A Unified Point-of-Interest Recommendation Framework in Location-Based Social Networks

CHEN CHENG, HAIQIN YANG, IRWIN KING, and MICHAEL R. LYU, The Chinese University of Hong Kong

Location-based social networks (LBSNs), such as Gowalla, Facebook, Foursquare, Brightkite, and so on, have attracted millions of users to share their social friendship and their locations via check-ins in the past few years. Plenty of valuable information is accumulated based on the check-in behaviors, which makes it possible to learn users' moving patterns as well as their preferences. In LBSNs, point-of-interest (POI) recommendation is one of the most significant tasks because it can help targeted users explore their surroundings as well as help third-party developers provide personalized services. Matrix factorization is a promising method for this task because it can capture users' preferences to locations and is widely adopted in traditional recommender systems such as movie recommendation. However, the sparsity of the check-in data makes it difficult to capture users' preferences accurately. Geographical influence can help alleviate this problem and have a large impact on the final recommendation result. By studying users' moving patterns, we find that users tend to check in around several centers and different users have different numbers of centers. Based on this, we propose a Multi-center Gaussian Model (MGM) to capture this pattern via modeling the probability of a user's check-in on a location. Moreover, users are usually more interested in the top 20 or even top 10 recommended POIs, which makes personalized ranking important in this task. From previous work, directly optimizing for pairwise ranking like Bayesian Personalized Ranking (BPR) achieves better performance in the top-*k* recommendation than directly using matrix matrix factorization that aims to minimize the point-wise rating error. To consider users' preferences, geographical influence and personalized ranking, we propose a unified POI recommendation framework, which unifies all of them together. Specifically, we first fuse MGM with matrix factorization methods and further with BPR using two different approaches. We conduct experiments on Gowalla and Foursquare datasets, which are two large-scale real-world LBSN datasets publicly available online. The results on both datasets show that our unified POI recommendation framework can produce better performance.

CCS Concepts: • **Information systems → Location based services**; • **Personalization**; • **Human-centered computing → Collaborative filtering;**

Additional Key Words and Phrases: Data mining, recommender systems, location-based social networks, location recommendation

ACM Reference Format:
Chen Cheng, Haiqin Yang, Irwin King, and Michael R. Lyu. 2016. A unified point-of-interest recommendation framework in location-based social networks. ACM Trans. Intell. Syst. Technol. 8, 1, Article 10 (September 2016), 21 pages.
DOI: http://dx.doi.org/10.1145/2901299

The work described in this paper was partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (No. CUHK 413213 and No. CUHK 14205214 of the General Research Fund).
The authors are aslo with Shenzhen Key Laboratory of Rich Media Big Data Analytics and Applications, Shenzhen Research Institute.
Authors' addresses: C. Cheng, I. King, and M. R. Lyu, Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong; emails: {ccheng, king, lyu}@cse.cuhk.edu.hk; H. Yang, Hang Seng Management College, Hong Kong; email: hqyang@ieee.org.

DOI: http://dx.doi.org/10.1145/2901299

1. INTRODUCTION

Recently, with the rapid development of mobile devices and ubiquitous Internet access, location-based social services become prevalent. Online LBSNs such as Gowalla, Foursquare, and so on, have attracted millions of users to share their social friendship, experiences and tips of POIs via check-in behaviors. These information pieces embed abundant hints of users' preferences on locations. The information not only can be utilized to help a specific user explore new places of the city but also can facilitate third-parties such as advertisers to provide specific advertisements for the recommended positions. Hence, POI recommendation becomes a significant task in LBSNs.

To solve the POI recommendation task in LBSNs, matrix factorization is a promising tool because it is a widely adopted method in traditional recommender systems such as movie recommendation [Salakhutdinov and Mnih 2007]. We first construct the user-location matrix, whose entry is the visiting frequency of a user to a location. Then we can obtain the user's preference on locations by performing matrix factorization on the user-location matrix. However, the extreme sparsity of the user-location matrix makes it difficult to capture the user's preference accurately. In our crawled Gowalla dataset, for example, the density of the user-location matrix is only 2.08×10^{-4}.

Fortunately, due to the availability of geographical information (i.e., latitude and longitude) of POIs, researchers can study users' moving patterns and leverage this geographical influence to help improve POI recommendation. In Ye et al. [2011], geographical influence is considered by assuming a power-law distribution between the check-in probability and the distance along the whole check-in history. The parameters of the power-law distribution are learned based on all users' histories; thus, they are not personalized. In this article, we carefully study each user's movement and find that users tend to check in around several centers and different users have different number of centers. We refer to this as multi-center check-in behavior. Based on this finding, we propose a Multi-center Gaussian Model (MGM) to capture this movement pattern. For each user, we will extract the centers based on his or her check-ins. Then for a new location to the user, we define the probability based on the user's centers.

Moreover, in real mobile app recommendation scenarios, users are usually more interested in the top 20 or even top 10 recommended POIs, which makes personalized ranking important in this task. Most of previous work on POI recommendation was mainly based on matrix factorization that minimized the point-wise prediction error for each entry in the user-location matrix. From previous work [Rendle et al. 2009], directly optimizing for pairwise ranking like Bayesian Personalized Ranking (BPR) produces better performance in the top-k recommendation than directly using matrix factorization. To address the top-k ranking as well as the geographical influence, we propose two methods based on BPR, a state-of-the-art personalized ranking method, with different integration approaches.

To our best knowledge, this is the first article to combine the MGM with matrix factorization and BPR into a unified framework in LBSNs, which explores users' preferences, geographical influence and personalized ranking in POI recommendation. Our contributions are threefold. First, we mine a large-scale dataset crawled from Gowalla and extract the characteristics to find out the multi-center check-in behavior. Second, based on the data properties, we model the probability of a user's check-in on a location as an MGM. This is different from the early POI recommendation work in LBSNs [Ye et al. 2011], which assumed a power-law distribution of the check-in probability with respect to the distance within the whole check-in history. Third, we propose a unified POI recommendation framework to fuse users' preferences, geographical influence, and personalized ranking together. Our experimental results on two large-scale real-world online LBSN datasets show that the unified POI recommendation framework presented in this article can achieve significantly better performance than other state-of-the-art methods.

2. RELATED WORK

The work in this paper is closely related to POI recommendation and ranking-oriented collaborative filtering (CF). In the following, we briefly review the related work.

2.1. Point-of-Interest Recommendation

Location-based service (LBS) research became prevalent due to a wide range of potential applications, for example, personalized marketing strategy analysis [Yang et al. 2011a], personalized behavior study [Lu et al. 2011], and POI recommendation [Zheng et al. 2011]. In particular, POI recommendation has attracted much research interest in recent years [Kang et al. 2006; Horozov et al. 2006; Zheng et al. 2009, 2010b; Leung et al. 2011]. In the following, we review several main approaches in collaborative filtering community.

One line of research is to solve POI recommendation based on the extracted stay points from GPS trajectory logs of several hundred monitored users [Zheng et al. 2009, 2010a, 2010b; Leung et al. 2011; Zheng and Xie 2011; Cao et al. 2010]. In Zheng et al. [2010b], three matrices (i.e., location-activity matrix, location-feature matrix and activity-activity matrix) were constructed. Based on the three matrices, a collective matrix factorization method was proposed to mine POIs and activities. Zheng et al. [2010a] explored a tensor factorization on the user-location-activity tensor to provide POI recommendation. In Leung et al. [2011], a memory-based method called the Collaborative Location Model (CLM) was proposed to incorporate activity to facilitate the recommendation.

The other line of work centers on POI recommendation based on the LBSN data [Ye et al. 2010, 2011; Zhang and Chow 2013]. All of these work leverage the geographical influence when providing recommendations and different models were proposed. A pioneer task of POI recommendation in LBSNs debuted in Ye et al. [2010]. The work has been extended and further studied in Ye et al. [2011]. More specifically, geographical influence is considered by assuming a power-law distribution between the check-in probability and the distance along the whole check-in history [Ye et al. 2011]. However, the paper ignored the user's multi-center check-in behavior. Moreover, the proposed method had to compute all pairwise distances of the whole visiting history, which was very time-consuming. Temporal information has also been considered to improve POI recommendation. Zhang and Chow [2013] proposed a kernel density estimation approach, which used kernel function to estimate the distribution for each user. Compared to our MGM model, the method assumed personalized distance distribution. However, when generating recommendations, the complexity for generating the geographical influence value is $O(n^3)$, where n is the number of a user's check-ins. It is not efficient compared to our model whose computational complexity is $O(1)$. Lian et al. [2014] proposed a GeoMF model, which assumed that both users' activity areas and POIs' influence area had effect on the check-in probability. In Gao et al. [2013], temporal non-uniformness and temporal consecutiveness were addressed to model temporal cyclic patterns of check-ins. Geographical and temporal information were incorporated together in Yuan et al. [2013]. Apart from temporal information, content information has been studied as well. Liu and Xiong [2013] employed an aggregated LDA model to study the effect of POI related tags. Gao et al. [2015] investigated three types of content information and modeled them into a unified POI recommendation framework.

