INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 19, Issue: 1, Month: February, Year: 2024
Article Number: 6149, https://doi.org/10.15837/ijccc.2024.1.6149

CCC Publications

Improving Short-Term Traffic Flow Prediction using Grey Relational Analysis for Data Filtering and Stacked LSTM Modeling

Zhizhu Wu, Mingxia Huang, Zhibo Xing, Tao Yang

Zhizhu Wu, Mingxia Huang*, Zhibo Xing
School of Transportation and Geomatics Engineering,
Shenyang Jianzhu University, Shenyang, 110168, China
*Corresponding author: mingxia@sjzu.edu.cn

Tao Yang
China railway Shenyang Bureau Group Co, Ltd, Shenyang, 110000, China

Abstract

Traffic flow prediction is one of the critical measures to alleviate traffic congestion. Currently, traffic flow prediction research has made some achievements, but there are still some deficiencies. In order to solve the problems of low prediction accuracy, poor real-time performance, and high data dimensions. This paper proposes a new traffic flow prediction method that employs Grey Relation Analysis (GRA) to detect the correlation between detection points, remove insignificant or uncorrelated traffic flow data points, and hence reduce the data dimensionality of the prediction model. Multiple Long Short-Term Memory (LSTM) models are then stacked to establish the traffic flow prediction model, considering that traffic flow is affected by multi-dimensional spatiotemporal factors, incorporating vehicle speed, occupancy, and traffic volume as inputs. We conducted experiments on real datasets, and the results showed that our GRA-SLSTM model improved prediction accuracy by 3.6% compared to other models, while reducing model prediction time by 27.33%. The proposed model's generalization ability is validated by predicting other detection points, which provides significant references for traffic flow prediction research and practical applications.

Keywords: Traffic flow prediction, GRA-SLSTM, Grey Relation Analysis, Long Short-Term Memory Network, Deep Learning.

1 Introduction

With the increasing number of vehicles each year, the limited road resources cannot meet the growing demand for transportation [1]. This leads to traffic congestion, which not only affects people's travel efficiency but also increases the risk of traffic accidents [2]. Real-time traffic flow prediction based on the analysis of traffic big data helps traffic management authorities grasp accurate road conditions and changing trends, enabling effective traffic guidance and control. Therefore, accurate Short-Term traffic flow prediction plays a crucial role in alleviating traffic congestion and reducing traffic accidents.

Nowadays, the methods for traffic prediction mainly include statistical-based methods, machine learning methods, and deep learning methods. The statistical-based methods utilize historical data to predict future traffic data. These historical data can include traffic volume, speed, congestion status, traffic accidents, etc. Time series modeling is applied to these data, and statistical methods are used to predict future values. Popular statistical theory prediction methods include Autoregressive Integrated Moving Average (ARIMA) models and their variants [1, 4], Kalman filtering algorithm [5, 6], etc. With the development of related theories, people have started to consider using machine learning algorithms to model and predict traffic data [7]. These algorithms mainly include K-Nearest Neighbor (KNN) algorithm [8, 9], Support Vector Machine (SVM) [10, 11, 12], etc. Deep learning, as a variant of machine learning methods, can explore the spatio-temporal information of traffic flow in-depth [13, 14, 15]. When deep learning emerged in the field of transportation, many scholars utilized convolutional neural networks to extract spatial features between traffic data [16, 17, 18]. Since convolutional neural networks can only perform convolution operations on Euclidean structures to extract features, researchers divided the road network topology into grids for subsequent processing. However, real-life traffic structures often exhibit non-Euclidean structures. Therefore, dividing the road network topology into grid forms may to some extent undermine the road network structure and affect the extraction of complex features and it is also the limitations of current state-of-the-art approaches. At this point, graph neural networks [19] emerged. Subsequently, more and more research has been conducted using graph neural networks to solve traffic flow prediction problems. Combined models that simultaneously extract spatio-temporal features have also achieved remarkable results in the field of traffic data [20, 21, 22].

