

Freight Volume Prediction for Logistics Sorting Centers Using an Integrated GCN-BiLSTM-Transformer Model

Weiyan Tian^{1,a,*}, Shujia Wu^{2,b}, and Ke Ma^{3,c}

¹Guangdong Veterans Service Center, Guangzhou, China;

²Shenzhen Frontier Science and Innovation Research Center, Shenzhen, China;

³University of California, Santa Cruz, Santa Cruz, CA 95064, USA.

^a chenrj@mail.ustc.edu.cn, ^b 2529313124@qq.com, ^c kma41@ucsc.edu

Abstract. In this paper, a comprehensive learning model is proposed for predicting cargo volume at sorting centers, which integrates GCN (Graph Convolutional Network), BiLSTM (Bidirectional Long Short-Term Memory Network), ARIMA (Autoregressive Integrated Moving Average model), and Transformer models. To achieve this, a directed weighted graph is constructed considering the transport network and average cargo volume of each sorting center. The GCN model is employed to extract spatial features from the transport connection information of the sorting centers and these features are then fed into the BiLSTM network. The BiLSTM network leverages bidirectional information flow to learn the temporal characteristics of the data. Subsequently, the GCN-BiLSTM model, which combines spatial and temporal features, is used to predict the daily cargo volume for the next 30 days. The results demonstrate that the GCN-BiLSTM model and the ARIMA-BiLSTM integrated model significantly enhance prediction performance compared to single-model approaches.

Keywords: Ensemble learning; GCN model; BiLSTM model.

1. Introduction

In the logistics network, the sorting center is the key node, and the accurate prediction of its cargo volume is of great significance to improve the overall efficiency and reduce the operating cost. Traditional forecasting methods are mostly based on time series analysis, but these methods often ignore the spatial characteristics and complex time dependence. By synthesizing GCN, BiLSTM, ARIMA and Transformer models, this study makes full use of the spatial information, time series features and series modeling capabilities of the sorting center to improve the accuracy of cargo volume prediction.

In 2007, Ju Songdong et al. put forward the theory of logistics network [1]. With the passage of time, a group mainly led by Yan Lijun et al. studied the logistics network of cities [2]. As for the cognition of logistics network, many famous scholars in China have conducted researches in various aspects: Xu Jie et al provided the theoretical research method of logistics network [3], pointing out that the study of logistics network is conducive to the further development of logistics enterprises. Chen Guorong et al. proposed a logistics network modeling method named NGM [4], and Huang Jicong published a method of applying mathematical modeling to logistics networks [5], which introduced logistics network models and established corresponding model algorithms to further promote the development of logistics management systems. Recently, Du Xiaohui et al. 's research on resource allocation of logistics network under random environment has made the logistics network model more effective [6].

In the multiple links of logistics network, the existence of sorting center is undoubtedly the core of the whole network, and the research on sorting center has become the academic focus. A large number of scholars have studied the logistics prediction and the distribution of manpower: Hou Rui et al. MLP neural network model was proposed [7], Zhang Jin established prediction theory to study logistics demand [8], Geng Yong et al proposed neural network to analyze and forecast logistics

demand [9], Zhang Peng further studied the logistics prediction method of neural network and proposed a three-layer neural network learning algorithm [10].

2. Correlation methodology

2.1 Establishment of GCN-BiLSTM model

2.1.1 GCN model

When GCN processes graph structure data to extract feature information and relationships between nodes, we use ChebNet convolution operation. The core idea of GCN is to extract information from adjacency matrices and node feature matrices through multi-layer graph convolution operations. GCN processes graph data containing nodes whose features form a dimensional feature matrix X , and the relationships between nodes form a dimensional adjacency matrix. See Equation (1) for the specific formula.

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (1)$$

Finally, according to the node characteristics and connection relations learned by the GCN model and the edge weight information, we can evaluate the importance and load of different transport lines, so as to analyze and forecast the sorting center network.

2.1.2 BiLSTM model

LSTM (Long Short-Term Memory network) solves the problem of long-term dependence in traditional recurrent neural networks and makes them perform better in processing time series data [2]. However, since the transmission of the cell state in LSTM is one-way from front to back, only the information of the past moment can be used, and the information of the future moment cannot be directly obtained. To solve this problem, BiLSTM (Bidirectional LSTM, Bidirectional Long short-term memory Network) was proposed. BiLSTM takes advantage of the flow of information in two directions, forward and backward, enabling it to consider both past and future information. Through recursion and feedback, BiLSTM can learn the characteristics of future volumes while using past information.

