Spatiotemporal Sequential Influence Modeling for Location Recommendations: A Gravity-based Approach

Jia-Dong Zhang, City University of Hong Kong
Chi-Yin Chow, City University of Hong Kong

Recommending users with personalized locations is an important feature of location-based social networks (LBSNs), which benefits users to explore new places and businesses to discover potential customers. In LBSNs, social and geographical influences have been intensively used in location recommendations. However, human movement also exhibits spatiotemporal sequential patterns, but only few current studies consider spatiotemporal sequential influence of locations on users' check-in behaviors. In this paper, we propose a new gravity model for location recommendations called LORE to exploit the spatiotemporal sequential influence on location recommendations. First, LORE extracts sequential patterns from historical check-in location sequences of all users as a Location-Location Transition Graph (L²TG), and utilizes the L²TG to predict the probability of a user visiting a new location through the developed additive Markov chain that considers the effect of all visited locations in the check-in history of the user on the new location. Further, LORE applies our contrived gravity model to weigh the effect of each visited location on the new location derived from the personalized attractive force (i.e., the weight) between the visited location and the new location. The gravity model effectively integrates the spatiotemporal, social, and popularity influences by estimating a power-law distribution based on (1) the spatial distance and temporal difference between two consecutive check-in locations of the same user, (2) the check-in frequency of social friends, and (3) the popularity of locations from all users, respectively. Finally, we conduct a comprehensive performance evaluation for LORE using three large-scale real-world data sets collected from Foursquare, Gowalla, and Brightkite. Experimental results show that LORE achieves significantly superior location recommendations compared to other state-of-the-art location recommendation techniques.

Categories and Subject Descriptors: H.2.8 [Database Management]: Database Applications—Data mining; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—Information filtering

General Terms: Design, Algorithms, Experimentation, Performance

Additional Key Words and Phrases: Location recommendation, spatiotemporal sequential influence, social influence, popularity influence, additive Markov chain, gravity model

ACM Reference Format:
DOI:http://dx.doi.org/10.1145/0000000.0000000

1. INTRODUCTION

With the rapid advancement of mobile devices and location acquisition technologies, location-based social networks (LBSNs), such as Foursquare, Gowalla and Brightkite,
have attracted millions of users. In an LBSN (Figure 1), users can establish social links with each other and share their experiences of visiting some specific locations, also known as points-of-interest (POIs), e.g., restaurants, stores, and museums, through performing check-in operations to these POIs in the LBSN via their handheld device. LBSNs generate plenty of community-contributed data including the social links between users and spatiotemporal check-in sequences of users to POIs which reflect the users’ preferences on the POIs. In the LBSNs, it is crucial to recommend personalized POIs to users based on their preferences learned from the community-contributed data, which benefit for users to know new POIs and discover a city while for businesses to deliver advertisements to targeted users.

Most existing methods (e.g., [Bao et al. 2012; Gao et al. 2012; Lian et al. 2014; Liu et al. 2014c; Wang et al. 2013; Ye et al. 2011; Yin et al. 2014; Ying et al. 2014; Zhang and Chow 2013; Zhang et al. 2014a; Zhang and Chow 2015a; Zhao et al. 2014]) make location recommendations for users by employing the memory- or model-based collaborative filtering (CF) techniques. Besides using the historical user-POI check-in frequencies, these CF-based methods also derive the user similarity from the social links between users and/or estimate the user check-in distribution over POIs from the geographical or spatial information of POIs, in terms of the two facts that the social friends are more likely to share common interests on POIs and the geographical proximity of POIs significantly influences the check-in behaviors of users. For example, friends often go to some POIs like movie theaters or restaurants together, and users usually visit POIs close to their homes, offices or visited locations. However, these CF-based methods do not consider the influence of spatiotemporal sequential patterns of check-in POIs on the check-in behaviors of users, called spatiotemporal sequential influence hereafter, or sequential influence for short. In reality, human movement exhibits spatiotemporal sequential patterns (Cho et al. 2011; Gómez-Bombarelli et al. 2008; Song et al. 2010). The sequential patterns may associate with the time of a day (e.g., people usually visit museums or libraries at daytime, go to restaurants for dinner in the evening, and then relax in cinemas or bars at night), the geographical proximity of POIs (e.g., tourists often orderly visit London Eye, Big Ben, Downing Street, Horse Guards, and Trafalgar Square [Yin et al. 2011]), the place nature and human preference (e.g., checking in stadiums first and then restaurants is better than the reverse way because it is not healthy to exercise right after a meal [Hsieh et al. 2014]).

To observe the sequential patterns in depth, we conduct analysis on three publicly available real data sets collected from popular LBSNs: Foursquare [Gao et al. 2012], Gowalla and Brightkite [Cho et al. 2011]. As an example, here we focus on the two-gram sequential patterns. Specifically, we randomly select a POI from a data set and calculate the probability or percentage of each of next POIs that are immediately visited by users after they visited the selected POI in the data set. Figure 2 depicts the
cumulative distribution of the next visited POIs for three selected POIs in each data set, in which the next visited POIs are sorted in descending order based on their probabilities. As depicted in Figure 2, a POI mainly transits to a certain set of POIs, i.e., each selected POI transits to the top hundred POIs out of several hundred thousand POIs with the probability of larger than 0.5. Thus, the distribution of a POI to another POI is not uniform and it implicitly incorporates some sequential patterns. Further, for each selected POI in Figure 2, we calculate its distance to the next visited POIs and obtain the rank of these next visited POIs in ascending order based on their distances to the selected preceding POI. Figure 3 shows the rank of distance closeness (the smaller, the closer to the selected POI) of the next visited POIs with the same order in the horizontal axis as in Figure 2. In term of Figure 3, some selected POIs are close to their top-100 next visited POIs, while other selected POIs are not. Importantly, there is no simple proportional relation between the distance closeness rank and the transition probability of the top-100 next visited POIs, and the rank of the top-100 next visited POIs is usually higher than 100. The reason is that the sequential patterns are resulting from comprehensive effects of different factors, e.g., time of a day, geographical proximity, place nature, and human preference.

Such sequential patterns become increasingly important in location recommendations. Current works [Chen et al. 2011; Cheng et al. 2011; Cheng et al. 2013; Kurashima et al. 2010; Zheng et al. 2012] extract the sequential patterns from check-in location sequences of users and use them to determine the probability of a user visiting a new POI given her historical check-in POI sequence. Unfortunately, these works suffer from three major limitations. (1) **First-order sequential influence.** They utilize the sequential influence based on the first-order Markov chain for efficiency that only considers the latest visited location of a user to estimate the probability of the user
visiting a new location. Nevertheless, in reality such an estimated visiting probability depends on not only the latest visited location but also the earlier visited locations of the user. (2) **Lack of spatiotemporal influences.** They often ignore the spatiotemporal influences embodied in the check-in POI sequences. For example, in practice the distance between consecutive check-in POIs of a user is usually not far and the recently visited POIs have stronger influence on the new location than the relatively old check-in POIs. (3) **No integration with the social influence and the popularity of POIs.** They do not integrate the sequential influence with the social influence and the popularity of POIs. Nonetheless, in practice people often go to popular places or POIs recommended by their friends.

