2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)
March 29–31, 2018, Xiamen, China

A Multi-Factor Influencing POI Recommendation Model Based on Matrix Factorization

YING XU, YING LI, and WEI YANG
College of Computer Science and Electronic Engineering
Hunan University
Changsha, China
hnxy@hnu.edu.cn, lying@hnu.edu.cn, weeyang@163.com

JIN ZHANG
Key Laboratory of High Performance Computing and Stochastic Information Processing of the Ministry of Education of China,
Hunan Normal University
Changsha, China
jinzhang@hunnu.edu.cn

***Abstract*—How to make recommendation for personalized users by using the available sparse data is a hot research topic in the area of big data and has wide application prospects. In this work, we investigate the POI (Point of Interest) recommendation of LBSN (Location Based Social Network) to provide users with personalized POI preference, such as attractions, hotels and shops and so on. A new POI recommendation model based on matrix factorization by considering the influences of both the geographical factor and the user factor, namely GeoUMF (Geographical and User Matrix Factorization), has been proposed in this paper. In GeoUMF, the objective function considers the difference between the ranking produced in the recommendation model and the actual ranking in the check-in data. In addition, an approximation method that considers the difference of visiting frequency of POI is defined in the objective function. Experimental results on real world LBSN data set demonstrate that GeoUMF obtained better performance in terms of the recommendation precision and the recall rate compared with some state-of-the-art algorithms in the literature.**

Keywords—recommendation system; point of interest (POI); matrix factorization; LBSN

I. INTRODUCTION

With the rapid development of Location Based Social Network (LBSN), Point of interest (POI) recommendation of LBSN is a hot research topic, which can be widely used in tourism recommendation, location navigation, advertising, etc. POI recommendation aims to mine user's preferences or other information based on user's historical travel data to recommend POIs to users. LBSN based POI recommendation is complicated since it needs to consider the user's social information and current location information simultaneously. POI recommendation has thus attracted much researcher's attention. For POI recommendation, some crucial issues need to be solved. For example: 1) The sparsity of the collected user data; 2) The effect of different factors within user's trajectory data need to be considered, such as the user's social information, geographical information, etc.

A widely used method in recommendation system is Collaborative Filtering (CF), which makes predictions (Filtering) of a user's interests by collecting preferences or information from multiple users (Collaborating). The basic idea of CF is that if two persons A and B have the same opinion on one item, A is more likely to have the same opinion of B on another item. CF can be divided into three categories: 1) Memory-based CF [1], which uses user rating data to measure the similarity between users or items; 2) Model-based CF [2,3], where models can be various data mining, machine learning algorithms to predict users' rating of items; 3) Hybrid CF [1, 4], which combines both memory-based and model-based CF algorithms to improve prediction accuracy of classic memory-based or model-based CF.

In order to deal with highly sparse data, Matrix Factorization (MF) has been proved to have high prediction accuracy and good flexibility for a variety of practical applications [5, 6]. The principle of MF is to decompose user-POI rating matrix into the product of the user latent factor matrix and the POI latent factor matrix. Two matrices are then trained to recalculate the user-POI rating matrix. Obviously, geographical distance has impact on POIs. For example, after the user visited POI *A*, the probability of visiting the POI around POI *A* will be relatively high. The relation-based MF model needs to consider the relationship between users, so it is important to mine user's relationship [7, 8]. For instance, the information that users visit the same POI may indicate the similar behavior or friendship between users. However, most traditional CF algorithms ignore the relationship between users.

In this work, we propose a multi-factor influencing POI recommendation method based on MF, where the effects of multi-factors including the geographical factor and the user factor have been considered. The main contributions of our work are as follows:

- A recommendation model based on MF is proposed to solve the sparsity of check-in data by MF which uses both visited POIs and unvisited POIs. The influence of both the geographical factor and the user factor has been considered in the model.
- A ranking-based objective function is designed in the proposed model, which considers the inconsistency between the ranking of the recommendation results and the ranking in the real data. The aim is to improve the sensitivity of the objective function to the error of POI recommendation.
- To deal with the problem of discontinuous of the objective function, an approximation method that

considers the difference of visiting frequency of POI is used in the objective function. Because a POI with higher visiting frequency should has a higher ranking.

