Geographic and Temporal Deep Learning Method for Traffic Flow Prediction in Highway Network

Tianpu Zhang[1,2], Weilong Ding[1,2](✉), Mengda Xing[1,2], Jun Chen[1,2], Yongkang Du[1], and Ying Liang[3]

[1] School of Information Science and Technology, North China University of Technology, Beijing 100144, China
dingweilong@ncut.edu.cn
[2] Beijing Key Laboratory on Integration and Analysis of Large-Scale Stream Data, Beijing 100144, China
[3] Research Center for Ubiquitous Computing Systems, Institute of Computing Technology, Chinese Academy of Science, Beijing, China

Abstract. Traffic congestion has become an inevitable situation faced by all countries and the prediction accuracy of traffic flow, as one of the means to solve this problem, still needs to be improved. Most studies lack the consideration of the influence of multiple factors such as spatial factors, time series factors and other external factors, which makes the prediction effect of traffic flow unsatisfactory. In this paper a method is proposed based on deep learning that can capture the geographic spatial relationship among toll stations, the dynamic temporal relationship of historical traffic flow, extreme weather and calendar types. On the three metrics of MAPE, MAE, and RMSE, the prediction effect of our model has increased by 30% compared with KNN, GBRT and LSTM models.

Keywords: Deep learning · Graph convolutional network · Long short-term memory · Traffic flow prediction · Spatio-temporal data

1 Introduction

With the growth of population and economy, many serious problems have been brought to modern cities, and traffic congestion has undoubtedly become one of the main problems faced by most countries in the world. Especially in the rush time of weekdays or holidays, the large number of vehicles brought huge challenges to the current capacity of highway. Therefore, it is indispensable to plan highway reasonably and provide correct guidance for people to travel. In recent years, with the development of artificial intelligence, more and more researchers are more inclined to use artificial intelligence technology to solve highway planning and guild people to travel. Intelligent transportation system (ITS) [2,37] is one of the important products brought about by the development of artificial intelligence. Among them, traffic flow prediction, as an important part of the

H. Gao and X. Wang (Eds.): CollaborateCom 2021, LNICST 407, pp. 385–400, 2021.
https://doi.org/10.1007/978-3-030-92638-0_23

intelligent transportation system and the basic index to solve traffic congestion problem, undoubtedly plays an important role. By analyzing and predicting the traffic flow, it is possible to learn about the distribution of vehicles in certain node at highway network, and then to regulate and divert the traffic flow so as to achieve the effect of alleviating traffic congestion. Traffic flow prediction is to predict the number of vehicles passing through the certain area in a certain periodic time. According to the length of periodic time, traffic flow prediction can be divided into two types, namely short-term traffic flow prediction that its time period is less than one day and long-term traffic flow prediction that its time period is more than one day. At the same time, according to the different spatial scopes, traffic flow prediction can be divided into single-site traffic flow prediction and the entire network traffic flow prediction.

Although traffic flow prediction has been developed for decades, in the actual situations, the traffic flow prediction at highway toll stations still has been facing many challenges. These problems can be divided into three points. The first point is the lack of consideration of the complex spatial dependence of traffic flow at highway toll stations. Because the highway is a kind of network structure where the vehicle in any toll station can reach the designated location, different toll stations have different influences on the predicted toll station. The second point is the lack of consideration of the temporal dependence of traffic flow at highway toll stations. Through analysis [1,39,40], the traffic flow at the same toll station shows a non-linear change over time. That means, different times of the same toll station have different effect on the moment to be predicted. The third point is the lack of consideration of the impact of extreme weather and calendar types on the prediction of traffic flow at highway toll stations. Under different weather conditions, the traffic flow will change differently. For example, in extreme weather conditions, highway may close for individual safety. By doing so, it may cause a sudden drop in traffic flow in this area. In addition, the date type is another influencing factor. According to our observation, there is a significant increase in traffic flow can on weekends and holidays.

In this paper, a GTM (Geographic and Temporal Model) is proposed, which based on deep learning method. Taking into account the spatial relationship among the toll stations, the historical temporal characteristics and the calendar types and the weather conditions of the toll stations, the GTM model can obtain a more accurate prediction of the traffic flow at the highway toll stations. The rest of this paper is organized as follows. At Sect. 2, we describe related work with our paper. Then at Sect. 3, we describe construction process of our prediction method of traffic flow at highway toll stations. At Sect. 4, we introduce two experiments to illustrate excellent effect of our method. Finally at Sect. 5, we conclude our work and discuss our future research direction.

2 Related Work

Nowadays, a large amount of work uses advanced technologies such as big data to process large amounts of historical traffic data to obtain useful information and functions, and then use artificial intelligence methods to obtain prediction results. We will introduce these methods from three aspects.

