Integrating External Attributes for Short-term Traffic Flow Prediction: A Postprint

Authors: Wang Qingrong, Wu Yuyu, Zhu Changfeng, Wang Yuan

Date: 2022-06-06T00:00:00+00:00

Abstract

Most existing traffic flow prediction algorithms only consider predictions under normal conditions, without accounting for the impact of weather attributes and surrounding geographic attributes on prediction results. This paper proposes a combined prediction model that integrates external attributes (A-STIGCN). First, external attributes are incorporated as attributes of road segments in the road network, while simultaneously modeling both segment attributes and traffic features to obtain enhanced feature vectors. Second, graph wavelet transform and adaptive matrix are utilized to extract local and global spatial feature information of traffic flow, respectively, and the long-term memory capability of Gated Recurrent Unit (GRU) is leveraged to extract temporal characteristics. Finally, an attention mechanism is employed to capture spatio-temporal dynamic variability for traffic flow prediction. The model is validated using Shenzhen taxi trajectory data, corresponding weather data, and POI data. The experimental results demonstrate that the A-STIGCN combined model achieves superior prediction performance compared to traditional linear models and their variants. Specifically, compared with the ASTGCN model without attention mechanism, the MAE is reduced by approximately 0.131 and accuracy is improved by 0.068. In comparative analysis with the TGCN model without external factors, the MAPE is decreased by approximately 0.637% and accuracy is enhanced by 0.079, thereby providing more effective guidance for traffic management.

Full Text

Short-term Traffic Flow Prediction Integrating External Attributes

Wang Qingrong[1], **Wu Yuyu**[1], **Zhu Changfeng**[2], **Wang Yuan**[1] [1]School of Electronics & Information Engineering, [2]School of Transportation Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract

Most existing traffic flow prediction algorithms consider only normal conditions and overlook the influence of weather attributes and surrounding geographical attributes on prediction results. This paper proposes a combined prediction model (A-STIGCN) that integrates external attributes. First, external attributes are incorporated as properties of road segments in the traffic network, and both segment attributes and traffic characteristics are modeled to obtain enhanced feature vectors. Second, graph wavelet transform and an adaptive matrix are employed to extract local and global spatial feature information from traffic flow, respectively, while a Gated Recurrent Unit (GRU) leverages its long-term memory capability for temporal information to extract temporal characteristics. Finally, an attention mechanism captures spatiotemporal dynamic variability for traffic flow prediction. Experiments conducted using Shenzhen taxi trajectory data, corresponding weather data, and POI data demonstrate that the A-STIGCN combined model outperforms traditional linear models and variant models. Compared with the ASTGCN model without attention mechanism, MAE decreases by approximately 0.131 and accuracy improves by 0.068. Compared with the TGCN model without external factors, MAPE decreases by approximately 0.637% and accuracy improves by 0.079, thereby providing better guidance for traffic management.

Keywords: traffic flow prediction; graph wavelet transformation; adaptive matrix; external factors; gated recurrent mechanism; attention mechanism

0 Introduction

Short-term traffic flow prediction has long been a research focus in intelligent transportation systems. Existing studies on traffic flow prediction primarily employ statistical theory, nonlinear theory, and neural network methods, establishing numerous prediction models such as time series models, Kalman filtering models, non-parametric regression models, and neural network models [1]. With continuous model improvements, shallow machine learning methods —including Bayesian estimation, Support Vector Regression (SVR), and random forest models [2]—have demonstrated promising application prospects in short-term traffic flow prediction due to their generalization capabilities and low maintenance costs. However, these models exhibit limitations in feature extraction and large-scale data processing. Deep learning represents one of the most prominent branches of data-driven methods and offers superior advantages in high-dimensional data processing and spatiotemporal relationship analysis compared to other data-driven approaches [3].

Since future traffic states depend not only on historical states but also on various static and dynamic external factors, accurate traffic prediction remains challenging. Static factors do not change continuously over time but affect traffic states within certain periods, while dynamic factors vary with time and cause changes in traffic conditions. These factors introduce randomness into traffic states and increase the complexity of accurate traffic prediction [4]. Consequently, integrating multi-source data to achieve precise traffic flow prediction represents an urgent problem requiring solution.

