# LORE: Exploiting Sequential Influence for Location Recommendations

Jia-Dong Zhang
Dept. of Computer Science
City University of Hong Kong
jzhang26-c@my.cityu.edu.hk

Chi-Yin Chow
Dept. of Computer Science
City University of Hong Kong
chiychow@cityu.edu.hk

Yanhua Li
HUAWEI Noah's Ark Lab
Hong Kong Science Park
li.yanhua1@huawei.com

## ABSTRACT

Providing location recommendations becomes an important feature for location-based social networks (LBSNs), since it helps users explore new places and makes LBSNs more prevalent to users. In LBSNs, *geographical influence* and *social influence* have been intensively used in location recommendations based on the facts that geographical proximity of locations significantly affects users' check-in behaviors and social friends often have common interests. Although human movement exhibits sequential patterns, most current studies on location recommendations do not consider any *sequential influence* of locations on users' check-in behaviors. In this paper, we propose a new approach called LORE to exploit sequential influence on location recommendations. First, LORE incrementally mines sequential patterns from location sequences and represents the sequential patterns as a *dynamic Location-Location Transition Graph* ($\mathrm{L^2TG}$). LORE then predicts the probability of a user visiting a location by *Additive Markov Chain* (AMC) with $\mathrm{L^2TG}$. Finally, LORE fuses *sequential influence* with *geographical influence* and *social influence* into a unified recommendation framework; in particular the geographical influence is modeled as *two-dimensional check-in probability distributions* rather than *one-dimensional distance probability distributions* in existing works. We conduct a comprehensive performance evaluation for LORE using two large-scale real data sets collected from Foursquare and Gowalla. Experimental results show that LORE achieves significantly superior location recommendations compared to other state-of-the-art recommendation techniques.

## Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications – Spatial Databases and GIS; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – Information Filtering; I.5 [**Computing Methodologies**]: Pattern Recognition


*SIGSPATIAL'14,* November 04 - 07 2014, Dallas/Fort Worth, TX, USA
Copyright 2014 ACM 978-1-4503-3131-9/14/11 ...$15.00
http://dx.doi.org/10.1145/2666310.2666400.

## General Terms

Algorithms, Experimentation.

## Keywords

Location-based social networks, location recommendations, geographical influence, social influence, sequential influence, additive Markov chain

## 1. INTRODUCTION

With the advancement of mobile devices and location acquisition technologies, location-based social networks (LBSNs) (e.g., Foursquare and Gowalla), have attracted millions of users [34]. In the LBSNs, it is crucial to utilize a variety of community-contributed data to make personalized location recommendations to users, which help them explore new places and make LBSNs more attractive to them.

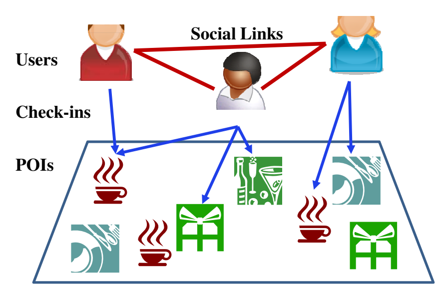



**Figure 1: A location-based social network**

In an LBSN (Figure 1), users can establish social links and share their experiences of visiting some specific locations, also known as *points-of-interest* (POIs), e.g., restaurants, stores and museums. These visits are also known as *check-in* activities that reflect users' preferences on locations. A promising way for location recommendations is to apply geographical-social location recommendation techniques [1, 4, 8, 11, 15, 22, 24, 28, 30]. These techniques recommend locations (i.e., POIs) by utilizing *geographical influence* derived from geographical information of check-in POIs and *social influence* from social links, since the geographical proximity of locations significantly influences users' check-in behaviors on locations and social friends are more likely to share common interests on locations. However, all these studies do not consider **the influence of sequential patterns of check-in locations on users' check-in behaviors**, called **sequential influence** hereafter, although in reality human movement exhibits sequential patterns [6, 9, 19]. For example, people usually go to cinemas or bars after restaurants since they would like to relax after dinner.

Therefore, this paper is motivated to enhance the quality of location recommendations in LBSNs by leveraging sequential influence on users' check-in behaviors. To this end, we propose a new **LO**cation **RE**commendation approach with sequential influence based on additive Markov chain (AMC), called **LORE**. (1) In LORE, we first **incrementally** mine sequential patterns from check-in location sequences of all users as a **dynamic** *location-location transition graph* ($L^2TG$), where a location sequence consists of check-in locations of the same user ordered by check-in time. $L^2TG$ incorporates not only *transition counts between locations* but also *outgoing counts of locations to other locations* in order to incrementally update the obtained sequential patterns; transition probabilities can be dynamically calculated through dividing transition counts by outgoing counts. (2) Then, instead of employing the well-known first-order Markov chain that is widely used to discover the most popular location sequence patterns for users [2, 3, 5, 12, 36], we develop an efficient $n$th-order additive Markov chain to predict the sequential probability of a user visiting a new location given $L^2TG$ and her visited location sequence. The main reason is that the new location depends on not only the latest visited location but also the earlier visited locations in the sequence. (3) We finally fuse *sequential influence* with *geographical influence* and *social influence* by integrating their derived *sequential probability* with *geographical probability* and *social rating* of a user visiting a new location into a unified score and rank the score to recommend the top-$k$ new locations.

The main contributions of this paper can be summarized:

- We exploit *the sequential influence of locations* on users' check-in behaviors for location recommendations in LBSNs through the proposed dynamic $L^2TG$ and additive Markov chain. (Section 3)
- We integrate the *sequential influence* with *geographical influence* and *social influence* into a unified recommendation framework. In particular, to model the geographical influence, we estimate *two-dimensional (latitude and longitude) check-in probability distributions* of users to locations, which is more reasonable and intuitive than using *one-dimensional distance probability distributions* in existing works. (Section 4)
- We conduct extensive experiments to evaluate the performance of LORE using two large-scale real data sets collected from Foursquare and Gowalla. Experimental results show that LORE outperforms other state-of-the-art recommendation techniques including the first-order Markov chain [2, 3, 5, 12, 36], geographical-social location recommendation methods [1, 4, 8, 11, 15, 22, 24, 28] and their combination in terms of recommendation accuracy. (Sections 5 and 6)

The remainder of this paper is organized as follows. Section 2 highlights related work. Section 3 exploits sequential influence for location recommendations using additive Markov chain with $L^2TG$. We present the fusion framework for integrating *sequential preference* with *geographical influence* and *social influence* in Section 4. In Sections 5 and 6, we evaluate LORE and analyze experimental results. Finally, we conclude this paper in Section 7.

## 2. RELATED WORK

In this section, we highlight related work on location recommendations. In general, there are four main categories for existing location recommendation approaches: *collaborative filtering*, *social influence*, *geographical influence*, and *sequential influence*.

