# A Markov Chain Position Prediction Model Based on Multidimensional Correction 

Sijia Chen ( ${ }^{\text {D }}$, Jian Zhang, Fanwei Meng (ㄹ), and Dini Wang (D)<br>School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China<br>Correspondence should be addressed to Fanwei Meng; mengfanwei@neuq.edu.cn

Received 2 November 2020; Revised 13 November 2020; Accepted 19 December 2020; Published 5 January 2021
Academic Editor: Wei Zhang
Copyright © 2021 Sijia Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.


#### Abstract

User location prediction in location-based social networks can predict the density of people flow well in terms of intelligent transportation, which can make corresponding adjustments in time to make traffic smooth, reduce fuel consumption, reduce greenhouse gas emissions, and help build a green cycle low-carbon transportation green system. This paper proposes a Markov chain position prediction model based on multidimensional correction (MDC-MCM). Firstly, extract corresponding information from the user's historical check-in position sequence as a position-position conversion map. Secondly, the influence of check-in period, space distance, and other factors on the position prediction is linearly weighted and merged with the position prediction of the $n$-order Markov chain to construct MDC-MCM. Finally, we conduct a comprehensive performance evaluation of MDC-MCM using the dataset collected from Brightkite. Experimental results show that compared with other advanced location prediction technologies, MDC-MCM achieves better location prediction results.


## 1. Introduction

With the development of the world's industrial economy, the rapid increase in population, and the unrestrained production and lifestyles, the world climate faces more serious problems. Greenhouse gas emissions are increasing, and the earth's ozone layer is suffering from unprecedented crises. Catastrophic climate changes have repeatedly appeared globally, which have seriously endangered the living environment and health and safety of human beings. The communication network and the positioning system are combined to form a new type of social network-locationbased social network [1]. In a location-based social network, people can share their location and location information at any time through communication devices, also known as sign-in. These data can be used for user location prediction, friend relationship prediction, and personal behavior patterns [2-4]. The user's location prediction is of great use in intelligent transportation. It can predict the density of people flow and make corresponding adjustments in time to make traffic smooth [5], reduce fuel consumption, reduce greenhouse gas emissions, and help build a green cycle low-
carbon transportation green system. In addition, it also plays an important role in smart cities and epidemiological communication research.

Currently, many methods of position prediction have emerged. Among them, Yuan et al. [6] explored the influence of time and space on location prediction in locationbased social networks. Ye et al. [7] used power law distribution to model spatial factors and combined user preferences and friend relationships to predict location. Cheng et al. [8] used a first-order Markov chain based on the influence of the most recently visited location on the next location and integrated the matrix decomposition method to predict the location. Based on the high-order influence of $n$ order weighted Markov chain, Zhang and Chow [9] combined time and space with friend relationship and popularity factors for location prediction.

In this paper, we adopt the $n$-order Markov chain [10] and then consider the period of check-in, space distance, friend relationship, and popularity of check-in points and propose a Markov chain position prediction model based on multidimensional correction (MDC-MCM), which realizes the position prediction for LBSNs.

In short, our contribution to this research work has three aspects.

Firstly, we link user location prediction in locationbased social networks with intelligent transportation to help build a green, circular, low-carbon transportation green system.
Secondly, Markov chain position prediction model based on multidimensional correction (MDC-MCM) comprehensively considers the check-in time period, spatial distance, friend relationship, and check-in point popularity. The dimensions considered are more comprehensive.
Finally, we evaluated the proposed location prediction method on the Brightkite dataset. The experimental results show that our proposed location prediction method has better prediction performance compared with other methods.
The rest of the paper is organized as follows. Section 2 describes the Markov chain position prediction model based on multidimensional correction in detail. In the third section, we will experiment with the proposed model in the Brightkite dataset to get the results and discuss further. Finally, in the fourth section, the conclusion is drawn and the future work arrangements are described.

## 2. MDC-MCM

2.1. LLTG Diagram. Figure [11]. A data structure composed of a set of vertices and a set of relations between vertices defined as Graph $=(V, E)$.

Out Degree [11]. The number of edges associated with a vertex is called a degree. In a directed graph, a vertex is the end of the arc and the number of arcs starting from the vertex.

