

# Learning Geographical Preferences for Point-of-Interest Recommendation

Bin Liu, Yanjie Fu, Zijun Yao, Hui Xiong<sup>\*</sup>  
Rutgers Business School, Rutgers University  
{binben.liu, yanjie.fu, zijun.yao, hxiong}@rutgers.edu

## ABSTRACT

The problem of point of interest (POI) recommendation is to provide personalized recommendations of places of interests, such as restaurants, for mobile users. Due to its complexity and its connection to location based social networks (LBSNs), the decision process of a user choose a POI is complex and can be influenced by various factors, such as user preferences, geographical influences, and user mobility behaviors. While there are some studies on POI recommendations, it lacks of integrated analysis of the joint effect of multiple factors. To this end, in this paper, we propose a novel geographical probabilistic factor analysis 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, the user mobility behaviors can be effectively exploited in the recommendation model. Moreover, the recommendation model can effectively make use of user check-in count data as implicitly user feedback for modeling user preferences. Finally, experimental results on real-world LBSNs data show that the proposed recommendation method outperforms state-of-the-art latent factor models with a significant margin.

## Categories and Subject Descriptors

H.2.8 [Database Management]: Database Applications—*Data mining*

## General Terms

Algorithms, Experimentation

## Keywords

Location-based Social Networks, Point-of-Interest, Recommender Systems, Human Mobility, User Profiling

## 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.

<sup>\*</sup>Contact Author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

KDD '13, August 11–14, 2013, Chicago, Illinois, USA.

Copyright 2013 ACM 978-1-4503-2174-7/13/08 ...\$15.00.

LBSNs allow users to explore Point-of-Interest (POIs), such as restaurants and shopping malls, for better services by sharing their check-in experiences and opinions on the POIs which they have checked in. Indeed, the task of POI recommendation is to provide personalized recommendations of places of interests to mobile users. It plays an important role in providing better location based services in location based social networks. In fact, both LBSN users and POI owners can benefit from effective POI recommendations. For mobile users, they can identify favorite POIs and have better user experiences by the right POI recommendations. Also, for POI owners, they could exploit POI recommendation for acquiring more target customers.

There are several unique characteristics of LBSNs which distinguish POI recommendations from traditional recommendation tasks.

- Tobler's first law of geography. "Everything is related to everything else, but near things are more related than distant things" [21]. This indicates that geographically proximate POIs are more likely to share similar characteristics. Also, the probability of a user for a POI is inversely proportional to geographic distance.
- Regional popularity. Two POIs with similar or the same semantic topics can have different popularities if they are located in different regions.
- Dynamic user mobility. In LBSNs, a user may check in POIs at different regions. For example, a LBSN user may travel to different cities. Dynamic user mobility imposes huge challenges on POI recommendations.
- Implicit user feedback. In the study of POI recommendations, the explicit user ratings are usually not available. The recommender system has to infer user preferences from implicit user feedback in terms of user check-in frequency data.

Indeed, traditional recommender systems usually assume that the user and item data are independent and identically distributed. This assumption ignores the interrelationships among items. In fact, the decision process of a user choose 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, the propensity of a user for a POI is inversely proportional to the distance between the user and the POI, which is similar to the probability of purchasing an item is inversely proportional to the cost. Second, utility matters. In economics, utility is an index of preferences over some sets of items and services when a user makes purchase decisions. In other words, a user may prefer a far away POI

than a nearby one for better satisfaction. Third, popularity affects purchase behaviors. In fact, the probability of a user purchasing a product is largely affected by the word-of-mouth about the product. Finally, the LBSN users have dynamic mobility behaviors, which impose the challenges on the modeling of check-in decision.

Another unique characteristic of POI recommendations is that user check-in count data follow a power law distribution. This is different from the rating data in traditional recommender systems. In other words, a user can visit a POI only once and another POI hundreds of times. Since we do not have explicit user rates for POIs, we can only make use of implicit user behavior data in the check-in records for POI recommendations. All the above demands a reconsideration of the recommendation model.

While there are some studies on POI recommendations, it lacks of integrated analysis of the joint effect of the above factors, such as user preferences, geographical influences, and user mobility behaviors. Therefore, in this paper, we propose a geographical probabilistic factor analysis framework which can strategically takes various factors into consideration. This framework allows to capture the geographical influences on a user's check-in behaviors, effectively model the user mobility patterns, and deal with the skewed distribution of check-in count data. Specifically, 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. Also, we 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 a collective influence. Moreover, a latent model is developed to capture user preferences. This model takes geographical influence, regional popularity and user mobility into consideration, and can effectively handle the skewed data distribution of POI count data.

Finally, experimental results on real-world LBSNs and POI data show that the proposed POI recommendation method outperforms state-of-the-art probabilistic latent factor models with a significant margin in terms of Top- $K$  recommendation. Also, we provide a location-marketing segmentation study by comparing the latent regions identified by our model with those derived from traditional marketing segmentation methods.

## 2. THE FRAMEWORK FOR POI RECOMMENDATION

In this section, we introduce a geographical probabilistic factor analysis framework for POI recommendations.

### 2.1 Problem Definition

The problem of personalized POI recommendation is to recommend POIs to a user. Let  $U = \{u_1, u_2, \dots, u_M\}$  be a set of LBSN users, where each user has a location  $l_i$  and observable properties  $x_i$  (e.g. a user's profile). The user location  $l_i$  is usually unknown due to user mobility. Let  $V = \{v_1, v_2, \dots, v_N\}$  be a set of POIs, where each POI has a location  $l_j = \{lon_j, lat_j\}$  in terms of longitude and latitude, observable properties  $x_j$  (e.g. a textual description of a POI), and a popularity score  $\rho_j$ . Typically, a tiny proportion of check-in records of users for POIs are given with  $y_{ij}$  as the number of times user  $u_i$  visited POI  $v_j$ . For convenience, we refer  $i$  as user and  $j$  as POI unless stated oth-

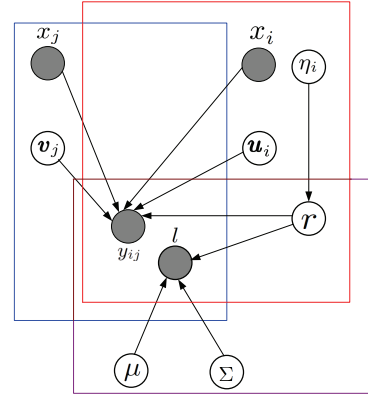


Figure 1: A graphical representation of the proposed model. The model priors have been excluded for the simplicity of description.

