Journal of Information Hiding and Privacy Protection
DOI:10.32604/jihpp.2020.016243
Article

Tech Science Press

A Location Prediction Method Based on GA-LSTM Networks and Associated Movement Behavior Information

Xingxing Cao[1], Liming Jiang[1,*], Xiaoliang Wang[1] and Frank Jiang[2]

[1]Hunan University of Science and Technology, Xiangtan, 411201, China
[2]Deakin University, Geelong, 3216, Australia
[*]Corresponding Author: Liming Jiang. Email: njustjlm@163.com

Received: 20 December 2020; Accepted: 01 January 2021

Abstract: Due to the lack of consideration of movement behavior information other than time and location perception in current location prediction methods, the movement characteristics of trajectory data cannot be well expressed, which in turn affects the accuracy of the prediction results. First, a new trajectory data expression method by associating the movement behavior information is given. The pre-association method is used to model the movement behavior information according to the individual movement behavior features and the group movement behavior features extracted from the trajectory sequence and the region. The movement behavior features based on pre-association may not always be the best for the prediction model. Therefore, through association analysis and importance analysis, the final association feature is selected from the pre-association features. The trajectory data is input into the LSTM networks after associated features and genetic algorithm (GA) is used to optimize the combination of the length of time window and the number of hidden layer nodes. The experimental results show that compared with the original trajectory data, the trajectory data associated with the movement behavior information helps to improve the accuracy of location prediction.

Keywords: Location prediction; information association; feature selection; GA-LSTM

1 Introduction

With the development of GPS and the prosperity of the taxi industry, a large amount of trajectory data is generated from moving vehicles every day. The trajectory data not only reflects the driving path of the vehicle, but also reflects the behavior of residents and urban traffic characteristics [1]. The application research of GPS trajectory data has attracted the attention of academia and industry. The main research directions are location-based services (LBS) [2] and intelligent transportation (ITS) [3]. Location prediction is the core and underlying support of LBS. Predicting the behavior of vehicles and users through trajectory features can provide more accurate and professional services [4]. Therefore, how to effectively and accurately predict the next location or target location has become a hot issue in the research field of location prediction.

In recent years, there have been many research results in location prediction. Reference [5] proposed a destination prediction method based on frequent pattern mining. However, frequent pattern mining is only suitable for specific situations, and the maintenance of the decision tree is complicated, and it will take a lot of time to process a large amount of trajectory data. The Markov-based location prediction method has also made remarkable achievements. Reference [6] predicted the destination of the taxi through the low-order Markov method, but the Markov prediction method cannot solve the long-term dependence of a large amount of trajectory data. The emergence of Recurrent Neural Networks (RNN) has solved the above problems and has become a general model in recent years. For time series data, the RNN model can capture the correlation between data [7], so it has become a general model for time series data prediction. However, RNN cannot solve the problem of long-term dependence. The LSTM model is a variant of RNN, which solves the long-term dependency problem [8–10]. For time series data, each record is a context sequence, so time and sequence are especially important in trajectory data. The time window is introduced in RNN to alleviate the over-fitting problem and enhance the correlation between time series data. For the length of the time window, too small or too large will have a great impact on the prediction results. Commonly used methods always use time interval mode, average value or empirical value to determine the length of the time window [11]. However, the above methods are suitable for data sets with good data quality and time distribution, but not for data sets with uneven time intervals. Reference [12] optimized the window size in the LSTM model by genetic algorithm. At the same time, network parameters such as hidden node numbers can also be optimized by GA [13,14].

In addition, trajectory data preprocessing can improve data quality, but data processing should not only focus on improving data quality, but also pay attention to improving data expression ability through trajectory data analysis and mining [15,16]. For trajectory data, it is important to express the movement behavior of the trajectory. Reference [17] associated the individual movement behavior information of the vehicle with the original data. In the traffic environment, a certain amount of individual movement behaviors will lead to group movement behaviors, and then affect individual movement behaviors. Therefore, vehicle movement behavior should not be limited to individuals. References [18,19] considered the scene features and traffic environment of the trajectory. The above two methods both improve the prediction accuracy by adding associated information, but the choice of associated information is subjective, and the effectiveness of the added information is not analyzed. In trajectory expression, it is not that the more relevant information contained the better. On the contrary, high-correlation features and low-contribution features will cause model complexity and increase training time. Therefore, when adding relevant information, their importance and relevance should be considered.

