A General Geographical Probabilistic Factor Model for Point of Interest Recommendation

Bin Liu, Hui Xiong, *Senior Member, IEEE*, Spiros Papadimitriou, Yanjie Fu, and Zijun Yao

Abstract—The problem of point of interest (POI) recommendation is to provide personalized recommendations of places, such as restaurants and movie theaters. The increasing prevalence of mobile devices and of location based social networks (LBSNs) poses significant new opportunities as well as challenges, which we address. The decision process for a user to choose a POI is complex and can be influenced by numerous factors, such as personal preferences, geographical considerations, and user mobility behaviors. This is further complicated by the connection LBSNs and mobile devices. While there are some studies on POI recommendations, they lack an integrated analysis of the joint effect of multiple factors. Meanwhile, although latent factor models have been proved effective and are thus widely used for recommendations, adopting them to POI recommendations requires delicate consideration of the unique characteristics of LBSNs. To this end, in this paper, we propose a general geographical probabilistic factor model (Geo-PFM) framework which strategically takes various factors into consideration. Specifically, this framework allows to capture the geographical influences on a user's check-in behavior. Also, user mobility behaviors can be effectively leveraged in the recommendation model. Moreover, based our Geo-PFM framework, we further develop a Poisson Geo-PFM which provides a more rigorous probabilistic generative process for the entire model and is effective in modeling the skewed user check-in count data as implicit feedback for better POI recommendations. Finally, extensive experimental results on three real-world LBSN datasets (which differ in terms of user mobility, POI geographical distribution, implicit response data skewness, and user-POI observation sparsity), show that the proposed recommendation methods outperform state-of-the-art latent factor models by a significant margin.

Index Terms—Recommender systems, point of interest (POI), probabilistic factor model, location-based social networks

1 Introduction

Recent years have witnessed the increased development and popularity of location-based social network (LBSN) services, such as Foursquare, Gowalla, and Facebook Places. LBSNs allow users to share their check-ins and opinions on places they have visited, ultimately helping each other find better services. Data collected through LBSN activity can enable better recommendations of places, or Points of Interest (POIs) such as restaurants and malls. This can drastically improve the quality of location-based services in LBSNs, simultaneously benefiting not only LBSN users but also POI owners. On one hand, mobile users can identify favorite POIs and improve their user experience via good POI recommendations. On the other hand, POI owners can leverage POI recommendations for better targeted acquisition of customers. In this paper we address exactly the problem of POI recommendation. We first identify the key challenges specific to geographical settings. Then, we propose a general framework to address these, as well as two instantiations of this framework.

Challenges. While latent factor models, such as matrix factorization [19], probabilistic matrix factorization (PMF) [27], [28], and many other variants [1], [3], [17], [18], [22], [36], have been proved effective and are widely used in diverse recommendation settings, adapting them to POI recommendations requires delicate consideration of unique characteristics of LBSNs. Indeed, there are several characteristics of LBSNs which distinguish POI recommendation from traditional recommendation tasks (such as movie or music recommendations). More specifically:

- Geographical influence. Due to geographical constraints and the cost of traveling large distances, the probability of a user visiting a POI is inversely proportional to the geographic distance between them.
- Tobler's first law of geography. The law of geography states that "Everything is related to everything else, but near things are more related than distant things" [32]. In other words, geographically proximate POIs are more likely to share similar characteristics.
- User mobility. Users may check into POIs at different regions; e.g., an LBSN user may travel to different cities. Varying user mobility imposes huge challenges on POI recommendations, especially when a user arrives at a new city or region.
- Implicit user feedback. In the study of POI recommendations, explicit user ratings are usually not available. The recommender system has to infer user preferences from implicit user feedback (e.g., check-in frequency).

The first three mutually related challenges due to geography imply interrelationships among items. However, traditional recommender systems usually ignore these, assuming that the items are independent and identically

• *The authors are with the Department of Management Science and Information Systems, Rutgers Business School, Rutgers University, Piscataway, NJ 08854.*
E-mail: {binben.liu, hxiong, s.papadim, yanjie.fu, zijun.yao}@rutgers.edu.
Manuscript received 4 May 2014; revised 22 Sept. 2014; accepted 24 Sept. 2014. Date of publication 8 Oct. 2014; date of current version 27 Mar. 2015. Recommended for acceptance by J. Wang.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TKDE.2014.2362525

distributed. In fact, the decision process of a user choosing a POI is complex and can be influenced by many factors. First, geographical distance plays an important role. According to the Tobler's first law of geography and the law of demand, a user's propensity for a POI is inversely proportional to the distance between them. This is similar to the observation that the probability of purchasing an item is inversely proportional to its cost. Second, utility matters. In economics, utility is an index of preferences over sets of items and services when a user makes purchasing decisions. In other words, a user may still prefer a remote POI to a nearby one, if higher satisfaction (utility) outweighs the overhead of travel. Finally, LBSN users have varying mobility behaviors, which further impose challenges on modeling check-in decisions.

An additional fourth challenge is that user check-in counts follow a distribution with power-law form. This is different from ratings in traditional recommender systems, in which explicit ratings are available to reflect users' item preferences. In other words, in LBSNs a user can visit a POI only once and another POI hundreds of times. Since we do not have explicit user ratings for POIs, we can only make use of implicit user behavior data in the check-in records for POI recommendations.

POI recommendation framework. All the above challenges demand a reconsideration of the recommendation model, to achieve effective POI recommendation in LBSNs. While there are some studies on POI recommendations, they lack an integrated analysis of the joint effect of the above factors, such as user preferences, geographical influences and user mobility behaviors.

To address these challenges, we propose a framework for geographical probabilistic factor modeling (Geo-PFM) which can strategically take various factors into consideration. This framework can capture the geographical influences on a user's check-in behaviors, can effectively model the user mobility patterns, and can deal with the skewed distribution of check-in count data. Specifically, we introduce a latent region variable and use a multinomial distribution over latent regions to model user mobility behaviors over different activity regions. These latent regions reflect the activity areas for all the users through collective actions. A Gaussian distribution is used to represent a POI over a sampled region. This can reflect the first law of geography; that is, similar POIs are more related than distant POIs. Moreover, geographical influence can be effectively modeled in the latent region. Finally, implicit user feedback in the form check-in counts is taken into account.

In our earlier work [23], we introduced Geo-PFM by specifically instantiating a geographical Bayesian non-negative matrix factorization(Geo-BNMF), to model user preferences. As a result, this model is capable of taking personal preferences, geographical influence, and user mobility into consideration, and can effectively handle the skewed distribution of POI count data.

In this paper, we study the Geo-PFM framework in more detail and we further develop a Poisson Geo-PFM, which is also able to capture the geographical influences on a user's check-in behavior and effectively model the user mobility patterns. In addition, the Poisson Geo-PFM provides much more flexibility and interpretability than Geo-PFM based on non-negative matrix factorization [23]. First, the Poisson Geo-PFM provides a rigorous probabilistic generative process for the model, while the NMF-based Geo-PFM uses an approximation solution. Second, the nature of Poisson distribution is more suitable and effective for modeling the skewed user check-in count data, which provide implicit feedback, for better POI recommendations.

Finally, we provide extensive experimental results on three real-world LBSNs data, which differ in terms of user mobilities, POI geographical distributions, implicit response data skewness and user-POI observation sparsity. The experimental results show that the proposed POI recommendation method consistently outperforms state-of-the-art probabilistic latent factor models with a significant margin in terms of Top-N recommendation. Moreover, the proposed Poisson Geo-PFM outperforms Geo-BNMF [23] even further.

2 Background

Latent factors models aim to characterize user-item interactions assuming that each user and each item can be expressed as a user and item latent vector $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ respectively. Consequently, the *response* (rating, like, or implicit frequency) is modeled as $p(y_{ij} \mid i, j) = p\big(y_{ij} \mid \boldsymbol{u}_i^\top \boldsymbol{v}_j; \Theta\big)$. In this section we summarize two types of latent factor models: probabilistic matrix factorization methods which are widely used for recommendations when explicit user feedback (e.g., item ratings) is available, and the Poisson factor model (PoiFM) which is more effective when user feedback is implicitly provided via heavily skewed frequency counts (as in the case of POI recommendation).

