

Article

A new way of airline traffic prediction based on GCN-LSTM

Jiangni Yu

¹ School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, China; yu-jiangni@163.com

* Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials) +xx-xxxx-xxx-xxxx (F.L.)

† Current address: Affiliation 3

‡ These authors contributed equally to this work.

Version January 23, 2021 submitted to Journal Not Specified

Abstract: With the development of society and the improvement of people's material level, airplanes are becoming more and more a means of transportation for people to travel. If an airline can predict the passenger flow in advance, it can be used as an important decision-making basis for its flight route planning, crew scheduling planning and ticket price formulation in the process of management and operation. However, due to the high complexity of aviation network, the existing traffic prediction methods generally have the problem of low prediction accuracy. In order to overcome this problem, this paper makes full use of graph convolutional neural network and long - short memory network to construct a prediction system with short - term prediction ability. Specifically, this paper uses the graph convolutional neural network as a feature extraction tool to extract the key features of air traffic data, and solves the problem of long term and short term dependence between data through the long term memory network, and builds a high-precision air traffic prediction system based on it. Finally, we design a comparison experiment to compare the algorithm with the traditional algorithm. The comparison results show that the algorithm has obvious advantages in air flow prediction.

Keywords: Graph Convolutional Network; Long Short Term Memory Network; Flow; Airlines; Predict

1. Introduction

In recent years, as an important industry in national economic and social development and an advanced mode of transportation, the demand for civil aviation passenger transport has been growing rapidly along with the rapid development of national economy and the substantial increase of people's income. Air fare has always been the focus of attention because it is related to the development of civil aviation industry, the profitability of airlines and the vital interests of passengers. Pricing strategies of airlines are mostly based on revenue management theory [1][2], in which air traffic forecasting plays an important role. The forecast results can be used to dynamically adjust the ticket prices, so that the airlines can get the maximum benefits. However, the pricing mechanism of each airline is complex, and the real-time ticket price is constantly changing under the influence of many factors, which has the characteristics of trend, randomness and volatility. Therefore, how to forecast air passenger flow accurately and reasonably has become an important content of air transport management in China.

Air passenger flow prediction has been studied at home and abroad. In [3], an ensemble learning binary classification algorithm, Hamlet, is proposed based on Q-learning. Hamlet applies rule learning, reinforcement learning and time series techniques, and combines their results through superposition generalization to produce the final decision. The accuracy can reach 74.5%. There is still much room

for improvement in traffic prediction. Riedel et al. [4] used autoregressive model (AR) to realize the prediction of air traffic. An AR (autoregressive model) is the regression of a variable relative to itself, using past observations to make predictions about current values. Moving average model (MA) is a regression model based on past forecasting errors. Sickles et al. [5] proposed a model ARMA that combines autoregressive model and moving average model. ARIMA[6] and Bayesian structured time series BSTS model[7][8] are used to forecast the air passenger and cargo demand of the Indian aviation industry. Wang et al [9] used gray scale model to forecast passenger flow, applied a function to the observation data set, converted it into a monotonically increasing data set, and obtained the forecast results. Nieto et al. [10] proposed a new damping trend grey model (DTSM) based on dynamic seasonal damping factor, which is used to predict airline passenger demand (PAX) in the air transport industry. This model is called Sarima Damped Trend Grey Prediction Model (SDTGM), which can effectively improve the accuracy of prediction.

The classical time series model is based on the assumption that there is a linear relationship between the past and the future value, but the time series model to predict the traffic demand data must have correlation, because the passenger flow is affected by various factors change, trend and randomness and volatility characteristics, will also affect the forecast of air flow. Causal models attempt to establish an explanatory equation that adequately describes the observation as the output of one or more underlying causative factors. Liu et al [11] proposed to use the probability logarithmic model to predict passenger flow, establish factors that can explain past demands, and then use the model to provide predicted values. Logit models, such as multinomial models, nested models and cross-nested models, are widely used in demand prediction of passenger selection models [12].