In order to solve the above problems, First, we pre-correlates the movement behavior features of vehicles from the individual movement behavior and group movement behavior, and obtains the behavior features values through statistics, calculation, visualization and other methods. Second, analyzes the importance and relevance of pre-associated movement behavior features to remove redundant and low contribution features, and then obtains the final association features. Third, the original data is incorporated with the association features and put into LSTM, and GA was used to optimize the length of window and the number of hidden layer node. Finally, the experiment is designed to verify the influence of trajectory data associated with movement behavior information on prediction accuracy.

2 Trajectory Expression Based on Associating Movement Behavior Information

2.1 Expression of Trajectory

The data set used in this paper contains the GPS trajectories of 10,357 taxis in Beijing from February 2 to February 8, 2008. The data set includes vehicle's time and location information. Tab. 1 shows a sample of track points. A track point contains attributes as follows: 1) Taxi ID; 2) date time; 3) longitude; 4) latitude.

The trajectory can be expressed as: $\mathrm{Tra1} = (\mathrm{tra1_i} | \mathrm{i} = 1,2,\cdots,\mathrm{n}), \mathrm{tra1_i} = (\mathrm{t_i}, \mathrm{lon_i}, \mathrm{lat_i})$, where $\mathrm{t_{i,}lon_i}, \mathrm{lat_i}$ represent the time, longitude and latitude in $\mathrm{i} - \mathrm{th}$ trajectory point $\mathrm{tra1_i}$.

Table 1: Example of original trajectory data

id	time	longitude	latitude
10	2008-02-02 16:31:43	116.39407	39.84887
10	2008-02-02 16:32:50	116.39412	39.84417
10	2008-02-02 16:34:35	116.39452	39.83635

2.2 Information Pre-Association Based on Movement Behavior

The movement behavior of vehicles is composed of individual movement behavior and group movement behavior. The individual movement behavior is the movement behavior of a single moving object, such as velocity and direction of a single vehicle. The group movement behavior shows the aggregation characteristics of a large number of moving objects in the movement law and trend and its research object is region. In this section we extract the moving behavior features of vehicles from individual and group movement features and show the calculate method.

2.2.1 Individual Movement Behavior

- Velocity feature

The velocity of the vehicle is different every moment. The next location is related to the speed of the previous moment, and the velocity can reflect the driving state of the vehicle. The velocity of the $\mathrm{i-th}$ trajectory point $\mathrm{V_i}$ is calculated as follows:

$$V_i = \frac{2*R\ arcsin\sqrt{sin^2\left(\frac{(lon_i-lon_{i-1})}{2}\right)+cos(lat_{i-1})*cos(lat_i)*sin^2\left(\frac{(lat_i-lat_{i-1})}{2}\right)}}{t_i-t_{i-1}} \tag{1}$$

- Direction feature

The direction of the vehicle in driving is dynamic, and the direction can reflect the dynamic change trend .The direction is the angle between the two locations and the north direction. The velocity of the $\mathrm{i-th}$ trajectory point $\mathrm{D_i = 1}$ is calculated as follows:

$$D_i = arctan\left(log\left(tan(lat_i/2+\pi/4)/tan(lat_{i-1}/2+\pi/4)\right), mod(|lon_{i-1}-lon_{i-1}|, 180)\right) \tag{2}$$

2.2.2 Group Movement Behavior

- Traffic Rush feature

During rush hours, the vehicle will slow down and may even remain stationary. Therefore, the characteristics of traffic rush is a great significance to predict the next location. Reference [16] analyzed the peak period characteristics based on Beijing taxi data. The results are as follows: On weekdays, the peak traffic hours are 7:00-10:00 and 17:00-20:00. On weekends, peak traffic hours are 13:00-15:00. The traffic rush feature value of the $\mathrm{i-th}$ trajectory point $\mathrm{T\_R_i = 1}$, if $\mathrm{t_i}$ in rush hours. Otherwise, $\mathrm{T\_R_i = 0}$.

