Personalized Point-of-Interest Recommendation on Ranking with Poisson Factorization

Yijun Su[1,2,3], Xiang Li[1,2,3], Wei Tang[1,2,3], Daren Zha[2,3 *], Ji Xiang[2,3], Neng Gao[2,3]
[1]*School of Cyber Security, University of Chinese Academy of Sciences*
[2]*State Key Laboratory of Information Security, Chinese Academy of Sciences*
[3]*Institute of Information Engineering, Chinese Academy of Sciences*
Beijing, China
{suyijun, lixiang9015, tangwei, zhadaren, xiangji, gaoneng } @iie.ac.cn

***Abstract*—The increasing prevalence of location-based social networks (LBSNs) poses a wonderful opportunity to build personalized point-of-interest (POI) recommendations, which aim at recommending a top-N ranked list of POIs to users according to their preferences. Although previous studies on collaborative filtering are widely applied for POI recommendation, there are two significant challenges have not been solved perfectly. (1) These approaches cannot effectively and efficiently exploit unobserved feedback and are also unable to learn useful information from it. (2) How to seamlessly integrate multiple types of context information into these models is still under exploration. To cope with the aforementioned challenges, we develop a new Personalized pairwise Ranking Framework based on Poisson Factor factorization (PRFPF) that follows the assumption that users' preferences for visited POIs are preferred over potential POIs, unvisited POIs are less preferred than potential POIs. The framework PRFPF is composed of two modules: candidate module and ranking module. Specifically, the candidate module is used to generate a series of potential POIs from unvisited POIs by incorporating multiple types of context information (e.g., social and geographical information). The ranking module learns the ultimate order of users' preference by leveraging the potential POIs. Experimental results evaluated on two large-scale real-world datasets show that our framework outperforms other state-of-the-art approaches in terms of various metrics.**

***Index Terms*—Location-based Social Networks, POI Recommendation, Unobserved Feedback, Ranking**

I. Introduction

In recent years, location-based social networks (LBSNs) such as Foursquare, Yelp, and Facebook Places are becoming increasingly popular as users can easily post their real location and location-related contents in the physical world via these online systems. In LBSNs, users can establish social links with others to share their experiences of visiting some Point-of-Interests (POIs), e.g., parks, restaurants and cinemas, through making check-ins at these POIs via their mobile devices. The huge volume of data in LBSNs contain valuable information about POIs and users, which can be exploited for building personalized POI recommender systems. POI recommendation has become an important research task in LBSNs, as it not only helps users to explore new interesting places, but also benefits for LBSNs businesses to place advertisements to targeted customers.

* Corresponding author.

The aim of POI recommendation is to generate a top-N ranked list of POIs that a user might be interested in but has not visited before. Most of the existing recommendation methods [1], [2], [3], [4] mainly apply Collaborative Filtering (CF) technique to suggest novel POIs to users. Among these CF techniques, memory-based (e.g., user-based) and model-based (e.g., matrix factorization) are two widely adopted approaches for POI recommendation. The two approaches first predict a rating by modeling the user's preference on a POI. Then the top-N ranked list of POIs suggestions is obtained by sorting the predicted user-POI ratings. However, in practice, POI recommender systems pay more attention to the top-N ranked list of POIs rather than the predicted ratings, hence ranking-based models (i.e., learning-to-rank) that aim to generate accurate ranked lists of POIs are more useful than rating prediction-based models [5], [6]. In other words, previous works on CF are not as effective as the ranking-based models in addressing the recommendation task related to POIs [7]. Hence, there seems a large marginal space left to improve the performance of POI recommendation by extending ranking-based models. In addition, there are two obvious challenges that need to be addressed urgently in previous works on CF. (1) These works usually ignore missing values and are unable to exploit and learn useful contribution information from unobserved feedback, because they cannot distinguish the real negative feedback and missing values [5]. (2) How to seamlessly utilize multiple types of context information, e.g., users' social links [1], geographical distance of POIs [8] and category information of POIs [9], is still under exploration.

In view of the aforementioned challenges, we propose a new Personalized pairwise Ranking Framework based on Poisson Factor factorization (PRFPF) to build the personalized POI recommender system. In the PRFPF, we make the generic assumption that the user's preferences for visited POIs are preferred over potential POIs, unvisited POIs are inferior to potential POIs. The potential POIs can be deemed as potential feedback, which is treated as weak preference relative to positive feedback while as strong preference in comparison to other unobserved feedback. More specifically, our recommendation framework consists of two modules, one of which is the **candidate module** and the other is the **ranking module**. The candidate module is to learn a series of potential POIs (potential feedback) from unvisited POIs for each user by exploiting multiple types of contextual information, e.g., social and geographical information.

For the ranking module, we use the potential feedback generated by the candidate module as input. According to our assumption, the module learns total preference ranking for each user by utilizing two pairwise preferences comparison: visited POIs and potential POIs, and potential POIs and the remaining unvisited POIs. To be specific, by introducing our assumption, we augment the ranking function of Bayesian Personalised Ranking (BPR) [5], which only considers pairwise preference comparison over observed and unobserved feedback. Since Poisson distribution is more suitable for fitting check-in frequency data than Gaussian distribution [1], we propose the Poisson factor factorization to model the difference of two preference prediction. Furthermore, we design a mini-batch gradient descent (MBGD) with the bootstrap sampling algorithm to optimize its objective function. Finally, experimental results on two large-scale real-world datasets demonstrate the effectiveness of our proposed framework compared to several state-of-the-art methods.

