Advances in Computer and Communication, 2024, 5(3), 183-189
https://www.hillpublisher.com/journals/acc/
ISSN Online: 2767-2875

Traffic Flow Prediction Based on Graph Convolutional Network and Transformer

Xingbo Zhao

University of Science and Technology of China, Hefei 230026, Anhui, China.

How to cite this paper: Xingbo Zhao. (2024) Traffic Flow Prediction Based on Graph Convolutional Network and Transformer. *Advances in Computer and Communication*, 5(3), 183-189.
DOI: 10.26855/acc.2024.07.005

Received: June 7, 2024
Accepted: July 4, 2024
Published: August 1, 2024

***Corresponding author:** Xingbo Zhao, University of Science and Technology of China, Hefei 230026, Anhui, China.

Abstract

In recent years, the rapid development of global artificial intelligence technology has become an important force in promoting the accelerated development of science and technology, and industry. From finance, medical care, and education to intelligent manufacturing and smart cities, artificial intelligence technology is penetrating into all walks of life, bringing convenience to people's lives. Through traffic flow forecasting, people can better plan their own travel routes and reduce the risk of traffic congestion. Accurate traffic forecast information can help to improve people's travel efficiency and improve urban service level, which is the core technology in the construction of intelligent transportation systems. The task of traffic flow forecasting aims at forecasting the future traffic flow trend according to the historical data information in the traffic network. The key challenge is how to model the temporal and spatial dependence in traffic networks. In the time domain, traffic flow has obvious periodicity characteristics, especially based on the daily periodicity and weekly periodicity characteristics of human social laws. We use the time series Transformer with predefined periodicity information to extract time dependence. In the spatial domain, there are similar traffic patterns in some areas on the macro level and the relationship between adjacent road sections on the micro level. We use the graph convolutional network based on the geographical adjacency matrix and the graph convolutional network based on the sequence semantic similarity matrix to extract the micro and macro spatial dependence respectively. We combine the two deeply and design a traffic flow forecasting model based on a graph convolutional network and Transformer to extract the deep temporal-spatial correlation. Finally, we experiment on several real-world traffic data sets to verify the effectiveness of the model.

Keywords

Traffic Flow Prediction; Spatial-Temporal Sequence Data; Graph Convolutional Network; Transformer

1. Introduction

In recent years, with the development of the social economy and the improvement of people's living standards, more and more people choose to buy private cars to travel. As shown in Figure 1, according to the data of the National Bureau of Statistics, the number of private cars in China has reached 278.73 million in 2022, a year-on-year increase of 6.2%. At the same time, with the acceleration of urbanization, a large number of people and vehicles gather in cities, which greatly increases the burden of road traffic. Traffic congestion has become a common phenomenon, especially in first-tier cities and some second-tier cities, which has caused many troubles for people's travel. Therefore, in the face of the increasing traffic congestion, how to improve the road utilization rate and the upper limit of traffic load is a key problem that we urgently need to solve. With the rapid development of the Internet of Things (IoT) and Cyber-Physical Systems, massive spatio-temporal data are continuously generated from traffic sensors, car navigation, and people's mobile phones. Based on these data, traffic flow forecasting has become the core technology in the construction of Intelligent Transportation System (ITS), which is of great significance to the development of the city. In this regard, domestic and foreign scholars have done a lot of research. From the beginning of traditional time series forecasting methods, such as the ARIMA model, to the gradual introduction of machine learning and deep learning methods into the field of traffic flow forecasting, related technologies are constantly innovating and developing, and the forecasting accuracy is also constantly improving. In recent years, the research hotspots in the field of traffic flow forecasting mainly focus on the model based on graph convolutional network and the model based on Transformer. In this paper, we will introduce these related works and design and propose new model methods on this basis.

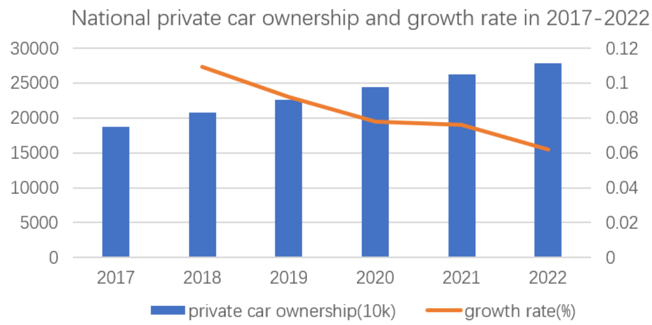

Figure 1. National private car ownership and growth rate.

