Real-time Route Recommendations for E-Taxies Leveraging GPS Trajectories

Wei Tu, *Member, IEEE*, Mai Ke, Yatao Zhang, Yang Xu, Jincai Huang, Min Deng,
Long Chen, *Senior Member, IEEE*, Qingquan Li

Abstract—Electric vehicles (EVs) currently face formidable challenges in promotion, i.e., short driving ranges, long charging times, and few charging stations, thereby limiting their acceptability to taxi drivers. Leveraging massive-scale taxi GPS trajectory data, we present a novel real-time route recommendation system for electric taxi (ET) drivers. Taxi travel knowledge, including the probability of picking up passengers and the distribution of destinations, is learned from the raw GPS trajectories. Considering the cascading effect of route decision making, consecutive ET actions are modeled with an action tree. The corresponding expected net revenue is estimated based on the learned knowledge. A prototype online system is developed for providing route recommendations, e.g., when to go to a charging station or cruise on certain roads. An experiment in Shenzhen demonstrates that the average daily net revenue of ET drivers is better than that of 76.2% of gasoline taxi drivers. The presented approach not only increases the revenue of ET drivers in the short term but also improves the viability of EVs in the long run.

Index Terms—Electric taxies, Action tree search, Taxi recommendation, GPS trajectories.

I. INTRODUCTION

ELECTRIC vehicles (EVs), along with autonomous vehicles [1] and connected vehicles [2], are bringing disruptive innovations to urban transportation [3]–[5]. Taxies, the most used vehicles in cities, have attracted considerable attention for electrification [6]. Several cities, such as New York, Shenzhen, and Beijing, have launched initiatives to promote the use of EVs in the taxi industry. In comparison to gasoline vehicles, however, EVs still have unignorable disadvantages at present, such as short driving ranges [4]

and long charging times (0.5-10 hours) [3], [7]–[9]. The long charging durations in particular will sharply reduce the on-road service times of electric taxies (ETs) [6], [7], [10]. These shortcomings, accompanied with low gas price, are likely to demotivate people and corporations from adopting EVs. Hence, the replacement of gasoline vehicles in the taxi community is still slow because taxi drivers tend to pursue high daily revenue with continuous driving. Li-ion battery and super-charging technologies can greatly improve EVs' performance in the future. But widely commercializing those technologies and deploying new charging infrastructure take time. Consequently, the widespread adoption of ETs is currently facing great challenges, thus highlighting the need for effective policies and strategies to promote ETs [11], [12].

Many efforts have been made to improve the acceptance of ETs, such as EV subsidy policies [10], optimally locating ET charging stations [6], and encouraging smart taxi operations [13]. In terms of taxi operations, traditional studies [14]–[18] have focused on intelligent route recommendation strategies for gasoline taxies. For example, Yeun et al. [19] developed an intelligent ride-sharing route recommendation framework that suggests the best route with the highest probability of finding compatible customers. Furthermore, knowledge of typical urban taxi travel, e.g., travel distance, travel time, and the spatial distribution of taxi demand [20], [21], can be used to improve route recommendations for taxi drivers. Kong et al. [22] revealed hidden human mobility patterns and the pick-up/drop-off relationship based on raw trajectories. They developed a taxi service recommendation model based on a Gaussian regression process to improve taxi drivers' profits and enhance passengers' travel experience. Yuan et al. [23] recommended routes matching vacant taxies and clients for gasoline taxi drivers. Qu et al. [24] recommended profitable taxi driving routes with minimal expected driving distances by estimated the probabilities of picking up passengers and the capacities at different locations from taxi GPS trajectories. Although these methods improve the performance of gasoline taxi drivers, they do not perform well for ETs, mainly for the following reasons: 1) currently, EVs require more than 30 minutes to fully recharge, thereby lowering their efficiency in serving clients [10], and 2) charging stations are spatially sparse, thus increasing ET drivers' anxiety about range [25].

Several studies have been devoted to route recommendation for ETs based on certain effective strategies, such as maximizing drivers' revenue by considering the dynamic price of electricity [26] and balancing the utilization of charging facilities to reduce social cost [27]. Yang et al. [13] proposed

a game-theory-based method to reduce the potential service income loss of ETs, which considered the queuing time and the long charging duration at charging stations. Tian et al. [25] predicted ET drivers' recharging intentions and the states of charge of ETs to estimate the queuing time at a charging station. They further developed a charging station recommender to decrease the waiting and detour times for ET recharging. However, those studies mainly focused on the interaction between ETs and charging facilities, neglecting the question of optimal routes for finding passengers. Recently, Tseng et al. [12] presented an ET service strategy to immediately maximize the expected net income of drivers' passenger-finding and recharging actions, but they underestimated the influence of the recharging action on drivers' revenue. These achievements focused on the interaction between ETs and charging facilities. However, they overlooked the potential influence of taxi service over the next several hours. From the daily cycle perspective, the actions of cruising on the road to find passengers and recharging at stations should be further coordinated. Thus, a comprehensive route recommendation method is necessary for ET drivers.

Massive-scale historical taxi GPS trajectory data contain rich spatial-temporal knowledge of taxi travel in a city, such as the distributions of taxi demand and clients' destinations and the taxi travel rhythm [23], [28], [29]. These data provide useful insights for deeply understanding taxi travel. Therefore, they enable us to simultaneously coordinate passenger delivery, cruising and recharging for ETs.

Motivated by current shortcomings and damped promotion of ETs, this study presents a comprehensive real-time route recommendation system for ET drivers to improve their net revenue. Knowledge of taxi travel is learned from massive-scale raw GPS trajectory data. The sequential actions of ET drivers, including the cruising on the road and the recharging at stations, are modeled by *Markov Decision Process* (MDP). The spatial-temporal expected net revenue (ENR) of potential ET actions is estimated using the learned knowledge. A spatial-temporal ET action tree is proposed to model the ETs' consecutive decision to maximize the expected net revenue. A prototype of an online route recommendation system is developed. The experimental results show that the developed system significantly outperforms baseline methods.

The main contributions of this study are as follows:

- **A comprehensive route recommendation system** for ET drivers is developed, which incorporates the cruising on the road and the recharging at stations.
- **A probabilistic action-based tree (ABT) recommendation method** is proposed to improve the ENRs of ET drivers over a long time period, which is learned from massive gasoline taxi GPS trajectories.
- **A speed-up strategy** is developed to accelerate the computing process and thus facilitate real-time ET route recommendations to meet the needs of practical application.
- **An intensive experiment** demonstrates that the presented approach is superior to baseline methods. Various ET settings and a sensitivity analysis are also considered to evaluate the performance of the proposed approach.
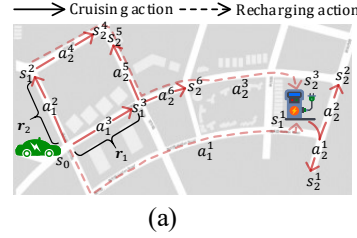

Fig. 1. Cruising, recharging and consecutive decisions of ET drivers in two steps.