The first type of methods that is used to predict traffic flow at highway toll stations is traditional statistical manners. ARIMA (Autoregressive Integrated Moving Average model) [3,4,33], VAR (Vector autoregressive models) [5] and HA (Historical Average) [6] are all classical statistical methods that used to predict traffic flow. Definitely, it is the advantage of statistical methods to depend on single factor, easily implement and quickly compute result. In the early developing stage of intelligent transportation system, these methods are used to predict traffic flow. However, these methods have obvious weaknesses, that is, they only consider historical temporal factors without considering other influencing factors. Since, especially facing the problem of traffic flow prediction that affected by temporal, spatial and external factors, statistical methods can not get better prediction accuracy. The second type of methods that is used to predict traffic flow at highway toll stations is machine learning methods. These models are usually used to predict traffic flow including SVR (Support Vector Regression) [7], KNN (K-Nearest Neighbor) [1,8], Bayesian model [9,10] and GBRT (Gradient BoostRegression Tree) [32]. With the appearance of machine learning methods, the problem of single feature dependence has been successfully solved. These methods get final predicted result by analysing the inner relationship of temporal and other features. The ability to measure the weight relationship between different features is the advantage of machine learning. However, machine learning methods too depend on handcraft feature engineering to get better predicted accuracy, especially when face massive and complicated features, such as prediction traffic flow problem. The third type of methods that is used to predict traffic flow at highway toll stations is deep learning methods. Recently, as a branch of machine learning, deep learning methods have become state-of-the-art technology and have been applied to various fields [11–13]. At the same time, due to the characteristics of deep learning methods that are good at capturing temporal and spatial features, deep learning has also been widely used in traffic flow prediction problem [14,15]. Traffic flow prediction has been studied for decades as a common time series problem. In these studies [16–19], people have achieved better prediction results by using the ability of LSTM (long short-term memory) and GRU (Gated Recurrent Unit) to model complex functions and the characteristics of dynamically capturing time relationships. However, these studies ignore the fact that traffic flow prediction is not a simple time series problem. It is also affected by spatial relationship. Therefore, under the assumption of ignoring the impact of spatial relationship, the prediction results are often not very good. In response to this problem, related researchers use CNN (Convolutional Neural Network) to obtain spatial relationship. Researchers use grids to divide the highway network, and then use CNN to extract the spatial relationship between adjacent toll stations to obtain prediction results, such as these [20–23,38]. However, CNN always maps the traffic flow prediction problem in non-Euclidean space to Euclidean space, which leads to the loss of spatial information. More recently, with the popularity of GCN developed by [24,25], more researchers prefer to choose GCN to obtain spatial information, such as [15,26–31]. Because GCN extends the convolution operation to the non-Euclidean space, this operation is more in line with the real graph structure data.

Therefore, in this paper, we propose a geographic and temporal deep learning method that belongs to the third type of methods we mentioned above for traffic flow prediction in highway network. In order to overcome the shortcomings of the above models that lack multi-dimensional features considerations, our model uses the advantages of LSTM to capture temporal series features and GCN to capture spatial features, and fully considers the spatio-temporal factors, extreme weather features and calendar type to achieve more accurate prediction results in real cases.

3 Feature Pre-processing and Prediction Method

3.1 Feature Engineering

In order to acquire correlation among toll stations, we first need to construct highway graph. The graph definition as follow:

Definition 1 (Highway Graph). *A highway graph is represented by a undirected graph $G = (V, E)$. Where V is a collection representing all toll stations on highway graph. E is also a collection representing all the edges on the highway. $e_{ij} \in E$ represents correlation between toll station v_i and v_j, here $v_i, v_j \in V$ and i, j are a positive integer.*

In this paper, we constructed a highway graph from the perspective of geographic relations. Meanwhile, highway graph is undirected graph, because in actual situations any points in the highway road network is connected to other points. At the same time, the graph is constructed from a geographical perspective, which conform to the business logic, so that the final prediction results have better interpretability.

Definition 2 (Daily Traffic Flow of Highway Toll Stations). *For any toll station v_i, the daily traffic flow is expressed as $s_{v_i}^t$. It represent the total amount of vehicles that passed this toll station v_i, on the day t. Here, t represent current date, $v_i \in V$.*

From daily traffic flow, we construct historical traffic flow S_{v_i} of toll station v_i. It is represented by the formula (1). Here, d represents the length of the time window of historical data we need.

$$S_{v_i} = \begin{bmatrix} s_{v_i}^1 \\ s_{v_i}^2 \\ \cdots \\ s_{v_i}^d \\ \cdots \end{bmatrix}^{\mathrm{T}} \tag{1}$$

Definition 3 (External Factors). *We use $P_t^{v_i} = (W_t^{vi}, D_t)$ to represent the external factors of the v_i toll station on the day t. Here, W_t^{vi} represents weather condition of the v_i toll station on the day t, D_t represents calendar type of the v_i toll stations on the day t.*

Firstly, we introduced extreme weather. Through analysis, we found that the traffic flow at highway toll stations fluctuates significantly under extreme weather conditions, while under good weather conditions, the traffic flow tends to stabilize. We use the label encoder method to divide the weather conditions into two types. The extreme weather conditions include heavy rain, heavy fog, and strong wind that we analysis from the raw data. We defined this factor as formula (2).