Literature [5,6] utilized deep learning theory to capture long-term dependencies and introduced attention mechanisms to adjust the weights of encoded vectors, thereby further improving prediction accuracy, but did not deeply consider spatial feature extraction. Literature [7] proposed STGCN (Spatio-Temporal Graph Convolutional Network), which can extract spatiotemporal features and capture spatial dependencies. Literature [8] proposed DCRNN (Diffusion Convolutional Recurrent Neural Network), employing diffusion convolution and recurrent networks to capture spatial and temporal dependencies of traffic flow on directed graphs, with results demonstrating favorable prediction accuracy. Literature [9] constructed a hybrid convolutional long short-term memory neural network model based on critical road sections to estimate future traffic changes, first identifying critical sections with maximum impact on sub-networks using spatiotemporal correlation algorithms, then using traffic speeds from these critical sections as input to ConvLSTM to predict future traffic states across the entire network. Experimental results validated the model' s predictive capability under different critical sections.

Although these methods achieved satisfactory results, most experiments were conducted under single influencing factors or specific conditions. Literature [10] modeled spatiotemporal and dynamic correlations of traffic data based on graph wavelet convolution operators, capturing spatial correlations from traffic network graph node neighborhoods by stacking multiple graph wavelet neural network modules. However, literature [11,12] did not consider multi-feature information. In response, literature [13] proposed a multi-feature attention mechanism to capture complex feature relationships in traffic data, designing an information-enhanced transmission mechanism to capture complex dynamic spatial dependencies and employing linear-nonlinear data fusion mechanisms to effectively improve prediction accuracy. Literature [14] selected the GBDT algorithm to model and predict highway traffic flow, focusing on analyzing the impact of weather changes, time period characteristics, and weekday features on the traffic state of originally closed highway systems, with different vehicle types exhibiting strong uncertainty and randomness, which holds significant implications for subsequent model development.

Building upon this foundation, this paper proposes treating external attributes as property information of road segments, combining weather factors and surrounding geographical factors to predict the impact of external factors on traffic flow. By integrating attribute information and traffic characteristics, the model' s perception of external information is enhanced. Graph wavelet networks can globally process road network topology information, mine spatial characteristics of short-term traffic, extract temporal characteristics through GRU' s long-term memory capability for temporal information, and employ attention mechanisms to allocate sufficient weights to important information distributions, thereby improving traffic flow prediction accuracy.

1 Traffic Attribute Definition

The objective of traffic flow prediction is to forecast traffic information over a certain period based on historical traffic data on roads. Initial variables include traffic volume in the urban road network and external factors affecting traffic, which are used to predict future traffic flow.

1.1 Road Network Feature Attributes

Definition 1: Road Network G. Using the installation location of traffic flow detection sensor i in the urban road network as nodes, denoted as v_i, and using adjacent sensor j on neighboring road segments as neighbor nodes, we construct a traffic road network graph model G defined as [15]:

$$G = (V, E, A)$$

where V represents the set of nodes, which corresponds to the set of traffic flow detection sensors in the road network. E is a set of edges representing connections in graph G, reflecting connectivity between road segments. Element $e_{i,j} = \langle v_i, v_j \rangle \in E$ indicates a connection relationship between node v_i and node v_j. In the graph construction process, edges are typically established between nodes within the neighborhood range N_i of node v_i. $A \in \mathbb{R}^{N\times N}$ is the adjacency matrix of graph G, containing only elements of 0 and 1. If no correlation exists between roads, the element is 0; 1 indicates correlation. The traffic road network graph G constructed in this paper is an undirected graph.

Definition 2: Feature Matrix X. Traffic flow on road segments is treated as attributes of network nodes, represented by the feature matrix $X \in \mathbb{R}^{N\times P}$, where P is the number of node attribute features, i.e., the length of the historical time series. X_i represents the traffic flow of each road segment at time i.