In this paper, we are therefore motivated to propose a gravity-model-based LOCation REcommender system by exploiting the spatiotemporal sequential influence in order to tackle the three limitations in the current works, called LORE. (1) **Higher-order sequential influence.** In LORE, we first mine sequential patterns from check-in location sequences of all users as a Location-Location Transition Graph (L²TG), and then develop an efficient n-th-order additive Markov chain to predict the sequential probability of a user visiting a new location based on L²TG. The additive Markov chain considers all visited locations in the check-in history of the user to determine her visiting probability on the new location, instead of only using her latest visited location adopted by the first-order Markov chain in the existing works. (2) **Gravity model for weighing the sequential influence by leveraging the spatiotemporal, social and popularity influences.** The effect of each visited location on the new location is weighed based on our contrived gravity model that effectively integrates the spatiotemporal, social and popularity influences to determine the attractive force (i.e., the weight) between the visited location and the new location. Gravity models have been widely used in various spatial interactions, e.g., traffic flows, migration flows and trade flows [Liu et al. 2014a]. All these gravity models are derived from the well-known Newton’s law of universal gravitation which states that: “Any two bodies in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.” In LORE, we devise a new gravity model to describe the personalized mimic gravity interaction between two bodies, i.e., a visited POI of a user and a new POI for the user. Specifically, (a) we identify the mass of the visited POI of a user with the personalized check-in frequency of the user on the POI. (b) We compare the mass of the new POI for a user to the check-in frequency from her social friends and the popularity from all users on the new POI, and deduce its directly proportional relation with the attractive force in terms of the power-law distributions of the social check-in frequency and popularity of all POIs. (c) We compare the physical distance between two bodies to the spatiotemporal distance of the two POIs, and determine its inversely proportional relation with the attractive force based on the power-law distributions of the spatial distance and temporal difference of two consecutive check-in POIs of the same user. (d) All the distributions of social check-in frequency, popularity, spatial distance, and temporal difference are learned from historical check-in data through maximum likelihood estimation.

This study is a significant extension to our previous work [Zhang et al. 2014b] by proposing a new gravity-based approach for location recommender systems. Its main contributions can be summarized as follows.

— In this study, we extend the additive Markov chain developed in our previous work [Zhang et al. 2014b] through a gravity model in order to exploit the spatiotemporal higher-order sequential influence. To the best of our knowledge, this is the first
We propose a new gravity model that effectively integrates the spatiotemporal, social, and popularity influences to find the personalized attractive force between a visited location of a user and a new location for the user as the weight of the visited location affecting the new location. Moreover, the spatiotemporal, social, and popularity influences are modeled as a distribution from the historical check-in data, respectively. (Section 4)

We conduct extensive experiments to evaluate the performance of LORE using three large-scale real-world data sets collected from Foursquare, Gowalla, and Brightkite. Experimental results show that LORE outperforms other state-of-the-art location recommendation techniques in terms of recommendation accuracy. (Section 5)

The remainder of this paper is organized as follows. Section 2 defines the research problem and introduces the overview of LORE. We extend the additive Markov chain in Section 3 and propose the gravity model in Section 4. In Section 5, we present our experiment settings and analyze the performance of LORE. Section 6 highlights related work. Finally, we conclude this paper in Section 7.

2. PRELIMINARIES AND OVERVIEW

Table I summarizes the key symbols used in this paper. In this section, we present preliminaries, the problem definition, and the overview of LORE.

**Definition 2.1 (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.

**Definition 2.2 (Check-in collection).** A check-in collection is a set of check-ins of all users visiting all locations in an LBSN, denoted as \( D = \{ \langle u, l, t \rangle \} \), in which \( |D| \) represents the number of check-ins in \( D \), the same hereafter.

**Definition 2.3 (Spatiotemporal check-in sequence).** From the check-in collection \( D \), we can extract the spatiotemporal check-in location sequence for each user \( u \), denoted by \( L_u = \{ \langle l_1, t_1 \rangle \rightarrow \langle l_2, t_2 \rangle \rightarrow \cdots \rightarrow \langle l_n, t_n \rangle \} \) such that user \( u \) goes through location \( l_i \) at time \( t_i \), in which \( t_1 \leq t_2 \leq \cdots \leq t_n \) and \( l_i \) is associated with a pair of latitude and longitude coordinates. We also use \( L_u = \{ l_1 \rightarrow l_2 \rightarrow \cdots \rightarrow l_n \} \) for short when the time information is not involved in the context.
Definition 2.4 (Transition, predecessor and successor). Given a check-in location sequence of user \( u \), \( L_u = \{l_1 \rightarrow l_2 \rightarrow \cdots \rightarrow l_n\} \), each two-gram subsequence consisting of two consecutive check-in locations \( l_i \rightarrow l_{i+1} \) is also called a \textit{transition} representing user \( u \) visiting \( l_i \) immediately before \( l_{i+1} \), where \( l_i \) is a transition \textit{predecessor} of \( l_{i+1} \) and \( l_{i+1} \) is a transition \textit{successor} of \( l_i \).

Definition 2.5 (Check-in frequency matrix). Given the check-in collection \( D \), we can build a check-in frequency matrix \( R_{\{U\} \times \{L\}} \), in which each entry \( R_{u,l} \) represents the frequency of user \( u \in U \) visiting location \( l \in L \) in \( D \), i.e., \( R_{u,l} = |\{u_i, l_i\} \in D \land u_i = u \land l_i = l| \). Note that most entries in \( R \) are zero, since users have only visited a very small proportion of locations in an LBSN.

Definition 2.6 (Social link matrix). In a social link matrix \( S_{\{L\} \times \{U\}} \), if there exists a social link between two different users \( u, u' \in U \), \( S_{u,u'} = 1 \); otherwise, \( S_{u,u'} = 0 \).

Problem definition. Given a social link matrix \( S \) and a check-in collection \( D \) that can be used to determine the check-in frequency matrix \( R \) and the spatiotemporal check-in location sequence \( L_u \) for each user \( u \), the goal is to predict the probability \( \Pr(l|L_u) \) of any user \( u \) visiting a \textit{new} location \( l \in L \land l \notin L_u \) after \( L_u \), and then return the top-\( k \) locations with the highest visiting probability \( \Pr(l|L_u) \) for \( u \).