The remainder of this paper is organized as follows. Section I defines the route recommendation problem for ET drivers. Section III describes the proposed method. Section IV reports the experiment and presents the comparison of results. We conclude this work in Section V.

II. PROBLEM DEFINITION

Taxies naturally traverse roads to find and deliver passengers. Compared to refueling of gasoline taxies, ETs spend more time in recharging at charging stations. Based on the ET driving cycle, we assume the following:

- An ET driver cruises on the road to find potential passengers.
- After picking up a passenger, An ET driver will immediately travel to the passenger's destination.
- An ET driver can recharge his or her ET only at charging stations.
- After dropping off a passenger, an ET driver will either cruise to find his or her next passenger or go to a charging station.

Fig. 1a presents an example of an ET at a road junction without any passenger aboard. The ET driver has three possible actions: (i) cruising on road $r1$, (ii) cruising on road $r2$, or (iii) going to a charging station while rejecting passenger pick-up.

A. MDP Definition

We utilize an MDP to describe the driving and recharging decisions of ET drivers.

State: The state of an ET is defined as a triple $s = (i, e, t) \in S$, where $i \in I$ is a road junction, denoting the current position; e is the remaining battery capacity of the ET; and $t \in T$ is the time in a day.

Action: An action is defined as $a = (i \to j, \theta) \in A$, where $i \to j$ denotes the travel from position $i \in I$ to $j \in I$; $\theta \geq 0$ denotes the recharging duration if the ET recharges at a station located at junction j. There are two types of actions:

i. The *cruising* action, denoted as $a \in A_c$, indicates that the vacant ET cruises, finds and delivers passengers. The movement $i \to j$ of a cruising action is confined in neighbor junctions, that is, there must be a road r_{ij} that connects i and j. The recharging duration is always zero, $\theta = 0$.

ii. The *recharging* action, denoted as $a \in A_r$, indicates that the driver will go to a charging station at j and recharge for $\theta > 0$ minutes. The locations of all available charging stations form a set, $\mathbf{C}$.

In accordance with the real-world taxi service scenario, an ET driver will refuse to pick up any passenger while on the way to a charging station. Here, we only focus on cruising routes and charging options.

Furthermore, we define $n \geq 1$ consecutive actions as an action sequence $\vec{A}_n = (a_1 \to a_2 \to ... \to a_n)$. Fig. 1b shows the example of all consecutive actions of the ET in two steps.

State Transition: The next state s' of an ET depends on the current state s and the action a it takes. For example, the cruising action will change the position and deplete the battery, while the recharging action replenishes the power for the ET battery. Under the assumptions that an ET in state $s = (i, e, t)$ is performing a cruising action $a = (i \to j, \theta = 0)$ and that the probability of successfully picking up a passenger on road r_{ij} at time t is $P_t^s(a) = P_t^s(r_{ij})$, the transition to the next state s' can be described as follows:

i. If the ET fails to pick up any passenger, $s' = (j, e - E(i, j), t + T_t(i, j))$, where $E(i, j)$ and $T_t(i, j)$ are the electricity and time, respectively, consumed when traveling from i to j. In this study, we calculate $T_t(i, j)$ using historical taxi trajectories. Note that $T_t(i, j)$ can be refined by real-time traffic to achieve better accuracy [20].

ii. If the ET successfully picks up a passenger, $s' = (k, e - E(i, j) - E(j, k), t + T_t(i, j) + T_t(j, k))$, where k is the destination of the passenger. Note that the probability of picking up a passenger with a specific destination $k \in I$ is $P_t^s(a, k) = P_t^s(r_{ij}, k)$, and $P_t^s(a) = \sum_{k \in I} P_t^s(a, k)$.

iii. If the ET performs a recharging action, $s' = (j, e - E(i, j) + \theta \cdot \alpha, t + T_t(i, j) + \theta)$, where α is charging rate (kW).

B. Expected Net Revenue (ENR)

An ET driver will earn rewards by taking actions. Here, the reward is defined as the net revenue. For a trip from i to k, the net revenue is as follows:

$$R(i, k) = R^*(i, k) - C(i, k) \tag{1}$$

where $R^*(i, k)$ is the revenue generated by traveling from i to k, which can be calculated based on the time and total distance, and $C(i, k) = E(i, k) \cdot \beta$ is the product of the consumed electricity and the electricity price β (CNY/kWh). Clearly, a trip without a passenger earns no profit, i.e., $R^*(i, k) = 0$.

The net revenue of an action varies with different destinations of the passenger. Considering the pick-up probability and the distribution of passengers' destinations, the ENR of one ET trip can be represented by (2):

$$E[R(s, a)] = E[R(i, j, t, a)] = \sum_{k \in I} [P_t^s(a, k) R(j, k)] - C(i, j) \tag{2}$$

where $E[]$ denotes the expectation.

Furthermore, the ENR of a set of sequential actions $\vec{A}_n = (a_1 \to a_2 \to ... \to a_n)$ can be calculated as shown in (3):

$$\begin{aligned} R_n(s_0, a_1) = &\sum_{k \in I} [P_t^s(a_1, k) R_{n-1}(s_k, a_k)] \\ &+ (1 - P_t^s(a_1)) R_{n-1}(s', a_2) + E[R(s_0, a_1)] \end{aligned} \tag{3}$$

where $R_0(s, a) = 0$; s_k and s' are the next states of s_0 if the ET successfully and unsuccessfully picks up and delivers

a passenger to k by a_1, respectively; and a_k denotes an action that the driver may take in state s_k. $E[R(s, a)]$ can be denoted as $R_1(s, a)$.

To recommend the optimal route for ET drivers, we should find the best action sequence to maximize the ENR, in other words, a path with maximum ENR in Fig. 1.

C. Electricity Constraints

Due to the limitation of the driving range and the spatial distribution of charging stations, some actions are infeasible for an ET when the remaining battery capacity e is insufficient to reach some station $j \in \mathbf{C}$. In addition, ET drivers need to reserve a certain level of remaining battery capacity, denoted as $\underline{e}$, in case of an emergency, e.g., traffic congestion. To guarantee at least $\underline{e}$ before arriving at a charging station, the feasibility of one action is represented by (4).

$$e - E(i, j) - \min_{h \in \mathbf{C}} \{E(j, h)\} > \underline{e} \tag{4}$$

Hence, the feasible actions are necessarily confined to a subset of A, which contains all actions that satisfy (4).

On the other hand, some specific destinations k of passengers are not reachable when e is insufficient for driving to any charging station after arriving at k. The ET driver has to reject those passengers. The reachable passenger destinations for an ET in state $s = (i, e, t)$ taking action $a = (i \to j, \theta = 0)$ should satisfy (5):

$$e - E(i, j) - E(j, k) - \min_{h \in \mathbf{C}} \{E(k, h)\} > \underline{e} \tag{5}$$

We denote the left side of (5) by $L(s, j, k)$ for simplicity.

Therefore, the probability of successfully picking up and delivering a passenger with an unreachable destination k when taking an action a is zero, that is, $P_t^s(a, k) = 0$.