The main contributions of this paper can be summarized as follows:

- We propose a new Personalized pairwise Ranking Framework based on Poisson Factor factorization (PRFPF) for POI recommendation. The PRFPF follows the assumption that users' preferences for visited POIs are preferred over potential POIs, unvisited POIs are inferior to potential POIs. Moreover, the PRFPF consists of two modules: candidate module and ranking module.
- The candidate module in PRFPF is designed for generating a series of potential POIs from unvisited POIs by exploiting multiple types of context information (i.e., users' social links, geographical information).
- The ranking module in PRFPF learns the ultimate order of users' preference by leveraging two pairwise preference comparison: visited POIs and potential POIs, and potential POIs and the remaining unvisited POIs.
- We conduct extensive experiments on two large-scale real-world datasets to evaluate the performance of PRFPF. Experimental results show that our framework outperforms other state-of-the-art methods in terms of various metrics.

The rest of this paper is organized as follows. Section II presents related work on POI recommendation. Section III formulates the problem and introduces the necessary background information. Section IV describes the recommendation framework in details. Section V provides an experimental evaluation of the PRFPF. Finally, we draw some conclusions of this study in Section VI.

II. Related Work

In this section, we review relevant studies based on collaborative filtering and ranking-based methods for POI recommendations.

Collaborative Filtering in POI Recommendations. Memory-based and matrix factorization (MF) [10] are two widely used collaborative filtering methods in POI recommendations. Memory-based methods can be grouped into user-based CF and item-based CF. The user-based CF generally uses two steps to make POI suggestions. It first finds similar users to the target user by using a similarity measure based on users' ratings or check-ins, such as Cosine similarity or Pearson correlation. Then the prediction score is calculated by weighting average of all the ratings from similar users. Similarity, the item-based CF works according to the user's preferences on other similar items. The friend-based CF [8] is a variant of the user-based CF, which can be realized by leveraging check-ins of social friends. MF [10] has gained a large popularity due to its effectiveness in dealing with the user-POI check-in matrix. Cheng et al. [1] first integrated social information into MF to improve the quality of POI recommendation. Weighted matrix factorization (WMF) [11], [3] by assigning different weights for positive and negative examples is also designed for POI recommendations. Besides, Lian et al. [3] incorporated the geographical information into MF to enhance recommendation performance. All the methods mentioned above are essentially rating prediction-based models, and performance improvements are limited.

POI Recommendation Approaches Based on Ranking. Bayesian personalized ranking (BPR) [5] is a pairwise ranking method, which focus on modeling the ranking of the feedback. It learns the ranking based on pairwise preference comparison over observed and unobserved feedback such that the Area Under the ROC Curves (AUC) can be maximized [12]. From the perspective of ranking tasks, these CF-based methods mentioned above can be viewed as pointwise methods [13]. Empirical studies [13], [6], [4] have demonstrated that pointwise methods are generally less effective than pairwise ranking methods. Matrix factorization-based BPR (BPR-MF) [5] is the most commonly used and effective ranking-based model for POI recommendation. Li et al. [14] proposed a ranking method based geographical factorization to make POI recommendations. Yuan et al. [12] proposed GeoBPR model that injects users' geo-spatial preference.

In this paper, the framework PRFPF we proposed differs from the existing CF and ranking approaches in three aspects. First, PRFPF makes a generic assumption that users' preferences for visited POIs are preferred over potential POIs, unvisited POIs are less preferred than potential POIs. Second, PRFPF is capable of integrating multiple types of context information, i.e., not limited to social or geographical ones. Third, PRFPF models the check-in frequencies by using the Poisson distribution instead of Gaussian distribution.

III. Preliminaries

In this section, we first formulate the POI recommendation problem in LBSNs. Then we provide background on Bayesian Personalised Ranking, which serves as the building block for our recommendation framework.

A. Problem Definition

The problem of personalized POI recommendation is to generate a top-N ranked list of POIs that a user might be interested in but has not visited before by leveraging users' historical check-ins and other available context information in LBSNs. Let $\mathcal{U} = \{u_1, u_2, ..., u_m\}$ be a set of users, where each user u_i checked in some POIs $\mathcal{L}_{u_i}^+$. Let $\mathcal{L} = \{l_1, l_2, ..., l_n\}$ be a set of POIs, where each POI has a location $l_j = \{lon_j, lat_j\}$ in terms of longitude and latitude. For convenience, we term i as user u_i and j as POI l_j, unless stated otherwise. The user-POI check-in matrix is represented as $F \in \mathbb{R}^{m \times n}$, where each entry f_{ij} denotes the check-in frequency of user i on POI j.

In this paper, we consider three different types of feedback, namely positive, potential and negative feedback. The positive feedback is defined as a set of POIs previously visited by user i: $P_i = \mathcal{L}_i^+$. The potential feedback $LP_i = \{l_1, ..., l_c\}$ is learned from unvisited POIs and divided into social and geographic feedback in detail. The remaining unvisited POIs are viewed as the negative feedback $N_i = \{l_1, ..., l_h\}$. Here *negative* only means no explicit feedback can be observed from the user and does not denote users' dislike of the POIs.

B. Bayesian Personalized Ranking

Bayesian personalized ranking (BPR) learns the ranking based on pairwise preference comparison over observed and unobserved feedback. Its optimization criterion by maximizing posterior estimator with Bayesian theory is proposed by [5], [12]. For user i, the ranking order of his preference is defined as follows:

$$j \succ_i h \Leftrightarrow j \in \mathcal{L}_i^+ \wedge h \in \mathcal{L} \setminus \mathcal{L}_i^+, \tag{1}$$

where $\succ_i$ is the total order, which denotes latent preference structure desired by the user.

