J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 8433--8447

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Short-term highway traffic speed prediction based on spatio-temporal analysis and attention mechanism

Dedong Xiao[1], Yitong Wang[2], Gang Li[3,✉], Lanjing Xing[1], Jinpeng Zhu[1], Zhongyi Han[2], Tong Li[1], Guangbin Shang[1], Chao Wang[2], Yizhe Han[1]

[1] *Engineering and Technology Department, Shandong Hi-speed Infrastructure Construction Co., Ltd., Jinan, Shandong, 250000, China*

[2] *Digital R & D Center, Shandong Provincial Communications Planning and Design Institute Group Co., Ltd., Jinan, Shandong, 250000, China*

[3] *School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan, Hubei, 430063, China*

ABSTRACT
Short-term traffic speed prediction in Intelligent Transportation System (ITS) provides an important idea for solving traffic problems. To capture the spatio-temporal properties of traffic speed prediction, we proposes a Graph Convolutional Network-Gated Recurrent Units with Attention (GCN-GRUA) mode for expressway. The Graph Convolutional Network (GCN) and Gated Recurrent Unit (GRU) were used to extract the spatial and temporal features of traffic speed, and the attention mechanism was introduced to improve the prediction performance of the model. Experimental results from the real traffic data set of Qingyin Expressway show that the proposed model has a significant improvement in prediction accuracy compared with GCN, GRU and GCN-GRU models. In addition, the importance of speed characteristic variables and exogenous variables on the traffic speed prediction accuracy show that the speed data with the closest time interval has the greatest influence on the traffic speed prediction, followed by the daily cycle characteristics of traffic speed. As the prediction time increases, the relative importance of the velocity characteristic variable remains above 0.6, while the relative importance of the exogenous variable keeps rising.

Keywords: Intelligent Transportation Systems; Traffic Speed Prediction; Spatio-temporal Analysis; Graph Convolutional Network; Gated Recurrent Unit; Attention

1. Introduction

With the rapid development of China's economy and the strong rise of modern information technology, China has entered a new period of rapid economic development in the 21st century [1]. Transportation, as an indispensable and important part of economic development and social operation, is playing an

✉ Corresponding author.
E-mail address: 333174@whut.edu.cn (G. Li).
Received 01 February 2024; Revised 20 April 2024; Accepted 10 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-461

increasingly important role [2]. As a product of social development, highway plays an important role as a bridge in realizing the rapid transportation of passenger and cargo flows between cities, and its characteristics of "safety, efficiency, convenience and comfort" make it become the lifeblood of the national economy and play an important role in promoting the rapid development of the national economy [3-4]. However, with the explosive growth of automobile ownership, traffic congestion is becoming more and more serious, and the rate of traffic accidents remains high, which leads to traffic safety, congestion delays, and other problems to the state and society has brought about significant economic losses [5-6].

In the face of increasingly serious traffic problems, intelligent transportation system came into being, which integrates advanced modern management technology and superior control technology, as well as advanced information technology, communication technology, sensing technology and computer processing technology, collects real-time information of traffic participants, vehicles, traffic management departments of road systems and other traffic elements, and uses excellent control algorithms to scientifically organize, induce and control traffic flow, so as to reduce traffic congestion, reduce traffic accident rate, and improve the traffic environment. Improve the operational efficiency of the transportation system [7-10]. And traffic prediction technology, as an important part of advanced traffic management system and advanced traveler information system in intelligent transportation technology, plays a key role in obtaining the future short-term traffic state and posture [11-12].

Accurate and efficient traffic information prediction is a prerequisite for advanced and proactive traffic management and control methods, and an important part of building ITS and vehicle-road collaboration environments [13-14]. China's highways have long mileage and bear large traffic pressure, how to make full use of the massive traffic data collected by the detection equipment deployed on the highway to make accurate traffic prediction, and to consolidate the foundation for the active traffic management and control methods, is the research direction of the current transportation field focus on. Highway traffic flow has a strong correlation in time and space dimensions [15-16]. In terms of time, the traffic flow information shows obvious periodicity by hourly, daily, weekly and yearly units, while fluctuating according to the specific situation, which can be modeled as time series data to speculate the future traffic information and status. In terms of space, because the traffic status of the highway section is directly related to its upstream and downstream, and influenced by the nearby ramp entrances and exits and toll stations, it has a certain spatial correlation at the roadway network level has certain spatial correlation [17-18]. Considering only the temporal or spatial characteristics of traffic data will lead to the extraction and mining of traffic information is not comprehensive enough, At the same time, if different research methods are taken to analyze the spatio-temporal characteristics of traffic data separately, it will lead to the spatio-temporal characteristics of the traffic information can not be accessed as a whole, and it may lead to the loss of some important information, which will not be able to accurately predict the traffic status [19-20]. Therefore, it is of great significance to adopt suitable research methods to dissect the spatio-temporal characteristics of traffic data and make accurate, efficient and scientific prediction of the traffic state at the highway information detection points for the development of related traffic control measures and intelligent transportation [21-22].

In this paper, we propose GCN-GRUA (Graph Convolutional Network-Gated Recurrent Units with Attention) model based on spatio-temporal analysis for expressway traffic speed prediction. The contributions of this paper are as follows.

1) We use GCN and GRU to extract spatial features and temporal features according to the spatio-temporal characteristics of traffic speed prediction. In addition, we add an attention mechanism to improve the prediction performance of the model.

2) We conduct experiments using real traffic data sets. Compared with GCN, GRU and GCN-GRU models, the model has a good improvement in prediction performance.

3) We analyze the importance of the velocity feature variable and three exogenous variables on the prediction accuracy of the model and find the traffic speed is most affected by the data of the latest time interval. The second is the diurnal periodic characteristics of traffic speed.

The rest of this paper is organized as follows. Section II details the proposed method. Section III presents the preparation of the experiment. The experimental results are reported in Section IV. Section V concludes the paper.

