ISSN: 1813-4890

Research on traffic flow prediction of augmented multi-component transformer based on dynamics and periodicity

Linlong Chen, Hongyan Wang

School of Computer & Communication, Lanzhou University of Technology, Lanzhou, 730050, China

Abstract

Traffic flow prediction is the basis of urban traffic guidance and control, and accurate traffic flow prediction is very important for road users and traffic management departments. However, traffic flow is highly dynamic, periodic and complex in time and space, which makes accurate traffic flow prediction a challenging task. To solve this problem, this paper proposes an augmented multi-component transformer model (AMR-GAT) based on dynamics and periodicity. Firstly, this paper proposes a traffic flow forecasting method based on Transformer encoder-decoder, and introduces augmented multi-component module into the traffic flow forecasting model to solve the periodicity problem in traffic flow. Secondly, a Temporal Module including convolutional neural network (CNN) and short-term memory network (LSTM) is designed to capture the long-term and short-term characteristics of traffic flow. In addition, the dynamic spatial characteristics of traffic flow are captured by graph attention network (GAT). Finally, TM and GAT are combined to dynamically extract temporal and spatial correlation. Experiments on two data sets show that the proposed method has better prediction performance.

Keywords

Traffic Flow Forecasting, Graph Attention Networks, Augmented Multi-Component, Periodicity, Dynamic Correlation.

1. Introduction

With the development of society and economy, the increasingly serious traffic congestion has become an important reason to hinder economic development. Therefore, how to improve the utilization efficiency of public transport through emerging technologies has attracted more and more attention[1]. Traffic flow forecasting is one of the effective tools to solve traffic problems. Through the prediction of traffic flow, the transportation department can allocate resources in advance and intervene in traffic, thus alleviating congestion and reducing traffic accidents.

Traffic flow will change with the change of time and space, and the traffic flow has the following remarkable characteristics: 1) Dynamic: Because all kinds of vehicles on the road randomly determine the path and are influenced and interfered by various factors of the external environment, it presents dynamic characteristics. 2) periodicity: there are 7 days in a week, including 5 working days and 2 weekends, and people's driving out also shows regularity. For the traffic flow, the traffic flow of the same road section shows the phenomenon of periodic patency and congestion, and the traffic flow state of the same road section shows repetitiveness at regular intervals. 3) Temporal and spatial characteristics: the traffic flow at the next moment is influenced by the previous traffic flow, that is, the historical traffic flow has an influence on the current and future traffic flow, and the traffic flows in adjacent sections will also influence each other[2]. 4) Network characteristics: roads are crisscrossed. For example, the key transportation hubs of the city are connected with each other by main roads, and other small sections are connected to the main roads by auxiliary traffic routes, which makes the whole road network look like a complicated network.

To sum up, the traffic network is affected by various factors, and the traffic flows in different time periods and different roads are interrelated and interact with each other, showing dynamics and complexity, so it is difficult to accurately predict the traffic flow. Aiming at the complex spatio-temporal characteristics of traffic flow, this paper proposes an augmented multi-component transformer model (AMR-GAT) based on dynamics and periodicity. Firstly, the short-term and long-term periodic information is comprehensively considered to predict the traffic, and the time series dependence is obtained by LSTM. In addition, the dynamic spatial characteristics of traffic flow are captured by graph attention network (GAT). Finally, Temporal Module is combined with GAT to dynamically extract temporal-spatial correlation. The contributions of this paper are summarized as follows:

(1) This paper proposes an augmented multi-component Transformer model (AMR-GAT) based on dynamics and periodicity, which solves the periodicity problem in traffic flow through the transformer encoder-decoder.

(2) In this paper, a Temporal Module including convolutional neural network (CNN) and short-term memory network (LSTM) is designed to capture the long-term and short-term characteristics of traffic flow. It is combined with graph attention network (GAT) to capture the dynamic spatio-temporal characteristics of traffic flow.

(3) A large number of experiments have been carried out on real-world traffic data sets, and the experimental results show that this model has better performance than SOTA model.

2. Related Work

In recent years, the development of high-performance intelligent traffic flow forecasting system has been widely concerned by people. Early research on traffic flow forecasting generally adopted traditional theoretical statistical methods, such as autoregressive integral moving average (ARIMA) and vector autoregressive (VAR)[3]. Among them, ARIMA is the most widely used time series model. Van Hinsbergen et al. [4] used Kalman filter to complete the traffic prediction task by minimizing the variance of the optimal solution. Although the traditional statistical methods are relatively simple and convenient to calculate, they are all based on the stationarity assumption of time series[5]. Compared with the above model, deep RNN and its successors, such as LSTM network[6] and GRU network[7], perform better in capturing the time correlation of traffic conditions. They can memorize information from a large number of sequence information and learn complex models. However, the simple RNNs model can not use the spatial information of traffic data, which has become a problem to be solved. In order to simulate spatial correlation, researchers use CNN to capture the differences in Euclidean space, but their limitation is that they can only be applied to standard grid data[8]. Recent studies have begun to use graph-convolution network (GCN) to simulate the non-Euclidean internal relations in road networks[9].