Location to Location Transition Graph (LLTG graph, Location to Location Transition Graph). This contains a series of vertices $L$ and edges $E=L^{2}$. Each vertex $l_{i}\left(l_{i} \in L\right)$ represents a point of interest, and each vertex $l_{i}$ has an out-degree, denoted as $O$ Count $\left(l_{i}\right)$, and the transition frequency from $l_{i}$ to $l_{j}$ is denoted as $T$ Count $\left(l_{i}, l_{j}\right)$. For example, in Figure 1, the out-degree of location node $l_{1}$ is 8 , the out-degree of location node $l_{2}$ is 3 , the out-degree of location node $l_{3}$ is 7 , and the out-degree of location node $l_{4}$ is 0 .

It can be seen from Figure 1 that the LLTG graph describes the transfer frequency from one location node to another location node and the outgoing degree of each node.

Transition probability represents the probability of one location node to another location node, and the transition probability from $l_{i}$ to $l_{j}$ is recorded as $\operatorname{TP}\left(l_{i} \longrightarrow l_{j}\right)$. And, considering that the out-degree of the location node may be 0 , we assume that the transition probability of the out-degree of the location node is 1 :

$$
\operatorname{TP}\left(l_{i} \longrightarrow l_{j}\right)= \begin{cases}\frac{T \operatorname{Count}\left(l_{i} \longrightarrow l_{j}\right)}{O \operatorname{Count}\left(l_{i}\right)}, & O \operatorname{Count}\left(l_{i}\right) \neq 0  \tag{1}\\ 1, & O \operatorname{Count}\left(l_{i}\right)=0\end{cases}
$$

2.2. $n$-Order Markov Chain. Markov chain [12] is a sequence of random variables $X_{1}, X_{2}, X_{3}$, and so on. The range of these variables, the set of all their possible values, is called the state space. If the state corresponding to time $n$ is $X_{n}$, then $X_{n+1}$ is regarded as a function of $X_{1}, \ldots, X_{n}$, also known as an n-order Markov chain [10], which has n-order memory. The matrix composed of transition probabilities is the transition probability matrix.

Assuming that user $u$ has $m$ location nodes and is now at time $n$ and the location node is $l_{n}$, the transition probability matrix is as follows:

$$
R=\left[\begin{array}{ccc}
P_{11} & \ldots & P_{1 m}  \tag{2}\\
\ldots & \ldots & \ldots \\
P_{m 1} & \ldots & P_{m m}
\end{array}\right]
$$

Among them, $P_{i j}=\operatorname{TP}\left(l_{i} \longrightarrow l_{j}\right)$.
The probability distribution vector of the initial state is as follows: $P_{1}=[1,0, \ldots, 0]$. Then, the probability distribution vector of user $u$ going to each location node at time $n+1$ is as follows:

$$
\begin{equation*}
P 0_{u}^{n+1}=P_{1} \cdot R^{n} \tag{3}
\end{equation*}
$$

2.3. Time Zone When Signing in. Studies have shown that the user's sign-in behavior largely meets the regularity of time [13]. Therefore, analyzing the data from the perspective of time is essential to improve the accuracy of position prediction. We select the Brightkite dataset and make a map of the week distribution and hour distribution of user sign-in (Figures 2 and 3).

From Figure 2, it is found that the proportion of check-in times varies periodically with the week. The number of check-ins from Monday to Thursday is relatively even, the number of check-ins on Friday and Saturday has increased significantly, and the number of check-ins on Saturday is the highest, and the number of check-ins on Sunday and Monday to Thursday is similar.

From Figure 3, it is found that the proportion of the number of check-ins changes periodically with the hour. From 0:00 in the morning, the number of user check-ins showed a downward trend, until the lowest peak of check-ins appeared at about 10 am . As the number of check-ins


Figure 1: Example of position-position transition diagram (LLTG diagram).


Figure 2: Weekly distribution of user sign-in.


Figure 3: Hourly distribution of user check-in.
increased, the highest peak appeared at about 7 pm , after which the number of check-ins fluctuated within a small range. According to the law of change, a day is divided into three-time intervals: interval 1 , interval 2, and interval 3. Let $T=\{$ interval 1, interval 2, interval 3\}, then corresponding time range is $0: 00-10: 00,10: 00-19: 00$, and $19: 00-24: 00$.

Consider the week and time interval comprehensively to study user sign-in location prediction.

