Hindawi
Journal of Advanced Transportation
Volume 2023, Article ID 2941035, 16 pages
https://doi.org/10.1155/2023/2941035

Research Article

CPT-DF: Congestion Prediction on Toll-Gates Using Deep Learning and Fuzzy Evaluation for Freeway Network in China

Tongtong Shi,[1,2] Ping Wang,[1,2] Xudong Qi,[2] Jiacheng Yang,[3] Rui He,[2] Jingwen Yang,[2] and Yu Han[1,4]

[1]*School of Intelligent System Engineering, Sun Yat-Sen University, Guangzhou 518000, China*
[2]*School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China*
[3]*School of Automation, Southeast University, Nanjing 210096, China*
[4]*Guangdong Province Key Laboratory of Fire Science and Technology, Guangzhou 510006, China*

Correspondence should be addressed to Ping Wang; wang0372@e.ntu.edu.sg and Yu Han; hanyu25@mail.sysu.edu.cn

Received 26 April 2022; Accepted 3 September 2022; Published 10 April 2023

Academic Editor: Dong-Kyu Kim

Toll-gates are crucial points of management and key congestion bottleneck for the freeway. In order to avoid traffic deterioration and alleviate traffic congestion in advance, it is necessary to predict and evaluate the congestion in toll-gates scattering in large-scale region of freeway network. In this paper, traffic volume and operational delay time are selected from various traffic indicators to evaluate congestion considering the particular characteristics of the traffic flow within the toll-gate area. The congestion prediction method is designed including two modules: a deep learning (DL) prediction and a fuzzy evaluation. We propose a modified deep learning method based on graph convolutional network (GCN) structure in the fusion of dilated causal mechanism and optimize the method for spatial feature extraction by constructing a new adjacency matrix. This new AI network could process spatiotemporal information of traffic volume and operational delay time, that extracted from large-scaled toll-gates spontaneously, and predict key indicators in 15/30/60 min future time. The evaluation module is proposed based on these predicted results. Then, fuzzy C-means algorithm (FCM) is further modified by determining coupling weight for these two key indicators to detect congestion state. Original traffic data are obtained from the current 186 toll-gates served on the freeway network in Shaanxi Province, China. Experimental tests are carried out based on historical data of four months after preprogressing. The comparative tests show the proposed CPT-DF (congestion prediction on toll-gates using deep learning and fuzzy evaluation) outperforms the current-used other models by 6-15%. The successful prediction could extend to the real-time prediction and early warning of traffic congestion in the toll system to improve the intelligent level of traffic emergency management and guidance on the key road of disasters to some extent.

1. Introduction

Freeway is the backbone of long-distance transport because of its low disruption and excellent road conditions [1]. As the demand for long-distance transport increases with economic development, the number of freeway mileage in China is growing exponentially, as shown in Figure 1(a). At present, researchers have made a large quantity of work on traffic state of urban road sections, but there are relatively few studies on toll-gate of freeway [2]. Since the freeway is closed management and toll-gates scattering in large-scale region of freeway network, characteristics of the traffic flow within the toll-gate area and other roads are different. There are multiple steps such as deceleration, lane changing, rendezvous, and toll payment when a vehicle enters the toll-gate area [3]. The entire traffic efficiency is strongly affected by passing time through the toll-gates. Thus, toll-gates are crucial points of management and key congestion bottleneck for the freeway network, especially in China. However, serious traffic congestion often occurs in the toll-gates area. As shown in Figure 1(b), in the top 10 toll-gates ranked by congestion in China, the vehicle speed may drop to 10 km/h and the congestion index may even be up to 50-60. It might additionally cause traffic accidents, energy

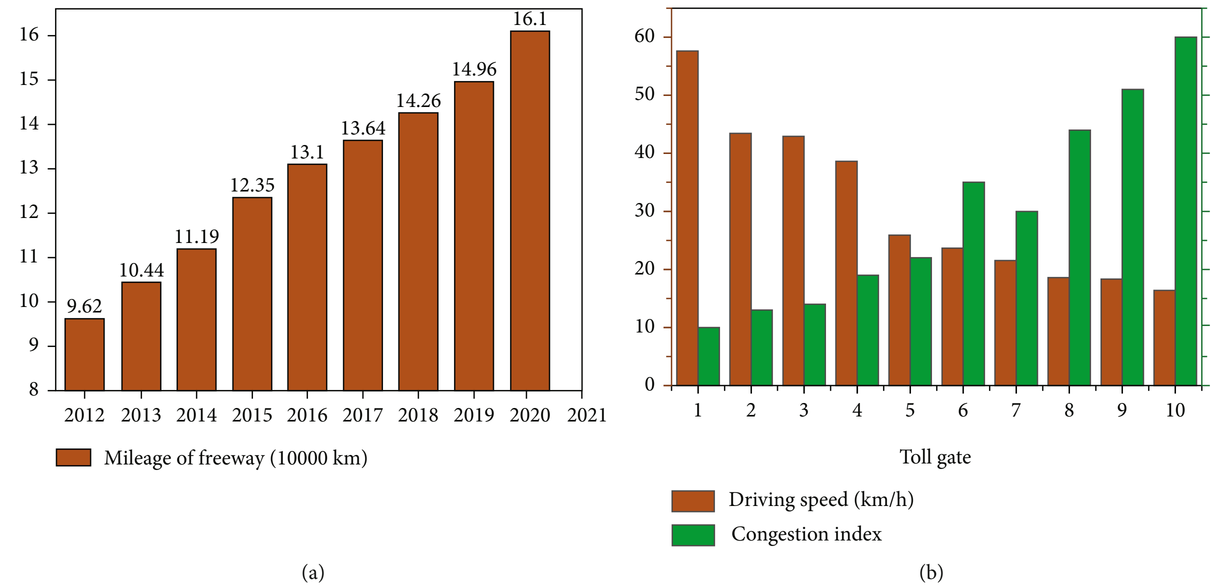


Figure 1: (a) Mileage of freeways in China. (b) Top 10 toll-gates ranked by congestion index (11 : 10 April 9, 2022). (Data source from reference: https://report.amap.com/congest.do#).

consumption, and environmental pollution in the areas of these toll-gates [4]. In order to alleviate or avoid the occurrence of these problems, an increasing number of researchers are working on two aspects: traffic prediction and congestion evaluation of the toll-gates area, in which further studies on quantitively predicted traffic indicators might be an alternative way to solve the problem to evaluate the congestion. The success prediction and evaluation would provide the possibility for the advance management of the toll-gate and the guidance of traffic routes. The traffic perception will be faster and more accurate with the development and breakthrough of artificial intelligence (AI) and deep learning (DL), providing more effective intelligent technology for alleviating congestion and traffic emergency management.

In terms of traffic prediction, the freeway network is a topological structure with dynamic spatiotemporal features, which is manifested by the periodicity of traffic flow and the spatial correlation of toll-gates between upstream and downstream. Initially, due to the limitation of intelligent technology, most of the prediction work of traffic congestion focused on the temporal dimension while ignoring the information of the large-scale spatial dimension. In recent years, graph neural network (GNN) based on deep learning was first proposed by Gori et al. [5] and Scarcella et al. [6], which introduced graph structure in the field of spatial correlation to skillfully simulate the spatial correlation between objects. Hence, various algorithms based on graph models have been widely used in different fields, including social network, biomedicine, and knowledge graph [7]. Similarly, the toll-gates and ordinary road sections of the freeway network can be mapped to the relationship between points and edges in the graph structure. Recently, as a branch of GNN, Graph Convolutional Networks (GCN) [8, 9] were introduced to traffic work and efficiently implement congestion prediction from a spatiotemporal perspective. However, there are also unsatisfactory results in long-term prediction work due to the defects of the model. For example, the disappearance and explosion of gradients in Recurrent Neural Networks (RNNs) lead to the loss of historical data and cannot be processed in parallel on a large scale, resulting in slow computation. Therefore, a graph convolutional network fusing the dilated causal mechanism was introduced in this paper to compensate for this deficiency. In addition, most existing graph-based methods usually only build spatial features based on distances or correlations between toll-gates [10–12]. But there will be two situations as shown in Figure 2: (1) Nearby Euclidean distances between gates, but little spatial correlation of traffic flows (Region A: A_1 and A_2 are not directly relatable), (2) Long Euclidean distances between gates, but strong spatial correlation of traffic flows (Region B). This means that only referring to a single factor cannot accurately grasp the large-scale spatial features and even has a great interference on the prediction accuracy. Based on this, a new spatial adjacency matrix that combines the two features of toll-gates including correlation and distance is proposed in this paper, which can more accurately explain the spatial characteristics of toll-gates in a large scale.

In terms of congestion evaluation, the work of congestion evaluation of freeway toll-gates is to further analyze and summarize the prediction results (single or multiple traffic indicators). The content of this part includes the selection of indicators and evaluation methods to measure traffic congestion. Firstly, researchers mostly use traffic indicators such as speed, traffic volume, occupancy, operational delay time, and queuing length as the detection criteria for urban traffic congestion. The freeway is a closed system consisting of “toll entrances-road sections-toll exits”. Under normal circumstances, the vehicle will drive at a high and uniform speed on ordinary road sections. When reaching the toll-gate area, the vehicle speed will

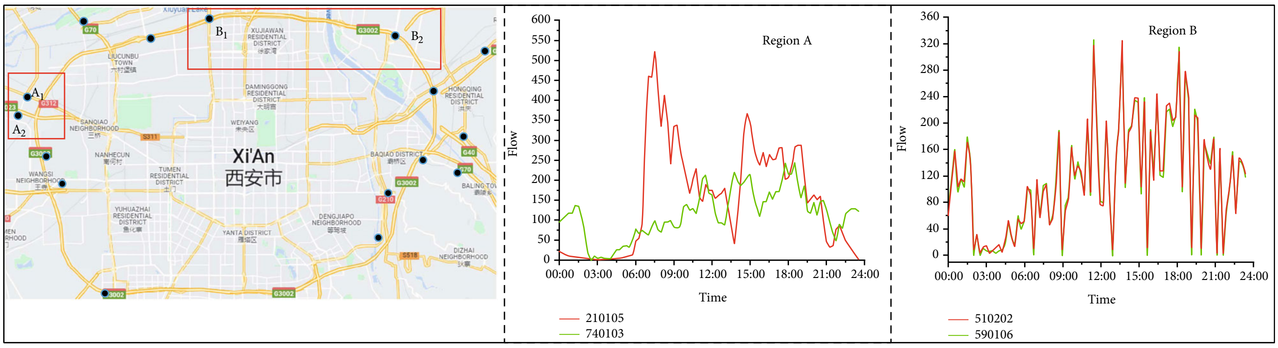


Figure 2: The special situation of toll-gate location and traffic flow distribution.