II. RELATED WORK

POI recommendation has attracted many researchers' attention, thus a variety of POI recommendation methods have been proposed in the literature. For POI recommendation, many factors need to be considered, such as the personal interest preferences, the impact of geographic distance, etc. The City Voyager system proposed in [9] can recommend a satisfying shop or bar to users, where the user's preferences and habits are extracted based on the user's historical location information. Gao et al. [10] proposed a location-based and preference-aware recommend system. In their system, user preferences are mined from the location history and social opinions are learned from local experts' data. Wei et al. [11] proposed the Route Inference framework based on Collective Knowledge (RICK) model, the city map is gridded to find the path with the highest popularity among different regions. However, because the proposed model focuses on the hot paths between regions, hot paths for smaller areas are ignored, so that the recommendation accuracy is low.

Ye et al. [12] used the linear interpolation to combine social and geographical impacts into a user-based CF framework. Levandoski et al. [13] proposed an extended project-based filtering approach that improves the performance of the proposed scheme. Their scheme considers spatial properties of both users and items to make the rating. Noulas et al. [14] applied the traditional MF method to feedback data in POI recommendation. Cheng et al. [15] proposed a multi-center Gaussian model to compute the geographical impact. However, this multi-center Gaussian model based MF method cannot effectively deal with sparse data. Liu et al. [16] proposed a Geographical-Topical Bayesian Non-negative Matrix Factorization (GTBNMF) method that incorporates the geographical impact and the text impact. Lian et al. [17] proposed a GeoMF model, which decomposes Weighted Matrix Factorization (WMF) and introduces the geographical influence. To deal with the sparse check-in data, nonzero check-ins are set large weights and zero check-ins (unvisited POIs) are set smaller weights. Yuan et al. [18] proposed a method called User-based CF with Temporal preference and smoothing Enhancement with Spatial influence based recommendation with popularity Enhancement (UTE+SE), which introduces temporal and geographical effects using the user-based filtering. Gao et al. [19] proposed a normalized nonnegative MF method without considering the geographical impact. A Breadth-first Preference Propagation has been proposed in [20], which is superior to the UTE+SE method because it uses preference propagation on a graph generated based on the check-in data to make recommendations.

Li et al. [21] proposed the Ranking based Geographical Factorization Method (Rank-GeoFM). The loss (objective) function is obtained by comparing the ranking of the recommendation results with the ranking of the actual check-ins data. But they only analyzed the impact of those POI that are misplaced in front of the current POI without considering the impact of the POI that are misplaced behind the current POI to the loss function. In addition, the difference of visiting frequency of the POIs has not been considered, which actually affects the loss function.

In order to solve these problems, in this paper, we propose a recommendation model based on MF. A Decomposition Method Considers the influence of geographical factors and user factors, namely Geographical and User Matrix Factorization (GeoUMF) algorithm. The recommendation model takes into account both the POI that is misplaced in front of the current POI and the POI misplaced behind the current POI in the objective function. The idea is to improve the sensitivity of objective function to the misplaced POI. In addition, to make the objective function reflect the real check-in data, the difference of visiting frequency between POIs has been considered. Two influence factors, including the geographical factor and the user factor, have been introduced to improve the accuracy of recommendation.

III. THE PROPOSED GEOUMF ALGORITHM

We propose a new POI recommendation method GeoUMF based on the ranking-based MF recommendation by considering both the geographic factor and the user factor. To formally define the POI recommendation problem, the main notations used are listed in Table 1.

