HUMAN BIOLOGY
2024, VOL. 94, ISSUE 6
ORIGINAL ARTICLE

OPEN ACCESS

A Passenger Flow Prediction Model Based on Graph Convolutional Network with Multivariate Spatio-temporal Correlation

MA Ying*, LI Yang

ABSTRACT

Accurate prediction of short-term passenger flow is very important for rational planning and stable operation of cities, however, the problem of passenger flow prediction faces many challenges, including both the establishment of an effective spatio-temporal dynamic model structure and the necessity to comprehensively consider a variety of factors affecting the explicit and implicit passenger flow. So, a Multi-Variate Spatio-Temporal Correlation Graph Convolutional Network model (MVSTCGCN) is proposed. The model utilizes three kinds of spatially correlated graphs to construct a base graph, which is combined to capture spatio-temporal features globally; temporal attention mechanism, spatial attention mechanism, graph convolution operation, and spatio-temporal convolution constitute the spatio-temporal graph convolution module to capture local spatio-temporal features; meanwhile, the core module of graph convolution network is improved by being integrated wavelet transformation operators. The model is validated by New York taxi YellowTrip dataset and self-built dataset respectively; the simulation experiments show that the performance of our algorithm has more obvious advantages compared with other excellent algorithms.

INTRODUCTION

In recent years, the rapid development of intelligent transportation technology solutions has greatly promoted the development of smart cities and brought great convenience to citizens. However, the increasing passenger flow has brought great pressure to the transportation system, which is in conflict with the bottleneck of the existing urban infrastructure. In order to promote the efficient and safe operation of intelligent transportation system, the importance of scientific prediction of short-term passenger flow changes is highlighted. So, short-term passenger flow prediction (STPFP) is one of the research hotspots in the field of transportation planning and intelligent transportation system Wu et al. (2023), mainly because: on one hand, accurate prediction of the recent passenger flow can help operators of the transportation or service industry to understand the passenger travel demands in time, and develop scientific and reasonable operation schedules to make the transportation system operation more convenient and efficient; on the other hand, according to the results of passenger flow prediction timely deployment of vehicles, to ease the congestion in the peak period, to ensure the safety of passenger travel.

Early traffic flow prediction methods mainly are time series analysis method, predict the future flows by mining the change rules in historical data. The representative methods include: autoregressive integral moving average model (ARIMA) Williams et al. (2003), Zhang et al. (2018), Wen et al. (2022), vector autoregressive model (VAR) Lu et al. (2016), Kalman filtering model Lippi et al. (2013), and etc. These time-series-based methods of the short-term traffic flow prediction in the actual task show the relatively lower accuracy, therefore are not widely used and promoted. In order to extract the nonlinear features in the passenger flow, some machine learning methods have been applied to short-term traffic flow prediction, such as the simple Bayesian algorithm (Naive) Zhang et al. (2017), the support vector machine (SVM) Zhang et al. (2018), and etc. However, the performance of these algorithms depended on the quality of the manual feature extraction, which is inefficient and error-prone. With the advancement of technical means, deep learning models are gradually introduced into the field for automatic extraction of more potential features. Deep learning models represented by convolutional neural network (CNN) and recurrent neural network (RNN) Wu et al. (2021), which perform spatial feature extraction and temporal modeling of traffic flow data respectively, have achieved a large improvement in prediction performance compared with previous schemes. However, limited by the algorithmic mechanism, the topological relationship between regional networks and the spatio-temporal connection between different nodes are not fully explored.

School of Electronic and Information Engineering, Xi'an Technological University, Xi'an,710032, China.

Correspondence to: Dr. MA Ying, School of Electronic and Information Engineering, Xi'an Technological University, Xi'an,710032, China. E-mail: innovator@163.com.

Keywords: Short-term passenger flow prediction, Graph convolution network, Wavelet transformation, Temporal attention mechanisms, Spatial attention mechanisms.

Graph Convolutional Networks (GCN), which have a strong modeling ability for non-Euclidean geometric structures and can well capture the spatio-temporal correlation properties of topological data, have begun to be applied to traffic flow prediction. Bruna et al. (2013) firstly proposed Graph Convolutional Neural Network, which applied convolutional layers to graph data. Yu et al. (2017) proposed Spatial-Temporal Graph Convolutional Networks (STCBN), which consists of multiple spatio-temporal convolutional blocks in a "sandwich" structure in order to establish spatio-temporal relationships. Wu et al. (2023) proposed the AMGC-T model to discriminate the similarity of each city, so the cities with less data resources can use the transfer learning method. Tang et al. (2022) proposed the SPRNN model to extract spatial features by using the structural information of the roads, such as road connections, road density, etc. Wu et al. (2019) proposed the Graph WaveNet model, which integrates the Gated TCN and the GCN , and can capture spatial and temporal information in the ST data analysis to capture spatio-temporal information. Bai et al. (2020) proposed an AGCRN model, which can automatically capture spatio-temporal correlation.

Among them, the spatial temporal graph neural network (ST-GNN) Wu et al. (2021) is relatively achieved better effects and more recognized, which can predict the trend of passenger flow by combining the spatial graph convolution block and time series model.

However, accurate prediction of short-term passenger still faces some challenges. This is due to the fact that short-term passenger flow variations are affected by a variety of complex factors, and existed models are difficult to take into account the influence of external factors, such as weather, temperature, air quality, and other factors. In addition, it is not possible to take advantage of the various cyclical characteristics of traffic flow and the inability to establish deeper spatial and temporal relationships between nodes in the transportation network that are far away from each other, and the performance of most of the models is susceptible to the effects of the embedding functions and the hyperparameters.

In order to solve the above problems, a Multi-Variate Spatio-Temporal Correlation Graph Convolutional Network (MVSTCGCN) is proposed. The model is composed of two parts: the first part is the spatio-temporal information extraction module, which consists of temporal attention (TA), spatial attention (SA), and graph convolution for extracting local spatio-temporal features; and the second part is the fusion module, which combines local features extracted from different graphs to obtain global spatio-temporal features for predicting the passenger flow. The main innovations of our model are as follows:

(1) A short-term passenger flow prediction model based on deep learning framework is proposed to utilize the heterogeneous spatial correlation between regions for regional level passenger flow prediction, along with input features. For example, weather, temperature, etc. More specifically, we construct three different graphs for each city region, i.e., distance graph, interaction graph and correlation graph. Those graphs are fused to fully exploit the dynamic correlation information between their individual regions. Meanwhile, the model utilizes three kinds of periodic (recent, daily, and weekly) time series information on the input side, and fuses data from multiple sources, including historical foot traffic, weather, temperature, etc.