It is a good trend to apply machine learning technology to solve nonlinear time series problems. Various machine learning algorithms have been studied and deployed. Neural Network model [13][14] is widely used in long-term demand prediction. Tsai et al. [15] use them for short-term demand forecasting. They compared two advanced formulas, MTUNN and Penn, with the general multilayer perceptron model. Wei and Chen[16] used hybrid neural network model to predict passenger flow in rapid transit systems. A Bayesian state-space model can be created by generating the current state matrix from the observed data in the time series [17]. SVM support vector machine is regarded as a classification tool, and support vector regression aims to identify and optimize the error range of regression. Jiang et al. [18] applied grey SVM combined with empirical mode decomposition (EMD) to high-speed railway passenger flow prediction. Xie et al. [19] used EMD to model passenger flow prediction of airport terminals by least squares support vector machine. Weng et al [20] applied the hybrid model combining LSTM and LightTGBM to air ticket sales prediction. Artificial intelligence or machine learning models are computation-intensive self-learning algorithms that iteratively modify and fine-tune the interpretation model through the evaluation results, and reduce the margin of error. However, the above machine learning based method is based on the historical data of each air station forecast, this method does not take into account the impact of ticket prices and the ridership flow of other stations and the current ridership flow of the station, so the prediction performance is poor.

Graphic Convolutional Neural Network (GCN), as a Neural Network that can extract unstructured data, has attracted a lot of attention in solving the relationship between adjacent points [21][22]. In view of the low accuracy of air passenger flow prediction and the trend, randomness and volatility of air traffic affected by many factors, we built a graph convolution-long short-term memory model based on graph convolutional neural network and long short-term memory (LSTM) neural network. In this model, the GCN is used to map the features of the data set. Then the LSTM model are used to process the matrix data set, and the fare prediction is realized. The method considers the different effects of various factors on the ticket price, and combines the trend and fluctuation characteristics of the ticket price. The experimental results show that the graph convolution model of long memory and short memory can predict the air ticket price well. Our contributions are summarized as

(1) In order to improve the feature extraction ability of aviation data, we have a deep understanding of the data characteristics of aviation data. By constructing a GCN feature extractor, we

can transform non-European spatial data into concise and efficient features. This method improves the feature expression ability of the data.

(2) In order to solve the problem of long term and short term dependence of time series data, in this paper, we introduce LSTM network to solve the data dependence of samples through logical gating unit, and further improve the performance of prediction.

(3) The air flow prediction scheme based on GCN-LSTM shows excellent performance on the existing aviation data. The experimental results show that the prediction performance of the proposed algorithm is obviously better than that of other existing algorithms.

The remaining part of this paper is organized as follows. The main contribution of this paper is described in Section 2. In Section 3, two comparative experiments are used to prove the effectiveness of the algorithm. And conclusions are drawn in Section 4.

2. Main result

In airlines, the spatial distribution of aircraft stations is a non-euclidean structure, that is, the number of stations around each station is uncertain, and even if two stations are adjacent, they may not actually communicate with each other, resulting in no spatial relationship between their traffic. Therefore, traditional convolutional neural network (CNN) cannot accurately obtain their spatial information. At this point, multiple sites can be abstracted into A graph (see Figure 1). Features are extracted from the original input data to obtain the result of feature mapping of multiple channels. The intercommunication relationship between each site is represented by adjacency matrix A.

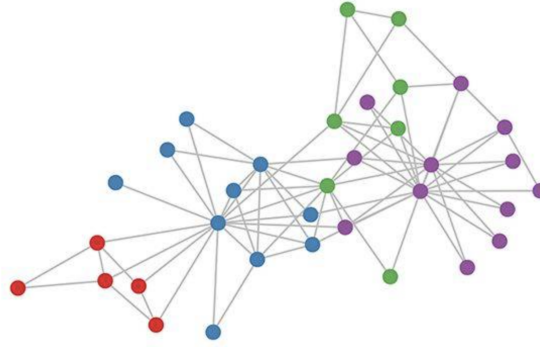


Fig. 1. The distribution of the stations.