$$W_t^{vi} = \begin{cases} 0, & otherwise \\ 1, & extreme\,weather \end{cases} \tag{2}$$

Secondly, the calendar type plays an important role in the prediction of the traffic flow at the highway toll stations. We also use the label encoder method to divide the date types into three categories, namely holidays, weekends and others. During holidays, the charging mode of highway will be adjusted and people will have more leisure time to travel so that traffic flow can be effect by this factors. We define this feature as formula (3).

$$D_t = \begin{cases} 0, & otherwise \\ 1, & if\,t\,is\,a\,holiday \\ 2, & if\,t\,is\,a\,weekend \end{cases} \tag{3}$$

Definition 4 (Traffic Flow of Upstream Toll Stations). *In order to get better spatial relationship among toll stations, we use Vol^{v_it} that is a vector represents the traffic flow of upstream toll station related to vi on the day t. Upstream toll stations are generally considered to be toll stations where vehicles enter the highway. The specific analysis method can refer to our previous work [1]. This vector can be represented by the formula (4).*

$$Vol^{v_it} = \begin{bmatrix} vol_1^{v_it} \\ vol_2^{v_it} \\ \cdots \\ vol_k^{v_it} \\ \cdots \end{bmatrix}^{\mathrm{T}} \tag{4}$$

3.2 Graph Construction

The creation of the highway graph is very important for us to extract the spatial correlation among toll stations. A reasonable graph construction method can greatly improve prediction accuracy of our model. Therefore, we created our highway topology graph based on three rules. The three rules are defined as follows.

Connectivity rule. This rule stipulates that the graph we construct is a connected graph. This rule is in line with the actual situation of China's highway, because in real situation, any point is connected to other points. At the same time, this rule also guarantees that we can achieve the traffic flow prediction for all toll stations in highway.

Neighborhood rules. This rule specifies how we choose toll stations adjacent to node v_j. This means that we will select adjacent toll stations based on the geographic topology of the highway.

Bidirection rules. This rule ensures that the highway graph we create is an undirected graph. This rule is in line with the actual Chinese highway scene, because any two toll stations are bidirection on the highway. You can reach v_j from v_i, and you can also reach v_i from v_j.

3.3 Traffic Flow Prediction

After the feature processing, the highway graph G we constructed, daily traffic flow $S_{v_i}^t$, external factors $P_t^{v_i}$ and upstream traffic flow $Vol^{v_i t}$ will be input into our model, the specific model structure can refer to the Fig. 1. This method is divided into four stages, namely feature engineering stage, feature extraction stage, feature concatenate stage and linear stage.

In the feature engineering stage, in order to obtain the spatial topological attributes of the highway, we create a highway graph G according to the rules specified above. In our work we use adjacency list to represent highway graph as shown in Fig. 2.

In this example, the stations that can be reached directly from Anyangbei toll station are connected by arrows, such as Anyang toll station. At the same time, in order to make better predictions, we also extract historical traffic flow features, traffic flow of upstream toll station and external features that include calendar types and extreme weather.

In the spatio-temporal feature extraction stage, we use GCN that is proposed by Kipf and Welling [36] to extract the spatial relationship of the highway graph and LSTM to extract the temporal relationship of daily traffic flow. In the GCN part, we use two layers of GCN, the formula is as follows.

$$X^{(1)} = Relu(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}XW) \tag{5}$$

$$X^{(2)} = Relu(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}X^{(1)}W) \tag{6}$$

Here, X represents the feature of toll stations in graph, including historical traffic flow S_{v_i} and $Vol^{v_i t}$ traffic flow of upstream toll stations mentioned in the definition 2 and definition 4 respectively. $\hat{A} = A + I_N$, A is the adjacency matrix of highway network G, I_N represents identity matrix of size N. And $\hat{D}_{ii} = \sum_j \hat{A}_{ij}$, $i, j \leq N$, $\hat{D}_{ii}$ represents the value of the i-th row and i-th column in $\hat{D}$. The weight matrix W is a learnable parameter. $X^{(1)}$ and $X^{(2)}$ represent the output of the first layer GCN and the second layer GCN respectively. They have the same feature count of 64.

At the same time, we input historical traffic flow data into the LSTM to obtain the temporal characteristics of the toll stations. As a kind of recurrent

