bCharge: Data-Driven Real-Time Charging Scheduling for Large-Scale Electric Bus Fleets

Guang Wang[1], Xiaoyang Xie[1], Fan Zhang[2], Yunhuai Liu[3], Desheng Zhang[1]
[1]Rutgers University, [2]SIAT, Chinese Academy of Sciences, [3]Peking University
gw255@cs.rutgers.edu, xx88@cs.rutgers.edu, zhangfan@siat.ac.cn, yunhuai.liu@pku.edu.cn, desheng.zhang@cs.rutgers.edu

Abstract*—We are witnessing a rapid growth of electrified vehicles because of the ever-increasing concerns over urban air quality and energy security. Compared with other electric vehicles, electric buses have not yet been prevailingly adopted worldwide due to the high owning and operating costs, long charging time, and the uneven distribution of charging facilities. Moreover, the highly dynamic environment factors such as the unpredictable traffic congestions, different passenger demands, and even changing weather, can significantly affect electric bus charging efficiency and potentially hinder further development of large-scale electric bus fleets. To deal with these issues, in this paper, we first analyze a real-world dataset including massive data from 16,359 electric buses, 1,400 bus lines and 5,562 bus stops, which is obtained from the Chinese city Shenzhen, who has the first and the largest full electric bus network for public transit. Then we investigate the electric bus network to understand its operating and charging patterns, and further verify the feasibility and necessity of a real-time charging scheduling. With such understanding, we design *bCharge*, a real-time charging scheduling system based on Markov Decision Process to reduce the overall charging and operating costs for city-scale electric bus fleets, taking the time-variant electricity pricing into account. To show the effectiveness of *bCharge*, we implement it with the real-world streaming dataset from Shenzhen, which includes GPS data of the electric bus fleet, the bus lines and stops data, coupled with the 376 electric bus charging stations data. The evaluation results show that *bCharge* can dramatically reduce the charging cost by 23.7% and 12.8% electricity usage simultaneously.

I. Introduction

With the growing concerns over the air quality and energy security, more countries such as China and U.K. have started their electric vehicle initiatives to reduce emissions and energy consumption [1]. It is reported that the worldwide sales of electric vehicles have been nearly quadrupled since 2014, and half of the vehicle sales will be electric vehicles by 2027 [2].

As one of the most common mass transportation, buses play an important role in people's daily life [3]–[6]. Because of the long travel distance and high-frequency services, electric buses (e-bus) have greater potentials to reduce the carbon dioxide and nitrogen oxides emissions [1] compared to other electric vehicles, e.g., electric private vehicles (e-pvs) [7] and electric taxis (e-taxis) [8]–[11]. Yet till date, e-buses have not been extensively adopted in the worldwide because of the following distinctive characteristics: (i) lack of spacious charging infrastructures for large-scale e-bus fleets, e.g., large charging stations with lots of parking spaces and charging stations [12]; (ii) high purchase costs due to the relatively new technologies, e.g., a BYD e-bus is around $263,000 [12] in 2017, which is 2.5 times of a diesel bus; (iii) relatively high operating cost caused by the charging fees (i.e., electricity) compared to one-time costs of infrastructure construction and ownership. For example, the yearly charging cost for one e-bus is about $18,000 [13] in Shenzhen, resulting in the day-to-day charging costs being one of the key concerns that hinder the e-buses to release their potentials fully [1].

There are many works have been conducted on how to reduce the charging costs for e-taxis and e-pvs [8], [9], [14], and some other works [15]–[22] have built the theoretical models and simulations for e-buses. However, few works, if any, have been done on the data-driven modeling and optimization for real-world e-bus fleets charging. More importantly, results for e-taxis [8], [9], [14], [23] can hardly be directly applied to the e-buses because of the following fundamentally different features. (i) The charging activities of e-taxis are directly related to the income of e-taxi drivers, so their charging and routes will be incentive-based; whereas e-bus drivers are not. (ii) The charging activities of e-taxis are mostly distributed and flexible. An e-taxi driver can decide when and where to charge, while the e-bus network is based on centralized operating and charging management. These two key differences lead to different charging incentives and optimization goals [7].

E-buses are centrally managed with fixed timetables, which makes it possible to design offline charging schedules and operating strategies. However, such offline strategies are not always optimal. The high dynamics of the real-world environments bring great challenges for optimal solutions, making e-buses very different from the flexible e-taxis and e-pvs. Such dynamics may include the unexpected break-downs of e-buses, unpredictable traffic congestions, time-variant electricity rates, as well as the changing weather/temperature and the traffic-light conditions [13]. For example, both too hot or too cold weather will require the e-buses to open their air conditioners, which drain their energy quickly. These dynamic factors will lead to both (i) the non-deterministic departure and arrival time for each e-bus and (ii) unpredictable State of Charge (SOC, i.e., the remaining battery level) of e-buses when they arrive at terminals. Note that the fixed timetables only require e-buses to leave the terminal in time. As a result, the off-line charging schedule and operation strategies, i.e., all e-bus lines follow the predesigned and fixed operating and charging patterns, may be far from the high operating and charging efficiency. For example, some e-buses may wait too long in the charging stations for available charging points, while other

2576-3172/18/$31.00 ©2018 IEEE
DOI 10.1109/RTSS.2018.00015

charging stations have a lot of unoccupied charging points.

To address these real-time issues, we develop *bCharge*, a data-driven real-time charging scheduling system for large-scale e-bus fleets based on real-world data. The dataset is obtained from the Chinese city Shenzhen, a pilot city that promotes e-buses in China. Shenzhen has electrified 100% of its public transit buses and became the first and the only city with a full e-bus network in the world. Moreover, Shenzhen also has the largest installation base of e-buses, e.g., 16,359 e-buses [24] in December 2017. Based on Shenzhen e-bus data, we perform a set of data-driven analyses to understand the behaviors of e-bus fleets and then design a data-driven real-time scheduling strategy to reduce the overall charging cost of the fleets. Our key contributions are as follows:

- To our best knowledge, we are the first to conduct the city-scale data-driven investigation of the real-time scheduling for electric bus fleets. Our investigation has two key features based on four-year real-world e-bus data including (i) the largest e-bus fleet in the world with more than 16,000 e-buses; (ii) the largest number of e-bus charging stations and charging points, e.g., 376 charging stations for e-buses. Such a large-scale data-driven investigation enables us to identify the real-world e-bus operating and charging issues, which are challenging to reveal by using simulation studies.
- We design a data-driven real-time charging scheduling system called *bCharge* taking factors of the e-bus daily and per-charge operating distances, charging spatial-temporal and cost distribution, charging station utilization rate into account. *bCharge* is based on a thorough data-driven analysis, which reveals some novel insights including e-bus adoption process, e-bus demand/supply, charging network distribution, charging activities distribution, charging cost, etc. To our knowledge, this is the first time that such detailed analyses were performed for a large-scale e-bus fleet to support data-driven scheduling.
- Given these data-driven insights, *bCharge* is designed based on the Markov Decision Process by considering the status of the e-bus fleet and the contextual factors. In particular, we consider both revenues and charging cost, along with timetables and time-of-use electricity prices to schedule e-buses among different bus lines. We also theoretically investigate the real-time features of our MDP-based scheduling by analyzing its time complexity. Finally, we elaborate on how our scheduling strategy guarantees timetables for the real-time requirement.
- We implement and evaluate *bCharge* based on the real-world data in Shenzhen. The results show *bCharge* reduces 23.7% of the overall charging cost and 12.8% of the electricity usage. Besides, some lessons learned and experience were reported, which are helpful for other cities to promote and optimize their e-bus fleets.

The rest of the paper is organized as follows. Section II introduces our dataset and conducts the detailed analyses. Section III presents the design and implementation of *bCharge*. Section IV evaluates the performance of *bCharge*. The lessons learned and related works are summarized in Sections V and VI. Finally, we conclude this paper in Section VII.

II. *bCharge*: Datasets and Analyses

In this section, we first describe our large-scale real-world dataset generated from the Shenzhen e-bus fleet. Based on the dataset, we then comprehensively investigate the operating, charging and cost patterns of the Shenzhen e-bus fleet to motivate our real-time charging scheduling.