2.1. The definition of the problem

The problem of airline passenger flow prediction can be described as follows: the historical flow data of each station $X_{t-s}, X_{t-s+1}, X_{t-2}, \dots, X_{t-1}$ (s is the time step) can be used to predict the flow X_t of the next period. The formula is described as

$$X_t = F([X_{t-s}, X_{t-s+1}, \dots, X_{t-2}, X_{t-1}]), \quad (1)$$

where, X is the site characteristics at each time step, and F is a nonlinear function.

In the actual traffic system, the network is regarded as a directed graph $G = (Q, V, A)$. Each sensor in the network is regarded as a node v_i and its value $Q \in R$ is a scalar. $V \in R^N$ and N is the number of sensors. The flow relationship between nodes consists of adjacency matrix A that is, the element A_{ij} in A represents the connection relationship between node V_i and V_j .

107 2.2. The description of the GCN

When dealing with the structure of the graph, it is necessary to obtain its Laplace matrix L , which is generally defined in the following ways:

$$L = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}} = I_N - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}, \quad (2)$$

108 where, I_N is the identity matrix of NN ; Degree matrix D is defined as $D_{ii} = \sum_j A_{ij}$. Decompose
 109 the eigenvalue of L to get $L = \Lambda U^T U$. Λ is made up of L eigenvalues of diagonal matrix.
 110 $U = \{u_1, u_2, \dots, u_N\}$ is composed of the eigenvector L , and it is an orthonormal basis for \mathbb{R}^N .

The spectral convolution theory in the graph structure has been supplemented and perfected in the paper. The convolution operation of convolution kernel G and input signal X in the time domain can be converted into the inner product form in the frequency domain.

$$g^*x = U \left(U^T g \right) \odot \left(U^T x \right) = U_{g_\theta}(A) U^T x, \quad (3)$$

where $g_s(\Lambda) = U^T g = \text{diag}(\theta)$, $\theta \in \mathbb{R}^N$, \odot represents the hadamar product, $U^T g$ means mapping g to the frequency domain space based on U . Due to g_θ high computational complexity, so hierarchical linear model constraints and the chebyshev polynomial are used to approximate calculation. In this paper, the simplified first order polynomial form of g^*x is adopted.

$$g^*x = U_{g_\theta} U_x^T \approx \theta \left(I_N + D^{-1/2} A D^{-1/2} \right) x, \quad (4)$$

where $\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} = I_N + D^{-1/2} A D^{-1/2}$ $\tilde{A} = I_N + A$, $D = \sum_i \tilde{A}_{ij}$, Therefore, the output of layer L is

$$H^{(l)} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l-1)} W^{(l)} \right), \quad (5)$$

where δ is the activation function, $\tilde{W}^{(l)} = \theta^{(l-1)} W^{(l)}$, $\theta^{(l-1)} \in \mathbb{R}^{c^{(1-l)} \times F^{(l-1)}}$, $W^{(l)} \in \mathbb{R}^{F^{(l-1)} \times c^{(l)}}$, $C^{(L-1)}$ is the output dimension of the $(L-1)$ layer, and $F^{(L-1)}$ is the characteristic vector size of each dimension. Therefore

$$H^{(l)} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l-1)} \tilde{W}^{(l)} \right). \quad (6)$$

At present, there is no effective measurement method for the calculation of adjacency matrix A . Most scholars use heuristic methods, that is, based on the Euclidean distance or Markov distance between sensors to determine the element value corresponding to the adjacency matrix. However, these methods all require manual calculation of the distance relationship between the sensors in advance. In this paper, the data-driven method is adopted to calculate the adjacency matrix, and $A = D$. The formula can be written as:

$$H^{(l)} = \sigma / \tilde{A} H^{(l-1)} \tilde{W}^{(l)} \quad (7)$$

111 The element value of matrix A is learned from the sample data, that is, the matrix is composed of
 112 trainable parameters. The data-driven approach is more realistic than the heuristic approach. Therefore,
 113 the L layer of the convolutional neural network is constructed in accordance with Formula 7. It should
 114 be noted that the initial matrix A is the same for each layer of the convolutional network, and the
 115 parameters are updated only when the error is propagated backwards.