D. The ET Route Recommendation Problem

The objective of the ET route recommendation problem considered in this work is to improve the net revenue of ET drivers by recommending routes that maximize the ENR of multiple future actions, considering both paths for finding passengers and going to charging stations.

The ET Route Recommendation Problem: Given an ET with a state s_0 and a fixed number of future actions $n \geq 1$, find and recommend the best action sequence $\vec{A}_n^*$ that maximizes (3) for the ET driver.

III. METHOD

We propose a comprehensive real-time ET recommendation system to receive the real-time requests from ET drivers and provides timely route recommendation. The recommended route is a sequence of cruising and recharging actions $\vec{A}_n^*$, as shown in Fig. 2. This system consists of an offline process and an online process. In first process, massive-scale historical taxi GPS trajectory data are mined to extract essential taxi knowledge. By integrating the knowledge with the topology of the road network, an offline ABT is built to find the optimal action by considering both cruising and recharging
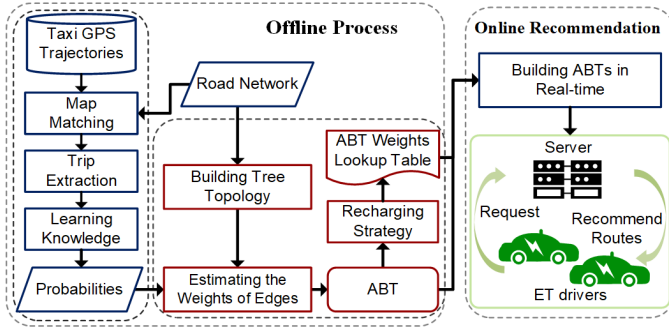
Fig. 2. Framework of the proposed system.

Fig. 4. The cumulative probability distribution of the travel distance in typical area of Shenzhen. (a) 8-9 AM. (b) 5-6 PM.

actions. The edge weights of the ABT are pre-estimated and stored in a lookup table for online recommendation. In the second process, the developed system receives ET drivers' real-time requests, and then, provides routes recommendation by building ABTs immediately.

A. Learning Knowledge from Taxi Trajectories

Learning knowledge from taxi trajectories includes three steps: map matching, taxi trip extraction, and statistical learning.

1) Map matching:

Using the parallel map matching technique [30], GPS points are mapped to the road network. By arranging each taxi's GPS records in time order, the spatiotemporal trajectories of each taxi are constructed.

2) Trip extraction:

Both the occupied trips and unoccupied trips of each taxi are extracted from the trajectories based on the occupied status (whether the taxi is carrying passengers or vacant) [6]. Using the occupied trips, the locations and times of the pick-up and drop-off activities are then extracted.

Fig. 3 shows the temporal variation and spatial distribution of passenger pick-up in Shenzhen. This figure illustrates that roadside taxi demands change with time and vary in space. Therefore, considering the long charging times of ETs, the actions of cruising on the road and recharging at stations should be suitably coordinated.
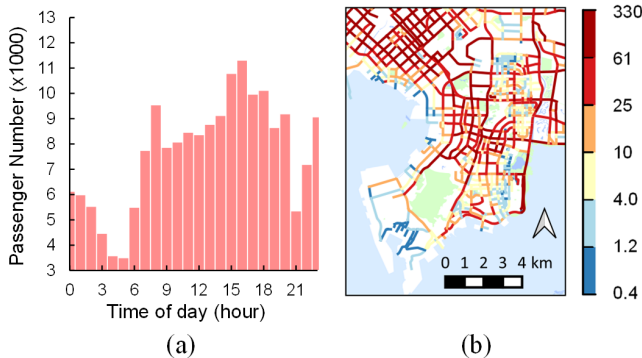

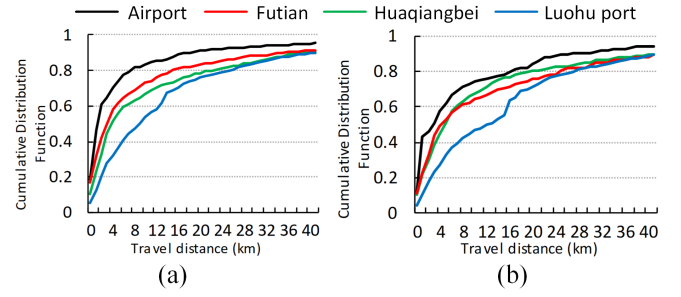Fig. 3. (a) Temporal variation and (b) spatial distribution of passenger pick-up.

3) Statistical learning:

Two probabilities of travel demand can be learned from taxi trajectories. Both the taxi demand and traffic vary from minute to minute. Hence, we divided the entire day into even 30 minute time slots and calculated these probabilities.

i. The probability of successfully picking up a passenger on road r_{ij} at time t:

$$P_t(r_{ij}) = \frac{Occ(r_{ij})}{Cru(r_{ij}) + Occ(r_{ij})} \qquad (6)$$

where time t is within time slot T and $Occ(r_{ij})$ and $Cru(r_{ij})$ are the numbers of occupied trips and cruising (unoccupied) trips, respectively, on road r_{ij} within time slot T.

ii. The probability of successfully picking up a passenger with destination k on road r_{ij}:

$$P_t(k|r_{ij}) = \frac{Occ(r_{ij}, k)}{Occ(r_{ij})} \qquad (7)$$

where $Occ(r_{ij}, k)$ is the number of occupied trips with the origin at r_{ij} within T and the destination at k.

The knowledge learned from gasoline taxi trajectories is then transferred to the ET domain. Meanwhile, the ETs are constrained by the battery and the available charging stations. Considering the constraint of the reachable passenger destinations, if an ET in state $s = (i, e, t)$ takes a cruising action $a = (i \rightarrow j, \theta = 0)$,

i. The probability of successfully picking up a passenger is

$$P_t^s(a) = P_t^s(r_{ij}) = P_t(r_{ij}) \sum_{k \in I : L(s,j,k) > 0} P_t(k|r_{ij}) \qquad (8)$$

where the probabilities of picking up passengers with unreachable destinations are excluded.

ii. The probability of successfully picking up a passenger with a reachable destination k is

$$P_t^s(a, k) = P_t^s(r_{ij}, k) = P_t^s(r_{ij})P(k|r_{ij}) \qquad (9)$$

Fig. 4 shows the cumulative probability of picking up passengers with different travel distance. This figure shows that the ability of ETs to pick up passengers not only varies in time and space but also is affected by the remaining battery capacity.

B. Action-based Tree (ABT)

We construct an ABT to model the consecutive actions and potential net revenue of an ET driver in n steps, including his/her *cruising* and *recharging* actions and corresponding

Fig. 5. Illustration of building an action-based tree. (a) Building the tree topology. (b) Estimating the weight of an edge.

ENRs. A node of an ABT is a possible state of the ET in the next n steps. A downward edge is a feasible action that incurs the transition to the next state. Each edge in the ABT is labeled with a weight equal to the maximum ENR of the corresponding action with the state. A path from the root to a leaf contains sequential edges, indicating consecutive feasible actions. For example, the red path shown in Fig. 5a denotes two consecutive feasible actions $(a_1^2 \rightarrow a_2^4)$. By choosing a path with the maximum weight among all the following paths, the best action sequence $\vec{A}_n^*$ can be obtained.

