Research and Implementation of Key Techniques for Indoor Movement Object Trajectory Prediction

Ping Zhang¹, Qianqian Jiang², Ruilin Zhang² and Bo Li²*

¹The Second Research Institute of CAAC, South No.17 Second Ring Road, 610041 Chengdu, Sichuan, P.R. China
²School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, No.2006 Xiyuan Ave, West Hi-Tech Zone, 611731 Chengdu, Sichuan, P.R. China
libo@uestc.edu.cn

Abstract. With the popularity of smart mobile terminals and Wi-Fi signals, people’s demand for indoor location services has also increased. However, in the indoor space, GPS positioning is inaccurate, and the Wi-Fi signal may also have signal instability even no signal in some areas. This paper proposes a prediction method based on improved HMM model combined with historical trajectory clustering. The experimental results on UJIIndoorLoc data set show that the predicting trajectories can greatly improve the real-time performance of location services and the proposed method has great improvement in accuracy and scalability comparing with another model.

1. Introduction

With the rapid development of wireless interconnect technology (Wi-Fi), indoor positioning technology and smart mobile terminals, Wi-Fi based indoor location services have been widely used. Although Wi-Fi technology can provide location information for mobile devices, if there are large indoor environments with a high population density, it is difficult to accurately obtain the position of the moving object, so it is necessary to predict the position to which the moving object may move in the future. In order to satisfy the user experience, it is necessary to improve the real-time positioning of indoor mobile object prediction. The main schemes are as follows:

(1) establishing an efficient index for trajectory data of indoor moving objects in order to query quickly [1];
(2) combining the features of indoor space model, the trajectory data is projected and converted, and then compressed. The aim is to save storage space [2].

Some scholars use spatial density clustering such as DBSCAN algorithm [3] and the STING algorithm based on grid analysis [4] to excavate the activity characteristics of moving objects according to the historical trajectory of mobile objects; Song et al. proposed a state-based moving object model [5], using Markov model performs state changes. Shaojie Qiao segmented the trajectory by using segmented segment information as the hidden state of the Markov model [6], and then performed position prediction. At the same time, he proposed improved parameters on the basis of the original Markov model to adapt to trajectory prediction methods. However, these methods are based on the classical HMM model, the classical HMM model method for the location prediction of indoor moving objects when there is a state of stay may lead to prediction failure.
Mean time, the indoor space contains a large number of spatial entity elements such as walkway elements, corridors, and elevators [7]. Therefore, it is necessary to construct an efficient model to express the interior space, which is the basis for improving all related technologies of indoor positioning. At present, methods for modelling indoor space mainly include object-based feature model [8], geometric model method [9] and symbol model method [10]. This paper mainly uses the symbol model to represent the indoor space model. Therefore, the indoor space model needs to be gridded.

In view of the shortcomings of the above methods, this paper proposes a prediction method for indoor moving trajectory data based on HMM model to achieve trajectory prediction of indoor moving objects. The grid track time series is formed and based on the matching moving trajectory data of Wi-Fi signal. The historical track data is mined and analyzed, and the trajectory location point is predicted with HMM model, and then the trajectory path sequence is compressed. This prediction can solve the problem of prediction failure.

The main work of this paper is as follows:
Firstly, the space model is divided into meshes, and then the trajectory data is projected on the model to generate the grid trajectory sequence according to the spatial model. The generated grid locus sequence is pre-processed and the historical grid locus eigenvalues are extracted and clustered to form an important location cluster database.
Secondly, the HMM model is used to predict the grid trajectory sequence according to the clustering database using the clustered cluster object as the hidden state of HMM. At the same time, the HMM model is used to generate the most probable hidden state sequence corresponding to the grid time series, in other words, the grid trajectory sequence represented by the cluster sequence, the grid trajectory sequence is compressed, and the compressed code table is constructed. Optimize the trajectory database of moving objects.
Finally, a simulation platform based on Canvas and NODE is established to verify the accuracy of trajectory location prediction and compression efficiency, which provides a new way of thinking and method for indoor trajectory prediction.

2. Framework
Prediction model based on HMM is divided into five modules, as shown in Figure 1.

![Figure 1. Trajectory prediction model](image)

Data acquisition collects trajectory data of users by client devices. The spatial model is based on the historical trajectory data to calculate the unit edge length and the grid division. Data pre-processing module filter and formats data which projected by the grid space model and provides reasonable data for trajectory clustering. In trajectory clustering module, we identify trajectory clusters based on the historical data of users to prepare for trajectory prediction model based on HMM. Trajectory prediction module predicts the next trajectory according to the user’s current location. In this paper, the classical HMM model is improved by combining the state residence time, and the problem of predicting failure in the classical HMM model is solved. This paper mainly focuses on trajectory data in an indoor environment where the signal is unstable pre-processing, trajectory clustering, and trajectory prediction algorithms.
3. Pre-processing and clustering

3.1. Definition
In order to make a better description, the related definitions are given as follows before data processing.