**Collaborative filtering techniques.** Some studies provide POI recommendations by using the conventional collaborative filtering techniques on users' check-in data in LBSNs [14, 17, 23], GPS trajectory data [13, 35, 33], or text data [10]. However, the performance of all these techniques is considerably limited due to no consideration for the *social influence*, *geographical influence*, or *sequential influence*.

**Social influence.** Based on the fact that nearby friends are more likely to share common interests, social link information has been widely utilized to improve the quality of recommender systems in LBSNs [1, 4, 8, 11, 15, 22, 24, 28, 30]. Specifically, these works derive the similarities between users in terms of their social links or/and residence distances, and then integrate them into the collaborative filtering techniques.

**Geographical influence.** The geographical proximity between POIs significantly affects the check-in behaviors of users on the POIs. Thus, the influence of geographical information of locations on users' check-in behaviors also has been intensively used in location recommendations. For instance, the studies [1, 7, 22, 25] 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. More sophistically, the studies [4, 11, 15, 16, 21, 24, 27] 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 [28, 29, 30] personalize the geographical influence by modeling a personalized nonparametric distribution for each user.

**Sequential influence.** In terms of the fact that human movement exhibits sequential patterns [6, 9, 19, 31], various sequential mining techniques [26] have been developed 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. The current studies exploiting sequential influence for location recommendations can be classified into three groups. (1) Some researchers mined the most popular location sequence patterns from travel histories to guide users to plan a trip [2, 36]. However, they did not take the personalization into account as their approaches just return the same sequence patterns for all users. (2) In contrast, other researchers personalized the sequence patterns through modeling users' profiles based on facial attributes (gender, age and race) [3] or a mixture of topics in which a topic is a probability distribution over POIs [12]. These facial attributes and topic models are extracted from community-contributed photos. Nonetheless, the photo data are not often available in LBSNs. (3) The researchers in [5] utilized the sequential influence to recommend locations for users by learning a personalized model for each user based on her own check-in location sequence only. Nevertheless, the method in [5] requires a user with more than one hundred check-in locations so as to learn sequential patterns from them, which is not applicable to most users since they usually check in a few POIs in LBSNs. Further, all these studies [2, 3, 5, 12, 36] exploit the sequence influence based on the first-order Markov chain that only uses the latest visited location in a sequence of a user to recommend a new location for the user, but in reality the new location relies on not only the latest visited location but also the earlier visited locations.

We can distinguish our work from previous studies in twofold: (1) We take full advantage of the sequential influence to recommend new POIs for users based on the proposed location-location transition graph (L$^2$TG) and $n$th-order additive Markov chain. (2) We integrate sequential, geographical and social influences through a unified framework to enhance the quality of location recommendations.

# 3. MODELING SEQUENTIAL INFLUENCE

In this section, we describe how to represent the sequential patterns of check-in location sequences as a location-location transition graph (L$^2$TG) in Section 3.1 and how to determine the probability of a user visiting a location based on the $n$th-order additive Markov chain over L$^2$TG in Section 3.2.

## 3.1 Location-Location Transition Graph

**Statement of problem 1.** We extract sequential patterns from the location sequences of all users and model them as a concise L$^2$TG.

We first present some basic definitions for L$^2$TG.

Definition 1. ***Sequence.*** *A location sequence of user $u$ denoted by $S_u = \langle (l_1, t_1), (l_2, t_2), \ldots, (l_n, t_n)\rangle$ is a path such that user $u$ goes through location $l_i$ at time $t_i$ ($t_1 \le t_2 \le \cdots \le t_n$). We also use $S_u = \langle l_1, l_2, \ldots, l_n\rangle$ for short.*

Definition 2. ***Transition, predecessor and successor.*** *Given two consecutive locations $l_i$ and $l_{i+1}$ in the location sequence $S_u = \langle (l_1, t_1), (l_2, t_2), \ldots, (l_n, t_n)\rangle$ and a certain threshold $\Delta T$, if $t_{i+1} - t_i \le \Delta T$, there is a transition from $l_i$ to $l_{i+1}$, denoted by $l_i \to l_{i+1}$, where $l_i$ is a transition predecessor of $l_{i+1}$ and $l_{i+1}$ is a transition successor of $l_i$.*

By Definition 2, given a location sequence $S_u = \langle (l_1, t_1), (l_2, t_2), \ldots, (l_n, t_n)\rangle$ and a certain threshold $\Delta T$, if $t_{i+1}-t_i > \Delta T$, we do not say a transition occurs from $l_i$ to $l_{i+1}$ since the large time interval between $l_i$ and $l_{i+1}$ may indicate they are irrelevant. Moreover, $l_i$ can be a transition successor as well as a transition predecessor at the same time if and only if $t_i - t_{i-1} \le \Delta T$ and $t_{i+1} - t_i \le \Delta T$. In Section 6, we study the impact of $\Delta T$ on location recommendation quality.

Definition 3. ***L$^2$TG.*** *A location-location transition graph (L$^2$TG) $G = (L, E)$ consists of a set of nodes $L$ and a set of edges $E \subseteq L \times L$. Each node $l_i \in L$ represents a location associated with an outgoing count of $l_i$ as a transition predecessor to other locations denoted by $OCount(l_i)$. And each edge $(l_i, l_j) \in E$ represents a transition $l_i \to l_j$ associated with a transition count denoted by $TCount(l_i, l_j)$.*

**Example.** Figure 2 shows an example of L$^2$TG, where nodes (circles) and edges (arrows) denote locations and transitions between locations, respectively, and the number along with a location node or a transition edge is its outgoing count or transition count, respectively. Note that the outgoing count of a location is the number of times that the location is visited as a transition predecessor. For instance, the outgoing count of location $l_1$ is 5 because $l_1$ has been a transition

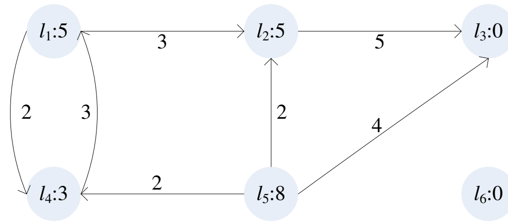



**Figure 2: An example of L$^2$TG**

predecessor of $l_2$ three times and $l_4$ twice. Although $l_3$ has been a transition successor nine times, its outgoing count is 0 because it has never been a transition predecessor.

In terms of *transition counts* and *outgoing counts* associated with L$^2$TG, *transition probabilities* can be determined based on Definition 4.