(2) The wavelet transform and Chebyshev fusion operator are proposed to replace the graph Laplace operator. The Chebyshev graph convolution (ChebNet) approximates the graph Laplace operator with Chebyshev polynomials, but ignores the multiscale information of spatial structure. Since the wavelet transform is able to extract the graph information from a multiresolution perspective, we fused wavelet transform to realize graph convolution, which is capable of capturing the higher-order nonlinear features of the graph data and also can pay attention to the multiscale information of spatial structure, as compared to ChebNet alone.

(3) A novel spatial and temporal attention mechanism is proposed to capture spatio-temporal features. When performing graph convolution, we combine the adjacency matrix A with the spatial attention matrix $S' \in R^{N\times N}$ In combination, the weights between nodes can be adaptively adjusted, and spatio-temporal relationships can be better established for nodes that are far away from each other.

In this paper, the effectiveness of our algorithm is validated by using publicly available datasets. The results show that the MVSTCGCN algorithm proposed in this paper has more significant advantages over other recent algorithms.

INTRODUCTION TO GCN

Definition of the problem

In urban passenger flow researches, crowd flow has great differences in different areas of the city. In order to evaluate and analyze the characteristics of the heat map of crowd flow in each region of the city, we form N regional networks based on longitude and latitude, and according to the administrative subdivisions of the city regions in the dataset.

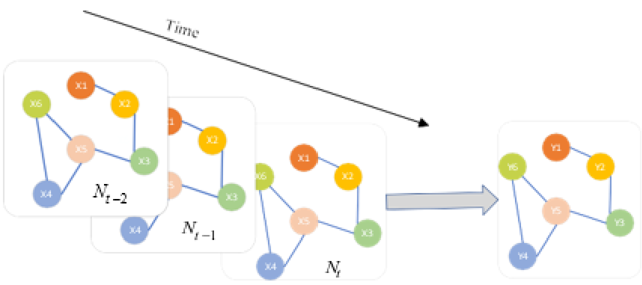
Taking Xi'an city as an example, its passenger flow at a certain moment is schematically shown in Fig. 1.

Figure 1: Heat Map of Xian's Pedestrian Flow

Definition 1: Based on the topology of the transportation network, define the inflow graph G_I = (U, E, A_I) and outflow graph G_O = (U , E, A_O) ,where U = (N_1 , N_2 ...N_n) is a set of N regions. E is an edge between two regions. $A_1 \in R^{N\times N}$ is the adjacency matrix of the inflow graph and $A_0 \in R^{N\times N}$ is the adjacency matrix of the outflow graph. In each time period, the eigenvalues of each region are represented by Yao et al. (2018), which mainly contains weather, temperature, and pedestrian flow.

Definition 2 : $x_t^{kf} \in (x_t^{k1},..., x_t^{kF}) \in RF$, k ∈ U , f ∈ F . f represent the eigenvalues of the k_{th} region at time t. x_t = (x_t^1,..., x_t^N) $\in R^{N\times F}$ represents the values of all features of all nodes at time t. x = (x_1 ,..., x_m) $\in R^{m\times N\times F}$ represents all eigenvalues of all nodes in the past m time periods. y = (y_1 ,..., y_τ) $\in R^{\tau\times N}$ represents all eigenvalues of all nodes for the predicted time period τ .The short-time passenger flow prediction utilizes the feature correlation of each node prediction as shown in Figure 2.

Figure 2: Short-term Patronage Forecasts



Graph Convolutional Networks (GCN)

Let D $\in R^{N\times N}$ be the degree matrix of inflow graph G_I and outflow graph G_O , where $d_i = \sum_j A_{ij}$ denotes the degree of node i ∈ 1,..., N . The theoretical basis of graph convolution in spatial dimension is to generalize the traditional data convolution operation based on grid structure to a wider range of graph structure data. According to Shuman et al. (2013), the traffic network can be essentially abstracted into a graph structure, in which each traffic node can be regarded as a signal on the graph. In order to fully exploit the feature information of the traffic network in terms of topology, we use the graph convolution operation based on spectral graph theory to directly process the signal data of each node at each time step.

Specifically, the algorithm exploits the signal correlation between the nodes of the traffic network in the spatial dimension. In addition, the spectral graph method is able to analyze the important properties of the graph structure, such as node connectivity, by representing the graph as an algebraic form, providing theoretical support for the subsequent exploration of the spatial structural relationships of the traffic network using the graph convolution network model. The symmetric normalized Laplace matrix of graph G is defined as: L =$D^{\frac{1}{2}}$ (D - A) $D^{\frac{1}{2}}$ 2 = I_N – $D^{\frac{1}{2}}$ 2 A$D^{\frac{1}{2}}$, The eigenvalue decomposition of Laplace matrix L is: L = $U\Lambda U^T$. Given the input x $\in R^N$, the Fourier transform is defined as: $\hat{x}= U^T x$, The inverse Fourier transform is defined as: x = $U_{\hat{x}}$.

