N}$
    Sample random minibatch of transitions $(s_t, a_t, R(s_t, a_t), s_{t+1})$ from $\mathbb{N}$
    if episode terminates at s_{t+1} **then**
      Set the target to $R(s_t, a_t)$
    else
      Set the target to $R(s_t, a_t) + \gamma \max_{a_{n+1}} [Q^*(s_{n+1}, a_{n+1})$
    Perform a gradient descent and update θ by SGD
    Every C steps, update: $\theta^* \leftarrow \theta$

//Stage 3: Get Optimal Charging Scheduling Policy

$\pi^* \leftarrow \underset{i \in \{1,2,....y\}}{\arg\max_a} \; Q^*(s_i, a_i)$
return π^*

3) EVs Energy Consumption Data: EV energy consumption data collected from the connected electric autonomous vehicles operating in the Columbus area.[3] The data includes vehicle locations, routes traveled, miles traveled, battery performance, etc.

B. Data Generation and Pre-Processing

Data Generation and Pre-processing takes the road network, the EVs charging piles data, and EVs energy consumption data as input, and performs the following three tasks to prepare the data for further processing:

1) Data Cleaning: Data Cleaning cleans the EVs charging piles data, and EVs energy consumption data. In this step, we clean the raw charging piles location data by filtering the noisy GPS points with a heuristic-based outlier detection method [39], and makes the coordinate conversion of the charging station GPS points matching the coordinate system of the OSM road network.

[3]https://ckan.smartcolumbusos.com/dataset/

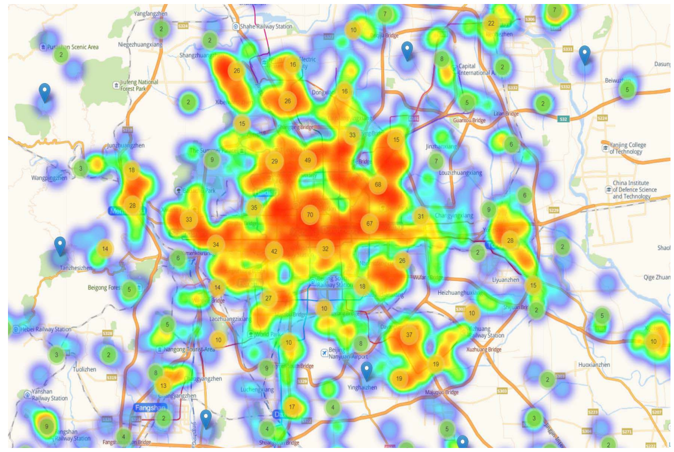


Fig. 5. Distribution of charging stations in Beijing.

2) Data Generation: In this module, we randomly generate GPS coordinate data of EVs in the area of Beijing. Then we map the GPS points onto the corresponding segments in road networks, which is crucial for simulating the position of driving EVs. The randomly generated GPS data can be at any location in the city, where the location of EVs are usually along the road network. This step maps the GPS points of each EV to the nearest corresponding segments in road networks with a global map matching method [40]. In order to avoid the generation of inappropriate GPS points in the process of generating EVs' driving destination, we select the POI (Point of Interest) point in the OSM map within a certain range of the EVs as the destination. The reason we choose POI as the destination is that the POI entities represent the place that is closely related to people's lives, such as schools, banks, restaurants, hospitals, supermarkets, etc. At the same time in generating the initial electricity of EVs, we ensure that each EV is able to reach at least one charging station.

3) Travel Distance Assessment: EVs travel distance is calculated based on the shortest navigation distance between EVs and charging stations in the OSM road network. We calculate the shortest navigation distance from the OSM road network using the Dijkstra [41] algorithm. Here, we suppose that for any charging piles in the same charging station, the distance between EVs to the charging piles is identical. Therefore, we only calculate the distance between EVs to the charging station as the travel distance.

