TICRec: A Probabilistic Framework to Utilize Temporal Influence Correlations for Time-aware Location Recommendations

Jia-Dong Zhang, Chi-Yin Chow, *Member, IEEE*

Abstract—In location-based social networks (LBSNs), time significantly affects users' check-in behaviors, for example, people usually visit different places at different times of weekdays and weekends, e.g., restaurants at noon on weekdays and bars at midnight on weekends. Current studies use the temporal influence to recommend locations through dividing users' check-in locations into time slots based on their check-in time and learning their preferences to locations in each time slot separately. Unfortunately, these studies generally suffer from two major limitations: (1) the loss of time information because of dividing a day into time slots and (2) the lack of temporal influence correlations due to modeling users' preferences to locations for each time slot separately. In this paper, we propose a probabilistic framework called TICRec that utilizes Temporal Influence Correlations (TIC) of both weekdays and weekends for time-aware location recommendations. TICRec not only recommends locations to users, but it also suggests when a user should visit a recommended location. In TICRec, we estimate a time probability density of a user visiting a new location without splitting the continuous time into discrete time slots to avoid the time information loss. To leverage the TIC, TICRec considers both user-based TIC (i.e., *different users' check-in behaviors to the same location at different times*) and location-based TIC (i.e., *the same user's check-in behaviors to different locations at different times*). Finally, we conduct a comprehensive performance evaluation for TICRec using two real data sets collected from Foursquare and Gowalla. Experimental results show that TICRec achieves significantly superior location recommendations compared to other state-of-the-art recommendation techniques with temporal influence.

Index Terms—Location-based social networks, location recommendations, time-aware location recommendations, continuous temporal influence, temporal influence correlations, kernel density estimation.

1 Introduction

Location-based social networks (LBSNs), such as Foursquare, Gowalla, and Facebook Places, have been growing rapidly in recent years. For example, as of May 2014, Foursquare had over 50 million people worldwide, over 6 billion check-ins with millions more every day, and over 1.7 million businesses using its Merchant Platform. Therefore, it is important to recommend relevant locations for users, since it can bring a lot of benefits to the society. For instance, location recommendations help people explore new places to enhance the quality of their daily life and enable businesses to provide relevant advertisements for their potential customers.

In an LBSN (Fig. 1), users can establish social links and share their experiences or tips of visiting or checking in some interesting locations, e.g., restaurants, stores, and museums. Such check-in locations are also known as points-of-interest (POIs). To make location recommendations for users, most existing studies infer their preferences to POIs through utilizing *collaborative filtering techniques* based on users' check-in data [1], [2], [3], [4], [5], *social influence* between users in terms of their social links [6], [7], [8], [9], and/or *geographical influence* of locations according to their geographic coordinates [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. However, all these studies cannot suggest appropriate time for users to visit a recommended POI, because they do not consider **the influence of the temporal context when users visiting locations on their check-in behaviors**, called **temporal influence** for short hereafter. In reality, time is a very important factor influencing human activities at different times on weekdays and weekends [16], [17], [18], [19]. For example, users often visit restaurants at noon on weekdays and bars at midnight on weekends. These weekday and weekend patterns reflect the temporal check-in preferences of users to locations [20], which can be used to make time-aware location recommendations by suggesting properly visiting time on weekdays or weekends.

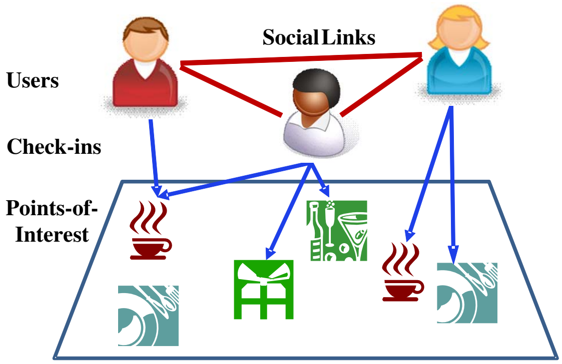


Fig. 1. A location-based social network

• *J.-D. Zhang and C.-Y. Chow are with Department of Computer Science, City University of Hong Kong, Hong Kong. Email: jzhang26-c@my.cityu.edu.hk, chiychow@cityu.edu.hk*

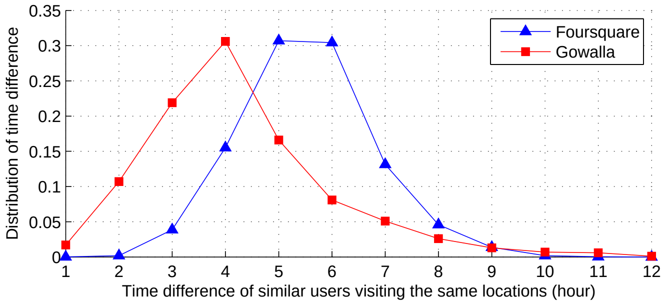


Fig. 2. User-based temporal influence correlations

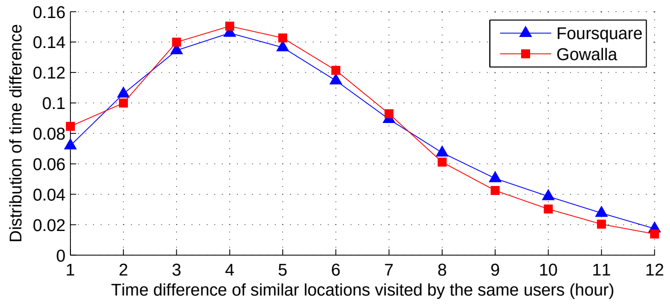


Fig. 3. Location-based temporal influence correlations

Existing methods using temporal influence and their limitations. To the best of our knowledge, although there are some studies [16], [17], [18], [19] that investigate the importance of temporal dynamics in human activities, only three existing methods [21], [22], [23] consider the temporal influence to recommend POIs for users in LBSNs. They split a day into time slots, e.g., 24 hours, and apply the memory or model based collaborative filtering recommendation techniques to infer users' preferences on locations at each time slot separately. Unfortunately, these studies generally suffer from two major limitations. (1) **Time information loss.** In these three existing methods, the exact time information of users visiting locations is lost, since they transform the continuous time into discrete time slots, e.g., one hour per slot. Consequently, they learn the user preference at a predefined time slot based on discrete time models that only aggregate the check-in data of users at the corresponding time slot. However, it is better to integrate the user preference within a target time interval based on continuous time models that derive the user preference at any continuous time by considering all check-in data of users. (2) **Lack of temporal influence correlations (TICs).** Even worse, these three existing methods employ the temporal influence separately for each time slot when making location recommendations, although they also utilize the temporal consecutiveness of *the same user visiting the same location at different time slots*. As a result, they are unable to correlate the temporal influences of different time slots on *different users' check-in behaviors to different locations*. However, in reality there exist two important kinds of TICs. (a) **User-based TIC: The check-in behaviors of different users to the same location at different times may be correlated.** For example, similar users (e.g., friends) may check in a location at different hours instead of the same hour, since they may have the common interest on the location but with different available time. (b) **Location-based TIC: The check-in behaviors of the same user to different locations at different times may be also correlated.** For instance, similar locations (e.g., locations belong to the same category) may be visited by a user at different times, because she may visit the locations sequentially.

Real-world motivating examples. To observe the two important kinds of temporal influence correlations: user-based TIC and location-based TIC, we conduct analysis on two publicly available real data sets collected from Foursquare [7] and Gowalla [24], which are the two of the most popular LBSNs. (1) For the user-based TIC, we select ten thousands of user pairs (similar users) from each data set who have checked in the most common locations, and calculate the time difference of two users in a pair visiting the same location. Fig. 2 plots the distribution of time difference: in both data sets more than 95% user pairs who visited the same location have a more than one hour time difference; in other words, different users usually check in the same location at different hours. This validates that different users at different times correlate to the same location. As a result, it is undesirable to separate the temporal influence on users at one time slot from another time slot as in the existing methods [21], [22], [23]. (2) For the location-based TIC, we select ten thousands of location pairs (similar locations) which have been visited by the most common users, and calculate the time difference of two locations in a pair visited by the same user. Fig. 3 plots the distribution of time difference: more than 90% location pairs visited by the same user in both data sets have a time difference longer than one hour; in other words, different locations are usually visited by the same user at different hours. This validates that different locations at different times correlate to the same user. Hence, it is also undesirable to separate the temporal influence on locations at one time slot from another time slot as in the existing methods [21], [22], [23].

Our approach. In this paper, we propose a probabilistic framework to utilize **T**emporal **I**nfluence **C**orrelations for time-aware location **Rec**ommendations, called TICRec that overcomes the two aforementioned limitations. (1) To avoid the loss of time information, we estimate a probability density over the continuous time of a user visiting a location rather than transforming the continuous time into discrete time slots. Specifically, we model the continuous time probability densities based on a non-parametric density estimation method, i.e., the popular kernel density estimation (KDE) [25], since time densities of users visiting locations are very diverse and we cannot assume their forms. (2) To estimate the time probability density of a user visiting a location, we collect (a) the *different time* history of *different users* visiting the same location based on the **user-based TIC** and (b) the *different time* history of the same user visiting *different locations* based on the **location-based TIC**. Accordingly, the time history is weighed by the similarity between users (resp. locations) for the user-based (resp. location-based) TIC. It is worth noting that: (i) TICRec not only recommends locations to users, but it also suggests when a user should visit a recommended location. (ii) TICRec differentiates users' check-in activities on weekdays and weekends, since their check-in patterns are distinct from each other. For example, users often go to office on weekdays while visit tourist attraction on weekends.