2.1 Probabilistic Matrix Factorization

Matrix factorization models [19] have been generalized into probabilistic matrix factorization [28], which is a Bayesian version. In PMF the response y_{ij} of user u_i for item v_j is assumed to follow a Gaussian distribution $y_{ij} \sim \mathcal{N}(y_{ij} \mid \boldsymbol{u}_i^\top \boldsymbol{v}_j, \sigma^2)$. When response y_{ij} is not normalized to a standard rating score, one solution is to scale the discrete response to a value between $(0, 1]$ by using $f(x) = (x-1)/(x_{max}-1)$, where x_{max} is the maximum response value for each user [28]. Furthermore, a zero-mean Gaussian prior is placed toon the user and item latent spaces

$$P\big(U \mid \sigma_u^2\big) = \prod_{i=1}^{M} \mathcal{N}\big(\boldsymbol{u}_i \mid 0, \sigma_u^2 \mathbf{I}\big),$$

$$P\big(V \mid \sigma_v^2\big) = \prod_{j=1}^{N} \mathcal{N}\big(\boldsymbol{v}_j \mid 0, \sigma_v^2 \mathbf{I}\big).$$

Then the latent factors $\boldsymbol{u}$ and $\boldsymbol{v}$ can be inferred by maximize thing likelihood over the observed ratings

$$P(Y \mid U, V, \sigma^2) = \prod_{i=1}^{M} \prod_{j=1}^{N} \big[\mathcal{N}\big(y_{ij} \mid \boldsymbol{u}_i^\top \boldsymbol{v}_j, \sigma^2\big)\big]^{I_{ij}}, \tag{1}$$

where I_{ij} is the indicator function. Maximizing the log-posterior over user and item latent factors with

hyperparameters is equivalent to minimizing the sum-of-squared-errors objective function:

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}\big(y_{ij} - \boldsymbol{u}_i^\top \boldsymbol{v}_j\big)^2 + \frac{\lambda_U}{2}\sum_{i=1}^{M}||\boldsymbol{u}_i||_F^2 + \frac{\lambda_V}{2}\sum_{j=1}^{N}||\boldsymbol{v}_j||_F^2, \tag{2}$$

where $\lambda_U = \sigma^2/\sigma_u^2$, $\lambda_V = \sigma^2/\sigma_v^2$, and $||\cdot||_F^2$ is the Frobenius norm. Gradient descent can be applied to infer the latent factors with partial derivatives $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ respectively,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{u}_i} = -\sum_{j=1}^{N} I_{ij}\big(y_{ij} - \boldsymbol{u}_i^\top \boldsymbol{v}_j\big)\cdot \boldsymbol{v}_j + \lambda_U \boldsymbol{u}_i$$
$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_j} = -\sum_{i=1}^{M} I_{ij}\big(y_{ij} - \boldsymbol{u}_i^\top \boldsymbol{v}_j\big)\cdot \boldsymbol{u}_i + \lambda_V \boldsymbol{v}_j. \tag{3}$$

2.2 Poisson Factor Model

The Poisson distribution is a more appropriate choice for response variables y_{ij} that represent frequency counts. The Poisson probabilistic factor model (Poi-PFM) [6], [12], [26] factorizes the user-item count matrix Y as $Y \sim \text{Poisson}(UV)$. More specifically, for each user-item response y_{ij}, we assume a Poisson distribution over the mean f_{ij}: $y_{ij} \sim \text{Poisson}(f_{ij})$. The mean matrix F is factorized into two matrices $U_{M\times K}$ and $V_{N\times K}$. Each element $u_{ik} \in U$ encodes the preference of user i for "topic" k, and each element $v_{ik} \in V$ reflects the topical affinity of item j to topic k. Further, u_{ik} and v_{ik} can be assigned empirical priors following Gamma distributions. We then have the following generative process.

1. Generate user latent factor $u_{ik} \sim \text{Gamma}(\alpha_U, \beta_U)$.
2. Generate item latent factor $v_{jk} \sim \text{Gamma}(\alpha_V, \beta_V)$.
3. Generate $y_{ij} \sim \text{Poisson}(\boldsymbol{u}_i^\top \boldsymbol{v}_j)$.

Given user latent factor $\boldsymbol{u}_i$ and item latent factor $\boldsymbol{v}_j$, the probability of response y_{ij} is

$$P(y_{ij} \mid \boldsymbol{u}_i, \boldsymbol{v}_j) = \big(\boldsymbol{u}_i^\top \boldsymbol{v}_j\big)^{y_{ij}} \exp\big\{-\boldsymbol{u}_i^\top \boldsymbol{v}_j\big\}/y_{ij}!.$$

We can apply maximum a posteriori (MAP) estimation over the observed data and priors to infer the latent vectors. Specifically,

$$P(U, V \mid Y, \alpha_U, \beta_U, \alpha_V, \beta_V) \propto p(Y \mid U, V) P(U \mid \alpha_U, \beta_U) p(V \mid \alpha_V, \beta_V),$$

where

$$p(Y \mid U, V) = \prod_{i=1}^{M}\prod_{j=1}^{N} \big(\boldsymbol{u}_i^\top \boldsymbol{v}_j\big)^{y_{ij}} \exp\big\{-\boldsymbol{u}_i^\top \boldsymbol{v}_j\big\}/y_{ij}!$$

$$P(U \mid \alpha_U, \beta_U) = \prod_{i=1}^{M}\prod_{k=1}^{K} \frac{u_{ik}^{\alpha_U - 1}\exp(-u_{ik}/\beta_U)}{\beta_U^{\alpha_U}\Gamma(\alpha_U)}$$

$$p(V \mid \alpha_V, \beta_V) = \prod_{j=1}^{N}\prod_{k=1}^{K} \frac{u_{ik}^{\alpha_V - 1}\exp(-v_{jk}/\beta_V)}{\beta_V^{\alpha_V}\Gamma(\alpha_V)}.$$

The log of the posterior distribution over the user and item latent factors is given by

TABLE 1
Mathematical Notations

Symbol	Size	Description
R	$1 \times \lvert R\rvert$	latent region set, r is a region in R
$\boldsymbol{\eta}$	$M \times \lvert R\rvert$	user level region distribution
$\boldsymbol{\mu}$	$\mathbb{R}^2$	location mean of a latent region
$\boldsymbol{\Sigma}$	$\mathbb{R}^{2\times 2}$	covariance matrix of a latent region
U	$M \times K$	user latent factor
V	$N \times K$	item latent factor
y_{ij}	$\mathbb{R}$	response of user i for item j
l_j	$\mathbb{R}^2$	location of item j

$$\begin{aligned}\mathcal{L}(U, V, \mid \mathcal{D}, \alpha_U, \beta_U, \alpha_V, \beta_V) &= \sum_{i=1}^{M}\sum_{k=1}^{K}\big((\alpha_U - 1)\ln u_{ik} - u_{ik}/\beta_U\big) \\ &+ \sum_{j=1}^{N}\sum_{k=1}^{K}\big((\alpha_V - 1)\ln v_{jk} - v_{jk}/\beta_V\big) \\ &+ \sum_{i=1}^{M}\sum_{j=1}^{N}(y_{ij}\ln f_{ij} - f_{ij}) + \text{const.}\end{aligned} \tag{4}$$

Taking derivatives on $\mathcal{L}$ with respect to u_{ik} and u_{jk}, we have

$$\frac{\partial \mathcal{L}}{\partial u_{ik}} = \frac{\alpha_U - 1}{u_{ik}} - \frac{1}{\beta_U} + \sum_{j=1}^{N}\left(\frac{y_{ij}}{f_{ij}} - 1\right)v_{jk}$$
$$\frac{\partial \mathcal{L}}{\partial v_{jk}} = \frac{\alpha_V - 1}{v_{jk}} - \frac{1}{\beta_V} + \sum_{i=1}^{M}\left(\frac{y_{ij}}{f_{ij}} - 1\right)u_{ik}. \tag{5}$$

Again, gradient ascent method can be applied to infer the latent factors.

3 Geographical Probabilistic Factor Model for POI Recommendation

In this section, we first formulate the problem of POI recommendation and then introduce a general geographical probabilistic factor analysis framework for this problem, addressing the challenges described previously.

3.1 Problem Definition

The problem of personalized POI recommendation is to recommend POIs to a user given user POI check-in records and other available side information. Let $U = \{u_1, u_2, \ldots, u_M\}$ be a set of LBSN users, where each user has a location l_i. The user location l_i is usually unknown due to user mobility. Let $V = \{v_1, v_2, \ldots, v_N\}$ be a set of POIs, where each POI has a location $l_j = [lon_j, lat_j]^\top$ represented by longitude and latitude. Throughout the paper we use indices i for users and indices j for POIs, unless stated otherwise. The number of times user u_i visited POI v_j is represented by the *response variable* y_{ij}. The check-in records for a particular user are sparse (most y_{ij} values are zero), with non-zeros following a power law. In the paper we use the terms "POI" and "item" interchangeably. Key notations are listed in Table 1.

(a) All POIs in different regions. (b) A user's check-ins in different regions. (c) User check-ins in San Francisco.

Fig. 1. An example of a typical user check-in pattern: (a) all the POIs; (b) the user's check-ins over different regions: San Francisco, Los Angeles, San Diego, Las Vegas, Houston, and New York City; (c) the user's check-ins in San Francisco area.

3.2 The General Idea

We aim to capture how different factors such as user preference, geographical influence and user mobility affect user POI check-in decisions. The key idea is that overall user preferences are the result of the interplay between geographical preferences and interest preferences. Our models aim to effectively capture that interplay.

Geographical preferences. To learn geographical user preferences, we need a model to encode the spatial influence and user mobility into the user check-in decision process. As shown in Fig. 1, LBSN users are most likely to check into a number of POIs and these POIs are usually limited to certain geographical regions. This observation has two implications: first, a user's mobility always happens across a limited number regions but these regions could be different among different users; second, user check-in activities happen in a given region and the activity patterns could be different given different regions. Based on this observation, we propose to introduce a set of $|R|$ latent regions R which are inferred based on the collective actions of *all* users, reflecting activity areas for the entire population.

Although the overall distribution of POIs is irregular, we can however assume a Gaussian geographical distribution of POIs *within* each region $r \in R$. The location l_j for POI j is characterized by $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$, where μ_r and Σ_r are the mean vector and covariance matrix of the region, respectively [14], [35]. Latent regions also reflect Tobler's first law of geography, which states that POIs with similar characteristics are likely to be clustered into the same geographical area. Once a region is fixed, geographical influence can be effectively modeled and applied to overall user preference profiling.

We finally model individual user mobility over the collectively inferred latent regions R by applying a multinomial distribution, $r \sim p(r \mid \eta_i)$, where η_i is a user-dependent distribution over latent regions for user i.

