A Robust Spatio-Temporal Graph Attention Network for Traffic Flow Forecasting Under Sudden Disturbances

Limi Chen[1,2,*], Zhihao Jiang[1,2], Jing Yang[1]
[1]Hainan Vocational University of Science and Technology, Haikou 571126, China
[2]Institute for Mathematical Research, Universiti Putra Malaysia, Serdang 43400, Malaysia

Corresponding author: Limi Chen (e-mail: chenlimi256@126.com)

ABSTRACT The prediction of traffic flow has evolved from an empirical extrapolation method to a spatiotemporal correlation-based method due to the development of artificial intelligence and intelligent algorithms. In order to remedy the problem that the disturbance in the future causes the prediction error to increase sharply, we put forward the robust spatiotemporal graph attention network. Based on graph attention and gated temporal modeling, we provide a framework which brings together a road network topological encoding, a disturbance factor mapping, a temporal dependency extraction and a strong loss constraint mechanism. Experimental results show that under normal conditions, the model achieves an MAE of 13.42%, an RMSE of 20.74%, and a MAPE of 8.31%, outperforming STGCN's 14.19%, 21.86%, and 8.64%, respectively; under disturbance conditions, the MAE is 17.96%, with a robustness index of 0.874, demonstrating good prediction stability.

INDEX TERMS Traffic flow prediction; Sudden disturbances; Spatio-temporal graph attention; Robust constraints; Road network modeling

I. INTRODUCTION

THE use of Artificial Intelligence (AI) and smart algorithms for traffic condition sensing, road network operation analysis, and short-term traffic flow prediction continues to be progressed, moving the paradigm of traffic flow prediction from time-series extrapolation to joint spatiotemporal modeling. Road network operations in urban areas have unique periodic and nonlinear as well as spatial propagation features. Unexpected events like road accidents, construction works, extreme weather conditions and temporary traffic controls can change the normal traffic flow evolution and can quickly transform a local congestion into a much larger problem, which can severely decrease the reliability of traffic prediction models.

Zheng et al. constructed a graph multi-attention network to enhance the representational capacity of spatiotemporal features in traffic forecasting [1]; Wang et al. utilized spatiotemporal graph neural networks to characterize dynamic dependencies between traffic nodes [2]; Song et al. proposed a synchronous graph convolutional framework to enhance the collaborative modeling capability of spatiotemporal network data [3]; Bai et al. mitigated the expressive limitations imposed by fixed adjacency relationships through an adaptive graph convolutional recurrent structure [4]. Wu et al. established graph-structured associations from a multivariate time series perspective, providing methodological support for predicting complex traffic conditions [5]; Li and Zhu emphasized the role of spatial and temporal feature fusion in improving the accuracy of traffic flow prediction [6]; Ta et al. further introduced an adaptive spatiotemporal graph structure, improving the model's adaptability to dynamic road network relationships [7]; Chen et al. considered multiple dynamic factors, expanding the scope of feature fusion research in traffic forecasting [8].

Although significant research has been done that forms the basis for spatiotemporal road network forecasting, there are still several gaps in the explicit representation of sudden disturbances, representation of disturbance propagation and development of robust constraints. To solve these issues, we turn to the traffic flow prediction task in sudden disturbance situations and build a spatiotemporal graph attention network with the consideration of the graph structure of the road network, the temporal dependence, disturbance information, and robust loss functions. In this way, prediction accuracy in regular traffic is increased and prediction stability in non-stationary traffic situations (e.g. accidents, construction, weather anomalies) is increased.

II. MODELING THE TRAFFIC FLOW PREDICTION PROBLEM UNDER SUDDEN DISTURBANCES

A. DEFINITION OF THE TRAFFIC NETWORK GRAPH STRUCTURE

A traffic network can be abstracted as a weighted directed graph $G = (V, E, A)$, where $V = \{v_1, v_2, \ldots, v_N\}$ represents the set of traffic detection nodes, E denotes the road connections between nodes, and $A \in \mathbb{R}^{N\times N}$ is the adjacency matrix, used to describe the spatial topology of the network and the intensity of traffic flow propagation. If nodes v_i and v_j are directly connected, then $A_{ij} > 0$; otherwise, $A_{ij} = 0$.

To account for the attenuating effects of road distance and connectivity on traffic, the adjacency weight can be defined as:

$$A_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), \quad e_{ij} \in E \tag{1}$$

where d_{ij} is the road distance between nodes v_i and v_j, and σ is the distance attenuation coefficient [9]. The traffic state of a node at time t is represented as $X_t \in \mathbb{R}^{N\times F}$, where F denotes the traffic characteristic dimensions, including flow, speed, and occupancy rate. The mapping relationship between the actual road network and the graph structure is shown in Fig. 1.