Definition 4. ***Transition probability.*** *If the outgoing count of $l_i$ is non-zero, i.e., $OCount(l_i) > 0$, the transition probability of $l_i \to l_j$, denoted $TP(l_i \to l_j)$, is calculated by*

$$TP(l_i \to l_j) = \frac{TCount(l_i, l_j)}{OCount(l_i)}. \tag{1}$$

*Otherwise, i.e., $OCount(l_i) = 0$, it is given by*

$$TP(l_i \to l_j) = \begin{cases} 1, & l_j = l_i; \\ 0, & l_j \ne l_i. \end{cases} \tag{2}$$

By Definition 4, if the outgoing count of $l_i$ is non-zero, the transition probability of $l_i \to l_j$ is defined as the proportion of $TCount(l_i, l_j)$ to $OCount(l_i)$ in Equation 1, which is essentially the relative frequency definition of probability. On the other hand, if $OCount(l_i) = 0$ that means all users do not check in any other locations after $l_i$ within the given time interval $\Delta T$; accordingly we define the transition probability of $l_i$ to itself is one for simplicity.

In Definition 3, L$^2$TG is associated with *transition counts* and *outgoing counts* instead of *transition probabilities* so that L$^2$TG can be incrementally updated in an online fashion.

**Online incremental maintenance of L$^2$TG.** We cannot process the location sequences $S_u$ with the order of user-by-user, since users' check-in locations will continuously arrive as time goes on. In fact, the continuously arriving check-in locations with timestamps from users, denoted by $(u_i, l_i, t_i)$, constitute an unbounded data stream, denoted by $(u_i, l_i, t_i)_{i=1}^{+\infty}$. As check-in locations possess the general characteristics of data streams: massive volume of data and temporal correlations [32], it is required to process the check-in locations according to their arriving order and incrementally update the constructed L$^2$TG.

**Algorithm.** Algorithm 1 depicts the pseudo code of the online incremental maintenance of L$^2$TG. It first initializes four variables, i.e., setting *latest_location* and *latest_time* to $\emptyset$, and setting all *outgoing counts* and *transition counts* to 0 (Lines 1 to 4). When Algorithm 1 receives a check-in location $l_i$ with timestamp $t_i$ of user $u_i$, i.e., $(u_i, l_i, t_i)$, it calculates the time interval between $t_i$ and the latest timestamp stored in *latest_time* for user $u_i$. If the time interval does not exceed the given threshold $\Delta T$, the corresponding *outgoing count* and *transition count* are increased by 1 (Lines 7 to 11). Algorithm 1 also updates the latest visiting location and timestamp for user $u_i$ with $l_i$ and $t_i$, respectively (Lines 13 and 14).

**Algorithm 1** The online incremental maintenance of L$^2$TG

**Input:** The numbers of users and locations in the system: $M$ and $N$; a certain threshold $\Delta T$; check-in stream $(u_i, l_i, t_i)_{i=1}^{+\infty}$.
**Output:** The incrementally updated L$^2$TG.

$latest\_location[M] \leftarrow \emptyset$
$latest\_time[M] \leftarrow \emptyset$
$OCount[N] \leftarrow 0$
$TCount[N][N] \leftarrow 0$
**while** detect a check-in activity of user $u_i$ at location $l_i$ and time $t_i$, i.e., $(u_i, l_i, t_i)$ **do**
  **if** $latest\_location[u_i] \neq \emptyset$ **then**
    **if** $t_i - latest\_time[u_i] \leq \Delta T$ **then**
      $l_{i-1} \leftarrow latest\_location[u_i]$
      $OCount[l_{i-1}] \leftarrow OCount[l_{i-1}] + 1$
      $TCount[l_{i-1}][l_i] \leftarrow TCount[l_{i-1}][l_i] + 1$
    **end if**
  **end if**
  $latest\_location[u_i] \leftarrow l_i$
  $latest\_time[u_i] \leftarrow t_i$
**end while**

**Complexity analysis.** Algorithm 1 works over the data stream $(u_i, l_i, t_i)_{i=1}^{+\infty}$ in only **one pass** that is an essential requirement for processing a data stream in real-time. Further, Algorithm 1 processes each update in a constant time $O(1)$. The space complexity of Algorithm 1 is $O(M + N^2) = O(N^2)$ that is the same as the item-based collaborative filtering method. It is important to note that: (1) the time and space complexities are independent of the size of a data stream, which is a nice property since the size is potentially infinite; (2) the space complexity of $TCount$ is $O(N^2)$, but in practice most entries of $TCount$ are zero-values and can be implemented by a sparse matrix.

## 3.2 Predicting Sequential Probabilities with Additive Markov Chain on L$^2$TG

**Statement of problem 2.** Based on the obtained L$^2$TG, this section focuses on predicting the sequential probability $p^{seq}(l_{n+1}|S_u)$ of a user $u$ visiting a new location $l_{n+1}$ given $u$'s visited location sequence ordered increasingly by their check-in timestamps $S_u = \langle l_1, l_2, \ldots, l_n \rangle$.

**First-order Markov chain as a baseline.** The current works [2, 3, 5, 12, 36] derive the sequential probability by employing the first-order Markov chain, which can be represented by Equation 1 as follows:

$$p^{seq}(l_{n+1}|S_u) = TP(l_n \rightarrow l_{n+1}) = \frac{TCount(l_n, l_{n+1})}{OCount(l_n)}. \quad (3)$$

The first-order Markov chain assumes the probability of visiting a new location $l_{n+1}$ only relies on the latest visited location $l_n$. Nevertheless, in reality the probability may depend on all the visited locations $l_1, l_2, \ldots, l_n$ in the sequence $S_u$. Hence, we are inspired to exploit the additive Markov chain [20], a highly important class of the $n$th-order Markov chain. Note that it is not efficient to apply the classical $n$th-order Markov chain, because its complexity (i.e., the number of states) increases exponentially with $n$.

**LORE's $n$th-order additive Markov chain.** We first define the general additive Markov chain (in Definition 5), and then we describe how to use it for LORE.

**Algorithm 2** Computation of the $n$th-order additive Markov chain on L$^2$TG

**Input:** L$^2$TG, $S_u = \langle l_1, l_2, \ldots, l_n \rangle$, $l_{n+1} \notin S_u$, $\alpha \geq 0$.
**Output:** $p^{seq}(l_{n+1}|S_u)$.