The core and key challenge of traffic flow forecasting is to model the temporal and spatial dependence. Traffic flow changes with the change of time and space, and is influenced by vehicles, pedestrians, and other interference factors, showing strong randomness and uncertainty, which brings great challenges to our forecasting task. However, this does not mean that the changing law of traffic flow is untraceable. On the one hand, traffic flow has obvious periodic characteristics, which is consistent with the activity law of human society. For example, there are obvious characteristics of early peak and late peak on working days, but there are different laws on non-working days. On the other hand, traffic flow has an obvious correlation in the neighboring streets, and sometimes two distant locations will show similar trends, such as Shopping Mall A and Shopping Mall B, and the streets near them often have similar traffic flow waveforms. At present, the model based on a graph convolutional network makes it difficult to capture the long-distance dependence, because the multi-layer graph convolutional network will bring the problem of smoothness. The model based on the Transformer has obvious advantages in long-time series forecasting, but it is not effective in short-time traffic flow forecasting. In view of the above problems, we launched the research of this paper and proposed a traffic flow forecasting method based on graph convolutional network and Transformer. By combining the advantages of the convolutional network in dealing with road network structure and a Transformer in dealing with sequence data, our model has achieved good results in traffic flow forecasting. At the same time, aiming at the problems mentioned above, we also introduce semantic analysis methods to help the model capture long-distance dependencies.

2. Related Work

In recent years, the task of traffic flow forecasting has been widely discussed by domestic and foreign researchers, and related research results have emerged one after another. Firstly, there are some classical methods based on statistical learning and machine learning, including the historical average model (HA), autoregressive moving average model (ARIMA), K nearest neighbor algorithm (kNN), and support vector machine (SVM) [1, 2]. These algorithms can only consider time information and cannot model spatial characteristics. With the development of artificial intelligence and depth model, in order to better describe the temporal and spatial characteristics of traffic flow forecasting tasks, people put forward the depth neural network model. Nowadays, deep learning has achieved great success in capturing complex temporal and spatial dependencies and has made great breakthroughs compared with traditional methods [3, 4]. Considering that traffic data is a typical structured graph data, people propose to use graph neural network (GNN) to model the spatial relationship. For example, Bing Yu, Haoteng Yin, and others put forward a spatio-temporal graph convolutional network model (STGCN)[5]. The spatial relationship was extracted by graph convolutional neural network (GCN) module and the temporal relationship was extracted by time convolutional module (TCN), and a spatio-temporal data model was constructed. The experimental results of SOTA at that time were obtained on PeMSD7 and BJER4 data sets. Yaguang Li, Rose Yu, and others regard traffic flow as a two-way random walk diffusion process [6], and use this to model and capture spatial dependence, and put forward a diffusion convolutional cyclic neural network (DCRNN) model. Bai, Yao, and others put forward two methods to enhance the current GCN [7]. In order to learn the unique traffic mode of each node, the NAPL module uses separate weights and offsets when performing GCN feature transformation, and at the same time, the node embedding vector is used to adaptively learn the adjacency matrix representation in a data-driven way without being limited by the prior road network structure, and an adaptive graph volume product cyclic network (AGCRN) model is constructed. Chao Song, Youfang Lin, and others constructed an adjacency matrix by combining space blocks and time blocks, designed the method of spatio-temporal synchronization modeling, and proposed the spatio-temporal synchronization graph convolutional network (STSGCN) model. Li, Zhu, and others put forward a data-driven method to generate a "time map" to compensate for the correlation that the space map may not reflect [9]. By integrating the fusion map module and a novel gated convolutional module into a unified layer, we hope to learn more about spatial-temporal dependence and construct a spatio-temporal fusion map neural network (STFGNN) model.

