POI Recommendation: A Temporal Matching between POI Popularity and User Regularity

Zijun Yao†, Yanjie Fu‡, Bin Liu§, Yanchi Liu†, Hui Xiong†*
†Rutgers University ‡Missouri University of Science and Technology §IBM Thomas J. Watson Research Center
Email: †{zijun.yao, yanchi.liu, hxiong}@rutgers.edu ‡fuyan@mst.edu §bin.liu1@ibm.com

***Abstract*—Point of interest (POI) recommendation, which provides personalized recommendation of places to mobile users, is an important task in location-based social networks (LBSNs). However, quite different from traditional interest-oriented merchandise recommendation, POI recommendation is more complex due to the *timing* effects: we need to examine whether the POI fits a user's availability. While there are some prior studies which included the temporal effect into POI recommendations, they overlooked the compatibility between time-varying popularity of POIs and regular availability of users, which we believe has a non-negligible impact on user decision-making. To this end, in this paper, we present a novel method which incorporates the degree of temporal matching between users and POIs into personalized POI recommendations. Specifically, we first profile the temporal popularity of POIs to show when a POI is popular for visit by mining the spatio-temporal human mobility and POI category data. Secondly, we propose latent user regularities to characterize when a user is regularly available for exploring POIs, which is learned with a user-POI temporal matching function. Finally, results of extensive experiments with real-world POI check-in and human mobility data demonstrate that our proposed user-POI temporal matching method delivers substantial advantages over baseline models for POI recommendation tasks.**

I. Introduction

The rapid development of GPS equipped mobile devices (e.g., smartphones) has powered large location-based social networks (LBSNs) (e.g., Foursquare), raised the number of mobile users, and enabled various location-based services (LBS). Using these LBS, users share their experiences of places, also known as Point of interests (POIs) such as restaurants or museums. Meanwhile, data collected through LBS activity enable better personalized recommendations of POIs. As a result, POI recommendation, which suggests personalized POIs to users, becomes an important component to improve user experiences and services provided by LBS.

Different from traditional interest-oriented merchandise recommendation (e.g., books, films, etc.), POI recommendation is more complex and challenging due to the unique characteristics of LBS. Firstly, besides personal interest, the timing of recommended POIs should be compatible with users' personal availability. For example, if a user is usually available to explore POIs during morning hours, he would be more likely to visit POIs with morning popularity (e.g., coffee shops, brunch restaurants). Similarly, if a POI is more popular during night hours (e.g., bars), it is more rational to recommend it to users who are available at nights. Secondly, area activity (or volume of people in an area) changes over time as people concentration to different places at different times throughout a day (e.g., work, entertainment). The area where recommended POIs reside should be active at a given time to increase the chance of visiting. For example, in the morning of weekdays, users are concentrated surrounding office/business locations, while at night time of weekends, nightlife and restaurant regions are most active.

Recent studies have considered temporal influences on POI recommendation, such as time-aware POI recommendation which recommends different POIs to users at different time. For example, [1] applies user-item matrix factorization for each time slot and assumes every user has similar preferences in consecutive time slots for regularization. [2] computes user similarity via the same spatio-temporal check-ins in the past and conducts a user-based recommendation approach. [3] adds the time dimension to user-item matrix and applies tensor factorization for recommendations. However, these studies overlooked temporal regularity of users, and time-varying popularity of POI. They also didn't consider the influence of temporal compatibility between users and POIs. In addition, they solely depended on the time input of history check-ins, and suffered from the sparsity problem of check-in data. Last, these studies didn't fully utilize spatio-temporal human mobility patterns which reflect the changes of areas' activity over time. In order to address these limitations, in this paper, we introduce a novel model which incorporates the temporal compatibility between user regularities and POI popularities into POI recommendation, and utilize human mobility data to boost recommendation performances.

In this paper, we propose a Temporal Matching Poisson Factorization Model (TM-PFM) to profile the popularity of POIs, model the regularity of users, and incorporate the temporal matching between users and POIs into overall recommending consideration. We first present a new framework to profile a time-varying popularity of POIs (e.g., hourly visiting change) in a day. Traditional methods usually capture this temporal variation by counting POIs' check-in frequencies therefore suffer from check-in data sparsity. Previous studies [4], [5] and [6] have demonstrated that human mobility is highly regular and predictable, and human mobility data from heterogeneous sources display similar patterns. Therefore, we utilize heterogeneous human mobility data to evaluate POI popularity. The benefits of employing human mobility data

* Contact author.

2374-8486/16 $31.00 © 2016 IEEE
DOI 10.1109/ICDM.2016.137

IEEE computer society

include (i) it is more abundant and less biased than check-in data, and (ii) it reveals which areas are currently active which is a determinant of POI popularity. Moreover, we further analyze POIs by categories and adopt a mixture model to obtain the final POI temporal popularity pattern. Secondly, except some particular events (e.g., parties, concerts), people's availability is usually determined by their routines, thus there is a predictable regularity. Therefore, we consider temporal regularity of each user which describes their regular available time every day for POI exploration. We propose to learn the latent regularity patterns of users by finding the best match with the popularity patterns of visited POI based on check-in frequencies. Finally, with the learned user regularity, we are able to match users with POIs they have not visited yet, and evaluate the temporal matching degree and the general user-POI interest to make recommendations.

In summary, in this paper we propose a novel temporal matching method between users and POIs for POI recommendation, and strategically leverage rich spatio-temporal human mobility data to boost the performance of the model. We highlight our key contributions as follows:

- We propose a factorization based POI recommendation model which incorporates the temporal matching between user regularity and POI popularity to improve POI recommendations.
- We present a novel framework which utilizes heterogeneous human mobility data to profile time-varying popularity of POIs which bypass the check-in data sparsity issue. Meanwhile, we model users' temporal regularity by incorporating user-POI temporal matching into preference estimation.
- We validate our proposed method with real-world LBSN check-in and human mobility datasets. The effectiveness of temporal matching in POI recommendation is proven by extensive experiments and a substantial improvement in recommendation performances over baseline methods is demonstrated.

II. Methodology Overview

We first provide some basic concepts in LBS, then formulate the problem of POI recommendation, and finally show the overview of the proposed temporal pattern matching based framework.

A. Preliminary

DEFINITION 1: (**Check-in**) A **check-in** is an event that a LBSN user reports his/her physical visit to a POI. Generally, a check-in contains the following information: LBSN user, check-in POI with location (e.g., longitude and latitude), category (e.g., Italian restaurant), and check-in timestamp.

DEFINITION 2: (**Taxi trip**) A **taxi trip** is a route that a taxi delivers passengers from one location to another. Every taxi trip starts with a passenger **pick-up** event and ends with a passenger **drop-off** event. Each pick-up and drop-off contains the information of location and timestamp.

TABLE I: Mathematical Notations.