B. CHARACTERIZATION OF SUDDEN DISTURBANCES

Sudden disturbance factors primarily include traffic accidents, road construction, severe weather, temporary traffic controls, and large-scale events; their impacts are time-varying, localized, and spatially diffusive [10]. Let the set of disturbance events at time t be denoted as $D_t = \{d_t^1, d_t^2, \ldots, d_t^M\}$. For each disturbance category m, a feature vector is constructed as follows:

$$z_t^m = [c_m, l_m, \tau_m, \rho_m] \tag{2}$$

where c_m denotes the disturbance type code, l_m is the disturbance location, τ_m is the duration, and ρ_m is the disturbance intensity [11]. Considering the attenuation effects of disturbances on different traffic nodes, the disturbance weight for node v_i can be expressed as:

$$P_{i,t} = \rho_m \exp\left(-\frac{\mathrm{dist}(v_i, l_m)}{\lambda}\right) \tag{3}$$

where $\mathrm{dist}(v_i, l_m)$ is the network distance from the node to the disturbance location, and λ is the influence attenuation coefficient [12]. This yields the disturbance feature matrix P_t, which is used to characterize the impact of sudden events on the state of network nodes. The mapping relationship between disturbance factors and traffic node features is shown in Fig. 2.

C. DESCRIPTION OF THE SPATIOTEMPORAL TRAFFIC FLOW FORECASTING TASK

Suppose the road network consists of N traffic nodes. The traffic state of each node at time t is $X_t \in \mathbb{R}^{N\times F}$, where F represents the feature dimensions, which primarily include traffic volume, average speed, road occupancy, and temporal attributes [13]. Given a continuous historical time window $\{X_{t-T+1}, X_{t-T+2}, \ldots, X_t\}$, an adjacency matrix A, and a disturbance feature matrix P_t, the prediction of traffic volume for the next H time steps can be expressed as:

$$\hat{Y}_{t+1:t+H} = f_\theta(X_{t-T+1:t}, A, P_t) \tag{4}$$

where T is the length of the historical input window, H is the prediction step size, f_θ represents the robust spatiotemporal prediction model to be learned, and $\hat{Y}_{t+1:t+H}$ is the prediction result [14].

III. CONSTRUCTION OF THE ROBUST SPATIO-TEMPORAL GRAPH ATTENTION NETWORK MODEL

A. OVERALL MODEL ARCHITECTURE

The robust spatiotemporal graph attention network is composed of the input layer, the spatial graph attention layer, the temporal dependency modeling layer, the disturbance detection layer, the robust constraint layer and the prediction output layer. The historical traffic state matrix, the road network adjacency matrix and the disturbance feature matrix are first input to the model. Then, it takes the dynamic relationships between nodes by the spatial graph attention module, and short-term fluctuation and periodic change of traffic flow by temporal modeling module [15]. The disturbance detection layer picks out the effect of sudden events on local nodes as well as the neighboring nodes, and the robust constraint layer filters the error amplification caused by abnormal fluctuations, to finally obtain multiple steps traffic flow forecasts. The general scheme of the model is in Fig. 3.

B. SPATIAL GRAPH ATTENTION FEATURE EXTRACTION

The spatial graph attention layer is used to characterize the strength of dynamic relationships between traffic nodes, thereby avoiding the inadequate representation of spatial propagation relationships that results from relying solely on a fixed adjacency matrix [16]. Let the input features of node v_i at time t be h_i^t. After a linear mapping, the latent representation is obtained as:

$$\tilde{h}_i^t = W_s h_i^t \tag{5}$$

where W_s is the spatial feature mapping matrix [17]. For nodes v_j adjacent to node v_i, their attention correlation can be expressed as:

$$e_{ij}^t = \mathrm{LeakyReLU}\left(a^T[\tilde{h}_i^t \parallel \tilde{h}_j^t]\right) \tag{6}$$

where a is the attention parameter, and $[\cdot \parallel \cdot]$ denotes feature concatenation [18]. Subsequently, Softmax is applied to normalize the weights of neighboring nodes:

$$\alpha_{ij}^t = \frac{\exp(e_{ij}^t)}{\sum_{k\in\mathcal{N}_i} \exp(e_{ik}^t)} \tag{7}$$

Finally, the spatial aggregated feature of node v_i is obtained by weighting the features of neighboring nodes [19]. This