1) Building ABT:

Given an ET driver in state s_0 and a fixed number of future actions $n \geq 1$, the process of building and using an ABT to find the best action sequence is described as follows:

i. Building the tree topology:

According to the road network, locations of charging stations, and remaining battery capacity, all consecutive feasible actions of the ET driver are deduced to build the topology of the tree. Generally, an action a taken in state s can produce an edge to a child node representing the next state. When all consecutive actions in n steps are explored, the topology of an ABT is produced. Fig. 5a shows an example of building a tree's topology for the ET driver shown in Fig. 1a. Note that all infeasible actions are removed by constraint (4).

ii. Estimating the edge weights:

The weight of each edge is the maximum ENR of the corresponding state-action pair $\langle s, a \rangle >$, where s is the parent node and a is the action. As (3) suggests, calculating the ENR of a state-action pair relies on the computation of the ENRs of the succeeding states and actions. Hence, a leaf-to-root approach is used to estimate the weights. The weights of the edge linking to the leaf nodes are first computed using the normal ENR as (2). Regarding the edge linking to a nonleaf node, its weight is the sum of the maximum ENRs of passenger delivery of the current action and the maximum of succeeding state-action pairs in the next node.

For example, the blue arrows in Fig. 5b show the order of calculating the weights of three edges, where the weight $R_2(s_0, a_1^3)$ of edge $\langle s_0, a_1^3 \rangle$ is the sum of the ENR of passenger delivery and the maximum value between $R_1(s_1^3, a_2^5)$ and $R_1(s_1^3, a_2^6)$. Consequently, we implement *max* operators at

nodes and *sum* at edges to calculate the maximum ENR, i.e., the weight.

By recursively calculating the ENR from the leaf node to the root, the maximum ENRs of all state-action pairs $\langle s, a \rangle$ are estimated. Figuratively speaking, each edge of the tree is labeled with a weight.

iii. Action selection:

Given the current state of an ET, the corresponding ABT can be built. From the root node, by sequentially selecting the next edge with maximum weight in the current reached node until reaching a leaf node, optimal sequential actions $\vec{A}_n^*$ are obtained.

2) Recharging strategy:

Compared to refueling, ET recharging generally takes a long time and therefore significantly affects drivers' profit. Reasonable recharging decisions will increase the efficiency of finding and delivering passengers, thereby improving the net revenue of ET drivers over the next several hours rather than focusing on immediate rewards. However, it is difficult to measure the potential profit of a recharging action with the ENR because the ENR converges when many steps are considered [31]. Consequently, the edge weight of the recharging action is not comparable with the weights of the cruising edge. We proposed a simple strategy, named *recharging strategy*, to align the long-term profit of recharging actions with ABT. In this strategy, after all the weights of an ABT are estimated, the weights of the edges that represent recharging actions are adjusted.

Assume that the expected accumulated net revenue (EANR) of an ET driver in state s in the next $\Delta T > 0$ hours is $R_{\Delta T}(s)$. If the driver takes a recharging action $a_r \in A_r$ in the current state s, we assume that the EANR changes into $R_{\Delta T}(s, a_r)$. Using the learned taxi knowledge, the Monte Carlo method is used to estimate the EANR.

Clearly, in the following ΔT hours, if the recharging action tends to gain more EANR, as (10) shows, the ET driver should go to recharge.

$$\max_{a_r \in A_r} \{R_{\Delta T}(s, a_r)\} > R_{\Delta T}(s) \tag{10}$$

For a downward edge (s, a_r) representing a recharging action a_r, we denote its weight as $R_{\Delta T}(s, a_r)$. Similarly, $R_{\Delta T}(s)$ of each node is recorded to calculate whether (10) holds. Fig. 6 shows a node with four child edges, where the weights of the two edges on the left are replaced with the corresponding EANR. An additional EANR of the node is also recorded.

After an ABT is built and the corresponding EANRs are estimated, in the selection stage, if an edge that represents a recharging action has the maximum EANR among all child edges, in other words, (10) holds, it will be selected; otherwise, we select a cruising edge with the maximum ENR. Eventually, considering the long-term profit of the recharging, optimal sequential actions are obtained from the root to a leaf.

Considering the spatial distribution of charging stations, two typical charging station candidate sets are shown as follows:

- *NCS*: Only the nearest charging station to the current ET location is considered. The NCS minimizes the detour distance to reach a charging station.
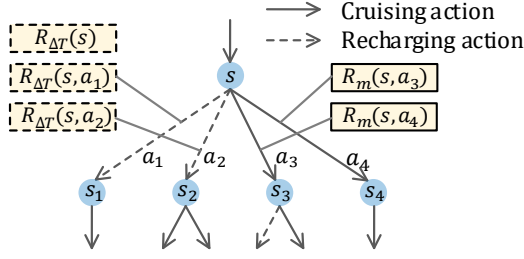
Fig. 6. Adjusting the weights of some child edges of a node.

- *ACS*: All charging stations are considered. This option may increase the detour cost but offers the opportunity to find the best recharging action.

3) Speed-up strategy:

Suppose that the number of feasible actions in a particular state is $\leq N$ and that the number of reachable passenger destinations is $\leq M = |I|$. Given the state s of an ET and the fixed length $n \geq 1$, estimating all the weights of the corresponding ABT has a computational complexity of $O((N + NM)^n)$, as Proof 1 shows.

Proof 1: Assume that the number of all possible actions in $1 \leq m \leq n$ steps is $f(m)$. If $n = 1$, $f(1)$ is $\leq N$. Clearly, the number of possible states after m steps is $\leq f(m)(M + 1)$; thus, the number of possible actions in $m + 1$ steps $f(m + 1)$ is $\leq f(m)(M + 1)N \leq f(1)(M + 1)^m N^m$. Therefore, the total number of actions within n steps is $\leq N \frac{1 - N^n (M+1)^n}{1 - N(M+1)}$. Because the ENR of each state-action pair is calculated recursively only once, the computational complexity of estimating the weights is $O((N + NM)^n)$.

Intuitively, the full computation process would require excessive computational resources; therefore, it would be difficult to support real-time route recommendation in this way. Here, we instead employ a lookup table to store the maximum ENRs of possible state-action pairs, that is, the weights of the edges in the ABTs, to increase the computing efficiency.

Note that a state-action pair $\langle s, a \rangle$ consists of five elements, $\langle i, e, t, j, \theta \rangle$. As the origin and destination of a trip and the recharging duration are fixed by $i \to j$ and θ, the maximum ENR of the state-action pair depends only on the remaining battery capacity e and the current time t.

Hence, to reduce the size of the lookup table, a sparse state space is used by dividing the remaining battery capacity and the time into larger intervals, e.g., 1 kWh and 5 minutes, respectively. The maximum ENR in one step, (2), can be rapidly calculated using the knowledge learned from taxi GPS trajectories. Furthermore, according to (3), the calculation of the maximum ENR over $m > 1$ steps is based on the estimated maximum ENR after $m - 1$ steps. Thus, the calculation of the maximum ENR is conducted step by step from 1 to n.