Table 1: Mathematical Notations

Symbol	Size	Description
$R$	$1 \times \mathbb{R}$	latent region set, $r$ is a region in $R$
$\eta$	$M \times \mathbb{R}$	user level region distribution
$\rho$	$1 \times N$	item popularity
$\mu$	$\mathbb{R}^2$	mean location of a latent region
$\Sigma$	$\mathbb{R}^{2 \times 2}$	covariance matrix of a latent region
$U$	$M \times K$	user latent factor
$V$	$N \times K$	item latent factor
$x$	$(\cdot) \times \mathbb{K}$	user or item observable properties
$\theta, \pi$	$(\cdot) \times \mathbb{K}$	user or item topic distribution

erwise. Also, we use POI and item interchangeably. Some important notations used in this paper are listed in Table 1.

### 2.2 The General Idea

To learn geographical user preferences, we need a model to encode the spatial influence and user mobility into the user check-in decision process. By Tobler's first law of geography, POIs with similar services are likely to be clustered into the same geographical area. To capture this, we assume that the geographical locations have been clustered into  $\mathbb{R}$  latent regions and denoted as  $R$ . Also, as shown in Figure 2, LBSN users are most likely to check in a number of POIs and these POIs are usually limited to some geographical regions. We apply a multinomial distribution,  $r \sim p(r|\eta_i)$ , to model user mobility over the regions  $R$  with  $\eta_i$  being a user dependent distribution over latent regions for user  $i$ . Then we assume a Gaussian geographical distribution for each region  $r \in R$ , and the location for POI  $j$  is characterized by  $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$  with  $\mu_r$  and  $\Sigma_r$  being the mean vector and covariance matrix of the region.

To model a user's propensity of a POI, we assume the following factors that will affect the user check-in decision process: (1) each user  $i$  is associated with an interest  $\alpha(i, j)$  with respect to POI  $j$ ; (2) each POI  $j$  has popularity  $\rho_j$ ; and (3) the distance between the user and the POI  $d(i, j)$ . As a result, the probability of observing a pair  $(i, j)$  is directly proportional to the user interest and the popularity of the entity, and monotonically decreases with the distance between them.

$$p(i, j) \propto \alpha(i, j) \rho(j) (d_0 + d(i, j))^{-\tau} \quad (1)$$

We use a power-law like parametric term  $(d_0 + d(i, j))^{-\tau}$  to model the distance factor in the decision making process.

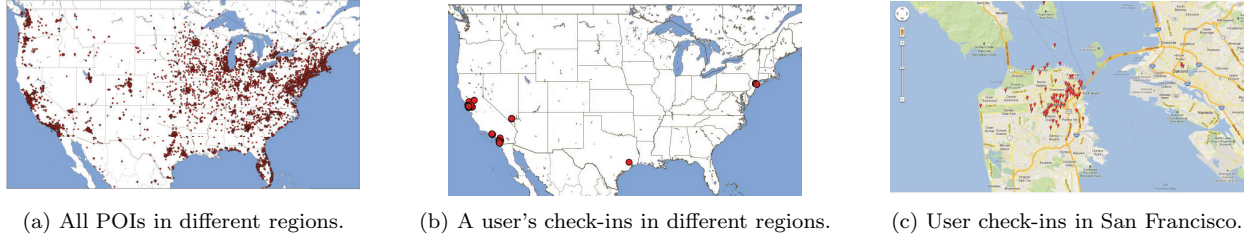


Figure 2: An example of a typical user check-in pattern: (a) all the POIs; (b) the user's check-ins over different regions: San Francisco, Los Angeles, San Diego, Las Vegas, Houston, and New York City; (c) the user's check-ins in San Francisco area.

This motivated by the following observations: the probability user  $u_i$  checking-in at POI  $v_j$  decays as the power-law of the distance between them.

Because we use implicit user check-in count data to model user preferences and the distribution of check-in counts are usually skewed, we adapt a Bayesian probabilistic non-negative latent factor model:  $y_{ij} \sim P(f_{ij})$  where  $f_{ij} = \alpha(i, j) \cdot \rho(j) \cdot (d_0 + d(i, j))^{-\tau}$  for encoding both the spatial influence and personalized preferences.

### 2.3 Geographical Probabilistic Factor Model

The check-in decision process of user  $i$  for POI  $j$  can be viewed in a generative way. First, a user samples a region from all  $\mathbb{R}$  regions following a multinomial distribution  $r \sim \text{Multinomial}(\eta_i)$ , and a POI is draw from the sampled region  $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$ . Then, depending on user preferences, the popularity of the selected POI in that sampled region, and the distance between the user and the POI, the user makes a check-in decision following certain distribution. The user  $i$ 's preference for POI  $j$  can be represented as a linear combination of a latent factor  $\mathbf{u}_i^\top \mathbf{v}_j$  and a function of user and item observable properties  $x_i^\top W y_j$ . As shown in Figure 1, the geographical probabilistic generative process for a user  $i$  to choose a POI  $j$  can be expressed as follows:

1. Draw a region  $r \sim \text{Multinomial}(\eta_i)$ .
2. Draw a location  $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$ .
3. Draw a user preference
  - a Generate user latent factor  $\mathbf{u}_i \sim P(\mathbf{u}_i; \Psi_{\mathbf{u}_i})$ .
  - b Generate item latent factor  $\mathbf{v}_j \sim P(\mathbf{v}_j; \Psi_{\mathbf{v}_j})$ .
  - c User-item preference  $\alpha(i, j) = \mathbf{u}_i^\top \mathbf{v}_j + x_i^\top W y_j$ .
4.  $y_{ij} \sim P(f_{ij})$  where
 
$$f_{ij} = \left( \mathbf{u}_i^\top \mathbf{v}_j + x_i^\top W y_j \right) \cdot \rho(j) \cdot (d_0 + d(i, j))^{-\tau}$$

In the generative model,  $\mathbf{u}$  and  $\mathbf{v}$  are user and item factors,  $x_i$  and  $x_j$  are user and item observable properties respectively, and  $W$  is a matrix that is used to transfer the observable prosperity space into the latent space to capture the affinity between the observed features and the user-item pair. Note that the proposed model is general and can be extended by different factor models. The model details will be described in Section 2.4.

## 2.4 Model Specification

This section introduces detailed model specifications.