- Grid Area feature

In order to study the traffic characteristics, this paper adopts the grid division method for analysis. Divide the area enclosed by the track points into a 500 m grid. The location of the $\mathrm{i-th}$ trajectory point will be converted into grid area $\mathrm{G\_id_i}$. On the one hand, point movement can be converted into regional movement; on the other hand, regional characteristics can be analyzed from group vehicles.

- Grid Velocity feature

The grid velocity value of the $\mathrm{i-th}$ trajectory point $\mathrm{G\_V_i}$ is a macro expression of $\mathrm{V_i}$, which can reflect the traffic environment where the vehicle is located. When $\mathrm{t_i}$ belongs to the peak period, the $\mathrm{G\_V_i}$ is the average velocity of m trajectory points entered into the grid $\mathrm{G\_id_i}$ in rush hour, otherwise, the average velocity is trajectories' velocity in the off-peak period.

- Grid Congestion feature

According to the classification standard of urban traffic congestion by the Ministry of Public Security, the grid congestion feature value of the $\mathrm{i-th}$ trajectory point $\mathrm{G\_C_i}$ is calculated as Eq. (3):

$$G\_C_i = \begin{cases} 1, \cdots\cdots\cdots\cdots\cdots G\_V_i \le 20km/h \\ 2, \cdots\cdots\cdots\cdots\cdots G\_V_i \le 30km/h \\ 3, \cdots\cdots\cdots\cdots\cdots G\_V_i \ge 20km/h \end{cases} \tag{3}$$

After adding pre-association features of individual and group movement behavior information, the trajectory points will contain more state information. The trajectory expression is changed as：$\mathrm{Tra2} = (\mathrm{tra2}_i | i = 1,2,\cdots,n), \mathrm{tra2}_i = \{t_i, \mathrm{lon}_i, \mathrm{lat}_i, V_i, D_i, T\_R_i, G\_id_i, G\_V_i, G\_C_i\}$. The example is shown in Tab. 2.

Table 2: Example of trajectory data after pre-association

Time	Longitude	Latitude	Velocity	Direction	Traffic-rush	Grid- area	Grid-velocity	Grid-congestion
2008-02-02 16:31:43	116.39407	39.84887	18.361	176.387	0	198414	24.243	2
2008-02-02 16:32:50	116.39412	39.84417	28.113	179.532	0	197448	23.258	2
2008-02-02 16:34:35	116.39452	39.83635	29.869	177.750	0	195516	29.845	2

2.3 Feature Selection

After adding pre-associated features, the trajectory data dimension will be expanded to 9. However, the current features are not necessarily optimal expression for location prediction. This paper uses random forest algorithm to estimate the importance of features. The basic idea is to rearrange the order of eigenvalues and observe how much accuracy is reduced. The correlation coefficient is calculated by Eq. (4). The feature correlation degree is obtained from Tab. 3. Based on the above results, low contribution and high correlation features will be removed, and the final correlation feature will be selected. The results are as follows:

$$\rho_{X,Y}=\frac{COV(X,Y)}{\delta_X\delta_Y}=\frac{E(XY)-E(X)E(Y)}{\sqrt{E(X^2)-E^2(X)}\sqrt{E(Y^2)-E^2(Y)}} \tag{4}$$

Table 3: Correlation degree table

Coefficient	Correlation degree
0.8–1.0	pole strength
0.6–0.8	strong
0.4–0.6	moderate
0.2–0.4	weak
0.0–0.2	pianissimo

According to Fig. 1, we can see that the most important for location prediction is location information. The velocity feature's contribution to the location prediction is the lowest, which is marked as low contribution. The reason may be that the time interval is large, the average speed cannot express the vehicle state, and the relative grid speed can better express it. In Fig. 2, there is a strong correlation between grid congestion and grid velocity. With reference to their importance evaluation, grid velocity has a greater contribution to prediction accuracy, so grid congestion features are removed. Therefore, seven features are selected: time, longitude, latitude, direction, traffic rush, grid area and grid velocity .After feature selection, the final trajectory of this model is expressed as：$\mathrm{Tra3} = (\mathrm{tra3}_i | i = 1,2,\cdots,n), \mathrm{tra3}_i = \{t_i, \mathrm{lon}_i, \mathrm{lat}_i, D_i, T\_R_i, G\_id_i, G\_V_i\}$. The example is shown in Tab. 4.