slow down to an extremely low state until the toll is completed and finally accelerate the process of driving away. Therefore, the speed itself varies greatly in the toll-gate area, so it is suitable as a congestion indicator in the ordinary road section rather than in the toll-gate area. In addition, since there are often nondeployed detector sections or detector damaged sections upstream and downstream of the toll-gate, the queuing length and occupancy rate cannot be completely and directly detected. In view of this, the traffic volume based on toll-data is preferred in this paper as one of the indicators to identify the congestion state of freeway toll-gates. Besides, due to the difference in the scale of toll-gates, the evaluation results of toll-gate i and toll-gate j may be different under the same traffic volume. Therefore, this paper selects operational delay time as another indicator and designs a calculation method based on toll-data. Secondly, there is no standard congestion evaluation method because the Electronic Toll Collection (ETC) channel has been added to the toll-gates of China's freeways in recent years, which is an important renovation of the toll-gates [13]. References include the "Highway Capacity Manual (HCM)" issued by United States and the "Urban Road Traffic Operation Evaluation Indicator System" issued by China, both of which quantify the service level into n intervals according to road characteristics. Therefore, the idea of clustering and the currently more popular methods based on fuzzy clustering are introduced in this paper to perform traffic state classification well. Currently, there are some studies in this area. A congestion evaluation model based on the fuzzy C-means (FCM) algorithm was proposed [14]. This method of unsupervised learning based on fuzzy clustering to analyze hidden data patterns is helpful for the congestion evaluation of toll-gates. However, the traditional FCM algorithm does not consider the influence of each traffic index on the clustering results, and the algorithm is prone to fall into the local minimum. Based on these problems, this paper refers to and improves the FCM method, which combines the coupling weights of the two traffic indicators to optimize the FCM algorithm and accurately realize the congestion evaluation. The main contributions can be summarized as follows.

(i) Based on toll-data, a new calculation method for "traffic volume" and "operational delay time" is proposed in this paper to complete the congestion prediction and evaluation of toll-gates more accurately

(ii) A new AI network (CPT-DF) for congestion prediction of freeway toll-gates using deep learning and fuzzy evaluation is proposed. The GCN of the prediction module is used to capture the spatial information of the road network by weighting the neighbourhood node features and embedding them into the graph structure, and the improved dilated causal mechanism preserves the nonlinear ability of the model by adding residual connections and gated linear units, so that temporal dimension information can be better captured. The evaluation module uses the improved FCM algorithm for accurate interval classification of toll-gate traffic state

(iii) A new adjacency matrix that combines both the correlation and distance features of toll-gates is constructed in this paper, which can optimize the existing methods to more accurately extract the spatial features of large-scale toll-gates

(iv) This paper verifies the proposed CPT-DF model based on toll data, and some toll-gates are selected to complete the work of congestion prediction, which could efficiently improve the intelligent level of traffic emergency management and guidance on the key road of disasters

2. Related Work

2.1. Prediction Algorithms. In recent years, with the increase in the amount of data and the development of fusion technology, data-driven prediction algorithms represented by statistical models, traditional machine learning and deep learning models have become more and more popular in various research fields. In the field of traffic prediction, ARIMA model [15] and Kalman filter model [16] can rely on statistical methods to simply model the relationship between traffic parameters to predict road traffic state. However, such models are usually based on linear assumptions, which cannot strongly explain the high-dimensional and nonlinear of traffic data. Later, machine learning methods such as Support Vector Regression (SVR) [17] proposed by researchers can solve the nonlinear relationship in traffic data well, so such methods are widely used in freeways and urban roads. However, the changes in road environment and traffic flow near toll-gates are more complicated than the temporal characteristics of ordinary road sections, and there are few studies on toll-gates at present. Wang et al. [18] fuse vehicle detector data, long-range microwave sensor data, and toll-data and employ Deep Belief Network (DBN) to successfully predict small-scale ring roads at time intervals of 30/60/120-minute traffic flow at the toll-gate. Shuai et al. [19] adopted the modified Long Short-Term Memory (LSTM) and predicted the traffic volume of the 51 screened toll-gate. These deep learning-based methods are good at capturing traffic trends in complex toll-gate environments or exploring spatial connectivity between one or more road segments in a single temporal dimension, but the spatial characteristics between each toll-gate on a large scale are not considered. Therefore, the researchers extracted the spatial features of the road network by introducing the convolutional neural network (CNN) [20] to convert the structure of the traffic road network into a standard graph structure, but CNN is not suitable for processing non-European data such as toll-gates. Recently, GCNs have been widely used for many graph-based tasks, and many studies have further explored the use of GCNs to model the topology of road networks. The STGCN model [10] combines GCN and CNN for the first time to model the traffic network and spatiotemporal sequence. This paper also refers to the fusion principle of STGCN in the model construction part, but the model can only use CNN to process the signal of each layer of the network. Propagating to the upper layer, the processing of samples is independent at each moment, so it cannot cope with long-term prediction well; the DCRNN model [11] models the spatial correlation as a diffusion process on a directed graph to establish a traffic flow transformation model, we propose to develop a diffuse convolutional recurrent neural network capable of capturing the spatial and temporal dependencies between long-term sequences using the seq2seq framework. This paper is inspired by DCRNN in the construction of time series model. The ASTGCN model [21] introduces an attention mechanism (GAT) in GCN to effectively model temporal and spatial correlations. However, this model cannot capture spatial and temporal dependencies simultaneously and only considers low-order neighbourhood relationships between nodes, ignoring the correlations between different historical periods. The T-GCN model [12] integrates GCN and GRU to capture traffic spatiotemporal features, which can be well derived from spatiotemporal features. These models can well capture the information between adjacent nodes from the perspective of spatial and temporal, so as to complete short-term traffic prediction. However, these models have not been successfully applied in the traffic prediction work of the toll station, so this paper proposes a new model and uses the toll-data to carry out the application work of the actual scene.

2.2. Evaluation Algorithms. The methods of congestion evaluation can be divided into traffic theory and data-driven algorithms [22]. The former is based on physical and mathematical theory to describe the characteristics of traffic behaviour, and then evaluate the traffic congestion, but it is not suitable for scenarios with high complexity and unknown degree, such as toll-gates. The latter is based on neural networks and clustering algorithms to evaluate traffic congestion [23]. Neural network-based methods need to manually classify traffic states in advance, but it is difficult to accurately classify states in unknown scenarios. In the contrast, clustering methods are very suitable for special scenarios of toll-gates without standard evaluation. The clustering algorithm is the process of dividing a collection of research objects into multiple classes consisting of similar objects. Initially, the K-means clustering algorithm [24] combined with three parameters of traffic volume, speed, and occupancy was applied to achieve a simple state evaluation. However, the algorithm cannot evaluate the critical threshold of the sample. Later, considering that traffic state is a fuzzy concept, the algorithm of fuzzy clustering is also applied to traffic state evaluation. The FCM algorithm was first proposed by Dunn [25], and later the improved algorithm based on FCM has been widely used in the field of traffic evaluation [26, 27].

3. Methodology

3.1. Problem Definition

3.1.1. Congestion Domain of Toll-Gates. To study the traffic congestion of the toll-gates, this paper constructs a "toll-gate congestion domain" according to the traffic characteristics of the toll-gates, which includes the three areas (upstream section, deceleration section, and toll section), as shown in Figure 3. After the vehicle passes through the congestion domain, it will quickly resume normal driving through the acceleration section and the downstream section.

Upstream Section. A fixed distance section before the vehicle enters the toll-gate. The first vehicle detector (VD) in Figure 3(1) is installed upstream of the toll-gate in upstream section, which is used to detect parameters such as the speed of passing vehicles and the current time.

Deceleration Section. The vehicle starts to slow down and enters the toll-gate and selects a different toll lane.

Toll Section. The toll lanes can be divided into Manual Toll Collection (MTC) and Electronic Toll Collection (ETC) based on the tolling method. The vehicle decelerates through the railing locomotive detector (RLD) as shown in Figure 3(2) and records the current time.

3.1.2. Calculation of Indicator. Firstly, the original data collected in this paper will be counted every 5 minutes and traffic volume can be counted directly (the details are presented in Section 4.1.1). Secondly, the operational delay time on ordinary roads can be obtained according to "speed-acceleration-distance". But the limitation of this method will bring a large error. According to the characteristics of toll-gates, the operational delay time is calculated as

$$D = \frac{1}{n}\sum_{i=1}^{n}\left(T_i - \frac{1}{K}\left(\sum_{i=1}^{k} T_{Pk}\right)\right), \tag{1}$$

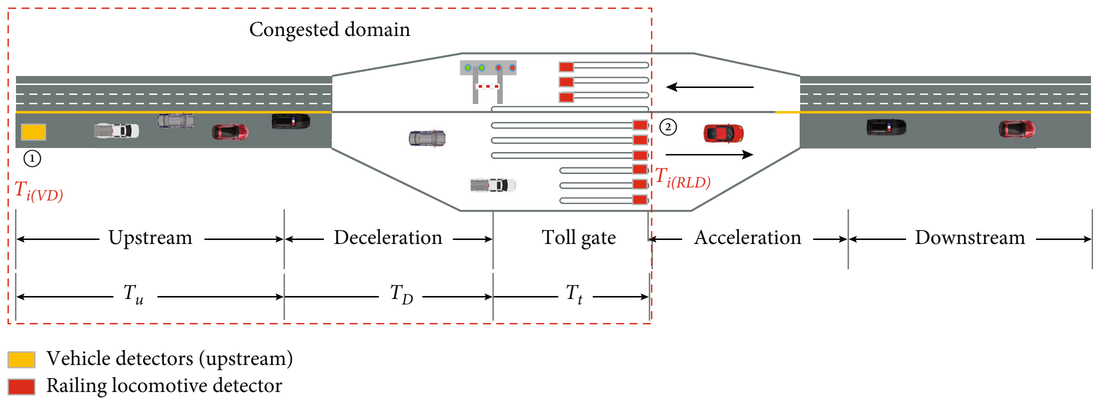


Figure 3: Composition of toll-gates and construction of congestion domain.

$$T_i = T_u + T_D + T_t = \left(T_{i(\mathrm{RLD})} - T_{i(\mathrm{VD})}\right), \quad (2)$$

$$T_{Pk} = \frac{1}{N}\sum_{j=1,\mathrm{volume}<Q}^{N}\left(T_{jk(\mathrm{RLD})} - T_{jk(\mathrm{VD})}\right), \quad (3)$$

where D is the calculated average operational delay time per five minutes. T_i is the passing time of the i-th vehicle passing through the congestion domain. T_u, T_D, and T_t are the upstream section and the deceleration section as shown in Figure 3, respectively and the travel time of the toll road. T_{Pk} is the travel time of the vehicle through the toll channel K in the congestion domain. Q is the flow threshold when the traffic state is unblocked. $T_{i(\mathrm{VD})}$ and $T_{i(\mathrm{RLD})}$ are the times detected by the i-th vehicle passing the upstream vehicle detector and the railing locomotive detector, respectively.