4) Energy Consumption & Charging Time Evaluation: We extract features from EVs energy consumption data and train the EVs' electricity consuming model $\mathscr{F}$. EVs travel distance are used as input, and we use $\mathscr{F}$ to calculate the amount of electricity consumption, when they travel to their scheduling charging station. Here we use the $R2$ coefficient of determination to calculate the accuracy of model $\mathscr{F}$. And the accuracy of the model inferring the energy consumption based on the distance traveled by the EVs is 80.38%. According to the literature [32], the rapid charger keeps the charging current constant. So here we calculate the charging time of each EV by Equations 1 and 2.

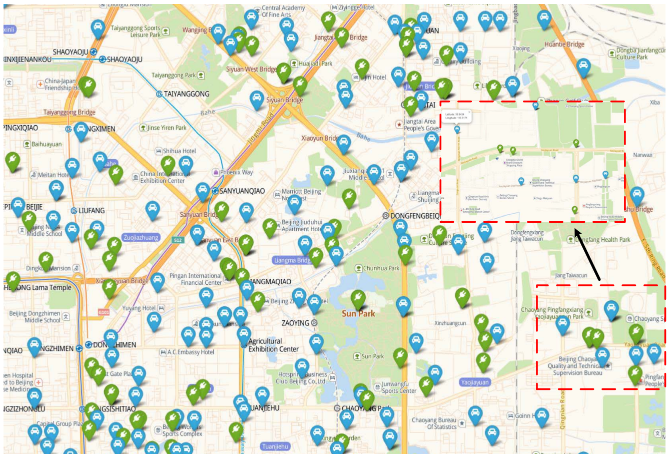


Fig. 6. Illustration: Distribution of EVs requesting charging.

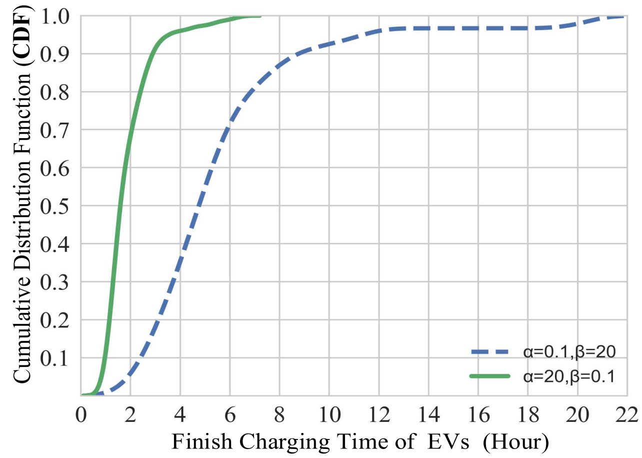


Fig. 7. The cumulative distribution function of EVs' charging time with different tuning parameter α and β.

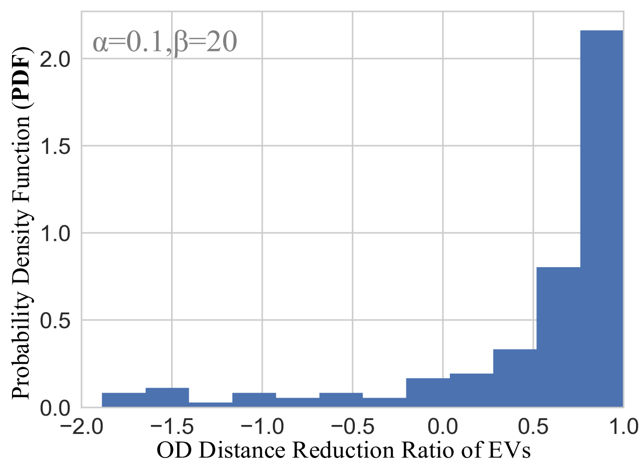


Fig. 8. The probability density function of EVs' OD distance reduction with tuning parameter $\alpha = 0.1$ and $\beta = 20$.

5) Queuing Time Evaluation: For assessing the EVs queuing time, we adopt the number of charging piles arriving at the charging station and the charging time of the preceding arrival EVs to estimate. The arrival time between different EVs determines their position in the queue, which could be obtained in the simulation system.

