JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 35, 959-975 (2019)
DOI: 10.6688/JISE.201909_35(5).0002

Toward Pattern and Preference-Aware Travel Route Recommendation over Location-Based Social Networks

LIANG ZHU[1,2], LIPING YU[1], ZENGYU CAI[1] AND JIANWEI ZHANG[1]
[1]*School of Computer and Communication Engineering*
Zhengzhou University of Light Industry
Zhengzhou, 450001 P.R. China
[2]*State Key Laboratory of Networking and Switching Technology*
Beijing University of Posts and Telecommunications
Beijing, 100876 P.R. China
E-mail: {lzhu; ing}@zzuli.edu.cn; yulipingl520@163.com; 8949934@qq.com

Travel route recommendation for Location-based Social Networks (LBSNs) has been received much attention to research people's activity patterns and personalized preferences. Existing travel route recommendation schemes in literature are confronted with three problems: (1) the location is limited in practical environment, and the data sparsity is always happened when they recommend travel route services based on the location information; (2) they fail to consider the order of mobile trajectory, which is valuable to reflect the interest and preference of users for travel route recommendation; (3) they can't be adapted to different kinds of POI category, which causes the extendibility is low. In this paper, we propose *PP-TRR*, a pattern and preference-aware travel route recommendation scheme to tackle the above problems. First, we construct the system architecture of our proposed travel route recommendation. Then, we model the movement pattern of each user. Finally, we present the travel route recommendation scheme to recommend personalized services for targeted users. The experimental results show that our method outperforms the existing method.

Keywords: user model, location prediction, friends discovery, route recommendation, LBSNs

1. INTRODUCTION

Location-Based Social Networks (LBSNs), such as Foursquare, Facebook Place, Twitter and Jie Pang, are that online social network and physical location are associated by making use of "check-in" to achieve the sharing and propagation of location-based services in virtual world. Recent years, lots of sensor-embedded smart mobile devices have been appeared to promote the development of LBSNs. Among these devices, smart phones are being favored by many users. Owners of smart phones can access Internet and use location-based applications ubiquitously in order to experience various kinds of services, such as multimedia entertainment [1-3], real-time news [4] and traffic information [5], *etc.* The forecast of Nokia shows that the amount of mobile data will increase 1000 folds with the drastic increase of mobile subscribers' footprint in near future. According to a recent report by International Telecommunications Union (ITU), there will be approximately 25 billion connected devices by 2020. Because of the rapid increasing of data volume, the first ITU standard on big data has been approved by 2015. Large numbers of services over internet have been pushed to mobile subscribers with many

Received August 20, 2018; accepted October 24, 2018.
Communicated by Gabriel-Miro Muntean

kinds of choices. However, there are many useless services in the current state of subscribers. Even some interested information has been submerged by the massive network services, and subscribers may give up using the corresponding services. Thus, the challenge for future service recommendation systems is not to recommend more services "*to anyone*, *at anytime*, *from anywhere*", but to recommend "*the suitable service*, *at the suitable time*, *in the suitable place*, *to the suitable person*".

Travel route recommendation is a new type of service recommendation task that comes along with LBSNs. Different with traditional recommendation methods, such as collaborative filtering (CF) [6] and content-based recommendation (CBR) [7], travel route recommendation concerns more on recommending personalized and quality of experience (QoE)-guaranteed services [8]. By making use of the context information (*e.g.* time, location and social relationship, *etc.*), travel route recommendation system can provide automatic and customized selection to subscribers. Research topics covered in this area mainly include location prediction [9], user behavior analysis [10-12], movement pattern modeling [13], and social relationship detection [14, 15], *etc.* Among all of these topics, user behavior analysis and movement pattern modeling are received much attention due to the high value in both research and academy.

Although the growing interest in travel route recommendation has resulted in thousands of peer-reviewed publications, there is still significant ongoing work addressing many challenges. There are three problems for current travel route recommendation schemes. First, it is limited for travel route recommendation by only considering the location information in practical environment, and the data sparsity is always happened by this way. Second, they fail to consider the order of mobile trajectory, which is valuable to reflect the interest and preference of users for route recommendation. Third, they can't be adapted to different kinds of Point of Interest (POI) category, which causes the extendibility is low.

Motivated by the above consideration, we propose a pattern and preference-aware travel route recommendation scheme (*PP-TRR*) for LBSNs. And the contributions of our work can be divided into three aspects as following:

(1) We introduce a semantic trajectory translation method to translate mobile trajectory from geographical space to semantic space.
(2) We construct the user model for each user by considering location trajectory, semantic trajectory, location popularity and user familiarity.
(3) We study a potential friend discovery algorithm to find the similar users for targeted user, and extract the candidate routes according to movement pattern and preference for personalized travel route recommendation.

2. RELATED WORKS

In this section, we briefly review some related works on user model, location prediction and route recommendation.

