Traffic flow prediction based on temporal multigraph convolutional neural network

Wenying GUAN *[a], Jiajia ZHENG [b], Yetao WANG [b], Zhengyuan LI [b], Hao LIU [b]
[a]College of Transportation Engineering, Chang' an University, Xi'an 710064, Shaanxi, CHINA
[b]China Road and Bridge Corporation, Beijing 100011, CHINA
*Corresponding author: guan516510@163.com

ABSTRACT

Traffic flow prediction plays an important role in intelligent transportation systems(ITS), but is challenged by the spatio-temporal complexity of traffic flow connections. In order to integrate the traffic temporal correlation, spatial correlation and semantic correlation in road network models, we propose a deep learning framework, Temporal Multi-Graph Convolutional Neural Network(T-MGCN) for traffic flow prediction. Firstly, we identify several semantic correlations and encode the non-Euclidean spatial correlations and heterogeneous semantic correlations between roads into multiple graphs. These correlations were modeled by a multi-graph convolutional neural network. Next, a recurrent neural network model is used to learn the dynamic characteristics of the traffic flow to obtain temporal correlations; Finally, a fully connected neural network is used to fuse spatio-temporal correlations and semantic correlations.

Key words: Intelligent Transportation Systems, traffic prediction, graph convolutional neural network, graph fusio.

1. INTRODUCTION

As one of the key research area in Intelligent Transportation Systems (ITS), traffic flow prediction is a process of analyzing urban road networks, mining traffic patterns, and predicting future traffic conditions (e.g. speed, traffic density) on road networks. Traffic flow prediction enables a variety of intelligent applications, for example, it can help private drivers with route planning and departure scheduling, and help traffic managers improve traffic efficiency and safety. However, traffic flow is a challenging task due to complex spatial, temporal and semantic correlations[1]: (a)Spatial correlation: Urban traffic flow is controlled by the topology of the underlying road network, and the traffic flow of a road has a great impact on the traffic flow of its neighboring roads, which is very intuitive, Furthermore, it's important to note that spatial correlation possesses directionality, with future traffic conditions being more influenced by downstream traffic as opposed to upstream traffic; (b)Temporal relevance: Traffic conditions change over time, hence temporal relevance can be reflected in distance and periodicity. Proximity implies that traffic conditions in recent time periods are more relevant than those in distant time periods, and periodicity refers to the pattern of periodic changes in traffic conditions at certain time intervals; (c)Semantic relevance: Distant roads, due to some potential semantic relevance, may also have some relevance, e.g., functionally similar downtown roads (e.g., residential and commercial areas) usually have similar traffic patterns.

Early traffic flow prediction methods primarily considered the temporal correlations of traffic flow, encompassing approaches based on Kalman filtering, auto-regressive moving averages, and deep learning techniques. However, these methods ignored spatial correlation and therefore, were not able to optimize the prediction performance of the entire road network. In order to characterize spatial correlation, some work has applied convolutional neural networks to spatial modeling. but convolutional neural networks were originally designed for the spatial structure of Euclidean space (e.g., 2d images and regular grids), so they cannot fully adapt to the complex topology of the road network. To address this problem, some recent works have investigated the application of graphical convolutional neural networks to the spatial modeling of the road network. However, these methods only consider the topological relationships between roads to construct the road map, and ignore the semantic factors such as traffic behavior and local functions that can measure the correlation between roads. To overcome the problems existing in traffic flow prediction methods, this paper presents a deep learning framework T-MGCN that combines spatial, temporal, and semantic correlations used for traffic flow prediction on road networks with various global features.

International Conference on Smart Transportation and City Engineering (STCE 2023),
edited by Miroslava Mikusova, Proc. of SPIE Vol. 13018, 130180Z

2. LITERATURE REVIEW

Traffic flow prediction is one of the main research contents of ITS. Traditional methods rely on a knowledge-driven approach, involving an analysis of the physical characteristics of transportation systems and the construction of models through traffic simulation and prior knowledge[2]. Representative approaches include queuing theory models, cellular transport models and microscopic fundamental graph models. Nevertheless, the complex nature of traffic flow, influenced by numerous factors, is difficult to accurately model. Additionally, these models often necessitate hard-to-obtain parameters in real-world settings.

