IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 24, 2019, accepted July 5, 2019, date of publication July 16, 2019, date of current version August 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929313

A Multi-Element Hybrid Location Recommendation Algorithm for Location Based Social Networks

REN YUE-QIANG, WANG ZE, SUN XIAO-NA, AND SUN SHI-MIN
School of Computer Science and Technology, Tianjin Polytechnic University, Tianjin 300387, China

Corresponding author: Wang Ze (wangze@tjpu.edu.cn)

This work was supported in part by the Key Project Foundation of Tianjin under Grant 15ZXHLGX003901, in part by the National Natural Science Foundation of China under Grant 61702366, and in part by the Natural Science Foundation of Tianjin under Grant 16JCYBJC41500.

ABSTRACT In the environment of data explosion, how to make an effective and accurate personalized point of interest (POI) recommendation in location-based social networks (LBSNs) is a challenging and meaningful task. Fortunately, there is a lot of information that we can use. We can make recommendations by mining the rich information hidden in user check-in records. In this paper, we propose a recommend system named GFP-LORE. Specifically, we have designed a framework, which integrates various influencing factors. First, we modeled friend sign-in frequencies and POI popularity as a *power-law distribution* and the experiment proved that it is effective. Then, we got the influence of geographic information by modeling it according to the longitude and latitude of the user's check-in location. After that, we sorted the historical check-in records of all users according to time and then mine an overall pattern of location transfer-order pattern. Then, we combine it with each user's own unique location transfer record to get the possibility of the user going to the next POI. Finally, we synthesize the above four influence factors into a unified correlation probability rating and recommend a new location by this probability rating. We tested our system on the open real check-in data set, and the results of our simulation experiments show that the recommendation effect of our system is better than those algorithms used in the contrast test.

INDEX TERMS LBSNs, recommendation, location prediction, point of interest, Markov chain, power-law distribution.

I. INTRODUCTION

The check-in data obtained from LBSNs contains a wealth of knowledge about user interests, and this kind of knowledge can be used in many applications, such as location recommendation [1]–[4], friend recommendation [5], etc. Among them, points of interest(POI) recommendation is an important field in LBSNs. It not only helps users explore new points of interest, enriches user experience, but also helps companies to acquire more potential customers. Nowadays, the most popular POI recommendation algorithm is to use the traditional collaborative filtering technology for POI recommendation like [6], [7]. However, as no other useful information is considered, the recommendation accuracy of these algorithms is not high enough. A relatively good POI recommendation algorithm should reasonably use geographic and social information like [8]–[12]. Through modeling the user's check-in geo-location information and calculate check-in location similarity between friends, these studies can obtain and utilize the *geographical influences* and *social influences* to make recommendations [13]–[16]. In addition, because the location information about POI has a great influence on the user's check-in behavior and it's easy to have the similar interests in friends, so this kind of method have been proved to be very effective. But there's still a lot of useful information that these algorithms don't take into account. Some studies [8], [17]–[19] have shown that people prefer more popular locations and the human moving trajectory shows some kind of sequential pattern. For instance, people are used to going to the cinema to relax after work and the more popular cinemas will be easier to be chosen. Therefore, when researching recommendation algorithm, it is necessary to consider not only *geographical influences* and *social influences*, but also the *popularity influences* and *sequence influences* needs to be focused on. But it is not enough to consider adding just one influence factor. Although there will be some improvements,

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan Bu.

they also have some flaws and not satisfactory enough. So, this paper aims to reasonable use both *sequence influences* and *popularity influences* to integrate geographic information and social relations information to promote the quality of POI recommendation in LBSNs. For that, we develop a new location recommendation algorithm named GFP-LORE. Our main contribution is that:

1) We aggregate four recommendation impact factors into a unified framework, and integrate their influence into a relevant score, and finally recommend new location for users according to the level of this relevant score.
2) First, we modeled *popularity influences* and *social influences* as *power-law distributions* [20]. On the one hand, modeling based on power-law distribution has been proven to be an effective method in many studies [21]–[23]. On the other hand, we use simulation experiments and analysis real data to verify that POI popularity and friend check-in frequencies really conform to power law distribution. So our modeling method is reasonable.
3) Then, we use a method called kernel density estimation (KDE) [24] to utilize geographic information. Different from previous models based on *one-dimensional distance probability distribution* [10], [13], [25], we use KDE to modeled geographic location information as *two-dimensional check-in probability distribution* based on longitude and latitude of each userąŕs check-in location.
4) After that, we get user's overall access order patterns by location-location transition graph, and combine it with the Additive Markov Chain [26], train it through user history check-in data set, and ultimately obtain *sequential influence*.
5) Finally, we aggregate the above influence into a unified correlation score by product fusion rules. We rank the scores from high to low and recommend the *top-k* POI candidates to the user. We tested it by using the public check-in data set Gowalla [27] and the results show that our system actually improve the recommendation accuracy of the original algorithm.

II. RELATED WORK

A. TRADITIONAL RECOMMENDATION ALGORITHM

In the 1990s, research on personalized recommendation systems was proposed [28]. In general, the recommendation methods can be basically divided into four types: collaborative filtering, content-based recommendation, Location-based social network recommendation and mixed recommendation.

1) COLLABORATIVE FILTERING

The collaborative filtering recommendation method is the most widely used recommendation method. It was originally proposed to be applied to the mail filtering system [29]and it can be divided into the following two categories: User-based collaborative filtering (User-based CF) and Item-based collaborative filtering (Item-based CF) [30]. But this approach does not take into account the impact of geographic information and Social correlations for points of interest. Moreover, the method of collaborative filtering is only applicable to the case where the amount of data is small. If the amount of data is large, the cost of calculating matrix similarity is large.

2) CONTENT-BASED RECOMMENDATION

Content-based recommendation is a recommendation mechanism widely used in recommendation engines. By analyzing metadata of project content, it recommends to users the items with similar metadata information and user preference. The usual content-based recommendation process is: project keyword extraction, project to project similarity calculation, and recommendation using Item CF. Although this recommendation algorithm can relatively well model the user's behavior preferences, it also has certain drawbacks [31]. For example, common problems are: missing information, cold start and classification, and labels are difficult to control.

3) LOCATION-BASED SOCIAL NETWORK RECOMMENDATION

Location-based social network recommendations (LBSNs) added the location to the traditional social network. The location-based social network in [32] is introduced as follows: In addition to the location attribute added to the social network, the LBSNs allow the user to share the location. More importantly, the location information shows the location of the user at a specific time, and can also reveal User's historical access record.

B. POPULAR RECOMMENDATION ALGORITHMS

Here are some typical recommendation algorithms:

FMC [2], [17], [33]. *First-order Markov Chains*(FMC), which use the influence of the most recently visited location on check-in sequences for prediction and recommendation. FMC cleverly uses the *sequential influence* to improve the accuracy of the recommendation. However, since the FMC only considers the impact of recently visited locations, the recommended effect is not satisfactory.

AMC [26]. *Additive Markov Chain*(AMC) is an extension based on FMC. AMC approach uses the *nth-order Markov chain* and predicts where the user wants to go. Although AMC considers the impact of historical visit records, because AMC does not consider other factors, such as geographic factors, the recommendation effect is still not ideal.

