Large-Scale Intelligent Taxicab Scheduling: A Distributed and Future-Aware Approach

Yang Wang, *Member, IEEE,* Zhengyang Zhou, *Student Member, IEEE,* Kai Liu, Xike Xie, *Member, IEEE,* and Wenhua Li

Abstract—Intelligent taxicab scheduling systems on smartphones continue to gain popularity as they offer prominent conveniences for urban travelling as well as increase potential profits for taxicab drivers, hence inject prosperity and vitality into intelligent transportation and urban business. Existing scheduling approaches usually fall into biases and myopia due to their single target perspective of satisfying the immediate order acceptance, or maximizing the global business success rates. However, the highly complex spatiotemporal dependencies among multiple factors and the efficiency bottleneck of massive order flows make the taxicab scheduling issue still challenging. To this end, in our paper, we propose an integrated scheduling algorithm with both future-aware and context-aware mechanisms. In particular, we first present a generalized graph-based framework which aims to capture traffic dependencies, providing precise and quantified supply-demand prediction in taxicab scheduling. Then, we develop a measurement of regional supply-demand context to perform cooperative and distributed scheduling. To tackle the efficiency bottleneck in massive order flow scenario, we correspondingly design a model to learn business patterns with considering benefits from both drivers and passengers, and further promote service delivery rates via a novel bi-incentive strategy. Extensive numerical studies illustrate the remarkable significance of our method, in terms of both service delivery rates and global driver revenues. The promising results and brand-new perspectives enable our algorithm to be a paradigm in general spatiotemporal scheduling tasks.

Index Terms—Intelligent taxicab scheduling; future-aware; intelligent transportation; context-aware.

✦

1 INTRODUCTION

Recent advances in wireless sensors and networking technologies in mobile vehicles, e.g., 4G/5G, Wi-Fi, RFID, and GPS, give prominence to the business opportunity and prosperity for the market of taxicab industries [17], [30], [38], [39]. According to the report of Analysis International [13], online taxicab calling services such as Didi [1] and Kuaidi [2] have more than 150 million active users in China, and are able to collect about 12 million taxicab trips every single day. Meanwhile, public controversy arises due to the safety concerns on existing commercial taxi service applications, which request taxi drivers to get orders actively with actual operations on mobile terminals during driving. The operating process for getting active orders, as shown by the example of Figure 1, can be a source of distraction for taxi drivers and therefore cause an ill effect on public traffic safety. More, due to the competition of taxicab drivers, such systems lead to inefficiency in scheduling and on the contrary, automatic order dispatch and taxicab scheduling system will directly contribute to a more than 10% improvement on the service delivery rate [29].

Existing data-driven approaches related to taxicab scheduling are mostly on finding the optimized routes for taxicab drivers [7], [14], [22], [33] in urban navigation, and on equilibrating the needs of both passengers and taxicab drivers [32], [34], [35]. Only a few address efficient scheduling problems of taxicabs, but with the optimization in terms of energy [16], [21], [37], profits [25], [36], and transportation QoS (Quality of Service) [4], [15], [20] in carpooling services. Most of these algorithms focus on processing historical trajectory records for the retrieval of taxicab mobility and passenger occupation patterns without considering dependencies among taxicabs and order-taxicab interactions. Advanced developments on taxicab order dispatching regard taxicab drivers as multiple agents in the road network and formulize a series of value functions to perform deep reinforcement learnings [29], [43]. However, such seemingly advanced techniques may degrade the interpretability of the model, hence increase the uncertainty of the scheduling results. In summary, existing solutions fail to consider time-varying taxicab demand and supply patterns as well as maximize the service delivery rates simultaneously, hence deteriorate the system performance in real scenarios involving large-scale order-taxicab pairs.

To this end, in this work, we propose an intelligent future-aware taxicab scheduling system, Taxicab Scheduling with Distributed and Future-Aware mechanisms (TS-DFA), to address the challenges of both interactive supply-demand effects and large-scale order flows during rush hours.

Specifically, we first forecast the future distributions of city-wide traffics and taxicab demands for subsequent scheduling by mining the spatiotemporal patterns of historical taxicabs and demands. Second, to help improve the efficiency of order dispatching, we develop a learning-based method to rank a series of relative values of all current orders to each driver. The main research challenges of our system include: i) the dependencies among taxicabs in different road segments and the dependencies between taxicabs and scheduling strategies; ii) the efficiency bottleneck if a large number of orders appear during rush hours; and iii) the algorithm-insolubility of scheduling multiple taxicabs

---

- *Prof. Y. Wang is the corresponding author.*
- *Y. Wang*, Z.Y. Zhou, K. Liu, X.K. Xie are with University of Science and Technology of China, Suzhou 215000, P.R.China.*
- *W.H. Li is with Traffic Administrative Agency of Suzhou Industry Park, Suzhou 215000, P.R.China.*

- *Email: {angyan*, xkxie}@ustc.edu.cn; xqgsh080504@sina.com; {zzy0929, liukai33}@mail.ustc.edu.cn.*
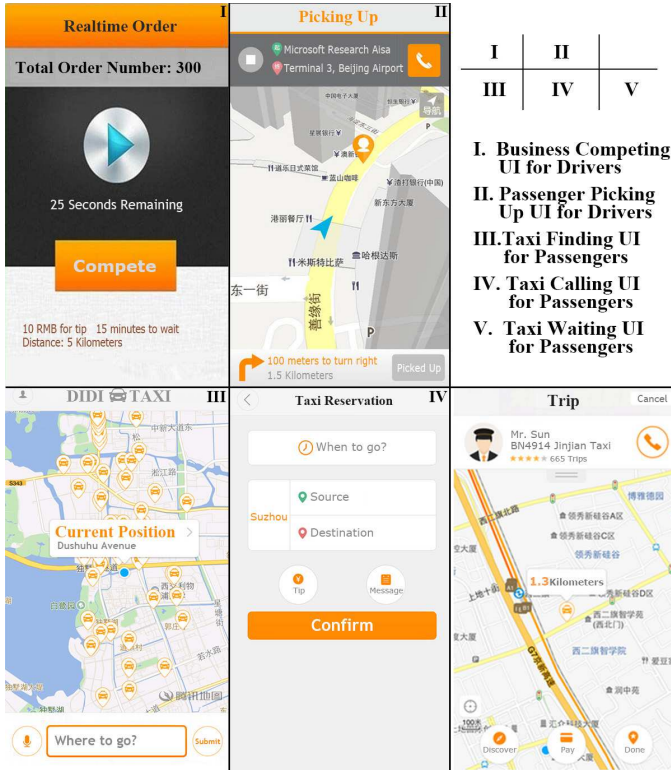
Fig. 1: Online taxicab Calling Apps (Didi as an example)

cooperatively to maximize the overall profits and service delivery rates simultaneously within polynomial time complexity. To our best knowledge, none of the existing taxicab scheduling systems address the above challenges systematically and comprehensively. In order to address these challenges, we propose a novel algorithm which contains two parts, offline learning and online scheduling. First, we build several spatiotemporal submodels to capture the time-varying traffic and business patterns. Specifically, by taking advantage of the region-wise spatial correlations among the road segments and temporal dependencies among time slots, we propose a graph-based approach to learn the dynamic taxi demand patterns and a regression-based method to estimate the order-taxicab value score with regard to each individual order-taxicab pair. Further, we introduce a context-aware feature named regional driving value (*RDV*), to model both taxicab supply and demand of each road segment. With these carefully-designed models, we are able to integrate the learned spatiotemporal demand-and-supply patterns with the real-time taxicab calling information for optimizing the service delivery rates as well as the total profits of taxicab drivers simultaneously. Next, in the online scheduling part, given a series of calling information and taxicab trajectories, we estimate the values for each order-taxicab pair and broadcast the appropriate orders to drivers after careful selections. Meanwhile, to alleviate the supply-demand imbalance in those taxicab-free places, we develop a bi-incentive strategy to stimulate both drivers and passengers in the case of long pickup distance and low-expected revenue orders. In our system, the tradeoff between vacant taxicabs and taxicab demands is equilibrated with the context-aware *RDV* and future-aware prediction by scheduling the idle taxicabs to most taxicab-needed urban regions. For evaluating the proposed TS-DFA, we compare it with the state-of-the-art solutions Cost-Effective Recommender System

(CERS) [25], Taxi Order Dispatch model with Combinatorial Optimization (TODCO) [41] and Large-scale Order Dispatch in on-demand Ride-Hailing Platform (LOD-RHP) [29] via two real-world datasets including Suzhou and New York City (NYC). The results show that our TS-DFA algorithm increases the average daily profit of taxicab by 10.2 (RMB) and 16.5 (USD) in Suzhou and NYC respectively, compared with the best baseline. Moreover, our TS-DFA achieves maximum service delivery rates and relatively less waiting time for online taxi calling passengers on both two datasets.

The main contributions of this work can be summarized as follows:

- To our best knowledge, this is the first work to optimize taxicab service delivery rates and global taxicab revenues simultaneously by considering the real-world massive taxicab business flows in rush hours as well as the complex multiple dependencies among the factors of taxicab supplies, demands and scheduling strategies.
- To jointly consider the mutual influences among taxicab peers and order-taxicabs in cooperative taxicab scheduling, we design the novel context-aware regional driving values and order-taxicab value scores to measure the road network taxicab supply-demand patterns as well as potential revenues between order-taxicab pairs respectively.
- We evaluate the proposed algorithm on two real-world datasets (Suzhou and NYC). Compared to the state-of-the-art machine learning-based method, TS-DFA can still enhance the global revenues and overall service delivery rates by up to 2.43%-6.90% and 3.93%-4.85%, respectively. The cross-validated experiments demonstrate the absolute effectiveness of our TS-DFA in both two cities.

The rest of this paper is organized as follows. Section 2 presents related works. Section 3 provides the preliminaries and formulates the problem. Section 4 analyzes the time-varying patterns of traffic flows and then models spatiotemporal dependencies among taxicab demands and moving taxicabs, respectively. Section 5 describes several submodels as well as our intelligent future-aware taxicab scheduling algorithm. Section 6 evaluates the proposal and Section 7 concludes the paper.