Since the Google research team first proposed the transformer model in the paper "Attention Is All You Need" in 2017 [10], the transformer has been widely used by researchers all over the world. The core of transformer lies in the self-attention mechanism, which enables the model to capture the context in the input sequence, so as to better understand and process the sequence data. Compared with the traditional recurrent neural network (RNN) and convolutional neural network (CNN), the transformer has the advantages of better parallel training and capturing long-distance dependencies, so it performs well in various tasks. With the application of transformers in the fields of natural language processing (NLP) and computer vision (CV), many scholars have begun to apply transformers to traffic flow forecasting. Transformer can better capture the long-term time dependence and solve the problem of error accumulation in the past autoregressive methods. Xu, Dai, and others put forward a spatio-temporal block [11] composed of spatial transformer and temporal transformer, jointly extracted the spatio-temporal features of context dynamic dependence, and constructed a spatio-temporal traffic flow prediction transformer(STTN) model. Jiang, Han and others designed different graph masking methods for long-distance spatial dependence to model local geographical neighborhood and global semantic neighborhood [12], which made attention more focused. At the same time, through the traffic delay perception feature conversion module, K-shape clustering was used to attach traffic modes, and a delay-aware dynamic long-distance transformer (PDFormer) model for traffic flow prediction was designed.

3. Traffic Flow Forecasting Model Based on Graph-convolutional Network and Transformer

1) Definition of traffic flow forecasting problem In this paper, we mainly study the traffic flow data obtained based on the sensors in the traffic network. Each sensor will collect a time series of data over a period of time. By aligning the data of all sensors according to time, a two-dimensional matrix will be obtained, and its horizontal and vertical coordinates are the number of sensors and the length of the time slice respectively. Our research work will mainly focus on this time series data matrix. Below we give a formatted definition of the traffic flow forecasting problem.

Definition: Given N interrelated nodes and their historical observations, we need to predict the future observations of these nodes. We express the correlation of n nodes as a weighted graph G=(V, E, A), where V is the set of nodes, |V|=N, E is the set of edges, and E is a weighted adjacency matrix, which indicates the correlation between nodes. At time t, the observed values of all nodes in the graph can be expressed as, where c is the characteristic dimension of a single node, usually 1. Our goal is to learn a function f, which maps the historical observation values at t moments and graph G to the observation values at t' moments in the future:

$$\left[X^{(t-T+1):t}, G\right] \xrightarrow{f} \left[X^{(t+1):(t+T')}\right]$$

Among them,

$$X^{(t-T+1):t} \in R^{N*C*T}, \quad X^{(t+1):(t+T')} \in R^{N*C*T'}$$

After getting the predicted results, we need some evaluation indicators to measure the gap between the predicted results of the model and the real values. In the task of traffic flow forecasting, the commonly used model evaluation indicators mainly include mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). Below we give the calculation formulas for these three kinds of errors.

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

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

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

Where n is the number of samples, the real value and the predicted value. The average absolute error of the above three error indicators is insensitive to outliers, but it cannot reflect the distribution of prediction errors; The root mean square error is sensitive to outliers, but it can reflect the distribution of errors. The average absolute percentage error can reflect the relative error, but it can't reflect the absolute error. These three evaluation indexes complement each other and can make a comprehensive evaluation of the forecast results of traffic flow.

2) Model design For spatial dependence, we adopt a two-pronged approach of graph convolutional network and semantic similarity analysis. On the one hand, the graph convolutional network can learn the close-range spatial dependence well according to the adjacency matrix of the traffic network, but it is difficult to model the long-distance spatial relationship because of the smoothness of the multi-layer graph convolutional network, so we add the method of analyzing the semantic similarity of the node time series data to learn the nodes with similar traffic patterns on the macro level. Specifically, we use the dynamic time warping algorithm (DTW) to calculate the similarity of the sequence data of two nodes, so as to obtain a semantic similarity matrix, in which k nodes with the highest similarity are set to 1 in each row, and the rest positions are set to 0. The input time series data are sent to the graph convolutional module based on the adjacency matrix and the graph convolutional module based on the semantic similarity matrix respectively, and the output results of the two modules are fused by weighted summation, thus taking into account both macro and micro spatial similarities.

For the time dependence, we use the Transformer module with predefined periodic information to extract the time dependence. Considering the obvious characteristics of daily periodicity and weekly periodicity of traffic flow, we add the position of each moment in daily and weekly to embedding, so that the model can learn the periodic information of the sequence better. Taking a data set with a time interval of 5 minutes as an example, one day corresponds to (60/5)*24=288 pieces of data, so we take 1-288 as the numbers of these data respectively, and there are seven days in the same week, corresponding to numbers 1-7 respectively. We add these two numbers as predefined period information into embedding and add them with the original position information coding of the Transformer to get the final embedding.