TABLE I
Specifications of one type e-bus in Shenzhen

Model	Capacity	Length	Charging	Max. Speed	Max. Dist
BYD K9	324 kWh	12m	3h	90 km/h	250 km

A. Data Description

Our *bCharge* is based on large-scale e-bus datasets obtained from Shenzhen, the 4th-largest city in the Chinese mainland. The time span of these datasets is from the year 2014 to 2018, during which Shenzhen has experienced a very fast growth of e-buses, i.e., the percentage of e-buses among all buses (i.e., e-buses and diesel buses) has raised from 7.8% to 100% as shown in Fig. 1. One of the most popular e-bus vehicle models in Shenzhen is BYD K9, whose specifications are shown in Tab. I. The battery capacity and maximum traveling distance of BYD K9 are 324 kWh and 250 km, respectively [25]. The four-year datasets include five different types of data, and the details are given as follows.

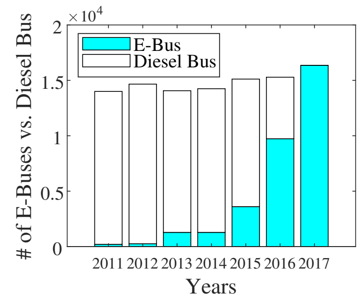

Fig. 1. Number of e-buses and diesel buses.

- **GPS Data** include 1.92TB historical and real-time GPS records of all buses in Shenzhen from July 2014 to May 2018. Each record includes 19 fields describing the status of a bus, e.g., the time-stamps, the bus ID, the bus line ID, the GPS location (i.e., longitude and latitude), the current speed, the direction and the total mileage (i.e., odometer data). The GPS data are collected by an onboard device with a cellular connection in real time.
- **Bus Stop Data** include all bus stops' information of 1,400 bus lines (including inbound and outbound directions) with 5,562 unique bus stops. For each bus stop, there are seven key fields including the route ID, the line direction, the stop name, the GPS location, etc.
- **Bus Transaction Data** include all transaction records of passengers' trip fares. The average daily number of passengers taking buses using smartcards is about 2.4 million, and 5 million in total. Each transaction has six key fields including the route ID, the line direction, the station ID, the station name and the GPS location.
- **Bus Charging Station Data** include the station names, the station IDs, the GPS locations and the number of

charging points in each station. There are 376 e-bus charging stations in Shenzhen as of the end of 2017.

- **Electricity Rate Data** include the time-variant electricity pricing within 24 hours of Shenzhen. Shenzhen adopts the time-of-use rating, which breaks up 24 hours of a day into several intervals and charges a different price for each interval [26]. The rates in Shenzhen are divided into three types, i.e., off-peak prices (low rates), semi-peak prices (medium rates, also called flat rates) and peak prices (high rates), and the corresponding electricity rates are 0.049, 0.121, and 0.173 \$/kWh, respectively. The time-variant electricity pricing in Shenzhen is shown as Fig. 2.

Based on these data, we perform an intensive data-driven analysis to understand the operating and charging patterns of the Shenzhen e-bus fleet. The details are shown below.

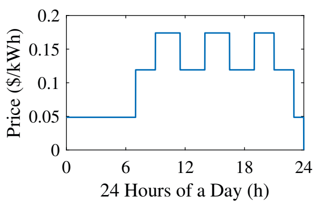

Fig. 2. Time-variant industrial electricity prices in Shenzhen.

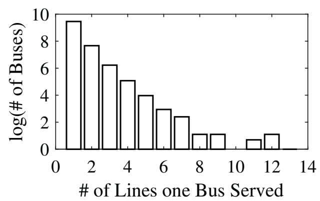

Fig. 3. The number of lines served for each bus in Shenzhen.

B. Operating Patterns

A visualization of Shenzhen e-bus network is shown as Fig. 4, where the yellow and red parts stand for the bus lines across Shenzhen. There are more passengers in the yellow part and fewer passengers on the red lines.

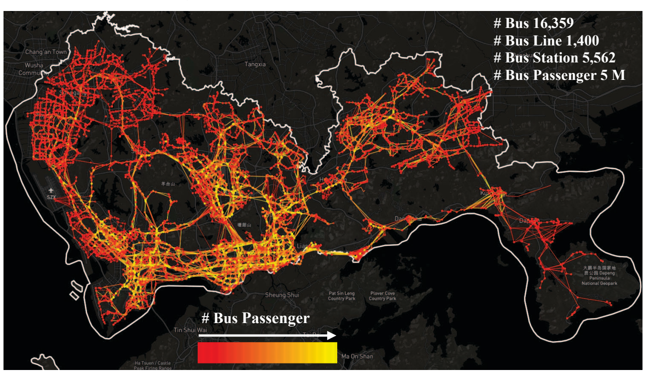

Fig. 4. Station and line distribution of Shenzhen bus network.

We found that the bus lines reach to the most remote areas although the highest density of lines is in the central business district area. We also found that the bus in Shenzhen may serve different lines at different times of a day, e.g., rush hour and non-rush hour. The number of lines served by each bus can be seen from Fig. 3. We found that about one-fifth of buses in Shenzhen serve for more than one fixed line, which indicates it is feasible for us to schedule e-buses to serve other lines.

We further investigate the daily operating distances and the distances between two charges of Shenzhen e-buses. The Cumulative Distribution Function (CDF) of operating distances between two charges of all e-buses is in Fig. 5.

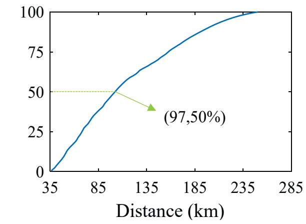

Fig. 5. Distance betw. two charges.

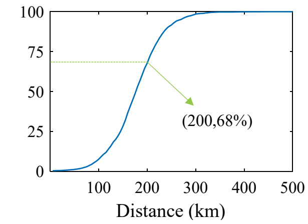

Fig. 6. Daily operating distance.

We found that about 50% e-buses operate no more than 97 km between two charges, even though their maximum operating distance is around 250 km, which is caused by many real-world factors, e.g., the availability of charging points, the electricity prices, and the range anxiety. Further, we show the CDF of total daily operating distances of each e-bus in Fig. 6. We found that the daily operating distances of 68% e-buses are less than 200 km, which is the maximum practical distance most e-buses would travel before a charge. Besides, 32% e-buses travel more than 200 km per day, which means they need at least two charges per day. Considering 50% e-buses' operating distance between two charges is no more than 97 km, most e-buses need to charge at least two times per day.

C. Charging Patterns

We utilize long-term bus data and charging station data to fully understand the overall e-buses charging patterns in Shenzhen. The spatial distribution of the e-bus charging stations is in Fig. 7, where the sizes of the circles indicate the number of charging points in each station.

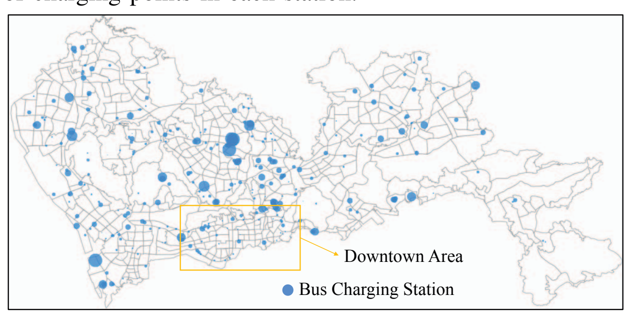

Fig. 7. Spatial distribution of Shenzhen e-bus charging network.

We found that most large stations are located in suburban areas, which is because there are more available and cheaper land resources in these areas than that in downtown areas.

Fig. 8 shows the charging start time distribution of the e-bus fleet. We found that there are two clustering charging time durations, i.e., 00:00-6:00 and 16:00-20:00, while very few charging events occur around 12:00 due to high electricity prices. The probability of charging events starting at various types of electricity rates is shown in Fig. 9. We also found that about 60% charging events start at off-peak hours but there is about 18% charging occurring during peak hours.