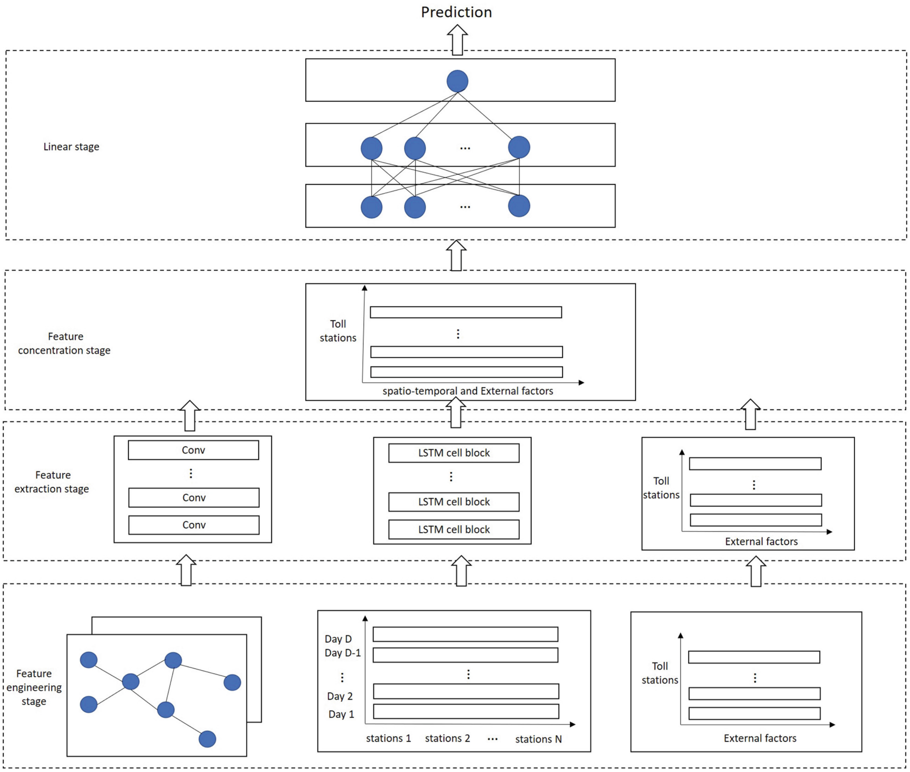

Fig. 1. Traffic flow predictoin method

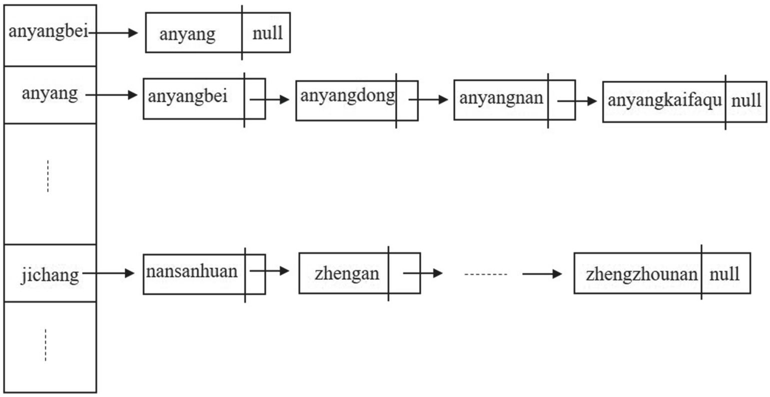

Fig. 2. Adjacency list of highway graph

neural network, LSTM solves the problem of gradient disappearance and explosion caused by too deep sequence by adding a gating mechanism. LSTM consists of input gate, forget gate, output gate and cell state. These components make LSTM memorize the time series and forget the unimportant parts of the time series. The specific calculation steps can be referred to as follows.

$$f_t = \sigma(W_f[h_{t-1}, S_V^t] + b_f) \tag{7}$$

$$i_t = \sigma(W_i[h_{t-1}, S_V^t] + b_i) \tag{8}$$

$$o_t = \sigma(W_o[h_{t-1}, S_V^t] + b_o) \tag{9}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot tanh(W_c[h_{t-1}, S_V^t] + b_c) \tag{10}$$

$$h_t = o_t \odot tanh(C_t) \tag{11}$$

In this part, our input is S_V^t, which represents the traffic flow of all toll stations on day t. σ and $\odot$ are a logistic sigmoid function and an elementwise multiplication, respectively. f_t, i_t, and o_t represent the forget gate, input gate and output gate respectively. In these formulas, W_* is the weight matrix of different gates, and b_* is the bias of different gates. At the same time, both W_* and b_* are learnable parameters. C_t represents the cell state on day t, which will be passed to the next cell on day $t+1$. h_t represents the temporal relationship learned by LSTM up to day t. The time window of historical traffic flow as same long as GCN part. We use historical traffic flow feature matrix for training, and use the next day's traffic flow as a label, to train LSTM to learn the temporal features of traffic flow. Finally, we get the temporal feature matrix h_t.

In feature concatenate stage, we merge the spatial feature matrix $X^{(2)}$, temporal feature matrix h_t and the external factor matrix P_t^V to form a new feature matrix M, and then enter the matrix into the linear stage. We use a three-layer fully connected network, including an input layer, a hidden layer, and an output layer to form a linear stage. In the input layer and the hidden layer, the number of neurons is 64. According to our task of predicting traffic flow in the output layer, the number of neurons is 1. The specific calculation steps can refer to the following.

$$M^{(1)} = Relu(MW_1 + b_1) \tag{12}$$

$$M^{(2)} = Relu(M^{(1)}W_2 + b_2) \tag{13}$$

$$M^{(3)} = M^{(2)}W_3 + b_3 \tag{14}$$

Here, W_* and b_* represent weighted matrix and bias respectively and both are learnable parameters. $M^{(*)}$ represents output of fully connected network. $M^{(3)}$ is the final traffic flow prediction of highway toll stations.