C. Effectiveness Evaluation

In this section, we study the effectiveness of EVs charging scheduling model. The test data are collected from the charging station data and EVs data in Beijing, which have 50 charging stations and 200 EVs in these area. Some of them is shown in Figure 6. Unless mentioned otherwise, the default parameters used in the experiments are: 1) the average speed of the EVs is $v = 60km/h$; 2) the battery is able to carry totally 1000Ah of electricity. In electrochemistry, it is recommended to use the the state of charge (SOC) from 90% - 20% of a battery to prolong its service life which we also implement in our case with the higher bound $E_{higher} = 900Ah$ and lower bound $E_{lower} = 200Ah$. The SOC is calculated by the ratio between the current electricity and the total electricity. So the charging ratio at the fast pile is $\triangle e_{fast} = 700Ah/h$ or slow pile is $\triangle e_{slow} = 70Ah/h$, and $\gamma = 1$. The neural networks of DQL consist of four densely connected layers with 100, 1024, 512 and action dimension neurons, where the last layer corresponds to the actions. The activation functions are Rectified Linear Units (ReLU). The optimizer used for training is Adam with a learning rate 10^{-4}.

1) EVs Charging Scheduling Route Planning: In this EVs charging scheduling problem, we study the effect of different tuning parameter α. Figure 7 provides the cumulative distribution function (CDF) results with different α and β settings. It shows that, when α is relative large($\alpha = 20$) and $\beta = 0.1$, more than 90% of the EVs finish charging in four hours. While, α is small ($\alpha = 0.1$) and $\beta = 20$, about 70% EVs have a total charging time more than four hours. The reason behind these phenomena is that with the large α and small β the total charging time of EVs will be more attention and the charging scheduling action with a short total charging time will get more rewards. However, the parameters α and β also affect the OD distance reduction value of EVs. Figure 8 and figure 9 give the probability density function of EVs' OD distance reduction with different tuning parameter. We observe that the reduction ratio probability of EVs' OD distance less than 0 for the small α (0.1) and relative large β (20) is much lower than the relative large α (20) and small β (0.1). when α is small ($\alpha = 0.1$) and $\beta = 20$ is relative large, the reduction ratio of EVs' OD distance is mainly concentrated between $0 - 1$. As shown in figure 10, When the value of α is equal to 20 and the value of β is equal to 0.1, the average ratio of EVs' OD distance reduction is -0.575. While the average ratio of EVs' OD distance reduction is 0.492 when α is equal to 0.1 and β is equal to 20. Where the negative reduction ratio of EVs' OD distance indicates that the selected charging station is further from the initial destination of the EV. While the ratio is closer to 1, the selected charging station is closer to the initial destination of the EV. So when the value of α is small and β is large, it is more preferable to select the charging scheduling action which can cause large OD distance reduction of EVs. Due to the influence of the initial remaining electricity of each EV, the choice of charging

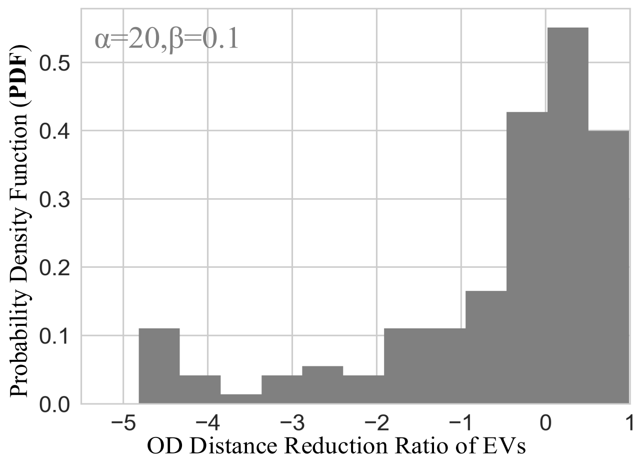


Fig. 9. The probability density function of EVs' OD distance reduction with tuning parameter $\alpha = 20$ and $\beta = 0.1$.

