

Fusing Spatio-Temporal Contexts into DeepFM for Taxi Pick-Up Area Recommendation

Yizhi Liu^{1,3}, Rutian Qing^{1,3}, Yijiang Zhao^{1,3,*}, Xuesong Wang^{1,3}, Zhuhua Liao^{1,3}, Qinghua Li^{1,2} and Buqing Cao^{1,3}

¹School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan, 411201, China

²Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL, 60208, USA

³Key Laboratory of Knowledge Processing and Networked Manufacturing in Hunan Province, Xiangtan, 411201, China

*Corresponding Author: Yijiang Zhao. Email: zhaoyj@hnust.edu.cn

Received: 09 July 2021; Accepted: 09 August 2021

Abstract: Short-term GPS data based taxi pick-up area recommendation can improve the efficiency and reduce the overheads. But how to alleviate sparsity and further enhance accuracy is still challenging. Addressing at these issues, we propose to fuse *spatio-temporal contexts into deep factorization machine (STC_DeepFM)* offline for pick-up area recommendation, and within the area to recommend pick-up points online using factorization machine (FM). Firstly, we divide the urban area into several grids with equal size. Spatio-temporal contexts are distilled from pick-up points or points-of-interest (POIs) belonged to the preceding grids. Secondly, the contexts are integrated into deep factorization machine (DeepFM) to mine high-order interaction relationships from grids. And a novel algorithm named *STC_DeepFM* is presented for offline pick-up area recommendation. Thirdly, we devise the architecture of offline-to-online (O2O) recommendation respectively based on DeepFM and FM model in order to trade-off the accuracy and efficiency. Some experiments are designed on the DiDi dataset to evaluate step by step the performance of spatio-temporal contexts, different recommendation models, and the O2O architecture. The results show that the proposed *STC_DeepFM* algorithm exceeds several state-of-the-art methods, and the O2O architecture achieves excellent real-time performance.

Keywords: Location-based service (LBS); trajectory data mining; offline-to-online (O2O) recommendation; deep factorization machine (DeepFM); spatio-temporal context

1 Introduction

Taxis are Global Positioning System (GPS) recorders of urban mobility. While offering online booking services, ride-hailing platforms such as DiDi and Uber also record the GPS data of taxis. These data have prompted lots of location-based services (LBS) [1], for example, pick-up areas or points recommendation. It is very useful for taxi-drivers to increase profits and decrease consumptions. Different from traditional recommender systems, pick-up area recommendation faces two new challenges: (1) Since more and more



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

GPS data occupy a lot of storage resources, the recommendation efficiency further goes down; (2) Due to the continuous renovation and updating of urban roads, too much historical data may contain too much noise and consequently reducing the recommendation performance. Therefore, some researchers presented taxi pick-up area recommendation based on short-term GPS data [2]. Nevertheless, the data sparseness arises owing to the difficulty of visiting many areas within a short time for taxis. Meanwhile, the recommendation accuracy may decrease because of too less data.

To alleviate data sparsity of recommender systems [3], a few state-of-the-art approaches have successively presented, such as matrix factorization (MF) [4], factorization machine (FM) [5], data fusion [6], attention mechanism [7], Neural Network [8], deep factorization machine (DeepFM) [9]. The traditional MF method [10,11] has been widely used due to its good performance and high scalability. It lacks effective use of contexts although more and more contextual features can be easily captured. Consequently, FM and DeepFM are put forward one after another and become popular owing to their excellent performance. Especially, DeepFM integrates the power of factorization machines for recommendation with deep learning for feature learning [12].

Nowadays, it is still challenging to alleviate sparsity and further enhance accuracy of recommend taxi pick-up areas or points recommendation. Researchers and corporations often concentrate on spatio-temporal contexts [13–15] and recommendation models [16–20]. Besides, a few two-stage recommender systems have been proved to obtain the balance between accuracy and efficiency [21–30].

Motivated by the outstanding behavior of DeepFM and two-stage recommender systems, we propose to fuse spatio-temporal contexts into DeepFM (*STC_DeepFM*) for taxi pick-up area recommendation. Firstly, we divide the urban area into several grids with equal size. Spatio-temporal contexts are distilled from pick-up points or points-of-interest (POIs) belonged to the preceding grids. Secondly, the contexts are integrated into deep factorization machine (DeepFM) to mine high-order interaction relationships from grids. And a novel algorithm named *STC_DeepFM* is presented for offline pick-up area recommendation. Thirdly, we devise the architecture of offline-to-online (O2O) recommendation respectively based on DeepFM and FM model in order to tradeoff the accuracy and efficiency. Our main contribution lies in:

- To tradeoff the accuracy and efficiency, we devise the architecture of offline-to-online (O2O) recommendation respectively based on DeepFM and FM model described in Fig. 1 and Section 4. In offline recommendation, we adopt the proposed *STC_DeepFM* algorithm to recommend a pick-up area. In online recommendation, we adopt the FM model to recommend a pick-up point based on the spatio-temporal analysis (STA) within the preceding pick-up area.
- We propose to learn sophisticated spatio-temporal interaction relationships among grids by the DeepFM model to further improve the prediction accuracy. Moreover, a novel algorithm *STC_DeepFM*, which is based on spatio-temporal interaction relationships, is put forward. Some works [31,32] have approved its effectiveness of integrating the spatial or temporal contexts into the FM model.
- Spatio-temporal contexts are proposed to fuse into the FM model or the DeepFM model in the way of feature engineering. Sophisticated feature interactions behind user behaviors are learned by the DeepFM model to improve the recommendation accuracy. And we integrate spatio-temporal analysis (STA) [33] with the FM model during pick-up points recommendation. Experiments on the DiDi dataset show that it can obviously improve the prediction accuracy and outperforms some state-of-the-art methods.

2 Related Work

Pick-up point or area recommendation is one of the eye-catching applications in urban computing and intelligent transportation systems. It is still challenging on how to accurately and quickly recommend pick-up

points or areas for taxi-users. Existing approaches mainly focus on spatio-temporal contexts [13–15] and recommendation models [16–20]. And there are some works on the framework of two-stage or two-step recommendation [21–30].