GS2D [19]. GS2D predicts where the user wants to go by modeling the latitude and longitude coordinate information of a user's check-in location, and modeling the user's social relationship information at the same time. The user is recommended for location by *geographic and social influence*. And compared to iGSLR [25] which apply the *geographical influence* to recommendation algorithm through modeling distance distribution for each user, GS2D is more reasonable and effective.

TABLE 1. Key notations in the paper.

Symbol	Meaning
U	Collection of all users in the LBSN
u	A user:$u \in U$
L	Collection of all POI in the LBSN
L_u	Collection of all check-in locations for user u, i.e., $L_u = \{l_1, l_2, \cdots, l_n\}$
l	A POI with coordinate information:$l \in L$
t	User access time
t_c	User recent access time
$\langle u, l, t\rangle$	User u accesses to POI l at time t
D	The set of all sign-in records, i.e., $D = \{< u_i, l_i, t_i >\}$
S_u	The user u's sign-in sequences sorted by time t, i.e., $S_u = \{< l_1, t_1 >, < l_2, t_2 >, \cdots, < l_n, t_n >\}$
$l_i \rightarrow l_j$	User u transfer from l_i to l_j
$R_{\|U\|\times\|L\|}$	Sign-in frequency matrix, $R_{u,l}$ represents the number of visits to POI $l(l \in L)$ by user $u(u \in U)$.
$F_{\|U\|\times\|U\|}$	Social relationship matrix, $F_{u,u'}$ represents the social relations between u and u'
$P_{\|U\|\times\|L\|}$	Popularity matrix, $P_{u,l}$ represents the attraction of the POI l $(l \in L)$) to u $(u \in U$

LORE [19]. This model combines sequence factors, geographic factors, and friend relationships to recommend points of interest to users. It's the same algorithm we came up with, it is a hybrid model. It shows good recommendation effects, but it is not reasonable enough to calculate *social influence* based on the distance between users. So, we improved it by using other modeling approaches. In addition, we also added the *popularity influence* to improve the recommendation accuracy and finally got better recommendation effect.

C. BASIC SYMBOLS AND DEFINITIONS

This subsection will describe the data structure and basic definition of this paper. These symbols are extracted from the check-in data set from LBSNs, including users' historical access information, social relationships between users, and geographic locations information. Table 1 lists the key symbols used in this article and the relevant definitions are listed at the same time.

Definition 1 (Sign-in Record Tuple): Once sign-in record is expressed as $\langle u, l, t\rangle$ which means that the user u checks in at the point of interest l at time t. The $u \in U$, U is a set of all users in the LBSN and $l \in L$, L is a collection of points of interest.

Definition 2 (Sign-in Record Collection): The sign-in collection means that all users u_i access all POIs l_i at different time t_i. It is a collection of all user check-in records.

Definition 3 (Spatiotemporal Sequence Collection): The sign-in order of a user u is expressed as $S_u = \{< l_1, t_1 >, < l_2, t_2 >, \cdots, < l_n, t_n >\}$, where the user u accesses the POI l_i at time t_i ($t_1 \leqslant t_2 \leqslant, \cdots, \leqslant t_n$), abbreviated as $S_u = \langle l_1, l_2, \cdots, l_n\rangle$.

Definition 4 (Transition, Predecessor and Successor): Two consecutive POI l_i and l_{i+1} and a certain time threshold ΔT are given in the spatiotemporal sequence $S_u = \{< l_1, t_1 >, < l_2, t_2 >, \cdots, < l_n, t_n >\}$. If $t_{i+1} - t_i \leqslant \Delta T$,the l_i to l_{i+1} is a transition, and it is defined as $l_i \rightarrow l_{i+1}$. The *transition predecessor* of l_{i+1} is l_i and the *transition successor* of l_i is l_{i+1}.

Definition 5 (Sign-in Frequency Matrix): Each element in the matrix represents the sign-in frequency of the user $u(u \in U)$ at the POI $l(l \in L)$, i.e., the $R_{u,l}$ means the number of times a user u accesses a POI l. Because user check-in locations often account for only a small fraction of all locations, most of the elements of the matrix R are 0.

Definition 6 (Social Relationship Matrix): For two different users $u(u \in U)$,$u'(u' \in U)$, if u and u' are friends, then in the friend relationship matrix $F_{|U|\times|U|}$, $F_{u,u'} = 1$, otherwise $F_{u,u'} = 0$.

Definition 7 (Popularity Matrix): We think that the number of times user signed in at the location can indicate the attraction of this location $l(l \in L)$ to user $u(u \in U)$. So, for a better understanding, although Popularity matrix and Check-in frequency matrix are numerically the same, the meanings are different. $P_{u,l}$ represent the attraction of POI l to u. Notice that, we think the higher the number of the total frequency of all users' check-in, the POI has a better popularity, so we define $\sum_{u'(u' \in U)} P_{u',l}$ as the prevalence of POI l.

III. POI RECOMMENDATIONS

A. POPULARITY AND SOCIAL CORRELATIONS ANALYSIS

In the recommendation service of LBSNs, the social relationship between users and the popularity of POI both will influence user's habits about choice the point of interest to a large extent. Users more like to go to more popular places or some places where friends prefer to go. In this article, we introduced an effective modeling method to take advantage of the impact of popularity and social relationship factors. In order to establish the model of *social influence* and *popularity influence*, we first aggregate the social relations of user u and all users 'check-in frequencies into *friend check-in frequencies*. Then we model *social influence* and *popularity influence* as power-law distribution through *friend check-in frequencies* and *POI popularity* to calculate the influence of them on recommendation. Note that, as defined in Definition 7, we believe that the popularity of POIs with a higher frequency of sign-in frequencies is higher and if the user checks in frequently at a POI, the POI is more attractive to this user. In addition, we define the *popularity influence* and *social relationship influence* as F_{pop} and F_{fri} and calculate them by training our model with a large amount of real historical data.

Power Law. It means that the product of the number of connections a node has and the number of such nodes is a fixed value, that is, the geometric average is a fixed value. For example, there are 10 people with $1000, 100 people with $100, and 1000 people with $10. Drawing it in logarithmic coordinates gives you an oblique downward line.

Power-Law Distribution. The form of power law distribution is $y = kx^{-r}$, where k is a constant and r is the law's exponent and always greater than zero. Here, we take the logarithm of both sides of the formula, we get that lny and

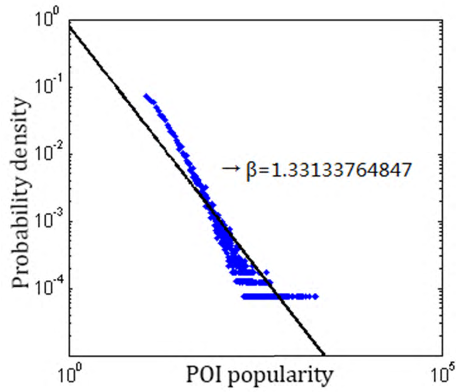


FIGURE 1. POI popularity distribution in real data.