### 2.4.1 A User Mobility Model

A user may check in POIs in different geographical areas as shown in Figure 2. Indeed, a LBSN user may travel to different regions (Figure 2b), and the recommendation problem

becomes difficult when this user travels to a new place. Also, it is important to capture users' mobility behaviors and find users' preferences. As shown in Figure 2, LBSN users are most likely to check in a number of POIs and these POIs are usually limited to some geographical regions. To model a user's propensity of a POI, we first need to choose the region where the POI is associated. We apply a multinomial model [10] to model the region  $r \sim p(r|\eta_i)$ , where  $\eta_i$  is a user dependent distribution over latent regions for user  $i$ . Each POI location  $l_j$  is drawn from a region-dependent multivariate normal distribution:  $l_j \sim \mathcal{N}(\mu_r, \Sigma_r)$ .

Also, due to the mobility pattern, the explicit location  $\ell(\cdot)$  of a user is unknown. We use the region  $r$  to represent the user activity area and model the user-POI distance as

$$d(i, j) = \sqrt{\|\mu_r - l_j\|_2}. \quad (2)$$

### 2.4.2 Popularity

Popularity can affect the user check-in behaviors to a great extent. Usually, an individual's decision to check in a POI is largely affected by the word-of-mouth opinions, which can be represented as the popularity of the POI. We propose to normalize the popularity score in a region level. On the one hand, two POIs with similar or the same semantic terms can be rated differently in a region. On the other hand, as a user choose a POI from a pool of choices in a region, the popularity score represents the user satisfaction level. A POI with a higher popularity score potentially provide better user experiences according to word-of-mouth opinions.

For example, in Foursquare, we know the total number of people have visited a specific POI  $j$ ,  $\text{totalPeo}_j$ , and total times they check in there,  $\text{totalCk}_j$  (refer the example in Table 2). We normalize the popularity score for a POI  $j$  in area  $r$  to  $[0, 1]$  by the following equation [15].

$$\rho_j = \frac{1}{2} \left\{ \frac{\text{totalPeo}_j - 1}{\max_{j \in r} \{\text{totalPeo}_j\} - 1} + \frac{\text{totalCk}_j - 1}{\max_{j \in r} \{\text{totalCk}_j\} - 1} \right\} \quad (3)$$

where  $\max_{j \in r} \{\text{totalPeo}_j\}$  and  $\max_{j \in r} \{\text{totalCk}_j\}$  are the maximum total number of people and total check-in number in the region respectively.

### 2.4.3 Geographical-Topical Bayesian Non-negative Matrix Factorization

Latent factor models aim on learning latent factors to model user-item interactions by associating user latent factor  $\phi_i$  for each user and item latent factors  $\phi_j$  for each item. Then, the rating is modeled as  $p(y_{ij}|i, j) = p(y_{ij}|\phi_i^\top \phi_j; \Theta)$ . Latent factor models, such as probabilistic matrix factorization (PMF) [18], have become popular in recommendation. One drawback of latent factor models is the so called cold-start problem, in which latent factor models perform poorly when there are many unseen users or items. Another

drawback of previous work on using latent factors for recommendation is that they assume a normal distribution for  $p(y_{ij}|\phi_i^\top \phi_j; \Theta)$  and apply a logistic function to bound the predicted values. Therefore, it is not suitable for recommendation problems where the user-item data have skewed frequency values. For POI recommendation, the user check-in frequency data actually follows the power law distribution.

To circumvent these two drawbacks of latent factor models for POI recommendations, we propose a Geographical-Topical Bayesian Non-negative Matrix Factorization (GT-BNMF) model by incorporating the geographical influence and textual information in terms of topic distribution.

### An Aggregated LDA Model.

To better explore the observable features, we first profile users and POIs through an aggregated LDA model proposed in our earlier work [15]. The Latent Dirichlet Allocation (LDA) model [4] is a popular technique to identify latent topic from a large document collection. In aggregated LDA, we aggregate all the documents of the POIs, user  $u_i$  have checked in, into a user document  $d_{u_i}$ . The topics of  $d_{u_i}$  can represent user  $u_i$ 's preferences in terms of interest topics. Each document corresponds to a LBSN user. As a result, the topic distribution of document  $d_{u_i}$  represents the interests of  $u_i$ . Each user  $u$  is associated with a multinomial distribution over topics, represented by  $\theta$ . Each interest topic is associated with a multinomial distribution over check-in textual terms, represented by  $\phi$ . The generation process of the region aware user interest topic is as follows.

1. For each topic  $z \in \{1, \dots, K\}$ , draw a multinomial distribution over terms,  $\phi_z \sim \text{Dir}(\beta)$ .
2. For the document  $d_{u_i}$  given a user  $u_i$ 
  - (a) Draw a topic distribution,  $\theta_{d_{u_i}} \sim \text{Dir}(\alpha)$
  - (b) For each word  $w_{d,n}$  in document  $d_{u_i}$ :
    - i. Draw a topic  $z_{d,n} \sim \text{Mult}(\theta_{d_{u_i}})$
    - ii. Draw a word  $w_{d,n} \sim \text{Mult}(\phi_{z_{d,n}})$

Then, we have: (1) Matrix  $\Theta_{M \times K}$ , where  $M$  is the number of users and  $K$  is the number of topics.  $\theta_{ij}$  represents the probability that user  $i$  is interested in topic  $t_j$ . (2) Matrix  $\Phi_{K \times V}$  where  $K$  is the number of topics and  $V$  is the number of unique terms in the dataset. Vector  $\phi_i$  is the probability distribution of topic  $i$  over the  $V$  terms. We further infer the topic distribution  $\pi_j$  of POI  $v_j$  based on the learned user topic term distribution  $\Phi_{K \times V}$ .

**Comments:** In our newly proposed geographical probabilistic factor model as described in Section 2.3, we assume both user and item observable properties  $x_i$  and  $x_j$ . Here, we apply an aggregated LDA model to mine the topic distribution, and replace  $x_i$  and  $x_j$  with topic distributions  $\phi_i$  and  $\pi_j$ . The aggregated LDA here does not improve the computation complexity of the proposed model.

### GT-BNMF.

The implicit feedback user-item count matrix  $Y$  can be factorized as  $Y \sim P(\mathbf{UV})$ , and priors  $P(\mathbf{u}_i; \Psi_{\mathbf{u}_i})$  and  $P(\mathbf{v}_j; \Psi_{\mathbf{v}_j})$  can be further specified. Consider the count data characteristic, we expect user factor  $\mathbf{U}$  and item factor  $\mathbf{V}$  to be non-negative. Some previous work such as [17, 5] assumed a Poisson distribution for  $Y \sim P(\mathbf{UV})$  and further placed Gamma priors on  $\mathbf{U}$  and  $\mathbf{V}$ . Actually, the Poisson factor analysis can be equal to NMF in certain condition [14]. Euclidean distance based optimization leads to Gaussian NMF, and KL-

divergence based optimization leads to Poisson NMF. For computation efficiency purpose, we here adopt a Bayesian non-negative matrix factorization (BNMF) method [19]; but our framework is flexible, thus it is also suitable for Poisson factor model among others.