We further investigate the daily charging time of each e-bus. As shown in Fig. 10, only 13% e-buses spend less than 3 hours for charging each day but the charging time of over 80% e-buses is no more than 5 hours, which indicates most e-buses spend 3-5 hours for at least two charges in each day.

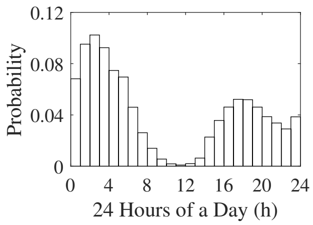

Fig. 8. Charging time distribution.

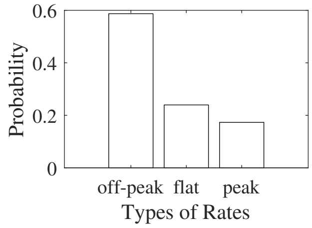

Fig. 9. Charge rate distribution.

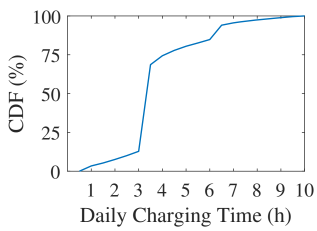

Fig. 10. Daily total charging time.

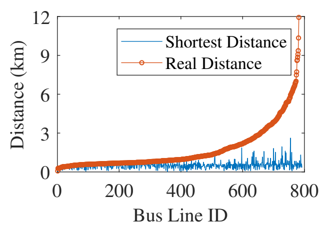

Fig. 11. Dist. to charging stations.

In the Shenzhen bus network, all the e-buses for the same line typically utilize the same charging stations, but they may not be the closest charging stations to their terminals. It leads them to take detours to charge due to parking issues even though there are closer charging stations. As a result, we investigate the distances between bus lines' terminals and the closest charging stations to them. Besides, we also study the distances between bus lines' terminals and the charging stations they actually went to charge. As shown in Fig. 11, we rank the bus lines according to the actual traveling distances to charge in ascending order. Even though 52% bus lines have charging stations within one kilometer, there are still 7% bus lines need to go to charging stations beyond 5 km away.

To study the station effectiveness, we define **Charging Station Utilization Rate** at a station s_i in a day as follows.

$$UR(s_i) = \frac{CE(s_i)}{CP(s_i)}, \tag{1}$$

where $CE(s_i)$ is the daily number of charging events in the station s_i; $CP(s_i)$ is the number charging points in station s_i.

Fig. 12 shows the daily charging station utilization rates in Shenzhen. We found that the utilization rates of 50% charging stations are no more than 4. However, there are also some charging stations with very high utilization rates, e.g., 12. These unbalanced utilization rates lead to the charging resources waste in some stations while crowded and long-waiting charging events in some other stations.

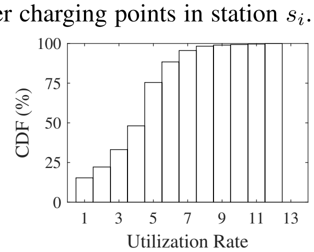

Fig. 12. Daily station utili. rate.

Besides, we have conducted a series of field studies in 2018 to investigate the charging patterns of e-buses in Shenzhen. Fig. 13 shows the e-buses and a charging station we visited in Shenzhen. More details will be given in Section V.

（a）E-buses in Shenzhen （b）Charging station in Shenzhen

Fig. 13. Electric buses and charging stations in Shenzhen.

D. Cost Patterns

Based on these charging station utilization rates, we further study charging start time and charging distribution over time of day to understand the electricity usage and charging costs. As shown in Fig. 14, the highest electricity usage for e-bus charging occurs in the off-peak hours, accounting for 63% of the total electricity usage. However, there is still 13.6% electricity usage during the peak hours. Although the percentage of electricity usage in the peak hours and flat hours is much lower than the usage in the off-peak hours, the usage in these two durations can cause more charging costs than the peak-hour charging due to the time-of-use pricing strategy.

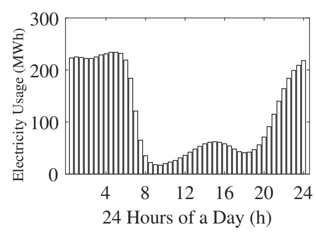

Fig. 14. Electricity usage.

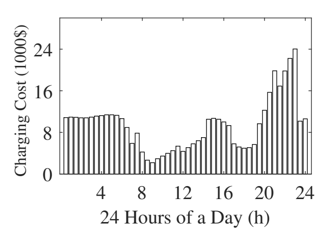

Fig. 15. Charging cost distribution.

In Fig. 15, the charging cost distribution is different from the electricity usage distribution. In particular, we found the electricity cost gap between 8:00-20:00 and 20:00-8:00 in Fig. 15 is much smaller than the electricity usage gap between the same two periods in Fig. 14. This indicates that even though the bus fleet does not charge much during 8:00-20:00, the costs are almost as high as 20:00-8:00, during which the e-buses charge substantially. As a result, it motivates us to answer a question that whether we can reduce the charging cost by scheduling more charges in off-peak hours, i.e., further increasing the electricity usage gap between 8:00-20:00 and 20:00-8:00 since a small usage decrease during peak hours will result in a huge cost decrease.

III. *bCharge*: Scheduling Design

We first present our problem formulation by showing the existing operating and charging scheduling for Shenzhen e-buses and our scheduling idea. Then, we introduce the detailed design of our *bCharge* in terms of scheduling formulation, scheduling design, and scheduling complexity analyses. Finally, we describe the timetables guarantee of our scheduling.

A. Problem Formulation

1) Existing Operating and Charging Patterns: Fig. 16 shows the operating patterns of buses in Shenzhen. In Shenzhen, a bus line generally has two terminals, e.g., Terminal *A* and Terminal *B* in Fig. 16. Based on the timetable for this line, a Bus 1 travels from Terminal *A* to Terminal *B* (or from *B* to *A*) through some intermediate bus stops, and then goes back to the *Terminal A* (or *B*). During the same time, multiple buses are serving for this line with same or different directions, e.g., Bus 2. For e-buses, they need to charge when their SOC decrease to a pre-defined low threshold. Besides, different from other electric vehicles (e.g., e-taxis and e-pvs), e-buses generally make charging decisions (i.e., to charge or to continue to serve) only after they arrive at bus terminals. This is because of their operational feature, i.e., they normally cannot charge in the middle of a trip with bus passengers onboard. As a result, based on this feature, the charging stations in Shenzhen are usually deployed near terminals or at terminals, which also addresses the parking issues since e-buses need lots of space for parking compared to taxis and private vehicles. Moreover, in Shenzhen, nearly half of the e-buses in the fleets charge at their closest terminals as Fig. 11 shows, (e.g., the terminals they just arrived) which may potentially decrease the charging efficiency of the entire e-bus charging network due to the unbalanced charging points placement (e.g., fewer charging points in the downtown and more in the suburb), resulting in no charging points available when e-buses arrive at their charging stations but many unoccupied charging points in some other charging stations.

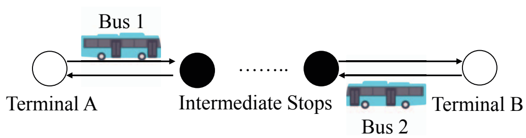

Fig. 16. Existing bus operating and charging patterns.

2) bCharge Operating and Charging Patterns: In this work, we focus on the real-time charging scheduling problem for e-bus fleets, considering the overall operational cost and revenues of all e-buses in the fleet, instead of individual vehicles, given its centralized management mode. Compared to conventional diesel buses, e-buses are less flexible due to their limited operating ranges and reliance on the charging infrastructures, which makes it challenging to schedule the e-bus fleet to operate and charge, especially for a large-scale fleet. Besides, the lower flexibility of e-buses potentially makes it necessary to (i) purchase more extra buses for contingency plans to cover additional ranges on routes or (ii) redesign the lines to accommodate e-buses. These two actions have been taken in Shenzhen. Moreover, the time-variant electricity rates compared with the 24-hour stable diesel price also cause the charging issues of e-bus fleets more complicated.