Then the BPR pairwise ranking function can be given:

$$\hat{r}_{ijh}(\Theta) := \hat{x}_{i,j} - \hat{x}_{i,h}, \tag{2}$$

where Θ represents a set of parameters, $\hat{r}_{ijh}(\Theta)$ is the ranking function that user i prefers POI j over POI h, $\hat{x}_{i,j}$ and $\hat{x}_{i,h}$ are the predicted check-in frequencies.

IV. POI Recommendation Framework

A. Candidate Module

The candidate module consists of two building blocks: social and geographic models. Each model learns a series of potential feedback from unvisited POIs for each user by exploiting the corresponding information.

1) Social Model:

Social friends are acting as an important role in people's life. Users always turn to friends they trust for movie or restaurant recommendations, and their tastes or behaviors can be easily affected by these friends. Many previous studies [4], [15] have indicated that social friends can help improve the performance of POI recommendation to a certain extent. That is, users might be interested in those POIs which have been checked in by their friends. Inspired by this research [16], we use social relationships to generate a set of social feedback instead of directly predicting real-valued scores to recommend POIs, which is different from traditional approaches on CF [17], [9], [1]. The friend-based CF (FCF) [8] is an effective algorithm in LBSNs, which mainly relies on the similarity between the user and his friends. However, the similarity used in this method is only leverages the 1-hop friendship [15] relation and ignores the local structure information of the user node in LBSNs. For instance, 2-hop neighbors of user node (i.e., friends of friends) are also likely to affect the user's check-in behavior but are overlooked.

Hence, we propose a novel friend-based CF based on network representation learning technique (i.e., Struct2vec [18]), which aims at learning low-dimensional vector representation for each user by capturing local structural information of nodes in LBSNs. The effectiveness of Struct2vec has been verified by classification task [18]. For the sake of convenience, the Friend-based CF based on Struct2vec is named as FCFS.

The FCFS can be described as follows.

$$\hat{c}_{ij} = \frac{\sum_{s \in SF_i} SI_{i,s} \cdot f_{s,j}}{\sum_{s \in SF_i} SI_{i,s}}, \tag{3}$$

where $\hat{c}_{ij}$ is the predicted score of user i at POI j, $SI_{i,s}$ is the similarity between user i and friend s, SF_i is a set of the user's social friends, and $f_{s,j}$ is the check-in frequency of friend s at POI j.

We define $SI_{i,s}$ using a combination of Kernel function and similarity of check-in, which differs from previous work [8].

$$SI_{i,s} = \psi exp(-\frac{||\mathbf{g_i} - \mathbf{g_s}||^2}{\sigma^2}) + (1-\psi)\frac{|\mathcal{L}_i^+ \cap \mathcal{L}_s^+|}{|\mathcal{L}_i^+ \cup \mathcal{L}_s^+|}, \tag{4}$$

where $\mathbf{g_i} \in \mathbb{R}^k$ and $\mathbf{g_s} \in \mathbb{R}^k$ are two 5-dimensional vectors learned by Struct2vec [18], ψ is a tuning parameter ranging with [0,1], and σ is a scale parameter that can be tuned by a local scaling technique. The first term of (4) measures the social relations in LBSNs. The second indicates the common check-in ratio between the user and his friends.

Finally, we sort all POIs that users' friends visited in accordance with their scores to acquire the Top-t_1 as social feedback for each user, where t_1 is the number of potential POIs we defined. There are two distinct advantages of using the FCFS in comparison with the FCF. First, the FCFS can capture local structure information of user node. Second, the FCFS alleviates the sparsity problem of social links to a certain extent and is friendly to the user with fewer friends.

2) Geographic Model:

Unlike traditional items (e.g., books, films and music) recommendation, the POI recommendation task needs to physical interactions between users and locations. Therefore, the geographical information of POIs represented by longitudes and latitudes is a significant factor that affects users' check-in decision-making. In this paper, we are committed to studying how geographical distance influences users' check-in choices. First, the work [1] clustered the user's whole historical check-ins and found these locations are around several centers (e.g., home, office and travel places). Enlightened by this, the distance between each center and POIs should be taken into account because it measures the distance cost of the user's check-in. For instance, one user has a small probability to visit a distant POI, even if he is interested. In other words, the user tends to visit POIs close to his centers. Second, this research [8] showed that the user may be interested in exploring nearby POIs of a POI that he likes, even though it is far away from home. For example, one user has a great chance to go to nearby famous restaurants to eat after shopping. Thus, we argue that the distance among POIs also has a significant influence on users' check-in behavior due to the user's preference on nearby POIs.

For the first distance, we use the following distance measure to find the t_2 POIs closest to each center of the user, where t_2 is the number of potential POIs.

$$d(j, o) = ||l_j - l_o||^2, \tag{5}$$

where $||.||$ denotes the Euclidean norm in the geographical space and $l_o \in M_i$, where M_i represents multi-centers of the user i. We apply a greedy clustering algorithm [1] to find the multi-centers M_i. More details are shown in [1].