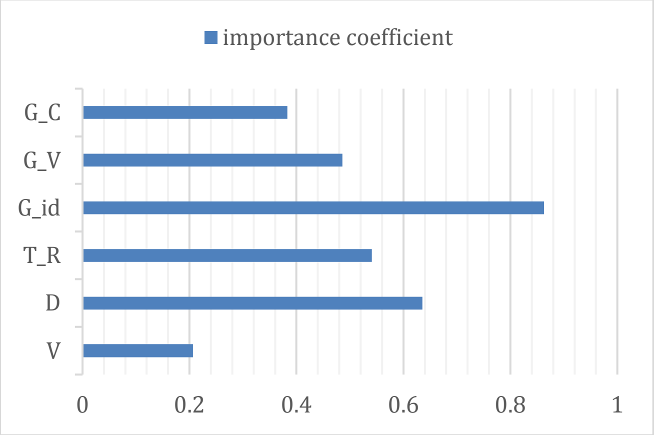
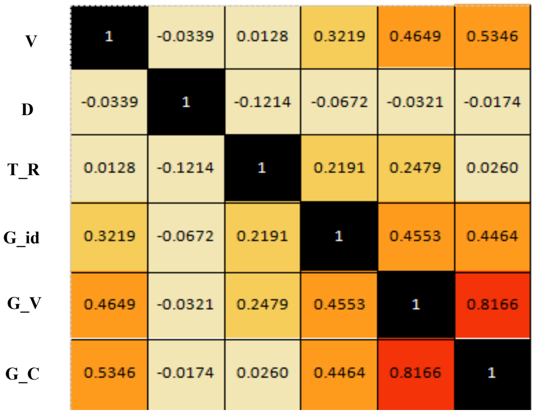


Figure 1: Importance coefficient of features



V	1	-0.0339	0.0128	0.3219	0.4649	0.5346
D	-0.0339	1	-0.1214	-0.0672	-0.0321	-0.0174
T_R	0.0128	-0.1214	1	0.2191	0.2479	0.0260
G_id	0.3219	-0.0672	0.2191	1	0.4553	0.4464
G_V	0.4649	-0.0321	0.2479	0.4553	1	0.8166
G_C	0.5346	-0.0174	0.0260	0.4464	0.8166	1

Figure 2: Correlation coefficient of features

Table 4: Example of trajectory data after feature selection

Time	Longitude	Latitude	Direction	Traffic rush	Grid area	Grid velocity
2008-02-02 16:31:43	116.39407	39.84887	176.387	0	198414	24.243
2008-02-02 16:32:50	116.39412	39.84417	179.532	0	197448	23.258
2008-02-02 16:34:35	116.39452	39.83635	177.750	0	195516	29.845

3 Next Location Prediction

The overall framework of GA-LSTM is composed of data processing module, optimization algorithm module and prediction module, as shown in Fig. 3. The data set is standardized and classified into training data set and test data set. Put the training data set into the LSTM model optimized by GA to train the network, and finally put the test data set into the network to predict the next location.