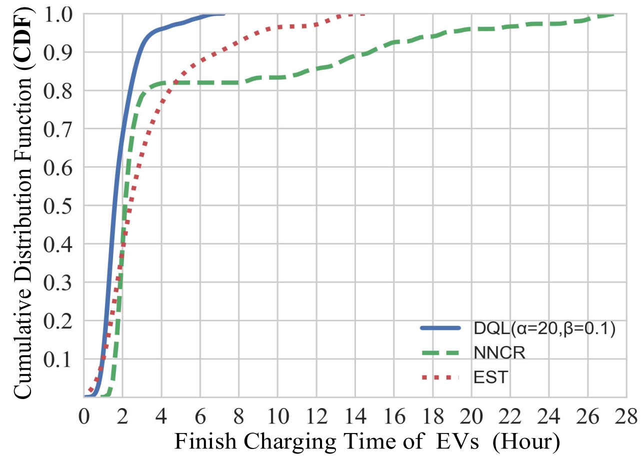


Fig. 11. The cumulative distribution function of EVs' charging time with different charging scheduling baseline methods.

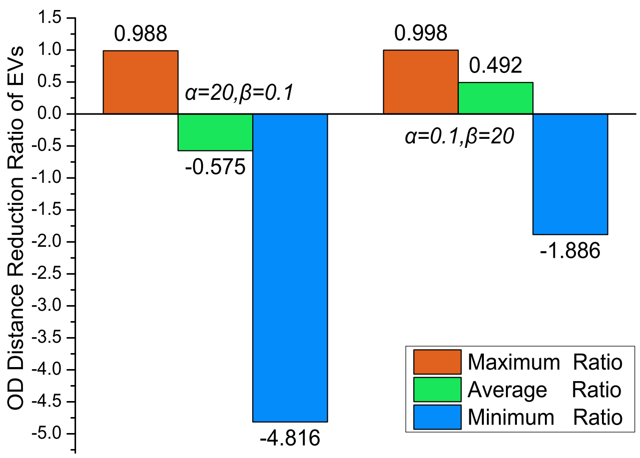


Fig. 10. The OD distance reduction of EVs with different tuning parameter α and β.

TABLE II
CHARGING TIME RESULTS OF EVS WITH DIFFERENT CHARGING SCHEDULING BASELINE METHODS

Algorithm	$NNCR$	EST	DQL ($\alpha = 20\&\beta = 0.1$)
Average Charging Time(h)	4.672	3.222	1.848
Maximun Charging Time(h)	27.0243	13.016	6.363
Minimum Charging Time(h)	1.553	0.941	0.941
Total Charging Time(h)	934.483	644.361	369.683

stations that can be arrived by EVs is limited. Therefore, the algorithm's optimization of the travel distance is limited, but adjusting the parameter α and β can make the charging strategy more preferred to choose the charging stations, which are closer to the destination of EVs in the limited selection.

2) Finish Charging Time Performance of Different Methods in EVs Charging Scheduling Route Planning: We compare our method, DQL-based Charging Scheduling, with two other baselines.

- **Baseline 1: Nearest Neighbor Charging Routing (NNCR).** Schedule each EV to its nearest charging station, then according to their arrival time choose the charging pile at this station which can let it earliest start charging.
- **Baseline 2: Earliest Start Time (EST).** EST algorithm aligns the currently earliest start charging time in a non-decreasing order each EV to each charging station, among which EST algorithm chooses the earliest one. Such that one of the EVs is scheduled to one charging station which provides the earliest start charging time [42].

Figure 11 illustrates the CDF of the EVs' finish charging time using different algorithm. From this figure, we observe that the finish charging time using NNCR, and EST are longer than using our algorithm DQL when the value of α is equal to 20 and β is equal to 0.1. More than 90% of the EVs using DQL are capable to finish charging in four hours, while it is about 20% of the EVs using EST have to spend more than four hours to finish charging. When we use NNCR algorithm, it will take more than 27 hours for all EVs to finish their charging and about 20% of the EVs have to spend more than ten hours. Table II shows the EVs charging time results that inclues average charging time, maximum charging time, minimum charging time and total charging time of three algorithms. The average finish charging time of DQL ($\alpha = 20, \beta = 0.1$) is 1.848 hour, which is 1.374-hour shorter than EST and 2.824-hour shorter than NNCR. The maximum finish time of DQL($\alpha = 20, \beta = 0.1$) is 6.653-hour shorter than EST and 20.6613-hour shorter than NNCR. Due to the NNCR algorithm only considers the closest charging station to the current charging EV (that is, the fastest arriving charging station), while does not consider the number of charging piles at the charging station and the charging rate of different charging piles. So the finish charging time of EVs is longer than our algorithm. The EST algorithm only focuses on the earliest start charging time of EVs, while it ignores the actual charging rate of different charging pile and EVs finish charging time. For example, when the EV arrives at a charging station, the fast charging piles have been occupied so that the slow charging piles are able to use, but the fast charging pile will be released in 1 minutes. In order to meet the shortest waiting time, the EST algorithm will choose slow charging, while our method will choose the charging schedule policy (fast charging pile) according to the final charging time. Therefore, the charging time of EVs will be further shortened.