The advent of data collection devices has spurred interest in data-driven traffic flow prediction methods. Among these, statistical models, shallow machine learning models, and deep learning models are notable. In addition, some researchers have proposed a distributed data-driven model predictive control strategy, which integrates the data-driven model with the distributed model predictive control (MPC) algorithm to alleviate the adverse effects of heterogeneous traffic flow with uncertain dynamics[3]. Statistical models predict future values based on time series analysis of previous observations, such as the historical average model[4], the Kalman filter model[5], the ARIMA model[6] and its variations[7][8]. However, statistical models tend to rely on linear assumptions, which fail to adequately capture the nonlinear characteristics of traffic flow. In contrast, machine learning methods, such as the SVR model[9], Bayesian model[10] and K-nearest neighbor model[11], have demonstrated an ability to learn nonlinear traffic patterns and external factors, offering improved prediction results. However, the performance of these machine learning methods is largely dependent on manually designed features, limiting their performance in complex prediction tasks.

To overcome these limitations, researchers have harnessed the power of GCNs to model the spatial intricacies of road networks, extending convolutional operations into non-Euclidean domains rooted in spectral graph theory [12]. For instance, Zhao et al. proposed T-GCN, a framework that leverages GCNs to comprehend road network topology and combines them with Gated Recurrent Units(GRUs) to gain insights into traffic flow dynamics[13]. Zhang et al. took a sequence-to-sequence approach, integrating GCNs into their framework to facilitate multi-step speed prediction[14]. However, it is worth noting that these studies primarily emphasized the utilization of road network topology for constructing graphs. The incorporation of semantic connections between urban roads, such as historical behavioral patterns, local functions and road types, has the potential to further enhance prediction performance. Furthermore, GCNs have found application across a spectrum of spatio-temporal data prediction tasks, encompassing areas such as pedestrian flow prediction[15], and passenger demand prediction[16].

In summary, from traditional knowledge-driven approaches to data-driven techniques, including shallow machine learning and deep learning models, there is a drive to capture the complexities of traffic flow dynamics. The incorporation of GCNs has shown promise in modeling spatial aspects, and there is growing recognition of the importance of semantic correlations between urban roads. As traffic flow prediction continues to advance, it’s necessary to integrate the traffic temporal correlation, spatial correlation and semantic correlation in road network.

3. MODELLING STRUCTURE

3.1 GCN-GRU modeling framework

One of the key issues in traffic flow prediction is to obtain spatial correlation and temporal correlation. GCN can be connected to the data structured as a graph to perform direct convolutional operations, and the data can be used to build topological relationships to capture the structural features of the graph, so as to efficiently extract spatial features on topological graphs for learning. The GRU is capable of recording neural networks with information about the previous time slices. Through the GRU the traffic flow of a past time slice with the demand within the next time slice to obtain the temporal correlation of the slices, and constructs a mapping relationship between the input data and the output data.The GRU incorporates the hidden states passed down from the previous slices when processing the present information. Therefore this paper proposes a spatio-temporal graphical convolutional neural network based on GCN and GRU for traffic flow prediction.

Figure 1 illustrates the architecture of the GCN-GRU combinatorial model used in this paper, which consists of four layers, i.e., input, convolutional, cyclic, and output layers. Initially, given a sample traffic flow S_t, the input layer is processed using a sliding window In_t to obtain a segment order, thus dividing a long matrix of input into multiple segments. Then, for each segment, the convolutional layer performs a convolution operation on the constructed adjacency and functional similarity graphs, capturing the eigenvalues among the neighbors of each vertex in the input graph and sending the eigenvalues that it has to other neighboring vertices, in order to obtain the spatial correlations embedded in the graph and fusing the results to obtain the feature matrix. Again, the loop layer applies GRU to process the sequence of feature matrices obtained from the order of segments, i.e., the temporal correlation is fused into the feature matrix by fusing the hidden states passed down from the previous time segment, and the output is used as the input to the output layer. Finally, the output layer outputs the above processed feature matrix after processing the activation function.

