International Journal of Digital Earth

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjde20

Predicting urban signal-controlled intersection congestion events using spatio-temporal neural point process

Jianlong Wang, Xiaoqi Duan, Peixiao Wang, A.-Gen Qiu & Zeqiang Chen

To cite this article: Jianlong Wang, Xiaoqi Duan, Peixiao Wang, A.-Gen Qiu & Zeqiang Chen (2024) Predicting urban signal-controlled intersection congestion events using spatio-temporal neural point process, International Journal of Digital Earth, 17:1, 2376270, DOI: 10.1080/17538947.2024.2376270

To link to this article: https://doi.org/10.1080/17538947.2024.2376270

Published online: 30 Jul 2024.

INTERNATIONAL JOURNAL OF DIGITAL EARTH
2024, VOL. 17, NO. 1, 2376270
https://doi.org/10.1080/17538947.2024.2376270

Predicting urban signal-controlled intersection congestion events using spatio-temporal neural point process

Jianlong Wang [a], Xiaoqi Duan [b], Peixiao Wang [c], A.-Gen Qiu [d] and Zeqiang Chen [e]

[a]Changjiang Spatial Information Technology Engineering Co., Ltd. (Wuhan), Wuhan, 430010, People's Republic of China; [b]State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang, People's Republic of China; [c]State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Nature Resources Research, Chinese Academy of Sciences, Beijing, People's Republic of China; [d]Chinese Academy of Surverying and Mapping, Beijing, People's Republic of China; [e]National Engineering Research Center for Geographic Information System, China University of Geosciences, Wuhan, People's Republic of China

ABSTRACT

The urban traffic signal-controlled intersections are of great significance for solving the problem of urban road congestion. Previous research on congestion prediction mainly aggregated data at the level of road segments or traffic flow at a coarse regulated time interval. Fine-grained prediction of congestion events at the lane-level and cycle-level enables detailed a understanding of spatio-temporal dependencies, leading to congestion reduction, improved efficiency. This paper presents a Spatio-Temporal Neural Point Process (STNPP) model that combines Graph Neural Networks and Neural Temporal Point Process to predict congestion events at urban intersections. The proposed model allows for complete prediction of congestion events, including their occurrence, development, dissipation. In the process of spatial correlation modeling, graph neural networks are used to model the spatial relationships between both region and intersections. The current intersection and its upstream/downstream areas are modeled separately. To model the temporal correlations at individual intersections, we focus on a specific lane and capture the evolution of congestion events using the Neural Point Process Gated Recurrent Unit (NPPGRU), which captures the temporal granularity changes of signal-controlled cycles in congestion events. Using actual traffic speed and signal-controlled data from Hangzhou city, we validate that the proposed method achieves stable predictive performance.

ARTICLE HISTORY

Received 1 March 2024
Accepted 30 June 2024

KEYWORDS

Signal-controlled intersections; congestion prediction; spatio-temporal dependencies; temporal point process; congestion events

1. Introduction

According to the 'World Migration Report 2022' by the United Nations Migration Agency, approximately 55% of the world's population currently lives in urban areas. It is estimated that by 2050, the global urbanization rate will reach 70%, with the global population expected to increase from 7.7 billion to 9.7 billion. The rapid urbanization and increased vehicle ownership have created challenges in urban mobility and traffic management. Limited road resources and growing traffic demand have resulted in widespread urban traffic congestion, hindering the sustainable development of cities.

CONTACT Xiaoqi Duan duanxq@gzu.edu.cn

Urban signal-controlled intersections, as fundamental units constituting urban road networks, have always been focal points for congestion occurrences, making congestion prediction a hot topic in the field of intelligent transportation systems (ITS) (Jin, Liang, et al., 2023; Jin, Liu, et al., 2023; Li and Shahabi 2018; Wang 2023; Yu, Yin, and Zhu 2017). In recent years, research on the prediction of future traffic states has proliferated due to the powerful capabilities of modern machine learning methods in capturing complex spatio-temporal traffic dynamics and dependencies (Duan et al. 2023; Gong et al. 2024). Most of these studies aim to predict specific traffic parameters, such as traffic flow, speed, or travel time. Since variations in traffic parameters tend to exhibit regular spatial and temporal patterns, the evolution of congestion on signalized urban road networks is highly volatile, making congestion prediction a challenging task. Unlike the short-term regression-based continuous numerical prediction of traditional traffic parameters, predicting the development and changes of congestion, as a complete non-rigid spatio-temporal event (Kharaghani, Etemadfard, and Golmohammadi 2023), while predicting continuous traffic parameters is more regular and easily discoverable for data-driven models in terms of spatio-temporal dependencies in continuous time domains, modeling and predicting discrete-distributed congestion events are more relevant for the majority of the people in practical applications, and the accuracy of predicting traffic parameters under free-flow or uncongested conditions is less critical. An important characteristic of modeling congestion events is the ability to capture triggering effects (Zhang et al. 2024; Zhu et al. 2021). The content of congestion event prediction can be more comprehensive than predicting traffic parameters, including the occurrence and ending times of congestion events, congestion indices (e.g. congestion levels), propagation effects, subsequent congestion propagation areas, and starting points, among others.

Figure 1 provides a comparison between the temporal dual granularity distributions of congestion events and traditional traffic flow. The red bars represent the irregular temporal granularity of congestion events, with their heights indicating the congestion index. The blue bars represent the continuous regular temporal granularity of traffic flow, with numerical values indicating traffic flow parameters. Congestion events can be seen as 'disturbances' within continuous traffic flow, leading to the existence of two different temporal granularities within continuous and regular traffic flow.

This study investigates the problem of short-term traffic congestion event prediction for signalized road networks. Although some attempts have been made, significant challenges still remain:

(1) The traffic environment of signalized road networks is highly stochastic, where traffic congestion can be caused by a variety of complex factors such as spatio-temporal variation of traffic demand, road capacity, signal control effects, weather changes, accidents, and road maintenance. Explicitly accounting for these factors and their interactions can be challenging, making it difficult to build a comprehensive prediction model.
(2) Traffic congestion event is a complex spatio-temporal process characterized by constantly changing spatial coverage and temporal extent. Even at the same location, congestion events

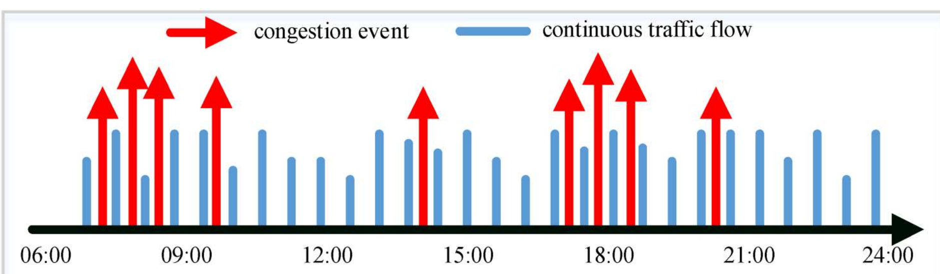


Figure 1. Continuous traffic flow and congestion events.

occurring at different times may exhibit distinct evolutionary patterns and durations, often influenced by spillover effects. Predicting the precise spatio-temporal extent of congestion events over their duration remains a challenging task.

(3) Fine-grained prediction of congestion events at the lane-level and signal light cycle-level in urban traffic signal-controlled intersections presents a profoundly challenging task. Congestion events in urban traffic are dynamic and non-linear. They can be influenced by factors such as traffic volume, road conditions, weather, and driver behavior. Modeling and predicting congestion accurately at the lane-level and signal light cycle-level requires accounting for these complex interactions and capturing the temporal dynamics.

In light of the above issues, we propose a STNPP model for predicting fine-grained congestion events at the lane-level and signal light cycle-level in urban traffic signal-controlled intersections. Traffic congestion is considered as a non-rigorous spatio-temporal extreme event and an event-oriented modeling method is developed. The proposed approach extends the point process model with the ability to emulate the evolution of congestion using a congestion embedding network that learns expressive spatio-temporal embeddings of historical congestion events with specially designed spatial and temporal graph convolutions, respectively. These embeddings encode the spatial and temporal evolutionary patterns of congestion, which are then integrated with the conditional intensity function of marked spatio-temporal processes to predict congestion events severities. The highlights of this work can be summarized as follows:

(1) A STNPP model is developed for fine-grained congestion event prediction, being capable of predicting each congestion event severity at the lane-level and signal light cycle-level within signal-controlled intersections with end-to-end deep learning framework;
(2) A dual temporal granularity pattern-aware NPPGRU is developed to generate temporal embeddings of historical congestion events based on traffic congestion graphs at a individual signal-controlled intersection;
(3) Extensive experiments have been conducted on real-world datasets from intelligent traffic signal-controlled intersections in Hangzhou, China. The experiments predict the complete cycle of congestion events, including their occurrence, development, and dissipation at a fine-grained spatio-temporal scale.

2. Related work

Deep Learning Congestion Prediction Model, Temporal Point Process Model, and Integration of Physical Models with Deep Learning Models.

2.1. Deep learning congestion prediction model

In the past few years, with the development of big data technology, many scholars have applied theories and methods related to big data in congestion prediction (Chahal et al. 2023; Sharma 2023; Tseng et al. 2018) and continuously improved and innovated to improve prediction accuracy. Graph Convolutional Neural Networks (GCNs) have been widely used to model urban traffic congestion prediction problems (Jin et al. 2022; Xiao et al. 2017). Currently, there are two main streams of GCNs, mainly based on spatial domain methods and spectral domain methods. Spatial domain methods attempt to perform convolutional filtering on graph nodes and their neighbors. Defferrard (2016) introduced a spectral domain method based on the graph Laplacian operator, considering spatial location on graph vertices and significantly reducing computational complexity. Furthermore, in modeling temporal correlations for congestion, Recurrent Neural Networks (RNNs) (Abdullah et al. 2023) and their variants (Huang, Wang, and Chao 2020) have been widely used to model time series patterns in traffic congestion data. Moreover, many neural network models have been proposed for predicting congestion based on graph convolutional neural networks and recurrent neural network architectures, addressing traffic status parameters such as traffic flow, speed, density, estimated time of arrival (ETA), and more (Chen, Yu, and Liu 2018; Zafar and Ul Haq 2020). However, most models have not directly modeled congestion events themselves. Given that the spatio-temporal patterns of congestion events are different from traffic flow, the lack of congestion propagation patterns to guide model learning and training limits the model's generalization ability.

2.2. Temporal point process model

The spatio-temporal dynamics of urban congestion exhibit strong randomness and are influenced by both other spatial locations and their own historical records, making it suitable for event-based descriptions. In previous research, Temporal Point Process (TPP) models (Daley and Vere-Jones 2008) have been widely used to predict the occurrence times of discrete events. Typically, traditional TPP models are based on a set of observed historical time-series events, modeling historical temporal information of event point processes by conditional intensity functions to obtain probability distributions on variable-length sequences within a time interval [0, T]. Du further introduced a general mathematical framework for Marked Temporal Point Processes (MTPPs) (Du et al. 2016), which can predict simultaneously event times and value information simultaneously, allowing the modeling of time-series event data with covariates. For example, in seismology, MTPPs were initially widely used to simulate earthquakes and aftershocks (Du et al., 2021). MTPPs can be represented as an event sequence $E = \{(t_1, m_1), (t_2, m_2), \cdots, (t_n, m_n)\}$, where N represents the random variable for event occurrence counts, and $0 < t_1 < t_2 < \cdots < t_n$ represents the event occurrence times, with $m_i \in \mathcal{M}$ representing event values. Typically, $\mathcal{M}$ uses categorical values, such as $\mathcal{M} = \{1, 2, \cdots K\}$, but other continuous value spaces (Chauhan et al. 2021) $\mathcal{M} \in \mathcal{R}^D$ can also be selected. In this context, the congestion index is used as the event value and defined in the continuous value space domain $\mathcal{M}$. However, parameter-based temporal point processes often pre-specify the occurrence of congestion events to follow a certain prior distribution. Relative to the underlying dynamics, the stochastic variability of real congestion can lead to underfitting.