The overall structure of our model is shown in Figure 2. The input traffic flow data and road network data will first pass through the data embedding layer and then be input into the multi-layer space-time block. Each space-time block contains two graph convolutional modules and a Transformer module with different functions. Finally, the output of each spatio-temporal block is jointly input into the output layer through Skip Connection to extract the spatio-temporal dependencies at different levels, and finally the final prediction result is obtained after passing through the 1*1 convolutional layer. By using a Transformer to expand the receptive field of the graph convolutional network, we can capture the related nodes which are far away from the central node in the graph. Traditional convolutional networks are usually limited by local receptive fields, and it is difficult to obtain information between distant nodes. However, by introducing Transformer into a graph convolutional network, we can effectively solve this problem. Specifically, we can achieve this goal by introducing a Transformer module into each layer of the graph convolutional network. These Transformer modules can learn the long-distance dependencies between nodes and integrate these relationships into the aggregation process of graph-convolutional networks. In this way, even if some related nodes are far away from the central node, the information between them can be transmitted and aggregated through the Transformer module.

At the same time, a graph convolutional network can also help the Transformer capture complex graph topology information. Graph convolutional network extracts the feature representation of nodes by aggregating local neighbor information in the graph. This method of local aggregation can effectively capture the local structural information in the diagram and pass it to the Transformer module. In this way, Transformer can better understand the complex topological structure in the graph, so as to better transfer and aggregate information between nodes. In addition, the graph convolutional network can also efficiently aggregate related nodes from adjacent areas. The traditional Transformer model usually needs to calculate the self-attention of all input nodes, which will lead to an increase in computational complexity. However, by taking advantage of the local aggregation ability of the graph-convolutional network, we can focus on the adjacent areas, thus reducing the amount of calculation and improving the efficiency of the model.

Fourth, the experiment 1. Data set In this paper, several data sets from the real world are used to train and evaluate our designed model, so as to better verify the effect of the model. The following is the data set we used: (1) METR-LA traffic data set. The data set comes from the highway in Los Angeles, USA, and is collected by the ring detector on the highway. The time interval is five minutes, from March 1, 2012 to June 30, 2012, for a total of four months. It contains 207 sensor nodes. (2) PeMS-BAY traffic data set. This data set comes from California, USA, and is collected by PeMS, a performance measurement system of CalTrans, a California transportation agency. The time interval is five minutes, from January 1, 2017, to May 31, 2017, for a total of six months. It contains 325 sensor nodes. (3) PeMSD7M traffic data set. The data set is also from California, USA, and was collected by PeMS with an interval of five minutes, from May 1, 2012, to June 30, 2012, for a total of two months. It contains 228 sensor nodes. We use 70% of the data set as a training set, 10% as a verification set, and 20% as a test set. 2. Experimental results We trained and tested on three data sets, and predicted the results of 15 minutes (3 steps), 30 minutes (6 steps), and one hour (12 steps) respectively. According to the prediction results of the model, three evaluation indexes, MAE, RMSE, and MAPE, are calculated and compared with the baseline method. The following figure shows our experimental results on METR-LA and PeMS-BAY data sets respectively.

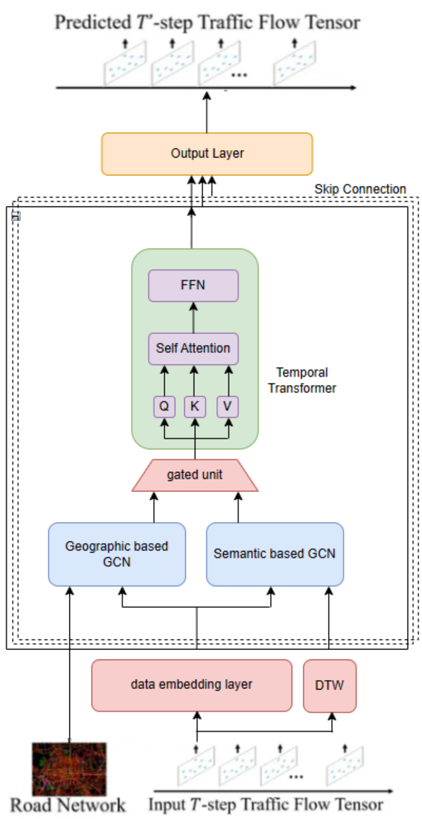

Figure 2. Overall structure diagram of the mode.