TABLE I. THE DEFINITION OF NOTATIONS

Symbol	Description
U	the user set $\{u_1, u_2, u_3 ..., u_{\lvert U \rvert}\}$
L	the POI set $\{\ell_1, \ell_2, \ell_3, \ldots, \ell_{\lvert L \rvert}\}$
L_u	the POI set that user u has visited
$X=[x_{u\ell}]$	a user-POI check-in matrix $\lvert U \rvert \times \lvert L \rvert$ $x_{u\ell}$ is the visiting times that user u visits the POI ℓ
D_1	the user-POI pairs: $D_1 = \{(u,\ell) \mid x_{u\ell} > 0\}$
$dis(\ell, \ell')$	the distance between two POIs ℓ and ℓ'
$N_{k(\ell)}$	the set of k nearest POIs of ℓ
$y_{u\ell}$	The recommendation score of user u for POI ℓ

A. The Preference Ranking Objective Function

Firstly, we analyze the ranking of user's preferences to all POIs using check-in data. Assuming X is the user-POI check-in matrix, for a user u with two POIs ℓ and ℓ', if $x_{u\ell} > x_{u\ell'}$, the rating of ℓ should be higher than ℓ'. Li et al. (Li et al., 2015) considered that the inconsistent of the rating in the recommendation model, when the visiting frequency of the ℓ is more than the visiting frequency of ℓ'. However if the visiting frequency of ℓ is less than ℓ', but the ranking of POI ℓ is before ℓ' in the recommendation model, this inconsistent is not considered. Therefore, we define a new objective function (loss function) which considers the inconsistency ranking of two cases: $x_{u\ell} > x_{u\ell'}$ and $x_{u\ell} < x_{u\ell'}$, the inconsistency can be calculated as follow:

$$Incomp\left(y_{u\ell},\varepsilon\right)=\sum_{\ell'\in L} I\left((x_{u\ell}-x_{u\ell'})((y_{u\ell'}+sign(x_{u\ell}-x_{u\ell'})\varepsilon)-y_{u\ell})\right) \quad (1)$$

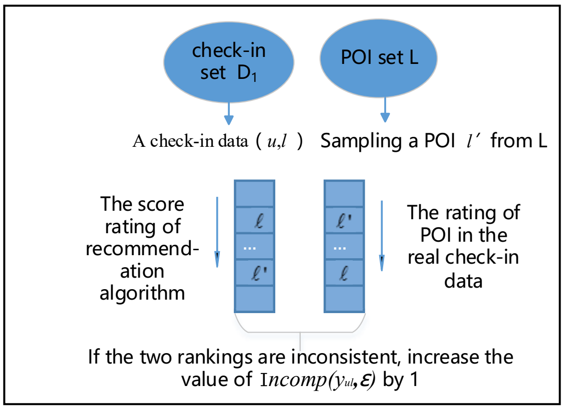


Fig. 1. The Inconsistency of ranking.

Where $I(\bullet)$ is the indicator function, if the internal value is true, $I(\bullet)=1$. Otherwise, $I(\bullet)=0$. $sign(\bullet)$ is a function which returns the parameter's signal. ε is a constant and ε >0. $x_{u\ell}$ is the number of visiting times that user u has visited ℓ. $y_{u\ell}$ is the rating score of ℓ. $Incomp\left(y_{u\ell},\varepsilon\right)$ is the number of inconsistencies between the recommended and actual ranking. When the visiting frequency of user u to POI ℓ is greater than the visiting frequency of u to ℓ', it means user u prefers ℓ. However, if the rating score of ℓ of the recommendation model is less than ℓ', here we define a constant ε, if the visiting frequency of ℓ is more than the visiting frequency of ℓ', the corresponding rating score should satisfy $y_{u\ell} > y_{u\ell'}$, and the difference should also no less than ε. Otherwise, the ratings of two POIs are considered to be inconsistent. The more inconsistency exists, the greater the value of $Incomp\left(y_{u\ell},\varepsilon\right)$. So we can calculate the inconsistent for the entire user dataset D_1. The objective function can be defined as follow:

$$\mathrm{O}=\sum_{(u,\ell)\in D_1} E\left(Incomp\left(y_{u\ell},\varepsilon\right)\right) \quad (2)$$

Where $D_1:\{u,\ell \mid x_{u\ell}>0\}$, L is the POI set, and $E(r)=\sum_{i=1}^{r}\frac{1}{i}$. As defined in [22, 23], $E(\cdot)$ is used to make $Incomp\left(y_{u\ell},\varepsilon\right)$ smoother, here we set $E(0)=0$.