2.3. Integrating physical models with deep learning models

Recently, some research efforts have begun to integrate statistical learning models or physical mechanism models with deep learning models, using statistical priors to guide deep learning models and improve their generalization to congestion (De Bézenac, Pajot, and Gallinari 2019; Saha, Dash, and Mukhopadhyay 2021; Willard et al. 2020). Given the strong randomness in the occurrence and spatio-temporal propagation of urban congestion, congestion can be seen as a stochastic event. Recent work has used temporal point process models to model congestion events, leading to the development of various efficient architectures and applications of Neural Temporal Point Processes (NTPPs) (Du et al. 2021; Jin, Liang et al. 2023; Jin, Liu, et al. 2023; Wu et al. 2020; Zhu et al. 2021). In comparison, NTPPs models can learn more complex dependencies and are often computationally more efficient than similar classical models. However, it is a challenge to effectively integrate physical models with deep learning models. Suitable structures and algorithms need to be designed to ensure that the two models work together effectively and to avoid information loss during the integration process.

In conclusion, inspired by the integration of deep learning with traditional temporal point process models (Du et al. 2016; Omi and Aihara 2019; Zhang et al. 2024), this paper proposes a neural temporal point process model for predicting congestion events at the lane level and signal cycle scale in urban intersections.

3. Spatio-temporal point process neural network for urban road intersection congestion prediction

3.1. Formal definition of urban road intersection congestion prediction

Definition 1. Regional Road Network Graph Structure. The regional road network graph structure, denoted as $\boldsymbol{G_g} = (\boldsymbol{V_g}, \boldsymbol{E_g}, \boldsymbol{A_g})$, represents the global road network configuration. In this representation, $\boldsymbol{V_g}$ signifies the set of road intersections intersections during the t-th signal cycle, and $\boldsymbol{\chi_g} = (\boldsymbol{X}_1^g, \cdots, \boldsymbol{X}_P^g)^T \in R^{N \times P \times F_g}$ encapsulates all historical features for N intersections over the past P signal cycles.

Definition 2. Intersection Graph Structure. The intersection graph structure, denoted as $\boldsymbol{G_l} = (\boldsymbol{V_l}, \boldsymbol{E_l}, \boldsymbol{A_l})$, represents the specific intersection's graph configuration. In this representation, $\boldsymbol{V_l}$ denotes all entry and exit lanes of the particular intersection ($|\boldsymbol{V_l}| = M$),$\boldsymbol{E_l}$ represents the connections between lanes, and $\boldsymbol{A_l} \in R^{M \times M}$ signifies the adjacency matrix of the graph $\boldsymbol{G_l}$. For the j-th lane of the i-th intersection during the t-th signal cycle, $\boldsymbol{x}_{i,j,t}^l \in R^{F_l}$ represents the traffic state feature vector, where $\boldsymbol{F_l}$ represents the number of features characterizing the intersection-level traffic state. Similarly, $\boldsymbol{X}_t^l = (\boldsymbol{x}_{i,1,t}^l, \ldots, \boldsymbol{x}_{i,M,t}^l)^T \in R^{M \times F_l}$ includes all features for all lanes during the t-th signal cycle, and $\boldsymbol{\chi_l} = (\boldsymbol{X}_1^l, \cdots, \boldsymbol{X}_P^l)^{\boldsymbol{T}} \in R^{M \times P \times F_l}$ covers all historical features for M lanes over the past P signal cycles.

Definition 3. Urban Intersection-Level Congestion Events. Urban intersection-level congestion events refer to a collection of specific intersection's congestion occurrence times and associated marked values. Based on the city's specified speed limit regulations for particular road intersections, any traffic state at a given moment with a speed lower than the standard speed is defined as a congestion event. The essential information for each event includes its occurrence timestamp (time) and speed (value). For instance, the congestion event set for the j-th lane of the i-th intersection is $\{t_c, e_{i,j}^{t_c}\}_{c=1}^{m_j}$. When there is no congestion event at a particular moment, it is represented as $\{0, 0\}$. Here, $t_c \in (0, P]$ denotes the timestamp of the c-th congestion event occurrence. Therefore, the historical time features input consists of a continuous time sequence formed by P event timestamps t_c, where $t_c/0$.

Distinct from historical input timestamps, the future occurrence times of congestion predicted in this study are represented discretely as $\{0, 1\}$, where 0 and 1 indicates the absence and presence of congestion at the current moment. $e_{i,j}^{t_c}$ represents the value information of this congestion event (e.g. congestion level, speed, and other traffic state features; in this paper, speed is referred to). When there is no congestion at a particular moment, the value is filled with 0. $m_j \in [1, 2, \cdots, \mathrm{K}]$, $\mathrm{K} \in Z^+$, represents the cumulative count of congestion events occurring within the historical time interval (0, P]. Consequently, within the historical interval (0, P], for all congestion event sets across the M lanes of a fixed i-th intersection: $\boldsymbol{E}_{0 \to P} = \cup_{j=1}^{M} \{t_c, \boldsymbol{e}_{i,j}^{t_c}\}_{c=1}^{m_j}$, with $t_c \in (0, P]$. To formulate the traffic forecasting problem, the main notations are summarized in Table 1.

This paper considers both the topological structures of the regional road network graph and intersection graph in congestion prediction. Given that traffic congestion patterns exhibit spatial heterogeneity across different intersections, the urban road network congestion prediction problem can be decomposed into subproblems for individual intersections. Since this study aims to simultaneously predict the occurrence times and values of congestion events, based on the symbols mentioned above, the congestion event prediction for the i-th intersection is defined as a multi-task learning problem, as shown in Equation (1):

$$\boldsymbol{E}_{P \to P+T_P} = f(\boldsymbol{G}_g, \boldsymbol{G}_l, \boldsymbol{\chi}_g, \boldsymbol{\chi}_l, \boldsymbol{E}_{0 \to P}, TP) \tag{1}$$

TP signifies the number of signal light cycles that are to be predicted. This parameter is used to

Table 1. Math symbols and descriptions.

Symbol	Domain	Description
$\boldsymbol{G_g}/\boldsymbol{G_l}$	$R^{N\times N}/R^{M\times M}$	The global/local road network/ intersection graph structure
$\boldsymbol{V_g}/\boldsymbol{V_l}$	N/M	The number of intersections /lanes
$\boldsymbol{E_g}/\boldsymbol{E_l}$	N/M	The collection of road segments
$\boldsymbol{A_g}/\boldsymbol{A_l}$	$R^{N\times N}/R^{M\times M}$	The adjacency matrix of the graph
$\boldsymbol{x}^g_{i,t}/\boldsymbol{x}^l_{i,j,t}$	R^{F_g}/R^{F_l}	The traffic state feature vector
$\boldsymbol{X}^g_t/\boldsymbol{X}^l_t$	$R^{N\times F_g}/R^{M\times F_l}$	Features for all intersections/lanes
$\boldsymbol{\chi_g}/\chi_l$	$R^{N\times P\times F_g}/R^{M\times P\times F_l}$	Historical features for N intersections/ M lanes over the past P signal cycles
t_c	$(0, P]$	The timestamp of the c-th congestion event occurrence
$e^{t_c}_{i,j}$	R	The marked value information of this congestion event
m_j	$[1, 2, \cdots, K]$	The cumulative count of congestion events occurring within the historical time interval $(0, P]$
$\boldsymbol{X}^g_l/\boldsymbol{X}^g_t$	$R^{N\times F_g}$.	Input features
$\boldsymbol{\Lambda}^g$	$R^{N\times N}$	Diagonal matrix of the regional road network graph
$\boldsymbol{A}/\boldsymbol{A_s}$	$R^{M\times M}/R^{M\times M\times P}$	The strength of correlations between lanes within each time slice of a signal cycle
$\boldsymbol{W}^l_{t,f}/\boldsymbol{W}^l_f/\boldsymbol{b_l}$	$R^{P\times(F_l+2)}/R^{F_l+2}/R^{M\times M}$	Network learning parameters
$\boldsymbol{M_{t_c}}/\boldsymbol{\Delta_{t_c}}\ /\boldsymbol{N_{t_c}}$	R	A masking / time interval of adjacent consecutive events/ cumulative vector
$[\odot]$	/	Element-wise multiplication
$[*]$	/	Matrix multiplication
$["\vert"]$	/	The concatenation symbol

determine the prediction horizon for forecasting traffic conditions and congestion events at a traffic signal intersection or network. T_P represents the future time span over which predictions regarding signal timing and traffic behavior are sought.

3.2. Urban intersection fine-Grained congestion prediction model

In this section, we introduce the STNPP model, which aims to provide fine-grained predictions of the occurrence times and values of traffic congestion events (e.g. speed or congestion index) at the signal cycle and lane levels of traffic signal-controlled intersections. Figure 2 illustrates the basic structure of the model, where F_g and F_l represent the number of traffic state features at the regional and intersection levels, respectively. T denotes the number of signal cycles, while N and M represent the counts of regional intersections and intersection lanes, respectively. The model takes into consideration both external and internal factors. External factors refer to the spatial influences of the regional road network and the temporal variations caused by signal cycle changes. Internal factors encompass the characteristics of congestion events between intersection lanes and across signal cycles, as well as the dynamics of traffic flow.

The model consists of two main components: the Spatial Correlation Module and the Dual-Granularity Temporal Correlation Module. Firstly, it uses the original spatial–temporal traffic data input, derived from the regional road network and local intersection graph structure, to learn the complex spatial dynamics among intersections through spatial correlation modeling. Then, the output of the spatial module is fed into the sequence-to-sequence module of the dual-granularity temporal correlation.

Inspired by the integration of deep learning and traditional temporal point process models, the model is extendedto include a NPPGRU. The NPPGRU is designed to fuse features from both the 'continuous' traffic flow sequence with regular temporalgranularity and the 'discrete' irregularly temporalcongestion event intervals. It models the discrete congestion event time correlations using traditional intensity functions while stacking NPPGRU units to model continuous traffic flow parameter sequences. The integrated NPPGRU not only harnesses the nonlinear modeling capabilities of the original GRU units but also improves the prediction capabilities of traditional temporal point process models for congestion events. This allowsthe prediction model to be more sensitive to 'macro' trends in event time granularity and 'micro' variations in traffic flow time granularity. The final encoding output from the last layer is used as input to the decoder

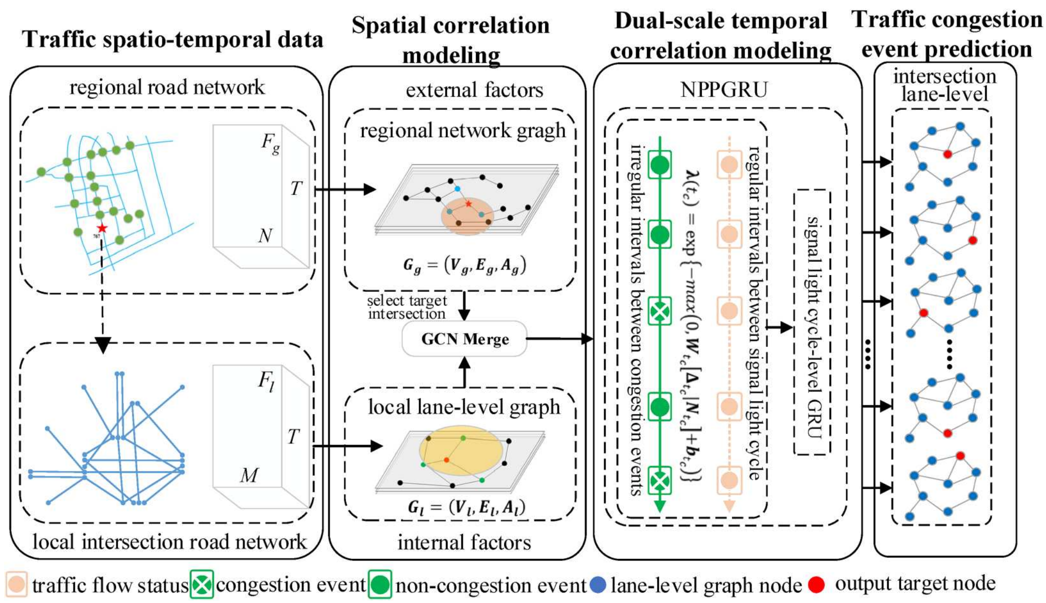


Figure 2. The framework for urban signal-controlled intersections congestion event prediction.

structure, allowing the model to predict the occurrence times and values of traffic congestion events for multiple future signal cycles.