Symbol	Size	Description
$\boldsymbol{Y}$	$M \times N$	user-POI check-in count matrix
$\boldsymbol{\mathcal{T}}$	1×2	day type = $\{wd(\text{weekday}), we(\text{weekend})\}$
$\boldsymbol{Q^*}$	$N \times S$	POI temporal popularity matrix, $* \in \boldsymbol{\mathcal{T}}$
$\boldsymbol{P^*}$	$M \times S$	user temporal regularity matrix, $* \in \boldsymbol{\mathcal{T}}$
$\boldsymbol{U}$	$M \times K$	user latent factor matrix
$\boldsymbol{V}$	$N \times K$	item latent factor matrix
$\boldsymbol{\mu^*}$	$1 \times M$	user temporal regularity parameter vector, $* \in \boldsymbol{\mathcal{T}}$

B. Problem Definition

Let $U = \{u_1, u_2, ..., u_M\}$ be a set of LBSN users and $V = \{v_1, v_2, ..., v_N\}$ be a set of POIs where each POI has a location (e.g., latitude and longitude). Consider the existence of the historical check-ins where each record indicates a user u_i checked into a POI v_j once, we can extract the check-in number that u_i preformed check-in to v_j, named y_{ij}. **The objective** of personalized POI recommendation is to recommend POIs to users based on personal check-in history. In addition, we integrate the large-scale spatio-temporal taxi trip data, where each trip ends with a drop-off event which indicates human arrivals with location and timestamp. We refer i as user and j as POI in following sections for simplicity. The important notations used in this paper are listed in Table I.

C. General Framework

We propose a two-step method which includes (i) profiling temporal patterns of POI popularity and (ii) modeling temporal matching of user-POI pairs.

Step 1: Profiling Temporal Patterns of POI Popularity. We aim at profiling the temporal popularity of POIs which describes how the popularity of a POI varies during a day. Specifically, we split a day into S equal-sized time slots (e.g., 24 hours), and each time slot is associated with a probability describing the ratio of the in time slot visit volume to the whole-day visit volume. We package these S probabilities chronologically as a vector which is the temporal pattern of popularity to be profiled. To achieve this, we first extract the area activity (e.g., how many active people in the area) around POI locations by utilizing human mobility data. Subsequently, we extract the category popularity by aggregating the check-in frequencies at the POI category level to refine the profiling. Lastly, we use a mixture model to smooth and further characterize the temporal pattern of POI popularity.

Step 2: Modeling Temporal Matching of user-POI pairs. We aim to develop a user-POI temporal matching model to infer the temporal regularity of users. First, we consider that each user has regular available times every day due to personal routines. Meanwhile, users are more likely to visit a POI at its popular times. We associate S equal-sized time slots with probabilities to show how likely a user may explore POIs during a specific time slot in a day. By vectorizing these probabilities, a user's temporal regularities are defined. Furthermore, we present a function which matches a user's latent regularity with a POI's profiled popularity. We combine the temporal matching degree with the general interest as the overall preferences. Finally, we learn the users' temporal regularity by optimizing the distance between the estimated preferences and the frequencies of history check-in at POIs.

III. Profiling Temporal Patterns of POI Popularity

In this section, we introduce how to profile the temporal popularity for POIs. Intuitively, counting the check-in frequency during each time slot for a POI can complete this job. However, the POI level check-in records are too few to provide valid results. Thus we propose to alternatively analyze the temporal popularity in an implicit way. Generally, the current popularity of a POI is affected by two aspects: (i) how many active people are around the POI, and (ii) what type of service this POI provides. For the former one, we assess the area activity by mining how many people come to a POI's area during a time slot with taxi trip data. For the later aspect, we profile the category popularity by answering how many people visit a POI category in a time slot with check-in and POI category data. We combine these two effects to generate rough popularity patterns for POIs. Last, we utilize the mixture Gaussian model to smooth and characterize the popularity variations to obtain the final popularity patterns.

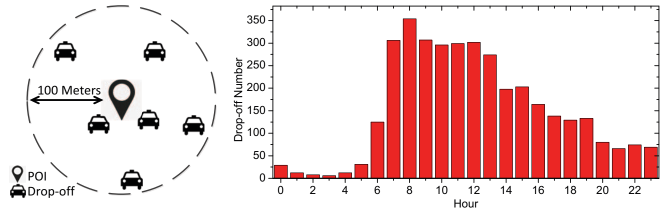
At the beginning, let us define the temporal popularity. We assign a unique popularity pattern to every POI to describe the visit volume changes over time slots every day. To profile this temporal pattern, we use a size-S vector to represent the ratio of each time slot's visits to the whole day's visits. All the ratios are organized chronologically and their sum for a day equals to 1. Usually, the temporal pattern of a POI's popularity changes largely from weekdays (Monday to Friday) to weekends (Saturday, Sunday). Therefore, for each POI, we identify two types of temporal pattern: (1) weekday pattern $\boldsymbol{q_j^{wd}}$ and (2) weekend pattern $\boldsymbol{q_j^{we}}$. Formally, we denote temporal patterns of a POI's popularity as following:

$$\boldsymbol{q_j^*} = \{q_{j,1}^*, ..., q_{j,S}^*\}, \quad * \in \{wd, we\}, \tag{1}$$

where $q_{j,s}^*$ represents the probability that visitors will check-in to the POI j in the time slot s with respect to weekday wd or weekend we. For each $\boldsymbol{q_j^*}$ we have $\sum_{s=1}^{S} q_{j,s}^* = 1$ and $q_{j,s}^* \geq 0$.

A. Assessing Temporal Patterns with Area Activity

Every day, people concentrate to different places at different times for daily purposes (e.g., working, entertaining). Given a particular time, if a POI's is in the area where contains high volume of people, the POI is expecting to have more visits. Taxi is a fundamental transportation tool for people who live in large cities (e.g., New York City). Since each taxi trip ends with a destination, given massive and comprehensive taxi trips of a city, we are able to know where concentrates high volume of people at different times. Therefore, we collect taxi drop-offs which happened within walking distance (e.g., 100 meters) of each POI's location as shown in Figure 1a. The reason of choosing 100-meter for drop-off collection is that a longer distance makes the collection area too large that the unique characteristic of POI location can not be captured, meanwhile, a shorter distance may not cover the nearest street crossing or road segment, thus POI locations may not be able to collect enough drop-offs to profile area activity. We count



(a) (b)

Fig. 1: (a) Method of collecting taxi drop-offs for a POI, (b) time-varying taxi drop-offs around an office POI.

taxi drop-offs by time slots and day types (e.g., the drop-offs during 10AM-11AM in weekend days) as the example shown in Figure 1b. Through this taxi data processing, we profile the temporal pattern of area activity around POIs and denote them as:

$$\boldsymbol{\mathcal{D}_j^*} = \{\mathcal{D}_{j,1}^*, ..., \mathcal{D}_{j,S}^*\}, \quad * \in \{wd, we\}, \tag{2}$$

where $\mathcal{D}_{j,s}^*$ represents the portion of taxi drop-offs around POI j during s-th time slot in a type of day. For each $\boldsymbol{\mathcal{D}_j^*}$ we have $\sum_{s=1}^{S} \mathcal{D}_{j,s}^* = 1$ and $\mathcal{D}_{j,s}^* \geq 0$.

B. Refining Temporal Patterns with Category Popularity

At the same time, the popularity pattern of a POI is not only dominated by area activity but also related to its category. For example, at mid-night, even though an area may be highly active by having many visits, a museum at this place can not be popular. Therefore, the profiled patterns based on area activity need to be further refined by integrating the category popularity. At the level of POI category (e.g., department stores), the sparsity problem of check-in data is alleviated. Therefore, we count check-ins frequencies for categories over time slots. We denote the category pattern as:

$$\boldsymbol{\mathcal{C}_v^*} = \{\mathcal{C}_{v,1}^*, ..., \mathcal{C}_{v,S}^*\}, \quad * \in \{wd, we\}, \tag{3}$$

where $\mathcal{C}_{v,s}^*$ represents the portion of check-in at the category v during s-th time slot in a type of day. For each $\boldsymbol{\mathcal{C}_v^*}$ we have $\sum_{s=1}^{S} \mathcal{C}_{v,s}^* = 1$ and $\mathcal{C}_{v,s}^* \geq 0$.