In summary, current research in the field of traffic flow prediction mainly focuses on modeling complex spatio-temporal correlations to improve prediction accuracy. However, as the complexity of models increases, the computational time cost also increases multiple times, which is not suitable for the practical application of short-term traffic flow prediction. Therefore, taking into account both prediction accuracy and computational time cost, we propose a short-term traffic flow prediction method called Grey Relation Analysis with Long Short-Term Memory Network (GRA-SLSTM). Specifically, we use GRA to analyse the correlation between the traffic volume at the target detection point and the traffic volume at surrounding detection points. We select the detection points with higher correlation ranks to achieve data compression. Then, considering the correlation among traffic flow parameters, we construct a dataset using the traffic volume, vehicle speed, and occupancy rate data at the target detection point. We input this dataset into a stacked LSTM model to predict the future traffic volume at the target detection point. By comparing with existing baseline models, we have demonstrated that our proposed method not only achieves good prediction performance but also reduces the computational time cost of the model, making it more practical and applicable.

2 Methodology

2.1 Grey relation analysis

Grey Relational Analysis (GRA) is a method used to study the correlation and influence factors between variables. The improved GRA model and grey correlation degree are used to calculate the indicators in China, and suggestions are put forward based on the public data and traffic status of Hebei provincial transportation department. It determines the contribution of these factors to the system by comparing their degree of correlation. In the same sample data column, if two factors have similar trends in terms of direction, magnitude, speed, etc., their correlation degree is high. The specific steps are as Figure 1:

We utilize GRA to analyse the correlation between various detection points. We select the flow monitoring values of the detection points to construct the reference sequence X_0 and the compared sequence X_n, specifically as shown in Equation 1 and Equation 2.

$$X_0 = (x_0(1), x_0(2), \dots, x_0(m)) \tag{1}$$

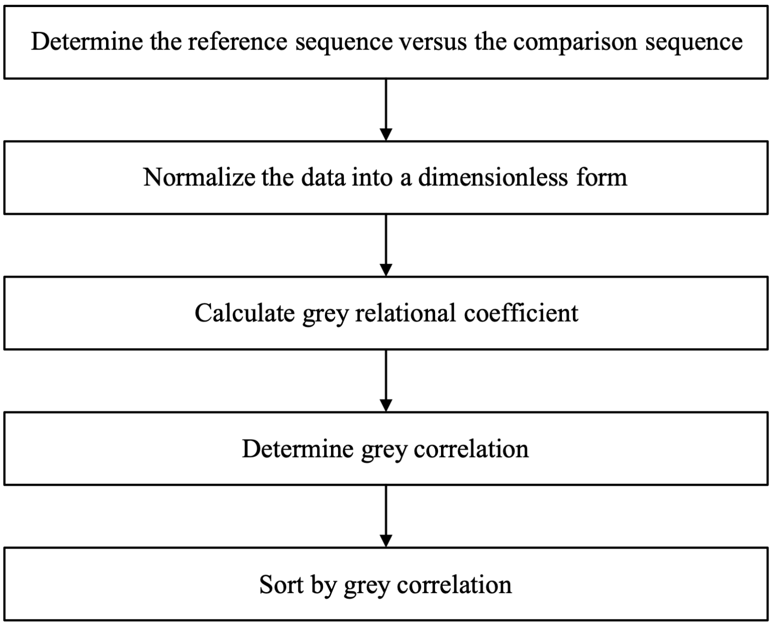

Figure 1: Specific Steps of Grey Relation Analysis Method

$$(X_1, X_2 \cdots, X_n) = \begin{pmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(m) & x_2(m) & \dots & x_n(m) \end{pmatrix} \tag{2}$$

In the equation, m represents the number of indicators and n represents the number of compared sequences.

