

Research Article

A Location-Based Business Information Recommendation Algorithm

Shudong Liu^{1,2} and Xiangwu Meng^{1,2}

¹Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing 100876, China

²School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Shudong Liu; bupt.mymeng@gmail.com

Received 5 August 2014; Accepted 1 December 2014

Academic Editor: Jong-Hyuk Park

Copyright © 2015 S. Liu and X. Meng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recently, many researches on information (e.g., POI, ADs) recommendation based on location have been done in both research and industry. In this paper, we firstly construct a region-based location graph (RLG), in which region node respectively connects with user node and business information node, and then we propose a location-based recommendation algorithm based on RLG, which can combine with user short-ranged mobility formed by daily activity and long-distance mobility formed by social network ties and sequentially can recommend local business information and long-distance business information to users. Moreover, it can combine user-based collaborative filtering with item-based collaborative filtering, and it can alleviate cold start problem which traditional recommender systems often suffer from. Empirical studies from large-scale real-world data from Yelp demonstrate that our method outperforms other methods on the aspect of recommendation accuracy.

1. Introduction

Rapid technology development has brought an increasing number of mobile devices equipped with GPS capacities, such as laptops, PDAs, and mobile phones. It makes check-in behavior become a new life style of millions of users who share their locations, tips, and experience about points of interest (POI) with their friends in location-based social networks. Recently, how to provide timely and personalized information and sharing services based on users' location information is gradually attracting a lot of attention both from the industry and from the research community. It also forms a known and independent research area named the location-based services (LBS). In particular, personalized information recommendation is more important since it is beneficial for users to know new POIs or special promotions in marketplace and explore their city and for advertisers to launch advertisements to targeted users.

Recently, many researches on information (e.g., POI and ADs) recommendation based on location have been done

in both research and industry [1–3]. Collaborative filtering (CF) is the mainstream of algorithm to solve this task. Both memory-based and model-based collaborative filtering methods have been proposed and investigated to learn users' preferences on the LBS from users' location check-in data [1, 4, 5]. However, previously proposed methods consider all check-ins in a whole and mobile users' basic laws governing human motion and dynamics are usually overlooked as well as rare researches on cold start problem which results from users' rarely rating on items in location-based recommender systems. As shown in [6], humans experience a combination of periodic movement that is geographically limited and seemingly random jumps that are correlated with their social networks. About 50% to 70% of all human movements are short-ranged and periodic both spatially and temporally and are not affected by the social network structure and about 10% to 30% of all human movements are long-distance and random and are usually influenced by social network ties. Hence, location-based recommendation should be sensitive to range of users' movement, and we will show how to

alleviate cold start problem in location-based recommender system using user's basic movement laws in this paper.

Hence in this paper, unlike the previous works, our goal is to provide users with information recommendation within the scope of users' movement in a very sparse rating system. The task is much harder than traditional location-based recommendation or prediction, because it recommends some interesting business information in the scope of user's daily movement. However, this task is more significant since it can provide various personalized favorite local pieces of information combined with long ranged travel information which is close to their friends' home. Thus, if we could divide users' movement region into two parts, the local part and remote part, then we can recommend their favorite business information in each part to them, but most of all, we should determine each user's scope of movement by exploiting their check-in log. In location-based services or social networks, the places users check in every time are often some points of interest or parkland they are visiting and we cannot obtain users' successive location trajectory from their check-in data, and more importantly, this results in a very sparse dataset in location-based recommender system.

Based on these two properties of check-in dataset and studies in [6], we focus on explicitly modeling users' local movements and long-distance travel preferences for recommendation in their check-in dataset. There are two challenges: (1) how to determine users' local movement region and remote movement region and (2) how to find users' favorite business information in each of their movement regions. To address these challenges, we propose a region-based location graph (RLG) and design new algorithms to make accurate top- N recommendation on RLG. The uniqueness of the proposed model is the introduction of users' movement region nodes which could help users find out their local neighbors and remote friends for collaborative information recommendation, including users' local movement region and remote movement region; furthermore, it captures users' local visit interest through user-local movement region connections and captures long-distance travel interest through user-remote movement region connections. As the two users' local movement regions are intersecting, we call the two users local neighbors and as the local movement region of a user and the remote movement region of the other user are intersecting, we call the two users remote friends.

To summarize, our main contributions are as follows.

- (1) We construct a region-based location graph (RLG), in which region node connects with user node and business node, respectively.
- (2) While the two regions of users' local movement are intersecting, we call them local neighbors and while the local movement region of a user and the remote movement region of the other user are intersecting, we call the two users remote friends. Based on RLG framework, we propose a novel location-based business information recommendation algorithm.
- (3) We compare our approach with other methods on a real dataset and show the performance of our

approach on alleviating cold start users problem in location-based recommender systems.