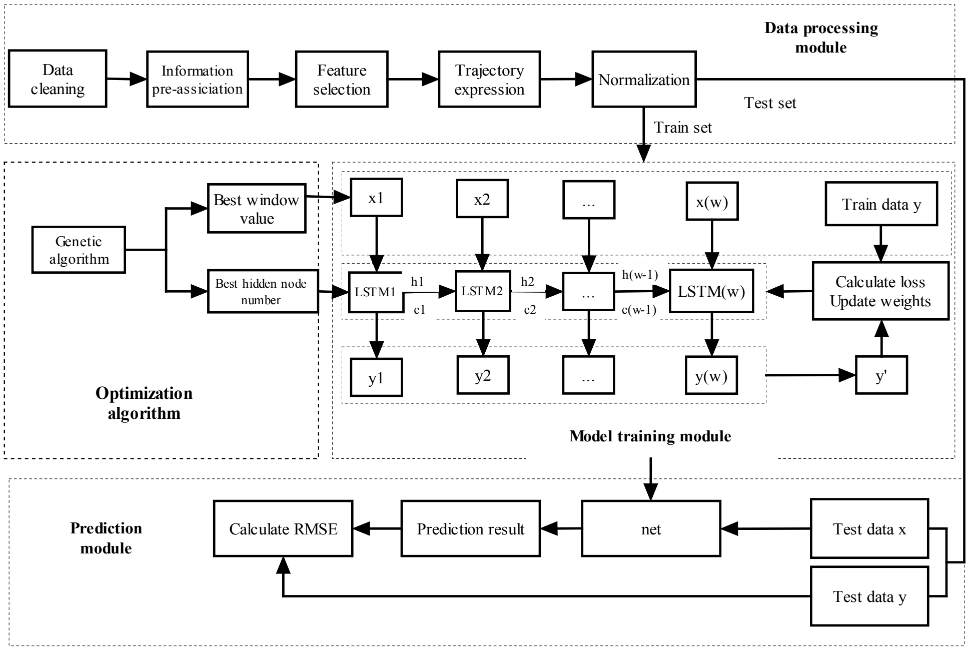


Figure 3: Location prediction framework

3.1 LSTM Networks

The difference between the LSTM model and the original RNN model is that it contains a gate structure, which consists of a forget gate, an input gate and an output gate. The structure of LSTM is shown in Fig. 4. According to the structure of the gate, choose the information to forget and remember. The calculation method is shown in the Eqs. (5) and (6).

At current time t, LSTM has three inputs: Input value at current time $\mathrm{x_t}$、LSTM output value at last time $\mathrm{h_{t-1}}$、Cell state at last moment $\mathrm{C_{t-1}}$; LSTM has two outputs：LSTM output value at current time $\mathrm{h_t}$、Cell state at current time $\mathrm{C_t}$.

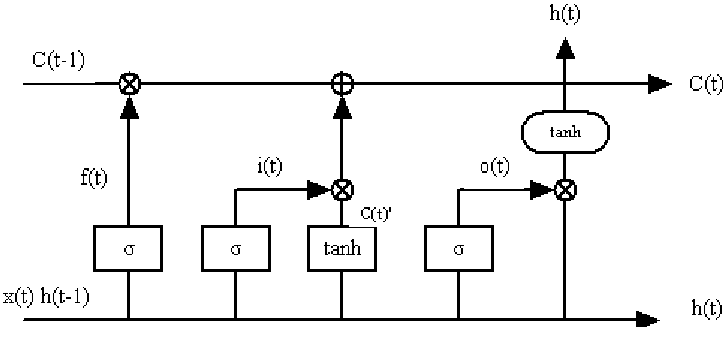


Figure 4: LSTM unit structure

$$f_t = \sigma\left(w_f.[h_{t-1}, x_t] + b_f\right), i_t = \sigma(w_i.[h_{t-1}, x_t] + b_i), o_t = \sigma(w_o.[h_{t-1}, x_t] + b_o) \quad (5)$$

$$C'_t = \tanh(w_c.[h_{t-1}, x_t] + b_c), C_t = f_t * C_{t-1} + i_t * C'_t, h_t = o_t * \tanh(C_t) \quad (6)$$

$\mathrm{w_f}$, $\mathrm{w_i}$, $\mathrm{w_o}$ are weight matrices for the corresponding inputs of the network activation functions σ. In this paper, we use square loss function given by the following Eq. (7), Where y is observed value, $\mathrm{y'}$ is predict value and $\mathrm{n_1}$ is the number of all predicted values.

$$loss = \sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1}(y_i - y'_i)^2} \tag{7}$$

3.2 GA Based LSTM Optimization

Genetic algorithm originated from the computer simulation of biological system. It is a stochastic global search and optimization method developed by imitating the evolution mechanism of natural organisms, and is often used to solve combinatorial optimization problems. In this paper, we select GA to obtain the optimal combination of window size w and hidden layer node h. Specific algorithm is shown in Algorithm 1.

