ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0086

Research on Four-dimensional Trace Prediction Method based on Civil Aviation Flight Big Data

Kaicheng Zhang, Boye Lu, and Youming Dou

School of Science, North University of Science and Technology, Tangshan, 063000 China

Abstract

The data in this paper adopt the annual data of aircraft flying between airports in 2019.10-2020.10 in the past year.Data processing was performed on the 3 D coordinates with time as an independent variable.To verify the effectiveness of Bi-LSTM for trajectory prediction, compare the real trajectories and BP neural network, the Bi-LSTM algorithm has small errors and outthan BP neural network.

Keywords

Four-dimensional Trace; The Bi-LSTM Neural Network Model; Civil Aviation Flight Big Data.

1. Introduction

In the current environment, the growing economy, the plane demand is more and more, the rapid rise of civil aviation makes the airline economic scale, lead to air control is increasingly difficult, crowded and some routes conflict problems increasingly apparent, in order to cope with the aviation flight challenges to the existing air traffic control system.Four-dimensional trace prediction technology arises at the historic moment. The so-called four-dimensional space adds the fourth dimensional time dimension to judge the flight track of aircraft, and also consider the height, weight, speed and other flight parameters, so as to obtain real-time, accurate and continuous four-dimensional trace information.Compared with the traditional prediction technology, the four-dimensional trace technology can greatly shorten the prediction time, so that the use efficiency of space resources can be greatly improved.

2. Modeling of Flight Trajectory Prediction based on the Bi-LSTM Neural Network

2.1 The LSTM Neural Network Model

Recurrent Neural Network (RNN) is a recurrent neural network (Recursive Neural Network) with input as sequence data and feedback connection, and is transmitted recursively in the direction of sequence propagation, and all nodes are connected by chain.Because its implicit layer adds a feedback structure, that is, the output of the current moment is not only related to the input of the current moment, but also related to the output of the previous moment, equivalent to a deep neural network expanded on the time series, enabling it to compare snakes and process the time series data.Therefore, RNN is often used to handle natural language processing problems with context containing temporal information with memory and parameter sharing.

The RNN structure is shown in Figure 1.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0086

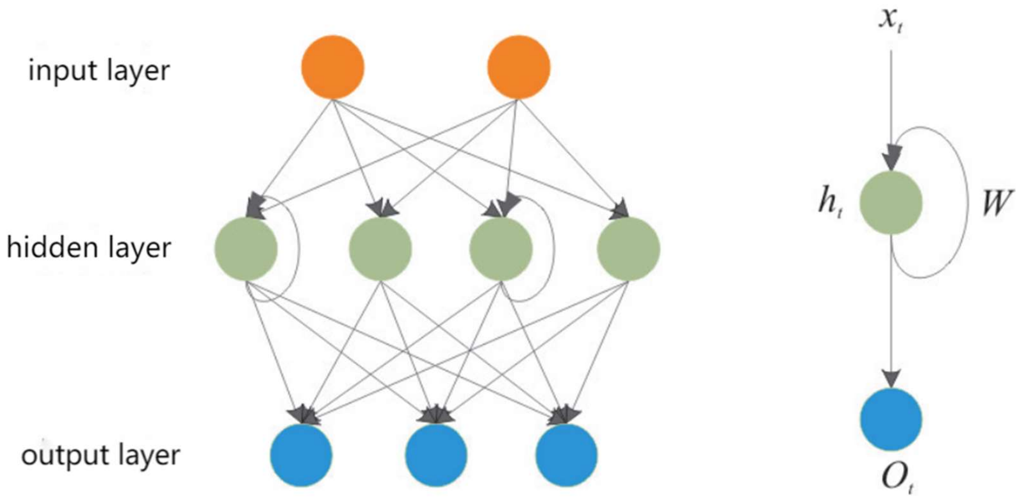

Figure 1. Structure diagram of the R N N

LSTM is an evolutionary network of traditional recurrent neural networks that can improve the gradient explosion and extinction problems due to the reverse ship error during regression neural network learning, capture long-distance dependencies and learn effectively from sequences of different lengths.The LSTM memory unit in the LSTM model is the evolution of recurrent neurons with hidden layers in a traditional RNN, and the unit can indicate when the network forgets historical information and when to update the storage unit with new input information.Each LSTM unit is equipped with three gates to control information flow, each containing a Sigmoid layer and a point multiplication operation, and can control the LSTM information flow to avoid gradient explosion and disappearance. The gates to control information flow are: the forgetting gate determines which information to remember or forget, which information is important for memory, and the output gate determines the information to be transmitted.

