An Entropy-Based Model for Recommendation of Taxis' Cruising Route

Yizhi Liu[1, 2], Xuesong Wang[1, 2], Jianxun Liu[1, 2, *], Zhuhua Liao[1, 2], Yijiang Zhao[1, 2] and Jianjun Wang[1, 2]

Abstract: Cruising route recommendation based on trajectory mining can improve taxi-drivers' income and reduce energy consumption. However, existing methods mostly recommend pick-up points for taxis only. Moreover, their performance is not good enough since there lacks a good evaluation model for the pick-up points. Therefore, we propose an entropy-based model for recommendation of taxis' cruising route. Firstly, we select more positional attributes from historical pick-up points in order to obtain accurate spatial-temporal features. Secondly, the information entropy of spatial-temporal features is integrated in the evaluation model. Then it is applied for getting the next pick-up points and further recommending a series of successive points. These points are constructed a cruising route for taxi-drivers. Experimental results show that our method is able to obviously improve the recommendation accuracy of pick-up points, and help taxi-drivers make profitable benefits more than before.

Keywords: Trajectory data mining, location-based services (LBS), optimal route recommendation, pick-up point recommendation, information entropy.

1 Introduction

Nowadays, the advances in mobile computing and location-acquisition techniques have enabled us a group of location-based services (LBS). LBS based on trajectory data mining mostly aims to alleviate urban traffic congestion and reduce environmental pollution. For instance, trajectory data are applied for road infrastructure monitoring [Zheng (2011); Li, Liu, Liu et al. (2011); Lili and John (2009)], traffic status probing [Chen and John (2010); Pablo, Zhang and Li (2012)], auxiliary urban planning [Pan, Qi, Wu et al. (2013)], transportation services improving [Rajesh, Nguyen and Jiang (2011); Yuan, Zheng, Xie et al. (2011); Yuan, Zheng, Zhang et al. (2010); Qu, Zhu, Liu et al. (2014)], etc.

Traveling by taxis is a major means of public transportation in modern cities. But taxis spend 35%-60% working time to search passengers by cruising along the roads [Dong, Zhang, Dong et al. (2014)]. Therefore, many researchers focus on pick-up point

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

[2] Key Laboratory of Knowledge Processing and Networked Manufacturing in Hunan Province, Xiangtan, 411201, China.

[*] Corresponding Author: Jianxun Liu. Email: ljx529@gmail.com.

Received: 14 March 2020; Accepted: 06 June 2020.

recommendation from the path-optimization or the profit-maximization perspective. Obviously, there are some defects: (1) They may result in sending all cruising taxis to the same location to compete for the same group of passengers. (2) Single attribute of trajectory data will lead to subjective one-sidedness.

Addressing at above problems, we propose an entropy-based model for recommendation of taxis' cruising route. It is able to provide an optimal cruising route recommendation service via measuring the value of historical pick-up points based on information entropy. This paper is an in-depth study based on previous work [Liu, Liu, Liao et al. (2016)]. Our contribution lies in three aspects:

(1) Manifold attributes of taxis' trajectories are measured based on information entropy theory. And they are applied to select a sequence of optimal pick-up points for taxis' cruising route recommendation.

(2) The passengers-finding points are obtained during a kind of pruning algorithm after mining the spatiotemporal pattern from historical GPS data.

(3) The model of entropy-based cruising route recommendation (ECRR) is evaluated by experiments. The results show that it outperforms the traditional method of top-k recommendation.