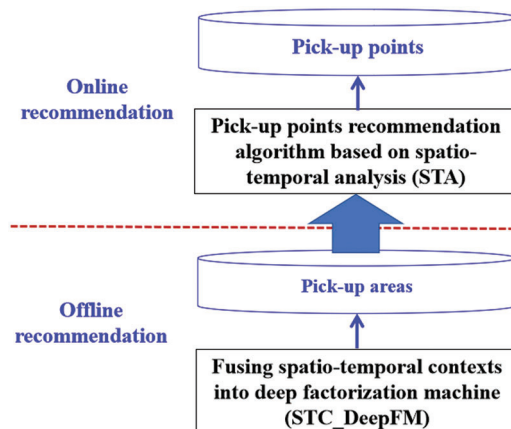


Figure 1: The proposed architecture of offline-to-online (O2O) recommendation

2.1 Recommendation Models

To cope with data sparseness in the field of recommender system [3], a few state-of-the-art approaches, such as Neural Network [8], matrix factorization [4], factorization machine (FM) [5], attention mechanism [7], deep factorization machine (DeepFM) [9], data fusion [6], have successively been put forward. The traditional matrix factorization method has been widely applied owing to its high scalability and good performance. In the big data era, more and more contextual features can be obtained easily, while the traditional matrix factorization approach lacks effective use of context information. Thus, FM is proposed and becoming popular. To further improve the recommendation accuracy, DeepFM is presented and achieves more excellent performance [12]. It integrates the power of factorization machines for recommendation with deep learning for feature learning. Other models in the literature [34–40] are also worthy to be considered in the future work.

Some researchers divided the study area into multiple same size grids, and partitioned the original data to the grids as their attributes, and then recommended hot areas based on their attributes to taxi drivers. This kind of area recommendation method is robust to noisy data and can find a better solution on the overall distribution of data. Kong et al. [2] predicted the passenger distribution of every grid by combining statistical learning with Gaussian process regression, and the regions with densely distributed passengers can be recommended for taxi drivers. Our team presented the GeoLFM model integrating geographic information with a latent factor model for hot pick-up area recommendation [20].

Most of the works concentrated on hotspot recommendation for taxi drivers. Li et al. suggested taxi driver whether they should wait or cruise based on their current location and time by L1-Norm support vector machines (SVM) [41]. Mu et al. proposed a hotspots recommendation model considering some spatial or temporal factors [42]. Huang et al. proposed a recommendation framework of searching passengers efficiently with multi-task deep learning [43].

2.2 Spatio-Temporal Contexts

Spatio-temporal context plays an important role except some recent works [20,31,32,44]. Liu et al. mined candidate pick-up points based on GPS data by spatio-temporal analysis, and created personalized

hot spots by integrating a probabilistic optimization model with project-based collaborative filtering method [33]. To maximize the possibility of getting passengers at shorter travel distances, they also applied road network to provide a range of hotspots for taxi drivers [45]. Zou et al. put forward a dynamic probability model for travel routes recommendation with high probability of carrying passengers to no-load taxis [46]. Liu et al. combined spatial scanning statistics into spatio-temporal hotspots analysis, and further recommended hotspots by searching potential areas with high reservation demand distribution [47].

Data mining technology is widely used for capturing spatio-temporal contexts. Hwang et al. proposed an OFF-ON graph model to represent the location relationship between the current drop-off passenger and the next get-on passenger [14]. Tang et al. used DBSCAN algorithm to cluster pick-up and drop-off points and then applied the Huff model to illustrate the attractiveness of pick-up regions [19]. Jain et al. devised an improved density-based clustering algorithm, which combined kd-tree with grid partitions to discover passenger hotspots [48]. Zhang et al. put forward a regional recommendation method by taking into account historical distributions of pick-up points and drivers' preferences [49].

2.3 Two-Stage or Two-Step Recommendation

The two-stage or two-step recommendation strategy has already applied in news recommendation [21,22], movies recommendation [23,24], e-commerce [25,26], Automatic playlist continuation [27], conversational recommendation [28], Session-based Recommendation [29], Cross-market Recommendation [30], etc. Knox et al. [21] proposed a scalable two-stage personalized news recommendation approach with a two-level representation in order to balance the novelty and diversity of the recommended result. Niu et al. [22] proposed a user based two-step recommendation algorithm with popularity normalization to improve recommendation diversity and novelty. Zhao, et al. [23,24] proposed the two-step recommendation approach of firstly recommending items and secondly finding high quality items which users like. Two-Stage Approach [25,26] achieved well-performed performance on online e-commerce retailer. Some other recommendation systems [27–30] devise the two-stage framework to obtain both effectiveness and efficiency.

Offline-to-online (O2O) recommendation is one of solutions of the two-stage or two-step recommendation strategy. Yuan et al. [16,17,50] proposed or adopted offline modeling and online recommendation to provide taxi drivers with some locations and the routes to these locations to pick up passengers quickly. Rossetti et al. [51] presented an approach of evaluating recommendation algorithms by contrasting offline and online results. Ding et al. [52] applied O2O recommendation to dynamically capture variation of users' purposes across time and location. Dong et al. [53] implemented a online and offline hybrid platform for English teaching recommendation based on reinforcement learning.

Some works [31,32] have approved its effectiveness of fusing spatial or temporal contexts into FM. These methods can be used to increase the prediction accuracy to some degree. However, they ignored low-order and high-order feature interaction which can further improve the accuracy. And a few two-stage recommender systems have been proved to obtain the balance between effectiveness and efficiency. One of its typical type is the framework of O2O recommendation. Furthermore, O2O recommendation has not been used for taxi pick-up area recommendation.

3 Our Proposed Algorithm “STC_DeepFM”

In this section, we illustrate our proposed recommendation algorithm “STC_DeepFM” (the abbreviation of fusing *Spatio-Temporal Context into DeepFM for taxi pick-up area recommendation*). It includes three parts: the framework of STC_DeepFM, spatio-temporal features extraction, spatio-temporal contexts and the proposed algorithm STC_DeepFM.

3.1 The Framework of *STC_DeepFM*

To accurately recommend pick-up areas for taxis, we propose to fuse spatio-temporal contexts into DeepFM (*STC_DeepFM*) for taxi pick-up area recommendation. The framework of *STC_DeepFM* is given in Fig. 2 and is illustrated as follows. In the framework, we focus on two parts, spatio-temporal contexts construction and fusion with recommendation model. Our dataset consists of GPS data, POI data, geographic information, and the current latitudes and longitudes of taxis.