2. Related Work

In recent years, the technologies of mobile communication and mobile location have achieved great development, especially the usage of social network sites, which brings a new chance for social application of geospatial information. The willingness of users to share their current locations and experience originally facilitates the creation of location-based recommender systems (LBRS) based on users generated content and makes it receive much attention from the academic community and industry.

Currently, there are two lines of work to solve the task of location-based recommendation [5]. One line of research is conducted based on the GPS trajectory logs [7–11]. The GPS trajectory data usually consist of small number of users but dense records [12, 13]. Many collaborative filtering algorithms have been proposed and deemed location or POIs as item in traditional recommender systems, such as collective matrix factorization [8], tensor factorization [9], memory-based collaborative location model [10], and pattern recognition model [11]. The other line of work focuses on location-based social networks data, which is very sparse and large-scale [1, 4, 14]. Currently, geographical influences, for example, modeling the check-in probability to the distance of the whole check-in history by power-law distribution [1], modeling users' multicenter check-in behaviors via multicenter Gaussians [4], and mining user check-in behaviors [15], have been addressed and fused with traditional CF algorithms.

The crucial point of location-based recommender systems enables users to read or ask for recommendations in the vicinity of a specified location users are visiting or used to visit, so it is important that recommendations in location-based recommender systems must have strong binding to users' movement region. However, there are rarely previous researches on this issue and cold start problem because of the fact that very sparse data in location-based recommender systems is not yet well studied. As a user can only visit a limited number of locations, especially when a user travels to a new city, the user's locations matrix is very sparse, leading to a big challenge to traditional collaborative filtering-based location recommender systems. To this end, Bao et al. [16] proposed a location recommender system, which consists of two main parts: offline modeling and online recommendation. The offline modeling part models each individual's personal preference with a weighted category hierarchy and infers the expertise of each user in a city with respect to different category of locations according to their location histories using an iterative learning model. The online recommendation part selects candidate local experts in a geospatial range that matches the user's preferences using a preference-aware candidate selection algorithm and then infers a score of the candidate locations based on the opinions of the selected local experts. The significance of the recommender systems in location-based services and the promising solution motivate us to investigate further in this paper.

3. Data Model and Problem Definition

In this section, we briefly introduce the related data model and define users' favorite business information finding problem in their different movement regions.

3.1. Data Model. Unlike GPS trajectory data, users' check-in data are not continuous in both special and temporal dimensions in location-based social network. The places users check in are often some points of interest or parkland they are visiting. For example, when a user has dinner in a restaurant, he may share some information about this restaurant and his experience with his friends at a social network or review websites. So such check-in data indicate that the scope of user's daily motion can cover all businesses reviewed by him; in other words, the business information reviewed by a user is limited in the scope of his daily motion.

Suppose that $U = \{u_i \mid i \in [1, N]\}$ is a mobile user set, where N is the total number of mobile users. Each user has some essential attributes, such as gender, age, and occupation, which is denoted by the form of $\{u_{i,x} \mid x \in [1, X]\}$. For a user's reviewed and favorite businesses $I = \{i_j \mid j = 1, 2, \dots\}$, which can be formed as the following triple $\langle \text{user}, \text{business}, \text{rating} \rangle$, each business i_j has some basic attributes, including location (e.g., longitude and latitude, denoted by $l_j = (l_{j,x}, l_{j,y})$) and service categories (e.g., restaurant, hotel, bar, and shopping mall, denoted by $s_j = \{s_{j,k} \mid k \in [1, K]\}$, $s_{j,k}$ is 1 or 0).

Our data are in the form of $\langle \text{User}, \text{Business}, \text{Location} \rangle$ triple which can be modeled by tripartite graph [17] or a tensor [18]. Although both tripartite graph and tensor treat location as a universal dimension shared by all users, as matter of fact, there is one-to-one correspondence between the business and location in users' check-in data and some users could never review lots of businesses which are out of their movement region in their daily lives. As argued in [6] most of the users' motions are composed of short-ranged daily travel between their homes and workplaces which is periodic both spatially and temporally and long-distance travel which is more influenced by social network ties. In a recommender system, the fixed correlation between business and location is typically not significant, while the movement region plays an important role in recommendation generating process, and the correlation between user and his movement regions is more relevant than that between user and location of business reviewed by them.

Therefore, according to the geographical position distribution of all businesses reviewed by each user, we divide a user's movement region into local movement region and remote movement region. Provided that (x_k, y_k) represents the geographical center of a user's movement region, D_{\max} is the farthest distance between the center $o : (l_{x,o}, l_{y,o})$ and the position the user can reach in his daily life. If there exists a number D_{int} ($D_{\max} \geq D_{\text{int}} > 0$), the percentage of businesses reviewed by a user in a circle region around the pointer of (x_k, y_k) with D_{int} as its radius can reach a fixed number ρ , and we call this circle region the user's local movement region

R_{loc} and the other region his remote movement region R_{rem} . Namely,

$$\rho = \frac{|I'|}{|I|}, \quad (1)$$

where $I' \subset I$, $I' = \{i_j \mid d(l_j, o) \leq D_{\text{int}}, j = 1, 2, \dots\}$, and $d(l_j, o)$ is the Euclidean distance between two points, and according to the conclusion in [6], we set ρ to 0.7. In addition, we call two users as local neighbors if their local movement regions are overlapping, denoted by N_{loc} , whereas we call them remote friends, if the remote movement region of a user and local movement region of the other user are overlapping, denoted by F_{rem} .