4 Evaluation

4.1 Settings

Our experimental data comes from real system Henan Highway Management System [34,35]. The system displays real-time traffic flow of highway toll stations in Henan Province. Based on this system, we collected traffic data, weather conditions, and calendar types at 269 toll stations from May 2017 to September 2017 to perform this traffic flow prediction task. For historical traffic flow S_{v_i}, we choose time window $d = 15$ and for $Vol^{v_i t}$ traffic flow of upstream toll stations, we choose $k = 3$. The system is built on a big data framework, and HBase 1.6.0 is a database for storing traffic flow data. We built the storage system with 3 servers each of witch server have 4 cores CPU, 22 GB RAM and 700 GB storage. Our method implementation is based on python 3.6 and the open source deep learning framework pytorch 1.7.0. The configuration of machine used to train our model is Intel (R) Core (TM) CPU i7-9750 2.59 GHz, 16 GB RAM, 1TB storage and one NVIDIA GeForce GTX 1660 Ti.

In order to evaluate the prediction effect of our model on traffic flow, we selected three evaluation indicators, namely root mean square error (RMSE), mean absolute percentage error (MAPE) and mean absolute error (MAE). Here, n is an integer, representing the number of all toll stations, $\hat{y}_i$ represents the traffic flow prediction for the v_i toll station and y_i represents the ground truth of traffic flow for the v_i toll station.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{15}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{16}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{17}$$

4.2 Experiments

We first evaluate the effect of temporal features in our proposed model GTM. We designed an experiment to compare the prediction result gained by our model with variant model without LSTM.

Experiment 1: The Effect of Temporal Characteristics. In this experiment, we used the GTM model and its variant model $\mathrm{GTM_{noL}}$ to predict the traffic flow of 269 toll stations. We selected Xuchangdongqu from all the toll stations of Henan highway to compare and analyze the differences between the two models. Because the traffic flow of most toll stations is similar to this station, this station can represent other stations with generality. The comparison chart of prediction results can refer to Fig. 3.

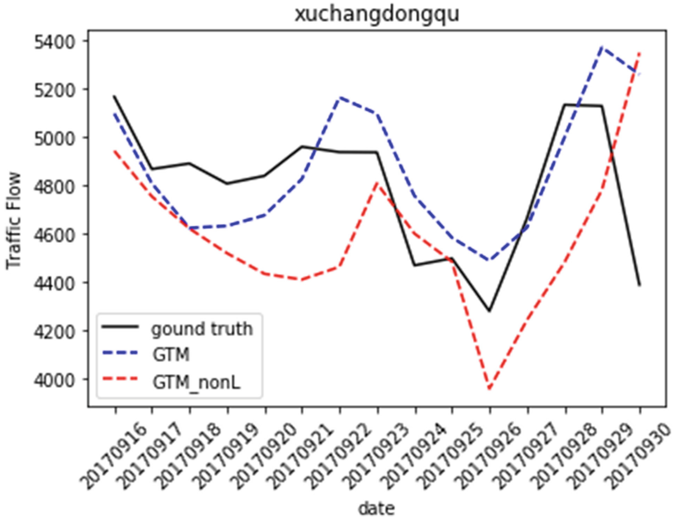

Fig. 3. The prediction traffic flow with GTM and its variant

GTM_{noL} is a variant of our GTM model that does not use LSTM to obtain the temporal characteristics of the traffic flow at the toll station. In Fig. 3, it can be observed that our model GTM, represented by the blue line, is more in line with the real situation represented by the black line. At the same time, at the wave peaks and wave troughs of traffic flow, our model also has good learning ability and can accurately predict. In order to compare our models more closely, we have given a Table 1 and a Table 2.

Table 1. Prediction performances at Xuchangdonqu

	GTM	GTM_{noL}
RMSE	282.025	421.997
MAE	207.4274	351.864
MAPE(%)	4.437	7.369

Table 2. Prediction performances at all toll stations

	GTM	GTM_{noL}
RMSE	573.256	682.393
MAE	372.092	403.861
MAPE(%)	28.069	23.185

The Table 1 shows the prediction results of the GTM model and its variant model GTM_{noL} for Xuchangdongqu. It can be seen from the table that the GTM model is higher than the GTM_{noL} model in all three metrics. The GTM model is improved by 33.16% on the RMSE metric. Through this comparison, it can be shown that we have achieved better results using LSTM to extract temporal features. Therefore, the temporal feature is an important role in the prediction of the traffic flow of highway toll stations. At the same time, the Table 2 is the prediction average result for all toll stations. From the network-wide perspective, it also shows that the accuracy of the model GTM prediction is higher than GTM_{noL} model. The GTM model is 15.99% higher than the GTM_{noL} model on the RMSE metric. However, the MAPE metric did not improve as we expected to the RMSE and MAE metrics. We guess that the reason for this problem is that the model GTM_{noL} without LSTM has a strong learning ability for toll stations that are more associated with other toll sites, because it mainly relies on GCN to obtain the spatial relationship of toll stations. For the GTM model, after adding LSTM to extract temporal feature, our model learns more comprehensively, thereby reducing the learning ability of toll stations that are more associated with other toll sites. In general, the MAPE metric of the two models still belong to same level. Therefore, from the three indicators and the line cart, it can be seen that the temporal features extracted by the LSTM model play an important role in predicting the traffic flow on the highway.