This study is a significant extension to our previous works [14], [15] on *non-time-aware* recommendations by proposing a new probabilistic framework to utilize temporal influence correlations for *time-aware* location recommendations; the new probabilistic framework tackles the aforementioned limitations in the existing location recommendation methods that use temporal influence. The main contributions of this paper can be summarized as follows:

- We propose TICRec — a probabilistic framework for time-aware location recommendations. TICRec is a general framework for integrating the temporal influence with other important information, e.g., the social influence (social links between users) and geographical influence (geographic coordinates of locations) in our previous work [15]. TICRec focuses on the problem of time-aware location recommendations that is significantly distinct from and much more challenging than the problem of non-time-aware recommendations studied in the work [15]. In TICRec, the proposed time models for tackling the "time-aware" challenge are totally new to the work [15].
- In TICRec, we estimate a time probability density over continuous time on weekdays and weekends for each user visiting a location rather than splitting the continuous time into discrete time slots to avoid the loss of time information. The proposed continuous time models in TICRec are adaptive to different notions of time, e.g., time of a day, time of a week, etc. Moreover, our continuous time models are essentially different from the discrete time models used in the works [21], [22], [23]; the differentiation is not only about the temporal granularity but also the way or methodology of modeling the temporal influence.
- We develop an approach to model the time probability density of a user visiting a location from the check-in time history of other users visiting the same location and the same user visiting other locations in order to exploit the user-based TIC and location-based TIC, respectively. To the best of our knowledge, this is the first approach to correlate the temporal influences of different users and different locations, which can enhance the predictive ability of the time probability density and improve the quality of location recommendations.
- We conduct extensive experiments to evaluate the performance of TICRec using two large-scale real data sets collected from Foursquare and Gowalla. Experimental results show that TICRec outperforms the state-of-the-art time-aware recommendation techniques [21], [22], [23] in terms of recommendation precision, recall, and time error.

The remainder of this paper is organized as follows. Section 2 highlights related work. Section 3 describes the proposed probabilistic framework to exploit the temporal influence correlations for location recommendations. In Sections 4 and 5, we present our experiment settings and analyze the performance of TICRec, respectively. Finally, we conclude this paper in Section 6.

2 Related Work

Recently with the rapid growth of LBSNs, like Foursquare, Gowalla, Facebook places, etc., recommending locations (i.e., POIs) for users becomes prevalent [26]. In general, there are four main categories for existing location recommendation approaches: *collaborative filtering*, *social influence*, *geographical influence*, and *temporal influence*.

Collaborative filtering techniques. Although there are a few works that recommend POIs through the content-based techniques [27], [28], [29], most studies provide POI recommendations by using the conventional collaborative filtering techniques based on users' check-in data [2], [3], [30], travel tour data [31], GPS trajectory data [32], [33], [34], [35], [36], [37], [38], or text data [39]. In particular, some techniques [1], [4], [5] employ users' residence to derive their similarity weights as an input of the conventional collaborative filtering techniques [40], [41], [42]. However, the performance of all these techniques is considerably limited due to no consideration for the *social influence*, *geographical influence*, or *temporal influence*.

Social influence. Since friends are more likely to share common interests, social link information has been widely utilized to improve the quality of recommender systems in the conventional social networks like Twitter [43], [44], [45] and the LBSNs [6], [7], [8], [9], [10], [14] by deriving the similarity between users based on their social friendships and integrating it into the collaborative filtering techniques.

Geographical influence. The geographical proximity between POIs significantly affects the check-in behaviors of users on the POIs. To exploit geographical influence for improving the quality of location recommendations, the studies [8], [10], [13], [46] view locations as ordinary non-spatial items and consider the geographical influence of locations by predefining a range; locations only within this range will be possibly recommended to users. The literature [11] presents a geo-topic model by assuming that if a location is closer to the locations visited by a user or the current location of a user, it is more likely to be visited by the same user. More sophistically, the studies [6], [9], [12], [22], [23], [47], [48] model the distance between two locations visited by the same user as a common distribution for all users, e.g., a power-law distribution or a multi-center Gaussian model. In particular, our previous papers [14], [15], [49], [50] personalize the geographical influence by modeling the distance between locations visited by the same user as a personalized distribution for each user.

Temporal influence. The time-dependent recommendation techniques can be divided into five main categories. (1) ***Absolute time factor.*** The time factor has been widely used for the conventional recommendations (e.g., books, music and movies) by considering the time gap between the occurring time of a previous rating and the recommendation time as a decaying factor to weigh the rating [51], which is different from the periodic time factor for location recommendations. (2) ***Sequential time factor.*** The time sequence has also been utilized to recommend next POI for

TABLE 1
Key notations in this paper

Symbol	Meaning
U	Set of all users in an LBSN
u	Some user: $u \in U$
L	Set of all locations (or POIs) in an LBSN
l	Some location: $l \in L$
T	A time interval
$\langle u, l, t\rangle$	Check-in or visit that describes user u visiting location l at time t
D	Collection of check-ins of all users visiting all locations in an LBSN: $D = \{\langle u_i, l_i, t_i\rangle\}_{i=1}^{\|D\|}$
$D_{u,l}$	Check-in time sample of user u visiting location l: $D_{u,l} = \{t_i \| \langle u_i, l_i, t_i\rangle \in D \wedge u_i = u \wedge l_i = l\}$ (note that usually $D_{u,l} = \emptyset$)
$S_{u,l}$	$D_{u,l}$'s extended time sample that is derived from D based on the temporal influence correlations
$W_{u,l}(t_i)$	Weight of the sample point $t_i \in S_{u,l}$
$P(l\|u, T, D)$	Predicted probability of u visiting **new** location l at time interval T given D
$P(l\|u, D)$	Prior probability of user u visiting location l
$f(t\|u, l, D)$	Time probability density conditioned on user u and location l

users [52], which is distinct from the time slots in a day. Accordingly, the method in [52] cannot generate time-aware recommendations for users. (3) ***Using time information for location predictions.*** The literatures [53], [54], [55], [56], [57] study the relationship between visited locations and temporal information for location predictions that refer to predicting an *existing* location. It is not straightforward to apply these techniques in location recommendations that refer to recommending a *new* location. (4) ***Periodic time pattern discovering.*** The works [16], [17], [18], [19] only show the temporally periodic patterns of users visiting locations without using the patterns to make location recommendations. (5) ***Periodic time pattern deducing.*** To the best of our knowledge, there only exist three literatures [21], [22], [23] that utilize the temporal influence in location recommendation. Specifically, they split a day into time slots, e.g., 24 hours, divide the user-location check-in data according to the check-in time and the time slots, and apply matrix factorization [21], user-based collaborative filtering [22], or graph-based method [23] to infer users' preferences on locations at each time slot.

3 Continuous Temporal Influence

In this section, we define some important concepts and the research problem of this paper in Section 3.1, propose a probabilistic framework with temporal influence based on KDE for location recommendations in Section 3.2, and develop an approach to exploit the user-based and location-based temporal influence correlations (TICs) in Section 3.3.

3.1 Problem Statement

TABLE 1 summarizes the key symbols used in this paper. We first present some basic concepts and the problem definition as follows.

Definition 1. Check-in or visit. A check-in or visit is a triple $\langle u, l, t\rangle$ that describes user $u \in U$ visiting location $l \in L$ at time t, in which U and L are the sets of users and locations in an LBSN, respectively.

It is worth emphasizing that the proposed probabilistic framework in this paper is applicable to different notions of time, e.g., time of a day and time of a week. However, since humans show strong daily and weekly periodic behavior [17], [18], [24], in this work we differentiate the temporal influence on weekdays from weekends, but we can model the temporal influence on weekdays and weekends based on the same process. Following, we focus on the temporal influence of weekdays, i.e., $t \in [0{:}00, 24{:}00)$ is always a time instant of a day in weekdays. For example, $t = 0{:}00$ and $t = 12{:}00$ represent midnight and noon on weekdays.

Definition 2. Check-in collection. A check-in collection is a set of check-ins of all users visiting all locations in an LBSN, denoted as $D = \{\langle u_i, l_i, t_i\rangle\}_{i=1}^{|D|}$, in which $|D|$ represents the number of check-ins in D, the same hereafter.

Definition 3. Check-in time sample. Given a check-in collection D, the check-in time sample of user u visiting location l is denoted as $D_{u,l} = \{t_i | \langle u_i, l_i, t_i\rangle \in D \wedge u_i = u \wedge l_i = l\}$. Note that $|D_{u,l}|$ represents the number of records in the sample, i.e., the frequency of u visiting l.

Problem definition. In the problem of time-aware location recommendations with the temporal influence, given a check-in collection D, a user u and a time interval T, the goal is to predict the probability of user u visiting **new** location $l \in L$ at time interval T, denoted as $P(l|u, T, D)$, and then return the top-k locations with the highest visiting probability $P(l|u, T, D)$ for u at T.

It is important to note that: (1) In the problem of *time-aware* location recommendations, it is required to not only recommend interesting locations to users based on their preferences but also suggest proper time for users to visit recommended locations. This task is much more challenging than the traditional problem of *non-time-aware* location recommendations that can recommend users with their preferred locations but not advise appropriate visit time. (2) If we straightforwardly apply the non-time-aware recommendation techniques to make time-aware location recommendations by randomly suggesting visiting time, their recommendation quality is pretty poor because of failing to advise proper visit time, as shown in our experimental results in Section 5.

3.2 A Probabilistic Framework with Continuous Temporal Influence

Existing time-aware location recommender systems transform the continuous time into discrete time slots and then learn a user's preference on locations in each time slot separately [21], [22], [23]. Such a discretization transformation leads to loss of continuous time information. In contrast, our TICRec does not require any discretization processing, so it can address this limitation. More specifically in terms of probability theory, we infer the probability $P(l|u, T, D)$

of user u visiting new location $l \in L$ at time interval T by

$$\begin{aligned}P(l|u,T,D) &= \frac{P(l|u,D)P(T|u,l,D)}{P(T|u,D)} \\ &\propto P(l|u,D)P(T|u,l,D), \forall l \in L \\ &= P(l|u,D)\int_{t\in T} f(t|u,l,D)dt, \qquad (1)\end{aligned}$$

where $P(l|u,D)$ is the prior probability of user u visiting location l that is independent of time interval T, and $f(t|u,l,D)$ is the time probability density conditioned on user u and location l that is essential to utilize the temporal influence.