2 RELATED WORK

Plenty of efforts have been invested on tasks related to taxicab scheduling. Taxi route recommendation is one of the widely studied problems targeting maximizing individual taxicab revenues and the probability of picking up passengers. For example, some recent algorithms focus on recommending optimized paths to taxicabs in order to enhance energy efficiency. [8] first extracts driving patterns of successful taxicab drivers in terms of revenue against fuel usage, then clusters pick-up locations for these drivers in a certain time period. It proposes two algorithms based on a candidate route evaluation function as well as a mobile recommender system to provide pick-up points for taxicab drivers. [36] proposes a new mathematical concept "cruising graph" by using intersections of road networks as vertices, road segments as edges and expected numbers of arrival passengers in road segments as weights. Based on this graph, the proposed pCruise algorithm selects the shortest cruising routes for taxicabs to find passengers, and recommends optimized routes for taxicab drivers to pick up passengers with the maximum profits. Based

on fundamental works above, the online scheduling task, which schedules taxicabs to order requests, is further studied for diversity objectives. [21] proposes a new algorithm to search and schedule candidate taxicabs to satisfy potential taxicab requests from passengers with minimum additional travel distance incurred, so as to save fuel consumptions. In [37], a carpooling service system named *coRide* is presented to reduce total mileage and achieve fewer energy consumptions. After formulating an NP-hard route calculation problem under different actual constrains, a linear programming optimized algorithm and a 2-approximation polynomial complexity algorithm are proposed with a win-win fare model for both taxicab drivers and passengers. Also, in terms of learning-based methods in mobile computing and traffic pattern learning, many existing studies contribute to promising results [10], [23], [24]. And [6] designs a dynamic programming algorithm to optimally solve the unlimited expansion of the size of the dispatch region while given a specific taxicab request by taking the benefits of both drivers and passengers into consideration. With the increasing prosperity of machine learning, few newest works invest on addressing the large-scale order-dispatching problem by formulating the drivers as multiple agents based on reinforcement learning [29], [43], which may lack the model interpretability.

In conclusion, all the above researches either focus on optimizing the routes for both cruising and on-business drivers, or provide order-taxicab dispatch methods by targeting one-side overall profits. Directly ignoring spatiotemporal dependencies may significantly influence the efficiency of scheduling. In contrast, our algorithm allows more intelligent scheduling factors for multiple taxicabs by taking the predicted future demands, the influences of previous scheduling strategies and spatiotemporal correlations into account. Further, although some previous works have advanced that the traffic recommender system should be traffic sensitive, none of them has jointly considered both the supply-demand balance as well as the massive order flows simultaneously, which is of great significance in the overall performance and stability of scheduling solutions.

3 Background and Problem Description

3.1 Preliminaries and Basic Concepts

We hereby present the preliminaries of our algorithm, including problem assumptions and formalizations of several important concepts.

Assumptions. For the intelligent taxicab scheduling, we assume that a taxicab follows the scheduling strategy once it accepts the task assignment. There is a threshold on the waiting time for both passengers and taxi drivers. The passenger would opt for other taxis and the driver would transmit its state from busy to free, if the waiting threshold is met. This means that if one taxicab is assigned with an order, it will immediately transmit its state from free to busy. And if it rejects this order within the allowable time, the state will then transmit to free. Further, we do not assume any specific mobility patterns for unoccupied taxicabs, they can either drive around or wait at a position for new passengers.

Concepts. Given the intersections and road segments, the road network can be denoted as a directed *Road Network Graph* $G(V, E)$. Here the vertex set V means all intersections and the edge set E represents road segments, i.e., $\forall i, j \in [1, |V|]$, $I_i \in V$ and directed road segment r_{ij} from intersection I_i to I_j must have $r_{ij} \in E$. Note that r_{ij} and r_{ji} are the road segments between the same two intersections with opposite directions.

The entire taxicab fleet running in the road network can be denoted as: $TF = \{n_1, n_2, \ldots, n_{|TF|}\}$ where n_i is the ith taxicab. The historical taxicab trajectory set is denoted as $\mathcal{TR}$. For each taxicab n_i in TF, we can write its trajectory in one specific day d as: $TR_{n_i}(d) = \{start = r_{k_1 k_2}(t_1), r_{k_2 k_3}(t_2), \ldots, r_{k_{l-1} k_l}(t_{l-1}) = end\}$, which is an element in set $\mathcal{TR}$. Here the taxicab starts its journey at time t_1 and road segment $r_{k_1 k_2}$, then ends this journey at time t_{l-1} and road segment $r_{k_{l-1} k_l}$.

To better solve the taxicab scheduling task, we introduce the cost $f_c(r_{ij}, TP)$ and profit $f_p(r_{ij}, TP)$ for taxicabs running through road segment r_{ij} during time slot TP.

3.2 Primary Goal of Taxicab Scheduling

As discussed, the two primary goals of our distributed taxicab scheduling task are to help taxicab fleet maximize their total profits and improve the service delivery rates simultaneously by recommending them the optimized driving routes, i.e.,

Definition 1 (Taxicab Fleet Scheduling Problem).

Assume the taxicabs within our online scheduling system are denoted as a taxicab set $\mathcal{A}$. For one specific day d, the total profits of taxicabs and global service delivery rates are abbreviated as $f_p(d)$ and $R(d)$, which can be formalized by:

$$f(d) = \sum_{i=1}^{|\mathcal{A}|} f_p^{n_i}(d) \tag{1}$$

$$R(d) = \frac{N_a(d)}{N_o(d)} \tag{2}$$

where $f_p^{n_i}(d)$ denotes the profits of taxicab n_i on day d, $N_a(d)$ and $N_o(d)$ are the numbers of accepted orders and total online requests on day d, respectively. Given the historical trajectory set $\mathcal{TR}$ and the corresponding business records, our target is to find the optimized joint taxicab scheduling function $F(order, driver)$ for order dispatching and taxicab route recommendation by maximizing the profits $f_p(d)$ and global service delivery rates $R(d)$ at the same time, i.e.,

$$\{\mathcal{TR}, \mathcal{A}\} \xrightarrow{F(order, driver)} \text{Optm}\{R(d), f_p(d)\} \tag{3}$$

Here $\text{Optm}\{R(d), f_p(d)\}$ is the closely optimized scheduling scheme calculated by our algorithm.

4 Urban Taxicab System Analysis

In this paper, the task of large-scale taxicab scheduling is decomposed into several subtasks including demand prediction, order revenue calculation, order dispatch and supply-demand balancing, which are of great difficulty to be solved by deep neural networks in an end-to-end way. Therefore, in this section, we first model the general traffic conditions and then, analyze the urban taxicab system in terms of taxicab running patterns and settings of the urban taxicab system, for designing our intelligent TS-DFA algorithm.

4.1 Analysis of Road Traffic Conditions

Urban traffic conditions are able to help calculate the expected revenues of potential taxicab businesses, which guides recommending optimized routes to taxicabs and further scheduling. As the assignment and scheduling are determined by potential gains in the future, the foreseeability of traffic conditions plays an increasingly important role in scheduling tasks. In general, traffic
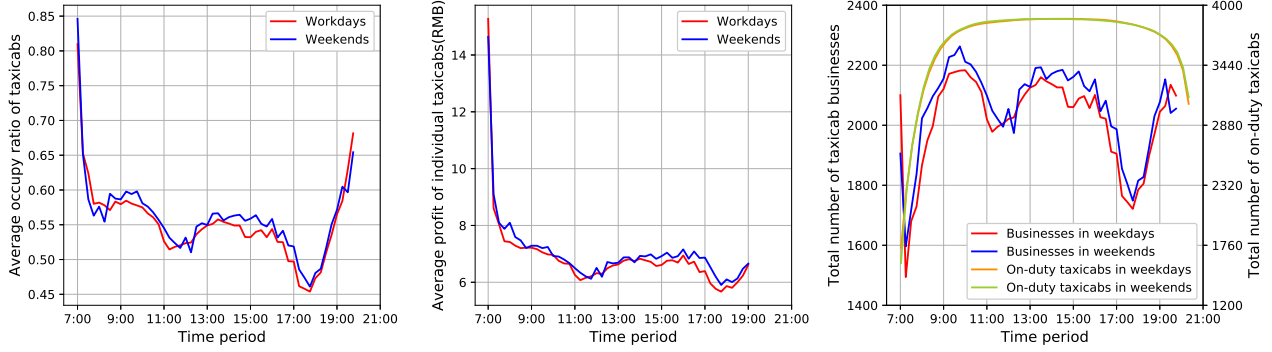
Fig. 2: Analysis of urban taxicab running during different time periods in March, 2012, Suzhou.

conditions can be depicted by speeds, traffic volumes and other dynamic elements in the road network. Here, to perform fine-grained predictions, we first divide one day into time slots of $\mathcal{T}$ minutes [1], and aggregate all traffic attributes into corresponding time slots. Second, for better capturing the road segment-wise traffic propagations, we extend our *Road Network Graph* $G(V, E)$ to $G(V, E, A)$. Then, we redefine the road segment set E as $r_i \in E$ where $1 \leq i \leq |E|$, and the matrix $A \in \mathbb{R}^{|E| \times |E|}$ is a linkage matrix. The element $a_{ij} \in A$ denotes the linkage between contiguous road segment r_i and r_j, namely, $a_{ij} = 1$ if and only if the terminated intersection of r_i should also be the initial intersection of r_j. The linkage matrix A enables the graph to carry richer information, hence help the modelling of traffic propagations [?]. Existing traffic-related studies demonstrate that a sequence of past observations of traffic conditions may imply many possible futures through mining their spatiotemporal correlations [31], [40], [44]. Hereby, we present a General Graph-based Traffic Forecasting framework (*GGTF*) which has the potential to foresee near future's traffic statuses via *Road Network Graph* $G(V, E, A)$, hence supports our future-aware intelligent taxicab scheduling. Given the historical observations of citywide traffic conditions $\mathcal{F}(\Delta t) \in \mathbb{R}^{|E| \times 1}(\Delta t = 1, 2, 3.., T)$ and linkages matrix A, we can predict the traffic conditions $\mathcal{F}(TP)$ of time slot TP. The related time slots for predicting $\mathcal{F}(TP)$ are in set $RT(TP)$ which consists of three temporal perspectives [31], [40], [45], hourly closeness (i.e. adjacent $\mathcal{P}$ time slots with TP), daily periodicity (i.e. the last $\mathcal{P}$ days of the same time slots with TP) and distant trend (i.e. the last $\mathcal{P}$ weeks of the same day time slots with TP) [2]. Considering the spatial correlations with regard to road segment-wise propagations as well as the temporal influences, we predict the traffic statuses by:

$$\mathcal{F}(TP) = \sum_{k \in RT(TP)} A * \mathcal{F}(k) \odot W_k^{\mathcal{F}} \qquad (4)$$

Here, $W_k^{\mathcal{F}} \in \mathbb{R}^{|E| \times 1}$ is the learnable traffic condition parameter which determines different spatiotemporal importances of each road segment to the final results, and $\odot$ denotes the element-wise product. Obviously, the general framework can be trained offline in

a back-propagation manner and the traffic conditions are predicted in a citywide way.