Define the graph convolution operation in the Fourier domain as:

$$m *_G n = U (U^T m) \square (U^T n)$$

where ⊙ denotes the element-by-element Hadamard product,and $*_G$ represents the graph convolution operation. The parameterized Fourier filter is defined as:

$$g_\theta *_G = g_\theta(L)x = g_\theta (U\Lambda U^T)x = Ug_\theta (\Lambda)U^T x$$

Here the parameter $\theta \in R^n$. When large-scale graphs are subjected to Laplacian eigenvector computation, it is generally difficult to compute the eigenvalue decomposition of the global Laplacian matrix directly, in order to speed up the convolution process, the Chebyshev polynomials T_k (**Λ**) of truncated expansions up to order k Defferrard et al. (2018), and the approximation can be expressed as:

$$g_\theta(\Lambda) \approx \sum_{k=0}^{k-1} \theta_k T_k(\tilde{\Lambda})$$

Here $\tilde{L} = \frac{2}{\lambda max}$L-$I_n$, where λ_{max} represents the largest eigenvalue of L, $\theta = (\theta_0 ,...,\theta_{k-1})^T \in R^K$. The Chebyshev polynomial realization of GCN can be defined as:

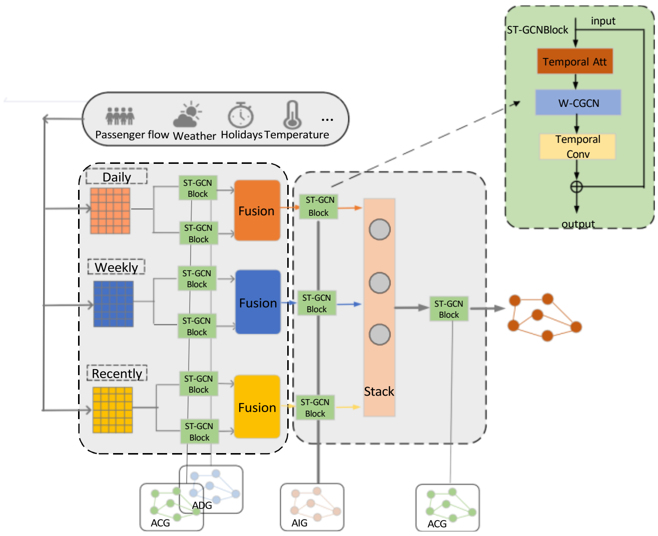
$$g_\theta *_G x = g_\theta(L)x = \sum_{k=0}^{k-1} \theta_k T_k(\tilde{L})x$$

Here $\tilde{L} = \frac{2}{\lambda max}$L-$I_n$, the Chebyshev polynomials are $T_k(x) = 2xT_{k-1}(x) - T_{K-2}(x)$.

Multivariate Spatio-Temporal Correlation Graph Convolution (MVSTCGCN) Modeling

The architectural diagram of the MVSTCGCN model proposed in this paper is shown in Figure 3:

Figure 3: MVSTCGCN Model Diagram



Overview of the Framework

(1) Graph construction: The division of the grid has significant effects on the prediction results of the algorithm. It has been shown that the prediction based on a single station is not effective Zhou et al. (2020). Therefore, we propose to describe the urban traffic flow system by utilizing the regional composition graph. The nodes in the graph represent the divided regions, and the edges connect different regions to indicate the spatial relationship between them. In order to make full use of various relevant information that may affect the prediction, we constructed three graph structures, namely, the distance graph, the interaction graph, and the correlation graph, which are mainly used as the adjacency matrix of GCN in the spatio-temporal convolution block (ST-GCNBlock), reflecting the relationships between regions from different perspectives. Since the features extracted from different graph structures are quite different, the distance and correlation graphs are fused using the adaptive gating mechanism, and then input into the interaction graph to fully correlate the features of different graphs and let the graphs interact with each other, and finally the prediction results are obtained from the correlation graph, which is a kind of multi-graph fusion that can fully mine the regional correlation laws, and is conducive to the time-series prediction of the distributed regions.

(2) Spatio-Temporal Graph Convolution Module (ST-GCNBlock): This module is the core component of this model, which consists of three parts, similar to a "sandwich structure". The first part is temporal attention, which focuses on the time series information in traffic data. The second part is the Wavelet Chebyshev Graph Convolution (W-CGCN) which incorporates the spatial attention mechanism. This part focuses on high-order nonlinear and multi-scale capturing of traffic data in terms of topology to fully explore the spatial correlation among its various regions. The third part is the temporal convolution module, STPFP is a typical temporal prediction problem, which is finally convolved in the time dimension to further update the temporal information of each node.

Model Inputs

We construct the input to MVSTCGCN to contain three time series segments with lengths T_r , T_d and T_w along the time axis, i.e., the current time input, the daily interval input and the weekly interval input, respectively. Let q be the sampling frequency (q times per day), t_0 is the current time, and T_p is the size of the prediction window, then these three time series fragments can be calculated as follows:

$$x_r = (X_{t_0-T_c+1},\dots,X_{t_0}) \in R^{F\times N\times T_c}$$

$$x_d = (X_{t_0-T_d*q+1},\dots,X_{t_0-T_d*q+T_p},\dots,X_{t_0-q+T_p}) \in R^{F\times N\times(T_dT_p)}$$

$$x_w = (X_{t_0-7*T_w*q+1},\dots,X_{t_0-7*T_w*q+T_p},\dots,X_{t_0-7*q+T_p}) \in R^{F\times N\times(T_wT_p)}$$

Constructing Draph

The graph representation method can fully explore the non-Euclidean relationship of the problem to be optimized. The rational graph structure can effectively optimize the learning parameters and improve the model prediction ability. Here, three kinds of graphs are constructed, namely, distance graph, interaction graph, and association graph.

Attributed distance graph (ADG): In the model, the ADG is constructed by the distance relationship between nodes firstly. The distance graph adopts the reciprocal of the spatial distance between nodes as the connection weight, which makes the expression between nodes with closer spatial distance closer. The ADG is defined as:

$$A = \begin{pmatrix} 0 & \cdots & \frac{1}{dis_{0N}} \\ \vdots & \ddots & \vdots \\ \frac{1}{dis_{N0}} & \cdots & 0 \end{pmatrix}$$

where $dis_{ij} = distance(N_i, N_j)$, represents the distance between region *i* and region *j*.

Attributed correlation graph (ACG): The ACG is constructed based on the functional correlation of nodes. The connection weights are set according to the correlation degree of functional attributes of nodes, which makes the expression between nodes with stronger functional correlation closer.

$$A = \begin{pmatrix} 0 & \cdots & r_{ON} \\ \vdots & \ddots & \vdots \\ r_{N0} & \cdots & 0 \end{pmatrix}$$

where

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

is the Pearson coefficient, representing the degree of correlation between region i and region j.

