Real-Time Charging Station Recommendation System for Electric-Vehicle Taxis

Zhiyong Tian, Taeho Jung, *Student Member, IEEE*, Yi Wang, Fan Zhang, Lai Tu, Chengzhong Xu, *Fellow, IEEE*, Chen Tian, and Xiang-Yang Li, *Fellow, IEEE*

***Abstract*—Electric vehicle (EV) taxis have been introduced into the public transportation systems to increase EV market penetration. Different from regular taxis that can refuel in minutes, EV taxis' recharging cycles can be as long as one hour. Due to the long cycle, the bad decision on the charging station, i.e., choosing one without empty charging piles, may lead to a long waiting time of more than an hour in the worst case. Therefore, choosing the right charging station is very important to reduce the overall waiting time. Considering that the waiting time can be a nonnegligible portion to the total work hours, the decision will naturally affect the revenue of individual EV taxis. The current practice of a taxi driver is to choose a station heuristically without a global knowledge. However, the heuristical choice can be a bad one that leads to more waiting time. Such cases can be easily observed in current collected taxi data in Shenzhen, China. Our analysis shows that there exists a large room for improvement in the extra waiting time as large as 30 min/driver. In this paper, we provide a real-time charging station recommendation system for EV taxis via large-scale GPS data mining. By combining each EV taxi's historical recharging events and real-time GPS trajectories, the current operational state of each taxi is predicted. Based on this information, for an EV taxi requesting a recommendation, we can recommend a charging station that leads to the minimal total time before its recharging starts. Extensive experiments verified that our predicted time is relatively accurate and can reduce the cost time of EV taxis by 50% in Shenzhen.**

***Index Terms*—Electric vehicle (EV), charging station, recommendation, taxis.**

I. INTRODUCTION

UNLIKE traditional Internal Combustion Engine Vehicles (ICEV), Electric Vehicles (EV) have less air pollution and are more environment friendly, and due to their contribution to carbon dioxide reduction, EVs are becoming increasingly popular nowadays. With the support from governments, many countries have already partially adopted EVs in their public transit systems [1], [2]. For example, many cities in Unites States are in the transition period by gradually replacing ICEVs to hybrid vehicles and finally to EVs [3]. Shenzhen, a city in China, is one of the cities who have achieved the most successful EV public transit system. In our previous research, we found that its EV taxi fleet has passed the business break-even point since 2013 [4]. Note that business success is critical to the sustainable development of public EVs.

Though successful it is in Shenzhen, we identified a large room for improvement in EV taxis' operational efficiency. EV Taxis need to recharge their batteries during the work hours since the battery capacity is not large enough. Considering the EV taxis' long daily travel distances, it is hardly possible that EV taxis will only need to recharge once outside the work hours in the near future no matter how battery-efficient EV taxis become. Then, different from ICEVs, the inherently long recharging cycles of EV taxis can become a bottleneck in the EV taxis' business if the idle hours (without working) in the stations become significantly long when compared to the total hours of operation per day. As a result, choosing a good charging station with short cost time will improve the efficiency of EV taxis business by increasing the duty ratio defined as the ratio of true working hours to total hours of operation.

Currently, without live information of charging stations, taxi drivers do not know the expected cost time at a station before they arrive at it. Even such information could be reported to the drivers, the information is time-dependent in the sense that the information will become stale every time a new driver arrives, and the number of idle charging piles in a station at the time a driver chooses the station will be different from the number of idle piles at the time the driver arrives at the station. We depict an example of such a case in Fig. 1. When a driver of the EV taxi v is seeking for a charging station, he may select the closer station S_1 instead of S_2 since both have empty piles. Then, the driver v may spend more time in waiting since v_1 arrives earlier

Manuscript received April 5, 2015; revised August 30, 2015, November 28, 2015, and January 30, 2016; accepted February 29, 2016. Date of publication April 14, 2016; date of current version October 28, 2016. This work was supported in part by the China National Basic Research Program (973 Program) under Grant 2015CB352400; by the National High Technology Research and Development Program of China (863 Program) under Grant 2014AA01A702; by the National Natural Science Foundation of China under Grant U1401258; by the U.S. National Science Foundation under Grants CMMI 1436786 and CNS 1526638; by the Research Program of Shenzhen under Grants JSGG20150512145714248, KQCX2015040111035011, and CYZZ20150403111012661; by the Natural Science Foundation of Hubei Province under Grant 2014CFB1007; and by the Fundamental Research Funds for the Central Universities. The Associate Editor for this paper was B. Chen. *(Corresponding author: Fan Zhang.)*

Z. Tian, Y. Wang, and L. Tu are with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: tianzhiyong@hust.edu.cn; ywang@hust.edu.cn; tulai.net@gmail.com).

T. Jung is with the Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: tjung@hawk.iit.edu).

F. Zhang is with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: zhangfan@siat.ac.cn).

C. Xu is with the Wayne State University and Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: cz.xu@siat.ac.cn).

C. Tian is with the State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China (e-mail: tianchen@nju.edu.cn).

X.-Y. Li is with the School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China (e-mail: xiangyang.li@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2016.2539201

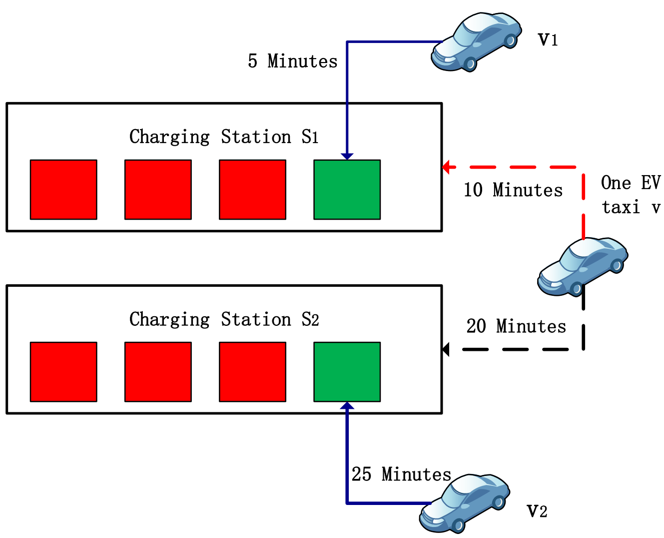


Fig. 1. A toy example of wrong station selection.

and uses the last available pile. On the other hand, if the driver had chosen S_2 instead, although driving to S_2 would lead to more travel time, he would reach S_2 earlier than v_2 and thus having zero extra cost time at the station.Similar situations are commonly observed in our real EV taxi data set.

In conclusion, the lack of global knowledge including all drivers' station selection may lead to EV taxis' longer cost time at the charging stations. Therefore, how to give *customized* recommendation for each EV taxi driver based on others' possible choices becomes extraordinarily challenging in the cases similar to that in Fig. 1. The problem is also crucial in improving the efficiency of EV taxis business. Notice that the optimization of cost time of a given individual EV taxi is our main concern.

In this paper, we provide a real-time charging station recommendation system for EV taxis by mining large-scale GPS data and taxis operation data. Unlike the recommendation system for ICEV taxis [5], due to the longer recharging cycle of EVs, EV taxis' cost time in different stations varies more than those of ICEVs in different refueling stations. Thus such a system for EVs is even more challenging due to its less tolerance to an improper recommendation.

To our best knowledge, this is the first recommendation system for EV taxis in a large city. The contributions of this paper are summarized as follows:

1) By exploiting data from multiple sources (e.g., taxi gps, taxi deal and charging station information), we successfully understand EV taxi drivers' recharging behavior patterns.
2) Based on the analysis of EV taxi drivers' recharging behavior patterns in Shenzhen, we managed to identify their recharging intentions, i.e., whether and how likely they will recharge at different stations at certain time given their locations.
3) By combining EV taxi drivers' recharging intentions and the conditions of charging stations, we propose a real-time recommendation system for EV taxi drivers in order to minimize the extra cost time in the station.