2.2. Ranking-Oriented Collaborative Filtering

Top-k recommendation has been studied in collaborative filtering in the past few years. CofiRank [Weimer et al. 2007] was the first proposed ranking-oriented CF approach, which introduced structured ranking loss into the collaborative filtering framework. Bayesian personalized ranking (BPR) [Rendle et al. 2009] was proposed as a state-of-the-art recommendation algorithm for situations with binary relevance data. The optimization criterion of BPR was essentially based on pairwise comparisons between relevant and a sample of irrelevant items. Several methods were explored to optimize directly the ranking metrics. In Shi et al. [2012b], the CF model directly maximized the Mean Reciprocal Rank (MRR) and [Shi et al. 2012a] proposed a model that directly maximized Mean Average Precision (MAP) with the aim of creating an optimally ranked list of items for individual users under a given context. Learning to rank techniques have also been applied in ranking-oriented CF. In Balakrishnan and Chopra [2012], the authors proposed to use user and item-latent vectors as the feature vector in a learning-to-rank framework. Volkovs and Zemel [2012] further proposed an efficient method to extract a good feature vector, which was used by the learning-to-rank framework later with only 17 parameters.

2.3. Ranking-Oriented Method in POI Recommendation

Most recently, Li et al. [2015] developed a ranking-based geographical factorization method called Rank-GeoFM in POI recommendation. In their work, they assumed that the check-in probability is determined by a POI's nearby locations as well as users' preference. And they proposed to use a ranking-based loss to learn the model. The score function is defined as:

$$F_{ul} = {U_u^{(1)}}^T L_l + {U_u^{(2)}}^T \sum_{l^\star \in \mathcal{N}_k(l)} w_{ll^\star} L_{l^\star}, \tag{1}$$

where $U^{(1)}$,$U^{(2)}$ and L are latent matrices. $\mathcal{N}_k(l)$ is the k-nearest neighbors of POI l. $w_{ll^\star} = (0.5 + d(l, l^\star))^{-1}$ is the influence weight of $l^\star$ to l, which is related to their distance $d(l, l^\star)$.

In their method, one iteration takes $O(k \cdot \#observations)$, where k is the neighbor location number and is set to several hundred. Besides, the total training space is $O(|U||L||L_u|)$, where $|L_u|$ is the user's average check-ins. As a result, it takes quite a lot of time to update the parameters to converge (in our Foursquare dataset, it takes around 2 days and it has not converged after 1 week on our Gowalla dataset). Moreover, the geographical influence in Rank-GeoFM is determined by the weight w, which is simply measure by the distance making it is not very precise. When there are few check-ins, both of the user preference score and geographical score in Rank-GeoFM is not accurate. Different from their work, our proposed two ranking methods in this article are more efficient and more effective when users' check-ins is few, which is verified in our experimental part.

In summary, the GPS dataset is usually in small scale with about 100 or 200 users, but the data are very dense. Contrarily, the LBSN's dataset is in large scale with thousands of users, but the data are very sparse [Noulas et al. 2011; Scellato et al. 2011]. To solve large-scale recommendation problems, matrix factorization is a promising tool due to its success in Netflix competition [Bell et al. 2007; Koren 2009]. However, the data sparsity of LBSN data makes the results of matrix factorization inaccurate. Moreover, traditional matrix factorization approaches do not consider the geographical influence, which has a great effect on POI recommendation. Besides, the final purpose of POI applications is to recommend a few top locations, where the ranking performance is important in this task, while previous work does not emphasize personalized

Table I. Basic Statistics of the Gowalla and Foursquare Dataset

(a) Gowalla

$\#U$	$\#L$	$\#E$
53,944	367,149	306,958
$\#\widetilde{U}$	$\#\widetilde{L}$	$\#\widetilde{E}$
51.33	7.54	11.38
#max. U	#max. L	#max. E
2,145	3,581	2,366

(b) Foursquare

$\#U$	$\#L$
6,084	37,976
$\#\widetilde{U}$	$\#\widetilde{L}$
35.98	5.76
#max. U	#max. L
182	985

	l_1	l_2	l_3	l_4	l_5	l_6	$\cdots$	$l_{\lvert\mathcal{L}\rvert-1}$	$l_{\lvert\mathcal{L}\rvert}$
u_1	?	?	164	?	1	?	$\cdots$	?	1
u_2	40	2	?	?	?	1	$\cdots$	?	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$u_{\lvert\mathcal{U}\rvert-1}$	?	?	1	1	?	?	$\cdots$	2	?
$u_{\lvert\mathcal{U}\rvert}$	?	2	?	?	1	?	$\cdots$	?	10

Fig. 1. User-location check-in frequency matrix.

ranking in POI recommendation. In this article, we propose a unified POI recommendation framework that incorporates user preference, geographical influence, as well as personalized ranking together.

3. CHECK-IN DATA CHARACTERISTICS

In this article, we conduct experiments on two publicly available online LBSN datasets: Gowalla[1] and Foursquare.[2] Gowalla is an LBSN website created in 2009 for users to check in to various locations through mobile devices. We collect a complete snapshot, including users' profile, users' check-in locations, check-in timestamps, users' friend lists, and location details, from Gowalla during the period from February 2009 to September 2011 via the provided public API. To reduce noise in data, we remove users with less than 10 check-ins and locations with less than 20 visits. Foursquare is another LBSN website similar to Gowalla. We use the 4-month Foursquare dataset, which spans from May 2010 to August 2010 provided by Cheng et al. [2011]. Similarly, in order to remove noise, we require that all users should have at least 10 check-ins. But we do not have the social information in the provided Foursquare dataset. The basic statistics of the datasets are summarized in Table I. In the table, we use a tilde to denote the average count.

Details of the data are depicted in the following:

—The Gowalla dataset has 4,128,714 check-ins from 53,944 users on 367,149 locations and a total of 306,958 edges are in the whole users' social graph. The density of the user-location matrix in the Gowalla dataset is about 2.08×10^{-4}. Figure 1 is an illustration of the user-location matrix. On the other hand, the Foursquare dataset consists of $6,084$ users, $37,976$ locations with $218,935$ check-ins. The density of the Foursquare dataset is about 9.48×10^{-4}.

—The average number of visited locations of a user is 51.33 and 35.98 for the Gowalla and Foursquare dataset, respectively. The average number of visiting users for a location is 7.54 in the Gowalla dataset and 5.76 in the Foursquare dataset. The average number of friends of a user is 11.38 in the Gowalla dataset.

[1]http://www.gowalla.com.
[2]http://www.foursquare.com.

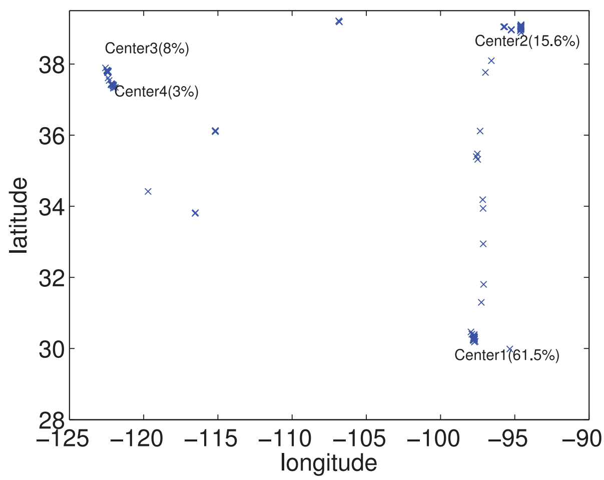
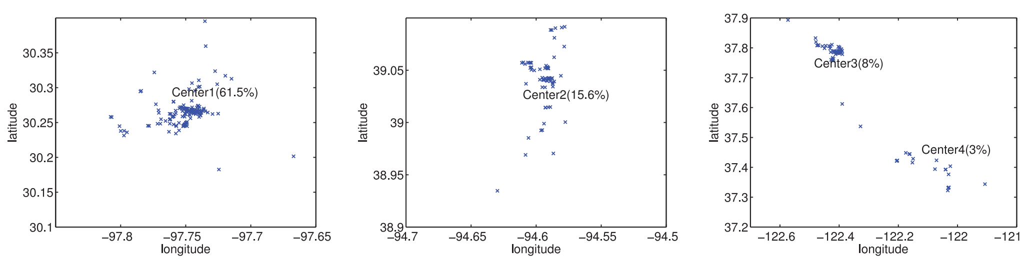


Fig. 2. Multi-center overview.