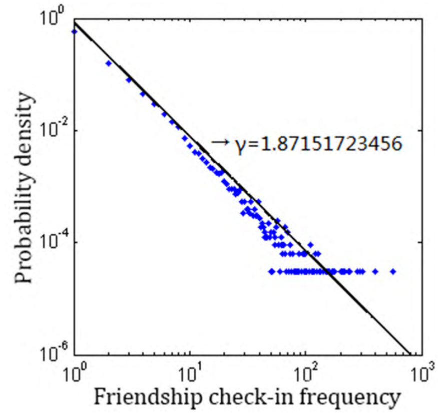


FIGURE 2. Friends check-in frequency distribution in real data.

lnx satisfy the linear relationship $lny = lnk - rlnx$. That is to say, on the logarithmic axis, the $Power - LawDistribution$ is represented by a straight line with a negative slope, which is also the basis for judging whether the random variable in a given event obeys the $Power - LawDistribution$(PD) [20].

1) POPULARITY CORRELATIONS ANALYSIS

Here, we presuppose that the POI popularity p is subject to the PD. The probability density function is:

$$f_{pop}(p) = (\beta - 1)(p + 1)^{-\beta} p \geq 0, \quad \beta \geq 1, \tag{1}$$

where β can be obtained by applying the Maximum Likelihood Estimate(MLE) in the Popularity matrix P, and the calculation formula is:

$$\beta = 1 + |L|([\sum_{l(l \in L)} ln(\sum_{u(u \in U)} P_{u,l} + 1)])^{-1}, \tag{2}$$

where $\sum_{u(u \in U)} P_{u,l}$ is the popularity of POI l (DEFINITION 7).

To verify our hypothesis, we get the β by training with real data set Gowalla [27] and substituted the β into Formula 1. After taking logarithms on both sides of the equation and presented it on the logarithmic graph, we get a straight line with a negative slope. After that, we analyzed the real public sign-in data set Gowalla and obtained the result shown in the Figure 1, which reflects that POI popularity (points in the graph) is fits with the power-law distribution (lines in the figure) that we estimated before. This result verifies that it is feasible to model popularity as a PD.

As a result of the POI popularity influence increases as the popularity increase, the cumulative distribution function of f_{pop} is used to obtain the influence of POI popularity (F_{pop}), defined as:

$$F_{pop}(P_{u,l}) = \int_0^{P_{u,l}} f_{pop}(p)\, dp = 1 - (P_{u,l} + 1)^{1-\beta}, \tag{3}$$

where $P_{u,l}$ means the POI l's attraction to User u (DEFINITION 7).

2) SOCIAL CORRELATIONS ANALYSIS

Similar to the above, for the distribution of the check-in frequency of the friends, we presuppose that the friend sign-in frequency z obeys the PD, and its probability density function is:

$$f_{fri}(z) = (\gamma - 1)(z + 1)^{-\gamma} z \geq 0, \quad \gamma \geq 1, \tag{4}$$

where γ is estimated by applying the Maximum Likelihood Estimate(MLE) in the sign-in frequency matrix R and the friend relationship matrix F:

$$\gamma = 1 + |L||U|([\sum_{u(u \in U)} \sum_{l(l \in L)} ln(\sum_{u'(u' \in U)} F_{u,u'} R_{u',l} + 1)])^{-1}, \tag{5}$$

where $\sum_{u'(u' \in U)} F_{u,u'} R_{u',l}$ is the total sign-in frequency of u's friend on the location l.

As before, we first compute the γ by training with real data and bring it into the Formula 4. After taking logarithms on both sides of the equation, we get a straight line with a negative slope. As shown in Figure 2, we compare this line with our analysis of the public sign-in data set, and prove that the friends' check-in frequency (points in the graph) also conforms to the power-law distribution (the line in the graph) we assumed before. The results show that our experiment is effective.

Because the influence of the friends' check-in frequency increases as the frequency of the friend's check-in increases, it is similar to the popularity factor, we obtain the influence of social relations (F_{fri}) through the cumulative distribution function of f_{fri}, express as Formula 6:

$$F_{fri}(z_{u,l}) = \int_0^{z_{u,l}} f_{fri}(z)\, dz = 1 - (z_{u,l} + 1)^{1-\gamma}, \tag{6}$$

where $z_{u,l}$ means the check-in frequency of the user's friends on the POI l, given by

$$z_{u,l} = \sum_{u'(u' \in U)} F_{u,u'} R_{u',l}. \tag{7}$$

B. GEOGRAPHIC CORRELATIONS ANALYSIS

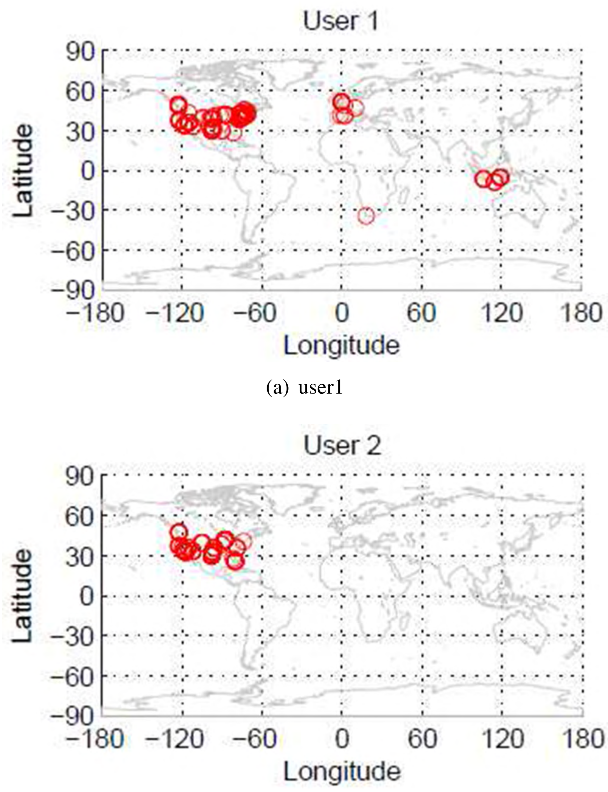
In LBSNs, point of interest is different from other non-spatial items because the user needs to physically interact with the location. Therefore, geographic information (like position coordinates) has a great impact on users' visit behavior. Some researchers turned to explore how to use geographic information to serve users. One way is that because the nearby friends share more commoner places to check in than others [34], [35], we can use the distance between the user's social friendsąŕ residence places to adjust their similarity weights. However, users often migrate from one location to another, so their home address sometimes can not reflect their real physical location. Not only that, the improvement of location recommendation quality by incorporating user residence information is also very limited.

There is an another better way to take advantage of this geographical information, which is to model the check-in behavior of all users by using the distance between each pair of POI accessed by the user. On the one hand, some scholars presuppose that the distance distribution is obeyed to PD (like the two influences in 3.1) [35], where the model parameters are derived from the entire check-in record history data set. On the other hand, the other scholars clustered on the entire historical check-in data set, finding the most popular point of interest as the center and assumed that the distance between the location and their centers was followed *Multi-center Gaussian Model*(MGM) [12].

In order to observe the user's unique check-in behavior, we analyzed the public data set collected from Gowalla [27]. Specifically, we extract the check-in location information of two users in the data set and analyze it. As shown in Figure 3, the geographic location information that we get is different among the two user check-in behaviors: User 1 (Figure 3(a)) travels around the world, and User 2(Figure 3(b)) only move in the United States. To further understands the impact of geographic factors on the check-in behavior of these two types of users, Figure 4 depict the check-in distributions based on the distance between each pair of POI they access. Their distance distributions are also unique and different, so we think the *geographic influence* should not be modeled as general distributions, such as PD or MGM.