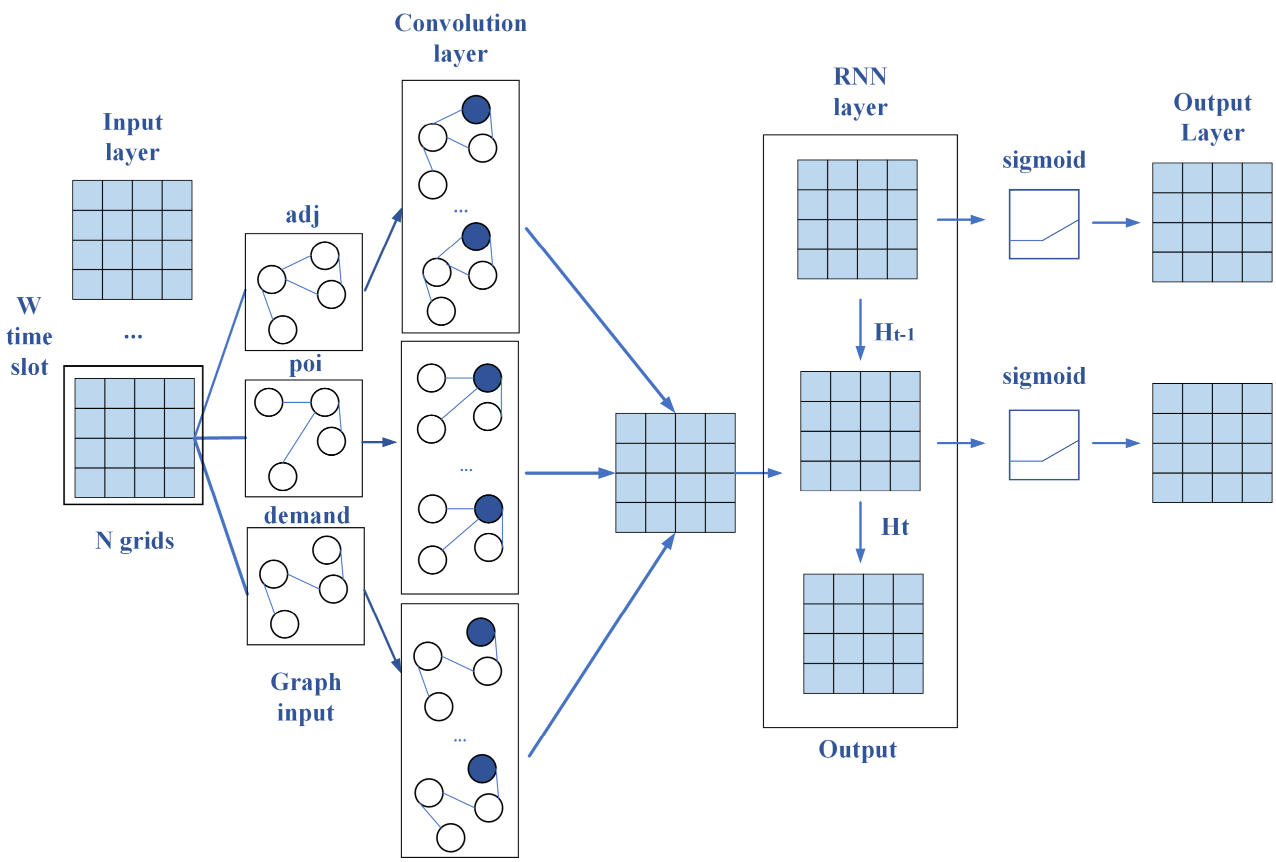

Figure 1 Traffic Flow Prediction Framework Based on GCN-GRU Model.

The input layer is given a sample traffic flow S_t that its input portion In_t can be viewed as $N \times W$ matrix, where N is the number of grids, and W is the number of time segments. In this paper, we use a window of size w and a step size of d sliding window to process In_t. Then, this paper will get k sequence of segments $\left(S_{t1}, S_{t2}, \ldots, S_{tk}\right)$, each of which is a $N \times W$ matrix. The rationale for using this processing strategy is as follows: Firstly, each segment should contain multiple time segments containing the number of traffic flows($w > 1$), because time series data have similar characteristics (i.e., data in adjacent time gaps have strong local interactions). For sparse segments, local interactions in adjacent time gaps are captured by the cyclic layer of each segment respectively . Therefore, if $w = 1$, the local interactions will be ignored. Secondly, the input portion of the traffic flow sample should be divided into multiple segments ($w < W$) in order to facilitate the capture of temporal dynamics by the recurrent layer. If $w = W$, the temporal dynamics will be ignored.