(a) Center 1 (b) Center 2 (c) Center 3 and 4

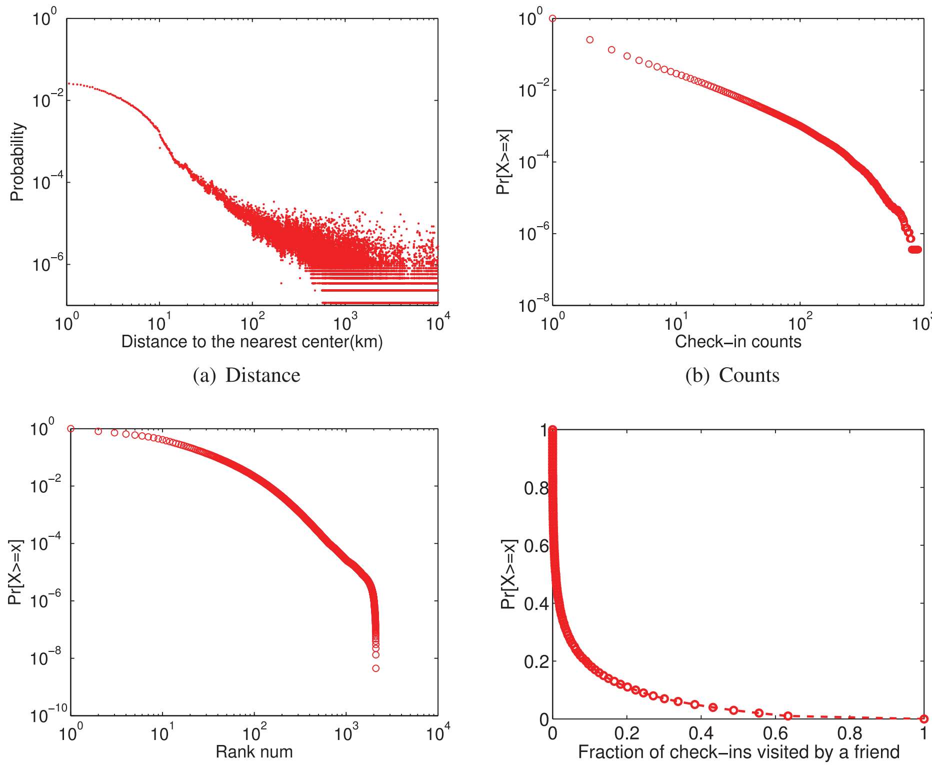
Fig. 3. A typical user's multi-center check-in behavior.

—In the Gowalla dataset, the maximum number of locations for a user is 2,145; in the Foursquare dataset, the maximum number is 182. The maximum number of visiting users for a location is 3,581 for Gowalla and 985 for Foursquare. The maximum number of friends of a user is 2,366.

In the following, we further study the location distribution, frequency distribution and the social relationship among users' check-ins. Because Gowalla and Foursquare share similar characteristics, we only show the results from Gowalla.

3.1. Location Distribution

Figure 2 shows the longitude and latitude of a typical user's check-in locations, where the locations form four centers. The details of each center are further shown in Figures 3(a) through 3(c). This observation reaches our assumption different from the power-law distribution on users' check-in histories in Ye et al. [2011]. In addition, our statistics are also a little differently from the two states ("home" and "office") check-in behavior mentioned in Cho et al. [2011]. After examining the comments of locations, we find that other than the centers of "home" and "office" (counting above half of a user's check-ins), other centers count at least 10% of the check-ins. These centers may be a user's usual business travel places, for example, an office of a branch of a large company or vocation places, which provide abundant information that needs to be differentiated. This means for each user, there may exist several centers around which the user would



(a) Distance (b) Counts

(c) Top-k locations (d) Friends

Fig. 4. Check-ins probability vs. distance, counts, top-k locations, common check-ins of friends.

like to conduct activities. Note that the POIs near these centers have a higher chance to be checked in than the POIs that are far away. It reflects the fact that most of the time human beings hang out around several familiar areas.

3.2. Frequency Distribution

Figure 4(b) plots the Complementary Cumulative Distribution Function (CCDF) for each user's check-in numbers at each location. It is shown that about 74% of locations are only visited once and only about 3% of locations are visited more than 10 times. This means that users usually visit several important places (e.g., home, office and some stores) with very high frequency, while most other places are seldom visited. Overall, these places are around several centers. Figure 4(c) further shows the CCDF function of top-k frequently visited locations. The most visited location accounts for about 18.8% of all users' check-ins. The top 10 most visited locations account for 68% of all check-ins and the ratio increases to 80.5% for the top-20 most visited locations, followingthe *Pareto principle* (a.k.a. 80-20 rule) [Hafner 2001].

3.3. Social Influence

In the dataset, we find that the average overlap of a user's check-ins to his or her friends' check-ins is only 9.6%. This indicates that less than 10% of a user's check-ins are also visited by the user's friends, which is similar to the statistics reported in Cho et al. [2011]. Figure 4(d) plots the CCDF of the fraction of a user's check-ins that are visited by his or her friends. It is known that for about 38% of users, their check-in locations are not checked in by their friends, while almost 90% of users contain less than 20% of common check-ins with their friends. The statistics are a little different from that in Cho et al. [2011], but the overall trend is similar. These observations imply that social relationship has a limited effect on users' check-ins, which also illustrated in the experimental part.

4. UNIFIED POINT-OF-INTEREST RECOMMENDATION FRAMEWORK

The problem of personalized POI recommendation is defined as follows: given a partially observed user-location check-in frequency matrix (users in $\mathcal{U}$ and locations in $\mathcal{L}$), the task is to recommend top-k locations to a user that the user does not visit before. To solve this problem, we first propose a personalized MGM to capture the geographical influence on a user's check-ins. Then we depict the matrix factorization method and propose a fused MF framework to include geographical influence. Finally, we introduce the unified framework, which incorporates geographical influence and matrix factorization to directly optimize the ranking loss for POI recommendation.

4.1. Multi-Center Gaussian Model (MGM)

A significant characteristic of check-in locations is that they are usually located around several centers as shown in Figure 3. The second characteristic of check-in locations is that the probability of a user visiting a location is inversely proportional to the distance from its nearest center (see Figure 4(a)).

These two characteristics indicate that geographical information plays a strong influence on the user's check-in behavior. Based on the statistics from Figures 3 and 4(a), we adopt Gaussian distribution to model the user's check-in behavior and propose the MGM. That is, the probability of a user u, visiting a POI l, given the multi-center set C_u, is defined by:

$$P(l|C_u) = \sum_{c_u=1}^{|C_u|} P(l \in c_u) \frac{f_{c_u}^{\alpha}}{\sum_{i \in C_u} f_i^{\alpha}} \frac{\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})}{\sum_{i \in C_u} \mathcal{N}(l|\mu_i, \Sigma_i)}. \quad (2)$$

Here, l denotes the longitude and latitude of a position, and C_u is the set of centers for the user u. For each center, calculating Equation (2) consists of the multiplication of three terms:

—$P(l \in c_u) \propto 1/dist(l, c_u)$ determines the probability of the location l, which belongs to the center c_u, which is inversely proportional to the distance between the location l and the center c_u.
—The second term denotes the normalized effect of check-in frequency f_{c_u}, on the center c_u. The parameter $\alpha \in (0, 1]$ is introduced to maintain the frequency aversion property, where very high check-in frequency does not play too significant effect.
—The third term denotes the normalized probability of a location belonging to the center c_u, where $\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})$ is the probability density function of the Gaussian distribution, and μ_{c_u} and Σ_{c_u} correspond to the mean and covariance matrices of regions around the center c_u.

Next we introduce how to find the centers for each individual user. We propose a greedy clustering algorithm among the check-ins due to the Pareto principle [Hafner 2001], which is very efficient. The computational complexity is linear to the number of observations in the user-location matrix. This property can be observed from Figures 3 and 4(c). There are several more advanced techniques to calculate data similarity, which can be referred to Yang et al. [2011b].

In our proposed MGM model, we are able to capture the personalized geographical influence for each user. Compared to iGSLR method in Zhang and Chow [2013], our

ALGORITHM 1: Multi-Center Discovering Algorithm

```
for all user $i$ in the user set $\mathcal{U}$ do
    Rank all check-in locations in $|\mathcal{L}|$ according to visiting frequency
    $\forall l_k \in L$, set $l_k.center = -1$;
    Center_list = $\emptyset$; center_no = 0;
    for $i = 1 \to |L|$ do
        if $l_i.center == -1$ then
            center_no++; Center = $\emptyset$; Center.total_freq = 0;
            Center.add($l_i$); Center.total_freq += $l_i$.freq;
            for $j = i + 1 \to |L|$ do
                if $l_j.center == -1$ and $dist(l_i, l_j) \leq d$ then
                    $l_j.center$ = center_no; Center.add($l_j$);
                    Center.total_freq += $l_j$.freq;
                end if
            end for
            if Center.total_freq $\geq |u_i|$.total_freq * $\theta$ then
                Center_list.add(Center);
            end if
        end if
    end for
    RETURN Center_list for user $i$;
end for
```

method is much more efficient when providing recommendations. The time complexity of our method is $O(1)$ for computing one new POI, because we can store the center information beforehand. While in iGSLR, the time complexity is $O(n^3)$, where n is a user's check-in number. Because we try to recommend top-*k* POIs, the total time complexity for recommend top-*k* POIs in iGSLR is $O(|\mathcal{L}|n^3$. When n is larger than 100, it costs more than an hour. As a result, we do not compare it with our MGM.