Table 1. Experimental results on METR-LA dataset

	15min			30min			60min		
Model	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
HA	14.737	11.103	23.34%	14.737	11.010	23.34%	14.736	11.005	23.33%
ARIMA	14.190	10.522	21.14%	14.317	10.619	21.38%	14.338	10.625	21.43%
LSTNet	8.067	3.914	9.27%	10.181	5.219	12.22%	11.890	6.335	15.38%
STGCN	7.918	3.469	8.57%	9.948	4.263	10.70%	11.813	5.079	13.09%
DCRNN	7.509	3.261	8.00%	9.543	4.021	10.12%	11.854	5.080	13.08%
GraphWavenet	7.512	**3.204**	**7.62%**	9.445	3.922	9.52%	11.485	4.848	11.93%
ASTGCN	7.977	3.624	9.13%	10.042	4.514	11.57%	12.092	5.776	14.85%
GMAN	8.869	4.139	10.88%	9.917	4.517	11.77%	11.910	5.475	14.10%
Our Model	**7.482**	3.209	7.65%	**9.429**	**3.916**	**9.48%**	**11.277**	**4.821**	**11.79%**

Table 2. Experimental results on PeMS-BAY dataset

	15min			30min			60min		
Model	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
HA	6.687	3.333	8.10%	6.686	3.333	8.10%	6.685	3.332	8.10%
ARIMA	6.426	3.085	7.29%	6.458	3.174	7.43%	6.672	3.259	7.91%
LSTNet	3.224	1.643	3.47%	4.375	2.383	5.04%	5.515	2.974	6.86%
STGCN	2.827	1.327	2.79%	3.887	1.698	3.81%	4.748	2.055	5.02%
DCRNN	2.867	1.377	2.96%	3.905	1.726	3.97%	4.798	2.091	4.99%
GraphWavenet	**2.759**	**1.322**	**2.78%**	**3.737**	1.660	3.75%	4.562	1.991	4.75%
ASTGCN	3.057	1.435	3.25%	4.066	1.795	4.40%	4.770	2.103	5.30%
GMAN	4.219	1.802	4.47%	4.143	1.794	4.40%	5.034	2.186	5.29%
Our Model	2.821	1.325	2.79%	3.792	**1.656**	**3.73%**	**4.480**	**1.926**	**4.51%**

According to the above experimental results, it can be seen that our model has been improved compared with the baseline model in many data sets, and has achieved good results in short-term traffic flow forecasting.

References

[1] Jeong Y S, Byon Y J, Castro-Neto M M, et al. Supervised weighting-online learning algorithm for short-term traffic flow prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(4): 1700-1707.

[2] Williams B M, Hoel L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results [J]. Journal of Transportation Engineering, 2003, 129(6): 664-672.

[3] Yin X, Wu G, Wei J, et al. Deep learning on traffic prediction: Methods, analysis, and future directions [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(6): 4927-4943.

[4] Pan Z, Liang Y, Wang W, et al. Urban traffic prediction from spatio-temporal data using deep meta learning [C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019: 1720-1730.

[5] Yu B, Yin H, Zhu Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting [A]. 2017.

[6] Li Y, Yu R, Shahabi C, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [A]. 2017.

[7] Bai L, Yao L, Li C, et al. Adaptive graph convolutional recurrent network for traffic forecasting [J]. Advances IN Neural Information Processing Systems, 2020, 33: 17804-17815.

[8] Song C, Lin Y, Guo S, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting [C]//Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 34. 2020: 914-921.

[9] Li M, Zhu Z. Spatial-temporal fusion graph neural networks for traffic flow forecasting [C]//Proceedings of the AAAI conference on artificial intelligence: Vol. 35. 2021: 4189-4196.

[10] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.

[11] Xu M, Dai W, Liu C, et al. Spatial-temporal transformer networks for traffic flow forecasting [A]. 2020.

[12] Jiang J, Han C, Zhao W X, et al. Pdformer: Propagation delay-aware dynamic long-range transformer for traffic flow prediction [C]//Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 37. 2023: 4365-4373.