VII. Conclusion

In this paper, we present a novel approach based on DQL to schedule the EVs charging and recommend the appropriate charging road path to the EVs based on the real charging stations data (Beijing, China) collected from ChargeBar. Our EVs charging scheduling model can address the problem of EVs' efficiency charging scheduling policy in a more realistic fashion, considering the constraints and requirements from EVs users' perspective: 1) travel distance limitations of EVs, 2) diverse unbalanced charging rate constraints of different kind of charging model, and 3) the EVs charging request at an arbitrary spatial location. In order to adjust preferences between the finish charging time and the OD distance reduction of EVs, we propose a flexible reward function. Then, we perform extensive experiments on a real test data and demonstrate the effectiveness of our proposed EVs charging scheduling model, where our model recommends the charging scheduling policy of EVs that the average finish charging time is 1.374-hour shorter than EST method and 2.824-hour shorter than NNCR method, reduced by 74.3% and 152.8% respectively. If we adjust the parameters α and β, they affect users the preference for reducing the total charging time or the OD distance of EVs. As future work, we plan to utilize the real-time traffic and charging station status data to better improve the charging scheduling policy of our EVs charging scheduling model.

References

[1] R. C. Green, L. Wang, and M. Alam, "The impact of plug-in hybrid electric vehicles on distribution networks: A review and outlook," *Renew. Sustain. Energy Rev.*, vol. 15, no. 1, pp. 544–553, Jan. 2011.

[2] W. Su, H. Eichi, W. Zeng, and M.-Y. Chow, "A survey on the electrification of transportation in a smart grid environment," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 1–10, Feb. 2012.

[3] T. Bunsen *et al.*, *Global EV Outlook 2018: Towards Cross-Modal Electrification*. Paris, France: International Energy Agency, 2018.

[4] J. Li, X. Sun, Q. Liu, W. Zheng, H. Liu, and J. A. Stankovic, "Planning electric vehicle charging stations based on user charging behavior," in *Proc. IEEE/ACM 3rd Int. Conf. Internet-Things Design Implement. (IoTDI)*, Apr. 2018, pp. 225–236.

[5] D. Steen, L. A. Tuan, O. Carlson, and L. Bertling, "Assessment of electric vehicle charging scenarios based on demographical data," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1457–1468, Sep. 2012.

[6] T. D. Chen, K. M. Kockelman, and M. Khan, "Locating electric vehicle charging stations: Parking-based assignment method for Seattle, Washington," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2385, no. 1, pp. 28–36, Jan. 2013.

[7] T. Zhang, X. Chen, Z. Yu, X. Zhu, and D. Shi, "A Monte Carlo simulation approach to evaluate service capacities of EV charging and battery swapping stations," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 3914–3923, Sep. 2018.

[8] S. H. Chung and C. Kwon, "Multi-period planning for electric car charging station locations: A case of Korean expressways," *Eur. J. Oper. Res.*, vol. 242, no. 2, pp. 677–687, Apr. 2015.

[9] Y. Xiong, J. Gan, B. An, C. Miao, and A. L. C. Bazzan, "Optimal electric vehicle fast charging station placement based on game theoretical framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2493–2504, Aug. 2018.

[10] J. Dong, C. Liu, and Z. Lin, "Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data," *Transp. Res. C, Emerg. Technol.*, vol. 38, pp. 44–55, Jan. 2014.