1: $X \leftarrow 0$; $Y \leftarrow 0$
2: **for** each $l_i \in S_u$ **do**
3:   $tp \leftarrow TCount(l_i, l_{n+1}) / OCount(l_i)$
4:   $w \leftarrow 2^{-\alpha \cdot (n-i)}$
5:   $X \leftarrow X + w \cdot tp$
6:   $Y \leftarrow Y + w$
7: **end for**
8: **return** $p^{seq}(l_{n+1}|S_u) \leftarrow X/Y$

Definition 5. ***General additive Markov chain [20].*** *Given a location sequence $S_u = \langle l_1, l_2, \ldots, l_n \rangle$, the additive Markov chain generally defines the sequential probability of visiting a new location $l_{n+1}$ by*

$$p^{seq}(l_{n+1}|S_u) = \sum\nolimits_{i=1}^{n} f(l_{n+1}, l_i, n+1-i), \quad (4)$$

*where $f(l_{n+1}, l_i, n+1-i)$ is the **additive contribution** of the location $l_i$ to the sequential probability $p^{seq}(l_{n+1}|S_u)$.*

**Weights of additive contribution.** In Definition 5, the key challenge is to determine the additive contribution of the location $l_i$ for a specific application. As done in the first-order Markov chain, the additive contribution of $l_i$ can be computed based on the transition probability of $l_i$ to $l_{n+1}$. Furthermore, the transition probability of $l_i$ to $l_{n+1}$ should be weighed through leaning towards recently visited locations, since the locations with recent check-in timestamps usually have stronger influence on a new possibly visiting location than the locations with old timestamps [6, 9, 19]. Thus, for the transition probability of $l_i$ to $l_{n+1}$, we develop a method to compute the weight of $l_i$ given by

$$W(l_i) = 2^{-\alpha \cdot (n-i)}, \quad (5)$$

where $2^{-\alpha \cdot (n-i)}$ represents the sequence decay weight with the decay rate parameter $\alpha \geq 0$ and the larger $\alpha$ is, the higher is the decay rate.

**Implementation of additive Markov chain.** Based on Equations 1 and 5, the additive contribution of the location $l_i$ is given by

$$f(l_{n+1}, l_i, n+1-i) = \frac{W(l_i) \cdot TP(l_i \rightarrow l_{n+1})}{\sum_{j=1}^{n} W(l_j)}. \quad (6)$$

We can derive the sequential probability of visiting a new location $l_{n+1}$ conditioned on the sequence $S_u$ based on the additive Markov chain in Definition 5 through

$$\begin{aligned} p^{seq}(l_{n+1}|S_u) &= \sum\nolimits_{i=1}^{n} f(l_{n+1}, l_i, n+1-i) \\ &= \sum\nolimits_{i=1}^{n} \frac{W(l_i) \cdot TP(l_i \rightarrow l_{n+1})}{\sum_{j=1}^{n} W(l_j)} \\ &= \frac{\sum_{i=1}^{n} W(l_i) \cdot TP(l_i \rightarrow l_{n+1})}{\sum_{i=1}^{n} W(l_i)}. \quad (7) \end{aligned}$$

**Algorithm.** To compute $p^{seq}(l_{n+1}|S_u)$ based on Equation 7, Algorithm 2 iteratively calculates the *transition probability* and *weight* for each $l_i \in S_u$ according to Equations 1 and 5, respectively (Lines 3 and 4), and then it sums them together (Lines 5 and 6).

**Complexity analysis.** Algorithm 2 runs over the sequence $S_u$ in one pass, so its time complexity is $O(n)$. Thus, Algorithm 2 is more efficient than collaborative filtering algorithms with the time complexity of $O(M)$ or $O(N)$ for prediction because of $M \gg n$ and $N \gg n$. As well, the space complexity of Algorithm 2 is $O(N^2)$.

# 4. FUSING *SEQUENTIAL* WITH *SOCIAL* AND *GEOGRAPHICAL* INFLUENCES

We can exploit the sequential influence only to make location recommendations for $u$ by returning the top-$k$ new locations $l_{n+1}$ with the highest probability (i.e., $p^{seq}(l_{n+1}|S_u)$) according to Equation 7. To further improve the quality of location recommendations, we integrate the *sequential* influence with *social* and *geographical* influences for LORE.

## 4.1 Using Social Influence

In the real world, friends often go to some places like movie theaters or restaurants together, or a user may travel on spots highly recommended by her friends. Thus, a user's preference on POIs can be influenced by her close friends or a group of friends that are likely to share some common interests. Based on these facts, we use the social friendships between users to recommend POIs based on the social location recommendation technique in our previous work [28].

Let $U$ be a set of users, $L$ be a set of locations (i.e., POIs), and $R$ be a user-location rating matrix derived from check-in activities, where each entry $r_{u,l}$ denotes the frequency of user $u \in U$ visiting location $l \in L$. Given a certain entry $r_{u,l} = 0$ (i.e., $u$ has not visited location $l$), the social rating of $u$ to unvisited location $l$, denoted as $\hat{r}^{soc}_{u,l}$, can be estimated:

$$\hat{r}^{soc}_{u,l} = \frac{\sum_{u' \in U \wedge u' \neq u} SocSim(u,u') \cdot r_{u',l}}{\sum_{u' \in U \wedge u' \neq u} SocSim(u,u')}, \qquad (8)$$

where $SocSim(u,u')$ in Equation 8 is the similarity measure between users $u$ and $u'$ derived from their social links and residence distance:

$$SocSim(u,u') = \begin{cases} 1 - \frac{distance(u,u')}{\max_{u'' \in F(u)} distance(u,u'')}, & u' \in F(u); \\ 0, & u' \notin F(u); \end{cases} \qquad (9)$$

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 $u'$.

## 4.2 Using Geographical Influence

In LBSNs, POIs are distinct from other non-spatial items, such as books, music and movies in conventional recommendation systems, because physical interactions are required for users to visit locations. Thus, the geographical information (i.e., latitude and longitude coordinates) of locations plays a significant role on influencing users' check-in behaviors. For example, users tend to visit locations close to their homes or offices and also may be interested in exploring the nearby places of their visited locations.

To this end, we use the geographic information of locations visited by a user to derive a probability of the user visiting a new location. Specifically, we model a *two-dimensional check-in probability distribution* over the latitude and longitude coordinates for each user instead of estimating the *one-dimensional distance probability distribution* in existing works, e.g., power-law distribution [11, 15, 16, 21, 24, 27] or nonparametric distance distribution based on the kernel density estimation [28, 30]. The two-dimensional check-in probability distribution is more reasonable and intuitive than the one-dimensional distance distribution. The main reason is that: It is hard to accurately compute a visiting probability for a new location based on a distance probability distribution, since it is intractable to find a reference location to derive a reasonable distance for the new location to the reference location in the first place. Conversely, it is considerably intuitive to employ a two-dimensional check-in probability distribution to compute a visiting probability for any location with latitude and longitude coordinates.