Deriving the prior probability with social and geographical influences. First, to obtain the prior probability $P(l|u,D)$, we can apply a variety of recommendation techniques. Specifically, since the social and geographical influences play a significant role in the check-in behaviors of users on locations, we employ the geo-social location recommendation technique proposed in our previous works [14], [15] to compute the prior probability of a user visiting a location independently of the temporal influence. In general, this technique has three main steps.

(1) The social links between users and distances between their residences are used to derive their similarities:

$$sim^{soc}(u,u') = \begin{cases} 1 - \frac{distance(u,u')}{\max_{u''\in F(u)} distance(u,u'')}, & u' \in F(u); \\ 0, & u' \notin F(u), \end{cases} \qquad (2)$$

where $F(u)$ is the set of users having social links with user u in an LBSN and $distance(u,u')$ is the distance between the residences of u and u'. Accordingly, we derive the social rating $r^{soc}_{u,l}$ of user u to location l based on social collaborative filtering:

$$r^{soc}_{u,l} = \frac{\sum_{u'\in U} sim^{soc}(u,u')|D_{u',l}|}{\sum_{u'\in U} sim^{soc}(u,u')}, \qquad (3)$$

where $|D_{u',l}|$ is the frequency of user u' visiting location l given in Definition 3.

(2) We use the geographic latitude and longitude coordinates of locations that a user has visited to derive a probability of the user visiting a new location. Formally, let L_u be the set of locations visited by user u, i.e., $L_u = \{l_i|\langle u_i, l_i, t_i\rangle \in D \wedge u_i = u\}$, the geographical probability $P^{geo}(l|L_u)$ of user u to new location l is given by:

$$P^{geo}(l|L_u) = 1 - \prod_{l_i\in L_u}\left(1 - \frac{1}{|X_u|\sqrt{2\pi}}\sum_{x\in X_u} e^{-[distance(l,l_i)-x]^2/(2\sigma^2)}\right), \qquad (4)$$

where $distance(l,l_i)$ is the geographical distance between l and l_i, $X_u = \{distance(l_i,l_j)|\forall l_i,l_j \in L_u\}$ is the set of distances between every pair of locations in L_u, and σ is approximately the standard deviation of X_u divided by $|X_u|^{1/5}$ [15].

(3) The geo-social prior probability in Equation 1 of user u to location l is given by

$$P(l|u,D) = \frac{r^{soc}_{u,l}\cdot P^{geo}(l|L_u)}{\sum_{l\in L} r^{soc}_{u,l}\cdot P^{geo}(l|L_u)}. \qquad (5)$$

Estimating the time probability density with kernel density estimation. To exploit the temporal influence, we can obtain the time probability density $f(t|u,l,D)$ in Equation 1 using density estimation techniques. Because the time densities $f(t|u,l,D)$ of different users to different locations are very diverse, in this paper we apply a general non-parametric method, known as the kernel density estimation [25] (KDE), which can be used with arbitrary distributions and without any assumption on the form of the underlying distribution. $f(t|u,l,D)$ is computed by:

$$f(t|u,l,D) = \frac{\sum_{t_i\in S_{u,l}} W_{u,l}(t_i)\frac{1}{h}K\left(\frac{t\ominus t_i}{h}\right)}{\sum_{t_i\in S_{u,l}} W_{u,l}(t_i)}, \qquad (6)$$

where $S_{u,l}$ is the time sample for estimating $f(t|u,l,D)$ and $W_{u,l}(t_i)$ is the weight of the sample point t_i that are derived from D based on the temporal influence correlations. The details of deriving $S_{u,l}$ and $W_{u,l}(t_i)$ will be presented in Section 3.3. In addition, $t\ominus t_i$, $K(\cdot)$, and h in Equation 6 are defined as follows:

- $t \ominus t_i$ is the time difference between two time instances t and t_i. Since time is periodic, we cannot directly subtract t_i from t to obtain their difference; hence, their difference is determined by:

$$t\ominus t_i = \begin{cases} |t-t_i|, & |t-t_i| \le 12:00; \\ 24:00 - |t-t_i|, & |t-t_i| > 12:00. \end{cases} \qquad (7)$$

For example, when $t = $ 4:10 and $t_i = $ 1:05, $t\ominus t_i = $ 3:05; when $t = $ 23:10 and $t_i = $ 1:05, $t\ominus t_i = $ 24:00 $-$ 22:05 $=$ 1:55.

- $K(\cdot)$ is a kernel function satisfying the following two conditions:

$$\forall x, K(x) \ge 0 \text{ and } \int_{-\infty}^{+\infty} K(x)dx = 1. \qquad (8)$$

In this paper, we apply the most popular normal kernel:

$$K(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}. \qquad (9)$$

- h is a smoothing parameter, called the bandwidth. In terms of the well-known three-sigma rule [58], nearly all values lie within three standard deviations of the mean in a normal distribution, i.e.,

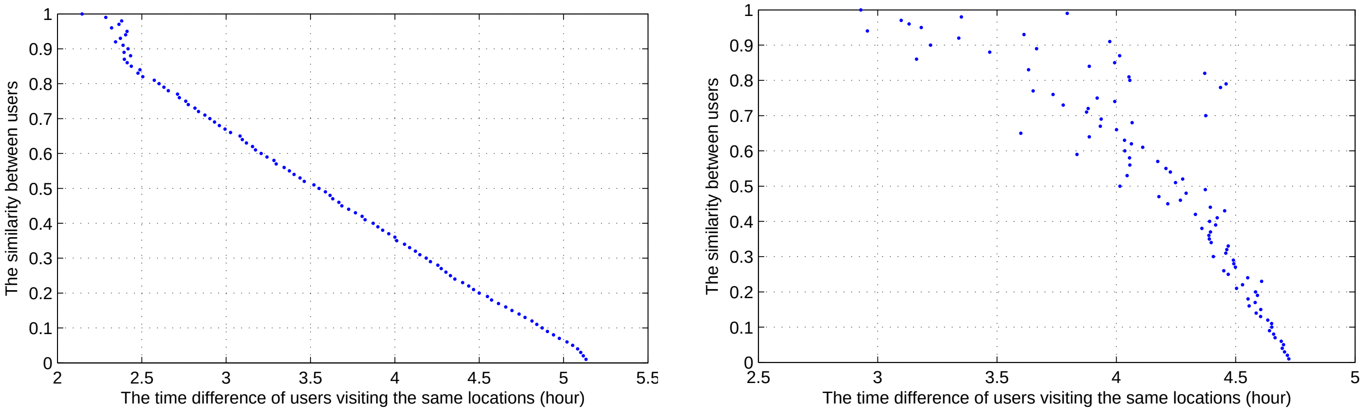
$$\int_{-3h}^{3h} \frac{1}{\sqrt{2\pi}h}K\left(\frac{x}{h}\right)dx \approx 1, \qquad (10)$$

and t always lies in [0:00, 24:00). Thus, $6h \approx 24$, i.e.,

$$h \approx 4. \qquad (11)$$

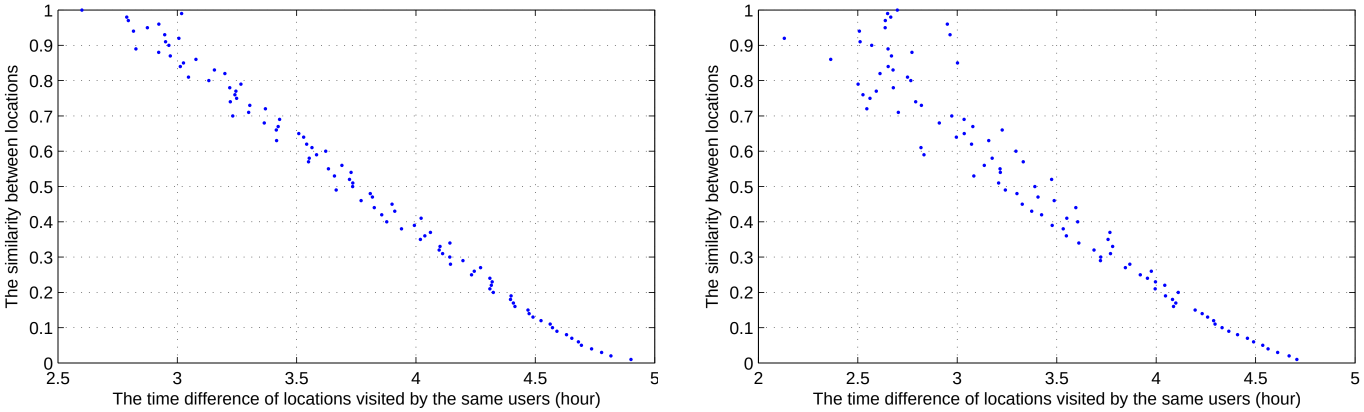
3.3 Exploiting User-based and Location-based Temporal Influence Correlations

In the proposed probabilistic framework with temporal influence described in Section 3.2, the key task is to derive sample $t_i \in S_{u,l}$ and weight $W_{u,l}(t_i)$ in order to estimate the time probability density $f(t|u,l,D)$ in Equation 6. Note that we cannot simply regard $S_{u,l}$ as $D_{u,l}$, since we are inferring the probability of user u visiting new location l, i.e., u has not checked in location l before and $D_{u,l}$ is empty.



(a) Foursquare (b) Gowalla

Fig. 4. The relationship between the time difference with the similarity of users



(a) Foursquare (b) Gowalla

Fig. 5. The relationship between the time difference with the similarity of locations

To this end, we exploit the user-based TIC and location-based TIC to derive the sample $t_i \in S_{u,l}$ and weight $W_{u,l}(t_i)$ for estimating the time probability density $f(t|u,l,D)$ in Equation 6.