In our greedy algorithm, we first scan from the most visited POIs and combine all other visited check-in locations, whose distance is less than d kilometers from the selected POI, into a region. If the ratio of the total check-in number of this region to the user's total check-in amount is greater than a threshold θ, we set these check-in positions as a region and determine its center. Algorithm 1 shows the procedure of discovering multiple centers. In our experiments, by trial on the training dataset, we set θ to 0.02, the the distance threshold d to 15 and the frequency control parameter α to 0.2.

4.2. Matrix Factorization

Matrix Factorization (MF) is one of the most popular methods for recommender systems [Salakhutdinov and Mnih 2007, 2008; Bell et al. 2007; Koren 2009]. It has been shown to be particularly effective in recommender systems as well as in the well-known Netflix prize competitions[3] [Bell and Koren 2007]. Given the partially observed entries in a $|\mathcal{U}| \times |\mathcal{L}|$ frequency matrix F, the goal of MF is to find two low-rank matrices $U \in \mathbb{R}^{K\times|\mathcal{U}|}$ and $L \in \mathbb{R}^{K\times|\mathcal{L}|}$ such that $F \approx U^T L$. The predicted probability of a user u who is likely to visit a location l is determined by

$$P(F_{ul}) \propto U_u^T L_l. \tag{3}$$

4.2.1. Probabilistic Matrix Factorization (PMF). PMF is one of the most famous MF models in collaborative filtering, which is proposed in Salakhutdinov and Mnih [2007]. It

[3]http://www.netflixprize.com.

assumes that the conditional distribution over the observed rating is:

$$p(F|U, L, \sigma_R^2) = \prod_{i=1}^{|\mathcal{U}|} \prod_{j=1}^{|\mathcal{L}|} \left[\mathcal{N}\left(F_{ij}|U_i^T L_j, \sigma_R^2\right)\right]^{I_{ij}^R}, \tag{4}$$

where $\mathcal{N}(x|\mu, \sigma^2)$ is the probability density function of the Gaussian distribution with mean μ and variance σ^2. I_{ij}^R is the indicator function that equals to 1 if user u_i has visited location l_j and equals to 0 otherwise. The zero-mean spherical Gaussian priors are also placed on user and location latent feature vectors:

$$p(U|\sigma_U^2) = \prod_{i=1}^{|\mathcal{U}|} \mathcal{N}\left(U_i|0, \sigma_U^2 \boldsymbol{I}\right),\ p(L|\sigma_V^2) = \prod_{j=1}^{|\mathcal{L}|} \mathcal{N}\left(L_j|0, \sigma_V^2 \boldsymbol{I}\right). \tag{5}$$

Through Bayesian inference, we have the following objective function:

$$\min_{U,L} \frac{1}{2} \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} I_{ij}^R \left(F_{ij} - U_i^T L_j\right)^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|L\|_F^2, \tag{6}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. In practice, we can use the sigmoid function $g(x) = 1/(1 + \exp(-x))$ to convert the rating into $(0, 1)$. Now the objective functions becomes:

$$\min_{U,L} \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} I_{ij} \left(g(F_{ij}) - g(U_i^T L_j)\right)^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|L\|_F^2. \tag{7}$$

Note: The observed frequency data are all positive, which makes the data biased. Consequently, it is a standard one-class collaborative filtering problem [Pan et al. 2008; Pan and Scholz 2009; Hu et al. 2008]. We sample the same number of unobserved data from the rest matrix and deem their frequency as 0.

4.2.2. PMF with Social Regularization (PMFSR). Social information is available in our Gowalla dataset. The social information has been shown to be useful in recommender systems [Ma et al. 2008; Zhou et al. 2009; Ma et al. 2011a, 2011b]. Although we illustrate in Section 3.3 that social information has limited influence, here we adopt one of the existing method to verify it.

We adopt the PMF with Social Regularization (PMFSR) [Ma et al. 2011a], where the Individual-based Regularization Model proposed to impose constraints between one user and his or her friends individually. The objective function is defined as follows:

$$\begin{aligned} \min_{U,L} \Omega(U, L) = & \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} I_{ij} \left(g(F_{ij}) - g\left(U_i^T L_j\right)\right)^2 \\ & + \beta \sum_{i=1}^{|\mathcal{U}|} \sum_{f \in \mathcal{F}(i)} Sim(i, f) \|U_i - U_f\|_F^2 \\ & + \lambda_1 \|U\|_F^2 + \lambda_2 \|L\|_F^2, \end{aligned} \tag{8}$$

where $\mathcal{F}(i)$ is the set of friends for user u_i, and $Sim(i, f)$ is the similarity between user u_i and his or her friend u_f. The similarity between a user and the user's friends can be computed by measuring the check-ins of them.

4.2.3. Probabilistic Factor Models (PFM). The PMF model makes assumption on the Gaussian distribution, which may not be appropriate when applied to the frequency data. This is demonstrated in our later experiment results. Because the check-in data in LBSNs are naturally frequency, we turn to Probabilistic Factor Models (PFM) [Chen et al. 2009; Ma et al. 2011], which can model the frequency data directly.

PFM places Gamma distributions as priors on the latent matrices U and L, while it defines a Poisson distribution on the frequency. This leads to seek U and L by minimizing $\Psi(U, L; F)$:

$$\begin{aligned}\Psi(\cdot, \cdot; \cdot) \; &= \sum_{i=1}^{|\mathcal{U}|} \sum_{k=1}^{K} ((\alpha_k - 1) \ln(U_{ik}/\beta_k) - U_{ik}/\beta_k) \\ &+ \sum_{j=1}^{|\mathcal{L}|} \sum_{k=1}^{K} ((\alpha_k - 1) \ln(L_{jk}/\beta_k) - L_{jk}/\beta_k) \\ &+ \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} (F_{ij} \ln(U^T L)_{ij} - (U^T L)_{ij}) + c, \end{aligned} \tag{9}$$

where c is a constant term.

4.3. A Fusion Framework with User Preference and Geographical Influence

We can observe that either PMF, PMFSR, or PFM only models users' preferences on locations. They do not explore the geographical influence. As observed from Figure 4(a), users tend to check in locations around their centers. It can be very helpful for POI recommendation, especially when we have very few check-ins, where matrix factorization does not perform very well. Hence, we fuse users' preferences on a POI and the probability from MGM together to determine the probability of a user u visiting a location l, which is defined as follows:

$$P_{ul} = P(F_{ul}) \cdot P(l|C_u), \tag{10}$$

where $P(l|C_u)$ is calculated by Equation (2) via MGM, and $P(F_{ul})$ encodes users' preferences on a location determined by Equation (3). After we get the final predicted value P_{ul}, we can obtain a ranked list of recommended POIs for user u. Finally, we recommend the top k locations to the user.

4.4. A Final Fusion Framework

Because our final goal is to recommend a ranking POI list to users, directly optimizing the ranking loss is desirable. Bayesian Personalized Ranking (BPR) [Rendle et al. 2009] is a state-of-the-art method that tries to minimize the pairwise ranking loss over user rated items and unrated items. On the other hand, geographical influence has a great effect on POI recommendation; therefore, we propose two methods to incorporate MGM with BPR, which combine pairwise ranking with geographical effect together. In the following, we describe the BPR model first, then we detail the two combined location ranking methods.

4.4.1. Bayesian Personalized Ranking (BPR). In LBSNs, all the check-ins are implicit feedback data, which means we only observe the positive data. The unobserved data, that is, the missing user-location pairs, are a mixture of real negative feedback (the user is not interested in visiting the location) and missing values (the user might want to check in the location but has not visited there).

In BPR, the task is to derive a personalized ranking $>_u$ over locations for each user u. The basic assumption is that if user u checks in location i while not checking in location

j, we say the user prefers location i over location j, denoted as $i >_u j$. We assume that there is an estimator $\hat{f} : U \times L \to \mathbb{R}$, which is used to define the ranking:

$$i >_u j \Leftrightarrow \hat{f}_{ui} > \hat{f}_{uj}. \tag{11}$$

The estimator $\hat{f}$ is usually calculated through matrix factorization:

$$\hat{f}_{ui} = U_u^T L_i. \tag{12}$$

The Bayesian formulation of finding the correct personalized ranking for all locations in $\mathcal{L}$ is to maximize the following posterior probability:

$$p(\Theta| >_u) \propto p(>_u |\Theta) p(\Theta), \tag{13}$$

where Θ represents the parameters.