2 Related work

2.1 Pick-up point recommendation for drivers

Sun et al. [Sun, Guan and JinHong (2017)] developed a novel clustering-based scheme that can exploit multi-source information for taxi pick-up points recommendations. In the literature [Zhang, Liu, Liu et al. (2012); Liu, Liu, Wang et al. (2016)], the passenger pick-up point recommendation is obtained by spatiotemporal analysis of the trajectory data. Davis et al. [Davis, Raina and Jagannathan (2018)] employed classical time-series tools to model the spatio-temporal demand. Lai et al. [Lai, Lv, Li et al. (2019)] collected useful information from trajectories, then calculated the traffic forces for cruising taxis, based on which taxis are routed to optimal road segments to pick up desired passengers. Li et al. [Li, Bao, Li et al. (2016)] mined the most influential set of locations by traversing the taxi trajectories to find k locations in a given spatial region. The single pick-up point recommendation may cause multiple cruise taxis to gather at the same location to compete for the same group of passengers. The singularity and smallness of the pick-up point attributes will lead to subjective one-sidedness of the recommendation results.

2.2 The sequence of pick-up points recommendation for drivers

Xu et al. [Xu, Zhou, Liu et al. (2015)] used a preference trajectory scanning algorithm to identify taxis Preference trajectory to build offline graph trajectory model to recommend hunting trajectory for taxis. Huang et al. [Huang, Shan, Cheng et al. (2018)] took an experienced taxi driver as a learning object, proposed an efficient recommendation system based on the wide and deep model. Dong et al. [Dong, Zhang, Dong et al. (2014)] used a system of linear equations to calculate the score of each road and recommended a profitable cruising route for taxi drivers. Yang et al. [Yang, Wang, Rahimi et al. (2015)] recommends profitable routes based on assigned potential profitable grids and updates the profitable route constantly based on taxis' movements. Zhang et al. [Liao, Zhang, Liu et al. (2019)] fused the driver's location, distance, and other geographic information into the hidden semantic model. Most of these above methods focus on taxi passenger recommendation from the perspective of path optimization or profit maximization, and lack of analysis of the current spatiotemporal information of the driver.

As a multi-dimensional measurement tool, information entropy can be used to determine how much valuable the information is and to measure the weight of each evaluation index in the estimation of information value. It is more meaningful to combine multiple indexes to measure the value of pick-up points. Therefore, we propose an entropy-based model for recommendation of taxis' cruising route. By considering multiple evaluation indexes, the pick-up points' selection model is constructed for empty taxis. Information entropy is applied to weight each evaluation index of pick-up points. Multiple pick-up points are successively recommended to create the optimal cruising route.

3 The framework of our approach

The framework is shown in Fig. 1. Firstly, pick-up points are mined from taxis' trajectories. Pick-up points mean the location carrying passengers. Passengers-finding points mean the gathering areas of passengers which are reckoned as those of pick-up points in this paper. The density of passengers-finding points or pick-up points indirectly reflects the passengers' carrying demands in this region or at this position.

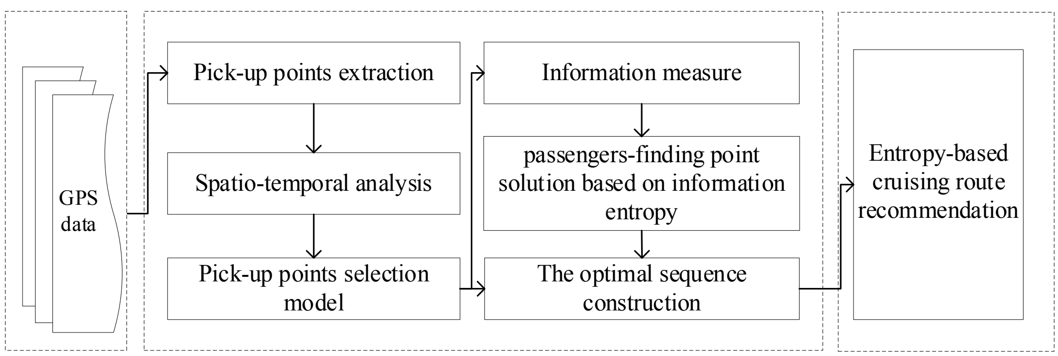

Figure 1: The framework of our approach

Secondly, a kind of spatial-temporal analysis method based on density clustering is proposed to capture the passengers-finding point. By investigating deeply pick-up points' spatial-temporal distribution characteristics, we can recommend passengers-finding routes for taxi-drivers and support the taxicab dispatching system.