In *bCharge*, we consider different real-world factors for e-buses including time-variant electricity rates and scheduling between different lines/extra e-buses. Weather conditions, congestion, time of day, and demographic features as contexts are implicitly considered since we leverage both the historical and real-time GPS records to predict energy consumption to serve a particular line with detailed routes. Based on previous research [27]–[29], both periodic congestion and static demographic features remain stable for the same spatial temporal combination, e.g., for the same road segment during the same time of different days. As a result, our historical GPS records implicitly contains the period congestion and basic demographic features. Our key idea for *bCharge* is that we schedule some e-buses to serve other bus lines which share the same terminals with it when they arrive at their terminal based on some real-world factors. These factors we considered include: (i) the real-time SOC of e-buses; (ii) the availability of charging points in charging stations; (iii) the expected energy consumption of different lines, which is related to both lengths and travel time related to traffic; (iv) the expected charging cost at particular terminals for different time slots of a day, which is related to both charging point availability (e.g., staying without charging if no point available) and electricity prices. Our final consideration for the scheduling is that we have to guarantee the timetables of all lines, which we will clarify in Sec. III-C.

Based on this key idea, we make streaming scheduling decisions after each e-bus arrives at a terminal and drops off all passengers since all e-buses arrive at terminals in an online fashion [7]. In particular, *bCharge* has four potential scheduling modes for an e-bus when it is in a terminal based on the four factors we discussed in last paragraph: (i) stay at the terminal; (ii) charge at the terminal immediately; (iii) keep serving the current line, i.e., go back to the terminal it came from; (iv) serve another line, i.e., go to a different terminal.

3) Scheduling Objective: The objective of our *bCharge* is to optimize the e-bus fleet by reducing the overall operational cost and increasing the profits by collecting fares for serving passengers, which can be formulated as follows.

$$F_s - C_c = \sum_{t \in 24h} \sum_{n=1}^{N_{eb}} (F_n^t - R^t \cdot C_n^t) \tag{2}$$

where F_s is the collected fare for serving passengers of the whole e-bus fleet; C_c is the charging cost for operating the e-bus fleet; N_{eb} is the number of e-buses in the fleet; F_n^t is the fare collected by the n^{th} e-bus during time t by serving passengers for a particular line; R^t is the electricity rate at time t; C_n^t stands for the electricity consumed (i.e., charged) by the n^{th} e-bus during time t. As Eq. 2 shows, the overall optimization objective depends on three items, e.g., the real-time electricity rates, the energy charged by each e-bus and the fares collected by each e-bus. Compared to the taxi and private vehicles, bus passenger demand is more stable and the fare is flatter in Shenzhen [30], so the expected fare can be obtained by historical data given a time and a bus line. As a result, we focus on deciding on scheduling e-buses to serve which lines and when to charge for reducing the overall charging cost.

An intuitive idea to achieve this goal is to have enough buses and to schedule all the e-buses to serve their original lines and to charge only during the off-peak hours, i.e., 23:00-7:00. However, given the limited buses in Shenzhen and a large number of lines as indicated by Fig. 6, around 32% e-buses cannot accomplish the daily operating task with one charge during nights alone. Note that for practical consideration, we only focus on scheduling existing buses for improving the charging efficiency, instead of adding new buses in this paper.

B. Charging Scheduling

Since e-buses arrive at bus terminals in an online fashion, we schedule the charging task for e-buses one by one. One of our major technical contributions in this paper is that we formulate the e-bus charging scheduling problem as a Markov Decision Process (MDP) problem to reduce the charging cost.

An MDP is a discrete-time state transition system, which aims to find an optimal policy to maximize the expected utility. Formally, an MDP is defined as a 5-tuple (S, A, T, R, β) [31], [32]. The MDP framework of *bCharge* is shown as Fig. 17.

- S is a set of states. For the charging scheduling scenario, we define four different states according to the SOC of e-buses. As Fig. 17 shows, the four states in *bCharge* are (i) SOC$_f$, which indicates an e-bus is at the Full SOC; (ii) SOC$_c^f$, which indicates the SOC of an e-bus is lower than the Full SOC but higher than a SOC with which it can serve the Current line, i.e., go back to the original terminal; (iii) SOC$_l^c$, which indicates the SOC of an e-bus is lower than the required SOC to serve the Current line but higher than a mandatory charging threshold, i.e., it may still have SOC to serve other lines sharing the same terminal with it; (iv) SOC^l, which indicates the SOC of an e-bus is below than the mandatory charging threshold, i.e., it needs to charge and cannot serve any lines.
- A is a set of actions. In *bCharge*, there are four actions: (i) A$_S$: Staying at this terminal but not to charge; (ii) A$_O$: going back to the Original terminal, i.e., serving the current line; (iii) A$_N$: serving a New line; (iv) A$_C$: Charging at this terminal. For the charging scheduling, the results of the former three actions follow the energy non-increasing principle, i.e., the energy in the next state does not exceed the current state; whereas, for the state A$_C$, i.e., charging, it increases the SOC of e-buses.
- T is a state transition matrix, which consists of the probability transition from one state to another state by taking an action. For example, T(SOC$_f$, A$_O$, SOC$_c^f$) =P$^{A_O}_{SOC_f \rightarrow SOC_c^f}$ = p$_{01}$ means the probability of an e-bus transferring from the full battery capacity SOC$_f$ to SOC$_c^f$ by serving its own line (i.e., action A$_O$) is p$_{01}$.
- R is a reward function. For each action, the scheduling strategy will cause a corresponding reward. In the charging scheduling problem, if an e-bus stays at the terminal (i.e., A$_S$), it will have no passengers (i.e., no revenue) and also have no energy consumption (i.e., no cost), so the current reward is 0. When e-buses serve for passengers (i.e., the action A$_O$ or A$_N$), they will have passenger fares for this line, so they will have a positive reward V$_i$, which is based on which line they serve. But if they take the action A$_C$, there is a charging cost for electricity, so the reward is negative, which is denoted as the -W$_j$ in Fig. 17, where $1 \leq j \leq 3$. The different negative reward values depend on the real-time electricity rates and the current SOC, as well as their full battery capacity.
- β is the discount factor, which captures the fact that an immediate reward might be worth more than the same reward in the future. The value of β is generally selected from [0,1), so the final expected utility will be convergent and bounded to a finite number. β is set up to 0 if and only if we do not consider the future reward.

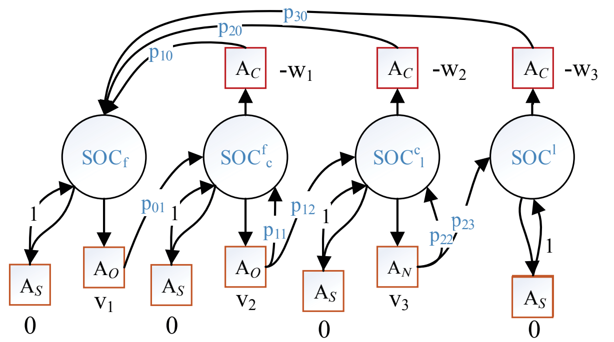

Fig. 17. Markov decision process for charging scheduling.

Definition 2: a **policy** π is defined as a distribution over actions given states, which gives the e-bus an action to execute at each state so as to maximize the expected utility.

$$\pi\left(a|s\right) = P\left[A = a|S = s\right] \tag{3}$$

where $a \in$ A(s) = {A$_S$, A$_O$, A$_N$, A$_C$} and $s \in$ {SOC$_f$, SOC$_c^f$, SOC$_l^c$, SOC^l}.

Definition 3: an **utility** of a state for a given policy is defined as $U^\pi(s)$, which can be formulated as

$$U^\pi(s) = E\left[\sum_{t=0}^{\infty} \beta^t \cdot R(s_t)|\pi, s_0 = s\right] \tag{4}$$

where s_0 is the initial state; s_t is the state of the e-bus after executing the policy π for t actions; β^t is the discount after t actions; and $R(s_t)$ is the immediate reward at each state.