Attributed interaction graph (AIG): In order to represent the interaction relationship between nodes, the AIG is constructed. The AIG sets connection weights based on the correlation of interaction behaviors between nodes, highlighting the correlation of interaction behaviors.

$$A = \begin{pmatrix} 0 & \cdots & d_{ON} \\ \vdots & \ddots & \vdots \\ d_{N0} & \cdots & 0 \end{pmatrix}$$

Where d_{ij} represents the history of interaction between region i and region j, i.e., thefoot traffic in region i may affect the foot traffic in region j, then the record of interaction between the two regions d_{ij} increases.

Temporal and Spatial Attention Mechanisms

The spatio-temporal attention mechanism is introduced in the model to capture the dynamic spatio-temporal correlation in the transportation network, focusing on the spatio-temporal relationship between nodes that are far away from each other. It consists of two parts: spatial attention and temporal attention.

Spatial Attention Mechanism: Spatial attention is used to learn the dynamic relationships between traffic regional points in the spatial dimension. Since the traffic conditions at different locations interact with each other, we use the attention mechanism to adaptively capture the strength of the relationship between nodes, and assign different importance degrees to them, and use the sigmoid activation function to perform a nonlinear mapping of the graph structure data in the spatial dimension to enhance the spatial representation ability. Let $Z \in R^{Batch \times F \times T \times N}$ be the input of ST-GCNBlock, $S' \in R^{N \times N}$ is defined:

$$Z_S\,[N, B, F, T] = Z[B, F, T, N]$$

$$S = sigmoid(((Z_s\, W_1\,)W_2\,)(W_3\,(W_4\, Z_s'\,)) + a_1\,)$$

$$S'_{jk} = \frac{\exp(S_{jk})}{\sum_{k=1}^{N}\exp(S_{jk})}$$

where$W \in R^T$, $W \in R^F$,$W \in R^F$,$W \in R^T$, $a_1 \in R^{N \times N}$ are the corresponding parameters of spatial attention mechanism, along with j ∈ Batch , k ∈ N .

Temporal Attention Mechanism: In temporal dimension, the correlation between nodes is dynamically changing, as the traffic characteristics of the same location vary at different times. For this reason, we use a similar attention mechanism to autonomously learn and differentiate the level of importance of features. The temporal attention mechanism $E' \in R^{N \times N}$ is defined as:

$$Z_e\,[N, B, F, T] = Z[B, F, T, N]$$

$$E = sigmoid(((Z_e\, U_1\,)U_2\,)(U_3\,(U_4\, Z_e^T)) + b_1\,)$$

$$E'_{jk} = \frac{\exp(E_{jk})}{\sum_{k=1}^{N}\exp(E_{jk})}$$

Where $U_1 \in R^T$,$U_2 \in R^F$,$U_3 \in R^F$, $U_4 \in R^T$, $b_1 \in R^{N \times N}$are the corresponding parameters of spatial attention mechanism, along with j ∈ Batch , k ∈ N .

Wavelet-Chebyshev Gragh Convolution (W-CGCN)

Graph convolution neural networks have been widely adopted as a powerful framework for unstructured data processing in many tasks. Traditional harmonic graph convolution methods such as Chebyshev graph convolution (ChebNet) approximate the graph Laplace operator with Cheby polynomials, but it ignore the multi-scale information of spatial structure. The wavelet transform is capable of extracting graph information from a multi-resolution viewpoint, but it is difficult to express high-order nonlinear relationships when used alone. A harmonic graph convolution method based on wavelet-Chebyshev polynomial approximation is proposed to compensate for the shortcomings of the existing methods by utilizing both wavelet transform and Chebyshev polynomial techniques, which is equivalent to fusing multiscale information by learning the weights of its components at different scales. The principle of convolution kernel realization using wavelet transform and Chebyshev is as follows:

Let the graph Laplace matrix be L, whose maximum eigenvalue is λ_{max} . Obtain the low-frequency components of L:

$$\tilde{L} = LowPass(L)$$

where LowPass denotes frist level of wavelet low-pass filtering:

$$LowPass(L) = \sum_j L[j]\varphi_j$$

where φ_j denotes the low-pass basis function in first level wavelet analysis.

That is, $\tilde{L}$ is obtained by first-level decomposition of the original Laplace matrix L by wavelet transform, being keeped the low-pass component and multiplied the low-pass basis. The original Laplace matrix L contains all the frequency information, so it would take excessive noise components if be used directly. The low-frequency component $\tilde{L}$ is taken to filter out the high-frequency noise, and a smooth approximation matrix is obtained. Normalize the transformation of $\tilde{L}$, and get $\tilde{L}'$. Normalize the eigenvalues of $\tilde{L}'$ to facilitate the subsequent approximation using the Chebby harmonics. At the same time, $\tilde{L}$ is mapped into the definition domain [-1,1], which matches the Chebyshev polynomial term Tk (x) definition domain term, and will not change the $\tilde{L}$ structure information in the original space. The standard variation formula is:

$$\tilde{L}' = \frac{2}{\lambda_{max}}\tilde{L} - I$$

The W-CGCN filter is defined using Chebyshev polynomial approximation as:

$$g_\theta(\Lambda) \approx \sum_{k=0}^{k-1} \theta_k T_k(\tilde{L}')$$

where T_k is a k_{th} order Chebyshev polynomial, θ is a filter parameter; wavelet variation of the Chebyshev term:

$$C_k^{wt} = DWT(T_k(\tilde{L}'))$$

Among them, DWT denotes the first-level discrete wavelet transform, and the specific equations are:

$$c_k^{wt}[j] = \sum_n T_k(\tilde{L}')[n]\varphi_{j,n} \quad d_k^{wt}[j] = \sum_n T_k(\tilde{L}')[n]\psi_{j,n} \quad .$$

where $\varphi_{j,n}$ and $\varphi_{j,n}$ denote the impulse response function and wavelet function composed of Hal's wavelet basis, respectively.