The rest of this paper is organized as follows. After a brief review of related work in Section II, we introduce how to preprocess the data and observe the opportunity of a recommendation system through data analysis in Section III. Section IV proposes the solution to the real-time recommendation system, including how to capture EV taxi drivers' recharging intentions and the model of the system. In Section V, we evaluate the experiment results and conclude the paper in Section VI.

II. Related Work

Recently, both industrial and academic communities started to have great interest to EVs and charging station deployment. The mainly studied issues are siting of charging stations and scheduling of EVs. For example, Kuby *et al.* [6] propose the Flow Refueling Location Model(FRLM) for alternative-fuel vehicles and Kim *et al.* [7], Capar *et al.* [8] extend raw FRLM by adding new features to solve the siting problem of charging stations by using exploiting graph theory. Jung *et al.* [9] use an activity-based model to analyze the queue delay of charge stations and offer decision support for choosing locations of undeployed charge stations. Besides, Gharbaoui *et al.* [10] also use activity-based models and find that in urban areas, public charging stations can be under-utilized and location selecting of charging stations should be considered to reduce EV owners' range anxiety. In [11], the FRLM model has been extended to solve the problem of allocation of fast charging stations in a case study of German autobahn.

Different from the above siting problems, our recommendation system aims at reducing the extra waiting time of EV taxi drivers at charging station based on the given charging stations with given capacity (i.e., number of charging piles) and distribution (e.g., locations). In this sense, our work is what should be pursued after charging stations are constructed based on the solutions to the siting problems since each individual driver's charging experience not only depends on the capacity and distribution of the charging stations but also depends on other drivers' charging behaviors.

Besides recommendation strategies, scheduling strategies can also reduce EVs' cost, including the time and money spent at charging stations. These studies are usually based on existing deployment of charging stations and use the theory of load balance. For example, Ma *et al.* [12] and Gan *et al.* [13], [14] use different decentralized control strategies to reduce the price of recharging by avoiding the peak electricity-using hours. Sundstrom *et al.* [15] propose a novel method to schedule EVs for easing the overloading problem in the power grid. Kim *et al.* [16] introduce a reservation-based scheduling system to respond to multiple recharging requests, which aims to reduce EV drivers' waiting time at charging stations. Qin *et al.* [17] and Lu *et al.* [18] propose different dispatching strategies to reduce EV drivers' waiting time. In [19], the behavior of charging stations and EVs are investigated based on the game theory, aiming to find a balanced price between EV users and power supply merchants.

Different from the above scheduling strategies, our recommendation system provides suggestions for EV taxi drivers, allowing them to make their own choices. In addition, our recommendation system takes the real-time charging conditions including charging stations and EV taxis into account, instead of only considering historical charging events data of EV taxis, thus improving the accuracy of our recommendation results. Besides, note that our recommendation system aims at reducing the time of a part of EV taxi drivers at charging stations instead of balancing the waiting time of all the EV taxi drivers, which is usually a goal of a scheduling system by utilizing load balance theory.

Electricity price and distribution grid may also have impact on the individual users. For example, Yang *et al.* [20] investigates the profit maximization problem of a plug-in electric taxi by identifying appropriate charging time slots to reduce expense on electricity prices. Yang *et al.* [21] concentrates on the optimal scheduling of EV by taking operation income, regulation revenue and the driver's charging habit into account to minimize users' cost on charging. In [22], a new charging scheduling mechanism named V2V has been proposed to take full advantage of every EV's battery energy and help relieve burden on distribution grid.

There is also other work dedicated to the analysis of transportation systems based on mining large scale data, such as taxis [23]–[25], public buses [26], [27], freight trunks [28] and metro [29], [30]. Our previous work [4] has studied the operational patterns and recharging behaviors of EV taxi drivers. However, none of these analysis aims to develop a real time recommendation system for EV taxis.

III. Data Preprocessing and Behavior Analysis

In this section, in order to capture EV taxi drivers' recharging intentions, two data preprocessing tasks need to be done.

Individual EV Taxi Recharging Events Detection: Most EV taxis in Shenzhen recharge at least three times per day due to the limitation of battery capacity for operation. As there is no direct data source indicating an EV taxi's recharging state, it is necessary to detect when and where the taxi recharges. By combining an EV taxi's trace data and charging stations' coordinates, individual EV taxi's recharging events in different time slots can be parsed.

Individual EV Taxi Recharging Features Analysis: In order to predict an EV taxi driver's recharging intention, we exploit its historical recharging events data to analyze its recharging features. Through interviews on EV taxi drivers and field investigations at charging stations, we acquire supportive evidence for our empirical findings, which will also be used to predict the recharging intentions in the recommendation system.

In the following sections, we first give an overview of EV taxi deployment in Shenzhen, introduce the data set used in this study and then present how to deal with the above two problems.

A. EV Taxi Specifications & Charging Station Deployment

The study is performed on the EV taxi deployment in Shenzhen, where the number of EV taxis was approximately 850 in 2014 and has been growing till now. The specifications of the EV taxis in Shenzhen (Table I) indicate that the distance

TABLE I
Specifications of EV Taxis in Shenzhen

Model	Dist./Charge	Max. Speed	Capacity	Charging Cycle
BYD e6	300Km	160Km/h	72KWh	1h

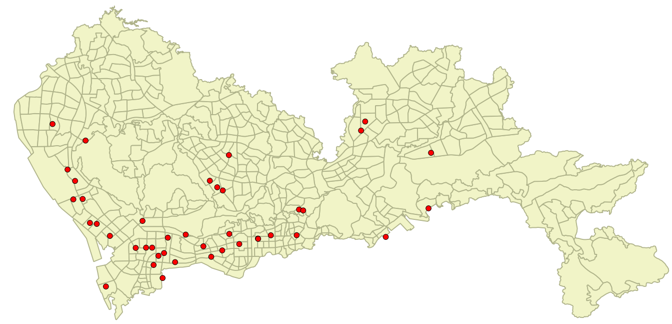


Fig. 2. Distribution of charge stations in Shenzhen (one dot per station).

which a fully charged EV can travel is shorter than that of a fully fueled ICEV (according to the fuel economy and the tank capacity specifications of the ICVE taxis in Shenzhen, a fully fueled ICVE can travel approximately 600 km). Meanwhile, the distribution of deployed charging stations for public EVs is shown in Fig. 2. There are two kinds of charging stations deployed in Shenzhen. A majority of the stations are exclusively for EV taxis while a few are shared by EV taxis and electric buses. Neither type of stations are open for private cars. Since electric buses usually have fixed schedule and recharge at a certain fixed time at late night, the available charging stations and piles for EV taxis can be considered as static resource. They will be negligibly affected by other type of electric vehicles.

Although the charging stations shown in Fig. 2 are mostly only for EV taxis, they are still in severe shortage to supply all EV taxis. Both above battery specifications and the current charging station deployment can lead to EV drivers' long time cost for recharging and can be obstacles in promoting the usage of EV. Infrastructure constructions and battery technology improvement can be an option. However, the initial cost of additional infrastructure is considerably large and the construction may also be limited by some constrains such as land use and power grid. The battery technology also needs further breakthrough to be widely applied in EVs. Therefore, the status quo of the EV specifications and charging station deployment raise the emergency of an efficient charging recommendation solution.

B. Data set Description

We analyze the GPS trace data set of over 800 EV taxis in Shenzhen during the period 07/01/2014–07/31/2014, which includes the GPS records collected by EV taxis every 30 seconds (Table II) and the taxi transaction records during the same period (Table III).

Besides the taxi data, we also collect information of charging stations as shown in Table IV. The field *PileNumber* indicates the total number of charging piles in every charging station. The information of charging stations is considered to be static in our analysis.