Define the probability $P_{u, l_{\text {I }}}^{h}$ of the user $u$ checking in at location $l_{i}$ in the time interval $h$ as

$$
\begin{equation*}
P_{u, l_{i}}^{h}=\frac{N_{u, l_{i}}^{h}}{\sum_{j=1}^{m} N_{u, l_{j}}^{h}} . \tag{4}
\end{equation*}
$$

Among them, $h$ is the element in the previously defined time interval $T, m$ is the size of the location set
$L=\left\{l_{1}, l_{2}, \ldots, l_{m}\right\}$, and $N_{\mu, l_{i}}^{h}$ indicates that the user $u$ checked in at the location $l_{i}$ in the time interval $h$ frequency.

Therefore, the check-in probability $P_{u, l_{i}}^{t}$ of the user $u$ at the location node $l_{i}$ on the $t$ day of the week can also be obtained:

$$
\begin{equation*}
P_{u, l_{i}}^{t}=\frac{M_{t}^{h} P_{u, l_{i}}^{h}}{M_{t}}, \quad h \in T, l_{i} \in L . \tag{5}
\end{equation*}
$$

Among them, $M_{t}^{h}$ is the number of check-ins in interval $h$ on the $t$ day of the week and $M_{t}$ is the total number of check-ins on the $t$ day of the week.

To simplify the calculation, the obtained probability is subjected to min-max normalization processing [14]:

$$
\begin{equation*}
P_{u, l_{i}}^{t^{*}}=\frac{P_{u, l_{i}}^{t}-\min \left\{P_{u, l_{1}}^{t}, P_{u, l_{2}}^{t}, \ldots, P_{u, l_{m}}^{t}\right\}}{\max \left\{P_{u, l_{1}}^{t}, P_{u, l_{2}}^{t}, \ldots, P_{u, l_{m}}^{t}\right\}-\min \left\{P_{u, l_{1}}^{t}, P_{u, l_{2}}^{t}, \ldots, P_{u, l_{m}}^{t}\right\}}, \quad 1 \leq i \leq m . \tag{6}
\end{equation*}
$$

Then, the probability distribution vector of user $u$ going to each location node at time $n+1$ is as follows:

$$
\begin{equation*}
P 1_{u}^{n+1}=\left[P_{u, l_{1}}^{t^{*}}, P_{u, l_{2}}^{t^{*}}, \ldots, P_{u, l_{m}}^{t^{*}}\right] . \tag{7}
\end{equation*}
$$

2.4. Spatial Distance. Since the spatial distances of the two consecutive check-in points are different, it is necessary to
estimate the distribution of the two consecutive check-in points with the spatial distance.

The sampling data of the space are collected from the check-in set $D$ as shown in the following:

$$
\begin{equation*}
D=\left\{\operatorname{Haversine}\left(l_{i}, l_{j}\right)\right\} . \tag{8}
\end{equation*}
$$

Among them, the Haversine distance formula [15] is as follows:

$$
\begin{equation*}
\operatorname{Haversine}\left(l_{j}, l_{j+1}\right)=2 r \cdot \arcsin \left(\sqrt{\sin ^{2}\left(\frac{\operatorname{lat}_{j+1}-\operatorname{lat}_{j}}{2}\right)+\cos \left(\operatorname{lat}_{j}\right) \cdot \cos \left(\operatorname{lat}_{j+1}\right) \cdot \sin ^{2}\left(\frac{\log _{j+1}-\log _{j}}{2}\right)}\right), \tag{9}
\end{equation*}
$$

where $r$ is the radius of the Earth, about 6371 km .
Assuming that the spatial distance $d$ between two consecutive check-in points approximately obeys the power law distribution [16], the probability density formula of the power law distribution is as follows:

$$
\begin{equation*}
f(d)=(\gamma-1)(d+1)^{-\gamma}, \quad d \geq 0, \gamma>1 \tag{10}
\end{equation*}
$$

According to the maximum likelihood estimation method [17], we can estimate from sample $D$

$$
\begin{equation*}
\gamma=1+|D|\left(\left[\sum_{d \in D} \ln (d+1)\right]\right)^{-1} . \tag{11}
\end{equation*}
$$

The Brightkite dataset is selected to plot the probability density and the spatial distance of two consecutive check-in points, as shown in Figure 4.

In Figure 4, we find that the spatial distance between the probability density and two consecutive check-in points is very similar to the estimated power law distribution. It shows that our hypothesis is reasonable and effective. The spatial distance $d$ of two consecutive check-in points can be regarded as obeying the power law distribution.