More specifically, we assume a rectified normal<sup>1</sup> distribution on  $Y \sim P(\mathbf{UV})$  with variance  $\sigma^2 \mathbf{I}$ .

$$Y \sim \mathcal{N}^R(Y|\mathbf{UV}, \sigma^2 \mathbf{I}), \text{ subject to } \mathbf{U} \geq 0, \mathbf{V} \geq 0 \quad (4)$$

with non-negative constrains. Further, we place an exponential distribution, which is a special case of Gamma distribution with shape equals to 1, on  $\mathbf{U}$  and  $\mathbf{V}$ , and an inverse gamma distribution on  $\sigma^2$  with shape  $a$  and scale  $b$ .

$$\begin{aligned} P(\mathbf{U}|\alpha) &= \prod_{i=1}^M \prod_{k=1}^K \alpha \cdot \exp(-\alpha \cdot u_{ik}), \\ P(\mathbf{V}|\beta) &= \prod_{j=1}^N \prod_{k=1}^K \beta \cdot \exp(-\beta \cdot v_{jk}) \\ P(\sigma^2) &= \frac{b^a}{\Gamma(a)} (\sigma^2)^{-a-1} \exp(-\frac{b}{\sigma^2}) \end{aligned} \quad (5)$$

Each element  $u_{ik} \in \mathbf{U}$  encodes the preference of user  $i$  to "topic"  $k$ , and each element  $v_{jk} \in \mathbf{V}$  reflects the topical affinity of item  $j$  to topic  $k$ .

Also, we utilize the user topic distribution  $\theta_i$  and item topic distribution  $\pi_j$  to ease the cold-start drawback of latent factor models. We add another term depends on  $\theta_i$  and  $\pi_j$ . Then we have the following generative process.

1. Generate user latent factor  $u_{ik} \sim \text{Exp}(\alpha_k)$ .
2. Generate item latent factor  $v_{jk} \sim \text{Exp}(\beta_k)$ .
3. User-item preference  $\alpha(i, j) = \mathbf{u}_i^\top \mathbf{v}_j + \theta_i^\top W \pi_j$ .

By combining the geographical influence and latent factor, we have GT-BNMF as

$$\begin{aligned} y_{ij} &\sim \mathcal{N}^R(y_{ij}|f_{ij}, \sigma^2) \\ f_{ij} &= (\mathbf{u}_i^\top \mathbf{v}_j + \theta_i^\top W \pi_j) \cdot \rho(j) \cdot (d_0 + d(i, j))^{-\tau} \end{aligned} \quad (6)$$

Here,  $W$  is a matrix that used to transfer the topic space into the latent space to capture the affinity between the observed features and the user-item pair.

## 2.5 Parameter Estimation

Let us denote all parameters by  $\Psi = \{\mathbf{U}, \mathbf{V}, \sigma^2, W, \eta, \mu, \Sigma\}$ , and  $\Omega = \{\alpha, \beta, a, b\}$  be the hyperparameters. 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. The posterior probability is

$$P(\Psi, \alpha, \beta; \mathcal{D}) = \prod_{\mathcal{D}} P(y_{ij}, l_j | \Psi, \Omega) \quad (7)$$

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

<sup>1</sup>[http://en.wikipedia.org/wiki/Rectified\\_Gaussian\\_distribution](http://en.wikipedia.org/wiki/Rectified_Gaussian_distribution)

### 2.5.1 E-step

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

$$r \sim P\left(r|\{y_{\cdot j}, l_j\}, R^{(t)}, \Psi^{(t)}\right) \times P(r|\boldsymbol{\eta}^{(t)}) \quad (8)$$

where

$$P\left(\{y_{\cdot j}, l_j\} | r, \Psi^{(t)}\right) = P\left(l_j | r, \Psi^{(t)}\right) \times P\left(y_{\cdot j} | r, \Psi^{(t)}\right) \quad (9)$$

$$P\left(l_j | r, \Psi^{(t)}\right) = \mathcal{N}\left(l_j | \mu_r^{(t)}, \Sigma_r^{(t)}\right) \quad (10)$$

$$P\left(y_{\cdot j} | r, \Psi^{(t)}\right) = \prod_{i=1}^M \mathcal{N}^R(f_{ij}, \sigma^2) \quad (11)$$

$P(r|\boldsymbol{\eta}^{(t)})$  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 collaborative user effects.

### 2.5.2 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  $\mu_r, \Sigma_r, \boldsymbol{\eta}$  directly from the samples.

$$\begin{aligned} \mu_r &= \frac{1}{\#(j, r)} \sum_{j=1}^D \mathbb{I}(r_j = r) l_j \\ \Sigma_r &= \frac{1}{\#(j, r) - 1} \sum_{j=1}^D \left( (l_j - \mu_r)^\top (l_j - \mu_r) \right) \end{aligned} \quad (12)$$

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  $\boldsymbol{\eta}_{ir}^{(t+1)} = \frac{C_{ir}^{(t+1)} + \gamma}{C_i^{(t+1)} + \mathbb{R}\gamma}$  where  $C_{ir}$  is the number of POIs being assigned to region  $r$  for user  $i$ , and  $C_i$  is all the the number of POIs of all the regions for user  $i$ .

After updating region  $R^{(t+1)}$ , we update  $\Psi^{(t+1)}$  that maximizes the expectation  $Q(\Psi|\Psi^{(t)})$ .

$$\begin{aligned} \mathcal{L}(\mathbf{U}, \mathbf{V}, \sigma^2, W | R^{(t+1)}, D) \\ = \sum_{i=1}^M \sum_{j=1}^N \left[ \frac{1}{2} \ln \sigma^2 - \frac{(y_{ij} - [\rho_j \delta_j \cdot \mathbf{U}\mathbf{V}_{ij} + \rho_j \delta_j \cdot \theta_i^\top W \pi_j])^2}{2\sigma^2} \right] \\ - \sum_{i=1}^M \sum_{k=1}^K \alpha \cdot U_{ik} - \sum_{j=1}^N \sum_{k=1}^K \beta \cdot U_{jk} - \left( (a+1) \ln \sigma^2 + \frac{b}{\sigma^2} \right) \end{aligned} \quad (13)$$

Here  $\delta_j$  is the distance factor of user  $i$  for POI  $j$  as described in Section 2.4.1.