We further assume that all users are independent and the ordering of each location pairs (i, j) for a specific user is also independent. Thus, the likelihood function for all users can be defined as:

$$\prod_{u \in \mathcal{U}} p(>_u |\Theta) = \prod_{(u,i,j) \in S} p(i >_u j|\Theta), \tag{14}$$

where $S = \{(u, i, j)|u \in \mathcal{U}, i \in \mathcal{L}_u^+ \wedge j \in \mathcal{L} \setminus \mathcal{L}_u^+\}$, and $\mathcal{L}_u^+$ is the set of locations visited by user u.

The individual probability of user u preferring location i to location j is defined as:

$$p(i >_u j|\Theta) = \sigma(\hat{f}_{uij}(\Theta)), \tag{15}$$

where σ is the logistic sigmoid function $\sigma(x) = 1/(1 + \exp(-x))$, and

$$\hat{f}_{uij} = \hat{f}_{ui} - \hat{f}_{uj}. \tag{16}$$

We further place a Gaussian prior over the parameters:

$$p(\Theta) \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}). \tag{17}$$

We use maximum a posterior (MAP) to estimate the parameters:

$$\arg\max_{\Theta} \ln \prod_{u \in \mathcal{U}} p(>_u |\Theta) p(\Theta). \tag{18}$$

Substituting Equations (15) and (16) into Equation (18), we have the final objective function:

$$\arg\max_{\Theta} \sum_{(u,i,j) \in S} \ln(\sigma(\hat{f}_{ui} - \hat{f}_{uj})) - \lambda_\Theta \|\Theta\|^2. \tag{19}$$

Stochastic gradient descent (SGD) can be applied to learn the model parameters Θ. We denote $\mathcal{F}$ as the objective function in Equation (19). The gradient of $\mathcal{F}$ with respect to the model parameters is:

$$\frac{\partial \mathcal{F}}{\partial \Theta} = \sum_{(u,i,j) \in S} \frac{\partial}{\partial \Theta} \ln(\hat{f}_{ui} - \hat{f}_{uj}) - \lambda_\Theta \frac{\partial}{\partial \Theta} \|\Theta\|^2 \tag{20}$$

$$\propto \sum_{(u,i,j) \in S} (1 - \sigma(\hat{f}_{uij})) \cdot \frac{\partial}{\partial \Theta} (\hat{f}_{ui} - \hat{f}_{uj}) - \lambda_\Theta \Theta \tag{21}$$

$$= \sum_{(u,i,j) \in S} (1 - \sigma(\hat{f}_{uij})) \cdot \frac{\partial}{\partial \Theta} \left(U_u^T L_i - U_u^T L_j\right) - \lambda_\Theta \Theta. \tag{22}$$

Here, $\Theta = \{U, L\}$. Note that

$$\frac{\partial}{\partial U_u}\left(U_u^T L_i - U_u^T L_j\right) = L_i - L_j, \tag{23}$$

$$\frac{\partial}{\partial L_i}\left(U_u^T L_i - U_u^T L_j\right) = U_u, \tag{24}$$

$$\frac{\partial}{\partial L_j}\left(U_u^T L_i - U_u^T L_j\right) = -U_u. \tag{25}$$

For each triple (u, i, j) we draw from S, the update rule is:

$$\Theta \leftarrow \Theta + \alpha \left((1 - \sigma(\hat{f}_{uij})) \cdot \frac{\partial}{\partial \Theta}(\hat{f}_{ui} - \hat{f}_{uj}) - \lambda_\Theta \Theta \right), \tag{26}$$

where α is the step size.

4.4.2. Ranking in POI Recommendation. We propose two methods to incorporate BPR with geographical influence. The first method is the same as the fuse framework in Section 4.3. The final probability that user u visits a location l is consequently defined as

$$P_{ul} = \hat{f}_{ul} \cdot P(l|C_u), \tag{27}$$

where $\hat{f}_{ul}$ is estimated from BPR. We refer to this method as BPR Location Recommendation 1 (BPRLR1).

In the second method, we borrow the idea from Cheng et al. [2013]. Instead of maximizing the difference between visited locations and all unvisited locations, we focus on maximizing the difference between visited locations and unvisited locations that are near users' centers. This idea is very intuitive—because users tend to check in locations near their activity centers, we do not consider the far away locations, which may introduce noise otherwise.

We denote N_u as the set of locations in the nearby activity area for user u. We define $N_u = \{l | P(l|C_u) > 0\}$, which requires that location l has a chance to be checked in by the MGM model. Then we define the trained pairwise location set $S' = \{(u, i, j) | u \in \mathcal{U}, i \in \mathcal{L}_u^+ \wedge j \in N_u \setminus \mathcal{L}_u^+\}$. Now the objective function is:

$$\arg\max_{\Theta} \sum_{(u,i,j)\in S'} \ln(\sigma(\hat{x}_{ui} - \hat{x}_{uj})) - \lambda_\Theta \|\Theta\|^2. \tag{28}$$

After we get the learned parameters, we employ the estimator $\hat{x}_{ui}$ to obtain the ranking list. We refer to this method as BPR Location Recommendation 2 (BPRLR2). The learning algorithm is shown in Algorithm 2.

ALGORITHM 2: Learning Algorithm for BPRLR2

draw U,L from $\mathcal{N}(0, \sigma^2)$
repeat
  draw (u, i, j) uniformly from S'
  Calculate $\sigma(\hat{f}_{uij})$
  Update U_u, L_i, L_j according to:
  $U_u = U_u + \alpha\left((1 - \sigma(\hat{f}_{uij}) \cdot (L_i - L_j) - \lambda_\Theta U_u\right)$
  $L_i = L_i + \alpha\left((1 - \sigma(\hat{f}_{uij}) \cdot (U_u) - \lambda_\Theta L_i\right)$
  $L_j = L_j + \alpha\left((1 - \sigma(\hat{f}_{uij}) \cdot (-U_u) - \lambda_\Theta L_j\right)$
until convergence
return U,L

4.5. Complexity Analysis

The computation cost consists of the calculation of matrix factorization models and calculating the probability of a user visiting a POI. The training time for the matrix factorization models scales linearly with respect to the number of observations [Salakhutdinov and Mnih 2007; Ma et al. 2011a]. For the probability computation, the cost is to calculate the centers. This also scales linearly with respect to the number of observations. Hence, the proposed fused framework in Section 4.3 is linear with respect to the number of observations. We use SGD to learn parameters in BPRLR1 and BPRLR2. In each iteration, we update the parameters U_u, L_i, and L_j. The cost of the iteration is $O(K)$, where K is the latent dimension and is usually very small. In practice, the convergence iteration number is a few times of the observations. So both BPRLR1 and BPRLR2 are efficient and can scale up to very-large-scale datasets.

5. EXPERIMENTS

The experiments address the following three questions:

(1) How do our approaches compare with the baseline and the state-of-the-art algorithms?
(2) How do the geographical influence and ranking loss affect the performance?
(3) What is the performance on users with different check-in frequency? This is a scenario for cold-start users whose check-ins are few.

5.1. Setup and Metrics

The experimental data include user-location check-in records, users' friendship list, and geographical information (longitude and latitude of check-in locations). We split the crawled Gowalla dataset and Foursquare dataset into two non-overlapping sets: a training set and a test set, where the proportion of training data and test data is 70% and 80%, respectively. Here, 70%, for example, means we randomly select 70% of the observed data for each user as the training data to predict the remaining 30% data. The random selection was carried out five times independently, and we report the average result. The hyper-parameters are tuned by cross-validation. For all experiments, we set the regularization term λ to 0.1 and the step size α to 0.2.

POI recommendation is to recommend the top-N highest ranked positions to a targeted user based on a ranking score from a recommendation algorithm. To evaluate the model performance, we are interested in finding out how many locations in the test set are recovered in the returned POI recommendation. Hence, we use the Precision@N and Recall@N as the metrics to evaluate the returned ranking list against the check-in locations where users actually visit. These two metrics are standard metrics to measure the performance of POI recommendation [Ye et al. 2011]. Precision@N defines the ratio of recovered POIs to the N recommended POIs, while Recall@N defines the ratio of recovered POIs to the size of the test set. In the experiments, N is set to 5 and 10, respectively.

5.2. Comparison

In the experiments, the compared approaches include:

(1) **Multi-center Gaussian Model (MGM)**: this method recommends a position based on the probability calculated by Equation (2).
(2) **PMF**: this is a well-known method in matrix factorization [Salakhutdinov and Mnih 2007]. We describe the details in Section 4.2.1. Its objective function is shown in Equation (7).