The objective of the data-driven charging scheduling strategy is to derive an optimal policy π^* that achieves the maximum utility $U^*\left(s\right)$ for all states, which is formulated as the Bellman Equation [33] as Eq. 5.

$$\begin{aligned} U^*\left(s\right) &= \max_{\pi} U^\pi\left(s\right) \\ &= \max_{a \in A(s)} \left[R_{sa} + \beta \cdot \sum_{s'} P^a_{ss'} \cdot U^*\left(s'\right)\right] \end{aligned} \tag{5}$$

where R_{sa} is the immediate reward after taking the action *a* in the state *s*, which is also the operational revenue or charging cost after taking different actions. $\sum_{s'} P^a_{ss'} \cdot U^*\left(s'\right)$ is the expected future utility. As Eq. 5 shows, the immediate reward R_{sa}, the discount factor β, and the transition probability $P^a_{ss'}$ are required in order to obtain the best policy.

As shown in Eq. 2, the objective of our *bCharge* is to minimize the overall charging cost of the e-bus fleet and maximize fares collected for serving passengers. The fares would be reduced if the e-buses do not keep the timetable, e.g., some passengers may take taxis. In this work, we envision that the fares are maximized if the timetables of all lines are kept as much as possible, which is ensured by our scheduling policy, which we will explain in Section III-C. As a result, the objective of Eq. 2 is equivalent to the objective of Eq. 5, which indicates we can leverage an existing MDP solver, i.e., MDP

Toolbox, to solve the e-bus charging scheduling problem by considering both charging cost and revenues of collected fares. In the next section, we will theoretically investigate if this MDP-based scheduling can satisfy the real-time requirement by studying its scheduling complexity.

Scheduling Complexity: There are two common approaches to solve the MDP optimization problem, i.e., value iteration and policy iteration [34]. There are *n* unknowns needed to solve in the equations when there are *n* states. If we leverage the value iteration, we need to use *max* in the equations, which is nonlinear, resulting in complexity of $O\left(m \cdot n^2\right)$ for each iteration for *m* actions. However, the policy iteration has the operation $\sum$ instead of the operation *max* in the equations, which implies that the equations are linear. Thus, solving these *n* linear equations is with a complexity of $O\left(n^3\right)$. Since we have four states and four actions in our scheduling process, the time complexity of the two approaches is similar for each iteration. However, the policy iteration searches a finite policy space instead of an uncountably infinite value space, indicating the policy iteration converges much faster than the value iteration. Now, we have shown that the policy iteration satisfies the real-time requirement for scheduling. In the worst case, the policy iteration for *bCharge* needs to search all the policy space, which is limited by the number of our states and the limited number of lines sharing the same terminals. Then the total time cost of the policy iteration for *bCharge* is linear, which can be easily realized at the second level for a normal PC, so the policy iteration based MDP is fast enough for the real-time charging scheduling requirement. Therefore, we leverage the policy iteration to find the best charging scheduling policy for e-buses.

C. Timetable Guarantee

Since keeping the timetable is critical for the bus fleet operation, our algorithm considers the timetable constraint and effectively guarantees timetables for e-bus lines when performing the scheduling. Our scheduling can be regarded as a priority-based scheduling and the priority is decided by the utilities that different strategies can achieve, which means a higher utility strategy has a higher priority. As follows, we show how our scheduling algorithm guarantees timetables for bus lines with an example. As shown in Fig. 18, the line L_1 and line L_2 share the same terminal T_1, and there are many e-buses serving for different lines, e.g., B_{L11} for L_1 and B_{L21} for L_2. There are three different scenarios when *bCharge* makes scheduling: (i) If the next e-bus B_{L11} for line L_1 can have a better scheduling to serve another line, e.g., L_2, compared to the original line L_1 and another e-bus B_{L12} or B_{L13} is available for L_1, then we schedule B_{L11} to serve another line (e.g., L_2) and have B_{L12} or B_{L13} to take the place of B_{L11}. We utilize the charging deadline to break the tie if multiple e-buses were available, which means our scheduling strategy will choose the bus with the highest SOC to replace B_{L11}. In this way, the timetable of L_1 is guaranteed. (ii) If we cannot find the next bus B_{L12} and B_{L13} for L_1 to guarantee the timetable of B_{L11}, we find all other lines that share the terminals with L_1, i.e., L_2 in this example, for an extra backup bus for L_1 (which do not affect their timetables). If there is an e-bus available from B_{L21}, B_{L22} and B_{L23}, then the timetable of L_1 can be guaranteed while we schedule B_{L11} for other lines. (iii) if there are no available buses from all other lines (e.g., L_2) to guarantee the departure timetable of B_{L11}, we keep bus B_{L11} on L_1 to guarantee L_1's timetable. E-buses prioritize serving the original lines over new lines when they can achieve the same performance by serving different lines, which is also for guaranteeing the timetables of all bus lines. For example, at the beginning of a day, all e-buses have enough energy to serve current lines and other lines so that they will keep their current lines. In this case, the timetable is always kept.

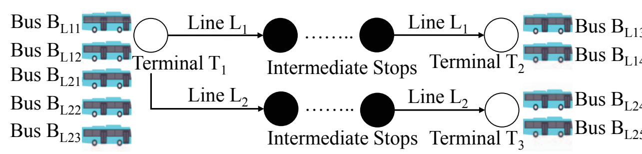

Fig. 18. Scheduling with timetable guarantee.

IV. Evaluation

In this section, we extensively evaluate the performance of our *bCharge* based on the massive e-bus data from Shenzhen by three different metrics, i.e., temporal distribution of charging events, the spatial distribution of charging events, and the most important electricity usage & charging cost.

A. Experimental Setup

Data Management: Due to the data-driven nature of our *bCharge*, we introduce how we manage our multi-source data related to e-buses as follows. In this project, we are working with Shenzhen Transportation Committee, and we utilize various data processing frameworks. The streaming data from Shenzhen e-buses require significant efforts for efficient management, querying, and processing. We employ a high-performance cluster with Spark for data processing. The details are given as follows: (i) 12 HP machines with 2 Tesla K80c each; (ii) 10 Dell machines with 4 Tesla K80c each; (iii) 4 Xeon E5-2650 with a half TB memory each; (iv) A series of 800GB SSD and 15TB of spinning-disk spaces; (v) 2 PB additional disk space. Due to the large size of our bus data, we performed a detailed cleaning process to filter out duplicate, error, and incomplete GPS/transaction data. More importantly, the key challenge in bus transaction data processing, compared to bus GPS data processing, is to protect the privacy of smartcard users and ensure the utility of the models at the same time. Due to space limitation, we briefly mention our privacy consideration in Sec. V.

Evaluation Data: In this evaluation part, as introduced in Sec. II, we utilize one-week GPS records generated by 16,359 e-buses from January 13th -19th 2018, in Chinese city Shenzhen. More than 69.6 million GPS records are generated by the e-bus fleet during this period. In addition to GPS data, the evaluation dataset also includes the static data of 376 e-bus charging stations, 1,400 bus lines, and 5,562 bus stops.

Parameter Setting: In our charging scheduling strategy *bCharge*, four parameters are needed to decide, as follows:

- **Four Different States**: For SOC^l, we decide it based on the real-world interaction with Shenzhen e-bus drivers. Based on our field studies in Shenzhen, drivers of e-buses are very conservative, and they normally stop serving passengers and go to find charging points if the SOC declines to 30%. As a result, we set the SOC^l to be 30%. Based on the expected real-time traffic and current line serving, we can calculate the energy consumption for serving the current line and other lines for each bus, which gives us the SOC$_c^f$ and SOC$_l^c$ since we have SOC^l.
- **Discount Factor** β: We empirically choose the discount factor β as 0.9 to guarantee the convergence of the algorithm similar to the previous work [35].
- **Immediate Reward** R_{sa}: For the immediate reward, the V$_i$ ($1 \leq i \leq 3$) is the expected revenue by serving different lines, which are calculated based on historical bus passenger demand; W$_j$ ($1 \leq j \leq 3$) is the expected charging cost, which is calculated based on the time of scheduling and the real-time SOC of e-buses.
- **Transition Probability** $P^a_{ss'}$: For the state transition probability, if an e-bus takes action A_S, it will stay in the same state, so the probability is set to 1. Since there is also no electricity consumption and operational profits, the reward is also set to 0 when taking action A_S. If there is only one possible transition, the probability is also 1 since the sum of the transition probabilities is 1, so p_{01} is 1 in Fig. 17. Other transition probabilities are calculated in real time based on policies given in Sec. III-B.