After constructing the required sequences, calculate the correlation coefficient according to Equation 3.

$$\text{☼}_i(k) = \frac{\min\limits_i |x_0(k) - x_i(k)| + \rho \cdot \max\limits_i |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot \max\limits_i |x_0(k) - x_i(k)|} \tag{3}$$

In the equation, $k = 1, 2, \dots, m$; $i = 0, 1, 2, \dots, n$; ρ represents the resolution coefficient, with a value range of $0 < \rho < 1$. Generally, a value of 0.5 is commonly used.

Calculate the average correlation coefficient between the remaining indicators of each compared sequence and the corresponding elements of the reference sequence. The calculation method is shown in Equation 4.

$$r_{0i} = \frac{1}{m} \sum_{k=1}^{m} \text{☼}_i(k) \tag{4}$$

2.2 Long Short-Term Memory

Long Short-Term Memory is designed to solve the problem of long sequence memory and gradient disappearance or explosion in traditional recurrent neural networks. It is a special type of recurrent neural network (RNN) . LSTM can effectively process long sequence data because it controls the information flow through gate mechanisms, allowing the network to preserve or discard information. The structure of an LSTM includes an input layer, a forget layer, an output layer, and a memory layer. The input layer receives external inputs, the forget layer controls whether or not the information in the memory cell is retained, the output layer generates the network's output, and the memory layer stores the current state information. LSTM is typically trained using backpropagation, with the goal of minimizing the error between the predicted result and the actual result. The LSTM structure unfolded over time is shown in Figure 2, and the structure of the hidden layer units is shown in Figure 3.

Assuming the number of hidden units is h, and given a mini-batch input $x_t \in R^{n\times d}$ (with n samples and n inputs) at time step t, along with the previous time step's hidden state $h_{t-1} \in R^{n\times h}$, the formulas for the forget gate $F_t \in R^{n\times h}$, input gate $I_t \in R^{n\times h}$, and output gate $O_t \in R^{n\times h}$ are calculated as shown in Equations 5, 6 and 7, respectively.

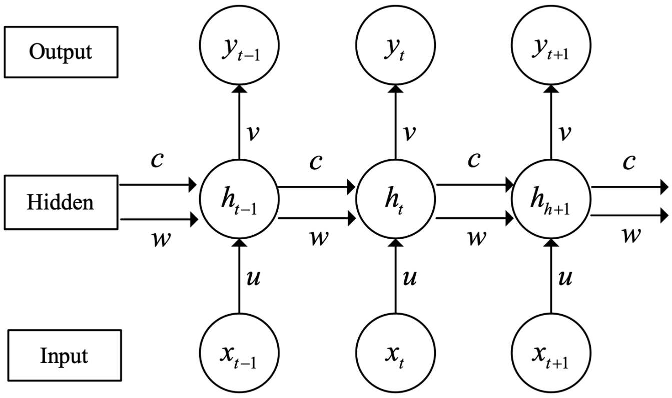

Figure 2: LSTM Unfolding over Time Line

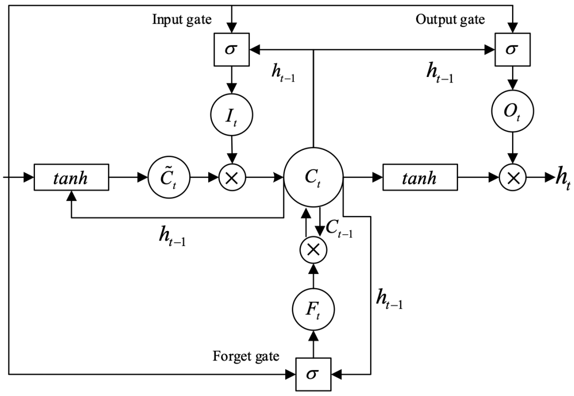

Figure 3: Cell Structure of LSTM

$$F_t = \sigma\left(x_t W_{x_f} + h_{t-1} W_{h_f} + b_f\right) \tag{5}$$

$$I_t = \sigma\left(x_t W_{x_i} + h_{t-1} W_{h_i} + b_i\right) \tag{6}$$

$$O_t = \sigma\left(x_t W_{x_o} + h_{t-1} W_{h_o} + b_o\right) \tag{7}$$

In the equation, $W_{x_f}, W_{x_i}, W_{x_o} \in R^{d\times h}$ and $W_{h_f}, W_{h_i}, W_{h_o} \in R^{h\times h}$ are weight parameters, and $b_f, b_i, b_o \in R^{l\times h}$ is a bias parameter.