Definition 3: Attribute Matrix K. External factors affecting traffic flow are incorporated as enhanced attributes of road segments in the urban road network. These factors form an attribute matrix $K \in \mathbb{R}^{N\times l}$, where l is the number of auxiliary information categories. The categories of auxiliary information can be expressed as $K_j = \{k_{1j}, k_{2j}, ..., k_{tj}\}$, where k_{tij} is the j-th auxiliary information for the i-th road segment at time t. Therefore, spatiotemporal dependency modeling for traffic prediction can be viewed as a mapping function f based on road network G, feature matrix X, and attribute features K [16]. Future traffic speed over time period T is calculated as:

$$[x_{t+1}, ..., x_{t+T}] = f(X, G, K|A)$$

1.2 External Attribute Features

To comprehensively consider factors influencing traffic states, external factors are categorized into dynamic (D) attributes and static (S) attributes. The traffic feature matrix X and attribute matrix $K = \{S, D\}$ are then used as inputs to obtain the augmented matrix [17].

a) Static Attribute Fusion. $S \in \mathbb{R}^{n \times p}$ refers to a set of p different static attributes. Since attribute values do not change over time, matrix S is used continuously, while only the corresponding feature matrix X columns are extracted during the augmented matrix generation process at each timestamp. Specifically, the matrix augmented by static attributes at time t is formed as $[X_t, S]$.

b) Dynamic Attribute Fusion. Unlike static attributes, $D \in \mathbb{R}^{n \times w \times t}$ refers to a set of w different dynamic attributes. Considering the cumulative impact of dynamic factors on traffic states, instead of selecting attribute values corresponding to time t, the window size is extended to m+1 when forming E_t, i.e., $D_{t-m}, ..., D_{t-1}, D_t$. Finally, through the attribute augmentation unit, the augmented matrix containing static and dynamic external attributes along with traffic feature information at time t is formed as:

$$E_t = [X_t, S, D_{t-m}, ..., D_{t-1}, D_t]$$

2 Methodology

2.1 Method Description

This method combines traffic flow features with an attribute augmentation unit, utilizing graph wavelet convolutional layers and an adaptive matrix to extract spatial road network feature information, employing Gated Recurrent Unit layers to extract temporal characteristics, and using an attention mechanism for global spatiotemporal information extraction to facilitate traffic flow prediction. The overall framework is shown in [Figure 1: see original paper]. This paper names it A-STIGCN (Attribute-Augmented Spatiotemporal Graph Convolutional Network Integrating Attention and Graph Wavelet).

2.2 Spatial Feature Extraction

This paper performs convolution operations based on graph wavelet transform, which can capture signal mutations. Similar to graph Fourier transform, graph wavelet transform also maps graph signals to the frequency domain for convolution operations. However, Fourier transform can only decompose time-domain signals into combinations of frequency-domain signals, whereas wavelet transform can reveal the location and time of frequencies with maximum phase variation during signal evolution [17]. Graph wavelet employs a mother wavelet function to decompose signals into different frequency components, which can be represented in graph networks as a set of wavelet bases $\psi_s = \{\psi_{s1}, \psi_{s2}, ..., \psi_{sn}\}$, where ψ_{si} represents the wavelet basis related to node i, s is the scale parameter controlling wavelet size, and λ represents eigenvalues of L.

Graph wavelet bases can be expressed as:

$$\psi_s = U g_s(\Lambda) U^T$$

where $g_s(\lambda) = e^{-\lambda s}$. Graph wavelet can map traffic flow data X to the frequency domain space, obtaining the data representation in the frequency domain as $\hat{X} = X\psi_s^{-1}$. Graph wavelet convolution operation can be defined as:

$$F_{gss} = [\psi_s((F_L) \circledast X)]\psi_s^{-1}$$

Compared with Fourier transform, graph wavelet bases ψ_s and ψ_s^{-1} are more sparse, making computation more efficient. Additionally, graph wavelet transform aggregates local node information to characterize node features, thereby improving method interpretability.

However, spatial relationships in traffic networks are uncertain. Road nodes are not only closely related to neighbor nodes but also to important road nodes located at network centers, which may be far apart. Adjacency matrices based on distance calculation cannot reflect correlations between such roads [18]. Therefore, this paper adopts an adaptive matrix for global spatial feature learning, with the expression:

$$A_{adp} = \mathrm{softmax}(\mathrm{ReLU}(E_c E_c^T))$$

where $E_c \in \mathbb{R}^{N \times c}$ consists of the first c eigenvectors of A. The adaptive matrix and graph wavelet convolution can extract local and global spatial feature information, respectively. Since spatial dependencies are typically nonlinear, a nonlinear activation function is required; this paper employs the ReLU function.