For the second distance, we present the method based on Gaussian Kernel function to discover the t_3 POIs that are closest to locations the user has checked in.

$$Sco(i, j) = \sum_{l_k \in \mathcal{L}_i^+} exp(-\frac{\Upsilon_i}{2}||l_j - l_k||^2), \tag{6}$$

where $Sco(i, j)$ denotes the distance relevance score of i at j, and Υ_i is a personalized adaptive bandwidth that depicts the user's activity area.

$$\Upsilon_i = max\left\{||l_k - h_i||^2\right\}, l_k \in \mathcal{L}_i^+, \tag{7}$$

where h_i indicates the user's home.

In the end, we merge the potential POIs produced by the above two ways and regard them as the final geographic feedback.

B. Ranking Module

In this section, we first formalize our model assumption regarding positive, potential, and negative feedback. Then we describe the proposed ranking model. Finally, we present the process of parameter estimation.

1) Model Assumption:

As mentioned in Section III-B, the basic assumption of BPR is that the user prefers an observed POI over all unobserved POIs. However, this assumption suffers from an obvious drawback, namely cannot mine more contribution information from unobserved POIs. Now our candidate module regards potential feedback as intermediate feedback, which can alleviate the situation. For user i, the ranking order of his preference for positive, potential and negative feedback can be formulated as follows:

$$j \succ_i c \wedge c \succ_i h \Leftrightarrow j \in \mathcal{P}_i \wedge c \in \mathcal{LP}_i \wedge h \in \mathcal{N}_i, \tag{8}$$

The pairwise ranking function can be further given:

$$\hat{r}_{i,j,c,h}(\Theta) := \underbrace{\hat{x}_{i,j} \succ \hat{x}_{i,c}}_{:=\hat{r}_{i,j,c}} \wedge \underbrace{\hat{x}_{i,c} \succ \hat{x}_{i,h}}_{:=\hat{r}_{i,c,h}}, \tag{9}$$

where Θ denotes a set of parameters, $\hat{r}_{i,j,c,h}(\Theta)$ is the ranking function that user i prefers POI j over POI c and prefers POI c over POI h, and $\hat{x}_{i,j}$, $\hat{x}_{i,c}$ and $\hat{x}_{i,h}$ are the predicted check-in frequencies.

Our proposed assumption is more general because any feedback generated by context information in LBSNs can be used as potential feedback. In addition, an opposite assumption that potential POIs are likely to be unattractive is presented. However, due to its poor performance in previous experiments [12], we are not conducting research here.

2) Model Formulation:

Based on the above assumption, we can find accurate personalized ranking for the user i by using the maximum posterior estimator:

$$p(\Theta| \succ_i) \propto p(\succ_i |\Theta)p(\Theta), \tag{10}$$

where $p(\succ_i |\Theta)$ represents the likelihood function and $p(\Theta)$ is the prior distribution of parameters Θ.

We use the same assumptions as [12]: (1) All users' actions are independent of each other. (2) The preference ordering of each triple of items (j, c, h) for a specific user is independent

Algorithm 1 Learning Algorithm

Require: feedback data: user $i \in \mathcal{U}$, positive feedback $\mathcal{P}_i$, potential feedback $\mathcal{LP}_i$, and negative feedback $\mathcal{N}_i$
hyperparameters: sampling times st, batch size bs, and learning rate η

Ensure: model parameters $\Theta = \{U, L\}$

Initialization Θ with Gamma distribution:
$U \sim Gamma(\alpha_U, \beta_U), L \sim Gamma(\alpha_L, \beta_L)$

$s = 0$

for $t = 1$ to st **do**
  Uniformly sample a user i from $\mathcal{U}$
  Uniformly sample a positive feedback j from $\mathcal{P}_i$
  Uniformly sample a potential feedback c from $\mathcal{LP}_i$
  Uniformly sample a negative feedback h from $\mathcal{N}_i$
end for
while $(s + 1) * bs \leq st$ **do**
  for $j = 1$ to bs **do**
    $u_{ik} \leftarrow u_{ik} - \eta(\frac{\partial J}{\partial u_{ik}})$
    $l_{jk} \leftarrow l_{jk} - \eta(\frac{\partial J}{\partial l_{jk}})$
    $l_{ck} \leftarrow l_{ck} - \eta(\frac{\partial J}{\partial l_{ck}})$
    $l_{hk} \leftarrow l_{hk} - \eta(\frac{\partial J}{\partial l_{hk}})$
  end for
  $s = s + 1$
end while
return Θ

of the ordering of every other triple. Hence, the likelihood function for all users can be given:

$$\prod_{i\in\mathcal{U}} p(\succ_i |\Theta) = \prod_{(i,j,c,h)\in\mathcal{U}\times\mathcal{L}\times\mathcal{L}\times\mathcal{L}} p(\hat{x}_{i,j} \succ \hat{x}_{i,c} \wedge \hat{x}_{i,c} \succ \hat{x}_{i,h}|\Theta)^{\zeta((i,j,c,h)\in D_s)} (1 - p(\hat{x}_{i,j} \succ \hat{x}_{i,c} \wedge \hat{x}_{i,c} \succ \hat{x}_{i,h}|\Theta))^{\zeta((i,j,c,h)\notin D_s)}, \tag{11}$$

where D_s is a poset of $>_i$:

$$D_s = \{(i,j,c,h)|j \in \mathcal{P}_i \wedge c \in \mathcal{LP}_i \wedge h \in \mathcal{N}_i\}$$

, and $\zeta(b)$ is an indicator function that equals to 1 if b is true, otherwise equals to 0.