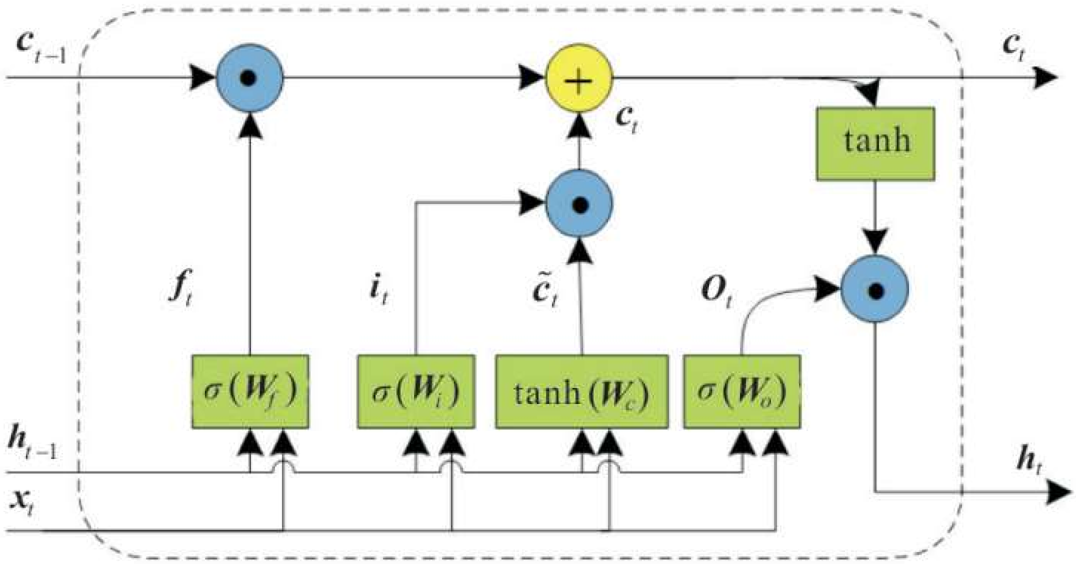

Figure 2. The L S T M structure

According to the LSTM structure, we obtain:

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{(t-1)} + b_{hi})$$
$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{(t-1)} + b_{hf})$$
$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{(t-1)} + b_{hg})$$
$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{(t-1)} + b_{ho})$$
$$c_t = f_t * c_{(t-1)} + i_t * g_t$$
$$h_t = o_t * \tanh(c_t)$$

Figure 3. Correlation function formula

In formula, i, f, c, o represent the input gate, forgetting gate, cell state, output gate; b represents the corresponding bias term; W represents the weight matrix between layers and gates; sigmoid activation function; and tanh is the hyperbolic tangential activation function.

The LSTM has the advantage that the current unit has information about all the units before the unit, and the disadvantage is that the information after the unit cannot be obtained.Considering a set of temporal data, information is included not only in past data but also in future data.If the network can learn the rules in forward and backward data, the performance of the network can be further improved.Thus, the Bi-LSTM came into being.

2.2 The Bi-LSTM Neural Network Model

The Bi-LSTM neural network structure model is divided into two independent LSTM, and the input sequence is input to two LSTM neural networks in forward and reverse order for feature extraction. The two word vectors (i. e., the extracted feature vectors) are expressed as the final features of the word.The model design concept of Bi-LSTM is to make the feature data obtained at time t have the information between the past and the future. Experimentally proves that this neural network structure model has better efficiency and performance on text feature extraction than a single LSTM structure model.It is worth mentioning that the 2 LSTM neural network parameters in the Bi-LSTM are independent of each other, and they only share the word-embedding word vector list.