Assuming that user $u$ has $m$ location nodes and is now at time $n$ and the location node is $l_{n}$, the probability of going to each node is as follows:

$$
\begin{equation*}
F\left(d_{l_{n}, l_{i}}\right)=\int_{d_{l_{n} l_{i}}}^{+\infty} f(d) \mathrm{d} y=\left(d_{l_{n} l_{i}}+1\right)^{1-y}, \quad 1 \leq i \leq m . \tag{12}
\end{equation*}
$$

Among them, $l_{i}$ is the $i$-th sign-in point.
To simplify the calculation, the obtained probability is subjected to min-max normalization processing [14].

$$
\begin{equation*}
F^{*}\left(d_{l_{n} l_{i}}\right)=\frac{F\left(d_{l_{n} l_{i}}\right)-\min \left\{F\left(d_{l_{n}, l_{1}}\right), F\left(d_{l_{n}, l_{2}}\right), \ldots, F\left(d_{l_{n} l_{m}}\right)\right\}}{\max \left\{F\left(d_{l_{n} l_{1}}\right), F\left(d_{l_{n} l_{2}}\right), \ldots, F\left(d_{l_{n} l_{m}}\right)\right\}-\min \left\{F\left(d_{l_{n} l_{1}}\right), F\left(d_{l_{n} l_{2}}\right), \ldots, F\left(d_{l_{n} l_{m} l_{m}}\right)\right\}}, \quad 1 \leq i \leq m \tag{13}
\end{equation*}
$$



Figure 4: The spatial distance distribution of two consecutive check-in points.

Then, the probability distribution vector of user $u$ going to each location node at time $n+1$ is as follows:

$$
\begin{equation*}
P 2_{u}^{n+1}=\left[F^{*}\left(d_{l_{n}, l_{1}}\right), F^{*}\left(d_{l_{n}, l_{2}}\right), \ldots, F^{*}\left(d_{l_{n} l_{m}}\right)\right] \tag{14}
\end{equation*}
$$

2.5. Friendship. Based on the previous research [9], user sign-in points are related to friends. Different friends have different influences. In order to measure the influence of different friends, we have introduced the Jaccard coefficient to measure the similarity and difference between different friends.
2.5.1. Jaccard Coefficient. Jaccard coefficient [18] is widely used in the field of information retrieval. It is often used as an index to measure the similarity of two objects, that is, to judge the probability that a certain characteristic is shared by two objects. Here, a certain characteristic is defined as the number of common friends, that is, the number of common friends owned by two user accounts for the sum of the number of friends owned by two users. The formula is as follows:

$$
\begin{equation*}
S_{\text {jaccard }}=\frac{|\Gamma(i) \cap \Gamma(j)|}{|\Gamma(i) \cup \Gamma(j)|} \tag{15}
\end{equation*}
$$

Among them, $\Gamma(i)$ is the set of neighbors of user node $i$ and $\Gamma(j)$ is the set of neighbors of user node $j$. The larger the Jaccard coefficient value, the higher the similarity between friends and the closer the relationship.

Assuming that user $u$ has $m$ location nodes and $p$ friends and he is now at time $n$ and his location node is $l_{n}$, then the probability that a friend will influence user $u$ 's check-in at location node $l_{i}$ is as follows:

$$
\begin{equation*}
P_{\text {Jaccard }}\left(l_{i}\right)=\sum_{k=1}^{p} S_{\text {Jaccard }}(u, k) \cdot T_{k}\left(l_{i}\right), \quad 1 \leq i \leq m . \tag{16}
\end{equation*}
$$

Among them, $T_{k}\left(l_{i}\right)$ represents the check-in frequency of the $k$-th friend of the user $u$ at the location node $l_{i}$.

To simplify the calculation, the obtained influence probability is subjected to min-max normalization processing [14].

$$
\begin{equation*}
P_{\text {Jaccard }}^{t^{*}}\left(l_{i}\right)=\frac{F\left(l_{i}\right)-\min \left\{P_{\text {Jaccard }}\left(l_{1}\right), P_{\text {Jaccard }}\left(l_{2}\right), \ldots, P_{\text {Jaccard }}\left(l_{m}\right)\right\}}{\max \left\{P_{\text {Jaccard }}\left(l_{1}\right), P_{\text {Jaccard }}\left(l_{2}\right), \ldots, P_{\text {Jaccard }}\left(l_{m}\right)\right\}-\min \left\{P_{\text {Jaccard }}\left(l_{1}\right), P_{\text {Jaccard }}\left(l_{2}\right), \ldots, P_{\text {Jaccard }}\left(l_{m}\right)\right\}}, \quad 1 \leq i \leq m . \tag{17}
\end{equation*}
$$