In addition, the candidate value for cell update C_t is determined by the tanh function, and the calculations for C_t and C_t are as shown in Equations 8 and 9, respectively.

$$C_t = \tanh\left(x_t W_{x_c} + h_{t-1} W_{h_c} + b_c\right) \tag{8}$$

$$C_t = I_t C_t + F_t C_{t-1} \tag{9}$$

The calculation of the input h_t for the fully connected layer is shown in Equation 10.

$$h_t = O_t \tanh\left(C_t\right) \tag{10}$$

The above formulas describe the basic operations of a single LSTM unit. Multiple LSTM units can be stacked together to form an LSTM network, which can be used for tasks like traffic flow prediction by learning to model the temporal dynamics of the data.

2.3 Stacked LSTM

Based on this, this study stacks multiple LSTM layers to form a Stacked LSTM for traffic flow prediction. By stacking multiple LSTM layers, the model's expressive power can be further enhanced, thereby improving its predictive performance in time series data and reducing the risk of overfitting to some extent. The Stacked LSTM model is suitable for accurately predicting traffic flows on both weekdays and non-working days. The stacked LSTM model in this study is illustrated in Figure 4, where the hidden state from the previous layer is used as input for the next layer. The weight and bias parameters are the same for each layer of LSTM.

The output $h_T^{(n)}$ of the model's output layer is also the input of the fully connected layer. W and b represent the weight matrix and bias term between the fully connected layer and the output layer of LSTM, respectively. The calculation method for the predicted traffic flow value $\hat{h}_T$ is shown in Equation 11.