3.2 Multi-graph construction

The key to improving the prediction performance of GCNs is the construction of the graphs. If the constructed network graphs do not effectively encode the correlations between individual grid regions, not only will they not be helpful for the learning of the model, but they are even likely to make the model's prediction performance poorer. In this section, this paper will show how to use multiple graphs to encode different types of correlations between regions and how to model these relationships using the proposed multi-graph convolution. Two types of correlations between regions are modeled using graphs, including: (1) encoding the adjacency of spatial proximity; (2) encoding the similarity between distant regions. Note that the approach in this paper can be easily extended to model new types of correlations by constructing correlation maps.

Neighborhood graph A : A graph consists of an exhaustive nonempty set of vertices and a set of edges between vertices, denoted by $G_r = (V, E)$, G represents the graph, V is the set of vertices, and E is the set of edges. A graph is undirected if the connections between vertices have no direction, otherwise it is directed. In this paper, we use the adjacency matrix to represent the connection relationship between vertices, which is denoted by the letters A , $W_r(i, j)$ denotes whether the lattice i is adjacent to lattice j . The adjacency matrix of the undirected graph can be expressed as:

$$A = \begin{pmatrix} 0 & W_r(1,2) & & W_r(1,N) \\ W_r(2,1) & 0 & & W_r(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ W_r(N,1) & W_r(N,2) & \dots & 0 \end{pmatrix} \tag{1}$$

Functional similarity graph X_f : Cities that share similar functions often have similar traffic patterns, for example, an industrial area may have heavy traffic flows during weekday rush hours while the downtown area may be congested on weekends. Functional similarity graph is denoted by $G_f = (V, E)$, where $X_f(i, j)$ the cosine distance between two grids: Firstly, previous studies have demonstrated that POI distribution can measure the functionality of urban areas. If an area l_i is given, then calculate the POI density from l_i around eight categories of POIs, i.e., dining, working, business, residential, science, education, healthcare, transportation, entertainment, scenic beauty. We compute the POI densities, which form a feature vector pl_i ,where $pl_i(j)$ represents the j class POI density surrounding gird l_i , which is calculated by Equation(2), In this equation, m_{ij} is the number of j class POIs surrounding grid l_i , m_i is the total number of POIs around l_i , m_j is the total number of POIs in the POI dataset. The formula refers to the TF-IDF algorithm in the field of natural language processing[17], which is designed to assign higher weights to POI categories with fewer overall number of categories. In addition, for given grid l_i and l_j ,this paper utilizes the cosine distance to measure the similarity of pl_i and pl_j , denoting $W_f(i, j)$. Then the adjacency matrix of X_f can be represented by Equation(3).

$$pl_i[j] = \frac{m_j^i}{m^i} \times \log \frac{M}{M_j} \tag{2}$$

$$X_f = \begin{pmatrix} 0 & W_f(1,2) & \dots & W_f(1,N) \\ W_f(2,1) & 0 & & W_f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ W_f(N,1) & W_f(N,2) & \dots & 0 \end{pmatrix} \tag{3}$$

4. EXPERIMENTS

4.1 Experimental dataset

HZJTD is a sampling of 202 roads in the main urban area of Hangzhou using a ring detector by the Hangzhou Comprehensive Traffic Research Center1. The large time period is from October 16, 2013 to October 3, 2014. HZJTD has traffic conditions including travel speed and traffic congestion index. Traffic speed was used as the traffic condition for prediction. We calculated the travel speed on the roadway every 15 minutes. Finally, HZJTD contains 30,353 records for each road. Based on Baidu map API, we collect POIs data of Hangzhou.