By querying the lookup table, the computational complexity of obtaining the weights in an ABT decreases to $O(N^n)$, as shown in Proof 2.

Proof 2: For an ABT with a height of n, the total number of its edges is $\leq N \frac{1 - N^n}{1 - N}$. Because each weight is obtained from the lookup table only once, the computational complexity is $O(N^n)$.

C. Online Recommendation Process

The online module processes real-time requests from ET drivers and calculates and recommends optimal cruising or recharging actions. After dropping off passengers or leaving a charging station, the on-board device in an ET will automatically send the current location i and the remaining battery capacity e to the recommendation server. Then, the server will build an ABT in a real-time manner. The edge weights of the ABT are associated with querying the lookup table and are adjusted when the recharging strategy is performed. Next, the best sequential actions will be found by using the ABT. The server matches the optimal sequential actions with the real-world road network to generate a practical route for the driver, which is then sent back to the on-board device in the ET. As shown in Fig. 7, the recommendation system provides the ET driver with a practical driving route, which is shown on the screen and the on-board device. Eventually, the ET driver can follow the recommended route to hunt for clients and schedule recharging.
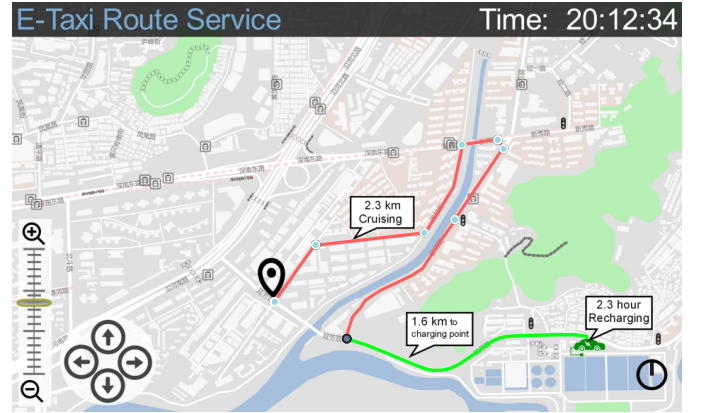

Fig. 7. The user interface illustration of the proposed system.

IV. EXPERIMENT AND RESULTS

A. Experimental Setup

An experimental ET route recommendation system was developed for ETs in Shenzhen, China. Three real-world datasets were used:

- **Road network**: The road network of Shenzhen contains 6,224 road segments and 5,724 junctions.
- **Taxi trajectories**: The GPS trajectories of 16,416 gasoline taxies were collected in June 2016. The collected data include the location, time, driving speed, and occupied status.
- **Charging stations**: This dataset includes 78 fast charging stations for ETs.

A simulation was conducted to evaluate the net revenue of ETs adopting this system. The simulated ETs were randomly distributed in the city and followed the recommended actions. Because the most popular ET model in Shenzhen is BYD E6, the following settings were adopted for the ETs: battery capacity = 60 kWh and driving range = 294 km. The battery energy consumption was modeled by a function that is linear in the driving distance [6]. The charging rate α was set to

TABLE I
STATISTICS FOR ETs AND GASOLINE TAXIES

	ET	Gasoline taxi
Working time (h)	18.02	17.40
Recharging time (h)	2.69	-
Total travel distance (km)	427.40	402.59
Total fare distance (km)	311.06	240.41

30 kW. The electricity price β was set to 0.8 CNY/kWh in accordance with the current cost of electricity. There are several parameters to be set in the proposed framework:

- **Number of sequential actions** n. We first set it to 5 and evaluated the trade-off with the ENR.
- **Recharging duration** ΔT. We set it to 2 hours.
- **Recharging stations**. We set all charging stations as candidates. In other words, the ACS strategy was adopted.

The extracted trips were used to simulate real-world taxi travel. The performance of the presented framework, including the average daily net revenue, recharging rhythm and efficiency, was evaluated. Three baseline methods were implemented for comparison:

- **Random**: one feasible route is randomly selected and recommended to each ET.
- **TaxiExp**: the optimal taxi travel routes based on the travel experience of gasoline taxi drivers are recommended.
- **MaxProb**: routes that maximize the probability of picking up passengers are recommended [23].

B. Results for the Net Revenue of the ETs

Tab. I reports the statistics of the trips of ETs adopting the recommendation. It suggests that the ETs will travel an average of 427.40 km in 18.02 hours per day, with a fare distance of 311.06 km (72.8%) for delivering passengers. Meanwhile, it will cost 2.69 hours to recharge. By comparison, the gasoline taxies travel an average of 402.59 km in a day, only 240.41 km (59.7%) of which is with passengers. Therefore, the proposed recommendation approach significantly improves the net revenue of the ETs with fewer cruising trips, thus promoting the viability of ETs.

C. Comparison with Gasoline Taxies

Fig. 8 presents the comparison with gasoline taxies. The results indicate that the daily net revenue of ET drivers mainly varies between approximately 550 and 950 CNY, with an average of 754 CNY. The daily net revenue of the ETs is greater than that of most of the gasoline taxies. The average daily net revenue of the ETs outperforms that of 76.2% of the gasoline taxi drivers. However, the best gasoline driver can earn more than 900 CNY, more than the net revenue of the best ET driver. This is due to the long charging time of ETs, which limits the full potential profit of ET drivers.

D. Comparison with Baseline Methods

Fig. 9 presents the comparison with the baseline methods. The results demonstrate that the proposed system outperforms all baseline methods. The naive *Random* method results in the lowest daily net revenue for ET drivers, with a mean value of 466 CNY. The *TaxiExp* method shows only slightly better performance, with an average daily net revenue of 471 CNY. The *MaxProb* method offers the second-largest average daily net revenue for ET drivers, namely, 665 CNY.

E. The Coordination of ET Actions

Fig. 10 displays the recharging rhythm of the ETs. Few ETs recharge during the morning (7:00-9:00 AM) and evening (5:00-7:00 PM) peak hours. A large proportion of ETs recharge from 4:00-5:00 PM. It is reasonable for the ET drivers to recharge and reserve more battery capacity for the evening peak. Most EVs are recommended to recharge at night (10:00 PM to 4:00 AM), which is asynchronous with the rhythm of the taxi demand.

The red line in Fig. 10 further displays the net revenue of ETs in each hour. Three peaks of the net revenue appear at the periods 5:00-7:00 AM, 10:00-12:00 AM, and 5:00-6:00 PM. Because of the relatively long-term vision of the recharging strategy, more ETs are recommended to recharge before the taxi demand peak in order to produce more net revenue.

F. Effect of the Recharging Strategy and the ET Setting

Both the available charging stations and the ET setting will impact the performance of ET drivers. Fig. 11a compares the results of the ACS and NCS strategies. By choosing the best charging stations and recharging duration among all options, the proposed system can achieve higher daily net revenue for ET drivers by improving from 692 CNY (NCS) to 754 CNY (ACS). This is because more charging stations will increase the long-term trade-off between recharging and possible passenger delivery.