Wavelet inversion transformation is defined as:

$$T_k(\tilde{L}')_{recons} = IDWT(C_k^{wt})$$

Where, IDWT denotes the inverse wavelet transformation, which is given by

$$T_k(\tilde{L}')[n] = \sum_j c_k^{wt}[j]\varphi_{j,n} + \sum_j d_k^{wt}[j]\psi_{j,n}$$

The W-CGCN graph convolution is defined as.

$$\xi = g_\theta(L)x = \sum_{k=0}^{k-1} \theta_k T_k(\tilde{L}')_{recons}\, x$$

Spatial-Temporal convolution block (ST-GCNBlock) the spatial-temporal convolution module consists of temporal attention mechanism (Tatt), spatial attention mechanism (Satt), wavelet Chebyshev graph convolution (W-CGCN), spatio-temporal convolution. Specifically, the input feature is denoted as $Z \in R^{Batch \times F \times T \times N}$, Z' will be obtained by using the temporal attention mechanism:

$$Z' = Z \square E'$$

Then, we analyze the graph structure using wavelet-Chebyshev graph convolution to extract the features in the Spatial domain, and integrate a spatial attention mechanism in order to dynamically adjust the weights between nodes.

$$Z'' = Z'\xi S' = g_\theta(L)S'Z' = \sum_{k=0}^{k-1} \theta_k T_k((\tilde{L}')recons \square S')Z'$$

$Z'' \in R^{Batch \times F1 \times N \times T}$ is a standard convolutional layer stacked on the time dimension of all nodes features extracted after graph convolution Guo et al. (2019). Using ELU for feature mapping, the model can learn the information of the nodes on the neighboring time segments, thus updating each node features. The time convolution operation is defined as:

$$Z''' = ELU(\varphi * Z'') \in R^{Batch \times F_1 \times N \times T_1}$$

Where φ represents the convolution kernel size, * represents the convolution operation, and the ELU is the activation function.

Multi-Graph fusion

As shown in Figure 3, the model consists of two main parts, corresponding to spatio-temporal information extraction and spatio-temporal information interaction. The first part consists of two graph structures, ADG and ACG, which constitute the ST-GCNBlock module to extract spatio-temporal information, and produces a total of three groups of results, with two graphs in each group, such as $Z_{i1(ADG)}$, ie [1,2,3] denotes the first graph of group i (ADG). The traffic flow features generated by different graph structures are different, and the use of adaptive gating mechanism to fuse the features of the two graphs generated in each group can effectively play the role of each graph. In the second part, the ST-GCNBlock composed of AIG graphs is used for the interaction of features between different graphs, and finally the results of the three groups graphs are stacked, and the ACG graphs are used for the feature correlation to obtain the

OPEN ACCESS

passenger flow prediction results.

$$\alpha_i = \sigma(Z'^{''}_{i1(ADG)} + Z'''_{i2(ACG)})$$

$$M_i' = \alpha_i \cdot Z''_{i1(ADG)} + (1-\alpha_i) \cdot Z''_{i2(ACG)}$$

$$M_i'' = Z''_{AIG}(M_i')$$

$$M''' = Z''_{ADG}(Stack(M_i''))$$

where , $\alpha \in (0,1)$ $M''' \in R^{Batch \times N \times F_1 \times T_1}$

Experimentation and Analysis

This section will analyze the experimental results in order to verify the validity of the algorithm through the experimental dataset.

Experimental data

1. Xi'an Tourism Shuttle Bus Data Set

The dataset used in this experiment is derived from the travel records of Xi'an tour group travelers provided by a technology company, which is a partner of the project. The city of Xi'an is located in Shaanxi Province, China, with a geographic location ranging from 107.40°to 109.49°east longitude and 33.42°to 34.45°north latitude. The plan is divided into 20 × 10 equal-sized tracts. The dataset covers the period from August 1, 2019 to October 31, 2019, and traffic flow were sampled and recorded at equal time intervals (1 hour) to form time series data. The dataset contains 24 attribute fields with 1,975,997 records, mainly recorded the information of taxi ride types, business types, departure time, departure locations, driving time and destinations. In addition to the traffic flow data, we also collected the weather and air quality monitoring data of Xi'an during the same period. Finally, through the discrete sampling of region and time, the complex information from multiple sources is fused to form a multidimensional time series data structure.

2. New York Yellow Taxi Trip Dataset

This dataset records data on New York City cab trips, including the following fields: date and time of pick-up and drop-off, location of pick-up and drop-off, distance traveled, sequential fare rate, type of rate, method of payment, and number of passengers reported by the driver. These taxi operation data are captured and provided to the NYC Taxi and Limousine Commission (TLC) by a licensed technology vendor under the authority of the NYC Taxi and Professional Vehicle Passenger Enhancement Program (TPEP/LPEP). The time frame of the dataset is from January 1, 2022 to February 28, 2022, with a total of 18,867,724 data, which were sampled and recorded using equal time intervals (1 hour). In summary, the dataset covers real-time taxi operation data within New York City, providing rich information for modeling and forecasting studies.

Experimental setup

The simulation environment is Linux system, CPU is Intel i9-13900H, GPU is NVIDIA GeForce RTX 4060Ti, and the deep learning framework is MXNet. The recent time series Tr is set to 23 hours, the daily cycle time series Td is set to 3 days, the periodic time series Tw is 2 set to weeks, the prediction time is Tp is set to 3 hours, the Epochs is set to 200, and the optimizer is Adam.

Assessment methodology

The performance of our algorithm was evaluated using root mean square error (RMSE) and mean absolute error (MAE).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - y_i)^2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y_i|$$

Where N represents the N regions predicted, y_i is the value to be predicted and y_i is the true value.

Model Comparison

The MVSTCGCN was compared with the following newer research achievements. The experiment runs 10 times, and selects the best convergence results of each algorithm. The results are shown in Table 1 and Figure 4.