B. The Matrix Factorization Method

The main idea of MF is to decompose the user-POI matrix $\hat{X}\in\mathbb{R}^{|U|\times|L|}$ into the product of the user latent factor matrix $U^{(1)}\in\mathbb{R}^{|U|\times K}$ and the POI latent factor matrix $L^{(1)}\in\mathbb{R}^{|L|\times K}$. In order to define $U^{(1)}$ and $L^{(1)}$, we introduce two K dimensional vector $u_u^{(1)}\in\mathbb{R}^{1\times K}$ and $l_\ell^{(1)}\in\mathbb{R}^{1\times K}$ for each user and each POI, respectively. We define $\hat{X}=U^{(1)}\times L^{(1)T}$ as the prediction scoring matrix obtained from $U^{(1)}$ and $L^{(1)}$. The purpose of the matrix decomposition model is to make this matrix $\hat{X}\in\mathbb{R}^{|U|\times|L|}$ consistent with the user-POI visiting frequency matrix $X\in\mathbb{R}^{|U|\times|L|}$ as much as possible, i.e. $\hat{X}=U^{(1)}\times L^{(1)T}\approx X$. In this way, the matrix decomposition model maps the user and the POI to a common latent factor space, and the product of the two vectors represents the user's POI score $y_{u\ell}=u_u^{(1)}\cdot l_\ell^{(1)}$, i.e. the user's basic preference score for POIs.

C. Multi-factors in GeoUMF

In this work, we consider two factors, the geographical factor and the user factor in the basic MF recommendation model to propose a new recommendation algorithm GeoUMF.

1) Geographical Factor

In order to add the geographical factor [24], we introduce another user latent factor matrix $U^{(2)}\in\mathbb{R}^{|U|\times K}$ by calculating the geographical factor impact score $y_{u\ell}^{G}=u_u^{(2)}\cdot\sum_{\ell*\in N_k(\ell)} w_{\ell\ell*}l_{\ell*}^{(1)}$. Here $u_u^{(2)}$ is the user latent factor vector of user u in $U^{(2)}\in\mathbb{R}^{|U|\times K}$, $N_k(\ell)$ is the k nearest POIs of ℓ, i.e. top-k POIs selected from $w_{\ell\ell_1},w_{\ell\ell_2},\dots,w_{\ell\ell_{|L|}}$ in descending order. $w_{\ell\ell*}=\frac{1}{0.5+dis(\ell,\ell*)}$ is the probability that the user will visit ℓ after the user visited $\ell*$. Here, $dis(\ell,\ell*)$ refers to the distance between ℓ and $\ell*$. Calculate $w_{\ell\ell*}$ for all POIs, the weight matrix of the geographical factor $W\in\mathbb{R}^{|L|\times|L|}$ can be obtained.

2) User Factor

In order to find this potential relationship, we define W_{ij} as the weight vector between user u_i and u_j, which counts the number of u_i and u_j visit the same POI. The user correlation coefficient matrix $T\in\mathbb{R}^{|U|\times|U|}$ can be obtained by calculating each item $T_{ij}=\frac{W_{ij}}{|S(u_i)\cup S(u_j)|}$ which is the correlation coefficient between user u_i and u_j. $S(u_i)$ and $S(u_j)$ represent the set of POIs visited by u_i and u_j, respectively. The u-th row in T includes the nearest users of user u, the top-k nearest users set $N_k(u)$ is selected based on the top-k values of T_{ij}. In order to add user factors, the user latent factor matrix $L^{(2)}\in\mathbb{R}^{|L|\times K}$ is defined, in which each item is the user factor's score $y_{u\ell}^{U}=l_\ell^{(2)}\cdot\sum_{v\in N_k(u)} T_{uv}u_v^{(1)}$.

IV. THE RECOMMENDATION MODEL

A new recommendation model GeoUMF is proposed based on the matrix factorization framework introduced in Section III, which considers the influence of the geographical factor and the user factor, and then the optimization and learning strategies have been designed.