[11] Q. Wang, X. Liu, J. Du, and F. Kong, "Smart charging for electric vehicles: A survey from the algorithmic perspective," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1500–1517, 2nd Quart., 2016.

[12] E. Celebi and J. D. Fuller, "Time-of-use pricing in electricity markets under different market structures," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1170–1181, Aug. 2012.

[13] E. Celebi and J. D. Fuller, "A model for efficient consumer pricing schemes in electricity markets," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 60–67, Feb. 2007.

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

[15] Y. Cao *et al.*, "An optimized EV charging model considering TOU price and SOC curve," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 388–393, Mar. 2012.

[16] Y. Wang, H. Nazaripouya, C.-C. Chu, R. Gadh, and H. R. Pota, "Vehicle-to-grid automatic load sharing with driver preference in micro-grids," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol., Eur.*, Oct. 2014, pp. 1–6.

[17] Y. Xiong, B. Wang, Z. Cao, C.-C. Chu, H. Pota, and R. Gadh, "Extension of IEC61850 with smart EV charging," in *Proc. IEEE Innov. Smart Grid Technol.-Asia (ISGT-Asia)*, Nov. 2016, pp. 294–299.

[18] Y. Wang, O. Sheikh, B. Hu, C.-C. Chu, and R. Gadh, "Integration of V2H/V2G hybrid system for demand response in distribution network," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 812–817.

[19] Z. Wan, H. Li, H. He, and D. Prokhorov, "Model-free real-time EV charging scheduling based on deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5246–5257, Sep. 2019.

[20] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 940–951, May 2013.

[21] A.-M. Koufakis, E. S. Rigas, N. Bassiliades, and S. D. Ramchurn, "Offline and online electric vehicle charging scheduling with V2V energy transfer," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 5, pp. 2128–2138, May 2020.

[22] H. Qin and W. Zhang, "Charging scheduling with minimal waiting in a network of electric vehicles and charging stations," in *Proc. 8th ACM Int. Workshop Veh. Inter-Netw. (VANET)*, 2011, pp. 51–60.

[23] C. Liu, M. Zhou, J. Wu, C. Long, and Y. Wang, "Electric vehicles en-route charging navigation systems: Joint charging and routing optimization," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 2, pp. 906–914, Mar. 2019.

[24] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1095–1105, Sep. 2012.

[25] H.-J. Kim, J. Lee, G.-L. Park, M.-J. Kang, and M. Kang, "An efficient scheduling scheme on charging stations for smart transportation," in *Security-Enriched Urban Computing and Smart Grid*, T.-H. Kim, A. Stoica, and R.-S. Chang, Eds. Berlin, Germany: Springer, 2010, pp. 274–278.

[26] M. M. de Weerdt, E. H. Gerding, S. Stein, V. Robu, and N. R. Jennings, "Intention-aware routing to minimise delays at electric vehicle charging stations," in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013, pp. 83–89.

[27] F. Baouche, R. Billot, R. Trigui, and N. E. El Faouzi, "Electric vehicle green routing with possible en-route recharging," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 2787–2792.

[28] M. Schneider, A. Stenger, and D. Goeke, "The electric vehicle-routing problem with time windows and recharging stations," *Transp. Sci.*, vol. 48, no. 4, pp. 500–520, Nov. 2014.

[29] F. Alesiani and N. Maslekar, "Optimization of charging stops for fleet of electric vehicles: A genetic approach," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 3, pp. 10–21, Jul. 2014.

[30] X. Chen, S.-C. Wong, and C. K. Tse, "Adding randomness to modeling Internet TCP-RED systems with interactive gateways," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 4, pp. 300–304, Apr. 2010.

[31] P. Zhang, X. Kang, X. Li, Y. Liu, D. Wu, and R. Wang, "Overlapping community deep exploring-based relay selection method toward multi-hop D2D communication," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1357–1360, Oct. 2019.

[32] H. Hoimoja, A. Rufer, G. Dziechciaruk, and A. Vezzini, "An ultra-fast EV charging station demonstrator," in *Proc. Int. Symp. Power Electron. Power Electron., Electr. Drives, Autom. Motion*, Jun. 2012, pp. 1390–1395.