$$\hat{h}_T = W h_T^{(n)} + b \tag{11}$$

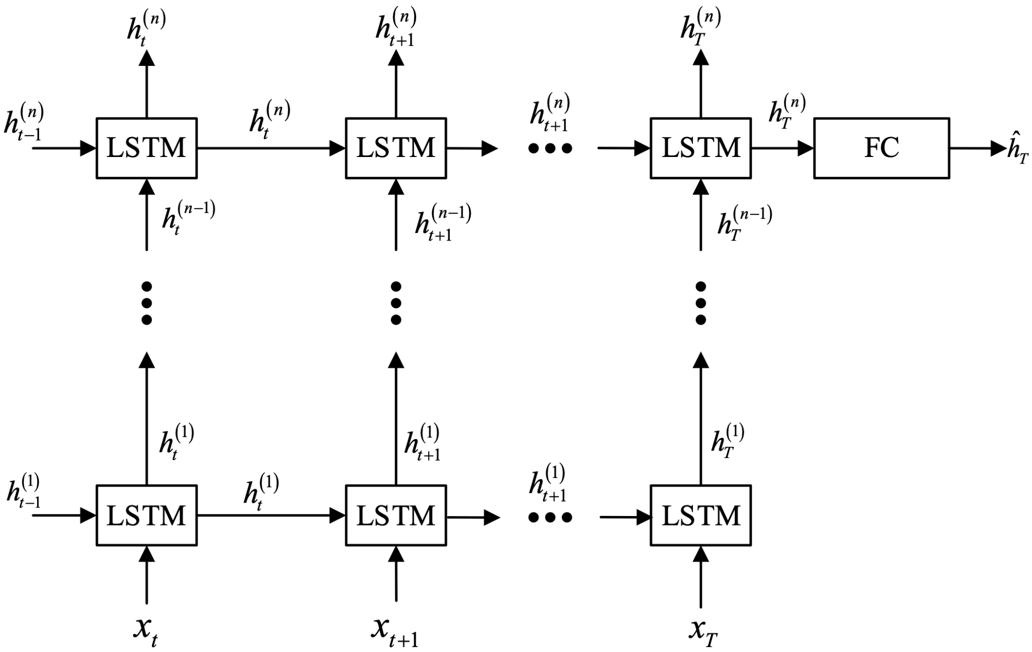

Figure 4: Stacked LSTM

3 Data processing

After fully considering factors such as the source, accuracy, and completeness of the data, we ultimately selected the traffic volume, speed, and occupancy data recorded by the detector numbered "VDS 1209076" and its surrounding 18 detectors in the PeMS dataset from March to April 2022 as the initial sample. These detectors upload data every 5 minutes, resulting in a total of 288 data points per day. From March to April, a total of 316,224 data points were generated. We used the GRA to analyse the spatial correlation of the detectors.

To avoid the influence of different ρ values on the correlation, we next compared and analysed the correlations between each detector and the target detector under different ρ values. The correlation results are shown in Figure 5. In order to compare the variations in correlation among the detectors, the ranking of the correlation for each detector under different ρ values is given in Table 1.

Table 1: Ranking table of correlation degree of each detection

Detector ID	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0.4$	$\rho = 0.5$	$\rho = 0.6$	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$
VDS1203124	12	12	12	12	11	11	11	10	12
VDS1209092	9	9	9	9	9	9	9	9	9
VDS1220773	7	7	7	7	7	7	7	7	7
VDS1220790	3	3	3	3	3	3	3	3	3
VDS1205088	1	1	1	1	1	1	1	1	1
VDS1204694	12	12	12	12	11	11	11	10	12
VDS1210173	17	17	17	17	17	17	16	16	16
VDS1210661	16	16	16	16	16	16	16	16	16
VDS1212333	1	1	1	1	1	1	1	1	1
VDS1204716	4	4	4	4	4	4	4	4	4
VDS1211309	9	9	9	9	9	9	9	10	9
VDS1204442	18	18	18	18	18	18	18	18	18
VDS1202337	12	12	12	12	11	14	11	10	12
VDS1221399	11	11	11	9	11	11	11	10	9
VDS1202889	8	8	7	7	7	8	8	7	7
VDS1205368	15	15	15	15	15	15	15	15	15
VDS1202942	6	5	5	5	5	5	5	5	5
VDS1208898	5	5	5	5	5	5	5	5	5

From Table 1 and Figure 5 , it can be observed that within the normal range of $\rho(0 < \rho < 1)$, the correlation between each detector and the target detector increases with an increase in ρ. However, in overall terms, the ranking of correlation for each detector remains almost unchanged. Among them,