The iterative formula for a graph convolutional layer obtained through a diagonal weight matrix A is:

$$H^{(l+1)} = \mathrm{ReLU}(\psi_s^{-1}(\psi_s H^{(l)} W_1^{(l)}) \oplus A H^{(l)} W_2^{(l)})$$

2.3 Temporal Feature Extraction

Gated Recurrent Unit (GRU) was developed to address the long-term dependency problem of Recurrent Neural Networks (RNN) and represents a simpler recurrent neural network than Long Short-Term Memory (LSTM) networks. GRU uses gating mechanisms to control input and memory information for predictions at the current time step. Gated recurrent units do not clear previous information over time; instead, they retain relevant information and pass it to the next unit, thereby utilizing all information and avoiding the vanishing gradient problem. The network structure is shown in [Figure 2: see original paper].

GRU has two gates: a reset gate and an update gate. The reset gate determines how to combine new input information with previous memory, while the update gate defines the amount of previous memory preserved to the current time step. The expressions are defined as follows:

$$
\begin{aligned}
z_t &= \sigma(W_z x_t + U_z h_{t-1}) \\
r_t &= \sigma(W_r x_t + U_r h_{t-1}) \\
\tilde{h}_t &= \tanh(W_h x_t + r_t \odot U_h h_{t-1}) \\
h_t &= z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t
\end{aligned}
$$

where x_t is the input vector at time step t, W and U are weight matrices, h_{t-1} is information from the previous time step, z_t is the update gate, r_t is the reset gate, $\tilde{h}_t$ represents the candidate state at the current time, σ is the activation function, and h_t represents the final memory at the current time step.

2.4 Spatiotemporal Feature Extraction with Attention Mechanism

In real-world traffic networks, temporal dependency relationships are often complex and not merely sequentially correlated. For example, if a traffic accident occurs at a certain moment, its impact on future time points will be maximal and persist for a long duration, rather than being primarily influenced by the immediate previous moment. This paper introduces an attention mechanism layer to automatically capture different input features, allocating sufficient weights to important information distributions using probability distribution theory [4,15] to enhance traffic flow prediction accuracy. The structure is shown in [Figure 3: see original paper].

The original feature matrix dimension is extended by combining static and dynamic external attributes. Here, X_t is extracted from the original flow feature matrix X, $\{D_{t-m}, ..., D_t\}$ is the collection of dynamic information from time *t-m* to t, and S represents static attributes that remain constant across different timestamps. E_t is the augmented matrix fusing external attributes with original traffic flow features. Using the augmented matrix as input to the spatiotemporal model yields n hidden states (h) containing spatiotemporal features [20]. Graph wavelet networks and the adaptive matrix capture road network topology features to obtain spatial dependencies, while GRU captures dynamic changes in node attributes to obtain local temporal trends in traffic states. These are then input into the attention model, calculated as follows:

$$
e_t = \tanh(W_e h_t + b_e)
$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^{H} \exp(e_i)}$$

$$Q = \sum_{t=1}^{H} \alpha_t h_t$$

where α_t represents the attention probability distribution value determined by output vector h_t at time t, e_t is the weight coefficient at time t, Q represents the output value, and W and b are weights and biases, respectively.

Through data preprocessing, combination of traffic features and external attributes, A-STIGCN model construction, parameter setting, and error evaluation, all extracted features are output using a fully connected Dense layer. The system model flowchart is shown in [Figure 4: see original paper].

3 Experiments

3.1 Experimental Data

a) Traffic Flow Data. SZ-taxi: This dataset contains trajectory data from January 2015 covering 156 major road segments, with connectivity modeled using a 156$×$156 adjacency matrix. Traffic speed time series for selected road segments were calculated to form a feature matrix with rows indexed by road segment and columns by timestamp.

b) Weather Data. SZ-weather: This auxiliary information includes weather conditions for the study area recorded every 15 minutes in January 2015. Weather conditions are categorized into five types: sunny, cloudy, foggy, light rain, and heavy rain. Using time-varying weather condition information, a dynamic attribute matrix of size 156$×$2,976 was constructed.

c) POI Data. SZ-POI: This dataset provides information about Points of Interest (POI) around selected road segments. POI categories are divided into nine types: dining services, enterprises, shopping services, transportation facilities, education services, living services, medical services, and accommodation. After calculating POI distributions for each road segment, the POI type with the largest proportion was used as that segment’ s feature, resulting in a static attribute matrix of size 156$×$1.