Formally, let $S_u$ be the set of locations visited by user $u$, i.e., $S_u = \langle l_1, l_2, \ldots, l_n \rangle$. Based on a general nonparametric distribution estimation technique, i.e., kernel density estimation [18] which does not assume a fixed distribution form in advance but instead learns the distribution form from historical data, the geographical probability $p^{geo}(l_{n+1}|S_u)$ of user $u$ to new location $l_{n+1}$ is given by:

$$p^{geo}(l_{n+1}|S_u) = \frac{1}{n\sigma^2} \sum_{i=1}^{n} K\left(\frac{l_{n+1} - l_i}{\sigma}\right), \qquad (10)$$

where each location $l_i = (lat_i, lon_i)^T$ is a two-dimensional column vector with the latitude ($lat_i$) and longitude ($lon_i$) coordinates, $K(\cdot)$ is the kernel function and $\sigma$ is a smoothing parameter, called the bandwidth. In this paper we apply the widely used standard two-dimensional normal kernel [18]

$$K(\mathbf{x}) = \frac{1}{2\pi} \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{x}), \qquad (11)$$

and the optimal bandwidth [18]

$$\sigma = n^{-\frac{1}{6}} \sqrt{\frac{1}{2}\hat{\sigma}^T\hat{\sigma}}, \qquad (12)$$

where $\hat{\sigma}$ is the marginal standard deviation vector of the latitude and longitude values in $S_u$, given by

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (l_i - \hat{\mu})^2} \qquad (13)$$

together with

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} l_i. \qquad (14)$$

## 4.3 Integrating Influences

As shown in the previous works [4, 28, 30], the product fusion rule is effective to integrate different factors, e.g., social and geographical influences, in which it is no need to normalize each factor since the normalization cannot affect the ranking of results. Here, we apply the product fusion rule to combine $p^{seq}(l_{n+1}|S_u)$ derived from *sequential influence* (Equation 7) with $\hat{r}^{soc}_{u,l_{n+1}}$ from *social influence* (Equation 8) and $p^{geo}(l_{n+1}|S_u)$ from *geographical influence* (Equation 10), given by

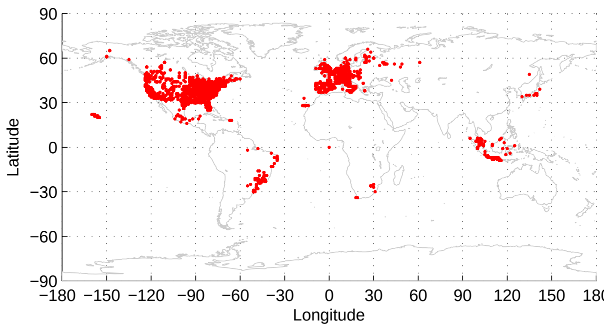
$$\hat{s}_{u,l_{n+1}} = p^{seq}(l_{n+1}|S_u) \cdot \hat{r}^{soc}_{u,l_{n+1}} \cdot p^{geo}(l_{n+1}|S_u), \qquad (15)$$

where $\hat{s}_{u,l_{n+1}}$ is the unified score of user $u$ to location $l_{n+1}$ embodying sequential, social and geographical influences. By this end, we can recommend locations for $u$ by returning her top-$k$ new locations $l_{n+1}$ with the highest unified score.

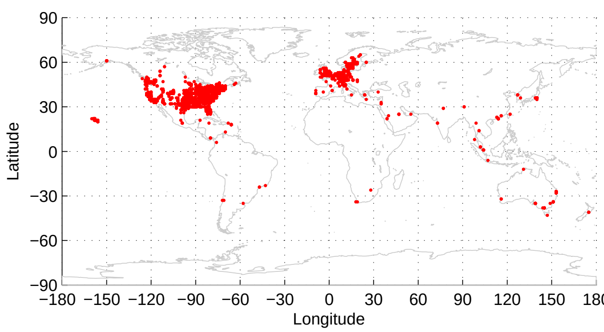
# 5. EXPERIMENTAL EVALUATION

Table 1: 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 | $2.3 \times 10^{-4}$ | $2.9 \times 10^{-5}$ |
| Avg. No. of visited POIs per user | 42.44 | 37.18 |
| Avg. No. of check-ins per location | 2.63 | 3.11 |




(a) Foursquare




(b) Gowalla

**Figure 3: Distribution of locations on a world map**

In this section, we describe our experiment settings for evaluating the performance of LORE against the state-of-the-art location recommendation techniques including the first-order Markov chain [2, 3, 5, 12, 36], geographical-social recommendation methods [1, 4, 8, 11, 15, 22, 24, 28] and their combination.

## 5.1 Two Real Data Sets

We use two publicly available large-scale real check-in data sets[1] that were crawled from Foursquare [8] between Jan. 2011 and Jul. 2011 and Gowalla [6] between Feb. 2009 and Oct. 2010. The statistics of the data sets are shown in Table 1. Figure 3 depicts the distribution of the locations in the data sets on a world map.

## 5.2 Evaluated Recommendation Techniques

[1]The large-scale real check-in data sets used for our experiments can be downloaded from `http://www.public.asu.edu/~hgao16/Publications.html` and `http://snap.stanford.edu/data/loc-gowalla.html`.

The recommendation techniques implemented in our experiments are classified into three categories based on the information they use in the location recommendations.

**Sequential category:**

- First-order Markov Chain (FMC): The existing FMC utilizes the sequential influence by deriving the sequential probability $p^{seq}(l_{n+1}|S_u)$ of user $u$ to new location $l_{n+1}$ based on **only the latest visited location** $l_n$ in the sequence $S_u$ (Equation 3). This technique is widely used in existing works [2, 3, 5, 12, 36].
- $n$th-order Additive Markov Chain (AMC): The proposed AMC exploits the sequential influence through deducing the sequential probability $p^{seq}(l_{n+1}|S_u)$ of user $u$ to location $l_{n+1}$ based on **all the visited locations** $l_1, l_2, \ldots, l_n$ in the sequence $S_u$ (Equation 7).

**Geographical-social category:**

- iGSLR: This existing technique employs the social influence based on Equation 8 and the geographical influence by estimating a **one-dimensional distance probability distribution** for each user. There are a variety of geographical-social location recommendation techniques [1, 4, 8, 11, 15, 22, 24, 28], but iGSLR shows the best performance in our previous work [28]. Thus, iGSLR is selected as a baseline in this category.
- GS2D: The proposed GS2D uses the social influence in terms of Equation 8 and the geographical influence by modeling a **two-dimensional check-in probability distribution** over the latitude and longitude coordinates for each user as shown in Equation 10.

Note that both iGSLR and GS2D also adopt the product fusion rule to combine social and geographical influences as done in the literatures [4, 28, 30].