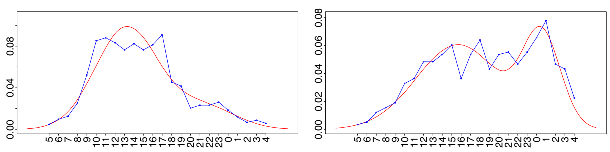
Next, by combining the effects of category popularity with the effects of area activity, we obtain the refined POI temporal popularity which is more close to the reality. We denote the combined temporal popularity of POIs as $\boldsymbol{q'^*_j}$, whose probability for each time slot is:

$$q'^*_{j,s} = \varphi \mathcal{D}_{j,s}^* + (1-\varphi)\mathcal{C}_{c(j),s}^*, \quad * \in \{wd, we\}, \tag{4}$$

where $c(j)$ is the operation to get the category v of POI j, $0 < \varphi < 1$ controls the weights.

C. Enhancing Temporal Patterns with Mixture Model

In the last part of temporal popularity profiling, we want to describe each temporal pattern with a proper distribution. The first motivation is to smooth the visit probability over time slots because artificial spiting of drop-offs into time slots may cause volatile patterns especially in adjacent time slots as shown in Figure 2. Another motivation which is more important is that we want to strategically characterize the



(a) (b)

Fig. 2: Example of two POI popularity patterns in hours of day. Blue and Red: before and after GMM smoothing.

popularity pattern to be more discriminative by weakening the idle time slots and highlighting the popular time slots.

To achieve the requirements raised by above motivations, we propose to adopt Gaussian Mixture Model (GMM) to model popularity patterns for several advantages. First, GMM can express one or more visit peaks in a day as a POI usually behaves in reality. Second, the Gaussian distribution can well simulate the process of visit changes of POIs. For example, a POI's popularity often starts from idle to busy and gets back to idle. Usually one process last for several time slots and the popularity changes smoothly. Third, for the idle times, the visit probability are weaken. Meanwhile, for the busy times, probability are enhanced and concentrated to the peak point. Therefore, we formally define the probability over time slots of a POI popularity pattern $\boldsymbol{q_j^*}$ with GMM as:

$$q_{j,s}^* = \sum_{r=1}^{R} w_{j,r}^* \cdot \mathcal{N}(s|\mu_{j,r}^*, \sigma_{j,r}^{*2}), \quad * \in \{wd, we\}, \tag{5}$$

where s represents the s-th time slots and R represents the number of Gaussian components in a daily temporal pattern. $w_{j,r}^*$ represents the mixture weight of rth Gaussian distribution. In our observation, most of the POIs have no more than two visiting peaks in a day such as restaurants. Therefore we predefine the number of Gaussian components $R = 2$ for all GMM modeling. Input the POI temporal patterns $\boldsymbol{q'^*_j}$ we obtained in previous step, we apply Expectation–Maximization (EM) algorithm to estimate the GMM for each pattern as shown in Figure 2. Last, we obtain the final popularity patterns $\boldsymbol{q_j^*}$ for each POI.

IV. Recommendations via Temporal Matching

In this section, we first introduce how to model the temporal matching between user and POI, then we present the parameter estimation of the model.

A. Model Specification

To generate recommendation of a POI j for a user i, we assume the overall preference on the user-POI pair f_{ij} is impacted by (i) the user-POI general interest score, $\delta(i, j)$, and (ii) the user-POI temporal matching score, $m(i, j)$:

$$f_{ij} = \delta(i, j) \cdot m(i, j). \tag{6}$$

The user-POI general interest score $\delta(i, j)$ is learned from classic matrix factorization methods, by combining K-dimensional user latent factor vector $\boldsymbol{u_i}$ and POI latent factor vector $\boldsymbol{v_j}$ as follows: $\delta(i, j) = \boldsymbol{u_i}^\top \boldsymbol{v_j}$. The user-POI temporal matching score $m(i, j)$ is the degree of matching between users and POIs, based on S-dimensional user temporal regularity vectors $\boldsymbol{\rho_i^*}$ and POI temporal popularity vectors $\boldsymbol{q_j^*}$, where $* \in \{wd, we\}$, wd and we respectively represent the day type of weekday and weekend. Next, we present the detailed temporal matching modeling.

Capturing User Daily Temporal Regularity Except some special events, the available hours for exploring POIs are usually regular for users due to personal daily routines. For example, if a user always have a long lunch break, thus he may regularly explore POIs during 12PM to 2PM. Therefore, we propose that every LBSN user has a latent daily-repeated personalized temporal regularity which decides when he/she is likely to explore POIs every day. Usually a individual's temporal regularities is different in weekday and weekend, we define two types of daily temporal regularities for each user:

$$\boldsymbol{\rho_i^*} = \{\rho_{i,1}^*, ..., \rho_{i,S}^*\}, \quad * \in \{wd, we\}, \tag{7}$$

where $\rho_{i,s}^*$ represent user i's exploring probabilities during time slot s for weekdays wd or weekend we. For each regularity pattern $\boldsymbol{\rho_i^*}$, we have $\sum_{s=1}^{S} \rho_{i,s}^* = 1$ and $\rho_{i,s}^* \geq 0$.

At the same time, we also want to regularize user's availability distribution over time slots. In reality, users usually plan one trip in a day and their availability does not fluctuate largely in adjacent time slots, therefore we assume one window per day for each user for POI exploration. We exploit a Gaussian distribution to regularize each regularity pattern. For $\boldsymbol{\rho_i^*}$, we have the probability in each time slots as:

$$\rho_{i,s}^* = \mathcal{N}(s|\mu_i^*, \varepsilon_i^{*2}), \quad * \in \{wd, we\}. \tag{8}$$

Here we model the check-in probability of s-th time slot as the probability density at s (e.g., $s = 5$).

Modeling User-POI Temporal Matching Here we present how we match the user's temporal regularities with the POI's temporal popularities. The objective of temporal matching for a user-POI pair is to examine if the POI is well-timed for the user's temporal regularity. For example, for a user who explores POI in the morning time, a coffee shop is more well-timed than a bar. Since the popularity pattern can indicate the optimum time slots of POIs, our method is to find out if the regularity pattern of users has any common time slots to favor a POI's popularity. We define the temporal matching score $m(i, j)$ for user i and POI j as following:

$$m(i, j) = \gamma \boldsymbol{\rho_i^{wd\top}} \boldsymbol{q_j^{wd}} + (1 - \gamma) \boldsymbol{\rho_i^{we\top}} \boldsymbol{q_j^{we}}, \tag{9}$$

where $0 < \gamma < 1$ controls the weights of temporal matching score on weekday and weekend. For example, we can assume that the importance of each day of a week would be the same for each user, therefore $\gamma = \frac{5}{7}$ for five days of weekday and the rest $\frac{2}{7}$ for two days of weekend.