Interest preferences. Interest preferences are modeled using a latent factor model, generating a user item preference $\alpha(i,j)$ based on user latent factor variable $\boldsymbol{u}_i$ and and item latent factor variable $\boldsymbol{v}_j$.

Overall user preferences. Finally, to model a user's propensity for a POI, we assume the following factors that will affect the overall user check-in decision process: (1) the personal preference $\alpha(i,j)$ of each user i with respect to POI j; and (2) geographical influence in terms of travel distance, namely, the distance $d(i,j)$ between the user and the POI as a geographical cost. As a result, the probability of observing a user-POI pair (i,j) is directly proportional to the user interest, and monotonically decreases with the distance between them,

$$p(i,j) \propto \mathbb{F}\left(\alpha(i,j), \left[\frac{d_0}{d_0 + d(i,j)}\right]^{\tau}\right),$$

where $\mathbb{F}(\cdot)$ is a function that combines user interest preference and geographical influence. We model the distance factor in the decision making process using a parametric term $[\frac{d_0}{d_0+d(i,j)}]^{\tau}$ with a power-law form. This motivated by the observation that the probability of user i choosing POI j decays exponentially with respect to the distance between them.

3.3 Geographical Probabilistic Factor Model Framework

Based on above discussion, we proposed a geographical probabilistic factor model to capture user mobility, and geographical influence in user profiling for POI recommendation. The complete graphical model is shown in Fig. 2.

The corresponding generative process to draw pairs (i,j) representing user i choosing POI j can then be expressed as follows. First, a user u_i samples a region r_i from all $|R|$ regions following a multinomial distribution $r_i \sim \text{Multinomial}(\eta_i)$, on which a conjugate Dirichlet prior $\text{Dir}(\boldsymbol{\gamma})$ can be further imposed. Here η_i is a user-dependent parameter, capturing user i's mobility pattern over the latent

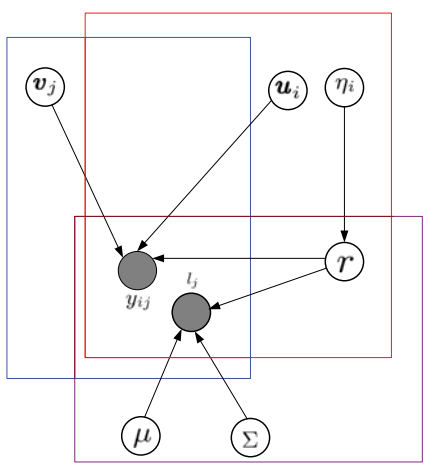


Fig. 2. A graphical representation of the proposed geographical probabilistic factor model, where the red plate represents users, the blue plate represents POIs, and the purple plate represents latent regions. The model priors have been excluded for simplicity.

regions. A POI is drawn from the sampled region $l_j \sim \mathcal{N}(\mu_{r_i}, \Sigma_{r_i})$. The interest preference $\alpha(i,j)$ of user i for POI j can be represented by combining latent factors $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$, specifically, $\alpha(i,j) = \boldsymbol{u}_i^\top \boldsymbol{v}_j$. Finally, the user-POI response y_{ij} (check-in frequency count) is assumed to follow certain distribution $y_{ij} \sim P(f_{ij})$ where f_{ij} depends on user preferences and the distance between the user and the POI. Summarizing:

1. Draw a geographical preference
   a. Draw region $r_i \sim \text{Multinomial}(\eta_i)$.
   b. Draw a POI j with location $l_j \sim \mathcal{N}(\mu_{r_i}, \Sigma_{r_i})$.
2. Draw an interest preference
   a. Draw user latent factor $\boldsymbol{u}_i \sim P(\boldsymbol{u}_i; \Psi_{\boldsymbol{u}_i})$.
   b. Draw item latent factor $\boldsymbol{v}_j \sim P(\boldsymbol{v}_j; \Psi_{\boldsymbol{v}_j})$.
   c. Draw user-item preference $\alpha(i,j) = \boldsymbol{u}_i^\top \boldsymbol{v}_j$.
3. For each user-POI pair (i,j) draw the response $y_{ij} \sim P(f_{ij})$, where

$$f_{ij} = \mathbb{F}\left(\boldsymbol{u}_i^\top \boldsymbol{v}_j, \left[\frac{d_0}{d_0 + d(i,j)}\right]^\tau\right).$$

Note that the proposed model is general and can be extended with different factor models, since we limit neither the user and item latent factor distribution, nor the user-item response distribution. $\mathbb{F}(\cdot)$ is a function of personalized preferences $\boldsymbol{u}_i^\top \boldsymbol{v}_j$ and of distance cost $[\frac{d_0}{d_0+d(i,j)}]^\tau$. User-item response $y_{ij} \sim P(f_{ij})$ can be: (i) Gaussian when explicit ratings are available, (ii) Bernoulli for binary response such as *liking*, or (iii) Poisson when count or frequency data is to be modeled.

3.4 Model Components

This section describes the model components of Geo-PFM in detail.

3.4.1 User Mobility and Geographical Influence

As discussed earlier, user mobility and geographical influence are among the most predominant factors that distinguish POI recommendation from traditional recommendation (e.g., for movies), and these two factors can interact with each other. Geographical influence has been exploited for POI recommendation due to the fact that geographical proximity could significantly affect a user's check-in decision [34]. However, check-in behavior can change as the user travels from one region to another, and little has been done to consider user mobility for POI recommendation. Capturing user mobility is important to understand user preferences in different regions, and it becomes even more important when a user travels to a new place.

To this end, as described earlier, we introduce a set of $|R|$ latent regions R, and model user mobility using multinomial distribution [14] $r \sim \text{Multinomial}(\eta_i)$, where η_i is a user-dependent distribution over latent regions for user i. The explicit location $\ell(\cdot)$ of a user is not observed. We use the region r with center μ_r to represent the user activity area and model the geographical influence as a parametric and power-law like term $[\frac{d_0}{d_0+d(i,j)}]^\tau$, with $d(i,j) = ||\mu_r - l_j||_2$, where μ_r approximates the current user activity area center. As a result, both user mobility and geographical influence can be effectively captured by the proposed Geo-PFM model.

3.4.2 Modeling Count Response

In most existing latent factor models, represented by PMF [28], the response $P(y_{ij} \mid \boldsymbol{u}_i^\top \boldsymbol{v}_j; \Theta)$ is assumed to follow a Gaussian distribution, namely, $y_{ij} \sim \mathcal{N}(\boldsymbol{u}_i^\top \boldsymbol{v}_j, \sigma^2)$. However, a Gaussian distribution is not suitable when the response variable is implicit count data, which are heavily skewed. Therefore, it is not suitable for the POI recommendation problem, since check-in counts follow a power-law like distribution.

We need to ensure our model is suitable for count responses. By combining geographical influence with latent factors, we model user-POI response as a geographical probabilistic factor model:

$$y_{ij} \sim P(y_{ij} \mid f_{ij}, \Theta), f_{ij} = \mathbb{F}\left(\boldsymbol{u}_i^\top \boldsymbol{v}_j, \left[\tfrac{d_0}{d_0+d(i,j)}\right]^\tau\right).$$

In the above, $\mathbb{F}(\cdot)$ is a suitably chosen function that captures the joint effect of personal interest preferences $\boldsymbol{u}_i^\top \boldsymbol{v}_j$ and distance cost $[\frac{d_0}{d_0+d(i,j)}]^\tau$. Also, the response function $P(\cdot)$ suitably chosen to model count data. Potential response function distributions include Poisson (see Section 4).

4 Model Specification

This section introduces detailed model specifications of the Geo-PFM model. In particular, we introduce a Poisson Geo-PFM model, which takes into account the characteristics of count response values.

4.1 Poisson Geo-PFM Model

As we use count response to infer user preferences, we expect the latent vectors are constrained to be non-negative. In our earlier work [23], we applied a rectified normal Bayesian non-negative matrix factorization model to capture the count response feature. Specifically, we assumed a rectified normal distribution on $Y \sim P(\boldsymbol{UV})$ with variance $\sigma^2\mathbf{I}$ and non-negativity constraints,

$$Y \sim \mathcal{N}^R(Y \mid \boldsymbol{UV}, \sigma^2\mathbf{I}), \quad \text{subject to } \boldsymbol{U} \geq 0, \boldsymbol{V} \geq 0. \tag{6}$$

We further placed an exponential distribution on $\boldsymbol{U}$ and $\boldsymbol{V}$, and an inverse gamma distribution on σ^2 with shape a and scale b.

However, a Poisson factor model is a better alternative. First, the Poisson distribution is a more appropriate choice for modeling skewed count data. Fig. 3 shows a typical distribution of check-in count distribution, for a randomly selected user in the Foursquare dataset. A Poisson distribution approximates this distribution well, and can also provide a response that is non-negative. More importantly, a Poisson factor guarantees a rigorous probabilistic generative process for the model, while the rectified normal BNMF provides a probabilistic approximation. Therefore we propose a Poisson Geo-PFM model which incorporates both user interest preference and geographical influence. More specifically, for each user-item frequency y_{ij} we assume a Poisson distribution over mean f_{ij}: $y_{ij} \sim \text{Poisson}(f_{ij})$ with

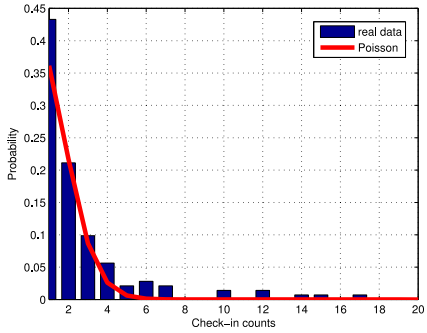


Fig. 3. The check-in counts distribution of a randomly selected user and a Poisson approximation of this distribution (Foursquare dataset).