3.2. Problem Definition. Intuitively, location-based recommendation is trying to find potential favorite business information within users' entire motion range. With data model defined above, we formally define this problem as follows: there are some basic datasets of all users, including review log set $I = \{i_j \mid j = 1, 2, \dots\}$, local neighbor set N_{loc} and remote friend set F_{rem} . For a specific query user, one recommendation method should return a ranked list of businesses which the user would like, and what is more, the ranking score in the process should consider both user's different movement region and social relationship.

4. RLG Construction

In this section, we treat the two movement regions of users as new nodes, which enable new linkages between users and the location of their reviewed business and construct a graph and name it region-based location graph (RLG), which contains three types of nodes: user node, movement region node, and business node. In this way, we can transform $\langle \text{user}, \text{business}, \text{location} \rangle$ into $\langle \text{user}, \text{region}, \text{business} \rangle$ by formation of users' two movement regions.

RLG is a bipartite graph $G(U, R, I, E, \omega)$, where U denotes the set of all user nodes, R is the set of users' movement region nodes, and I is the set of business nodes. $\omega : E \rightarrow R^*$ is a nonnegative weight function for all edges. Figure 1 is a simple example of RLG containing two user nodes, 4 region nodes, and 6 business nodes. It shows that user u_1 interacts with his two movement regions R_{11} and R_{12} ; likewise, the user node u_2 interacts with his two movement region nodes R_{21} and R_{22} . Furthermore, region node R_{11} is also linked to business nodes i_1, i_2 , and i_4 because user node u_1 reviewed business nodes i_1, i_2 , and i_4 in his local movement region node R_{11} , and region node R_{12} is connected with business nodes i_3 because user node u_1 reviewed businesses node i_3 in his remote movement region node R_{12} . Similarly, region node R_{21} is connected with business nodes i_4 and i_5 and region node R_{22} is connected with business node i_6 .

In RLG, each user node connects with two movement regions, and the two movement regions only connect with some business nodes which were reviewed by user. If two users coreviewed a business, then their two movement regions would be overlapping; for example, in Figure 1, user

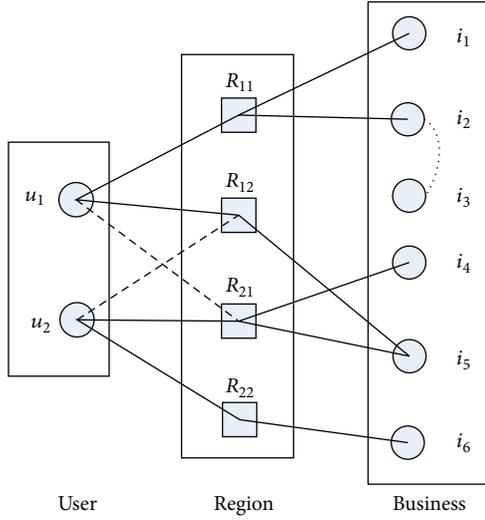


FIGURE 1: A simple example of RLG.

u_1 and user u_2 coreviewed, respectively, business i_4 in their local movement regions. This means that user u_2 is the local neighbor of user u_1 and the two regions are overlapping; that is to say, user u_1 could reach u_2 's local movement region and user u_2 could reach u_1 's local movement region, and they both would like some businesses in the region. Thus, based on the above empirical observations, we connect user node u_1 with region node R_{21} and connect user node u_2 with region node R_{11} in RLG (dotted lines), and if we start working from a user node (u_1), passing through a region node (R_{11}), we will find out his local neighbor (u_2) and we can reach an unknown business (i_5) in his local region; namely, $u_1 \rightarrow R_{11} \rightarrow u_2 \rightarrow R_{21} \rightarrow i_5$. In the same way, we can obtain another path $u_2 \rightarrow R_{21} \rightarrow u_1 \rightarrow R_{11} \rightarrow i_1(i_2)$ which connects user node u_2 with business node i_1 or i_2 . Hence, in this way, we can help user search for favorite business information in his movement region from local neighbors or remote friends.