We apply an iterated conditional modes (ICM) [19] to update  $\mathbf{U}, \mathbf{V}, \sigma^2$

$$\begin{aligned} u_{ik} &= \frac{\sum_{j=1}^N \left[ \left( y_{ij} - \sum_{k' \neq k} u_{ik'} v_{jk'} - \rho_j \delta_j \cdot \theta_i^\top W \pi_j \right) \rho_j \delta_j \cdot v_{jk} \right] - \alpha \cdot \sigma^2}{\sum_{j=1}^N (\rho_j \delta_j \cdot v_{jk})^2} \\ v_{jk} &= \frac{\sum_{i=1}^M \left[ \left( y_{ij} - \sum_{k' \neq k} u_{ik'} v_{jk'} - \rho_j \delta_j \cdot \theta_i^\top W \pi_j \right) \rho_j \delta_j \cdot u_{ik} \right] - \beta \cdot \sigma^2}{\sum_{i=1}^M (\rho_j \delta_j \cdot u_{ik})^2} \\ \sigma^2 &= \frac{\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^N (y_{ij} - [\rho_j \delta_j \cdot \mathbf{U}\mathbf{V}_{ij} + \rho_j \delta_j \cdot \theta_i^\top W \pi_j])^2 + b}{\frac{MN}{2} + a + 1} \end{aligned} \quad (14)$$

Since the elements in each column of  $\mathbf{U}$  and  $\mathbf{V}$  are conditionally independent, we can update entire column of  $\mathbf{U}$  and  $\mathbf{V}$  simultaneously. When  $M \times N \times K \gg (M+N) \times K$ , it can save a lot of computation. A gradient based method is applied to update  $W$ :  $W^{(t+1)} = W^{(t)} - \epsilon \times \frac{\partial \mathcal{L}}{\partial W}$  where  $\epsilon$  is the learning rate.

**Comments:** Note that the region  $R$  is updated in each E-step, so as the popularity and user-POI distance. The model parameters are updated upon the new region information.

## 2.6 Recommendation

After parameters  $\Psi$  are learned, the GT-BNMF model predicts the number of check-ins of a user for a given POI as  $\mathbb{E}(y_{ij}|u_i, v_j) = (\mathbf{u}_i^\top \mathbf{v}_j + \theta_i^\top W \pi_j) \cdot \rho(j) \cdot (d_0 + d(i, j))^{-\tau}$ . We make recommendations based on the predicted check-in times, the larger the predicted value, the more likely the user will choose this POI. Since POI recommendation in LBSNs is highly location sensitive, the recommendation list should be within a user's activity region. Our GT-BNMF model learn the latent region. Therefore, we make recommendations based on the learned latent region.

## 3. EXPERIMENTAL RESULTS

Here, we provide an empirical evaluation of the performances of the proposed model. All the experiments were performed on a large real-world LBSN dataset collected from Foursquare ([www.foursquare.com](http://www.foursquare.com)), one of the most popular LBSN communities.

### 3.1 The Experimental Data

The dataset is formulated as follows [15, 7]. Foursquare users usually report their check-ins of POIs via Twitter. When a LBSN user posted a Tweet, which shows a check-in of a POI, we consider it as the user has checked in the POI physically. As shown in Table 2, for POIs in Foursquare, we have geographical description information such as location information, textual information, and popularity information in terms of the total number of people and the total number of check-ins. The dataset includes POIs across the United States (except Hawaii and Alaska), and the geographical distribution of all POIs is shown in Figure 2a.

**Name:** Otto Enoteca Pizzeria  
**Address:** 1 5th Ave, New York, NY 10003  
**Tags:** pizza wine bar italian olive oil cheese mario batali meat wine pasta gelato gluten free menu zagat rated pizza  
**Total people:** 3,127, **Total check-ins:** 4,770.

Table 2: An example POI and its associated information.

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%; it is very sparse. Each user checked in 59.44 POIs on average, only a very small fraction of all the POIs. The number of check-ins for a POI is ranged from 1 to 786. This range is very wide as show in Figure 3.

### 3.2 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 in



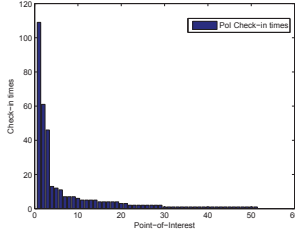


Figure 3: An example of wide range user check-in count data for a randomly chosen user.

a region, precision and recall are defined as  $\text{Precision@}N = \frac{|S_{N,\text{rec}} \cap S_{\text{visited}}|}{N}$ ,  $\text{Recall@}N = \frac{|S_{N,\text{rec}} \cap S_{\text{visited}}|}{|S_{\text{visited}}|}$  where  $S_{\text{visited}}$  are the POIs a user have visited in the test data. The precision and recall for the entire recommender system are computed by averaging all the precision and recall for all the users respectively.

**r-Precision and r-Recall.** In addition to using absolute precision and recall for evaluation, we also compare the models in terms of relative precision and recall. Let  $C$  denote the candidate POIs, the precision and recall in a top- $N$  list of a random recommender system are  $\frac{|S_{\text{visited}}|}{|C|}$  and  $\frac{N}{|C|}$ . Then, the relative precision and recall [24] are defined as

$$\begin{aligned} \text{rPrecision@}N &= \frac{\text{Precision@}N}{\frac{|S_{\text{visited}}|}{|C|}} = \frac{|S_{N,\text{rec}} \cap S_{\text{visited}}| \cdot |C|}{|S_{\text{visited}}| \cdot N} \\ \text{rRecall@}N &= \frac{\text{Recall@}N}{\frac{N}{|C|}} = \frac{|S_{N,\text{rec}} \cap S_{\text{visited}}| \cdot |C|}{|S_{\text{visited}}| \cdot N} \end{aligned} \quad (15)$$

The relative precision and recall have the same value. As a result, we only need to show one of them, and we will name it as the relative performance, which measures the improvement over a random recommendation.

### 3.3 Baseline Algorithms

We first compare our proposed GT-BNMF model with four traditional recommendation methods: singular value decomposition (SVD) [12], probabilistic matrix factorization (PMF) [18], Nonnegative matrix factorization (NMF) [14], and Bayesian nonnegative matrix factorization (BNMF) [19]. SVD and PMF are the standard algorithms for rating based recommendations. As mentioned before, the check-in count data have a wide range, so we also use NMF and BNMF for comparison. Moreover, we also compare our model with the fused method [6], which fuse the geographical influence into factor models. Since BNMF also exploits the count check-in characteristics, we fuse the geographical influence into BNMF and denote it as F-BNMF.