TABLE II
TAXI GPS DATA DESCRIPTION

CarId	Unique ID of a taxicab
CarType	Flag indicating EV or ICEV
Company	Company which a taxi belongs to
Location	Longitude and latitude of the location
Time	Second-level timestamp (*e.g.,* 2014-07-01 04:52:15.000)
Occupied	Flag indicating meter on/off

TABLE III
TAXI TRANSACTION DATA DESCRIPTION

CarId	Unique ID of a taxicab
On/off time	Time when the taxis pick up/drop off passengers
On/off location	Location of a picking up/droping off
Cruise distance	Non operating cruise distance before picking up
Distance	Operating distance from a picking up to its dropping off
Fare	Fare of a metered trip

TABLE IV
CHARGING STATION DATA DESCRIPTION

StationId	Unique ID of a charging station
Name	Name of a charging station
Location	Longitude and latitude of a charging station location
PileNumber	The number of charging piles at a charging station

C. Individual EV Taxi Recharge Events Detection

As the recharging cycle in Table I, we can assume that an EV taxi that is recharging or waiting for recharging at any charging station should stay almost still and close to the station for a long period of time. This should also be reflected in the GPS trace data of every EV. Therefore, based on such assumption, we detect an EV taxi's recharging events by fusing taxi GPS and charging station data in a similar way described in [5].

We performed three individual field investigations at different charging stations and time on July 20, 2014, Nov. 9, 2015 and Nov. 11, 2015 respectively. Totally 92 EV taxis were involved in the field investigations. We recorded their behaviors near the charging station and get the facts of whether they were charging as the ground truth. After that, we compared the recharge events detected by aforementioned method with the ground truth by varying r and θ, and out of 92 EV taxis from the ground truth, 88 EV taxis' recharge events are correctly detected, yielding the recall of 95.7%.

We further investigate on the 4 misjudge samples. We found that there exists cases of EV taxis parking at positions nearby a charging station not for recharging but identified as a recharging event by our detection method. The GPS records of them follow a similar pattern like normal charging events, which is difficult to distinguish. So at second and third time of the field investigation, we talked with those EV taxi drivers to learn about their recharging behavior patterns during the field investigations. We found that the misjudge samples come from the taxis whose drivers happen to live near the charging station. These portions of drivers are very limited and can be found out by long time observation with a few more field investigations. Therefore, if we exclude such exceptional samples, the accuracy of the charging event detection can be even high and meet the requirement for recommendation.

TABLE V
EV TAXI RECHARGE EVENT DATA DESCRIPTION

CarId	Unique ID of an EV taxi
StationID	Unique ID of a charging station
ReachTime	Time of an EV taxi reaches a charging station
BeginTime	Time of a recharge event begins
EndTime	Time of a recharge event ends
DropTime	Time of last dropping off a passenger before a recharge event

D. Behavior Analysis and Opportunities for a Recommendation System

Utilizing the aforementioned method, we obtain individual EV taxi recharging events data as in Table V. Through the analysis of those recharging events, we can observe the room for improvement by reducing EVs' extra cost time at charging stations.

The fields *CarID* and *StationID* come from Tables II and IV, respectively, indicating which charging station an EV taxi chooses to recharge. Then,$(BeginTime - ReachTime)$ is the extra cost time at the station where BeginTime refers to the time the vehicle starts the recharging and the ReachTime refers to the time when the vehicle arrives at the station. Besides, in conjunction with the taxi transaction data, we identify the $DropTime$ which refers to the time the taxi drops off the last customer, and then the value of $(ReachTime - DropTime)$ refers to the travel time an EV taxi travels to a charging station. It's easily understood that at $DropTime$, EV taxi drivers intends to recharge their taxis and we should be able to identify this kind of recharging intentions for our later real-time recommendation system.

Most EV taxis in Shenzhen have two drivers operating the same vehicle in different time periods, i.e., one driver works in day time and the other works in night time. Through interviews with EV taxi drivers in Shenzhen, we are informed that they usually recharge within 4 time slots: 03:00–05:00, 11:00–13:00, 16:00–18:00 and 20:00–22:00 because of their operation cycle. In general, the shift handover usually takes place within 03:00–05:00, and another shift handover takes place within 16:00–18:00. From the statistical results of recharging event data set, we can obtain the temporal distribution of collected EV taxis' recharging events in Shenzhen. The results shown in Fig. 3(a) verify aforementioned drivers' answers in the interview. The proportion of recharging event frequency in different time slots, which is depicted in blue line, shows four prominent peak periods for recharging which is consistent with our interviews. Besides, taxi pick-up events are also gathered from taxi transaction data including both EV and ICEV taxis from the field *On/off time*. The dotted red line in Fig. 3(a) illustrates the proportion of passenger riding requests (i.e., taxi pick-up events) in different time slots, indicating that EV taxi drivers always choose time of less passenger riding requests to recharge their taxis to minimize the lost during the recharging hours.

Fig. 3(a) describes the temporal distribution of collected EV taxis' recharging events. Besides, we concern more concerned about the temporal feature of individual EV taxi recharging events. Based on the recharging event data set, we can analyze every individual EV taxi's recharging events by grouping the data records by the field of *CarID*. We observe that most EV

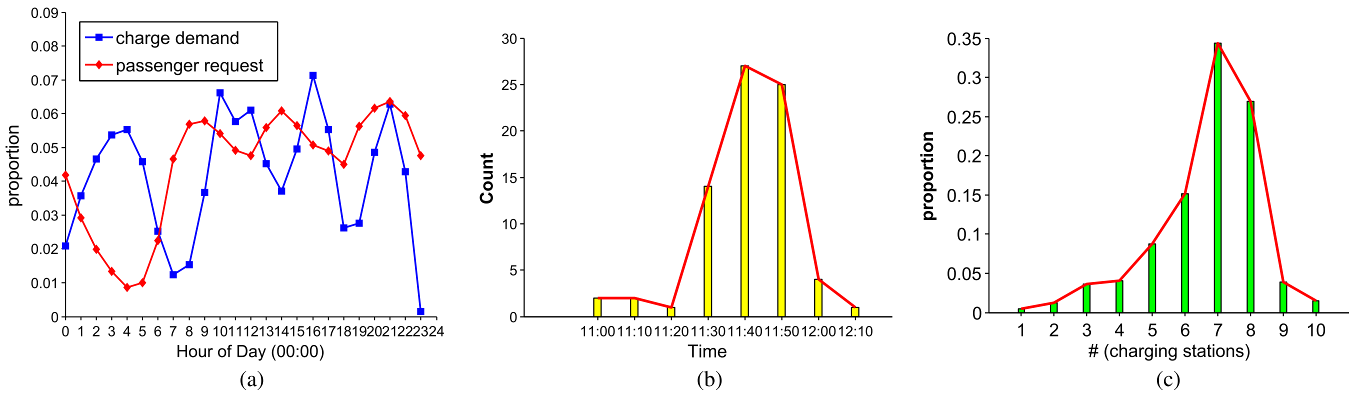


Fig. 3. Temporal and spatial statistics on recharging station usage. (a) Temporal distribution of collected recharging events; (b) recharge event distribution of individual EV taxi at noon; (c) visited charging station distribution of individual EV taxi.

taxi drivers are likely to recharge at some fixed time slots and we randomly select one EV taxi to illustrate this observation. Fig. 3(b) depicts an EV taxi's recharging events distribution at noon during the period 07/01/2014–07/31/2014, where one bar indicates the frequency that this EV taxi intends to recharge at a certain time throughout the whole month. It is obvious that the period when this EV taxi driver intends to recharge concentrates around 11:40. This observation is very common among EV taxi drivers in the statistical results in Shenzhen, which is beneficial for us to capture their recharging intentions.

From the recharging event data set, spatial feature of EV taxi recharging events can also be revealed, which we denote as EV taxi drivers' *adherence* of station selection, just as shown in Fig. 3(c): although over 50 charging stations are deployed in Shenzhen, most of EV taxi drivers choose 6–8 stations among them regularly. As for an individual EV taxi, we obtain a list of charging stations it choose to recharge and rank charging stations by the frequency this taxi has visited.

This spatio-temporal analysis in Fig. 3 proves that EV taxis' long time cost at charging stations results from the fact that EV taxi drivers always rush to the same charging stations during the same period. This is partly because they do not have access to the real-time loads of the charging stations (i.e., the number of vehicles using the stations), and it also results from the lack of accurate recommendation system.