3.3. Modeling spatial correlations of urban intersection-related factors

As shown in Figure 2, during the spatial correlation modeling phase, intersection congestion problems are influenced by both internal and external factors. Internal factors include the combined effects of congestion event features represented by the set $\boldsymbol{E}_{0\rightarrow P}$ and traffic flow characteristics denoted as $\boldsymbol{\chi}_{\boldsymbol{g}}$. External factors primarily encompass the dynamic variations in signal light strategies and the impact of congestion patterns within the regional road network on local intersections.

(1) Modeling global spatial correlations of regional lever factors

In the context of regional spatial dependencies, the congestion patterns at local intersections are influenced by the overall traffic patterns at the regional level. This results in intricate spatial and dynamic correlations in traffic congestion. Therefore, a regional road network graph structure, denoted as $\boldsymbol{G_g} = (\boldsymbol{V_g}, \boldsymbol{E_g}, \boldsymbol{A_g})$, is established with intersections as nodes, focusing on the target intersection as the central node. We aggregate the congestion patterns from multiple adjacent spatial intersections upstream and downstream of a target intersection, in order to consider the impact of regional congestion on that particular intersection.

Taking advantage of graph convolutional networks (Defferrard, Bresson, and Vandergheynst 2016), for each signal cycle, the intersection serves as the basic unit of the regional road network graph structure. Graph convolutional networks are employed to model the spatial dependencies between$N$ intersections in the regional road network, using the input features represented by $\boldsymbol{X}_l^g = \boldsymbol{X}_t^g \in R^{N\times F_g}$. The objective is to capture the global spatial correlations among the regional intersections,, taking into account the influence of external factors.

$$\boldsymbol{X}_{l+1}^g = ReLU(\boldsymbol{U}^g\boldsymbol{g}(\boldsymbol{\Lambda}^g)(\boldsymbol{U}^g)^T\boldsymbol{X}_l^g) \tag{2}$$

In equation (2), $X_{l+1}^g, X_l^g \in R^{N\times F_g}$, represent the input data blocks of the l-th and $(l+1)$-th layers of the regional road network graph convolution. U^g is an orthogonal matrix, while $g(\cdot)$ denotes a polynomial kernel function applied to the diagonal matrix $\Lambda^g \in R^{N\times N}$. For the regional road network graph G_g, its Laplacian matrix L can be decomposed as $U^g g(\Lambda^g)(U^g)^T$.

By applying the convolution equation (2), spatial structural information is obtained among the embedded regional intersections. This information is then used to calculate the historical data over P signal cycles for the entire regional road network, resulting in $X^g \in R^{P\times N\times F_g}$. Then, through an index query, the index of the target congestion intersection$n$is selected from the$N$intersections, and its corresponding features are represented as $X^{g(n)} \in R^{P\times F_g}$, further yielding the regional road network impact features over P signal cycles for all lanes of the target congestion intersection, denoted as $Z^g \in R^{P\times M\times F_g}$. These features serve as the spatio-temporal input for the local intersection, aiding in modeling the internal factors influencing congestion.

(2) Modeling the local impact of external signal-control strategies factors at local intersection

To account for external factors affecting congestion, we integrated the environmental context of signal-controlled intersections with the modeling of variations in lane-level external signal-controlled strategies in the spatial domain. Research has shown that modeling external factors as an adaptive matrix can significantly capture hidden dynamic spatial dependencies in specific traffic scenarios (Diao et al. 2019; Hu et al. 2019). Therefore, for a given intersection in a traffic network comprising M lanes, by considering the dynamic changes in signal-controlled strategies, we constructed an external signal-controlled perceptual transition matrix, denoted as $A \in R^{M\times M}$, representing the strength of correlations between lanes within each time slice of a signal cycle. Furthermore, the obtained traffic history over a period of P time intervals is accumulated to form $A_s \in R^{M\times M\times P}$. This matrix quantifies the lane-to-lane dependencies in different time slices based on the different sizes of signal cycles between upstream and downstream lanes.

However, due to the inherent sparsity of congestion events themselves, relying solely on external event features to model spatial dependencies cannot fully reflect the true dynamics of congestion, and is insufficient in capturing genuine inter-lane dependencies. Therefore, it is necessary to consider modeling the internal factors that play a critical role in traffic congestion.

(3) Modeling local lane-level spatial correlations at local intersections

When dealing with local intersections, establishing an explicit model to represent the influence of internal factors on lane-level spatial modeling is a complex task. Therefore, we employed attention mechanism to jointly consider the congestion-related internal event features set, denoted as $E_{0\to P}$, and the traffic flow feature χ_ι. This enables the calculation of lane assignment weights that reflect the combined impact of internal factors on the current congested lanes. Within the intersection, an adaptive spatial attention mechanism (Fang et al. 2021) is used to dynamically learn the internal spatial associations of congestion events.

Taking a specific intersection in a regional road network as an example, suppose we have observed traffic congestion data for P historical signal cycles. The traffic flow χ_l for M lanes at the intersection and the events $E_{0\to P}$ are provided as inputs to neural network layers. Within each signal cycle, we compute the weighted matrix for the inputs:

$$A_l = \sigma[((\chi_l|E_{0\to P})W_t^l)W_{t,f}^l(W_f^l(\chi_l|E_{0\to P}))^T + b_l] \tag{3}$$

Where σ represents the activation function, '|' denotes concatenation, $W_t^l \in R^P$, $W_{t,f}^l \in R^{P\times(F_l+2)}$, $W_f^l \in R^{F_l+2}$, and $b_l \in R^{M\times M}$ are the network learning parameters. The input features for congestion events consist of event time and a label value.

For the i-th congested lane, its spatial correlation $\alpha_{i,j}$ with adjacent and upstream/downstream lanes can be calculated based on $\boldsymbol{A_l}$ as described in equation (4):

$$\alpha_{i,j} = \frac{exp((\boldsymbol{A_l})_{i,j})}{\sum_{k=1}^{M} \exp((\boldsymbol{A_l})_{i,k})} \tag{4}$$

Where j is the index of the lane associated with the i-th lane, and $\alpha_{i,j}$ represents the spatial correlation strength between j-lane and i-lane within the intersection.

At local signal-controlled intersections, there are geometric topological associations between lanes, and this spatial and physical correlation is crucial for the transmission of congestion information across fine-grained lanes. Similar to the structure of regional road networks, a graph convolutional network (GCN) is used to model the structure of local intersections. Traditional graph convolutional networks are combined with the aforementioned attention mechanism and signal control perception transition matrix to adaptively capture dynamic spatial dependencies between lanes at the current intersection and lanes at upstream and downstream intersections. Within each signal cycle slice, a graph convolutional operator is used to capture spatial correlations between intersection lanes. The equation (5) describes this operation:

$$\boldsymbol{X_{l+1}} = ReLU(\boldsymbol{U}\boldsymbol{g}(\boldsymbol{\Lambda})\boldsymbol{U}^T\boldsymbol{X_l}) \tag{5}$$

Where, $\boldsymbol{X_{l+1}}$ and $\boldsymbol{X_l}$ represent the l-th and $(l+1)$-th layers of input data blocks for the graph convolutional network. $\boldsymbol{X_l} = (\boldsymbol{X}_t^l|\boldsymbol{Z}_t^g) \in R^{M\times(F_l+F_g)}$ (Note: historical event data is used as spatial attention input for calculating lane-to-lane spatial influences, but it is not used as an input layer for spatial convolution). $\boldsymbol{Z}_t^g \in R^{M\times F_g}$ represents the influence of the global regional road network at time t. $\boldsymbol{U}$ is an orthogonal matrix. $\boldsymbol{g}(\cdot)$is a polynomial kernel function applied to the diagonal matrix $\boldsymbol{\Lambda} \in R^{M\times M}$. For local intersection-level graph structures, the graph Laplacian matrix $\boldsymbol{L}$ can be decomposed into $\boldsymbol{U\Lambda U}^T$. The ReLU activation function introduces non-linearity. To facilitate computation for large-scale graphs, Chebyshev polynomials are used to approximate graph operations (Simonovsky and Komodakis 2017).

To capture dynamic lane-level spatial correlations between adjacent intersections, the spatial attention weights calculated in equation (4) and the signal-controlled perceptual transition matrix $\boldsymbol{A} \in R^{M\times M}$ are integrated into the graph convolutional network as described in equation (6):

$$\boldsymbol{X_{l+1}} = ReLU\left\{\left(\sum_{k=1}^{K-1}[\boldsymbol{\beta_k}(C_k(\tilde{\boldsymbol{L}})]] \odot (\boldsymbol{A_l}|\boldsymbol{A})\right)\boldsymbol{X_l}\right\} \tag{6}$$

Where $\boldsymbol{\beta_k}$ is a vector of polynomial coefficients. K is the truncation order. C_k is Chebyshev polynomial functions recursively computed on $\tilde{\boldsymbol{L}}$, $\tilde{\boldsymbol{L}} = \frac{2}{\lambda_{max}}\boldsymbol{L} - \boldsymbol{I_M}$ (where λ_{max} is the maximum eigenvalue of $\boldsymbol{L}$, and $\boldsymbol{I_M}$ is an identity matrix). $\odot$ represents the Hadamard product operator, i.e. element-wise multiplication.

This integrated design has two main advantages: (1) it enhances the sensitivity of the graph convolutional network to spatial correlations between upstream and downstream lanes by explicitly considering the impact of external signal control strategies on the spatial modeling; (2) it allows the model to measure the interactive influence between lanes at intersections, improving the model's interpretability.

3.4. Modeling traffic congestion at urban intersections with pattern-aware neural point process gated recurrent unit

After capturing dynamic lane-level correlations between adjacent intersections in spatial modeling, our focus shifts to temporal modeling. Specifically, we investigated the time-related correlations of congestion for a specific lane between the current observation and its neighboring signal cycles. Considering that traffic congestion is influenced by multiple temporal granularities, we divide the traffic time series into dual temporal granularity: an 'irregular' time granularity based on congestion event intervals and a 'continuous' regular time granularity based on signal cycles.

Since congestion events occur at different timestamps, the time gap between the durations of two congestion events is not fixed. Therefore, it is crucial to effectively utilize irregular delays between congestion event timestamps. Past some studies have begun integrating deep learning with traditional temporal point process models (Du et al. 2016; Shchur et al. 2021; Zhang et al. 2024).

A dual temporal granularity pattern-aware NPPGRU is developed to generate temporal embeddings of historical congestion events, which modeled discrete congestion events temporal correlation using traditional intensity functions and stacked GRU units to model continuous traffic flow parameter sequences. Figure 3 illustrates the differences between NPPGRU and traditional GRU units. The integrated NPPGRU units not only harness the nonlinear time modeling capabilities of the original GRU units but also enhance the predictive power of traditional point process models. This allows the model to handle not only the global trends of congestion events at the macro-event time granularity but also to capture finer details and variations of congestion at the micro-traffic flow time granularity.