116 GCN introduces the spatial features of the graph by convolving the Laplace matrix with the input.
 117 In this paper, the model takes flight segments as nodes and the association between flight segments as
 118 edges to build a graph. According to the graph, the adjacency matrix is obtained and the demand of
 119 future flight segments is predicted by combining the price and demand of historical flight segments.

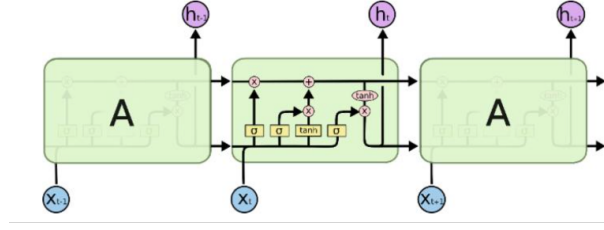


Fig. 2. The structure of LSTM.

2.3. The problem of time series

It is found that Recurrent Neural networks (RNN) are widely used in sequential data such as natural language and image processing, which have a significant effect. Since then, various types of Circulating Neural Networks have been used. Aiming at the problem of air passenger flow prediction, this paper introduces the LSTM, which can extract the characteristic information of the input sequence and find its internal relation, so as to improve the prediction accuracy of the model. In order to make use of the spatial-temporal characteristic of the airborne data information at the same time, the GCN model is combined with the LSTM model, and it is added to the output of the upper level.

The LSTM network structure used in this paper is shown as Figure 2.

The network model mainly accepts three inputs: X , H and C represent the current state, hidden layer state and cell state, respectively.

LSTM mainly realizes the management of long and short term memory through three gating units. Firstly, the LSTM needs to determine what information needs to be thrown out. This step is decided by the layer called the "Forget Gate". Input x and h , output a number between 0 and 1. The value of 1 means "keep the value completely", while 0 means "throw the value away completely". The formula of forgetting gate is as follows:

$$f_t = \sigma \left(W_f \cdot [h_{t-1}, x_t] + b_f \right), \quad (8)$$

where W_f and b_f are the parameters to be learned, and σ is the sigmoid activation function.

Secondly, the LSTM needs to determine what information need to store in the cell state. There are two stage to this question. First, the layer of "Input Gate" determine which data needs to be updated. Then, the vector C_1 was created by a tanh layer .

$$\begin{aligned} i_t &= \sigma (W_i \cdot [h_{t-1}, x_t] + b_i), \\ \tilde{C}_t &= \tanh (W_C \cdot [h_{t-1}, x_t] + b_C). \end{aligned} \quad (9)$$

After deciding what needs to be forgotten and what needs to be added, C_{t-1} can be updated to C_t .

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t. \quad (10)$$

Finally, we need to decide what to export. This output is based on our cell state. The final output is part of the cell state. First, we run an output gate to determine which part of the cell state we are going to output. Then we put the cell state into the tanh (pressing the value between -1 and 1). Finally we multiply it by the output of the output gate.

$$\begin{aligned} o_t &= \sigma (W_o [h_{t-1}, x_t] + b_o), \\ h_t &= o_t * \tanh (C_t). \end{aligned} \quad (11)$$

2.4. The description of algorithm

The network structure based on GCN-LSTM model proposed in this paper is shown in the Fig. 3. The model mainly adopts encoder-decoder structure. In the encoder, multiple parallel GCN modules

are used to extract the key features of the graph network with different time series. Then, the extracted time series features are transmitted to LSTM, and feature analysis and further feature extraction are carried out on the sequence data through LSTM to solve the long-term and short-term dependencies between the data. Finally, the encoder generates an encoded pair vector and sends it to the decoder. In the decoder, the multi-layer feedforward neural network is used to further process the features of the coding vector. Finally, the processed data is transmitted to a GCN network to produce predicted values.