In conclusion, the preliminary observations and analysis of recharging behavior patterns suggest the necessity of the charging station recommendation system. Therefore, we propose a real-time recommendation system for EV taxi drivers. Note that although our observations are achieved from the data set collected in Shenzhen, none of them are dependent on special characteristics of Shenzhen, and all of the spatio-temporal patterns are likely to hold in any EV taxi system under other circumstances.

IV. Real-Time Recommendation System

A. System Overview

In this section, we propose the solution to the real-time recommendation system. The framework of the real-time recommendation system is shown in Fig. 4.

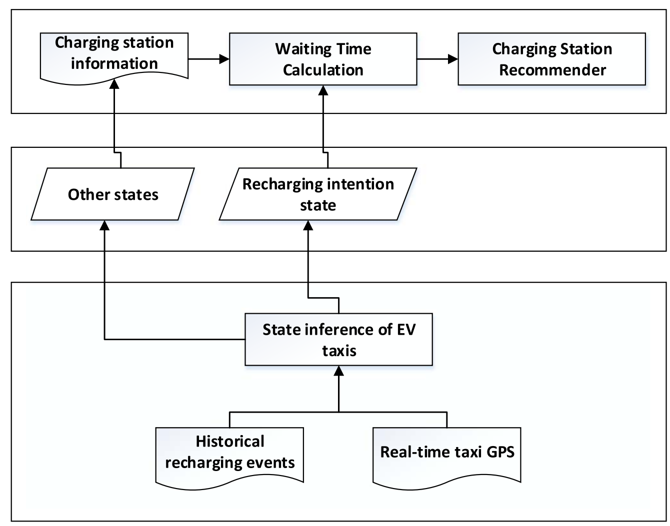


Fig. 4. Solution framework of our real-time recommendation system.

We firstly perform the EV taxi state inference, focusing on recharging intention formulation. Specifically, we investigate EV taxi drivers' recharging intentions from two perspectives: 1) at a time t, how to identify whether an EV taxi v has a recharging intention; 2) if v has a recharging intention, which charging station it chooses to go. Thereafter, we will combine EV taxi drivers' recharging intentions and the occupancy of charging piles in each station to calculate waiting time at each charging station. Finally, based on the calculated waiting time, we propose a model for our real-time recommendation system. Therefore, charging stations where cost time is minimal can be recommended for EV taxi drivers who use the system, improving their operational time on roads.

As a recommendation system, the solution presented in this paper is a user centric approach, rather than a global scheduling. Recommendation is computed upon users' requests. The recommendation workflow runs with incrementally update when a new recommendation request comes or a new real time GPS record comes from an involved taxi. As an EV reports its GPS record every 30 seconds, in case that the new GPS record shows a change of an EV's state or intention, the recommendation decision can be recalculated and issued within 30 seconds. As a

TABLE VI
TAXI STATE DESCRIPTION

state	notation	descrpition
to-station	s_t	driving towards a charging station
recharging	s_r	recharging at a charging station
operating	s_o	working on the road

driver should usually request a recommendation much earlier than 30 seconds before he arrives at a charging station, the 30 seconds computing round based on the real time data should be acceptable for updating the recommendation decision.

B. Recharging Intention Identification

The first objective of recharging intention identification is to identify whether an EV taxi v has a recharging intention at a time t. The identification is accomplished by combining historical recharging event data (Table V) and real-time taxi GPS data (Table II). As we know, there exists three states for EV taxis: *to-station*, *recharging* and *operating* as in Table VI, and *to-station* corresponds to EV taxis' recharging intentions.

We partition one day into K small time intervals, and the length of each interval is τ (we empirically set $\tau = 20$ min), thus the k_{th} interval is $I_k = [(k-1)\tau, k\tau)$ for any k. Hereafter, we apply this time partition to v's recharging event data throughout July, 2014. Therefore, for any time interval I_k, we can count the number of times different states occurred, then the probability of v's in *to-station* state (i.e., having recharge intention) in I_k can be approximated with the following *empirical probability*:

$$P_k(s_t) = \frac{N_k(s_t)}{N_k(s_t) + N_k(s_r) + N_k(s_o)} \tag{1}$$

where $N_k(s_t)$ denotes the number of times that v is in *to-station* state during the interval I_k and similar definitions for $N_k(s_r)$ and $N_k(s_o)$. The state may change in an interval as well, e.g., an EV taxi begins to recharge after it arrives at a charging station, which means this process contains *to-station* and *recharging* states. In our study, for simplicity, we suppose only one state exists in one interval and for those multiple states cases, state selection is based on the state's duration in the interval. For above instance, if the duration of *to-station* is longer than *recharging*, then we regard this EV taxi's state as *to-station* in the interval. Therefore, given an EV taxi v at time t, we firstly map t into an interval I_k and then obtain the maximum value from $\{P_k(s_t), P_k(s_r), P_k(s_o)\}$ as our preliminary predicted result.

However, there is another factor that may influence EV taxi drivers' recharge intentions: the travel distance after the last recharging event. This can be computed by analyzing the sum of *Cruise distance* and *Distance* between two consecutive recharge events in the taxi transaction data set. Although we mainly focus on time slots when predicting states of EV taxis, it follows from our observation that an EV taxi driver do not drive to a charging station for recharging if he did not travel a long distance after the last recharging event. Thus we take this distance data set into consideration when predicting their states to avoid the occurrence of this exception.

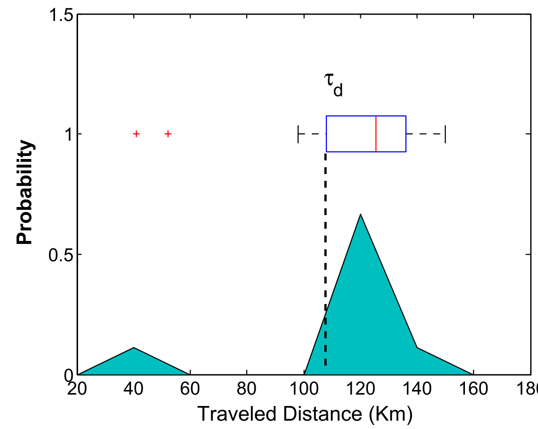


Fig. 5. The distribution of the travelled distance after last recharging event.

Specifically, for an EV taxi, we firstly obtain its data set of travelled distance after last recharging event, and then analyze its distribution, as shown in Fig. 5. Observing that they mainly fall on a same distance slot, we choose the first quartile value in the distribution as the threshold τ_d in judging the intention. Besides, we will also update this distribution once a week after collecting these values of travelled distance.

As for above instance, if the maximum value from state distribution is $P_k(s_t)$ but v's travel distance d from the time when last recharging event finishes to t deviates from the distance slot where its distance densely distributed, we suppose v will resume operating instead of having recharging intention. Actually, if an EV taxi has not travelled a long distance which means the battery consumption is relatively low, it is unnecessary to drive to a charging station for recharging.

Meanwhile, to enhance the accuracy of recharging intention identification, we also integrate taxi GPS data (Table II) into recharging event data (Table V) to accomplish this task. Specifically, we firstly extract v's GPS points from $t - \sigma$ to t (we empirically set $\sigma = 20$ min) to generate a time sequence trajectory $trj = \langle p_1, p_2, \ldots, p_m \rangle$, where p_i denotes a GPS point. If *Occupied* field of p_m equals 1, it means v is occupied by passengers, thus v will continue its carry task which tells us v has no recharging intention at t. Besides, we also apply trj into our recharging event detection method by adjusting parameters (let $r = 100$ meter and $\theta = 5\,\text{min}$) to extract v's stay points during $(t - \sigma, t)$. Here, we use $sp = \langle \hat{p}_1, \hat{p}_2, \ldots, \hat{p}_m \rangle$ to denote the sub-trajectory extracted from trj, where $\hat{p}_i$ is a stay point. If $p_m = \hat{p}_m$, it means v has stayed at a charging station for at least 5 minutes before the time t and we incline to predict v has been recharged. The Algorithm 1 summarizes above discussions to accomplish an EV taxi's state inference.

Note that in algorithm 1, states of EV taxis including operating, recharging and to-station are predicted, but to-station state is our prior concern because how to capture their recharging intention is critical in our recommendation. In this regard, historical state probability distribution possess same or even higher priority than real-time taxi GPS data since it mainly relies on state distribution to capture the recharging intentions of EV taxi drivers.