For a specific lane m at intersection n, given a collection of discrete event lists $\{t_c, e^{t_c}_{n,m}\}^{m_j}_{c=1}$ occurring within a historical time window P, as shown in Table 2: Let $t_c \in (0, P]$ represent the timestamp when the c-th congestion event occurred, assuming the first observation occurred at timestamp 0. We introduce a masking vector $\boldsymbol{M}_{t_c}$ to indicate whether a congestion event occurred at time

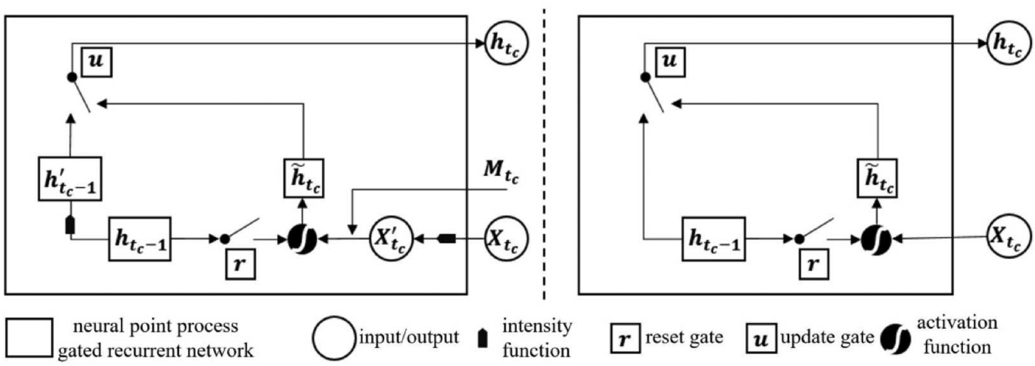


Figure 3. Comparison between the proposed Neural Point Processed Gated Recurrent Unit and the traditional Gated Recurrent Unit.

Table 2. Vector case $\boldsymbol{X}$ of congestion events (where '/' represents not occurring, and C represents occurring) Time stamp $\boldsymbol{t_c}$ of congestion events, mask vector $\boldsymbol{M}_{t_c}$, time interval $\boldsymbol{\Delta}_{t_c}$ of adjacent consecutive congestion events, cumulative vector $\boldsymbol{N}_{t_c}$ of congestion events.

$\boldsymbol{X}$	/	C	C	/	/	/	C	/	C
$\boldsymbol{M}_{t_c}$	0	1	1	0	0	0	1	0	1
$\boldsymbol{t_c}$	0	1	5	9	12	15	18	25	36
$\boldsymbol{\Delta}_{t_c}$	0	1	4	4	7	10	13	7	18
$\boldsymbol{N}_{t_c}$	0	1	2	2	2	2	3	3	4

t_c, with values being 1/0, respectively. Specifically:

$$\boldsymbol{M}_{t_c} = \begin{cases} 1, \ Congestion\ occur \\ 0, \ Congestion\ does\ not\ occur \end{cases} \tag{7}$$

$$\boldsymbol{\Delta}_{t_c} = \begin{cases} t_c - t_{c-1} + \Delta_{t_{c-1}}, \ t_c > 1, \ M(t_{c-1}) = 0 \\ t_c - t_{c-1}, \ t_c > 1, \ M(t_{c-1}) = 1 \\ 0, \ t_c = 1 \end{cases} \tag{8}$$

Inspired by the intensity function modeling of event history influenced by Poisson processes, in order to capture time-related patterns of congestion events under dual temporal granularity, this study introduces an adaptively learnable prior intensity function model (Equation 9). This model is then integrated into the traditional gated recurrent neural network:

$$\boldsymbol{\lambda}(t_c) = \exp\{-\max(0, \boldsymbol{W}_{t_c}[\boldsymbol{\Delta}_{t_c}|\boldsymbol{N}_{t_c}] + \boldsymbol{b}_{t_c})\} \tag{9}$$

Where ["|"] represents the concatenation symbol, and $\boldsymbol{W}_{t_c}$ and $\boldsymbol{b}_{t_c}$ are model learning parameters. The selection of the external function $\boldsymbol{\lambda}(\cdot)$ needs to consider two key criteria: (1) the intensity function should be positive, and (2) as mentioned earlier, the intensity function should vary with the time intervals between congestion events and the cumulative count of events.

After obtaining the result from Equation (9), to model the temporal correlations of historical congestion events and the cumulative congestion count, the original features of time $[\boldsymbol{X}_{t_c}|\boldsymbol{h}_{t_c-1}]$ are updated by element-wise multiplication with $\boldsymbol{\lambda}(\boldsymbol{t_c})$. This updated information is then combined with the GRU model, resulting in the NPPGRU.

The primary idea behind the NPPGRU is to introduce a time point process model tailored for congestion events. This allows for the dynamic extraction of event-related temporal patterns from traffic event data, rather than relying solely on data-driven modeling. Additionally, the masking vector $\boldsymbol{M}_{t_c}$, based on congestion events, is directly input into the model to assist in modeling the temporal correlations of events. Below is the update mechanism for NPPGRU:

$$[\boldsymbol{X}'_{t_c}; \boldsymbol{h}'_{t_c-1}] = \boldsymbol{\lambda}(\boldsymbol{t_c}) \odot [\boldsymbol{X}_{t_c}|\boldsymbol{h}_{t_c-1}] \tag{10}$$

$$\boldsymbol{r} = \sigma(\boldsymbol{W_r}[\boldsymbol{X}'_{t_c}|\boldsymbol{h}'_{t_c-1}] + \boldsymbol{U_r}\boldsymbol{M}_{t_c} + \boldsymbol{b_r}) \tag{11}$$

$$\boldsymbol{u} = \sigma(\boldsymbol{W_u}[\boldsymbol{X}'_{t_c}|\boldsymbol{h}'_{t_c-1}] + \boldsymbol{U_u}\boldsymbol{M}_{t_c} + \boldsymbol{b_u}) \tag{12}$$

$$\tilde{\boldsymbol{h}}_{t_c} = \tanh(\boldsymbol{W}_{\tilde{h}_{t_c}} * [\boldsymbol{r} * \boldsymbol{h}_{t_c-1}|\boldsymbol{X}_{t_c}] + \boldsymbol{U}\boldsymbol{M}_{t_c} + \boldsymbol{b}_{\tilde{h}_c}) \tag{13}$$

$$\boldsymbol{h}_{t_c} = (1 - \boldsymbol{u}) \odot \boldsymbol{h}'_{c-1} + \boldsymbol{u} \odot \tilde{\boldsymbol{h}}_{\boldsymbol{c}} \tag{14}$$

Where $\boldsymbol{u}$ represents the update gate, $\boldsymbol{r}$ denotes the reset gate, $\tilde{\boldsymbol{h}}_{t_c}$ signifies the candidate hidden variable. σ is the activation function (sigmoid). $\boldsymbol{W_r}$, $\boldsymbol{W_u}$, $\boldsymbol{U_r}$, $\boldsymbol{U_u}$, $\boldsymbol{U}$ and $\boldsymbol{W}_{\tilde{h}_c}$ are the learned weight parameters. $\boldsymbol{b_r}$, $\boldsymbol{b_u}$ and $\boldsymbol{b}_{\tilde{h}_t}$ represent bias terms. [$\odot$] represents element-wise multiplication. [$*$] denotes matrix multiplication.

The time correlation module is built upon a sequence-to-sequence framework structure. It learns the encoded representation of historical congestion events through recursive computations and generates future multi-step congestion event outputs by selecting the optimal parameters θ (Equation 15), based on conditional probabilities. As depicted in Figure 4, the input consists of a collection of discrete event timestamps $\{(t_1, y_1), (t_2, y_2), \cdots, (t_p, y_p)\}$, along with a continuous regular traffic flow parameter sequence$[X_1, X_2, \cdots, X_p]$. After encoding with stacked NPPGRU units, a hidden representation $\boldsymbol{C}$ capturing the historical sequence is obtained. Subsequently, through the decoding process, future multi-step event predictions for values and timestamps are

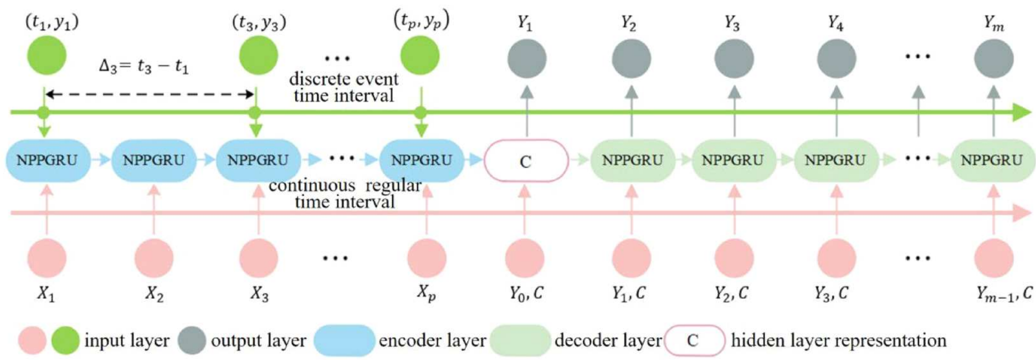


Figure 4. Dual time granularity neural point process gated recurrent Unit neural network structure.

generated as $[Y_1, Y_2, \ldots Y_m]$, where $Y_i = (t_i, y_i)$. This process computes whether congestion occurs and the congestion values at future time steps.

Algorithm 1. Training STNPP Model

Input: Intersection Graph Structure. $\boldsymbol{G_l} \in R^{M\times M}$; Regional Road Network Graph Structure. $\boldsymbol{G_g} \in R^{N\times N}$; Training traffic flow parameter feature set. $\{\boldsymbol{\chi_g}, \boldsymbol{\chi_l}\}$ and collection of traffic congestion events. $\{\boldsymbol{E}_{0\to P}, \boldsymbol{E}_{P\to P+T_P}\}$, batchsize b, Number of iterations K, number of historical and predicted signal cycles input P, T_P.
Output: Parameters of STNPP model $\boldsymbol{\theta}$.
// **Forward propagation calculation**
1: Randomly initialize $\boldsymbol{\theta}$
For $n = 1 \cdots K$ do:
2: Training sample input: Input b traffic flow spatio-temporal features $\{(\boldsymbol{G_g}, \boldsymbol{\chi_g}), (\boldsymbol{G_l}, \boldsymbol{\chi_l})\}$ and collection of congestion events each time $\{\boldsymbol{E}_{0\to P}, \boldsymbol{E}_{P\to P+T_P}\}$
3: Model spatial relation with single-layer GCN for event spatial correlation based on equation (6): $\{(\boldsymbol{G_g}, \boldsymbol{\chi_g}), (\boldsymbol{G_l}, \boldsymbol{\chi_l}, \boldsymbol{E}_{0\to P})\} \to \boldsymbol{X}_{l+1} \in R^{M\times P_l\times F_1}$
4: Model temporal relation of events using equations (10) – (14): $\boldsymbol{X}_{l+1} \to \boldsymbol{h}_{t_c} \in R^{M\times F_2}$
5: Model output: Predict if congestion events will occur in multiple time steps and the associated values: $\boldsymbol{h}_{t_p} \to \{\boldsymbol{E}_{P\to P+T_P}\}$
6: Calculate loss of congestion event time and associated values using equation (15):$\hat{\boldsymbol{\theta}}$
$= \underset{\theta}{\arcmin} 1/b \sum_{q=1}^{n} (\log P(\boldsymbol{e}_{i,j}^{t_q} | [\boldsymbol{\chi}_g, \boldsymbol{\chi}_l], \boldsymbol{E}_{0\to P}; \theta) + \log f(k_c | [\boldsymbol{\chi}_g, \boldsymbol{\chi}_l], \boldsymbol{E}_{0\to P}; \theta))$
// **Backward propagation calculation**
7: Update gradient $\nabla\hat{\boldsymbol{\theta}}$ with Adam backpropagation
8: Update model weight parameter with gradient: $\hat{\boldsymbol{\theta}} \leftarrow \hat{\boldsymbol{\theta}} + \nabla\hat{\boldsymbol{\theta}}$
9: Stop iteration.