The overview of LORE. Figure 4 demonstrates the overview of LORE, including three major parts: \textit{sequential pattern mining}, \textit{additive Markov chain}, and \textit{gravity model}. (1) \textit{Sequential pattern mining}. The sequential patterns are incrementally mined from check-in location sequences of all users and represented as a dynamic location-location transition graph \((L^2TG)\). The \( L^2TG \) incorporates not only \textit{transition counts} between locations but also \textit{outgoing counts} of locations to other locations in order to incrementally update the obtained sequential patterns. (2) \textit{Additive Markov chain}. We devise an efficient \( n \)-th-order additive Markov chain that takes into account all visited locations in the check-in history of a user to predict her visiting probability on a new location based on \( L^2TG \). The main reason is that the new location may rely on the latest visited location and the earlier visited locations in her check-in history. (3) \textit{Gravity model}. To further enhance the predictive ability of our additive Markov chain, we propose a gravity model to deduce the personalized attractive force between a visited location of a user and a new location for the user so as to weigh the effect of the visited location on the new location. The gravity model effectively integrates the spatiotemporal, social and popularity influences by two fundamental quantities: (i) The \textit{distance between the visited and new locations} is derived from their \textit{temporal d-
**L2TG** and **spatial distance**, since the effect of the visited location on the new location is related to not only their spatial distance but also their check-in time difference. That is, the visited locations with more recently check-in time and shorter distance to the new location have stronger influence on the preference of the user to the new location.

(ii) The mass of the visited location is determined based on the check-in frequency of the user to the visited location, while the mass of the new location is derived from the check-in frequency of all users to the new location as location popularity and the check-in frequency of the user’s friends to the new location as social popularity, because the check-in frequency or popularity of locations naturally simulate their masses.

### 3. Modeling Spatiotemporal Sequential Influence

In this section, we extract the sequential patterns from check-in location sequences of all users as a location-location transition graph in Section 3.1 and develop the nth-order additive Markov chain to predict the probability of a user visiting a location using sequential patterns in Section 3.2.

#### 3.1. Sequential Pattern Mining

In Markov chain models, sequential patterns are usually represented as a matrix of transition probabilities from one location or sequence to another location. To implement the incremental update of sequential patterns or transition probabilities, we define a concise and dynamic L2TG as follows.

**Definition 3.1 (L2TG).** A location-location transition graph (L2TG) G = (L, E) consists of a set of nodes L and a set of edges E ⊆ L × L. Each node li ∈ L represents a location associated with an outgoing count of li as a transition predecessor to other locations denoted by OCount(li). And each edge (li, lj) ∈ E represents a transition li → lj associated with a transition count denoted by TCount(li → lj).

The L2TG is only associated with the counts of transition between locations and outgoing of locations to other locations, and hence it can be easily updated for each newly arriving check-in ⟨ui, li, ti⟩ ∈ D in an incremental manner. Accordingly, the transition probabilities from one location to another location can be dynamically calculated through dividing transition counts by outgoing counts, as described in Definition 3.2.

**Definition 3.2 (Transition probability).** If OCount(li) > 0, the transition probability of li → lj, denoted as TP(li → lj), is calculated by

\[
TP(li → lj) = \frac{TCount(li → lj)}{OCount(li)}.
\]  

(1)

Otherwise, TP(li → lj) = 1 for lj = li and TP(li → lj) = 0 for lj ≠ li.

By Definition 3.2, if the outgoing count of li is non-zero, the transition probability of li → lj is defined as the proportion of TCount(li → lj) to OCount(li) in Equation (1), which is essentially the relative frequency definition of probability. On the other hand, if OCount(li) = 0 that means all users do not check in any other locations after li, accordingly we define the transition probability of li to itself is one for simplicity.

#### 3.2. Additive Markov Chain

Based on the obtained sequential patterns, this section focuses on predicting the sequential probability Pr(⟨L_u⟩) of a user u visiting a new location l given u’s visited location sequence Lu = {⟨t1, l1⟩ → ⟨t2, l2⟩ → · · · → ⟨tn, ln⟩}.

**The reasons of not using the first-order Markov chain or classical nth-order Markov chain.** For efficiency, most current works determine the sequential probabil-
LORE’s $n$th-order additive Markov chain. To this end in our recent work [Zhang et al. 2014b], we contrive an efficient $n$th-order additive Markov chain that significantly reduces the number of states from $O(|L|^{n+1})$ as in the classical $n$th-order Markov chain to $O(|L|^2)$. The original additive Markov chain estimates the sequential probability $\Pr(l|L_u)$ by summing the transition probabilities $TP(l_i \rightarrow l)(i = 1, 2, \ldots, n)$. In this study, we extend the additive Markov chain by a gravity model, given by

$$\Pr(l|L_u) \propto \sum_{i=1}^{n} TP(l_i \rightarrow l) \cdot Gravity((l_i, t_i); (l, t_c)), \forall l \in L,$$

(2)

where each transition probability $TP(l_i \rightarrow l)$ is weighed through the attractive force between visited location $l_i$ and new location $l$ of user $u$, given by our proposed gravity model $Gravity((l_i, t_i); (l, t_c))$ that combines the spatiotemporal, social and popularity influences, as detailed in Section 4.

4. THE GRAVITY MODEL

This section first presents our gravity model, including the gravity model framework (Section 4.1), distance derived from spatiotemporal influences (Section 4.2), and mass derived from popularity and social influences (Section 4.3). Then, we implement the gravity-based location recommender system in ALGORITHM 1 (Section 4.4) and demonstrate its applications in two typical scenarios (Section 4.5).

4.1. The Gravity Model Framework

Gravity models have been widely used in various spatial interactions (e.g., traffic, migration and trade flows) and are derived from the well-known Newton’s law of universal gravitation. Because there exist intrinsic spatial-temporal interactions between two POIs, in LORE we devise a new gravity model to describe the mimic gravity interaction between visited location $l_i$ and new location $l$ for user $u$, as depicted in Figure 5.
The framework of the new gravity model is defined by

$$\text{Gravity}(⟨l_i, t_i⟩, ⟨l, t_c⟩) \propto \frac{\text{Mass}(l_i) \cdot \text{Mass}(l)}{\text{Distance}(⟨l_i, t_i⟩, ⟨l, t_c⟩)},$$  \hspace{1cm} (3)

where we contain the elements of Mass and Distance lent to the metaphor of the well-known Newton’s law of universal gravitation, but they have different meanings, as specified in Sections 4.2 and 4.3.

Here we emphasize two important notes: (i) Personalized attractive force. The attractive force from the gravity model is personalized, i.e., it varies for the same pair of visited location l, and new location l of different users u. (ii) \textbf{Time \textit{t_c} setting for new location \textit{l}}. In Equation (3), it is required to estimate the time \textit{t_c} for new location \textit{l}, since user \textit{u} has not visited it yet. In LORE, we set it to the latest check-in time \textit{t_n} for each user instead of the current time for all users to improve personalization.

\textbf{4.2. Distance Derived from Spatiotemporal Influences}

As aforementioned, the check-in behaviors of users are significantly affected by the geographical proximity of POIs and the POIs with recent check-in time usually have stronger impact on a new possibly visiting location than the POIs with old check-in time. Thus, in Equation (3) the distance between visited location \textit{⟨l_i, t_i⟩} and new location \textit{⟨l, t_c⟩} is spatiotemporal, with two parts: temporal difference and spatial distance,

$$\frac{1}{\text{Distance}(⟨l_i, t_i⟩, ⟨l, t_c⟩)} = F_{\text{Tem}}(x_{t_i, t_c}) \cdot F_{\text{Spa}}(y_{1, l}),$$  \hspace{1cm} (4)

where \textit{x_{t_i, t_c}} is the temporal difference between \textit{t_i} and \textit{t_c} (note that \textit{t_c} is set to the latest check-in time \textit{t_n} of the underlying user \textit{u}), and \textit{y_{1, l}} is the spatial Haversine distance between \textit{l_i} and \textit{l} with latitude and longitude coordinates [Sinnott 1984].