Algorithm 1 An EV Taxi' State Inference

Input: an EV taxi v, time t, time interval I_k correspond to t, v's state distribution $\{P_k(s_t), P_k(s_r), P_k(s_o)\}$ in I_k, travel distance d, v's real-time trajectory $trj = \langle p_1, p_2, \ldots, p_m \rangle$
Output: predict v's state at t.
if *Occupied* field of p_m is 1 **then**
  v's state at t is operating;
else
  if $\exists sp \subset trj$ **and** $p_m = \hat{p}_m$ **then**
    v's state at t is recharging;
  else $\{\exists sp \subset trj$ **and** $p_m \neq \hat{p}_m\}$
    v's state at t is operating;
  else
    if $P_k(s_t) = \max\{P_k(s_t), P_k(s_r), P_k(s_o)\}$ **and** $d < \tau_d$ **then**
      v's state at t is operating;
    else $\{P_k(s_t) = \max\{P_k(s_t), P_k(s_r), P_k(s_o)\}$ **and** $d > \tau_d\}$
      v's state at t is to-station;
    else $\{P_k(s_o) = \max\{P_k(s_t), P_k(s_r), P_k(s_o)\}\}$
      v's state at t is operating;
    else $\{P_k(s_r) = \max\{P_k(s_t), P_k(s_r), P_k(s_o)\}\}$
      v's state at t is recharging;
    end if
  end if
end if

C. Charging Station Selection

After identifying whether an EV taxi v has a recharging intention at t, we should also predict which charging station v is most probably heading to. Specifically, for v, we firstly obtain a list of charging stations v choose to recharge and rank those charging stations by the frequency v has visited. We denote v's ordered list as $sta = \langle s_1, s_2, \ldots, s_m \rangle$, where s_i refers to a charging station. This ordered list reveals an EV taxi driver's preference to different charging stations and for every $s_i \in sta$, we denote c_i as the count that s_i has been visited by v for recharging.

Besides the preference to charging stations, we should also consider v's position at t since the investigations through EV taxi drivers indicate that the drivers always prefer to drive to nearby charging stations instead of remote ones, thus for every $s_i \in sta$, we denote $\text{dist}(p_m, s_i)$ as the distance between v's current location at t and charging station s_i, where p_m comes from trj and indicates v's location at the moment closest to t. Then for every $s_i \in sta$, we give it a score computed as

$$\text{score}_i = \frac{c_i}{\text{dist}(p_m, s_i)}. \tag{2}$$

Then we choose the charging station with the highest score as the EV taxi's station selection.

D. Recommendation System Model

Based on the results of EV taxi state inference achieved from Algorithm 1, we design the real-time recommendation system as follows. For an EV taxi v_0 which sends a recharging request at t, the system recommends a charging station s_0 where v_0 can obtain the minimum value of *cost time*. Here, v_0's *cost time* at any charging station s will include the travel time that v_0 gets to s from its current location, and the waiting time which depends on other EV taxis arriving at s earlier than v_0, then the recommendation problem can be formulated as

$$s^* = \arg\min_{s \in S} \left(T_{v_0}(s) + W_{v_0}(s)\right) \tag{3}$$

where S is the set of charging stations while s refers to a charging station. $T_{v_0}(s)$ and $W_{v_0}(s)$ refer to the functions of travel time and waiting time, respectively. Obviously, the function $T_{v_0}(s)$ merely relates to v_0's position thus can be computed easily. However, $W_{v_0}(s)$ indeed needs other EV taxis' information, including EV taxis recharging at s and especially those having recharging intentions for s. Notice that in our study, there is an assumption of electricity price and charging rate equality in all charging stations, therefore, we mainly focus on $W_{v_0}(s)$ function.

For $W_{v_0}(s)$, we firstly find those EV taxis having recharging intentions for s, and then partition them into two groups: one containing EV taxis which will arrive at s earlier than v_0, and the other containing those later than v_0. Since the latter group won't affect v_0's waiting time at S, we only consider the former one and the process of $W_{v_0}(s)$'s calculation is performed as Algorithm 2.

Algorithm 2 Waiting Time Calculation

Input: A charging station s, the number of charging piles N_s, current time t, the number of EVs being recharging and waiting at s is l and n, respectively, the number of EVs arriving at s earlier than v_0 is k;
Output: $W_{v_0}(s)$, i.e., v_0's waiting time at s;.
if $n = 0$ **then**
  if $l + k < N_s$ **then**
    $W_{v_0}(s) = 0$;
  else $\{l + k > N_s\}$
    $M = l + k$
    when an EV taxi finishes recharging, $M = M - 1$
    denote the time when $N_s - M = 1$ as t_1;
    $W_{v_0}(s) = t_1 - t$;
  end if
else $\{n > 0\}$
  when an EV taxi finishes recharging, $n = n - 1$
  denote the time when $n = 0$ as t_2;
  $M = l + k$
  after t_2, when an EV taxi finishes recharging, $M = M - 1$
  denote the time when $N_s - M = 1$ as t_3;
  $W_{v_0}(s) = t_3 - t$;
end if

Actually, our recommendation system will continue to monitor EV taxis after sending recommendations to respond their requests. Specifically, for those EV taxis following our recommendations, their states including where to recharge and how soon reaching a charging station are available, which is taken into account when recommending for a new EV taxi. And as for those not following our recommendations, their states can be correctly supposed as to-station, i.e., having recharging intentions, thus where to recharge is our main concern in this case.

TABLE VII
SAMPLE EVALUATIONS ON STATE PREDICTION OUT OF THE WHOLE EV FLEET

Car ID	Historical Probabilities			Real-time GPS	Predicted State	Real State
	Recharging	To-station	Operating			
v_1	0.32	**0.52**	0.16	$\nexists sp \subset trj$	to-station	to-station
v_2	0.30	0.05	**0.65**	$\nexists sp \subset trj$	operating	operating
v_3	**0.74**	0.00	0.26	$\exists sp \subset trj$ **and** $p_m = \hat{p}_m$	recharging	recharging
v_4	0.22	0.13	**0.65**	$\exists sp \subset trj$ **and** $p_m \neq \hat{p}_m$	operating	operating

E. Discussions on Mutual Influence of Multiple Charging Request

There do exist cases that multiple drivers request recommendations during adjacent time period. In such case, their recommendation may be correlated and have mutual influences. Generally we deal such influences according to the first request, first served policy. Without loss of generality, we consider a case where two EVs, v_1 and v_2, request for charge station recommendations successively. As we have taken account of the charging intention predictions for adjacent EVs of v_1, if the prediction for v_2's intention is correct, then v_2's recommendation won't affect nor be affected by v_1's recommendation result. In a worse case that we failed to predict v_2's charging intention and v_2 requests for charging recommendation, we still compute v_2's recommended charge station as usual except that we consider that v_1 will be given higher priority. More specifically, if we recommended s_1 to v_1 and then v_2 sends a request out of our expectation, we will enforce the time that v_2 travels to s_1 is larger than that of v_1 in computation, even physically it is not true. The policy is based on the assumption that the misjudgment of v_2's intention is due to v_2's decision at whim, i.e., v_2's charging demand is not urgent. So we still keep the v_1's reservation.

In the worst case of the misjudgment that v_2 still goes to s_1 for charging without a recommendation request or despite of the recommendation result, the system will update the decision to v_1 as soon as it detects v_2's intention or v_2's charging state. Such strategy also mocks the route recommendation in the navigation system, where we can never ensure 100% precision in traffic jam prediction, but we can always update the route recommendation based real time traffic.

V. EXPERIMENTS AND RESULTS

A. Evaluations on State Prediction

To evaluate our prediction on EV taxis' states, we firstly collect the real states of every EV taxi at 08/01/2014 11:30 A.M. as the ground truth. This time is selected since it's the peak time of recharge as Fig. 3(a) indicates. Thus many EV taxi drivers will have recharging intentions, beneficial for evaluating on state prediction, especially on recharging intention identification. The predicted states of the whole EV taxis in Shenzhen can be obtained by exploiting algorithm 1 and we select several typical EVs to illustrate the comparisons between our predicted states and real states, as shown in Table VII.