### 3.4 The Method for Comparison

Since POI recommendation in LBSNs is highly location sensitive, the recommendation should be within the user's current region. Therefore, we measure the Top  $N$  performances by considering the recommendation within a region. We first cluster all the POIs into  $\mathbb{R}$  regions. This is the initialization of the GT-BNMF model. For SVD, PMF, NMF, BNMF and F-BNMF methods, we make top- $N$  recommendation within the initialized  $\mathbb{R}$  clustered regions by  $K$ -means. For the proposed GT-BNMF model, we recommend top- $N$  POIs within each learned latent region.

### 3.5 Implementation Details

We randomly divided the data into training (80%) and test(20%) data. We set  $\lambda_U = 0.01$  and  $\lambda_C = 0.01$  for PMF.

K	Pre	SVD	PMF	NMF	BNMF	F-BNMF	GT-BNMF
10	@1	0.0041	0.0034	0.0125	0.0181	0.0192	<b>0.0347</b>
	@5	0.0066	0.0062	0.0169	0.0197	0.0208	<b>0.0288</b>
	@10	0.0081	0.0080	0.0202	0.0224	0.0237	<b>0.0306</b>
20	@1	0.0052	0.0029	0.0126	0.0147	0.0166	<b>0.0326</b>
	@5	0.0067	0.0059	0.0163	0.0160	0.0177	<b>0.0278</b>
	@10	0.0088	0.0079	0.0202	0.0197	0.0210	<b>0.0304</b>

Table 3: Precision @N with different latent dimensions  $K$ .

K	Recall	SVD	PMF	NMF	BNMF	F-BNMF	GT-BNMF
10	@1	0.0008	0.0008	0.0049	0.0077	0.0081	<b>0.0061</b>
	@5	0.0038	0.0036	0.0103	0.0121	0.0127	<b>0.0147</b>
	@10	0.0060	0.0060	0.0153	0.0167	0.0176	<b>0.0212</b>
20	@1	0.0011	0.0006	0.0046	0.0059	0.0068	<b>0.0060</b>
	@5	0.0037	0.0034	0.0098	0.0097	0.0107	<b>0.0144</b>
	@10	0.0065	0.0059	0.0151	0.0148	0.0158	<b>0.0210</b>

Table 4: Recall @N with two different latent dimensions  $K$ .

We set the topic number  $K_{\text{topic}} = 30$ ,  $\alpha = 50/K_{\text{topic}}$  and  $\beta = 0.1$  in the aggregated LDA model to derive the user and POI profile topic distribution. For the GT-BNMF model, we set  $\alpha = 1$  and  $\beta = 1$  for latent priors,  $a = 1$  and  $b = 1$  for prior  $\sigma^2$ ,  $1/\mathbb{R}$  for user region multinomial prior  $\gamma$ , and  $\tau = 1$  for the distance model. We set the number of regions  $\mathbb{R} = 49$ , which is the number of regions partitioned according to all the states in USA (except HI and AK).

### 3.6 Performance Comparisons

In this subsection, we present the performance comparison on recommendation accuracies between our model and the baseline methods. We compare the results by setting two latent dimensions  $K = 10$  and  $K = 20$ .

**Precision and Recall @N.** Table 3 and Table 4 show the precision and recall@N ( $N = 1, 5, 10$ ) of all the methods. It is clear that GT-BNMF is significantly better than all the baseline methods. For example, for top 10 recommendation when  $K = 10$ , SVD and PMF have precisions around 0.008, NMF and BNMF improve SVD and PMF about 1.5 times. F-BNMF can slightly improve BNMF, and GT-BNMF significantly outperforms F-BNMF. More clear comparisons can be observed in relative performance as described below.

**Relative performance @N.** There are two reasons that relative performance @N serves as a better metric than the absolute precision and recall. First, please note that for baseline methods, we make top- $N$  recommendation within the initialized  $\mathbb{R}$  clustered regions by  $K$ -means; for the proposed GT-BNMF model, we recommend top- $N$  POIs within each learned latent region. So the candidate POIs can be different. We need a metric can handle such a change of candidate POIs. Second, we need to see how the models actually outperform a universal baseline, the random recommendation. The relative performance as defined in Equation (15) can satisfy these two requirements. A relative performance value equal to 1 indicates a random recommendation.

The relative performances as shown in Table 5 give a more clear comparisons between our GT-BNMF and baseline methods. The relative performance values for SVD and PMF are around 1, which means that the performances of these two methods are close to a random guess. For top-1 recommendation, when latent dimension  $K = 10$ , GT-BNMF can achieve a 41.8 times improvements as compared 3.7 for NMF, 7.1 for BNMF, and 7.3 for F-BNMF. For top-5

K	@N	SVD	PMF	NMF	BNMF	F-BNMF	GT-BNMF
10	@1	0.8729	0.7280	4.7736	8.1243	8.3408	<b>42.8835</b>
	@5	0.9345	0.8251	2.5871	3.0713	3.2956	<b>21.0357</b>
	@10	0.9829	0.8856	2.4480	2.8588	2.9585	<b>16.1661</b>
20	@1	1.0148	0.7124	4.1618	5.8585	6.6864	<b>42.0183</b>
	@5	0.9287	0.8271	2.6095	2.6219	2.8506	<b>20.4742</b>
	@10	1.0084	0.9067	2.5131	2.4717	2.6043	<b>16.0662</b>

Table 5: The relative performance @N with two different latent dimensions (initiated by  $K$ -means). It measures the improvement over a random recommendation.

and top-10 recommendation, our proposed method can also achieve 20.0 times and 15.1 times improvements respectively, which are significantly better than all the baselines<sup>2</sup>. The improvements can be clearly observed.

**Summary.** Due to the assumption of a Gaussian over the product of user and item latent factors  $\mathbf{u}^\top \mathbf{v}$ , SVD and PMF perform poorly as they cannot handle the skewed count data. Some previous studies, such as [18], suggest to normalize the data through some transfer function and bound the results by a logistic function. This can work well if the data range is small. However, when the data is skewed, this method can make the normalized values indistinguishable. Thus, we call for a reconsideration of the recommendation assumption when we face the wide range count data in implicit feedback recommendation scenarios. We observe improvements by NMF and BNMF over SVD and PMF. The F-BNMF method which exploits an *ad hoc* two-steps to fuse the geographical influence into user preferences can only slightly improve BNMF. Therefore, an integrated analysis of multiple factors is needed for POI recommendations.