The edge weights of RLG between user node and region node are defined as

$$\omega(u_p, R_{q,k}) = \begin{cases} \omega_{*1}, & \text{if } p = q, k = 1, \\ \omega_{*2}, & \text{if } p = q, k = 2, \\ \omega_{*2} = 1 - \omega_{*1}, \omega_{*2} < \omega_{*1} \\ \text{sim}(u_p, u_q)_{R_{p*} \cap R_{qk}}, & \text{if } p \neq q, u_p \in NB(u_q) \\ & \text{or } u_p \in RF(u_q) \\ 0, & \text{otherwis.} \end{cases} \quad (2)$$

Given an edge $e(u_p, R_{q,k})$, if $p = q, k = 1$, its weight will be ω_{*1} and if $p = q, k = 2$, its weight will be ω_{*2} , and, $\omega_{*1} > \omega_{*2}$, and since each user's local movement region or remote region is a part of his whole movement region and the probability of the user being active in local movement region is much greater than that in the remote region, so we let $\eta = \omega_{*1}/\omega_{*2}$ as a learning parameter in

the following experiment, and in this way, we use different weights to model the influence of local mobility preferences and long-distance mobility preferences. $\text{sim}(u_p, u_q)_{R_{p*} \cap R_{qk}}$ is the similarity between the two users who coreviewed lots of businesses. When one of the two users only reviewed one business, we use Jaccard coefficient (as shown in formula (6)) between the businesses reviewed by them as their similarity $\text{sim}(u_p, u_q)_{R_{p*} \cap R_{qk}}$, while when the two users reviewed more than one business, we calculate $\text{sim}(u_p, u_q)_{R_{p*} \cap R_{qk}}$ by using the Cosine similarity (as shown in formula (7)).

The edge weights of RLG between region node and business node are defined as

$$\omega(R_{qk}, i_j) = \begin{cases} \frac{r(u_p, i_j) - r_{\min}}{r_{\max} - r_{\min}}, & \text{if } (l_{jx}, l_{jy}) \in R_{qk}, k = 1, 2 \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The definition means the higher the reviewing score is, the more the user likes the business in a movement region. We denote by $\text{Pr}(G)$ the transition probability matrix of RLG:

$$\text{Pr}(G) = \begin{pmatrix} \text{Pr}(U \rightarrow R), & 0 \\ 0, & \text{Pr}(R \rightarrow I) \end{pmatrix}, \quad (4)$$

where $\text{Pr}(U \rightarrow R)$ is an $|U| \times |R|$ matrix representing the transition probability between user nodes and region nodes, as defined in (2) and $\text{Pr}(R \rightarrow I)$ is an $|R| \times |I|$ matrix representing the transition probability between region nodes and business nodes, as defined in (3), and they are symmetric matrixes. We will choose random walk with restarting process to simulate location-based business recommendation process.

5. Making Recommendation on RLG

So far, several graph-based methods have been introduced into recommendation system [19–21] to model the interaction between users and items on a graph and to compute node similarity from a global perspective, instead of local pairwise computation of neighborhood [19]. They essentially transform the recommendation process into graph search problem in a graph. Random walk on graphs has shown that it has a rather good performance in graph-based recommendation systems. PageRank, one of typical random walk algorithms, has been widely used in search engines to rank items globally.

Now we describe the recommendation process as a graph search problem in RLG and use the example shown in Figure 1. Suppose that the system needs to recommend lots of businesses to an active user in one of his movement regions. We firstly determine the association between the user and each of businesses which have not been reviewed by this user. The businesses are then sorted according to the associations and top N businesses are chosen for recommendation.

In our model, the association between two nodes is determined by all paths connecting them. For the pair of a user node u_i and a business node i_j , we compute the association between them as the sum of weights of all

```

(1) Initialize  $a_k(0) = 0$  for all  $n_i \in N$ ,  $a_i(0) = 1$ 
(2) for  $t = 1, 2, 3, \dots, T$ , or until convergence do
(3)   for each node  $n_k \in N$  do
(4)      $a_k(t) \leftarrow e_{ik}$ 
(5)     for each node  $n_i \in N$  do
(6)       if  $e_{ki} > 0$  or  $t = T$ 
(7)          $a_k(t) \leftarrow a_k(t) + \gamma e_{ki} a_k(t-1)$ 
(8)       end for
(9)     end for
(10)  end for
(11) return  $a_k(T)$ ,  $a_k(T)$  is the association between an active  $u_i$  and a business via paths of the length  $T$ .

```

ALGORITHM 1

TABLE 1: The total number of users who reviewed N businesses.