Comparison results of over 800 EV taxis between our predicted states and real states show that the overall accuracy of our state prediction is 94.7%, among which the accuracy of to-station state prediction is 92.2%, 96.8% for recharging and 95.3% for operating. Note that for to-station states, we still need to predict which charging station is selected for recharging. Based on score$_i$ shown in Equation (2), we propose our predicted charging station selections and make a comparison with the real selections, obtaining that the accuracy of integral recharge intention identification, including to-station state and charging station selection, is 84.7%.

TABLE VIII
SAMPLE EVALUATIONS OF REAL RESULTS VERSUS RECOMMENDED RESULTS

Car ID	CS ID	RTT	PTT	RWT	PWT
ev_1	B	12	10	0	0
ev_2	D	6	5	20	15
ev_3	A	5	5	15	0
ev_4	E	15	14	0	0
ev_5	C	20	18	21	25

B. Evaluations on Recommendation Results

As mentioned previously, most of the EV taxis in Shenzhen have two drivers for each vehicle. Normally, the two drivers set up an agreement on when and where to hand over the EV taxi. Since the agreement constraint, an EV taxi driver usually chooses the charging station near the agreed place for his shift convenience, instead of our recommended charging stations even though it can save his *cost time*. Therefore, we choose time 11:30 A.M. around noon recharging peak period to evaluate our recommendation results. Another setting factor to consider is that we only choose a part of EV taxis from the whole fleet as the evaluation objectives for our recommendation system, different from the load balance problem involved the whole EV taxi fleet. Actually, when more drivers use our system, the system can acquire more exact real-time information to enhance the recommendation performance but incline to become a global optimization system. Therefore, we choose 80 EV taxis from the whole fleet supposed to be installed our recommendation system as the evaluation objectives.

Before evaluating our recommendation results, we verify the precision of our recommendation system. Assuming that for an EV taxi v_0, our system can recommend a charging station s for it, meanwhile, by exploiting recharging event detection method, v_0's real charging station selection $\hat{s}$ can be known. We choose the instances of $s = \hat{s}$ and regard their real charging station selections as the ground truth, then we compare our recommended results with the ground truth on travel time and waiting time. A part of comparison results are shown in Table VIII, where CS denotes charging station, RTT denotes real travel time, RWT denotes real waiting time, PTT denotes predicted travel time and PWT denotes predicted waiting time. The above four concepts of time use the minute as the unit.

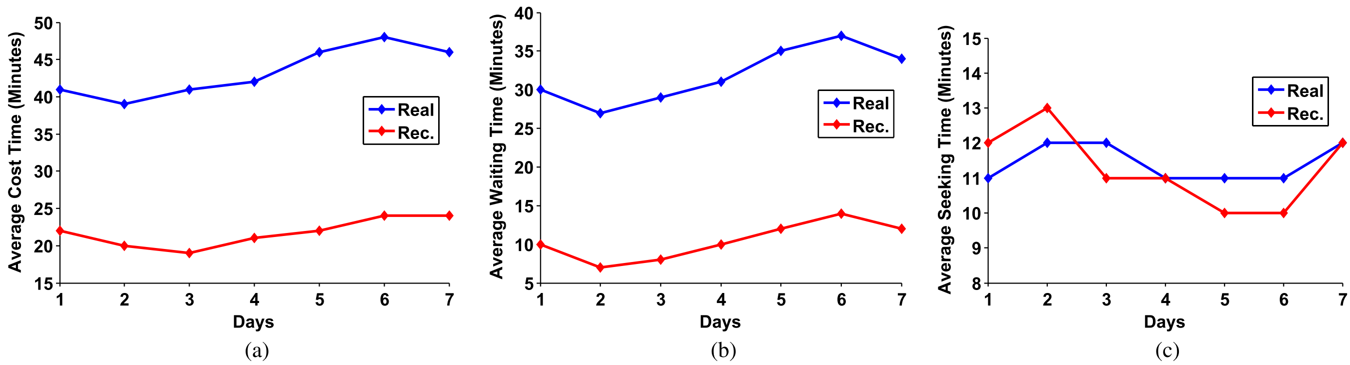


Fig. 6. Comparisons between recommended and real results. (a) Comparisons on average cost time; (b) comparisons on average waiting time; (c) comparisons on average travel time.

Our predicted results based on the recommendation system are close to real results except ev_3 whose real waiting time at A is 15 minutes while our prediction is 0 minutes. This occurs since another EV taxi drops off passengers after the time A is recommended to ev_3 and then reaches A earlier than ev_3 for recharge. Because our system only predicts EV taxis' states at the moment recommendation occurs and later state change is not observed by the system. This kind of exceptions is seldom in our application scene, thus overall recommendation results will be affected little.

We evaluate our recommended results from three perspectives: overall *cost time*, travel time and waiting time as shown in Fig. 6. We trace those 80 EV taxis as evaluation objectives for a week and then make a everyday comparison. Fig. 6(a) shows overall *cost time* comparisons between real condition and recommended condition among from those 80 EV taxis as evaluation objectives, indicating our recommendation system can reduce the *cost time* by half. Since *cost time* is composed of travel time and waiting time, thus we propose evaluations on them, respectively. Fig. 6(b) shows comparisons on average waiting time, similar with *cost time* while Fig. 6(c) displays comparisons on average travel time. Overall, the reduction of *cost time* by using our recommendation system is mainly achieved by the huge reduction in the waiting time instead of travel time, indicating that EV taxis usually don't choose their nearby charging stations of less waiting for recharging due to lack of global information about EVs and charging stations.

We also divide the typical 80 evaluation objectives into several subsets and, respectively evaluate their performances. We still use v_0's example to illustrate the rule of division. Assuming that for an EV taxi v_0, our system can recommend a charging station s for it, meanwhile, by exploiting recharge event detection method, v_0's real charging station selection $\hat{s}$ can be known. When $s = \hat{s}$, it means the driver of v_0 finds a optimal charging station (minimum *cost time*) by himself. As for $s \neq \hat{s}$, the driver of v_0 may choose a charging station nearest from him (minimum *travel time*) or with no waiting delay (*waiting time* equals 0). The performance result is shown in Table IX. It can be found that it's reasonable to take *travel time* and *waiting time* together into account rather than consider these two factors independently when searching for a optimal charging station.

TABLE IX
DIVISION AND PERFORMANCE OF OVERALL EVALUATION OBJECTIVES

Category	Condition	Proportion	RTT	RWT	PTT	PWT	Saving Time
optimal	$s = \hat{s}$	10%	12	10	10	12	0
nearest	RTT_{min}	20%	7	35	15	5	22
no waiting	$RWT = 0$	20%	38	0	7	7	24
remaining	the other	50%	14	31	12	13	20

C. Comparisons With Other Recommendation Strategies

In order to evaluate our recommendation system further, we compare it with different recommendation strategies. The recommendation strategies to be compared include the result without recommendation (*No-R* for short), with Nearest-Recommend (*Nearest-R* for short) and Shortest-Recommend (*Shortest-R* for short). No-R means no external recommendation strategy is applied at all, reflecting the current situation in Shenzhen. Then Nearest-R strategy means an EV taxi is always recommended to its nearest charging station; and finally, Shortest-R strategy always recommends an EV taxi to the charging station where the waiting time equals 0 regardless of its probable long distance from this EV taxi. Notice that the waiting time in Shortest-R strategy can be computed as the way that in our real-time recommendation system. We collect the real charging stations that the 80 evaluation objectives choose for recharging and analyze the change of waiting time at those charging stations under different recommendation strategies as shown in Fig. 7.