**Sequential+geographical-social category:**

- First-order Markov Chain + GS2D (FMC+GS2D): To recommend POIs for users, FMC+GS2D integrates the *sequential influence* (**Equation 3**) with the *social influence* (Equation 8) and *geographical influence* (Equation 10) through the product fusion rule (Equation 15).
- Additive Markov Chain + GS2D (LORE): Our LORE combines the *sequential influence* (**Equation 7**) with the *social influence* (Equation 8) and *geographical influence* (Equation 10) through the product fusion rule (Equation 15).

## 5.3 Performance Metrics

**Recommendation accuracy.** In general, recommendation techniques compute a score for each candidate item (i.e., a location or POI in this paper) regarding a target user and return POIs 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 actually visited by the target user in the testing data set are discovered by the recommendation technique. For this purpose, we employ two standard metrics: *precision* and *recall*:

- Precision defines the ratio of the number of discovered POIs to the $k$ recommended POIs, i.e.,

$$\text{precision} = \frac{\text{number of discovered locations}}{k}.$$

- Recall defines the ratio of the number of discovered POIs to the number of **positive POIs**, which have been visited by the target user in the testing set, i.e.,

$$\text{recall} = \frac{\text{number of discovered locations}}{\text{number of positive locations}}.$$

## 5.4 Experiment Settings

In the experiments, each data set is divided 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; the training set is used to learn the recommendation models of the evaluated techniques described in Section 5.2 to predict the testing data. Unless otherwise specified, the threshold $\Delta T$ in Definition 2 is set to one day and the decay rate parameter $\alpha$ in Equation 5 is set to 0.05.

In our experiments, we examine the precision and recall of evaluated recommendation techniques with respect to a large range of top-$k$ from 2 to 50 and a large range of the number of check-in locations in the training set, i.e., the **given-$n$** locations in $S_u$ from 2 to 50. Further, we also investigate the effect of varying the threshold $\Delta T$ and decay rate parameter $\alpha$ on recommendation quality.

# 6. EXPERIMENTAL RESULTS

This section analyzes the extensive experimental results. We first compare our LORE against the state-of-the-art location recommendation techniques in terms of the recommendation accuracy in Section 6.1. We then discuss some important findings in Section 6.2. Finally, we investigate the effect of the threshold $\Delta T$ and decay rate parameter $\alpha$ on the recommendation quality of LORE in Sections 6.3 and 6.4, respectively.

## 6.1 Comparison of Performance

In this section, we present the performance comparison among the evaluated location recommendation techniques with regard to a large range of top-$k$ values in Figure 4 and given-$n$ values in Figure 5. We conclude **the most important and general findings** in the experimental results on the two large-scale real data sets collected from Foursquare and Gowalla as follows.

**Sequential category:**

1) FMC [2, 3, 5, 12, 36]: FMC simply considers the sequential influence by utilizing the latest visited location in the check-in sequence of a user to derive her visiting probability to new POIs. As a result, it often cannot take full advantage of sequential patterns in location recommendations, since it ignores the impact of the earlier visited locations in the sequence on the new likely visiting POIs. Thus, FMC returns the most inaccurate POIs in terms of precision and misses most POIs actually visited by target users in terms of recall, as depicted in Figures 4 and 5.

2) AMC: To overcome the limitation of FMC, this paper proposes AMC that exploits the sequential influence through deducing the sequential probability of a user to new POIs based on all her visited locations and leaning the sequential influence towards recently visited locations. Accordingly, AMC significantly improves the precision and recall with an increment of around 100% in comparison to FMC on the two data sets, as depicted in Figures 4 and 5. These results verify the superiority of exploiting the influence of the whole location sequence for location recommendations proposed in this paper over only considering the latest visited location in the influence adopted by the current works [2, 3, 5, 12, 36].

**Geographical-social category:**

3) iGSLR [28]: As a representative geographical-social recommendation approach, iGSLR employs the geographical influence through modeling a one-dimensional distance distribution for each user and shows good recommendation quality in terms of precision and recall which also has been observed in our previous experimental results [28].

4) GS2D: However, this paper develops a new geographical-social recommendation method GS2D that leverages the geographical influence via modeling a two-dimensional check-in distribution over the latitude and longitude coordinates for each user and exhibits the superior performance to iGSLR. Our explanation is that: the two-dimensional check-in distribution is more intuitive and sophisticated than the one-dimensional distance distribution, since it avoids the difficulty of finding a reference location to derive a reasonable distance for a new location so as to compute the visiting probability for the new location based on the distance probability distribution.

**Sequential+geographical-social category:**

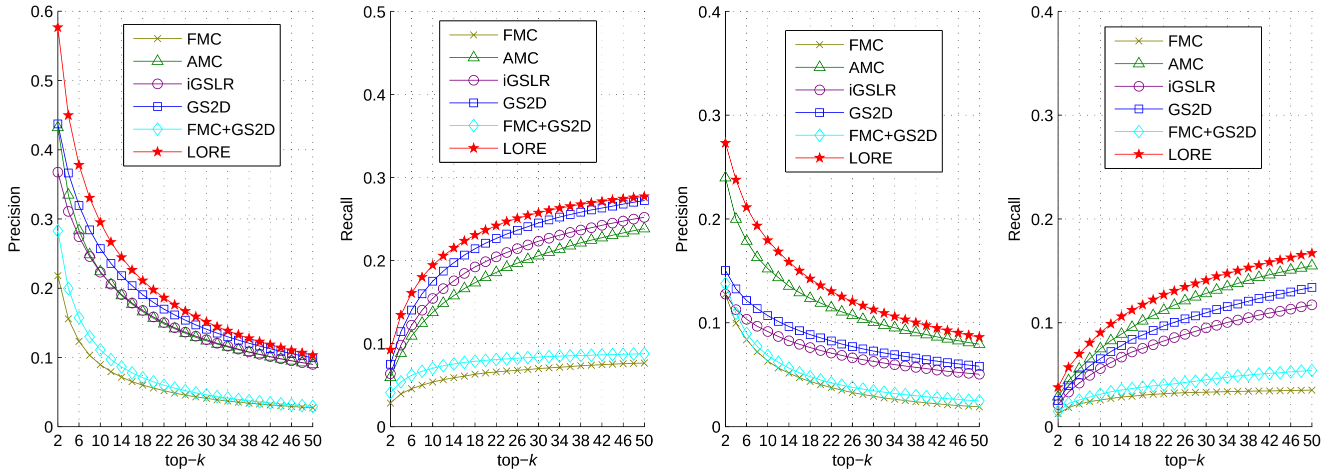
5) FMC+GS2D: FMC+GS2D shows a little better performance than FMC but much worse performance than GS2D. The reason is that FMC+GS2D suffers from the limitation of FMC that only uses the latest visited location to derive the sequential probability of users to POIs.

6) LORE (i.e., AMC+GS2D): In contrast, LORE inherits the superiorities of both AMC and GS2D including the $n$th-order additive Markov chain and two-dimensional check-in probability distribution. Thus, it generally exhibits the best performance regarding various top-$k$ and given-$n$ values on both the Foursquare and Gowalla data sets.