3.1.3. Preparing for Input/Output. As shown in Table 1, two sets of inputs/outputs are designed to perform both prediction and evaluation. For the prediction module, the input values are the two previously selected feature data (traffic volume and operational delay time) and graph data. The graph data represents the spatial relationship between each toll-gate. The two traffic indicators represent the temporal relationship between the historical traffic state and future traffic state of toll-gate. The output values are just two traffic indicators for future moments predicted by the model. For the evaluation module, the input value is the output value of the prediction module and the congestion threshold corresponding to each toll-gate, and the output value is the final traffic state. This part mainly introduces the construction of the spatiotemporal data input to the prediction module.

(1) A. Description of Feature Data. Traffic prediction is a typical spatiotemporal prediction problem. Given the previous M observations of historical traffic feature, the data measured at the N toll-gates at time step H can be viewed as a matrix of size $M \times N$. Then, the predicted value of the flow closest to the true value in the next H time steps is as

$$F_{t+1}, \cdots, F_{t+H} = \mathrm{argmax}\log P(F_{t+1}, \cdots F_{t+H} | F_{t-M+1}, \cdots, F_t), \quad (4)$$

where $F_t \in R^n$ is a vector of observations for n toll-gates at time step t, F_t here also refers to the traffic volume and operational delay time.

(2) Construction of Graph Data. For unordered gate network traffic data, the observations F_t are not independent and can be viewed as graph signals defined on an undirected graph G as shown in Figure 4, the graph is expressed in terms of $G_t = (F_t, E, W)$. E is a set of edges representing the connections between gates, and $W \in R^{n\times n}$represents the adjacency matrix of G_t.

Each of the toll-gates can be regarded as a vertex in the graph structure, and the road segments connecting the toll-gates can be regarded as edges. In order to represent the spatial relationship between each toll-gate (vertex), the Euclidean Distance is usually chosen, but there will be the defects shown in Figure 2 that were mentioned earlier. Therefore, we introduce Distance Matrix (D-Matrix) and Correlation Matrix (C-Matrix) to represent spatial features.

(i) D-Matrix is the Euclidean distance between each gate, which can be calculated using the latitude and longitude values of the gates by "Vincenty solutions of geodesics on the ellipsoid" [28].

(ii) C-Matrix is judged based on whether the gates are directly connected. As shown in Figure 5, $P_1$and P_4 are directly connected, but P_1 and P_2 are not directly connected, so $C_{P_1P_4} = 1$, $C_{P_1P_4} = 0$.

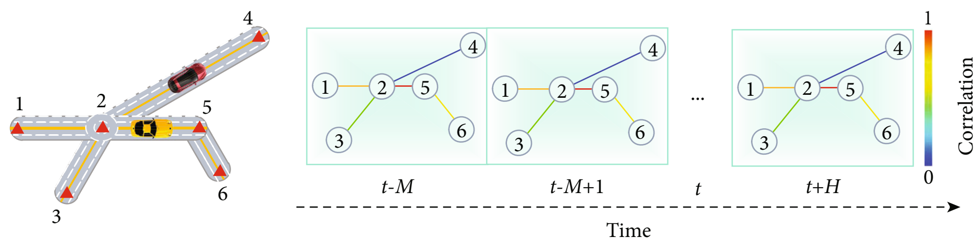
Finally, a novel type of Distance and Correlation Matrix (D&C-Matrix) is constructed to calculate the adjacency matrix $\boldsymbol{W}_{\boldsymbol{ij}}$ as

$$\boldsymbol{w}_{\boldsymbol{ij}} = \begin{cases} \exp\left(-\dfrac{\left([\boldsymbol{D}_{\boldsymbol{ij}}] \odot [\boldsymbol{C}_{\boldsymbol{ij}}]\right)^2}{\sigma^2}\right) & \geq \varepsilon, i \neq j, \\ 1, \text{others} & \end{cases} \quad (5)$$

where $[\boldsymbol{D}_{\boldsymbol{ij}}]$ and $[\boldsymbol{C}_{\boldsymbol{ij}}]$ are D-Matrix and C-Matrix, respectively, $\odot$ is the Hadamard product, and $\boldsymbol{D}_{\boldsymbol{ij}}$ and $\boldsymbol{C}_{\boldsymbol{ij}}$ are the

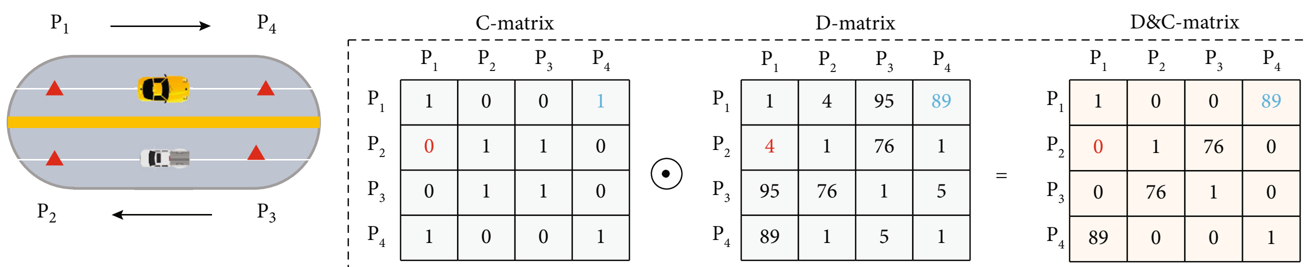
Table 1: The input/output values of the proposed model.

	Prediction module			Evaluation module	
Input values	Feature data (historical)		Graph data	Feature data	Threshold
	Traffic volume (number/n min)	Operational delay time (s)	D&C-matrix	Traffic volume (number/n min) (future)	Congestion boundaries
Output values	Traffic volume (number/n min) (future)			Traffic state	



(a) (b)

Figure 4: (a) Toll-gates in a road network. (b) Spatiotemporal correlation.



(a) (b)

Figure 5: Construction of adjacency matrix. (a) Location distribution of the four toll-gates (Yellow lines represent dividing fences that vehicles on the freeway cannot pass through). (b) The process of building an adjacency matrix.

distance and correlation between gates i and j. $\sigma^2$and ε is the threshold of control matrix distribution and sparsity.

3.2. Overview. In this paper, we propose an AI network (CPT-DF) of deep learning that integrates a fine-grained congestion evaluation mechanism, as shown in Figure 6. The CPT-DF network includes two modules: prediction module and evaluation module. The prediction module includes input/output layer 1 and spatiotemporal convolution layer, and the evaluation module includes input/output layer 2 and congestion evaluation layer. The output layer 1 and the input layer 2 are marked with green fonts, just because of the transmission relationship in the calculation process.

First, the preliminary work completed the construction of the congestion domain of the toll-gate, the selection of indicators, and the data required by the input layer. Then, the prediction module detects traffic indicators (traffic volume and operational delay time) for future periods based on the spatiotemporal convolutional layers constructed by GCN fused with dilated causal convolutions. Finally, the evaluation module combines the prediction indicators with the FCM clustering mechanism to realize the congestion detection of the toll-gate in the future period.

3.3. Prediction Module

3.3.1. Spatial Feature Extraction. GCN is a basic operation based on spectral decomposition method or spatial structure. The spectral decomposition-based method is to deal with the spectral domain correlation representation of the graph. In this paper, the spectral decomposition method is introduced to extract node spatial features given node information. As early as 2014, Bruna et al. [29] proposed Spectral Network to define convolution operations in the Fourier domain, which can be defined as the product of feature $\mathbf{x} \in \mathbb{R}^N$ and a convolution kernel $\mathcal{G}_\theta = \text{diag}\,(\theta)$ as

$$\mathbf{x} * \mathcal{G}_\theta = \boldsymbol{U}\mathcal{G}_\theta(\boldsymbol{\Lambda})\boldsymbol{U}^T\mathbf{x}, \tag{6}$$