A. The Multi-factors Recommendation Model

The basic matrix factorization recommendation method cannot fully mine the user's preference, we thus introduce the influence of the geographical factor and the user factor, the recommendation model can be defined as follows:

$$y_{ul} = u_u^{(1)} \times l_l^{(1)} + u_u^{(2)} \times \sum_{l^* \in N_k(l)} w_{ll^*} l_{l^*}^{(1)} + l_l^{(2)} \times \sum_{v \in N_k(u)} T_{uv} u_v^{(1)} \tag{3}$$

The latent factor matrix set that consists of all POIs and user latent factor vectors is defined as $\Theta := \{L^{(1)}, L^{(2)}, U^{(1)}, U^{(2)}\}$. Then it is learned by the learning algorithm (see details in the next section). In order to avoid over-fitting, using the method in [18], the regularization item is defined to limit the size of each latent factor vector as follows:

$$\left\| l_\ell^{(1)} \right\|_2 \le C \qquad \ell = 1,2,...,|L| \tag{4}$$

$$\left\| l_\ell^{(2)} \right\|_2 \le \beta C \qquad \ell = 1,2,...,|L| \tag{5}$$

$$\left\| u_u^{(1)} \right\|_2 \le C \qquad u = 1,2,...,|U| \tag{6}$$

$$\left\| u_u^{(2)} \right\|_2 \le \alpha C \qquad u = 1,2,...,|U| \tag{7}$$

Where $C > 0$, $0 \le \alpha \le 1$ and $0 \le \beta \le 1$ are hyper-parameters. The latent factor vector of $L^{(1)}$, $L^{(2)}$ and $U^{(1)}$ is confined within a sphere of radius C. So the score of $u_u^{(1)} \cdot l_\ell^{(1)}$ is always within the range $[-C^2, C^2]$. The latent factor vector of $U^{(2)}$ is confined within a sphere of radius αC. The parameter α is used to prevent over-fitting and limit the weight of geographic impact. Similarly, the latent factor vector of $L^{(2)}$ is confined within a sphere of radius βC. The introduction of β is to balance the influence of the user factor and prevent the over-fitting.

B. *The Optimization and Learning Algorithm*

The objective function defined in (2) needs to be smoothed to deal with the discontinuous problem. To make the objective function become continuous, $E(Incomp(y_{u\ell}, \varepsilon))$ can be calculated as in (8) by using the method in [21], so the partial derivative can be obtained. Since the greater gap between two POIs visiting frequencies, the more losses if these POIs are ranked error in the recommendation model. Therefore, the objective function is smoothed by considering the visiting frequency difference using $(x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell})$.

$$\begin{aligned} &E(Incomp(y_{u\ell}, \varepsilon)) = \\ &E(Incomp(y_{u\ell}, \varepsilon)) \frac{\sum_{\ell' \in L} I((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))}{Incomp(y_{u\ell}, \varepsilon)} \\ &= E(Incomp(y_{u\ell}, \varepsilon)) \frac{\sum_{\ell' \in L(u,\ell)} I((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))}{Incomp(y_{u\ell}, \varepsilon)} \\ &\approx E(Incomp(y_{u\ell}, \varepsilon)) \frac{\sum_{\ell' \in L(u,\ell)} S((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))}{Incomp(y_{u\ell}, \varepsilon)} \end{aligned} \tag{8}$$

According to the sampling learning method in [15], the minimization objective function can be directly substituted by minimizing the loss expectation of a single sample. The loss expectation for each sample is calculated as follows:

$$\overline{E} = E(Incomp(y_{u\ell}, \varepsilon)) s((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell})) \tag{9}$$

$Incomp(y_{u\ell}, \varepsilon) = \left| \frac{|L|}{n} \right|$ is the geometric distribution sampling estimation [23]. n is the sampling number, i.e. sample n times to get the misplaced sampled data. The partial derivation of the sample loss expectation is:

$$\begin{aligned} \frac{\partial \overline{E}}{\partial \Theta} &= E(Incomp(y_{u\ell}, \varepsilon)) \frac{\partial s((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))}{\partial \Theta} \\ &\approx E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} \frac{\partial((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))}{\partial \Theta} \end{aligned} \tag{10}$$

$$\delta_{u\ell\ell'} = s((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))(1 - s((x_{u\ell} - x_{u\ell'})((y_{u\ell'} + sign(x_{u\ell} - x_{u\ell'})\varepsilon) - y_{u\ell}))) \tag{11}$$

where $s(a) = \frac{1}{1 + e^{-a}}$, Θ is calculated by Stochastic Gradient Descent (SGD) as follows:

$$\Theta \leftarrow \Theta - \gamma \frac{\partial \overline{E}}{\partial \Theta} \tag{12}$$