Compared with No-R strategy, Nearest-R makes the waiting time of charging station A and B increase sharply. This trend is caused by the fact that A and B are located in the downtown area of Shenzhen, thus many EV taxis are close to them. Although Nearest-R reduces travel time since it always recommends an EV taxi to its nearest charging station, the waiting time increases tremendously which makes Nearest-R unfavorable. On contrary, the waiting time achieved from our recommendation system clearly outperforms both, and it is even close to the Shortest-R. In the Shortest-R, although the waiting time is reduced, EV taxis are directed to charging stations far from their current locations, leading to two drawbacks: large travel time is incurred by the long-distance travel, and whether the taxis can reach those stations is unknown due to the low state of charge (SoC) of batteries.

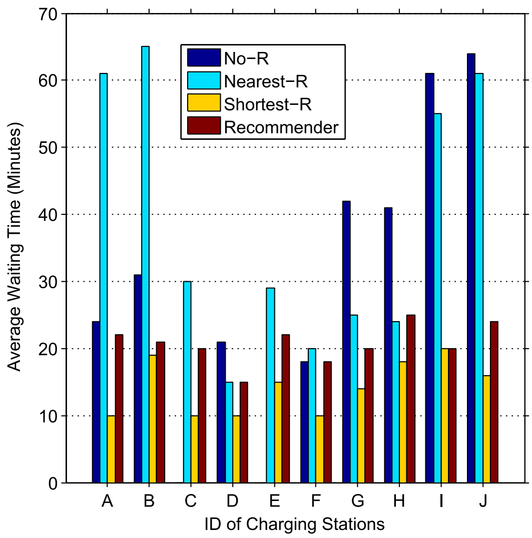


Fig. 7. Comparisons of waiting time at selected charging stations.

Actually, for EV taxi drivers, time spent on waiting at charging stations is much more than that on driving to charging stations, therefore, EV taxi drivers are more concerned about waiting time. Obviously, saving more time on waiting can make EV taxi drivers pick up more passengers, thus they don't care about more electricity consumption to find a proper charging station in this case. Therefore, we mainly focus on waiting time among these different recommendation strategies.

D. Evaluations on Utilization of Charging Piles

Through the publicly available information of the charging stations (Table IV), we can obtain the total number of charging piles at each station. Recharging efficiency also can be displayed by the utilization of the charging piles since when the amount of recharge requests is constant, smaller number of occupied posts means larger number of EV taxis waiting at charging stations. Thus a proper recommendations for EV taxis should make more charging piles occupied, improving the level of charging piles utilization. We use the ratio of the number of occupied charging piles over the number of total piles to reflect the utilization of charging piles. The utilization ratio is evaluated with the periods ranging from 9:30 A.M.–13:30 P.M., which can avoid work shift conditions (mentioned in Section V-B) and cover the noon recharging peak period. We choose 30 minutes as the size of the interval because recharge events always take up to hours. The average charging piles utilization across all charging stations during 9:30 A.M.–13:30 P.M. is presented in Fig. 8.

The total number of charging piles in Shenzhen is about 700, i.e., 5% improvement on charging pile occupancy means about 35 charging piles turn into occupied state from non-occupied, easing the seriousness of queuing delay. As shown in Fig. 8, the utilization achieved by our recommendation system outperforms No-R and Nearest-R almost in all periods.

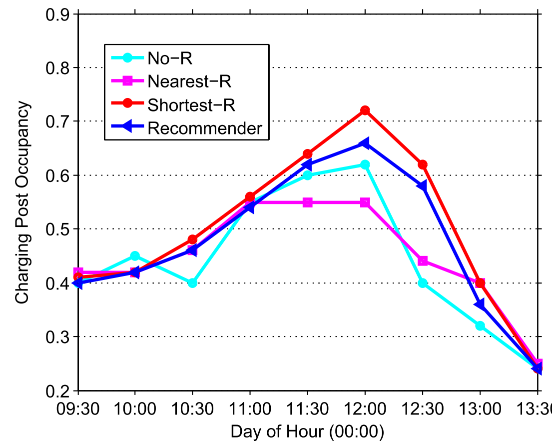


Fig. 8. Evaluations on utilization of charging piles.

VI. CONCLUSION

Taxi GPS traces can be exploited to investigate EV taxi drivers' recharging behavior patterns. In this paper, we have studied EV taxi drivers' recharging intention identification based on the activities of over 800 EV taxis in Shenzhen. To understand recharging intentions of EV taxi drivers, we first propose a method to detect their recharging events and then analyze their historical recharging behavior patterns by utilizing the detected recharging events data and field investigations through EV taxis drivers. The investigations are mainly focused on two perspectives: 1) EV taxi drivers usually have recharging intentions at a fixed period; 2) although over 50 charging stations are deployed in Shenzhen, most of EV taxi drivers choose 6–8 stations among them regularly. Based on our verifications for these investigations, we combine historical recharging event data and real-time taxi GPS data to identify drivers' recharging intentions, including when and where they will choose for recharging.

In this paper, we present a real-time recommendation system for them by linking charging stations' operational condition information to reduce their cost time for recharging. The system first predict EV taxi drivers' recharging intentions. Then, given the current location and time of an EV taxi that sending a recharging request, the system can recommend a charging station for the EV taxi driver, to which the driver's overall cost time for recharging is most likely to be minimal. Our extensive analysis on the real data set shows that our system can reduce the cost time by 50% in Shenzhen. Although the research is based on the study case in Shenzhen, we claim that most of our preliminary observations will be commonly observable among EV taxi drivers in general, and the theoretic model of our system is irrelevant to the city, which make it evident that our system is universally applicable in any city or country considering the adoption of EV taxi system.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for their valuable comments. This work was done while the first author was an intern in Shenzhen Institutes of Advanced Technology.

REFERENCES

[1] South Korea Releases Electric Public Transportation System. [Online]. Available: http://globalenergyinitiative.org/south-korea-releases-electric-public-transportation-system.html.

[2] Electric Bus. [Online]. Available: http://www.transitchicago.com/electricbus.

[3] More Than 35 Technology. [Online]. Available: http://www.apta.com/mediacenter/pressreleases/2013/Pages/130422_Earth-Day.aspx.

[4] Z. Tian *et al.*, "Understanding operational and charging patterns of electric vehicle taxis using GPS records," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 2472–2479.

[5] F. Zhang, D. Wilkie, Y. Zheng, and X. Xie, "Sensing the pulse of urban refueling behavior," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2013, pp. 13–22.

[6] M. Kuby and S. Lim, "The flow-refueling location problem for alternative-fuel vehicles," *Socio-Econ. Plan. Sci.*, vol. 39, no. 2, pp. 125–145, 2005.

[7] J.-G. Kim and M. Kuby, "The deviation-flow refueling location model for optimizing a network of refueling stations," *Int. J. Hydrogen Energy*, vol. 37, no. 6, pp. 5406–5420, 2012.

[8] I. Capar, M. Kuby, V. J. Leon, and Y.-J. Tsai, "An arc cover-path-cover formulation and strategic analysis of alternative-fuel station locations," *Eur. J. Oper. Res.*, vol. 227, no. 1, pp. 142–151, 2013.

[9] J. Jung, J. Y. Chow, R. Jayakrishnan, and J. Y. Park, "Stochastic dynamic itinerary interception refueling location problem with queue delay for electric taxi charging stations," *Transp. Res. C, Emerging Technol.*, vol. 40, pp. 123–142, 2014.

[10] M. Gharbaoui *et al.*, "Designing and evaluating activity-based electric vehicle charging in urban areas," in *Proc. IEEE IEVC*, 2013, pp. 1–5.

[11] P. Jochem, C. Brendel, M. Reuter-Oppermann, W. Fichtner, and S. Nickel, "Optimizing the allocation of fast charging infrastructure along the German Autobahn," *J. Business Econ.*, pp. 1–23, 2015.

[12] Z. Ma, D. Callaway, and I. Hiskens, "Decentralized charging control for large populations of plug-in electric vehicles," in *Proc. 49th IEEE CDC*, 2010, pp. 206–212.

[13] L. Gan, U. Topcu, and S. Low, "Optimal decentralized protocol for electric vehicle charging," in *Proc. IEEE 50th CDC-ECC*, 2011, pp. 5798–5804.

[14] 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.