Recently, graph convolution extends the traditional convolution method to data with graph structure[10], and there are two types of graph convolution methods. One is to generalize the spatial neighborhood by convolution filtering, with emphasis on node neighborhood selection. For example, Velickovic et al. [11] introduced the attention mechanism into the graphics field, allowing different neighboring nodes to be assigned different weights. The other type is to expand convolution to graphs in the spectral domain by searching the corresponding Fourier basis. For example, Defferrar et al. [12] used Chebyshev polynomial approximation to decompose eigenvalues, thus avoiding the calculation of eigenvectors of Laplacian operators, thus obtaining better results in graph theory. Yu et al. [13] proposed a gated GCN for traffic prediction based on this method, but the model failed to capture the dynamic temporal-spatial correlation of traffic data.

Because of its flexibility and effectiveness in dependency modeling, attention mechanism is now widely used in natural language processing, traffic flow prediction and speech recognition [14]. The core goal of attention mechanism is to filter out the information that is more critical to the current task goal from a large amount of information [15]. Liang et al. [16] introduced a multi-layer attention mechanism to simulate the dynamic spatio-temporal correlation between geographic sensors to solve the problem of spatio-temporal data prediction. However, it takes a lot of time to train a model separately for each time series. Guo et al. [17] proposed a graphic CNN with attention mechanism to achieve good performance in traffic flow prediction. Different from previous work, this paper considers the graph structure of traffic network and the dynamic spatio-temporal pattern and periodicity of traffic data. Therefore, in this paper, Transformer is extended to the spatio-temporal data prediction task of graph, and an Augmented Multi-Component and a Temporal Module including Convolutional Neural Network (CNN) and Long-term Memory Network (LSTM) are designed to capture the periodic characteristics of traffic flow, and combined with Graph Attention Network (GAT) to extract the dynamic spatio-temporal characteristics of traffic flow.

3. Methodology

3.1. Problem Definition

In this paper, the undirected graph $G=(V,E,A)$ is used to describe the traffic road network, and each sensor is considered as a node, and the connection of any two sensors is considered as an edge between two nodes, where $|V|=N$ denotes a set of road nodes, E denotes a set of edges, and $A\in R^{N\times N}$ denotes the adjacency matrix of graph G. The adjacency matrix contains only elements 0 and 1. In time step t, the traffic flow observed in graph G is represented as graph signal $X^{t}\in R^{N*F}$, where F denotes the feature dimension of each node. Historical traffic flow data can be defined as $X=\left(X^{t-H+1},X^{t-H+2},X^{t}\right)\in R^{H*N*F}$, and H indicates the length of historical data. Traffic forecasting is expressed as $Y=\left(X^{t+1},X^{t+2},\ldots,X^{t+P}\right)\in R^{P*N*F}$, and P represents the length of forecast data. The traffic forecasting problem aims to learn a function φ, which can accurately predict the future P map signal given the historical H map signal of the whole road network:

$$\left(X^{t+1},X^{t+2},\ldots,X^{t+P}\right)=f\left(\left(X^{t-H+1},X^{t-H+2},X^{t}\right),G\right) \tag{1}$$

3.2. Framework of AMR-GAT

As shown in Fig.1, this paper shows the framework diagram of AMR-GAT model, which mainly consists of three modules: (1) an augmented multi-component module, which is used for synchronously capturing the characteristics of periodicity and periodic time offset of traffic flow; Where X_h, X_{ds} and X_{ws} respectively represent recent component, daily augmented component and weekly augmented component; (2) The encoder module is used to capture the temporal and spatial correlation of traffic flow. The encoder includes Temporal Module and graph attention network, in which Temporal Module includes CNN and LSTM. (3) Decoder module, which is used to predict traffic flow from time-space series, including Temporal Module and Convolution. x_p stands for the p prediction time slice. $F(X)$ represents the output of the decoder, and Aggregation represents the aggregation operation.