Definition 1 Moving locus. Describes a single-track point of an indoor moving object whose data format can be represented by an object (1).

\[ p = \{ p \mid p = (x, y, \text{time}, \text{angle}, \text{level}) \} \]  

In this formula, \((x, y)\) is latitude and longitude of the trajectory point; \(\text{angle}\) is deflection Angle; \(\text{Level}\) is barometer value; \(\text{time}\) is the time stamp of the collection time is the Unix timestamp.

Definition 2 Trajectory sequence. A sequence composed of a single-track point in chronological order.

\[ \text{originalTrail} = \{ p_1, ..., p_n \} \mid \forall p_i, p_j (\exists i < j, p_i, \text{time} < p_j, \text{time}) \]  

In this formula, \(p\) is a single-track point, \(N\) is the number of trace points contained in a trajectory.

Definition 3 Unit length. Based on the historical trajectory data, the shortest distance between two adjacent non-repeat locus points in the interior space.

Definition 4 grid sequence. According to the length of the unit, the interior space model is meshed to generate the grid interior space model, and the points in a grid can be represented by its centre. For the sake of simplicity, take a square to represent a grid.

According to the grid model, the trajectory data is projected and the grid sequence (3) is generated.

\[ \text{gridSequence} = \{ g_1, g_2, ..., g_m \}, g_i = \{ id, \text{stoptime}, \text{angle}, \text{level} \} \]  

In this formula, \(id\) is grid number. \(\text{Stop time}\) is the duration of stay; \(\text{angle}\) is interval deviation angle; \(\text{level}\) is floor.

Definition 4 stop time. If a grid sequence is given such as \(\ldots, g_i, g_{i+1}, g_{i+2}, ..., g_j, ..., \) \(g_i\) and \(g_j\) of them have the same grid number, different time of grid track points, \(g_{j}, g_{i+1}\) is a grid of different grid number track points, time Settings, such as type (4).

\[ \text{stoptime} = \begin{cases} 0, & g_{i}[id] \neq g_{i+1}[id] \wedge g_{i+1}[time] - g_{i}[time] < \delta, \\ g_{i+1}[time] - g_{i}[time], & g_{i+1}[id] = g_{i}[id] \end{cases} \]  

In this formula, \(\delta\) is the time threshold. When the time interval between the continuous mesh point is greater than that, it is considered that the grid point has a stop state.

3.2. Data pre-processing
Users’ location information can be depicted with grid sequence. In this paper, trajectory data is pre-treated according to the above definition.

Step 1 according to the measurement value of the internal barometer of the MEMS inertial sensor and air pressure formula, the floor of the track point is calculated. The angle of deflection is obtained through the angle between two trajectory points. Traverse the generated grid sequence, calculate the dwell time of grid point according to definition 4.

Step 2 the generated grid trajectory sequences are complemented by redundancy and grid trajectory sequences.
Step3 the angle in the grid trajectory sequence formats. The Angle interval is divided into eight sub-intervals that do not intersect. The statistics are mapped to the number of angles of each sub-interval and the angle information object is generated.

3.3. Trajectory clustering
Before the trajectory of clustering, this paper first presents the following definition

Definition 5 corner. An angle that belongs to the range of $\left[\frac{\pi}{4}, \frac{7\pi}{4}\right]$

Definition 6. important location. The grid point meets the following condition:

(1) Stop time is greater than time threshold.
(2) The number of corners is greater than threshold.
(3) The number of track points through the grid point is greater than threshold.

In this paper, based on the definition of important points, the historical trajectory data is clustered by using DBSCAN algorithm, and finally different clusters are formed.

4. Trajectory prediction
In this paper, the trajectories were predicted by combining trajectory clustering and HMM model. Before trajectory prediction, this paper first presents the following definition.

Definition 7 Hidden states. In this paper, the trace cluster is used as a hidden state.

Definition 8 Observed sequence. The Observed sequence is grid Sequence (Def 4)

Definition 9 Initial state probability matrix.

\[ \lambda = [0, 0, 1, \frac{1}{k}, 0, \frac{1}{k}, ..., 0]^T. \quad (5) \]

Definition 10 Initial state transfer probability matrix A.

\[ A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{12} & \cdots & \cdots & a_{12} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}, \quad \sum_{j=1}^{m} a_{ij} = 1 \quad (6) \]

\[ a_s = \begin{bmatrix} \beta^{(ij)}_u \\ \beta^{(ij)}_s \\ \beta^{(ij)}_{\text{dis}} \end{bmatrix} \begin{bmatrix} \omega_u \\ \omega_s \\ \omega_{\text{dis}} \end{bmatrix}. \quad (7) \]

In this formula, \( a_s \) is the state transfer probability of hidden state \( C_i \) and \( C_j \); \( \beta^{(ij)}_a \) is the characteristic value of hidden state \( C_i \) and \( C_j \); \( \beta^{(ij)}_s \) is the state transfer probability of hidden state \( C_i \) and \( C_j \); \( \beta^{(ij)}_{\text{dis}} \) is the distance of \( C_i \) and \( C_j \); \( \omega \) is the weight of the three eigenvalues.