We evaluated the effect of the battery capacity on the recommendation. Fig. 11b illustrates the results obtained with different ET battery capacities. By increasing the driving range from approximately 200 km to 300 km, the average daily net revenue will be increased from 690 CNY to 754 CNY. However, for the same increase of 20 kWh (or approximately 100 km in driving range), a capacity increase from 60 kWh to 80 kWh produces an improvement in the average daily net revenue of only 3 CNY.

We also varied the charging rate to evaluate its impact on the ET performance. Fig. 11c displays the results. When faster charging points with a charging speed of 60 kW are adopted, the daily net revenue will be improved from 754 to 784 CNY. On the other hand, a lower charging rate (30 kW) will decrease the average daily net revenue by 116 CNY because more time must be spent recharging.

G. Sensitivity Analysis

1) The number of recommended actions:

The number of consecutive actions n represents the long-term vision of the route recommendation. Fig. 12a shows the average daily revenue of ETs with different n. This figure suggests that ET drivers will earn 625 CNY per day on average, with only one action considered. By increasing n from 1 to 5, the average net revenue per day increases to 747 CNY. However, as the more consecutive actions are
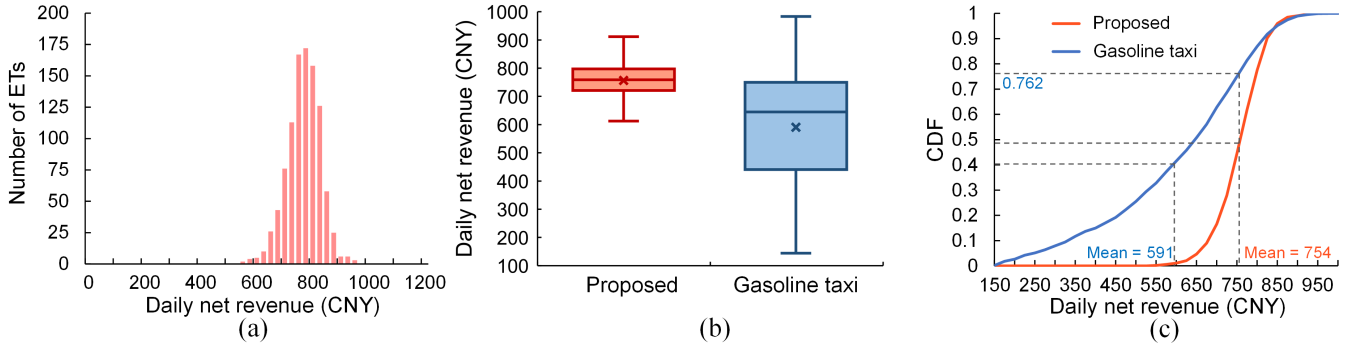
Fig. 8. Comparison with gasoline taxies: (a) Daily net revenue of ETs. (b) Comparison of daily net revenue. (c) Comparison of the cumulative distribution probability of the daily net revenue.
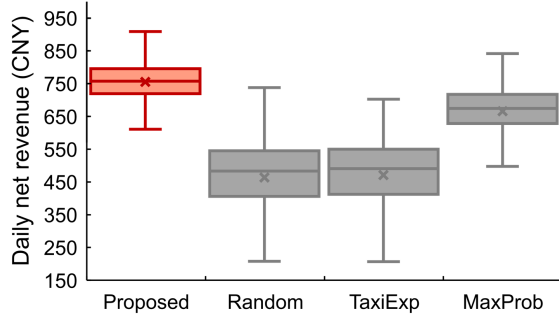
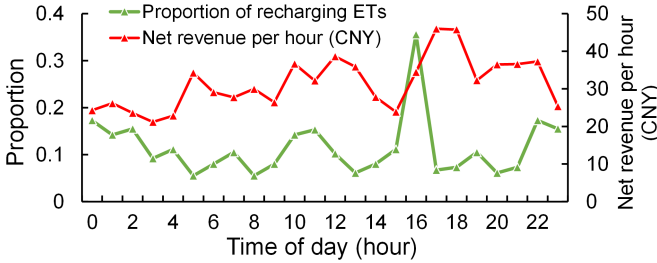

Fig. 9. Comparison with baseline methods.

Fig. 10. Coordination of travel and recharging.

considered, the increasing trend slowed, where the increasing rate of the average daily net revenue decreased from 13.1% to less than 0.1% as n grows from 2 to 8.

2) The duration of the recharging strategy:

The EANR was estimated to reflect the potential revenue of recharging actions in the following ΔT hours, reported in Fig. 12b. If ΔT is too small, the estimated EANR cannot uncover the potential net income of recharging actions in a longer period. Fig. 12b shows that a short duration, such as $\Delta T = 1h$, will produce a small daily net revenue, on average 725 CNY per day. Intuitively, a larger ΔT could generate a more precise estimation of the ENR. However, if ΔT is too large, more potential influence of other actions will be fused with the estimated EANR, thus affecting the coordination of recharging. It can be measured from Fig. 12b that a long duration, $\Delta T \geq 3h$, will generate less daily net revenue, less than 703 CNY per day. Hence, ΔT may have an optimal value at approximately 2 hours in this study.

3) Number of ETs served by the system:

Recommending the same route to many ET drivers may lead to the over-provision of taxies on certain roads. To balance taxi service and taxi demand, we adopted the road capacity strategy proposed by [12] to constrain the ETs on each road. If the number of ETs on a road exceeds a certain threshold, the system will not recommend that any ET enter this road. Under this strategy, we evaluated the effect of the ET account. Fig. 12c shows the effect on daily net revenue with different numbers of ETs. This figure demonstrates that more ETs relying on the recommendation system will reduce the net revenue. For example, with 4,000 ETs in the recommendation system, the average daily net revenue of ETs decreases from 754 CNY to 641 CNY. However, even in the extreme case (8,000 ETs), the average result for ETs that adopt our system still outperforms 50% of real-world gasoline taxi drivers.

V. CONCLUSION

Leveraging massive-scale taxi GPS trajectory data, this study presents a comprehensive real-time route recommendation system for ET drivers that integrates the cruising on the road and the recharging at stations decision. Taxi travel knowledge is learned from raw GPS trajectories of gasoline taxies and used to estimate the ENRs of sequential actions of ET drivers. An ABT is built and used to recommend the route choice of ET drivers with an effective speeding-up strategy. An online prototype system has been developed for high-efficiency real-time route recommendation. An experiment in Shenzhen demonstrates the effectiveness and efficiency of the developed system. The results show that the average daily net revenue of ET drivers using the developed system outperforms 76.2% of gasoline taxi drivers. The presented approach can not only increase the revenue of ET drivers in the short term, but also improve ETs viability in the long term. The proposed method is applicable to plug-in hybrid EVs and hydrogen EVs. In the future work, taxi ranks and real-time traffic information may be considered. The correlation between waiting time and net revenue will be explored in real applications.