In this part, we aim to forecast the average driving speed of every road segment during different time periods, then estimate the driving time between two specific locations in the road network during a given time period. For road segment speed prediction, the $\mathcal{F}$ is substituted by the historical speed observations TV. Then the citywide expected speed in time slot TP can be derived as

$$Ex[TV(TP)] = \sum_{k \in RT(TP)} A * TV(k) \odot W_k^V \qquad (5)$$

where the expected speed in a specific road segment r_i is denoted as $Ex[tv_{r_i}(TP)]$. $W_k^V \in \mathbb{R}^{|E| \times 1}$ is the speed-related learnable parameter and the estimated time cost for running through r_i can be achieved by:

$$Ex[tc_{r_i}(TP)] = \frac{L_{r_i}}{E[tv_{r_i}(TP)]} \qquad (6)$$

where L_{r_i} denotes the length of road segment r_i. With the time cost of running through road segment r_i in time slot TP as its weight W^E of E, the *Road Network Graph* can also be modified as $G(V, E, W^E)(TP)$, which becomes an $E - weighted$ graph for time slot TP. Based on this graph, we calculate the optimized driving route from road segment r_i to r_j during time slot TP by:

$$OR[tc_{r_i \to r_j}(TP)] = Dijkstra[G(V, E, W^E)(TP), r_i, r_j] \qquad (7)$$

where *Dijkstra* means Dijkstra's shortest path algorithm [5]. The expected driving time between these two road segments r_i and r_j can be further calculated by the summation of each road segment along the optimized route, which is written as:

$$Ex[tc_{r_i \to r_j}(TP)] = \sum_{r \in OR[tc_{r_i \to r_j}(TP)]} Ex[tc_r(TP)] \qquad (8)$$

4.2 Analysis of Taxicab Running

The running of urban taxicabs usually exhibits obvious time-varying patterns [28], [42]. On this account, we select four different weeks in March, 2012 from Suzhou dataset, and average them into period-wise running statistics in Figure 2. The spatial distributions of taxicab businesses are illustrated in Figure 3.

As shown in the first subfigure of Figure 2, the average occupied ratios [3] during 7:00 a.m. to 7:30 a.m. are the highest, due to the high service demands and the low taxicab on-duty

---

1. The setting of $\mathcal{T}$ should balance the tradeoff between the accuracies and temporal granularity. In our implementation, we slice the temporal information into slots of 15 minutes. Notice that such a setting may be related to the results of scheduling but is orthogonal to the generalities of our proposals

2. The hyperparameter $\mathcal{P}$ in our paper is set as 3 according to common settings in previous studies [31], [40].

3. It refers to the rate of occupied taxi number to all on-duty taxicab number.

I An example of pick-up points on workday mornings
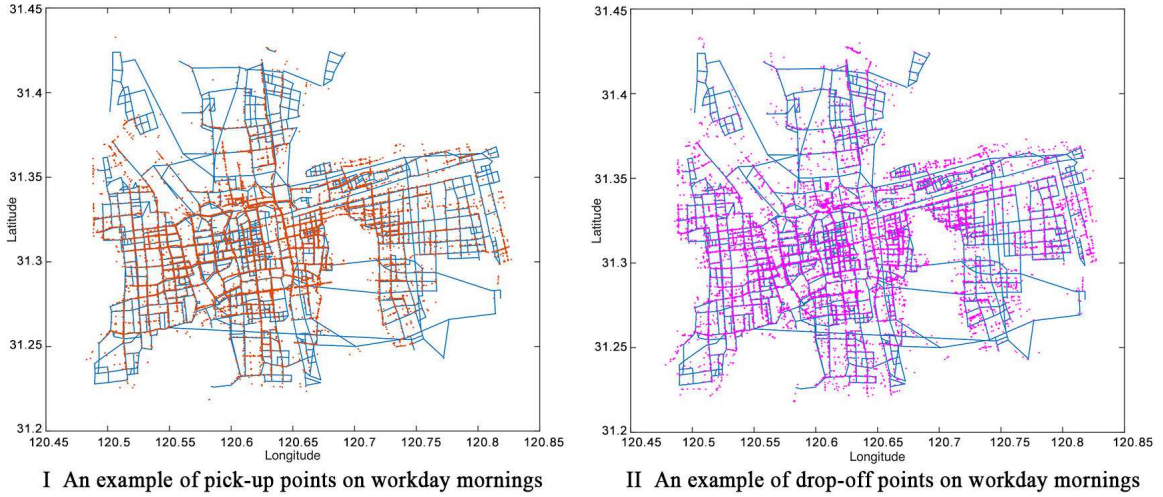

II An example of drop-off points on workday mornings

Fig. 3: Examples of pickup and drop-off points on workday mornings. Each red point represents a pick-up activity while each purple point indicates a drop-off activity.

rates. Also, the supply-demand ratio patterns on workdays and holidays are different, where occupied rates are relatively higher on workdays than weekends during morning hours while they are higher on weekends than workdays in late afternoon. From the second subfigure, we discover that the average profits of individual taxicabs are extremely high in the early morning, it is most likely that the number of on-duty taxicabs is relatively small and there are more long-distance taxicab trips with destinations of airports, train stations. In the third subfigure, the orange and green lines are for the total numbers of on-duty taxicabs of workdays and weekends, respectively. There are more taxicab businesses on weekends than workdays, because most people take buses or subways on workdays in rush hours, and they tend to take taxicabs for travelling to leisure places that are not distributed along the subway lines. Notice that there are also three valleys and a prominent peak in all three subfigures, which correspond to three mealtimes of a day and the disequilibrium between supplies and demands caused by the fact that driver shift periods of taxicabs are occasionally overlapped with the traffic rush hours in the morning.

Figure 3 visualizes the pick-up and drop-off points during 8:30 a.m. to 9:00 a.m. for ten randomly selected workdays. The point cluster of drop-off points in the peripheral urban area is much denser than that of pick-up points. Actually, most factories and companies are in the peripheral urban area, and therefore there are obvious traffic flows from residential places to business places in the morning and vice versa, reflecting the traffic can be influenced by hierarchically urban patterns.

Based on the visualized observations above, we can conclude that taxicab demands reveal an obvious period-wise and holiday-related pattern, which is also interacted with urban hierarchical structures. Thus, a more precise taxicab demand and volume prediction method considering spatiotemporal dependencies is subsequently developed to facilitate the description of taxi revenues, and further help vacant taxicabs hunt for new passengers.

4.2.1 Modeling of Taxicab Businesses in Road Segments

It is widely accepted that historical calling information can be obtained easily while real-time and near future demands cannot always be collected timely and foreahead [**?**], [45]. Therefore, the

capacity of real-time perception and future foresight in the taxicab scheduling system is the critical factor that allows the scheduling more intelligent and balanced.

In this section, we model the spatiotemporal dependencies of taxicab demands and supplies to predict their citywide distributions in the near future, based on the variants of *GGTF* proposed in Section 4.1. Regarding taxi demands, since the urban area reveals a distinct core-periphery structure, the observations of demands can be concluded that core regions tend to follow a periodical recurrent regularity while the orders on the periphery tend to be more random [9], [18], [31]. Therefore, we predict citywide demands by capturing spatiotemporal dependencies from historical records. In the demand prediction task, the linkage matrix A is replaced by the functional similarity matrix FS. Specifically, the Point of Interest (POI) distribution of each road segment will be aggregated into a $C-$dimension vector where C is the number of POI categories. FS becomes region-function similarities between road segments in terms of POI vectors. The prediction of demands can be simplified as a weighted combination of recent citywide demands and long-term expected demands. Given the historical sequence of calling information $\mathcal{D} = \{D_i, i = 1, 2, ..., T\}$, the expected citywide demands in time slot TP can be predicted by:

$$Ex(D(TP)) = \sum_{x \in DR(TP)} FS * D(x) \odot W_{dr_x}^D + \sum_{y \in DE(TP)} FS * D(y) \odot W_{de_y}^D \tag{9}$$

where $DR(TP)$ consists of last $\mathcal{P}$ time slots of TP while $DE(TP)$ contains last $\mathcal{P}$ weeks of average demands with regard to the day time slot TP in each corresponding week. Also, for one road segment r_i, $Ex_{r_i}[D(TP)]$ is the element in $Ex[D(TP)]$, denoting the total number of taxicab businesses during time slot TP. $W_{dr_x}^D, W_{de_y}^D \in \mathbb{R}^{|E| \times 1}$ represent the demand-related learnable weights, which are responsible for the combination of recent demands and historical long-term expected demands, respectively. Besides, we assume that the location at which the passenger becomes available is uniformly distributed along the road segment.

Similarly, we estimate the number of vacant taxicabs coming into road segment r_i within near future's time slot TP. Due to the dynamic but periodical transition patterns in cruising taxicabs, we apply the *GGTF* by modifying the propagation matrix A as a

transition matrix TM. The element $TM_{ij}(TP)$ in $TM(TP)$ represents the number of vacant taxicabs transfer from road segment r_i to r_j averagely in historical daily time slot TP, indicating the traffic patterns among adjacent road segments. The statistical time slot-varying transition matrices are stored as a tensor. We organize the historical data and learn the trainable weight W_T offline.

Then, we obtain the estimated number of citywide vacant taxicabs during time slot TP:

$$Ex[Tx(TP)] = \sum_{i \in RT(TP)} TM * Tx(i) \odot W_i^T \qquad (10)$$

Note that $Ex_{r_i}[Tx(TP)]$, the element in $Ex[Tx(TP)]$, is the number of vacant taxicabs coming into road segment r_i during time slot TP, and $W_i^T \in \mathbb{R}^{|E| \times 1}$ is the learnable weight in terms of vacant taxi numbers.

All parameters including statistics and learnable weights such as TM, FS, W_{dr}^D, W_{de}^D and W^T can be viewed as the derived traffic and taxi business patterns in our data-driven method.

4.2.2 Profit Analysis of Taxicab Businesses

Now, we present the profit model for taxicabs. In the revenue system studied before, the profit of a taxi driver is the total income minus all costs. Then, we detail the compositions of profits in terms of the cost and income respectively.

As for the cost of each driver, it is further classified into rental fees and energy costs. Since the rental fee is fixed, we can transform the daily rental fee into a rental fee rate, that is the average fee per time unit, denoted as α. Also, we transform the energy expense into an energy cost rate, that is the average energy cost per distance unit, denoted as β. Then, the costs for taxicabs running through road segment r_i are as follows.

$$f_c(r_i, TP) = Ex[tc_{r_i}(TP)] * \alpha + L(r_i) * \beta \qquad (11)$$

Here, $Ex[tc_{r_i}(TP)]$ represents the real-time estimated time arrival (ETA) costs for a taxicab running through road segment r_i in time slot TP, which can be calculated via Equation (5) and (6).

Next, we elaborate the taxi profits from the perspective of incomes. First, the potential profits of a taxicab are traffic-sensitive, e.g. fluctuations on ETA in accordance with the congestion conditions of different road segments.

Then, the route planning for taxi drivers has a direct effect on its profits. Hence, we devise a method for calculating the potential profits in different road segments and time slots.