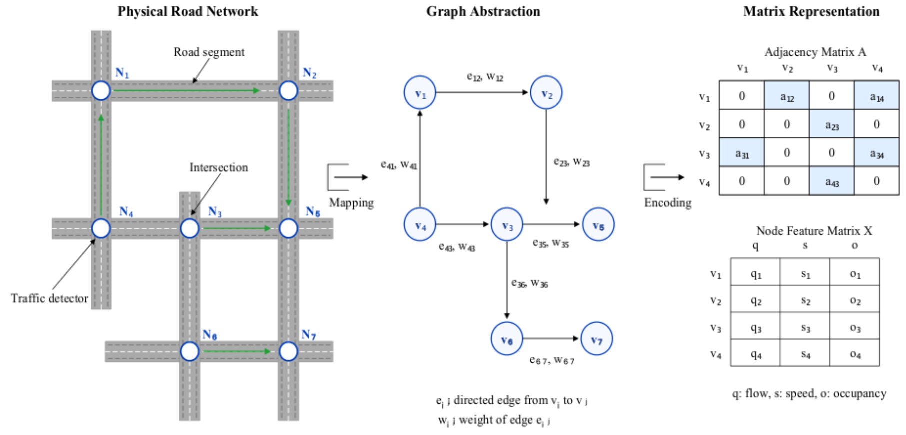

Fig. 1. Schematic diagram of the traffic network graph structure.

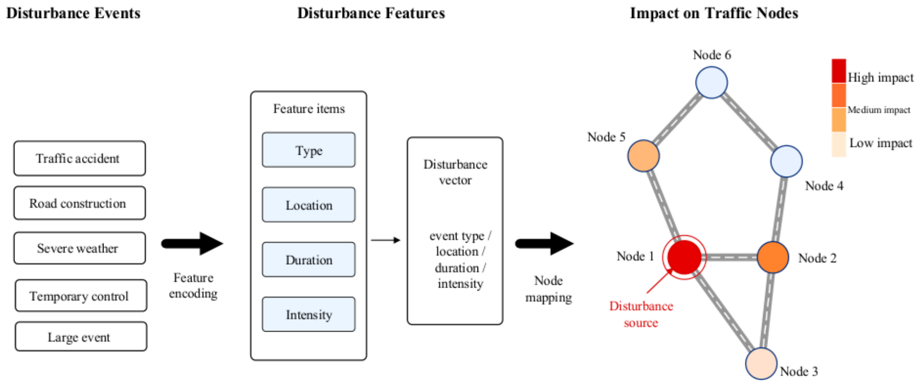

Fig. 2. Mapping of sudden disturbance factors.

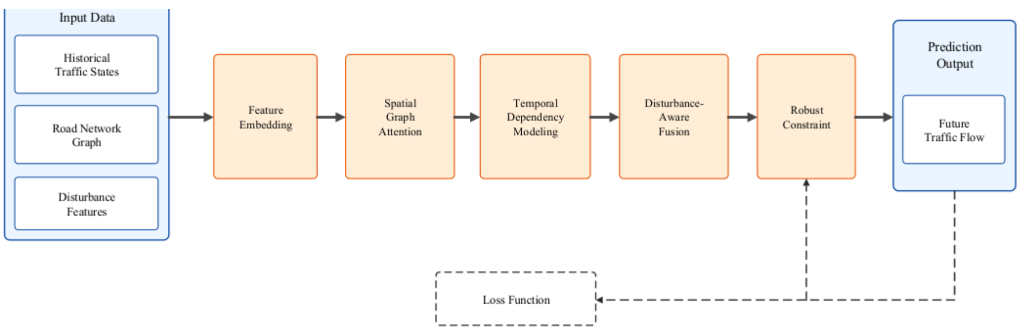

Fig. 3. Overall model architecture.

mechanism can adaptively adjust the influence weights of different road segments based on changes in traffic conditions, enhancing the model's ability to capture congestion propagation and disturbance propagation. The spatial attention mechanism is illustrated in Fig. 4.

C. TIME-DEPENDENT MODELING MECHANISM

The patterns of changes in the traffic flow are extracted in continuous time series using the time-dependent modeling module. A sampling interval of 5 minutes is used in this paper and 12 consecutive historical time steps are inputted to predict the next 3 time steps. The input sequence for each traffic node contains data like flow, speed, occupancy, time stamps, and describes the short time variations, periodic variations, and recovery from disturbances [20]. Based on the output from the spatial graph attention mechanism, the model incorporates gated recurrent units for recursive update of historical states to make sure that changes in traffic flow at critical moments are effectively retained.