In order to improve the prediction performance of the GCN-LSTM algorithm, in this paper, we use the final output through the network and the value of the real label to calculate L_1 loss, also known as the mean absolute error (MAE), and take the mean square error (MSE) as the evaluation index of the model. The specific calculation formula is as:

$$\text{loss} = \text{MAE} = \frac{1}{m} \sum_{i=1}^m |(y_i - \hat{y}_i)|. \quad (12)$$

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m ((y_i - \hat{y}_i))^2. \quad (13)$$

In this paper, the Adam-based batch gradient descent optimization algorithm is used to learn and update parameters through the loss function to minimize the loss function until the loss function converges. The trained model is used to predict the test set and calculate the MSE. The smaller the MSE value is, the closer the predicted value of the model is to the real value and the better the generalization ability of the model is.

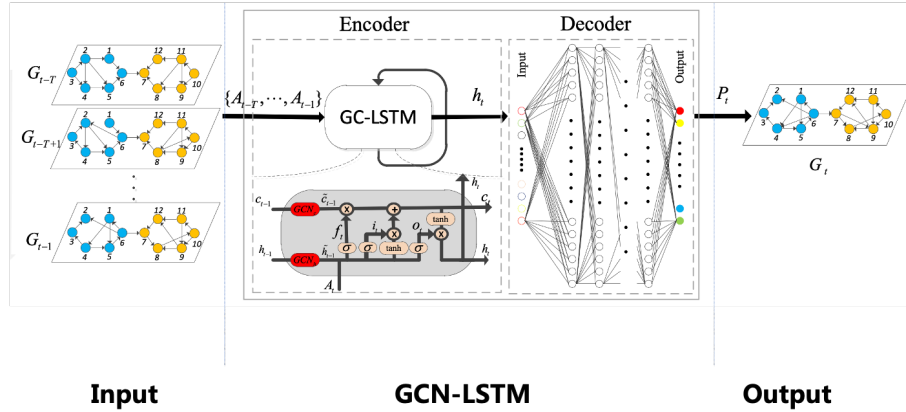


Fig. 3. Overall structure of GCN-LSTM model (original features of air passenger flow data are extracted by using first-order approximate GCN, and output features are analyzed by LSTM for long-term and short-term sequence characteristics, and then predicted values are obtained).

3. Experiments

This paper selects 6,911,332 ticket sales data examples of 28,809 flights of 17 domestic airlines between September 1, 2020 and October 31, 2020 to verify the feasibility and effectiveness of the GCN-LSTM combined prediction model. The selected data set is relatively complete without missing values. After obtaining a complete data set, an adjacency matrix is constructed according to the data set. Each flight segment is a node. Between nodes, if the origin is the same, it is considered to have a diverting effect on traffic, and the weight is set as the number of (0,1). If the destination is the same and the traffic is considered to be promoted, the weight is set as the number greater than 1 and the rest as 0.

Five sections of AAT_URC, CAN_PEK, CAN_CSX, CAN_CTU and CAN_CKG are taken as examples to construct an adjacency matrix, in which AAT, URC, CAN, PEK, CSX, CTU and CKG are city names. If the origin of two flight sections is the same, it is considered to have a diversion effect on

traffic, and the weight is set as 0.5. None of the five sections had the same destination, so the rest of the weights were set to 0.

Table 1. Adjacent Matrix.