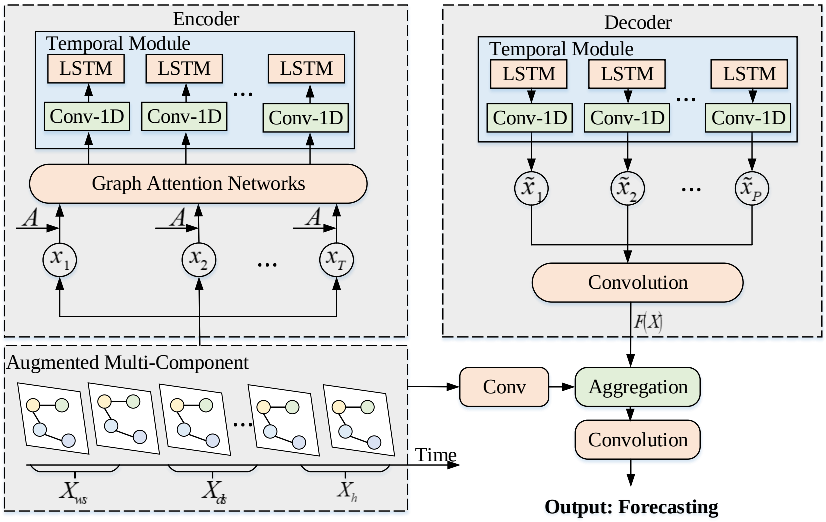

Fig.1 Framework diagram of AMR-GAT model

3.3. Augmented Multi-Component

The Multi-component module includes recent component, daily periodicity component and weekly periodicity component of traffic flow. Three time series segments of length T_h, T_d and T_w are intercepted in chronological order, which are used as the inputs of recent component, daily augmented component and weekly augmented component respectively. Augmented multi-component introduces daily augmented component and weekly augmented component for daily shift and weekly shift. T_{ds} and T_{ws} represent the length of daily augmented component and weekly augmented component respectively. S indicates that there is a periodic offset of the time step before and after the daily and weekly cycles. $T_h = N_h * T_p$, N_h is the traffic flow sequence of the past N_h hours. $T_d = N_d * T_p$, N_d is the traffic flow at the same time in the past N_d days. $T_w = N_w * T_p$, N_ω is the traffic flow at the same time in the past N_ω weeks.

The relationship between Augmented multi-component and multi-component is expressed as follows:

$$T_{ds} = T_d * (2*S+1) = N_d * \mathrm{T}_p * (2*S+1) \tag{2}$$

$$T_{ws} = T_w * (2*S+1) = N_w * \mathrm{T}_p * (2*S+1) \tag{3}$$

Where T_p represents the prediction window size T_p and S represents the periodic offset.

(1) Recent component

The recent component is a historical time series directly adjacent to the forecast period. The recent component is as follows:

$$X_h = \left(X^{t_c - T_h + 1}, X^{t_c - T_h + 2}, ..., X^{t_c} \right) \in R^{T_h * N * F} \tag{4}$$

Where t_c represents the current time, N represents the number of nodes in the road network, F represents the dimension represented by each node, and X^t represents the traffic flow at time t.

(2) Daily augmented component

The Daily augmented component represents the traffic flow in the same time period as the forecast window in the past few days. The daily augmented component is as follows:

$$X_{ds} = \Big(X^{t_c - N_d * f - S*T_p + 1}, \ldots, X^{t_c - N_d * f + S*T_p + T_p}, \\ X^{t_c - (N_d - 1)* f - S*T_p + 1}, \ldots, X^{t_c - (N_d - 1)* f + S*T_p + T_p}, \ldots, \\ X^{t_c - f - S*T_p + 1}, \ldots, X^{t_c - f + S*T_p + T_p} \Big) \in R^{T_{ds} * N * F} \tag{5}$$

Let $S = 1$, $T_d = 24$, then $N_d = 2$, $f = 288$ are the collection frequency of one day, and Equation (5) indicates that the traffic flow from 4: 00 pm to 7: 00 pm on the nearest Wednesday and Thursday is used. Then $X^{t_c - N_d * f - S*T_p + 1}, \ldots, X^{t_c - N_d * f + S*T_p + T_p}$ can represent the traffic flow from 4: 00 pm to 7: 00 pm on the last Wednesday, and $X^{t_c - (N_d - 1)* f - S*T_p + 1}, \ldots, X^{t_c - (N_d - 1)* f + S*T_p + T_p}$ can be the traffic flow from 4: 00 pm to 7: 00 pm on the last Thursday.

(3) Weekly augmented component

Weekly augmented component represents the traffic flow in the same time period as the forecast window in the past few weeks. The weekly augmented component is as follows:

$$X_{ws} = \Big(X^{t_c - N_w * f * 7 - S*T_p + 1}, \ldots, X^{t_c - N_w * f * 7 + S*T_p + T_p}, \\ X^{t_c - (N_w - 1)* f * 7 - S*T_p + 1}, \ldots, X^{t_c - (N_w - 1)* f * 7 + S*T_p + T_p}, \ldots, \\ X^{t_c - f * 7 - S*T_p + 1}, \ldots, X^{t_c - f * 7 + S*T_p + T_p} \Big) \in R^{T_{ws} * N * F} \tag{6}$$

Let $S = 1$, $T_w = 24$, and $N_w = 2$, and Equation (6) indicates that the traffic flow from 4: 00 pm to 7: 00 pm in the past two Thursdays is used. $X^{t_c - N_w * f * 7 - S*T_p + 1}, \ldots, X^{t_c - N_w * f * 7 + S*T_p + T_p}$ represents the traffic flow from 4: 00 to 7: 00 pm in the past two Thursdays, and $X^{t_c - (N_w - 1)* f * 7 - S*T_p + 1}, \ldots, X^{t_c - (N_w - 1)* f * 7 + S*T_p + T_p}$ represents the traffic flow from 4: 00 to 7: 00 pm in the past Thursday.