(A) User model

User model can well reflect the periodic behavior of users throughout certain periods of one day. For example, people usually visit "home" and "work" locations on weekdays, and "home" and social network driven locations on weekends. In order to construct user model, E. Cho *et al.* [14] considered three elements, *i.e.* spatial locations, temporal movement between these locations and social relationship. After that, J. C. Ying *et al.* [10] studied the user check-in behavior by considering a user's social-triggered intentions, preference-triggered intentions, and popularity-triggered intentions. W. Y. Zhu *et al.* [16] proposed two user movement models, *i.e.* Gaussian-based and distance-based movement models, to capture the check-in behavior of individual LBSNs user, based on which location-aware propagation probabilities can be derived respectively. Z. W. Yu *et al.* [17] proposed an approach which utilized data collected from LBSNs to model users and locations, and it determined users' preferred destinations using collaborative filtering approaches. However, those studies were all failed to consider the semantical information of locations. Therefore, this paper will utilize geographical information and semantical information to construct user model, in order to well reflect the user interest or preference for locations.

(B) Location prediction

Different facets of user behavior offer different predictive power. The users' movements can be inferred by analyzing historical location data. A. Noulas *et al.* [9] studied the location prediction based on transitions between types of places, mobility flows between venues, and spatio-temporal characteristics of user check-in patterns. After that, J. Ye *et al.* [18] proposed a framework which utilized a mixed hidden Markov model to predict the category of user activity at the next step and then predict the most likely location by considering the estimated category distribution. In practice, location prediction of users not only relied on the forecast of novelty-seeking tendency but also depended on how to determine unvisited candidate locations. In order to judge whether user would seek unvisited locations to visit, D. Lian *et al.* [19] constructed a Collaborative Exploration and Periodically Returning model (CEPR) based on exploration prediction. X. Zheng *et al.* [20] made a survey of location prediction on Twitter, which aimed at offering an overall picture of location prediction, especially concentrated on the prediction of user home locations, tweet locations, and mentioned locations. In this paper, the next location can be predicted according to user preference model, movement pattern and the continuity of trajectory sequence.

(C) Route recommendation

Route recommendation is a practical way to provide reasonable and valuable proposals for users. The personalized route can be generated by analyzing the location data based on GPS and users' social media. In [11], V. W. Zheng *et al.* proposed a collective matrix factorization scheme to mine interesting locations and activities, and use them to recommend to the users where they can visit if they want to perform some specific activities and what they can do if they visit some specific places. Then, B. Liu *et al.* [12] proposed a geographical probabilistic factor analysis framework which strategically took various factors into consideration, *e.g.* user preferences, geographical influences, and user mobility behaviors. In [21], Z. Yu *et al.* mined personalized travel packages by considering user preferences, POI characteristics, and temporal-spatial constraints such as travel time and starting location. In order to meet the need for automatic trip organization, more features of POIs should be extracted. Therefore, Y. T. Wen *et al.* [22] proposed an efficient keyword-aware representative travel route framework which utilized knowledge extraction from users' historical movement records and social interactions. In this paper, the personalized travel route can be recommended according to user preference model, location trajectory sequence and semantical trajectory sequence.

3. PROBLEM STATEMENT AND FRAMEWORK DESIGN

3.1 Problem Statement

In LBSNs, users own their historical location with "check-in" information. By making use of historical location, the activity in next location will be predicted by servers. LBSNs servers can find the similar users and recommend personalized travel route service for targeted user according to the predicted activity. However, because of the data sparsity, it is hard to mine similar users from geographical location. As shown in Fig. 1, there are three trajectories (*A*, *B*, and *C*) corresponding to three users. The left part is the location trajectory, and the corresponding semantic trajectory is described in right part. If the geographical distance between each location is only considered, user *A* is more similar than user *C* with user *B*. However, we can see that user *C* is more similar than user *A* with user *B* if the three trajectories are compared by considering the semantical distance in semantic space. And there is the same semantic trajectory as "*Hospital* $\rightarrow$ *Market* $\rightarrow$ *Park*" between *B* and *C*. Thus, user *C* can recommend the corresponding location as "*Hospital* 3 $\rightarrow$ *Market* 3 $\rightarrow$ *Park* 3" to user *B*.

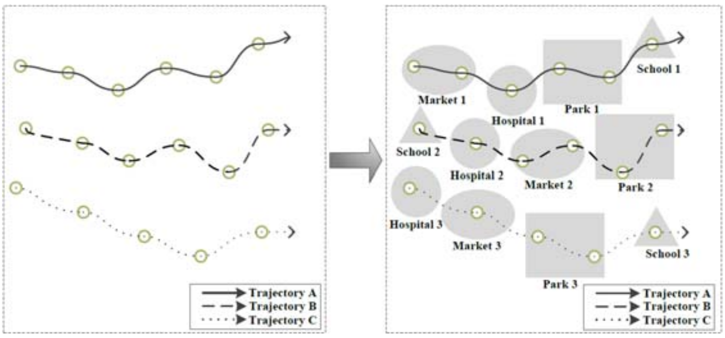

Fig. 1. An example of translation from location trajectory to semantic trajectory.

According to the above example, we research the pattern and preference-aware travel route recommendation in this paper. Movement pattern is a semantic trajectory that reflects the interest and visit order of one user. Preference means that one user is familiar with some special semantic location. By making use of the movement pattern and preference, potential similar users who may be in different geographical location (*e.g.* living in different cities, but they have similar movement pattern) can be mined. And if one user comes to a strange place, LBSNs server can recommend the travel route service to satisfy his/her interests and personalized needs.