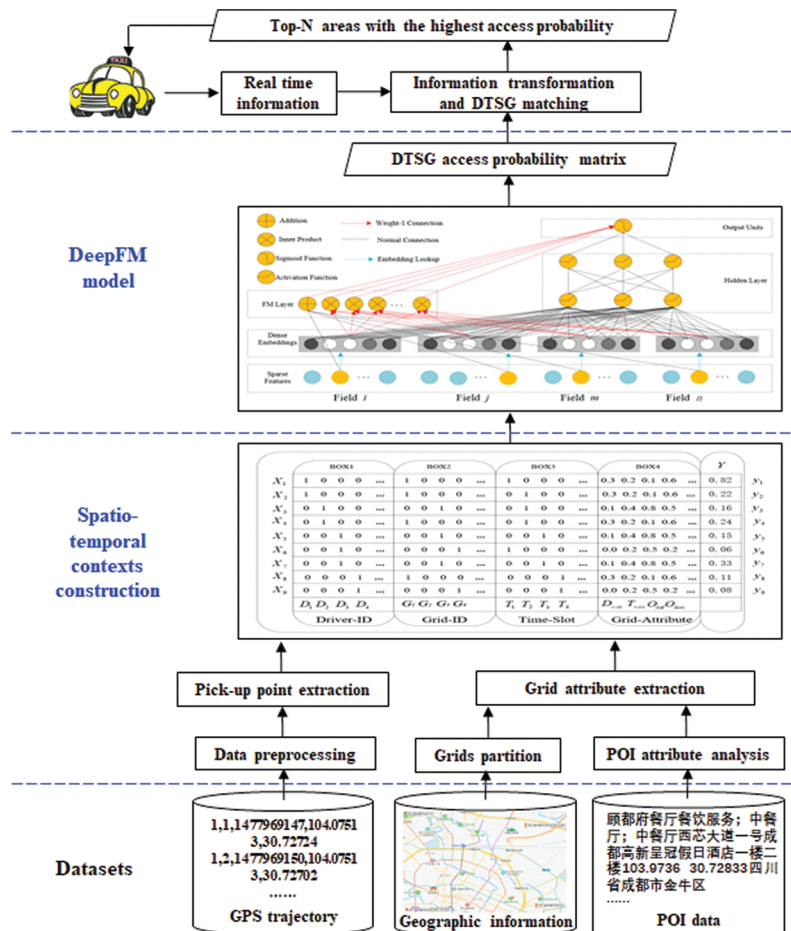


Figure 2: The framework of *STC_DeepFM*

- (1) Constructing spatio-temporal contexts. Firstly, the urban area is divided into the $n \times n$ regular grids [54] and the POI data of the city is classified into twelve types. Secondly, the GPS data are traversed and then the pick-up points are detected by computing the staying time of taxis. Next, those pick-up points and the POI data are mapped into one of the grids in light of their own longitudes and latitudes. Furthermore, we extract the spatio-temporal contexts from the grid's pick-up points and points-of-interest (POIs). The spatio-temporal contexts include the number of historical pick-up points, the number of POIs, the type value, the geometric center position of the grid, and the average driving time and distance after carrying passengers.
- (2) Fusing contexts into the recommendation model. The preceding spatio-temporal contexts are combined to be a multi-dimensional feature vector matrix and fused in the way of feature engineering into the DeepFM model. The DeepFM model integrates the power of factorization

machines for recommendation with deep learning for feature learning and achieves more excellent performance. It is able to learn the high-order interaction relationships among the grids. The multi-dimensional feature vector matrix consists of the driver-ID feature vector matrix, the grid-ID feature vector matrix, the time-slot vector matrix, and the grid-attribute vector matrix. It is put into the DeepFM model to learn the high-order interaction relationships among the grids. For example, the relations among drivers and time-slots or among time-slots and grids are deeply mined. Therefore, these relations can reflect the predicted value of the drivers' access probability to different grids in different time-slots. The driver-time slot-grid (DTSG) access probability matrix is generated by using the trained DeepFM model. The elements in the matrix are the driver's access probability prediction value in the corresponding slot and the corresponding grid. According to the probability matrix and the drivers' current information, the grids with the top-N results will be recommended to the taxi-drivers.

3.2 Spatio-Temporal Features Extraction

We adopt a map grid-based method to divide the study area into grid-like. The determination of grid size has a crucial impact on the accuracy and efficiency of taxi pick-up area recommendation. According to the grid size related research in the literature [55], we finally decide to divide the study area into grids of size 300 m × 300 m. Referring to the existing research [12,20,28,49] and analysis, there are some regional attributes of each grid that affect the accuracy of recommendation, such as the functional type of a region, the number of pickups in history, the number of POIs, the geometric center of an area, the average travel distance and time after pickup in an area. After obtaining the pickup points from taxi trajectory, both the pickup points and POIs are mapped into corresponding grids according to their location. Then, the attributes of each grid are calculated. The grid features selected for passenger-carrying area recommendation are as follows.

The first kind of grid features is timeslots and geographic locations. In the domain of LBS, timeslots and geographic locations are very important temporal and spatial factors, respectively. It is always easier for taxi drivers to find potential passengers in hotspots than in non-hotspots. Besides, near department stores or commercial districts where high demand for taxis usually occurs during the day, but usually near bars at night. Thus, timeslots and geographic locations are very important attributes of each grid. Herein, a day is divided into 24 timeslots with one timeslot per hour, and the geographic location of a grid is represented by its geometric center.

The second kind of grid features is the number of historical pickup points. The number of historical passengers in a region is a type of crucial information for taxi drivers to find passengers. Historical pickup numbers vary across grids. It reflects the difference in the number of passengers carried between different regions, that is, the impact of spatial factors on it. In general, areas with high historical pickup points are likely to have more potential passengers. Their quantity represents the load demand of the grids to a certain extent.

The third kind of grid features is the average travel time and distance after pickup by region. The taxi-drivers' income is mainly determined by the driving time and distance after carrying passengers. The average travel time and distance after carrying passenger in the area are respectively computed as follows:

$$T_{i-on} = \frac{1}{Num} \sum_{j=1}^{Num} (t_{j(on)} - t_{j(off)}) (psegment_{i-on}) \quad (1)$$

$$D_{i-on} = \frac{1}{Num} \sum_{j=1}^{Num} distance_j (psegment_{i-on}) \quad (2)$$

In the preceding formula, psegment represents the trajectory set of all taxis under load; psegmenti-on is all the trajectory segments that carried passengers in area Gi; Num means the number of these trajectories; tj(on) is the time when the j-th trajectory segment start; tj(off) is the time when the j-th trajectory segment end; distancej is the total length of the j-th trajectory segment.

The forth kind of grid features is the area type. The area type can be regarded as different functional area, such as residential and commercial area, by computing the numbers of various POIs' types in the area. For instance, if the number of residential POIs in the area is large, the area can be reckoned as residential. Furthermore, the number of passengers always emerges in the early and late peaks of the working day. The area type of a grid can be computed by the number distribution of every POI type in the grid. The ratio of the number of each POI's type to the total amount of all POIs is computed as its type index of the POI subtype in the area. The ratio of each subtype of POIs is computed as the following formula, in which $|POI_{typej}|$ means the number of POIs belonging to the type j .

$$type_{j(i)} = \frac{|POI_{typej}|}{\left| \sum_{i=1}^n POI_{typei} \right|} \quad (3)$$