Given a taxicab business trip $b(r_m, r_n, TP)$, where r_m and r_n represent the starting point and the destination of this trip, respectively. We assume the optimized driving route for this business can be unfolded as $OR[tc_{r_m \to r_n}(TP)] = (r_m, r_{m+1}, \cdots, r_n)(n > m)$. The potential profits with regard to the trip $b(r_m, r_n, TP)$ can be written as:

$$f_p(b(r_m, r_n, TP)) = M - \sum_{k=m}^{n} f_c(r_k, TP) \qquad (12)$$

where M is the business fare and can be computed by:

$$M = \begin{cases} \xi & if \sum_{k=m}^{n} L(r_k) \le l \\ \xi + \gamma * \text{ROUND}\left(\sum_{k=m}^{n} L(r_k) - l\right) & if \sum_{k=m}^{n} L(r_k) > l \end{cases}$$
$$(13)$$

Here, ξ is the taxicab starting fee, l is the taxicab starting distance, γ is the taxicab fare price per unit distance, and Function **ROUND** is to round the taxicab fee into an integer. In the case

of the total driving distance of a business $\sum_{k=m}^{n} L(r_k)$ is less than the taxicab starting distance l, the total fee is just ξ, otherwise the total fee is the sum of the taxicab starting fee and the subsequent service fee.

The above approach is to evaluate the potential profits of taxicab businesses while sources and destinations are given. It is suitable for online taxicab calling requests since the destinations are marked by passengers before the trip starts. Nevertheless, for roadside service requests, the destination is not known before and therefore the route cannot be planned in advance.

In our work, we calculate the possible profits of a roadside taxicab business appearing at r_m during time slot TP, by mining all historical taxicab businesses to find those occur in the same road segment and at the same time slot. More precisely, we can calculate it by the equation below i.e.,

$$f_p(r_m, TP) = \frac{\sum\limits_{b(r_m, r_n, TP) \in B} f_p(b(r_m, r_n, TP))}{|b(r_m, r_n, TP) \in B|} \qquad (14)$$

Note that B is the set of historical taxicab businesses and $|b(r_m, r_n, TP) \in B|$ denotes the number of elements in set B.

4.3 Analysis of Settings in Urban Taxicab System

Here, we discuss the settings of our intelligent taxicab systems. For the length of the time slot, people used to uniformly split one day into time slots of 30 minutes [28], [40]. However, from Figure 2, the average occupied ratio of taxicabs, the average profit of individual taxicabs, and the number of taxicab businesses all change rapidly during different specified time slots. Based on this, we find that 30 minutes is too long to depict the changes of the traffic conditions and taxicab business patterns. After a few rounds of testing, in our implementation, we set the length of time slot to be 15 minutes which is more accurate to capture the fluctuation of dynamics in the road network. In the case of increasing more and smaller time slots, a specific trip may be covered by multiple time slots, we need to perform a recursive multi-step predictions of traffic statuses. This way, to calculate a trip from road segment r_i to r_j, we associate the *GGTF* framework to predict the citywide statuses, then generate $Ex[tv(TP)]$, $Ex[tc(TP)]$, $Ex[D(TP)]$ and $Ex[Tx(TP)]$ for corresponding road segments.

5 ONLINE TAXICAB SCHEDULING

In this section, we design a context-aware based regional value to help quantify the potential profits of each possible driving route considering the impact of peer taxis and time-varying traffic conditions. To improve the order dispatch efficiency and maximize the service delivery rates of our taxicab scheduling system, the order-taxicab value score and bi-incentive strategy are well-designed, respectively. Finally, we present the integrated online scheduling algorithm.

5.1 Online Regional Values for Potential Driving Routes

In order to optimize the recommended route for a taxicab from all its possible driving routes, we need to compute the potential profits of this taxicab while it entering a road segment in a given route at a given time slot, which boils down to calculating the local potential profits that a taxicab is expected to earn. Here, we suppose the location where a passenger appears along the road is uniformly distributed, and all vacant taxicabs travel along the road segment at a fixed speed depending on traffic conditions.
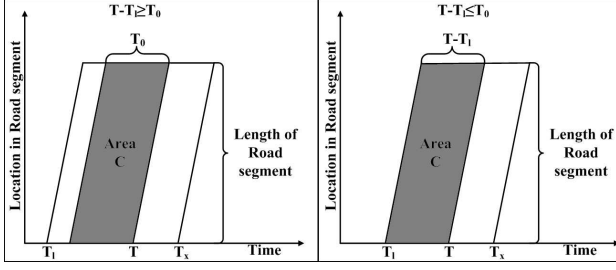
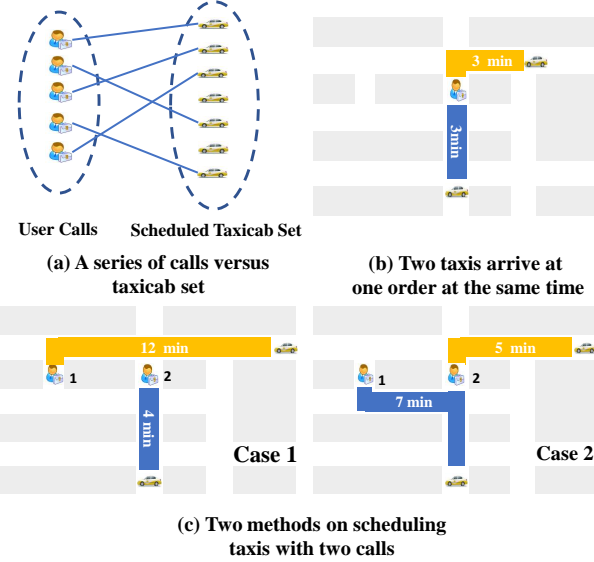Fig. 4: Temporal-spatial diagram of taxicabs in road segments

Fig. 5: Examples of taxicab scheduling

As discussed above, most previous studies were on recommending optimized paths to individual taxicabs [25], [33], [35]. However, the optimal scheduling strategies for individual taxicabs may not be optimal for the entire taxicab fleet. For instance, a scheduled vacant taxicab enters a road segment at a specific time, and the potential profit of this road segment will be influenced, because the potential business could be taken away by the taxicab. At this moment, if sending another taxicab to this road segment, the potential profit in this road segment would become very limited, because of the competition taken for taxi orders. Another example is shown in Figure 5. Figure 5(b) illustrates the scanerio that two drivers can arrive at the order departure place within the same time, it becomes difficult to decide which taxicab to pick up the passenger. And in Figure 5(c), if both two drivers can arrive at the order departure place within maximum waiting time, the system is required to dispatch the best order-taxicab pair. In case 1, imbalance arises when passenger 1 has to wait much longer than passenger 2, which deteriorates the user experience of passenger 1. However, in case 2, it balances the waiting time between two passengers even though passenger 2 has to wait for longer time. These two cases jointly indicate that the cooperative scheduling with context taxicabs and orders considered is of great significance in optimization.

In order to avoid collisions as well as biased solutions in taxicabs scheduling, we are encouraged to collect the orders subsequently as Figure 5(a), and take the spatiotemporal influences of previous taxicab scheduling strategies into account for calculating the potential profits of the road segments. This way, such scheduling strategies are arguably collision-avoided and context-aware.

It is accepted that the potential profits are mainly determined by the driving route of each driver [14], [36], [41]. Intuitively, according to the city structure and urban functionality, it follows different revenue patterns in each road segment, i.e. closer to downtown, the values are higher, which indicates a larger opportunity to meet with high-quality orders with long business distance and short pick-up distance. The regional revenues are dependent on the supplies provided by taxi drivers and demands launched by passengers in the road segment. To describe real-time prospective revenues of taxicabs cruising, we design a context-aware feature, named regional driving value (RDV), to perceive real-time supply-demand business status in each road segment.

Definition 2 (Regional Driving Value). Given a road segment r_i and the current time slot TP, the number of vacant taxicabs Tx_{r_i}, and the number of passenger callings D_{r_i}, the Regional Driving Value (RDV) of the road segment r_i can be defined as $\mathcal{V}_{RDV}(r_i, TP)$,

$$\mathcal{V}_{RDV}(r_i, TP) = 1 + \frac{1 + D_{r_i}(TP)}{1 + Tx_{r_i}(TP)} \tag{15}$$

where the factor '1' is designed to eliminate the impacts of zero values of $Tx_{r_i}(TP)$ or $D_{r_i}(TP)$. As described in [25], we don't have to consider the influences of the road segments which are 5 steps away from the taxicab driver when calculating the potential revenues, since the increasing rates of the expected revenues are less than 10% after more than 5 road segments. Thus, it is reasonable to set an upper bound $\Omega = 5$ for the length of possible route S. Without considering online taxicab calling information, the total expected values are indicated by the following five road segments which starts at the current one and ends within five steps in the corresponding route S. Then the regional driving value along the route S is computed by

$$\mathcal{V}_{RDV}(S, TP) = \prod_{r_i \in S} \mathcal{V}_{RDV}(r_i, TP) \tag{16}$$

As we obtain the predicted demands and taxicabs in Section 4.2.1 in future time slots, we are also able to foresee future potential values in all road segments. To better understand our proposed RDV, we show an example of citywide RDV in Figure 6. The subregions colored deeper represent they are with greater demand-supply ratios and more potential driving values. Meanwhile, the citywide RDV generally follows the core-periphery pattern except for some highlighted spots such as transportation center, which conforms to the real-world business patterns.

Definition 3 (Optimized Route Selection Problem). Given a taxicab, for a candidate set of driving routes Θ, the target is to recommend a route $S^* \in \Theta$ with the maximum expected revenue in the current time slot, i.e.,

$$S^* = \arg\max_{S \in \Theta}\{\mathcal{V}_{RDV}(S, TP)\} \tag{17}$$

Since we bound the length of possible routes by Ω, and determine that each intersection connects with constant outgoing road segments, the searching space of possible routes can thus be elaborated with constant time complexity. Then the quantified potential revenues along the route S^* can be calculated with the sum of profits in each involved road segment by Equation (12)-(13):

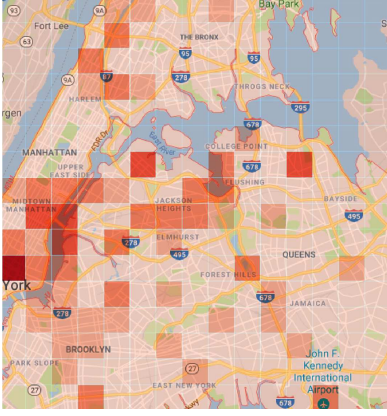$$f_p(S^*, TP) = \sum_{r_i \in S^*} f_p(r_i, TP) \tag{18}$$

Fig. 6: An example of the citywide RDV map

5.2 Order-Driver Value Prediction for Improvements on Order Dispatch Efficiency

By investigating real-life taxicab businesses, we find that broadcasting orders in a wide-spread fashion in most taxicab scheduling systems such as Didi deteriorates the efficiency of matching order-taxicab pairs. Since order-taxicab pairs are usually with various interacted attributes (e.g. pick-up distance) that contribute to unequal values to each driver, it is highly wanted to compute the potential revenue of each order specific to every individual taxicab. Inspired by [41], in our system, we design an order-taxicab value score, which can be estimated by the descriptions of orders and drivers. Further, to perform in a data-driven manner, the value score can be viewed as the predicted probability of a successful business within an order-taxicab pair. Thus, the training samples can be organized easily. The description of orders and drivers are classified as order related, including business distances, ETAs between the origin and destination and the destination category, taxicab related including the historical and recent order probability, and order-taxicab related including pick-up distance as well as contextual factors. All these features are aggregated into a feature vector $\mathbf{x}$ and fed into a regression function. Given the taxicab's road segment r_k, order's origin r_o and destination r_d, the score $\mathcal{V}_{O_{od}} = 1$ stands for acceptance while $\mathcal{V}_{O_{od}} = 0$ stands for non-acceptance, the value score can be formulated as:

$$\mathcal{V}_{O_{od}} = p\left(y = 1 | r_k, r_o, r_d\right) = \frac{1}{\exp\left(-\mathbf{w}^{\mathrm{T}}\mathbf{x}_{k,o,d}\right)} \qquad (19)$$

where $\mathbf{w}^{\mathrm{T}}$ represents the learnable parameters. This model can be trained offline and the value scores within order-taxicab pairs are calculated with online deployment. This mechanism captures the order-taxicab interactions and improves the coarse-grained order dispatch to a more precise recommendation, which not only avoids the order explosion in rush hours, but promotes the dispatching efficiency.