In Equation (4), we should not naively regard the product of the temporal difference \textit{x_{t_i, t_c}} and spatial distance \textit{y_{1, l}} as the final distance \textit{Distance}(⟨l_i, t_i⟩, ⟨l, t_c⟩), because the two new locations with small temporal difference and spatial distance will dominate the recommendation result. More sophisticatedly, we transform the temporal difference and spatial distance into a normalized score via the \textbf{decreasing functions} \textit{F_{\text{Tem}}} and \textit{F_{\text{Spa}}} with respect to \textit{x_{t_i, t_c}} and \textit{y_{1, l}}, respectively. \textit{F_{\text{Tem}}} and \textit{F_{\text{Spa}}} are deduced from the distributions of temporal difference and spatial distance that are learned from the check-in data. We assume that the temporal difference or spatial distance between consecutive check-ins of users follows a power-law distribution; this assumption has been validated in three publicly available real data sets collected from popular LBSNs: Foursquare [Gao et al. 2012], Gowalla and Brightkite [Cho et al. 2011].

\textbf{4.2.1. Temporal Difference Distribution Estimation for \textit{F_{\text{Tem}}}}. To estimate the distribution of temporal difference, we collect a sample set \textit{X} from check-in collection \textit{D}, given by

$$X = \{t_{j+1} - t_j\},$$  \hspace{1cm} (7)

where \textit{t_j} and \textit{t_{j+1}} are any two consecutive check-in time instants of a user in \textit{D}. We have observed that the temporal difference \textit{x} of consecutive check-in time instants approximately follows a power-law distribution; the probability density is defined by

$$f_{\text{Tem}}(x) = (\alpha - 1)(x + 1)^{-\alpha}, x \geq 0, \alpha > 1,$$  \hspace{1cm} (8)
in which $\alpha$ is estimated from the sample $X$ through maximum likelihood estimation, given by

$$\alpha = 1 + |X| \left[ \sum_{x \in X} \ln(x + 1) \right]^{-1}.$$  \hfill (9)

Figure 6 shows that the temporal differences (i.e., the dots) in the three real-world data sets fit a certain power-law distribution (i.e., the line) quite well that is estimated through Equations (8) and (9). Thus, modeling the temporal difference as a power-law distribution is reasonable and effective.

Further, because the attractive force of a gravity model is decreasing with regard to the temporal difference and it is helpful to normalize the effect of temporal difference, $F_{Tem}$ is defined as the complementary cumulative distribution of $f_{Tem}$:

$$F_{Tem}(x_{t_i,t_c}) = \int_{x_{t_i,t_c}}^{+\infty} f_{Tem}(x) dx = (x_{t_i,t_c} + 1)^{1 - \alpha},$$  \hfill (10)

in which clearly $F_{Tem}$ is a decreasing function regarding the temporal difference $x_{t_i,t_c}$ of $t_i$ and $t_c$ because of $1 - \alpha < 0$. That is, the attractive force of the gravity model is decreasing on the temporal difference in terms of Equations (3) and (4).

4.2.2 Spatial Distance Distribution Estimation for $F_{Spa}$. As the distribution estimation for temporal difference, we can use the similar process to build the distribution of spatial distance. First, we collect a spatial distance sample $Y$ from check-in collection $D$:

$$Y = \{Haversine(l_j, l_{j+1})\},$$  \hfill (11)

where $l_j$ and $l_{j+1}$ are any two consecutive check-in locations of a user in $D$ and $Haversine$ is a spatial distance function for two points with longitude and latitude coordinates [Sinnott 1984]. Then, we assume that the spatial distance $y$ of consecutive check-in locations has a power-law distribution,

$$f_{Spa}(y) = (\gamma - 1)(y + 1)^{-\gamma}, y \geq 0, \gamma > 1,$$  \hfill (12)

in which $\gamma$ is estimated from the sample $Y$:

$$\gamma = 1 + |Y| \left[ \sum_{y \in Y} \ln(y + 1) \right]^{-1}.$$  \hfill (13)

Figure 7 also shows that the probability density of the spatial distances (i.e., the dots) in the three real-world data sets approximately follows an estimated power-law distribution (i.e., the line) based on Equations (12) and (13). Thus, the assumption made here is reasonable and feasible as well.
4.3. Mass Derived from Popularity and Social Influences

Social links of users and popularity of POIs greatly affect the check-in behaviors of users to POIs as well. For example, friends often go to some places like movie theaters or restaurants together and people also prefer visiting popular places, e.g., the Statue of Liberty and the Eiffel Tower. To this end, we use the social and popularity influences to derive the mass of locations for our gravity model in Equation (3).

Mass of a visited location. The mass of visited location \( l_i \) for a given user \( u \) can be determined straightforwardly based on the historical check-in frequency of \( u \) to \( l_i \) due to two reasons: (i) The mass of a location should be personalized in order to learn the real preference of a user for personalized location recommendations. (ii) The historical check-in frequency of user \( u \) on location \( l_i \) actually reflects the personalized preference of \( u \) to \( l_i \). Accordingly, we define:

\[
\text{Mass}(l_i) = R_{u,l_i},
\]

where \( R \) is the check-in frequency matrix from the check-in collection \( D \), as described in Definition 2.5, and \( R_{u,l_i} \) is the frequency of user \( u \) visiting location \( l_i \).

Mass of a unvisited location. The mass of unvisited location \( l \) for a given user \( u \) can be deduced by combining the popularity of \( l \) from all users and the check-in frequency of \( l \) from the social friends of \( u \), given by

\[
\text{Mass}(l) = F_{\text{Pop}}(p_l) \cdot F_{\text{Soc}}(q_{u,l}),
\]

\[
p_l = \sum_{u' \in U} R_{u',l},
\]

\[
q_{u,l} = \sum_{u' \in U} S_{u,u'} \cdot R_{u',l},
\]

where \( R_{u',l} \) is the frequency of user \( u' \) visiting location \( l \) (Definition 2.5) and \( S_{u,u'} \) indicates whether there exists a social link between users \( u \) and \( u' \) (Definition 2.6);
hence, $p_l$ is the popularity, i.e., the total number of check-ins of all users on $l$ and $q_{u,l}$ is the social check-in frequency of $u$’s friends to $l$.

In Equation (16), it is not feasible to simply take the product of the popularity $p_l$ and social check-in frequency $q_{u,l}$ as $\text{Mass}(l)$, since the new locations with large popularity and social check-in frequency will dominate the recommendation result. Sophisticatedly, we map the popularity and social check-in frequency into a normalized score via the increasing functions $F_{Pop}$ and $F_{Soc}$ with regard to $p_l$ and $q_{u,l}$, respectively. $F_{Pop}$ and $F_{Soc}$ are derived from the distributions of popularity and social check-in frequency learned from the check-in data. In terms of Section 4.2, we are inspired to model the popularity or social check-in frequency as a power-law distribution, which also has been validated in three publicly available real data sets collected from popular LBSNs: Foursquare [Gao et al. 2012], Gowalla and Brightkite [Cho et al. 2011].