3.3 Spatio-Temporal Contexts and the Proposed Algorithm STC_DeepFM

This section shows how to extract implicit but valuable features in the raw dataset and convert the features as the input for DeepFM model. It contains five vectors: driver-ID vector, grid-ID vector, time-slot vector, grid-attribute vector, and target-score vector. (1) We assume that D_i represents a driver-ID, $D = \{D_1, D_2, D_3, \dots, D_N\}$ is a vector of one-hot code. N is the dimension of D is N . For instance, the one-hot code for no. 1 taxi driver is $D_1 = [1, 0, 0, \dots, 0]_{1 \times N}$. (2) We use G to represent the grid-ID, whose one-hot code can be denoted as $G = \{G_1, G_2, G_3, \dots, G_N\}$. In this paper, the grid can also be referred to as area. The grid is the place where the order occurs. It is also the recommendation target. (3) We use T to represent the timeslot, whose one-hot code denotes the driver's access time to the area, can be represented as $T = \{T_1, T_2, T_3, \dots, T_{24}\}$. (4) we define A as the grid-attribute, which includes several attributes in every grid: the number of historical pickup points, the average travel time and distance after pickup by region, POI amount, distribution indexes of 12 POI subtypes, geometric center (O_{lat} and O_{lon} represent its latitude and longitude, respectively). These attributes have different influences on the accuracy of taxi passenger carrying area recommendation. The data of this part will be normalized before used as input data of DeepFM. (5) Target score S of a region is calculated by normalizing the number of passengers picked up in their regions in each driver's slot. The output of DeepFM is the predicted grid access probability, which is regarded as the recommendation score of every pickup area. We distil spatio-temporal contexts from the original dataset and create the input matrix.

The DeepFM model mainly contains two parts: **deep neural network (DNN)** and **factorization machine (FM)**. Both the two parts share the same input vector [9]. DNN and FM can respectively learn feature interactions of high-order and low-order from raw data. Consequently, DeepFM can extract both low-order and high-order features between data. The DeepFM model is applied to predict the probability of passenger occupancy for all candidate regions. According to the pick-up probability, we can provide high-quality service of pickup points recommendation for drivers.

The proposed **STC_DeepFM** algorithm given in Algorithm 1 shows the process of fusing **spatio-temporal context into DeepFM** for recommending pick-up area. It is based on a geographical location. In order to make the driver make greater profits, set a threshold of θ for the search distance *dist* between the driver's current location and the recommendation grid, recommended for the driver when the search distance between the recommended grid is less than θ When the distance *dist* is less than θ , the grid can

be used as the candidate area recommended by the driver. Herein, we compute Euclidean distance between the geometric centers of any two grids.

Algorithm 1 STC_DeepFM

Input: Trajectory t_1, t_2, \dots, t_n ; side length of grid l ; POI set p_1, p_2, \dots, p_n ; the grid of current position g and time slot t of driver d

Output: the **driver-time slot-grid** (DTSG) access probability matrix, the recommendation list (L)

1. *GridSet* $G = \text{GridDivide}(l)$
 2. *for* t_i *in* *Trajectory*
 - for* G_j *in* *GridSet*
 - Traverse the POIset*
 - Construct Multidimensional features* O
 - end*
 - end*
 3. $\text{DTSG} = \text{DeepFM}(O)$
 4. $\text{DTSG} = \text{Filtrate}(\text{DTSG})$ when $D = d$ and $T = t$
 5. $L = \text{ListRecommend}(\text{DTSG})$ when $\text{dist} < \theta$
 6. *return* *Top* (*Sort* (L))
-

4 The Architecture of O2O Recommendation

With the increase of the grid size in the urban city, the trajectory points in each grid are increased so that the computation overhead on the spatio-temporal attributes of each grid increases. The increase of the precision of grids' attributes leads to the improvement of the recommendation accuracy. Meanwhile the fault tolerance range of the pick-up area recommendation enhances due to the increase of the grid size.

To further improve the performance of real-time recommendation, we devise a framework of offline-to-online (O2O) recommendation shown in Fig. 1. It takes part in offline recommendation and online recommendation. The principle of the architecture is represented in Fig. 3. In the part of offline recommendation, we propose a novel approach of *fusing spatio-temporal contexts into DeepFM (STC_DeepFM)* for taxi pick-up area recommendation, which is illustrated in Section 4. Nevertheless, it takes more time to recommend a smaller grid for drivers. Consequently, a bigger grid is recommended by offline recommendation (Fig. 3a), and then online recommendation is implemented in the bigger grid to recommend a point (Fig. 3b) more quickly and more accurately than ever before.

5 Experiments

We gather two types of data in the experiments: the GPS data from Didi GAIA Open Dataset, and about 340 thousand POI entries crawled from Amap. Each piece of POI entry consists of name, type, latitude, longitude, etc. The classification tags of POIs are also derived from Amap. There are totally twelve types: tourist attraction, food and beverages, communal facilities, science and culture, means of transportation, finance and insurance, serviced apartment, shopping service, life services, health care, sports leisure, and residential services. The GPS trajectory data collected from taxis within the second ring road in Chengdu in October 2016, with a sampling frequency of 2 to 4 s. Each GPS data entry mainly includes fields such as order ID, taxi ID, latitude, longitude and timestamp.

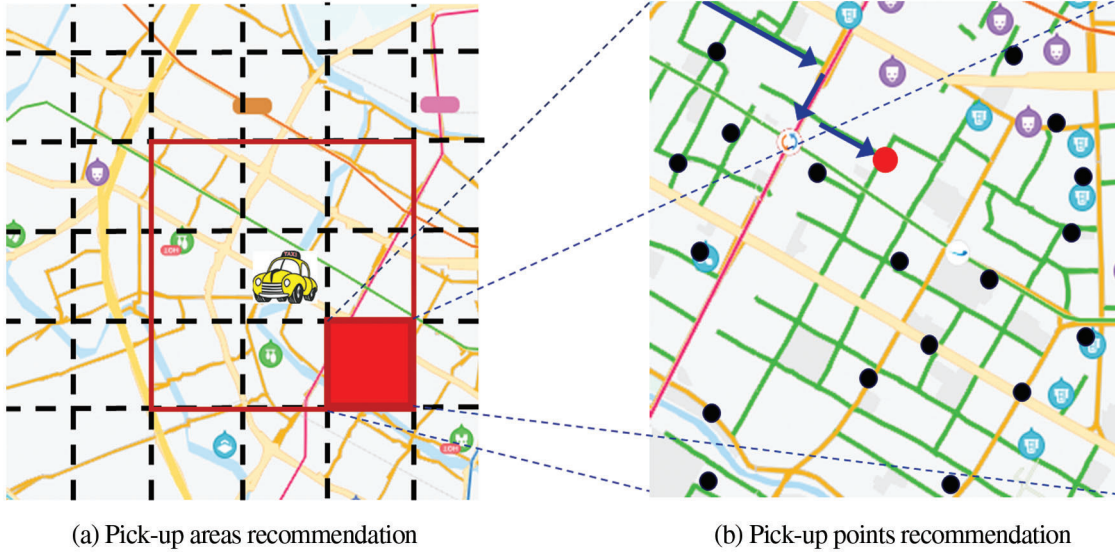


Figure 3: (a) Pick-up areas recommended by offline recommendation, and (b) pick-up points recommended by online recommendation

5.1 Evaluation on STC_DeepFM for Taxi Pick-up Area Recommendation

To promote the recommendation performance, they are preprocessed, for instance, eliminating redundant and noisy points, map offset processing. We divide the urban area into $300\text{ m} \times 300\text{ m}$ size grids by its latitude and longitude. We mainly adopt *mean absolute error (MAE)* and *root mean square error (RMSE)* as the metrics for performance evaluation. Both of them are used to evaluate the error between predicted value and ground truth. Therefore, if the MAE and RMSE value are smaller, the recommendation accuracy is higher, and vice versa.