The temporal state update can be expressed as:

$$H_t = \mathrm{GRU}(S_t, H_{t-1}) \tag{8}$$

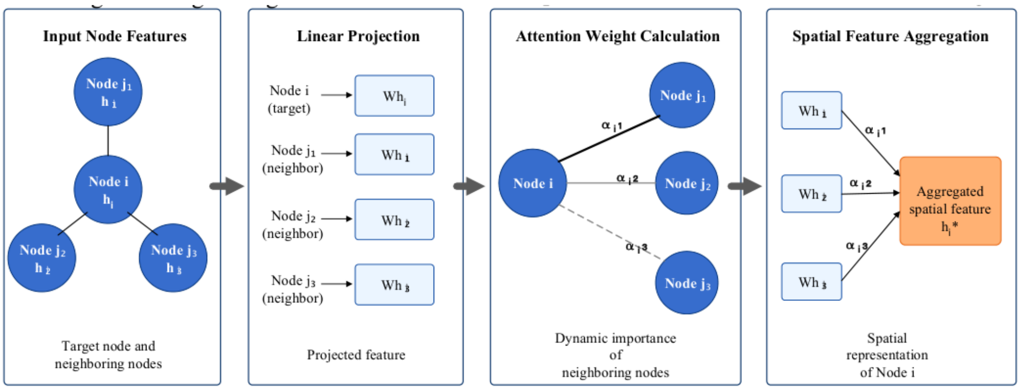

Fig. 4. Diagram of the spatial attention mechanism.

where S_t is the feature output from the spatial attention layer, and H_t is the hidden state at the current time step [21]. To enhance the ability to identify peak periods and disturbance periods, temporal attention weights are further introduced to assign differentiated influences to different historical time steps. The process of modeling temporal dependencies is shown in Fig. 5.

D. DISTURBANCE DETECTION AND ROBUST CONSTRAINT DESIGN

The disturbance detection module is used to describe the local impact and neighborhood propagation of sudden events on traffic nodes. The types of disturbance event are divided into five categories: accidents, construction, severe weather, temporary traffic controls and large-scale events. The sampling interval is 5 min, the disturbance influence radius is 0–3 km, and the duration is 15 min, 30 min, 60 min and 120 min [22].

Let the disturbance perception intensity at node v_i at time t be $r_{i,t}$; then:

$$r_{i,t} = \sum_{m=1}^{M} \rho_m \cdot \exp\left(-\frac{d_{i,m}}{\lambda}\right) \cdot \eta_{m,t} \tag{9}$$

where M represents the number of disturbance events, ρ_m represents the disturbance intensity of the mth class, $d_{i,m}$ represents the network distance from node v_i to the disturbance location, λ represents the spatial attenuation coefficient, and $\eta_{m,t}$ represents the temporal indicator variable indicating whether the disturbance is active at time t [23]. After fusing the disturbance characteristics with the traffic state, the model is expressed as:

$$\bar{h}_{i,t} = \sigma(W_h h_{i,t} + W_r r_{i,t} + b) \tag{10}$$

where $h_{i,t}$ represents the latent traffic features at a node, W_h, W_r are the learnable weight matrices, b is the bias term, and $\sigma(\cdot)$ is the activation function [24]. To reduce the interference of abnormal fluctuations on model training, a robust constraint term is introduced:

$$L_{\text{rob}} = \frac{1}{N}\sum_{i=1}^{N} \omega_{i,t}\,|\hat{y}_{i,t} - y_{i,t}| \tag{11}$$

where $\hat{y}_{i,t}$ and $y_{i,t}$ represent the predicted and true values, respectively; $\omega_{i,t}$ is the perturbation weight; and N is the number of nodes [25].

E. DESIGN OF PREDICTED OUTPUT AND LOSS FUNCTION

The prediction output layer maps the spatiotemporal features, after incorporating perturbations, to predicted future traffic flow values. Let the prediction result for node v_i over the next H time steps be $\hat{Y}_{t+1:t+H}$; its output can be expressed as:

$$\hat{Y}_{t+1:t+H} = W_o H_t + b_o \tag{12}$$

where H_t is the model's final spatiotemporal latent representation, W_o is the output mapping matrix, and b_o is the bias term [26].

To balance prediction accuracy in normal scenarios and stability in perturbed scenarios, the loss function consists of a prediction error term, a perturbation-weighted term, and a regularization term:

$$L = L_{\text{pre}} + \beta L_{\text{rob}} + \gamma L_{\text{reg}} \tag{13}$$

where L_{pre} measures the overall error between the predicted and true values, L_{rob} enhances robustness under perturbed nodes, and L_{reg} suppresses parameter overfitting; β and γ are weight coefficients.