$N \geq 400$	$300 \leq N < 400$	$200 \leq N < 300$	$100 \leq N < 200$	$50 \leq N < 100$	$30 \leq N < 50$	$10 \leq N < 30$	$5 \leq N < 10$	$N < 5$
3	8	33	152	418	534	2977	4808	34940

distinct paths that connect u_i and i_j . In this computation, we differentiate two types of paths—paths via local neighbor nodes and paths via remote friend nodes. The length of the two paths is 4. As we defined problem earlier, for a user u_i , if the larger score is on a business reviewed by most of his local neighbors, this user probably will review this business in his local movement region. We adapt modification [22] of PageRank algorithm to calculate the association between an active user and recommended business i_j . For the ease of the algorithms description, let N denote the set of nodes that can form paths from u_i to business nodes; let $n_k \in N$ denote a node regardless of this is a user node, region node, or business node; let e_{kl} denote the weight of the link between nodes $n_k \in N$ and $n_l \in N$; and let e_{il} denote the link weight between node u_i and $n_l \in N$; then, the matrix is formed from $\text{Pr}(G)$. Furthermore, let $a_k(t)$ denote the association degree between u_i and node $n_k \in N$ when considering paths of length t ($t \leq T$). The algorithm for computing association between u_i and business nodes is shown in Algorithm 1.

Here, $\gamma \in [0, 1]$ is a parameter that downweights longer paths. We fix γ with 0.5 in our experiments. The most time consuming part of this algorithm is from line 3 to line 9 which requires $O(N^2)$ computations over all e_{kl} . However, the matrix (e_{kl}) is very sparse with most elements that are equal to zero and are symmetric. This allows us to use sparse and triangular matrix representation for (e_{kl}) , which can reduce the complexity to $O(C \cdot N)$, where C is the maximum number of nonzero elements for each row of matrix (e_{kl}) .

6. Experimental Evaluation

6.1. Data Set. We use dataset from Yelp in our following experiment, and it is publicly available [23]. It is from a US city, Phoenix; each review has a location (being reviewed by users) that is associated with a unique pair of latitude and longitude coordinates. It contains 43873 users, 229907 reviews, and 11537 pieces of business information. About half of all users reviewed just only one business, and consequently

the dataset is very sparse (99.9545% sparsity). The other pieces of information about the dataset are given in Table 1.

In the following section, we will discuss the location distribution of businesses reviewed by all users which can reveal all users' movement regions. We firstly randomly select ten users from the dataset, and the total number of their reviewed businesses, respectively, is 1, 1, 14, 22, 28, 49, 69, 91, 102, and 112, and the location distribution of businesses reviewed by ten users is as shown in Figure 2. The data from Figure 2 indicates that almost all businesses reviewed by each user are located in a certain region, and most of regions are overlapping, and obvious zoning appeared in the whole region. We can get an observation which is the same as the observation in [6]. To further verify the above-mentioned conclusion, we do some statistical analyses on the percentage of businesses reviewed by each user in three circle regions, whose center is user's movement center $o : (l_{x,o}, l_{y,o})$ and radii are, respectively, $D_{\max}/3$, $D_{\max}/2$, and $2D_{\max}/3$ (as shown in Figures 3, 4, and 5), and we, respectively, call them $D_{\max}/3$ region, $D_{\max}/2$ region, and $2D_{\max}/3$ region.

We can see from the three figures that the larger the number of businesses reviewed by a user is, the higher the proportion of businesses in the central zone is. It also demonstrates that most mobility of all users is restricted in a local region by their daily activities. Furthermore, the percentage of all users who reviewed more than half of businesses in $D_{\max}/2$ region is, respectively, 48.95%, 86.03%, and 99.48% in the three figures. Therefore, we call $D_{\max}/2$ region the local movement region of each user and the other is the remote region.

6.2. Evaluation Metric and Compared Methods. We use Hit ratio [24] as the metric for the top- N recommendation. Our dataset is split into training part and testing part: for each user, the latest businesses he reviewed are selected as test data and other businesses are selected as training data. When a recommendation is made, we always generate a list of N ($N = 10$) businesses named $R(u, R)$ for every user in his

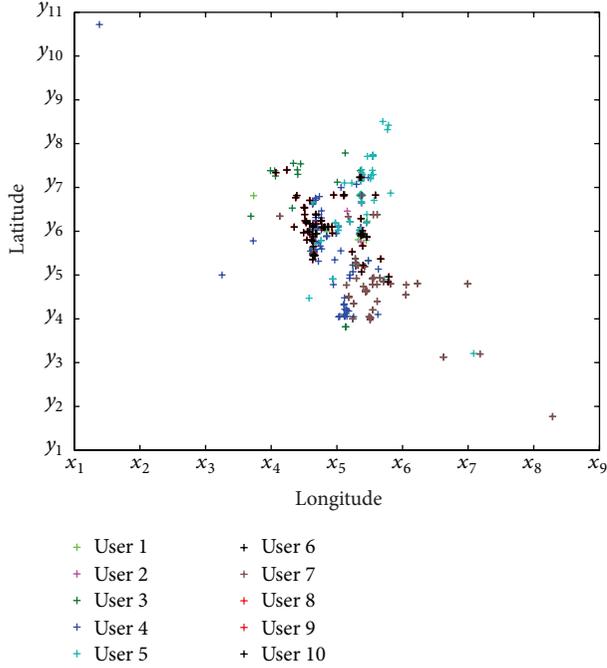


FIGURE 2: The location distribution of businesses reviewed by ten users.