2. Methodology

In this part, we analyze the spatio-temporal relationship of traffic speed prediction, establish GCN-GRUA model, and introduce the basic framework, the time feature extraction layer and spatial feature extraction layer of the GCN-GRUA model.

2.1. Spatio-temporal correlation for traffic speed prediction

Traffic speed prediction uses typical spatio-temporal data, which are embedded in continuous space and change dynamically with time. The spatio-temporal correlation of traffic speed prediction is shown in Figure 1. In the figure, there are three time slices from 10:00 am to 10:10 am in the time dimension, and seven nodes in the space dimension (A to G) represent the road network structure. The purple dotted line indicates that the nodes influence each other in the spatial dimension, the yellow dotted line indicates that the future node is affected by its own historical state, and the red dotted line indicates that the future node is affected by the historical state of its neighbors. Therefore, the traffic speed has obvious spatio-temporal correlation, which also lays the foundation for the prediction of highway traffic speed from the perspective of spatio-temporal analysis.

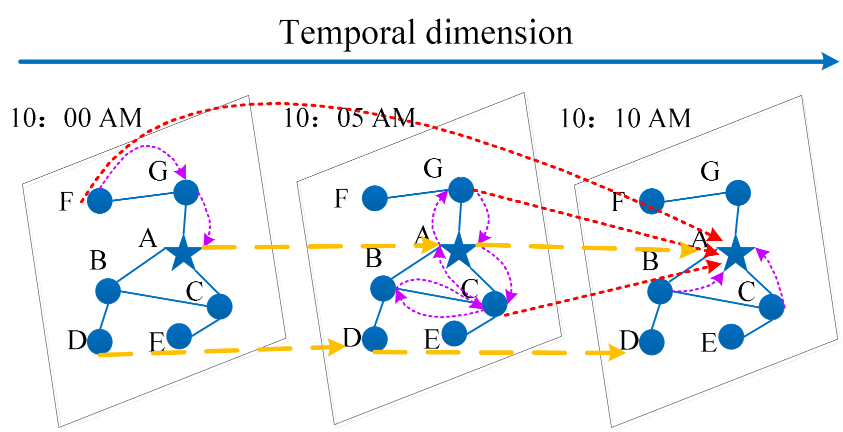


Fig. 1. Spatio-temporal correlation of traffic speed prediction

Traffic speed prediction can predict the speed of each road section in a certain period in the future by observing the historical speed data of the road section. The prediction problem can be described as follows:

$$V_{t+1}, V_{t+2}, \cdots, V_{t+n} = \underset{v_{t+1}, v_{t+2}, \cdots, v_{t+n}}{\arg\max} P\left(V_{t+1}, V_{t+2}, \cdots, V_{t+n} \middle| V_t, V_{t-1}, \cdots, V_{t-m}; G\right) \tag{1}$$

In the equation: $V_{t+1}, V_{t+2}, \cdots, V_{t+n}$ represents the predicted average speed for different time periods on each segment; $V_t, V_{t-1}, \cdots, V_{t-m}$ denotes the vector of observed traffic flow speeds for each road segment at time t, where each vector contains a speed value for every time stamp; $P(|)$ is the conditional probability function.

The GCN-GRUA model framework. We establish the GCN-GRUA model according to the spatio-temporal correlation of traffic speed. The framework of the model is shown in Figure 2, which consists of four parts: traffic speed data input layer, spatial feature extraction layer, temporal feature extraction layer, and prediction data output layer. Among them, the spatial feature extraction layer is constructed by the GCN network, and the temporal feature extraction layer is constructed by the GRU gated unit and the attention mechanism.