In this paper, we actively explore the personalized impact of geographic location on user check-in behavior, and model it by the individual check-in distribution of each user, rather than the public distance distribution of every users. In our study, we are modeling geographic information based on non-parametric method called kernel density estimation, so there is no need to assume the form of the unknown distribution.

Kernel Density Estimation. Kernel density estimation(KDE) is a non-parametric method used to estimate unknown density functions in probability theory, proposed by Rosenblatt and Emanuel Parzen [24]. Unlike parameter estimation, non-parametric estimation does not need to add any prior knowledge, but only fits the distribution according to the characteristics and properties of the data itself.



(a) user1

(b) user2

FIGURE 3. User person check-in distribution.

Based on previous studies [14], model *geographic influence* according to the coordinate information of individual visit location is more accurate and reasonable. So, in this paper, we choose estimating the *two-dimensional check-in probability distribution based on latitude and longitude coordinates* instead of *distance probability distribution* like some current studies [10], [13], [16], [21], [22], [35]. Based on *kernel density estimation*, since it learns the distribution function from historical data, rather than assuming a specific distribution form in advance, we can give the probability $p^{geo}(l|L_u)$ of the user u to a unvisited location l as

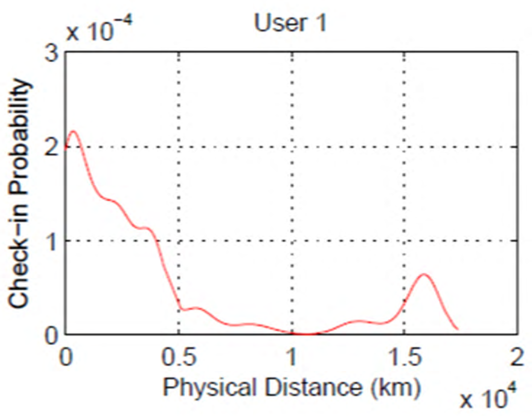
$$p^{geo}(l|L_u) = \frac{1}{n\delta^2}\sum_{i=1}^{n} K(\frac{l-l_i}{\delta}), \tag{8}$$

where each of the points of interest $l_i = (la_i, lo_i)^T$ is a two-dimensional vector with latitude and longitude (la_i and lo_i) coordinates, K(·) representative a non-negative function called kernel, L_u is a set of all visited locations for user u, i.e., $L_u = \{l_1, l_2, \cdots, l_n\}$. A smooth parameter called bandwidth is represented by $\delta(\delta \geq 0)$. Then we choose to use the standard two-dimensional normal kernel [24]

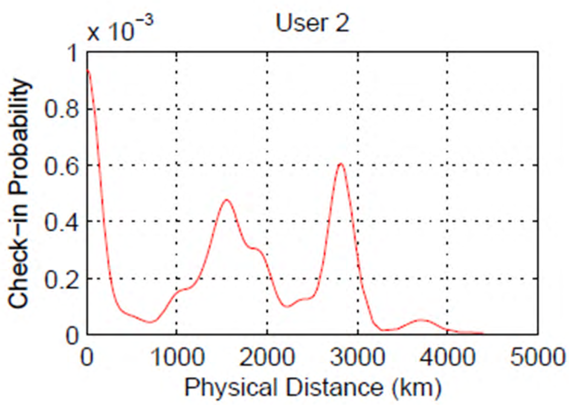
$$K(X) = \frac{1}{2\pi}exp(-\frac{1}{2}X^TX), \tag{9}$$

and optimal bandwidth [24]

$$\delta = n^{-\frac{1}{6}}\sqrt{\frac{1}{2}\hat{\delta}^T\hat{\delta}}, \tag{10}$$



(a) user1



(b) user2

FIGURE 4. User personalized check-in probability based on geographic distance.

where $\hat{\delta}$ is the edge standard deviation vector of L_u, which can be obtained by the following formula

$$\hat{\delta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(l_i - \hat{l})^2}, \tag{11}$$

where $\hat{l}$ is the mean of the latitude and longitude values

$$\hat{l} = \frac{1}{n}\sum_{i=1}^{n} l_i. \tag{12}$$

C. SEQUENTIAL CORRELATIONS ANALYSIS

As a matter of fact, user movement behavior is a space-time sequence transferred from (l_n, t_n) to another (l_{n+1}, t_{n+1}). To take advantage of the user's access pattern information, we first sorted the check-in records of each user in chronological order to get *Spatiotemporal Sequence Collection*, i.e., S_u. Then, We try to mine the sequence pattern information from the space-time sequence of all users and uniformly model the information into a *Location-Location Transition Graph*(L^2TG). This diagram model effectively reflects the overall access sequence pattern of all users in the user set. Based on this graph model, we can get the location transition

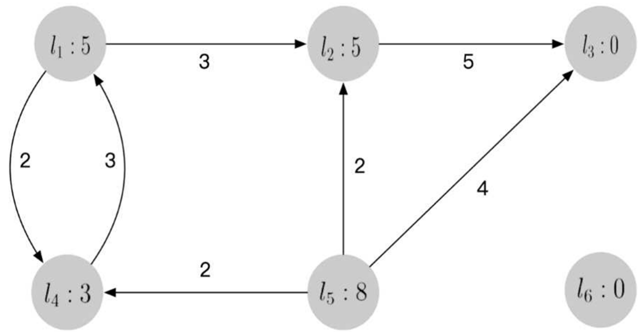


FIGURE 5. An example of location-location transition graph.

probability of users among all locations. Specifically, we can use this graph model to get the probability that the user accesses $l_{n+1} \in L$ after accessing $l_n \in L$. Even if l_{n+1} is a POI that the user has never accessed, we can still get the probability of user u accessing l_{n+1} by the overall access sequence pattern of all users and recently visited location l_n.

However, it is not scientific to simply use the above method to obtain the transfer probability, because it only considers the influence of the most recently visited place on the next to be visited, ignoring the impact of each user's unique historical access sequence. So, in order to solve one of the defects, after getting the position transfer diagram, we combine the user's own historical access order information with the location transfer graph by using the Additive Markov Chain, and finally obtained a more reasonable location transfer probability.

Location-Location Transition Graph [19]. Location-Location Transition Graph(L^2TG) is a kind of graph model, which is generated by combining the concept of graph with user's access pattern. As shown in Figure 5, L^2TG consists of a set of vertices and directed edges. Each node(circle) represents a POI $l_i(l_i \in L)$, and the figures in the circle represents the POI's out-degree, which means the number of POI l_i as a transition predecessor, denoted as $OCount(l_i)$. Each directed edge(arrow) represents a transition from POI l_i to POI $l_{i+1}(l_{i+1} \in L)$, denoted as $l_i \rightarrow l_{i+1}$. The number on each edge represents the frequency of transitions from POI l_i to POI l_{i+1}, denoted as $TCount(l_i, l_{i+1})$. For example, the l_1 transfer to l_2 three times as the transition predecessor of l_2 and transfer to l_4 two times as the transition predecessor of l_4, so the out-degree of l_1 are 5, that is, $OCount(l_1) = 5$. Although in the figure, there also are transitions between l_5 and l_3, but the number of times l_3 is used as a transition predecessor is 0, so the out-degree of l_3 are 0. Note that as defined in Definition 4, all of the above are discussed in a given time range. In other words, only when the difference between two POI (l_i and l_{i+1}) check-in time (t_i and t_{i+1}) are less than the given threshold (ΔT), i.e., $t_{i+1} - t_i \leqslant \Delta T$, we can consider l_i to l_{i+1} as an true transfer. In contrast, if the difference value is greater than the given threshold, i.e., $t_{i+1} - t_i > \Delta T$, we believe that these two POIs are irrelevant and the transfer between them is invalid. According to the out-degree and transfer frequencies in L^2TG, we can calculate the transfer probability from l_i to l_{i+1} denoted as $TP(l_i \to l_j)$. If out-degree of $l_i(OCount(l_i))$ is non-zero, we can calculate the transition probability from l_i to $l_(i+1)$ by the ratio of the transfer frequency between l_i and $l_{i+1}(TCount(l_i, l_j))$ and the out-degree of l_i, expressed as Formula 13:

$$TP(l_i \to l_j) = \frac{TCount(l_i, l_j)}{OCount(l_i)}, \tag{13}$$

If the out-degree of l_i is 0, it means that in a given time range, the user does not visit any other place again after visiting l_i, then we can express it as Formula 14:

$$TP(l_i \to l_j) = \begin{cases} 1, & l_i = l_j \\ 0, & l_i \neq l_j \end{cases}, \tag{14}$$

where record the transition probability from l_i to l_i as 1. In the actual check-in situation, the user's check-in records with time interval $< u_i, l_i, t_i >$ are continuously accumulated, which will form a set of infinite data streams $D = \{< u_i, l_i, t_i >\}_{i=1}^{+\infty}$. Therefore, it is necessary to process the incoming data according to the order of arrival, and gradually update the constructed L^2TG. Taking this into account, L^2TG is associated with the transfer frequency and the out-degree, rather than the transition probability, so that the L^2TG can be incrementally updated in an online manner. So far, we use the graph model method to construct the access order pattern of all users into a location-location transfer graph(L^2TG). We can mine the overall access order pattern through L^2TG and use it to make new POI recommendation for users. In order to use this information, we introduce the concept of nth-Markov Chain [36].

nth order Markov Chain. Markov Chain (MC) is a stochastic process which has Markov property and exists in discrete index set and state space. The Markov Chain must be "memory-less". It is means that the probability of future actions are only depend on the current state. It can be defined by transfer matrices or transition graphs. The conditional probability of a random variable satisfies the following relationship [36]:

$$p(X_{i+1}|X_i, X_{i-1}, X_{i-2}, \cdots, X_1) = p(X_{i+1}|X_i). \tag{15}$$

as for nth-order Markov chains, it has nth-order "memory", which can be regarded as the generalization of Markov chains. And it satisfies the following relationship:

$$p(X_{i+1}|X_i, X_{i-1}, \cdots, X_1) = p(X_{i+1}|X_i, X_{i-1}, \cdots, X_{i-n+1}). \tag{16}$$

What I want to show here is that the traditional Markov chain is the First order Markov chain. Many existing related studies [2], [17], [18], [33], [37] are based on the use of FMC to derive the prediction problem of sequence probability. Given a check-in sequence $S_u =< l_1, l_2, \cdots, l_n >$ and a unvisited l_{n+1}, it can be expressed as a formula:

$$p_{FMC}^{seq}(l_{n+1}|S_u) = TP(l_n \to l_{n+1}) = \frac{TCount(l_n, l_n + 1)}{OCount(l_n)} \tag{17}$$

FMC assumes that the probability of a user sign in at a new location l_{n+1} is only related to the most recently visited location l_n, and without considering the other locations previously visited at all. But, the actual situation is often not like this, the places which had been visited a long time ago sometimes contain a lot of useful information. Therefore, the low accuracy of FMC prediction is an inevitable problem. However, if the traditional nth order Markov model is used, the complexity of the algorithm will increase as n increases, and the cost and efficiency will be unsatisfactory. Based on the above considerations, we decided to use the improved Markov model - *Additive Markov Chain*(AMC) [26].

mth Additive Markov Chain. [26] *mth Additive Markov Chain*(AMC) is an extension of *nth order Markov Chain*. It having the following property: the probability of future actions X_{n+1} is affected by the previous m states, and the influence of previous states is additive. It can be expressed as Formula 18 and Formula 19:

$$p(X_{n+1}|X_n, \cdots, X_1) = p(X_{n+1}|X_n, \cdots, X_{n-m+1}), \tag{18}$$

together with

$$p(X_{n+1}|X_n, \cdots, X_{n-m+1}) = \sum_{i=1}^{m} f(x_{n+1}, x_{n-i+1}, i), \tag{19}$$

where $X_{n+1} = x_{n+1}, X_n = x_n, \cdots, X_{n-m+1} = x_{n-m+1}$.

In this paper, we can rewrite it to that: given a check-in sequence $S_u =< l_1, l_2, \cdots, l_n >$, the probability of access to new location l_{i+1} defined by AMC as:

$$p^{seq}(l_{n+1}|S_u) = \sum_{i=1}^{n} f(l_{n+1}, l_i, n - i + 1), \tag{20}$$

where $f(l_{n+1}, l_i, n + 1 - i)$ means the weighted contribution of POI l_i to the total probability of visit new location l_{n+1}, i.e., $p^{seq}(l_{n+1}|S_u)$. Since the POI recently visited are usually more influential than the POI previously visited for the probability of visiting to a new location [18], [23], [33], so, for the transition probability of l_t to l_{n+1}, [19] propose a method to calculate the weight of l_t. In combination with Formula 17 and Formula 20, the weighted contribution of position l_t is

$$f(l_{n+1}, l_t, n + 1 - t) = \frac{W(l_t)TP(l_t \to l_{n+1})}{\sum_{i=1}^{n} W(l_i)}, \tag{21}$$

together with

$$W(l_t) = 2^{-\alpha(n-t)}, \tag{22}$$

where $2^{-\alpha(n-t)}$ is the sequence attenuation weight of the attenuation parameter($\alpha > 0$),and the larger α, the higher the decay rate.

After integrating Formula 22 Formula 21 and Formula 20, we can get the probability of visit a new location $l_n + 1$ as:

$$p^{seq}(l_{n+1}|S_u) = \sum_{i=1}^{n} \frac{W(l_i)TP(l_i \to l_{n+1})}{W(l_i)}. \tag{23}$$

TABLE 2. User check in data set details.

Attributes	Number
Number of users	18737
Number of POIs	32510
Number of check-in records	1278274
Number of social ties	86985
User-POI check-in matrix density	1.3×10^{-4}

D. POI COMPREHENSIVE RECOMMENDATIONS

In this paper, we adopt the widely used product fusion rule to integrate the influence of different elements. The main reason why we use it is that the product fusion rule is less complex and has been proved to be highly robust in many previous studies [10], [12], [25]. We integrate $F_{pop}(P_{u,l})$, $F_{fri}(q_{u,l})$,$p^{geo}(l|L_u)$,$p^{seq}(l|S_u)$ into a unified relevant probability score $C(u, l)$. The final result is

$$C(u, l) = F_{pop} \times F_{fri} \times p^{geo} \times p^{seq}, \tag{24}$$

where $C(u, l)$ is a unified relevant probability score for user u to unvisited POI l, reflecting comprehensive influence of sequence, friend, popularity and geography. Then, we can recommend *top-k* POIs for users according to this score. The higher the score, the more likely the user u will go to the this POI.