Next, we choose the most determinant attributes by our experiences, and construct an entropy-based model of selecting passengers-finding points. Information entropy is applied to measure attributes values for taxis.

Finally, we repeat the entropy-based model K-1 times. Each time an optimal passengers-finding point is obtained, and it is used as the current location of taxis for the next round of calculations. Thus, we can get K optimal passengers-finding points by the depth-first method. These points are concatenated as a cruising route recommended to taxi-drivers.

4 Details of our approach

4.1 Attributes definition and computation

Historical GPS data of taxis are firstly preprocessed to distil the pick-up points and the passengers-finding points. The spatial-temporal features of these points are described by manifold attributes, such as the carrying probabilities, the waiting time of passengers-finding points, the travel time and the travel distance after carrying. According to the taxi's current location, the distance and the arrival time to the passengers-finding point can be computed. Through the entropy-based model of selecting passengers-finding points, several sequential points are successively obtained to create a cruising route for recommendation. The values of the four attributes are computed as follows:

(1) Carrying probabilities of passengers-finding points.

It is denoted as the number of pick-up points in unit time and in unit area. If it is bigger, the chance of finding passengers at this location is bigger. In the Eq. (1), $Area(i)$ means the given area range centered on the passengers-finding point, T_S is the given time range. $Number(i)$ means the number of pick-up points in the spatial and temporal range.

$$P(i) = \frac{Number(i)}{Area(i) \times T_S} \tag{1}$$

(2) Waiting time of passengers-finding points

It is a time-range for a vacant taxi from waiting to picking up passengers at this location. In the Eq. (2), T_S is the given time range, $Number(i)$ is the number of pick-up points in the spatial and temporal range. And $w(k)$ is the waiting time of the *k*-th pick-up points in the area $Area(i)$.

$$W(i) = \frac{\sum_{k=1}^{Number(i)} w(k)}{Number(i) \times T_S} \tag{2}$$

(3) Trip time after carrying

In the Eq. (3), T_S is the given time range, $Number(i)$ is the number of pick-up points in the spatial and temporal range. $ft(k)$ is the trip time after carrying of the *k-th* pick-up points in the area $Area(i)$.

$$Ft(i) = \frac{\sum_{k=1}^{Number(i)} ft(k)}{Number(i) \times T_S} \tag{3}$$

(4) Trip distance after carrying

In the Eq. (4), T_S is the given time range, $Number(i)$ is the number of pick-up points in the spatial and temporal range. And $fd(k)$ is the trip distance after carrying of the *k-th* pick-up points in the area $Area(i)$.

$$FD(i) = \frac{\sum_{k=1}^{Number(i)} fd(k)}{Number(i) \times T_S} \tag{4}$$

Generally speaking, the drivers' profit is proportional with the trip time and the trip distance, and is inversely proportional with the waiting time.

4.2 Pick-up points selection model

Assuming that the no-load taxi starts from point O, we select the passengers-finding points within a certain distance from O in the current slot from the historical passengers-finding points data. The spatial-temporal features of each passengers-finding point are described by N attribute values. Herein, we choose $N = 5$. The five attributes are respectively carrying probabilities of passengers-finding points, waiting time of passengers-finding points, trip time after carrying, trip distance after carrying, and the distance between the current location and the passengers-finding point. Those passengers-finding points and attribute values are respectively depicted with two sets $X = \{x_1, x_2, \cdots, x_k\}$ and $U = \{p_1, p_2, \cdots, p_N\}$.