$f_{ij} = \boldsymbol{u}_i^\top \boldsymbol{v}_j \cdot [\frac{d_0}{d_0+d(i,j)}]^\tau$. Furthermore, u_{ik} and v_{ik} are given Gamma distributions as empirical priors [6], [26], $u_{ik} \sim \text{Gamma}(\alpha_U, \beta_U)$ and $v_{jk} \sim \text{Gamma}(\alpha_V, \beta_V)$. Then, the generative process we introduced earlier to model user-item preference becomes specifically:

1. Draw a region $r \sim \text{Multinomial}(\eta_i)$.
2. Draw a POI j with location $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$.
3. Draw user latent factor $u_{ik} \sim \text{Gamma}(\alpha_U, \beta_U)$.
4. Draw item latent factor $v_{jk} \sim \text{Gamma}(\alpha_V, \beta_V)$.
5. Draw $y_{ij} \sim \text{Poisson}(f_{ij})$ with

$$f_{ij} = \boldsymbol{u}_i^\top \boldsymbol{v}_j \cdot \left[\frac{d_0}{d_0 + d(i,j)}\right]^\tau.$$

4.2 Parameter Estimation

Let $\Psi = \{U, V, \boldsymbol{\eta}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ denote all parameters, and let $\Omega = \{\alpha_U, \beta_U, \alpha_V, \beta_V, \gamma\}$ be the hyperparamters. We are given the observed data collection $\mathcal{D} = \{y_{ij}, l_j\}^{I_{ij}}$ where y_{ij} is the user check-in count and l_j is the location of v_j; and I_{ij} is the indicator function with $I_{ij} = 1$ when user u_i visited POI v_j, and $I_{ij} = 0$ otherwise. Then we aim to maximize the posterior probability given the observed data:

$$\begin{aligned}
P(\Psi; \mathcal{D}, \Omega) &\propto \prod_{\mathcal{D}} P(y_{ij}, l_j \,|\, \Psi, \Omega) P(\Psi \,|\, \Omega) \\
&\propto \prod_{\mathcal{D}} P(y_{ij}, l_j, \Psi \,|\, \Omega) P(U \,|\, \alpha, \beta) P(V \,|\, \alpha, \beta) P(\eta \,|\, \gamma) \\
&\propto \prod_{i=1}^{M} \left\{ \prod_{j=1}^{N_i} |\boldsymbol{\Sigma}_r|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(l_j - \boldsymbol{\mu}_r)^T \boldsymbol{\Sigma}_r^{-1} (l_j - \boldsymbol{\mu}_r)\right) \right. \\
&\quad \left. \frac{f_{ij}^{\,y_{ij}} \exp(-f_{ij})}{y_{ij}!} \right\} \times \eta_{i1}^{c_{i1}} \cdots \eta_{iR}^{c_{iR}} \times \prod_{i=1}^{M} \prod_{k=1}^{K} u_{ik}^{\alpha-1} \exp\left(\frac{-u_{ik}}{\beta}\right) \\
&\quad \times \prod_{j=1}^{N} \prod_{k=1}^{K} v_{jk}^{\alpha-1} \exp\left(\frac{-u_{jk}}{\beta}\right) \times \prod_{i=1}^{M} \prod_{r=1}^{R} \eta_{ir}^{\gamma_i - 1}.
\end{aligned}$$

To estimate the parameters Ψ, we use a mixing Expectation Maximization (EM) and sampling algorithm to learn all the parameters [2], [14]. We regions r as a latent variable and introduce the hidden variable $P(r \,|\, l_j, \Psi)$ [14], [35], which is the probability of $l_j \in r$, given POI location l_j and Ψ. The algorithm iteratively updates the parameters by mutual enhancement between Geo-clustering and Geo-PFM. The Geo-clustering updates the latent regions based on both location and check-in behaviors; and Geo-PFM learns the graphical preference factors.

4.2.1 E-step

In the **E-step**, we iteratively draw latent region assignments for all POIs. For each POI, a latent region r is first drawn from the following distribution:

$$r \sim P\big(r \,|\, \{y_{\cdot j}, l_j\}, R^{(t)}, \Psi^{(t)}\big) \times P(r \,|\, \boldsymbol{\eta}^{(t)}), \tag{7}$$

where

$$P\big(\{y_{\cdot j}, l_j\} \,|\, r, \Psi^{(t)}\big) = P(l_j \,|\, r, \Psi^{(t)}) \times P(y_{\cdot j} \,|\, r, \Psi^{(t)})$$

$$P(l_j \,|\, r, \Psi^{(t)}) = \mathcal{N}(l_j \,|\, \mu_r^{(t)}, \Sigma_r^{(t)})$$

$$P(y_{\cdot j} \,|\, r, \Psi^{(t)}) = \prod_{i=1}^{M} \text{Poission}(f_{ij} \,|\, \boldsymbol{U}^{(t)}, \boldsymbol{V}^{(t)}).$$

$P(r \,|\, \boldsymbol{\eta}^{(t)})$ updates region assignment in terms of user mobility, $P(l_j \,|\, r, \Psi^{(t)})$ is the location PDF function for multivariate normal distribution with mean vector and variance matrix obtained in last iteration, and $P(y_{\cdot j} \,|\, r, \Psi^{(t)})$ updates region assignment through collective actions.

4.2.2 M-step

In the **M-step**, we maximize the log likelihood of the model with respect to model parameters by fixing all regions obtained in the E-step. Since we sample the regions in the E-step, we can update $\boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r, \boldsymbol{\eta}$ directly from the samples,

$$\begin{aligned}
\boldsymbol{\mu}_r &= \frac{1}{\#(j,r)} \sum_{j=1}^{\mathcal{D}} \mathbb{I}(r_j = r) l_j \\
\boldsymbol{\Sigma}_r &= \frac{1}{\#(j,r) - 1} \sum_{j=1}^{\mathcal{D}} ((l_j - \mu_r)(l_j - \mu_r)^\top)
\end{aligned} \tag{8}$$

where $\#(j,r)$ is the number of POIs assigned to region r. Through imposing a conjugate Dirichlet prior $\text{Dir}(\boldsymbol{\gamma})$, we update $\boldsymbol{\eta}^{(t+1)}$ by

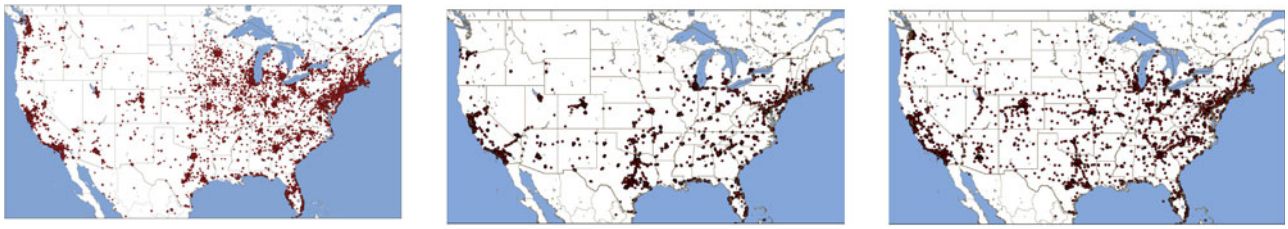
$$\eta_{ir}^{(t+1)} = \frac{C_{ir}^{(t+1)} + \gamma}{C_{i\cdot}^{(t+1)} + R\gamma}, \tag{9}$$

where C_{ir} is the number of POIs being assigned to region r for user i, and $C_{i\cdot}$ is the number of all POIs and all regions for user i.

After updating region $R^{(t+1)}$, we update $\Psi^{(t+1)}$ by maximizing the posterior with respect to latent factors $\boldsymbol{u}$ and $\boldsymbol{v}$. We use a gradient ascent method to find $\Psi^{(t+1)}$ that maximizes the posterior. Note that we already update R as $R^{(t+1)}$, and we here need to maximize the posterior with respect to latent factor variables $\boldsymbol{u}$ and $\boldsymbol{v}$. More specifically, we maximize the following objective function:

$$\begin{aligned}
\mathcal{L}(U, V \,|\, R^{(t+1)}) = &\sum_{i=1}^{M} \sum_{k=1}^{K} ((\alpha_U - 1) \ln u_{ik} - u_{ik}/\beta_U) \\
&+ \sum_{j=1}^{N} \sum_{k=1}^{K} ((\alpha_V - 1) \ln v_{jk} - v_{jk}/\beta_V) \\
&+ \sum_{i=1}^{M} \sum_{j=1}^{N} (y_{ij} \ln f_{ij} - f_{ij}) + \text{const.}
\end{aligned} \tag{10}$$

where $f_{ij} = \boldsymbol{u}_i^\top \boldsymbol{v}_j \cdot [\frac{d_0}{d_0+d(i,j)}]^\tau$.



(a) Foursquare. (b) Gowalla. (c) Brightkite.

Fig. 4. POI geographical distribution for the three different datasets.