Then, the probability distribution vector of user $u$ going to each location node at time $n+1$ is as follows:

$$
\begin{equation*}
P 3_{u}^{n+1}=\left[P_{\text {Jaccard }}^{t^{*}}\left(l_{1}\right), P_{\text {Jaccard }}^{t^{*}}\left(l_{2}\right), \ldots, P_{\text {Jaccard }}^{t^{*}}\left(l_{m}\right)\right] . \tag{18}
\end{equation*}
$$

2.6. Popularity of Check-In Points. The popularity of checkin points can greatly affect the user's prediction of the next check-in location. The popularity of the check-in point can be directly determined by the historical check-in frequency of the location $l_{i}$ of the user $u$.

Assuming that user $u$ has $m$ location nodes and is now at time $n$ and the location node is $l_{n}$, the probability of going to each node is as follows:

$$
\begin{equation*}
F\left(l_{i}\right)=T_{u}\left(l_{i}\right), \quad 1 \leq i \leq m . \tag{19}
\end{equation*}
$$

Among them, $T_{u}\left(l_{i}\right)$ represents the historical check-in frequency of the user $u$ at the check-in node $l_{i}$.

To simplify the calculation, the obtained probability is subjected to min-max normalization processing [14].

$$
\begin{equation*}
F^{*}\left(l_{i}\right)=\frac{F\left(l_{i}\right)-\min \left\{F\left(l_{1}\right), F\left(l_{2}\right), \ldots, F\left(l_{m}\right)\right\}}{\max \left\{F\left(l_{1}\right), F\left(l_{2}\right), \ldots, F\left(l_{m}\right)\right\}-\min \left\{F\left(l_{1}\right), F\left(l_{2}\right), \ldots, F\left(l_{m}\right)\right\}}, \quad 1 \leq i \leq m . \tag{20}
\end{equation*}
$$

Then, the probability distribution vector of user $u$ going to each location node at time $n+1$ is as follows:

$$
\begin{equation*}
P 4_{u}^{n+1}=\left[F^{*}\left(l_{1}\right), F^{*}\left(l_{2}\right), \ldots, F^{*}\left(l_{m}\right)\right] . \tag{21}
\end{equation*}
$$

The linear weighted fusion of the various predicted probabilities that affect the next check-in position proposed above is used to obtain a Markov chain position prediction model based on multidimensional correction (MDC-MCM). The probability distribution vector of each check-in point of user $u$ at time $n+1$ is as follows:

$$
\begin{align*}
P_{u}^{n+1}= & P 0_{u}^{n+1}+\mu_{1} \cdot P 1_{u}^{n+1}+\mu_{2} \cdot P 2_{u}^{n+1} \\
& +\mu_{3} \cdot P 3_{u}^{n+1}+\mu_{4} \cdot P 4_{u}^{n+1} \tag{22}
\end{align*}
$$

Among them, $\mu_{1}, \mu_{2}, \mu_{3}$, and $\mu_{4}$ are all correction coefficients.

## 3. Experiment

In this section, the proposed model is compared with the latest position prediction technology, and the accuracy and recall rates are obtained on the Brightkite dataset [2].
3.1. Brightkite Dataset. The Brightkite dataset is a dataset based on user sign-in data in the LBSN sign-in website. The data format for check-in in Brightkite dataset is <userid, check-in time, latitude, longitude, locationid $>$. Brightkite is the second-largest sign-in site after Foursquare. The statistics of the dataset are shown in Table 1.

In Table 1, we need to preprocess the data in Table 1 to ensure the quantity and quality of the data. In the preprocessing, to prevent the sparse data from affecting the experimental results, users with less than ten check-ins and points of interest with a total of fewer than ten check-ins are filtered out. According to the check-in time, the check-in data are divided into training set and test set. And, the first $80 \%$ of the check-in data are used as the training set, and the last $20 \%$ of the check-in data are used as the test set. In the experiment, the training set adopts the Markov chain position prediction model based on multidimensional correction to predict the test data.
3.2. Evaluation Technology. We will compare the Markov chain location prediction model (MDC-MCM) we built based on multidimensional corrections and previous location recommendation technologies, including the following:

STI. This method considers time and space factors, independently predicts the user's preference for location nodes in each time interval, and users are more inclined to visit nearby points of interest [6].
USG. This method uses comprehensive location prediction model spatial factors according to a power law distribution and combines user preferences and friend relationships [7].
FMC. This method is based on a first-order Markov chain, which uses the influence of the most recently
visited location on the next location and incorporates the matrix factorization method [8].
$A M C$. This method uses a sequence prediction algorithm based on an n-order weighted Markov chain, combined with a simple weight decay method, so that the recommendation results are more inclined to check-in to places that are closer [19].
LORE. This method uses a high-order sequential influence based on an n-order weighted Markov chain and combines time and space with friendly relations and popularity factors [9].
$M D C-M C M$. The MDC-MCM proposed in this paper is based on the high-order sequence influence of the n-order Markov chain and combines the check-in period, space distance, friend relationship, and checkin point popularity factors.
3.3. Performance Metrics. To evaluate the performance of each method, we selected two metrics, precision [20] and recall [20] as follows:

$$
\begin{equation*}
\text { precision }=\frac{1}{|U|} \sum_{u=1}^{|U|} \frac{H_{u}}{R_{u}} \text {, recall }=\frac{1}{|U|} \sum_{u=1}^{|U|} \frac{H_{u}}{C_{u}} . \tag{23}
\end{equation*}
$$

Among them, $|U|$ is the number of users to be predicted, $H_{u}$ is the predicted hit number of user $u, R_{u}$ is the number of location prediction sequences of user $u$, and $C_{u}$ is the set of locations visited by user $u$ in the test set.
3.4. Result. The number of next positions (top-k) for each prediction is set from 1 to 20 . Repeatedly adjust the correction coefficient in the training set, and finally get the current correction coefficient: $\mu_{1}=0.67, \mu_{2}=0.84$, $\mu_{3}=0.24$, and $\mu_{4}=0.13$. Better prediction results can be obtained in the test set, and draw precision and recall separately with other position prediction techniques curve. The results are shown in Figures 5 and 6.

### 3.5. Analysis. Here, we analyze the experimental results.

3.5.1. The Number of Check-In Points Recommended for Users Top-k. In Figures 5 and 6, it can be observed that as the number of recommended check-in point top-k increases, the accuracy rate gradually decreases and the recall rate gradually increases. This is also in line with expectations. As the number of recommended check-in point top- $k$ increases, if the location visited by the user is already in the recommended check-in point, it will change as the remaining recommended check-in points increase and the user will be at the recommended check-in point. The proportion of CM becomes lower, and the accuracy rate decreases; as the recommended check-in points increase, the more likely the place users visit is in the recommended check-in points, the greater the recall rate.

Table 1: Statistics of the Brightkite dataset.

| Index | Brightkite |
| :--- | :---: |
| Nodes | 58228 |
| Edges | 214078 |
| Nodes in largest SCC | $56739(0.974)$ |
| Edges in largest SCC | $212945(0.995)$ |
| Average clustering coefficient | 0.1723 |
| Number of triangles | 494728 |
| Fraction of closed triangles | 0.03979 |
| Diameter (longest shortest path) | 16 |
| 90-percentile effective diameter | 6 |
| Check-ins | $4,491,143$ |



Figure 5: Precision curve.


Figure 6: Recall curve.
3.5.2. The Effect of Different Factors on Recommendation Results. In Figures 5 and 6, through the prediction curve of the FMC method and the prediction curve of the STI method, it can be found that the time factor plays an important role in the position prediction. Through the prediction curve of the STI method and the prediction curve of the USG method, it can be found that the friend relationship plays an important role in location prediction. Through the prediction curve of the USG method and the prediction curve of the AMC method, it can be found that spatial distance plays an important role in position prediction. Through the prediction curve of the AMC method and the prediction curve of the LORE method, it can be found that the popularity of the check-in point plays an important role in the location prediction. MDC-MCM models the sequence influence based on the n-order Markov chain and considers the influence of check-in period, space distance, friend relationship, and check-in point popularity to ensure that MDC-MCM is superior to other location prediction algorithms. However, MDC-MCM uses an n-order Markov chain and has many correction parameters, which makes each run time very long; there are too many correction parameters, and parameter adjustment is cumbersome.