Due to the totality and antisymmetry of pairwise ordering scheme, the above likelihood function can be simplified to:

$$\prod_{i\in\mathcal{U}} p(\succ_i |\Theta) = \prod_{i\in\mathcal{U}, j\in\mathcal{P}_i, c\in\mathcal{LP}_i} p(\hat{x}_{i,j} \succ \hat{x}_{i,c}|\Theta) \prod_{i\in\mathcal{U}, c\in\mathcal{LP}_i, h\in\mathcal{N}_i,} p(\hat{x}_{i,c} \succ \hat{x}_{i,h}|\Theta), \tag{12}$$

We apply a differential function (e.g. a sigmoid function $\sigma(x) = \frac{1}{1+e^{-x}}$) to approximate the function $p(.)$ so that the likelihood function is differentiable. Based on this trick, we can obtain:

$$p(\hat{x}_{i,j} \succ \hat{x}_{i,c}|\Theta) = \sigma(\hat{x}_{i,j} - \hat{x}_{i,c}), \tag{13}$$

$$p(\hat{x}_{i,c} \succ \hat{x}_{i,h}|\Theta) = \sigma(\hat{x}_{i,c} - \hat{x}_{i,h}), \tag{14}$$

where $\hat{x}_{i,j} - \hat{x}_{i,c}$ denotes the difference of two predicted preference.

Since Poisson distribution is more suitable for fitting check-in frequency data than Gaussian distribution [1], [2], we propose the Poisson Factor Factorization to model the difference of two preference. Hence, $\Theta = (U, L)$ and $\hat{x}_{i,j} = \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_j}$ where $\boldsymbol{u_i}$ and $\boldsymbol{l_j}$ represent latent feature vector of user-specific and POI-specific.

For the $p(\Theta)$, we have:

$$\begin{aligned} p(\Theta) =& p(U|\alpha_U, \beta_U)p(L|\alpha_L, \beta_L) \\ & \prod_{i=1}^{m}\prod_{k=1}^{d} \frac{u_{ik}^{\alpha_U-1} exp(-u_{ik}/\beta_U)}{\beta_U^{\alpha_U}\Gamma(\alpha_U)} \times \\ & \prod_{j=1}^{n}\prod_{k=1}^{d} \frac{l_{jk}^{\alpha_L-1} exp(-l_{jk}/\beta_L)}{\beta_L^{\alpha_L}\Gamma(\alpha_L)}, \end{aligned} \tag{15}$$

where $\alpha_U, \alpha_L, \beta_U, \beta_L$ are parameters for Gamma distributions and $\Gamma(.)$ is the Gamma function.

Finally, we can give the objective function as follows:

$$\begin{aligned} J(\Theta) = \min_{U,L} - \sum_{i\in\mathcal{U}} \Bigg[& \sum_{j\in\mathcal{P}_i}\sum_{c\in\mathcal{LP}_i} ln\sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_j} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c}) \\ & + \sum_{c\in\mathcal{LP}_i}\sum_{h\in\mathcal{N}_i} ln\sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_h}) \Bigg] \\ & - \sum_{i=1}^{m}\sum_{k=1}^{d}((\alpha_U - 1)ln(u_{ik}/\beta_U) - u_{ik}/\beta_U) \\ & - \sum_{j=1}^{n}\sum_{k=1}^{d}((\alpha_L - 1)ln(l_{jk}/\beta_L) - l_{jk}/\beta_L), \end{aligned} \tag{16}$$

3) Parameter Estimation:

The matrices U and L can be learned by solving the optimization problem in (16). Here, we propose a Mini-batch Gradient Descent with the bootstrap sampling to optimize the objective function. The process of learning is to first sample and then iteratively update model parameters. More details of optimization are depicted in Algorithm 1. The gradients of the objective function with respect to u_{ik}, l_{jk}, l_{ck} and l_{hk} are:

$$\begin{aligned} \frac{\partial J}{\partial u_{ik}} =& \frac{1}{\beta_U} - \frac{\alpha_U - 1}{u_{ik}} - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_j} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c}))(l_{jk} - l_{ck}) \\ & - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_h})(l_{ck} - l_{hk}), \end{aligned} \tag{17}$$

$$\frac{\partial J}{\partial l_{jk}} = \frac{1}{\beta_L} - \frac{\alpha_L - 1}{l_{ik}} - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_j} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c}))u_{ik}, \tag{18}$$

$$\begin{aligned} \frac{\partial J}{\partial l_{ck}} =& \frac{1}{\beta_L} - \frac{\alpha_L - 1}{l_{ik}} - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_j} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c}))(-u_{ik}) \\ & - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_c} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_h})u_{ik}, \end{aligned} \tag{19}$$

$$\frac{\partial J}{\partial l_{hk}} = \frac{1}{\beta_L} - \frac{\alpha_L - 1}{l_{hk}} - (1 - \sigma(\boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_v} - \boldsymbol{u_i}^{\mathrm{T}}\boldsymbol{l_h}))(-u_{ik}), \tag{20}$$

Since a series of potential feedback generated in candidate module can be precomputed, the computational complexity of our framework PRFPF mainly consists of ranking module learning and predicting preference. The computation of each update gradient is $O(d)$ [6], where d is the number of latent feature dimensions. Thus the total complexity of ranking module learning is $O(st \cdot d)$, where st is the sampling times. Regarding predicting a user's preference on a specific POI, its complexity is $O(d)$. From this perspective, the computational complexity of PRFPF do not increase and is approximately equal to BPR-MF [5].