5.3 Bi-incentive Strategy for Maximizing the Service Delivery Rates

Nearly all studies on order dispatch or taxicab scheduling focus on maximizing the profits of drivers and very few consider to promote the service delivery rates which closely related to passengers' experience. Considering the two following scenarios, since the orders with long pick-up distances occur in peripheral and taxi-free regions, it becomes difficult to match the order-taxicab pair.

This is mainly because drivers wouldn't like to take external detour costs to pick up the passenger with limited revenues. And in another scenario, orders with shorter business distances are also unpopular to drivers due to the revenues are proportional to the order distances. In this case, there are still some taxicabs cruising unoccupiedly. In order to improve the overall occupied ratios and service delivery rates, we propose a bi-incentive mechanism to stimulate both drivers and passengers on the deal in the case of long pick-up distance and low expected revenue orders. Specifically, from the view of drivers, our system raises the order price by asking the passenger and providing additional allowances. From the view of passengers, we identify the nearest taxicab hotspots where passengers can easily arrive and recommend the hotspots to passengers for a more quick pick up.

5.4 Proposed Algorithm: Future-Aware Intelligent Taxicab Scheduling

The difference between the goal of intelligent taxicab scheduling and order dispatching lies in that the former one not only determines the best order-taxicab matching but recommends the vacant taxicabs to the regions with more potential revenues. In our problem settings, this goal can be interpreted as maximizing global future gains of all drivers as well as the overall service delivery rates in a coordinated way.

So far, we have proposed mechanisms to predict taxicab demands, taxicab supplies as well as order-taxicab interactive values for supporting our order dispatching and scheduling task. In this section, given the taxicabs within our online scheduling system $\mathcal{A}$ and the historical trajectory set $\mathcal{TR}$, we formally introduce our distributed and future-aware intelligent scheduling algorithm F.

The overall taxicab scheduling algorithm consists of two parts. The first part is the learning-based module which aims to learn the spatiotemporal patterns of traffic and taxicab business. This part only needs to be executed once and the model parameters will be restored for online taxicab scheduling. The second part is taxicab scheduling with online real-time taxicab statuses and passenger calling information. With above pre-trained models, our scheduling system can perceive the current traffic conditions and foresee the status in near future. Noticed that due to the seasonal fluctuations on traffic, the learnable parameters will be updated every three months, hence it enables the system to be more sensitive in different seasons. We estimate the real-time local values RDV in each road segment to sense the supply-demand conditions which indicate how much the taxicab will earn, and subsequently provide guidance to select the optimized route for taxicabs.

Next, we describe our proposed scheduling algorithm in the following scenarios.

Scenario 1: Our system monitors the business statuses in terms of both demand and supply sides regularly and updates every time slot. For time slot TP, it computes the real-time and near future's citywide RDV with our proposed future-aware method, to quantify the balance of demands and supplies. For any vacant taxicab in set $\mathcal{A}$, the system calculates the expected revenues of all possible driving routes in a current time slot with Equation (15)-(17) and recommends the corresponding optimized route S^*. Specifically, the system schedules the vacant taxicabs cruising on demand-overloaded road segments to the underloaded. In a demand-overloaded road segment, the system compares the timestamp of each taxicab entering and then schedules the latest-entering taxicab to other nearest potentially demand-underloaded

road segments. Notice that only vacant taxicabs can impact the next steps of other taxicabs that are currently vacant. In other words, the routes of occupied taxicabs should be ignored for route optimization.

Scenario 2: Once a vacant taxicab accepts a service request, its driving route is determined by the trip specified by the passenger and cannot be rescheduled before the trip ends. Then, the system deletes all recorded recommended future routes for this taxicab to make sure that the obsolete routes will not impact the scheduling strategies of other vacant taxicabs. After the trip, the occupied taxicab becomes a vacant taxicab and informs the system. Upon receiving the information, the system calculates and recommends an optimized route starting at the coming road intersection.

Scenario 3: Upon receiving a series of service requests from passengers, the system triggers the order dispatch task. For a service request originated from road segment r_o and destined at r_d, the system first calculates the expected driving time to the passenger r_o for all vacant taxicabs in set $\mathcal{A}$, based on current and predicted traffic conditions. Then the system organizes the order-taxicab attributes of those taxicabs can arrive in T_0 and evaluate their order-taxicab value scores $\mathcal{V}_{O_{od}}$ with our pre-trained model. The taxicabs with higher value scores tend to have a higher possibility to accept the orders, hence taxicabs with top-K scores are selected into a candidate taxicab set $\mathcal{H} = \{v_{h_1}, ..., v_{h_K}\}$. Among them, we select one taxicab to pick up this passenger and maximize the overall profits of all scheduled taxicabs at the same time. Specifically, given taxicab v_{h_k}, the next recommended route to taxicab v_{h_k} is $s^*_{v_{h_k}}$, the expected driving revenue of the given recommended route is $f_p(s^*_{v_{h_k}}, TP)$. To optimize the scheduling globally, we select the taxicab v^* in candidate taxicab set $\mathcal{H}$ each time to ensure the expected revenue of its potential route less than the expected revenue of the given online taxicab calling business, i.e.,

$$
\begin{aligned}
v^* &= \arg\min_{v_{h_k} \in \mathcal{H}} \left\{ f_p(s^*_{v_{h_k}}, TP) \right\} \\
s.t.\ & \\
& f_p(b(r_o, r_d, TP)) > f_p(s^*_{v_{h_k}}, TP)
\end{aligned}
\tag{20}
$$

It worth noting that, here we omit the taxicab detour cost from the cruising road segment to the passenger departure position due to the arrival time constrains the detour cost as well as the order-taxicab scores have taken pick-up distances into account. K is a hyperparameter in our algorithm and we set $K = 10$ according to the experiments.

If the condition is not met, we will relax above conditions to selecting the taxicab with minimum expected potential revenues. Then the taxicabs cruising on low expected-value road segments are more likely to be scheduled to pick up the most recent passengers. Otherwise, if no vacant taxicabs in set $\mathcal{A}$ can pick up this passenger within T_0, then the bi-incentive mechanism will be triggered to help facilitate the deal of this order. In this case, the order will be broadcast one round by one round dynamically until the maximum round meets or the passenger cancels.

Scheduling Method. Our taxicab scheduling system runs regularly as Scenario 1 described. Once a vacant taxicab in $\mathcal{A}$ arrives at an intersection, the optimized route is recommended for a global taxi scheduling. Such recommendation decides the direction of next step, once a vacant taxicab is approaching an intersection. When the next step of a taxicab is planned, the system records it and takes it as a possible precondition for the subsequent recommendation. Once the system receives a series of online callings, the pick-up time is calculated and order-taxicab value

TABLE 1: Evaluation setup

Parameters	Description
Length of time slot $\mathcal{T}$	15 minutes
Upper bound of route length	$\Omega = 5$
Percentage of taxis in set $\mathcal{A}$ account for total	25%
Number of days for evaluation	10
Waiting time limit of passenger	$T_0 = 10\ min$
Taxicab starting distance	$l = 3\ kilometers$
Taxicab starting fee	$\xi = 10\ RMB/8\ USD$
Taxicab fare price per unit distance	$\gamma = 2\ RMB/Kilometer$

scores are computed to help the order dispatch process, triggering Scenario 3. Once a vacant taxicab in $\mathcal{A}$ picks up a passenger, it will be released and labeled as an occupied taxicab, triggering Scenario 2. And once the order is canceled or taxi arrives at the destination, the system triggers Scenario 1.

To better understand the operation mechanism of our proposed algorithm, we demonstrate the technical process of TS-DFA in Figure 7.
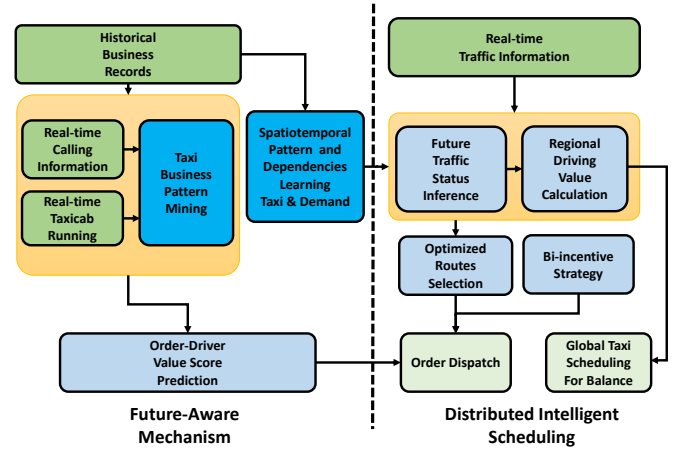

Fig. 7: The technical process of proposed algorithm TS-DFA

6   EVALUATION

In this section, we evaluate the effectiveness of our proposed taxicab scheduling algorithm by conducting extensive empirical studies on two cities, Suzhou and NYC.

6.1   Data Description

In the experiments, we use two real-world datasets Suzhou and NYC to implement our future-aware real-time taxicab scheduling.

Suzhou is a leading city in deploying taxicab trajectory monitoring systems, and the dataset is collected from the government of Suzhou, China [3]. By processing the dataset, we obtain $40,970,885$ pick-up and drop-off activities throughout the entire year in 2012 (365-day) including $4,303$-taxicab trajectories. We randomly select 10 days which cover $1,122,490$ records for test, and use the remaining 355 days to infer taxicab business patterns and taxicab running costs of road segments.

NYC is a top international metropolis with a highly dynamic taxicab system. We collect the taxi trip records from NYC Taxi &

Limousine Commision [4] and recover the trajectories with the pick-up and drop-off locations. Totally, we obtain 32, 540, 088 taxi trips containing detailed pick-up and drop-off information with time stamps and taxi fees between 01/01/2017 and 05/31/2017 (151 days). We randomly select 10 days including 3, 361, 620 pick-up and drop-off activities for test, and the remaining 141 days will be utilized for pattern learnings.