3.2 Framework Design

Based on the notion of movement pattern and preference, in this section, we propose a novel framework, namely *PP-TRR*, for personalized travel route recommendation. Different from conventional travel route recommendations based on geographical features of trajectories, we stress on the semantic information of trajectories, in order to find potential friends who may have same semantic trajectory and recommend corresponding location trajectory for targeted user. Fig. 2 shows the framework and flow of data processing within *PP-TRR*. During the offline phase, the original GPS data and POI data will be analyzing by LBSNs server, which mainly includes (1) path extraction and scoring; (2) pattern extraction and scoring. At the online stage, users firstly issue the query request. Then, LBSNs server recommends personalized travel route services to users by (1) potential friend discovery; (2) candidate route mining; and (3) LBSNs server query.

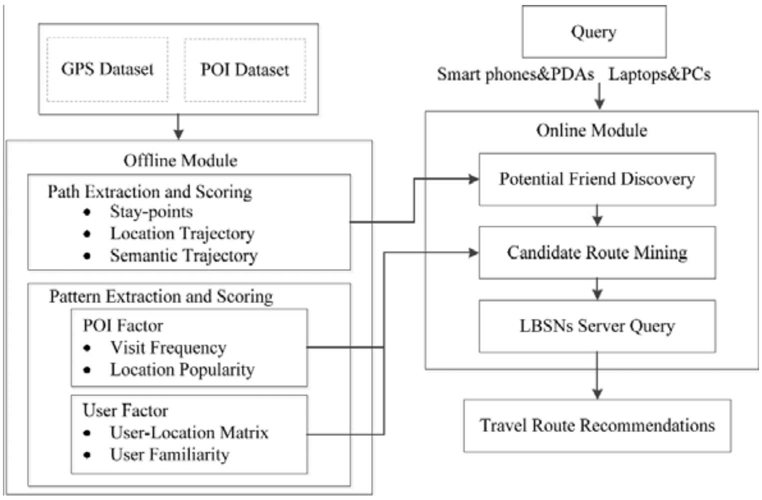

Fig. 2. The framework and flow of data processing within *PP-TRR*.

4. USER MODEL

In this section, we construct user model by making use of the raw Global Positioning System (GPS) trajectory of each user, in order to clearly reflect the location information, semantic information, location popularity, and user familiarity.

4.1 Semantic Trajectory Translation

Because raw GPS trajectory can't well reflect the activity behavior of users, we use stay-point to stand for a geographical region where a user stays for a while [23]. Let S be the set of stay-point, a stay-point s_i can be computed as

$$s_i(lon) = \sum_{j=m}^{n} \frac{p_j(lon)}{|P_i|}, s_i(lat) = \sum_{j=m}^{n} \frac{p_j(lat)}{|P_i|}, \tag{1}$$

where $p_j(lon)$ and $p_j(lat)$ are the longitude and latitude of each raw point p_j, P_i is the set of raw points in stay region s_i. And a trajectory consisted by the stay-points can be constructed as $Tra\_s = s_1 \rightarrow s_2 \rightarrow \ldots \rightarrow s_n$.

In order to mining the interest and preference of users, it is not enough to know the activity trajectory in geographical space only. We define a semantic space to describe the semantic information of each geographical position. The semantic information can help us discover the similar users that live in different geographical region.

We describe the semantic information by making use of Term Frequency-Inverse Document Frequency (TF-IDF) [24], and the weight of each type i for a stay region can be computed as

$$w_i = \frac{n_i}{N} \times \log \frac{|S|}{|S_i|}, \tag{2}$$

where N is the total number of POI appears in the region, n_i is the number of POI for type i, and S_i is the set of stay-points for type i.

Although stay-points have owned semantic information for describing the behavior of users stopping around this point, it can't fully explain the mobile behavior of users. By clustering the stay-points into locations, the coordinate of location L_i can be computed as

$$L_i(lon) = \frac{\sum_{s_j \in L_i} s_j(lon)}{|L_i|}, L_i(lat) = \frac{\sum_{s_j \in L_i} s_j(lat)}{|L_i|}. \tag{3}$$

And the location trajectory can be represented as $Tra\_L = L_1 \rightarrow L_2 \rightarrow \ldots \rightarrow L_n$.

However, it is difficult to reflect the interest and preference of users. So we translate location trajectory into semantic trajectory by using the semantic information of stay-points. The weight value of each type i can be computed as

$$\overline{w}_i = \frac{\sum_{w_i \in f_s} w_i}{|\{w_i \mid w_i > 0\}|}, \tag{4}$$

where f_s means the feature vector of each stay-point, and it can be normalized as

$$W_i = \frac{\overline{w}_i}{\sum_{i=1}^{k} \overline{w}_i}. \tag{5}$$

Based on the above method, each location has a feature vector $F_L = \langle W_1, W_2, \ldots, W_k \rangle$, and the semantic information of each location is described by clustering the feature vectors.

Thus, the semantic trajectory can be generated as $Tra\_C = C_1 \rightarrow C_2 \rightarrow \ldots \rightarrow C_n$.