## 6.2 Discussion on Performance

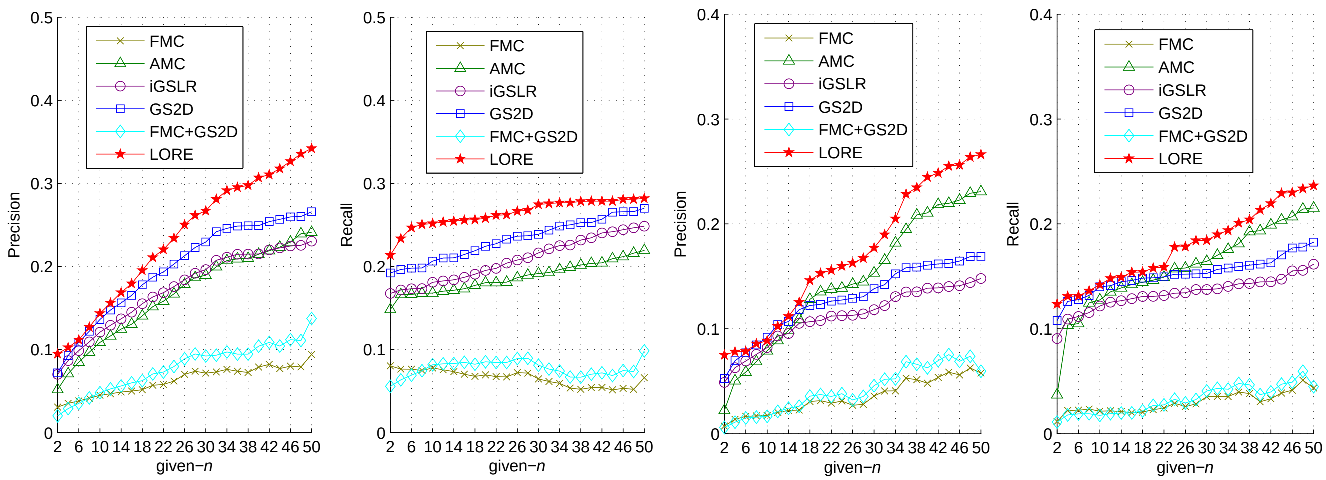
**Which influence (sequential vs. geographical-social) is better?** According to Figures 4 and 5, AMC with the sequential influence is inferior to GS2D with the geographical-social influence on the Foursquare data set. However, on the Gowalla data set with one order-of-magnitude lower density than that of the Foursquare data set, AMC outperforms GS2D, which shows that AMC with the sequential influence can better deal with the data sparsity problem. The main reason is that: AMC mines sequential patterns from all users' check-in location sequences as a common location-location transition graph ($\text{L}^2\text{TG}$) whereas GS2D separately models a two-dimensional check-in distribution for each user using her own check-in locations. Fortunately, both sequential and geographical-social influences are not conflicting to each other and can be integrated into a unified recommendation framework.

**When is the integration of influences helpful?** In terms of Figures 4 and 5, in most cases LORE is significantly superior to its components, i.e., AMC and GS2D; in other




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

**Figure 4: Recommendation accuracy with respect to top-$k$ values**




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

**Figure 5: Recommendation accuracy with respect to given-$n$ values**

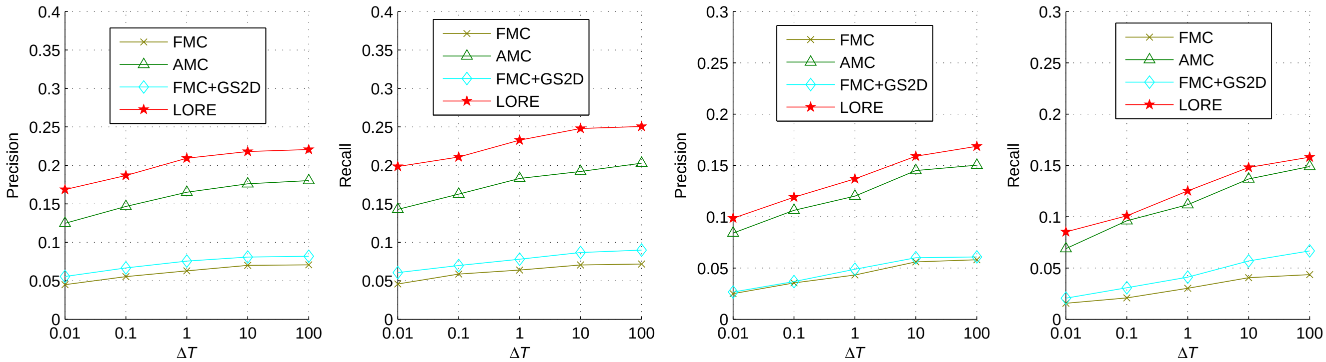
words, the integration of sequential and geographical-social influences is usually helpful for enhancing recommendation quality. On the contrary, for users who have checked in few locations, i.e., $n$<10 in Figure 5, LORE is a little better than its components AMC and GS2D, i.e., the benefit from the integration is very limited. The underlying reason is that: the individual components of LORE cannot accurately estimate the preference score on new locations for the users with a few check-in locations due to the data sparsity problem.

**Recommendation accuracy on various top-$k$ values.** From Figure 4, with the increase of $k$ the recall gradually gets higher but the precision steadily becomes lower on the two data sets. Our explanation is that, in general, by returning more locations for users, it is always able to discover more locations that users would like to visit. However, the extra recommended locations are less possible to be liked by users due to the lower visiting probabilities of these locations, since the recommendation techniques return the locations with the top-$k$ highest scores.

**Recommendation accuracy on various given-$n$ values.** In Figure 5, a measure at "given-$n$ = 2" is averaged on only the users who have checked in two locations in the training data set. From Figure 5, when users check in more locations, the performance of various recommendation techniques generally inclines, because they can estimate the sequential or/and geographical visiting probability of users to POIs more accurately through using more check-in data. For example, in general the longer sequence generally benefits for removing the uncertainty from obtained sequential patterns.