In this model, we have four latent variables to be learned: POI interest latent factors $\boldsymbol{v_j}$, user interest latent factors $\boldsymbol{u_i}$, and user daily temporal regularities $\boldsymbol{\rho_i^*}$, where $* \in \{wd, we\}$ for day types of weekday and weekend respectively. $\boldsymbol{v_j}$ and $\boldsymbol{u_i}$ are K-dimensional vectors while $\boldsymbol{\rho_i^*}$ are S-dimensional vectors. Since we model the regularity on every time slot s

to be $\rho^*_{i,s} = \mathcal{N}(s|\mu^*_i, \varepsilon^{*2}_i)$ for user temporal regularity $\boldsymbol{\rho^*_i}$ as Equation (8), we further translate user temporal regularity factors $\boldsymbol{\rho^*_i}$ into μ^*_i and ε^*_i. For reducing parameters to learn and improving computational efficiency, we predefine a unified ε for all user temporal regularities by referring a usual availability window of people (e.g., 4 hours). Therefore, we rewrite the temporal matching score $m(i,j)$ in Equation (9) as following:

$$m(i,j) = \gamma \sum_{s=1}^{S} \mathcal{N}(s|\mu^{wd}_i, \varepsilon^2) q^{wd}_{j,s} + (1-\gamma) \sum_{s=1}^{S} \mathcal{N}(s|\mu^{we}_i, \varepsilon^2) q^{we}_{j,s}. \tag{10}$$

Finally, to infer the latent factors $\boldsymbol{v_j}$, $\boldsymbol{u_i}$ and μ^*_i, we need to formulate the estimated user-POI preference f_{ij} to follow a probability distribution $\Pr(y_{ij}|f_{ij})$, where y_{ij} is the user-POI check-in count as the groundtruth of user preference. Also, since all the user-POI visit count y_{ij} are non-negative, we expect our estimated preference f_{ij} to be non-negative. We use a Bayesian non-negative latent factor model.

Given the heavy skewness and wide range of discrete check-in count data as shown in Figure 3b, we adopt a Poisson distribution to model $\Pr(y_{ij}|f_{ij})$:

$$\begin{aligned} y_{ij} &\sim \text{Poisson}(f_{ij}) \\ \Pr(y_{ij}|f_{ij}) &= (f_{ij})^{y_{ij}} \frac{\exp\{-f_{ij}\}}{y_{ij}!}, \end{aligned} \tag{11}$$

where $f_{ij} = \boldsymbol{u_i}^\top \boldsymbol{v_j} \cdot m(i,j)$ refers to Equation (6), $m(i,j)$ refers to Equation (10).

Furthermore, v_{jk}, u_{ik} can be given Gamma distributions while μ^*_i can be given Gaussian distribution as empirical priors. Therefore, the user-POI preferences can be modeled as a generative process:

1) For each POI j, generate K-dim POI latent factor:

$$v_{jk} \sim Gamma(\alpha_V, \beta_V), \tag{12}$$

2) For each user i, generate K-dim user latent factor:

$$u_{ik} \sim Gamma(\alpha_U, \beta_U), \tag{13}$$

Also, generate user temporal regularity factor for weekday and weekend:

$$\mu^*_i \sim \mathcal{N}(\alpha_\mu, \sigma^2_\mu), \quad * \in \{wd, we\}, \tag{14}$$

3) For each user-POI pair $< i, j >$, generate response:

$$Pr(y_{ij}|\boldsymbol{v_j}, \boldsymbol{u_i}, \mu^{wd}_i, \mu^{we}_i) = (f_{ij})^{y_{ij}} \frac{\exp\{-f_{ij}\}}{y_{ij}!}, \tag{15}$$

where $\Theta = \{\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}\}$ are parameters for estimation, and $\Phi = \{\alpha_V, \beta_V, \alpha_U, \beta_U, \alpha_\mu, \sigma^2_\mu\}$ are hyperparameters.

B. Parameter Estimation

Given the observations of user-POI check-in count $\boldsymbol{Y}$ and the hyperparameters Φ, according to Maximum a posteriori (MAP) estimation, we optimize parameters $\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}$ by maximizing the posterior probability:

$$\begin{aligned} &\Pr(\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}|\boldsymbol{Y}, \Phi) \\ &\propto \Pr(\boldsymbol{Y}|\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}) \Pr(\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}|\Phi). \end{aligned} \tag{16}$$

For $Pr(y_{ij}|\boldsymbol{v_j}, \boldsymbol{u_i}, \mu^{wd}_i, \mu^{wd}_i)$, we use Equation (15) to compute:

$$\begin{aligned} &\Pr(\boldsymbol{Y}|\boldsymbol{V}, \boldsymbol{U}, \boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}, \Phi) \\ &= \prod_{i=1}^{M} \prod_{j=1}^{N} (f_{ij})^{y_{ij}} \frac{\exp\{-f_{ij}\}}{y_{ij}!} \end{aligned} \tag{17}$$

For $\Pr(\boldsymbol{v_j}, \boldsymbol{u_i}, \mu^{wd}_i, \mu^{we}_i|\alpha_V, \beta_V, \alpha_U, \beta_U, \alpha_\mu, \sigma^2_\mu)$ which are the prior distributions of $\boldsymbol{V}$, $\boldsymbol{U}$, $\boldsymbol{\mu^{wd}}$, and $\boldsymbol{\mu^{we}}$, we use Equation (12, 13, 14) to generate:

$$\begin{aligned} \Pr(\boldsymbol{V}|\alpha_V, \beta_V) &= \prod_{j=1}^{N} \prod_{k=1}^{K} \frac{v_{jk}^{\alpha_V - 1} \exp(-v_{jk}/\beta_V)}{\beta_V^{\alpha_V} \Gamma(\alpha_V)} \\ \Pr(\boldsymbol{U}|\alpha_U, \beta_U) &= \prod_{i=1}^{M} \prod_{k=1}^{K} \frac{u_{ik}^{\alpha_U - 1} \exp(-u_{ik}/\beta_U)}{\beta_U^{\alpha_U} \Gamma(\alpha_U)} \\ \Pr(\boldsymbol{\mu^*}|\alpha_\mu, \sigma^2_\mu) &= \prod_{i=1}^{M} \frac{1}{\sigma_\mu \sqrt{2\pi}} \exp\{-\frac{(\mu^*_i - \alpha_\mu)^2}{2\sigma^2_\mu}\}, * \in \{wd, we\}. \end{aligned} \tag{18}$$

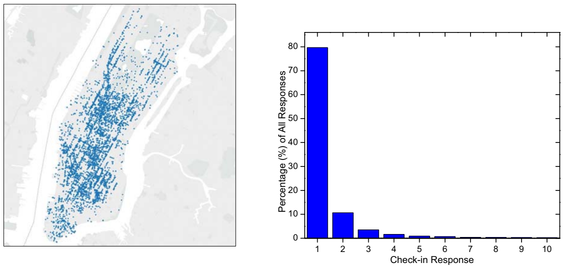
Then we have the log posterior of Equation (16) as:

$$\begin{aligned} \mathcal{L}(\boldsymbol{V}, \boldsymbol{U}, &\boldsymbol{\mu^{wd}}, \boldsymbol{\mu^{we}}|\boldsymbol{Y}, \Phi) = \sum_{i=1}^{M} \sum_{j=1}^{N} (y_{ij} \ln f_{ij} - f_{ij}) \\ &+ \sum_{j=1}^{N} \sum_{k=1}^{K} \Big((\alpha_V - 1) \ln v_{jk} - v_{jk}/\beta_V \Big) \\ &+ \sum_{i=1}^{M} \sum_{k=1}^{K} \Big((\alpha_U - 1) \ln u_{ik} - u_{ik}/\beta_U \Big) \\ &+ \sum_{i=1}^{M} \Big(-\frac{1}{2} \ln \sigma^2_\mu - \frac{(\mu^{wd}_i - \alpha_\mu)^2}{2\sigma^2_\mu} \Big) \\ &+ \sum_{i=1}^{M} \Big(-\frac{1}{2} \ln \sigma^2_\mu - \frac{(\mu^{we}_i - \alpha_\mu)^2}{2\sigma^2_\mu} \Big) + \text{const.} \end{aligned} \tag{19}$$