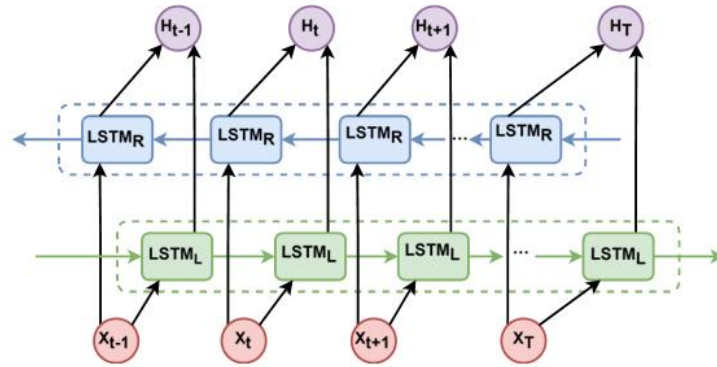


Fig.1 BiLSTM network structure

Assume that \vec{h}_t is the hidden layer state of the forward LSTM network at a certain time, and its calculation formula is shown in equation (2).

$$\vec{h}_t = LSTM(x_t, \vec{h}_t - j) \quad (2)$$

Here x_t is the input at this time t and \overleftarrow{h}_{t-1} is the hidden layer state $t - 1$ of the forward LSTM network at some point. h_t is the hidden layer state of the LSTM network at t certain time, and its calculation formula is shown in equation (3) :

$$\overleftarrow{h}_t = LSTM(x_t, \overleftarrow{h}_{t-1}) \quad (3)$$

Here x_t is the input t at this time, and \overleftarrow{h}_{t-1} is the hidden layer state $t - 1$ of the later LSTM network at a certain time. The output of the BiLSTM network is the combination of two hidden layer states, \overrightarrow{h}_t and \overleftarrow{h}_t , to form the entire hidden state network.

2.1.3 Transformer model

The Transformer model is a deep learning architecture that processes sequential data through a self-attention mechanism. It consists mainly of an encoder and a decoder. The encoder transforms the sequence data into fixed-dimension vector representations using input embeddings and positional encodings, and captures the dependencies between positions in the sequence through multi-head attention mechanisms. Feed-forward neural networks and layer normalization further enhance the stability and performance of the model.

2.1.4 GCN-BiLSTM-Transformer Model

First, the original cargo volume time series data of 57 sorting centers were pre-processed, including data cleaning, missing value handling, and standardization. The dataset was then divided into training, validation, and test sets. The spatial features of the transport connection information of the sorting centers were extracted using a two-layer GCN network, and these features were input into a BiLSTM network. The BiLSTM network leverages bidirectional information flow to capture temporal patterns in the time series data. During training, dropout regularization was employed to reduce overfitting, and the model was optimized using the Adam optimizer by minimizing the mean squared error (MSE) loss between the predicted and actual cargo volumes.

In addition to the GCN-BiLSTM model, a Transformer model was integrated to enhance the prediction performance. The Transformer model processes the time series data with its self-attention mechanism, capturing long-range dependencies and complex temporal patterns. The outputs from the Transformer model were combined with those from the GCN-BiLSTM and ARIMA models to improve prediction accuracy. The final model evaluation utilized root mean square error (RMSE) and mean absolute error (MAE) as performance metrics to assess the prediction accuracy. The RMSE and MAE values reflect the absolute errors between the actual and predicted cargo volumes.

3. Experimental analysis

Compared with traditional LSTM, BiLSTM has significant advantages in that its bidirectional information flow mechanism enables the model to consider both past and future cargo information at the same time, so as to capture long-term dependencies in sequence data more comprehensively and improve the forecasting ability of the model. We use an integrated learning model based on GCN and Bidirectional Long Short-term Memory Network (BiLSTM) to predict the amount of goods per hour over the next 30 days. When analyzing the Annex 2 data, we found that the data is clearly cyclical, with distinct peaks and troughs occurring at different points in time each day. Therefore, the integrated learning model based on GCN and BiLSTM is used for improvement. We determine the weight of the two models by minimizing the root mean square error (RMSE) of each sorting center to obtain the final prediction result.

Table 1. Experimental results of cargo volume prediction for the next 30 days at SC33 sorting point

Date	RMSE	MAE
2023/12/1	3.534	4.414
2023/12/2	3.459	4.358
2023/12/3	3.745	4.574
2023/12/4	3.862	4.682
2023/12/8	4.956	4.963
.....
2023/12/23	4.149	6.534
2023/12/26	6.368	6.862
2023/12/29	7.283	8.635
2023/12/30	9.579	6.956

According to the experimental results of the cargo volume prediction in the next 30 days at the SC33 sorting point of a sorting center, the GCN-BiLSTM model has better prediction performance in the prediction of different sorting centers. Both the RMSE error and MAE error of these models appear to be on the rise as the number of forecast days increases. On the forecast day, the GCN-BiLSTM model outperformed the LSTM method in terms of RMSE. The GCN-BiLSTM model, combined with spatial factors and bidirectional time flow, performs better than the LSTM model which only considers a single factor.