The decision matrix A is given as follows. In the matrix, $\alpha_{ij}\ (1 \le i \le k, 1 \le j \le n)$ means the *j-th* attribute value of the *i-th* passengers-finding point. Different attribute values have different effects on the decisions. To avoid the effects, we employ gravimetric transformation to normalize the decision matrix A , in which r_{ij} are normalized parameters of column vectors according to the Eqs. (6) and (7). For three benefit type attributes, carrying probabilities of passengers-finding points, trip time and trip distance after carrying, we apply the Eq. (6) to compute r_{ij} . For two cost type attributes, carrying probabilities of passengers-finding points and the distance between the current location and the passengers-finding point, we apply the Eq. (7) to compute r_{ij} . Thus, the normalized matrix R is created.

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{kN} \end{bmatrix} \tag{5}$$

$$r_{ij} = \frac{\alpha_{ij}}{\sum_{1 \le i \le k} \alpha_{ij}} \tag{6}$$

$$r_{ij} = \frac{1/\alpha_{ij}}{\sum_{1 \le i \le k} 1/\alpha_{ij}} \tag{7}$$

$$R = (r_{ij})_{k \times N} \tag{8}$$

4.3 Passengers-finding point solution model based on information entropy

After constructing the selection model of pick-up points, it is necessary to measure the points' values in the decision matrix according to the current location and the current time of taxis. The values are computed in light of the theory of information entropy.

Step 1 The entropy E_j is the information entropy of attribute values p_j in the matrix U.

$$E_j = -\frac{1}{\ln(k)}\sum_{i=1}^{k} r_{ij} \ln(r_{ij}) \tag{9}$$

Step 2 The weight vector $W(w_1, w_2, ..., w_N)$ is computed for the N attribute values.

$$\begin{cases} w_j = \dfrac{(1-E_j)}{\sum_{j=1}^{n}(1-E_j)} \\ w_1 + w_2 + \cdots + w_N = 1 \end{cases} \tag{10}$$

where w_j is the weight value of the *j-th* attribute p_j in U, E_j is the information entropy of p_j in U.

Step 3 The integrated attribute vector $Z(W)$ of $X = \{x_1, x_2, \cdots, x_K\}$ is computed via the following formula.

$$Z(W) = \sum_{i=1}^{k}\sum_{j=1}^{n} r_{ij} w_j \tag{11}$$

Step 4 To obtain the top-m optimal passengers-finding points, we sort these points by the integrated attribute vector $Z(W)$ and select them from high to low.

$$Z = MAX_m[Sort(Z(W))] \quad (1 \le \text{m} \le \text{k}) \tag{12}$$

The top-m optimal pick-up points can be computed within certain time-slot and in the distance scope via this model. Owing to considering manifold attributes, it effectively avoids the one-sidedness.

4.4 Entropy-based cruising route recommendation

Since the actual situation is dynamic, only a single pick-up point recommended for taxis may lead to the failure of carrying passengers at that position. If single factor such as the carrying probabilities is applied for top-k recommendation, taxi-drivers may be puzzled to find the next point when the passenger at the recommended point has been lifted. Therefore, we attempt to recommend not one point but an optimal sequence of pick-up points with the highest probabilities. Passengers release pick-up requirements with the current time and the location at first. The optimal sequence will be constructed via the selection model of pick-up points and the solution model based on information entropy.

According to the current time and the local position, the two models, the selection model of pick-up points and the solution model based on information entropy, are repeated (k-1) times. Each time the previous passengers-finding point is used as the initial point. Suppose the location O of a taxi is the initial point. Then the top-m optimal pick-up points $P_{i1}, P_{i2}, \cdots, P_{im}$ can be calculated from the pick-up data according to the time-slot and distance scope. In turn, one of the previous points is reckoned as the initial point, and then the $m \times m$ optimal pick-up points are computed successively. Thus, a k-layer tree is constructed with optimal pick-up points. Herein, we adopt k=3. Next, the pruning algorithm and the depth-first searching algorithm are explored to select points successively. Consequently, the optimal cruising route is generated and then recommended to taxis.