REFERENCES

[1] S. A. Bagloee, M. Tavana, M. Asadi, and T. Oliver, "Autonomous vehicles: challenges, opportunities, and future implications for transportation policies," *J. of Modern Transp.*, vol. 24, no. 4, pp. 284–303, 2016.
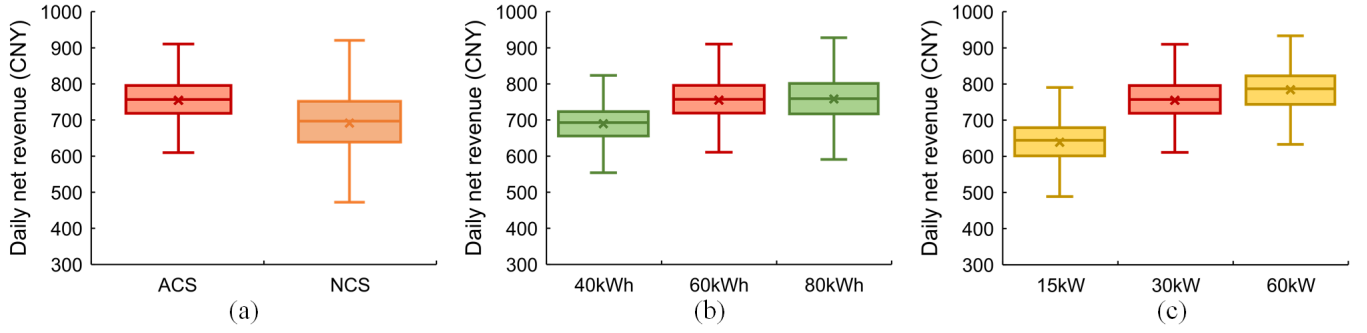
Fig. 11. Effect of the recharging strategy and ET settings: (a) ACS vs. NCS. (b) Battery capacity. (c) Charging rate.
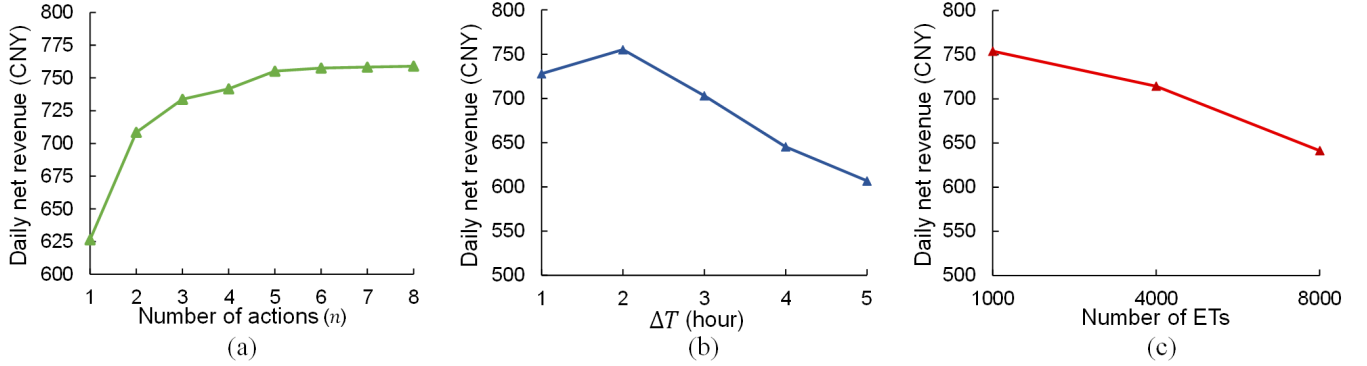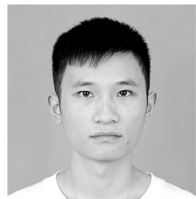

Fig. 12. Effect of the parameter settings: (a) The recommended number of actions, n. (b) The duration of the recharging strategy, ΔT. (c) The number of ETs.

[2] Z. Zhou, C. Gao, X. Chen, Z. Yan, S. Mumtaz, and J. Rodriguez, "Social big data based content dissemination in internet of vehicles," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 768–777, 2018.

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

[4] W. Kempton, "Electric vehicles: Driving range," *Nat. Energy*, vol. 1, no. 9, p. 16131, 2016.

[5] J. Jung and J. Y. J. Chow, "Effects of charging infrastructure and non-electric taxi competition on electric taxi adoption incentives in new york city," *Transp. Res. Record*, vol. 2673, no. 4, pp. 262 – 274, 2019.

[6] W. Tu, Q. Li, Z. Fang, S.-l. Shaw, B. Zhou, and X. Chang, "Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach," *Transp. Res. Pt. C-Emerg. Technol.*, vol. 65, pp. 172–189, 2016.

[7] J. Yang, J. Dong, Z. Lin, and L. Hu, "Predicting market potential and environmental benefits of deploying electric taxis in nanjing, china," *Transport. Res. Part D-Transport. Environ.*, vol. 49, pp. 68–81, 2016.

[8] M. Baum, J. Dibbelt, A. Gemsa, D. Wagner, and T. Zündorf, "Shortest feasible paths with charging stops for battery electric vehicles," in *Proc. 23rd SIGSPATIAL Int. Conf. Advances in GIS*, ser. SIGSPATIAL '15. New York, NY, USA: ACM, 2015, pp. 44:1–44:10.

[9] M. Reinthaler, J. Asamer, H. Koller, and M. Litzlbauer, "Utilizing mobility data to facilitate the introduction of e-taxis in vienna: Feasibility study of a decision support system for the introduction of battery electric vehicles as taxis," in *Int. Conf. Connected Vehicles and Expo (ICCVE)*, Nov 2014, pp. 516–517.

[10] J. Yang, J. Dong, and L. Hu, "Design government incentive schemes for promoting electric taxis in china," *Energy Policy*, vol. 115, pp. 1–11, 2018.

[11] L. Hu, J. Dong, Z. Lin, and J. Yang, "Analyzing battery electric vehicle feasibility from taxi travel patterns: The case study of new york city, usa," *Transp. Res. Pt. C-Emerg. Technol.*, vol. 87, pp. 91 – 104, 2018.

[12] C. Tseng, S. C. Chau, and X. Liu, "Improving viability of electric taxis by taxi service strategy optimization: A big data study of new york city," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 817–829, March 2019.

[13] Z. Yang, T. Guo, P. You, Y. Hou, and S. J. Qin, "Distributed approach for temporal-spatial charging coordination of plug-in electric taxi fleet," *IEEE Trans. Ind. Informat.*, 2018.

[14] H. A. Chaudhari, J. W. Byers, and E. Terzi, "Putting data in the driver's seat: Optimizing earnings for on-demand ride-hailing," in *Proc. 11th ACM Int. Conf. Web Search Data Mining*, ser. WSDM '18. New York, NY, USA: ACM, 2018, p. 90–98.