4.2 Location Popularity and User Familiarity

Although we get the location trajectory and semantic trajectory in geographical space and semantic space, it can't well explain the degree of familiarity of each user for the type of locations. So location popularity and user familiarity are necessary to be considered for constructing user model.

Based on the thought of Hypertext Induced Topic Search (HITS) [25], users and locations are taken as hub nodes and authority nodes, respectively. The value of hub nodes is user familiarity, and the value of authority nodes is location popularity. Thus, user familiarity can be computed by the sum of the value of authority nodes and location popularity can be computed by the sum of the value of hub nodes.

According to the different semantic information, we classify the location in order to get the user-location matrix M^C, and each element M^C_{ij} represents the visit frequency of location j for user i. Also, the user familiarity i with each type C is defined as $u_i^C(f)$, and the location popularity j with these types is defined as $L_j^C(p)$. Because a location only belongs to a type, each location only has a value of popularity. Corresponding to each type, it has

$$u_i^C(f) = \sum_{L \in C} (M_{ij}^C \times L_j^C(p)) \tag{6}$$

and

$$L_j^C(p) = \sum_{u} (u_i^C(f) \times M_{ij}^C). \tag{7}$$

By using the iterative method, we define F_n^C is the vector of the familiarity of users with same type C, and P_n^C is the vector of the popularity of locations with type C. The final iterative process is

$$F_n^C = M^C \cdot (M^C)^T \cdot P_{n-1}^C \tag{8}$$

and

$$P_n^C = (M^C)^T \cdot M^C \cdot F_{n-1}^C, \tag{9}$$

where n represents the iterative number and M^C represents the user-location matrix. Initializing $F_0^C = P_0^C = (1, 1, \ldots, 1)^T$, and it doesn't stop until $|F_n^C - F_{n-1}^C| + |P_n^C - P_{n-1}^C| < \varepsilon$. Thus, we get the user familiarity for different type and the location popularity for the corresponding type.

5. PATTERN AND PREFERENCE-AWARE TRAVEL ROUTE RECOMMENDATION

5.1 Potential Friends Discovery

The interest and preference of users can be mined from the historical locations in real world. For example, one people usually visits the stadium and gym, it is more possible that he/she likes sporting. Generally, shorter distance between two things, more related they are. And the shorter distance between two users in geographical space and semantic space, higher similarity have they in interest and preference. What's more, we do not need to consider the overlap area of historical location when we compute the user similarity in semantic space. Thus, we find potential friend based on the user model, which contains the location trajectory in geographical space and the semantic trajectory in semantic space.

Traditional user similarity computing method (*e.g.* cosine similarity) fails to consider the visit order, so it can't well reflect the interest and preference of users. In this paper, we take the following three factors into account to compute the similarity.

Popularity of Location and Type Similar with IDF, the higher is the popularity of visited location or type, the lower is the similarity among users. And the popularity can be computed as

$$pop(L_k) = \log(\frac{n_L}{N}), pop(C_k) = \log(\frac{n_C}{N}), \tag{10}$$

where N is the total number of locations, n_L is the number of location L, and n_C is the number of location with type C.

Activity Sequence We use path to represent the activity sequence of users. The definition of path is that it is a continuous l-length sub-sequence of location trajectory or semantic trajectory. Taking Fig. 1 as example, there is a path "*hospital* → *market* → *park*" of user B, and user C have the same path with user B. Thus, user C is more similar than user A.

Also, we consider the time interval to compute the similarity. It is the starting time of each path. The shorter is the time interval of two paths, the more similar are two users.

By making use of the popularity and time interval, the similarity of path can be computed as

$$sim(Path\_L) = \sum_{i=1}^{l} pop(Path\_L[i])^{-1} \times 2^{-|t_1 - t_2|} \tag{11}$$

$$sim(Path\_C) = \sum_{i=1}^{l} pop(Path\_C[i])^{-1} \times 2^{-|t_1 - t_2|}, \tag{12}$$

where l is the length of path, $|t_1 - t_2|$ is the time interval of two path.

Geographical Space and Semantic Space We take geographical space and semantic space into account to synthetically compute user similarity.

Based on the three factors, the similarity between u_1 and u_2 can be computed as

$$sim(u_1, u_2) = W_L \cdot \frac{\sum_{j=1}^{m} sim(Path\_L_j)}{|\{Path\_L\}_{u_1}| \times |\{Path\_L\}_{u_2}|} + W_C \cdot \frac{\sum_{j=1}^{m} sim(Path\_C_j)}{|\{Path\_C\}_{u_1}| \times |\{Path\_C\}_{u_2}|}, \tag{13}$$

where W_L and W_C are the weight of location path and semantic path, m is the number of common path.

Thus, we can find the potential friends for targeted user according to $sim(u_1, u_2)$.

5.2 Candidate Route Mining

The movement pattern can well reflect the interest and pattern of users. In order to extract movement pattern from the semantic trajectory, we define the length of continuous sub-sequence is n, and the occurrence number is ρ. Algorithm 1 shows the process of movement pattern extraction.