5 Experiment and analysis

5.1 GPS data preprocessing

The data set used in this paper comes from the open trajectory data set provided by the GeoLife project of Microsoft Research Asia, which contains the trajectory data set of 182 users in Beijing from April 2007 to August 2012. The information such as the latitude, longitude and time of the user's location is mainly collected, the status information is not recorded, that is, the no-load or the state of carrying passengers. Since the taxi trajectory data is collected by the GPS device, when a special situation occurs, such as signal occlusion, cold start, etc., the signal is weak or no signal may interfere with the data acquisition of the GPS device, we correspondingly performed some data cleaning operations, including outlier detection, trajectory smoothing, and stay area detection. In order to obtain the passenger record point therefrom, the trajectory data needs to be pre-processed. The trajectory data preprocessing steps are as follows：

5.1.1 Extracting pick-up points

The occurrence of passenger-carrying events usually means that taxis are stationed somewhere or cruise at low speed, so pick-up points can be extracted based on the detection of the residence area. The location where a taxi picks up passengers is usually the last point in the stay area or the end of a low-speed cruise and the starting point of the trajectory. The average speed before carrying passengers is less than the given speed threshold V_θ, and the average speed after carrying passengers exceeds the given speed threshold V_θ. And the driving distance exceeds the given distance threshold δ. The track segment $T\{p_1, p_2, ..., p_i, ..., p_n\}$, p_i is the stop point, and the criterion for defining p_i as the passenger point is:

$$\begin{cases} d(p_o, p_i) / (p_i.t - p_o.t) < V_\theta \\ d(p_i, p_n) / p_n.t - p_i.t) > V_\theta \\ d(p_i, p_n) > \delta \end{cases} \tag{13}$$

5.1.2 Passengers-finding points generation and evaluation

According to the time attribute and spatial attribute of pick-up points, this paper draws on the cluster analysis method mentioned in Yuan et al. [Yuan, Zheng, Zhang et al. (2013); Zhang (2013)], improves the Spatial Temporal Analysis (STA). This paper uses the density-based OPTICS [Ankerst, Breunig, Kriegel et al. (1999)] algorithm. This paper divides the time in hours and divides the day into 24 time slices.

Pick-up points are divided according to the experimental area. Spatial clustering analysis of pick-up points in small areas, then the whole area is analyzed. The results obtained by the spatial attribute analysis are divided according to the time attribute of pick-up points in the cluster, and the spatial analysis result is retained to the time period when pick-up points ratio of a certain period exceeds a given threshold. A candidate point set $ST[N]$ can be obtained.

Pick-up points are divided into various time segments, and then the spatial clustering analysis is carried out according to the spatial attributes of pick-up points in each time period. A candidate point set $TS[M]$ can be obtained.

The data sets $ST[N]$ and $TS[M]$ should be eliminated the duplicate data. Let candidate point i and j come from candidate sets $ST[N]$ and $TS[M]$, respectively. When the distance $dist(i,j)$ between candidate points i and j is less than the given distance threshold, it is considered that candidate points i and j represent the same region and are repeated candidates. Calculation of the score of repeated candidate points by Eq. (14).

$$Score(i) = \sum_{1 \le k \le K} (1 - \frac{dist(i,k)}{dist(i,j)}), dist(i,k) \in [0,d] \tag{14}$$

In the Formula, $Score(i)$ is the score of the repeated candidate point $dist(i,k)$, $dist(i,k)$ is the distance between pick-up point k and the candidate point i in the region represented by the candidate point, K is the total number of pick-up points in the region represented by the repeated candidate point, and d represents the distance threshold.