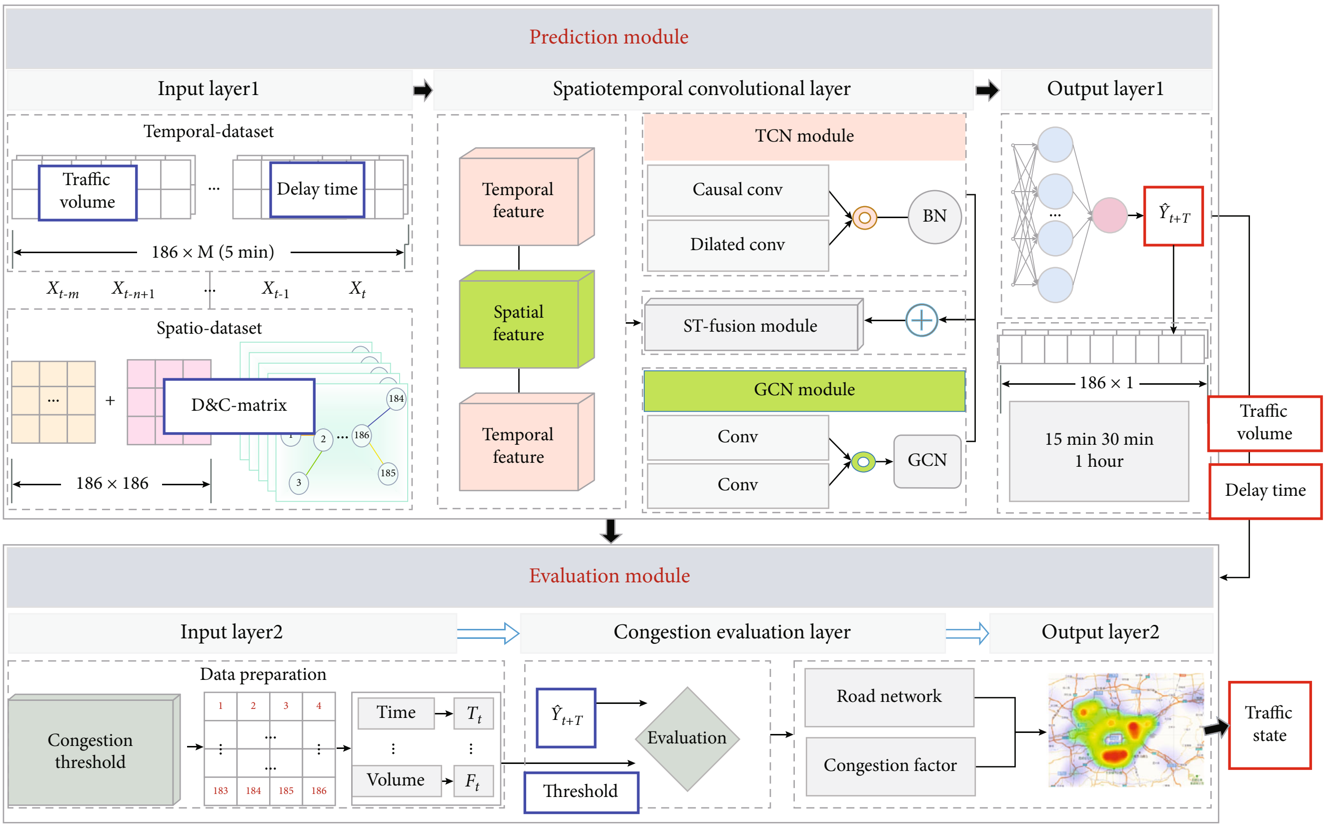


Figure 6: CPT-DF network: congestion prediction on toll-gates using deep learning and fuzzy evaluation. The blue line boxes represent the input values of the prediction module and the evaluation module, respectively, the red line boxes represent the output values, and the output value of the prediction module is also one of the input values of the evaluation module.

where $\boldsymbol{U}$ is a matrix composed of the eigenvectors of the normalized Laplacian matrix, and $\boldsymbol{\Lambda}$ is a diagonal matrix of traffic indicators. However, this method of convolution, which computes the Eigen decomposition of the Laplacian matrix of the graph, leads to potentially intensive computations and results in unsatisfactory locality of the convolution kernels.

In order to alleviate the problems of Spectral Network, in 2016, Michael et al. [30] proposed Cheb Nets K-jump convolution to define convolution on the graph, thus eliminating the time cost of calculating Laplacian matrix vectors. On this basis, this paper sets $K = 1$ to alleviate the local overfitting problem. Therefore, the graph convolution can be written as

$$\begin{aligned} \mathbf{x} * \mathcal{G}_{\theta'} \& &= \theta' \mathbf{x} - \theta' \left(\boldsymbol{D}^{-(1/2)} \boldsymbol{W} \boldsymbol{D}^{-(1/2)} \right) \mathbf{x} \\ &= \theta \left(\boldsymbol{I}_n + \boldsymbol{D}^{-(1/2)} \boldsymbol{W} \boldsymbol{D}^{-(1/2)} \right) \mathbf{x}, \end{aligned} \tag{7}$$

where the adjustable parameter is $\theta = \theta' = -\theta'$. $\boldsymbol{D}$ is the degree matrix. The GCN model constructs filters in the Fourier domain, constructs spatial features by stacking multiple local Covn layers, and extracts the structural information of the network in the form of convolutions. Therefore, a deeper structure can be constructed to deeply recover spatial information, which achieves a larger receptive field and reduces the number of convolution kernels.

3.3.2. Temporal Feature Extraction. As a derivative of CNN, Temporal Convolutional Networks (TCN) [31] is a network framework that can accurately process sequences or data containing time series. It aims to extract features across time steps by directly exploiting the powerful properties of convolutions and uses fully connected networks and dilated causal convolution to achieve corresponding outputs for each input, respectively, and ensure that no historical data is missed. In this paper, an improved TCN network is designed to extract temporal features by fusing dilation convolution, GLU, and residual blocks. The specific improved TCN structure is shown in Figure 7.

(1) Dilated Causal Convolution. Dilated causal convolution is used to solve the problem of the time dimension of big data. Among them, the expansion coefficient of the convolution kernel can be arbitrarily combined from the range of [1, 2, 4, 8, 16, 32]. Through comparative experiments, it is found that the experimental results obtained by [1, 2, 4], [1, 2, 4, 8, 16], and [8, 16, 32] are relatively stable. At the same time, in order to maintain the temporal relationship of historical information, the kernel is set to 2, and the

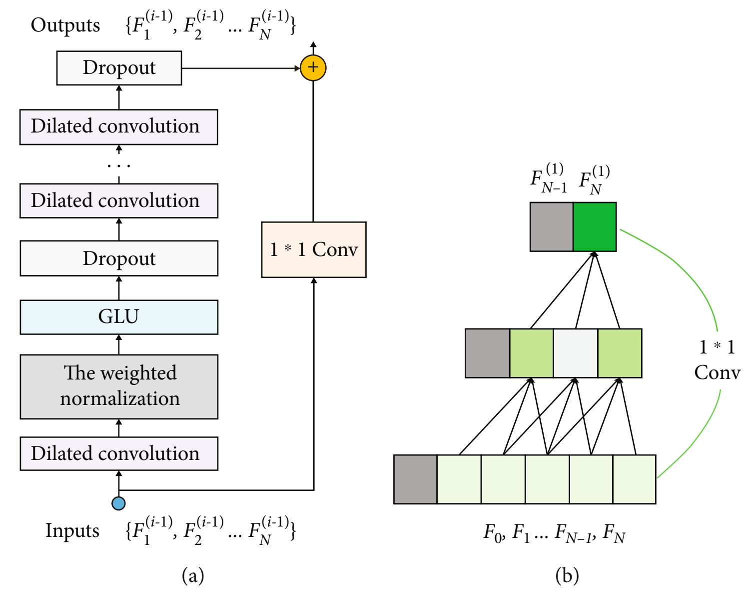


FIGURE 7: Structure of improved temporal feature extraction. (a) The composition structure of the overall framework. (b) Structure of Dilated Causal Convolutions.

expansion coefficient is used as the sliding jump value, and the receptive field is set to $2 \times 2^\wedge(4-1) = 16$.

Formally, for the one-dimensional sequence input $x \in \mathbb{R}^M$ and the kernel function $\phi \in \mathbb{R}^k$, d is the expansion coefficient, the flow data sequence input by time is $\{F_0, F_1, \cdots F_N\}$, and the output result is denoted as $\{g_0, g_1, \cdots g_N\}$, the mapping relationship S between F and g can be expressed as:

$$\widehat{g}_0, \widehat{g}_1, \cdots \widehat{g}_N = S(F_0, F_1, \cdots F_N). \tag{8}$$

The convolution operation S on the element F is

$$S(F) = (x * d\phi)(F) = \sum_{i=0}^{k-1} \phi(i) \cdot x_{F-d \cdot i}, \tag{9}$$

where k represents the kernel size, $F - d \cdot i$ maps the upper layer history information, and at the same time introduces the residual block in the TCN.

(2) Gated Linear Units (GLU). After adding the residual module, the TCN has 3 layers of dilated convolution, and the data distribution is normalized by weights, and then the GLU is used to replace the ReLU in the original structure to save the nonlinearity of the remaining blocks, at the same time to expand the volume every time add dropout after the product to prevent overfitting. Furthermore, $1 * 1$ Conv with a width of K_t is introduced to obtain the output Y of the time sequence as the input of the next stage. At this time, the time convolution input of each node can be regarded as a sequence of length M, and the number of channels is C_i, so $Y \in \mathbb{R}^{M \times C_i}$. The convolution kernel $\tau \in \mathbb{R}^{K_t \times C_t \times 2C_0}$ is used to map the input Y to a single output element $[P\,Q] \in \mathbb{R}^{(M-K_t+1) \times 2C_0}$. $1 * 1$ Conv keeps the remaining input and output dimensions the same, the convolution can be defined as:

$$\Gamma * \tau Y = P \odot \sigma(Q) \in \mathbb{R}^{(M-K_t+1) \times C_0}, \tag{10}$$

where P and Q are the input of the GLU gate, respectively, and $\odot$ represents the element-wise Hadamard product. $\sigma(Q)$ controls the dynamic change of the input P, and the added nonlinear link ensures the stacked input of the time layer, and the residual connection is realized in the time layer of the stack. Using the same convolution kernel Γ for each node $y_i \in \mathbb{R}^{M \times C_i}$ in the traffic graph, the time domain convolution $\Gamma * \tau y$ can be extended to the three-dimensional variable $y \in \mathbb{R}^{M \times n \times C_i}$.

(3) ST-Fusion Module. To fuse spatiotemporal features, a spatiotemporal fusion module (ST-Fusion Module) inspired by [32] is constructed in this paper. The modules can be stacked or expanded depending on the size and complexity of a particular case. As shown in Figure 8, each spatial convolutional layer bridges two temporal convolutional layers, which can achieve fast transition of the states of the temporal and spatial layers. In addition, this design scales the channel C through the graph convolution layer, which also helps the network to fully apply the bottleneck strategy and achieve scale and feature compression.