Algorithm 1: Movement Pattern Extraction

Input: semantic trajectory Tra_C, length of pattern n, occurrence number ρ
Output: set of movement pattern P //The last bit of each pattern is occurrence number
Initialize $P = \emptyset$, $num = 0$, $i = 1$;
Define L is the length of semantic trajectory Tra_C;
While $i \leq (L - n + 1)$
extract the n-length sequence *seq* begin with i;
If *seq* belongs to Tra_C //the sequence is occurred in one day
If *seq* doesn't belong to P
$num = 1$;
put *seq* into P;
Else
add *num* of sequence that same with *seq* in P;
End if
End if
$i = i + 1$;
End while
Delete the element of P if $num < \rho$;
Sort the elements of P according to *num*;

Based on the movement pattern, we can recommend the personalized travel route service with n-length continuous location path for current user. First, we sort the mobility pattern according to ρ and extract the first f pattern. Then, we sort the potential friends according to the similarity and extract the location path. Finally, we translate the location path into semantic path and compare the semantic path with the first f pattern. If there is a semantic path same with pattern, we put the corresponding location path as a candidate service. Algorithm 2 shows the process of candidate route mining.

Algorithm 2: Candidate Route Mining

Input: targeted user u, set of similar users $\{sim(u, u_i)\}$, set of location path $\{Path\_L\}$, set of semantic path $\{Path\_C\}$
Output: set of candidate route Γ, set of users U
Initialize $\Gamma = \emptyset$, $U = \emptyset$, $i = 1$, $j = 1$;
Extract the first f movement pattern of user u;
Sort $\{sim(u, u_i)\}$; //according to the order from big to small
While $|\Gamma| < k$ //the number of candidate route is no more than k
While $j \leq n$ // n is the number of semantic path of user i
If $\{Path\ C\}(j)$ belongs to the first f movement pattern
put $\{Path\ L\}(j)$ into Γ;
put user i into U;
End if
End while
$i = i + 1$;
End while

5.3 Personalized Travel Route Recommendation

In order to recommend the top-*k* route service for targeted user *p*, we propose a HITS-based similarity grading method. It includes three steps as following:

First, the higher is the similarity between user *p* and user *q*, the higher is the score of candidate service.

Second, we refine the user similarity according to the special type. Based on the global similarity, if the familiarity of user *p* and user *q* for each type is similar, the score of candidate service is high.

Third, we consider the location popularity of candidate path. The higher is the popularity of location, the higher is the score of candidate route.

Thus, the score of candidate route can be computed as

$$Score_p(S) = sim(p,q) \times \sum_{i=1}^{l} 2^{-|p^{C_i}(f) - q^{C_i}(f)|} \times \sum_{i=1}^{l} Path\_L[i]^C(p). \tag{14}$$

According to Eq. (14), we can recommend the top-*k* travel route for user *p*.

6. EXPERIMENTS

6.1 Datasets and Experimental Setup

In this section, we use two real-world datasets to verify the validity and efficiency of proposed *PP-TRR*. GeoLife datasets [26] has recorded the GPS trajectory of 182 users with 18670 trajectories in five years (from 2008/10 to 2012/8), which not only includes the daily activity (*e.g.* going home, working, *etc.*), but also includes the recreational activity (*e.g.* shopping, traveling, eating, sports, *etc.*). Most of the data in GeoLife datasets lies in Beijing and few of them in Europe or USA. POI datasets include the location information for all kinds of interest points in Beijing. As shown in Table 1, we classify the raw POI datasets with 20 types for the following step to well describe the semantic trajectory of users. Because the locations can be classified by clustering feature vectors with existing POI, so our method can be adapted to different kinds of POI category.

Table 1. 20 types of raw POI datasets.

Type	Name	Type	Name
1	Food & Beverages Service	11	Motorcycle Service
2	Road Ancillary Facilities	12	Car Service
3	Place Address Information	13	Car Maintenance
4	Scenic Spot	14	Car Sales
5	Public Facilities	15	Commercial Housing
6	Company	16	Life Service
7	Shopping Service	17	Sports Leisure Service
8	Transportation Service	18	Health Care Service
9	Financial Insurance Service	19	Governments Organizations
10	Education Culture Service	20	Accommodation Service

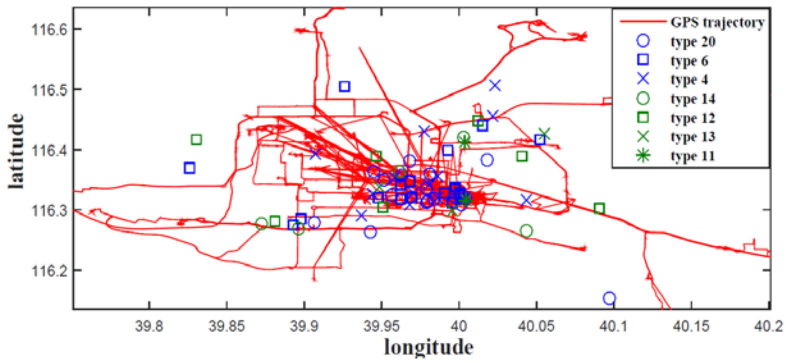

Fig. 3. Different types of each location.