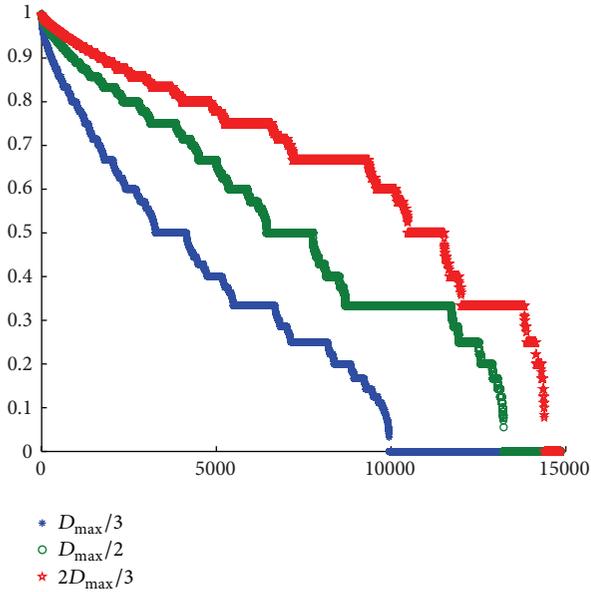


FIGURE 3: The percentage of businesses reviewed by 14583 users in three circle regions ($N > 3$).

whole movement region. If the test business appears in the recommendation list, we call it a hit, and then the Hit ratio can be calculated as follows:

$$\text{Hit_Ratio} = \frac{\sum_{u_k} h_{u_k}(i) \mid i \in R(u_k, R_{k*})}{N} \quad (5)$$

We compared the Top- N recommendation performance of our method with several existing methods: popularity-based

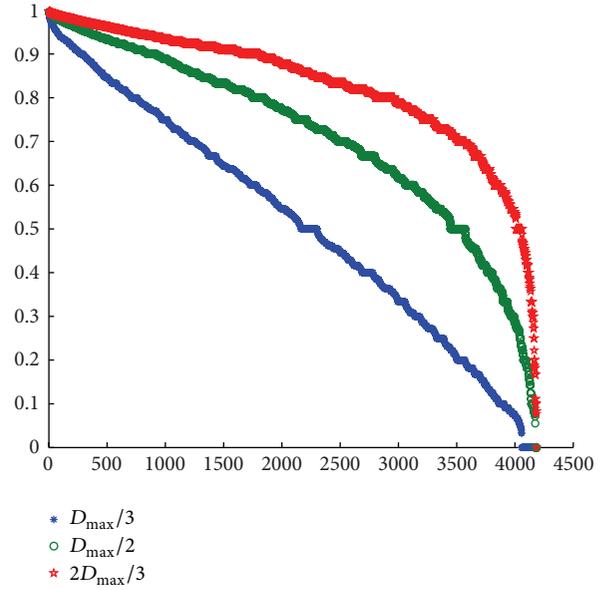


FIGURE 4: The percentage of businesses reviewed by 4185 users in three circle regions ($N > 10$).

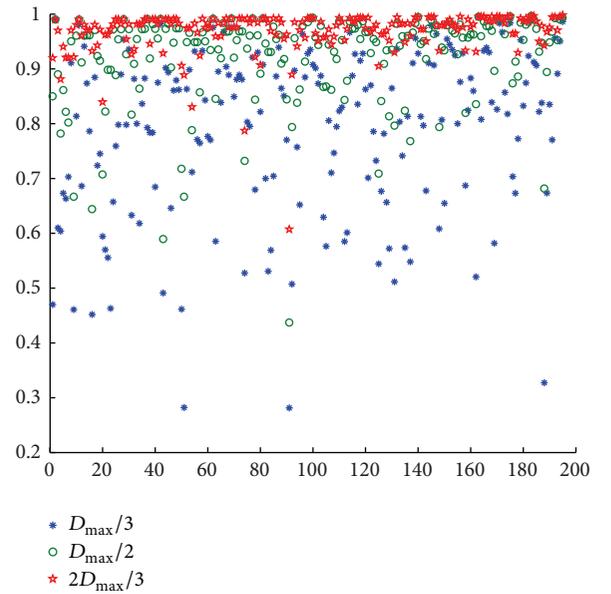


FIGURE 5: The percentage of businesses reviewed by 195 users in three circle regions ($N > 100$).

(Pop@1), item-based collaborative filtering (ItemKNN@1), user-based collaborative filtering (UserKNN@1), and their extended methods under condition of users' movement regions divided into two parts (Pop@2, ItemKNN@2, and UserKNN@2).

Popularity-Based (Pop) Method. Popularity-based method generates a ranking list based on the popularity of businesses in the training dataset. It is not a personalized method

and consequently generates the same list of recommended businesses for every user.