Visualize the prediction of SC43 sorting points, as shown in the figure 2 below.

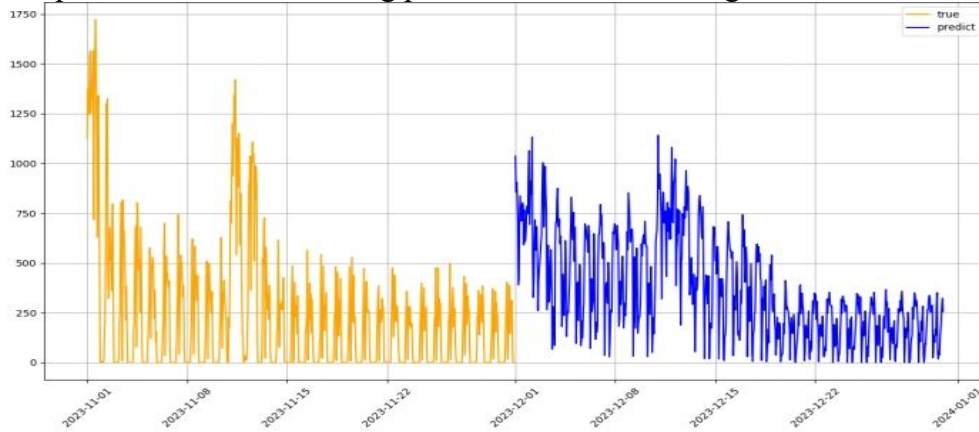


Fig. 2 SC43 sorting point volume forecast for the next 30 day

By comparing the evaluation indicators of multiple sorting centers, we found that the integrated learning model based on ARIMA and BiLSTM is better than the integrated learning model based on LSTM and ARIMA in terms of cargo volume prediction, reflecting the advantages of BiLSTM model. Meanwhile, the attention mechanism is used to improve the model's attention to the data in a specific period of time. The GCN-BiLSTM model predicts the cargo volume in the next 30 days after the change of the transportation line network of the sorting center, and the methods such as restricting and improving the cargo volume in the next 30 days and hours also improve the accuracy of the prediction accuracy.

4. Conclusion

In this paper, we propose a comprehensive learning model based on GCN-BiLSTM, ARIMA, and Transformer models, which shows significant advantages in improving the accuracy of cargo volume predictions for sorting centers. Firstly, the GCN network extracts spatial features of the transport connection relationships between sorting centers, and the BiLSTM network learns the sequential features of time series data, enabling the model to effectively capture the relationships

and structural features between sorting centers, thus enhancing the modeling ability of time series data. Additionally, integrating the ARIMA model and Transformer model further improves the prediction performance. The Transformer model excels in capturing long-term dependencies and sequence characteristics, providing more accurate future cargo volume forecasts.

References

- [1] Wang Zeyu, Zhang Zhiqing. Comparative study on performance of LSTM and GRU models for logistics demand forecasting in Hubei Province [J]. Logistics Engineering and Management, 2019,46(04):10-14.
- [2] Tong Linlin. Cost prediction of cross-border e-commerce logistics based on neural network [J]. Logistics Technology, 2019,46(04):48-51.
- [3] Zheng Yongyou, Du Yan. Research on Shanghai Air logistics forecasting based on grey GM(1,1) model [J]. China Storage and Transportation, 2022,(10):76-77.
- [4] Du Xiaohui, Meng Fanfeng. Logistics network traffic distribution based on random environments research [J]. Water transport in China, 2021 (8) : 97-99..
- [5] Max W ,Konstantin M ,Johannes F , et al.Towards predictive analytics in internal logistics - An approach for the data-driven determination of key performance indicators[J].CIRP Journal of Manufacturing Science and Technology, 2023,44:116-125.
- [6] I. V R ,A. V S .Logistic Forecasting of Aircraft Nitrogen-Charging Service of Air Units in Peacetime and in Wartime[J].Russian Aeronautics, 2022,64(4):606-612.
- [7] Li Minjie, Wang Jian. Research on cold chain logistics demand prediction of aquatic products based on RBF neural network [J]. China Agricultural Resources and Regional Planning, 2020,41(06):100-109.
- [8] Nie Yiwen, Wang Wei. Forecasting of logistics demand in Anhui Province based on PCA-GA-BP neural network model [J]. Journal of Fuyang Normal University (Natural Science Edition), 2023,40(04):96-102.
- [9] Jonna, Fang Qing. Research on demand prediction of agricultural products logistics in Hubei Province -- Based on GM(1,N) model [J]. Logistics Engineering and Management, 2019,45(10):66-69+74.
- [10] Jianjun W ,Yaqian X ,Shuo Z , et al.A Hybrid Total Logistics Forecasting Method Combined with ARIMA and BPNN[J].Journal of Physics: Conference Series, 2021,1769(1):012051-..