The input and output of ST-Fusion Module are both 3D tensors. For the input $F^l \in \mathbb{R}^{M \times n \times C^l}$ of block l, the output $F^{l+1} \in \mathbb{R}^{(M-2(K_t-1)) \times n \times C^{l+1}}$ is

$$F^{l+1} = \Gamma_1^l * \tau ReLU\left(\Theta^l * g\left(\Gamma_0^l * \tau F^l\right)\right), \tag{11}$$

where Γ_0^l, Γ_1^l are the upper and lower kernels of the temporal convolutional layer of the inclusion graph convolution. Θ^l is the spectral domain convolution kernel in graph convolution, and $\mathrm{Re}\,LU(\cdot)$ represents the activation unit.

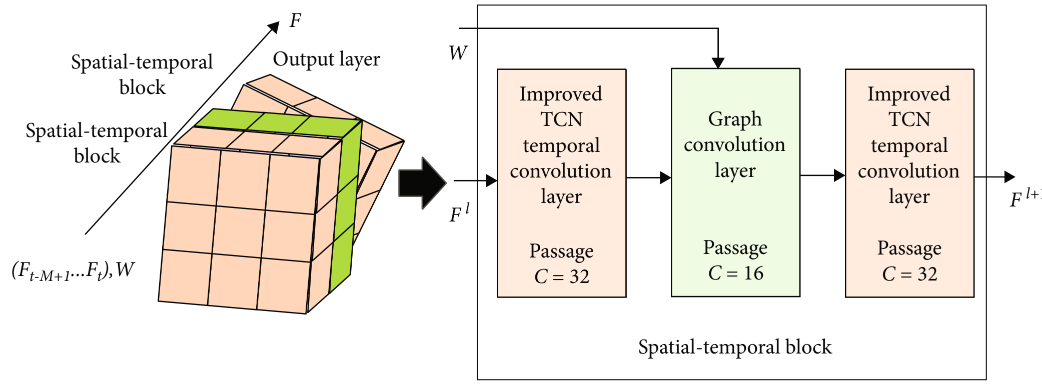


Figure 8: Spatiotemporal block and output connection diagram.

After fusion of temporal convolution and spatial convolution, apply linear transformation $\widehat{F} = Zw + b$ on channel C to obtain n nodes. The predicted value of traffic $w \in \mathbb{R}^c$ is the weight vector, b is the deviation, considering the convergence speed, and using L2-loss to measure the model performance, the flow loss function is expressed as

$$L(\widehat{v}\,; W_\theta) = \sum_t \|\widehat{v}(v_{t-M+1}, \cdots, v_t, W_\theta) - v_{t+1}\|^2. \tag{12}$$

Since the deepening of the spatiotemporal block mentioned above will gradually slow down during the training process, this paper introduces Batch Normalization (BN) before the hidden layer activation function to fix the distribution of the input and pull the distribution back to the normal distribution interval of [0, 1] to speed up convergence speed, while making the optimization smoother. The specific transformation is

$$\widehat{x}^{(k)} = \frac{x^{(k)} - E\left[x^{(k)}\right]}{\sqrt{Var\left[x^{(k)}\right]}}, \tag{13}$$

where $x^{(k)}$ represents the output of the activation function of the hidden layer, $E[\cdot]$represents the mean value, and $Var[\cdot]$represents the variance.

At the same time, two parameters γ, β are added to perform the inverse activation transformation.

$$y^{(k)} = \gamma^{(k)}\widehat{x}^{(k)} + \beta^{(k)}. \tag{14}$$

3.4. Evaluation Module

3.4.1. Entropy Weight Method (EWM). The traditional FCM algorithm does not consider the influence of traffic indicators and individual samples on the clustering results. The EWM uses the idea of entropy value to judge the discrete degree of an indicator and determines the weight of each indicator through the information entropy. In this paper, the EWM is used to determine the weights of different indicators, and the degree of influence of each sample on the clustering results is defined by designing a sample weighting method. First, calculate the proportion p_{ij} and entropy value E_j of sample i under indicator j. The weight w of each traffic indicator is further calculated as

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^{n} x'_{ij}}, \tag{15}$$

$$E_j = -\frac{1}{\ln(n)} \sum_{i=1}^{n} p_{ij} \cdot \ln\left(p_{ij}\right), j = 1, 2, \cdots z, \tag{16}$$

$$w_j = \frac{1 - E_j}{\sum_{j=1}^{z} 1 - E_j}. \tag{17}$$

The specific congestion evaluation process is shown in Figure 9, and the improved FCM algorithm introduced in this paper is an unsupervised fuzzy clustering method, which is a data clustering method based on the optimization of the objective function. The membership degree of the cluster center is represented by a numerical value. Input the feature prediction sample set $\mathrm{X} = \{x_1, x_2, \cdots x_n\}$ and the number of traffic state categories, and then the calculation formulas of the objective function of the improved FCM clustering algorithm for traffic state classification are

$$J(X, U, V) = \sum_{m=1}^{c} \sum_{i=1}^{n} (u_{im})^\sigma d_{im}^2, \tag{18}$$

$$\begin{aligned} t_i &= \exp\left(-\sum_{m=1}^{c} u_{im} \bullet d_{im}^w\right), \\ d_{im}^w &= \left(\sum_{j=1}^{z} w_j x_i - v_m\right). \end{aligned} \tag{19}$$

where U is the affiliation matrix of each sample belonging to different traffic states. V is the matrix composed of all clustering centres. n is the total number of samples. u_{im} is the affiliation degree of samples belonging to traffic state m. σ is the weighted index, indicating that the less fuzzy the algorithm is, the more accurately the state is divided. t is the

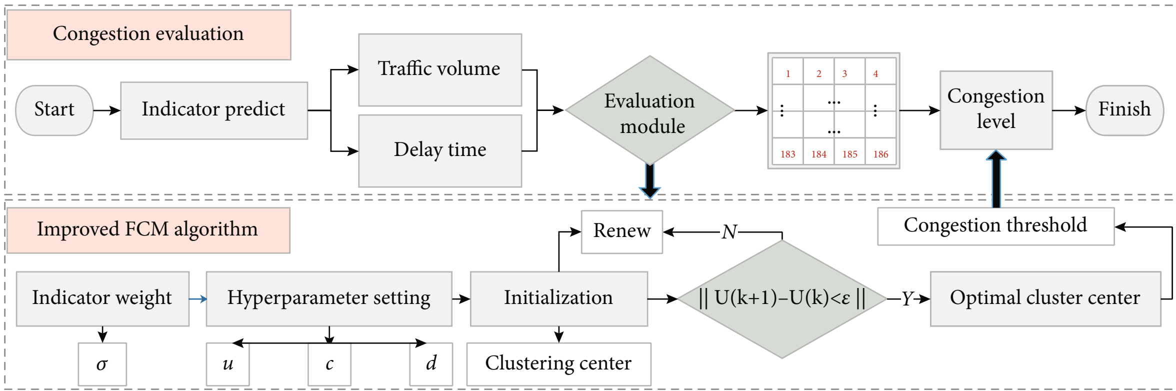
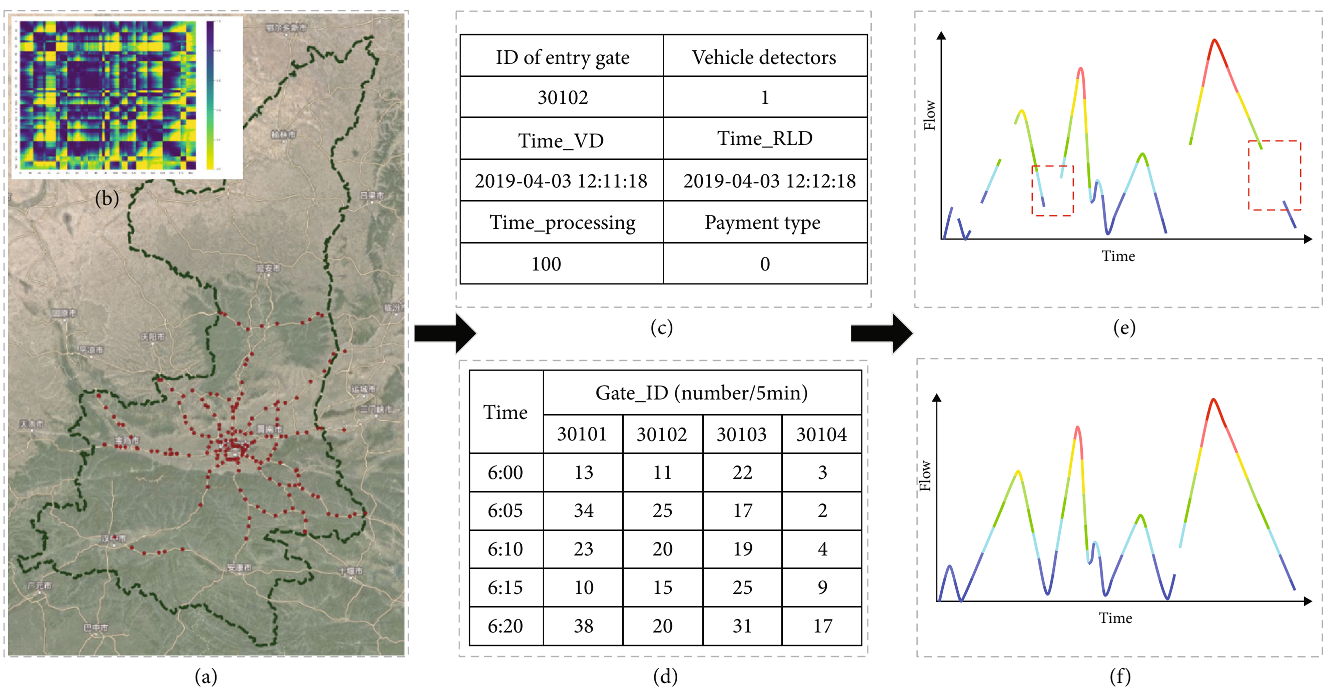


FIGURE 9: The algorithm flow of traffic congestion evaluation.