Baseline Setting: To show the effectiveness of our real-time charging scheduling for e-bus fleets, we compare the performance of our *bCharge* with Ground Truth and another real-time scheduling method called Earliest-deadline-first (EDF), which is a common method for electric vehicle charging scheduling [7], [8]. In EDF, we schedule the e-buses with the earliest timetable deadline to charge first. Further, we leverage three metrics to understand the performance of *bCharge*. They are (i) the temporal distribution of charging events, (ii) the spatial distribution of charging events, and (iii) most importantly, the electricity usage and charging cost.

B. Temporal Distribution of Charging Events

Fig. 19 and Fig. 20 show the distribution of the charging start time of *bCharge*, EDF and Ground Truth in different electricity rates durations, i.e., different time slots of a day. We found that there are more charging events happen in 23:00-7:00 under *bCharge*, resulting in less charging events during the daytime. Especially during 18:00-21:00, the gap between *bCharge* and the Ground Truth is more obvious. This is because *bCharge* schedules some e-buses that have been charged during 12:00-16:00 to serve other lines for some other e-buses, leaving these e-buses to charge after 23:00 for a lower rate. While during 16:00-18:00, more e-buses need to charge after the long-time operation, which results in the increase of charging events under EDF. Overall, *bCharge* has more charging events during the late night to early morning, i.e., off-peak hours, which potentially lead to lower charging cost.

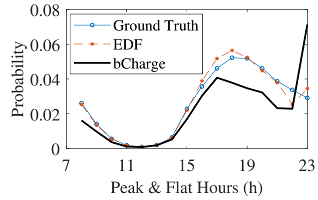

Fig. 19. Charging start time during the peak & flat hour.

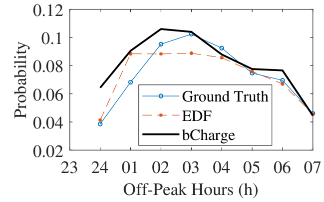

Fig. 20. Charging start time during the off-peak hour.

C. Spatial Distribution of Charging Events

We leverage the *Charging Station Utilization Rate* that we defined in Sec. II to describe the spatial distribution of e-bus charging. As shown in Fig. 21, there are fewer charging stations with too high or too low utilization rates with *bCharge* and baselines. The number of charging stations with the utilization rates between 3-6 accounts for 73% under *bCharge*, achieving 7% improvement compared to the Ground Truth. Since EDF does not change the charging locations of e-buses, the utilization rate of EDF is same as the Ground Truth. The balanced utilization of the charging infrastructure can effectively reduce the under-utilization and the overcrowded waiting phenomenon in the charging stations. The reason why *bCharge* balances the charges between stations is that the e-buses can be scheduled to serve for different lines under *bCharge*, which increases the flexibility of scheduling and potentially leads to better performance. For example, when an e-bus arrives at the terminal and it needs to charge, but no charging points are available at this charging station. Besides, its SOC is not enough for serving the current line, i.e., go back to the original terminal. In this case, *bCharge* schedules the e-bus to serve another shorter line with lower expected energy consumption, e.g., a shorter distance or a less congested route. Hence, *bCharge* can potentially improve the charging efficiency of two charging stations at the same time, i.e., reducing the utilization rate of one charging station and increasing the utilization rate of the other to balance their rates.

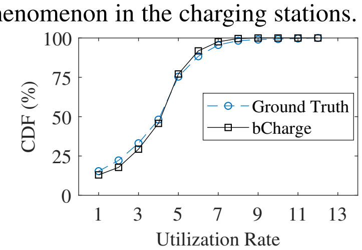

Fig. 21. Comparison of the utili. rate.

D. Electricity Usage & Charging Cost

Fig. 22 and Fig. 23 show the electricity usage for e-bus charging in different electricity rates durations. We found *bCharge* have more electricity usage during the midnight to early morning, especially during 23:00-23:59. This is because *bCharge* schedules some e-buses with enough energy to serve for other lines before this duration and then charge during this period, resulting in more electricity usage during this period. Besides, *bCharge* also reduces the electricity usage during daytime peak hours, e.g., 14:00-16:00. EDF increases the electricity usage during 16:00-18:00, due to the high charging demand in this duration. In total, our *bCharge* reduces 12.8%

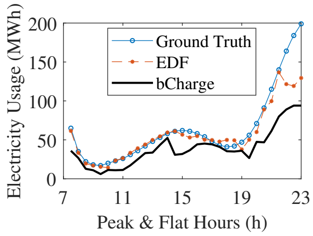

Fig. 22. Electricity usage during the peak & flat hour.

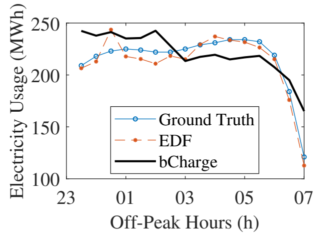

Fig. 23. Electricity usage during the off-peak hour.

(701 MWh) and 8.2% electricity energy for the e-bus fleet in Shenzhen per day compared with the Ground Truth and EDF.

Fig. 24 and Fig. 25 show the charging cost distribution. We found that the cost gap between *bCharge* and Ground Truth/EDF is more obvious than Fig. 19 during the daytime. This is because the electricity rates in the daytime are much higher than the price of off-peak hours at late night. Even though *bCharge* causes a slightly higher charging cost for the fleet during 23:00-2:00, it reduces 23.7% ($106,870) and 17.8% of the overall charging cost for one day compared with the Ground Truth and EDF, which indicates *bCharge* can potentially reduce 39 million dollars for the Shenzhen e-bus fleet per year based on the current Shenzhen e-bus budget.

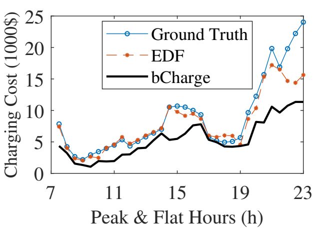

Fig. 24. Charging cost during the peak & flat hour.

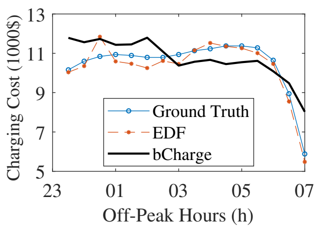

Fig. 25. Charging cost during the off-peak hour.

V. Lessons Learned

Based on our results, we have been conducting a few rounds of field studies to verify the patterns we found. In particular, we have been communicating with bus drivers, fleet managers, and charging station operators to fully understand the potential impacts of our study and the implication of our scheduling. We summarize a few lessons we learned from the project and field studies regarding the e-bus fleet in Shenzhen below.

Data Issues: The most unexpected lessons we learned is the data issues. In our field studies, we have been communicating with various interested parties for data collection, data quality, and data management. We summarize a few key insights as follows. (i) Since the Shenzhen e-bus fleet is operated by three companies, the data formats and access policies are very different. It takes us a long time to prepare the bus GPS data to understand the current operating patterns. (ii) Further, Shenzhen bus fare data are managed separately by a smartcard company, and the fare data regarding smartcards have personal information, e.g., cellphone number and addresses. The company staff at the smartcard company has removed all these personal information from the smartcard data in order to help us understand the bus passenger demand anonymously without privacy concerns. (iii) The detailed charging data at each charging point of all charging stations can provide more detailed insights into the charging behaviors for our scheduling. However, since the charging stations in Shenzhen are operated separately by several companies and they compete for each other, these charging data cannot be shared even after multiple negotiations. As a result, our charging activities are inferred by GPS data. (iv) Since Shenzhen just finished the process to replace all their regular buses with e-buses, the fleet management process is in transition. Lots of e-buses went to different charging stations based on some ad-hoc factors, e.g., charging station renovation and road work, which is hard to make sense and quantify with real-world data. We have reported many data issues to the fleet management team, and some of the problems they already knew but lots of issues are new to them. Based on our interactions with the fleet management team, these issues will be addressed soon.