All experiments are conducted on a computer with Intel i7-3770 3.40 GHz CPU and 4 GB RAM, running 64-bit Windows 7 OS. Taking one user for example, Fig. 3 shows the types of each location by translating location trajectory from geographical space to semantic space, and the blue points mean that the locations where he/she usually visits. Also, we set $n = 3$ and $\rho = 3$ to extract the path and movement pattern of each user.

In order to explain the performance of our proposed *PP-TRR* method, we compare it with the cosine similarity method. The cosine similarity can be defined by Eq. (15).

$$sim_{\text{cosine}}(u_{1,2}) = \frac{\sum_i (K_1^i \times K_2^i)}{\sqrt{\sum_i (K_1^i)^2} \times \sqrt{\sum_i (K_2^i)^2}}, \tag{15}$$

where K_1^i and K_2^i is the visiting number of location i of u_1 and u_2, respectively.

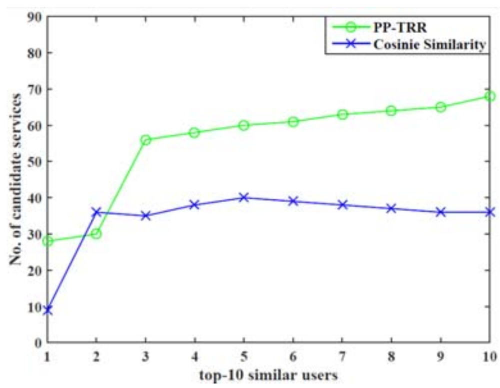

Fig. 4. Comparison of candidate route extraction.

Fig. 4 shows the comparison between cosine similarity method and *PP-TRR* method on the number of candidate route. We can see the performance of *PP-TRR* method is better than cosine similarity method in the aspect of candidate route extraction. The reason it that, cosine similarity method only considers location similarity to find the similar users, which causes the number of candidate services decrease as the number of similar users increases. However, our proposed *PP-TRR* method not only considers location similarity, but also takes semantical similarity into account to find the potential friends who have similar movement pattern with target user. Thus, for *PP-TRR* method, the number of candidate services increases with the number of similar users increases.

6.2 Evaluation Metrics

We define *precision*, *recall*, *k-cover* and *F-measure* to measure the performance of *PP-TRR*.

Precision Given a test user list $U_{u,rec}$, precision can be computed as

$$precision@u = \frac{|U_{u,rec} \cap U_{accepted}|}{u}, \tag{16}$$

where u is the number of test users, $U_{accepted}$ are the users who accept the recommended route.

***k*-cover and Recall** Given a top-k recommendation list $L_{k,rec}$, k-cover and recall can be computed as

$$k\text{-}cover = \frac{|L_{k,rec} \cap L_{used}|}{k}, recall@k = \frac{|L_{k,rec} \cap L_{used}|}{|L_{used}|}, \tag{17}$$

where L_{used} are the recommended route a user used in the test data.

F-measure It is the harmonic value of precision and recall. The higher is the value of F-measure, the better is the performance of route recommendation algorithm. F-measure can be computed as

$$F_{measure} = \frac{2 \times precision \times recall}{precision + recall}. \tag{18}$$

6.3 Results and Discussions

In this section, we evaluate the performance of four route recommendation methods, *i.e. PP-TRR*, *UF*, *LP* and *CS*. Thereinto, *PP-TRR* is the proposed method in this paper. *UF* method means that it only considers the user familiarity to recommend travel route for target user. *LP* method means the location popularity is only considered during the process of route recommendation. *CS* represents the cosine similarity method.

From the results of Fig. 5, we can see that our proposed *PP-TRR* method has the best precision by comparing with the other three methods. There are almost 60 percent in recommended routes are accepted for *PP-TRR* method. However, the precision of emphCS method is only 30%. The reason is that emphCS method fails to consider the visit order and semantic information of locations. By using *CS* method, the users are similar only when they visit the same location, therefore it can not be recommended by the users with low similarity. Moreover, even *UF* method and *LP* method take semantical information of locations into account to find potential friends, the parameters of user familiarity and location popularity are considered separately. It causes the performance of *UF* method and *LP* method is better than *CS* method, and worse than *PP-TRR* method. For our proposed *PP-TRR*, the visit order and semantic information of locations are both considered to find similar users with same movement pattern and preference. Although two users are in different area, the similarity between them can still be computed. According to discovered potential friends, LBSNs server can recommend suitable travel route for target user.

Fig. 6 shows the comparison of recall for four methods. We can see the value of recall increases as the number of recommended routes increases. However, our proposed *PP-TRR* method have the highest recall by comparing with the other three methods. The recall by using *CS* method is lowest in four methods. Thus, it proved that potential friends and suitable route services can be better mined by considered the visit order and semantic information of locations.