IV. MODEL TRAINING AND OPTIMIZATION METHODS

Before training the model, missing data was imputed, outliers were corrected and the traffic flow, speed, occupancy and disturbance features were normalized. All these features were scaled to [0, 1] to decrease the effect of the various units on model convergence. Two samples were made for predicting the 3-time-step traffic flow with the historical traffic data window equal to 12 time steps and the data sampling period equal to 5 minutes. Data was divided into training (70%), validation (20%) and test (10%) set. The validation set was used to tune hyper parameters and to make decisions on early stop, while using the test set for evaluation of final performance of the model, and training

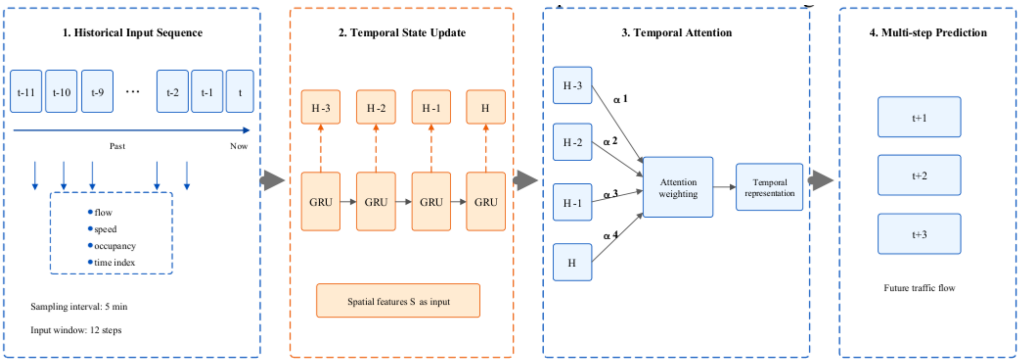

Fig. 5. Temporal dependency modeling diagram.

TABLE 1. Training parameters

Parameter Name	Parameter Setting	Description
Sampling interval	5 min	Data collection period of traffic states
Input window	12 steps	Historical data from the past 60 min
Prediction horizon	3 steps	Traffic flow prediction for the next 15 min
Data split	7:2:1	Training set, validation set, and test set
Batch size	64	Number of samples in each training batch
Learning rate	0.001	Initial learning rate of the Adam optimizer
Maximum epochs	200	Maximum number of training iterations
Early stopping	15	Stop if the validation error does not decrease
Dropout	0.3	Used to prevent model overfitting
Regularization coefficient	0.0001	Used to control model complexity

TABLE 2. Experimental environment and dataset configuration

Configuration Category	Configuration Item	Setting
	CPU	Intel Core i9-12900K
Hardware Environment	GPU	NVIDIA RTX 3090 24 GB
	Memory	64 GB
	Operating System	Ubuntu 20.04
Software Environment	Programming Language	Python 3.9
	Deep Learning Framework	PyTorch 2.0
	Data Type	Traffic flow, speed, occupancy, disturbance events
Data Configuration	Sampling Interval	5 min
	Input Window	12 steps
	Prediction Horizon	3 steps
	Data Split	7:2:1

set for parameter learning. The model is trained with Adam optimizer and a learning rate of 0.001, a batch size of 64, and a maximum of 200 epochs. To prevent overfitting, Dropout and L_2 regularization are introduced, and training is halted when the validation set error fails to decrease for 15 consecutive epochs. The main training parameters are shown in Table 1.

V. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENTAL ENVIRONMENT CONFIGURATION

The experiments were performed on Ubuntu 20.04 operating system with hardware platform consisting of Intel Core i9-12900K CPU, 64 GB RAM and NVIDIA RTX 3090 GPU. Model implementation was done in python 3.9 with PyTorch 2.0 and trained with the Adam optimizer. The experimental data collected were detection section data from urban expressways and arterials with a 5-minute sampling interval that included data on traffic volume, speed, and occupancy rate as well as sudden disturbance events. The data was then imputed for missing values, corrected for outliers and normalized, before being divided into Training, Validation and Test sets in the ratio of 7:2:1. The experimental environment and the configuration of the experimental datasets are presented in Table 2.

B. DESIGN OF COMPARATIVE EXPERIMENTS

For comparison, we chose HA, ARIMA, SVR, Random Forest, LSTM, GRU, TCN, GCN, GAT and STGCN as representatives of traditional time series model, machine learning model, deep time series model and spatio-temporal graph model respectively. The data splitting method, the input window, the prediction step size and the evaluation metrics were all the same for all models to ensure a fair comparison. The evaluation metrics selected were MAE, RMSE, MAPE, and R^2, used to measure prediction error, error volatility, relative deviation, and fitting ability, respectively. To test the robustness of the models under non-stationary traffic conditions, experiments were carried out under both normal traffic conditions and sudden disturbance scenarios which involved both accident and road construction, severe weather and temporary traffic control samples. The comparative analysis was carried out on the differences of the models related to spatial correlation modeling, temporal dependency capture and disturbance adaptability.