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

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial v_{jk}} &= \frac{\alpha_V - 1}{v_{jk}} - \frac{1}{\beta_V} + \sum_{i=1}^{M} \Big(\frac{y_{ij}}{f_{ij}} - 1 \Big) u_{ik} \cdot m(i,j) \\ \frac{\partial \mathcal{L}}{\partial u_{ik}} &= \frac{\alpha_U - 1}{u_{ik}} - \frac{1}{\beta_U} + \sum_{j=1}^{N} \Big(\frac{y_{ij}}{f_{ij}} - 1 \Big) v_{jk} \cdot m(i,j) \\ \frac{\partial \mathcal{L}}{\partial \mu^*_i} &= -\frac{\mu^*_i - \alpha_\mu}{\sigma^2_\mu} + \sum_{j=1}^{N} \Bigg((\frac{y_{ij}}{f_{ij}} - 1) \cdot \boldsymbol{u_i}^\top \boldsymbol{v_j} \\ &\cdot \sum_{s=1}^{S} \Big(\frac{\gamma^* q^*_{js} (s - \mu^*_i)}{\varepsilon^3 \sqrt{2\pi}} \exp\{-\frac{(s - \mu^*_i)^2}{2\varepsilon^2}\} \Big) \Bigg), * \in \{wd, we\}, \end{aligned} \tag{20}$$

where γ^{wd} and γ^{we} are γ and $1-\gamma$. We use gradient ascending method to infer the parameters. Specifically, we maximize the posterior by updating parameters as $\upsilon^{(t+1)} \leftarrow \upsilon^{(t)} + \epsilon \times \frac{\partial \mathcal{L}}{\partial \upsilon}$, where υ is an element in $\{\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{\mu^*}\}$, $\frac{\partial \mathcal{L}}{\partial \upsilon}$ is the derivatives according to Equation (20), and ϵ is the learning rate.



(a) (b)

Fig. 3: (a) POI geographical distribution. (b) Check-in response distribution.

V. Experiment

In this section, we empirically evaluate the performance of our proposed methods. We perform all the experiments on read-world datasets: LSBN data from Foursquare, human mobility data from taxi trip records of New York City.

A. Experimental Data

For LBSN dataset, we use the Foursquare dataset which is formulated in work [7]. The dataset includes the check-in data in New York City (NYC) for 10 months (April 2012 to February 2013). Each check-in contains the information such as user ID, POI ID, location, timestamp and POI category. To work with NYC taxis which mainly drive in the city area, we limit the POIs to the most densely populated borough - Manhattan. Also, we remove the users and POIs with too few check-ins (e.g., less than 3) from our dataset to avoid cold start problem. We finalized a dataset of 975 users for 4722 POIs with 64702 check-in observations. The user-POI check-in count matrix has a sparsity of 99.24 percent. Each user performs 66 check-ins to POIs on average. The number of check-ins for a POI ranges from 1 to 257. Figure 3 provides the geographical distribution of POIs as well as the distribution of user-POI check-in responses.

For human mobility data, we use yellow cab trip records from NYC taxi & limousine commission[1] covering the time range of check-in dataset. Due to the large size of taxi trips in NYC, we randomly sample 2 million trips in Manhattan. Each taxi trip contains an origin and a destination with information of location and timestamp.

B. Experimental Metrics

In our experiments, we recommend each user a list of N POIs which have the highest predicted values but are not visited in training set. Then we evaluate the lists based on the recommended POIs which are actually visited by users in testing set.

[1]http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml

Precision and Recall: Given a top-N recommendation list of POIs $L_{N,rec}$, precision and recall are defined as:

$$\begin{aligned} \text{Precision@}N &= \frac{|L_{\text{N,rec}} \cap L_{\text{visited}}|}{N} \\ \text{Recall@}N &= \frac{|L_{\text{N,rec}} \cap L_{\text{visited}}|}{|L_{\text{visited}}|}, \end{aligned} \tag{21}$$

where $L_{\text{N,rec}}$ represents the recommended list of N POIs for a user, and L_{visited} represents the visited POIs of the user in test set. By averaging the precision and call value of all users, we obtain the overall precision and recall for a recommender system.

F-measure: F-measure is the harmonic mean of precision and recall. We adopt a unbalance F-measure F_β which put more emphasis on precision than recall by setting $\beta = 0.5$:

$$F_\beta = (1+\beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{\beta^2 \text{Precision} + \text{Recall}}. \tag{22}$$

C. Baseline Algorithms

The experimental study compares our proposed temporal matching Poisson factor model (TM-PFM) with state-of-the-art factor-based models. Specifically, we compare our proposed TM-PFM model with following algorithms:

- Probabilistic Matrix Factorization (PMF) [8]: a widely used probabilistic factor-based model with Gaussian observation noise.
- Non-negative Matrix Factorization (NMF)[9]: a matrix factorization model with the constrain of non-negative latent variables.
- Bayesian Probabilistic Tensor Factorization (BPTF)[3]: a model which introduces time dimension to traditional user-item factor-based collaborative filtering method.
- Location Recommendation with Temporal effects (LRT)[1]: a factor-based model which learns users' time-aware preferences at separated time slots and use the preference similarity in consecutive times as regularization.

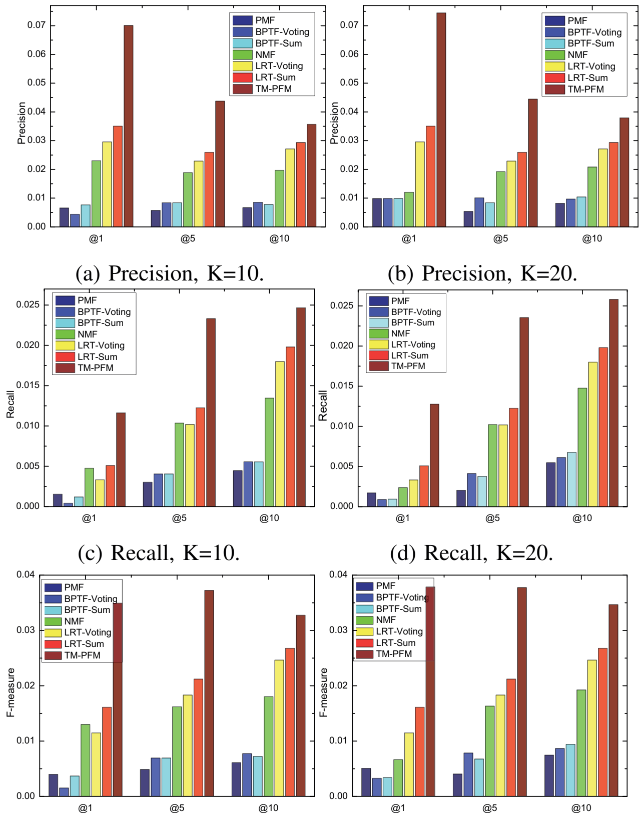
Since BPTF and LRT are temporal recommendation model, therefore, we need to obtain the overall preference for POIs. We aggregate the preference at each time slot by two ways.