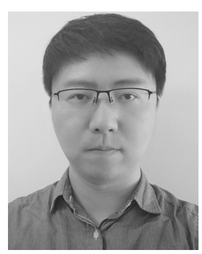
[33] J. K. Lenstra, A. R. Kan, and P. Brucker, "Complexity of machine scheduling problems," in *Annals of Discrete Mathematics*, vol. 1. Amsterdam, The Netherlands: Elsevier, 1977, pp. 343–362.
[34] Y. K. Lin, M. E. Pfund, and J. W. Fowler, "Heuristics for minimizing regular performance measures in unrelated parallel machine scheduling problems," *Comput. Oper. Res.*, vol. 38, no. 6, pp. 901–916, Jun. 2011.
[35] G. A. Seber and A. J. Lee, *Linear Regression Analysis*, vol. 329. Hoboken, NJ, USA: Wiley, 2012.
[36] R. S. Sutton, D. Precup, and S. Singh, "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning," *Artif. Intell.*, vol. 112, nos. 1–2, pp. 181–211, Aug. 1999.
[37] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, p. 529, 2015.
[38] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT*, Y. Lechevallier and G. Saporta, Eds. Berlin, Germany: Physica-Verlag HD, 2010, pp. 177–186.
[39] F. Meng, G. Yuan, S. Lv, Z. Wang, and S. Xia, "An overview on trajectory outlier detection," *Artif. Intell. Rev.*, vol. 52, no. 4, pp. 2437–2456, Dec. 2019.
[40] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-matching for low-sampling-rate GPS trajectories," in *Proc. 17th ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2009, pp. 352–361.
[41] A. Graser, M. Straub, and M. Dragaschnig, "Towards an open source analysis toolbox for street network comparison: Indicators, tools and results of a comparison of OSM and the official Austrian reference graph," *Trans. GIS*, vol. 18, no. 4, pp. 510–526, Aug. 2014.
[42] M. Zhu, X.-Y. Liu, L. Kong, R. Shen, W. Shu, and M.-Y. Wu, "The charging-scheduling problem for electric vehicle networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 3178–3183.

Cong Zhang (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the Beijing University of Posts and Telecommunications, China. His research interests include the Internet of Things, sensor search, spatial–temporal big data, urban intelligence, and data-driven cyber-physical systems (CPS).

Yuanan Liu (Member, IEEE) received the M.E. and Ph.D. degrees from the Chengdu University of Electronic Science and Technology, China, in 1989 and 1992, respectively. He was a Post-Doctoral Researcher with the Beijing University of Posts and Telecommunications, China, from 1992 to 1994. He has worked with Carleton University, Canada, from 1995 to 1997. He is currently the Director of the Beijing Key Laboratory of Safety Production Intelligent Monitoring and the Vice-Chairman of (TC9) in China Communication Standards Association (CCSA), majoring in EMC, mobile communications, and the Internet of Things. He is a fellow of IEE.

Fan Wu (Member, IEEE) received the Ph.D. degree from the Beijing University of Posts and Telecommunications, China, in 2009. She is currently an Associate Professor with the School of Electronic Engineering, Beijing University of Posts and Telecommunications. Her research interests include the Internet of Things, sensor search, and pervasive computing.

Bihua Tang received the M.S. degree from the Chengdu University of Electronic Science and Technology, China, in 1989. She has worked with Carleton University, Canada, in 1996. She spent one year as a Senior Visiting Scholar with the National Institute of Informatics, Japan, in 2002. She is currently a Professor with the School of Electronic Engineering, Beijing University of Posts and Telecommunications, China. Her research interests include the Internet of Things, mobile computing, and wireless sensor networks.



Wenhao Fan (Member, IEEE) received the Ph.D. degree from the Beijing University of Posts and Telecommunications, China, in 2013. He is currently an Associate Professor with the School of Electronic Engineering, Beijing University of Posts and Telecommunications. His research interests include mobile terminal, cloud computing, and the Internet of Things.