By obtaining the score of repeated candidate points i and j, the new candidate point position is calculated by Eq. (15). Where O is the position coordinate of the new candidate point, and $L(i)$ is the position coordinate of the candidate point i.

$$O = L(i) \sim \frac{Score(i) \times dist(i,j)}{Score(i) + Score(j)} \tag{15}$$

In order to verify that pick-up points obtained by spatial-temporal analysis method can indeed represent the gathering place of passengers. This paper draws on the method of judging the accuracy of docking sites in Yuan et al. [Yuan, Zheng, Zhang et al. (2013)], and uses this method to test the search point passengers-finding points of this paper. In this paper, the radius of the accuracy of the passengers-finding points is constantly adjusted from 5 meters to 50 meters, and the results are shown in Fig. 2, it can be seen that the accuracy of passengers-finding points obtained by the spatial-temporal analysis method reaches 90% at a radius of 25 meters. The spatial-temporal analysis (STA) method of obtaining pick-up points in this paper is compared with the method of MSRA for obtaining the Carrying-passenger docking point by the hierarchical clustering method in Yuan et al. [Yuan, Zheng, Zhang et al. (2013)], as shown in Tab. 1. It is clear that our method is superior to the hierarchical clustering analysis method mentioned.

5.2 Recommendation result evaluation

In order to evaluate the effectiveness of the entropy-based cruising route recommendation (ECRR), the experimental results are compared with the typical Top-K recommendations, and two different values, the ratio of travel time after carrying passengers and driving distance after carrying passengers, are used as the measure of recommended performance.

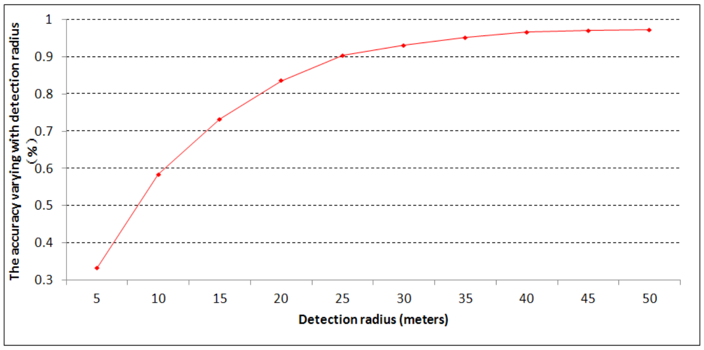

Figure 2: The accuracy of passengers-finding points recommendation varies with detection radius

Table 1: Comparison of spatial-temporal analysis with hierarchical clustering in literature [Yuan, Zheng, Zhang et al. (2013)]

Method	Radius of Buffer	Precision
MSRA	50	90.9%
STA	50	**98.7%**

The typical Top-K recommendation recommends the number of K passengers-finding points with the highest carrying passenger probability near the current location of the taxi. The ratio of driving distance after carrying passengers is the ratio of the driving distance after being carried to the passenger to the total running distance in the current time period. The ratio of travel time after carrying passengers is the ratio of the driving time after being carried to the passenger to the total running time in each period. Fig. 3 and Fig. 4 respectively compare the ratios of the trip time and the trip distance after carrying passengers between the Top-K approach and our ECRR method.

Fig. 3 shows that in the case of more demand for passengers during the day, the carrying-passenger travel time ratio of this method is higher than the Top-K method. In the early morning when the demand for passengers is small, the recommended carrying-passenger travel time ratio of this paper is inferior to Top-K. Because for periods when demand for taxis is small. Because of to the taxi demand small time interval, Top-K gives a high probability of carrying passengers at the location, and relatively more opportunities for carrying passengers. However, this method needs to consider a number of influencing factors, and the road cost largely restricts the choice of these methods for these locations. Therefore, the recommended passengers-finding point is only the nearest optimal, rather than the overall optimal.

Fig. 4 shows that the distance of carrying passengers in this method is higher than that of Top-K method in any time period. The reason is that the method takes into account the cost of the journey, while the Top-K method only considers where it is easier to pick up passengers, even the distance between passengers-finding points at which it is recommended to stay is far away. Since the distance after carrying passengers usually indicates the actual payment of passengers, from the perspective of actual revenue, the performance of this method is better than Top-K recommendation, which can bring good benefits to drivers.