The above three component constitute an augmented multi-component module, which considers the periodicity and periodic time offset in traffic flow prediction. Let $T = T_h + T_{ds} + T_{ws}$ be the length of the augmented multi-component module, and input the input data $X_{am} = (X_h, X_{ds}, X_{ws}) \in R^{T*N*F}$ into the encoder-decoder.

3.4. Encoder

The encoder is used to capture the temporal-spatial correlation of traffic flow. The encoder includes Temporal Module and graph attention network (GAT), in which Temporal Module includes convolutional neural network (CNN) and long-term and short-term memory network (LSTM). GAT and Temporal Module are used to learn the spatio-temporal representation of augmented multi-component.

(1) Graph Attention Network (GAT)

Graph attention network (GAT)[12] proposes a weighted summation method for the features of adjacent nodes by using attention mechanism, which overcomes the limitations of convolution network based on spectrogram and is easy to assign different learning weights to different neighborhoods. This paper uses multi-head attention mechanism to enhance the ability of the model, as shown below:

$$\beta_{i,j} = LeakyRelu\left(\alpha\left[Wh_{t,i} \,\|\, Wh_{t,j}\right]\right) \tag{7}$$

$$e_{i,j} = \frac{\exp\left(\beta_{i,j}\right)}{\sum_{v_k \in N(v_i)} \exp\left(\beta_{i,j}\right)} \tag{8}$$

$$x'_i = \sigma\left(\frac{1}{K}\sum_{K=1}^{K}\sum_{v_k \in N(v_i)} e_{i,j}^{(k)} W^{(k)} h_{t,j}\right) \tag{9}$$

Among them, W is the weight parameter of node feature transformation, $\sigma(\cdot)$ is the function to calculate the correlation between two nodes, $h_{t,i}$ and $h_{t,j}$ are the hidden layer outputs of LSTM network, $LeakyRelu$ is the activation function, $\beta_{i,j}$ is the attention score between two nodes, $e_{i,j}^{(k)}$ is the weight coefficient calculated by the K group attention mechanism, $W^{(k)}$ is the learnable parameter, and x'_i is the final result obtained by averaging K outputs.

(2) Temporal Module

In Temporal Module, the internal features of each sensor are integrated by adding Conv-1D in the feature dimension, and the time features of traffic flow are captured by using LSTM in the sequence dimension. As shown in Fig.2, this paper represents the spatial features of each moment captured by GAT as $G_t \in R^{N*C}$, Temporal Module integrates the spatial features through Conv-1D based on hidden states $H_{t-1} \in R^{N*H}$ and G_t, and transmits them to LSTM with cell memory state $C_{t-1} \in R^{N*H}$ to capture the temporal features, and H represents hidden size.

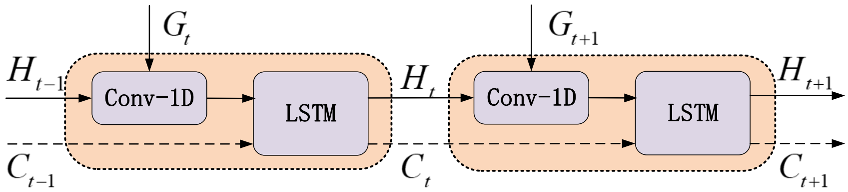

Fig.2 Framework diagram of temporary module

By inputting the spatial feature G_t as the Temporal Module input at each moment of the traffic flow sequence, the correlation of traffic flow is captured. Therefore, at the time step t, the Temporal Module is calculated as follows:

$$H_t, C_t = TCL(G_t; H_{t-1}; C_{t-1}) \tag{10}$$

Where $t \in \{1,...,T\}$, transfer G_t to Temporal Module, use input gate I_t, forgetting gate F_t and hidden state H_{t-1} to update cell memory state C_t, and use output gate O_t to update current hidden state H_t, as shown below:

$$I_t = \sigma\left(W_{gi} * G_t + W_{hi} * H_{t-1} + W_{ci} \odot C_{t-1}\right) \tag{11}$$

$$F_t = \sigma\left(W_{gf} * G_t + W_{hf} * H_{t-1} + W_{cf} \odot C_{t-1}\right) \tag{12}$$

$$C_t = F_t \odot C_{t-1} + I_t \odot \tanh\left(W_{gc} * G_t + W_{hc} * H_{t-1}\right) \tag{13}$$

$$O_t = \sigma\left(W_{go} * G_t + W_{ho} * H_{t-1} + W_{co} \odot C_t\right) \tag{14}$$

$$H_t = O_t \odot \tanh(C_t) \tag{15}$$

Where $W_{\alpha\beta}\left(\alpha \in (g,h,c), \beta \in (i,f,c,o)\right)$ stands for learnable parameters of Temporal Module, $*$ stands for one-dimensional convolution operation and $\odot$ stands for Hadamard product.