Figure 4: The Iterative process of different algorithms

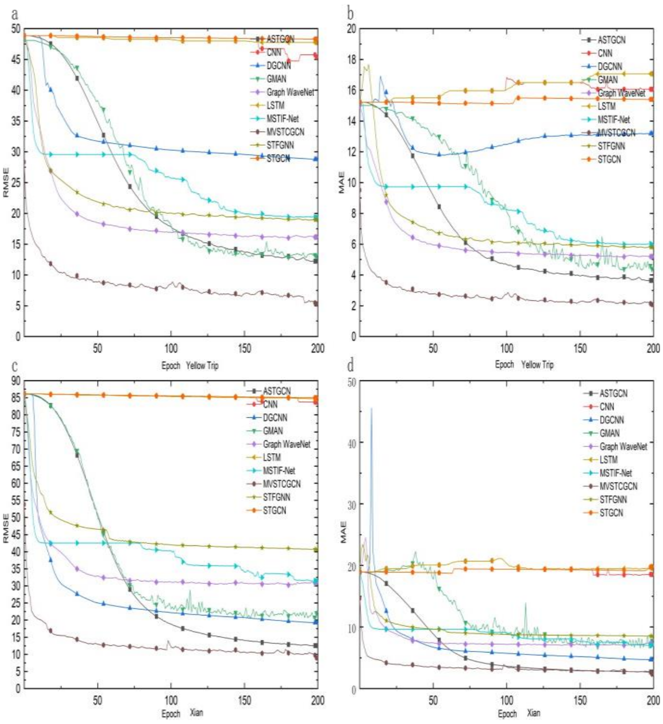


Table 1: The performance of different algorithms

Model	Xian		Yellow Trip	
	MAE	**RMSE**	**MAE**	**RMSE**
CNN	18.53	80.74	16.06	45.38
LSTN	19.25	84.75	17.64	47.67
STGCN	19.88	82.98	15.58	47.98
ASTGCN	2.72	12.5	3.55	12.47
GMAN	7.52	22.29	4.61	13.36
ASTGCN	2.72	12.5	3.55	12.47
Graph WaveNet	6.98	30.43	5.39	16.56
STFGNN	8.52	40.68	5.8	18.94
MSTIF-Net	7.05	31.49	5.96	19.39
DGCNN	4.88	20.27	13.38	28.45
MVSTCGCN	**2.32**	**8.97**	**2.1**	**5.31**

STGCN Yu et al. (2017) (Spatial-Temporal Graph Convolutional Network): uses a graph convolutional network and incorporates a temporal convolution module with a sandwich structure.

DGCNN Diao et al. (2019) (Dynamic Spatial-Temporal Graph Convolutional Neural Network):A method for estimating the dynamic Laplacian matrix.

ASTGCN Guo et al. (2019) (Attention-based Spatial-Temporal Graph Convolutional Network): introduces a variety of temporal attributes such as recent, periodic, and spatial-temporal attention mechanisms.

Graph WaveNet Wu et al. (2019) (Graph Wave Network): integrates gated loop cells and graph convolutional networks to model both temporal and spatial information.

MSTIF-Net Jin et al. (2020) (Multi-source Spatial-Temporal Information Fusion Network): combines GCN, VAE, Seq2Seq and other mechanisms to process traffic data.

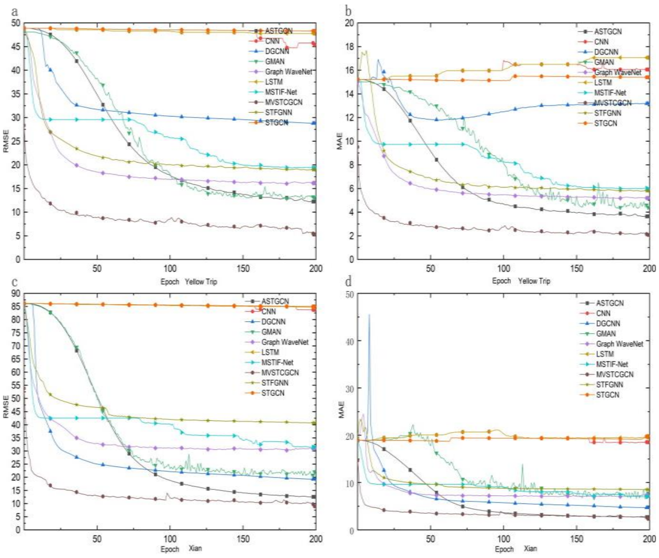
AGCRN Bai et al. (2020) (Adaptive Graph Convolutional Recurrent Network): automatic learning of spatial and temporal correlations.

GMAN Zheng et al. (2020) (Graph Multi-Attention Network):Learning graph structure and temporal information through node embedding and linear mapping.

STFGNN Li et al. (2021) (Spatial-Temporal Fusion Graph Neural Network): constructs temporal correlation graphs and designs spatial-temporal fusion mechanisms.

As shown in Table 1, in general, other non-graph neural network deep learning models (e.g., FC-LSTM, CNN, etc.), which ignore the correlation between nodes in spatial dimension, are not as good as graph neural network models designed using graph convolutional units in traffic prediction tasks. Although some GCN models using only a single graph structure (e.g., ASTGCN, etc.) have achieved relativelygood results in terms of convergence accuracy, they still need to be improved in terms of convergence speed and prediction accuracy. In contrast, those GCN models based on adaptive learning graph structure (e.g., DGCNN, GMAN, etc.) focus on the dynamic changes of node relationships while ensuring the prediction accuracy.

Figure 4: The Iterative process of different algorithms



Another class of models utilizing multiple graph structures (e.g., MSTIF-Net, etc.) provides richer spatial information for the model, which helps to uncover the implicit relations on the traffic features. In addition, the simultaneous collection of multi-source inputs from the time dimension can help the model to acquire more adequate time-dependent knowledge (e.g., ASTGCN, etc.), which performs well and further improves the prediction performance.

As shown in Table 1 and Fig. 4, MVSTCGCN utilizes multi-graph fusion to fully explore the relationships of spatio-temporal data in all dimensions and provides rich input features, which can establish deep spatio-temporal relationships between two nodes that are far away from each other, and the algorithm performs better compared to the other baseline models both in terms of the convergence accuracy and convergence speed.

Ablation experiments

Three graph structures are constructed in the MVSTCGCN model, and different graphs contribute differently to the model, so the permutations of different forms of graphs are used as hyperparameters of the model. In order to select the best combination of graphs with optimal performance and more interpretable, we choose a total of 14 combinations for experimentation and finally get (ACG,ADG,AIG,ACG) works best.

By analyzing Fig. 5 and Table 2, it can be seen that the AIG graph mainly carries out the interaction between different graphs, the ACG graph mainly carries out the correlation between different graphs, and the ADG establishes the topological structure, thus verifying the role of the model in designing these three types of graphs, which is in line with the expected results.