Experiment 2: Analysis of the Accuracy of Traffic Flow Prediction Results on Weekdays and Weekends. To show more clearly that our GTM model has obtained better results for the traffic flow prediction of highway toll stations, we further compared the traffic flow prediction results that are obtained by GTM with results that are obtained by variant GTM_{noL} model under different calendar types. Since our testing set uses September 16, 2019 to September 30, 2019, which contains two weeks, we randomly selected two group weekdays and weekends from these two weeks. 20170916 and 20170920 belong to the weekday and weekend of the first week, and 20170928 and 20170930 belong to the weekday and weekend of the second week.

The traffic flow prediction results in highway toll stations that are obtained by GTM and its variant model GTM_{noL} are illustrated in Fig. 4 and Table 3. From the scatter diagram, the GTM model can capture the real traffic flow of all toll stations well either on weekdays or on weekends. In addition, according to the metrics in the Table 3, we can observe that the GTM model obtains better results than the GTM_{noL} model in RMSE metric. This phenomenon shows that it is very necessary to comprehensively consider the temporal characteristics and spatial characteristics of toll stations. However, the scores of the GTM model on the MAE and MAPE metrics are not consistently better than GTM_{noL}, but the scores of the two models are almost at the same level.

Table 3. Prediction performances in different calendar types

	GTM	GTM_{noL}	GTM	GTM_{noL}
	20170916		20170920	
RMSE	488.494	525.594	410.185	417.627
MAE	334.164	320.777	322.028	283.176
MAPE(%)	25.151	18.435	27.18	16.553
	20170928		20170930	
RMSE	588.093	658.360	1114.432	1578.820
MAE	333.209	420.355	322.028	283.176
MAPE(%)	22.021	20.553	39.266	58.505

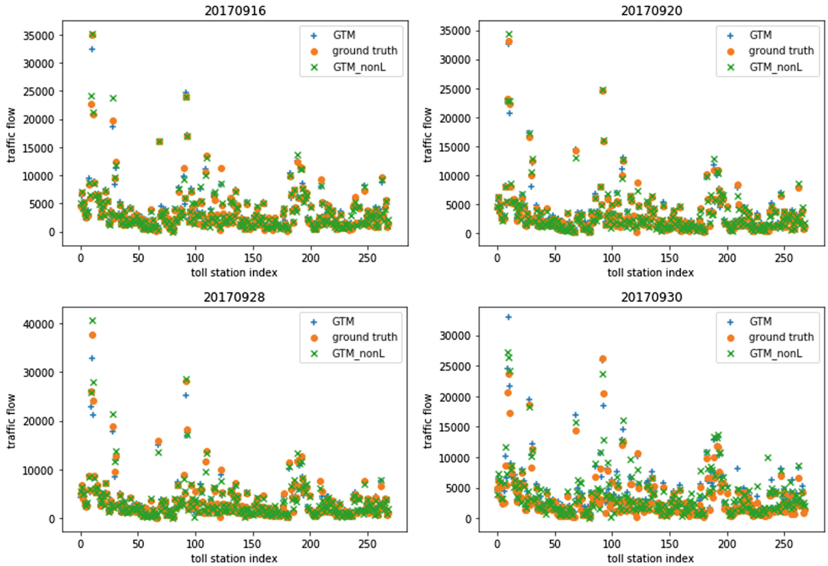

Fig. 4. The traffic flow prediction of toll stations in weekdays and weekends

Experiment 3: Prediction Effects Comparison with Other Models. In this experiment, we use the GTM model to compare with shallow machine learning models such as GBRT [32] and KNN [1], and deep model LSTM often used for time-series problem prediction. In order to clearly illustrate the superiority of our model, we selected two toll stations with a large difference in traffic flow and compared them. Refer to the Fig. 5.

We chose two toll stations, Luoyang and Zhengzhouxinqu. The daily traffic flow of their two stations is completely different. Luoyang's daily traffic flow is between 4000 and 7000, while Zhengzhouxinqu's daily traffic flow is between 15000 and 20000. By comparing the prediction results of these two stations, it can be seen that our model GTM has a good learning ability for all toll stations. At the same time, compared with other models, our model has the highest accuracy of prediction results. For detailed metrics comparison, please refer to the Table 4.