4.2 Evaluation indicators

Root Mean Square Error (RMSE): RMSE is a measure of variance that is used to compare the difference between a model's predicted and actual values. RMSE is usually greater than zero, with a value of zero indicating that the predicted values match the data exactly. Specifically, a larger RMSE is worse than a smaller RMSE. The formula for calculating RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \tilde{X}_i\right)^2} \tag{4}$$

Mean Absolute Error (MAE): MAE is the absolute value of the average error between the predicted value and the actual value, which is a number greater than or equal to 0. When the value is closer to 0, it means that the predictive performance of the model is better, and its calculation formula is calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \tilde{Y}_i\right| \tag{5}$$

4.3 Model parameters

To facilitate model training and validation, the dataset is divided into a training set and a test set in a ratio of 8:2. The hyperparameters of the model include learning rate, batch size, number of training iterations, and number of hidden neurons in the neural network unit. Among them, the learning rate determines the speed of weight update, the batch size is the number of training samples processed at one time during the training process, according to the experience and the adjustment in the repeated experiments, the learning rate is set to 0.001.

4.4 Traffic flow prediction results

In order to investigate the performance of the model on traffic flow prediction, this experiment is conducted. After setting appropriate parameters for the model, this paper inputs the traffic flow data in the time slices with fixed 20-minute intervals into the model, and then after 50 iterations of training on the model with this data, through the setting of hyper-parameters, we use the traffic flow data of the first W time slices for certain time periods with known traffic flow to predict the traffic flow for these time periods, so as to get the best prediction performance of the model. Now predict traffic flow for these time periods to get the prediction performance of the model. This experiment explores the case where the model has the best prediction performance by varying the value of W with the number of time slices predicted (i.e., predicting how long the demand will be in the future). The experimental results are shown in Table 1 and Table 2.

Table 1. Prediction results on RMSE.

historical time slice	**predicted time duration(min)**		
	20	40	60
36	2.17	3.77	7.20
72	2.18	3.99	6.67
144	2.03	3.68	6.56

Table 2. Prediction results on MAE.

historical time slice	predicted time duration(min)		
	20	40	60
36	1.61	2.87	5.54
72	1.60	3.03	4.96
144	1.48	2.70	4.97

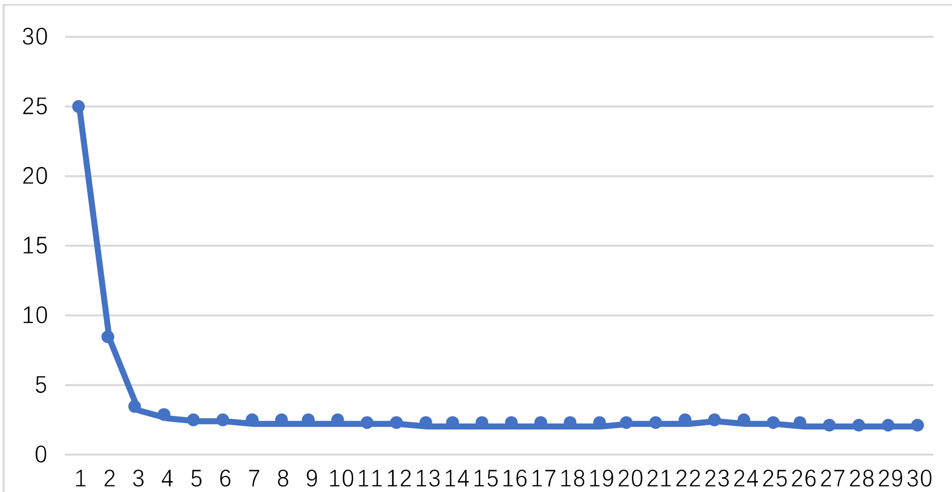

Figure 2. Training plot for 30 iterations.

According to Figure 2, the iterative graph of model training using 144 time slices to predict the demand 20 minutes in the future, it can be observed that the evaluation metric RMSE tends to stabilize and converge at around 2.0 when the model is trained for 27 times. Therefore, to ensure the performance of model prediction in the future, the number of iterations of model training is set to 50 in this paper.

As shown in Table 1 and Table 2, the evaluation metrics of the model's prediction results for the next 20, 40, and 60 minutes using 36, 72, and 144 time slices appear to be decreasing overall. This is due to the fact that the 144 time slices cover 2 days of data, and the model can well grasp the temporal correlation embedded in the data. At the same time it uses the neighbor graph constructed from the input data and the POI functional similarity relationship graph to merge with it, so that the model can make good use of the temporal, spatial and semantic correlation in the prediction process, thus obtaining a better prediction result. Comparatively speaking, the model utilizes 36 and 72 time slices for the prediction process to grasp the temporal correlation not well enough, so the results are slightly worse. At the same time, it can be observed that as the predicted time grows, the evaluation metrics increase, i.e., the worse the prediction performance, the proposed model has a better performance in predicting the traffic flow in the short future time.