(3) Decoder

The decoder is used for traffic flow prediction, and consists of Temporal Module and Convolution. By using the hidden state obtained by the encoder, the high-dimensional feature representation is generated from the spatio-temporal sequence. In this paper, Temporal

Module is used to decode the hidden state H_T and the cell storage state C_T from the encoder. Assuming that the size of the prediction window is P, the expression of the decoder is:

$$X_{t+1} = TCL\left(\vec{0}; H_t; C_t\right) \tag{16}$$

Where $t \in \{T, \ldots, T+P-1\}$, $\vec{0}$ represent all-zero arrays.

The spatio-temporal features captured by the encoder-decoder are input to the fusion module for traffic flow prediction. The fusion module fuses the residual information of the augmented multi-component module with the high-dimensional representation $F(X)$ of the decoder by using convolution residual connection to speed up model training. Finally, Convolution is used to ensure the prediction of the dimension and shape of $Y \in R^{P*N*F}$.

4. Experiment

4.1. Datasets

In this paper, the prediction performance of the proposed AMR-GAT model is verified on PeMSD04 and PeMSD08 data sets, which are collected by Caltrans Performance Measurement System (PeMS) every 30 seconds. The detailed information of the experimental data set is shown in Table 1:

Table 1 Description of experimental dataset

Datasets	Nodes	Edges	Interval	Time range	Time Steps
PEMSD04	307	340	5min	1/1/2018-2/28/2018	16992
PEMSD08	170	295	5min	7/1/2016-8/31/2016	17856

PEMSD04: This data set contains 3848 detectors on 29 roads, and the time span is from January to February, 2018. In this paper, the data of the first 50 days are selected as the training set, and the remaining 9 days are selected as the test set.

PEMSD08: This data set contains 1979 detectors on 8 roads, and the time span is from July to August 2016. In this paper, the data of the first 50 days are selected as the training set, and the data of the last 12 days are selected as the test set.

Before the data is input into the prediction model, the data is preprocessed by standardization, as shown below:

$$X = \frac{X}{Max(X)} \tag{17}$$

4.2. Model Parameter

Based on the deep learning framework of Pytorch, this experiment completes the model construction and training in the development environment of PyCharm. This paper predicts the next 15-minute, 30-minute and 60-minute traffic flows, that is $T_p = 3$, $T_p = 6$ and $T_p = 12$. The parameters of the three components are set to $T_h = 24$, $T_d = 12$, $T_w = 12$, and the augmented multi-component sequence is $T_D = 96$. The encoder uses two layers of GAT, and the convolution filters of Temporal Module are 64 hidden units. In the decoder, the Temporal Module has 64 hidden units, and the output sequence is T_p. In the fusion module, convolution filters are T_p. In this paper, Adam optimizer is used to train the model. epoch is set to 200, initial learning rate is 0.001, weight decay is 5×10^{-4}, and Mean Square Error (MSE) is used as the loss function.

4.3. Evaluation Metric

In this paper, root mean square error (RMSE) and mean absolute error (MAE) are used as evaluation indexes:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i\right)^2} \tag{18}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i\right| \tag{19}$$

Where y_i and y_i represent the actual traffic speed and the predicted traffic speed respectively, and n is the number of observations.

4.4. Baseline

This paper compares the AMR-GAT model with the following models:

(1) LSTM [18]: Long Short-term Memory network, RNN model for time series forecasting.

(2) GRU [19]: Gated Recurrent Unit network, an improved RNN model for time series forecasting.

(3) STGCN [13]: Spatial-temporal Graph Convolution Network, which combines graph convolution with 1D convolution.

(4) MSTGCN [20]: A multi-component network for traffic flow forecasting.

(5) ASTGCN [17]: Attention-based Spatio-temporal Graph Convolutional Network, which further integrates spatial and temporal attention mechanisms to STGCN for capturing dynamic spatial and temporal patterns.

(6) STSGCN [21]: Spatial-Temporal Synchronous Graph Convolutional Network that captures spatial-temporal correlations by stacking multiple localized GCN layers with adjacent matrix over the time axis.

4.5. Results and Analysis

In this paper, AMR-GAT model is tested on PeMSD04 and PeMSD08 data sets, and compared with six baseline methods. As shown in Table 2, the prediction performance of AMR-GAT model and different baseline models on PeMSD04 and PeMSD08 data sets is shown. It can be seen that the AMR-GAT model in this paper shows the best prediction performance on two sets of data sets.

For example, on PeMSD04 data set, compared with LSTM model, the MAE and RMSE of AMR-GAT model decreased by about 17.67% and 17.50% respectively in 15 minutes. Compared with GRU model, the MAE and RMSE of AMR-GAT model decreased by about 17.78% and 17.91% respectively. Because LSTM and GRU models only consider time dependence, they ignore the spatial correlation of traffic network. STGCN, MSGCN, ASTGCN, STGCN and the AMR-GAT model in this paper all consider the spatial correlation, so they have better prediction performance than the methods only used for time series prediction.