Table II. Performance Comparisons on the Gowalla Dataset with $K = 20$

Ratio	Metrics	Dimension = 20								
		MGM	PMF	PMFSR	PFM	FMFMGM	GeoMF	BPR	BPRLR1	BPRLR2
70%	P@5	0.0317	0.0140	0.0153	0.0173	0.0643	0.0660	0.0645	**0.0791**	0.0500
	Improve	149.53%	465.00%	416.99%	357.23%	23.02%	19.85%	22.64%		58.20%
	R@5	0.0113	0.0032	0.0035	0.0040	0.0202	0.0212	0.0187	**0.0264**	0.0168
	Improve	133.63%	725.00%	654.29%	560.00%	30.69%	24.53%	41.18%		57.14%
	P@10	0.0273	0.0166	0.0166	0.0172	0.0635	0.0638	0.0615	**0.0682**	0.0615
	Improve	149.82%	310.84%	310.84%	296.51%	7.40%	6.90%	10.89%		10.89%
	R@10	0.0194	0.0079	0.0078	0.0084	0.0395	0.0402	0.0355	**0.0445**	0.0396
	Improve	129.38%	463.29%	470.51%	429.76%	12.66%	10.70%	25.35%		12.37%
80%	P@5	0.0263	0.0106	0.0107	0.0114	0.0464	0.0476	0.0462	**0.0544**	0.0334
	Improve	106.84%	413.21%	408.41%	377.19%	17.24%	14.29%	17.75%		62.87%
	R@5	0.0141	0.0034	0.0034	0.0039	0.0207	0.0216	0.0194	**0.0258**	0.0160
	Improve	82.98%	658.82%	658.82%	561.54%	24.64%	19.44%	32.99%		61.25%
	P@10	0.0226	0.0120	0.0121	0.0117	0.0452	0.0458	0.0427	**0.0468**	0.0412
	Improve	107.08%	290.00%	286.78%	300.00%	3.54%	2.18%	9.60%		13.59%
	R@10	0.0244	0.0082	0.0084	0.0083	0.0404	0.0409	0.0358	**0.0442**	0.0382
	Improve	81.15%	439.02%	426.19%	432.53%	9.41%	8.07%	23.46%		15.71%

(3) **PMF with Social Regularization (PMFSR)**: this method is proposed to include the social friendship under the PMF framework [Ma et al. 2011a], which is introduced in Section 4.2.2. Its objective function is shown in Equation (8). We only report the results on Gowalla data, as we do not have the social information in our Foursquare data.
(4) **Probabilistic Factor Models (PFM)**: this method is a promising method to model frequency data [Ma et al. 2011]. Its objective function is shown in Equation (9), and the details are in Section 4.2.3.
(5) **FMF with MGM (FMFMGM)**: this is the Fused Matrix Factorization framework with the Multi-center Gaussian Model (FMFMGM). The user's preference on locations is calculated by the PFM model. Here, we select PFM because PFM can model the frequency data better than PMF.
(6) **BPR**: this method is a ranking-oriented method for implicit data [Rendle et al. 2009]. We introduced the details in Section 4.4.1.
(7) **GeoMF**: this method is proposed in Lian et al. [2014]. which incorporated location information into the weighted matrix factorization method. We report the results of GeoMF on our large Gowalla dataset.
(8) **Rank-GeoFM**: this method is the state-of-the-art method proposed in Li et al. [2015]. We briefly introduced the details in Section 2.3. We only report the results on Foursquare, as our Gowalla data is too large. It take too long time to converge in Gowalla data.
(9) **BRPLR1**: this method is the first scheme we proposed to incorporate BPR and geographical influence.
(10) **BPRLR2**: this method is the second scheme we proposed to incorporate BPR and geographical influence.

Tables II through V report the average of five-run results on the top 5 and top 10 recommendation by the competing models using 20 and 30 as the number of latent feature dimensions, respectively. The results show that:

—FMFMGM outperforms PMF, PMFSR, and PFM significantly in all metrics. For example, in Gowalla, FMFMGM attains 0.0643 in terms of P@5 when the latent

Table III. Performance Comparisons on the Gowalla Dataset with $K = 30$

Ratio	Metrics	Dimension = 30								
		MGM	PMF	PMFSR	PFM	FMFMGM	GeoMF	BPR	BPRLR1	BPRLR2
70%	P@5	0.0317	0.0148	0.0158	0.0173	0.0672	0.0683	0.0674	**0.0802**	0.0517
	Improve	153.00%	441.89%	407.59%	363.58%	19.35%	17.42%	18.99%		55.13%
	R@5	0.0113	0.0033	0.0035	0.0040	0.0212	0.0219	0.0199	**0.0270**	0.0175
	Improve	138.94%	718.18%	671.43%	575.00%	27.36%	23.29%	35.68%		54.29%
	P@10	0.0273	0.0162	0.0174	0.0173	0.0656	0.0660	0.0643	**0.0700**	0.0628
	Improve	156.41%	332.10%	302.30%	304.62%	6.71%	5.11%	8.86%		11.46%
	R@10	0.0194	0.0075	0.0080	0.0084	0.0408	0.0415	0.0382	**0.0465**	0.0408
	Improve	260.82%	833.33%	775.00%	733.33%	13.97%	12.05%	83.25%		13.97%
80%	P@5	0.0263	0.0106	0.011	0.0114	0.0486	0.0490	0.0488	**0.0551**	0.0348
	Improve	109.51%	419.81%	400.91%	383.33%	13.37%	12.45%	12.91%		58.33%
	R@5	0.0141	0.0035	0.0037	0.0039	0.0218	0.0222	0.0210	**0.0263**	0.0172
	Improve	86.52%	651.43%	610.81%	574.36%	20.64%	18.47%	25.24%		52.91%
	P@10	0.0226	0.0115	0.0117	0.0117	0.0472	0.0473	0.0450	**0.0479**	0.0432
	Improve	111.95%	316.52%	309.40%	309.40%	1.48%	1.27%	6.44%		10.88%
	R@10	0.0244	0.0079	0.0081	0.0085	0.0424	0.0430	0.0386	**0.0456**	0.0407
	Improve	86.89%	477.22%	462.96%	436.47%	7.55%	6.05%	18.13%		12.04%

Table IV. Performance Comparisons on the Foursquare Dataset with $K = 20$

Ratio	Metrics	Dimension = 20							
		MGM	PMF	PFM	FMFMGM	BPR	BPRLR1	Rank-GeoFM	BPRLR2
70%	P@5	0.0409	0.0591	0.0706	0.1190	0.1074	0.1447	0.1548	**0.1734**
	Improve	323.96%	193.40%	145.61%	45.71%	61.45%	19.83%	12.01%	
	R@5	0.0306	0.0258	0.0308	0.0588	0.0513	0.0749	0.0718	**0.0878**
	Improve	186.93%	240.31%	185.06%	49.32%	71.15%	17.22%	22.28%	
	P@10	0.0373	0.0610	0.0652	0.1157	0.1078	0.1501	0.1565	**0.1671**
	Improve	347.99%	173.93%	156.29%	44.43%	55.01%	11.33%	6.77%	
	R@10	0.0531	0.0550	0.0608	0.1152	0.1032	0.1545	0.1452	**0.1699**
	Improve	219.96%	208.91%	179.44%	47.48%	64.63%	9.97%	17.01%	
80%	P@5	0.0288	0.0448	0.0486	0.0830	0.0771	0.1031	0.1161	**0.1273**
	Improve	342.01%	184.15%	161.93%	53.37%	65.11%	23.47%	9.65%	
	R@5	0.0332	0.0311	0.0362	0.0645	0.0572	0.0826	0.0830	**0.0969**
	Improve	191.87%	211.58%	167.68%	50.23%	69.41%	17.31%	16.75%	
	P@10	0.0265	0.0466	0.0504	0.0812	0.0766	0.1042	0.1177	**0.1207**
	Improve	355.47%	159.01%	139.48%	48.65%	57.57%	15.83%	2.55%	
	R@10	0.0586	0.0647	0.0671	0.1245	0.1138	0.1648	0.1673	**0.1859**
	Improve	217.24%	187.33%	177.05%	49.32%	63.36%	12.80%	11.11%	

dimension is 20 and 70% of data are used for training, while PFM, the best current model without considering location information, achieves 0.0173 for the counterpart. This implies that geographical influence plays a significant role in POI recommendation. By utilizing the geographical influence, we can provide much more accurate POI recommendations to targeted users.

—FMFMGM achieves significantly better performance than MGM in both Gowalla and Foursquare datasets. That is, for the case of the latent dimension being 30 and 80% of data for training, the performance increases from 0.0141 for MGM to 0.0218 for FMFMGM. This verifies that the probability of a user visiting a POI is controlled by both the user's personal preference and the personal check-in location constraints.