- Sum: we consider a user's overall preference on a POI as the sum of his preference at each time slot.
- Voting: for each time slot, we make a separate recommendation list and give the recommended POIs a nomination. The overall preference on a POI is obtained by the number of nominations.

For the experiment setup, we randomly divided the user-item check-in count data into 80 percent for training and 20 percent for testing. We set $\lambda_U = \lambda_V = 0.005$ for PMF. We set $\nu_\alpha = W_\alpha = \beta = 1$ for BPTF. For TM-PFM, we set $\alpha_U = \alpha_V = 4$ when $K = 10$, and $\alpha_U = \alpha_V = 3$ when $K = 20$. For both K, we set $S = 24$, $\varphi = 0.6$, $\gamma = \frac{5}{7}$, $\beta_U = \beta_V = 0.2$, $\varepsilon = 3$, $\alpha = 11.5$, and $\sigma_\mu = 3.5$.

D. Overall Performances

Figure 4 shows the precision@N, recall@N and F_β measure@N ($\beta = 0.5$) of all compared approaches on our dataset.



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

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

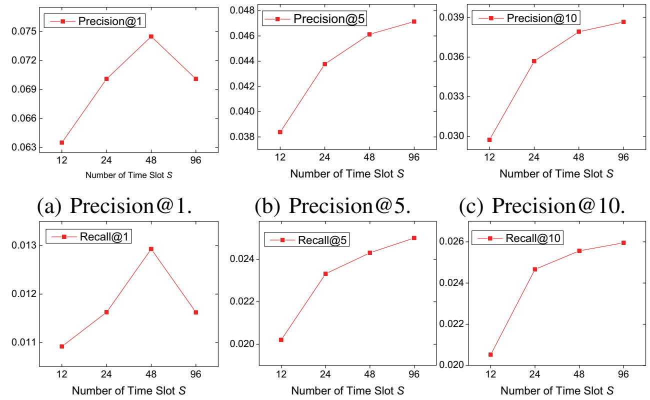
(e) F_β measure, K=10. (f) F_β measure, K=20.

Fig. 4: Precision, recall, and F_β measure @1, @5 and @10 with two different latent dimensions K.

For top-N position, we examine $N = 1, 5, 10$. For latent factor dimension, we explore $K = 10$ and $K = 20$.

Generally, we can see that our proposed approach TM-PFM consistently outperform baseline methods, including traditional recommendation models (PMF, NMF) as well as the temporal recommendation model (BPTF, LRT) for different N and different K. Specifically, we find that PMF performances similarly as BPTF approach with aggregation rule of voting or sum. NMF outperforms the previous three approach (PMF, BPTF-Voting, BPTF-Sum) by making latent variables non-negative. Furthermore, LRT approach (LRT-Sum, LRT-Voting) with either voting or sum rule outperforms NMF by learning time-aware preferences and assuming similarities for consecutive time slots. For temporal recommendation models BPTF and LRT, we find that the sum aggregation rule generally performs better than voting aggregation rule, especially on LRT with quilt significant differences. Last, our proposed TM-PFM model further outperforms LRT significantly on precision, recall and F_β measure, with respect to $K = 10$ and $K = 20$.

At the same time, from the experiment results we can see that the non-negative factor models (NMF, LRT and TM-PFM) preform better than the regular factor models (PMF, BPTF). One reason is that regular models are more suitable for explicit



(a) Precision@1. (b) Precision@5. (c) Precision@10.

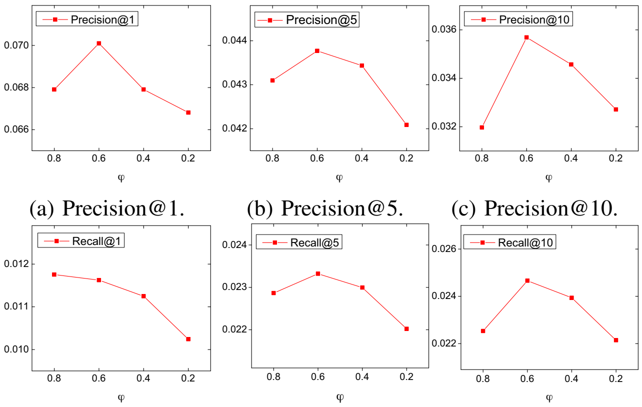
(d) Recall@1. (e) Recall@5. (f) Recall@10.

Fig. 5: Precision and recall of proposed model with different time slot number S of temporal patterns (K=10).

response (e.g., rating), but for implicit response such as check-in count data which is heavily skewed to 1, non-negative models provide better performance. Comparing our proposed model with other non-negative models (NMF, LRT), our model which adopts Poisson observation noise is more appropriate for modeling count data. Also, while the other non-negative models can only apply an approximation of probabilistic generative process, our proposed model provides a more authentic way. Compare to all baseline methods, our model demonstrates the effectiveness of incorporating user-POI temporal matching consideration into POI recommendations.

E. Performance with Different Time Slot Number

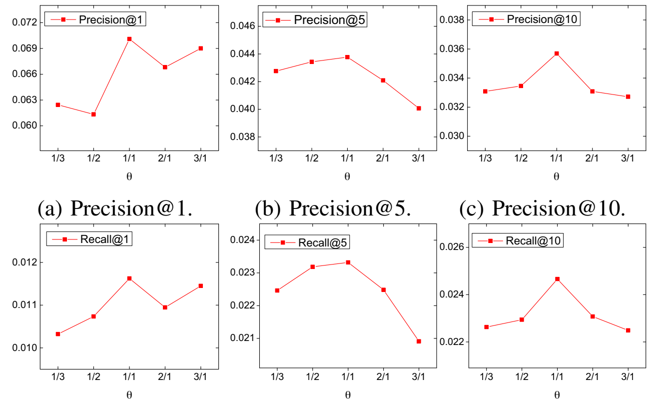
We study the model performance in different time slot numbers as shown in Figure 5. As we equally split one day into multiple time slots to construct temporal patterns, the number of time slots S decide the length of each single slot. We compare four different numbers of time slots in this study: 12 time slots for 2-hour/slot, 24 time slots for 1-hour/slot, 48 time slots for 30-minute/slot and 96 time slots for 15-minute/slot. The larger number of time slots means the patterns of popularity or regularity are more fine-grained. Figure 5 shows the precision and recall performance of model at top-N position 1,5, and 10. We have two observations. Firstly, we can see that, as the number goes higher, the model achieve better performance. The only exception appears at precision@1 and recall@1 where performance decrease a few from 48 time slots to 96 time slots. However, the overall increasing trend still exists. The reason is that the popularity patterns characterize every POI to be more distinctive as the number of time slot goes up. By matching users' regularity with their visited POIs, the regularity patterns can be inferred with finer resolution. Therefore, the performance can be boosted by larger time slot numbers generally. Secondly, we can find out that the increase slows down when time slot number goes large. The largest increase usually happened at 12 time slots. After that, the performance does not increase strongly as before. One reason is that each time slot starts to lack sufficient observations



(a) Precision@1. (b) Precision@5. (c) Precision@10.