## 4. Conclusion

This paper proposes a Markov chain position prediction model based on multidimensional correction (MDC-MCM). First, MDC-MCM utilizes the high-order sequence influence based on the n -order Markov chain to consider all positions and transition probabilities in the user's check-in history. In addition, MDC-MCM combines the influence of check-in period, space distance, friendship, and popularity of checkin points. Finally, the experimental results on the Brightkite dataset show that the MDC-MCM position prediction works well. In the future, we will consider using the community as a unit to make predictions and then make predictions in the community to reduce the workload of computer operations. In addition, consider deploying the model on a distributed computing platform, which greatly shortens the running time and makes it easier to adjust the correction parameters.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was partly supported by the Fundamental Research Funds for the Central Universities (no. N182304010), the Natural Science Foundation of Liaoning Province (no. 20170520333), and the Natural Science Foundation of Hebei Province (no. F2019501012).

## References

[1] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," Geoinformatica, vol. 19, no. 3, pp. 525-565, 2015.
[2] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in Proceedings of the 17 th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 10821090, San Diego, CA, USA, August 2011.
[3] R. Trasarti, R. Guidotti, A. Monreale, and F. Giannotti, "MyWay: location prediction via mobility profiling," Information Systems, vol. 64, pp. 350-367, 2017.
[4] X. Zheng, J. Han, and A. Sun, "A survey of location prediction on twitter," IEEE Transactions on Knowledge \& Data Engineering, vol. 30, no. 9, p. 1, 2018.
[5] L. Nie, X. Wang, L. Wan, S. Yu, H. Song, and D. Jiang, "Network traffic prediction based on deep belief network and spatiotemporal compressive sensing in wireless mesh backbone networks," Wireless Communications and Mobile Computing, vol. 2018, Article ID 1260860, 10 pages, 2018.
[6] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. Magnenat-Thalmann, "Time-aware point-of-interest recommendation," in Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 363-372, Dublin, Ireland, July 2013.
[7] M. Ye, P. Yin, W. C. Lee et al., "Exploiting geographical influence for collaborative point-of-interest recommendation," in Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 325-334, Beijing, China, July 2011.
[8] C. Cheng, H. Yang, M. R. Lyu et al., "Where you like to go next: successive point-of-interest recommendation," in Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, Beijing, China, August 2013.
[9] J.-D. Zhang and C.-Y. Chow, "Spatiotemporal sequential influence modeling for location recommendations," ACM Transactions on Intelligent Systems and Technology, vol. 7, no. 1, pp. 1-25, 2015.
[10] J. D. Zhang, G. Ghinita, and C. Y. Chow, "Differentially private location recommendations in geosocial networks," in Proceedings of the 2014 IEEE 15th International Conference on Mobile Data Management, pp. 59-68, Brisbane, Australia, July 2014.
[11] H. N. Chen, Data Structure: Description in C++, People's Posts and Telecommunications, Beijing, China, 2nd edition, 2008.
[12] N. Eagle and A. S. Pentland, "Eigenbehaviors: identifying structure in routine," Behavioral Ecology and Sociobiology, vol. 63, no. 7, pp. 1057-1066, 2009.
[13] Z. Yao, Y. Fu, B. Liu, Y. Liu, and H. Xiang, "POI recommendation: a temporal matching between POI popularity and user regularity," in Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM), pp. 549-558, Barcelona, Spain, December 2016.
[14] X. H. Chen, "Normalization method and case processing of four-dimensional seismic data," Acta Petrolei Sinica, vol. 20, no. 6, pp. 22-26, 1999.
[15] J.-D. Zhang, J. Xu, and S. S. Liao, "Aggregating and sampling methods for processing GPS data streams for traffic state estimation," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 4, pp. 1629-1641, 2013.
[16] H. B. Hu and L. Wang, "A brief history of power law distribution," Journal of Physics, vol. 34, no. 12, pp. 889-896, 2005.
[17] X. J. Tao, Theory and Application of Nonlinear Maximum Likelihood Estimation for Random Signals, Southeast University, Nanjing, China, 1995.
[18] H. Pham, C. Shahabi, and Y. Liu, "EBM: an entropy-based model to infer social strength from spatiotemporal data," in Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data, pp. 265-276, New York, NY, USA, June 2013.
[19] J. D. Zhang, C. Y. Chow, and Y. Li, "Lore: exploiting sequential influence for location recommendations," in Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 103-112, New York, NY, USA, 2014.
[20] G. Karypis, "Evaluation of item-based top- $n$ recommendation algorithms," in Proceedings of the Tenth International Conference on Information and Knowledge Management, pp. 247-254, Atlanta Georgia, October 2001.