	AAT_URC	CAN_PEK	CAN_CSX	CAN_CTU	CAN_CKG
AAT_URC	0	0.0	0.0	0.0	0.0
CAN_PEK	0	0.0	0.5	0.5	0.5
CAN_CSX	0	0.5	0.0	0.5	0.5
CAN_CTU	0	0.5	0.5	0.0	0.5
CAN_CKG	0	0.5	0.5	0.5	0.0

Time series features are constructed. The above five flight sections are also taken as examples. According to the take-off time, the price series 14 days before the take-off time is taken as the X feature. The flow of 1 day before the take-off was taken as Y, and such input and output were a sample. The take-off time was calculated forward to increase the number of samples. In this way, there are 14 days' time data, one piece of demand data, arranged in a time series according to time.

In this paper, the data set of two months is divided into training set and test set according to the ratio of 8:2, normalized, and trained by GCN-LSTM. The timing step size of the model is 15, and the window is constantly moved to predict. Under the premise of the same input of historical price time series, it is compared with the experimental results of traditional AR model, Moving Average(MA), Exponential Smoothing(ES), Long short-term memory(LSTM) and Support Vector Regression(SVR).

Table 2 lists five precise results of route prediction, The prediction mean square error (MSE) of the proposed GCN-LSTM model is smaller than that of LSTM, SVR, AR, moving average(MA) and exponential smoothing model(ES). The results show that the GCN-LSTM model can improve the accuracy of prediction.

Table 2. Prediction results MSE.

	GCN-LSTM	LSTM	SVR	AR	MA	ES
Average of five routes	6.86%	10.84%	12.72%	14.46%	13.73%	16.93%
AAT_URC	5.44%	8.71%	10.73%	11.37%	10.74%	14.39%
CAN_PEK	6.91%	9.74%	12.02%	12.75%	11.71%	16.47%
CAN_CSX	8.87%	12.03%	13.07%	17.13%	14.57%	18.29%
CAN_CTU	6.24%	11.73%	14.35%	15.01%	13.79%	17.48
CAN_CKG	6.84%	11.99%	13.43%	16.04%	17.84%	18.02%

AAT_URC a single route, for example, using the proposed GCN - LSTM traffic prediction model, the horizontal axis shows departure date from September 15, 2020 to October 31, 2020, 14 days before the use of the time sequence to forecast the traffic on the same day price, compare the renderings as shown in figure 4, dot says the actual passenger flow, square said according to 14 days before the price time series prediction of passenger flow.

Fig. 5 shows the average MSE values predicted by AR model, moving average, exponential smoothing, LSTM and SSR for 5 airlines. The horizontal axis represents five different airlines, the vertical axis represents square error (MSE), and the red line is the prediction results of the proposed GCN-LSTM model. It can be seen from the figure that the GCN-LSTM model is used for prediction, and the error fluctuation is small, and the MSE is basically in the range of 5%-9%. The prediction results of the proposed GCN-LSTM model are obviously better than those of other models.

4. Conclusion

In view of the problems existing in air flow prediction, an air flow prediction model based on graph convolutional neural network and long short-term memory network is proposed based on in-depth analysis of the influencing factors of air flow prediction. Firstly, based on the characteristics of