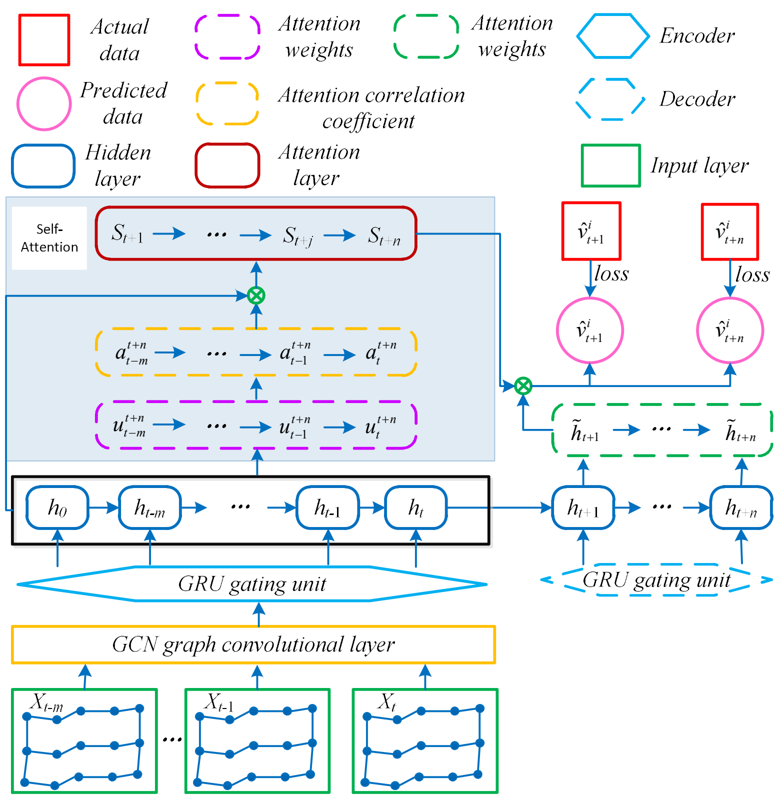


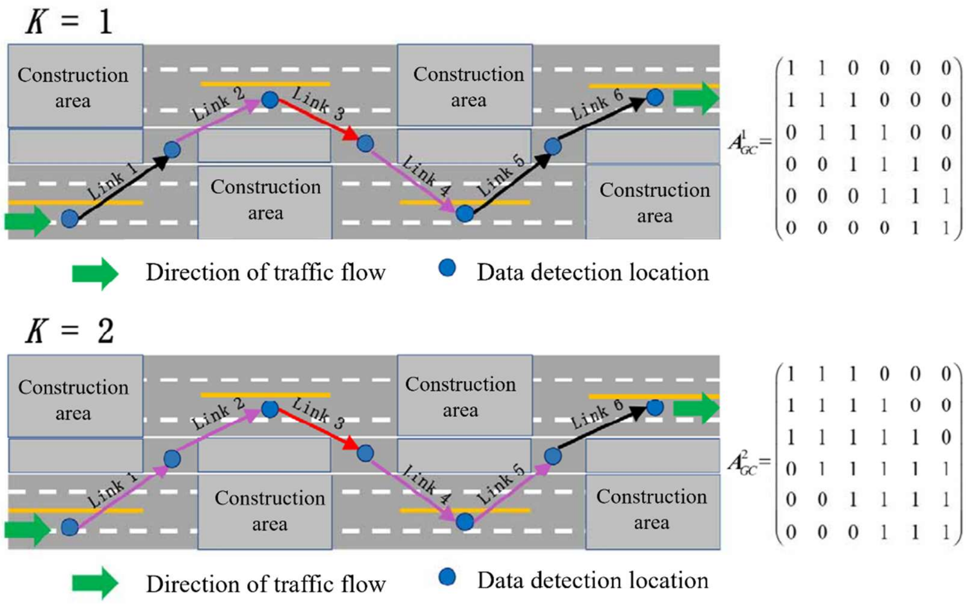
Fig. 2. GCN-GRUA Network Structure

2.2. *spatial feature extraction layer*

In general, graph convolutional networks use undirected graphs, where the adjacency matrix of an undirected graph is symmetric and can be used with the degree matrix to form the Laplacian matrix for performing graph convolution operations. However, for the directed graphs studied in this paper, we use the K-order graph convolution method. This method defines the K-order neighbors of each edge as $H_i(K)=\{l_i \in E | dis(l_i,l_j) \le K\}$, where $dis(l_i,l_j) \le K$ is the minimum number of connections from edge l_i to edg l_j.

To mimic the Laplacian matrix for K-order neighbors, unit diagonal elements are added to the adjacency matrix to obtain $A_{GC}^{K}=Ci\left[(A+I)^{K}\right]$, where Ci is a matrix function that modifies non-zero elements to 1. The assignment of A_{GC}^{K} is given by equation (2). The construction method of the adjacency matrix is illustrated in Figure 3.

$$A_{GC}^{K}=\begin{cases}1, l_j \in H_i(K) ori=j \\ 0, else\end{cases} \tag{2}$$



Link ID	1	2	3	4	5	6
$A^1_{GC}[3]$	0	1	1	1	0	0
$A^2_{GC}[3]$	1	1	1	1	1	0

Fig. 3. Construction of adjacency matrix

Based on this adjacency matrix, the spatial graph convolution expression is given by Equation 3:

$$V_t(K)=\left(W_{GC}\odot A_{GC}^K\right)\cdot V_t \tag{3}$$

where W_{GC} is a trainable parameter matrix of the same size as the adjacency matrix A_{GC}^K, and $\odot$ denotes the Hadamard product, which represents element-wise multiplication of matrices. Due to the characteristics of the adjacency matrix, parameter training occurs only at the K-th neighboring positions, which can be understood as the discrete convolution of the velocity vector V_t over time period t. The elements V_t^i within the vector represent the spatiotemporal fused velocity of the road segment i at time t, which includes information from adjacent road segments. Its one-dimensional convolutional representation is given by Equation 4.

$$\left[V_{t-m}^i(K),\cdots,V_t^i(K)\right]=\left(W_{GC}[i]\odot A_{GC}^K[i]\right)^T\cdot\left[V_{t-m},\cdots,V_t\right] \tag{4}$$

In the equation: $W_{GC}[i]$ and $A_{GC}^K[i]$ are the i row elements of matrices W_{GC} and A_{GC}^K, respectively.

This paper considers embedding external factors into the time dimension model input vector in the spatial dimension. The definitions of external factors E_{t-j} and the input vector X_{t-j} are given by Equations 5 and 6.

$$E_{t-j}=\left[N_{t-j};WT_{t-j};RC_{t-j}\right] \tag{5}$$

$$X_{t-j}=\left[v_{t-j}^i;E_{t-j}\right] \tag{6}$$

2.3. temporal feature extraction layer

In order to capture the time series characteristics of traffic speed, the GRU gated convolution unit, a variant of Recurrent Neural Network (RNN), is used and the Attention mechanism is added as the time dimension modeling. GRU is a kind of Recurrent Neural Network (RNN), which is used to solve the phenomenon of gradient disappearance or gradient explosion in RNN that cannot retain long-term information and backpropagation process. By setting reset gate and update gate, the invalid information is forgotten and the valid information is remembered. Its gating unit composition is shown in Figure 4.