3.5. Loss function and model training

Given a collection of irregular temporal traffic state data containing historical, regional-level, and intersection-level variables $\boldsymbol{\chi}_g$, and $\boldsymbol{\chi}_l$, as well as a discrete list of historical and future irregular congestion events denoted as $\boldsymbol{E}_{0\to P} = \bigcup_{j=1}^{M} \{t_c, \boldsymbol{e}_{i,j}^{t_c}\}_{c=1}^{m_j}$, where $t_c \in (0, P]$ and $\boldsymbol{E}_{P\to P+T_P}$. Here, $k_c = 1$ indicates the occurrence of congestion at the current moment, otherwise, it is 0. Model parameters are updated by maximizing the joint log-likelihood of $\boldsymbol{E}_{P\to P+T_P}$.

$$l(\theta) = \min_{\theta} \sum_{q=P}^{T_P} \log P(\boldsymbol{e}_{i,j}^{t_q} | [\boldsymbol{\chi}_g, \boldsymbol{\chi}_l], \boldsymbol{E}_{0\to P}; \theta) + \log f(k_c | [\boldsymbol{\chi}_g, \boldsymbol{\chi}_l], \boldsymbol{E}_{0\to P}; \theta) \tag{15}$$

For the training phase of the model, as Algorithm 1, input data includes traffic data collected from the regional road network and local intersections. For the regional road network graph $\boldsymbol{G_g} = (\boldsymbol{V_g}, \boldsymbol{E_g}, \boldsymbol{A_g})$, traffic features for all approach segments of each intersection are aggregated every 10 min, resulting in $\chi_g \in R^{N\times P\times F_g}$, where P represents data collection signal cycles, and F_g stands for the aggregated traffic features. $F_g = [v_{avg,}, v_{min}, v_{max}]$ for whether it is during peak hours (0/1), where $v_{avg,}, v_{min}, v_{max}$ represent average, minimum, and maximum speeds.

At the intersection level, a specific intersection graph $G_I = (V_I, E_I, A_I)$ is defined to capture the local topological relationships among lanes within that intersection. Each node in the graph corresponds to a lane. The lane-level local temporal traffic state data is represented as $\chi_l \in R^{M\times P\times F_l}$, where $F_l = [v_{i,avg,}, v_{i,min}, v_{i,max}, TP]$. Here, $v_{i,avg}, v_{i,min}, v_{i,max}$ represent the average, minimum, and maximum speeds at the i-th intersection over P cycles, respectively. TP refers to other relevant traffic parameters at the lane level.

4. Experimental evaluation and analysis

The proposed congestion prediction method was evaluated using actual traffic speed data from Hangzhou city and signal timing data from intersections. This study compared the proposed method with several common traffic congestion prediction methods and conducted ablation studies to demonstrate the effectiveness of each component in the prediction method. All experiments were conducted on a desktop server equipped with a 3.7 GHz Intel Core i7-8700 K processor, GeForce RTX®2080 Ti graphics card, and 32GB of memory.

4.1. Experimental data

The experimental data used in this study consisted of traffic speed data, which involved averaging the instantaneous speeds of vehicles at the lane level within each signal cycle. Similar to the definition of congestion events in (Zhu et al. 2021), this study determined the occurrence of traffic congestion events based on the latest industry standard in China, 'Road Traffic Congestion Evaluation Methods' (GAT 115-2020) issued by the Traffic Management Science Research Institute of the Ministry of Public Security in 2020.

According to this standard, the congestion status of a road is determined by comparing the speed limits and free-flow speeds at intersections. When the average speed of a lane at a given time falls within the congestion range, it is considered a congestion event; otherwise, it is classified as a non-congestion (free-flow) event. From the traffic dataset of Hangzhou city, 111,897 traffic congestion events were extracted for a 30-day period from December 1, 2018, to December 31, 2018. The maximum number of events in a single day was 4,235, while the minimum was 2,354. Additionally, geometric data describing lane counts, lane geometries, topological connections between lanes, and turning restrictions at each intersection were collected for the road network under investigation. Signal timing data for all intersections in the study area, including cycle lengths and effective green times, were also gathered. The road network studied consisted of 18 intersections and 154 interconnected lanes, as depicted in Figure 5.

4.2. Method comparison

This paper compared the proposed model with both deep learning spatio-temporal prediction methods and traditional methods to further validate the predictive capabilities. The compared models included:

4.2.1. Support vector regression (SVR)

SVR is known for its stable predictive ability on nonlinear time series data. SVR transforms input features into a high-dimensional feature space using a nonlinear function and then finds a linear function that accurately represents the relationship between input and output data.

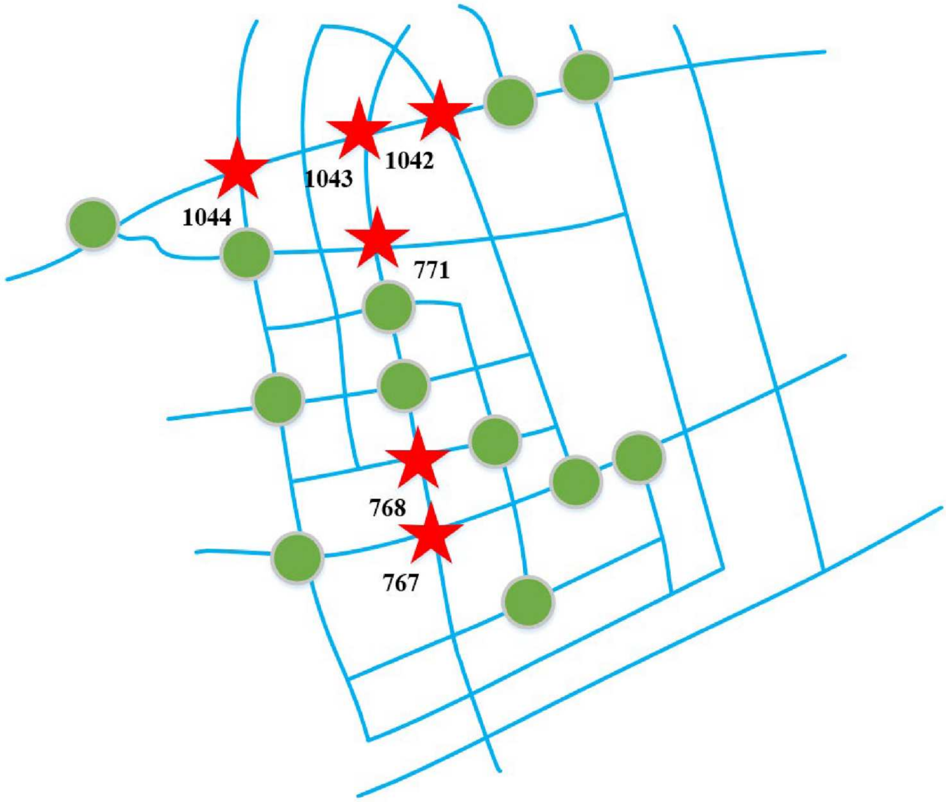


Figure 5. Road network structure in the experiment (Red stars represent the six multi-lane intersection locations to be evaluated).

4.2.2. Graph convolutional networks + attention (GCN + ATT)

The GCN model used in this test adopts a similar architecture for spatial correlation modeling as STNPP but does not consider temporal correlations within GCN. To ensure a fair comparison, the spatial attention mechanism from this paper was incorporated into the GCN structure.

4.2.3. Attention-Based spatial-temporal graph convolutional networks (Astgcn)

This model combines attention mechanisms with spatial–temporal graph convolution to model the spatio-temporal features among congestion events.

Additionally, two variant models of the proposed method were evaluated to assess the contributions of different designs (NPPGRU and spatial correlation modeling):

4.2.4. Variant 1 (S2S-GRU)

This variant is based on a Seq2Seq framework but does not consider spatial correlations. It retains the same architecture as the proposed method for time correlation modeling, validating the impact of remote spatial intersections.

4.2.5. Variant 2 (NP-STNPP)

This variant retains the spatial and temporal correlation modeling parts of STNPP but replaces the prior design of the time point process model (NPPGRU) with a conventional GRU.

4.2.6. Variant 3 (NR-STNPP)

This variant retains the local intersection spatial gragh and temporal correlation modeling components of STNPP, but removes regional spatial map structure and validates regional spatial dependencies, congestion patterns at local intersections are influenced by the regional traffic pattern.

4.3. Evaluation metric

Three performance metrics were used for evaluating the proposed congestion prediction approach: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE):

$$MAE_i(M, T_P) = \frac{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt}|y_{mt} - y'_{mt}|}{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt}} \tag{16}$$

$$RMSE_i(M, T_P) = \sqrt{\frac{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt}(y_{mt} - y'_{mt})^2}{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt}}} \tag{17}$$

$$MAPE_i(M, T_P) = 100 \times \frac{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt} \dfrac{|y_{mt} - y'_{mt}|}{y_{mt}}}{\sum_{m=1}^{M} \sum_{t=1}^{T_P} I_{mt}} \tag{18}$$

where i indicates the examined ith intersection, which consists of M lanes. The congestion prediction was performed over T_P signal cycles. y_{mt} is the ground truth congestion event value for lane i at signal cycle t, y'_{mt} is the congestion prediction value for y_{mt}. I_{mt} is an indicator function that returns '1' when y_{mt} occurs congestion event and '0' otherwise. The test results of these metrics were averaged over the selected six intersections.

4.4. Model details

The STNPP model is developed using the Apache Mxnet framework and is implemented in Python. The data is split into training, validation, and test datasets. In terms of time, 7 days of data are used for training, days 8 to 11 are used for validation, and the data for the last 21 days is used for testing. This means that one week of data is used for training, days 8 to 11 are used for validation, and the last three weeks of data are for testing. Optimal network parameters, including the number of network layers and hidden units, are determined based on the validation dataset. The spatial modeling employs a two-layer GCN structure. The number of hidden units in both GCN and NPPGRU is set to 64. The ADAM optimizer is used to train the network with a learning rate of 0.001. Training is performed for 100 epochs using mini-batches of size 64 and backpropagation of the loss function. Batch normalization is used to speed up the convergence of GCN and NPPGRU. All input data are normalized with a mean of 0 and a variance of 1. The Chebyshev polynomial truncation order (K) in Equation 6 is set to 3. The historical length (P) of the training input in terms of the number of signal cycles is set to 20. The model predicts the next 5 signal cycles because empirically, it empirically performs the best, indicating that NPPGRU captures the temporal correlations of traffic congestion events over 20 signal cycles. Once the total loss converges, the STNPP model is used to predict whether congestion will occur in future time steps and the congestion values.

4.5. Analysis and evaluation of experimental results

(1) Single Lane Congestion Prediction Accuracy Comparison

In the first set of experiments, the performance of STNPP is evaluated for predicting congestion in two single-lane scenarios. The prediction is made for future s signal cycles of 1/3/5, and the results are reported in Tables 3 and 4. It's observed that as the prediction horizon increases, the performance of all models decreases. The STNPP model has the lowest error in most cases and the prediction performance is more stable, NR-STNPP as a suboptimal model, performs closely to the proposed other variant model, but in some cases, (the third Prediction step size of southbound driving lane), ASTGCN gets better results. for single-lane congestion prediction. For predicting congestion in the southbound lane for the next 5 cycles, the MAE/RMSE/MAPE are reported as 2.60/3.51/8.44, and for the northbound lane: 2.29/2.97/6.51. This demonstrates that the design of the STNPP algorithm for modeling spatio-temporal correlations is better suited to the actual distribution pattern of congestion than other models.