More Buses vs. Effective Scheduling vs. More Charging Stations: Another lesson we learned from our analyses and field studies is that Shenzhen right now just finished the first stage of their e-bus plan, i.e., replacing all regular diesel buses with e-buses for a fully electric bus fleet. The second stage (yet more challenging and unclear) is the charging infrastructure upgrade, which is expected to take a much longer time due to various complicated issues, e.g., the land prices, the site survey, and security issues. For example, Shenzhen has one of the highest land prices in China and a charging station near the downtown area will cost more than 60 e-buses based on our interactions with the Shenzhen e-bus operators. As a result, for now, effective scheduling or even buying more buses is more practical than building new charging stations.

Additional Drivers Exclusively for Charging: The final unexpected lesson is the labor-intensive charging operation in the night. Fig. 26 shows a detailed charging station setting in our field studies and their status at two different times of a day. We can see that all charging points are occupied by e-buses at 23:05; whereas only one e-bus was charging at 17:06. This is because the electricity rate between 23:00-7:00 is much lower than the rate at 17:00. However, given limited charging points

（a）Charging station at 23:05 （b）Charging station at 17:06

Fig. 26. Charging station status in different times.

due to high costs, e.g., 80,000 US dollars for a charging point deployment, e-buses need to be moved around before or after charging, but the regular bus drivers will be off-duty after

23:00 and before 7:00. Hence, the Shenzhen e-bus network hires additional drivers just for moving buses during late night and early morning charging. In particular, one of the three e-buses operating companies in Shenzhen has 750 additional drivers just for moving e-buses before or after charging. By considering this labor cost factor, it may be more reasonable to charge more buses during the daytime since regular drivers can move the buses without cost for additional drivers. However, based on our interactions with the fleet management team, hiring additional drivers is a short-term and temporary issue since the charging points will be upgraded soon, e.g., longer and secure cables, to address this issue.

Implementation in Different Cities: The bus networks in different cities typically have different operating patterns due to geographic and demographic features, so it is extremely significant to implement $bCharge$ in different cities. Currently, we are in the process of obtaining bus data from other cities for dual-city modeling. But since only Shenzhen has a fully electric bus network, it is hard to find such a large fleet for a parallel study currently. One possible direction we are exploring is to design transfer learning models to transfer the knowledge (e.g., operating pattern, charging pattern, etc) from the Shenzhen e-bus network to bus networks in other cities for a "what if" investigation. For example, what if all regular buses in Beijing or New York City were replaced by e-buses, how much will it cost and how to schedule e-buses for these cities. It opens some very interesting research directions.

VI. Related Work

There are two different charging management modes for electric vehicles, i.e., centralized and decentralized ones [7], [36]. For the existing research, some works are based on large-scale real-world data, while others leverage a small dataset or experiment-based simulations. Based on these two factors, we divide the electric vehicle research into four different categories, which can be seen from Tab. II.

TABLE II
Categories of Related Work

Scheduling	Small-Scale	City-Scale ($>$ 500 vehicles)
Decentralized	[15], [16], [37]–[40]	[8]–[11], [14], [17]
Centralized	[7], [36], [41]–[45]	*bCharge*

A. Decentralized Scheduling

Small-scale Scheduling: The decentralized charging scheduling of electric vehicles has been widely studied by many researchers, but most of them are based on the small-scale or experiment-based simulations. [37] develops a reservation recommendation algorithm for electric vehicles considering the shortest distance and shortest waiting time. [16] presents a distributed power schedule framework based on Game Theory to obtain the optimal schedule for online electric vehicles. These works are based on small-scale data or theoretical models, which are hard to capture the dynamics of real-world large-scale electric vehicles operating and charging patterns.

City-scale Scheduling: [9] designs a recommendation system for e-taxis to reduce the total charging time cost for each driver. [14] develops a charging station deployment and charging point placement framework to minimize the overall charging time of e-taxis. PickaChu [11] provides a charger deployment scheme that maximizes the probability of picking up passengers for e-taxis and minimizes the deployment cost. They aim to schedule e-taxis or find the optimal locations to deploy charging infrastructures for e-taxis for reducing the operating cost or deployment cost based on real-world data. But they are based on a distributed charging nature; whereas in this work, we consider a centralized scheduling model.

B. Centralized Scheduling

Small-scale Scheduling: There are also some existing works for the centralized scheduling of electric vehicles based on small-scale data or experimental simulations. [43] introduces and analyzes the electric transit bus system with wireless based on one e-bus line in a research institute campus. [36] proposes a population-based heuristic approach to minimize the total charging cost, which is executed on a 20-bus test system. [42] proposes an effective charging rate control algorithm to optimize the social welfare of electric vehicles. However, these small systems cannot fully reveal the complexity and advantage of centralized scheduling for city-scale systems.

City-scale Scheduling: Different from the existing work, our work addresses a practical charging issue for e-bus fleets by centralized scheduling in a setting of city-scale systems. To our best knowledge, *bcharge* is the first work of city-scale data-driven investigation on studying the real-time scheduling for e-bus fleets. Such a data-driven investigation enables us to identify the real-world e-bus operating and charging issues, which are challenging to reveal using simulation studies, small-scale data or under a decentralized setting.

VII. Conclusion

In this paper, we conduct, to the best of our knowledge, the first study called *bCharge* for real-time charging scheduling of e-bus fleets based on a real-world dataset in Shenzhen, which includes the data from 16,359 e-buses, 1,400 bus lines, and 376 charging stations. In *bCharge*, we consider various real-world factors based on the long-term data including e-bus daily and per-charge operating distances, charging spatial-temporal distribution, charging station utilization rates and charging cost distribution, etc. More importantly, we have shown that with its effective scheduling, *bCharge* outperforms the ground truth by 23.7% and outperforms the baseline method by 17.8% regarding the total charging cost. For the immediate benefit, *bCharge* can reduce the operating cost for the Shenzhen e-bus network with its data-driven real-time scheduling. For the long-term benefit, our results in *bCharge* may be used for other cities to promote e-buses for green public transportation.

Acknowledgment

This work is partially supported by Rutgers Global Center, Rutgers Research Council, China 973 Program (2015CB352400), National Natural Science Foundation of China (NSFC) 41401470, 61772026 and 61332018, and National Key R&D Program of China 2018YFB0803400.

References

[1] B. N. E. Finance, "Electric buses in cities: Driving towards cleaner air and lower co_2," http://www.c40.org/c40_research, pp. 1–63, 2018.

[2] V. I. Roland Irle, Jos Pontes, "The electric vehicle world sales database," http://www.ev-volumes.com/, 2018.

[3] J. Zhang, D. Shen, L. Tu, F. Zhang, C. Xu, Y. Wang, C. Tian, X. Li, B. Huang, and Z. Li, "A real-time passenger flow estimation and prediction method for urban bus transit systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 3168–3178, 2017.

[4] J. Zhang, X. Yu, C. Tian, F. Zhang, L. Tu, and C. Xu, "Analyzing passenger density for public bus: Inference of crowdedness and evaluation of scheduling choices," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, 2014, pp. 2015–2022.

[5] D. Zhang, J. Huang, Y. Li, F. Zhang, C. Xu, and T. He, "Exploring human mobility with multi-source data at extremely large metropolitan scales," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 201–212.

[6] D. Zhang, J. Zhao, F. Zhang, and T. He, "comobile: Real-time human mobility modeling at urban scale using multi-view learning," in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2015, p. 40.