Observing temporal influence correlations by analyzing the check-in data. The user-based TIC states that *the check-in behaviors of different users to the same location at different times* may be correlated. For example, a group of friends may visit a POI at different times, because they have the common interest in the POI, but with different available time. On the other hand, the location-based TIC indicates that *the check-in behaviors of the same user to different locations at different times* may be correlated as well. For instance, the POIs belonging to the same category may be visited by a user at different times, because she could visit the POIs for different purposes (e.g., a user visits a restaurant for a breakfast, lunch or dinner). To further observe the user-based and location-based TICs, we analyze these TICs in two publicly available real data sets collected from Foursquare [7] and Gowalla [24]. Specifically, we calculate (a) the cosine similarity between every pair of users

$$sim(u,u') = \frac{\sum\limits_{l \in L} |D_{u,l}||D_{u',l}|}{\sqrt{\sum\limits_{l \in L} |D_{u,l}|^2}\sqrt{\sum\limits_{l \in L} |D_{u',l}|^2}}, \tag{12}$$

and the average time difference of them visiting the same location, and (b) the cosine similarity between every pair of locations

$$sim(l,l') = \frac{\sum\limits_{u \in U} |D_{u,l}||D_{u,l'}|}{\sqrt{\sum\limits_{u \in U} |D_{u,l}|^2}\sqrt{\sum\limits_{u \in U} |D_{u,l'}|^2}}, \tag{13}$$

and the average time difference of them visited by the same user. Figs. 4 and 5 show the relationships between the time difference with the similarity of users and the similarity of locations, respectively. In general, we have two findings: (1) Different users check in a POI at different times while different locations are visited by a user at different times as well, both with the time gap of larger than two hours. (2) With the increase of similarity of users or locations, the time difference decreases approximately linearly.

Exploiting temporal influence correlations for estimating time probability densities. The experimental results shown in Figs. 4 and 5 inspire us to correlate (a) the temporal influences of *different users to the same location at different times* and (b) the temporal influences of *the same user to different locations at different times*, in order to obtain the time sample $S_{u,l}$ for estimating the time probability density $f(t|u,l,D)$ in Equation 6. In other words, we can derive the time sample $S_{u,l}$ of user u to location l through combining the check-in samples: (i) $D_{u',l}$ of another user u' visiting l (i.e., $u, u' \in U \land u \neq u'$) and (ii) $D_{u,l'}$ of u visiting another location l' (i.e., $l, l' \in L \land l \neq l'$). Formally,

$$S_{u,l} = \left(\bigcup_{u' \in U} \{t_i | t_i \in D_{u',l}\} \right) \bigcup \left(\bigcup_{l' \in L} \{t_j | t_j \in D_{u,l'}\} \right), \tag{14}$$

where the time sample $S_{u,l}$ is a multiset that may contain duplicate sample points from $D_{u',l}$ and $D_{u,l'}$. It is important to note that the temporal influence correlations used in Equation 14 are different from the temporal consecutiveness that smooths the check-in activity of a user to a location at a time slot through the check-in activity of the same user to the same location at other time slots, as applied in [21], [22]. Thus, the methods in [21], [22] cannot correlate the temporal influence of different users or different locations.

Further, we consider the similarity between users or locations as the sample weight of the corresponding time sample points, since the higher the similarity is, the smaller is the time difference of users visiting locations. Actually, the relationship between the similarity of users or locations and the time difference is approximately linear, as shown in Figs. 4 and 5. Thus, $W_{u,l}$ is given by:

$$\begin{aligned} &\forall t_i \in D_{u',l}, W_{u,l}(t_i) = sim(u, u'); \\ &\forall t_j \in D_{u,l'}, W_{u,l}(t_j) = sim(l, l'). \end{aligned} \tag{15}$$

Therefore, by employing temporal influence correlations for estimating the time probability density $f(t|u,l,D)$ based on Equations 14 and 15, Equation 6 is rewritten into

$$\begin{aligned} f(t|u,l,D) = \frac{1}{C} \Bigg[&\sum_{u' \in U} sim(u,u') \sum_{t_i \in D_{u',l}} \frac{1}{h} K\left(\frac{t \ominus t_i}{h}\right) \\ &+ \sum_{l' \in L} sim(l,l') \sum_{t_j \in D_{u,l'}} \frac{1}{h} K\left(\frac{t \ominus t_j}{h}\right) \Bigg] \end{aligned} \tag{16}$$

together with

$$C = \sum_{u' \in U} sim(u,u') |D_{u',l}| + \sum_{l' \in L} sim(l,l') |D_{u,l'}|. \tag{17}$$

Note that all time sample points in $D_{u',l}$ or $D_{u,l'}$ have the same weight $sim(u,u')$ or $sim(l,l')$, respectively, according to Equation 15. Thus, in Equation 17, $sim(u,u')$ and $sim(l,l')$ are multiplied by $|D_{u',l}|$ and $|D_{u,l'}|$, respectively. Further, to obtain the time-aware probability $P(l|u,T,D)$ of user u visiting new location l at time interval T in Equation 1, we can compute the integral of $f(t|u,l,D)$ at time interval T through

$$\begin{aligned} &\int_{t \in T} f(t|u,l,D) dt = \\ &\frac{1}{C} \Bigg[\sum_{u' \in U} sim(u,u') \sum_{t_i \in D_{u',l}} \int_{t \in T} \frac{1}{h} K\left(\frac{t \ominus t_i}{h}\right) dt \\ &+ \sum_{l' \in L} sim(l,l') \sum_{t_j \in D_{u,l'}} \int_{t \in T} \frac{1}{h} K\left(\frac{t \ominus t_j}{h}\right) dt \Bigg], \end{aligned} \tag{18}$$

where $\int_{t \in T} \frac{1}{h} K(\frac{\cdot}{h}) dt$ is the integral of the normal probability density that can be easily calculated by the numerical integration methods [59].

Algorithm 1 TICRec: The computation of $P(l|u,T,D)$

Input: User set U, location set L, check-in collection D, and time interval T.

Output: $P(l|u,T,D)$ for each pair of (u,l), $u \in U$ unvisiting $l \in L$.

1: // **Step 1: The pre-computation step**

2: Compute user similarity matrix $UM[u,u']$ by Equation 12

3: Compute location similarity matrix $LM[l,l']$ by Equation 13

4: Compute integral of normal probability density at time interval T for each pair of (u,l):

$$\Phi[u,l] = \sum_{t_i \in D_{u,l}} \int_{t \in T} \frac{1}{h} K\left(\frac{t \ominus t_i}{h}\right) dt$$

5: // **Step 2: The prior probability calculation step**

6: Compute social rating $r^{soc}_{u,l}$ by Equation 3

7: Compute geographical probability $P^{geo}(l|L_u)$ by Equation 4

8: Compute prior probability $P(l|u,D)$ by Equation 5

9: // **Step 3: The time probability density estimation step**

10: **for** each $u \in U$ **do**

11:   **for** each unvisited location $l \in L$ **do**

12:     $B = \sum_{u' \in U} UM(u,u')\Phi[u',l] + \sum_{l' \in L} LM(l,l')\Phi[u,l']$

13:     $C = \sum_{u' \in U} UM(u,u')|D_{u',l}| + \sum_{l' \in L} LM(l,l')|D_{u,l'}|$

14:     $P(l|u,T,D) = P(l|u,D)B/C$ by Equations 1 and 18

15:   **end for**

16: **end for**

Algorithm and computational complexity. Algorithm 1 outlines the process for computing $P(l|u,T,D)$ through Equation 1.

Step 1: The pre-computation step. Algorithm 1 first computes the user and location similarities which are used in the time probability density estimation and need $\max(O(|U|^2|L|), O(|U||L|^2))$ work (Lines 2 and 3). The key idea of Algorithm 1 is to pre-compute the common integrals Φ (Line 4). Each $\Phi[u,l]$ has been accessed repeatedly in the time probability density estimation. To obtain the common integrals Φ, we only need to scan the check-in collection D, compute integral for each check-in, and group the integrals according to the pair of user and location in the check-ins. Thus, it requires $O(|D|)$ work. It is important to note that we can incrementally compute the common integrals Φ for newly arriving check-ins, which is a nice property.

Step 2: The prior probability calculation step. This step is based on the geo-social location recommendation technique [15] and needs $O(|U|^2|L|)$ work (Lines 6 to 8).

Step 3: The time probability density estimation step. This step calculates the visiting probability of user u to location l by summing all temporal influence correlations based on the user and location similarities and the common integrals. The computational complexity of this step is $\max(O(|U|^2|L|), O(|U||L|^2))$ work (Lines 10 to 16).

Overall computational complexity. The complexity of Algorithm 1 is $\max(O(|U|^2|L|), O(|U||L|^2)) + O(|D|) = \max(O(|U|^2|L|), O(|U||L|^2))$, in which $|D| \ll |U||L|$ since

TABLE 2
Statistics of the two data sets

	Foursquare	Gowalla
Number of users	11,326	196,591
Number of locations (POIs)	182,968	1,280,969
Number of check-ins	1,385,223	6,442,890
Number of social links	47,164	950,327
User-location matrix density	2.3×10^{-4}	2.9×10^{-5}
Avg. No. of visited POIs per user	42.44	37.18
Avg. No. of check-ins per location	2.63	3.11

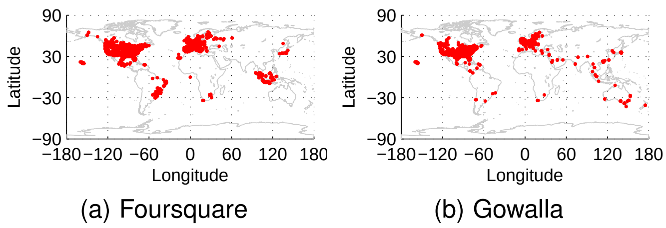


Fig. 6. Distribution of check-in locations on a world map

users only check in a small fraction of locations. More importantly, Algorithm 1 has the same computational complexity with the conventional collaborative filtering techniques (i.e., the computation of a user or location similarity matrix), although it takes full advantage of temporal influence correlations.