Meanwhile, we assume that all active taxicabs in set $\mathcal{A}$ follow the scheduling strategies of our proposal, and taxicabs in set $\mathcal{A}$ account for 25% of all taxicabs except in the experiment of investigating the impact of the percentage of taxicabs involved in set $\mathcal{A}$.

6.2 Implementation Details

In the offline learning part of the traffic forecasting framework, we set the length of past observation $\mathcal{P}$ in terms of trend, period and closeness as 3 for all perspectives, according to the settings in traffic flow prediction paradigm [40]. For online forecasting, we fetch the corresponding data and pass it through the model, the algorithm outputs the predicted future traffic statuses to support the scheduling. And in order dispatching module, online calling requests or taxicab pick-up points in real-world datasets will trigger the scheduling. The taxicabs with top-K order-taxicab value scores will be selected into the candidate set. Then, the system encourages the taxicab with lower potential revenue to accept the most recent order. In terms of the implementation of the proposed bi-incentive strategy, we simulate the conditions by extending the maximum waiting time of the passengers to 1.5 times and increasing the revenues to 1.05 times of the previous for simplicity. We fix the period of a time slot as 15 minutes and therefore there are 96 time slots in one day.

Regarding taxicab operations, we assume that the fuel price is 8 RMB (4 USD) per liter, taxicab consumes 10 liters of fuel per 100 kilometers, and a driver pays 200 RMB (100 USD) to the taxicab company every day for taxicab rents. The taxi fare is 10 RMB (5 USD) for the first 3 kilometers and 2 RMB/Kilometer (2 USD/Kilometer) afterwards. Unless otherwise specified, the default values are set according to Table 1.

6.3 Experimental Results

6.3.1 Baselines

Since the studies most closely related to ours are taxicab route recommendations and order dispatch tasks, we compare our taxi scheduling algorithm TS-DFA with state-of-the-art algorithms, CERS [25], LOD-RHP [29], and TODCO [41].

- **CERS** designs a profit objective function to evaluate the potential profits of all possible driving routes while minimizing the cruising time and recommending the optimal route for drivers [25].
- **TODCO** is a combinatorial optimization method which aims to maximize the global business success rate by predicting driver actions and passenger destinations simultaneously [41].
- **LOD-RHP** is an MDP-based method that formulates the order dispatch as a large-scale sequential decision-making problem and aims to optimize the platforms long-term efficiency as well as instant customer demands [29].

4. https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

The performances of CERS, TODCO and LOD-RHP algorithms are all evaluated with online taxicab calling information for fair comparisons.

6.3.2 Evaluation Metrics

The performances of a taxicab scheduling system are mainly concerned with drivers' profits and passengers' experience. In this case, the metrics designed for evaluating the performances of taxi scheduling algorithms are as follows: i) average daily profits, average passenger waiting time and overall service delivery rates of taxicabs; ii) average profits, passenger waiting time and service delivery rates in terms of different time slots; iii) average daily profits of taxicabs on different types of day and iv) profit distributions among various taxicabs. To be noted that the service delivery rate refers to the order acceptance rate in the calling system.

6.3.3 Numerical Results Comparison and Analysis

Daily evaluation. The numerical results in terms of daily average metrics are illustrated in Table 2. Overall, our proposed algorithm TS-DFA gains remarkable improvement on average profits and service delivery rates, which are two key metrics in taxicab business system. Obviously, the service delivery rates, namely business success rates, raise a lot compared with alternative methods such as CERS by 10% and 5% on Suzhou and NYC, respectively. It achieves closely and slightly higher performance with TODCO, which mainly aims to optimize global business success rates. The improvement of our integrated algorithm may stem from the proposed future-aware mechanism, which guides the inexperienced drivers to cruise on road segments with higher potential values and balances supply-demand dilemma ahead of time. And the predictions of order-taxicab value scores enable our algorithm to filter out the most appropriate taxicabs for best dispatching. At the same time, there is no doubt that overall profits are raised with the promotion of business rates. Noted that, CERS helps drivers earn more profits than TODCO, as CERS only optimizes the single objective of profits. These results coincide with our motivation that a farsighted and win-win business operation leads to more revenues and service success rates.

For the metric of passenger waiting time, LOD-RHP achieves better performance, as it tends to satisfy the instant customer demands. Even so, our method still performs best among them mainly because the order-taxicab value score and bi-incentive strategy make sense when in the face of extreme taxicab business environments.

From the perspective of different day types, we observe that the average daily profits of taxicabs during weekends are better than those during workdays in Table 3. This is due to the orders at weekends may be a little more than workdays. Another reason can lie in that traffic patterns are more easily captured at weekends, and taxicabs can run more efficiently under this circumstance.

Figure 9 presents the distributions of the average daily profits of individual taxicabs on Suzhou dataset. The distribution of our algorithm is roughly in the range of 350-500 RMB per day. It means that our algorithm is capable of equilibrating the daily profits among taxicabs. Also, the average daily profit of TS-DFA is more stable than others, especially compared with CERS, due to CERS is only to maximize the driver profits without considering the imbalanced traffic and demand patterns.

Another interesting phenomenon in Table 2 is that the average profits in NYC are higher and the waiting time is less than those of Suzhou consistently. The reason behind it may lie in that the
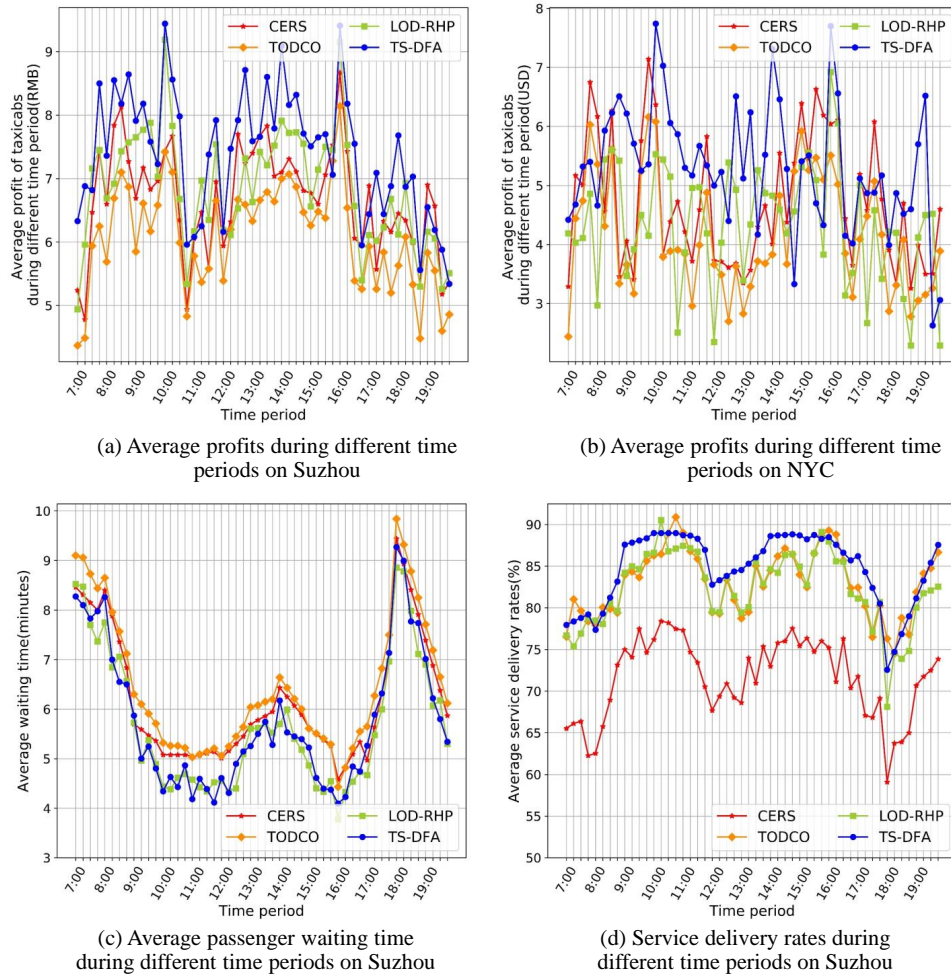
(a) Average profits during different time periods on Suzhou

(b) Average profits during different time periods on NYC

(c) Average passenger waiting time during different time periods on Suzhou

(d) Service delivery rates during different time periods on Suzhou

Fig. 8: Performance during different time periods

TABLE 2: Results comparisons in different cities

City	Methods	Average Profits per day (RMB/USD)	Passenger Waiting Time (min)	Service Delivery Rates (%)
Suzhou	CERS	406.0	6.29	74.82
	TODCO	395.6	6.42	82.45
	LOD-RHP	418.5	5.69	81.03
	TS-DFA (Ours)	428.7	5.76	84.96
NYC	CERS	223.5	4.87	78.94
	TODCO	218.0	4.66	82.42
	LOD-RHP	238.3	4.32	80.36
	TS-DFA (Ours)	254.8	4.48	83.17

TABLE 3: Profit performance on different types of days

Dataset Method	Suzhou(RMB)		NYC(USD)	
Day type	Workday	Weekeed	Workday	Weekend
CERS	398	408	241	251
TODCO	385	412	248	254
LOD-RHP	409	424	252	262
TS-DFA	422	432	262	268

average GDP values in NYC are much higher than those of Suzhou while the urban areas covered in NYC are smaller than Suzhou, then the drivers can arrive at the pick-up location more quickly.

Time slot-wise evaluation. To investigate the algorithm performance in terms of various time periods, we visualize the average profits between 7:00 a.m. and 8:00 p.m. in Figure 8(a) and (b) for Suzhou and NYC, respectively. Also, in order to evaluate the online calling service capacity, we take Suzhou as an example and present the performances on these two metrics in Figure 8(c) and (d). Overall, our proposed method is able to gain more profits and facilitate the deal of taxi business in most time slots. Worth noting that we observe the proposed algorithm performs remarkably better than others during most time periods from 7:00 a.m. to 9:00 a.m., 5:00 p.m. to 7:00 p.m. in terms of service delivery rates. This reveals our proposed method has
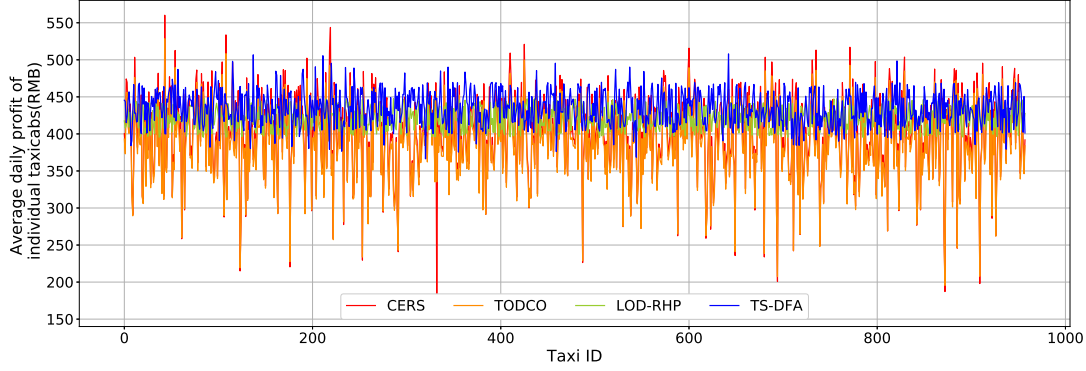
Fig. 9: Profit distribution of scheduled taxicabs

the superiority to deal with large-scale orders with order-taxicab value scores and can work better with insufficient taxicab supplies. At nights, our method schedules and stimulates the unoccupied taxis to areas with higher probability to pick up passengers under the circumstance of supply-demand imbalance, satisfying the passenger demands and improving overall revenues.