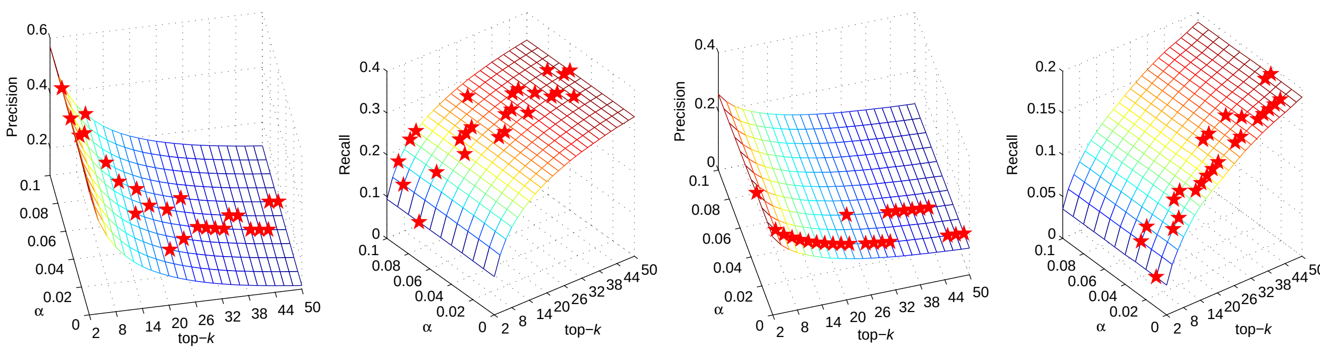
## 6.3 Effect of Threshold $\Delta T$

Figure 6 depicts the effect of the time interval $\Delta T$ for determining transitions in DEFINITION 2 on the recommendation accuracy of the sequential recommendation techniques including FMC, AMC, FMC+GS2D and LORE. Note that $\Delta T$ does not affect the performance of the geographical-social recommendation techniques like iGSLR and GS2D. As shown before, LORE always outperforms FMC, AMC and FMC+GS2D according to Figure 6. Moreover, when the value of $\Delta T$ changes from 0.01 to 100 days, the overall precision and recall of LORE dramatically increase at first and then




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

**Figure 6: Effect of $\Delta T$ on recommendation accuracy**




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

**Figure 7: Effect of $\alpha$ on recommendation accuracy of LORE**

remain steady. The reason is that there are more transitions when $\Delta T$ becomes larger according to DEFINITION 2, and then the number of transitions maintains the same when $\Delta T$ is larger than the maximum time interval between two consecutive locations in check-in sequences of users. Based on this finding, we should not split the location sequence of users into several parts with a small threshold $\Delta T$, because some users do not travel frequently and they plan trips in more high-level granularity. Hence, the accuracy of LORE has the potential to be uplifted using a larger value than the default value in the experiments.

## 6.4 Effect of Decay Rate Parameter $\alpha \geq 0$

Figure 7 depicts the effect of the decay rate parameter $\alpha$ on the precision and recall of LORE, in which the optimal value of $\alpha$ is marked with a red star for each value of $k$. When $\alpha = 0$ (i.e., no decay), the best performance cannot be achieved. Thus, it is necessary to decay the sequential influence of check-in locations on new possibly visiting locations. The optimal value of $\alpha$ should be small, always lying between 0.01 and 0.1, because the sequence decay weight $2^{-\alpha\cdot(n-i)}$ in Equation 5 exponentially decreases with the decrease of $i$, where $i$ denotes the order of locations visited by a user. The performance of LORE is stable when $\alpha$ lies in the optimal range [0.01, 0.1]. This important feature makes it possible to choose a default value for $\alpha$ instead of finding the optimal value that usually costs much effort and suffers from over-fitting. As in our experiments, the default value of $\alpha$ is set to 0.05. In general, the optimal value of $\alpha$ for the Foursquare data set is larger than that for the Gowalla data set. The reason is that the Foursquare data set is denser than the Gowalla data set as shown in Table 1.

# 7. CONCLUSION AND FUTURE WORK

In this paper, we have explored the sequential influence on users' check-in behaviors in LBSNs. We have proposed LORE to incrementally mine sequential patterns from a check-in location stream of all users as a dynamic $\mathrm{L^2TG}$ and derive a probability of a user visiting a new location based on the $n$th-order additive Markov chain (AMC). Unlike the first-order Markov chain, AMC does not assume that a new visiting location only depends on the latest visited location and hence can employ the sequential influence more comprehensively and sophisticatedly. Furthermore, we have integrated the *sequential influence* with the *geographical influence* and *social influence* into a unified location recommendation framework, in which the geographical influence is modeled as two-dimensional check-in distributions on locations instead of using one-dimensional distance distributions. Finally, we have conducted extensive experiments to evaluate the performance of LORE using two large-scale real data sets collected from Foursquare and Gowalla. Experimental results show that LORE provides much better location recommendations than all other recommendation techniques evaluated in our experiments.

We have two directions for future study: (1) how to recommend a trip of POIs and (2) how to take temporal influence into account to capture the change of users' preferences.

## 8. 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).

## 9. REFERENCES


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

[2] Z. Chen, H. T. Shen, and X. Zhou. Discovering popular routes from trajectories. In *IEEE ICDE*, 2011.

[3] 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.

[4] 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.

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

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

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

[8] 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.

[9] M. C. González, C. A. Hidalgo, and A.-L. Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

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

[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] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura. Travel route recommendation using geotags in photo sharing sites. In *ACM CIKM*, 2010.

[13] 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.

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

[15] B. Liu, Y. Fu, Z. Yao, and H. Xiong. Learning geographical preferences for point-of-interest recommendation. In *ACM KDD*, 2013.

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

[17] 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.

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

[19] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, 2010.

[20] O. V. Usatenko, S. S. Apostolov, Z. A. Mayzelis, and S. S. Melnik. *Random finite-valued dynamical systems: Additive Markov chain approach*. Cambridge Scientific Publishers, Cambridge, 2009.

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

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

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

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

[25] H. Yin, Y. Sun, B. Cui, Z. Hu, and L. Chen. LCARS: A location-content-aware recommender system. In *ACM KDD*, 2013.

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

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

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

[29] J.-D. Zhang and C.-Y. Chow. CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations. *Information Sciences, accepted to appear*, 2014.

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

[31] J.-D. Zhang, G. Ghinita, and C.-Y. Chow. Differentially private location recommendations in geosocial networks. In *IEEE MDM*, 2014.

[32] J.-D. Zhang, J. Xu, and S. S. Liao. Aggregating and sampling methods for processing GPS data streams for traffic state estimation. *IEEE TITS*, 14(4):1629–1641, 2013.

[33] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Towards mobile intelligence: Learning from GPS history data for collaborative recommendation. *Artificial Intelligence*, 184-185:17–37, 2012.

[34] Y. Zheng. Location-based social networks: Users. In *Computing with Spatial Trajectories*. Springer, 2011.

[35] Y. Zheng and X. Xie. Learning location correlation from GPS trajectories. In *IEEE MDM*, 2010.

[36] Y.-T. Zheng, Z.-J. Zha, and T.-S. Chua. Mining travel patterns from geotagged photos. *ACM TIST*, 3(3):56:1–56:18, 2012.