In 15 minutes, on PeMSD08 data set, compared with STGCN, MSGCN, ASTGCN and STGCN models, the MAE of AMR-GAT model in this paper decreased by about 14.70%, 8.05%, 5.42% and 1.69% respectively. RMSE decreased by about 15.50%, 7.19%, 4.77% and 5.29% respectively. Because STGCN, MSGCN, ASTGCN and STGCN models use two modules to model spatial correlation and time dependence respectively, they ignore the time dependence and periodic changes in traffic flow data. The AMR-GAT model in this paper captures the temporal-spatial correlation of traffic flow at the same time, and considers the time dependence and periodic change, which proves that the AMR-GAT model can better capture the temporal-spatial correlation and periodic time offset of traffic flow.

To sum up, the AMR-GAT model proposed in this paper uses augmented multi-component method to capture the periodic time offset characteristics of traffic flow, and combines Temporal Module with GAT at every moment to capture the temporal and spatial correlation of traffic flow. The AMR-GAT model achieves the best prediction performance, and it shows excellent prediction performance within 15 minutes, 30 minutes or 60 minutes. To sum up, the model in this paper has more advantages in capturing the temporal and spatial characteristics and periodic time offset characteristics of traffic flow.

Table 2 Performance comparison between AMR-GAT model and baseline model

Data	Method	15min		30min		60min	
		RMSE	MAE	RMSE	MAE	RMSE	MAE
PEMSD04	LSTM	34.00	22.02	35.81	23.34	38.81	25.58
	GRU	34.17	22.05	35.88	23.45	38.84	25.83
	STGCN	32.77	21.34	34.07	21.78	37.42	24.32
	MSTGCN	28.98	19.40	30.61	20.49	32.71	22.01
	ASTGCN	29.19	19.59	30.26	20.32	32.37	21.83
	STSGCN	29.74	18.52	31.52	19.73	33.63	21.06
	AMR-GAT	**28.05**	**18.13**	**29.44**	**18.75**	**31.42**	**20.40**
PEMSD08	LSTM	26.02	17.95	28.35	19.68	32.56	22.61
	GRU	25.92	17.97	28.35	19.71	31.80	22.18
	STGCN	24.58	16.33	27.31	17.91	31.24	20.85
	MSTGCN	22.38	15.15	23.90	19.09	25.46	17.11
	ASTGCN	21.81	14.76	23.33	15.71	24.40	16.33
	STSGCN	21.93	14.20	23.71	15.28	26.05	16.67
	AMR-GAT	**20.77**	**13.96**	**22.06**	**14.55**	**23.95**	**15.87**

Predictive performance of AMR-GAT model and variant model in 60 minutes

Dataset	Method	Augmented Multi-Component	
		RMSE	MAE
PEMSD08	AM-CNN-GCN	24.31	16.00
	AM-LSTM-GCN	26.85	18.19
	AMR-GAT	**23.95**	**15.87**

In order to further study the performance of the temporary module of AMR-GAT model, this paper replaces the temporary module and GAT with CNN-GCN and LSTM-GCN, studies the influence of the temporary module and GAT on the prediction performance of the model, and compares them on the PEMSD08 data set, and makes a 60-minute traffic flow prediction, as shown in0.

It can be seen from 0 that the prediction performance of AM-CNN-GCN model is better than that of AM-LSTM-GCN model. When AM-LSTM-GCN model is trained, there is error propagation, which leads to the decline of prediction performance. On the contrary, the AMR-GAT model captures the temporal and spatial correlation of traffic flow through the Temporal Module and GAT at the same time in the encoder. In each prediction time step, the temporal and spatial information of the last time step is considered, and the periodic time change of traffic flow is captured through the augmented multi-component, which effectively alleviates the spread of errors.

ISSN: 1813-4890

In order to better explain the AMR-GAT model, this paper visualizes the experimental results of AMR-GAT model and FNN, LSTM, ASTGCN and MSTGCN in PEMSD04 and PEMSD08 data sets, as shown in Fig.3 and Fig.4.