For the clarity of this paper, we estimate the power-law distribution of the popularity and social check-in frequency, respectively, since they are independent of each other, although in which the similar process is applied, as follows.

4.3.1. Popularity Distribution Estimation for $F_{Pop}$. We assume that the probability density of popularity $p$ is a power-law distribution:

$$f_{Pop}(p) = (\beta - 1)(p + 1)^{-\beta}, p \geq 0, \beta > 1,$$

where $\beta$ is estimated from check-in frequency matrix $R$ by maximum likelihood estimation:

$$\beta = 1 + |L| \left[ \sum_{l' \in L} \ln \left( \sum_{w' \in U} R_{w',l'} + 1 \right) \right]^{-1},$$

in which $\sum_{w' \in U} R_{w',l'}$ is the popularity of location $l'$. Figure 8 shows the probability density of popularity (i.e., the dots) in the three real-world data sets matches the power-law distribution (i.e., the line) very well, estimated by Equations (19) and (20). Thus, these results have validated the assumption of the power-law distribution.

In contrast to temporal difference and spatial distance in Section 4.2, the attractive force of a gravity model is increasing with respect to the mass of locations, i.e., the popularity here, in terms of Equations (3) and (16). Accordingly, we normalize the effect of popularity by the cumulative distribution of $f_{Pop}$ to obtain $F_{Pop}$, given by

$$F_{Pop}(p_l) = \int_0^{p_l} f_{Pop}(p) dp = 1 - (p_l + 1)^{1-\beta}.$$  

Clearly, $F_{Pop}$ is an increasing function with respect to the popularity $p_l$ due to $1-\beta < 0$. This reflects the directly proportional relation of the attractive force and popularity.
4.3.2. Social Check-in Frequency Distribution Estimation for $F_{soc}$. As the distribution estimation for popularity, we similarly construct the distribution of social check-in frequency $q$ as follows. First, we assume its probability density as a power-law distribution:

$$ f_{soc}(q) = (\eta - 1)(q + 1)^{-\eta}, \quad q \geq 0, \eta > 1, $$  \hfill (22)

in which $\eta$ is estimated by check-in frequency matrix $R$ and social link matrix $S$:

$$ \eta = 1 + |U||L| \left( \sum_{u' \in U} \sum_{l' \in L} \ln \left( \sum_{u'' \in U} S_{u',u''} R_{u'',l'} + 1 \right) \right)^{-1}, $$  \hfill (23)

in which $\sum_{u'' \in U} S_{u',u''} R_{u'',l'}$ is the social check-in frequency of the friends $u''$ of user $u'$ on location $l'$. Figure 9 presents the probability density of social check-in frequency (i.e., the dots) in the three real-world data sets that approaches to the power-law distribution (i.e., the line) estimated based on Equations (22) and (23). Hence, the assumption of the power-law distribution has been validated.

Furthermore, since the attractive force of a gravity model is increasing with respect to the social check-in frequency, in terms of Equations (3) and (16), we normalize the effect of social check-in frequency based on the cumulative distribution of $f_{soc}$:

$$ F_{soc}(q_{u,l}) = \int_{0}^{q_{u,l}} f_{soc}(q) dq = 1 - (q_{u,l} + 1)^{1-\eta}. $$  \hfill (24)

$F_{soc}$ is an increasing function concerning the social check-in frequency $q_{u,l}$ because of $1-\eta < 0$ that models the real relation between the attractive force and social influence.

4.4. Implementation

ALGORITHM 1 outlines our proposed gravity-model-based location recommender system. It mainly has three phases: (i) the data pre-processing phase to prepare check-in frequency matrix and location-location transition graph; (ii) the parameter estimation phase to learn the model parameters $\alpha, \gamma, \beta, \eta$ for the power-law distribution of temporal difference, spatial distance, popularity and social check-in frequency; and (iii) the visiting probability prediction phase to predict the probability of a user visiting any new location that consists of three steps including distance derivation, mass derivation, and weighing additive Markov chain with the gravity model. It is important to note that: (a) We can compute the visiting probabilities of users to new locations offline and then make online recommendations for users by ranking the obtained visiting probabilities and returning new locations with the top-$k$ highest visiting probability. (b) ALGORITHM 1 can be optimized easily, e.g., the popularity of locations independently of users can be pre-computed. For clarity, we have not showed these optimizations.
ALGORITHM 1: LORE: A gravity-model-based location recommender system

**Input:** Check-in collection $D$ and social link matrix $S_{|V|×|U|}$.

**Output:** Top-$k$ new locations for each user.

1: // **Phase 1: The data pre-processing phase**
2: Construct check-in frequency matrix $R_{|V|×|L|}$ from $D$
3: Construct location-location transition graph ($L^2$TG) from $D$ (Section 3.1)

4: // **Phase 2: The parameter estimation phase**
5: Estimate $\alpha, \gamma, \beta, \pi$ based on Equations (9), (13), (20), (23), respectively

6: // **Phase 3: The visiting probability prediction phase**
7: for each $u \in U$ do
8: Extract check-in sequence for $u$ from $D$: $L_u = \{(l_1, t_1) \rightarrow (l_2, t_2) \rightarrow \cdots \rightarrow (l_n, t_n)\}$
9: Set user-specific latest check-in time: $t_c = t_n$
10: for each new location $l \in L \land l \notin L_u$ do
11: Initialize $Pr(l | L_u) = 0$
12: for each visited location $l_i \in L_u$ do
13: // **Step 3.1: Distance derivation**
14: Compute temporal difference $x_{t_i, t_c} = t_c - t_i$ in Equation (5)
15: Compute spatial distance $y_{l_i, l} = Haversine(l, l_i)$ in Equation (6)
16: Compute $F_{Tem}(x_{t_i, t_c})$ based on Equation (10)
17: Compute $F_{Spa}(y_{l_i, l})$ based on Equation (14)
18: Compute $Distance(l_{t_i}, l_{t_c}) = 1 / \left[F_{Tem}(x_{t_i, t_c}) \cdot F_{Spa}(y_{l_i, l})\right]$ in Equation (4)

19: // **Step 3.2: Mass derivation**
20: Compute $Mass(l_i) = R_{u,i}$ in Equation (15)
21: Compute popularity $p_i$ based on Equation (17)
22: Compute social check-in frequency $q_{u,i}$ based on Equation (18)
23: Compute $F_{Pop}(p_i)$ based on Equation (21)
24: Compute $F_{Soc}(q_{u,i})$ based on Equation (24)
25: Compute $Mass(l_i) = F_{Pop}(p_i) \cdot F_{Soc}(q_{u,i})$ in Equation (16)

26: // **Step 3.3: Weighing additive Markov chain with the gravity model**
27: Compute transition probability $TP(l_i \rightarrow l)$ based on Equation (1)
28: Compute $Gravity(l_i, t_c)$ based on Equation (3)
29: $Pr(l | L_u) = Pr(l | L_u) + TP(l_i \rightarrow l) \cdot Gravity(l_i, t_c)$ based on Equation (2)
30: end for
31: end for
32: return Top-$k$ new locations with the highest visiting probability $Pr(l | L_u)$ for $u$
33: end for

4.5. Applications

LORE can be easily applied to recommend users with personalized POIs based on their preferences that are learned from the community-contributed data through ALGORITHM 1. Here we demonstrate the applications of LORE in two typical scenarios: (1) recommending POIs to a user without knowing her current location and (2) recommending nearby POIs (e.g., POIs located within a user-specified distance) to a user based on her current location. The two application scenarios are specified below.