Taking derivatives on $\mathcal{L}$ with respect to u_{ik} and v_{jk}, we have

$$\frac{\partial \mathcal{L}}{\partial u_{ik}} = \frac{\alpha_U - 1}{u_{ik}} - \frac{1}{\beta_U} + \sum_{j=1}^{N} \left(\frac{y_{ij}}{f_{ij}} - 1 \right) v_{jk} \left[\frac{d_0}{d_0 + d(i,j)} \right]^{\tau}$$
$$\frac{\partial \mathcal{L}}{\partial v_{jk}} = \frac{\alpha_V - 1}{v_{jk}} - \frac{1}{\beta_V} + \sum_{i=1}^{M} \left(\frac{y_{ij}}{f_{ij}} - 1 \right) u_{ik} \left[\frac{d_0}{d_0 + d(i,j)} \right]^{\tau}. \quad (11)$$

We use stochastic gradient ascent to update u_{ik} and u_{ik}. Stochastic gradient ascent (descent) have been widely used for many machine learning tasks [4]. The main process involves randomly scanning all training instances and iteratively updating parameters,

$$u_{ik} \leftarrow u_{ik} + \epsilon \times \frac{\partial \mathcal{L}}{\partial u_{ik}}, v_{jk} \leftarrow v_{jk} + \epsilon \times \frac{\partial \mathcal{L}}{\partial v_{jk}}, \quad (12)$$

where ϵ is a learning rate.

Remark. The region R is updated in each E-step. The latent factor model parameters are updated based on the new regions. We summarize the parameter estimation procedure for Geo-PFM in Algorithm 1.

Algorithm 1. Geo-PFM Estimation

```
Initialize region partition R^(0) by k-means (k = R)
for t ← 1 to N_iteration do
    Update region R^(t) according to Equ. (7)
    Update region mean μ_r^(t) and covariance Σ_r^(t) according to Eq. (8)
    Update user region preference distribution η_ir^(t) according to Eq. (9)
    Update u_ik^(t), v_jk^(t) by stochastic ascent
    while not converge do
        ε^nIter := ε ν/(ν+nIter−1) //annealing learn rate
        for each random {i, j} pair do
            for k ← 1 to K do
                u_ik ← u_ik + ε^nIter × ∂L/∂u_ik
                v_jk ← v_jk + ε^nIter × ∂L/∂v_jk
            end for
        end for
    end while
end for
```

4.3 Recommendation

After parameters Ψ are learned, the Geo-PFM model predicts the check-in counts of a user for a given POI as $\mathbb{E}(y_{ij}|u_i, v_j) = \boldsymbol{u}_i^{\top} \boldsymbol{v}_j \times [\frac{d_0}{d_0 + d(i,j)}]^{\tau}$. We make recommendations based on the predicted check-ins as well as the user mobility. One way to combine the predicted value and user mobility is $\hat{y}_{ij} = \mathbb{E}(y_{ij}|u_i, v_j) \times \eta_{ir}$ with $j \in r$, the larger the predicted value, the more likely the user will choose this POI.

5 Experimental Results

In this section we empirically evaluate the performance of our proposed methods. All experiments were performed on three real-world LBSN datasets, collected from Foursquare (one of the most popular LBSN communities), Gowalla, and Brightkite.

5.1 Datasets

Foursquare dataset. The Foursquare dataset is formulated as follows [8], [9]: Foursquare users usually report their check-ins at POIs via Twitter. When an LBSN user posts a Tweet check-in at a POI, we consider it as evidence that the user has physically checked into the POI. The dataset includes POIs across the Unites States (except Hawaii and Alaska), and the geographical distribution of all POIs is shown in Fig. 4a. According to the Twitter reports from Foursquare users, we finalized a dataset of 12,422 users for 46,194 POIs with 738,445 check-in observations. The user POI check-in count matrix has a sparsity of 99.87 percent; it is very sparse. Each user checked into 59.44 POIs on average, only a very small fraction of all the POIs. The number of check-ins for a POI ranges from 1 to 786. This range is very wide as shown in Fig. 5, in which the user check-in count of a randomly chosen user is plotted.

Gowalla dataset. Besides the Foursquare dataset, we also evaluate the proposed models on Gowalla [10]. In this dataset, we remove those POIs with less than 10 users, and remove users with less than 30 user-POI pairs. We finalize a dataset of 7,070 users for 30,755 POIs with 520,950 check-in observations. The user POI check-in count matrix has a

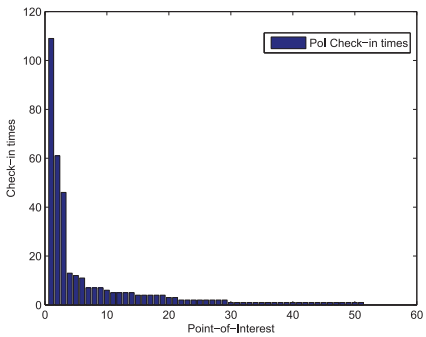


Fig. 5. An example of wide range user check-in counts for a randomly chosen user (Foursquare).

TABLE 2
Data Description

	# users	# POIs	# records	sparsity	avg POIs
Foursquare	12,422	46,194	738,445	99.87%	59.44
Gowalla	7,070	30,755	520,950	99.76%	73.68
Brightkite	2,192	9,865	72,543	99.66%	33.09

sparsity of 99.76 percent, with each user checked into 73.68 POIs on average. The number of check-ins for a POI ranges from 1 to 286, and the geographical distribution of all Gowalla POIs is shown in Fig. 4b.

Brightkite dataset. Finally, we evaluate the proposed models on Brightkite [10]. We finalize a dataset of 2,192 users and 9,865 POIs with 72,543 check-in observations. The user POI check-in count matrix has a sparsity of 99.66 percent, with each user checked into 33.09 POIs on average. The number of check-ins for a POI ranges from 1 up to more than one thousand, and the geographical distribution of all Brightkite POIs is shown in Fig. 4c. We summarize the data statistics for all datasets in Table 2.

5.2 Evaluation Metrics

Since there is no explicit rating for validation, we evaluate the models in terms of ranking. We present each user with N POIs sorted by the predicted values and evaluate based on which of these POIs were actually visited by the user.

Precision and recall. Given a top-N recommendation list $S_{N,\text{rec}}$ sorted in descending order of the prediction values, precision and recall are defined as

$$\begin{aligned} \text{Precision@}N &= \frac{|S_{N,\text{rec}} \bigcap S_{\text{visited}}|}{N} \\ \text{Recall@}N &= \frac{|S_{N,\text{rec}} \bigcap S_{\text{visited}}|}{|S_{\text{visited}}|}, \end{aligned} \tag{13}$$

where S_{visited} are the POIs a user has visited in the test data. The precision and recall for the entire recommender system are computed by averaging all the precision and recall values of all the users, respectively.

F-measure. F-measure combines precision and recall, and is the harmonic mean of precision and recall. Here we use the F_β measure with $\beta = 0.5$,

$$F_\beta = (1+\beta^2) \cdot \frac{\text{Precision} \times \text{Recall}}{\beta^2 \cdot \text{Precision} + \text{Recall}}. \tag{14}$$

The F_β measure with $\beta < 1$ indicates more emphasis on precision than recall.

5.3 The Method for Comparison

We experimentally compare our proposed Poisson Geo-PFM[1] model with state-of-the-art latent factor models. Specifically, we compare our proposed Poisson Geo-PFM model with following algorithms:

- *Probabilistic Matrix Factorization* [28]. PMF is a recommendation method widely used for different recommendation tasks, and the details of PFM are summarized in Section 2.1.
- *Bayesian Non-negative Factorization (BNMF)* [30]. This is the base model which our earlier work [23] adopted.
- *Poisson Factor Model* [26]. Poisson factor model provides an alternative for count data recommendation as Poisson is effective in modeling count data (more details in Section 2.2).
- *Fused Poisson factor model* (Fu-*PoiFM*). this method fuses the geographical influence into factor models by considering the multi-region of user check-in behaviors and the inverse distance in an *ad hoc* way [7]. Since Poison factor model also exploits the count check-in characteristics, we fuse the geographical influence into PoiFM and denote it as Fu-PoiFM.
- Geo-*BNMF*. This is the model we used in our earlier work [23].

In particular, we are interested in investigating the following questions:

- How the proposed Geo-PFM improves the non-geographical baseline models (PMF, BNMF, PoiFM) as well as the fused model (Fu-PoiFM).
- How the Poisson based model Geo-PFM improves its counterpart based on non-negative factorization, Geo-BNMF.

We randomly divided the data into 80 percent for training and 20 percent for testing. We set $\lambda_U = 0.005$ and $\lambda_V = 0.005$ for PMF. For Poisson factor based models used in this experiment, we set $\alpha_U = 5, \alpha_V = 20$ and $\beta_U = \beta_V = 0.2$. We set $1/\mathbb{R}$ for user region multinomial prior γ. We set $\tau = 1$ and $d_0 = 0.2$ for the distance model $[\frac{d_0}{d_0 + d(i,j)}]^\tau$. For Geo-PFM and Fu-PoisonFM, we first cluster all the POIs into $|R|$ regions. This is the initialization of the Geo-PFM model. We set the number of regions $|R| = 49$, which is the number of regions partitioned according to all the states in USA (except Hawaii and Alaska). All the latent factor models are implemented with stochastic gradient ascent/descent optimization method with an annealing procedure to discount learning rate ϵ at iteration nIter with $\epsilon^{\text{nIter}} := \epsilon \frac{\nu}{\nu + \text{nIter} - 1}$ by setting $\nu = 10$.

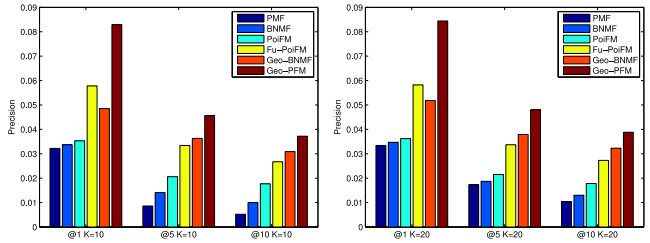
5.4 Performance Comparison[2]

In this section, we present the performance comparison on recommendation accuracy between our model and the baseline methods. We compare the results using both the Foursquare and the Gowalla dataset by setting latent dimensions to $K = 10$ and $K = 20$.