Furthermore, deep learning-based spatio-temporal prediction models (GCN + ATT, ASTGCN, S2S-GRU, NP-STNPP, STNPP) generally outperform the machine learning SVR model, demonstrating that deep learning models have a stronger ability to capture spatio-temporal features. However, in Table 4, the prediction performance of GCN + ATT for the northbound lane is not as good as SVR. This is mainly due to the presence of strong temporal dependencies in congestion events. Ignoring temporal features and focusing solely on spatial modeling may not fit congestion distribution well. GCN + ATT, S2S-GRU, and NP-STNPP achieve the worst prediction results, mainly because they do not simultaneously model both time and spatial correlations and do not consider the influence of multiple temporal granularities in congestion events.

(2) Performance Comparison for Peak Hours

The study further evaluates the model's congestion prediction performance during peak hours. The same two single-lane scenarios from (1) are used, and the models that did not perform well in (1) are excluded. The MAPE (%) results for predicting the next 1/3/5 cycles are shown in Figure 6. STNPP, along with its two variants, achieves the best prediction accuracy among all deep learning models. This is mainly because the model in this study fully considers the local, specific intersection-level spatial lane-level and upstream-downstream correlations and considers the congestion

Table 3. Performance comparison of southbound driving models.

Prediction step size	evaluating indicator	SVR	GCN + ATT	ASTGCN	S2S-GRU	NP-STNPP	NR-STNPP	STNPP
1	MAE	4.02	2.88	1.55	2.30	1.60	1.50	**1.30**
	RMSE	4.29	3.67	2.10	3.12	2.21	2.02	**1.71**
	MAPE(%)	12.40	9.58	5.02	7.52	5.16	4.52	**4.27**
3	MAE	4.36	3.40	2.55	4.57	3.40	2.32	**2.09**
	RMSE	4.37	4.36	3.32	5.77	4.05	3.15	**2.73**
	MAPE(%)	14.70	11.39	8.31	14.95	10.55	7.86	**6.72**
5	MAE	5.25	4.61	3.05	5.13	4.07	2.82	**2.60**
	RMSE	5.59	5.60	4.01	6.46	4.58	3.71	**3.51**
	MAPE(%)	16.85	15.34	10.06	16.75	12.53	9.30	**8.44**

Table 4. Performance comparison of northbound income models.

Prediction step size	evaluating indicator	SVR	GCN + ATT	ASTGCN	S2S-GRU	NP-STNPP	NR-STNPP	STNPP
1	MAE	2.24	3.05	1.36	1.85	1.18	1.20	**1.17**
	RMSE	2.25	4.00	1.90	2.79	1.66	1.69	**1.59**
	MAPE(%)	6.61	8.28	3.82	5.22	3.41	3.48	**3.37**
3	MAE	2.78	3.06	**1.93**	3.42	4.45	2.58	2.15
	RMSE	2.91	3.97	**2.67**	5.02	5.01	3.28	2.79
	MAPE(%)	7.69	8.24	**5.56**	10.08	12.07	6.79	6.12
5	MAE	3.97	3.30	2.36	3.94	4.48	2.34	**2.29**
	RMSE	4.19	3.89	3.26	5.79	5.08	3.16	**2.97**
	MAPE(%)	10.73	8.11	6.86	11.49	12.16	6.72	**6.51**

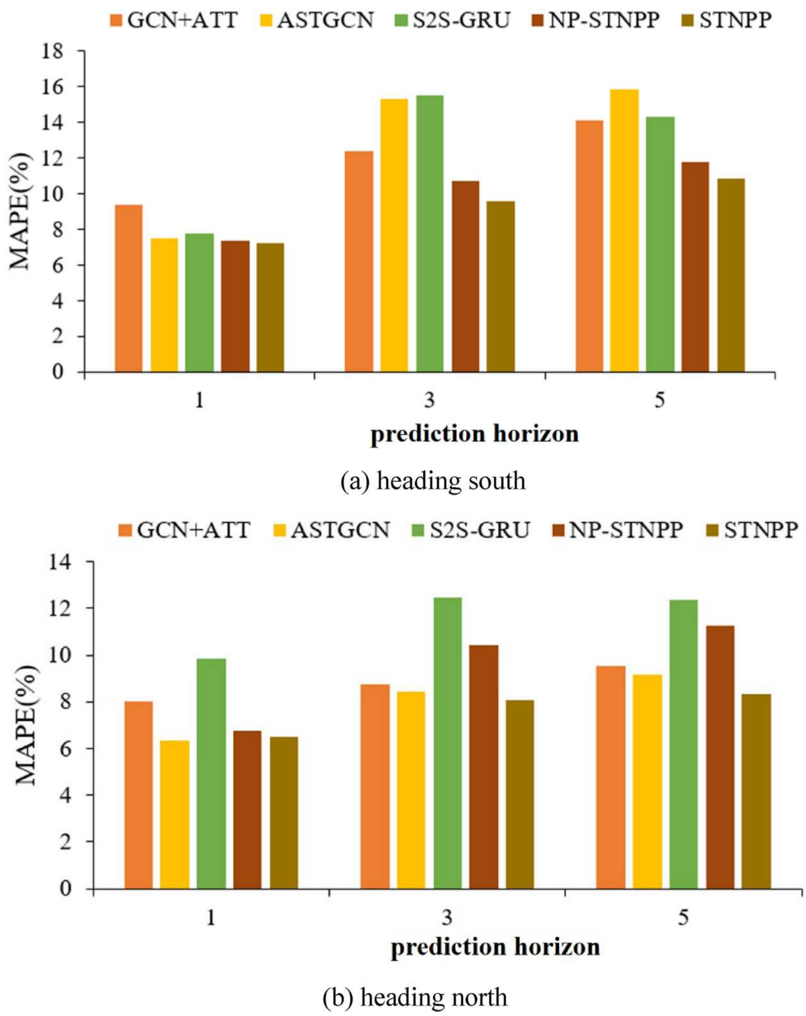


Figure 6. Congestion performance Evaluation during peak hours (MAPE index).

distribution patterns at two temporal granularities, effectively capturing the multi-step congestion change patterns during peak hours.

The model performs well in both (1) and (2), demonstrating the stability of the proposed STNPP model's predictive capabilities across different time periods. Based on the model results and analysis, the effectiveness of the STNPP model's design for congestion prediction is confirmed.

Additionally, the two variant models, NP-STNPP and S2S-GRU, which only consider time or do not account for congestion time granularity, perform poorly. This further confirms the effectiveness of the design in spatio-temporal correlation modeling. On the other hand, considering both spatial and temporal correlations in congestion problems generally leads to higher accuracy in most cases. Among them, the model proposed in this section achieves the best prediction performance. ASTGCN achieves suboptimal prediction accuracy for peak-hour congestion, possibly because the model incorporates an attention mechanism in convolutional neural networks, which allows it to capture spatiotemporal features over a wider time range. Next is the GCN + ATT model, which also performs well in multi-stsge prediction, possibly because the specific intersection lane has closer spatial correlations with upstream and downstream during each signal cycle, and the spatial distribution has a greater impact on prediction results.

(3) Comparison of congestion life cycle prediction during peak hours

To assess the overall congestion prediction performance of the model at specific intersections, we conducted a comprehensive evaluation of the refined congestion prediction over its entire life cycle. The life cycle of congestion was first divided into distinct phases: occurrence, development, and dissipation. Using the true scalar values (speed) of the test set, these stages were defined based on the temporal progression of congestion events. Specifically, congestion occurrence was identified by instances where the speed remained consistently below the standard speed per hour for multiple consecutive signal cycles at fixed intersections. Conversely, congestion dissipation was determined by sustained speeds above the standard, with the moment of departure from the standard speed regarded as the critical dissipation point. The period leading up to this point was defined as the congestion dissipation time period. The time interval between congestion occurrence and dissipation was designated as the congestion development stage. Subsequently, corresponding predicted scalar values were extracted, and an evaluation of the extracted true and predicted values of congestion events was conducted.

The daily peak congestion period was identified as occurring between 6:00 am to 9:00 am and 5:00 pm to 8:00 pm. Figures 7 (a) and (b) present a comparison of the predicted scalar accuracy during peak hours at intersection #767 at differentstages of the congestion life cycle. The results show that STNPP had the minimum RMSE and MAPE, consistently outperforming variants I and II in most cases. Although ASTGCN excelled in predicting congestion occurrence stages, its performance lagged behind in the development and dissipation stages. It is noteworthy that all models exhibited the least accurate predictions during the development stage, followed by the

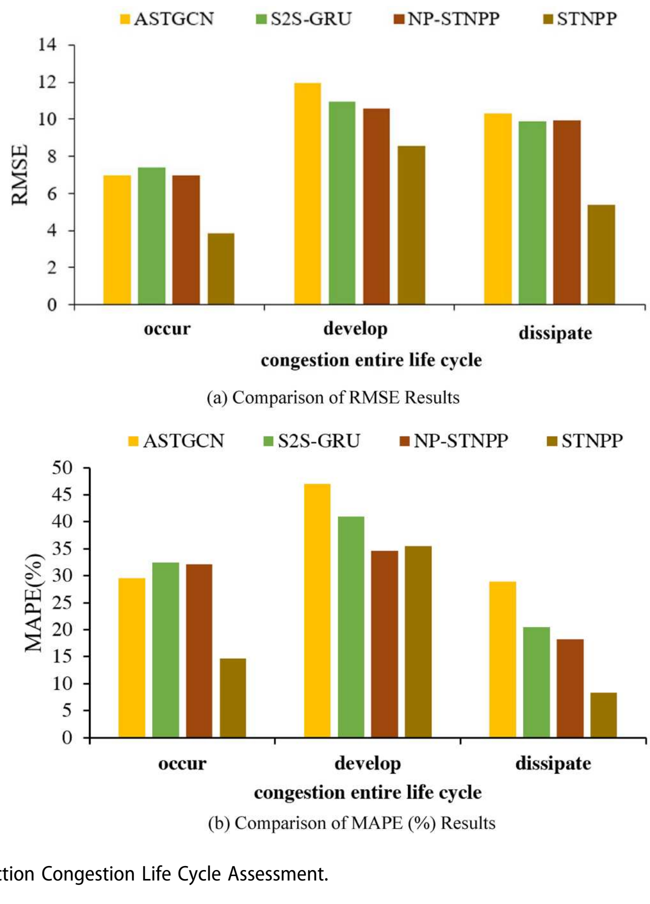


Figure 7. # 767 Intersection Congestion Life Cycle Assessment.

dissipation stage, while achieving the highest accuracy for predicting congestion occurrence. The challenge in predicting development was attributed to the broader spatio-temporal spread of congestion diffusion at intersections, introducing additional complexity. Moreover, congestion event marked values during the development stage displayed more pronounced and unpredictable fluctuations compared to general traffic flow parameters, further complicating the prediction process.

The prediction performance of STNPP surpassed that of both S2S-GRU and NP-STNPP, demonstrating the effectiveness of the model design. By incorporating intersection signal-cycle information into spatial considerations, STNPP outperformed the spatio-temporal prediction model ASTGCN, demonstrating the impact of signal cycle changes on spatial congestion distribution. Comparing STNPP with S2S-GRU and NP-STNPP, the model showed significant advantages in the occurrence, development, and dissipation stages. Although its predictive ability approached that of variants during the development stage, the model consistently exhibited absolute advantages in most cases, underscoring the pivotal role of spatial design and dual temporal granularity modeling.