IV. EXPERIMENTS AND ANALYSIS

A. REAL DATA SETS

The data set we used to test and train our framework is openly large-scale real check-in data sets which is crawled from Gowalla [27] between February 2009 and October 2010. The details of the data set after data cleaning are shown in Table 2.

B. EXPERIMENTAL ENVIRONMENT

In the experiment, since we must use the past check-in data to predict future check-in events, so the data set is divided into training set and test set according to check-in time, instead of using random partition method. The train set is half of the earlier data, and the other half is the test set(see Table 2 for more information about the dataset). We set the time interval(ΔT) in Definition 4 to one day and set the attenuation parameter α of Formula 22 to 0.05. The specific reasons will be analyzed later. In the test, we evaluated the *Precision* and *Recall* of recommendation techniques, covering the *top-k* only range from 2 to 20, because we do not think it makes sense to recommend too many locations to users.

C. PERFORMANCE METRICS

The recommendation algorithm usually calculates a target user's preference score for each unvisited location, and then recommends to the user the *top-k* location candidates by this return score. To evaluate the quality of the recommendation algorithm, the most important thing is the ratio of the number of POI actually accessed by the target user to the number of POI recommended by the recommendation algorithm. This paper uses the *Precision* and *Recall metrics* to assessment the recommendation results:

$$\begin{aligned} Precision &= \frac{\sum_{u\in U} |X_u| \cap |T_u|}{k} \\ Recall &= \frac{\sum_{u\in U} |X_u| \cap |T_u|}{\sum_{u\in U} |T_u|} \end{aligned} \tag{25}$$

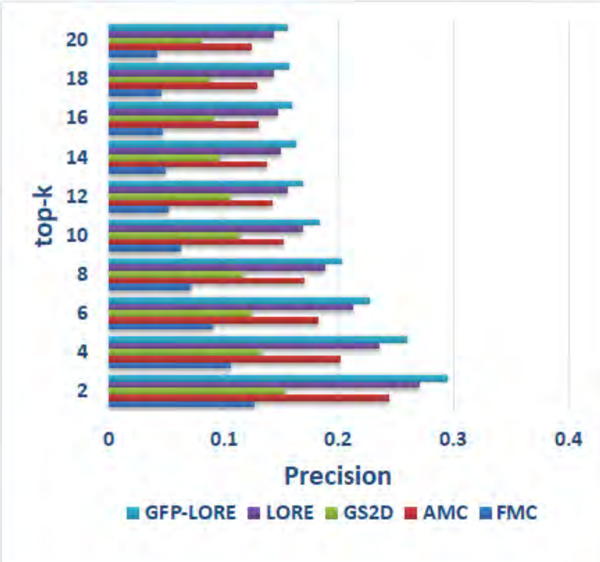
where X_u represents a set of *top-k* POI candidates recommended by the system for user u, and T_u represents a set of location really visited by the user u.

D. COMPARISON OF EXPERIMENTAL RESULTS

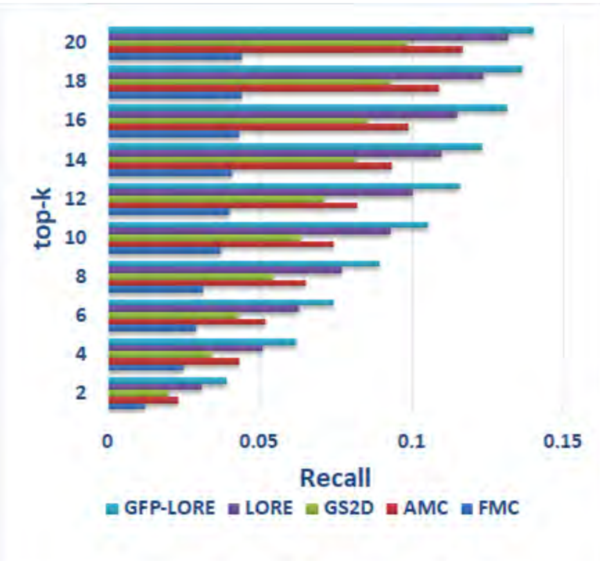
After considering a series of factors such as sequence factors, geographic factors, friend relationships and popularity of interest points, we proposed the GFP-LORE POI recommendation framework. In order to verify the quality of our proposed framework, we compared our framework with some other advanced recommendation algorithms. Such as: FMC [2], [17], [33] and AMC [26] which recommend new POI to users based on *sequence influence*; GS2D, a recommendation algorithm which modeling *geographic influence* as the *two-dimensional check-in probability distribution* and considering the influence of friends at the same time; LORE [19], a hybrid recommendation algorithms combines *sequence influence*, *geographic influence* and *social influence*. We have experimented with these four algorithms on the Gowalla data set and compared the results with our algorithms. Figure 6 show the *Precision rate* and *Recall rate* of each recommended algorithm and we have summarized the most important and most common findings after it.

The following are analysis and comparison of various algorithms: 1)

1) **FMC:** The FMC only considers the impact of the sequence by using ordinary Markov chain to obtain the probability of user access to the new location. It fails to take full advantage of the sequence pattern in the POI recommendation because it ignores the influence of previously visited POI in access sequence on predict new possible POIs. As shown in Figure 6, the FMC recommends the most inaccurate POI according to *Precision rate* display, and misses most of the POI that the target user actually visits in terms of *Recall rate*.
2) **AMC:** In order to make up for the deficiency of FMC, AMC use the nth order sequence to derive the user's access probability to the new POI based on sequential pattern extracted from historical records. In Figure 6, AMC has significantly improved *Precision* and *Recall* rate compared to FMC. These results validate the superiority of recommending POI using the whole check-in sequence, rather than just considering the impact of the most recently visited POI.
3) **GS2D:** GS2D utilizes geographic location information by analyzing the latitude and longitude coordinates distribution of the user's check-in location. This kind of modeling method is more complex but effective than the modeling method based on distance because it over-



(a) Precision



(b) Recall

FIGURE 6. The effect of top-k.

came the difficulty of finding the reference position and the difficulty of obtain a reasonable distance for the new location. However, GS2D does not consider the other useful information, so its *Precision* and *Recall* rate are not ideal. As shown in figure, it is better than FMC and inferior to other algorithm.

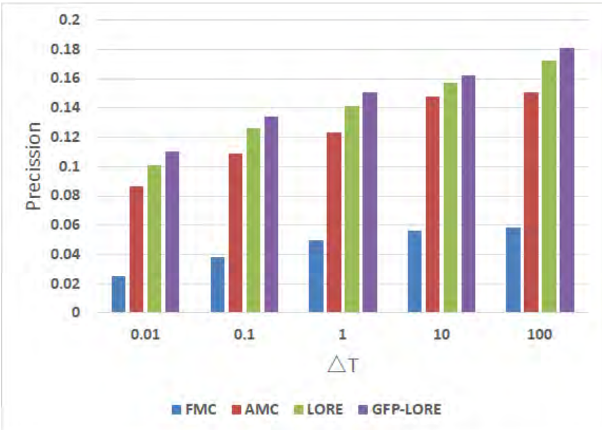
4) **LORE:** LORE inherits the advantages of AMC and GS2D. Therefore, its *Precision* and *Recall* rate are superior to the above three recommend algorithms. However, it is unreasonable for LORE to use the family residence distance to the calculation of the influence of the social, and it does not consider the influence of the popularity of the POI. So this recommendation algorithm is not as effective as what our proposed.