ID of entry gate	Vehicle detectors
30102	1
Time_VD	Time_RLD
2019-04-03 12:11:18	2019-04-03 12:12:18
Time_processing	Payment type
100	0

Time	Gate_ID (number/5min)			
	30101	30102	30103	30104
6:00	13	11	22	3
6:05	34	25	17	2
6:10	23	20	19	4
6:15	10	15	25	9
6:20	38	20	31	17

(a) (c) (d) (e) (f)

FIGURE 10: Data sources of freeway tolls and data processing of traffic indicators.

weight of sample i. d_{im}^{w} is the weighted Euclidean distance between sample i and cluster center m.

4. Experiments

This section contains the experimental settings and experimental results.

4.1. Experimental Settings

4.1.1. Study Site and Datasets. In this paper, 186 toll-gates on freeways in Shaanxi Province, China are selected as the study site. The locations of the toll-gates are shown in Figure 10(a), showing a radial distribution. Based on the calculation in Equation (5) of the D&C-Matrix, the spatial heat map relationship between the toll-gates obtained by further analysis is shown in Figure 10(b).

In addition, this paper collects the original toll-data for 4 months (December 2018-March 2019) and selects some fields as shown in Figure 10(c), including the toll-gate number, vehicle detector, time to VD and arrival to RLD, respectively (VD and RLD are the positions of (1) and (2) in Figure 3, respectively), travel time, and payment method (choice of ETC and MTC channels). Then, this paper converts the collected raw data into the traffic volume and running delay time required by the experiment. The collected traffic volume data are integrated at 5-minute intervals as shown in Figure 10(d), and the operational delay time is calculated according to Equations ((1)–(3)). Due to equipment failure and other reasons, there will be data missing in some periods as shown in Figure 10(e). Facing the problem of missing temporal data, we analyze the distribution characteristics of missing data and establish a reasonable complementary rule framework to interpolate

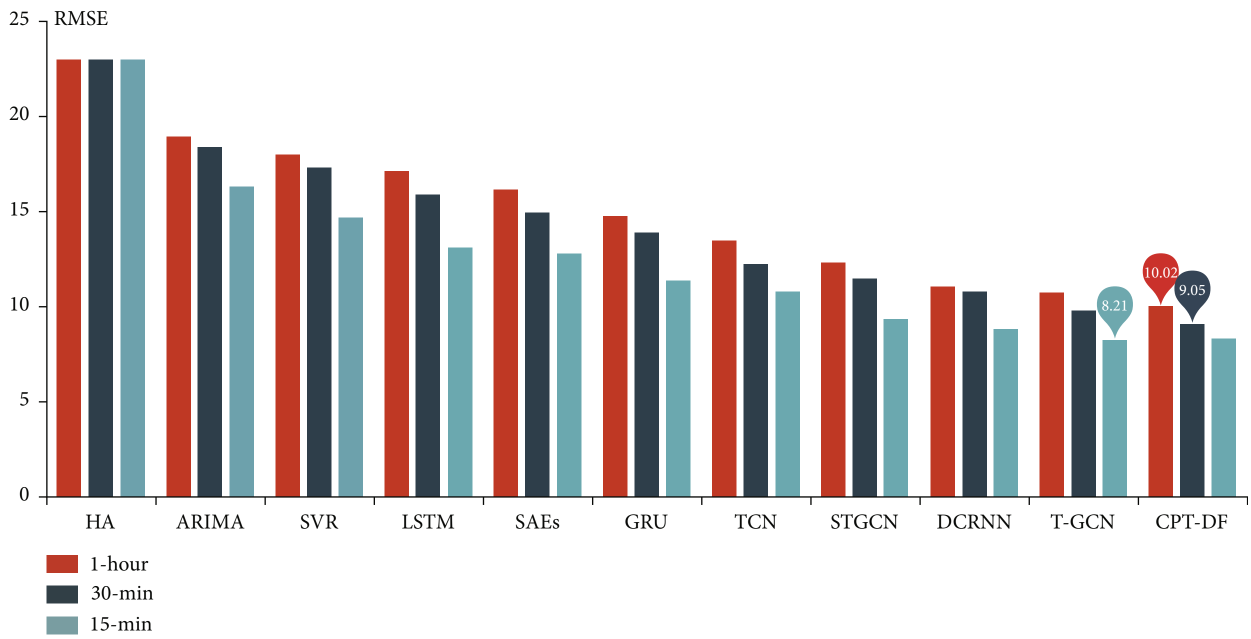


FIGURE 11: Visualization of prediction performance (RMSE) of different models at different time steps.

temporal data. Finally, the data completion effect shown in Figure 10(f) is achieved, and a relatively complete data set is prepared for the experimental part.

4.1.2. Parameters Setting. This experiment uses the ADAM optimizer for training, setting the learning rate every 5 epochs to 0.7, the initial learning rate to 0.001, and the batch size to 50. The channels of the spatiotemporal block are set to 32, 16, 32. In addition, the experiment selected the following three evaluation metrics as shown in Equations (20)–(22).

(i) *Root Mean Squared Error (RMSE).*

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}. \tag{20}$$

(ii) *Mean Absolute Error (MAE).*

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i|, \tag{21}$$

(iii) *Mean Absolute Percentage Error (MAPE).*

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}(|(x_i - \hat{x}_i)/x_i|) * 100\%, \tag{22}$$

where x_i is the true value, and $\hat{x}_i$ is the predicted value.

4.2. Experimental Results. This part uses the feature matrix and adjacency matrix datasets based on the Shaanxi Province toll-data to demonstrate the long-term prediction ability of the model proposed in this paper under large-scale networks. The experimental results are discussed from three aspects: prediction results, evaluation results, and ablation experiments.

4.2.1. Prediction Results. From the overall comparison results, as shown in Figure 11, all models based on temporal and spatial features have better prediction performance than models based on temporal features only. This also proves that there is a strong spatial correlation between various toll stations in the large-scale road network.

Table 2 lists the comparison results of the prediction performance at 15/30/60 minutes of the seven models that only consider the temporal feature. Here, MAE and RMSE have the same units as the quantity being estimated. When the predicted object is traffic, the unit is the number of vehicles/n min; when the predicted object is delay time, the unit is min. The error of the traditional method will be larger because the temporal storage capacity of the linear model is limited, the toll-gates of the freeway are more complex, and nonlinear traffic flow characteristics than ordinary road sections. Especially with historical average (HA) models, relying only on the average makes it difficult to predict accurate results. Compared with ARIMA and SVR models, it is more suitable for short-term prediction work. When the time step increases, the model will have difficulty converging. Deep learning methods (LSTM, SAE, GRU, and TCN) are affected by data distribution and will produce larger prediction errors for gates with large peak-to-valley fluctuations. Especially, the TCN model also shows better performance among the four methods. Because the causal convolution in TCN is different from the traditional time

Table 2: Comparison of the performance of temporal prediction.

Task	15-min			30-min			60-min		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
HA	9.32	17.85%	22.98	9.32	17.85%	22.98	9.32	17.85%	22.98
ARIMA	6.97	13.20%	16.29	7.44	14.09%	18.37	7.90	14.19%	18.89
SVR	6.11	12.91%	14.70	6.55	13.43%	17.21	7.21	13.83%	17.93
LSTM	5.65	12.02%	13.11	6.09	12.71%	15.92	6.95	13.01%	17.07
SAEs	5.23	11.69%	12.78	5.60	12.01%	14.91	6.26	12.61%	16.13
GRU	4.78	10.20%	11.30	5.22	10.92%	13.84	5.68	11.22%	14.72
TCN	4.57	9.21%	10.80	4.91	9.85%	12.18	5.17	10.15%	13.44

Table 3: Comparison of the performance of spatiotemporal prediction.

Task	15-min			30-min			60-min		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE
STGCN	4.06	8.02%	9.31	4.35	9.12%	11.43	4.61	9.62%	12.29
DCRNN	3.79	7.79%	8.73	4.13	8.07%	10.75	4.39	8.37%	11.01
T-GCN	3.69	7.06%	8.21	3.81	7.83%	9.79	4.07	8.03%	10.68
CPT-DF	3.71	7.11%	8.26	3.78	7.44%	9.05	3.97	7.82%	10.02

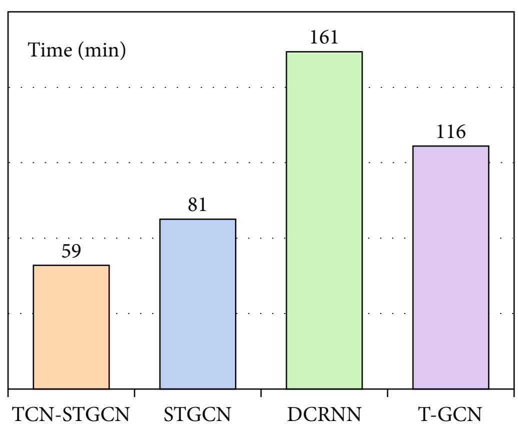


Figure 12: Comparison results of calculation time of different models.

series network, it has the characteristic of a one-way structure in which the value of the next moment only depends on the value of the previous multistep. Furthermore, the receptive field is enlarged by adding dilated convolutions to capture longer dependencies. Therefore, considering the long-term prediction performance, this paper selects and improves the optimal dilated causal convolutional network.

Table 3 lists the comparison results of the prediction performance at 15/30/60 minutes of the four models considering both spatial and temporal features. In graph-related models (T-GCN, DCRNN, STGCN, and CPT-DF), the information of node features and graph structure can be learned end-to-end by GCNs. Furthermore, the topological structure and spatial correlation features of the toll-gate are well captured. The obtained time series with spatial characteristics is further input into the unit model of the processing temporal module, and the dynamic changes are obtained through the information transfer between units to capture the temporal characteristics. Finally, the characteristics of regional gates will be better predicted. In contrast, in the work on the time series of highly nonlinear and complex toll-data, the capture ability and memory ability of CNN's temporal feature are slightly insufficient compared with RNN and its variants (LSTM and GRU). Similarly, the spatiotemporal prediction ability of the STGCN model based on GCN and CNN is slightly insufficient compared with the diffusion convolutional recurrent neural network (DCRNN) based on RNN and the T-GCN model based on GRU. In the short-term prediction (15-minute) work, the T-GCN model adopts GCN to learn complex topology for spatial correlation and GRU to learn dynamic changes of traffic data for temporal correlation. GRU solves the gradient disappearance and gradient explosion problems faced by RNN when training a large amount of data and retains the

Traffic condition	Unblocked	Basically unblocked	Lightly congested	Moderately congested	Severely congested
Congestion level	A	B	C	D	E
Colour					
Traffic volume	(0, 0.164)	(0.164, 0.363)	(0.363, 0.565)	(0.565, 0.753)	(0.753, 1)
Delay time	(0, 0.151)	(0.151, 0.324)	(0.324, 0.484)	(0.484, 0.669)	(0.669, 1)
Cluster centre	(0.066, 0.058)	(0.244, 0.220)	(0.434, 0.377)	(0.563, 0.504)	(0.784, 0.690)

Figure 13: The value range and cluster center corresponding to the two normalized indicators.