We carefully construct four feature matrices: *driver-ID D*, *grid-ID G*, *time-slot T*, and *grid-attribute A*. The driver-ID matrix D contains the ID numbers of corresponding drivers. The Grid-ID matrix G is the target grids of corresponding drivers. The two former matrices D and G are necessary to the input vector. Its target value is the matrix $Y_{D,G}$. $Y_{D,G}$ is calculated by normalizing the number of passengers carried by the driver D on the grid G . The input vector is described as $\{D, G, Y_{D,G}\}$, which means the driver D 's historical carrying passenger on the grid G . The matrix T represents the temporal contexts, and its target value and its input vector are respectively $Y_{D,G,T}$ and $\{D, G, T, Y_{D,G,T}\}$. $Y_{D,G,T}$ is obtained by figuring out the number of passengers carried by the driver D on the grid G within the time slot T . The input vector $\{D, G, T, Y_{D,G,T}\}$ represents driver D 's historical carrying passenger on grid G within the time-slot T . The matrix A describes the spatial contexts, and its target value and its input vector are respectively $Y_{D,G,A}$ and $\{D, G, A, Y_{D,G,A}\}$. The input vector $\{D, G, A, Y_{D,G,A}\}$ refers that driver D drives to grid G whose attributes are described by the matrix A . To observe the benefits of the temporal or spatial contexts, we respectively add T or A or both into the input vector. Fig. 3 gives an example of the combination of both spatial and temporal contexts. The input vector $\{D, G, T, A, Y_{D,G,T,A}\}$ refers that driver D drives to grid G whose attributes are represented by the matrix A within the time-slot T .

The experimental results after fusing different contexts combinations are shown in Tab. 1. The performance improves to varying degrees after adding T to $\{D, G, Y_{D,G}\}$ or $\{D, G, A, Y_{D,G,A}\}$. Similarly, the performance has been enhanced after respectively embedding A into $\{D, G, Y_{D,G}\}$ or $\{D, G, T, Y_{D,G,T}\}$. The performance of $\{D, G, T, A, Y_{D,G,T,A}\}$ is the best. And time-slot T does better than grid-attribute A . *Explained variance score* is adopted as the metric of rating index. It measures the ability

that our model interprets dataset fluctuations. The best score of a model could be 1.0, which means the model is perfect. Consequently, we can draw three conclusions: (1) Time-slot matrix T or grid-attribute matrix A is able to accurately describe the temporal or spatial contexts why the driver D chooses the grid G . (2) The spatio-temporal contexts can obviously improve the recommendation performance. (3) Our method obtains excellent performance, that is MAE and RMSE are respectively 1.06% and 6.89%. And the explained variance achieves 98.09%.

Table 1: Performance evaluation on different combinations among contexts

	$\{D, G, Y_{D,G}\}$	$\{D, G, A, Y_{D,G,A}\}$	$\{D, G, T, Y_{D,G,T}\}$	$\{D, G, T, A, Y_{D,G,T,A}\}$
MAE	18.43%	5.66%	1.35%	1.06%
RMSE	20.12%	10.11%	7.17%	6.89%
explained_variance	85.31%	95.91%	97.93%	98.09%

To evaluate the advantage of our approach, we compare with five typical or state-of-the-art methods in the domain, such as *User-based collaborative filtering (UBCF)* [18], *A two-layer approach (ATLA)* [19], *Recommendation based on time-location-relationship (RTLRL)* [2], *Fusing geographic information into latent factor model (GeoLFM)* [20], and *Taxi pick-up area recommendation based on factorization machine (FM)* [5].

Owing to data sparseness, the performances of RTLRL, UBCF, ATLA are still not good enough. To deal with data sparseness, the GeoLFM and FM model are successively adopted. The former one improves the utilization of data to a certain extent by matrix decomposition. The latter one has a certain anti-thinning characteristic owing to low-order features extraction. Thus, we merge the spatio-temporal contexts into the FM model and observe that it is better than all the other preceding methods. Because the DeepFM model utilizes not only end-to-end low-order but also high-order feature interactive learning, we fuse spatio-temporal contexts into DeepFM to further improve the recommendation accuracy. Multi-dimensional features, such as pickup points and POIs distribution, are combined to illustrate the spatial and temporal attributes of every region where taxi drivers pick up passengers in the real world. In light of the excellent effect of the time-slot matrix T , we divide all datasets into weekdays and weekends to further improve the recommendation performance. The (a) and (b) in Fig. 4 represent the MAE and RMSE of recommend results of the six methods during both weekdays and weekends, respectively. Comparing with UBCF, ATLA, RTLRL, GeoLFM, FM, the MAE and RMSE of our method is the lowest whether weekdays or weekends.

5.2 Evaluation on the O2O Recommendation Framework

To further improve the performance of real-time recommendation, we devise an architecture of offline and online recommendation respectively based on DeepFM and FM model. Some experiments are designed to evaluate its performance. We adopt the proposed STC_DeepFM algorithm to recommend pick-up area offline. It is combined with some typical pick-up point recommendation algorithms, such as *the top-k recommendation algorithm (Top-k)*, *the item-based collaborative filtering algorithm (IBCF)*, *the combination of the former two algorithms (Top-k+IBCF)*, and *pick-up point recommendation based on factorization machine (FM)*. We compare them by the accuracy and the cost time of pick-up point recommendation, as shown in Tabs. 2 and 3.

