



Multi-community passenger demand prediction at region level based on spatio-temporal graph convolutional network

Jinjun Tang^a, Jian Liang^a, Fang Liu^{b,*}, Jingjing Hao^c, Yinhai Wang^d

^a Smart Transportation Key Laboratory of Hunan Province, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China

^b School of Transportation Engineering, Changsha University of Science and Technology, Changsha 410205, China

^c School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

^d Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195-2007, USA



ARTICLE INFO

Keywords:

Passenger demand prediction
Region division
Deep learning
Graph convolutional network
Community detection

ABSTRACT

Region-level passenger demand prediction plays an important role in the coordination of travel demand and supply in the urban public transportation system. The complex urban road network structure leads to irregular shapes and arrangements of regions, which poses a challenge for capturing the spatio-temporal correlation of demand generated in different regions. In this study, we propose a multi-community spatio-temporal graph convolutional network (MC_STGCN) framework to predict passenger demand at a multi-region level by exploring spatio-temporal correlations among regions. Specifically, the gated recurrent unit (GRU) is applied to encode the temporal correlation in regions into a vector. On the other hand, the spatial correlations among regions are encoded into two graphs through the graph convolutional network (GCN): geographically adjacent graph and functional similarity graph. Then, a prediction module based on the Louvain algorithm is used to accomplish the passenger demand prediction of multi-regions. The two real-world taxi order data collected in Shenzhen City and New York City are used in model validation and comparison. The numerical results show that the MC_STGCN model outperforms both classical time-series prediction methods and deep learning approaches. Moreover, in order to better illustrate the superiority of the proposed model, we further discuss the improvement of prediction performance though spatio-temporal correlation modeling and analyzing, the effectiveness of community detection compared with random classification of regions, and the advantages of regional level prediction compared with grid-based prediction models.

1. Introduction

Realizing the balance between passenger demand and supply is an important issue to enhance travel efficiency in the transportation system. Due to the different development levels of the urban transportation system in different regions, it often happens that the areas with higher demand attract too many vehicles and result in an oversupply of the system. On the contrary, passengers in low-demand areas may spend more waiting time during their travels. It is generally believed that improving the service level and coverage of the public transport system, including urban transit, taxi, subway, and shared mobility, is an effective strategy to alleviate the spatial

* Corresponding author.

E-mail address: reliufang@163.com (F. Liu).

contradiction between regional demand and supply. The accurate passenger demand prediction over different spatial regions can play an important role in addressing this issue. Specifically, in the urban transit system, the routes and operating frequency can be updated quickly according to the future variation of passenger demand. Besides, accurate demand prediction can help taxi operator and ride-hailing platform to dynamically pre-allocate vehicles to meet the potential passenger demands, and thus improve the utilization rate of the registered cars and reduce passengers' waiting time.

However, a region-level passenger demand prediction is of a great challenge mainly due to the estimation of complex spatial and temporal correlation among different regions. According to Tobler's first law of geography (Tobler, 1970), the demand in the target region is greatly affected by the spatially adjacent regions. At the same time, two regions with a similar contextual environment but far apart may also have strong correlations (Chai et al., 2018; Ke et al., 2019). Moreover, the distribution of demand also expresses similar patterns at peak hours in the temporal dimension. How to effectively capture the spatio-temporal correlation of variables is a challenge faced in many studies focusing on traffic prediction, including traffic speed/volume prediction, human mobility pattern forecasting, and arrival time prediction. Taking the research object as time-series, various prediction methods have been widely used to model temporal correlativity, such as autoregressive integrated moving average (ARIMA) model and its variant (Ahmed and Cook, 1979; Hamed Mohammad M. et al., 1995), the Kalman filtering model (Okutani and Stephanedes, 1984), the Bayesian model (Sun et al., 2006), the Fuzzy-neural network (Tang et al., 2015, 2017), and the recurrent neural network (RNN) (Fu et al., 2016). To characterize the spatial relations, one common method is to conduct grid-based map segmentation over a study area, and then applies the convolutional neural network (CNN) and its variant to construct spatial correlation in the prediction. Based on the above two ideas, researchers focus on integrating CNN and RNN into a hybrid network to capture spatio-temporal correlation simultaneously (Shi et al., 2015; Yao et al., p. 2018).