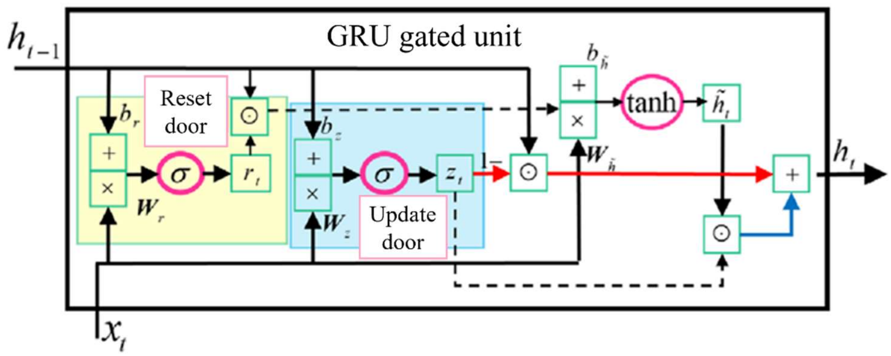


Fig. 4. GRU Gated Unit

After embedding the spatial feature extraction vectors with external factors to obtain the time dimension input vector, the GRU encoder module encodes the input vector X_{t-j} , and the GRU decoder module decodes the predicted output time series to obtain the hidden layer prediction at time $t+j$. The operation of the encoder is illustrated in Equation 7.

$$h_{t-j} = \begin{cases} GRU\left(X_{t-j}, h_0\right), j = m \\ GRU\left(X_{t-j}, h_{t-j-1}\right), j \in \{0, \cdots, m-1\} \end{cases} \tag{7}$$

In the equation: GRU encodes h_t and X_{t-j} using gated convolution; h_0 is initialized as a zero vector. Ultimately, h_t will store all information from the previous hidden layers $\left(h_{t-m}, h_{t-m+1}, \cdots, h_{t-1}\right)$ and the input vector $\left(X_{t-m}, X_{t-m+1}, \cdots, X_{t-1}\right)$, which will be used for decoding to achieve predictions.

The input to the decoder mainly depends on the training method. For Recurrent Neural Networks, there are currently Teacher Forcing and Scheduled Sampling methods. Teacher Forcing uses the ground truth of the training data for the previous step as the input for the next state, which allows for faster and better fitting results. However, this method has the drawback of poor model adaptability and inconsistent input forms during training and testing in multi-step predictions, making it unsuitable for time series models. Scheduled Sampling addresses this drawback by adjusting the probability of choosing the ground truth using the inverse sigmoid function, selecting the ground truth with probability α and the previous prediction with probability $1-\alpha$.

$$h_{t+j} = \begin{cases} GRU\left(v_{t+j}, h_t\right), j = 1 \\ GRU\left(v_{t+j}, h_{t+j-1}\right), j \in \{2, \cdots, n\} \end{cases} \tag{8}$$

In the equation: v_{t+j} represents the observed true data (km/h); h_{t+j} denotes the decoded hidden layer. This way, the predicted hidden layer retains the features of historical traffic flow data while ensuring synchronization of input distributions during training and testing, addressing the issue of inconsistent inputs in Teacher Forcing.

Since Graph Convolutional Networks (GCNs) assume graphs are undirected and treat all neighboring nodes equally, they cannot assign different weights based on node importance. To address the directionality in traffic flow prediction and the specific characteristics of different road segments, the attention mechanism is chosen. This mechanism relies on pairs of neighboring nodes rather than the specific network structure during training, better capturing the periodicity of traffic flow and node relevance. The self-attention mechanism formulas are given in Equations 9-11.

$$u_{t-i}^{t+j} = q^T \tanh\left(h_{t+j} W_f h_{t-i}\right), i = 0,1,\cdots,m \tag{9}$$

$$a_{t-i}^{t+j} = soft\max\left(u_{t-i}^{t+j}\right) = \frac{\exp\left(u_{t-i}^{t+j}\right)}{\sum_{r=1}^{m} \exp\left(u_{t-i}^{t+j}\right)}, i = 0,1,\cdots,m \tag{10}$$

$$S_{t+j} = \sum_{i=1}^{m} a_{t-i}^{t+j} h_{t-i} \tag{11}$$

In the equation: u_{t-i}^{t+j} measures the similarity between the query vector h_{t+j} and each historical state h_{t-i}, and the weights a_{t-i}^{t+j} are obtained through normalization. W_f is the weight matrix to be trained, and q^T is the weight vector to be trained. The attention vector S_{t+j} is obtained by a_{t-i}^{t+j} and weighted summation of the historical states h_{t-i} .

Ultimately, with the cooperation of the attention mechanism and the GRU gated convolution unit, the speed prediction formula of the AGC-GRU model is given in Equations 12-13.

$$h_{t+j} = \tanh\left(W_h \cdot \left[S_{t+j} h_{t+j}\right]\right) \tag{12}$$

$$\hat{v}_{t+j} = W_v \cdot h'_{t+j} + b_v \tag{13}$$

In the equations: $\hat{v}_{t+j}$ represents the predicted speed value (km/h) for the $t+j$ -th time period, W_v is the weight matrix to be trained, h'_{t+j} is the attention hidden layer, and b_v is the offset for generating the predicted speed.

3. Preliminary

In this section, we experimentally evaluate the proposed model and compare it with three other models. Then we analyze the importance of characteristic variables and exogenous factors.

3.1. Data source

The data set used in this study is collected from the construction stage of the reconstruction and expansion of the Qingyin Expressway, and the study area is the section of Henan Interconnect (K242) to Linzi Interconnect (K228), as is shown in Figure 5. The total length of Henan interconnect to Linzi interconnect is 14.65km, and the average daily traffic volume during the reconstruction and expansion period is 32035pcu/d. There are one parallel divisible road G306 around the construction section. Predicting the traffic speed of the construction section and releasing the congestion index in time is of great practical significance for alleviating traffic congestion and improving traffic efficiency.