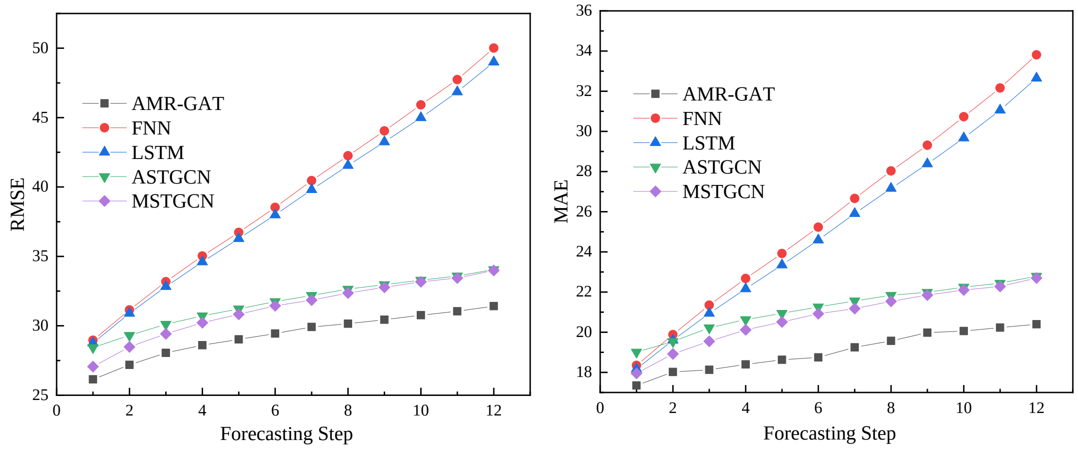

Fig.3 RMSE and MAE prediction results of different models on PEMSD04

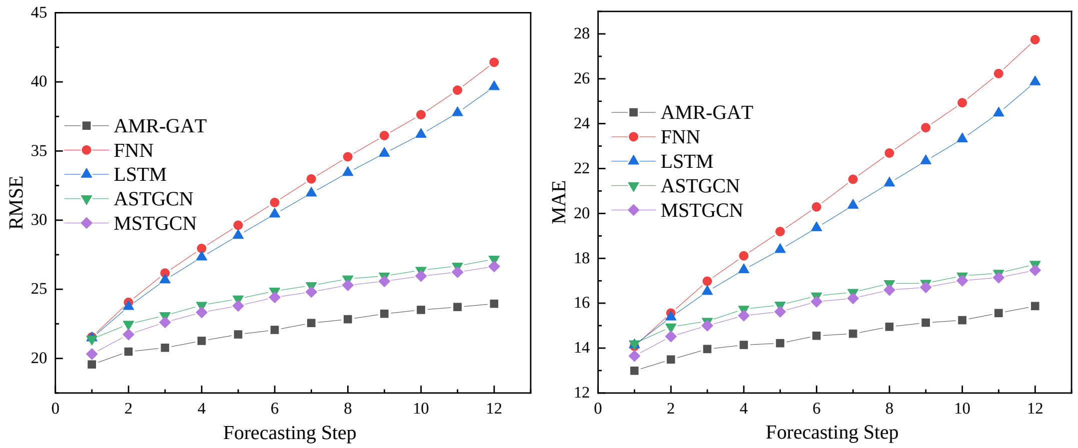

Fig.4 RMSE and MAE prediction results of different models on PEMSD08

It can be observed from the figure that the values of RMSE and MAE of AMR-GAT model are always lower than those of FNN, LSTM, ASTGCN and MSGCN, which shows that the prediction performance of AMR-GAT model is the best. In a word, AMR-GAT model can always get the best prediction results in different prediction time steps, which shows that AMR-GAT model can have good prediction performance in traffic prediction tasks. AMR-GAT model can not only effectively capture the temporal-spatial correlation and periodic time offset characteristics of traffic flow at the same time, but also capture the changing trend of traffic flow. AMR-GAT model can accurately predict traffic congestion, thus proving the effectiveness of AMR-GAT model in real-time traffic flow prediction.

5. Conclusion

In this paper, an augmented multi-component transformer model (AMR-GAT) based on dynamics and periodicity is proposed, which introduces an augmented multi-component module to capture the periodic time offset characteristics of traffic flow. In this paper, a traffic flow forecasting framework based on Transformer is proposed. Conv-1D and LSTM are combined through Temporal Module to capture the time characteristics of traffic flow. The spatial characteristics of traffic flow are captured by graph attention network (GAT). Then, the Temporal Module is combined with GAT to deal with the temporal-spatial correlation of traffic network. The decoder obtains high-dimensional representation from spatio-temporal sequence prediction through Temporal Module and convolutional neural network. Finally, experiments are carried out on two sets of data sets to verify the prediction performance of AMR-GAT model, and compared with the baseline model. The experimental results show that the prediction performance of AMR-GAT model is better than the baseline model in different data sets and different prediction time periods, which proves the effectiveness and accuracy of AMR-GAT model in traffic flow prediction. In the future work, external factors will be further considered to further improve the accuracy of traffic flow prediction.