3.2 Experimental Setup

3.2.1 Evaluation Metrics During experiments, three metrics were selected to evaluate the model: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Accuracy, defined as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

$$\text{Accuracy} = 1 - \frac{\|Y - \hat{Y}\|_F}{\|Y\|_F}$$

where y_i represents observed values, $\hat{y}_i$ represents predicted values, n is the number of selected samples, and $\|\cdot\|_F$ denotes the Frobenius norm.

3.2.2 Parameter Settings The traffic flow prediction model was implemented using the Google TensorFlow deep learning framework in the PyCharm development environment. To improve model convergence speed and prediction accuracy, the data partitioning strategy was set as: the first 80% of data as the training set and the remaining 20% as the test set. Model hyperparameters primarily include learning rate, epoch, hidden units, and batch size, with learning rate and batch size set to 0.001 and 64, respectively, while other parameters were searched experimentally.

First, training capabilities in the set [50, 100, 150, 200, 300, 350] were tested to analyze model performance variation. [Figure 5: see original paper] shows evaluation results under different training settings. As epoch values increase, evaluation metrics stabilize, with the turning point at 200. Then, with epoch fixed at 200, the number of hidden units was selected from the candidate set [8, 16, 32, 64, 100, 128]. As shown in [Figure 6: see original paper], the model stabilizes when the unit count reaches 100. Therefore, epoch was determined to be 200, and the number of hidden units was set to 100.

3.3 Results and Analysis

3.3.1 Comparison with Traditional Models To verify the effectiveness of the A-STIGCN model in traffic flow prediction tasks, it was compared with baseline models as shown in . Experimental results indicate that the combined prediction model achieves substantial improvement in prediction accuracy, with both MAE and MAPE lower than the average performance of the other three baseline models. Compared with traditional SVR and GRU models, MAE decreased by approximately 3.422 and 2.811, respectively. Compared with GCN, which focuses only on spatiotemporal relationships, MAPE decreased by approximately 1.777%, with significant improvements in other overall evaluation metrics. These comparison results validate the effectiveness of the A-STIGCN combined model.

3.3.2 Comparison with Variant Models As shown in , TGCN exhibits relatively poor prediction performance. Due to the randomness of traffic flow data caused by multiple external factors, compared with the TGCN model without external factors, A-STIGCN reduces MAE by approximately 0.173 and MAPE by approximately 0.637%. Compared with the ASTGCN model without attention mechanism, MAE decreases by approximately 0.131 and accuracy improves by 0.068. These results demonstrate that the A-STIGCN model maintains favorable prediction performance when considering external factors' impact on traffic flow.

3.3.3 Performance at Different Prediction Horizons This experiment tested the A-STIGCN model' s prediction performance at different horizons (15min, 30min, 45min, and 60min). Performance comparisons between the A-STIGCN model and baseline models are shown in . At the 15-minute prediction horizon, A-STIGCN's MAE is approximately 0.047 and 0.097 lower than TGCN and DCRNN, respectively. At the 30-minute horizon, A-STIGCN' s MAPE is approximately 0.107% and 1.12% lower than TGCN and DCRNN, respectively. For the 45-minute horizon, A-STIGCN' s accuracy is 0.015 and 0.024 higher than TGCN and DCRNN, respectively. When the prediction horizon is set to 60 minutes, A-STIGCN' s MAPE is approximately 0.077% and 1.073% lower than TGCN and DCRNN, respectively. These results demonstrate that the proposed model maintains robust performance and long-term prediction capability across different horizons.

3.3.4 Perturbation Analysis Gaussian noise and Poisson noise were added to the data to verify model robustness. The two noise types follow Gaussian distribution $N(0,\sigma)$ with $\sigma \in \{0.2, 0.4, 0.6, 0.8, 1\}$ and Poisson distribution $P(\lambda)$ with $\lambda \in \{1, 2, 4, 8, 16\}$, respectively. As shown in [Figure 7: see original paper], changes in evaluation metrics under different noise settings are negligible, verifying the robustness of the A-STIGCN model.

4 Conclusion

This paper integrates external attribute information and traffic characteristics to form an attribute-augmented matrix, enhancing the model' s perception of external information. By combining graph wavelet networks with an adaptive matrix, local and global spatial characteristics of short-term traffic flow are extracted, while GRU extracts temporal characteristics and attention mechanisms allocate sufficient weights to important information distributions. Comparative experiments were designed to test the prediction performance of the A-STIGCN model, thereby improving traffic flow prediction accuracy. However, the model inadequately considers vehicle type characteristics and time period features on traffic roads. Future work will refine prediction granularity, consider road network complexity to improve the model, and conduct related traffic flow prediction research.