6.4 Ablation Studies

In our work, three important schemes are proposed, traffic forecasting based future-aware mechanism, regional driving value metric and order-taxicab value score-based context-aware mechanism. To investigate how each newly proposed component contributes to the success of the entire algorithm, we remove the proposed strategy from the integrated algorithm to perform the ablation studies on our Suzhou dataset. The following strategies are removed successively as TS-V1 to TS-V4: (1) TS-V1: Future-aware mechanism, (2) TS-V2: Regional driving value, (3) TS-V3: Order-taxicab value score, (4) TS-V4: Bi-incentive strategy.

Table 4 shows the performances of these four variants of our algorithm. Obviously, the future-aware mechanism contributes to a large decrease of waiting time by 6% from 6.14 min to 5.76 min. This verifies that the mechanism enables our algorithm to precisely foresee the near future traffic statuses, and to improve the taxicabs utilization as well as global revenues. As for the bi-incentive strategy, it provides opportunities for drivers and passengers to reconsider the business and facilitate the deal success, promoting the overall profits and service delivery rates. If we remove it, the business success rates and profits will drop to 82.17% and 416.5 *RMB*. With *RDV* removed, the profits have an increases, and with bi-incentive strategy removed, the passenger waiting time decreases due to the discards of out-of-range orders. In summary, our integrated model is consistently superior to other variants on most metrics and the scheduling task needs the tradeoff between different factors.

6.5 Parameter studies

In this part, we study the impacts of an external parameter and an internal hyperparameter on performances. The results are illustrated in Figure 10.

As for the external factor, the percentage of taxicabs involved in $\mathcal{A}$, we compare the performance with the best baseline LOD-RHP. The profits decrease when the involved taxicabs increase,

TABLE 4: Performance on different variants of TS-DFA on Suzhou dataset

Variant	Average Daily Profits (RMB)	Passenger Waiting Time (min)	Service Delivery Rates (%)
TS-V1	426.7	6.14	81.44
TS-V2	431.3	5.93	82.26
TS-V3	422.8	5.72	81.55
TS-V4	416.5	5.50	82.17
Integrated	428.7	5.76	84.96

(a) Performance on different scheduled taxicabs

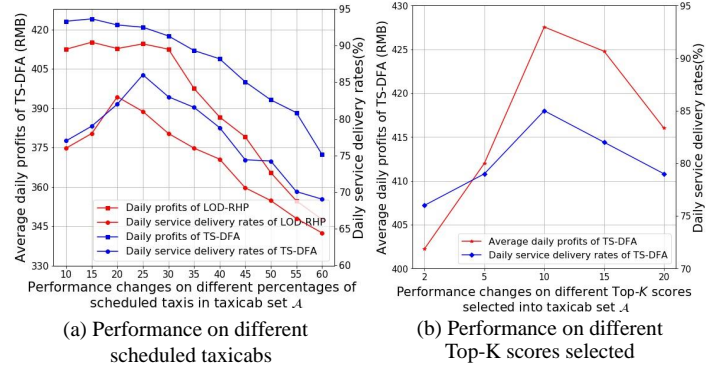(b) Performance on different Top-K scores selected

Fig. 10: Performance on different parameter settings

meanwhile the service delivery rates climb to a peak around 22% and then decline afterwards, indicating that the success of businesses requires a medium-sized volume of scheduled taxicabs. We also observe that our method has an edge on the robustness in this test when the percentage of taxicabs in set $\mathcal{A}$ increases from 15% to 60%.

In terms of the internal hyperparameter K, we adjust K in {2, 5, 10, 15, 20} and obtain the best performance when K equals 10. The reason lies in that a greater K declines the efficiency in order dispatch while a smaller K reduces the opportunity to select the best order-taxicab match. Also, the highly dynamic of supplies and demands inspires us that K can be a learnable time-varying parameter in different time slots and it remains as our future work.

6.6 Case Study and Procedure Visualization

Besides the numerical results, in order to provide a more intuitive understanding of the proposed future-aware and context-aware

TABLE 5: Order-driver value scores between each order-taxicab pair

Order (dist) / Driver	a (6.8km)	b (5.2km)	c (4.8km)	d (3.2km)
1	0.56	0.23	0.45	0.09
2	0.62	0.15	0.43	0.08
3	0.41	0.22	0.48	0.07
4	0.46	0.27	0.33	0.10
5	0.20	0.54	0.08	0.44
6	0.27	0.51	0.12	0.47

TABLE 6: Order-driver dispatch results

Driver	Assigned order	Order-driver value score
1	-	-
2	a	0.62
3	c	0.48
4	-	-
5	b	0.54
6	d	0.47

(a) Time-varying RDV maps of a selected day in NYC   (b) Example of order dispatch

Fig. 11: Cases of RDV Maps and order dispatch example

scheduling algorithm, we perform case studies in this subsection. We first analyze the dynamic *RDV* map generated by the real-world dataset. Then we select a subregion of NYC and establish a small set of order requests and taxicabs. We will analyze how the algorithm dispatches 4 orders to 6 drivers within the subregion.

Figure 11(a) exhibits the *RDV* map indicating the dynamic demand-to-supply ratios during different time slots. The circles colored yellow reveal the residential places near Jackson Heights and those colored red are located in Manhattan business blocks. We discover that Manhattan blocks tend to have higher regional values during almost all hours, indicating driving to the business areas usually gain more profits. Also, values in the morning in Jackson Heights are relatively higher than afternoon while values in Manhattan are relatively lower in the morning than afternoon. It conforms to the urban commuting regularity that people in residential areas will travel to urban business blocks in the morning, increasing the taxicab demands in regions circled yellow and the values in business blocks shrink in the morning due to the lower population intensity. Fortunately, these results correspond to our previous analysis and motivations, leading to a more accurate route optimization.

Then we randomly release 4 orders with 6 taxicabs involved to build the toy example in the selected regions as Figure 11(b) shows. We show the derived order-taxicab value scores in Table 5 via the learned business patterns. The order dispatch results are presented in Table 6. Specifically, the order *b* is assigned to taxicab 5 which is with longer pick-up distances than taxi 6 considering the balance of passenger waiting time within order *b* and *d*. And the order *a* is assigned to taxicab 2 rather than taxicab 1, probably because the taxicab 1 locates at regions with higher potential values than taxicab 2. The toy example results particularly verify the effectiveness and the aims of our context-aware and future-aware algorithm.

6.7 Discussion

In this subsection, we will discuss some practical issues related to our proposed algorithm.

Acceptable time complexity of TS-DFA. Regarding the algorithm complexity of our proposed method, it can be roughly
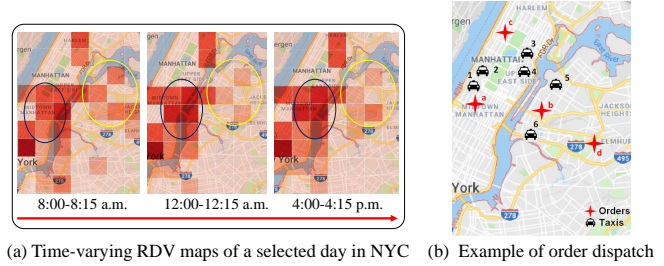
divided into three parts. The future-aware mechanism takes the graph-based model as the main component, which includes $3 \times \mathcal{P} \times |V|$ parameters and $\mathcal{P}$ is a constant number determined by the combined time slots. And for our context-aware scheduling, let Ω be the search bound in route optimization, it takes 4^Ω search space due to 4 intersections connected maximumly. The order-taxicab value evaluation process takes $|\mathbf{w}|$ parameters, which is a constant number. In the experiments, the forecasting process of our algorithm takes averagely 1.08 seconds and the scheduling process takes approximately 1.28 seconds when a series of orders and taxicabs come, sufficiently meeting the requirement of real-time taxicab scheduling. Therefore, our work can easily schedule multiple taxicabs cooperatively within polynomial time complexity with the proposed multi-stage learning-based mechanism.

The scalability of TS-DFA. The core idea in our method includes the future-aware mechanism which foresees the traffic statuses by mining historical patterns, the context-aware mechanism that measures the regional values, and the learnable order-taxicab value that measures the revenue of each order specific to every individual taxicab. Thus, it has a nice extension for encourage better optimized solutions with learnable mechanisms in other dynamic scheduling tasks such as work scheduling and assignment.

The limitations of TS-DFA. Despite the promising results, TS-DFA also suffers the unforeseebility of individual vehicles, hence the drivers within our system cannot be aware of the surrounding individual vehicle conditions during driving process. Thanks to edge computing techniques [19], [26], [27], there must be opportunities to deploy sensing devices on the road to facilitate local traffic the awareness, for supporting local driving decisions and route planning. In addition, with emerging applications of deep learning on vehicle technologies [11], [12], by defining revenues and introducing novel reinforcement deep learning, our model can be more powerful in adaptivity to more complex situations and reduce applying labor-intensive labeled data. This leads to another direction of future studies.

7 CONCLUSION

The increasing use of online taxicab calling services facilitates the development of taxicab industries and intelligent transportation systems. However, the large-scale order flows and short-term imbalanced supply-demand distributions pose a great challenge to further prosperity. In this paper, to achieve farsighted traffic awarenesses and quick order dispatches, we propose an integrated intelligent taxicab scheduling approach based on future-aware and context-aware mechanisms. Specifically, we propose a graph-based traffic forecasting framework to foresee the traffic statuses considering the spatiotemporal dependencies, and measure the

regional supply-demand contexts by proposed *RDV* for route selection. Then we design an order-taxicab value score to mine the historical order-taxicab business patterns in order to tackle the efficiency bottleneck in large-scale order flows. To facilitate the service delivery, a bi-incentive strategy is seamlessly coupled into our dispatch scheme. Experimental results on two real-world taxicab datasets demonstrate the significant improvement of service delivery rates and global revenues of our proposal. Therefore, this paper offers new solutions to urban dynamic traffic challenges and spatiotemporal scheduling tasks with future-aware and context-aware perspectives.

For future work, we will investigate the adaptive dynamic order dispatch system and extend our future-aware and context-aware mechanism to a deep reinforcement learning method while maintaining the interpretability, which will be of significance in dynamic scheduling tasks.

8 Acknowledgments

References

[1] Didi taxi. http://www.xiaojukeji.com/wensite/index.html, 2015.

[2] Kuaidi taxi. http://www.kuaidadi.com/, 2015.