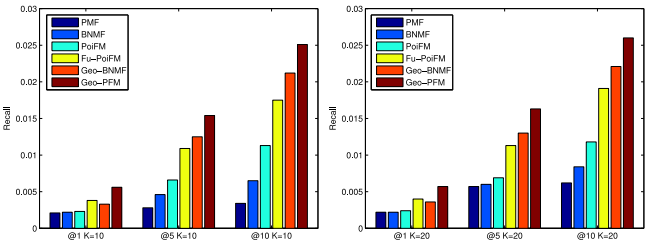
Foursquare dataset. Fig. 6 shows the precision and recall@N ($N = 1, 5, 10$) all the methods achieve on the Foursquare dataset, and Table 3 shows the F_β measure ($\beta = 0.5$). From the results, it is clear that the proposed Geo-PMF consistently outperforms all the baseline methods, including the non-geographical baseline models (PMF, BNMF, PoiFM) as well as the fused model (Fu-PoiFM). Specifically,

1. We will refer to Poisson Geo-PFM as Geo-PFM in this Section unless stated otherwise.

2. In the experiments of this paper, we rank the top-N recommendation globally, which is different from the regional way we used in [23]. Also we further tune some parameters. Therefore, the absolute experimental values in this paper may somewhat differ from those in [23].

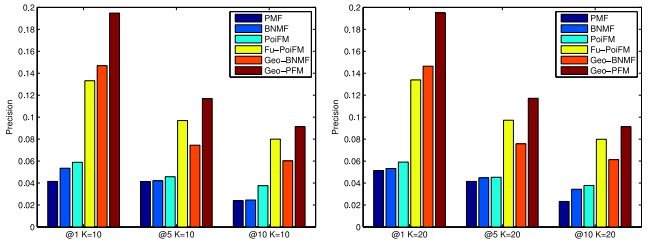


(a) Precision, K=10. (b) Precision, K=20.

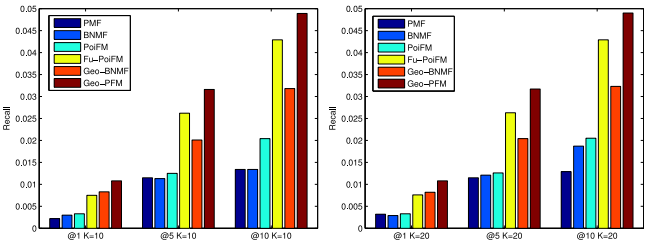


(c) Recall, K=10. (d) Recall, K=20.

Fig. 6. Precision and Recall with two different latent dimensions K (Foursquare dataset). Note that we focus on two comparisons: (1) How the proposed Geo-PFM improves the non-geographical baseline models (PMF, BNMF, PoiFM) as well as the fused model (Fu-PoiFM); (2) How the Poisson based model Geo-PFM improves its non-negative factorization based counterpart Geo-BNMF.



(a) Precision, K=10. (b) Precision, K=20.



(c) Recall, K=10. (d) Recall, K=20.

Fig. 7. Precision and Recall with two different latent dimensions K (Gowalla dataset).

TABLE 3
F_β Measure ($\beta = 0.5$) with Two Different Latent Dimensions K (Foursquare Dataset)

K	@N	PMF	BNMF	PoiFM	Fu-PoiFM	Geo-BNMF	Geo-PFM
10	@1	0.0083	0.0087	0.0091	0.0150	0.0130	**0.0220**
	@5	0.0061	0.0100	0.0145	0.0236	0.0263	**0.0328**
	@10	0.0047	0.0090	0.0159	0.0242	0.0339	**0.0339**
20	@1	0.0087	0.0088	0.0095	0.0157	0.0141	**0.0224**
	@5	0.0123	0.0131	0.0151	0.0241	0.0274	**0.0346**
	@10	0.0092	0.0117	0.0162	0.0251	0.0296	**0.0353**

TABLE 4
F_β Measure ($\beta = 0.5$) with Two Different Latent Dimensions K (Gowalla Dataset)

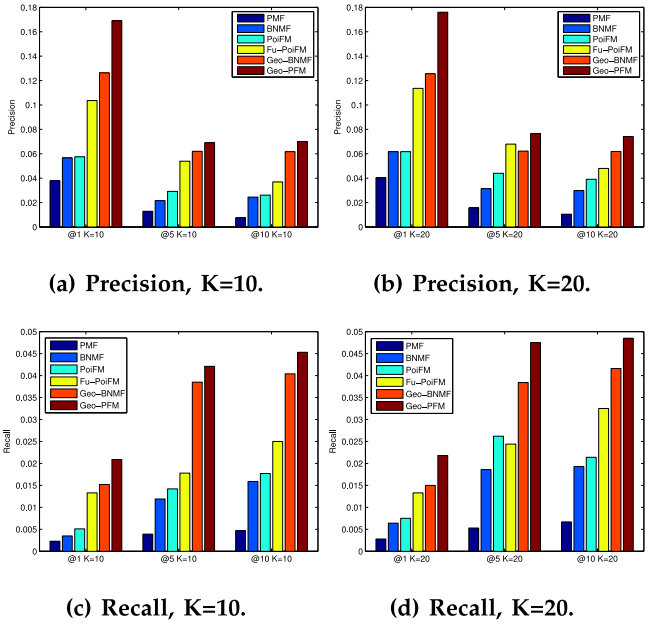
K	@N	PMF	BNMF	PoiFM	Fu-PoiFM	Geo-BNMF	Geo-PFM
10	@1	0.0091	0.0123	0.0135	0.0306	0.0338	**0.0442**
	@5	0.0272	0.0272	0.0298	0.0629	0.0483	**0.0759**
	@10	0.0207	0.0221	0.0322	0.0682	0.0551	**0.0778**
20	@1	0.0128	0.0119	0.0135	0.0310	0.0335	**0.0442**
	@5	0.0273	0.0290	0.0298	0.0632	0.0491	**0.0761**
	@10	0.0201	0.0295	0.0323	0.0681	0.0520	**0.0779**

nonnegative based Poison factor model and BNMF outperform PMF. Furthermore, PoiFM outperforms its counterpart BNMF by making Poisson assumptions. The fused method, Fu-PoiFM, improves PoiFM due to the fusion of geographical influence and multi-center user activity pattern into the latent factor model. Our proposed Geo-PFM further improves Fu-PoiFM significantly. From Table 3, we can observe an average of 0.0089 improvement in terms of F_β measure for Geo-PFM over Fu-PoiFM.

Meanwhile, from Fig. 6 we can see that the Poisson-based model Geo-PFM improves its counterpart based on non-negative factorization, Geo-BNMF, with an average of 0.0069 improvement in terms of F_β measure. This improvement can be ascribed to the following reasons. First, the Poisson-based latent factor is more appropriate for modeling count data. As shown, the improved performance of PoiFM over BNMF from Fig. 6, PoiFM can improve BNMF with an average of 0.0032 improvement in terms of F_β. Second, the Poisson Geo-PFM provides a more rigorous probabilistic generative process for the model, while the non-negative matrix factorization based Geo-PFM applied an approximation solution. As shown in the model estimation in Section 4.2, we need a rigorous probability for model inference. While the Poisson based model provides an exact probability representation, the Geo-BNMF applies a rectified normal distribution.

Gowalla dataset. Fig. 7 shows the precision and recall@N ($N = 1, 5, 10$) of all the methods evaluated on the Gowalla dataset, and the corresponding F_β measure values are shown in Table 4. We can clearly observe that the proposed Geo-PFM performs consistently better over all the baseline methods. From Table 4, we can observe an average of 0.0121 improvement in terms of F_β measure for Geo-PFM over Fu-PoiFM. We further observe that the Poisson-based Geo-PFM improves Geo-BNMF by an average of 0.0207 increase in terms of F_β measure.

Brightkite dataset. Fig. 8 shows the precision and recall@N ($N = 1, 5, 10$) of all the methods evaluated on the Gowalla dataset, and the corresponding F_β measure values are shown in Table 4. We can still observe consistent improvements of the proposed Geo-PFM over all the baseline methods. From Table 5, we can observe an average of 0.0246 improvement in terms of F_β measure for Geo-PFM over Fu-PoiFM. Again, we further observe that Poisson-based Geo-PFM improves Geo-BNMF with an average of 0.0129 increase in terms of F_β measure.



(a) Precision, K=10. (b) Precision, K=20.

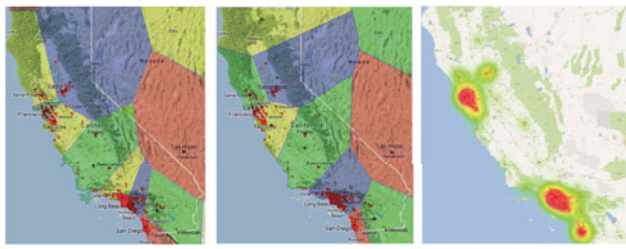
(c) Recall, K=10. (d) Recall, K=20.

Fig. 8. Precision and Recall with two different latent dimensions K (Brightkite dataset).