The GT-BNMF model not only considers the geographical information of POIs, regional popularity and user mobility patterns for recommendation, but also updates the latent regions by considering these three sources of information. The learned regions reflect the collaborative user activity pattern. As a result, we can observe significant improvements over all the baseline algorithms.

### 3.7 Implications of Latent Regions

In addition to improving recommendation performances, our proposed model also provides an unique perspective on POI marketing segmentation in the form of the learned regions. Traditional geographic segmentation for target market is segmented according to geographic criteria such as states [13] and  $K$ -means. Here, we intend to answer how the proposed method with learned latent regions improves the method based on traditional geographic segmentation according to administrative region division (states)? The recommendation comparisons initialized by  $K$ -means are shown in Section 3.6.

**Performance Comparisons.** We first evaluate whether the GT-BNMF model is robust to region initiations. We compare the relative performances initiated by  $K$ -means and initiated by administrative region partition (state). Instead of using  $K$ -means for initial region clustering, we initialize our model with region partition according to administrative region division of states. Table 6 shows the relative performance comparisons of our proposed GT-BNMF model

<sup>2</sup>The best relative performance reported in [24] is 11. This provides an average relative performance value even though we have a different recommendation task here.

K	@N	SVD	PMF	NMF	BNMF	F-BNMF	GT-BNMF
10	@1	1.4271	0.6106	4.4885	6.0672	6.1341	<b>48.4512</b>
	@5	0.9628	0.8550	2.5644	2.9042	3.1243	<b>17.4855</b>
	@10	0.9161	0.8354	2.5335	2.3370	2.5497	<b>18.9053</b>
20	@1	1.1717	0.5414	3.7472	6.1969	6.5094	<b>46.8855</b>
	@5	1.1172	0.7180	2.7691	2.9479	3.0560	<b>17.4738</b>
	@10	0.9642	0.8181	2.5067	2.5391	2.6679	<b>18.8852</b>

Table 6: The relative performance @N initialized according to administrative region partition.

and baseline methods. The comparison results are similar as discussed before. For baseline methods, we make top- $N$  recommendation within the initialized regions according to administrative region division (states). For the proposed GT-BNMF model, we recommend top- $N$  POIs within each learned latent region by the model. We can observe similar performance improvements of our proposed method over baselines as that initiated by  $K$ -means.

Figure 4 shows the relative performance initiated by two methods. We can see that GT-BNMF initiated by states outperforms GT-BNMF initiated by  $K$ -means for top 1 and top 10 recommendation, but GT-BNMF initiated by  $K$ -means outperforms GT-BNMF initiated by states for top 5 recommendation. In general, the GT-BNMF model can achieve a relative stable performance initiated by two methods. Note that the GT-BNMF model can update the latent regions towards a region segmentation that reflects geographical influence and user activity efforts.

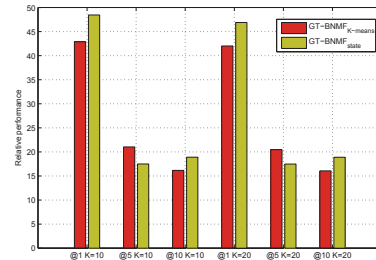


Figure 4: Relative performance of GT-BNMF initiated by  $K$ -means and initiated according to state partition.

**Latent Region Analysis.** We then analyze the latent regions identified by our model. Figure 6 shows the Voronoi visualization of regions identified by the proposed GT-BNMF model compared with two different initiations,  $K$ -means and according to state partition. Figure 6b shows the regions identified by GT-BNMF vs its initiation by  $K$ -means (Figure 6a). Figure 6d shows the region update by GT-BNMF vs its initiation according to state partition (Figure 6c).

We take a representative hotspot, California, as an example to analyze the region update learned by the GT-BNMF model. Figure 5 shows the zoom in visualization of latent regions (Figure 5b) learned from our model vs its initiation by  $K$ -means (Figure 5a) in California. Though we have no ground truth about an optimal POI region segmentation, we can infer the user activity regions in CA through the collaborative check-in behaviors of users who have visited CA and view the region clusters formulated by collaborative check-in effort as ground truth (as shown in Figure 5c). Through analyzing the collaborative check-in frequency by those users, as shown in Figure 5c, we can see two clear clusters in north

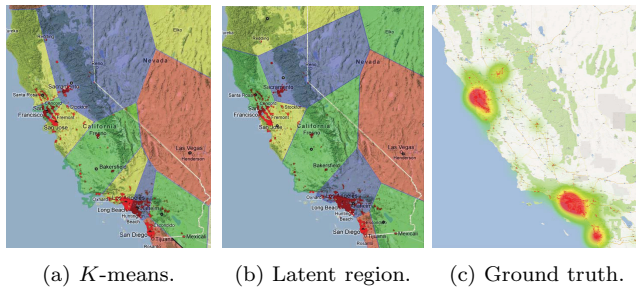


Figure 5: Voronoi visualization of POI segmentation in California area. (b) latent regions learned from our model and (a) initiation by  $K$ -means. (c) true user collaborative activity clusters. Deeper color (red) indicates more check-ins for a POI, as contrary to light color (green). Best view in color.

California area among other scattered POIs around, one in Los Angeles area, one in San Diego, and some scattered POIs between south and north California.  $K$ -means only depends on POI distances to cluster POIs for region segmentation. As shown in Figure 5a,  $K$ -means segments north California area 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 Figure 5b, which is more coherent to real user activity as shown in Figure 5c. Similar results can also be observed in Figure 6 when GT-BNMF is initiated according to administrative region partition. GT-BNMF initiated by  $K$ -means leads to better POI segmentation. We can see that GT-BNMF models not only improves recommendation performances, but also provides an interesting perspective on POI marketing segmentation in the form of the learned regions.

## 4. 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. In contrast, recommender systems can also be developed based on implicit user feedback [11], such as the search and click on a web site [17, 5], 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. For instance, NMF [14] based method is used in recommendation due to its effectiveness in modeling implicit user feedback. Some studies [17, 5] assumed a Poisson factor model with Gamma priors. Actually, the Poisson factor analysis can be equal to NMF in certain condition [14]. Euclidean distance based optimization leads to Gaussian NMF, and KL-divergence based optimization leads to Poisson NMF. For the purpose of efficiency, we adopt a Bayesian probabilistic non-negative matrix factorization (NMF) method [19]. One drawback of latent factor models is the so called cold-start problem, in which latent factor models perform poorly when there are unseen users or items. Recent works such as [22, 1, 2] incorporated side information into latent factor models by exploring user observable features and item observable features. In this paper, we applied a topic model to get the

observable user features and item features, and leverage the topic distribution for the cold-start problem.