C. ANALYSIS OF PREDICTION PERFORMANCE UNDER NORMAL TRAFFIC CONDITIONS

To evaluate the models' predictive capabilities under normal traffic conditions, multiple benchmark models were selected for comparison, using the same input window, prediction step size, and evaluation metrics. The results are shown in Table 3.

As shown in Table 3, the MAEs for HA and ARIMA are 24.86 and 22.43, respectively, indicating that traditional methods are insufficient for capturing complex traffic fluctuations. Among deep time-series models, the MAE for GRU

TABLE 3. Comparison of prediction performance among different models

Model	MAE	RMSE	MAPE (%)	R^2
HA	24.86	35.71	15.64	0.781
ARIMA	22.43	32.96	14.27	0.803
SVR	20.18	30.44	12.96	0.826
Random Forest	18.73	28.61	11.82	0.847
LSTM	16.92	25.83	10.47	0.872
GRU	16.35	24.96	10.12	0.881
TCN	15.68	23.91	9.64	0.894
GCN	15.21	23.47	9.38	0.901
GAT	14.67	22.58	8.92	0.913
STGCN	14.19	21.86	8.64	0.921
Proposed Model	13.42	20.74	8.31	0.934

TABLE 4. Comparison of robustness among different models

Model	Disturbance MAE	Disturbance RMSE	Error Growth Rate (%)	Robustness Index
HA	38.74	52.63	55.83	0.612
ARIMA	34.91	48.26	55.64	0.635
SVR	29.86	42.17	47.97	0.681
Random Forest	27.45	39.62	46.56	0.704
LSTM	24.13	35.48	42.61	0.742
GRU	23.21	34.06	41.96	0.758
TCN	21.94	32.38	39.92	0.779
GCN	21.16	31.74	39.12	0.792
GAT	20.27	30.48	38.17	0.811
STGCN	19.46	29.35	37.14	0.835
Proposed Model	17.96	27.14	33.83	0.874

TABLE 5. Comparison of ablation experiment results for different modules

Model Variant	Normal MAE	Disturbance MAE	RMSE	MAPE (%)	Robustness Index
Without Spatial Attention	15.76	20.82	31.46	11.37	0.804
Without Temporal Modeling	16.21	21.35	32.08	11.84	0.791
Without Disturbance Awareness	14.98	22.64	33.71	12.26	0.765
Without Robust Constraint	14.31	20.93	31.25	11.18	0.813
Without Spatial Smoothing	14.06	19.84	29.92	10.73	0.836
Complete Model	13.42	17.96	27.14	9.68	0.874

TABLE 6. Parameter sensitivity analysis

Parameter	Setting	MAE	RMSE	MAPE (%)	Training Time per Epoch (s)
Input Window	6 steps	15.08	23.46	9.42	18.7
	9 steps	14.16	21.92	8.76	21.4
	12 steps	13.42	20.74	8.31	24.9
	15 steps	13.59	20.96	8.44	29.6
Attention Heads	2	14.05	21.71	8.69	22.8
	4	13.42	20.74	8.31	24.9
	6	13.48	20.83	8.37	28.2
Hidden Dimension	32	14.23	22.18	8.84	21.6
	64	13.42	20.74	8.31	24.9
	128	13.37	20.69	8.29	33.5
Robust Weight β	0.2	13.86	21.32	8.57	24.5
	0.4	13.42	20.74	8.31	24.9
	0.6	13.71	21.08	8.49	25.1

drops to 16.35, outperforming the 16.92 achieved by LSTM. After introducing spatial structure, the MAE for STGCN is 14.19, with an R^2 of 0.921. The MAE of the model proposed in this paper was further reduced to 13.42, with an RMSE of 20.74 and a MAPE of 8.31%; the R^2 improved to 0.934, indicating its more stable ability to fit normal spatiotemporal traffic variations.

D. ROBUSTNESS VERIFICATION UNDER SUDDEN DISTURBANCE SCENARIOS

To assess adaptability under sudden disturbance conditions such as accidents, construction, severe weather, and temporary traffic controls, a disturbance test set was constructed for a robustness comparison; the results are shown in Table 4.

Table 4 indicates that sudden disturbances result in considerable higher prediction errors of all models. The MAE of the HA model under disturbances is 38.74 with an error growth rate of 55.83% which is weak. With the help of spatial modeling, the MAE of the STGCN model is reduced to 19.46 and the robustness index is 0.835. The MAE value of the model in this paper is 17.96, the RMSE value is 27.14, the error growth rate is controlled within 33.83% and the robustness index is improved to 0.874. This shows that disturbance awareness and strong constraints are able to successfully minimize the amplification of errors due to abnormal traffic situations.