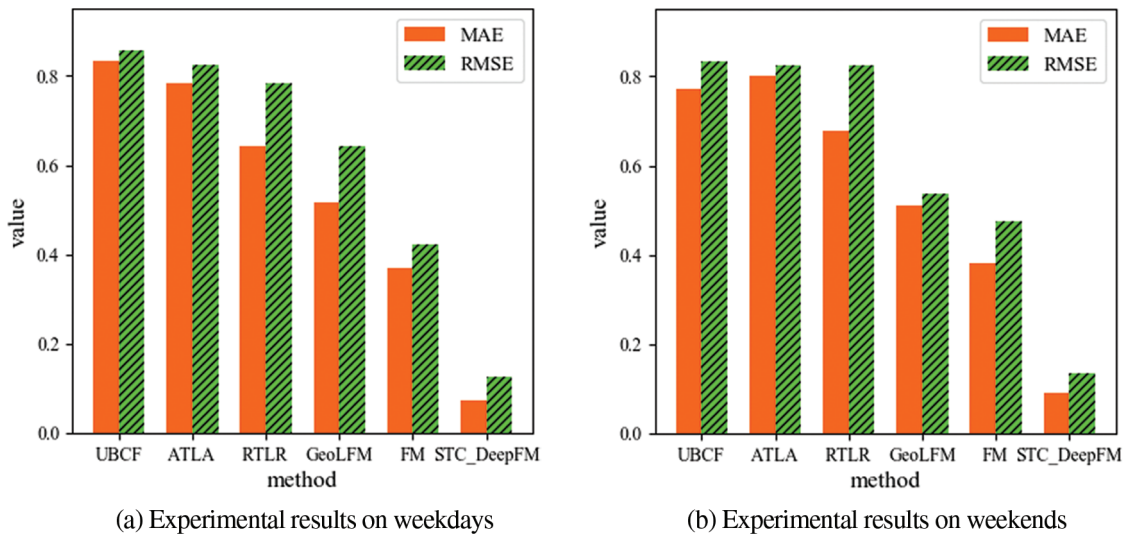


Figure 4: Comparisons of the performance among some state-of-the-art methods

Table 2: Comparisons among the accuracy of different online recommendation algorithms combined with FM or STC_DeepFM offline recommendation

	FM offline recommendation	STC_DeepFM offline recommendation
Top-k online recommendation	71.57%	76.46%
IBCF online recommendation	85.7%	83.57%
Top-k+IBCF online recommendation	88.47%	89.02%
FM online recommendation	90.26%	92.73%

Table 3: Comparisons among the cost time of different online recommendation algorithms combined with STC_DeepFM offline recommendation

Online recommendation	Top-k	IBCF	Top-k + IBCF	FM
The cost time (seconds)	1.046	3.126	3.473	1.852

6 Conclusion

In this paper, a novel approach of fusing spatio-temporal contexts into DeepFM is proposed to accurately recommend pick-up areas for taxi-drivers. Firstly, the urban region is divided into many grids with equal size, the pickup points are obtained from GPS data. The pickup points and the points-of-interest (POIs) are mapped to the grids. Then the spatial or temporal attributes are extracted from these grids to create a multi-dimensional feature vector. A multi-dimensional feature vector can be constructed by different feature vector combinations. Secondly, the multi-dimensional feature vector is input into the DeepFM model, which realize end-to-end and both low-order and high-order feature interactive learning. Therefore, it is able to greatly promote the quality of pickup area recommendation for taxi drivers. We estimate respectively the performance of spatio-temporal contexts and different recommendation methods on the real datasets. The experimental results show that the proposed *STC_DeepFM* algorithm greatly

improves the recommendation performance of pick-up areas. It is superior to many state-of-the-art methods, such as user-based collaborative filtering (UBCF), a two-layer approach (ATLA), recommendation based on time-location-relationship (RTLRL), fusing geographic information into latent factor model (GeoLFM), factorization machine (FM). To further improve the performance of real-time recommendation, we devise the O2O recommendation architecture respectively based on DeepFM and FM model. Experimental results show that the O2O recommendation architecture achieves excellent real-time performance. In future work, we will focus on mining more useful information from GPS data or via better recommendation models to improve the performance.

Funding Statement: This work was supported by the National Natural Science Foundation of China (41871320, 61873316), the Key Project of Hunan Provincial Education Department (19A172), the Scientific Research Fund of Hunan Provincial Education Department (18K060), and the Postgraduate Scientific Research Innovation Project of Hunan Province (CX20211000).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Y. H. Wang, Y. B. Sun, S. Su, Z. H. Tian, M. H. Li *et al.*, “Location privacy in device-dependent location-based services: Challenges and solution,” *Computers, Materials & Continua*, vol. 59, no. 3, pp. 983–994, 2019. <https://doi.org/10.32604/cmc.2019.05547>.
- [2] X. J. Kong, F. Xia, J. Z. Wang, A. Rahim and S. K. Das, “Time-location-relationship combined service recommendation based on taxi trajectory data,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1202–1212, 2017. <https://doi.org/10.1109/TII.2017.2684163>.
- [3] K. K. Zhao, L. F. Zhang, J. Zhang, C. P. Li and H. Chen, “Survey on factorization machines model,” *Journal of Software*, vol. 30, no. 3, pp. 799–821, 2019. (in Chinese) <https://doi.org/10.13328/j.cnki.jos.005698>.
- [4] S. B. Min, H. T. Yao, H. T. Xie, Z. Zha and Y. D. Zhang, “Multi-objective matrix normalization for fine-grained visual recognition,” *IEEE Transactions on Image Processing*, vol. 29, pp. 4996–5009, 2020. <https://doi.org/10.1109/TIP.2020.2977457>.
- [5] S. Rendle, “Factorization machines,” in *Proc. ICDM*, Sydney, Australia, pp. 995–1000, 2010. <https://doi.org/10.1109/ICDM.2010.127>.
- [6] W. Liang, L. J. Xiao, K. Zhang, D. F. Zhang, M. D. Tang *et al.*, “Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems,” *IEEE Internet of Things Journal*, pp. 1–1, 2021. <https://doi.org/10.1109/JIOT.2021.3053842>.
- [7] C. B. Liu, H. T. Xie, Z. J. Zha, L. Y. Yu, Z. N. Chen *et al.*, “Bidirectional attention-recognition model for fine-grained object classification,” *IEEE Transactions Multimedia*, vol. 22, no. 7, pp. 1785–1795, 2020. <https://doi.org/10.1109/TMM.2019.2954747>.
- [8] Y. H. Du, Y. Yang, W. Zhang and Y. Taketoshi, “DeepRec: A deep neural network approach to recommendation with item embedding and weighted loss function,” *Information Sciences*, vol. 470, pp. 121–140, 2019. <https://doi.org/10.1016/j.ins.2018.08.039>.
- [9] H. F. Guo, R. M. Tang, Y. M. Ye, Z. G. Li and X. Q. He, “DeepFM: A factorization-machine based neural network for CTR prediction,” in *Proc. IJCAI*, Melbourne, Australia, pp. 1725–1731, 2017. <https://doi.org/10.24963/ijcai.2017/239>.
- [10] H. T. Bai, X. Li, L. L. He, L. H. Jin, C. Wang *et al.*, “Recommendation algorithm based on probabilistic matrix factorization with adaboost,” *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1591–1603, 2020. <https://doi.org/10.32604/cmc.2020.09981>.
- [11] G. S. Liu, K. Meng, J. C. Ding, J. P. Nees, H. Y. Guo *et al.*, “An entity-association-based matrix factorization recommendation algorithm,” *Computers, Materials & Continua*, vol. 58, no. 1, pp. 101–120, 2019. <https://doi.org/10.32604/cmc.2019.03898>.