[15] O. Sundstrom and C. Binding, "Planning electric-drive vehicle charging under constrained grid conditions," in *Proc. IEEE Int. Conf. POWERCON*, 2010, pp. 1–6.

[16] 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*. New York, NY, USA: Springer-Verlag, 2010, pp. 274–278.

[17] 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.*, 2011, pp. 51–60.

[18] J.-L. Lu *et al.*, "Operating electric taxi fleets: A new dispatching strategy with charging plans," in *Proc. IEEE IEVC*, 2012, pp. 1–8.

[19] F. Malandrino, C. Casetti, and C.-F. Chiasserini, "A holistic view of its-enhanced charging markets," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1736–1745, Aug. 2014.

[20] Z. Yang *et al.*, "Profit maximization for plug-in electric taxi with uncertain future electricity prices," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3058–3068, Nov. 2014.

[21] J. Yang and Z. Yang, "Optimal scheduling of electric vehicle using dual decomposition," in *Proc. IEEE 26th CCDC*, 2014, pp. 2144–2149.

[22] P. You and Z. Yang, "Efficient optimal scheduling of charging station with multiple electric vehicles via v2v," in *Proc. IEEE Int. Conf. SmartGridComm*, 2014, pp. 716–721.

[23] Y. Li, C. Tian, F. Zhang, and C. Xu, "Traffic condition matrix estimation via weighted spatio-temporal compressive sensing for unevenly-distributed and unreliable GPS data," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 1304–1311.

[24] D. Zhang *et al.*, "Understanding taxi service strategies from taxi GPS traces," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 123–135, Feb. 2015.

[25] C. Chen *et al.*, "Real-time detection of anomalous taxi trajectories from GPS traces," in *Mobile and Ubiquitous Systems: Computing, Networking, and Services*. New York, NY, USA: Springer-Verlag, 2012, pp. 63–74.

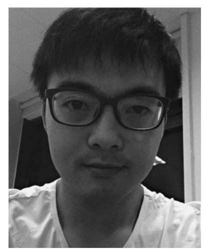
[26] J. Zhang *et al.*, "Analyzing passenger density for public bus: Inference of crowdedness and evaluation of scheduling choices," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 2015–2022.

[27] L. Yin, J. Hu, L. Huang, F. Zhang, and P. Ren, "Detecting illegal pickups of intercity buses from their gps traces," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 2162–2167.

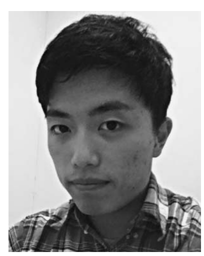
[28] J. Huang, L. Wang, C. Tian, F. Zhang, and C. Xu, "Mining freight truck's trip patterns from gps data," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 1988–1994.

[29] F. Zhang *et al.*, "Spatio-temporal segmentation of metro trips using smart card data," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1137–1149, Mar. 2016.

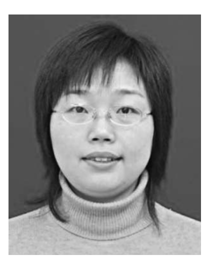
[30] J. Zhao, C. Tian, F. Zhang, C. Xu, and S. Feng, "Understanding temporal and spatial travel patterns of individual passengers by mining smart card data," in *Proc. IEEE 17th Int. Conf. ITSC*, 2014, pp. 2991–2997.



Zhiyong Tian received the B.S. and M.S. degrees from Huazhong University of Science and Technology, China, in 2010 and 2012 respectively. He is currently working toward the Ph.D. degree with Huazhong University of Science and Technology. His research interests include electric vehicles, big data analysis, big-data-driven systems, and spatio-temporal data mining.



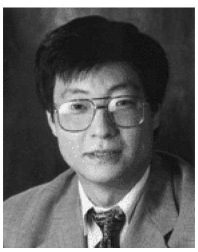
Taeho Jung (S'15) received the B.E. degree in computer software from Tsinghua University, Beijing, in 2011. He is currently working toward the Ph.D. degree in computer science with Illinois Institute of Technology, under the supervision of Dr. X.-Y. Li. He is currently working on the privacy-preserving computation in various applications and scenarios where big data is involved. His research interests include privacy and security issues in big data analysis and networking applications.



Yi Wang received the B.S., M.S., and Ph.D. degrees from Huazhong University of Science and Technology, China, in 2000, 2003, and 2009, respectively. She is currently a Lecturer with the School of Electronics Information and Communications, Huazhong University of Science and Technology, China. Her research interests include big data for smart transportation.

Fan Zhang received the Ph.D. degree in communication and information system from Huazhong University of Science and Technology in 2007. He is currently an Associate Professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. From 2009 to 2011, he was a Postdoctoral Fellow with the University of New Mexico and University of Nebraska-Lincoln. His research interests include big data processing, data privacy, and urban computing.

Lai Tu received the B.S. in communication engineering and the Ph.D. degree in information and communication engineering from Huazhong University of Science and Technology, China, in 2002 and 2007, respectively. From July 2007 to December 2008, he worked as a Postdoctoral Fellow with the Department of EIE, Huazhong University of Science and Technology. From January 2009 to October 2010, He worked as a Postdoctoral Researcher with the Department of CSIE, Nation Cheng Kung University, Taiwan. He is currently an Associate Professor with the School of Electronic and Information and Communications, Huazhong University of Science and Technology. His research interests include urban computing, human behavior study, mobile computing and networking.



Chengzhong Xu (F'16) received the Ph.D. degree from the University of Hong Kong in 1993. He is currently a Professor with the Department of Electrical and Computer Engineering, Wayne State University, USA. He also holds an adjunct appointment with the Shenzhen Institute of Advanced Technology of Chinese Academy of Science as the Director of the Institute of Advanced Computing and Data Engineering. He is the author of more than 200 papers in journals and conferences. His research interests include parallel and distributed systems and cloud computing. He serves on a number of journal editorial boards, including IEEE Transactions on Computers, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Cloud Computing, *Journal of Parallel and Distributed Computing* and *China Science Information Sciences*. He was the Best Paper Nominee of 2013 IEEE High Performance Computer Architecture (HPCA) and the Best Paper Nominee of 2013 ACM High Performance Distributed Computing (HPDC). He received the Faculty Research Award, Career Development Chair Award, and the Presidents Award for Excellence in Teaching of WSU. He was also a recipient of the "Outstanding Oversea Scholar award of NSFC.

Chen Tian received the B.S., M.S., and Ph.D. degrees from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, China, in 2000, 2003, and 2008, respectively. He was previously an Associate Professor with the School of Electronics Information and Communications, Huazhong University of Science and Technology, China. From 2012 to 2013, he was a Postdoctoral Researcher with the Department of Computer Science, Yale University. He is currently an Associate Professor with the State Key Laboratory for Novel Software Technology, Nanjing University, China. His research interests include data center networks, network function virtualization, distributed systems, Internet streaming, and urban computing.

Xiang-Yang Li (M'99–SM'08–F'15) received the Bachelor's degree with the Department of Computer Science and the Bachelor's degree with the Department of Business Management from Tsinghua University, China, in 1995, and the M.S. and Ph.D. degrees with the Department of Computer Science, University of Illinois at Urbana-Champaign, in 2000 and 2001, respectively. He is currently a Professor with the School of Computer Science and Technology, University of Science and Technology of China. He was previously a professor at Department of Computer Science, Illinois Institute of Technology. His research interests include wireless networking, mobile computing, security and privacy, cyber physical systems, and algorithms. He has served many international conferences in various capacities, including ACM MobiCom, ACM MobiHoc, IEEE MASS. He serves as an Editor for several journals, including IEEE Transaction on Mobile Computing, and IEEE/ACM Transaction on Networking. He is an ACM Distinguished Scientist. He and his students won five Best Paper Awards (IEEE GlobeCom 2015, IEEE HPCCC 2014, ACM MobiCom 2014, COCOON 2001, IEEE HICSS 2001) and one Best Demo Award (ACM MobiCom 2012), He is the author of the monograph *Wireless Ad Hoc and Sensor Networks: Theory and Applications*. He is a Co-Editor of several books, including *Encyclopedia of Algorithms*.