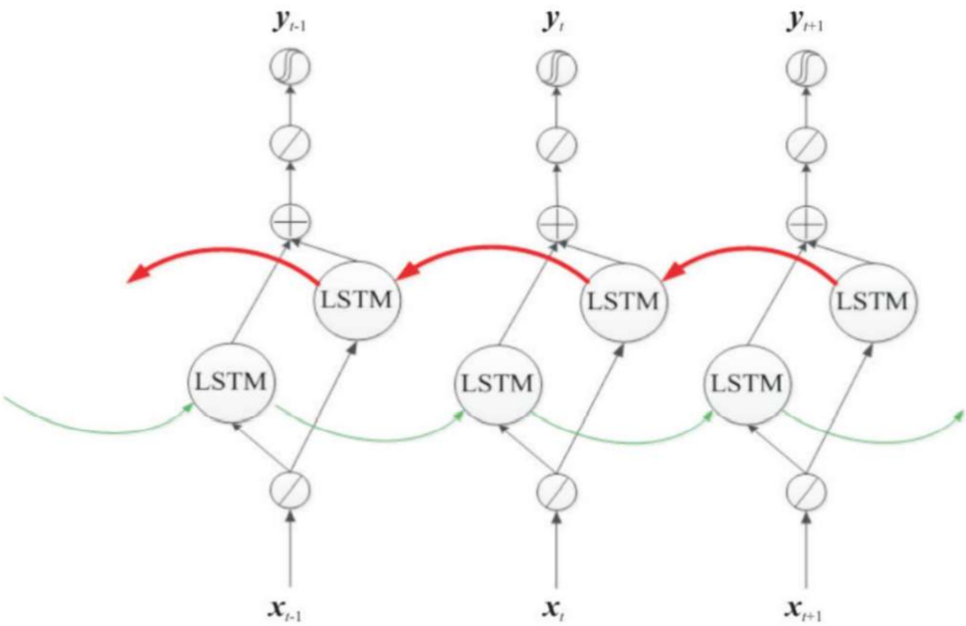

Figure 4. Schematic diagram of the B i-LSTM neural network model

International Core Journal of Engineering Volume 8 Issue 6, 2022
ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0086

3. Data Analysis

The data in this paper adopt the annual data of aircraft flying between domestic airports of 2019.10-2020.10 for nearly a year.

In recent years, trajectory data analysis has received much attention in the field of controlled traffic.Literature [3] uses principal component analysis and K mean algorithm for trajectory clustering, sampling flight locations, and clustering in space according to the given key components.In China, Wang Taotao uses the fuzzy clustering algorithm to extract the air trace feature points, and gives the spectral clustering to extract the prevailing traffic flow in the terminal area for analysis.

According to our data analysis, for such a large-scale data processing: first, we sampled the flight data through the clustering of the trajectory, and then aligned the true data with the measurement results of the test equipment to complete the corresponding coordinate transformation, so that it unified the two kinds of data in time and space.

In order to compare and analyze the strength of the three variables of longitude and latitude height, we express the part of the sampled data by simulating the navigation track.As shown in Figure 5 trajectory simulation.