- [12] X. S. Wang, Y. Z. Liu, Z. H. Liao and Y. J. Zhao, "DeepFM-Based taxi pick-up area recommendation," in *Proc. ICPR*, Milan, Italy, pp. 407–421, 2021. https://doi.org/10.1007/978-3-030-68821-9_36.
- [13] H. Chang, Y. Tai and J. Y. Hsu, "Context-aware taxi demand hotspots prediction," *International Journal of Business Intelligence and Data Mining*, vol. 5, no. 1, pp. 3–18, 2010. <https://doi.org/10.1504/IJBIDM.2010.030296>.
- [14] R. H. Hwang, Y. L. Hsueh and Y. T. Chen, "An effective taxi recommender system based on a spatiotemporal factor analysis model," *Information Sciences*, vol. 314, pp. 28–40, 2015. <https://doi.org/10.1016/j.ins.2015.03.068>.
- [15] J. W. Powell, Y. Huang, F. Bastani and M. Ji, *Towards Reducing Taxi-Cab Cruising Time using Spatio-Temporal Profitability Maps Advances in Spatial and Temporal Databases*, Springer-Verlag, Berlin, Heidelberg, Germany, pp. 242–260, 2011. https://doi.org/10.1007/978-3-642-22922-0_15.
- [16] J. Yuan, Y. Zheng, L. H. Zhang, X. Xie and G. Z. Sun, "Where to find my next passenger," in *Proc. UbiComp*, Beijing, China, pp. 109–118, 2011. <https://doi.org/10.1145/2030112.2030128>.
- [17] N. J. Yuan, Y. Zheng, L. H. Zhang and X. Xie, "T-Finder: A recommender system for finding passengers and vacant taxis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2390–2403, 2013. <https://doi.org/10.1109/TKDE.2012.153>.
- [18] C. Wang, Z. G. Zhu, Y. X. Zhang and F. F. Su, "Recommendation efficiency and personalized improvement of user-based collaborative filtering algorithm," *Mini-Micro Systems*, vol. 37, no. 3, pp. 428–432, 2016. (in Chinese).
- [19] J. J. Tang, H. Jiang, Z. B. Li, M. Li, F. Liu *et al.*, "A Two-layer model for taxi customer searching behaviors using GPS trajectory data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3318–3324, 2016. <https://doi.org/10.1109/TITS.2016.2544140>.
- [20] Z. H. Liao, J. Zhang, Y. Z. Liu, H. Xiao, Y. J. Zhao *et al.*, "Fusing geographic information into latent factor model for pick-up region recommendation," in *Proc. ICME Workshops*, Shanghai, China, pp. 330–335, 2019. <https://doi.org/10.1109/ICMEW.2019.00-65>.
- [21] D. Knox, L. Li, D. D. Wang, T. Li and B. Padmanabhan, "Scene: A scalable two-stage personalized news recommendation system," in *Proc. SIGIR*, Beijing, China, 2011. <https://doi.org/10.1145/2009916.2009937>.
- [22] K. Niu, W. Chen, N. Li, X. P. Peng, X. Y. Zhao *et al.*, "UTSP: User-based two-step recommendation with popularity normalization towards diversity and novelty," *IEEE Access*, vol. 7, pp. 145426–145434, 2016. <https://doi.org/10.1109/ACCESS.2019.2939945>.
- [23] X. Y. Zhao, Z. D. Niu and W. Chen, "Interest before liking: Two-step recommendation approaches," *Knowledge-Based Systems*, vol. 48, pp. 46–56, 2013. <https://doi.org/10.1016/j.knsys.2013.04.009>.
- [24] X. Y. Zhao, Z. D. Niu, W. Chen, C. Y. Shi, K. Niu *et al.*, "A hybrid approach of topic model and matrix factorization based on two-step recommendation framework," *Journal of Intelligent Information Systems*, vol. 44, no. 3, pp. 335–353, 2015. <https://doi.org/10.1007/s10844-014-0334-3>.
- [25] M. Volkovs, "Two-stage approach to item recommendation from user sessions," in *Proc. RecSys'15*, Vienna, Austria, no. 3, pp. 1–4, 2015. <https://doi.org/10.1145/2813448.2813512>.
- [26] B. Smith and G. Linden, "Two decades of recommender systems at Amazon.com," *IEEE Internet Computing*, vol. 21, no. 3, pp. 12–18, 2017. <https://doi.org/10.1109/MIC.2017.72>.
- [27] M. Volkovs, H. Rai, Z. Y. Cheng, G. Wu, Y. H. Lu *et al.*, "Two-stage model for automatic playlist continuation at scale," in *Proc. RecSys Challenge'18*, Vancouver, Canada, no. 9, pp. 1–6, 2018. <https://doi.org/10.1145/3267471.3267480>.
- [28] K. Christakopoulou, A. Beutel, R. Li, S. Jain and E. H. Chi, "Q&R: A two-stage approach toward interactive recommendation," in *Proc. KDD'18*, London, United Kingdom, pp. 139–148, 2018. <https://doi.org/10.1145/3219819.3219894>.
- [29] M. Khandwawala, J. A. S. Rodríguez and J. C. Wu, "Two-stage session-based recommendations with candidate rank embeddings," in *Proc. Fashion RECSYS Workshop recsysXfashion'19*, Copenhagen, Denmark, pp. 49–66, 2019. <https://doi.org/10.48550/arXiv.1908.08284>.
- [30] Z. Y. Chen, H. Wang, X. Y. Zhu, H. Y. Wu, C. C. Gu *et al.*, "A practical two-stage ranking framework for cross-market recommendation," in *Proc. WSDM'22*, Phoenix, AZ, USA, pp. 1–4, 2022. <https://doi.org/10.48550/arXiv.2204.12682>.