Table V. Performance Comparisons on the Foursquare Dataset with $K = 30$

Ratio	Metrics	Dimension = 30							
		MGM	PMF	PFM	FMFMGM	BPR	BPRLR1	Rank-GeoFM	BPRLR2
70%	P@5	0.0409	0.0621	0.0718	0.1201	0.1086	0.1484	0.1694	**0.1783**
	Improve	335.94%	187.12%	148.33%	48.46%	64.18%	20.15%	5.25%	
	R@5	0.0306	0.0277	0.0312	0.0594	0.0528	0.0763	0.0797	**0.0901**
	Improve	194.44%	225.27%	188.78%	51.68%	70.64%	18.09%	13.05%	
	P@10	0.0373	0.0638	0.0663	0.1166	0.1107	0.1522	0.1624	**0.1698**
	Improve	355.23%	166.14%	156.11%	45.63%	53.39%	11.56%	4.56%	
	R@10	0.0531	0.0574	0.0622	0.1166	0.1070	0.1568	0.1607	**0.1728**
	Improve	225.42%	201.05%	177.81%	48.20%	61.50%	10.20%	7.53%	
80%	P@5	0.0288	0.0450	0.0482	0.0833	0.0820	0.1050	0.1220	**0.1287**
	Improve	346.88%	186.00%	167.01%	54.50%	56.95%	22.57%	5.49%	
	R@5	0.0332	0.0306	0.0364	0.0640	0.0606	0.0834	0.0879	**0.0998**
	Improve	200.60%	226.14%	174.18%	55.94%	64.69%	19.66%	13.54%	
	P@10	0.0265	0.0478	0.0512	0.0811	0.0796	0.1053	0.1168	**0.1227**
	Improve	363.02%	156.69%	139.65%	51.29%	54.15%	16.52%	5.05%	
	R@10	0.0586	0.0657	0.0677	0.1242	0.1176	0.1658	0.1788	**0.1898**
	Improve	223.89%	188.89%	180.35%	52.82%	61.39%	14.48%	6.15%	

By utilizing users' personalized tastes captured by MF models, we can attain more accurate predictions.

—PMFSR attains a little better results than those of PMF. This shows that social influence is not so important in POI recommendation and it also coincides the fact that friends share very low, only 9.6%, common POIs.

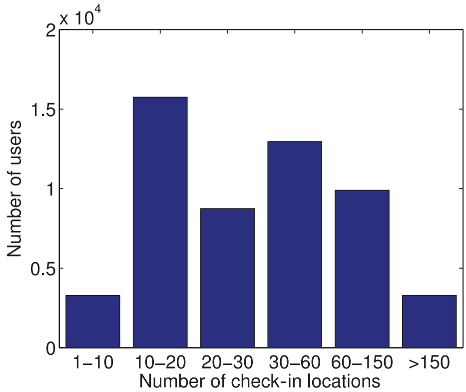
—BPR almost achieves comparable performance with FMFMGM, which verifies our assumption that ranking loss affects the final recommendation. An interesting result is that in Gowalla, BPRLR1 performs the best, while in Foursquare, BPRLR2 performs the best. The reason might be that the data in our Gowalla dataset are sparser than the Foursquare dataset. Note when we use the second scheme, that is, focusing on nearby POIs, it may not work well on Gowalla. We need to have sufficient negative pairs to learn parameters. In Gowalla, the nearby POIs is not enough to learn all parameters well. One scheme may be that we can further sample some far not-visited POIs as well. It may also the reason that BPRLR1 is a little worse than BPR in Gowalla.

—On the Gowalla dataset, GeoMF outperforms our FMFMGM sightly because GeoMF benefits from modeling weighted matrix factorization and incorporates the geographical as a unified model. However, our BPRLR1 performs better than GeoMF. One reason maybe that GeoMF fits both nonzero and zero check-ins with different weights, which is less reasonable than the ranking method. Zero values may be missing values or potential positive ratings just because the user has not noticed them yet.

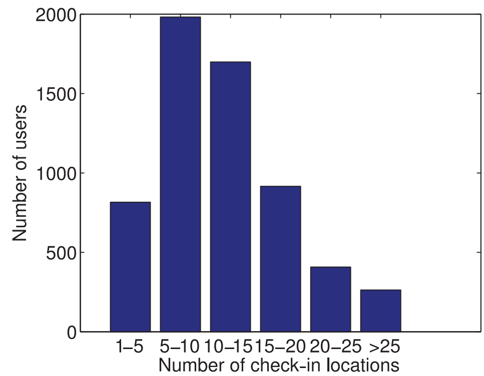
—Our BPRLR2 is still outperforms Rank-GeoFM, which is the state-of-the-art method. One possible reason is that our MGM model capture the geographical influence is more precise that the simple scheme in Rank-GeoFM. Moreover, our method is much more efficient than Rank-GeoFM. BPRLR2 only takes several minutes to converge, while the time for Rank-GeoFM is around 2 days.

5.3. Performance on Different Users

One challenge of the POI recommendation is that it is difficult to recommend POIs to those users who have very few check-ins. In order to compare our methods with the other methods thoroughly, we first group all the users based on the frequency of



(a) Gowalla



(b) Foursquare

Fig. 5. Distribution of user groups.

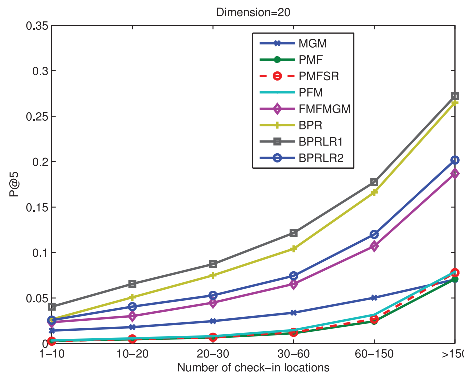
observed check-ins in the training set. Then we evaluate the model performances within different user groups. Here, users are grouped into 6 types: "1–10", "10–20", "20–30", "30–60","60–150" and ">150" for Gowalla; "1–5", "5–10", "10–15", "15–20","20–25" and ">25" for Foursquare. The number denotes the frequency range of users' check-ins in the training data.

Figure 5 summarizes the distribution on different ranges of users' check-in frequency in 70% of the training data. From Figure 6, we observe that when users' check-in frequency is small, MGM outperforms PMF, PMFSR, and PFM. But when users' check-in frequency becomes larger, PMF, PMFSR, and PFM performs better than MGM. It is reasonable because when users' check-in frequency is small, especially for cold-start users, it is difficult to learn users' preferences. Thus, geographical information plays more influence on the prediction. When more check-in information is available, both users' preferences and geographical influence can be learned more accurately, but users' preferences dominate the geographical influence. From Figures 6(c) and 6(d), we can observe that Rank-GeoFM performs poor when there are few check-ins. The reason may be that the geographical influence in Rank-GeoFM is not accurate. When taking the ranking loss into account with our MGM model, we achieve the best performance, especially when the dataset is denser, BPRLR2 consistently outperform other competing methods.

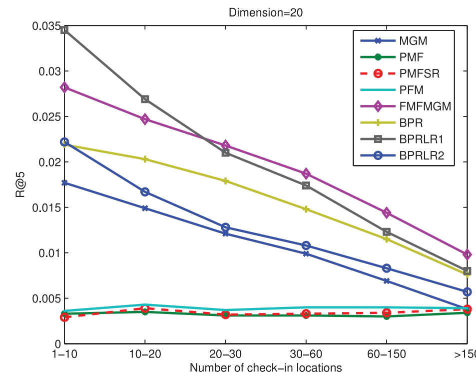
6. CONCLUSION

In this article, we have investigated the characteristics of the large-scale check-in data from two popular LBSNs websites, Gowalla and Foursquare. Based on the extracted properties of the data, we proposed a novel MGM to model the geographical influence of users' check-in behavior. We then propose a fused matrix factorization method to include the geographical influence of users' check-in locations. Furthermore, we proposed to incorporate ranking-oriented CF with all the information together into a unified framework. Results from extensive experiments showed that our proposed methods outperformed other state-of-the-art approaches significantly.

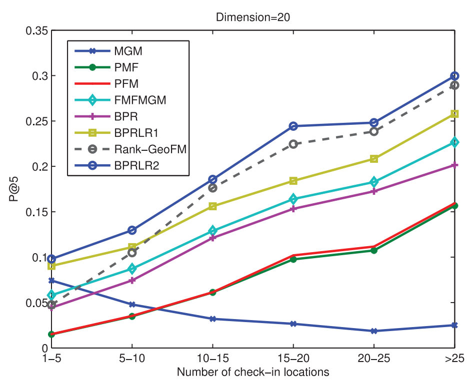
There are several directions worthy of consideration for future study: (1) how to model extremely sparse frequency data, for example, by designing more subtle sampling techniques, to improve MF methods; (2) how to include other information, for example, location category and activity, into our fused framework; (3) how to incorporate



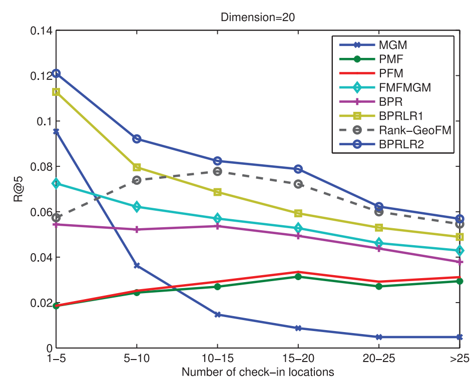
(a) Precision@5 on different user groups in Gowalla



(b) Recall@5 on different user groups in Gowalla



(c) Precision@5 on different user groups in Foursquare



(d) Recall@5 on different user groups in Foursquare

Fig. 6. Performance comparison on different user groups.

temporal effect on POI recommendation to capture the change of users' preferences. We will continue to explore these future directions.