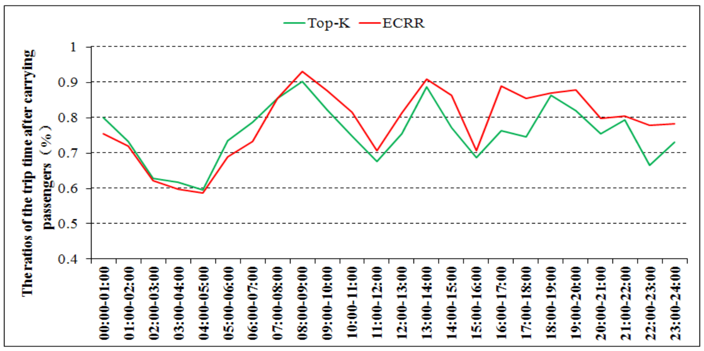

Figure 3: A comparison on the ratios of the trip time after carrying passengers

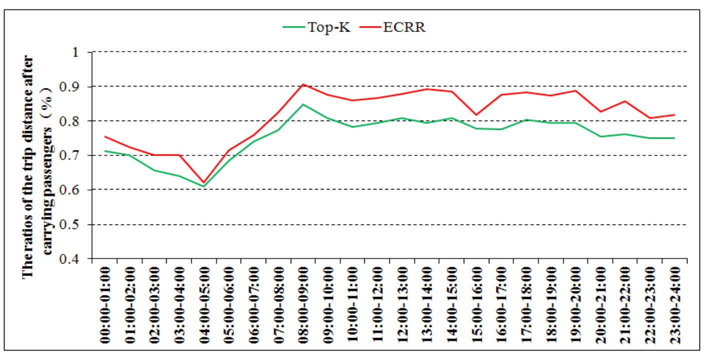

Figure 4: A comparison on the ratios of the trip distance after carrying passengers

6 Conclusions

According to different measurement standards of finding passengers, different drivers with no-load taxis choose different destinations so that it leads to different incomes. So it is challenging issue to find out a more profitable passenger-finding route by taking into account multiple factors. Therefore, we propose an entropy-based model for recommendation of taxis' cruising route. It applies information entropy to trade off manifold factors. And our approach is able to recommend an optimal cruising route for no-load taxi-drivers. By comparing the traditional Top-K recommendation algorithm, two different values of driving time and distance after carrying passengers are taken as the evaluation index of recommendation performance. Experimental results show that our method can deal with the problem of taxi no-load after comprehensive consideration of various factors.

Funding Statement: This project was funded by the National Natural Science Foundation of China (61872139, 41871320), Provincial and Municipal Joint Fund of Hunan Provincial Natural Science Foundation of China (2018JJ4052), Hunan Provincial Natural Science Foundation of China (2017JJ2081), the Key Project of Hunan Provincial Education Department (17A070, 19A172), the Project of Hunan Provincial Education Department (17C0646).

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

References

Ankerst, M.; Breunig, M. M.; Kriegel, H. P.; Jörg, S. (1999): Optics: ordering points to identify the clustering structure. *Stanford Research Inst Memo Stanford University.*

Chen, Y.; John, K. (2010): Probabilistic modeling of traffic lanes from GPS traces. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 81-88.

Davis, N.; Raina, G.; Jagannathan, K. (2018): Taxi demand forecasting: a hedge-based tessellation strategy for improved accuracy. *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-12.

Dong, H.; Zhang, X.; Dong, Y.; Chen, C.; Rao, F. (2014): Recommend a profitable cruising route for taxi drivers. *IEEE International Conference on Intelligent Transportation Systems*, Qingdao, China, pp. 2003-2008

Huang, Z.; Shan, G.; Cheng, J.; Sun, J. (2018): TRec: an efficient recommendation system for hunting passengers with deep neural networks. *Neural Computing and Applications*, vol. 31, pp. 209-222.