Some previous studies on POI recommendations mainly relied on user trajectory data to infer user preferences. For example, previous works [27, 26, 9, 16] applied collaborative filtering (CF) based method to recommend locations and travel packages based on user trajectory data. Actually, these works only consider one dimension factor of user POI check-in decision process. More recent works, such as [23, 3], began to explore user preferences, social influence, and geographical influence for POI recommendations. However, it only used a simple CF algorithm to fuse these information, and thus lacking a comprehensive way to model how these information collectively influence user POI check-in decision. Work [15] 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.

More recently, Cheng et al. [6] 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. It considered the multi-center pattern of user check-in behaviors, but it 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 would 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, word-of-mouth and skewed data for POI recommendation.

In addition, our work has a connection with recent works on mobility modeling [10, 8]. However, their tasks were different. Work [10] used similar multinomial assumption over different regions to model geographical topics in Twitter stream, paper [8] investigated human mobility for social network analysis. Also, people have used gaussian geographical distribution to model region over locations [20, 25, 10].

As described above, while there are some studies on POI recommendation, it lacks of 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 preferences, geographical influences, user mobility pattern, word-of-mouth, and the skewed implicit user check-in counts data. The proposed method strategically takes all these factors into consideration and presents a flexible probabilistic generative model for POI recommendations.

## 5. CONCLUSION

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 of geographical factors in LBSNs, such as the regional popularity and the Tobler's first law of geography. Second, the methods effectively modeled the user



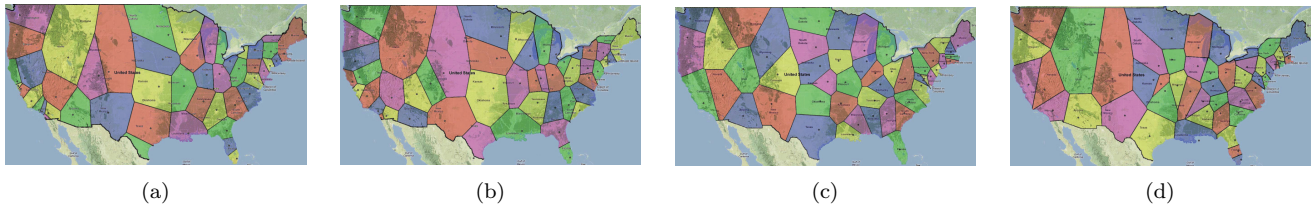


Figure 6: Voronoi visualization of POI segmentation comparisons. Latent regions (b) identified by our model GT-BNMF vs its initiation by  $K$ -means (a); Latent regions (d) updated by GT-BNMF vs its initiation according to administrative region partition (c). We exclude the POI points here, please refer to Figure 2a for geographical POI distributions.

mobility patterns, which are important for location based services. Third, the proposed approach extended the latent factor in 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 later factor models which are suitable for both explicit and implicit feedback recommendation settings. Finally, experimental results on real-world LBSNs data validated the performances of the proposed method.

## 6. ACKNOWLEDGEMENT

This research was partially supported by 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).

## 7. REFERENCES

- [1] D. Agarwal and B.-C. Chen. flda: matrix factorization through latent dirichlet allocation. *WSDM'10*, pages 91–100.
- [2] D. Agarwal and B.-C. Chen. Regression-based latent factor models. *KDD '09*, pages 19–28, 2009.
- [3] J. Bao, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. pages 199–208, 2012.
- [4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, Mar. 2003.
- [5] Y. Chen, M. Kapralov, D. Pavlov, and J. Canny. Factor modeling for advertisement targeting. *NIPS'09*, pages 324–332. 2009.
- [6] C. Cheng, H. Yang, I. King, and M. R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In *AAAI*, 2012.
- [7] Z. Cheng, J. Caverlee, K. Y. Kamath, and K. Lee. Toward traffic-driven location-based web search. *CIKM'11*, pages 805–814, 2011.
- [8] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. *KDD '11*, pages 1082–1090, 2011.
- [9] Y. Ge, H. Xiong, A. Tuzhilin, K. Xiao, M. Gruteser, and M. Pazzani. An energy-efficient mobile recommender system. In *KDD*, KDD '10, pages 899–908, 2010.
- [10] L. Hong, A. Ahmed, S. Gurumurthy, A. J. Smola, and K. Tsioutsoulis. Discovering geographical topics in the twitter stream. *WWW '12*, pages 769–778, 2012.
- [11] Y. Hu, Y. Koren, and C. Volinsky. Collaborative filtering for implicit feedback datasets. *ICDM '08*, pages 263–272, 2008.
- [12] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, Aug. 2009.
- [13] P. Kotler, K. Keller, M. Brady, M. Goodman, and T. Hansen. *Marketing Management*. Pearson Prentice Hall, 2006.
- [14] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *NIPS*, pages 556–562, 2000.
- [15] B. Liu and H. Xiong. Point-of-interest recommendation in location based social networks with topic and location awareness. *SDM'13*, 2013.
- [16] Q. Liu, Y. Ge, Z. Li, E. Chen, and H. Xiong. Personalized travel package recommendation. In *ICDM*, pages 407–416, 2011.
- [17] H. Ma, C. Liu, I. King, and M. R. Lyu. Probabilistic factor models for web site recommendation. *SIGIR '11*, pages 265–274, 2011.
- [18] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *NIPS*, volume 20, 2008.
- [19] M. N. Schmidt, O. Winther, and L. K. Hansen. Bayesian non-negative matrix factorization. *ICA'09*, pages 540–547, 2009.
- [20] S. Sizov. Geofolk: latent spatial semantics in web 2.0 social media. *WSDM '10*, pages 281–290, 2010.
- [21] W. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46(2):234–240, 1970.
- [22] C. Wang and D. M. Blei. Collaborative topic modeling for recommending scientific articles. In *KDD 2011*, pages 448–456, 2011.
- [23] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point of interest recommendation. *SIGIR '11*, pages 325–334.
- [24] P. Yin, P. Luo, W.-C. Lee, and M. Wang. App recommendation: a contest between satisfaction and temptation. *WSDM '13*, pages 395–404, 2013.
- [25] Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang. Geographical topic discovery and comparison. *WWW '11*, pages 247–256.
- [26] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. *WWW'10*, pages 1029–1038.
- [27] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW 2009*, pages 791–800, 2009.