**Scenario I without the current location of users.** When the current location of a target user is unknown, LORE considers each new POI of the target user as a recommendation candidate, estimates the visiting probability of the user to all candidates, and recommends the user with the candidates that have the top-$k$ highest visiting probability. For example, as depicted by “Back Points” in Figure 10(a), the target user has visited the Hollywood Hills in Los Angeles, Mob Museum in Las Vegas, Yellowstone National Park in Wyoming, Oak Street Beach in Chicago, Empire State Building and Statue of Liberty National Monument in New York City. Whence, LORE infers that the user would like to travel around the world to explore new POIs, takes into
account new POIs' popularity from all users, check-in frequency from social friends, spatial distance and temporal difference compared to the visited history of the user based on ALGORITHM 1, and then returns the user with the top-5 new POIs denoted by “Green Marks”, i.e., the Golden Gate Bridge in San Francisco, Grand Canyon National Park in Arizona, White House in Washington, DC, Niagara Falls and Banff National Park in Canada.

**Scenario II with the current location of users.** When the current location of a target user is detected, LORE only considers the recommendation candidates as the POIs that are within a certain range of the current location of the user. The range is a circle centered at the current location with a radius threshold that usually has a default value given by the recommender system and can be modified by the user. The remaining recommendation process is the same as that in Scenario I. For instance, as depicted in Figure 10(b), the current location (denoted by “Red Star”) of the same target user is detected at St. Patrick's Cathedral in New York City. In terms of her current location and visited history on POIs, LORE suggests the user to visit the Times Square, Museum of Modern Art, Plaza Hotel, Central Park and Metropolitan Museum of Art that are around her current location and satisfy her personal preference derived from ALGORITHM 1.

5. **EXPERIMENTS**

This section evaluates the recommendation accuracy of LORE compared to the state-of-the-art location recommendation techniques on three real-world data sets.

5.1. **Three Real Data Sets**

We use three publicly available large-scale real check-in data sets that were crawled from Foursquare [Gao et al. 2012], Gowalla and Brightkite [Cho et al. 2011], in which the locations are distributed all over the world. The statistics of the data sets are shown in Table II. In the pre-processing, we split each data set into a training set and a 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. The 80% of check-in data with earlier timestamps are used as the training set and the other 20% of check-in data are used as the testing set. In
Table II. Statistics of the Three Real Data Sets

<table>
<thead>
<tr>
<th></th>
<th>Foursquare</th>
<th>Gowalla</th>
<th>Brightkite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of users</td>
<td>11,326</td>
<td>196,591</td>
<td>58,228</td>
</tr>
<tr>
<td>Number of POIs</td>
<td>182,968</td>
<td>1,280,969</td>
<td>772,965</td>
</tr>
<tr>
<td>Number of check-ins</td>
<td>1,385,223</td>
<td>6,442,890</td>
<td>4,491,143</td>
</tr>
<tr>
<td>Number of social links</td>
<td>47,164</td>
<td>950,327</td>
<td>214,078</td>
</tr>
<tr>
<td>User-POI matrix density</td>
<td>$2 \times 10^{-4}$</td>
<td>$2.4 \times 10^{-9}$</td>
<td>$1.9 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

the experiments, the training set is used to learn the recommendation models of the evaluated techniques described in Section 5.2 to predict the testing data.

5.2. Evaluated Techniques

We compare our proposed gravity-model-based recommender system (LORE) with the state-of-the-art location recommendation techniques including:

— STI: This method utilizes the Spatio-Temporal Influences through separately inferring users' preferences to locations at each time slot [Yuan et al. 2013].
— USG: This method is a unified location recommendation framework that integrates User preferences, Social and Geographical (i.e., spatial) influences [Ye et al. 2011].
— CoRe: This method fuses social collaborative filtering with geographical check-in probability density over latitude and longitude coordinates [Zhang and Chow 2015a].
— LCARS: This method builds a Location-Content-Aware Recommender System based on the well-known topic model (i.e., latent Dirichlet allocation) to infer personal interest and local preference (i.e., local specialty) [Yin et al. 2014].
— DRW: This method is a Dynamic Random Walk model that combines social and popularity influences [Ying et al. 2014].
— FMC: This method exploits the sequential influence of the latest visited location only based on the First-order Markov Chain. None of existing methods using the first-order Markov chain can be applied in our experiments, as discussed in Section 6. Instead, FMC has been integrated with our gravity model to make fair comparison. Note that the classical $n$th-order Markov chain is not applicable as well due to its exponentially increasing computational cost with respect to the order $n$.
— AMC: This method employs the sequential influence based on the $n$th-order Additive Markov Chain with the simple decay weights through leaning towards recently visited locations [Zhang et al. 2014b].

5.3. Performance Metrics

To evaluate the quality of location recommendations, it is important to find out how many POIs actually visited by a user in the testing data set are discovered by the recommendation techniques. For this purpose, we employ two standard metrics:

\[
\text{Precision} = \frac{\text{No. of discovered POIs}}{\text{No. of recommended POIs for the user: } k}. \\
\text{Recall} = \frac{\text{No. of discovered POIs}}{\text{No. of POIs actually visited by the user in the testing set}}.
\]

5.4. Parameter Settings

The number of recommended POIs (top-$k$) is set to a range from 2 to 20 because a larger number of recommended POIs may not be helpful for users. We also examine the recommendation quality regarding the length $n$ of a user’s visited location sequence with a range from 2 to 50, i.e., the number of visited POIs by a user in the training set. Note that $\alpha$, $\beta$, $\gamma$, and $\eta$ are not free parameters; they are learned from check-in data.
5.5. Experimental Results and Discussion

Here we analyze and discuss the experimental results.

5.5.1. Comparison of Recommendation Accuracy. Figures 11 and 12 compare the recommendation accuracy of the state-of-the-art location recommendation techniques regarding the number of recommended POIs for users (top-\(k\)) and the number of visited POIs of users in the training set (given-\(n\)), respectively, on the three real data sets.

STI. This method [Yuan et al. 2013] utilizes the spatiotemporal influences through modeling the spatial influence as a power-law distribution and inferring the temporal influence at each time slot separately which suffers from time information loss and may not correlate temporal influences at different time slots due to time discretization. Moreover, it ignores the social influence of friends on users. As a result, STI does not perform well in comparison to other recommendation methods.