Definition 11 confusion matrix B. The confusion matrix B is like equation (8), which is the transfer probability matrix of hidden state and grid sequence, \( m \) is the number of hidden states, \( n \) is the state observation length, \( b_{ij} \) is the state transfer probability of the observed sequence and the hidden state, \( b_{ij} \) is derived from \( b_{ij}' \).
In this formula, \( b_{ij} \) is the state transfer probability of hidden state \( C_i \) and observed sequence \( g_j \); \( \gamma_a^{(ij)} \) is the normalized angular eigenvalues; \( \gamma_{dis}^{(ij)} \) is the distance of \( C_i \) and \( g_j \); \( w_a \) and \( w_{dis} \) are the weight of the two eigenvalues.

In this paper, the moving object mesh sequence is segmented by sliding window. Because of the moving objects movement patterns of randomness, can be thought of in the mobile mode, every point and turning point is the beginning of a new model, the input to the HMM model, which accords with the HMM model of short-term memory, and can avoid the excessive fitting model training. Based on the segmented processing of the mobile object grid sequence, the specific trajectory prediction steps are as follows.

Step 1 Segmenting the grid track points based on important points. Step 2 The HMM initial model is calculated according to definitions 7 to 11. Step 3 use the forward-backward algorithm and the Baum-Welch algorithm training model. Because there may be multiple consecutive grid sequence trajectories projected into the same cluster, generating successive cluster sequences. If the classical HMM definition of the state transfer matrix is adopted, the probability of the hidden state rotation is zero, which leads to the failure of prediction.

In the training model phase, when the rotation probability of matrix A is 0, use a linear smoothing method as equation (10), (11).

\[
\varepsilon_a = \text{Logistic}(\text{stoptime}(C_i)), 0 < \varepsilon_a \leq \max(a_{t,z})
\]  

\[
\text{stoptime}(C_i) = \max_{1 \leq s \leq k}(\text{stoptime}(g_{s}))
\]

In those formulas, \( \varepsilon_a \) is the value of the logarithmic trajectory of the maximum of stop time in the cluster. \( a_{t,z} \) is the value of row \( t \), column \( z \) in the state transition matrix \( A \).

Step 4 Based on the HMM model trained in Step 3, the Viterbi algorithm is used for prediction. Step 5 Based on the implicit state sequence group returned in Step 4, the current shift Angle is calculated based on the mobile terminal MEMS inertial navigation unit, and the Angle of the deflection Angle is standardized. Gets the most likely implicit state in this Angle interval, returning the nearest grid point in the cluster that is the most likely implicit state representation to the current grid location.

5. Experimental results and analysis

The experiment adopts the data-set that public data provided by UJIIndoorLoc. According to the data-set, the grid side length is 1m. The error threshold accuracy calculation method defined in this section is calculated as follows.

Definition 12 The acceptance value of the prediction position \( H^{(i)} \). When the prediction position is less than the actual position, the European distance (\( \text{dis}(i) \)) is less than the error threshold and the value is 1, otherwise it is 0.
\[ H(i) = \begin{cases} 1, & \text{dis}(i) \leq \delta_{\text{mis}} \\ 0, & \text{dis}(i) > \delta_{\text{mis}} \end{cases} \]  

(12)

Definition 13 Error threshold accuracy (AC). Denotes the accuracy of the prediction position within the acceptance range of the pre-set error threshold.

\[ AC = \frac{\sum_{i=1}^{N} H(i)}{N} \]  

(13)

In this formula, \( N \) is the predicted points.

Based on the offline and Wi-Fi signal environment, we compared the improved HMM model, the classical HMM model, the probability method and the HMM (HMM-KF) prediction algorithm based on Kalman filter. Based on the method of calculating prediction accuracy of 13, we compare and test three algorithms with different prediction error thresholds. Figure 2 shows the test results. When the improved HMM algorithm is below 1.5m in error value, it is 20% higher than the HMM-KF, and is more than 60% higher than the probability method. In addition, the improved HMM algorithm is more accurate than the other two algorithms.

![Figure 2. The comparison of prediction accuracy with different prediction error thresholds.](image)

6. Conclusion

This paper presents a scheme to analyse the indoor moving trajectory data, build the interior space, and predict trajectories based on trajectory cluster. The prediction model improves the classical HMM model through the time-dependent characteristic value of historical trajectory. Experiment shows that the method proposed has great improvement in accuracy and scalability as compared traditional methods. Besides, the scheme has a high practical value with its flexible application.

7. References

[1] Dequn Z, Chuanjie X and Tao P 2017 Semantic cell oriented indoor moving objects index based on Hbase *Journal of Geo-Information Science*

[2] Hao L, Shukui Z and Penghui S 2018 Trajectory compression algorithm for adaptive parameters *Computer application research* 35 685-688,716

[3] Ester M, Kriegel H P and Xu X 1996 A density-based algorithm for discovering clusters in large spatial databases with noise *International Conference on Knowledge Discovery & Data Mining* 1996 226-231
Acknowledgments
This work is supported by Sichuan Science & Technology Program under Grant 2017GZ0034 & 2018GZ0077 & 2018GZ0069.