References

[1] Liu Jing, Guan Wei. A Review of Traffic Flow Prediction Methods [J]. Highway Traffic Technology, 2004, (3): 82-85.

[2] Zhang S, Lin K P. Short-term traffic flow forecasting based on data-driven model [J]. Mathematics, 2020, 8 (2): 152-169.

[3] Zhou T, Jiang D, Lin Z, et al. Hybrid dual kalman filtering model for short-term traffic flow forecasting [J]. IET Intelligent Transport Systems, 2019, 13 (6): 1023-1032.

[4] SUN B, SUN T, ZHANG Y J, et al. Urban traffic flow online prediction based on multi-component attention mechanism [J]. IET Intelligent Transport Systems, 2020, 14 (10): 1249-1258.

[5] Cao Yang, MAO Yibo, Shi Quan, et al. Encoder-decoder Deep Learning Model Study for Daily Traffic Flow Prediction [J/OL]. Computer Engineering and Application: 1-11 [2021-11-22]. http://kns.cnki.net/kcms/detail/11.2127.TP.20210913.1903.020.html.

[6] Gui Zhiming, Li Zhuangzhuang, Guo Limin, et al. Short-time traffic flow prediction based on the ACGRU model [J]. Computer Engineering and Applications, 2020, 56 (21): 260-265.

[7] Yu B, Yin H, Zhu Z. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [C]// 27th International Joint Conference on Artificial Intelligence, 2018: 3634-3640.

[8] Li Y, Yu R, Shahabi C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [J]. arXiv preprint arXiv: 1707.01926, 2017.

[9] Yang G, Wang Y, Yu H, et al. Short-term traffic state prediction based on the spatiotemporal features of critical road sections [J]. Sensors, 2018, 18 (7): 2287.

[10] Jiang Shan, Ding Zhiming, Xu Xinrun, et al. Graph Neural Network Model for Road Network Traffic Flow Situation Prediction [J]. Computer Science and Exploration, 2021, 15 (06): 1084-1091.

[11] Dai Junming, Cao Yang, Shen Qinqin, et al. Traffic prediction based on multi-spatiotemporal graph convolutional networks [J]. Computer Application Research, 2022, 39 (03): 780-784. DOI: 10.19734/j.issn.1001-3695.2021.08.0361.

[12] Zhang L, Alharbe N R, Luo G, et al. A hybrid forecasting framework based on support vector regression with a modified genetic algorithm and a random forest for traffic flow prediction [J]. Tsinghua Science and Technology, 2018, 23 (4): 479-492.

[13] Ni Qingjian, Peng Wenqiang, Zhang Zhizheng, et al. Spatiotemporal graph neural network traffic flow prediction based on information-enhanced transmission [J]. Computer Research and Development, 2022, 59 (02): 282-293.

[14] Zhang Yaofang, Chen Jian. Short-time prediction model of highway subvehicle traffic flow based on GBDT algorithm [J]. Highway, 2022, 67 (01): 221-227.

[15] LI M, WANG Y W, WANG Z W, et al. A deep learning method based on an attention mechanism for wireless network traffic prediction [J]. Ad Hoc Networks, 2020, 107: 102258.

[16] Tang J. Sun Y, Sun. "Dynamic Spatial-Temporal Graph Attention Graph Convolutional Network for Short-Term Traffic Flow Forecasting," 2020 IEEE International Symposium on Circuits and Systems (ISCAS), 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9181234.

[17] Zhu Q. Wang C. Tao H, et al. "AST-GCN: Attribute-Augmented Spatiotemporal Graph Convolutional Network for Traffic Forecasting," in IEEE Access, vol. 9, pp. 35973-35983, 2021, doi: 10.1109/ACCESS.

[18] Wu Z, Pan S, Long G, et al., Graph WaveNet for Deep Spatial-Temporal Graph Modeling [C]// Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macao, China, Aug10-16, 2019.

[19] Zhao L, et al. "T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction," in IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 9, pp. 3848-3858, Sept. 2020, doi: 10.1109/TITS.

Note: Figure translations are in progress. See original paper for figures.