[7] F. Kong, Q. Xiang, L. Kong, and X. Liu, "On-line event-driven scheduling for electric vehicle charging via park-and-charge," in *Real-Time Systems Symposium (RTSS), 2016 IEEE*. IEEE, 2016, pp. 69–78.

[8] Z. Dong, C. Liu, Y. Li, J. Bao, Y. Gu, and T. He, "Rec: Predictable charging scheduling for electric taxi fleets," in *Real-Time Systems Symposium (RTSS), 2017 IEEE*. IEEE, 2017, pp. 287–296.

[9] Z. Tian, T. Jung, Y. Wang, F. Zhang, L. Tu, C. Xu, C. Tian, and X. Y. Li, "Real-time charging station recommendation system for electric-vehicle taxis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3098–3109, 2016.

[10] L. Yan, H. Shen, J. Zhao, C. Xu, F. Luo, and C. Qiu, "Catcharger: Deploying wireless charging lanes in a metropolitan road network through categorization and clustering of vehicle traffic," in *IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2017, pp. 1–9.

[11] L. Yan, H. Shen, Z. Li, A. Sarker, J. A. Stankovic, C. Qiu, J. Zhao, and C. Xu, "Employing opportunistic charging for electric taxicabs to reduce idle time," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (Ubicomp)*, vol. 2, no. 1, p. 47, 2018.

[12] People.cn, "Shenzhen is connected by electric buses," http://society.people.com.cn/n1/2018/0508/c1008-29970131.html, 2018.

[13] Sohu, "A survey of shenzhen electric buses," https://m.sohu.com/n/465555062/, 2016.

[14] Y. Li, J. Luo, C.-Y. Chow, K.-L. Chan, Y. Ding, and F. Zhang, "Growing the charging station network for electric vehicles with trajectory data analytics," in *Data Engineering (ICDE), 2015 IEEE 31st International Conference on*. IEEE, 2015, pp. 1376–1387.

[15] A. Sarker, H. Shen, and J. A. Stankovic, "Morp: Data-driven multi-objective route planning and optimization for electric vehicles," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (Ubicomp)*, vol. 1, no. 4, p. 162, 2018.

[16] A. Sarker, Z. Li, W. Kolodzey, and H. Shen, "Opportunistic energy sharing between power grid and electric vehicles: A game theory-based pricing policy," in *Distributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on*. IEEE, 2017, pp. 1197–1207.

[17] C. Liu, K. Deng, C. Li, J. Li, Y. Li, and J. Luo, "The optimal distribution of electric-vehicle chargers across a city," in *Data Mining (ICDM), 2016 IEEE 16th International Conference on*. IEEE, 2016, pp. 261–270.

[18] K. Valogianni, W. Ketter, J. Collins, and D. Zhdanov, "Effective management of electric vehicle storage using smart charging." in *AAAI*, 2014, pp. 472–478.

[19] Y. Xiong, J. Gan, B. An, C. Miao, and A. L. Bazzan, "Optimal electric vehicle charging station placement." in *IJCAI*, 2015, pp. 2662–2668.

[20] Y. Xu, S. Çolak, E. C. Kara, S. J. Moura, and M. C. González, "Planning for electric vehicle needs by coupling charging profiles with urban mobility," 2018.

[21] B. Du, Y. Tong, Z. Zhou, Q. Tao, and W. Zhou, "Demand-aware charger planning for electric vehicle sharing," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018, pp. 1330–1338.

[22] X. Wang, C. Yuen, N. U. Hassan, N. An, and W. Wu, "Electric vehicle charging station placement for urban public bus systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 1, pp. 128–139, 2017.

[23] C.-M. Tseng, S. C.-K. Chau, and X. Liu, "Improving viability of electric taxis by taxi service strategy optimization: A big data study of new york city," *IEEE Transactions on Intelligent Transportation Systems*, no. 99, pp. 1–13, 2018.

[24] P. D. Online, "Shenzhen becomes world's first city with all-electric public transportation," http://en.people.cn/n3/2017/1228/c90000-9309683.html, 2017.

[25] B. Y. Dreams, "Whole vehicle technologies and battery technology of byd k9," http://en.byd.com/la/auto/ebus.html, 2018.

[26] B. Spiller, "All electricity is not priced equally: Time-variant pricing," http://blogs.edf.org/energyexchange/2015/01/27/all-electricity-is-not-priced-equally-time-variant-pricing-101/, 2015.

[27] J. Yuan, Y. Zheng, X. Xie, and G. Sun, "Driving with knowledge from the physical world," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 316–324.

[28] Y. Zheng, Y. Chen, Q. Li, X. Xie, and W.-Y. Ma, "Understanding transportation modes based on gps data for web applications," *ACM Transactions on the Web (TWEB)*, vol. 4, no. 1, p. 1, 2010.

[29] G. Wang, T. Xu, T. Tang, T. Yuan, and H. Wang, "A bayesian network model for prediction of weather-related failures in railway turnout systems," *Expert systems with applications*, vol. 69, pp. 247–256, 2017.

[30] Y. Li, C. Zhan, M. de Jong, and Z. Lukszo, "Business innovation and government regulation for the promotion of electric vehicle use: lessons from shenzhen, china," *Journal of Cleaner Production*, vol. 134, pp. 371–383, 2016.

[31] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[32] X. Xie, F. Zhang, and D. Zhang, "Privatehunt: Multi-source data-driven dispatching in for-hire vehicle systems," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, p. 45, 2018.

[33] R. Bellman, *Dynamic programming*. Courier Corporation, 2013.

[34] E. A. Feinberg and A. Shwartz, *Handbook of Markov decision processes: methods and applications*. Springer Science & Business Media, 2012.

[35] E. B. Iversen, J. M. Morales, and H. Madsen, "Optimal charging of an electric vehicle using a markov decision process," *Applied Energy*, vol. 123, pp. 1–12, 2014.

[36] Q. Kang, J. Wang, M. Zhou, and A. C. Ammari, "Centralized charging strategy and scheduling algorithm for electric vehicles under a battery swapping scenario," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 3, pp. 659–669, 2016.

[37] C. J. Park, J. Lee, G. L. Park, and J. S. Hyun, "Development of reservation recommendation algorithms for charging electric vehicles in smart-grid cities," *International Journal of Smart Home*, vol. 8, no. 1, pp. 113–122, 2014.

[38] Y. Mou, H. Xing, Z. Lin, and M. Fu, "Decentralized optimal demand-side management for phev charging in a smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 726–736, 2015.

[39] E. Kim, K. G. Shin, and J. Lee, "Real-time discharge/charge rate management for hybrid energy storage in electric vehicles," in *Real-Time Systems Symposium (RTSS), 2014 IEEE*. IEEE, 2014, pp. 228–237.

[40] R. Gopalakrishnan, A. Biswas, A. Lightwala, S. Vasudevan, P. Dutta, and A. Tripathi, "Demand prediction and placement optimization for electric vehicle charging stations," *international joint conference on artificial intelligence (IJCAI)*, pp. 3117–3123, 2016.

[41] F. Kong and X. Liu, "Distributed deadline and renewable aware electric vehicle demand response in the smart grid," in *Real-Time Systems Symposium (RTSS), 2015 IEEE*. IEEE, 2015, pp. 23–32.

[42] F. Kong, X. Liu, Z. Sun, and Q. Wang, "Smart rate control and demand balancing for electric vehicle charging," in *International Conference on Cyber-Physical Systems (ICCPS)*, 2016, p. 4.

[43] Y. J. Jang, E. S. Suh, and J. W. Kim, "System architecture and mathematical models of electric transit bus system utilizing wireless power transfer technology," *IEEE Systems Journal*, vol. 10, no. 2, pp. 495–506, 2016.

[44] E. Kim, J. Lee, and K. G. Shin, "Modeling and real-time scheduling of large-scale batteries for maximizing performance," in *Real-Time Systems Symposium, 2015 IEEE (RTSS)*. IEEE, 2015, pp. 33–42.

[45] T. Paul and H. Yamada, "Operation and charging scheduling of electric buses in a city bus route network," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pp. 2780–2786.