5) **GFP-LORE:** In view of the deficiencies in LORE, GFP-LORE proposed in this paper models *social influence* and *POI popularity influence* according to *power law distribution*. And the simulation experiment verifies the validity of such modeling method. Then, it is fused with the *sequence influence* and the *geographical influence*, realize effective recommendations for new locations. In the experimental results, the *Precision* and *Recall* rate shows the best recommendation quality.

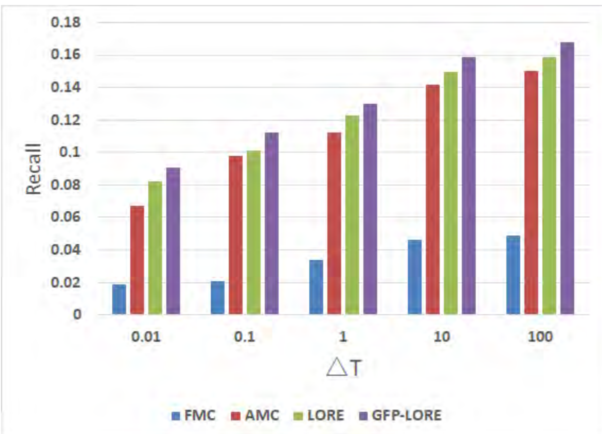
E. PERFORMANCE ANALYSIS

We not only analyze the influence of data sparsity and the number of *top-k* on the recommended quality, but also we investigated the effects of time interval ΔT and attenuation parameter α on the recommended quality of GFP-LORE.

1) **Impact of data sparsity:** Since the density of user check-in matrices on LBSNs is very low, so the accuracy of the recommend method is usually not very high. Because the check-in data set is very sparse(i.e., low check-in matrix density), the information we can mine from it is very limited. Notice that, what we're comparing here is the magnitude of the improvement in accuracy, not the magnitude of the accuracy number. So, as shown in Figure 6, the relatively low *Precision* and *Recall* rate are reasonable and consistent with expectations, and the *Precision* and *Recall* rate are improved compared with other recommendation algorithms in the experiment.
2) **The influence of *top-k*:** The experimental results in Figure 6 show that as the number of k increases, the *Recall* rate increases gradually and the *Precision* decreases. The reason is obvious. If we recommend more candidates for user, we can find more location that users actually visit. However, since other candidates have lower access probability, these POIs are less likely to be accessed by the user. The recommendation algorithm returns the *top-k* POIs of the highest score in turn, so the larger the value of k, the more POIs are returned, but the probability of accessing the POIs is gradually reduced.
3) **The influence of time interval:** Figure 7 depict the effect of time interval ΔT in Definition 4 on sequence-related recommendation system. Since the geo-social recommendation system (such as GS2D) are not influenced by time interval, so the comparison here does not include GS2D. As shown in Figure 7, the performance of GFP-LORE is always better than FMC, AMC and LORE. In addition, when the value of ΔT is gradually adjusted from 0.01 to 100, the *Precision* and *Recall* rate of GFP-LORE are both gradually increased first, and then tend to be stable. According to definition 4, The number of POI transfer becomes larger as ΔT becomes larger. But when time interval is greater than the maximum time interval between two sign-in location which are consecutive in the user check-in sequence, the num-

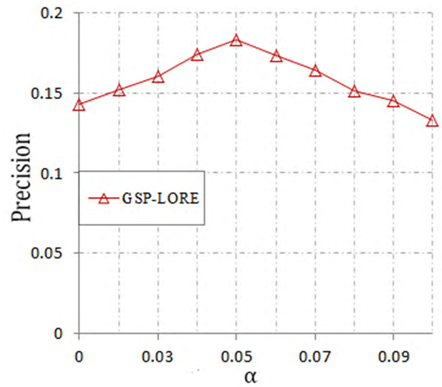


(a) Precision

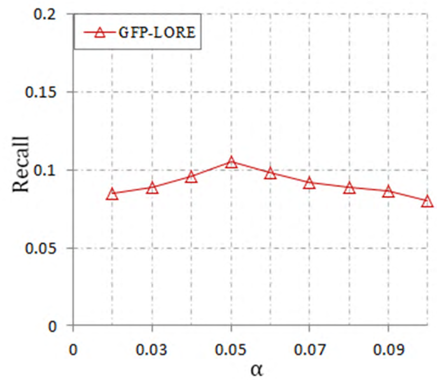


(b) Recall

FIGURE 7. **The effect of time interval.**



(a) Precision



(b) Recall

FIGURE 8. **The effect of attenuation parameter.**

ber of POI transfer remains the same. Based on this finding, we should not use the small threshold ΔT to segment the user's check-in sequence because some users do not travel often and they need more time intervals to plan the trip. So the accuracy of GFP-LORE is likely to be improved by using larger default values.

4) **The influence of attenuation parameter:** Figure 8 shows the influence of the attenuation parameter α on the recommended *Precision* and *Recall* of GFP-LORE. It is unreasonable that the influence of previous visits location on predicting new locations has not been weakened at all. Therefore, we have to weaken the sequential impact of the history check-in POI on the new POI but the value of α cannot be too large. The best value of α always between 0.01 and 0.1. As a result of the sequence attenuation weight in Formula 22 decreases with the decrease of i, and i represents the order of the user's visit location. Figure 8 shows the optimal range of performance of GFP-LORE is the best at [0.03, 0.07]. This important feature has made it reasonable to choose a default value, Because finding the optimal value is usually time consuming. In our experiments, the α was set to 0.05.

V. CONCLUSION AND FUTURE WORK

In this article, we propose a new hybrid recommendation algorithm framework called GFP-LORE. We integrate *social influence*, *popularity influence*, *geographic influence* and *sequential influence* into a unified framework, and prove that this method effectively improves the accuracy of recommendations. Firstly, we prove that the user's social correlations and POI's popularity are subject to power-law distribution, and modeled them according to this phenomenon, implemented through social factor and popularity factor recommend a new POI. Then, we analyzed the user's individual check-in distribution through the user's check-in history, explore the personalized geographic information in the user's check-in behavior, and calculate the probability of the user's arrival to the new location based on the method of *Kernel Density Estimation*(KDE). After that, our system mine the sequence pattern from the check-in data of all users in the form of dynamic L^2TG which can reflect overall transfer sequence pattern, and derive the probability of the user accessing the new POI based on the *Additive Markov Chain*(AMC), realize recommend to the user through the user's history sequence pattern. Finally, we integrate the above four influence factors into a unified recommendation framework, get a unified relevant probability score and recommend new locations to users based on this probability score.

We train and test our algorithms by using the publicly check-in data sets. The final result proves that the recommendation accuracy of our GFP-LORE algorithm is better than other recommendation algorithms in the experiment. In future work, we expect to combine the existing recommendation algorithm framework with mainstream *Deep Learning* algorithms. Extract the keywords in the user's evaluation of the POI, analyze the emotions contained in them, and add them as an impact factor to the framework to improve the accuracy of the recommendation.