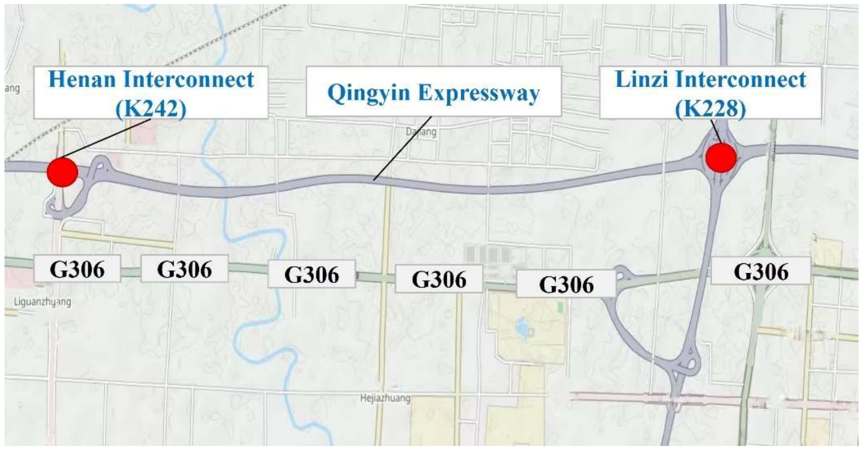


Fig. 5. the section of Henan Interconnect (K242) to Linzi Interconnect (K228)

3.2. Expressway network construction

Figure 6 shows the construction of the expressway network in the study area. There are four construction areas in the construction stage of expressway reconstruction and expansion. a directed graph $G=(V_G,E_G)$ is constructed with each data collection point as the graph node and the connection line along the direction of traffic flow as the edge. The direction of the graph is determined according to the direction of survey traffic flow. Where V_G represents the set of vertices and E_G represents the set of edges of the graph.

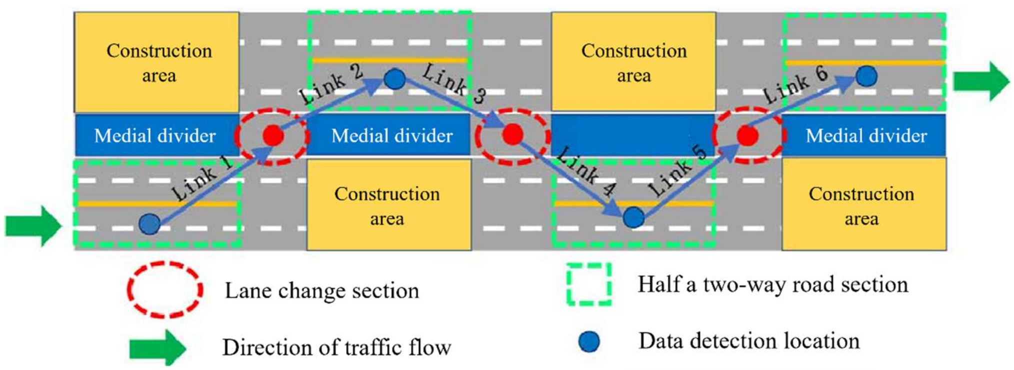


Fig. 6. The construction of the expressway network

The A is defined as the adjacency matrix of the graph, and the dummy variable $A(i,j)$ is used to determine whether nodes i and j are connected.

$$A(i,j)=\begin{cases}1, l_i \text{ and } l_j \text{ are connected} .\\ 0, else\end{cases} \tag{14}$$

The average speed v_t^i of the t^{th} time segment and the $l_i(\forall l_i \in E)$ road segment is defined as a vector $V_t \in R^{|E|}$ (where $|E|$ is the cardinality of the edge set E), and the i^{th} element of the vector is defined as $(V_t)_i = v_t^i$.

3.3. Data acquisition and processing

This study requires topological data of the road network, as well as traffic speed data for each time period. The traffic speed data of the study area are collected by radar observation, and the topology data of the road are collected by UAV observation. A total of 30 days of data were collected from November 2018 to December 2018. Traffic speed data are cleaned, including eliminating fault outliers and filling missing values by linear interpolation. Since the traffic network is a non-Euclidean space, the distance between the data collection points is normalized and used as the initial weight of the connection edge between the road nodes.

Traffic speed prediction is not only affected by the law of traffic speed, but also affected by exogenous factors, such as time factors, road conditions and other factors. Considering exogenous factors helps to provide more information to predict the future traffic speed. Considering the acquisition of traffic speed data, the traffic speed timestamp corresponding to the day time period N_t, the traffic speed timestamp corresponding to the week time period W_t, and the road condition RC of the data collection area were selected as exogenous factors. The first two factors consider the daily and weekly cycle characteristics of traffic speed, and the third factor considers the influence of road conditions on traffic speed prediction.

4. Experiment

In this section, we experimentally evaluate the proposed model and compare it with three other models. Then we analyze the importance of characteristic variables and exogenous factors.

4.1. Ealuation function

In order to evaluate the effectiveness of the prediction results, this paper adopts three error functions as the model error evaluation indicators: Root Mean square error (RMSE), Mean Absolute error (MAE) and Mean Absolute percentage error (MAPE).

The RMSE reflects the dispersion degree of the sample, as shown in Equation 15. Since RMSE is sensitive to abnormal points, RMSE is used as the loss function in the traffic speed prediction and the mini-batch gradient descent algorithm is used to minimize the loss function to update the training parameters.

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(v_i-\overline{v_i}\right)^2} \tag{15}$$

The MAE reflects the error level between the predicted results and the actual results, and the calculated error is shown in Equation 16.

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|v_i-\overline{v_i}\right| \tag{16}$$

The MAPE reflects the percentage of the prediction error to the true value, and the calculated error is shown in Equation 17.

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{v_i-\overline{v_i}}{v_i}\right| \tag{17}$$

In these three formulas: v_i represents the actual speed (km/h); $\overline{v_i}$ represents the predicted speed for future time points (km/h); and n represents the number of samples in the test set.

4.2. Analysis of prediction results

We performed experiments to obtained the traffic speed prediction results from the GCN+GRUA model. Then the comparison test is carried out with the following three models:

1) GCN: GCN model only considers the feature of spatial dimension.

2) GRU: GRU model only considers the feature of temporal dimension.
3) GCN+GRU: GCN+GRU is based on GCN and GRU, which is able to extract both spatial and temporal features to predict traffic speed.