[3] Suzhou city passenger transport traffic management department. http://www.szkg.gov.cn/, 2015.

[4] Kanika Bathla, Vaskar Raychoudhury, Divya Saxena, and Ajay D Kshemkalyani. Real-time distributed taxi ride sharing. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 2044–2051. IEEE, 2018.

[5] Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

[6] Yubin Duan, Ning Wang, and Jie Wu. Optimizing order dispatch for ride-sharing systems. In *2019 28th International Conference on Computer Communication and Networks (ICCCN)*, pages 1–9. IEEE, 2019.

[7] Nandani Garg and Sayan Ranu. Route recommendations for idle taxi drivers: Find me the shortest route to a customer! In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge discovery & Data Mining*, pages 1425–1434. ACM, 2018.

[8] Yong Ge, Hui Xiong, Alexander Tuzhilin, Keli Xiao, Marco Gruteser, and Michael Pazzani. An energy-efficient mobile recommender system. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 899–908. ACM, 2010.

[9] Xu Geng, Yaguang Li, Leye Wang, Lingyu Zhang, Qiang Yang, Jieping Ye, and Yan Liu. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3656–3663. AAAI, 2019.

[10] Hongzhi Guo, Jiajia Liu, Jie Zhang, Wen Sun, and Nei Kato. Mobile-edge computation offloading for ultradense iot networks. *IEEE Internet of Things Journal*, 5(6):4977–4988, 2018.

[11] Sheng Hao, Zhang Yang, Wu Yubin, Wang Shuai, Lyu Weifeng, Ke Wei, and Xiong Zhang. Hypothesis testing based tracking with spatio-temporal joint interaction modeling. *IEEE Transactions on Circuits and Systems for Video Technology*, pages 1–1, 2020.

[12] Sheng Hao, Zheng Yanwei, Ke Wei, Yu Dongxiao, Cheng Xiuzhen, Lyu Weifeng, and Xiong Zhang. Mining hard samples globally and efficiently for person re-identification. *IEEE Internet of Things Journal*, pages 1–1, 2020.

[13] Analysys International. China online taxi calling apps market size 2014. http://english.analysys.com.cn, 2014.

[14] Yongxuan Lai, Zheng Lv, Kuan-Ching Li, and Minghong Liao. Urban traffic coulombs law: A new approach for taxi route recommendation. *IEEE Transactions on Intelligent Transportation Systems*, 20(8):3024–3037, 2018.

[15] Yansong Li, Qian Luo, Jiajia Liu, Hongzhi Guo, and Nei Kato. Tsp security in intelligent and connected vehicles: Challenges and solutions. *IEEE Wireless Communications*, 26(3):125–131, 2019.

[16] Jiajia Liu, Hongzhi Guo, Jingyu Xiong, Nei Kato, Jie Zhang, and Yanning Zhang. Smart and resilient ev charging in sdn-enhanced vehicular edge computing networks. *IEEE Journal on Selected Areas in Communications*, 38(1):217–228, 2020.

[17] Qian Luo and Jiajia Liu. Wireless telematics systems in emerging intelligent and connected vehicles: Threats and solutions. *IEEE Wireless Communications*, 25(6):113–119, 2018.

[18] Xiao Luo, Liang Dong, Yi Dou, Ning Zhang, Jingzheng Ren, Ye Li, Lu Sun, and Shengyong Yao. Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: a case of shanghai, china. *Journal of cleaner production*, 142:926–935, 2017.

[19] Feng Lyu, Ju Ren, Nan Cheng, Peng Yang, Minglu Li, Yaoxue Zhang, and Xuemin Shen. Demystifying traffic statistics for edge cache deployment in large-scale wifi system. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, pages 965–975. IEEE, 2019.

[20] Yan Lyu, Victor CS Lee, Joseph Kee-Yin Ng, Brian Y Lim, Kai Liu, and Chao Chen. Flexi-sharing: A flexible and personalized taxi-sharing system. *IEEE Transactions on Vehicular Technology*, 68(10):9399–9413, 2019.

[21] Shuo Ma, Yu Zheng, and Ouri Wolfson. T-share: A large-scale dynamic taxi ridesharing service. In *Data Engineering (ICDE), 2013 IEEE 29th International Conference*, pages 410–421. IEEE, 2013.

[22] Gabor Nagy and Said Salhi. Heuristic algorithms for single and multiple depot vehicle routing problems with pickups and deliveries. *European journal of operational research*, 162(1):126–141, 2005.

[23] Yuanyuan Qiao, Yihang Cheng, Jie Yang, Jiajia Liu, and Nei Kato. A mobility analytical framework for big mobile data in densely populated area. *IEEE transactions on Vehicular Technology*, 66(2):1443–1455, 2016.

[24] Yuanyuan Qiao, Zhizhuang Xing, Zubair Md Fadlullah, Jie Yang, and Nei Kato. Characterizing flow, application, and user behavior in mobile networks: A framework for mobile big data. *IEEE Wireless Communications*, 25(1):40–49, 2018.

[25] Meng Qu, Hengshu Zhu, Junming Liu, Guannan Liu, and Hui Xiong. A cost-effective recommender system for taxi drivers. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 45–54. ACM, 2014.

[26] Ju Ren, Hui Guo, Chugui Xu, and Yaoxue Zhang. Serving at the edge: A scalable iot architecture based on transparent computing. *IEEE Network*, 31(5):96–105, 2017.

[27] Ju Ren, Deyu Zhang, Shiwen He, Yaoxue Zhang, and Tao Li. A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet. *ACM Computing Surveys (CSUR)*, 52(6):1–36, 2019.

[28] Yang Wang, Liusheng Huang, Tianbo Gu, Hao Wei, Kai Xing, and Junshan Zhang. Data-driven traffic flow analysis for vehicular communications. In *INFOCOM, 2014 Proceedings IEEE*, pages 1977–1985. IEEE, 2014.

[29] Zhe Xu, Zhixin Li, Qingwen Guan, Dingshui Zhang, Qiang Li, Junxiao Nan, Chunyang Liu, Wei Bian, and Jieping Ye. Large-scale order dispatch in on-demand ride-hailing platforms: A learning and planning approach. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 905–913. ACM, 2018.

[30] Yijie Xun, Jiajia Liu, Nei Kato, Yongqiang Fang, and Yanning Zhang. Automobile driver fingerprinting: A new machine learning based authentication scheme. *IEEE Transactions on Industrial Informatics*, 16(2):1417–1426, 2020.

[31] Huaxiu Yao, Fei Wu, Jintao Ke, Xianfeng Tang, Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye, and Zhenhui Li. Deep multi-view spatial-temporal network for taxi demand prediction. In *Thirty-Second AAAI Conference on Artificial Intelligence*, pages 2588–2595. AAAI, 2018.

[32] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. T-drive: Enhancing driving directions with taxi drivers' intelligence. *IEEE Transactions on Knowledge and Data Engineering*, 25(1):220–232, 2013.

[33] Jing Yuan, Yu Zheng, Chengyang Zhang, Wenlei Xie, Xing Xie, Guangzhong Sun, and Yan Huang. T-drive: driving directions based on taxi trajectories. In *Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems*, pages 99–108. ACM, 2010.

[34] Jing Yuan, Yu Zheng, Liuhang Zhang, Xing Xie, and Guangzhong Sun. Where to find my next passenger. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 109–118. ACM, 2011.

[35] Nicholas Jing Yuan, Yu Zheng, Liuhang Zhang, and Xing Xie. T-finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering*, 25(10):2390–2403, 2013.

[36] Desheng Zhang and Tian He. pcruise: Reducing cruising miles for taxicab networks. In *2012 IEEE 33rd Real-Time Systems Symposium*, pages 85–94. IEEE, 2012.

[37] Desheng Zhang, Ye Li, Fan Zhang, Mingming Lu, Yunhuai Liu, and Tian He. coride: carpool service with a win-win fare model for large-scale taxicab networks. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, pages 1–14. ACM, 2013.

[38] Haibin Zhang, Qian Zhang, Jiajia Liu, and Hongzhi Guo. Fault detection and repairing for intelligent connected vehicles based on dynamic bayesian network model. *IEEE Internet of Things Journal*, 5(4):2431–2440, 2018.

[39] Jie Zhang, Hongzhi Guo, Jiajia Liu, and Yanning Zhang. Task offloading in vehicular edge computing networks: A load-balancing solution. *IEEE Transactions on Vehicular Technology*, 69(2):2092–2104, 2019.

[40] Junbo Zhang, Yu Zheng, and Dekang Qi. Deep spatio-temporal residual networks for citywide crowd flows prediction. In *Thirty-First AAAI Conference on Artificial Intelligence*, pages 1655–1661. AAAI, 2017.

[41] Lingyu Zhang, Tao Hu, Yue Min, Guobin Wu, Junying Zhang, Pengcheng Feng, Pinghua Gong, and Jieping Ye. A taxi order dispatch model based on combinatorial optimization. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2151–2159. ACM, 2017.

[42] Yu Zheng, Yanchi Liu, Jing Yuan, and Xing Xie. Urban computing with taxicabs. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 89–98. ACM, 2011.

[43] Ming Zhou, Jiarui Jin, Weinan Zhang, Zhiwei Qin, Yan Jiao, Chenxi Wang, Guobin Wu, Yong Yu, and Jieping Ye. Multi-agent reinforcement learning for order-dispatching via order-vehicle distribution matching. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 2645–2653. ACM, 2019.

[44] Xian Zhou, Yanyan Shen, Yanmin Zhu, and Linpeng Huang. Predicting multi-step citywide passenger demands using attention-based neural networks. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 736–744. ACM, 2018.

[45] Zhengyang Zhou. Attention based stack resnet for citywide traffic accident prediction. In *2019 20th IEEE International Conference on Mobile Data Management (MDM)*, pages 369–370. IEEE, 2019.

Yang Wang is now an associate professor at USTC. He got his Ph.D. degree at University of Science and Technology of China in 2007, under supervision of Professor Liusheng Huang. He also worked as a postdoc at USTC with Professor Liusheng Huang. His research interest mainly includes wireless (sensor) networks, distributed systems, data mining, and machine learning.

Zhengyang Zhou is now the master degree candidate in the School of Computer Science and Technology, University of Science and Technology of China. His research interests include machine learning, spatiotemporal data mining as well as artificial intelligence in traffic applications.

Kai Liu is now the master degree candidate in School of Software Engineering, University of Science and Technology of China. His research interests include data mining and machine learning.

Xike Xie is currently a research professor in the School of Computer Science and Technology, University of Science and Technology of China. His research interests include data uncertainty, spatiotemporal databases, and mobile computing. He is a member of ACM and IEEE.
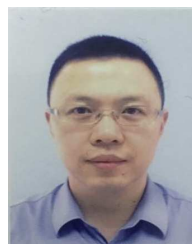
Wenhua Li is currently a traffic policeman in Traffic Administrative Agency of Suzhou Industry Park (SIP). He is now focusing on the traffic analysis and management of SIP.