4 Experimental Evaluation

In this section, we describe our experiment settings for evaluating the performance of TICRec.

4.1 Two Real Data Sets

We use two publicly available large-scale real check-in data sets that were crawled from Foursquare [7] between January 2011 and July 2011 and Gowalla [24] between February 2009 and October 2010. The statistics of the data sets are shown in TABLE 2. Fig. 6 depicts the distribution of the locations in the data sets on a world map. In the pre-processing, we split each data set into the training set and the testing set in terms of the check-in time rather than using a random partition method, because in practice we can only utilize the past check-in data to predict the future check-in events. A half of check-in data with earlier timestamps are used as the training set, and the other half of check-in data are used as the testing set that needs to be divided into different time slots for the purpose of evaluation. In the experiments, the training set is used to learn the recommendation models of the evaluated techniques described in Section 4.2 to predict the testing data.

4.2 Evaluated Recommendation Techniques

The location recommendation techniques implemented in our experiments are listed below.

- iGeoRec [15]: iGeoRec is presented in our recent work to make time-unaware location recommendations since it does not take into account the temporal influence. Here, we apply iGeoRec to recommend time-aware locations by advising a random time slot for a user to visit a recommended location.
- LRT [21]: The Location Recommendation framework with Temporal effects (LRT) uses the temporal influence through separately learning the user check-in preferences to locations at each time slot from the check-in user-location matrix at the corresponding time slot only based on matrix factorization with the temporal regularization term. Note that originally LRT does not suggest a time slot for a user to visit a recommended location, but we can straightforwardly apply LRT to make time-aware location recommendations without aggregating the learned user check-in preferences at different time slots. To make fair comparison, LRT also utilizes the geo-social influence to compute the prior visiting probability, as shown in Section 3.2.
- UTE [22]: The User-based collaborative filtering with Temporal influence and smoothing Enhancement method (UTE) utilizes the temporal influence through separately inferring users' preferences to locations at each time slot from the check-in user-location matrix at the corresponding time slot only based on the user-based collaborative filtering. Note that UTE also incorporates the spatial (i.e., geographical) influence and temporal popularity [22].
- GTAG [23]: The method constructs a Geographical-Temporal Aware Graph (GTAG) for each time slot separately from the check-in data at the corresponding time slot only. Then, GTAG injects preferences to user nodes and propagates preferences to candidate location nodes via various paths in the graph based on geographical and temporal influences in order to find the locations with large preferences for each user at the corresponding time slot.
- TICRec: Our technique exploits the temporal influence correlations based on the proposed probabilistic framework in Section 3.

4.3 Performance Metrics

Recommendation accuracy. In general, time-aware location recommendation techniques compute a score for each POI regarding a target user at time interval T and return locations with the **top-k** highest scores as a recommendation result to the target user. To evaluate the quality of location recommendations, it is important to find out how many locations that are being recommended are actually visited by the target user at the corresponding time interval T in the testing data set. Also, it is important to know how many locations actually visited at T were recommended by the evaluated technique for the corresponding time interval T. The former aspect is captured by *precision* and the latter by *recall*. Precision and recall are standard metrics used to evaluate the recommendation accuracy [6], [9], [22], [23]. We define a *discovered* location as a location that is both recommended and actually visited by the target user at the same time interval T. Formally,

- Precision at T defines the ratio of the number of discovered locations for T to the total number k of recommended locations, i.e.,

$$\text{precision}(T) = \frac{\text{No. of discovered locations for } T}{k}.$$

- Recall at T defines the ratio of the number of discovered locations for T to the total number of actually visited locations at T by the target user in the testing set, i.e.,

$$\text{recall}(T) = \frac{\text{No. of discovered locations for } T}{\text{No. of visited locations at } T}.$$

Recommendation time error. To comprehensively compare TICRec to LRT, UTE and GTAG, we also investigate the time deviation between *the recommended visit time* and *the actual visit time* of a user to a location which can be measured by the popular metric: mean absolute error (MAE), calculated by:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{t}(l_i) - t(l_i)|,$$

where $\hat{t}(l_i)$ is the time at which the evaluated technique recommends a user to visit location l_i while $t(l_i)$ is the actual time at which the same user visits location l_i.

4.4 Parameter Settings

We examine (1) the effect of the number of recommended locations for users (top-k from 1 to 10), the number of visited locations of users in the testing set (from 1 to 10), and the length of time slots on precision and recall; and (2) the effect of the time interval (T for each hour of a day) and the number of visited locations of users in the training set (given-n from 5 to 50) on precision, recall and MAE.

Note that: The number of recommended locations for users (top-k), the number of visited locations of users in the testing set, and the length of time slots have no effect on MAE. k is set to a smaller number than n because a large number of recommended locations may not be helpful for users, and the number of visited locations of users at an hour in the testing set is often less than 10. Since the three existing methods [21], [22], [23] split a day into 24 hour slots, for comparison we compute the probability $P(l|u, T, D)$ in Equation 1 for each hour of a day by default, and also study the effect of different lengths of time slots. Precision, recall and MAE are averaged on all users for the effect of T, k and length of time slots, or a subset of users having a given number of visited locations in the training or testing set.

5 Experimental Results

This section analyzes our extensive experimental results. We compare our TICRec against the state-of-the-art location recommendation techniques [15], [21], [22], [23] in terms of *recommendation accuracy* (Section 5.1) and *recommendation time error* (Section 5.2), using two large-scale real data sets collected from Foursquare and Gowalla.

It is worth emphasizing that, unlike prediction techniques for trajectory data, the accuracy of all recommendation techniques for LBSNs is usually not high, because the density of user-location check-in matrix is pretty low. For example, the reported maximum precision is 0.035 over a data set with 8.84×10^{-4} density in [21], and 0.03 over two data sets with 9.85×10^{-4} and 6.35×10^{-3} densities in [22]. Even worse, the two data sets used in our experiments have a lower density, 2.3×10^{-4} in the Foursquare data set and 2.9×10^{-5} in the Gowalla data set (TABLE 2), so the relatively low precision and recall values are common and reasonable in the experiments. Thus, we focus on the relative accuracy of TICRec compared to the state-of-the-art techniques, which we expect that TICRec can improve recommendation accuracy as more check-in activities are logged in Section 5.1. In addition to the precision and recall, we also study the time error in location recommendations generated by the evaluated techniques in Section 5.2.

5.1 Recommendation Accuracy

Here we compare the recommendation accuracy of iGeoRec, LRT, UTE, GTAG, and TICRec with respect to the effect of various time intervals of weekdays and weekends (Fig. 7), numbers of recommended locations for users (Fig. 8), numbers of visited locations of users in the training set (Fig. 9), numbers of visited locations of users in the testing set (Fig. 10), and lengths of time slots (Fig. 11). At first, we conclude **the most important and general findings** in all experiments on two large-scale real data sets collected from Foursquare and Gowalla as follows.

iGeoRec. iGeoRec does not use the influence of the temporal context when users visiting locations based on their check-in behaviors and simply suggests a random time slot for a user to visit a recommended location. As a result, it usually fails to advise a proper visiting time slot for the user to the recommended location. Accordingly, iGeoRec returns the most inaccurate locations in terms of precision and misses most locations actually visited by target users in terms of recall, especially on the Gowalla data set. This experimental result shows that it is ineffective to straightforwardly apply the *non-time-aware* location recommendation techniques to make the *time-aware* recommendations, even though they perform well in the non-time-aware recommendations.

LRT. LRT utilizes the temporal influence through dividing the check-in user-location matrix into sub-matrices for each time slot and learning the user preferences to locations at each time slot based on matrix factorization collaborative filtering techniques. In general, LRT is greatly superior to iGeoRec according to their recommendation accuracy, which shows that the temporal influence is an essential factor to recommend time-aware location recommendations. However, the overall performance of LRT still does not perform well in contrast to our TICRec. The reason is that LRT cannot correlate the temporal influences of *different users visiting different locations at different time slots* so as to deal with the sub-matrices with a very low density at each time slot, although it uses the regularization term in matrix factorization to model the temporal consecutiveness of *the same user visiting the same location at two neighboring time slots*.

UTE. By utilizing the smoothing technique on *every pair of time slots* instead of *only two neighboring time slots* in LRT to increase the density of the check-in user-location sub-matrices at each time slot, UTE improves the precision and recall to some extent in comparison to LRT. Nonetheless,