Where γ is the learning rate. The latent factor vector in $\Theta := \{L^{(1)}, L^{(2)}, U^{(1)}, U^{(2)}\}$ can be updated as follows:

$$u_u^{(1)} \leftarrow u_u^{(1)} - \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} \left(l_{\ell'}^{(1)} - l_\ell^{(1)} \right) \tag{13}$$

$$u_u^{(2)} \leftarrow u_u^{(2)} - \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} \left(\sum_{\ell^* \in N_k(\ell')} \omega_{\ell'\ell^*} l_{\ell^*}^{(1)} - \sum_{\ell^+ \in N_k(\ell)} \omega_{\ell\ell^+} l_{\ell^+}^{(1)} \right) \tag{14}$$

$$l_{\ell'}^{(1)} \leftarrow l_{\ell'}^{(1)} - \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} u_u^{(1)} \tag{15}$$

$$l_\ell^{(1)} \leftarrow l_\ell^{(1)} + \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} u_u^{(1)} \tag{16}$$

$$l_{\ell'}^{(2)} \leftarrow l_{\ell'}^{(2)} - \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} \sum_{v \in N_k(u)} T_{uv} u_v^{(1)} \tag{17}$$

$$l_\ell^{(2)} \leftarrow l_\ell^{(2)} + \gamma E\left(\left| \frac{|L|}{n} \right| \right) \delta_{u\ell\ell'} \sum_{v \in N_k(u)} T_{uv} u_v^{(1)} \tag{18}$$

Fig. 2 presents the pseudo code of our proposed GeoUMF.

Algorithm 1: GeoUMF

Input: check-in data D_1, parameters W, T, γ

Output: model parameters $\Theta := \{L^{(1)}, L^{(2)}, U^{(1)}, U^{(2)}\}$

Initialize Θ with a normal distribution;

Randomly disorganize check-in data D_1;

Repeat

  for $(u, \ell) \in D_1$ do

    Calculate the score $y_{u\ell}$, set n=0;

    Repeat

      Sampling a POI ℓ';

      Calculate the score $y_{u\ell'}$, set n=n+1;

    Until $I(x_{u\ell} > x_{u\ell'}) I(y_{u\ell} < y_{u\ell'} + \varepsilon) = 1$ or $n > |L|$;

    if $I(x_{u\ell} > x_{u\ell'}) I(y_{u\ell} < y_{u\ell'} + \varepsilon) = 1$ then

      Calculate based on (12)- (17);

    if $\left\| l_\ell^{(1)} \right\|_2 > C$ then $\left\| l_\ell^{(1)} \right\|_2 = C$;

      Check (4), (5), and (6);

Until convergence

Return $\Theta := \{L^{(1)}, L^{(2)}, U^{(1)}, U^{(2)}\}$

Fig. 2. The pseudo code of GeoUMF.

V. Performance Evaluation

In order to evaluate the performance of our proposed GeoUMF based on MF, we compare it with other two algorithms in the literature:

- Rank-GeoFM: a POI recommendation algorithm based on MF which considers the geographical factor.
- Rank-GeoFM/G: the Rank-GeoFM without the geographical factor.

A. Test Datasets

In the experiment, we use Singapore's Foursquare check-in data to test the algorithm. This data set contains 194108 check-in data of 2321 users in 5596 POIs from Aug. 2010 to July 2011. The data set is divided into three parts: 30% of the data is the training data, 20% of the data is the test data and the rest is used to tune parameters.