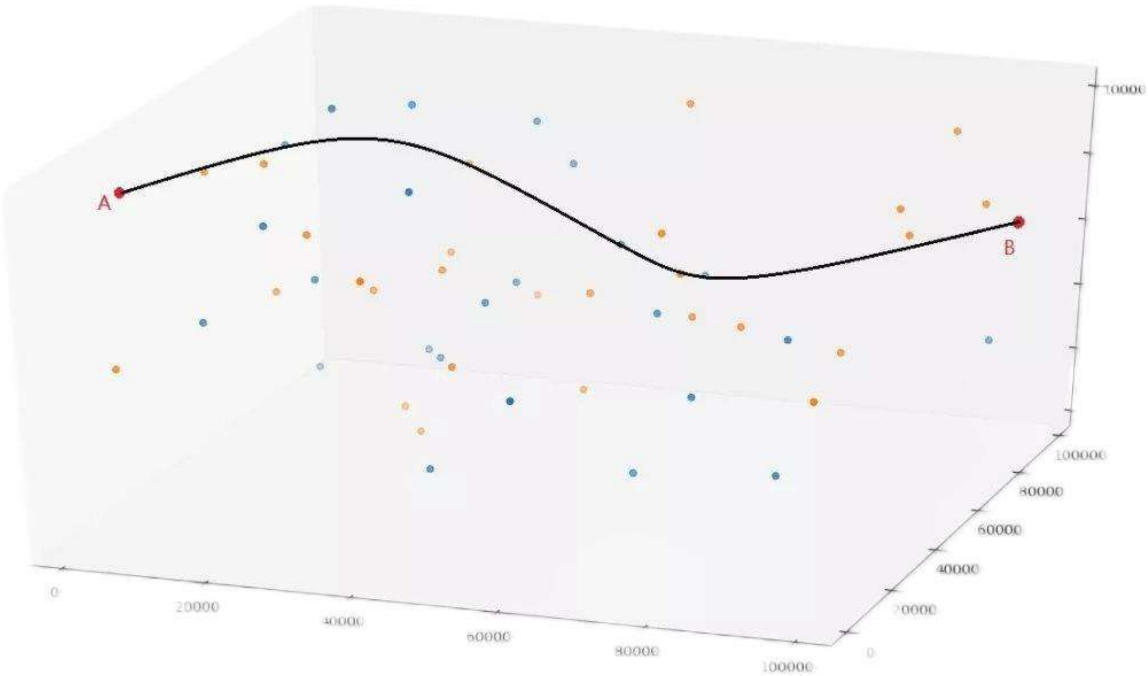

Figure 5. Schematic diagram of the trajectory simulation

To verify the effectiveness of the Bi-LSTM algorithm for trajectory prediction, we performed the trajectory prediction analysis at latitude and longitude by comparing the real trajectories as well as the BP neural network.

As shown in Figure 6, the Bi-LSTM algorithm has a smaller error in all directions and outperforms the BP neural networks.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202206_8(6).0086

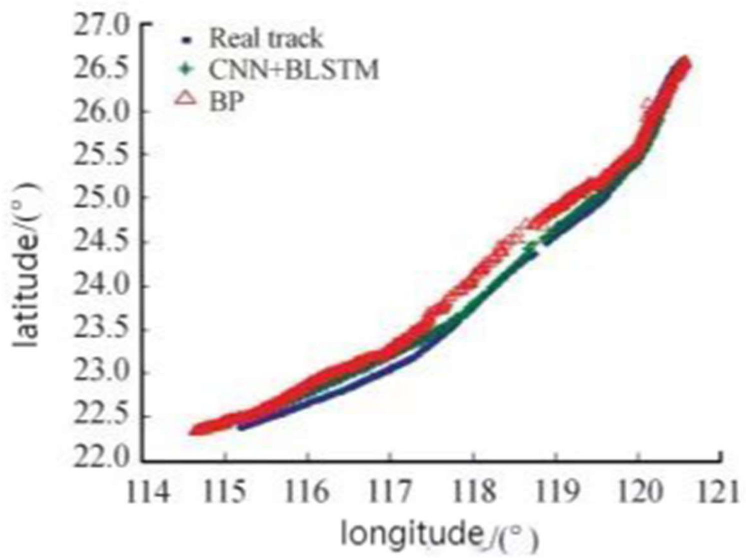

Figure 6. latitude and longitude trajectory prediction

4. Conclusion

In this paper, the flight trajectory in four dimensions: longitude, latitude, altitude and time.Based on a large amount of civil aviation flight data, this project conducts big data mining of civil aviation trajectory, and uses Matlab calculation of statistical data to generate a flight track training set.Using a large number of data sets obtained after training and clustering, the data model is constructed through the convolutional neural network to initialize the convolutional layer in the network, and then use the one-dimensional convolutional network to predict the latitude and longitude and height in the time series.Then, the data obtained from the convolutional network is hierarchically constructed through the deep belief network, and the optimal solution is established to obtain the most similar solution to the actual results.The predicted trajectory node arrival time was obtained using the SAE stacking autoencoder model.Finally, the final values of the three kinds of neural network prediction were weighted average, and the more accurate real-time prediction values were obtained after being manually adjusted. Finally, the dimension reduction process was performed to make the data three-dimensional and displayed on the visualization platform.In the subsequent application process, we will analyze and adjust the initial weights to learn to improve the neural network to strengthen the accuracy.

References

[1] Wang Yi, research on flight test data processing technology based on Matlab.Electronic testing. 2021, (12) 67-69.

[2] Hui Huihui, Wei Bin, four-dimensional track prediction technology overview.Scientific and technological innovation and application. 2020,(33)18-21.

[3] Adric Eckstein.Automated flight track taxonomy for measuring benefits from performance based navigation. In Integrated CommunicationNavigation and Surceillance Conference,2009.