Algorithm 1: GA-LSTM

```
Input: populationSize, max generation(maxgen)
Output: Optimal parameter combination (w,h)
gen = 0; //initialize the value of generation.
pop [gen] = initializePopulation (populationSize);//create primary population.
fitvalue = rank (LSTM (pop [gen])); //calculate population's fitness value.
While gen < maxgen //judge termination condition.
    parents = selectParents (fitvalue);//select excellent individuals.
    pop [gen + 1] = crossover (parents);
    pop [gen + 1] = mutate (pop [gen + 1]);//mutate and product a new population.
    (w,h) = min (LSTM (pop [gen + 1]));//record the optimal solution.
    fitvalue = rank (LSTM (pop [gen + 1]));//calculate new population's fitness value.
    gen++;
End
```

4 Experiments and Evaluations

4.1 Experimental Settings

Program Language: Python 3.7.7 and MATLAB; Integrated development environment: matlab2018a, spyder4.1.3; LSTM is implemented in Keras library with tensorflow as the back end.

4.2 Model Parameter Settings

The optimal combination of hidden layer node number and window size of LSTM is solved by GA. Set the population scale is 50, and the maximum iteration is 100. The error reached the minimum before 40 generation, and the corresponding combination value is (5, 80). The parameter configuration of LSTM model is shown in Tab. 5.

Table 5: LSTM parameter settings

Parameter	Parameter value
Time Window	5
Hidden Nodes	80
Epoch	200
Learning Rate	0.01

4.3 Evaluation Metric

In order to express the performance of the model accuracy, root mean square error (RMSE) is used as the evaluation metric. The smaller the RMSE is, the higher the accuracy is. The calculation method is shown in the Eq. (8), where p and p′ are the observed value and the predicted value, n_2 is the number of all predicted values.

$$RMSE = \sqrt{\frac{1}{n_2}\sum_{i=1}^{n_2}(p_i - p'_i)^2} \tag{8}$$

4.4 Experiment and Analysis

The experiment is designed to verify the influence of different trajectory expressions and parameters optimized based on genetic algorithm on the prediction model. The experiment results are as follows:

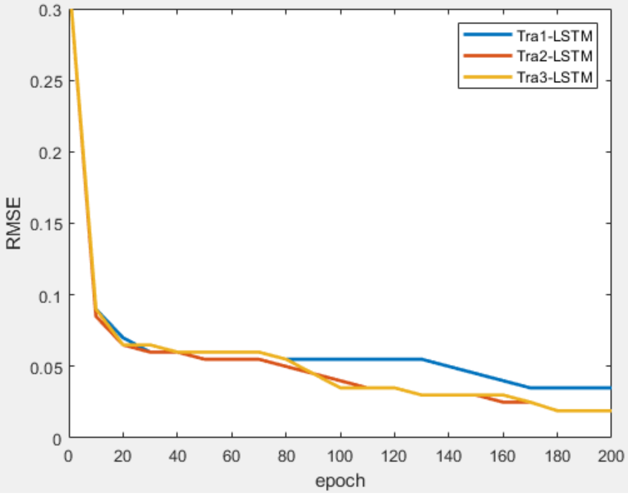


Figure 5: RMSE of LSTM in different trajectory expressions

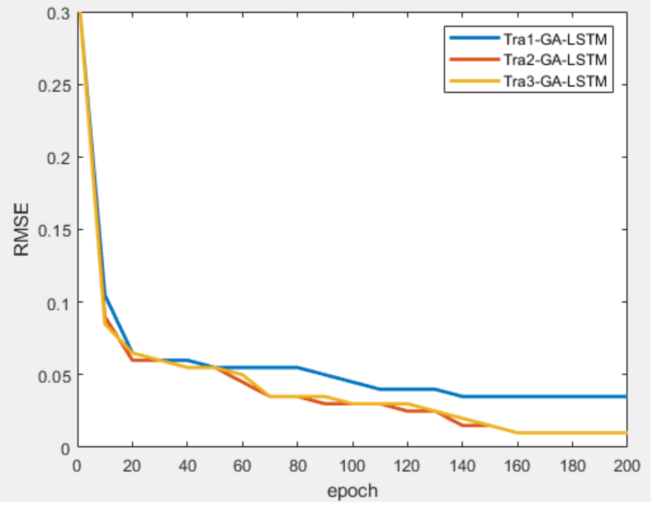


Figure 6: RMSE of GA-LSTM in different trajectory expressions

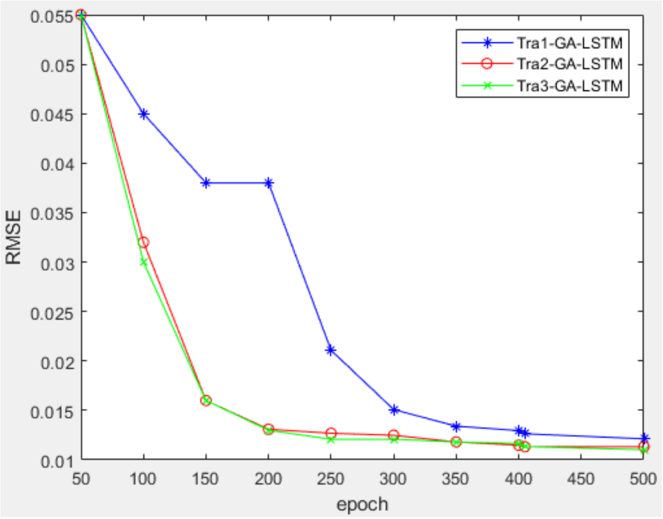


Figure 7: RMSE of GA-LSTM in different epochs

It can be seen from Fig. 5 that after the information is associated, the model is more complicated and the training time is longer due to the increase of the dimensionality. Therefore, compared with the model represented by the original trajectory, the convergence speed of the prediction model based on the information-related trajectory data is slightly slower, but the impact is not significant. As shown in Fig. 6, after using GA to optimize the number of hidden nodes and the window size, the error of the model reaches the minimum error in the range significantly faster than that without GA. This shows that proper window size and hidden layer nodes can improve the model learning efficiency, and GA is effective for optimization of combination parameters.

Fig. 7 increases the number of iterations. It can be seen that as the number of iterations increases, the prediction errors of tra1 and Tra1 are smaller than those of Tra2, indicating that the model using associated features still has an accuracy advantage under the condition of increasing g iterations.

Table 6: Performance of prediction model under different trajectory expressions

	GA-LSTM			LSTM		
	Training time (s)	Predict time (s)	RMSE	Training time (s)	Predict time (s)	RMSE
Tra1	51.60	2.86	0.04	55.09	2.96	0.04
Tra2	58.12	3.07	0.01	68.01	3.26	0.02
Tra3	56.50	2.93	0.01	60.22	3.17	0.02

According to Tab. 6, it can be seen that the RMSE of the prediction model expressed by the trajectory of Tra2 is lower than that of the prediction model expressed by the trajectory of Tra1. Therefore, the association of movement information features helps to improve the accuracy of location prediction. In addition, compared with the RMSE of the prediction model expressed using Tra2, the RMSE of the prediction model expressed using Tra3 hardly increased, and the training time is shorter. Therefore, feature selection can reduce the complexity of the model without increasing the prediction error. Under the same prediction model, the RMSE of the prediction model using genetic algorithm is obviously greater than that of the prediction model without GA optimization. Therefore, GA is effective for optimizing model parameters.

5 Conclusion

This paper designs a location prediction method that combines GA-LSTM and related movement behavior feature information, and uses Beijing taxi data to test the prediction error. The results show that associating movement behavior features and then selecting features will help reduce the accuracy of prediction and the complexity of the model, and GA can optimize parameters within the search range. The significance of this work is that accurate prediction of vehicle location can make the recommendation service more professional and provide technical support for ITS optimization.

Funding Statement: This work is supported by the Hunan University of Science and Technology Doctoral Research Foundation Project (E51873). The authors have no conflicts of interest to report regarding the present study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] H. R. Dai, Y. B. Tao and H. Lin, "Visual analytics of urban transportation from a bike-sharing and taxi perspective," in *Proc. 2019 12th Int. Sym. on Visual Information Communication and Interaction*, New York, NY, USA, pp. 1–8, 2019.

[2] M. M. Wu, C. C. Zhu and L. L. Chen, "Multi-task spatial-temporal graph attention network for taxi demand prediction," in *Proc. 2020 5th Int. Conf. on Mathematics and Artificial Intelligence*, New York, NY, USA, pp. 224–228, 2020.

[3] Y. Liu, Z. Li, W. Ai and L. Zhang, "SIGIR 2018 workshop on intelligent transportation informatics," in *Proc. The 41st Int. ACM SIGIR Conf. on Research & Development in Information Retrieval*, Ann Arbor, MI, USA, pp. 1441–1443, 2018.