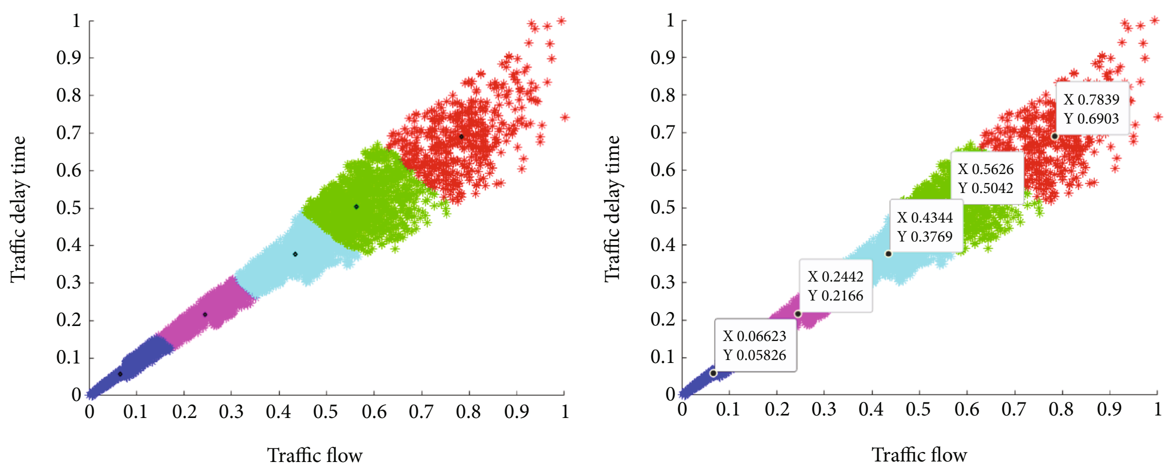


Figure 14: Improved congestion evaluation FCM algorithm clustering results.

trend of historical charging data, so it can better predict traffic feature in the future in the short-term prediction work combined with GCN. However, in the long-term prediction (30-minute and 60-minute) work, different from the RNN structure, the CPT-DF model proposed in this paper can be massively parallelized due to the dilated causal convolution. The expansion coefficient and the size of the filter change the receptive field and can also avoid the problems of gradient dispersion and gradient explosion in RNN. Therefore, the CPT-DF model achieves the best results in the long-term prediction work.

Furthermore, real-time traffic flow prediction is a basic requirement of intelligent transportation systems and has strict requirements on the total time cost of model training and testing [33]. Therefore, in this paper, taking 1 hour as the historical time window, the calculation time of the four graph-based models is calculated as shown in Figure 12. The comparison found that since the gated convolution in STGCN is replaced by the improved dilated causal convolution, the training time of the CPT-DF model is greatly reduced. DCRNN and T-GCN models take longer to train than diffuse causal convolutions when using RNN and its variants to capture time series. By comparing the STGCN model before the improvement, the training time of the improved CPT-DF model is reduced by 27.16%, which will provide advantages for real-time traffic flow prediction.

4.2.2. Evaluation Results. This paper selects the 3-month collection data of 30101 (toll-gate ID) as an example, and calculates the delay time in the congestion area of the toll-gate according to Equations (1), (2), and (3), and the traffic volume is used as the input of the clustering algorithm. According to the "Urban Road Traffic Operation Evaluation Indicator System" issued by China, an evaluation index called Road Traffic Performance Index (TPI) is proposed as shown in Figure 13. Traffic congestion is divided into 5 grades in TPI [34]: unblocked (A), basically unblocked (B), lightly congested (C), moderately congested (D), and severely congested (E). Therefore, set the clustering traffic state category $C = 5$, the fuzzy factor $\sigma = 2$, and the objective function change value $\mathrm{e} = \mathrm{le} - 5$. In order to speed up the training speed and optimize the clustering results to normalize the data, the calculation time of the algorithm is about 25-30 s. Figure 14 shows the visualization results of the traditional improved FCM algorithm on actual traffic data. From left to right, it corresponds to five traffic states (A)-(E), and according to the clustering results of the improved FCM algorithm, two indicators under various traffic states are given, the corresponding value range and cluster center.

Furthermore, this paper is verified based on the proposed clustering method and the actual and predicted values of working days (March 20, 2019) and holidays (March 24, 2019) are selected for comparison. As shown in Figure 15,

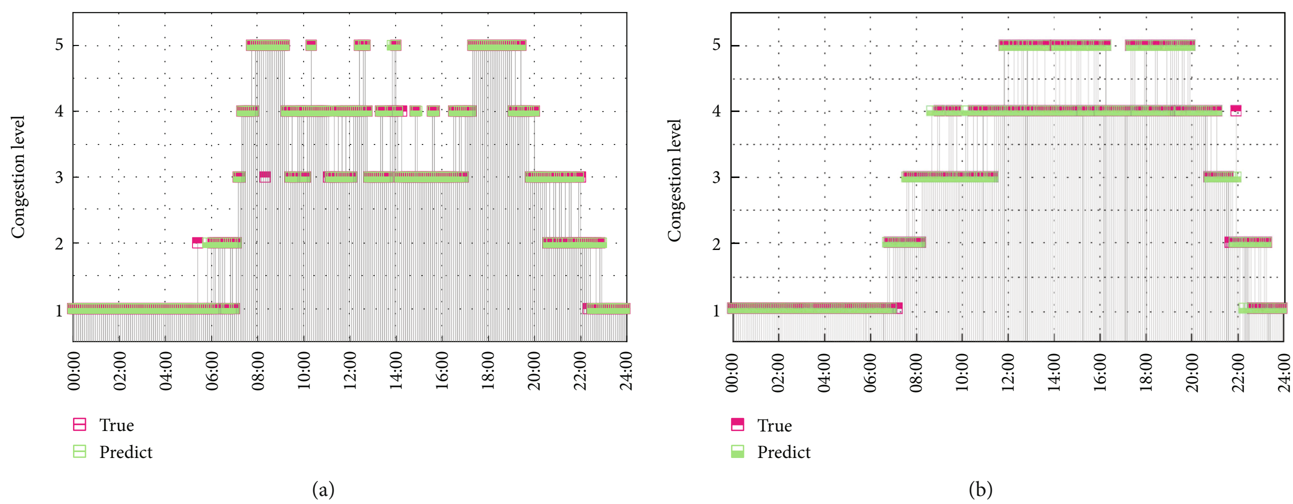


FIGURE 15: Comparison of actual and predicted traffic conditions on weekdays and holidays.

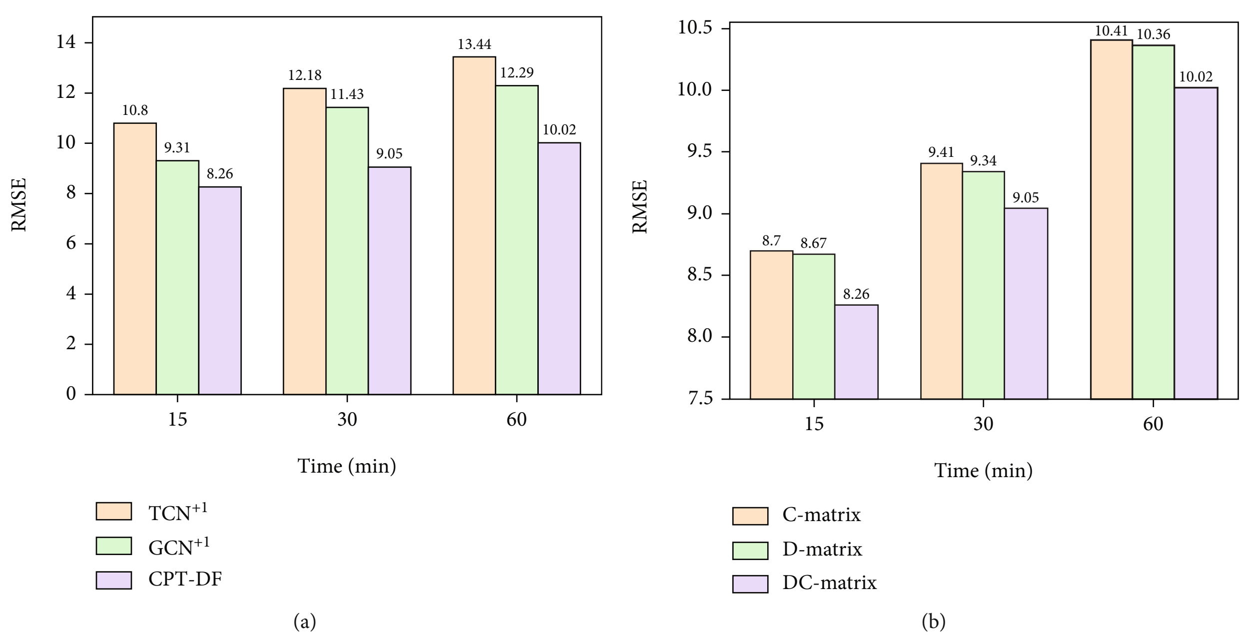


FIGURE 16: (a) Model improvement before and after comparison. TCN^{+1} represents the prediction model after removing the spatial module; GCN^{+1} represents the prediction model after replacing the temporal module. (b) Comparison of prediction accuracy under different matrices.

the clustering results of the improved FCM algorithm proposed in this paper are basically in line with the actual traffic conditions, and the CPT-DF algorithm proposed in this paper can accurately predict future traffic congestion except for special interference points.