Table 1. Results of single-step (5 min) traffic speed prediction

Model	RMSE	MAE	MAPE (%)
GCN	6.13	4.42	11.88
GRU	6.35	4.68	12.37
GCN+GRU	5.88	4.33	11.25
GCN+GRUA	5.64	4.21	10.25

Table 1 shows the results of single-step (5 min) traffic speed prediction using different algorithms. GCN mainly predicts traffic speed from spatial features, while GRU mainly predicts traffic speed from temporal features. It can be seen that the RMSE value of GCN is 6.13, which is lower than 6.35 of GRU, the MAE value is 4.42, which is relatively lower than 4.68 of GRU, and the MAPE value is 11.88%, which is lower than 12.37% of GRU. It can be seen that GCN performs better than GRU, which indicates that capturing spatial features alone for modeling is more applicable than capturing temporal features alone in single-step short-term traffic flow prediction. It can also be found from the table that GCN+GRU performs significantly better than GCN or GRU algorithms alone when performing short-term traffic speed prediction from both spatio-temporal perspectives. The RMSE value of GCN+GRU is 5.88, which is significantly lower than 6.13 of GCN and 6.35 of GRU. The MAE value of GCN+GRU is also only 4.33, which is lower than that of GCN and GRU. Its MAPE value is 11.25%, which is also lower than the previous two. In addition, through the comparative analysis of GCN+GRU and GCN+GRUA models, it is found that the addition of attention mechanism can effectively improve the prediction performance of the algorithm. The RMSE, MAP and MAEE of GCN+GRUA are 5.64, 4.21 and 10.25%, respectively. It is lower than 5.88, 4.33 and 11.25% of GCN+GRU. It shows that with the addition of the attention mechanism, the model takes more into account the dynamic characteristics of traffic speed data in the spatio-temporal dimension, so it has a more accurate prediction effect.

In real life, five-minute single-step traffic speed prediction time interval is too short to meet the actual needs of traffic travelers and managers. Therefore, it is more practical to realize multi-step prediction of traffic flow and observe the traffic speed information of the next half hour or even one hour. A multi-step traffic speed prediction experiment was carried out, and different models were used to predict the future traffic speed of 30 minutes and 60 minutes at a time interval of 5 minutes, and their prediction performance was compared. The experimental results are shown in Table 2:

Table 2. Results of multi -step (30min、60min) traffic speed prediction

model	RMSE	MAE	MAPE (%)
1. 30min (six -step prediction)			
GCN	8.97	6.89	16.82
GRU	10.40	7.23	17.74
GCN+GRU	7.62	6.02	15.95
GCN +GRUA	6.92	5.58	15.25
2. 60min (twelve-step prediction)			
GCN	13.28	9.98	23.88
GRU	14.69	10.05	22.97
GCN+GRU	11.43	8.61	21.24
GCN +GRUA	10.58	7.26	19.87

As can be seen from the table, the prediction performance of the same prediction model is decreasing as the prediction time increases. In the same prediction model, when the prediction time increases from 30min to 60min, the RMSE of the GCN algorithm increases from 8.97 to 13.28, the MAE increases from 6.89 to 9.89, and the MAPE increases from 16.82% to 23.88%. The RMSE of GRU algorithm increases from 10.40 to 14.69, the MAE increases from 7.23 to 10.05, and the MAPE increases from 17.74% to 22.97%. The RMSE of GCN+GRU algorithm increases from 7.62 to 11.43, MAE increases from 6.02 to 8.61, and MAPE increases from 15.95% to 21.24%. The RMSE of GCN+GRUA algorithm is increased from 6.92 to 10.58, MAE is increased from 5.58 to 7.26, and MAPE is increased from 15.25% to 19.87%. In addition, the prediction accuracy varies between different models as the prediction time changes. When the prediction time is extended from 30min to 60 min, the RMSE, MAE and MAPE of the GCN +GRUA model are improved by 3.66, 1.68 and 4.62%, respectively. The RMSE, MAE and MAPE of the GCN+GRU model are improved by 3.81, 2.59 and 5.29%, respectively. The RMSE, MAE and MAPE of the GCN model are increased by 4.29, 2.82 and 5.23%, respectively. The RMSE, MAE and MAPE of the GRU model are increased by 4.31, 3.09 and 7.06%, respectively. It can be found that the proposed GCN+GRUA has good prediction accuracy, and the error growth rate is the slowest as the prediction time expands. In summary, it is helpful to predict traffic speed using Spatio-temporal Analysis, and add attention mechanism on this basis to improve the prediction performance of the model.

4.3. Importance of characteristic variables and exogenous factors

To explore the importance of the characteristic variables $(V_{t-11}, V_{t-10} \cdots, V_t)$ and three exogenous factor variables (N_t, W_t, RC) considered in the GCN-GRUA model for the prediction accuracy of the model, the random forest method is used to calculate the feature importance by calculating the out-of-bag error of the order subset of the training set and replacing the j variable of the input set X, where the training set is set to $D=\left\{\left\{X^{(1)}, y^{(1)}\right\},\left\{X^{(2)}, y^{(2)}\right\},\cdots,\left\{X^{(m)}, y^{(m)}\right\}\right\}$ (X is the input set parameters, y is the real observation value, The time window of m is taken as 12).The calculation formula is as follows:

$$errO_k = \frac{1}{n}\sum_{i\in O_k}\left(y^{(i)} - \hat{y}^{(i)}\right)^2 \tag{18}$$

$$VI\left(X_j\right) = \frac{1}{K}\sum_{k}\left(errO_k{}^j - errO_k\right) \tag{19}$$

Where, $errO_k$ is the Out-of-bag error of order K for the input set X. $y^{(i)}$ is the true observed value, and $\hat{y}^{(i)}$ is the predicted output value. $VI\left(X_j\right)$ is the importance of the j-th feature variable. $errO_k{}^j$ is the out-of-bag error after permuting the j-th variable of the input set X.