E. ABLATION EXPERIMENTS AND PARAMETER SENSITIVITY ANALYSIS

To assess the contribution of each module to model performance, we sequentially removed the spatial attention, temporal modeling, perturbation awareness, robust constraint, and spatial smoothing modules. The ablation results are shown in Table 5.

Table 5 shows that the disturbance MAE increased to 22.64 and robustness index decreased to 0.765 when the disturbance perception module was removed, which means that the disturbance features play the most important role in predicting the sudden events. The normal MAE was 16.21 after the time modeling was removed, showing that the capacity to model short term changes was much reduced. Disturbance MAE was 20.82 after spatial attention was removed, which was not enough to model the propagation relationship between traffic in the road network. The complete model obtained the lowest normal MAE and the lowest disturbance MAE value of 13.42 and 17.96 respectively and the highest robustness index of 0.874 showing the best performance.

To assess the impact of key hyperparameters on model performance, sensitivity analyses were conducted on the input window, number of attention heads, hidden layer dimension, and robust loss weight; the results are shown in Table 6.

As shown in Table 6, when the input window was increased from 6 steps to 12 steps, the MAE decreased from 15.08 to 13.42, indicating that 60 minutes of historical data is more effective at capturing traffic changes; however, when the window was further increased to 15 steps, the MAE rose back to 13.59, suggesting that redundant information caused interference. The RMSE is lowest (20.74) when the number of attention heads is set to 4. When the hidden layer dimension is increased from 64 to 128, the MAE decreases only slightly from 13.42 to 13.37, but the training time per epoch increases from 24.9 s to 33.5 s. A robust weight of $\beta = 0.4$ yields optimal results, balancing accuracy and training efficiency.

VI. CONCLUSIONS

An understanding of the sequence of events in the past is not sufficient for traffic flow prediction when the surface of the roads is disturbed abruptly, but the topological propagation of the road network, the strength of the shocks, and the changes of spatiotemporal correlations all contribute. In order to solve this problem, a spatio-temporal graph attention network with strong spatial correlation, temporal evolution, disturbance factor characterization and robust constraints for traffic prediction is built, which has good stability in both normal traffic and sudden disturbance scenarios. It is innovative in that it applies dynamic graph attention to model the differences in the impact of road segments, it uses a disturbance sensing mechanism to capture the local impact and neighborhood diffusion of events like an accident, construction, weather, etc. and it has a strong loss function to reduce the error amplification due to abnormal fluctuations. However, the model is limited by the accuracy of disturbance event annotation, the extent of the data sources, and cross regional generalization capabilities; further work is still needed to characterize extreme congestion and compounding of disturbances, as well as to characterize long-term evolutionary trends. In the future, the model could be enriched with multi-source real-time event data, online update mechanisms, and cross-road-network transfer methods, further boosting the model's adaptability and applicability in complex traffic scenarios.

FUNDING

This work was supported by the Youth Program of the Natural Science Foundation of Hainan Province, China (No. 625QN349).

REFERENCES

[1] C. Zheng, X. Fan, C. Wang, et al., "Gman: A graph multi-attention network for traffic prediction," in Proc. AAAI Conf. Artif. Intell., vol. 34, no. 1, pp. 1234–1241, 2020.

[2] X. Wang, Y. Ma, Y. Wang, W. Jin, X. Wang, J. Tang, C. Jia, and J. Yu, "Traffic flow prediction via spatial temporal graph neural network," in Proc. Web Conf., pp. 1082–1092, 2020.

[3] C. Song, Y. Lin, S. Guo, and H. Wan, "Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting," in Proc. AAAI Conf. Artif. Intell., vol. 34, no. 1, pp. 914–921, 2020.

[4] L. Bai, L. Yao, C. Li, X. Wang, and C. Wang, "Adaptive graph convolutional recurrent network for traffic forecasting," in Advances in Neural Information Processing Systems, pp. 17804–17815, 2020.

[5] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, "Connecting the dots: Multivariate time series forecasting with graph neural networks," in Proc. 26th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining, pp. 753–763, 2020.

[6] M. Li and Z. Zhu, "Spatial-temporal fusion graph neural networks for traffic flow forecasting," in Proc. AAAI Conf. Artif. Intell., vol. 35, no. 5, pp. 4189–4196, 2021.

[7] X. Ta, Z. Liu, X. Hu, L. Yu, L. Sun, and B. Du, "Adaptive spatio-temporal graph neural network for traffic forecasting," Knowledge-Based Systems, vol. 242, Art. no. 108199, 2022.