Table 4. Prediction performances of different models

	GTM	GBRT	KNN	LSTM	GTM	GBRT	KNN	LSTM
	Luoyang				Zhengzhouxinqu			
RMSE	208.854	385.904	428.284	522.572	884.770	1566.178	1979.314	1193.320
MAE	176.025	382.159	423.263	411.560	793.238	1556.606	1969.376	1061.904
MAPE(%)	3.475	7.504	8.291	7.669	4.363	8.688	10.985	5.919

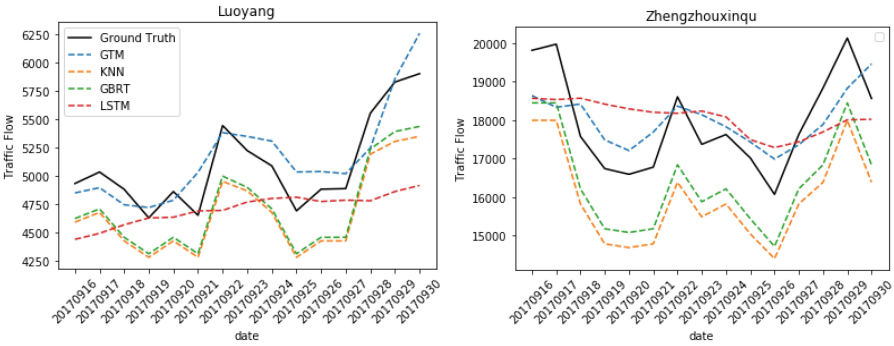

Fig. 5. Comparison of prediction accuracy of different traffic flow

From the Table 4, we can conclude that our model outperforms other algorithms in the three metrics of RMSE, MAE and MAPE. This also shows that our model can well capture the spatial and temporal characteristics of the traffic flow prediction problem at toll stations through the combination of GCN and LSTM.

5 Conclusion

In this paper, we proposed a deep learning method that can take into account spatio and temporal factors and has achieved good results in dealing with the prediction of highway toll stations. GTM is a model based on deep learning, combined with characteristics that GCN can capture the relationship between toll stations in the perspective non-Euclidean space and LSTM can capture the dynamic temporal factors of historical traffic flow. Experiments show that our model can capture the wave peaks and wave troughs of traffic flow at highway toll stations very well, and has improved three metrics compared with other models. Our model improves on the RMSE metric by an average of 30%. In future work, we will try to construct highway graoh from different perspectives to obtain spatial information from different perspectives.

Acknowledgments. This work was supported by National Natural Science Foundation of China (No. 61702014), Beijing Municipal Natural Science Foundation (No. 4192020) and 2020 "Shipei Plan" Project for the Cultivation of High-level Talents in Beijing Colleges and Universities (No. 21XN217).

References

1. Ding, W., Wang, X., Zhao, Z.: CO-STAR: a collaborative prediction service for short-term trends on continuous spatio-temporal data. Futur. Gener. Comput. Syst. **102**, 481–493 (2020)

2. Yuan, J., Zheng, Y., Xie, X., Sun, G.: Driving with knowledge from the physical world. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery Data Mining (KDD), pp. 316–324 (2011)
3. Hamed, M.M., Al-Masaeid, H.R., Said, Z.M.B.: Short-term prediction of traffic volume in urban arterials. J. Transp. Eng. **121**(3), 249–254 (1995)
4. Lingras, P., Sharma, S.C., Osborne, P., Kalyar, I.: Traffic volume time-series analysis according to the type of road use. Comput.-Aided Civil Infrastruct. Eng. **15**(5), 365–373 (2000)
5. Zivot, E., Wang, J.: Vector autoregressive models for multivariate time series. In: Zivot, E., Wang, J. (eds.) Modeling Financial Time Series with S-Plus R, pp. 385–429. Springer, New York (2006). https://doi.org/10.1007/978-0-387-32348-0_11
6. Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.Y.: Traffic flow prediction with big data: a deep learning approach. IEEE Trans. Intell. Transp. Syst. **16**(2), 865–873 (2015)
7. Wu, C.-H., Ho, J.-M., Lee, D.T.: Travel-time prediction with support vector regression. IEEE Trans. Intell. Transp. Syst. **5**(4), 276–281 (2004)
8. Zhang, X., He, G., Lu, H.: Short-term traffic flow forecasting based on K-nearest neighbors non-parametric regression. J. Syst. Eng. **24**(2), 178–183 (2009)
9. Sun, S., Zhang, C., Yu, G.: A Bayesian network approach to traffic flow forecasting. IEEE Trans. Intell. Transp. Syst. **7**(1), 124–132 (2006)
10. Anacleto, O., Queen, C., Albers, C.J.: Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. J. Roy. Stat. Soc. Ser. C (Appl. Stat.) **62**(2), 251–270 (2013)
11. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436 (2015)
12. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
13. Larochelle, H., Bengio, Y., Louradour, J., Lamblin, P.: Exploring strategies for training deep neural networks. J. Mach. Learn. Res. **10**(Jan), 1–40 (2009)
14. Guo, S., Lin, Y., Feng, N., Song, C., Wan, H.: Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 922–929 (2019)
15. Cui, Z., Henrickson, K., Ke, R., Wang, Y.: Traffic graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting. IEEE Trans. Intell. Transp. Syst. **21**(11), 4883–4894 (2019)
16. Luo, Y., Cai, X., Zhang, Y., Xu, J., et al.: Multivariate time series imputation with generative adversarial networks. In: Advances in Neural Information Processing Systems, pp. 1596–1607 (2018)
17. Rangapuram, S.S., Seeger, M.W., Gasthaus, J., Stella, L., Wang, Y., Januschowski, T.: Deep state space models for time series forecasting. In: Advances in Neural Information Processing Systems, pp. 7785–7794 (2018)
18. Cao, W., Wang, D., Li, J., Zhou, H., Li, L., Li, Y.: Brits: bidirectional recurrent imputation for time series. In: Advances in Neural Information Processing Systems, pp. 6775–6785 (2018)
19. Lai, G., Chang, W.-C., Yang, Y., Liu, H.: Modeling long-and short-term temporal patterns with deep neural networks. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 95–104 (2018)
20. Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X.: DNN-based prediction model for spatio-temporal data. In: Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 1–4 (2016)