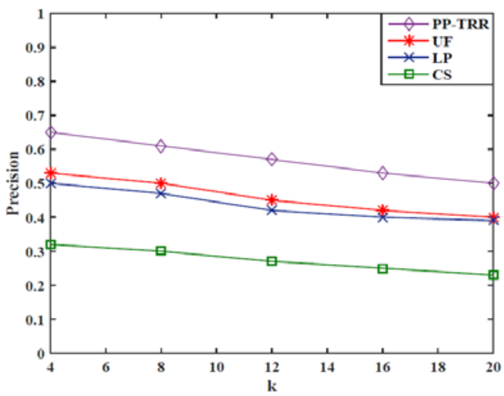

Fig. 5. Comparison of precision.

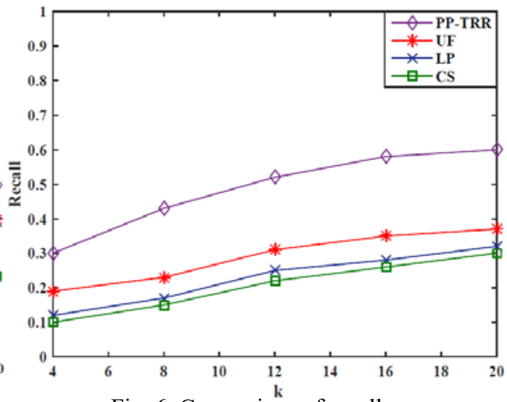

Fig. 6. Comparison of recall.

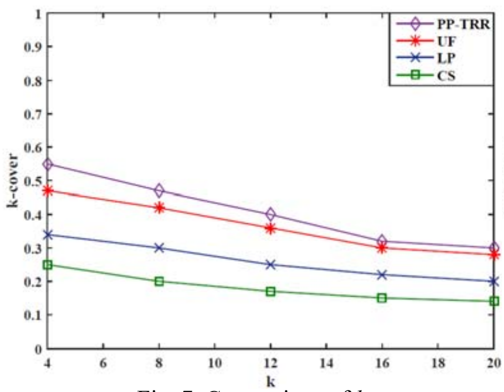

Fig. 7. Comparison of *k*-cover.

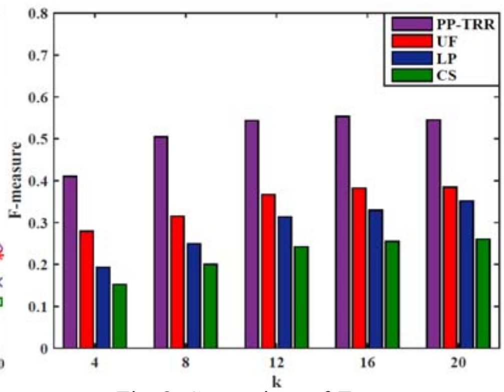

Fig. 8. Comparison of *F*-measure.

In addition, we compare the proposed *PP-TRR* with the other three methods on *k-cover* and *F-measure*. As shown in Fig. 7, the value of *k-cover* decreases as the number of recommended route services increases. However, from the comparison of *k-cover*, we can see that the number of accepted route services of *PP-TRR* is more than *UF* method, *LP* method and *CS* method. Meanwhile, the value of *F-measure* increases as the number of recommended route services increases in Fig. 8. Therefore, the total performance of our proposed *PP-TRR* method is the best compared with the other three methods.

What's more, we consider the time interval to compute the similarity of users in this paper. It ensures at more similar are the users, more similar are interest or preference of them.

7. CONCLUSIONS AND FUTURE WORK

In this paper we focus on the problem of personalized travel route recommendation by considering the interest or preference of user activities. We construct a user model to extract movement pattern of each user. The proposed potential friend discovery algorithm finds similar users according to the path in geographical space and semantic space We conduct extensive experiments over real-world GPS datasets and POI datasets. The experimental results show that our proposed *PP-TRR* method outperforms the existing method.

For the future work, it would be interesting to apply the proposed framework to pattern and preference-aware recommendation in other tasks, *e.g.* content recommendation on multimedia service, POI recommendation, *etc.* What's more，user privacy preserving is also necessary for the research of personalized service recommendation.

ACKNOWLEDGMENT

This work is supported by Open Foundation of State key Laboratory of Networking and Switching Technology (Beijing University of Posts and Telecommunications) (SKLNST-2019-2-04), Key Technologies R&D Program of Henan Province (172102210059 and 172102210060).

REFERENCES

1. C. Xu, S. Jia, L. Zhong, and G. Muntean, "Socially aware mobile peer-to-peer communications for community multimedia streaming service's," *IEEE Communications Magazine*, Vol. 53, 2015, pp. 150-156.
2. C. Xu, S. Jia, M. Wang, L. Zhong, H. Zhang, and G. Muntean, "Performance-aware mobile community-based VoD streaming over vehicular ad hoc networks," *IEEE Transactions on Vehicle Technology*, Vol. 64, 2015, pp. 1201-1217.
3. C. Xu, S. Jia, L. Zhong, H. Zhang, and G. Muntean, "Ant-inspired mini-community based solution for video-on-demand services in wireless mobile networks," *IEEE Transactions on Broadcasting*, Vol. 60, 2014, pp. 322-335.
4. Y. Shin, J. Lim, and J. Park, "Joint optimization of index freshness and coverage in real-time search engines," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24, 2012, pp. 2203-2217.
5. S. Gisdakis, V. Manolopoulos, S. Tao, A. Rusu, and P. Papadimitratos, "Secure and