[8] L. Chen, W. Shao, M. Lv, W. Chen, Y. Zhang, and C. Yang, "AARGNN: An attentive attributed recurrent graph neural network for traffic flow prediction considering multiple dynamic factors," IEEE Trans. Intell. Transp. Syst., vol. 23, no. 10, pp. 17201–17211, 2022.

[9] J. Huang, K. Luo, L. Cao, Y. Wen, and S. Zhong, "Learning multiaspect traffic couplings by multirelational graph attention networks for traffic prediction," IEEE Trans. Intell. Transp. Syst., vol. 23, no. 11, pp. 20681–20695, 2022.

[10] Z. Qiu, T. Zhu, Y. Jin, L. Sun, and B. Du, "A graph attention fusion network for event-driven traffic speed prediction," Information Sciences, vol. 622, pp. 405–423, 2023.

[11] W. Weng, J. Fan, H. Wu, Y. Hu, H. Tian, F. Zhu, and J. Wu, "A decomposition dynamic graph convolutional recurrent network for traffic forecasting," Pattern Recognition, vol. 142, Art. no. 109670, 2023.

[12] C. Wang, R. Tian, J. Hu, et al., "A trend graph attention network for traffic prediction," Information Sciences, vol. 623, pp. 275–292, 2023.

[13] G. Jin, L. Liu, F. Li, and J. Huang, "Spatio-temporal graph neural point process for traffic congestion event prediction," in Proc. AAAI Conf. Artif. Intell., vol. 37, no. 12, pp. 14268–14276, 2023.

[14] J. Feng, Y. Zhang, X. Piao, Y. Hu, and B. Yin, "Traffic anomaly detection based on spatio-temporal hypergraph convolution neural networks," Physica A: Statistical Mechanics and its Applications, vol. 646, Art. no. 129891, 2024.

[15] S. Yang, Q. Wu, and M. Li, "Decoupled multi-spatio-temporal fusion graph convolutional recurrent network for traffic prediction," Engineering Applications of Artificial Intelligence, vol. 163, Art. no. 112956, 2026.

[16] X. Yang, S. Tan, and P. Krähenbühl, "Long-term traffic simulation with interleaved autoregressive motion and scenario generation," in Proc. IEEE/CVF Int. Conf. Computer Vision, pp. 25305–25314, 2025.

[17] L. Feng, Q. Li, Z. Peng, et al., "Trafficgen: Learning to generate diverse and realistic traffic scenarios," arXiv preprint arXiv:2210.06609, 2022.

[18] Q. Li, Z. M. Peng, L. Feng, et al., "Scenarionet: Open-source platform for large-scale traffic scenario simulation and modeling," Advances in Neural Information Processing Systems, vol. 36, pp. 3894–3920, 2023.

[19] P. Cao, Y. Wang, and X. Liu, "Trajectory-based method for dividing lanes of vehicle trajectories collected by roadside LiDAR," Transportation Research Record, vol. 2678, no. 3, pp. 816–828, 2024.

[20] G. Reyes, R. Tolozano-Benites, L. Lanzarini, et al., "Method for the identification and classification of zones with vehicular congestion," ISPRS Int. J. Geo-Inf., vol. 13, no. 3, Art. no. 73, 2024.

[21] R. Jain, K. Okumura, M. Amir, et al., "Graph attention-guided search for dense multi-agent pathfinding," in Proc. AAAI Conf. Artif. Intell., vol. 40, no. 35, pp. 29504–29512, 2026.

[22] W. Lan, J. Liu, B. Yuan, et al., "AutoSceCraft: Generate various driving scenarios from scratch for autonomous driving systems," Tsinghua Science and Technology, vol. 31, no. 2, pp. 1282–1305, 2026.

[23] T. Phan, T. Phan, and S. Koenig, "Generative curricula for multi-agent path finding via unsupervised and reinforcement learning," Journal of Artificial Intelligence Research, vol. 82, pp. 2471–2534, 2025.

[24] G. Hess, J. Jaxing, E. Svensson, et al., "Masked autoencoder for self-supervised pre-training on lidar point clouds," in Proc. IEEE/CVF Winter Conf. Applications of Computer Vision, pp. 350–359, 2023.

[25] B. Chen, Z. Cao, and Q. Bai, "Satf: A scalable attentive transfer framework for efficient multiagent reinforcement learning," IEEE Trans. Neural Netw. Learn. Syst., vol. 36, no. 4, pp. 6627–6641, 2024.

[26] R. A. Alshamsi, I. M. Al Jawarneh, L. Foschini, et al., "ApproxGeoMap: An efficient system for generating approximate geo-maps from big geospatial data with quality of service guarantees," Computers, vol. 14, no. 2, Art. no. 35, 2025.