Lai, Y.; Lv, Z.; Li, K. C.; Liao, M. (2019): Urban traffic coulomb's law: a new approach for taxi route recommendation. *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 3024-3037.

Li, H.; Liu, J.; Liu, Y.; Jin, L. (2011): Evaluating roving patrol effectiveness by GPS trajectory. *Autonomic and Secure Computing (DASC)*, pp.832-837.

Li, Y.; Bao, J.; Li, Y.; Wu, Y.; Zheng, Y. (2016): Mining the most influential k-location set from massive trajectories. *The 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, no. 51 pp. 1-4.

Lili, C.; John, K. (2009): From GPS traces to a routable road map. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp.3-12.

Liu, Y.; Liu, J.; Wang, J.; Liao, Z. ; Tang, M. (2016): Recommending a personalized sequence of pick-up points. *Journal of Computational Scienc*e, pp. 278-291.

Liu, Y.; Liu, J.; Wang, J.; Liao, Z.; Tang, M. (2016): Recommending a personalized sequence of pick-up points. *Advances in Services Computing*, APSCC 2016, Springer, pp. 278-291.

Liao, Z.; Zhang, J.; Liu, Y.; Xiao, H.; Zhao, Y. et al. (2019): Fusing geographic information into latent factor model for pick-up region recommendation. *2019 IEEE International Conference on Multimedia & Expo Workshops*, pp. 330-335.

Pablo, S. C.; Zhang, D.; Li, S. (2012): Urban traffic modelling and prediction using large scale taxi GPS traces. *Pervasive Computing. Springer Berlin Heidelberg*, pp. 57-72.

Pan, G.; Qi, G.; Wu, Z.; Zhang, D.; Li, S. (2013): Land-use classification using taxi GPS traces. Intelligent Transportation Systems. *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 113-123.

Qu, M.; Zhu, H.; Liu, J.; Liu, G.; Xiong, H. (2014): A cost-effective recommender system for taxi drivers. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 45-54.

Rajesh, K. B.; Nguyen, X. K.; Jiang, L. (2011): Real-time trip information service for a large taxi fleet. *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*. pp. 99-112.

Sun, J.; Guan, C.: Jinhong, M. (2017): Exploiting optimization mechanism for pick-up points recommendations. *2017 International Conference on Computer Systems, Electronics and Control*, pp. 1180-1184.

Xu, X.; Zhou, J.; Liu, Y.; Xu, Z. Z.; Zha, X. W. (2015): Taxi-RS: taxi-hunting recommendation system based on taxi GPS data. *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1716-1727.

Yang, W.; Wang, X.; Rahimi, S. M.; Luo, J. (2015): Recommending profitable taxi travel routes based on big taxi trajectories data. *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, Cham, pp. 370-382.

Yuan, J.; Zheng, Y.; Xie, X.; Sun, A. G. (2011): Driving with knowledge from the physical world. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 316-324.

Yuan, J.; Zheng, Y.; Zhang, C.; Xie, W.; Xie X. et al. (2010): T-drive: driving directions based on taxi trajectories. *Proceedings of the 18th SIGSPATIAL International conference on advances in geographic information systems*, pp. 99-108.

Yuan, J.; Zheng, Y.; Zhang, L.; Xie, X. (2013): 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.

Zhang, M. (2013): Recommendation for pick-up points and hot zone based on taxi trajectories. *Hunan University of Science and Technology.*

Zhang, M.; Liu, J.; Liu, Y.; Hu, Z.; Yi, L. (2012): Recommending pick-up points for taxi-drivers based on spatio-temporal clustering. *International Conference on Cloud & Green Computing*, Xiangtan, China, pp. 67-72

Zheng, Y. (2011): Urban computing with taxicabs. *Proceedings of the 13th International Conference on Ubiquitous Computing*, pp. 89-98.