(d) Recall@1. (e) Recall@5. (f) Recall@10.

Fig. 6: Precision and recall of proposed model with different weight φ of area-activity for mixing with category popularity $(1-\varphi)$ (K=10).



(a) Precision@1. (b) Precision@5. (c) Precision@10.

(d) Recall@1. (e) Recall@5. (f) Recall@10.

Fig. 7: Precision and recall of proposed model with different weekday/weekend ratio θ for computing temporal match scores (K=10).

(e.g., taxi drop-offs) for popularity profiling as the number of time slot becomes large. On the efficiency aspect, larger time slot number means more computation in temporal matching analysis, thus increase the training time. Therefore, 24 (1 hour per slot) and 48 (30-minute per slot) are relatively optimal time slot numbers which take account of both model performance and computation efficiency.

F. Tuning the Weight φ for Area Activity

As shown in Figure 6, we tune the weight φ of area activity pattern to test the performances of our model. For profiling the final popularity patterns of POIs, we propose to combine the patterns of both area activity and category popularity with a mixing parameter φ. Recalling Equation (4), φ decides the mix ratio of the two patterns. For example, $\varphi = 0.8$ means we combine 0.8 times area activity effects and 0.2 times POI category effects to generate popularity patterns. Here we study what φ value gives good performances with four φ configuration: from $\varphi = 0.8$ which emphasizes more on area activity to $\varphi = 0.2$ which favors more on POI category. Here we can see that $\varphi = 0.6$ provides the best performance and $\varphi = 0.4$ gets the second place. From the observation, we can conclude that both the area activity and the category popularity provide the important knowledge for POI popularity profiling. Specifically, area activity studied by taxi trips makes the larger contribution in POI profiling.

G. Tuning the Ratio θ of Weekday to Weekend

In our original configuration, we assume that each day of the week has the same weight for modeling user-POI temporal matching. Therefore, we set the matching score of a user-POI pair $m(i,j)$ in Equation (9) to have $\gamma = \frac{5}{7}$ because of 5 weekdays in a week while $1-\gamma = \frac{2}{7}$ comes from two days in weekend. However, people may have unbalance weight on the days of weekdays or weekends. In this study, we want to tune the trade-off between weekdays and weekend to explore the day importance of LBSN users in New York City. We use the ratio θ to denote the trade-off:

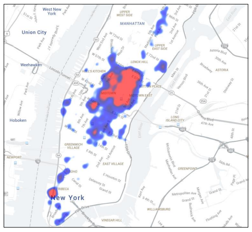
$$\theta = \frac{Day^{wd}}{Day^{we}}, \tag{23}$$

where Day^{wd} denotes the weight of a day of weekday and Day^{we} denotes the weight of a day of weekend. For example, if $\theta = \frac{1}{2}$ which means a day of weekend is twice important than a day of weekday, we have $\gamma = \frac{5}{9}$ for weekday and $1-\gamma = \frac{4}{9}$ for weekend by considering 5 days as weekday and 2 days as weekend in a week.

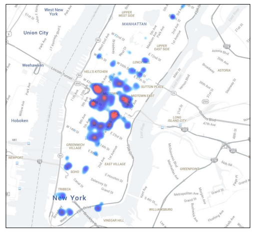
Figure 7 shows the performance comparison of different weight ratios θ. We test θ from $1/3$ which means a day of weekend is three times more important than a day of weekday to $3/1$ which means the opposite. The top-N performances in terms of precision, recall are visualized. We can observe that the performance at $\theta = 1/1$ achieves the highest, which means the importance of weekdays and weekends are almost the same for modeling the user-POI temporal compatibility. Also, as the ratio θ goes more and more unbalance, the performance becomes worse generally, except precision and recall @1 from $\theta = 2/1$ to $\theta = 3/1$. From this study, we can see that large city such as NYC provides rich lifestyles in weekdays as in weekends.

H. Correlation between Taxi Trips and LBSN Check-ins

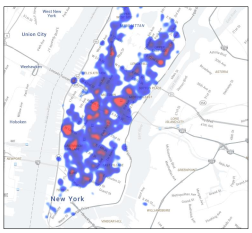
Here we conduct a case study to explore the spatio-temporal correlation between heterogeneous taxi rider mobility and LSBN user check-in behavior. We randomly sample taxi drop-offs and POI check-ins during different time period of weekday and plot their locations to make the heatmaps. Figure 8 shows the heatmaps of two different period: 8AM–9AM and 8PM–9PM in weekdays. For visualization purpose, we only show the heat color on relatively high density areas, therefore the plain areas do not mean there are no events of drop-off or check-in.



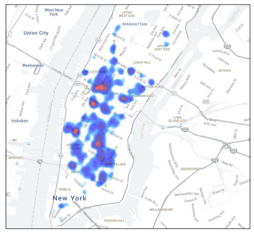
(a) Taxi drop-off 8–9AM.



(b) Check-in 8–9AM.



(c) Taxi drop-off 8–9PM.



(d) Check-in 8–9PM.

Fig. 8: Heatmap of mobility of taxi riders and LBSN users at different time slot of weekday.

Figure 8a, 8b show the human mobility of taxi riders and LBSN users in morning hour 8AM to 9AM. We can find that in the morning people are mainly concentrated in two business regions of NYC: Midtown (area around Rockefeller Center) and Financial District (area around Wall Street). Because of worse traffic, people who take taxi to Financial District are fewer than those take taxi to Midtown. Compare the taxi drop-off with check-in in morning hour, we can see that the taxi riders are relatively strong co-located with LBSN user mobility. Figure 8c, 8d show the taxi drop-offs and check-ins in night hour 8PM to 9PM. We can see that people leave business regions where they work in daytime. Meanwhile, they are going to residential areas (e.g., Upper Each, Upper West) where do not have many POIs for check-in as well as restaurant & nightlife areas (e.g., Greenwich Village, East Village, Fashion District) where have dense check-ins. Significantly, the check-in distribution is spatio-temporal correlated with the mobility trend of taxi riders. This case study supports our idea that the temporal patterns of LBSN users' check-in behavior are predictable via heterogeneous massive human mobility data, which can be utilized for boosting recommendation performance.

VI. Related Work

In this section, we introduce the related work from three research angle: personalized recommendation methodology especially latent factor model, temporal influence enhanced POI recommender system, and human mobility analysis especially in LBSN environment.

Collaborative filtering technique, especially the factorization based approach has shown its importance to the field of recommender system. It has been widely used for various classic recommendation algorithms. The basic factorization algorithms include matrix factorization [10], probabilistic matrix factorization [8] and its Bayesian version [11], as well as other variants [12], [13]. Most of these algorithms are majorly developed for explicit user response (e.g., rating), and assumes that the responses follow a Gaussian distribution over the predicted preferences. As more and more emerging recommendation applications which only have implicit user responses (e.g., count of web-click or check-in) came to the research filed, recommender systems are also required to infer user preferences from these heavy skew and wide range data. However, Gaussian-based latent factor models show their limitation on prediction performance. Under this circumstance, researchers developed latent factor models which is more suitable for implicit responses by setting non-negative constraints on latent variables [14], [9], [15], which aims to force the predicted preferences into a wider range to adapt implicit responses. Furthermore, by better modeling heavy skew data and providing rigorous probabilistic generative process, Poisson distribution became popular in recommendation modeling especially for implicit response [16], [17], [18], [19].