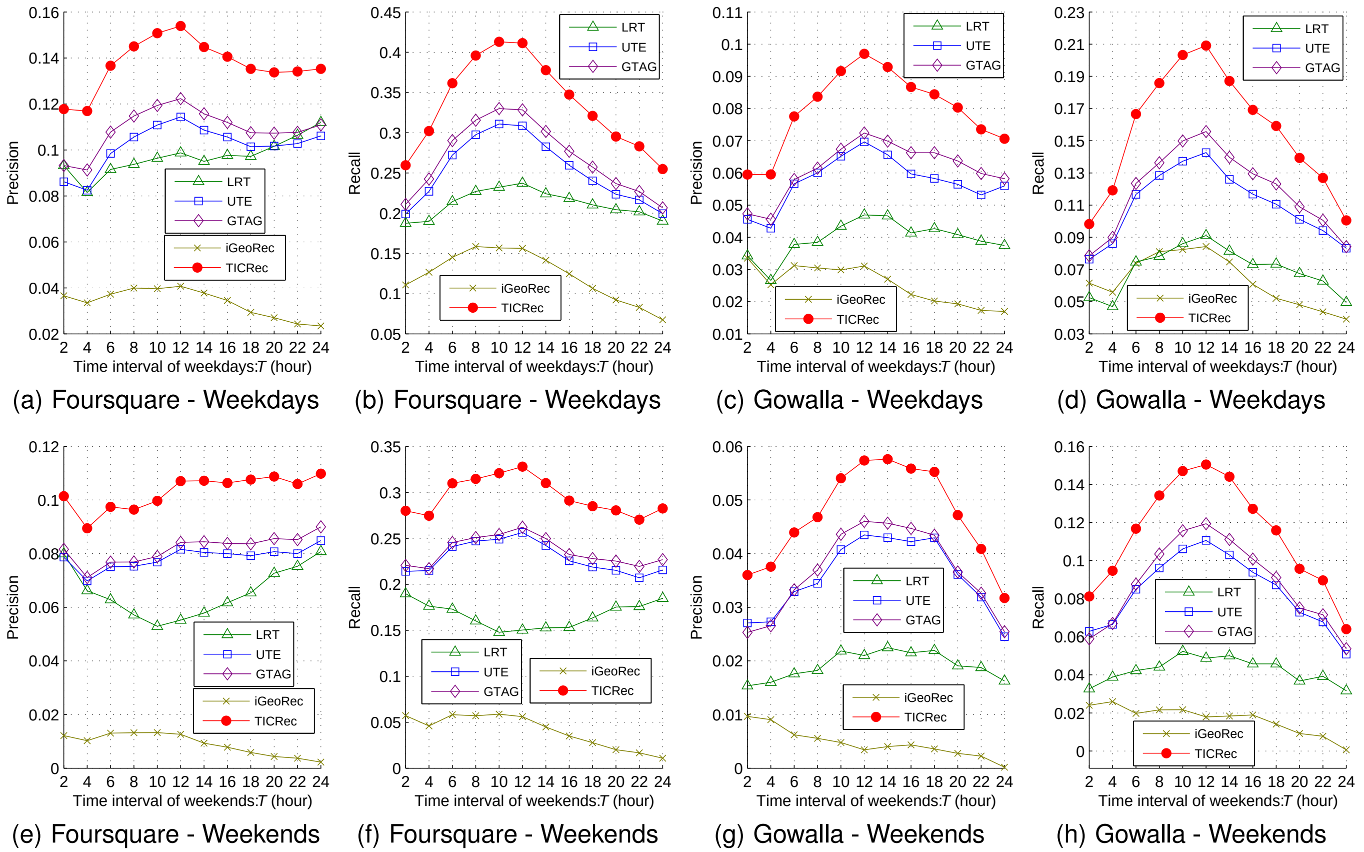


Fig. 7. Effect of hours of weekdays and weekends on recommendation accuracy

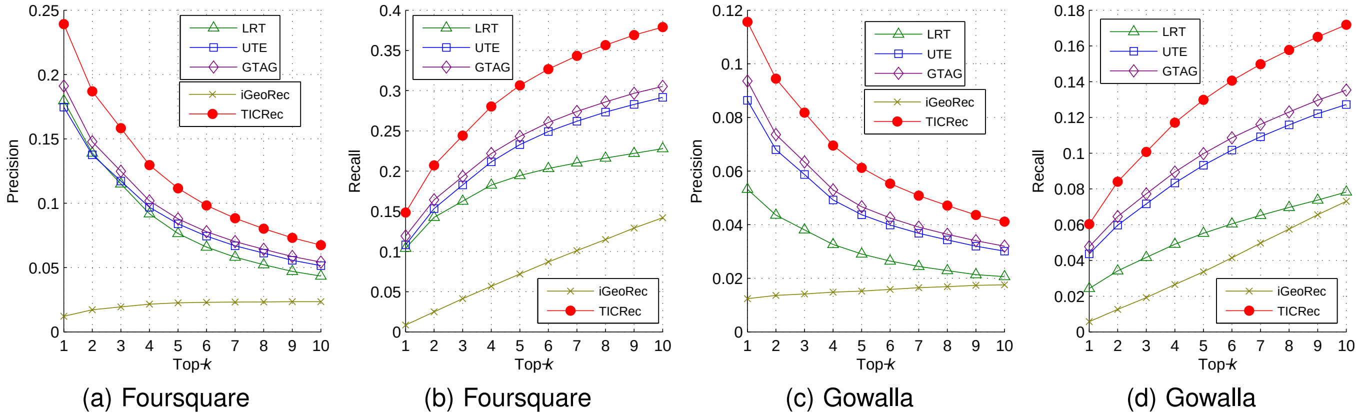
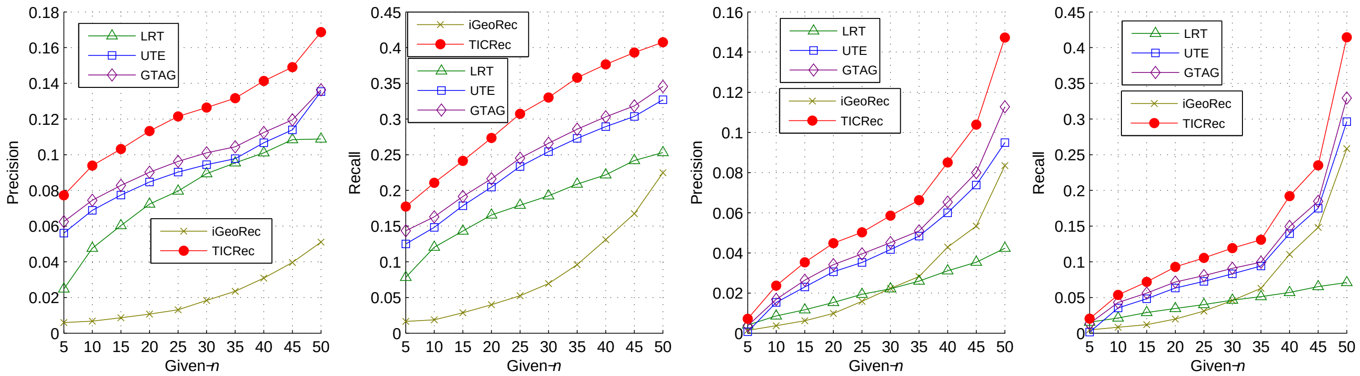


Fig. 8. Effect of numbers of recommended locations for users on recommendation accuracy

like LRT it still inherits the two major limitations: (a) the loss of time information because of dividing a day into time slots and (b) the lack of temporal influence correlations due to modeling users' preferences to locations for each slot separately.

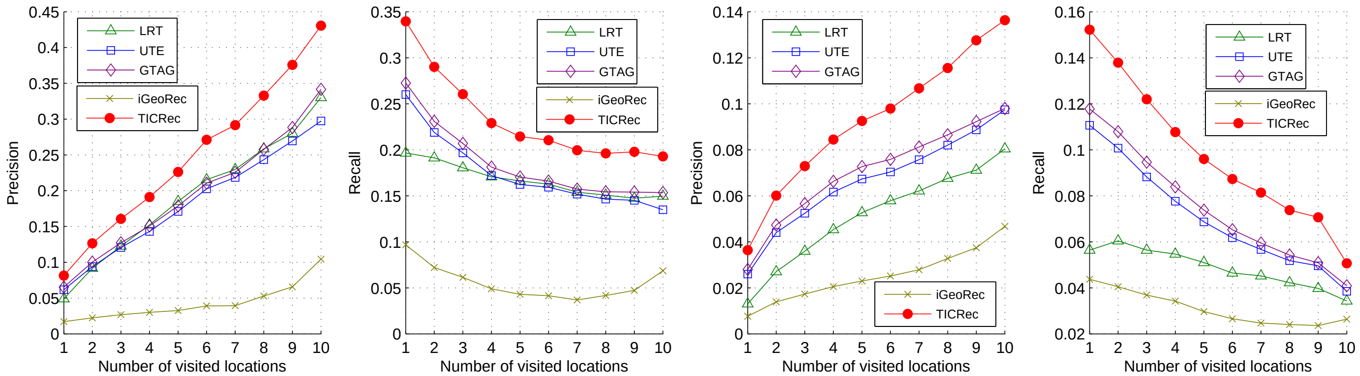
GTAG. As opposed to UTE that applies a *linearly weighted method* to combine the geographical and temporal influences into the final preference score for a user to a location at each time slot, GTAG employs a *graph-based preference propagation method* to integrate the geographical and temporal influences at each time slot. Consequentially, GTAG generally outperforms UTE based on their precision and recall, but the improvement is very limited since GTAG has the same two limitations as UTE.

TICRec. By exploiting the two important kinds of temporal influence correlations, our TICRec always exhibits the best recommendation quality in terms of precision and recall. The main reason is that when estimating the probability of a user visiting a new location at a given time, TICRec leverages (a) the time history of other users visiting the location based on the user-based TIC and (b) the time history of the user visiting other locations based on the location-based TIC, in which the time history is well weighed by the similarity between users or locations, respectively. These



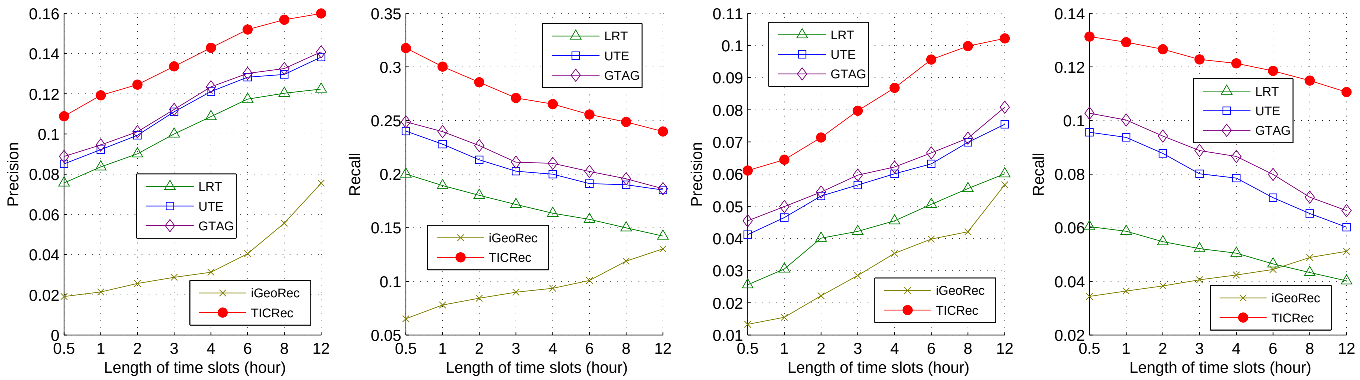
(a) Foursquare (b) Foursquare (c) Gowalla (d) Gowalla