- [31] Y. Koren, “Collaborative filtering with temporal dynamics,” *Communications of the ACM*, vol. 53, no. 4, pp. 89–97, 2010. <https://doi.org/10.1145/1721654.1721677>.
- [32] P. Symeonidis, A. Papadimitriou, Y. Manolopoulos, P. Senkul and I. H. Toroslu, “Geo-social recommendations based on incremental tensor reduction and local path traversal,” in *Proc. GIS-LBSN*, Chicago, IL, USA, pp. 89–96, 2011. <https://doi.org/10.1145/2063212.2063228>.
- [33] Y. Z. Liu, J. X. Liu, Z. H. Liao, M. D. Tang and J. J. Chen, “Recommending a personalized sequence of pick-up points,” *Journal of Computational Science*, vol. 28, pp. 382–388, 2018. <https://doi.org/10.1016/j.jocs.2017.05.004>.
- [34] L. Y. Xiang, S. H. Yang, Y. H. Liu, Q. Li and C. Z. Zhu, “Novel linguistic steganography based on character-level text generation,” *Mathematics*, vol. 8, pp. 1558, 2020. <https://doi.org/10.3390/math8091558>.
- [35] L. Y. Xiang, G. Q. Guo, Q. Li, C. Z. Zhu, J. R. Chen *et al.*, “Spam detection in reviews using LSTM-based multi-entity temporal features,” *Intelligent Automation & Soft Computing*, vol. 26, no. 6, pp. 1375–1390, 2020. <https://doi.org/10.32604/iasc.2020.013382>.
- [36] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He and X. Chen, “TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. <https://doi.org/10.1109/TITS.2021.3130403>.
- [37] W. Sun, L. Dai, X. R. Zhang, P. S. Chang and X. Z. He, “RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring,” *Applied Intelligence*, vol. 52, pp. 1–16, 2021. <https://doi.org/10.1007/s10489-021-02893-3>.
- [38] X. C. Liu, W. Liu, T. Mei and H. D. Ma, “PROVID: Progressive and multimodal vehicle reidentification for large-scale urban surveillance,” *IEEE Transactions on Multimedia*, vol. 20, no. 3, pp. 645–658, 2018. <https://doi.org/10.1109/TMM.2017.2751966>.
- [39] W. Sun, X. Chen, X. R. Zhang, G. Z. Dai, P. S. Chang *et al.*, “A Multi-feature learning model with enhanced local attention for vehicle re-identification,” *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3549–3561, 2021. <https://doi.org/10.32604/cmc.2021.021627>.
- [40] W. Sun, G. C. Zhang, X. R. Zhang, X. Zhang and N. N. Ge, “Fine-grained vehicle type classification using lightweight convolutional neural network with feature optimization and joint learning strategy,” *Multimedia Tools and Applications*, vol. 80, pp. 30803–30816, 2021. <https://doi.org/10.1007/s11042-020-09171-3>.
- [41] B. Li, D. Q. Zhang, L. Sun, C. Chen, S. J. Li *et al.*, “Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset,” in *Proc. PerCom Workshops*, Seattle, USA, pp. 63–68, 2011. <https://doi.org/10.1109/PERCOMW.2011.5766967>.
- [42] B. Mu and M. Dai, “Recommend taxi pick-up hotspots based on density-based clustering,” in *Proc. CCET*, Beijing, China, pp. 176–181, 2019. <https://doi.org/10.1109/CCET48361.2019.8989132>.
- [43] Z. H. Huang, J. Y. Tang, G. X. Shan, J. Ni, Y. W. Chen *et al.*, “An efficient passenger-hunting recommendation framework with multitask deep learning,” *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7713–7721, 2019. <https://doi.org/10.1109/JIOT.2019.2901759>.
- [44] Z. H. Liao, H. Xiao, S. L. Liu, Y. Z. Liu and A. P. Yi, “Impact assessing of traffic lights via GPS vehicle trajectories,” *ISPRS International Journal of Geo-Information*, vol. 10, no. 11, pp. 1–13, 2021. <https://doi.org/10.3390/ijgi10110769>.
- [45] Y. Z. Liu, R. T. Qing, J. X. Liu, Z. H. Liao, Y. J. Zhao *et al.*, “Extracting campus’ road network from walking GPS trajectories,” *Journal of Cyber Security*, vol. 2, no. 3, pp. 131–140, 2020. <https://doi.org/10.32604/jcs.2020.010625>.
- [46] Q. N. Zou, G. T. Xue, Y. Luo, J. D. Yu and H. Z. Zhu, “A novel taxi dispatch system for smart city,” in *Proc. DAPI (International Conference on Distributed, Ambient, and Pervasive Interactions)*, Las Vegas, USA, pp. 326–335, 2013. https://doi.org/10.1007/978-3-642-39351-8_36.
- [47] Y. Z. Liu, X. S. Wang, J. X. Liu, Z. H. Liao, Y. J. Zhao *et al.*, “An entropy-based model for recommendation of taxis’ cruising route,” *Journal on Artificial Intelligence*, vol. 2, no. 3, pp. 137–148, 2020. <https://doi.org/10.32604/jai.2020.010620>.
- [48] R. Jain, S. Garg, S. Gangal and M. K. Thakur, “TaxiScan: A scan statistics approach for detecting taxi demand hotspots,” in *Proc. IC3*, Noida, India, pp. 1–6, 2019. <https://doi.org/10.1109/IC3.2019.8844898>.

- [49] L. Zhang, C. L. Chen, Y. Y. Wang and X. P. Guan, "Exploiting taxi demand hotspots based on vehicular big data analytics," in *Proc. VTC-Fall*, Montreal, QC, pp. 1–5, 2016. <https://doi.org/10.1109/VTCFall.2016.7881010>.
- [50] J. Bao, Y. Zheng and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proc. SIGSPATIAL'12*, Redondo Beach, California, pp. 199–208, 2012. <https://doi.org/10.1145/2424321.2424348>.
- [51] M. Rossetti, F. Stella and M. Zanker, "Contrasting offline and online results when evaluating recommendation algorithms," in *Proc. RecSys'16*, Boston, MA, USA, pp. 31–34, 2016. <https://doi.org/10.1145/2959100.2959176>.
- [52] X. C. Ding, F. Shi, C. Xu, Q. X. Jiang, T. Liu *et al.*, "Infer implicit contexts in real-time online-to-offline recommendation," in *Proc. KDD'19*, Anchorage, AK, USA, pp. 2336–2346, 2019. <https://doi.org/10.48550/arXiv.1907.04924>.
- [53] D. L. Dong and L. B. Wu, "Implementation of English 'online and offline' hybrid teaching recommendation platform based on reinforcement learning," *Security and Communication Networks*, vol. 2021, pp. 1–10, 2021. <https://doi.org/10.1155/2021/4875330>.
- [54] L. Q. Liu, C. Z. Wu, H. Zhang, A. H. N. Hasan, W. H. Chu *et al.*, "Research on taxi drivers' passenger hotspot selecting patterns based on GPS data: A case study in Wuhan," in *Proc. ICTIS*, Banff, AB, Canada, pp. 432–441, 2017. <https://doi.org/10.1109/ICTIS.2017.8047802>.
- [55] Z. H. Huang, G. X. Shan, J. J. Cheng and J. Sun, "TRec: An efficient recommendation system for hunting passengers with deep neural networks," *Neural Computing and Applications*, vol. 31, no. 1, pp. 209–222, 2018. <https://doi.org/10.1007/s00521-018-3728-2>.