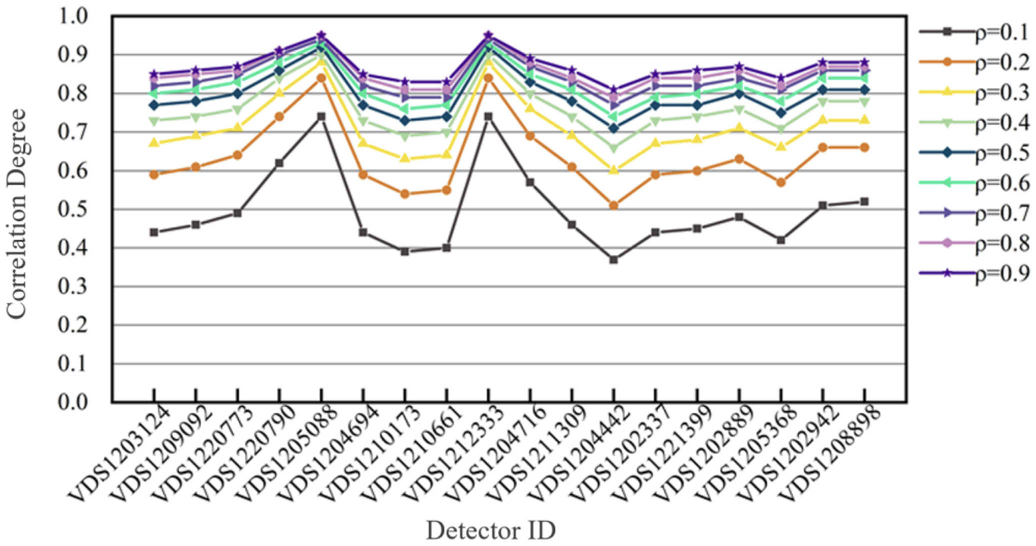

Figure 5: Correlation Between Each Detector and Target Detection Point When ρ Is Different

five detectors with the IDs VDS1220790, VDS1205088, VDS1212333, VDS1204716, and VDS1208898 consistently rank among the top 5 in terms of correlation with the target detector. Therefore, in this study, we preliminarily select the data from these five detectors for a total of 61 days from March 1, 2022, to April 30, 2022, as the input for the subsequent model.

4 Comparative experiments and analysis of results

4.1 Experiment setup

All experiments were conducted on an NVIDIA GeForce RTX 2060S GPU and an Intel i7 12700 CPU. The LSTM hidden layer was set to 32 with 128 neurons in the hidden layer. The batch size was set to 64, the number of epochs was set to 100, and the learning rate was set to 0.01. Dropout was employed to optimize the model training process, and Mean Squared Error (MSE) was chosen as the loss function. Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) were selected as the final evaluation metrics, calculated according to Formulas 12, 13, and 14, respectively.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|f_i - y_i| \tag{12}$$

$$MAPE = \sum_{i=1}^{n}\left|\frac{f_i - y_i}{y_i}\right| \times \frac{100}{n} \tag{13}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(h\left(x_i\right) - y_i\right)^2} \tag{14}$$

In the formulas, y_i represents the true value, and f_i represents the predicted value corresponding to the true value.

4.2 Experiment results

The comparison between the predicted values and the true values for April 29 and April 30, 2022 obtained from the model is shown in Figure 6.

GRA-SLSTM was compared with the following models: ARIMA, GRU, LSTM, and SLSTM. The specific evaluation metrics and comparison results are shown in Table 2.

According to Table 2, GRA-SLSTM demonstrates better performance compared to traditional forecasting models. Compared to the standalone SLSTM model, GRA achieves data compression and

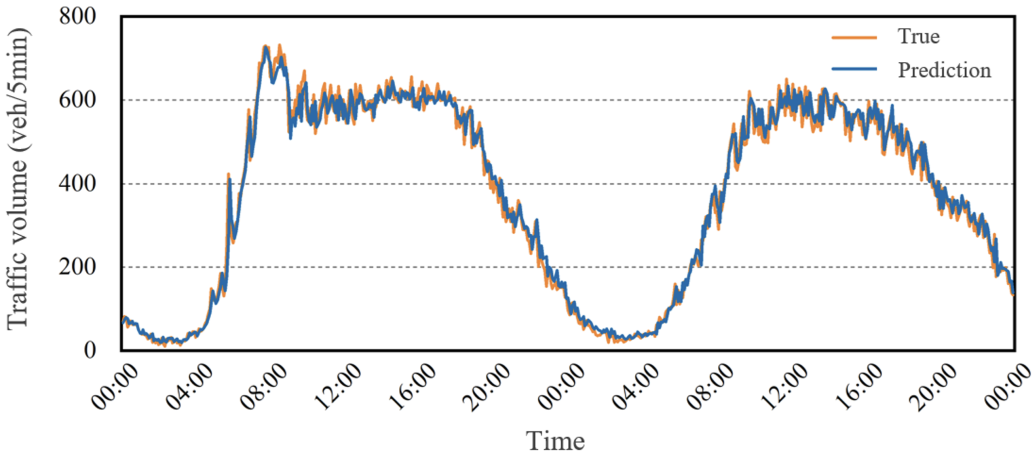

Figure 6: Predicted and Real Data

Table 2: Comparison of prediction results of different models

Model	MAE	RMSE	MAPE	Running time(s)
ARIMA	27.95	38.75	15.23	1.54
GRU	25.23	35.22	10.77	2.01
LSTM	24.57	32.31	9.76	2.83
SLSTM	23.75	29.86	9.35	3.22
GRA-SLSTM	**23.41**	**30.13**	**9.01**	**2.34**