Figure 5: Comparison of the performance of different structure diagrams

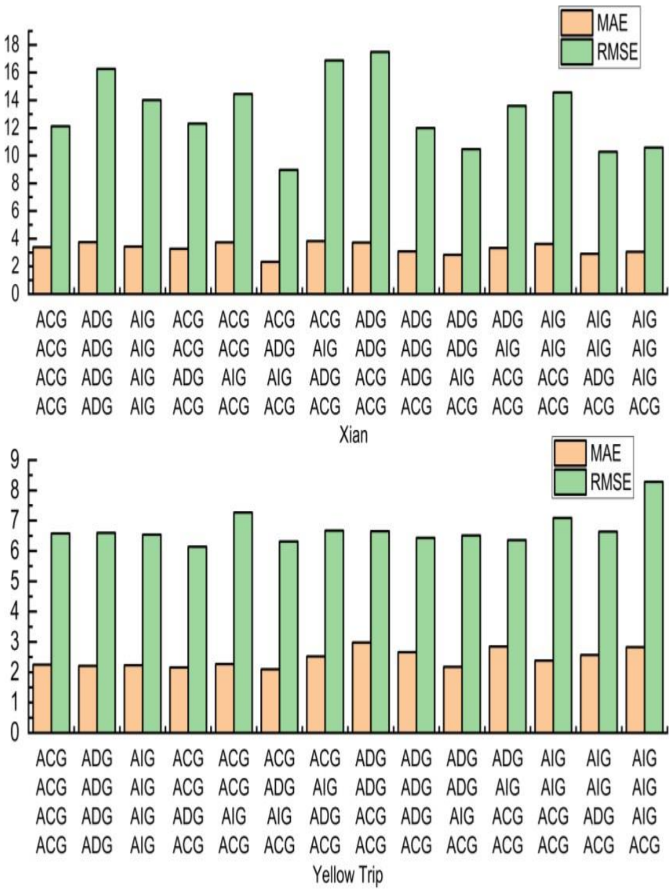
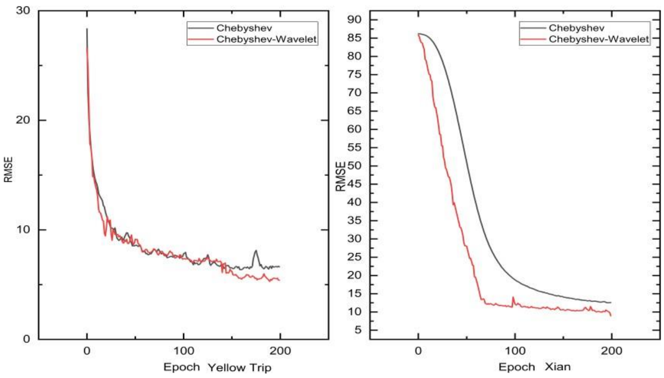


Table 2: The performance of different structure graph

Model	Xian		Yellow Trip	
	MAE	RMSE	MAE	RMSE
ACG ACG ACG ACG	3.39	12.12	2.25	6.58
ADG ADG ADG ADG	3.75	16.27	2.21	6.6
AIG AIG AIG AIG	3.44	14.02	2.23	6.54
ACG ACG ADG ACG	3.28	12.32	2.16	6.14
ACG ACG AIG ACG	3.73	14.44	2.27	7.27
ACG ADG AIG ACG	2.32	8.97	2.1	6.31
ACG AIG ADG ACG	3.82	16.88	2.52	6.67
ADG ADG ACG ACG	3.72	17.49	2.98	6.65
ADG ADG ADG ACG	3.08	12	2.66	6.43
ADG ADG AIG ACG	2.84	10.47	2.18	6.51
ADG AIG ACG ACG	3.33	13.59	2.85	6.36
AIG AIG ACG ACG	3.62	14.56	2.38	7.09
AIG AIG ADG ACG	2.91	10.27	2.57	6.64
AIG AIG AIG ACG	3.05	10.59	2.83	8.28

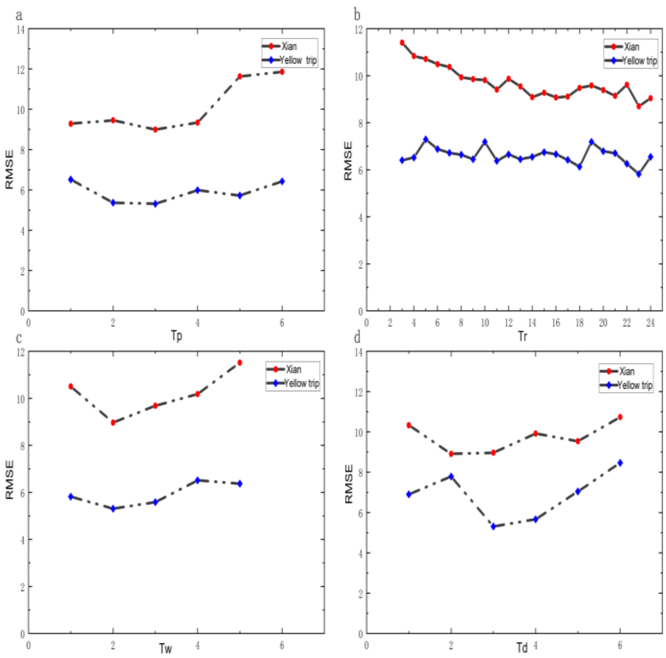
One of the crucial factors in the model is the use of Wavelet-Chebyshev graph convolution. In order to verify the role of wavelet transform and Chebyshev polynomials in graph convolution process, we do the corresponding experiments, as shown in Figure 6. The convergence speed with wavelet transform is faster and the RMSE is lower than model without wavelet transform, which indicates that the graph convolution implemented by Chebyshev polynomial integrated with wavelet transform is more effective in extracting multi-scale features and further more accelerates the convergence speed.

Figure 6: Comparison of the Role of Wavelet Transform



The influence of historical cycle characteristics of different time granularity on the prediction effect is important, and it is also one of the important hyperparameters. The MVSTCGCN model is executed on different datasets with different values of T_r, T_d, T_w and T_p to observe the RMSE trends, as shown in Fig. 7. Overall, the RMSE trends on different datasets are similar, which verifies the good generalization ability of the model. In both datasets, the optimal value of T_p is 3 hours.