5. CONCLUSION

This paper has addressed the challenging task of traffic flow prediction within road networks. We have introduced a novel deep learning framework, T-MGCN, designed to capture the intricate spatio-temporal and semantic correlations present in traffic flow data. T-MGCN leverages multiple graphs to encode non-Euclidean spatial relationships and potential semantic associations among roads. These correlations are explicitly modeled through the fusion of multiple graph convolutional networks.

Through extensive experimentation utilizing real-world traffic datasets, we have demonstrated the effectiveness and feasibility of the T-MGCN approach in enhancing traffic flow prediction accuracy and efficiency. This work contributes to advancing our understanding of traffic prediction in complex road networks and highlights the potential of deep learning techniques in addressing these challenges.

REFERENCES

[1] Do L., Vu., Vo B., et al. An effective spatial-temporal attention based neural network for traffic flow prediction[J]. Transportation research part C: emerging technologies, 2019, 108: 12-28.

[2] Guo Z., Zhang Y., Lv J., et al. An online learning collaborative method for traffic forecasting and routing optimization[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(10): 6634-6645.

[3] Wu Y., Zuo Z., Wang Y., et al. Distributed Data-Driven Model Predictive Control for Heterogeneous Vehicular Platoon with Uncertain Dynamics[J]. IEEE Transactions on Vehicular Technology, 2023,72(8):9969-9983.

[4] El E., Daily bicycle traffic volume estimation: Comparison of historical average and count models[J]. Journal of Urban Planning and Development, 2018, 144(2): 04018011.

[5] Achar A., Bharathi D., Kumar B. A., et al. Bus arrival time prediction: A spatial Kalman filter approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(3): 1298-1307.

[6] Siripanpornchana C., Panichpapiboon S., Chaovalit P., Travel-time prediction with deep learning[C]//2016 ieee region 10 conference (tencon). IEEE, 2016: 1859-1862.

[7] Huang W., Jia W., Guo J., et al. Real-time prediction of seasonal heteroscedasticity in vehicular traffic flow series[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 19(10): 3170-3180.

[8] Sun B., Sun T., Zhang Y., et al. Urban traffic flow online prediction based on multi-component attention mechanism[J]. IET intelligent transport systems, 2020, 14(10): 1249-1258.

[9] Kwak S., Geroliminis N., Travel time prediction for congested freeways with a dynamic linear model[J]. IEEE transactions on intelligent transportation systems, 2020, 22(12): 7667-7677.

[10] Zechin D., Cybis H., Probabilistic traffic breakdown forecasting through Bayesian approximation using variational LSTMs[J]. Transportmetrica B: Transport Dynamics, 2023, 11(1): 1026-1044.

[11] Agafonov A. A., Yumaganov A. S., Myasnikov V., Big data analysis in a geoinformatic problem of short-term traffic flow forecasting based on ak nearest neighbors method[J]. Computer Optics, 2018, 42(6): 1101-1111.

[12] Defferrard M., Bresson X., Vandergheynst P., Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering[C]. Advances in neural information processing systems, 2016, 29.

[13] Zhao L., Song Y., Zhang C., et al. Temporal Graph Convolutional Network for Urban Traffic Flow Prediction Method[J]. IEEE transactions on intelligent transportation systems, 2019, 21(9): 3848-3858.

[14] Zhang Z., Li M., Lin X., et al. Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies[J]. Transportation Research Part C: Emerging Technologies, 2019, 105(AUG.):297-322.

[15] Sun J., Zhang J., Li Q., et al. Predicting Citywide Crowd Flows in Irregular Regions Using Multi-View Graph Convolutional Networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, PP(99):1-1.

[16] Ke J., Zheng H., Yang H., et al. Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach[J]. Transportation research part C: Emerging technologies, 2017, 85: 591-608.

[17] Qaiser S., Ali R., Text mining: use of TF-IDF to examine the relevance of words to documents[J]. International Journal of Computer Applications, 2018, 181(1): 25-29.