C. Unified framework

For further generic, we specially design a unified Personalized Ranking Framework (named PRFPF) to integrate candidate module and ranking module. In addition to exploiting social, geographical information, it can also integrate other information (e.g., category and time) in LBSNs into the framework. The overall architecture of the PRFPF is shown in Fig.1.

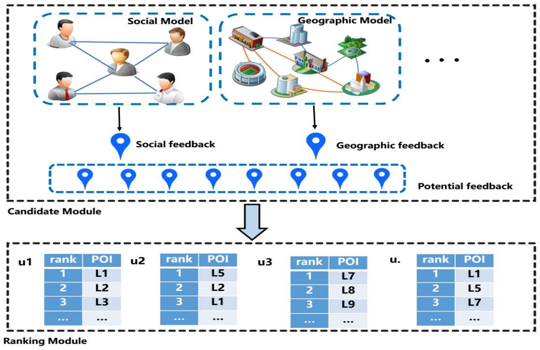


Fig. 1. The architecture framework of PRFPF. The candidate module is used to learn a series of potential feedback from unvisited POIs for each user. The ranking module recommends a personalized top-N list of POIs for each user by exploiting positive, potential, and negative feedback.

V. Experimental Evaluation

In this section, we systematically evaluate the recommendation performance of our framework PRFPF and compare our framework with some state-of-the-art POI recommendation algorithms. All experiments are conducted on two large-scale real-world LBSN datasets, collected from Yelp and Foursquare.

A. Datasets

We use two publicly available real-world check-in datasets that were collected from Yelp [19] and Foursquare [4]. Each check-in contains the user ID, location ID, check-in time, and geo-coordinates of the location. Users' social links are also provided in datasets. Since the user's home is used in the geographic model, we adopt the recursive grid method [20] to estimate its geo-coordinates. Note that, for Yelp and Foursquare datasets, we remove those users who have checked-in less than 10 locations and those POIs which are visited by less than 10 users. The statistics of these two datasets are shown in Table I.

In our experiments, we divide each dataset into training set, tuning set and test set in terms of the user's check-in time instead of choosing a random partition method. For each user, the earliest 70% check-ins are selected for training, the most recent 20% check-ins as testing, and the next 10% as tuning.

B. Evaluation Metrics

We utilize four popular metrics to evaluate the performance of the model we proposed: precision (Pre@N), recall (Rec@N) [9], mean average precision (MAP@N) [4] and normalized discounted cumulative gain (NDCG@N) [19], where N is the number of recommended POIs. For each metric, we calculate the average performance of all users. We omit detailed descriptions for saving space.

C. Baseline Methods

In order to demonstrate the benefits of our recommendation framework, we compare our model with the following baselines for POI recommendation.

TABLE I
Statistical Information of the Two Datasets

Statistical item	Yelp	Foursquare
Number of users	30,887	2,551
Number of POIs	18,995	13,474
Number of categories	624	10
Number of check-ins or ratings	860,888	124,933
Number of social links	265,533	32,512
User-POI matrix density	0.14%	0.291%

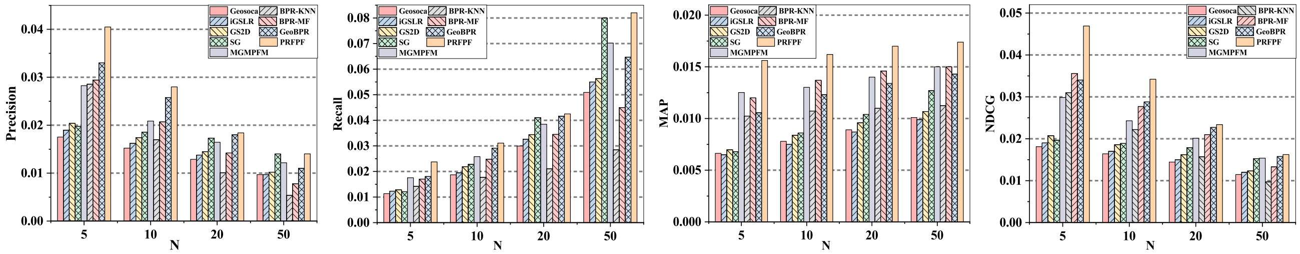
- **GeoSoCa**: This is a state-of-the-art personalized POI recommendation method that combines geographical, social, and categorical information [9].
- **SG**: This is a typical recommendation approach that integrates social and geographical information [8].
- **MGMPFM**: This is a recommendation framework based on Poisson Factor Factorization, which exploits geographical influence with Multi-center features to recommend POIs for users [1].
- **GS2D**: This method uses social and geographical influences to recommend POIs [21].
- **iGSLR**: This method integrates user preference, social influence and personalized geographical influence into a unified location recommendation framework [17].
- **BPR-KNN**: This is a ranking-based adaptive model, which employs item-based k-nearest-neighbor to predict ratings [5].
- **BPR-MF**: This is a classical pairwise ranking model based on matrix factorization [5].
- **GeoBPR**: This method is a state-of-the-art method for POI recommendation, which incorporates the geographic feedback into the BPR model [12].

D. Parameter settings

For all the compared baselines, we adopt the optimal parameter reported in their works. In our experiments, all critical parameters are tuned through cross-validation. Empirically, for the social model, the parameters σ and ψ are set to 0.1 and 0.05, respectively. In Foursquare dataset, d is set to 50, $\alpha_U = \alpha_L = 20$, $\beta_U = 0.05$, $\beta_L = 0.3$, $t_1 = t_2 = 10$, $t_3 = 5$ and $\eta = 0.005$. In Yelp dataset, $d = 100$, $\alpha_U = 30, \alpha_L = 15$, $\beta_U = 0.02$, $\beta_L = 0.15$, $t_1 = t_2 = 20$, $t_3 = 5$ and $\eta = 0.001$. The effect of latent factor dimension d will be detailed later.