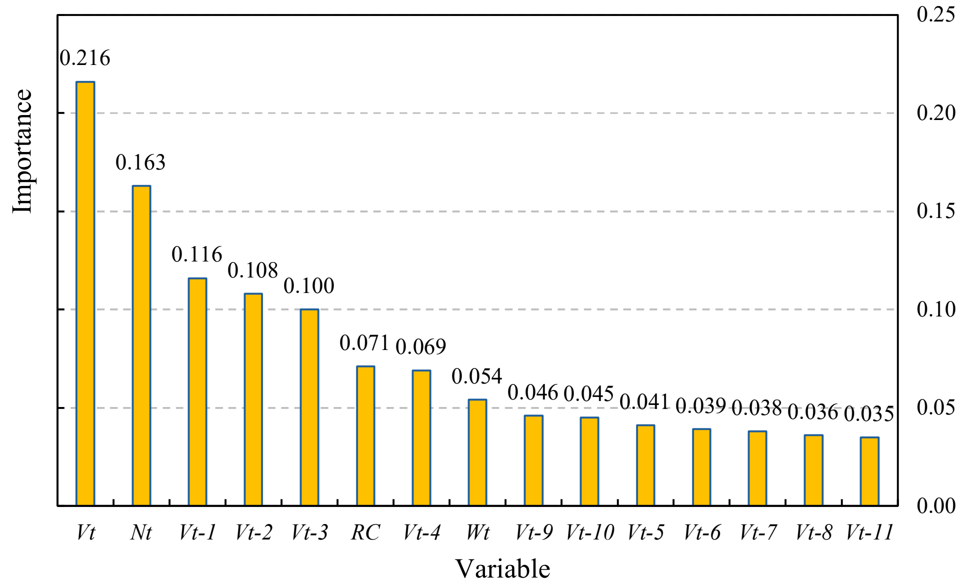


Fig. 7. The importance of every variable for 30min prediction

Figure 7 illustrates the importance of all variables for the 30min prediction with the GCN-GRUA model. It can be found that among all the variables, the importance of Vt is 0.216, which accounts for the most important role, indicating that the traffic speed is most affected by the data of the latest time interval. The second is the daily cycle characteristic of traffic speed Nt, with an importance of 0.163, indicating that the daily cycle characteristic of traffic speed has a great influence on the traffic speed prediction results. This is followed by the observations Vt-1, Vt-2 and Vt-3 of the most recent time interval, which illustrates that the influence of adjacent time interval data in traffic speed prediction is also very large. The importance value of road characteristic RC of traffic speed is 0.071, ranking sixth, indicating that the exogenous factor variable RC has a certain impact on traffic speed prediction. The importance value of Wt is 0.054, which is ranked after the exogenous factors Nt and RC, indicating that the influence of weekly cycle characteristics on traffic speed is less than that of daily cycle and road characteristics. The importance value of Vt-4 ~ Vt-11 is small and the difference is not large. The importance value of Vt-11 is 0.035, which is the lowest of the farthest separated observation data, indicating that in traffic speed prediction, the importance of data separated by longer time may be lower.

In order to explore the relative importance of the characteristic variable $\left(V_{t-11}, V_{t-10} \cdots, V_t\right)$ and the three exogenous factor variables $\left(N_t, W_t, RC\right)$, the set of characteristic variables is named F_1, and the set of three exogenous factor variables is named F_2. The relative importance changes of F_1 and F_2 under different prediction step sizes are explored. The calculation formula of relative importance is as follows:

$$VI\left(F_1\right)=\frac{\sum VI\left(V_{t-m}\right)}{\sum VI\left(V_{t-m}\right)+VI\left(N_t\right)+VI\left(W_t\right)+VI\left(RC\right)} \tag{20}$$

$$VI\left(F_2\right)=\frac{VI\left(N_t\right)+VI\left(W_t\right)+VI\left(RC\right)}{\sum VI\left(V_{t-m}\right)+VI\left(N_t\right)+VI\left(W_t\right)+VI\left(RC\right)} \tag{21}$$

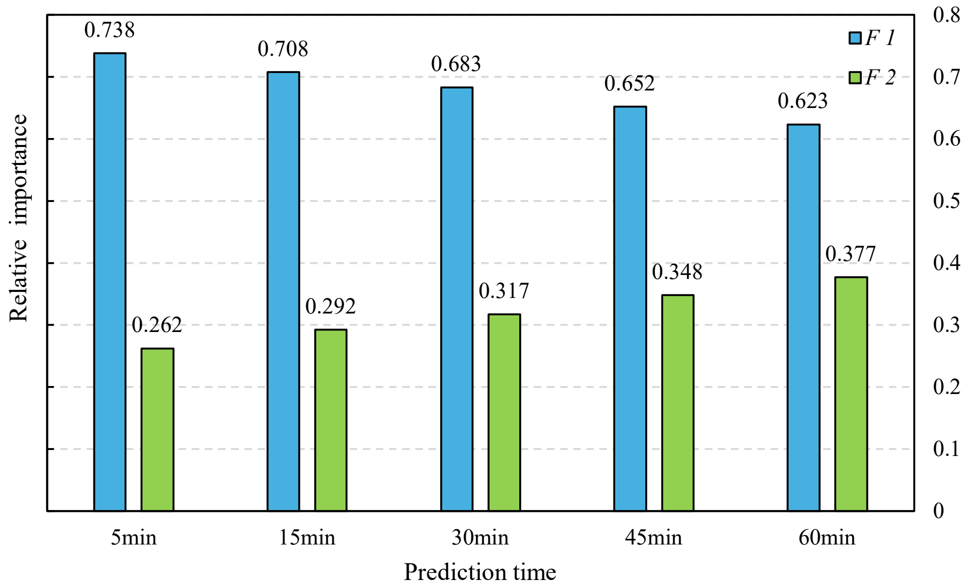


Fig. 8. The importance of F_1 and F_2 at different prediction times

Figure 8 shows the change of the relative importance of F_1 and F_2 at different prediction times. It can be found that F_1 always occupies a relatively important position under all prediction times, which is above 0.6, indicating that historical observation data plays a decisive role in traffic speed prediction. With the increase of prediction time, the relative importance of F_2 gradually increases from 0.262 to 0.377, indicating that exogenous factor variable plays an increasingly important role in traffic speed prediction with the increase of prediction time.