21. Zhang, J., Zheng, Y., Qi, D.: Deep spatio-temporal residual networks for citywide crowd flows prediction. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
22. Bai, L., Yao, L., Kanhere, S.S., Yang, Z., Chu, J., Wang, X.: Passenger demand forecasting with multi-task convolutional recurrent neural networks. In: Yang, Q., Zhou, Z.-H., Gong, Z., Zhang, M.-L., Huang, S.-J. (eds.) PAKDD 2019. LNCS (LNAI), vol. 11440, pp. 29–42. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16145-3_3
23. Yao, H., et al.: Deep multi-view spatial-temporal network for taxi demand prediction. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)
24. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: Advances in Neural Information Processing Systems, pp. 3844–3852 (2016)
25. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
26. Chen, C., et al.: Gated residual recurrent graph neural networks for traffic prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 485–492 (2019)
27. Li, Y., Yu, R., Shahabi, C., Liu, Y.: Diffusion convolutional recurrent neural networks: data-driven traffic forecasting. In: Proceedings of the International Conference on Learning Representations (2018)
28. Song, C., Lin, Y., Guo, S., Wan, H.: Spatial-temporal sychronous graph convolutional networks: a new framework for spatial-temporal network data forecasting (2020). https://github.com/wanhuaiyu/STSGCN/blob/master/paper/AAAI2020-STSGCN.pdf
29. Wu, Z., Pan, S., Long, G., Jiang, J., Zhang, C.: Graph wavenet for deep spatial-temporal graph modeling. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence, pp. 1907–1913. AAAI Press (2019)
30. Chen, W., Chen, L., Xie, Y., Cao, W., Gao, Y., Feng, X.: Multirange attentive bicomponent graph convolutional network for traffic forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence (2020)
31. Diao, Z., Wang, X., Zhang, D., Liu, Y., Xie, K., He, S.: Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, pp. 890–897 (2019)
32. Ding, W., et al.: An ensemble-learning method for potential traffic hotspots detection on heterogeneous spatio-temporal data in highway domain. J. Cloud Comput. Adv. Syst. Appl. **9**, 1–11 (2020)
33. Ding, W., Zhao, Z.: DS-harmonizer: a harmonization service on spatiotemporal data stream in edge computing environment. Wirel. Commun. Mob. Comput. **2018**, Article ID 9354273, 12 p (2018). https://doi.org/10.1155/2018/9354273
34. Ding, W., Zou, J., Zhao, Z.: A multidimensional service template for data analysis in highway domain. Int. J. Internet Manuf. Serv. **7**(4), 290 (2020)
35. Zhou, J., Ding, W.: An evolutionary service solution for spatio-temporal data analysis in highway domain. Int. J. Intell. Internet Things Comput. **1**(1), 43–52 (2019)
36. Kip, F.T.N., Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks (2016)
37. Zhou, J., Ding, W., Zhao, Z., Li, H.: SMART: a service-oriented statistical analysis framework on spatio-temporal big data (short paper). In: Wang, X., Gao, H., Iqbal, M., Min, G. (eds.) CollaborateCom 2019. LNICSSITE, vol. 292, pp. 91–100. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30146-0_7

38. Wang, Z., Ding, W., Wang, H.: A hybrid deep learning approach for traffic flow prediction in highway domain. In: Gao, H., Wang, X., Iqbal, M., Yin, Y., Yin, J., Gu, N. (eds.) CollaborateCom 2020. LNICSSITE, vol. 350, pp. 253–267. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-67540-0_15
39. Ding, W., Wang, Z., Zhao, Z.: A platform service for passenger volume analysis on massive smart card data in public transportation domain. In: Wang, X., Gao, H., Iqbal, M., Min, G. (eds.) CollaborateCom 2019. LNICSSITE, vol. 292, pp. 681–697. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30146-0_46
40. Ding, W., Zhao, Z., Li, H., Cao, Y., Xu, Y.: A passenger flow analysis method through ride behaviors on massive smart card data. In: Romdhani, I., Shu, L., Takahiro, H., Zhou, Z., Gordon, T., Zeng, D. (eds.) CollaborateCom 2017. LNICSSITE, vol. 252, pp. 374–382. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00916-8_35