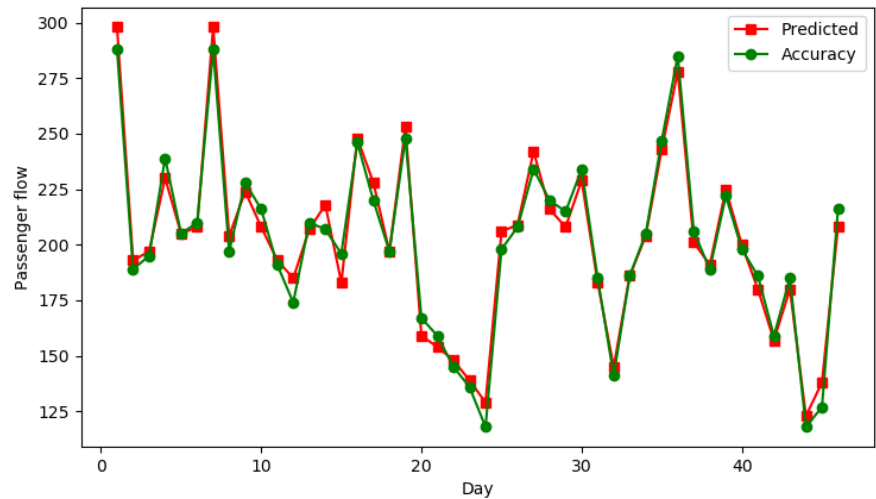


Fig. 4. OGCN-LSTM model prediction results of AAT URC route.

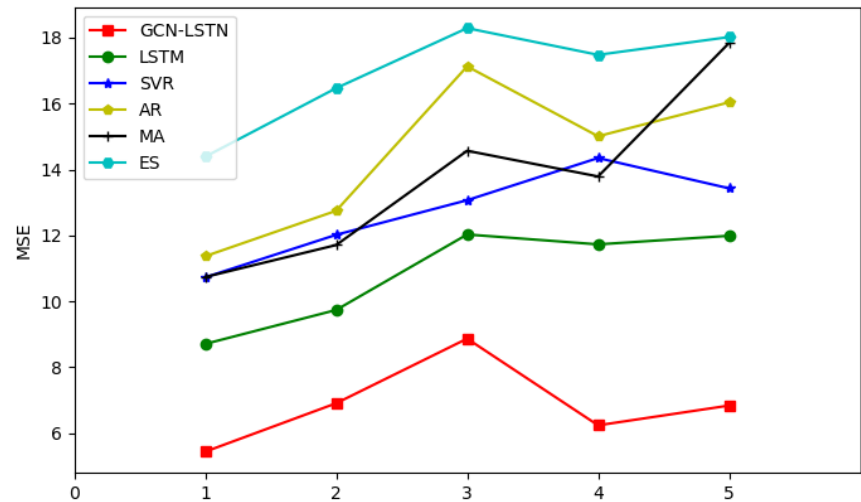


Fig. 5. Comparison of MSE results of five airlines.

air traffic data, a feature extraction network based on graph convolutional neural network is designed. Then combined with the long - short memory network to solve the problem of long - short - term data dependence; Finally, the prediction results are outputted based on the feedforward neural network. In the experimental part, we verify the performance of the GCNLSTM model on aviation data sets. The experimental results show that the prediction results of this model are obviously more accurate than the existing algorithms, and it has higher prediction performance.