privacy-preserving smartphone-based traffic information systems," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, 2015, pp. 1428-1438.
6. J. Cao, Z. Wu, Y. Wang, and Y. Zhuang, "Hybrid collaborative filtering algorithm for bidirectional web service recommendation," *Knowledge and Information Systems*, Vol. 36, 2013, pp. 607-627.
7. W. Liu and W. Wong, "Web service clustering using text mining techniques," *International Journal of Agent-Oriented Software Engineering*, Vol. 3, 2009, pp. 6-26.
8. C. Xu, F. Zhao, J. Guan, H. Zhang, and G.-M. Muntean, "Qoe-driven user-centric vod services in urban multi-homed p2p-based vehicular networks," *IEEE Transactions on Vehicle Technology*, Vol. 62, 2013, pp. 2273-2289.
9. A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *Proceedings of IEEE 12th International Conference on Data Mining*, 2012, pp. 1038-1043.
10. J. J.-C. Ying, W.-N. Kuo, V. S. Tseng, and E. H.-C. Lu, "Mining user check-in behavior with a random walk for urban point-of-interest recommendations," *ACM Transactions on Intelligent Systems and Technology*, Vol. 5, 2014, pp. 40:1-40:26.
11. V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *Proceedings of the 19th ACM International Conference on World Wide Web*, 2010, pp. 1029-1038.
12. B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning geographical preferences for point-of-interest recommendation," in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2013, pp. 1043-1051.
13. Y. Zheng, "Trajectory data mining: An overview," *ACM Transactions on Intelligent Systems and Technology*, Vol. 6, 2015, pp. 29:1-29:41.
14. E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011, pp. 1082-1090.
15. C. Biancalana, F. Gasparetti, A. Micarelli, and G. Sansonetti, "An approach to social recommendation for context-aware mobile services," *ACM Transactions on Intelligent Systems and Technology*, Vol. 4, 2013, pp. 10:1-10:31.
16. W.-Y. Zhu, W.-C. Peng, L.-J. Chen, K. Zheng, and X. Zhou, "Modeling user mobility for location promotion in location-based social networks," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1573-1582.
17. Z. Yu, H. Xu, Z. Yang, and B. Guo, "Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints," *IEEE Transactions on Human-Machine Systems*, Vol. 46, 2016, pp. 151-158.
18. J. Ye, Z. Zhu, and H. Cheng, "What's your next move: User activity prediction in location-based social networks," in *Proceedings of SIAM International Conference on Data Mining*, 2013, pp. 171-179.
19. D. Lian, X. Xie, V. W. Zheng, N. J. Yuan, F. Zhang, and E. Chen, "Cepr: A collaborative exploration and periodically returning model for location prediction," *ACM Transactions on Intelligent Systems and Technology*, Vol. 6, 2015, pp. 8:1-8:27.
20. X. Zheng, J. Han, and A. Sun, "A survey of location prediction on twitter," *IEEE Transactions on Knowledge and Data Engineering*, 2018, p. 1.

21. Z. Yu, Y. Feng, H. Xu, and X. Zhou, "Recommending travel packages based on mobile crowdsourced data," *IEEE Communications Magazine*, Vol. 52, 2014, pp. 56-62.
22. Y. Wen, J. Yeo, W. Peng, and S. Hwang, "Efficient keyword-aware representative travel route recommendation," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, 2017, pp. 1639-1652.
23. Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proceedings of the 18th ACM International Conference on World Wide Web*, 2009, pp. 791-800.
24. Q. He, K. Chang, E. Lim, and A. Banerjee, "Keep it simple with time: A reexamination of probabilistic topic detection models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, 2010, pp. 1795-1808.
25. J. M. Kleinberg, "Authoritative sources in a hyperlinked environments," *Journal of ACM*, Vol. 46, 1999, pp. 604-632.
26. Y. Zheng, X. Xie, and W. Ma, "Geolife: A collaborative social networking service among user, location and trajectory," *IEEE Data Engineering Bulletin*, Vol. 33, 2010, pp. 32-40.

Liang Zhu (朱亮) received Ph.D. degree in Computer Science and Technology from Beijing University of Posts and Telecommunications, Beijing, China, in October 2017. He is currently a Lecturer with the Institute of Computer and Communication Engineering at Zhengzhou University of Light Industry, Henan, China. His current research interests include mobile social networks, personalized service recommendation, and privacy preserving.

Liping Yu (余丽萍) received her master degree in Computer Application Technology from Henan Polytechnic University, Henan, China, in 2013. She is a Research Associate at Zhengzhou University of Light Industry. His research interests include distributed computing, user behavior analyzing and modeling.

Zengyu Cai (蔡增玉) received his master degree in Computer Application Technology from Northeast Normal University, Changchun, China, in 2006. He is an Associate Professor at Zhengzhou University of Light Industry. His research interests include trusted computing, plan recognition and information security.

Jianwei Zhang (张建伟) received his Ph.D. degree in Computer Application Technology from PLA Information Engineering University in 2010. He is a Professor at Zhengzhou University of Light Industry. His research interests include broadband information network and network security.