REFERENCES

Suhrid Balakrishnan and Sumit Chopra. 2012. Collaborative ranking. In *Proceedings of the 5th ACM International Conference on Web Search and Data Mining*. ACM, 143–152.

Robert M. Bell and Yehuda Koren. 2007. Lessons from the netflix prize challenge. *SIGKDD Explorations* 9, 2 (2007), 75–79.

Robert M. Bell, Yehuda Koren, and Chris Volinsky. 2007. Modeling relationships at multiple scales to improve accuracy of large recommender systems. In *KDD*. 95–104.

X. Cao, G. Cong, and C. S. Jensen. 2010. Mining significant semantic locations from GPS data. *Proceedings of the VLDB Endowment* 3, 1–2 (2010), 1009–1020.

Ye Chen, Michael Kapralov, Dmitry Pavlov, and John Canny. 2009. Factor modeling for advertisement targeting. In *NIPS*. 324–332.

Chen Cheng, Haiqin Yang, Michael R. Lyu, and Irwin King. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*. AAAI Press, 2605–2611.

Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z. Sui. 2011. Exploring millions of footprints in location sharing services. In *ICWSM*.

Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. Friendship and mobility: User movement in location-based social networks. In *KDD*. 1082–1090.

Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 2013. Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM Conference on Recommender Systems*. ACM, 93–100.

Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 2015. Content-aware point of interest recommendation on location-based social networks. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. Austin, Texas, USA, 1721–1727.

Arthur W. Hafner. 2001. Pareto's principle: The 80-20 rule. *Retrieved December* 26 (2001), 2001.

Tzvetan Horozov, Nitya Narasimhan, and Venu Vasudevan. 2006. Using location for personalized POI recommendations in mobile environments. In *Proceedings of the International Symposium on Applications on Internet (SAINT'06)*. IEEE Computer Society, Washington, DC, 124–129. DOI:http://dx.doi.org/10.1109/SAINT.2006.55

Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *Proceedings of the 2008 8th IEEE International Conference on Data Mining*. IEEE, 263–272.

Eui-young Kang, Hanil Kim, and Jungwon Cho. 2006. Personalization method for tourist point of interest (POI) recommendation. In *Knowledge-Based Intelligent Information and Engineering Systems*, Bogdan Gabrys, Robert Howlett, and Lakhmi Jain (Eds.). Lecture Notes in Computer Science, Vol. 4251. Springer, Berlin, Heidelberg, 392–400. DOI:http://dx.doi.org/10.1007/11892960_48

Yehuda Koren. 2009. Collaborative filtering with temporal dynamics. In *KDD*. 447–456.

Kenneth Wai-Ting Leung, Dik Lun Lee, and Wang-Chien Lee. 2011. CLR: A collaborative location recommendation framework based on co-clustering. In *SIGIR*. 305–314.

Xutao Li, Gao Cong, Xiao-Li Li, Tuan-Anh Nguyen Pham, and Shonali Krishnaswamy. 2015. 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*. ACM, 433–442.

Defu Lian, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. 2014. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 831–840.

Bin Liu and Hui Xiong. 2013. Point-of-interest recommendation in location based social networks with topic and location awareness. In *SDM*, Vol. 13. 396–404.

Eric Hsueh-Chan Lu, Vincent S. Tseng, and Philip S. Yu. 2011. Mining cluster-based temporal mobile sequential patterns in location-based service environments. *IEEE Trans. Knowl. Data Eng.* 23, 6 (2011), 914–927.

H. Ma, C. Liu, I. King, and M. R. Lyu. 2011. Probabilistic factor models for web site recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information*. ACM, 265–274.

H. Ma, H. Yang, M. R. Lyu, and I. King. 2008. Sorec: Social recommendation using probabilistic matrix factorization. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management*. ACM, 931–940.

H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King. 2011a. Recommender systems with social regularization. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*. ACM, 287–296.

H. Ma, T. C. Zhou, M. R. Lyu, and I. King. 2011b. Improving recommender systems by incorporating social contextual information. *ACM Transactions on Information Systems (TOIS)* 29, 2 (2011), 9.

A. Noulas, S. Scellato, C. Mascolo, and M. Pontil. 2011. An empirical study of geographic user activity patterns in foursquare. In *ICWSM11*.

Rong Pan and Martin Scholz. 2009. Mind the gaps: Weighting the unknown in large-scale one-class collaborative filtering. In *KDD*. 667–676.

Rong Pan, Yunhong Zhou, Bin Cao, Nathan Nan Liu, Rajan M. Lukose, Martin Scholz, and Qiang Yang. 2008. One-class collaborative filtering. In *ICDM*. 502–511.

Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 452–461.

Ruslan Salakhutdinov and Andriy Mnih. 2007. Probabilistic matrix factorization. In *NIPS*.

Ruslan Salakhutdinov and Andriy Mnih. 2008. Bayesian probabilistic matrix factorization using Markov chain Monte Carlo. In *ICML*. 880–887.

Salvatore Scellato, Anastasios Noulas, Renaud Lambiotte, and Cecilia Mascolo. 2011. Socio-spatial properties of online location-based social networks. In *Proceedings of the Fifth International Conference on Weblogs and Social Media*. Barcelona, Catalonia, Spain.

Yue Shi, Alexandros Karatzoglou, Linas Baltrunas, Martha Larson, Alan Hanjalic, and Nuria Oliver. 2012a. TFMAP: Optimizing MAP for top-n context-aware recommendation. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 155–164.

Yue Shi, Alexandros Karatzoglou, Linas Baltrunas, Martha Larson, Nuria Oliver, and Alan Hanjalic. 2012b. Climf: Learning to maximize reciprocal rank with collaborative less-is-more filtering. In *Proceedings of the 6th ACM Conference on Recommender Systems*. ACM, 139–146.

Maksims Volkovs and Richard S. Zemel. 2012. Collaborative ranking with 17 parameters. In *Advances in Neural Information Processing Systems*. 2294–2302.

Markus Weimer, Alexandros Karatzoglou, Quoc V. Le, and Alexander J. Smola. 2007. COFIRANK - Maximum margin matrix factorization for collaborative ranking. In *Advances in Neural Information Processing Systems 20, Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems*. Vancouver, British Columbia, Canada, 1593–1600.

Haiqin Yang, Shouyuan Chen, Michael R. Lyu, and Irwin King. 2011a. Location-based topic evolution. In *Proceedings of the 1st International Workshop on Mobile Location-Based Service*.

Haiqin Yang, Zenglin Xu, Jieping Ye, Irwin King, and Michael R. Lyu. 2011b. Efficient sparse generalized multiple kernel learning. *IEEE Transactions on Neural Networks* 22, 3 (March 2011), 433–446.

M. Ye, P. Yin, and W. C. Lee. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 458–461.

Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*. 325–334.

Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. 2013. Time-aware point-of-interest recommendation. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 363–372.

Jia-Dong Zhang and Chi-Yin Chow. 2013. iGSLR: Personalized geo-social location recommendation: A kernel density estimation approach. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 334–343.

Vincent Wenchen Zheng, Bin Cao, Yu Zheng, Xing Xie, and Qiang Yang. 2010a. Collaborative filtering meets mobile recommendation: A user-centered approach. In *Proceedings of the 24rd AAAI Conference on Artificial Intelligence*.

Vincent Wenchen Zheng, Yu Zheng, Xing Xie, and Qiang Yang. 2010b. Collaborative location and activity recommendations with GPS history data. In *WWW*. 1029–1038.

Yu Zheng and Xing Xie. 2011. Learning travel recommendations from user-generated GPS traces. *ACM TIST* 2, 1 (2011), 2.

Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. 2011. Recommending friends and locations based on individual location history. *TWEB* 5, 1 (2011), 5.

Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *WWW*. 791–800.

Tom Chao Zhou, Hao Ma, Irwin King, and Michael R. Lyu. 2009. TagRec: Leveraging tagging wisdom for recommendation. In *Proceedings of CSE*. 194–199.

Received July 2015; revised February 2016; accepted March 2016