B. Evaluation Metrics

Two metrics, i.e. the precision and the recall, are used to evaluate the performance of different recommendation algorithms. The precision for user u is defined as follows:

$$\mathrm{Pr}e_u @N = \frac{tp_u}{tp_u + fp_u} \tag{19}$$

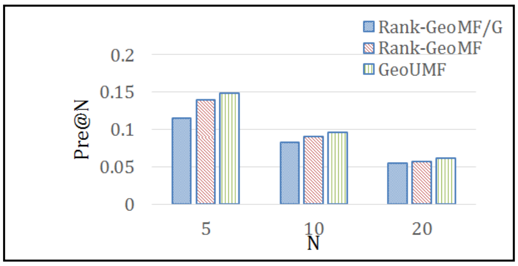
where $@N$ indicates the number of recommended POIs is *N*. tp_u represents the number of POIs which are correctly recommended in the top-*N* POIs compared with the real check-in data, fp_u is equal to the number of POIs contained in the top-*N* recommended POIs but not contained in the real check-in data. The recall of recommendation for user u is defined as follow:

$$\mathrm{Re}c_u @N = \frac{tp_u}{tp_u + tn_u} \tag{20}$$

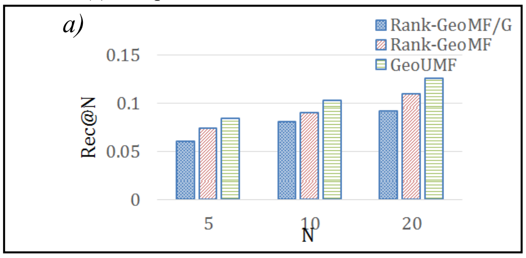
Where tn_u is the number of POIs which are not contained in the top-*N* recommended POIs but are visited in the real check-in data.

C. Experimental Results

In the experiments, we set the hyper-parameters $\varepsilon = 0.3$ and $C = 1$, the latent factor dimensions $k = 200$. In addition, in order to ensure sufficient accuracy, the learning rate should be small enough, so the learning rate is set as 0.0001. We set $\beta = 0.1$ through a number of initial tests on the tuning data set. We compare the precision and recall of the proposed GeoUMF with two state-of-the-art algorithms, including Rank-GeoMF/G and Rank-GeoMF on the Foursqure data set. Experimental results of $\mathrm{Pr}e@N$ and $\mathrm{Re}c@N$ for $N = 5$, $N = 10$ and $N = 20$ have been shown in Fig. 3.



(a) Comparison of Precision



(b) Comparison of Recall

Fig. 3. Performance Comparison on Foursqure.

For the Foursquare dataset, experimental results show that the Rank-GeoFM/G algorithm without considering the geographical factor or the user factor has the worst performance. The Rank-GeoFM algorithm, which introduces the geographical factor, has a better performance compared with Rank-GeoFM/G. Our proposed GeoUMF performs the best by considering both the geographical factor and the user factor. In addition, the difference of visiting frequency introduced in the objective function also help to improve the recommendation accuracy.

VI. Conclusions

In this work, a new multi-factor influencing matrix decomposition based POI recommendation model has been proposed by considering both the geographic factor and the user factor. A new objective function which fully considers the inconsistency between the scored ranking and the actual ranking in the check-in data has been designed. In addition, an approximation method that considers the difference of visiting frequency of POI is used in the objective function. Experimental results demonstrate our proposed GeoUMF obtained better performance compared with other two algorithms in the literature. More impact factors can be considered in the future. For example, in general, individuals have similar preferences for similar objects, so a similar set of POIs can be added to the recommendation model as another impact factor.

Acknowledgment

This research has been supported by Science and Technology Plan of Hunan Province (Project No. 2015GK3015); National Key Research and Development Plan (Project No. 2016YFB0200405); Project supported by the Education Department of Hunan Province, China (Grant No. 15k082).

References

[1] F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor, "Recommender Systems Handbook," Springer, Springer Science Business Media LLC. 2011.

[2] D. Christopher, P.M. Raghavan, and H. Schutze, "Introduction to Information Retrieval," Cambridge University Press. 2008.