[15] C. Chen, C. Liao, X. Xie, Y. Wang, and J. Zhao, "Trip2vec: a deep embedding approach for clustering and profiling taxi trip purposes," *Pers. Ubiquitous Comput.*, vol. 23, no. 1, pp. 53–66, 2019.

[16] R. Wang, C.-Y. Chow, Y. Lyu, V. C. Lee, S. Kwong, Y. Li, and J. Zeng, "Taxirec: recommending road clusters to taxi drivers using ranking-based extreme learning machines," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 3, pp. 585–598, 2017.

[17] N. Garg and S. Ranu, "Route recommendations for idle taxi drivers: Find me the shortest route to a customer!" in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, ser. KDD '18. New York, NY, USA: ACM, 2018, p. 1425–1434.

[18] M. Han, P. Senellart, S. Bressan, and H. Wu, "Routing an autonomous taxi with reinforcement learning," in *Proc. 25th ACM Int. Conf. Inform. Knowl. Management*, ser. CIKM '16. New York, NY, USA: ACM, 2016, p. 2421–2424.

[19] C. F. Yuen, A. P. Singh, S. Goyal, S. Ranu, and A. Bagchi, "Beyond shortest paths: Route recommendations for ride-sharing," in *The World Wide Web Conf.*, ser. WWW '19. New York, NY, USA: ACM, 2019, p. 2258–2269.

[20] Y. Cheng, Q. Zhou, and Y. Lai, "Taxic: A taxi route recommendation method based on urban traffic charge heat map," in *Service-Oriented Comput. – ICSOC 2018 Workshops*. Cham: Springer, 2019, pp. 322–334.

[21] D. Kumar, H. Wu, S. Rajasegarar, C. Leckie, S. Krishnaswamy, and M. Palaniswami, "Fast and scalable big data trajectory clustering for understanding urban mobility," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 11, pp. 3709–3722, 2018.

[22] X. Kong, F. Xia, J. Wang, A. Rahim, and S. K. Das, "Time-location-relationship combined service recommendation based on taxi trajectory data," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1202–1212, 2017.

[23] N. J. Yuan, Y. Zheng, L. Zhang, and X. Xie, "T-finder: A recommender system for finding passengers and vacant taxis," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 10, pp. 2390–2403, 2012.
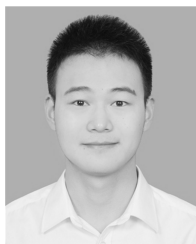
[24] B. Qu, W. Yang, G. Cui, and X. Wang, "Profitable taxi travel route recommendation based on big taxi trajectory data," *IEEE Trans. Intell. Transp. Syst.*, 2019.

[25] 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 Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3098–3109, 2016.

[26] Z. Yang, L. Sun, J. Chen, Q. Yang, X. Chen, and K. Xing, "Profit maximization for plug-in electric taxi with uncertain future electricity prices," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3058–3068, 2014.

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

[28] B. Li, D. Zhang, L. Sun, C. Chen, S. Li, G. Qi, and Q. Yang, "Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset," in *IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, March 2011, pp. 63–68.

[29] C. Silva, P. M. d'Orey, and A. Aguiar, "Interpreting traffic congestion using fundamental diagrams and probabilistic graphical modeling," in *IEEE Int. Conf. Data Mining Workshops*, Nov 2018, pp. 580–587.

[30] J. Xie, W. Tu, Q. Li, X. Chang, C. Ma, Z. Li, and L. Huang, "A parallel map-matching approach for large volume floating car stream data," *Geom. Inf. Sci. WHU*, vol. 42, no. 5, pp. 697–703, 2017.

[31] M. Qu, H. Zhu, J. Liu, G. Liu, and H. Xiong, "A cost-effective recommender system for taxi drivers," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, ser. KDD '14. New York, NY, USA: ACM, 2014, p. 45–54.

Yang Xu is currently an Assistant Professor in the Department of Land Surveying and Geo-Informatics at The Hong Kong Polytechnic University. His research lies at the intersection of GIScience, Transportation, and Urban Informatics. Leveraging big data, his work focuses on the quantification and modeling of human activities in cities, aiming to reveal their linkage with urban and technological developments, and their impact on future economic, social and transportation systems. Before joining PolyU, He worked as a joint postdoctoral associate at the MIT Senseable City Lab and the Singapore-MIT Alliance for Research and Technology.

Jincai Huang received the Ph.D. degree from Central South University, Changsha, China. His is currently a postdoctoral fellow in the College of Civil and Transportation Engineering, Shenzhen University, China. He is also a member of China Computer Federation and the reviewer of manuscripts for International Journal of Geographical Information Science, Concurrency and Computation-Practice & Experience and IEEE Access.

Wei Tu (M'19) received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2013. He is currently an Associate Professor with Shenzhen Key Laboratory of Spatial Smart Sensing and Service and Research Institute of Smart Cities, Department of Urban Informatics, Shenzhen University, China. He is also a visiting scholar at Senseable City Laboratory, Massachusetts Institute of Technology. His research interests include urban informatics, big data driven human activity and mobility, and trajectory analytic.

Min Deng received the Ph.D. degrees from Wuhan University, Wuhan, China, in 2003, and from the Asian Institute of Technology, Khlong Nueng, Thailand, in 2004. He is currently the Head of the Department of Geo-Information and a Doctoral Supervisor with the School of Geosciences and Info-Physics, Central South University, Changsha, China. His current research interests include geospatial data update, spatio-temporal data mining, analysis and modeling. Dr. Deng has hosted numerous major projects including a Key Project of National Natural Science Foundation of China.

Ke Mai received his M.S. degree in surveying mapping engineering from Wuhan University, Wuhan, China, and B.S. degree from Sun Yat-sen University, Guangzhou, China, in 2019 and 2017, respectively. He is currently a research assistant in Laboratory of Spatial Smart Sensing and Service and Department of Urban Informatics, Shenzhen University. His current research interests focus on spatio-temporal data mining and optimization.

Long Chen (M'13-SM'19) received the Ph.D. degree in signal and information processing from Wuhan University, Wuhan, China, in 2013. From October 2010 to November 2012, he was co-trained Ph.D. Student at the National University of Singapore, Singapore. He is currently an Associate Professor with the School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China. His areas of interest include autonomous driving, robotics, and artificial intelligence. He serves as an Associate Editor for the IEEE Transactions on Intelligent Transportation Systems and the IEEE Technical Committee on Cyber-Physical Systems Newsletter and a Guest Editor for the IEEE Transactions on Intelligent Vehicles.

Yatao Zhang received the B.S. degree from Sun Yat-sen University, Guangzhou, China. He is currently working toward the M.S. degree in GIS in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China. His major research interests include multi-source spatiotemporal data fusion and urban sensing.

Qingquan Li received the M.S. degree in survey engineering and the Ph.D degree in photogrammetry and remote sensing from Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1988 and 1998, respectively. He is the President of Shenzhen University, Shenzhen, China, and the Director of the Guangdong Key Laboratory of Urban Informatics, Shenzhen University. His research interests include photogrammetry, remote sensing, mobile mapping, and intelligent transportation systems.