Fig. 9. Effect of numbers of visited locations of users in the training set on recommendation accuracy



(a) Foursquare (b) Foursquare (c) Gowalla (d) Gowalla

Fig. 10. Effect of numbers of visited locations of users in the testing set on recommendation accuracy



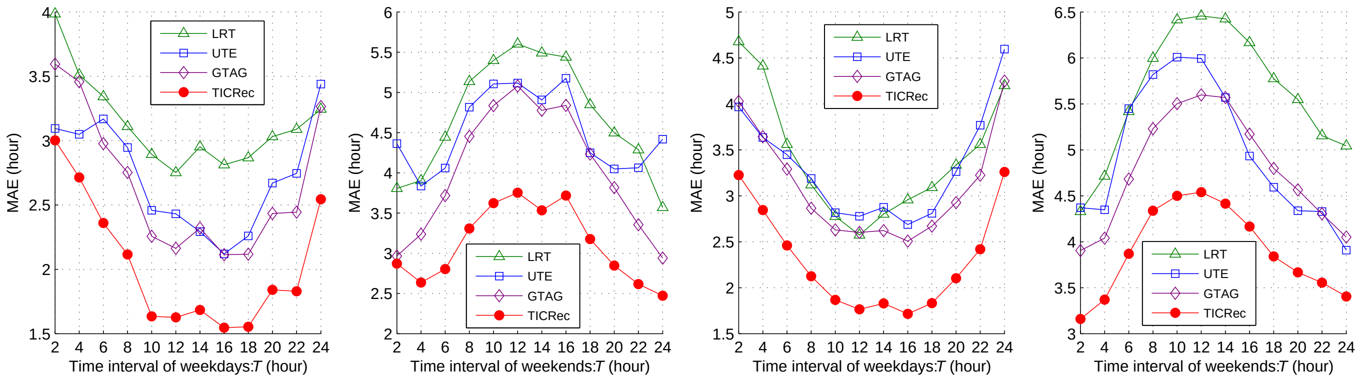
(a) Foursquare (b) Foursquare (c) Gowalla (d) Gowalla

Fig. 11. Effect of lengths of time slots on recommendation accuracy

promising results verify the superiority of exploiting the temporal influence correlations for location recommendations proposed in this paper over analyzing the temporal influence in a separated manner as in LRT, UTE and GTAG.

5.1.1 Effect of hours in weekdays and weekends

Fig. 7 depicts the recommendation accuracy with respect to varying the time intervals, i.e., hours in weekdays and weekends. (1) Interestingly, from 2:00 to 24:00 (or 0:00), the precision and recall of most time-aware location recommendation methods (e.g., TICRec) steadily increase at first and then gradually decrease. Usually, they achieve the best accuracy around noon. Our explanation is that at noon most users leave their office or home and visit some other places, e.g., restaurants for lunch; as a result, the density of check-ins at this hour is higher than the others. (2) The recommendation accuracy of all evaluated methods in the weekdays is higher than that in the weekends. The reason is that: the check-in behaviors of users are regular on weekdays, i.e., they usually travel between their home



(a) Foursquare - Weekdays (b) Foursquare - Weekends (c) Gowalla - Weekdays (d) Gowalla - Weekends

Fig. 12. Effect of hours of weekdays and weekends on time error

and office. In contrast, the check-in behaviors of users on weekends are diverse, e.g., indoorsy persons like visiting venues around their living areas while outdoorsy persons prefer traveling around the world to explore new interesting places.

5.1.2 Effect of numbers of recommended locations for users

Fig. 8 depicts the recommendation accuracy regarding the various numbers of recommended locations for users, i.e., k is increased from 1 to 10. Aside from the accuracy of iGeoRec, in general the precision steadily gets lower while the recall gradually becomes higher with the increase of k. The reason is that by returning more locations for users, it is always able to discover more locations that users would like to check in, but some recommended locations are less possible to be liked by users due to their lower visiting probabilities. Note that the recommendation techniques return the top-k locations based on the estimated visiting probability, for example, the second recommended location has the lower visiting probability than the first one.

5.1.3 Effect of numbers of visited locations of users in the training set

Fig. 9 depicts the recommendation accuracy regarding the change of the number of visited locations of users in the training set. For instance, a measure at "Given-n = 5" is averaged on all users who have checked in five locations in the training set. As users check in more locations, our TICRec can more accurately estimate the time probability density and predict the visiting probability of new locations for these users at the corresponding hour through using more check-in data. As a result, their precision and recall incline. We have the similar observation in our previous work [15].

5.1.4 Effect of numbers of visited locations of users in the testing set

Fig. 10 depicts the recommendation accuracy with respect to varying the number of visited locations of users in the testing set. For example, a measure at "Number of visited locations = 5" is averaged on all users who have checked in five locations in the testing data set. As the number gets larger, the precision generally increases but the recall usually decreases. The reason is that: the raise of the number of the visited locations in the testing set means that the recommendation techniques are more capable of discovering locations that users would like to visit but it is hard to discover all of this kind of locations. We also have the similar observation in our previous work [15].

5.1.5 Effect of lengths of time slots

Fig. 11 depicts the recommendation accuracy with regard to varying the length of time slots which determines the time granularity of time-aware location recommendations. (1) As the length of time slots gets larger, the precision of all recommendation methods gradually inclines in Figs. 11a and 11c. The reason is twofold: (i) A larger length of time slots indicates that the recommendation results will be less time-specific. (ii) The check-in matrix at the larger time slot becomes denser that benefits for recommendation models to estimate the accurate visiting probability of users to locations. (2) Nevertheless, the larger length of time slots brings in a larger number of ground truth locations, i.e., the locations visited by users in the testing set at each time slot, which will counteract the increase of discovered locations from higher precision under the longer time slots in Figs. 11a and 11c. Subsequently, the recall of time-aware recommendation techniques including LRT, UTE, GTAG, and TICRec drops in Figs. 11b and 11d, but the recall of iGeoRec still increases because its precision raises dramatically that dominates the effect of the larger number of ground truth locations. (3) More importantly, our TICRec consistently outperforms the state-of-the-art time-aware location recommendation technique for all lengths of time slots, since it models the continuous time probability densities of users visiting locations independently of the time granularity based on temporal influence correlations.

5.2 Recommendation Time Error

To further compare the performance of TICRec, LRT, UTE and GTAG, Fig. 12 depicts their recommendation time error on the time intervals from 2:00 to 24:00 (or 0:00) of weekdays and weekends. Interestingly, the recommendation time error on weekdays (Figs. 12a and 12c) is quite distinct from that on weekends (Figs. 12b and 12d). (1) On the weekdays,

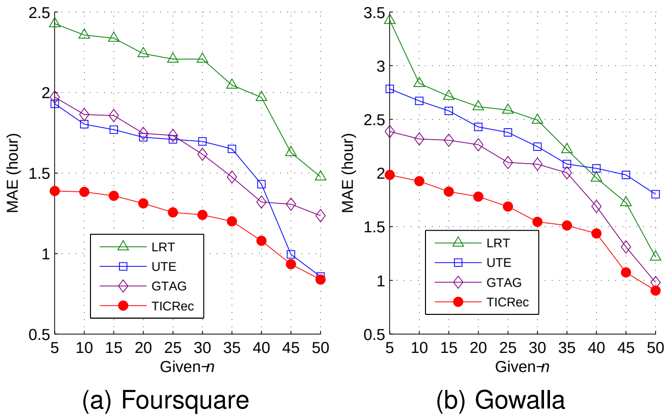


Fig. 13. Effect of numbers of visited locations of users in the training set on time error

the minimum MAE lies in the daytime, i.e., from 10:00 to 18:00, since most users regularly visit POIs around their office for different purposes, e.g., checking in restaurants at noon for dinner. From 20:00 to 6:00, after work users usually go out for relaxation and their check-in behavior has more diversity relative to the daytime. Thus, it is more difficult to recommend the correct visit time for users to POIs and then the MAE reaches the maximum. (2) Reversely on the weekends, the maximum MAE often occurs in the daytime, because the users' check-in behaviors on POIs are highly diverse during daytime, e.g., they usually explore different categories of POIs in a new city on weekends. Around midnight, i.e., from 22:00 to 2:00, the four methods record the minimum MAE, since the POIs visited at this time have the strong time indicator. For example, some bars open only at night and users often go there on weekends. (3) The time error on weekdays is lower than that on weekends, which is in accordance with Fig. 7.

Fig. 13 depicts the recommendation time error regarding the change of the number of visited locations of users in the training set. As expected, the MAE gradually decreases since these time-aware recommendation methods have more information about users' check-in behaviors and can infer users' time preferences more accurately as users check in more locations. Promisingly, TICRec accomplishes the time error of less than two hours when users check in more than five locations. In addition, the MAE of TICRec is still significantly lower than that of LRT, UTE and GTAG.

6 Conclusion and Future Directions

In this paper, we proposed TICRec; a probabilistic framework to utilize temporal influence correlations (TICs) for location recommendations in location-based social networks (LBSNs). TICRec overcomes two major limitations in existing time-aware location recommendation techniques. In our TICRec, we use kernel density estimation (KDE) method to estimate a continuous time probability density of a user visiting a new location to avoid the time information loss. To incorporate TICs in TICRec, our time probability density considers (1) **user-based TIC** by correlating the check-in behaviors of different users to the same location at different times and (2) **location-based TIC** by correlating the check-in behaviors of the same user to different locations at different times. Finally, we have conducted extensive experiments to evaluate the recommendation accuracy and time error of TICRec using the two real data sets collected from Foursquare and Gowalla. Experimental results show that TICRec provides significantly better performance than the state-of-the-art time-aware location recommendation techniques.