The second group is more specific to incorporate temporal influences into recommender system for better understanding users' temporal preferences. The first category can be summarized as time-aware recommendation which learns temporal preferences to recommend items for specific time slots (e.g., an hour of a day). The early work in [20], [21] discover the dynamic of user preference or interests over time. More recently, researchers start to investigate periodic patterns of user preferences (e.g., hourly interests of every day). One direct solution is to add an time dimension to user-item matrix and apply tensor factorization [3], [22]. The work in [1] considers a user's separated latent variables at different time slots, and preserves the similarity of personal preference in consecutive times. The work in [2] makes time-aware recommendations by a user-based collaborative filtering method which computes the similarity between users by finding the same POIs at the same times in their check-in history. The work in [23] learns temporal preferences by adopting topic model and training unique temporal features for each topic. Relevant work can also be found in [24], [25]. Our work is mostly related to this category. Meanwhile, there exists the other category which can be concluded as successive POI recommendation. The objective of this category aim to learn sequential patterns to predict user preferences for next POI, such as the work in [26], [27] which train personalized Markov chain to capture sequential check-in preferences. More relevant work can be found in [28], [29].

The last group of research concentrates on human mobility analysis of LBSN users. The work in [30] shows that users' mobility similarity is strongly correlated with their social proximity. The work in [5] utilizes Gaussian mixture model to capture users' periodic mobility at different states (e.g., home/work). The work in [31] explores the connectivity among urban places via the mobility of LBSN users. The work in [6] uses heterogeneous mobility data to measure a static connectivity among areas for boosting the performance of user location prediction. The work in [32] explores and categorize urban lifestyles with the mobility of LBSN users. More relevant work can be found in [33], [34].

VII. Conclusions

In this paper, we developed a POI recommendation model by considering the temporal matching between users and POIs. Firstly, we presented a method to profile the temporal popularity of POIs by (i) mining area activity patterns with taxi trips, (ii) integrating category popularity pattern with POI category level check-ins, and (iii) enhancing patterns with mixture mode. Moreover, we learned the latent temporal regularity of users by incorporating the temporal matching degrees of user-POI pairs into user overall preference estimation. Finally, we conducted extensive experiments with POI check-in and human mobility data. As demonstrated by the experimental results, the consideration of temporal matching between users and POIs can better model LBSN users' choosing processes. The performance improvement of our proposed method is substantial compared to benchmark methods.

VIII. Acknowledgments

This research was partially supported by National Science Foundation (NSF) via the grant number IIS-1648664. Also, it was supported in part by Natural Science Foundation of China (71329201) and the Rutgers 2015 Chancellor's Seed Grant Program. Finally, the contact author is Professor Hui Xiong (hxiong@rutgers.edu).

References

[1] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring temporal effects for location recommendation on location-based social networks," in *RecSys*. ACM, 2013, pp. 93–100.

[2] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *SIGIR*. ACM, 2013, pp. 363–372.

[3] L. Xiong, X. Chen, T.-K. Huang, J. G. Schneider, and J. G. Carbonell, "Temporal collaborative filtering with bayesian probabilistic tensor factorization." in *SDM*, vol. 10. SIAM, 2010, pp. 211–222.

[4] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on gps data," in *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 2008, pp. 312–321.

[5] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 1082–1090.

[6] Y. Wang, N. J. Yuan, D. Lian, L. Xu, X. Xie, E. Chen, and Y. Rui, "Regularity and conformity: Location prediction using heterogeneous mobility data," in *Proceedings of the 21th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2015, pp. 1275–1284.

[7] D. Yang, D. Zhang, V. W. Zheng, and Z. Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns," *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, vol. 45, no. 1, pp. 129–142, 2015.

[8] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *NIPS*, vol. 20, 2008.

[9] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *NIPS*, 2001, pp. 556–562.

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

[11] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using markov chain monte carlo," in *ICML*. ACM, 2008, pp. 880–887.

[12] D. Agarwal and B.-C. Chen, "Regression-based latent factor models," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 19–28.

[13] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2008, pp. 426–434.

[14] Q. Gu, J. Zhou, and C. H. Ding, "Collaborative filtering: Weighted nonnegative matrix factorization incorporating user and item graphs." in *SDM*. SIAM, 2010, pp. 199–210.

[15] S. Zhang, W. Wang, J. Ford, and F. Makedon, "Learning from incomplete ratings using non-negative matrix factorization." in *SDM*, vol. 6. SIAM, 2006, pp. 548–552.

[16] H. Ma, C. Liu, I. King, and M. R. Lyu, "Probabilistic factor models for web site recommendation," in *SIGIR*. ACM, 2011, pp. 265–274.

[17] Y. Chen, M. Kapralov, J. Canny, and D. Y. Pavlov, "Factor modeling for advertisement targeting," in *NIPS*, 2009, pp. 324–332.

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

[19] B. Liu, D. Kong, L. Cen, N. Z. Gong, H. Jin, and H. Xiong, "Personalized mobile app recommendation: Reconciling app functionality and user privacy preference," in *WSDM*. ACM, 2015, pp. 315–324.

[20] Y. Koren, "Collaborative filtering with temporal dynamics," *Communications of the ACM*, vol. 53, no. 4, pp. 89–97, 2010.

[21] Y. Ding and X. Li, "Time weight collaborative filtering," in *Proceedings of the 14th ACM international conference on Information and knowledge management*. ACM, 2005, pp. 485–492.

[22] P. Bhargava, T. Phan, J. Zhou, and J. Lee, "Who, what, when, and where: Multi-dimensional collaborative recommendations using tensor factorization on sparse user-generated data," in *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2015, pp. 130–140.

[23] J. McInerney, J. Zheng, A. Rogers, and N. R. Jennings, "Modelling heterogeneous location habits in human populations for location prediction under data sparsity," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 2013, pp. 469–478.

[24] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Who, where, when and what: discover spatio-temporal topics for twitter users," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2013, pp. 605–613.

[25] H. Gao, J. Tang, X. Hu, and H. Liu, "Modeling temporal effects of human mobile behavior on location-based social networks," in *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*. ACM, 2013, pp. 1673–1678.

[26] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, ser. IJCAI '13. AAAI Press, 2013, pp. 2605–2611.

[27] W. Mathew, R. Raposo, and B. Martins, "Predicting future locations with hidden markov models," in *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, 2012, pp. 911–918.

[28] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *Data mining (ICDM), 2012 IEEE 12th international conference on*. IEEE, 2012, pp. 1038–1043.

[29] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks." in *ICWSM*, 2012.

[30] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi, "Human mobility, social ties, and link prediction," in *KDD*. ACM, 2011, pp. 1100–1108.

[31] A. Noulas, B. Shaw, R. Lambiotte, and C. Mascolo, "Topological properties and temporal dynamics of place networks in urban environments," in *Proceedings of the 24th international conference on world wide web companion*. International World Wide Web Conferences Steering Committee, 2015, pp. 431–441.

[32] N. J. Yuan, F. Zhang, D. Lian, K. Zheng, S. Yu, and X. Xie, "We know how you live: exploring the spectrum of urban lifestyles," in *Proceedings of the first ACM conference on Online social networks*. ACM, 2013, pp. 3–14.

[33] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[34] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.