References

- Donovan, A.W. Yield management in the airline industry. *Journal of Aviation/Aerospace Education & Research* **2005**, *14*, 9.
- Klein, R.; Koch, S.; Steinhardt, C.; Strauss, A.K. A review of revenue management: Recent generalizations and advances in industry applications. *European Journal of Operational Research* **2020**, *284*, 397–412.
- Etzioni, O.; Tuchinda, R.; Knoblock, C.A.; Yates, A. To buy or not to buy: mining airfare data to minimize ticket purchase price. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2003, pp. 119–128.
- Riedel, S.; Gabrys, B. Adaptive mechanisms in an airline ticket demand forecasting system **2003**.
- Sickles, R.; Good, D.; Postert, A.; Getachew, L. A model of world aircraft demand. Airport Facilities: Innovations for the Next Century. Proceedings of the 25th International Air Transportation Conference. American Society of Civil Engineers, 1998.
- Li, Y.; Han, H.; Liu, X.; Li, C. Passenger flow forecast of sanya airport based on ARIMA model. International Conference of Pioneering Computer Scientists, Engineers and Educators. Springer, 2018, pp. 442–454.
- Madhavan, M.; Ali Sharafuddin, M.; Piboonrungraj, P.; Yang, C.C. Short-term Forecasting for Airline Industry: The Case of Indian Air Passenger and Air Cargo. *Global Business Review* **2020**, p. 0972150920923316.
- Kulkarni, G.E.; Muley, A.A.; Deshmukh, N.K.; Bhalchandra, P.U. Autoregressive integrated moving average time series model for forecasting air pollution in Nanded city, Maharashtra, India. *Modeling Earth Systems and Environment* **2018**, *4*, 1435–1444.
- Wang, Y.; Chen, X.; Han, Y.; Guo, S. Forecast of passenger and freight traffic volume based on elasticity coefficient method and grey model. *Procedia-Social and Behavioral Sciences* **2013**, *96*, 136–147.
- Carmona-Benítez, R.B.; Nieto, M.R. SARIMA damp trend grey forecasting model for airline industry. *Journal of Air Transport Management* **2020**, *82*, 101736.
- Liu, R.; Li, A. Forecasting high-speed rail ridership using a simultaneous modeling approach. *Transportation Planning and Technology* **2012**, *35*, 577–590.
- Leng, N.; Nie, L.; Guo, G.; Wu, X. Passenger Flow Forecasting for Chinese High Speed Rail Network. 2015 International Conference on Mechatronics, Electronic, Industrial and Control Engineering (MEIC-15). Atlantis Press, 2015, pp. 675–678.
- Tsai, T.H.; Lee, C.K.; Wei, C.H. Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Systems with Applications* **2009**, *36*, 3728–3736.
- Babai, M.; Tsadiras, A.; Papadopoulos, C. On the empirical performance of some new neural network methods for forecasting intermittent demand. *IMA Journal of Management Mathematics* **2020**, *31*, 281–305.
- Grimme, W.; Bingemer, S.; Maertens, S. An analysis of the prospects of ultra-long-haul airline operations using passenger demand data. *Transportation Research Procedia* **2020**, *51*, 208–216.
- Chen, D.; Zhang, J.; Jiang, S. Forecasting the Short-Term Metro Ridership with Seasonal and Trend Decomposition Using Loess and LSTM Neural Networks. *IEEE Access* **2020**, *8*, 91181–91187.
- Zhang, Z.; Wang, C.; Gao, Y.; Chen, Y.; Chen, J. Passenger flow forecast of rail station based on multi-source data and long short term memory network. *IEEE Access* **2020**, *8*, 28475–28483.
- Jiang, X.; Chen, X.; Zhang, L.; Zhang, R. Dynamic demand forecasting and ticket assignment for high-speed rail revenue management in China. *Transportation Research Record* **2015**, *2475*, 37–45.
- Xie, G.; Wang, S.; Lai, K.K. Air passenger forecasting by using a hybrid seasonal decomposition and least squares support vector regression approach. Proceedings of the 59th world statistics congress of the international statistical institute, Hong Kong, 2013.

- 239 20. Weng, T.; Liu, W.; Xiao, J. Supply chain sales forecasting based on lightGBM and LSTM combination model.
240 *Industrial Management & Data Systems* **2019**, *ahead-of-print*.
- 241 21. Wei, Y.; Wang, X.; Nie, L.; He, X.; Hong, R.; Chua, T.S. MMGCN: Multi-modal graph convolution network
242 for personalized recommendation of micro-video. Proceedings of the 27th ACM International Conference
243 on Multimedia, 2019, pp. 1437–1445.
- 244 22. He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; Wang, M. LightGCN: Simplifying and Powering Graph
245 Convolution Network for Recommendation. *arXiv preprint arXiv:2002.02126* **2020**.

246 **Sample Availability:** Samples of the compounds are available from the authors.

247 © 2021 by the authors. Submitted to *Journal Not Specified* for possible open access publication
248 under the terms and conditions of the Creative Commons Attribution (CC BY) license
249 (<http://creativecommons.org/licenses/by/4.0/>).