Item-Based Collaborative Filtering (ItemKNN) Method. Item-based collaborative filtering method finds the N businesses which are similar to some businesses reviewed by each user. The similarity between two businesses can be calculated as follows:

$$\text{sim}(i_m, i_n) = \frac{|s_m \cap s_n|}{|s_m \cup s_n|}. \quad (6)$$

User-Based Collaborative Filtering (UserKNN) Method. User-based collaborative filtering method finds the N similar users to the active user and aggregates the ratings of the similar users. The similarity between two users and aggregation function are as follows:

$$\begin{aligned} \text{sim}(u, u') &= \cos(u, u') \\ &= \frac{\sum_{i \in R(u) \cap R(u')} r_{u,i} \cdot r_{u',i}}{\sqrt{\sum_{i \in R(u) \cap R(u')} r_{u,i}^2 \cdot \sum_{i \in R(u) \cap R(u')} r_{u',i}^2}}, \end{aligned} \quad (7)$$

$$r_{u,i} = \bar{r}_u + \frac{\sum_{u' \in NB} (r_{u',i} - \bar{r}_{u'}) \cdot \text{sim}(u', u)}{\sum_{u' \in NB} |\text{sim}(u', u)|}. \quad (8)$$

6.3. Experimental Results. In this section, we illustrate the results of all methods and show performance of our method for all cold users who do not review any business information. We will firstly investigate the impact of parameter η and then compare the Hit ratios of the four methods recommending cold users.

6.3.1. The Impact of Parameter η . We focus on analyzing parameter η which governs the influence of users' local mobility preference formed by their daily activity and long-distance mobility preference formed by their social network ties in location-based businesses recommendation. When tuning η , the results of how Hit ratios change against all algorithms are shown in Figure 6. Pop@1, ItemKNN@1, and UserKNN@1 do not have parameter η ; thus, their Hit ratios are drawn as a straight line. The results show that Pop@2, ItemKNN@2, and UserKNN@2, respectively, outperform Pop@1, ItemKNN@1, and UserKNN@1, and our method always outperforms others whatever parameter η is. Moreover, when we set parameter η to 3/2, we can get the best results of Hit ratio for most of the methods, so we simply fix parameter η to 3/2 in the following experiments.

6.3.2. Make Recommendation for Cold Users. One challenge for most of existing methods is that the recommendation accuracy suffers when the user-business matrix is very sparse. From Table 2, we can see that over half users reviewed just one business in the dataset. Traditional user-based collaborative filtering cannot recommend any business to these users because of the fact that it is difficult to find out the nearest neighbors for these users in it. We regard the location of business reviewed by a user as the center of the user's local

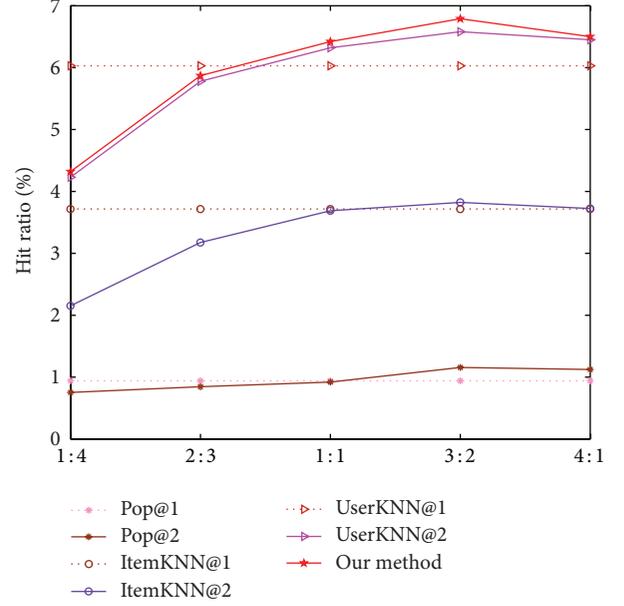


FIGURE 6: Hit ratios of all algorithms with different η .

TABLE 2: The number of users who reviewed N businesses ($1 \leq N \leq 5$).

$N = 5$	$N = 4$	$N = 3$	$N = 2$	$N = 1$
1625	2303	3617	7088	21932

region and the average $D_{\max}/2$ of other users as the radius of his local region. We use (6) to calculate $\text{sim}(u_p, u_q)_{R_{p^*} \cap R_{q^*}}$, and thus our method possesses the advantages of user-based collaborative filtering and item-based collaborative filtering. Our method can alleviate the sparsity problem by exploiting movement region to find out users' local neighbors and remote friends. To verify this hypothesis, we use four methods that recommend some businesses in their local region to all cold users, and the results are shown in Table 3. Apparently, our method has better performance than other methods.