4.2.3. Ablation Experiment. In order to further study the optimization effect of different modules on the CPT-DF model, under the same conditions, this paper cancels the temporal module and the spatial module, respectively, to complete the ablation experiment. In addition, this paper also discusses the effect of three adjacency matrices on the prediction performance.

Figure 16(a) shows the comparison of the prediction performance of each module of the CPT-DF model. The performance of the improved CPT-DF model is about 20% higher than that of the TCN^{+1} model that only considers the temporal dimension. The performance of the GCN^{+1} model is reduced by about 12% after removing the TCN^{+1} model. The improved CPT-DF model has stronger spatio-temporal prediction ability and higher traffic flow prediction efficiency. Therefore, the experimental results show that both the temporal module and the spatial module have great effects on the model proposed in this paper.

Figure 16(b) shows the prediction results obtained by using three different matrices (C-Matrix, D-Matrix, and D&C-Matrix). It can be seen that the improved D&C-Matrix has better prediction accuracy than C-Matrix and D-Matrix. It is about 3%-4%, so it can be used as a better construction method for spatial adjacency matrix.

5. Conclusions and Discussion

Large-scale congestion occurs frequently at toll-gates on freeways, especially during holidays or daily peak times. In order to alleviate the congestion of the toll-gate and prevent the occurrence of additional traffic accidents, environmental pollution, etc., it is necessary to select the toll-gate as the study site and effectively predict and evaluate the future congestion based on historical traffic data. In this paper, the topology of the freeway network is modelled as a graph structure. Toll-gates and the ordinary road segments between gates are regarded as vertices and edges in the graph structure, and the traffic volume and operational delay time are selected as feature matrices at the vertices. Then, by analysing the compositional characteristics of toll-gates and constructing a congestion domain, a new AI network of toll-gate congestion based on GCN fusing dilated causal mechanism in DL and FCM clustering integrating coupling weight is proposed in this paper. Further, 4-month toll data of freeways in China's Shaanxi Province are collected to complete the experimental work.

In the experiment, congestion detection of toll-gate is realized from two aspects: prediction and evaluation.

Analysing the prediction results: the performance of the graph-based models is about 8%-35% better than other non-graph models in the long-term prediction (60-min) work. The reason is that the graph-based algorithm additionally verifies the correlation between each toll-gate from the global spatiotemporal dimension and quantifies it using the D&C-Matrix. It provides the possibility for the advance management and traffic guidance for toll-gates of large-scale freeways.

Analysing the evaluation results: the traffic state is reasonably divided into five levels and the congestion of the toll-gate is accurately evaluated using the fuzzy clustering method. It provides a possibility to accurately release the congestion information and avoid wrong alarm of the toll-gates.

The successful prediction could extend to the real-time prediction and early warning of traffic congestion in the toll system to improve the intelligent level of traffic emergency management and guidance on the key road of disasters. On the one hand, when the management department of the freeway receives the accurate traffic indicators of the toll-gate in the future period, it is not only to grasp the regional traffic evolution from global toll-gate but also to adjust the service instructions from a single toll-gate. For example, based on the predicted information, the manager adjusts the optimal ratio of opening and closing for the toll lanes of the freeway in advance to ensure the smooth flow of vehicles. On the other hand, the optimal route and travel time by the guidance is chosen by the driver based on the accurate congestion state of the toll-gate predicted in advance. For example, a driver will have n routes that can be selected from the origin A to the destination B, and then based on the congestion evaluation and the calculation of the travel time, multiple options (the least delay, the shortest distance, or the least congestion, etc.) are provided for the road user, which is especially effective in the emergency of key road of disasters.

Furthermore, if real-time and multisource traffic data based on toll-gates or other sensors are collected, the work of traffic prediction and evaluation of multitime span of day/month/year and multiscale (road network, local lanes, vehicle, etc.) can be further analysed in a more fine-grained manner using the AI graph network proposed in this paper. Of course, each method will have some unexpected problems when it is applied in practice. For example, although our proposed model can predict future traffic congestion in spatiotemporal dimension, it cannot intuitively explain the rules for the evolution of traffic flow outside the "congestion domain" of toll-gates, including how congestion flow forms and dissipates. Therefore, in the follow-up work, we will combine the evolution rules of actual traffic flow to better control the change of traffic to improve the intelligent level of traffic emergency management.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

The research is financially supported by National Key Research and Development Program of China, Key technologies for derivative composite disaster assessment and emergency adapting in the Guangdong-Hong Kong-Macao Greater Bay Area (2021YFC3001000).

References

[1] Y. Wang, X. Yu, S. Zhang et al., "Freeway traffic control in presence of capacity drop," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1497–1516, 2021.

[2] S. Zahedian, A. Nohekhan, and K. F. Sadabadi, "Dynamic toll prediction using historical data on toll roads: case study of the I-66 inner beltway," *Transportation Engineering*, vol. 5, article 100084, 2021.

[3] L. Shen, J. Lu, D. Geng, and L. Deng, "Peak traffic flow predictions: exploiting toll data from large expressway networks," *Sustainability*, vol. 13, no. 1, pp. 260–318, 2021.

[4] L. Yan, P. Wang, J. Yang, Y. Hu, Y. Han, and J. Yao, "Refined path planning for emergency rescue vehicles on congested urban arterial roads via reinforcement learning approach," *Journal of Advanced Transportation*, vol. 2021, no. 1, Article ID 8772688, p. 12, 2021.

[5] M. Gori, G. Mandarina, and F. Scarcella, "A new model for earning in graph domains," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2, no. 2, pp. 729–734, 2005.

[6] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.

[7] J. Zhou, G. Cui, S. Hu et al., "Graph neural networks: a review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020.

[8] B. Yu, Y. Lee, and K. Sohn, "Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (GCN)," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 189–204, 2020.

[9] X. Shi, H. Qi, Y. Shen, G. Wu, and B. Yin, "A spatial–temporal attention approach for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 4909–4918, 2021.

[10] B. Yu, H. Yin, and Z. Zhu, "Spatiotemporal graph convolutional networks: a deep learning framework for traffic forecasting," in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 3634–3640, Stockholm, 2018.

[11] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," 2017, https://arxiv.org/abs/1707.01926.

[12] L. Zhao, Y. Song, C. Zhang et al., "T-GCN: a temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2020.

[13] J. Zhao, Y. Gao, Z. Bai, H. Wang, and S. Lu, "Traffic speed prediction under non-recurrent congestion: based on LSTM method and BeiDou navigation satellite system data," *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 70–81, 2019.

[14] Z. Lv, L. Qiao, K. Cai, and Q. Wang, "Big data analysis technology for electric vehicle networks in smart cities," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1807–1816, 2021.

[15] B. Williams and L. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.

[16] H. Yang, P. Jin, B. Ran, D. Yang, Z. Duan, and L. He, "Freeway traffic state estimation: a Lagrangian-space Kalman filter approach," *Journal of Intelligent Transportation Systems*, vol. 23, no. 6, pp. 525–540, 2019.

[17] C. Wu, J. Ho, and D. Lee, "Travel-time prediction with support vector regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 276–281, 2004.

[18] P. Wang, W. Hao, and Y. Jin, "Fine-grained traffic flow prediction of various vehicle types via fusion of multisource data and deep learning approaches," *IEEE Transactions on Intelligent Transportation Systems.*, vol. 22, no. 11, pp. 6921–6930, 2021.

[19] C. Shuai, W. Wang, G. Xu, M. He, and J. Lee, "Short-term traffic flow prediction of expressway considering spatial influences," *Journal of Transportation Engineering, Part A: Systems*, vol. 148, no. 6, 2022.

[20] W. Zhang, Y. Yu, Y. Qi, F. Shu, and Y. Wang, "Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning," *Transportmetrica A Transport Science*, vol. 15, no. 2, pp. 1688–1711, 2019.

[21] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33no. 1, pp. 922–929, 2021.

[22] M. Akhtar and S. Moridpour, "A review of traffic congestion prediction using artificial intelligence," *Journal of Advanced Transportation*, vol. 2021, Article ID 8878011, 18 pages, 2021.

[23] T. Afrin and N. Yodo, "A probabilistic estimation of traffic congestion using Bayesian network," *Measurement*, vol. 174, article 109051, 2021.

[24] R. Esfahani, F. Shahbazi, and M. Akbarzadeh, "Three-phase classification of an uninterrupted traffic flow: a k-means clustering study," *Transportmetrica B: transport dynamics*, vol. 7, no. 1, pp. 546–558, 2019.

[25] J. Dunn, "Well-Separated clusters and optimal fuzzy partitions," *Journal of cybernetics*, vol. 4, no. 1, pp. 95–104, 1974.

[26] Z. Cheng, W. Wang, J. Lu, and X. Xing, "Classifying the traffic state of urban expressways: a machine-learning approach," *Transportation Research Part A: Policy and Practice*, vol. 137, pp. 411–428, 2020.

[27] S. Gan, S. Liang, K. Li, J. Deng, and T. Cheng, "Trajectory length prediction for intelligent traffic signaling: a data-driven approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 2, pp. 426–435, 2018.

[28] T. Vincenty, "Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations," *Survey Review*, vol. 23, no. 176, pp. 88–93, 1975.

[29] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," 2014, https://arxiv.org/abs/1312.6203.

[30] E. Michael, K. Hermann, and B. Dagmar, "Large-scale quantum networks based on graphs," *New Journal of Physics*, vol. 18, no. 5, article 053036, 2016.

[31] S. Bai, J. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modelling," 2018, https://arxiv.org/abs/1803.01271.

[32] J. Wang, Y. Zhang, Y. Wei, Y. Hu, X. Piao, and B. Yin, "Metro passenger flow prediction via dynamic hypergraph convolution networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 12, pp. 7891–7903, 2021.

[33] J. Yang, P. Wang, W. Yuan, Y. Ju, W. Han, and J. Zhao, "Automatic generation of optimal road trajectory for the rescue vehicle in case of emergency on mountain freeway using reinforcement learning approach," *IET Intelligent Transport Systems*, vol. 15, no. 9, pp. 1142–1152, 2021.

[34] J. Jiang, Q. Chen, J. Xue, H. Wang, and Z. Chen, "A novel method about the representation and discrimination of traffic state," *Sensors*, vol. 20, no. 18, p. 5039, 2020.