[4] J. Liang, L. Jiang, J. C. Nieble, A. Hauptmann and L. Fei-Fei, "Peeking into the future: Predicting future person activities and locations in videos," in *Proc. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 5725–5734, 2019.

[5] Y. H. Lu, Z. C. He and L. L. Luo, "Learning trajectories as words: a probabilistic generative model for destination prediction," in *Proc. 16th EAI Int. Conf. on Mobile and Ubiquitous Systems: Computing, Networking and Services*, New York, NY, USA, pp. 464–472, 2019.

[6] M. Chen, X. Yu and Y. Liu, "Mining moving patterns for predicting next location," *Information Systems*, vol. 54, no. C, pp. 156–168, 2015.

[7] Z. N. Zou, H. P, L. Liu, G. Xiong and D. Li, "Deep convolutional mesh RNN for urban traffic passenger flows prediction," in *Proc. 2018 IEEE Smart World, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, San Francisco, CA, USA, pp. 1305–1310, 2018.

[8] H. Al-Theiabat, M. Al-Ayyoub, M. Alsmirat and M. Aldwair, "A deep learning approach for amazon EC2 spot price prediction," in *Proc. 2018 IEEE/ACS 15th Int. Conf. on Computer Systems and Applications (AICCSA)*, Jordan, pp. 1–5, 2018.

[9] P. Shu, Y. Sun, Y. Zhao and G. Xu, "Spatial-temporal taxi demand prediction using LSTM-CNN," in *Proc. 2020 IEEE 16th Int. Conf. on Automation Science and Engineering (CASE)*, Hong Kong, China, pp. 1226–1230, 2020.

[10] B. Yan, X. Tang, J. Wang, Y. Zhou and G. Zheng, "An improved method for the fitting and prediction of the number of Covid-19 confirmed cases based on LSTM," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1473–1490, 2020.

[11] H. Huang, T. Wang, J. Liu and S. Xie, "Predicting urban rail traffic passenger flow based on LSTM," in *Proc. 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, China, pp. 616–620, 2019.

[12] S. S. Vaitheeswaran and V. R. Ventrapragada, “Wind power pattern prediction in time series measuremnt data for wind energy prediction modelling using LSTM-GA networks,” in *Proc. 2019 10th Int. Conf. on Computing, Communication and Networking Technologies (ICCCNT)*, Kanpur, India, pp. 1–5, 2019.

[13] P. MKumar and S. Batra, “Meta-heuristic based optimized deep neural network for streaming data prediction,” in *Proc. Int. Conf. on Advances in Computing, Communication Control and Networking*, Greater Noida, India, pp. 1079–1085, 2018.

[14] Q. Wang and X. Wang, “Parameters optimization of the heating furnace control systems based on BP neural network improved by genetic algorithm,” *Journal of Internet of Things*, vol. 2, no. 2, pp. 75–80, 2020.

[15] R. Ibrahim and M. O. Shafiq, “Mining trajectory data and identifying patterns for taxi movement trips,” in *Proc. 2018 Thirteenth Int. Conf. on Digital Information Management (ICDIM)*, Berlin, Germany, Germany, pp. 130–135, 2018.

[16] C. Liu, S. Y. Wang, S. Cuomo and G. Mei, “Data analysis and mining of traffic features based on taxi GPS trajectories: A case study in Beijing,” *Concurrency & Computation: Practice & Experience*, vol. 31, no. 9, pp. 1–13, 2019.

[17] S. M. Cui, L. Zhang and Y. Li, “Deep learning method for taxi destination prediction,” *Computer Engineering & Science*, vol. 42, no. 1, pp. 185–190, 2020.

[18] J. Xu, R. Rahmatizadeh, L. Bölöni and D. Turgut, “Real-time prediction of taxi demand using recurrent neural networks,” *IEEE Transactions on Intelligent Transportation Systems,* vol. 19, no. 8, pp. 2572–2581, 2018.

[19] X. Fan, L. Guo, N. Han, Y. Wang, J. Shi *et al.,* “A deep learning approach for next location prediction,” in *Proc. 2018 IEEE 22nd Int. Conf. on Computer Supported Cooperative Work in Design (CSCWD)*, Nanjing, China, pp. 69–74, 2018.