REFERENCES

[1] F. Yu, Z. Li, S. Jiang, and S. Lin, "Point-of-Interest recommendation for location promotion in location-based social networks," in *Proc. 18th IEEE Int. Conf. Mobile Data Manage. (MDM)*, 2017, pp. 344–347.
[2] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013, pp. 2605–2611.
[3] W. Luan, G. Liu, C. Jiang, and L. Qi, "Partition-based collaborative tensor factorization for POI recommendation," *IEEE/CAA J. Automatica Sinica*, vol. 4, no. 3, pp. 437–446, Jul. 2017.
[4] J. He, X. Li, and L. Liao, "Category-aware next point-of-interest recommendation via listwise Bayesian personalized ranking," in *Proc. IJCAI*, Aug. 2017, pp. 1837–1843.
[5] M. Wu, Z. Wang, H. Sun, and H. Hu, "Friend recommendation algorithm for online social networks based on location preference," in *Proc. 3rd Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, Jul. 2016, pp. 379–385.
[6] A. Deng, Z. Zuo, and Y. Zhu, "Collaborative filtering recommendation algorithm based on item clustering," *Mini-Micro Syst.*, vol. 25, no. 9, pp. 1665–1670, Jul. 2004.
[7] A. Deng, Y. Zhu, and B. Shi, "A collaborative filtering recommendation algorithm based on item rating prediction," *J. Softw.*, vol. 14, no. 9, pp. 1621–1628, 2003.
[8] J. Zhang and C. Chow, "Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations," in *Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2015, pp. 443–452.
[9] L. Hu, A. Sun, and Y. Liu, "Your neighbors affect your ratings: On geographical neighborhood influence to rating prediction," in *Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2014, pp. 345–354.
[10] J. Zhang, C. Chowmember, and Y. Li, "iGeoRec: A personalized and efficient geographical location recommendation framework," *IEEE Trans. Services Comput.*, vol. 8, no. 5, pp. 701–714, Oct. 2015.
[11] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proc. 20th Int. Conf. Adv. Geographic Inf. Syst.*, Nov. 2012, pp. 199–208.
[12] C. Cheng, H. Yang, I. King, and M. R. Lyu, "Fused matrix factorization with geographical and social influence in location-based social networks," in *Proc. 26th AAAI Conf. Artif. Intell.*, 2012, pp. 17–23.
[13] B. Liu, Y. Fu, Z. Yao, and X. Hui, "Learning geographical preferences for point-of-interest recommendation," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2013, pp. 1043–1053.
[14] J. Zhang and C. Chow, "CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations," *Inf. Sci.*, vol. 293, pp. 163–181, Feb. 2015.
[15] H. Gao, J. Tang, X. Hu, and H. Liu, "Content-aware point of interest recommendation on location-based social networks," in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 1721–1727.
[16] Y. Zhao, L. Nie, X. Wang, and T. Chua, "Personalized recommendations of locally interesting venues to tourists via cross-region community matching," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, Sep. 2014, Art. no. 50.
[17] Y. Zheng, Z. Zha, and T. Chua, "Mining travel patterns from geotagged photos," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, May 2012, Art. no. 56.
[18] Z. Chen, H. T. Shen, and X. Zhou, "Discovering popular routes from trajectories," in *Proc. IEEE 27th Int. Conf. Data Eng.*, 2011, pp. 900–911.
[19] J. Zhang, C. Y. Chow, and Y. Li, "Lore: Exploiting sequential influence for location recommendations," in *Proc. ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, Nov. 2014, pp. 103–112.
[20] H. Hu and W. Lin, "A brief history of power law distributions," *Physics*, vol. 34, no. 12, pp. 889–896, 2005.
[21] 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 *Proc. 6th ACM Int. Conf. Web Search Data Mining*, Feb. 2013, pp. 375–384.
[22] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2013, pp. 363–372.
[23] X. Liu, Y. Liu, K. Aberer, and C. Miao, "Personalized point-of-interest recommendation by mining users' preference transition," in *Proc. 22nd ACM Int. Conf. Inf. Knowl. Manage.*, Nov. 2013, pp. 733–738.
[24] P. Green, A. H. Seheult, and B. W. Silverman, "Density estimation for statistics and data analysis," *Appl. Statist.*, vol. 37, no. 1, p. 120, 1988.
[25] J. Zhang and C. Chow, "iGSLR: Personalized geo-social location recommendation: A kernel density estimation approach," in *Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, Nov. 2013, pp. 334–343.
[26] O. V. Usatenko, S. S. Apostolov, Z. A. Mayzelis, and S. S. Melnik, *Random Finite-Valued Dynamical Systems: Additive Markov Chain Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[27] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1082–1090.
[28] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," in *Proc. ACM Conf. Comput. Supported Cooperat. Work*, Oct. 1994, pp. 175–186.
[29] G. Feng and X. Liang, "Review of collaborative filtering recommender," *Library Inf. Service*, vol. 55, no. 16, pp. 126–130, 2011.
[30] J. Wang, A. P. De Vries, and M. J. T. Reinders, "Unifying user-based and item-based collaborative filtering approaches by similarity fusion," in *Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Aug. 2006, pp. 501–508.
[31] L. Shi, "Reserch and implementation on weibo user-oriented content and friends recommendation algorithm," M.S. thesis, School Comput. Sci., Beijing Univ. Posts Telecommun., Beijing, China, 2014.
[32] Y. Zheng, "Tutorial on location-based social networks," in *Proc. 21st Int. Conf. World Wide Web*, 2012, p. 12.
[33] C. Song, Z. Qu, N. Blumm, and A. Barabasi, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
[34] J. J. Ying, E. H. C. Lu, W. N. Kuo, and V. S. Tseng, "Urban point-of-interest recommendation by mining user check-in behaviors," in *Proc. ACM SIGKDD Int. Workshop Urban Comput.*, Aug. 2012, pp. 63–70.
[35] M. Ye, P. F. Yin, W. Lee, and D. L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2011, pp. 325–334.
[36] R. Serfozo, *Basics of Applied Stochastic Processes*. Springer, 2009.
[37] A. Cheng, Y. Chen, Y. Huang, W. H. Hsu, and H. M. Liao, "Personalized travel recommendation by mining people attributes from community-contributed photos," in *Proc. 19th ACM Int. Conf. Multimedia*, Nov. 2011, pp. 83–92.

REN YUE-QIANG received the bachelor's degree from Changsha University, China, in 2018. He is currently pursuing the master's degree with the Tianjin Polytechnic University. His primary research interests include data mining and recommendation algorithm.

WANG ZE received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, in 1998 and 2001, respectively, and the Ph.D. degree in computer applications technology from Northeastern University, China, in 2004. Since December 2016, he has been a Professor with the School of CSSE, Tianjin Polytechnic University, China. His primary research interests include mobile computing, network security, and the Internet applications.

SUN XIAO-NA received the bachelor's and master's degrees from Tianjin Polytechnic University, China, in 2016 and 2019, respectively. Her primary research interests include machine learning and data mining.

SUN SHI-MIN received the M.S. and Ph.D. degrees in computer and information communication engineering from the Konkuk University of Korea, in 2009 and 2016, respectively. He is currently a Lecturer with Tianjin Polytechnic University, China. His research interests include software-defined networking, the future network architecture, cloud computing, edge computing, and network security. He is the Chair of the Natural Science Foundation of China and the Natural Science Foundation of Tianjin on future network architecture.