4.6. Presentation of experimental results

(1) Visualization Results of Lane-Level Congestion Marked value Prediction

In this section, the results of fine-grained lane-level marked congestion value prediction are presented for the two methods that performed well. The results are shown in Figure 8. It can be observed that:

According to the city's traffic speed standards, some lanes at the current intersection are in slow-moving or severely congested states.

Even in cases where traffic speeds undergo significant changes, our model still achieves predictions that closely approximate the real values, with the smallest residuals. This is primarily due to the pattern-aware NPPGRU units capturing the historical impact and cumulative triggers of congestion over dual temporal granularity. As a result, more attention is given to locations with dramatic changes, allowing for a flexible response to real traffic fluctuations and avoiding overly smooth predictions.

STNPP leverages both spatial factors, considering the influence of traffic signals between lanes, and temporal factors, modeling the effects of continuous traffic flow and 'jump' events at dual temporal granularities. This approach results in smaller residuals (the difference between predicted and actual values) compared to other baseline models. The residuals for STNPP exhibit a clear distribution pattern centered around zero, without significant skewness.

(2) Visualization of the Entire Life cycle of Congestion Events

To qualitatively illustrate the predictive performance of STNPP for congestion, this section selects a full-life-cycle visualization of a congestion event at intersection #767 during the morning peak hours. By choosing a congestion event that occurred on Monday, December 10, 2018, during the morning rush hour, the event's life cycle is visualized in three stages: congestion initiation (8:15), development (8:45), and dissipation (9:30), as shown in Figure 9.

During the initial stages of the congestion event at 8:15 at the intersection, certain lanes, such as 5, 10, and 11, experienced slow-moving traffic (as depicted in Figure 9(a)). This period corresponds to the time when 'commuters' leave their homes to go to work, considering that most workplaces start at 9 in the morning.

As the morning rush hour progressed, the number of vehicles and pedestrians on the road continued to increase, and congestion began to spread spatially. By 8:45, lanes 0, 4, 5, 9, 10, and 11 experienced slow-moving and congested conditions, as severe congestion became apparent (as

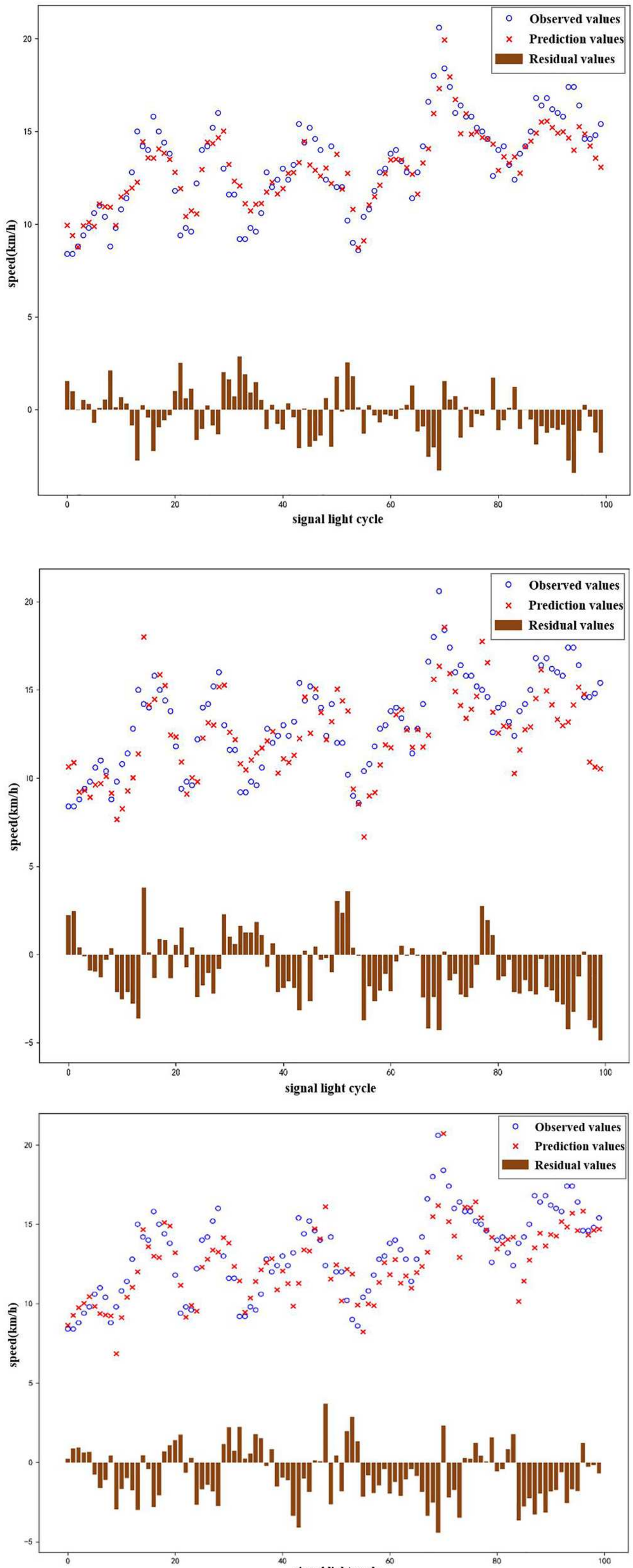
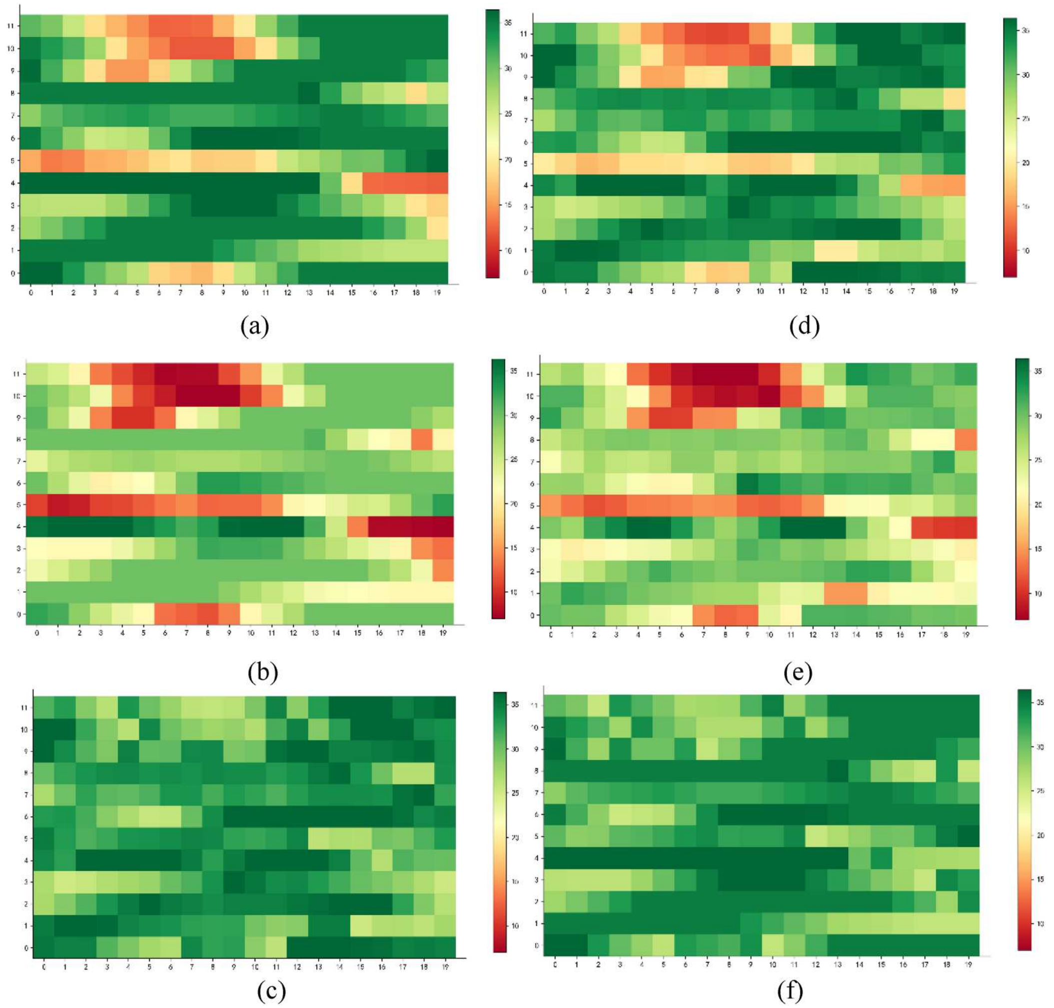


Figure 8. Comparison of congestion prediction results of northbound lane at Intersection #767.



Figuer 9. illustrates the complete life cycle prediction results of congestion at intersection #767 in the road network. The horizontal axis represents signal light cycles, while the vertical axis represents lane numbers at the intersection. This section presents the predictive performance of congestion threshold values for 20 signal cycles and 12 lanes at the intersection. In this context, (a), (b), and (c) represent the ground truth for congestion initiation, development, and dissipation, while (d), (e), and (f) represent the predictions from our model in this study.

shown in (b)). Comparing it with (e), it becomes evident that the STNPP model can accurately predict the distribution of congestion events. This highlights the model's advantage in predicting the development stage of congestion events, reaffirming the effectiveness of the model in modeling spatio-temporal correlations.

Finally, at 9:30, as most people had already arrived at their workplaces, the morning rush hour ended, and traffic congestion began to dissipate. The lanes at the intersection returned to a smooth-flowing state (as depicted in (c)). Comparing (c) with (f), the model's predictions for the dissipation stage align well with the actual values. While there are some overestimations in certain lanes, such as predicting smooth traffic as slow-moving, the overall prediction errors are minimal and meet the basic requirements for congestion prediction at the intersection.

In conclusion, through the visualization and analysis of the geometric morphology of congestion events at intersections over their entire life cycle, it is demonstrated that the proposed STNPP can accurately predict the entire process of congestion, from its initiation, development, to dissipation. From a qualitative perspective, this further illustrates the significance of our model in spatio-

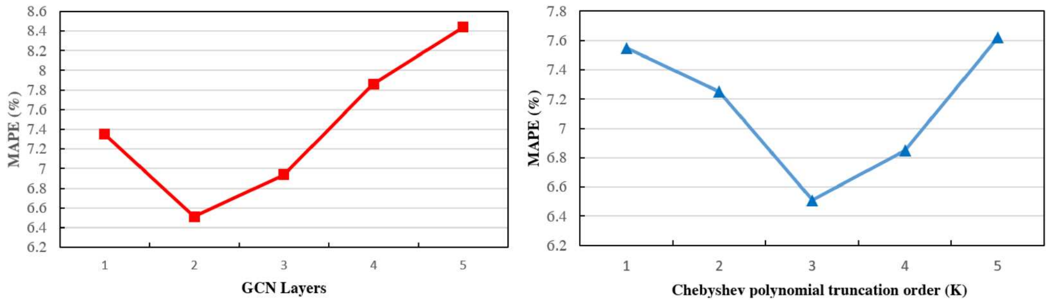


Figure 10. Studies on hyper-parameters (northbound driving lane, five signal cycle prediction).

temporal modeling design, taking into account the influence of both internal and external spatial factors and incorporating dual-granularity temporal modeling.

4.7. Parameters study

Since certain parameters can have a significant impact on the learning ability, we conducted a parametric study to further explore the effectiveness of the model. We chose the number of GCN layers and the Chebyshev polynomial truncation order (K). The experimental results are shown in Figure 10. We can see that the optimal number of GCN layers is 2, and as the number of layers increases, the MAPE becomes larger and larger. The reason may be that the overfitting problem of GCN limits its improvement. For the Chebyshev polynomial truncation order (K), the optimal value is 3. As K increases, the results become worse. This is because a larger K enhances the representativeness of the model while incurring an increase in computational complexity, a decrease in numerical stability, and increasing the risk of overfitting.