References

[1] N.G. Polson and V.O. Sokolov: Deep Learning for Short-term Traffic Flow Prediction, Transportation Research Part C: Emerging Technologies, Vol. 79 (2017), p.1-17.
[2] T. Wang, Q. Meina, A. Zhu, G.C. Shan and H. Snoussi: Abnormal Event Detection Via the Analysis of Multi-Frame Optical Flow Information, Frontiers of Computer Science, Vol. 14 (2020) No. 2, p. 304-313.
[3] M.M. Hamed, H.R. Al-Masaeid and Z.M. B.Said: Short-term Prediction of Traffic Volume in Urban Arterials, Journal of Transportation Engineering, Vol. 121 (1995) No. 3, p. 249-254.
[4] C.P.I.J. Van Hinsbergen, T. Schreiter, F.S. Zuurbier, J.W.C. Van Lint and H.J. Van Zuylen: Localized Extended Kalman Filter for Scalable Real-time Traffic State Estimation, IEEE Transactions on Intelligent Transportation Systems, Vol. 13 (2012) No. 1, p. 385-394.
[5] X.Y. Yin, G.Z. Wu, J.Z. Wei, Y.M. Shen, H. Qi and B. Yin: Deep Learning on Traffic Prediction: Methods, Analysis and Future Directions, IEEE Transactions on Intelligent Transportation Systems, Vol. 23 (2021) No. 6, p. 4927-4943.
[6] K. Greff, R.K. Srivastava, J. Koutník, B.R. Steunebrink and J. Schmidhuber: LSTM: A Search Space Odyssey, IEEE Transactions on Neural Networks and Learning Systems, Vol. 28 (2017) No. 10, p. 2222-2232.
[7] R. Dey and F.M. Salem: Gate-Variants of Gated Recurrent Unit (GRU) Neural Networks: *2017 IEEE 60th International Midwest Symposium on Circuits and Systems* (Boston, MA, USA, August 06-09, 2017), p. 1597-1600.
[8] H.X. Yao, X.F. Tang, H. Wei, G.J. Zheng and Z.H. Li: Revisiting Spatial-Temporal Similarity: A Deep Learning Framework for Traffic Prediction, *Proceedings of the AAAI Conference on Artificial Intelligence* (January 07-17, 2019) vol. 33, p. 5668-5675.
[9] Z.J. Lv, J.J. Xu, K. Zheng, H.Z. Yin, P.P. Zhao and X.F. Zhou: LC-RNN: A Deep Learning Model for Traffic Speed Prediction, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence* (July 2018), p. 3470-3476.
[10] Z.H. Wu, S.R. Pan, F.W. Chen, G.D. Long, C.Q. Zhang and S.Y. Philip: A Comprehensive Survey on Graph Neural Networks, IEEE Transactions on Neural Networks and Learning Systems, Vol. 32 (2021), No. 1, p. 4-24.
[11] P. Velickovi´c, G. Cucurull, A. Casanova, A. Romero, P. Lio and Y. Bengio: Graph Attention Networks. *In International Conference on Learning Representations* (2018).
[12] M. Defferrard, X. Bresson and P. Vandergheynst: Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering, *Proceedings of the 30th International Conference on Neural Information Processing Systems* (December, 2016), Vol. 29, p. 3844-3852.
[13] B. Yu; H. Yin; Z. Zhu: Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting, *Proceedings of the International Joint Conference on Artificial Intelligence* (Stockholm, Sweden, July 13-19, 2018), p. 3634-3640.
[14] T. Shen, T. Zhou, G.D. Long, J. Jiang, S.R. Pan and C.Q. Zhang: DiSAN: Directional Self-Attention Network for RNN/CNN-Free Language Understanding, *Proceedings of the AAAI Conference on Artificial Intelligence* (New Orleans, Lousiana, USA, February 2-7, 2018), Vol. 32, p. 668.
[15] W.Y. Cheng, Y.Y. Shen, Y.M. Zhu and L.P. Huang: A Neural Attention model for Urban Air Quality Inference: Learning the Weights of Monitoring Stations, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence* (New Orleans, Louisiana USA February 2-7, 2018), p. 2151-2158.
[16] Y.X. Liang, S.Y. Ke, J.B. Zhang, X.W. Yi and Y. Zheng: GeoMAN: Multi-Level Attention Networks for Geo-Sensory Time Series Prediction, *Proceedings of the 27th International Joint Conference on Artificial*

Intelligence (Stockholm Sweden, July 13-19, 2018), p. 3428-3434.
[17] S.N. Guo, Y.F. Lin, N. Feng, C. Song and H.Y. Wan: Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting, *AAAI Conference on Artificial Intelligence* (Honolulu, Hawaii USA, 27 January 2019-1 February, 2019), Vol. 33, No. 01, pp. 922-929.
[18] S. Hochreiter, J. Schmidhuber: Long Short-Term Memory, Neural Computation, Vol. 9 (1997), p. 1735-1780.
[19] J.Y. Chung, C. Gulcehre, K. Cho, Y. Bengio: Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling, arXiv 2014. https://doi.org/10.48550/arXiv.1412.3555.
[20] N. Feng, S.N. Guo, C. Song, Q.C. Zhu, H.Y. Wan: Multi-component Spatial-Temporal Graph Convolution Networks for Traffic Flow Forecasting, Journal of Software, Vol. 30 (2019), p. 759-769.
[21] Song, C.; Lin, Y.; Guo, S.; Wan, H. Spatial-Temporal Synchronous Graph Convolutional Networks: A New Framework for Spatial-Temporal Network Data Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence* (New York, NY, USA, February 7-12, 2020), Vol. 34, p. 914-921.