In the future, we plan to study two directions of location recommendations to extend TICRec: (a) how to integrate the temporal influence with the social and geographical influences more effectively, and (b) how to incorporate the category information of locations into the probabilistic location recommendation framework.

Acknowledgements

Jia-Dong Zhang and Chi-Yin Chow were partially supported by Guangdong Natural Science Foundation of China under Grant S2013010012363 and a research grant (CityU Project No. 9231131).

References

[1] J. J. Levandoski, M. Sarwat, A. Eldawy, and M. F. Mokbel, "LARS: A location-aware recommender system," in *IEEE ICDE*, 2012.

[2] E. H.-C. Lu, C.-Y. Chen, and V. S. Tseng, "Personalized trip recommendation with multiple constraints by mining user check-in behaviors," in *ACM SIGSPATIAL*, 2012.

[3] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "A random walk around the city: New venue recommendation in location-based social networks," in *IEEE SocialCom*, 2012.

[4] M. Ye, P. Yin, and W.-C. Lee, "Location recommendation for location-based social networks," in *ACM SIGSPATIAL*, 2010.

[5] J. J.-C. Ying, W.-N. Kuo, V. S. Tseng, and E. H.-C. Lu, "Mining user check-in behavior with a random walk for urban point-of-interest recommendations," *ACM TIST*, vol. 5, no. 3, pp. 40:1–40:26, 2014.

[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] H. Gao, J. Tang, and H. Liu, "gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks," in *ACM CIKM*, 2012.

[8] H. Wang, M. Terrovitis, and N. Mamoulis, "Location recommendation in location-based social networks using user check-in data," in *ACM SIGSPATIAL*, 2013.

[9] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *ACM SIGIR*, 2011.

[10] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *ACM SIGSPATIAL*, 2012.

[11] T. Kurashima, T. Iwata, T. Hoshide, N. Takaya, and K. Fujimura, "Geo topic model: Joint modeling of user's activity area and interests for location recommendation," in *ACM WSDM*, 2013.

[12] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," *IEEE TKDE, accepted to appear*, 2014.

[13] H. Yin, B. Cui, Y. Sun, Z. Hu, and L. Chen, "LCARS: A spatial item recommender system," *ACM TOIS*, vol. 32, no. 3, pp. 11:1–11:37, 2014.

[14] J.-D. Zhang and C.-Y. Chow, "iGSLR: Personalized geo-social location recommendation - a kernel density estimation approach," in *ACM SIGSPATIAL*, 2013.

[15] J.-D. Zhang, C.-Y. Chow, and Y. Li, "iGeoRec: A personalized and efficient geographical location recommendation framework," *IEEE TSC, accepted to appear*, 2014.

[16] S. Bannur and O. Alonso, "Analyzing temporal characteristics of check-in data," in *WWW Companion*, 2014.

[17] H. Gao, J. Tang, X. Hu, and H. Liu, "Modeling temporal effects of human mobile behavior on location-based social networks," in *ACM CIKM*, 2013.

[18] M. Ye, K. Janowicz, C. Mülligann, and W.-C. Lee, "What you are is when you are: The temporal dimension of feature types in location-based social networks," in *ACM SIGSPATIAL*, 2011.

[19] K. Zhang, Q. Jin, K. Pelechrinis, and T. Lappas, "On the importance of temporal dynamics in modeling urban activity," in *ACM UrbComp*, 2013.

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

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

[22] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann, "Time-aware point-of-interest recommendation," in *ACM SIGIR*, 2013.

[23] Q. Yuan, G. Cong, and A. Sun, "Graph-based point-of-interest recommendation with geographical and temporal influences," in *ACM CIKM*, 2014.

[24] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *ACM KDD*, 2011.

[25] B. W. Silverman, *Density estimation for statistics and data analysis*. Chapman and Hall, 1986.

[26] J. Bao, Y. Zheng, D. Wilkie, and M. F. Mokbel, "A survey on recommendations in location-based social networks," *submitted to GeoInformatica*, 2014.

[27] A.-J. Cheng, Y.-Y. Chen, Y.-T. Huang, W. H. Hsu, and H.-Y. M. Liao, "Personalized travel recommendation by mining people attributes from community-contributed photos," in *ACM MM*, 2011.

[28] K. Lee, R. Ganti, M. Srivatsa, and P. Mohapatra, "Spatio-temporal provenance: Identifying location information from unstructured text," in *IEEE PERCOM Workshops*, 2013.

[29] A. X. Zhang, A. Noulas, S. Scellato, and C. Mascolo, "Hoodsquare: Modeling and recommending neighborhoods in location-based social networks," in *IEEE SocialCom*, 2013.

[30] D. Zhou, B. Wang, S. M. Rahimi, and X. Wang, "A study of recommending locations on location-based social network by collaborative filtering," in *Canadian AI*, 2012.

[31] Y. Ge, Q. Liu, H. Xiong, A. Tuzhilin, and J. Chen, "Cost-aware travel tour recommendation," in *ACM KDD*, 2011.

[32] C.-C. Hung, W.-C. Peng, and W.-C. Lee, "Clustering and aggregating clues of trajectories for mining trajectory patterns and routes," *VLDB Journal, accepted to appear*, 2011.

[33] K. W.-T. Leung, D. L. Lee, and W.-C. Lee, "CLR: A collaborative location recommendation framework based on co-clustering," in *ACM SIGIR*, 2011.

[34] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Smart itinerary recommendation based on user-generated GPS trajectories," in *Ubiquitous Intelligence and Computing*, Z. Yu, R. Liscano, G. Chen, D. Zhang, and X. Zhou, Eds. Berlin: Springer, 2010, pp. 19–34.

[35] V. W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang, "Collaborative filtering meets mobile recommendation: A user-centered approach," in *AAAI*, 2010.

[36] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with GPS history data," in *WWW*, 2010.

[37] ——, "Towards mobile intelligence: Learning from gps history data for collaborative recommendation," *Artificial Intelligence*, vol. 184-185, pp. 17–37, 2012.

[38] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma, "Recommending friends and locations based on individual location history," *ACM TWEB*, vol. 5, no. 1, pp. 5:1–5:44, 2011.

[39] B. Hu and M. Ester, "Spatial topic modeling in online social media for location recommendation," in *ACM RecSys*, 2013.

[40] X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, "Personalized QoS-aware web service recommendation and visualization," *IEEE TSC*, vol. 6, no. 1, pp. 35–47, 2013.

[41] H. Sun, Z. Zheng, J. Chen, and M. R. Lyu, "Personalized web service recommendation via normal recovery collaborative filtering," *IEEE TSC*, vol. 6, no. 4, pp. 573–579, 2013.

[42] Z. Zheng, H. Ma, M. R. Lyu, and I. King, "Collaborative web service QoS prediction via neighborhood integrated matrix factorization," *IEEE TSC*, vol. 6, no. 3, pp. 289–299, 2013.

[43] Y.-C. Chen, W.-Y. Zhu, W.-C. Peng, W.-C. Lee, and S.-Y. Lee, "CIM: Community-based influence maximization in social networks," *ACM TIST*, vol. 5, no. 2, pp. 25:1–25:31, 2014.

[44] H. Ma, T. C. Zhou, M. R. Lyu, and I. King, "Improving recommender systems by incorporating social contextual information," *ACM TOIS*, vol. 29, no. 2, pp. 9:1–9:23, 2011.

[45] Y. Shen and R. Jin, "Learning personal + social latent factor model for social recommendation," in *ACM KDD*, 2012.

[46] G. Ference, M. Ye, and W.-C. Lee, "Location recommendation for out-of-town users in location-based social networks," in *ACM CIKM*, 2013.

[47] X. Liu, Y. Liu, K. Aberer, and C. Miao, "Personalized point-of-interest recommendation by mining users' preference transition," in *ACM CIKM*, 2013.

[48] C. Wang, M. Ye, and W.-C. Lee, "From face-to-face gathering to social structure," in *ACM CIKM*, 2012.

[49] J.-D. Zhang, C.-Y. Chow, and Y. Li, "LORE: Exploiting sequential influence for location recommendations," in *ACM SIGSPATIAL*, 2014.

[50] J.-D. Zhang and C.-Y. Chow, "CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations," *Information Sciences*, vol. 293, pp. 163–181, 2015.

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

[52] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *IJCAI*, 2013.

[53] A. Monreale, F. Pinelli, R. Trasarti, and F. Giannotti, "WhereNext: A location predictor on trajectory pattern mining," in *ACM KDD*, 2009.

[54] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *IEEE ICDM*, 2012.

[55] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and Z. Xu, "Destination prediction by sub-trajectory synthesis and privacy protection against such prediction," in *IEEE ICDE*, 2013.

[56] J. J.-C. Ying, W.-C. Lee, and V. S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction," *ACM TIST*, vol. 5, no. 1, pp. 2:1–2:33, 2013.

[57] J. J.-C. Ying, W.-C. Lee, T.-C. Weng, and V. S. Tseng, "Semantic trajectory mining for location prediction," in *ACM SIGSPATIAL*, 2011.

[58] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994.

[59] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, 2nd ed. New York: Dover Publications, 2007.

Jia-Dong Zhang received the M.Sc. degree in computer science and engineering from Yunnan University, China, in 2009. He is currently working toward the Ph.D. degree with the Department of Computer Science, City University of Hong Kong. His research work has been published in premier conferences (e.g., *ACM SIGSPATIAL*) and journals (e.g., *IEEE TDSC*, *IEEE TSC*, *IEEE TITS*, *Pattern Recognition* and *Information Sciences*). His research interests include location-based services and location privacy.

Chi-Yin Chow received the M.S. and Ph.D. degrees from the University of Minnesota-Twin Cities in 2008 and 2010, respectively. He is currently an assistant professor in Department of Computer Science, City University of Hong Kong. His research interests include spatio-temporal data management and analysis, GIS, mobile computing, and location-based services. He is the co-founder and co-organizer of ACM SIGSPATIAL MobiGIS 2012, 2013, and 2014.