USG. This method [Ye et al. 2011] linearly integrates user preference from user-based CF, social influence from social CF, and geographical (i.e., spatial) influence from
a power-law distribution. It improves the precision and recall compared to STI. However, it is difficult to determine the linear weights for user preference, social influence, and geographical influence. Moreover, the weights should not be unified, since some users are affected by social friends more and other users may rely on the geographical influence more. Subsequently, USG only gives the fourth best recommendation accuracy.

CoRe. This method [Zhang and Chow 2015a] employs the social influence in the same way as USG, but it models a personalized geographical check-in probability distribution over latitude and longitude coordinates for each user and combines the social and geographical influences by a more robust product rule rather than using the linear sum rule. Accordingly, CoRe outperforms USG to some extent and generates the third best recommendation precision and the second best recommendation recall.

LCARS. This method [Yin et al. 2014] exploits the well-known topic model, i.e., latent Dirichlet allocation, to infer personal interest and local preference (i.e., local specialty). The personal interest of a user or local preference of a region (e.g., a city) is represented
as a mixture of topics, in which each topic is a distribution over POIs and learned from the check-in data. Nonetheless, LCARS does not take into account the social and geographical influences; it suffers from low recommendation accuracy.

DRW. This method [Ying et al. 2014] adopts a dynamic random walk model to fuse the social and popularity influences. Unfortunately, like LCARS, it does not consider the unique characteristic of location recommendations for LBSNs, i.e., the influence of geographical information of POIs on the check-in behaviors of users, for instance,indoorsy persons like visiting POIs around their living areas while outdoorsy persons prefer traveling around the world to explore new POIs. Consequently, DRW also reports the low recommendation accuracy.

FMC. This method leverages the sequential influence based on the first-order Markov chain that only uses the latest visited POI of a user to determine the new POIs possibly visited by the user. Then, FMC generates the worst result at most cases, although it has been integrated with our gravity model. The reason is that the first-order sequential influence is not sufficient and in reality the new POIs may count on all visited history.

AMC. To overcome the limitation of FMC, AMC exploits the \( n \)th-order sequential influence. As a result, AMC greatly improves the recommendation accuracy of FMC. At the same time, AMC is considerably competitive to CoRe, i.e., AMC reports higher precision but lower recall in comparison to CoRe. These results show that both the geo-social influence and the higher-order sequential influence are very useful for location recommendations.

LORE. Our proposed LORE always exhibits the best recommendation quality in terms of precision and recall. In particular, it achieves the significant improvement compared to the second best recommendation techniques, i.e., CoRe and AMC. We attribute the promising results to three reasons: (1) LORE takes full advantage of the higher-order sequential influence based on the developed additive Markov chain. (2) The weight of each visited location in the historical check-in sequence of a user is determined by the devised gravity model that integrates a widely range of information, including spatiotemporal, social and popularity influences. (3) These influences are modeled as power-law distributions that are validated by and learned from the check-in data.

5.5.2. Discussion. We discuss the general trends and important findings as follows.

Effect of the sparsity of data. It is worth emphasizing that the accuracy of location recommendation techniques for LBSNs is usually not very high, because the density of a user-POI check-in matrix is pretty low. For example, the reported maximum precision is 0.06 over a data set with \( 2.72 \times 10^{-4} \) density in [Ye et al. 2011], and 0.03 over two data sets with \( 9.85 \times 10^{-4} \) and \( 6.35 \times 10^{-3} \) densities in [Yuan et al. 2013]. Thus, the relatively low precision and recall values are common and reasonable in the experiments. Instead, we focus on the relative accuracy of our LORE compared to the state-of-the-art location recommendation techniques and expect LORE can improve recommendation accuracy as more check-in activities are recorded. Fortunately, as depicted in Figures 11 and 12, the recommendation accuracy generally increases from Brightkite to Gowalla to Foursquare, because their density correspondingly raises from \( 1.9 \times 10^{-5} \) to \( 2.4 \times 10^{-5} \) to \( 2.3 \times 10^{-4} \) (Table II).

Effect of the number of recommended POIs for users. In Figure 11, we examine the recommendation quality respecting the recommended number \( k \) of POIs; note that recommending too many POIs is not helpful for users. As expected with the increase of \( k \), the recall gradually gets higher but the precision steadily becomes lower on the three data sets. The explanation is pretty straightforward: in general, by returning more POIs for users, it is always able to discover more POIs that users would like to
visit. However, the extra recommended POIs are less possible to be liked by users due to the lower visiting probabilities of these POIs, since the recommendation techniques return the POIs with the top-k highest scores. For example, the second returned POI has the lower visiting probability than the first one.

**Effect of the number of visited POIs by users in the training set.** In Figure 12, we investigate recommendation quality regarding the number n of visited POIs by users in the training set. When users check in more POIs, the performance of various recommendation techniques generally inclines. The reason is that they can learn the preference models of users to POIs more accurately through using more check-in data. For instance, the more check-in data are helpful for removing the uncertainty from obtained power-law distributions or sequential patterns.

**Role of different influences.** (1) Since human movement exhibits sequential patterns, the higher-order sequential influence can bring a lot of benefits for location recommendations, as indicated in AMC. (2) 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 POIs at some time. Thus, the spatiotemporal influences play a significant role on users' check-in behaviors. For example, CoRe records the second best recommendation quality as AMC, but LCARS and DRW suffer from low recommendation accuracy due to the lack of spatiotemporal influences. (3) The social influence is essential to model the correlation of check-in behaviors between friends. As an example, SfI fails to use the social influence and hence presents low recommendation quality. (4) The check-in data are highly sparse, since users have only visited a very small proportion of POIs in an LBSN. It is common that none of friends of a user visits a POI. Therefore, it is helpful to carefully use the popularity of POIs from all users, not just friends. (5) The method of modeling and integrating these influences is also the key to provide high quality of location recommendations. LORE models the sequential influence based on the n-th-order additive Markov chain and the spatiotemporal, social and popularity influences as power-law distributions from check-in data, and further fuses them based on the gravity model, which guarantee that LORE is significantly superior to the state-of-the-art location recommendation methods.

6. RELATED WORK

In general, there are six main categories for existing location recommendation approaches in LBSNs. Note that some works belong to more than one category, since they combine different recommendation methods with different input data.