TABLE 5
F_β Measure ($\beta = 0.5$) with Two Different Latent Dimensions K (Brightkite Dataset)

K	@N	PMF	BNMF	PoiFM	Fu-PoiFM	Geo-BNMF	Geo-PFM
10	@1	0.0092	0.0140	0.0188	0.0439	0.0513	**0.0699**
	@5	0.0088	0.0186	0.0241	0.0383	0.0553	**0.0612**
	@10	0.0067	0.0221	0.0238	0.0337	0.0558	**0.0632**
20	@1	0.0110	0.0226	0.0252	0.0453	0.0508	**0.0729**
	@5	0.0112	0.0276	0.0387	0.0501	0.0553	**0.0682**
	@10	0.0094	0.0269	0.0336	0.0438	0.0564	**0.0670**

Comparisons across different datasets. First, we observed consistent improvements of the proposed Geo-PFM over all the baseline methods, though the three dataset differ in terms of user-POI observation sparsity, response skewness, and POI geographical distributions (see Fig. 4). Second, the Poisson-based Geo-PFM improves its counterpart based on non-negative factorization, Geo-BNMF. Third, user-POI observation sparsity, response skewness and POI geographical distributions could affect the algorithm performances. For example, the results on the Gowalla dataset and the Brightkite dataset are better than those on the Foursquare dataset. The Gowalla dataset is much denser than the Foursquare dataset. Note that Gowalla dataset has a sparsity of 99.76 percent, and an average of 59.44 user-POI observations; while the Foursquare dataset has a sparsity of 99.87 percent, an average of 73.68 user-POI observations. Although Brightkite dataset has fewer user-POI observations, on average, than Foursquare dataset, its sparsity is the lowest among the three datasets. Further, the Gowalla dataset is less skewed than the Foursquare dataset. These two factors could allow the latent factor models, both PMF and PoiFM, to achieve better performances. Also, the Gowalla dataset is more geographically centralized than the Foursquare dataset. As a result, the performances of Geo-PFM would be more obvious compared to Fu-PoiFM when applied to more geographically distributed circumstances.



(a) K-means. (b) Latent region. (c) Ground truth.

Fig. 9. Voronoi visualization of POI segmentation in California area (Foursquare): (b) latent regions learned by Geo-PFM, (a) initiation by K-means, and (c) true user collaborative activity clusters. Deeper color (red) indicates more check-ins for a POI, as contrary to light color (green). Best view in color.

Latent region analysis. In addition to improving recommendation performance, our proposed model also provides a unique perspective on POI marketing segmentation, in the form of the learned regions. We take a representative area, California, as an example to analyze the regions learned by the Geo-PFM model. Fig. 9 visualizes the latent regions (Fig. 9b) learned from our model in versus its initialization by K-means (Fig. 9a) in California. Though we have no ground truth about an optimal POI region segmentation, we can infer the user activity regions in California through the collective check-in behaviors of users who have visited California and view the region clusters formulated by collective check-ins as ground truth (see Fig. 9c). Through analyzing the collaborative check-in frequency by those users, as shown in Fig. 9c, we can see two clear clusters in northern California among other scattered POIs, one cluster in the Los Angeles area, one in San Diego, and some scattered POIs between southern and northern California. K-means only depends on POI distances to cluster POIs for region segmentation. As shown in Fig. 9a, K-means segments northern California into four different regions, and segments Los Angeles into two regions. However, by considering the user check-in behaviors and geographical factors, our model identified a more meaningful region partition as shown in Fig. 9b, which is more coherent to real user activity as shown in Fig. 9c. Geo-PFM initiated by K-means leads to better POI segmentation. We can see that Geo-PFM models not only improve recommendation performance, but also provide an interesting perspective on POI marketing segmentation in the form of the learned regions.

Summary. Geographical influence and user mobility are two of the most important characteristics for LBSNs, and play an important role in POI recommendation. The fused method (Fu-PoiFM) which exploits an *ad hoc* two-step process to fuse the geographical influence and multi-center user activity pattern into user preferences can improve pure latent factor model (PoiFM). However, an integrated analysis of multiple factors for POI recommendations lead to further improvements. The proposed Geo-PFM model not only considers the geographical information of POIs and user mobility patterns for recommendation, but also updates the latent regions by considering these sources of information. The learned regions reflect the collaborative

user activity pattern. As a result, we can observe obvious improvements over all the baseline algorithms. Also, as shown in the performance of Poisson factor model compared to its Gaussian counterpart PFM, we observe improvements by Poison factor model, as Poisson distribution is more suitable for modeling count data. Further evidence of this is the fact that the Poisson based model Geo-PFM improves its non-negative factorization based counterpart Geo-BNMF in all the evaluation datasets, though Geo-BNMF imposes a non-negativity constrain.

6 Related Work

Recommender systems can be developed based on explicit user feedback. In other words, users rate items and the user-item preference relationship can be modeled on the basis of the user ratings. Latent factor models, such as as matrix factorization [19], probabilistic matrix factorization [28], its non-parametric version [27], and other other variants [1], [3], [17], [18], [22], [36], have become popular and widely used in recommendation. Most of the latent factors along this line of work assume that the response follows a Gaussian distribution over the product of user and item latent factors. The Gaussian-based latent factor models can achieve good prediction performance when explicit ratings are available. In contrast, recommender systems can also be developed based on implicit user feedback [16], such as the search and click behaviors on a web site [26], advertisement targeting [6], and the check-in behaviors in LBSNs, as we discussed in this paper. In this case, the recommender system has to infer user preferences from implicit user feedback. Here, latent factor models which are suitable for implicit user feedback are preferred. One option is to set non-negative constraints on latent factors to force the response variable into a wider range than the rating-based response. As a result, methods based on non-negative matrix factorization are widely used [13], [21], [25], [37]. However, the Poisson distribution is suitable for modeling count data. As a result, Poisson factor models are widely used for count based feedback recommendation settings [5], [6], [12], [26].

Some previous studies on POI recommendation, or more precisely location recommendation, mainly relied on user trajectory data to infer user preferences. For example, previous works [11], [38], [39], [40], [41] applied collaborative filtering (CF) methods to recommend locations and taxi pick-up locations based on user trajectory data. However, POI recommendation provide exact POIs a user would be interested rather than a "location". Due to the development and popularity of location-based social networks, more recent works, such as [33], [34], began to explore user preferences, social influence, and geographical influence for POI recommendations. However, these used a simple CF algorithm to fuse this information, and thus lack a comprehensive way to model how all this information collectively influence user POI check-in decision. The work in [24] tried to explore side information to improve POI recommendations, but it does not explore user mobility information and does not take the skewed data characteristics of implicit user check-in counts into the consideration. Kurashima et.al [20] extended the latent Dirichlet allocation (LDA) model to include geographical influence to profile user location preference, but it did not consider user mobility and the user activity areas modeled in this paper are constrained only to areas that a user has traveled to.

More recently, Cheng et al. [7] considered the geographical influence, the multi-center of user check-in patterns, the skewed user check-in frequency and social networks for POI recommendation. However, this work applied an *ad hoc* two-step method to fuse the geographical influence into user preferences, and did not really consider the user mobility and lacked an integrated consideration of factors that can influence POI recommendation. Moreover, the greedy clustering method applied to derive the personalized multi-centers could easily lead to overfitting problems in that it focuses on the regions a user has visited. Instead, our work is an integrated analysis of geographical influences, user mobility, and skewed data for POI recommendation. Hu and Ester [15] proposed a spatial topic model by considering the spatial and textual aspects of posts published by mobile users, and predict future user locations as POI recommendation. This is the work most closely related to ours in terms of the way to account for geographical influence and user mobility. However, their work is more similar to a location prediction problem than a POI recommendation task. Moreover, the Poisson model used in this paper could be equivalent conditioned on the per-user sums and where the item weights are constrained to sum to one [12], [42], [43]. However, our proposed Geo-PFM is more flexible and can be extended to different latent factor settings.

In addition, our work has a connection with recent works on mobility modeling [10], [14]. However, their tasks were different. Work [14] used a similar multinomial assumption over different regions to model geographical topics in Twitter stream, and the work in [10] investigated human mobility for social network analysis. Also, people have used Gaussian distribution to model region over locations [14], [31], [35].

As described above, while there are some studies on POI recommendation, they lacks an integrated analysis of the joint effects of multiple factors that influence the decision process of a user choosing a POI. These factors include user interest preferences, geographical influences, user mobility pattern, and the skewed implicit user check-in count data. The proposed method strategically takes all these factors into consideration and presents a flexible probabilistic generative model for POI recommendations.

7 Conclusion and Discussion

In this paper, we presented an integrated analysis of the joint effect of multiple factors which influence the decision process of a user choosing a POI and proposed a general framework to learn geographical preferences for POI recommendation in LBSNs. The proposed geographical probabilistic factor analysis framework strategically takes all these factors, which influence the user check-in decision process, into consideration. There are several advantages of the proposed recommendation method. First, the model captures the geographical influence on a user's check-in behavior by taking into consideration the geographical factors in LBSNs, such as the Tobler's first law of geography. Second, the

methods effectively modeled the user mobility patterns, which are important for location-based services. Third, the proposed approach extended the latent factors from explicit rating recommendation to implicit feedback recommendation settings by considering the skewed count data characteristic of LBSN check-in behaviors. Last but not least, the proposed model is flexible and could be extended to incorporate different latent factor models, which are suitable for both explicit and implicit feedback recommendation settings. Finally, extensive experimental results on real-world LBSNs data validated the performance of the proposed method.