Figure 7: Effects on different hyperparameter values



CONCLUSION

In this paper, we proposed the MVSTCGCN model for the traffic flow prediction task. The MVSTCGCN model utilizes spatio-temporal convolutional blocks composed of three different graph structures (distance graph, interaction graph, and correlation graph) to capture the features representation of traffic nodes in the spatio-temporal domain.

In order to realize multi-scale learning, we integrate wavelet transform with the Chebyshev graph convolution. In addition, we design a temporal attention mechanism and spatial attention mechanism so that the model can learn the spatio-temporal dependency relationship between more distant nodes. The input fuses multiple information and relies on spatio-temporal convolutional blocks composed of three graph structures to capture the spatio-temporal features of traffic flow data. We simulate comparative experiments with some existed traffic prediction models. The experimental results show that our proposed MVSTCGCN model achieves better prediction performance on test datasets and also converges faster.

In future work, we will further improve the MVSTCGCN model structure to further increase the granularity of prediction. More feature information on the input side can be fused to get further faster convergence speed of the algorithm. On constructing different graph structures, we can mine more valuable graph structures and establish diversified topology information, so as to achieve richer characterization of transportation networks.

DECLARATIONS

Corresponding Author:

MA Ying, E-mail address: innovator@163.com

Acknowledgements

This work was supported in part by Science and Technology Program Project of Xi'an Science and Technology Bureau, Xi'an, Shaanxi Province, China, under Grants No. 22GXFW0049.

Authors contribution statement

MA Ying: Conceptualization, Methodology, Writing - Review & Editing.

LI Yang: Programming, Experiments and data analysis, Writing - Original Draft.

Data Availability

The datasets generated or analysed during the current study are available from the corresponding author on reasonable request.

OPEN ACCESS

Ethical and informed consent for data used

All data in the database is anonymized and does not contain any personally identifiable information. All data is accessed and used in accordance with the relevant policy use policy and data protection guidelines.

Conflicts of interest

The authors declare that they have no conflict of interest.

REFERENCES

1.Wu J, Li X, He D, et al. 2023. Learning spatial-temporal dynamics and interactivity for short-term passenger flow prediction in urban rail transit. Applied Intelligence. 1-22.

2.Williams B M, Hoel L A. 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. Journal of transportation engineering. 129(6): 664-72.

3.Zhang H, Wang X, Cao J, et al. 2018. A hybrid short-term traffic flow forecasting model based on time series multifractal characteristics. Applied Intelligence. 48: 2429-40.

4.Wen K, Zhao G, He B, et al. 2022. A decomposition-based forecasting method with transfer learning for railway short-term passenger flow in holidays. Expert Systems with Applications.189: 116102.

5.Lu Z, Zhou C, Wu J, et al. 2016. Integrating Granger Causality and Vector Auto-Regression for Traffic Prediction of Large-Scale WLANs. KSII Transactions on Internet & Information Systems. 10(1).

6.Lippi M, Bertini M, Frasconi P. 2013. Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning. IEEE Transactions on Intelligent Transportation Systems. 14(2): 871-82.

7.Zhang H, Lv Y, Liu D, et al. 2017. A Naive Bayes Approach for Short-Term Traffic Flow Prediction. IEEE Access. 5(1): 21650-659.

8.Zhang X, Dong L. 2018. Short-term passenger flow prediction based on Naive Bayes and ARIMA. Journal of Intelligent Transportation Systems. 22(2): 98-108.

9.Wu Z, Pan S, Chen F, et al. 2021. A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst. 32(1): 4-24.

10.Bruna J, Zaremba W, Szlam A, et al. 2013. Spectral networks and locally connected networks on graphs. arXiv preprint arXiv. 1312.6203.

11.Yu B, Yin H, Zhu Z. 2017. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. arXiv preprint arXiv. 1709.04875.

12.Wu F, Zheng C, Zhang C, et al. 2023. Multi-View Multi-Attention Graph Neural Network for Traffic Flow Forecasting. Applied Sciences. 13(2): 711.

13.Tang G, Li B, Dai H N, et al. 2022. SPRNN: A spatial-temporal recurrent neural network for crowd flow prediction. Information Sciences. 614: 19-34.

14.Wu F, Souza A, Zhang T, et al. 2019. Simplifying graph convolutional networks, in:International Conference on Machine Learning. PMLR. 6861-71.

15.Bai L, Yao L, Li C, et al. 2020. Adaptive graph convolutional recurrent network for traffic forecasting in: Advances in Neural Infor mation Processing Systems. Curran Associates Inc. 17804-815.

16.Yao H, Wu F, Ke K, et al. 2018. Deep multi view spatial-temporal network for taxi demand prediction in: Proceedings of the AAAI Conference on Artificial Intelligence. 2588-95.

17.Shuman D I, Narang S K, Frossard P, et al. 2013. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE signal processing magazine. 30(3): 83-98.

18.Defferrard M, Bresson X, Vandergheynst P. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. Adv. Neural Inf. Process.Syst. 3844-52.

19.Zhou J, Cui G, Hu S, et al. 2020. Graph neural networks: a review of methods and applications. AI Open 1. 57 -81.

20.Guo S, Lin Y, Feng N, et al. 2019. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. Proceedings of the AAAI conference on artificial intelligence. 33(01): 922-29.

21.Diao Z, Wang X, Zhang D, et al. 2019. Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting. Proceedings of the AAAI conference on artificial intelligence.33(01): 890-97.

22.Wu F, Souza A, Zhang T, et al. 2019. Simplifying graph convolutional networks international conference on machine learning. Pmlr. 6861-71.

23.Jin G, Cui Y, Zeng L, et al. 2020. Urban ride-hailing demand prediction with multiple spatio-temporal information fusion network. Transp Res. 117:102665.

24.Zheng C, Fan X, Wang C, et al. 2020. Gman: A graph multi-attention network for traffic prediction. Proceedings of the AAAI conference on artificial intelligence. 34(01): 1234-41.

25.Li M, Zhu Z. 2021. Spatial-temporal fusion graph neural networks for traffic flow forecasting. Proceedings of the AAAI conference on artificial intelligence. 35(5): 4189-96.