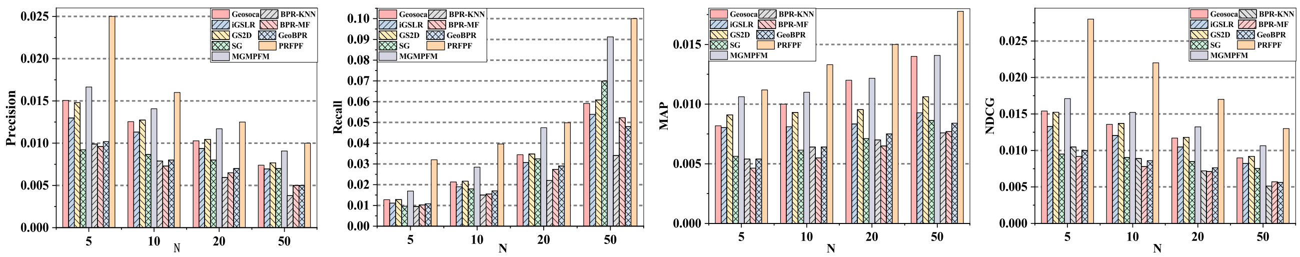
E. Experimental Results

Performance Comparisons. The experimental results of each recommendation algorithm in terms of Pre@N, Rec@N, MAP@N and NDCG@N on Foursquare and Yelp are reported in Fig.2 and Fig.3. From the performance comparison of all algorithms in Fig.2, we can see that our framework achieves the best performance in terms of all four metrics, which illustrates the superiority of our framework. On the one hand, compared with non-ranking algorithms Geosoca, iGSLR, GS2D and MGMPFM, our ranking framework presents an absolute advantage. For instance, Pre@5, Rec@5, MAP@5 and NDCG@5 are improved by around 43%, 35%, 25% and 57%, comparing to the best non-ranking approach MGMPFM. In addition, the reason we guess MGMPFM performs well



(a) Precision (b) Recall (c) MAP (d) NDCG

Fig. 2. Varying N on Foursquare



(a) Precision (b) Recall (c) MAP (d) NDCG

Fig. 3. Varying N on Yelp

in these non-ranking models is that it employs the Poisson distribution to fit the check-in frequency.

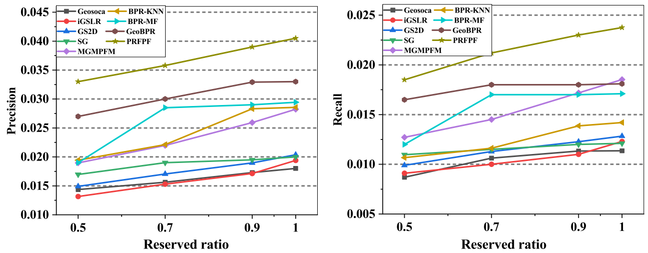
On the other hand, our framework significantly outperforms other three ranking algorithms BPR-KNN, BPR-MF and GeoBPR. For example, PRFPF improves the second best recommendation algorithm GeoBPR by 22.7%, 31.6%, 47.9% and 37.8% in terms of Pre@5, Rec@5, MAP@5 and NDCG@5, respectively. Based on the results of ranking algorithms, two interesting observations are revealed. First, the basic BPR methods like BPR-KNN or BPR-MF performs the worst among all ranking algorithms. One possible explanation is that the basic BPR methods only exploit preference order between observed and non-observed feedback to learn to rank. Second, the extended BPR methods GeoBPR has better performance than the basic BPR methods. This is because it uses geographic feedback to facilitate learning.

The performance on Foursquare is similar to that of Yelp, and the specific analysis is omitted here. Finally, we summaries two advantages in our framework by comparing all the above eight baselines. (1) PRFPF effectively learns a series of potential POIs from unvisited POIs by exploiting incorporating social and geographical information so that assist ranking learning. (2) PRFPF employs Poisson factor factorization to model the difference of two preference prediction as Poisson distribution is more suitable for fitting check-in frequency data than Gaussian distribution.

Impact of Data Sparsity. Here, we study the effectiveness of our model on sparse problems. In order to generate check-in matrix with different sparsity, we we randomly reserve $x\%$ (x= 50,70,90,100) of check-ins from each user's visited records. Fig.4 shows the overall results of all recommendation algorithms on Foursquare under different sparsity. Here, the smaller the reserved ratio x is, the sparser the check-in matrix is. From the Fig.4, we find that the Pre@5 and Rec@5 of all algorithms are increasing with the increase of the reserved ratio x. A reasonable explanation is that, with the increase of the proportion of the training set, the number of positive examples increases, and then contributes to the improvement on the Pre@5 and Rec@5. By comparing the results of ranking and non-ranking models, we can observe that all the ranking models consistently better than the non-ranking models in terms of all four metrics under different sparsity. We guess the reason is that the training set of ranking models is composed of a large number of positive and negative examples. This alleviates the sparsity of non-ranking methods to a certain extent.