dimensionality reduction in the initial stage, reducing the prediction time by 27.33%. Additionally, due to the consideration of spatial correlation in GRA, the prediction accuracy is improved by 3.6%. In order to validate the generalization ability of the proposed model, target detectors were randomly selected from District 11, District 7, and District 4, and GRA-SLSTM was used to predict their traffic flow. The results are shown in Table 3, which indicates that the GRA-SLSTM model exhibits stable performance and possesses strong generalization ability across different road segments.

Table 3: Prediction effect of GRA-SLSTM

Detector ID	MAE	RMSE	MAPE
VDS 1111542	22.56	30.05	9.33
VDS 718421	23.39	31.02	10.01
VDS 400438	22.92	30.54	9.78

5 Conclusion

This paper focuses on optimizing the runtime of traffic flow prediction models by considering data scaling-down and dimensionality reduction operations on the initial data, while preserving the spatial features of the data. The aim is to reduce the time required by the model while maintaining prediction accuracy. Based on these considerations, the following conclusions were drawn:

(1) A new traffic flow prediction model, called GRA-SLSTM, which combines grey relational analysis (GRA) and LSTM, is proposed in this paper. Experimental results demonstrate that the proposed model outperforms the selected baseline models in terms of prediction accuracy. Specifically, the maximum reduction in Mean Absolute Percentage Error (MAPE) is 16.24%, the maximum decrease in Root Mean Square Error (RMSE) is 22.24%, and the maximum improvement in prediction accuracy reaches 21.86%.

(2) After the data is filtered through GRA, it retains spatial correlation to the maximum extent while compressing the data dimensions, reducing the time required by the model during actual prediction. Therefore, compared to other models with similar prediction accuracy, the GRA-SLSTM model reduces the prediction time cost by 27.33%.

(3) Many prediction models suffer from the problem of poor performance on other datasets. To validate the performance of the proposed GRA-SLSTM model, data from other regions were randomly selected to construct a dataset for testing. The results show that the GRA-SLSTM model exhibits stable performance and strong generalization ability.

References

[1] Yuan„ Z.; Yuan, X.; Yang, Y.; et al. (2023). Greenhouse gas emission analysis and measurement for urban rail transit: a review of research progress and prospects, *Digital Transportation and Safety*, 1(1), 37–52, 2023.

[2] Yang, Y.; Yin, Y. X.; Wang, Y. P.; et al. (2023). Modeling of freeway real-time traffic crash risk based on dynamic traffic flow considering temporal effect difference, *Journal of Transportation Engineering Part A Systems*, 149(7), 04023063, 2023.

[3] Ahmed, M. S.; Cook, A. R. (2023). *Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques*, 1979.

[4] Williams, B. M.; Hoel, L. A. (2003). Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results, *Journal of Transportation Engineering*, 129(06), 664–672, 2003.

[5] Kumar, S. V. (2017). Traffic flow prediction using Kalman filtering technique, *Procedia Engineering*, 187, 582–587, 2017.

[6] Zhang, S. Y.; Yang, Y. H.; Chen, C.; Yang, Y. C.; Li, F. L. (2021). Spatio-temporal prediction of ship traffic flow in waterway regulation area based on regression Kalman Filter combination model, *Journal of Dalian Maritime University*, 47(1), 37–44, 2021.

[7] Yang, Y.; He, K.; Wang, Y. P.; Yuan, Z. Z.; Yin, Y. H.; Guo, M. Z. (2022). Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods, *Physica A: Statistical Mechanics and its Applications*, 595(2022), 127083, 2022.