7. Conclusion

User mobility often exhibits long- and short-distance factors which, respectively, formed daily activity and social network ties. Tracking and leveraging these factors for location-based business information recommendation pose great challenges. In this paper, we construct a region-based location graph (RLG), which can combine with user short-ranged mobility formed by daily activity and long-distance mobility formed by social network ties and sequentially can recommend local business information and long-distance business information to users. Moreover, it can combine user-based collaborative filtering with item-based collaborative filtering and can be successful in generating recommendation for cold start users, and, consequently, it can alleviate cold start problem which traditional recommender systems often suffer from.

TABLE 3: Hit ratios of algorithms for cold users ($\eta = 3 : 2$).

Method	Hit ratio (%)	Improvement
UserKNN@2	0	0
Pop@2	0.94	—
ItemKNN@2	1.428	51.9%
Our method	2.58	80%

The experiments on real dataset confirm that the effectiveness of the proposed method is better than that of other methods.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 60872051) and the Mutual Project of Beijing Municipal Education Commission, China.

References

- [1] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*, pp. 325–334, July 2011.
- [2] J. Sang, T. Mei, and J. Sun, "Probabilistic sequential POIs recommendation via check-in data," in *Proceedings of 20th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 402–405, November 2012.
- [3] T. H. Dao, S. R. Jeong, and H. Ahn, "A novel recommendation model of location-based advertising: context-aware collaborative filtering using GA approach," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3731–3739, 2012.
- [4] C. Cheng, H. Yang, I. King, and M. R. Lyu, "Fused matrix factorization with geographical and social influence in location-based social networks," in *Proceedings of 26th AAAI Conference on Artificial Intelligence*, pp. 17–23, July 2012.
- [5] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: successive point-of-interest recommendation," in *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI '13)*, pp. 2605–2611, August 2013.
- [6] 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 (KDD '11)*, pp. 1082–1090, August 2011.
- [7] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proceedings of the 18th International World Wide Web Conference (WWW '09)*, pp. 791–800, April 2009.
- [8] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *Proceedings of the 19th International World Wide Web Conference (WWW '10)*, pp. 1029–1038, April 2010.
- [9] V. W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang, "Collaborative filtering meets mobile recommendation: a user-centered approach," in *Proceedings of the 24th AAAI Conference on Artificial Intelligence and the 22nd Innovative Applications of Artificial Intelligence Conference (AAAI '10)*, pp. 236–241, July 2010.
- [10] K. W.-T. Leung, D. L. Lee, and W.-C. Lee, "CLR: a collaborative location recommendation framework based on co-clustering," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*, pp. 305–314, July 2011.
- [11] X. Yu, A. Pan, L.-A. Tang, Z. Li, and J. Han, "Geo-friends recommendation in GPS-based cyber-physical social network," in *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining (ASONAM '11)*, pp. 361–368, July 2011.
- [12] Y. Zheng and X. Xie, "Learning travel recommendations from user-generated GPS traces," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 1, article 2, 2011.
- [13] X. Cao, G. Cong, and C. S. Jensen, "Mining significant semantic locations from GPS data," *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 1009–1020, 2010.
- [14] M. Ye, P. Yin, and W. C. Lee, "Location recommendation for location-based social networks," in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 458–461, 2010.
- [15] J. J. Ying, E. H. Lu, and W. Kuo, "Urban point-of-interest recommendation by mining user check-in behaviors," in *Proceedings of the ACM SIGKDD International Workshop on Urban Computing (KDD '12)*, pp. 63–70, August 2012.
- [16] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proceedings of the 20th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '12)*, pp. 199–208, November 2012.
- [17] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag recommendations in social bookmarking systems," *AI Communications*, vol. 21, no. 4, pp. 231–247, 2008.
- [18] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Towards mobile intelligence: learning from GPS history data for collaborative recommendation," *Artificial Intelligence*, vol. 184-185, pp. 17–37, 2012.
- [19] L. Xiang, Q. Yuan, S. Zhao et al., "Temporal recommendation on graphs via long- and short-term preference fusion," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10)*, pp. 723–731, July 2010.
- [20] S. Lee, S.-I. Song, M. Kahng, D. Lee, and S.-G. Lee, "Random walk based entity ranking on graph for multidimensional recommendation," in *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys '11)*, pp. 93–100, October 2011.
- [21] Z. Zhang, D. D. Zeng, A. Abbasi, J. Peng, and X. Zheng, "A random walk model for item recommendation in social tagging systems," *ACM Transactions on Management Information Systems*, vol. 4, no. 2, article 8, 2013.
- [22] J. Weston, A. Elisseeff, D. Zhou, C. S. Leslie, and W. S. Noble, "Protein ranking: from local to global structure in the protein similarity network," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 17, pp. 6559–6563, 2004.

- [23] B. Hu and E. Martin, "Spatial topic modeling in online social media for location recommendation," in *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys '13)*, pp. 25–32, October 2013.
- [24] G. Karypis, "Evaluation of item-based top-N recommendation algorithms," in *Proceedings of the 10th International Conference on Information and Knowledge Management*, pp. 247–254, November 2001.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