5. Conclusion

In this study, we propose an innovative model, Graph Convolutional Network-Gated Recurrent Unit Based on Spatio-temporal Analysis (GCN-GRUA), for the highway traffic speed prediction problem. The model combines Graph Convolutional Network (GCN) and Gated Recurrent Unit (GRU) to effectively extract the spatial and temporal features of traffic speed data, and introduces an attention mechanism to further improve the prediction performance. Through experiments using real traffic data sets, we find that the GCN-GRUA model significantly outperforms the traditional GCN, GRU and their combination model (GCN-GRU) in terms of prediction accuracy. This result shows that the fusion of spatio-temporal features and the introduction of attention mechanism are of great significance in complex traffic flow prediction.

In addition, we conduct an importance analysis of the influence of the velocity characteristic variable and exogenous variables on the prediction accuracy of the model. The results show that the traffic speed is most significantly affected by the recent time interval data, followed by the daily periodic characteristics of traffic speed. We find that with the increase of prediction time, the speed feature variable always occupies a more important position, but the relative importance of exogenous variables is increasing, which suggests that we should not only pay attention to the speed feature variable, but also consider the exogenous variables to improve the accuracy of long-term prediction in the future model construction. On this basis, future research can further explore the influence of different exogenous variables and the direction of model optimization.

References

[1] Banerjee, A., Duflo, E., & Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. Journal of Development Economics, 145, 102442.

[2] Giuliano, G., & Hanson, S. (Eds.). (2017). The geography of urban transportation. Guilford Publications.

[3] Karami, Z., & Kashef, R. (2020). Smart transportation planning: Data, models, and algorithms. Transportation Engineering, 2, 100013.

[4] Guerrero-Ibáñez, J., Zeadally, S., & Contreras-Castillo, J. (2018). Sensor technologies for intelligent transportation systems. Sensors, 18(4), 1212.

[5] Zhu, L., Yu, F. R., Wang, Y., Ning, B., & Tang, T. (2018). Big data analytics in intelligent transportation systems: A survey. IEEE Transactions on Intelligent Transportation Systems, 20(1), 383-398.

[6] de Dios Ortúzar, J., & Willumsen, L. G. (2024). Modelling transport. John wiley & sons.

[7] Alomari, A. H., Khedaywi, T. S., Marian, A. R. O., & Jadah, A. A. (2022). Traffic speed prediction techniques in urban environments. Heliyon, 8(12).

[8] Yao, B., Chen, C., Cao, Q., Jin, L., Zhang, M., Zhu, H., & Yu, B. (2017). Short‐term traffic speed prediction for an urban corridor. Computer‐Aided Civil and Infrastructure Engineering, 32(2), 154-169.

[9] Yi, H., & Bui, K. H. N. (2020). An automated hyperparameter search-based deep learning model for highway traffic prediction. IEEE Transactions on Intelligent Transportation Systems, 22(9), 5486-5495.

[10] Rahman, R., & Hasan, S. (2018, November). Short-term traffic speed prediction for freeways during hurricane evacuation: a deep learning approach. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 1291-1296). IEEE.

[11] Zhao, J., Gao, Y., Bai, Z., Wang, H., & Lu, S. (2019). Traffic speed prediction under non-recurrent congestion: Based on LSTM method and BeiDou navigation satellite system data. IEEE Intelligent Transportation Systems Magazine, 11(2), 70-81.

[12] Xie, Z., Lv, W., Huang, S., Lu, Z., Du, B., & Huang, R. (2019). Sequential graph neural network for urban road traffic speed prediction. IEEE Access, 8, 63349-63358.

[13] Zou, G., Lai, Z., Wang, T., Liu, Z., Bao, J., Ma, C., ... & Fan, J. (2024). Multi-task-based spatiotemporal generative inference network: A novel framework for predicting the highway traffic speed. Expert Systems with Applications, 237, 121548.

[14] Shaygan, M., Meese, C., Li, W., Zhao, X. G., & Nejad, M. (2022). Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities. Transportation research part C: emerging technologies, 145, 103921.

[15] Wang, J., Chen, R., & He, Z. (2019). Traffic speed prediction for urban transportation network: A path based deep learning approach. Transportation Research Part C: Emerging Technologies, 100, 372-385.

[16] Li, L., Qu, X., Zhang, J., Wang, Y., & Ran, B. (2019). Traffic speed prediction for intelligent transportation system based on a deep feature fusion model. Journal of Intelligent Transportation Systems, 23(6), 605-616.

[17] Jeong, M. H., Lee, T. Y., Jeon, S. B., & Youm, M. (2021). Highway speed prediction using gated recurrent unit neural networks. Applied Sciences, 11(7), 3059.

[18] Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2021). Deep learning on traffic prediction: Methods, analysis, and future directions. IEEE Transactions on Intelligent Transportation Systems, 23(6), 4927-4943.

[19] Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: A deep convolutional

neural network for large-scale transportation network speed prediction. sensors, 17(4), 818.

[20] Cao, M., Li, V. O., & Chan, V. W. (2020, May). A CNN-LSTM model for traffic speed prediction. In 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring) (pp. 1-5). IEEE.

[21] Tang, J., Liu, F., Zou, Y., Zhang, W., & Wang, Y. (2017). An improved fuzzy neural network for traffic speed prediction considering periodic characteristic. IEEE Transactions on Intelligent Transportation Systems, 18(9), 2340-2350.

[22] Meng, X., Fu, H., Peng, L., Liu, G., Yu, Y., Wang, Z., & Chen, E. (2020). D-LSTM: Short-term road traffic speed prediction model based on GPS positioning data. IEEE Transactions on Intelligent Transportation Systems, 23(3), 2021-2030.