Limitations and discussion. Geographical influence and user mobility are among the most important characteristics in LBSNs and could greatly affect POI recommendation. The proposed Geo-PFM model captures these two factors by introducing latent regions, which represent the collective user activity areas. This method coarsely captures the geographical influence and user mobility. However, the geographical influence and user mobility can be subtle [10], [29]. A possible future direction is to combine both the macroscopic and microscopic effects of geographical influence and user mobility.

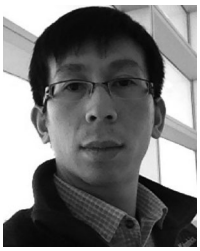
ACKNOWLEDGMENTS

This is a extended and revised version of [23], which appears in the *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD 2013)*. This research was partially supported by US National Science Foundation (NSF) via grant numbers CCF-1018151 and IIS-1256016. Also, it was supported in part by Natural Science Foundation of China (71028002). The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

REFERENCES

[1] D. Agarwal and B.-C. Chen, "Regression-based latent factor models," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2009, pp. 19–28.
[2] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, nos. 1/2, pp. 5–43, 2003.
[3] R. Bell, Y. Koren, and C. Volinsky, "Modeling relationships at multiple scales to improve accuracy of large recommender systems," in *Proc. 13th ACM SIGKDD Conf. Knowl. Discov. Data Min.*, 2007, pp. 95–104.
[4] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. 19th Int. Conf. Comput. Stat.*, 2010, pp. 177–187.
[5] J. Canny, "Gap: A factor model for discrete data," in *Proc. 27th ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2004, pp. 122–129.
[6] Y. Chen, M. Kapralov, D. Pavlov, and J. Canny, "Factor modeling for advertisement targeting," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 324–332.
[7] C. Cheng, H. Yang, I. King, and M. R. Lyu, "Fused matrix factorization with geographical and social influence in location-based social networks," in *Proc. 26th AAAI Conf. Artif. Intell.*, 2012, p. 1.
[8] Z. Cheng, J. Caverlee, K. Y. Kamath, and K. Lee, "Toward traffic-driven location-based web search," in *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage.*, 2011, pp. 805–814.
[9] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, "Exploring millions of footprints in location sharing services," in *Proc. 5th Int. Conf. Weblogs Social Media*, 2011, pp. 81–88.
[10] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2011, pp. 1082–1090.
[11] Y. Ge, H. Xiong, A. Tuzhilin, K. Xiao, M. Gruteser, and M. Pazzani, "An energy-efficient mobile recommender system," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2010, pp. 899–908.
[12] P. Gopalan, J. M. Hofman, and D. M. Blei, "Scalable recommendation with poisson factorization," *CoRR*, vol. abs/1311.1704, 2013.
[13] Q. Gu, J. Zhou, and C. H. Ding, "Collaborative filtering: Weighted nonnegative matrix factorization incorporating user and item graphs," in *Proc. 10th SIAM Int. Conf. Data Mining*, 2010, pp. 199–210.
[14] L. Hong, A. Ahmed, S. Gurumurthy, A. J. Smola, and K. Tsioutsiouliklis, "Discovering geographical topics in the twitter stream," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 769–778.
[15] B. Hu and M. Ester, "Spatial topic modeling in online social media for location recommendation," in *Proc. 7th ACM Conf. Recommender Syst.*, 2013, pp. 25–32.
[16] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Proc. 8th Int. Conf. Data Mining*, 2008, pp. 263–272.
[17] Y. Koren, "Factorization meets the neighborhood: A multifaceted collaborative filtering model," in *Proc. 14th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 2008, pp. 426–434.
[18] Y. Koren, "Collaborative filtering with temporal dynamics," *Commun. ACM*, vol. 53, no. 4, pp. 89–97, 2010.
[19] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.
[20] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura, "Geo topic model: Joint modeling of user's activity area and interests for location recommendation," in *Proc. 6th ACM Int. Conf. Web Search Data Mining*, 2013, pp. 375–384.
[21] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2000, pp. 556–562.
[22] T.-K. Huang, J. Schneider, J. G. CCarbonell, L. Xiong, and X. Chen, "Temporal collaborative filtering with bayesian probabilistic tensor factorization," in *Proc. SIAM Data Mining*, 2010, pp. 211–222.
[23] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning geographical preferences for point-of-interest recommendation," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2013, pp. 1043–1051.
[24] B. Liu and H. Xiong, "Point-of-interest recommendation in location based social networks with topic and location awareness," in *Proc. SIAM Int. Conf. Data Mining*, 2013, pp. 396–404.
[25] C. Liu, H.-C. Yang, J. Fan, L.-W. He, and Y.-M. Wang, "Distributed nonnegative matrix factorization for web-scale dyadic data analysis on mapreduce," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 681–690.
[26] H. Ma, C. Liu, I. King, and M. R. Lyu, "Probabilistic factor models for web site recommendation," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2011, pp. 265–274.
[27] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo," in *Proc. 25th Int. Conf. Mach. Learn*, 2008, pp. 880–887.
[28] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2008, vol. 20, pp. 1257–1264.
[29] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2011, pp. 1046–1054.
[30] M. N. Schmidt, O. Winther, and L. K. Hansen, "Bayesian non-negative matrix factorization," in *Proc. 8th Int. Conf. Independent Component Anal. Signal Separation*, 2009, pp. 540–547.
[31] S. Sizov, "Geofolk: Latent spatial semantics in web 2.0 social media," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2010, pp. 281–290.
[32] W. Tobler, "A computer movie simulating urban growth in the detroit region," *Econ. Geography*, vol. 46, no. 2, pp. 234–240, 1970.
[33] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, 2010, pp. 458–461.
[34] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point of interest recommendation," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2011, pp. 325–334.

[35] Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang, "Geographical topic discovery and comparison," in *Proc. Int. Conf. World Wide Web*, 2011, pp. 247–256.
[36] L. Zhang, D. Agarwal, and B.-C. Chen, "Generalizing matrix factorization through flexible regression priors," in *Proc. 5th ACM Conf. Recommender Syst.*, 2011, pp. 13–20.
[37] S. Zhang, W. Wang, J. Ford, and F. Makedon, "Learning from incomplete ratings using non-negative matrix factorization," in *Proc. 6th SIAM Int. Conf. Data Mining*, 2006, pp. 549–553.
[38] V. W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang, "Collaborative filtering meets mobile recommendation: A user-centered approach," in *Proc. 26th AAAI Conf. Artif. Intell.*, 2010, pp. 236–241.
[39] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *Proc. 10th Int. Conf. World Wide Web*, 2010, pp. 1029–1038.
[40] Y. Zheng and X. Xie, "Learning travel recommendations from user-generated GPS traces," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 1, pp. 2:1–2:29, Jan. 2011.
[41] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proc. 10th Int. Conf. World Wide Web*, 2009, pp. 791–800.
[42] M. Zhou and L. Carin, "Negative binomial process count and mixture modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*
[43] M. Zhou, L. Hannah, D. B. Dunson, and L. Carin, "Beta-negative binomial process and poisson factor analysis," in *Proc. 15th Int. Conf. Artif. Intell. Stat.*, 2012, pp. 1462–1471.



Bin Liu is currently working toward the PhD degree in the Department of Management Science and Information Systems, Rutgers Business School, Rutgers University. His general area of research is data mining and business analytics, with a focus on recommender systems, mining rich user-generated content, and location-based services.

Hui Xiong (SM'07) received the BE degree from the University of Science and Technology of China (USTC), China, the MS degree from the National University of Singapore (NUS), Singapore, and the PhD degree from the University of Minnesota (UMN). He is currently a professor and vice chair of the Management Science and Information Systems Department, and the director of Rutgers Center for Information Assurance at the Rutgers, the State University of New Jersey, where he received a two-year early promotion/tenure in 2009, the Rutgers University Board of Trustees Research Fellowship for Scholarly Excellence in 2009, and the ICDM-2011 Best Research Paper Award in 2011. His general area of research is data and knowledge engineering, with a focus on developing effective and efficient data analysis techniques for emerging data intensive applications. He has published prolifically in refereed journals and conference proceedings (three books, more than 40 journal papers, and more than 60 conference papers). He is a coeditor-in-chief of *Encyclopedia of GIS*, an associate editor of *IEEE Transactions on Data and Knowledge Engineering (TKDE)*, and the *Knowledge and Information Systems (KAIS)* journal. He has served regularly on the organization and program committees of numerous conferences, including as a program co-chair of the Industrial and Government Track for the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining and a program co-chair for the 2013 IEEE International Conference on Data Mining (ICDM-2013). He is a senior member of the ACM and IEEE.

Spiros Papadimitriou is an assistant professor in the Department of Management Science & Information Systems at Rutgers Business School. Previously, he was a research scientist at Google, and a research staff member at IBM Research. His main interests are large scale data analysis, time series, graphs, and clustering. He has published more than 40 papers on these topics and has three invited journal publications in best paper issues, several book chapters and he has filed multiple patents. He has also given a number of invited talks, keynotes, and tutorials. He received the Siebel Scholarship in 2005 and the Best Paper Award in SDM 2008.

Yanjie Fu received the BE degree from the University of Science and Technology of China, China, 2008, the MS degree from the Chinese Academy of Sciences, China, 2011. He is currently working toward the PhD degree in the Management Science and Information Systems Department at Rutgers University. His research interests include data mining, business analytics, geoeconomics, and customer targeting.

Zijun Yao received the BE degree in electrical engineering from the Guangdong University of Technology in 2009, the MS degree in computer engineering from Northeastern University in 2011. He is currently working toward the PhD degree in information technology at Rutgers, The State University of New Jersey. His research interests include data mining and business analytics.