[3] H. Yin, B. Cui, Z. Huang, W. Wang, X. Wu, and X. Zhou, "Joint Modeling of Users' Interests and Mobility Patterns for Point-of-Interest Recommendation," In Proceedings of MM'15, pp. 26-30, Brisbane, Australia. 2015.

[4] C.X. Yin and Q.K. Peng, "A New Hybrid Collaborative Filtering Algorithm," Applied Mechanics & Materials, vol. 135-136, pp. 80-86, 2011.

[5] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui, "GeoMF: Joint Geographical Modeling and Matrix Factorization for Point-of-interest Recommendation," In Proceedings of SIGKDD, pp. 831-840, ACM, 2014.

[6] X. Li, C. Gao, X.L. Li, T.A. Pham, Nguyen, and S. Krishnaswamy, "Rank-GeoFM: A Ranking Based Geographical Factorization Method for Point of Interest Recommendation," In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 433-442, New York, NY, USA: ACM, 2015.

[7] M. Jamali and M. Ester, "A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks," In Proceedings of ACM Conference on Recommender Systems, pp. 135-142. Barcelona, Spain, 2010.

[8] Z. Zhang and H. Liu, "Social Recommendation Model Combining Trust Propagation and Sequential Behaviors," Applied Intelligence, vol. 43(3), pp. 695-706, 2015.

[9] A. Majid, L. Chen, and G. Chen, et al. "A Context-aware Personalized Travel Recommendation System based on Geotagged Social Media Data Mining," International Journal of Geographical Information Science, vol.27(4), pp.662-684, 2013.

[10] J. Gao, 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, vol. 2012, pp. 199-208, USA: ACM, 2012.

[11] L. Wei, Y. Zheng, and W.C. Peng, "Constructing Popular Routes from Uncertain Trajectories," In Proceedings of the 18th ACM SIGKDD, pp. 195-203, 2012.

[12] M. Ye, P. Yin, W.C. Lee, and D.L. Lee. "Exploiting geographical influence for collaborative point-of-interest recommendation," In Proceedings of SIGIR, ACM, pp. 325-334, 2011.

[13] J.J. Levandoski, M. Sarwat, A. Eldawy, and M.F. Mokbel. "LARS: A location-aware recommender system," In Proceedings of IEEE ICDE, pp. 450-461, 2012.

[14] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "A random walk around the city: New venue recommendation in location-based social networks," In Proceedings of IEEE Conference of PASSAT, pp. 144-153, 2012.

[15] 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 AAAI, 2012.

[16] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning Geographical Preferences for Point-of-interest Recommendation," In Proceedings of SIGKDD, pp. 1043-1051, ACM, 2013.

[17] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui, "GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation," In Proceedings of SIGKDD, pp. 831-840, 2014.

[18] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N.M. Thalmann, "Time-aware point-of-interest Recommendation," In Proceedings of SIGIR, pp. 363-372, ACM, 2013.

[19] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring Temporal Effects for Location Recommendation on Location-based Social Networks," In Proceedings of ACM Conference on Recsys, pp. 93-100, 2013.

[20] Q. Yuan, G. Cong, and A. Sun, "Graph-based Joint-of-interest Recommendation with Geographical and Temporal Influences," In Proceedings of CIKM, pp. 659-668, ACM, 2014.

[21] X.T. Li, C. Gao, X.L. Li, T.P. Nguyen, and K. Shonali. "Rank-GeoFM: A Ranking Based Geographical Factorization Method for Point of Interest Recommendation," In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '15, pp. 433-442, ACM, 2015.

[22] N. Usunier, D. Buffoni, and P. Gallinari, "Ranking with Ordered Weighted Pairwise Classification," In Proceedings of ICML, pp. 1057-1064, 2009.

[23] J. Weston, S. Bengio, and N. Usunier, "Large Scale Image Annotation: Learning to Rank with Joint Word-image Embeddings," Machine Learning, vol. 81(1), pp. 21-35, 2010.

[24] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning Geographical Preferences for Point-of-interest Recommendation," In Proceedings of SIGKDD, pp. 1043-1051, ACM, 2013.

978-1-5386-4362-4/18/$31.00 ©2018 IEEE