[8] Davis, G. A.; Nihan, N. L. (1991). Nonparametric regression and short-term freeway traffic forecasting, *Journal of Transportation Engineering*, 117(2), 178–188, 1991.

[9] Liang, Y. P.; Mao, Z. Y.; Zou, W. B.; Xu, R. (2018). Short-term traffic flow prediction based on similar data aggregation and KNN with varying K-value, *Journal of Geo-information Science*, 20(10), 1403–1411, 2018.

[10] Lin, H.; Li, L. X.; Wang, H. (2020). SPGAPSO-SVM: a prediction algorithm of urban bus passenger flow, *Journal of Chinese Computer Systems*, 41(11), 2458–2464, 2020.

[11] Fu, C. H.; Yang, S. M.; Zhang, Y. (2019). Promoted short-term traffic flow prediction model based on deep learning and support vector regression, *Journal of Transportation Systems Engineering and Information Technology*, 19(4), 130–134, 2019.

[12] Hu, W. B.; Yan, L. P.; Lin, K. Z.; et al. (2016). A Short-term traffic flow forecasting method based on the hybrid PSO-SVR, *Neural Processing Letters*, 43(1), 155–172, 2016.

[13] Chen, X. Q.; Zhou, L. X.; Cao, Z. (2020). Short-term network-wide traffic prediction based on graph convolutional network, *Journal of Transportation Systems Engineering and Information Technology*, 20(4), 49–55, 2020.

[14] Li, G.; Peng, L.; Li, X.; et al. (2019). Short-term traffic forecast of urban bus stations based on long short-term memory, *Journal of Highway and Transportation Research and Development (English Edition)*, 13(2), 65–72, 2019.

[15] Luo, X. L.; Li, D. Y.; Yang, H. (2018). Short-term traffic flow prediction based on KNN-LSTM, *Journal of Beijing University of Technology*, 44(12), 1521–1527, 2018.

[16] Yan, Z.; Yu, C. C.; Han, L.; Su, W. J.; Liu, P. (2019). Short-term traffic flow forecasting method based on CNN+LSTM, *Computer Engineering and Design*, 40(9), 2620–2624+2659, 2019.

[17] Luo, W. H.; Dong, B. T.; Wang, Z. S. (2017). Short-term traffic flow prediction based on cnn-svr hybrid deep learning model, *Journal of Transportation Systems Engineering and Information Technology*, 17(5), 68–74, 2017.

[18] Li, L.; Zhang, Q. M.; Zhao, J. H.; Nie, Y. W. (2021). Short-term traffic flow prediction method of different periods based on improved CNN-LSTM, *Journal of Applied Sciences*, 39(2), 185–198, 2021.

[19] Yu, B.; Yin, H. T.; Zhu, Z. X. (2018). Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting, *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18)*, 3634–3640, 2018.

[20] He, W. W.; Pei, B. Y.; Li, Y. T.; Liu, X. Y.; Xu, S. B. (2023). Traffic flow forecasting based on bi-directional adaptive gating graph convolutional networks, *Journal of Transportation Systems Engineering and Information Technology*, 23(1), 187–197, 2023.

[21] Cui, W. Y.; Gu, Y. L.; Zhao, S. L.; Rui, X. P. (2023). A method of predicting short-term traffic flows based on a DGC-GRU model, *Journal of Transport Information and Safety*, 41(2), 121–128, 2023.

[22] Gao, H.; Jia, H.; Yang, L.; Li, R. (2022). An improved CEEMDAN-FE-TCN model for highway traffic flow prediction, *Journal of Advanced Transportation*, 1–20, 2022.

Cite this paper as:

Wu, Z.; Huang, M.; Xing, Z. (2024). Improving Short-Term Traffic Flow Prediction using Grey Relational Analysis for Data Filtering and Stacked LSTM Modeling, *International Journal of Computers Communications & Control*, 19(1), 6149, 2024.
https://doi.org/10.15837/ijccc.2024.1.6149