5. Conclusion

This paper addresses the issue of fine-grained traffic congestion prediction at urban signal-controlled intersections by proposing STNPP intersection congestion event prediction model based on combined graph neural networks and neural temporal point process.

In the spatial domain, the model focuses on intersection environments and constructs a fine-grained graph structure network model at the lane-level and cycle-level. It considers both global and local topological correlations while considering the influence of external signal control strategies.

In the temporal domain, the model addresses the dual-granularity temporal effects observed in congestion problems. It combines traditional point process models with gated recurrent neural network units, resulting in a novel NPPGRU.

The proposed congestion prediction method was validated using actual traffic data from Hangzhou City. Experimental scenarios included single-lane congestion prediction at intersections and performance evaluation during peak hours. The results showed that the proposed model outperformed the compared models in terms of MAE/RMSE, and MAPE metrics.

Future research will include analyzing congestion propagation processes at the urban road network level to extract congestion patterns, accurately define spatio-temporal congestion events, distinguish between sporadic and recurrent congestion mechanisms, and investigate the causes of congestion due to external factors such as weather changes and unexpected traffic accidents.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Key Research and Development Program of China [grant 2021YFB3101100], the Guizhou University talent introduction project [No. (2022) 49], Basic research project of Guizhou University [No. [2024] 16], 2024 Basic Research Program (Natural Science) Youth Guidance Project, Chinese Academy of Surveying and Mapping Basic Research Fund Program [grant number AR2204], Open Fund of National Engineering Research Center of Geographic Information System, China University of Geosciences, Wuhan 430074, China [grant number 2023KFJJ09], China Postdoctoral Science Foundation [grant number 2023M743454].

Data availability statement

Data not available due to commercial restrictions.

ORCID

Jianlong Wang http://orcid.org/0009-0000-6919-3216
Xiaoqi Duan http://orcid.org/0000-0002-6879-5634
Peixiao Wang http://orcid.org/0000-0002-1209-6340
A.-Gen Qiu http://orcid.org/0009-0007-2195-9252
Zeqiang Chen http://orcid.org/0000-0001-6624-6693

References

Abdullah, S. M., M. Periyasamy, N. A. Kamaludeen, S. K. Towfek, R. Marappan, S. Kidambi Raju, A. H. Alharbi, and D. S. Khafaga. 2023. "Optimizing Traffic Flow in Smart Cities: Soft GRU-Based Recurrent Neural Networks for Enhanced Congestion Prediction Using Deep Learning." *Sustainability* 15 (7): 5949. https://doi.org/10.3390/su15075949.

Chahal, A., P. Gulia, N. S. Gill, and I. Priyadarshini. 2023. "A Hybrid Univariate Traffic Congestion Prediction Model for IoT-Enabled Smart City." *Information* 14 (5): 268. https://doi.org/10.3390/info14050268.

Chauhan, A. V. S., S. Reddy, M. Singh, K. Singh, and T. Bhowmik. 2021. "Deviation-Based Marked Temporal Point Process for Marker Prediction." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 289–304. Cham: Springer. https://doi.org/10.1007/978-3-030-86486-618m

Chen, M., X. H. Yu, and Y. Liu. 2018. "PCNN: Deep Convolutional Networks for Short-Term Traffic Congestion Prediction." *IEEE Transactions on Intelligent Transportation Systems* 19 (11): 3550–3559. https://doi.org/10.1109/TITS.2018.2835523.

Daley, D. J., and D. Vere-Jones. 2008. *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. New York: Springer.

De Bézenac, E., A. Pajot, and P. Gallinari. 2019. "Deep learning for Physical Processes: Incorporating Prior Scientific Knowledge." *Journal of Statistical Mechanics: Theory and Experiment* 2019 (12): 124009. https://doi.org/10.1088/1742-5468/ab3195.

Defferrard, M., X. Bresson, and P. Vandergheynst. 2016. "Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering." *Advances in Neural Information Processing Systems* 29 (Nips 2016). https://doi.org/10.48550/arXiv.1606.09375.

Defferrard, M, X Bresson, and P Vandergheynst. 2016. "Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering." *Advances in Neural Information Processing Systems* 29.

Diao, Z., X. Wang, D. Zhang, K. Xie, and S. He. 2019. "Dynamic Spatial-Temporal Graph Convolutional Neural Networks for Traffic Forecasting." *Proceedings of the AAAI Conference on Artificial Intelligence* 33: 890–897. https://doi.org/10.1609/aaai.v33i01.3301890

Du, N., H. Dai, R. Trivedi, U. Upadhyay, M. Gomez-Rodriguez, and L. Song. 2016. "Recurrent Marked Temporal Point Processes: Embedding Event History to Vector." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1555–1564. https://doi.org/10.1145/2939672.2939875.

Du, H., Y. Zhou, Y. Ma, and S. Wang. 2021. "Astrologer: Exploiting Graph Neural Hawkes Process for Event Propagation Prediction with Spatio-Temporal Characteristics." *Knowledge-Based Systems* 228: 107247. https://doi.org/10.1016/j.knosys.2021.107247.

Duan, X., T. Zhang, Z. Xu, Q. Wan, J. Yan, W. Wang, and Y. Tian. 2023. "Discovering Urban Mobility Structure: A Spatio-Temporal Representational Learning Approach." *International Journal of Digital Earth* 16 (2): 4044–4072. https://doi.org/10.1080/17538947.2023.2261769

Fang, M., L. Tang, X. Yang, Y. Chen, C. Li, and Q. Li. 2021. "FTPG: A Fine-Grained Traffic Prediction Method with Graph Attention Network Using Big Trace Data." *IEEE Transactions on Intelligent Transportation Systems* 23 (6): 5163–5175. https://doi.org/10.1109/TITS.2021.3049264.

Gong, S., J. Liu, Y. Yang, J. Cai, G. Xu, R. Cao, C. Jing, and Y. Liu. 2024. "Self-Paced Gaussian-Based Graph Convolutional Network: Predicting Travel Flow and Unravelling Spatial Interactions through GPS Trajectory Data." *International Journal of Digital Earth* 17 (1): 2353123. https://doi.org/10.1080/17538947.2024.2353123.

Hu, J., C. Guo, B. Yang, and C. S. Jensen. 2019. "Stochastic Weight Completion for Road Networks Using Graph Convolutional Networks." *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 1274–1285. https://doi.org/10.1109/ICDE.2019.00116.

Huang, F.-R., C.-X. Wang, and C.-M. Chao. 2020. "Traffic Congestion Level Prediction Based on Recurrent Neural Networks." In *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 248–252. https://doi.org/10.1109/ICAIIC48513.2020.9065278.

Jin, G., F. Li, J. Zhang, M. Wang, and J. Huang. 2022. "Automated Dilated Spatio-Temporal Synchronous Graph Modeling for Traffic Prediction." *IEEE Transactions on Intelligent Transportation Systems* 24: 8820–8830. https://doi.org/10.1109/TITS.2022.3195232.

Jin, G., Y. Liang, Y. Fang, Z. Shao, J. Huang, J. Zhang, and Y. Zheng. 2023. "Spatio-Temporal Graph Neural Networks for Predictive Learning in Urban Computing: A Survey." *IEEE Transactions on Knowledge and Data Engineering*, 1–20. https://doi.org/10.1109/TKDE.2023.3333824.

Jin, G., L. Liu, F. Li, and J. Huang. 2023. "Spatio-Temporal Graph Neural Point Process for Traffic Congestion Event Prediction." *Proceedings of the AAAI Conference on Artificial Intelligence* 37 (12): 14268–14276. https://doi.org/10.1609/aaai.v37i12.26669

Kharaghani, H., H. Etemadfard, and M. Golmohammadi. 2023. "Spatio-Temporal Analysis of Precipitation Effects on Bicycle-Sharing Systems with Tensor Approach." *Journal of Geovisualization and Spatial Analysis* 7 (2): 30. https://doi.org/10.1007/s41651-023-00161-1.

Li, Y., and C. Shahabi. 2018. "A Brief Overview of Machine Learning Methods for Short-Term Traffic Forecasting and Future Directions." *SIGSPATIAL Special* 10 (1): 3–9. https://doi.org/10.1145/3231541.3231544.

Omi, T., and K. Aihara. 2019. "Fully Neural Network Based Model for General Temporal Point Processes." *Advances in Neural Information Processing Systems*, 32.

Saha, P., S. Dash, and S. Mukhopadhyay. 2021. "Physics-Incorporated Convolutional Recurrent Neural Networks for Source Identification and Forecasting of Dynamical Systems." *Neural Networks* 144: 359–371. https://doi.org/10.1016/j.neunet.2021.08.033.

Sharma, P. 2023. "Congestion Aware Traffic Prediction System Based on Pipelined Time Variant Feature Selection for Improving Transportation of Real Time Service." *2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 1–6. https://doi.org/10.1109/ICDCECE57866.2023.10150903.

Shchur, O., A. C. Türkmen, T. Januschowski, et al. 2021. "Neural Temporal Point Processes: A Review" *arXiv preprint arXiv:2104.03528*.

Simonovsky, M., and N. Komodakis. 2017. "Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3693–3702. https://doi.org/10.48550/arXiv.1704.02901.

Tseng, F. H., J. H. Hsueh, C. W. Tseng, Y. T. Yang, H. C. Chao, and L. D. Chou. 2018. "Congestion Prediction with Big Data for Real-Time Highway Traffic." *IEEE Access* 6: 57311–57323. https://doi.org/10.1109/ACCESS.2018.2873569.

Wang, Y. 2023. "Advances in Spatiotemporal Graph Neural Network Prediction Research." *International Journal of Digital Earth* 16 (1): 2034–2066. https://doi.org/10.1080/17538947.2023.2220610.

Willard, J., X. Jia, S. Xu, M. Steinbach, and V. Kumar. 2020. "Integrating Physics-Based Modeling With Machine Learning: A Survey." *arXiv preprint arXiv*:2003.04919. https://doi.org/10.48550/arXiv.2003.04919.

Wu, Z., S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip. 2020. "A Comprehensive Survey on Graph Neural Networks." *IEEE transactions on neural networks and learning systems* 32 (1): 4–24. https://doi.org/10.1109/TNNLS.2020.2978386.

Xiao, S., M. Farajtabar, X. Ye, J. Yan, L. Song, and H. Zha. 2017. "Wasserstein Learning of Deep Generative Point Process Models." *Advances in Neural Information Processing Systems* 30 (Nips 2017): 30.

Yu, Bing, Haoteng Yin, and Zhanxing Zhu. 2017. "Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting." *arXiv preprint arXiv:1709.04875*. https://doi.org/10.48550/arXiv.1709.04875.

Zafar, N., and I. Ul Haq. 2020. "Traffic Congestion Prediction Based on Estimated Time of Arrival." *PLoS ONE* 15 (12): e0238200. https://doi.org/10.1371/journal.pone.0238200.

Zhang, T., J. Wang, T. Wang, Y. Pang, P. Wang, and W. Wang. 2024. "A Deep Marked Graph Process Model for Citywide Traffic Congestion Forecasting." *Computer-Aided Civil and Infrastructure Engineering* 39 (8): 1180–1196. https://doi.org/10.1111/mice.13131.

Zhu, S., R. Ding, M. Zhang, P. Van Hentenryck, and Y. Xie. 2021. "Spatio-Temporal Point Processes with Attention for Traffic Congestion Event Modeling." *IEEE Transactions on Intelligent Transportation Systems* 23 (7): 7298–7309. https://doi.org/10.1109/TITS.2021.3068139.