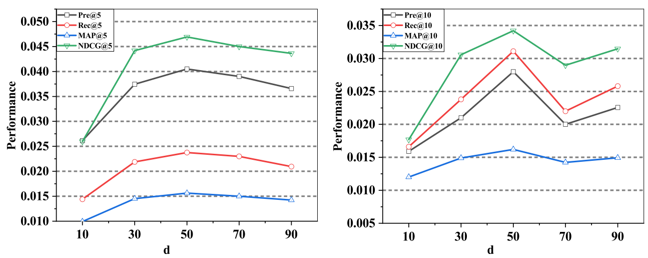
We can further see that our framework PRFPF consistently outperforms all ranking and non-ranking baselines at different densities, which demonstrates the effectiveness of PRFPF. This may be attributed to the two pairwise preference assumption in our framework: users' preferences for visited POIs are preferred over potential POIs, unvisited POIs are less preferred than potential POIs.

Study of Influence of Latent Factor Dimension d. In this study, we employ Poisson factor factorization (a matrix factorization technique) to predict the difference between the two scores of preference for users. Hence, it is necessary to study the effect of sensitivity parameter d, where d is the number of latent feature dimension. In our experiment, we set d to 10, 30, 50, 70 and 90, respectively. We choose Foursquare dataset to observe the performance of the recommendation. Fig.5 shows that the recommended quality for different values of d. From the figure we can observe that the performance of the PRFPF increases with the increase of the d at the beginning, then hits the highest recommended quality when d = 50, and eventually tends to decline. The above trend



(a) Pre@5-Foursquare (b) Rec@5-Foursquare

Fig. 4. Impact of Data Sparsity



(a) N=5 (b) N=10

Fig. 5. Influence of Latent Factor Dimensions d

indicates that the performance is best at d=50, and so we finally choose the optimal parameter d=50.

VI. Conclusions

In this paper, we propose a novel personalized ranking framework based on Poisson factor factorization, PRFPF, for POI recommendation. The framework is composed of two modules: candidate module and ranking module. The candidate module is used to generate a series of potential feedback from unvisited POIs by exploiting social and geographical information. Based on potential feedback, the ranking module learns the personalized top-N ranked list of POIs by comparing two pairwise preferences: visited POIs and potential POIs, potential POIs and unvisited POIs. PRFPF has good scalability and can integrate other information besides the two kinds of information used in this paper. Experimental results on real-world Foursquare and Yelp datasets showed that our framework is effective and significantly outperforms other state-of-the-art approaches.

Acknowledgement

This work is supported by National Key Research and Development Program of China, National Natural Science Foundation of China (No.U163620068).

References

[1] 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 the Twenty-Sixth AAAI Conference on Artificial Intelligence*. AAAI Press, 2012, pp. 17–23.

[2] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 5, pp. 1167–1179, 2015.

[3] 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 the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2014, pp. 831–840.

[4] H. Li, Y. Ge, R. Hong, and H. Zhu, "Point-of-interest recommendations: Learning potential check-ins from friends," in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 975–984.

[5] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, ser. UAI 2009. AUAI Press, 2009, pp. 452–461.

[6] J. Manotumruksa, C. Macdonald, and I. Ounis, "A personalised ranking framework with multiple sampling criteria for venue recommendation," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 1469–1478.

[7] Y. Shi, A. Karatzoglou, L. Baltrunas, M. Larson, N. Oliver, and A. Hanjalic, "Climf: Learning to maximize reciprocal rank with collaborative less-is-more filtering," in *Proceedings of the Sixth ACM Conference on Recommender Systems*, ser. RecSys 2012. ACM, 2012, pp. 139–146.

[8] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2011, pp. 325–334.

[9] J.-D. Zhang and C.-Y. Chow, "Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 443–452.

[10] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[11] Y. Liu, W. Wei, A. Sun, and C. Miao, "Exploiting geographical neighborhood characteristics for location recommendation," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 2014, pp. 739–748.

[12] F. Yuan, J. M. Jose, G. Guo, L. Chen, H. Yu, and R. S. Alkhawaldeh, "Joint geo-spatial preference and pairwise ranking for point-of-interest recommendation," in *2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)*, 2016, pp. 46–53.

[13] T.-Y. Liu, "Learning to rank for information retrieval," *Found. Trends Inf. Retr.*, vol. 3, no. 3, pp. 225–331, Mar. 2009.

[14] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy, "Rankgeofm: 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, 2015, pp. 433–442.

[15] Y. Su, X. Li, W. Tang, J. Xiang, and Y. He, "Next check-in location prediction via footprints and friendship on location-based social networks," in *2018 19th IEEE International Conference on Mobile Data Management (MDM)*, vol. 00, 2018, pp. 251–256.

[16] T. Zhao, J. McAuley, and I. King, "Leveraging social connections to improve personalized ranking for collaborative filtering," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 2014, pp. 261–270.

[17] J.-D. Zhang and C.-Y. Chow, "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, 2013, pp. 334–343.

[18] L. F. Ribeiro, P. H. Saverese, and D. R. Figueiredo, "Struc2vec: Learning node representations from structural identity," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 385–394.

[19] Y. Liu, T.-A. N. Pham, G. Cong, and Q. Yuan, "An experimental evaluation of point-of-interest recommendation in location-based social networks," *Proc. VLDB Endow.*, vol. 10, no. 10, pp. 1010–1021, Jun. 2017.

[20] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, "Exploring millions of footprints in location sharing services." in *ICWSM*. The AAAI Press, 2011.

[21] J.-D. Zhang, C.-Y. Chow, and Y. Li, "Lore: Exploiting sequential influence for location recommendations," in *Proceedings of the 22Nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2014, pp. 103–112.