**Collaborative filtering.** Most studies provide POI recommendations using the memory or model based collaborative filtering techniques on users' check-in data [Bao et al. 2012; Levandoski et al. 2012; Lian et al. 2015; Liu and Xiong 2013; Lu et al. 2012], GPS trajectories [Leung et al. 2011; Zheng et al. 2011; Zheng et al. 2012], or geotagged photos [Shi et al. 2013]. These studies usually concentrate on measuring the similarities among users or locations, for instance, the study [Zheng et al. 2011] takes into account three factors: (a) the sequence property of people's outdoor movements, (b) the visited popularity of a geographic region, and (c) the hierarchical property of geographic spaces. Specifically, Levandoski et al. [2012] applied the item-based collaborative filtering method, Bao et al. [2012], Lian et al. [2015], Lu et al. [2012] and Zheng et al. [2011] employed the user-based collaborative filtering methods, Liu and Xiong [2013], Zheng et al. [2012] and Shi et al. [2013] utilized the matrix factorization methods, and Leung et al. [2011] designed a new similarity-based approach and co-clustering technique. These classical collaborative filtering techniques have been extensively extended to integrate with other information such as social links between users, geographical coordinates and textual contents of POIs, as discussed below.
Social influence. Based on the fact that nearby friends are more likely to share common interests [Cho et al. 2011], social link information has been widely utilized to improve the quality of recommender systems in LBSNs. Current works usually compute the similarities between users from social links and put them into the collaborative filtering techniques. For example, most literatures [Gao et al. 2012; Lian et al. 2015; Lu et al. 2012; Ying et al. 2012; Ye et al. 2011; Zhang and Chow 2013; Zhang et al. 2014a] naturally fuse the similarity of users into the user-based collaborative filtering methods, some papers [Cheng et al. 2012; Yang et al. 2013; Zhao et al. 2014] combines the similarity of users as a regularized term into the matrix factorization or tensor models, and other articles [Wang et al. 2013; Ying et al. 2014] integrates the similarity of users into the random walk approaches.

Geographical or spatial influence. The geographical information of POIs has also been intensively used in location recommendations. For instance, the studies [Bao et al. 2012; Ference et al. 2013; Wang et al. 2013; Yin et al. 2014] consider the geographical influence of current locations; only POIs within a certain distance from the current locations will be possibly recommended to users. More sophisticatedly, the studies [Cheng et al. 2012; Hu and Ester 2013; Kurashima et al. 2013; Lian et al. 2014; Liu et al. 2013a; Liu et al. 2014b; Liu et al. 2014b; Liu et al. 2014d; Yao et al. 2014; Ye et al. 2011; Yuan et al. 2013; Yuan et al. 2014] model the geographical influence of check-in locations as a common distance distribution for all users, e.g., a power-law distribution or a multi-center Gaussian distribution. Further, the papers [Lian et al. 2015; Zhang and Chow 2013; 2015a; 2015b; Zhang et al. 2014a] personalize the geographical influence by modeling a personalized nonparametric distribution for each user.

Temporal influence. The time factor has been widely used for the conventional recommendations (e.g., books, music and movies) by considering the time difference between the occurring time of a previous rating and the recommendation time as a decaying factor to weigh the rating [Koren 2010]. In LBSNs, the check-in behaviors of users on POIs show temporally periodic patterns. For example, users visit restaurants rather than bars in the morning, but a bar rather than a library at midnight; and some users would like to visit stadiums in the daytime while others prefer at the night. These periodic patterns have been employed to provide location recommendations for users. Current works [Gao et al. 2013; Hu et al. 2013; Yuan et al. 2013; Yuan et al. 2014; Zhao et al. 2014] transform continuous time into discrete time slots and use the temporal influence separately for each time slot based on collaborative filtering techniques. Nonetheless, these techniques suffer from time information loss and may not correlate temporal influences at different time slots because of time discretization. To this end, the recent study [Zhang and Chow 2015c] develops a continuous temporal model based on the kernel density estimation method to build the continuous time probability density of a user visiting a new location.

Sequential influence. In terms of the fact that human movement exhibits sequential patterns [Cho et al. 2011; González et al. 2008; Song et al. 2010], various sequential mining techniques [Lian et al. 2015; Ying et al. 2013] 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 four groups. (1) Some researchers mined the most popular location sequence patterns from travel histories to guide users to plan a trip [Chen et al. 2011; Hsieh et al. 2014; Zheng et al. 2012]. 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 [Cheng et al. 2011] or a
mixture of topics in which a topic is a probability distribution over POIs [Kurashima et al. 2010]. 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 [Cheng et al. 2013] 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, this method 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. Moreover, all these studies [Chen et al. 2011; Cheng et al. 2011; Cheng et al. 2013; Kurashima et al. 2010; Zheng et al. 2012] exploit the sequential 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. (4) Further, in the previous work [Zhang et al. 2014c], we applied the classical nth-order Markov chain to utilize the higher-order sequential influence for the coarse-grained area or region recommendations instead of the fine-grained location or POI recommendations, because the classical higher-order Markov chain is prohibitively expensive due to its exponentially increasing number of states regarding the order. To this end, in the recent study [Zhang et al. 2014b], we devised an efficient nth-order additive Markov chain with the quadratic complexity of states that can be applied in the fine-grained location recommendations.

Content-aware methods. There are a few works that use the content information of POIs for location recommendations. For example, the work [Kurashima et al. 2013] employs the tags of POIs to provide interpretable representations for latent topics extracted from check-in data, the study [Zhao et al. 2014] represents each POI as a vector of words in its textual comments and utilizes the word vector to derive the similarity between POIs as a regularized term of tensor models over check-in data, and the research [Gao et al. 2015] leverages the user-by-word, POI-by-word and user-by-POI frequency matrices to deduce the latent preferences of users and latent features of POIs based on matrix factorization. Sophistically, some other methods [Hu and Ester 2013; Liu et al. 2013a; Liu and Xiong 2013; Wang et al. 2015; Yin et al. 2013; Yuan et al. 2015] apply the well-known latent Dirichlet allocation (LDA) over tags of POIs or comments of users to mine the topic profiles for users and POIs, in which each topic is a distribution over words; then the topic profiles and topic distributions are exploited to determine the preference score of users to POIs. Due to the difficulty of understanding the real meaning implied in the content information of POIs, the benefit from the use of the content information is considerably limited in these literatures.

7. CONCLUSION AND FUTURE WORK
This paper proposes a gravity-model-based location recommender system, called LORE. First, LORE exploits the higher-order sequential influence based on the nth-order additive Markov chain that considers all visited locations in the check-in history of a user to find out her probability of visiting new locations. Further, LORE weighs the effect of each visited location on a new location based on the gravity model that effectively integrates the spatiotemporal, social and popularity influences to determine the attractive force between the visited location and the new location. Finally, experimental results on three real-world data sets show that LORE achieves significantly better recommendation performance than the state-of-the-art location recommendation techniques.

In the future, we plan to extend LORE to recommend a trip of POIs for users, investigate the sequential patterns between location categories at the coarse level of
granularity, and develop methods to exploit the coarse-grained sequential patterns in fine-grained location recommendations.

REFERENCES


Bin Liu, Hui Xiong, Spiros Papadimitriou, Yanjie Fu, and Zijun Yao. 2014c. A general geographical probabilistic factor model for point of interest recommendation. IEEE Transactions on Knowledge and Data Engineering, accepted to appear (2014).


Zhijun Yin, Liangliang Cao, Jiawei Han, Jiebo Luo, and Thomas S. Huang. 2011. Diversified trajectory pattern ranking in geo-tagged social media. In Proceedings of the SIAM International Conference on Data Mining. Mesa, AZ, 980–991.


Jia-Dong Zhang, Chi-Yin Chow, and Yanhua Li. 2014b. LORE: Exploiting sequential influence for location recommendations. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. Dallas, TX, 103–112.