Recently, in the passenger demand prediction, grid-based method combining with CNN or convolutional long short term memory network (ConvLSTM) produces high prediction accuracy and attracts the attention of a large number of researchers (Ke et al., 2018; Liu et al., 2019; Zhang et al., p. 2016.). It is still difficult or unreasonable for this kind of prediction methods in extracting and capturing the spatial correlation of passenger demand among different regions: (1) It is unreasonable to divide the area of the city into numerous equal-size grids. On the one hand, the grid division will destroy the complex structure of the urban road network, specifically, a school or shopping mall may be divided into two small units. On the other hand, treating the demand as an image and applying convolution operation will include the regions with weak correlations in the demand prediction of the target region, which may result in the decline of prediction performance; (2) The heterogeneity of spatial correlations is ignored. The passenger demand between two regions with a similar contextual environment but far apart with each other could still express a high correlation, while the grid-based approaches do not fully consider this dependency.

Inspired by the successful application of graph convolutional network (GCN) on non-Euclidean data, we propose an end-to-end deep learning architecture named multi-community spatio-temporal graph convolutional network (MC_STGCN) to implement urban passenger demand prediction under region level. In the MC_STGCN, the geographical relationship and functional similarity between regions are encoded into two graphs to represent the heterogeneous spatial correlation. The gated recurrent unit (GRU) is used to model temporal correlation and the GCN is applied to capture spatial correlations. Moreover, inspired by the work of Yu et al. (2018a), Yu et al. (2018b), a community detection algorithm is enabled to deeply fuse two graphs and perform simultaneous prediction for all regions. Our main contributions in this study are summarized as follows: (1) The MC_STGCN obtains the demand intensity based on regions instead of regular grids, and then combines GCN and GRU for region-level demand forecasting. (2) The MC_STGCN encodes the geographical location and traffic attributes between regions into different graphs to characterize the heterogeneous spatial correlations. (3) The community detection algorithm is designed to classify regions with strong spatio-temporal correlations into several communities to achieve the simultaneous prediction of multi-regions. The regions belonging to the same community can learn together and share the loss function. (4) The proposed architecture is validated using two sets of real-world passenger demand data, and numerical experiment results show that MC_STGCN outperforms benchmark models.

The remainder of this paper is organized as follows. Section 2 reviews relevant research about demand forecasting, and then briefly introduces applications of graph neural network. Section 3 provides preliminaries including the background of the core components and problem definition. We further describe the details of the proposed MC_STGCN in this section. In section 4, we evaluate the prediction performance of the proposed method based on two real-world datasets. Finally, conclusions and future research are drawn in Section 5.

2. Related works

Passenger demand prediction is essential for matching the number of waiting passengers and vacant resources from the urban public transportation system in a specific area (Moreira-Matias et al., 2012), so it has attracted extensive attention in the past decades. Sequence learning methods are often used to predict passenger demand based on historical observations, including traditional statistical models and recurrent neural networks. Li et al. (2012) presented an improved ARIMA based model to forecast passenger demand in a hotspot. Moreira-Matias et al. (2013) predicted passenger demand by combining several different time-series forecasting techniques, i.e., the time-varying Poisson model, the ARIMA model, and the weighted time-varying Poisson model. Davis et al. (2016) employed the multi-level clustering technique to improve the performance of demand prediction using the time-series model. Chen et al. (2017) proposed a framework based on recurrent neural networks to predict rental and return demand for bike-sharing stations. Xu et al. (2018) trained a Long Short Term Memory network (LSTM: a special recurrent neural network) with historical taxi demand data to predict its future variation patterns. The common disadvantage of these sequential learning methods is that they ignore the effect of the spatial correlations of passenger demand among different regions or zones. Due to the natural mobility of travelers, the

passenger demand in one particular area will be affected by its neighbors and other areas.

To explore spatial interaction effectively, various approaches have been proposed. Tong et al. (2017) extracted spatial features from point of interest (POI) data and processed them with a unified linear regression model to predict the unit taxi demand. Zhang et al. (2019) applied a spatio-temporal dynamic time warping algorithm to quantify the spatial correlations between different regions. Yu et al. (2019) presented a modified density-based spatial clustering algorithm with noise to construct many sub-networks considering the spatial correlation. Moreover, an interesting way to understand the local and global correlations is to treat the urban traffic states as an image, and then apply CNN on it to finish passenger demand prediction (Chen et al., 2016; Ma et al., 2017). On this basis, researches began to model spatial and temporal correlations simultaneously in an end-to-end deep learning framework. For instance, Wu et al. (2018) improved the traffic flow prediction accuracy with a deep neural network-based model, in which CNN was used to mine spatial features and RNN was applied to learn temporal dynamics in traffic flow. Yu et al. (2017) incorporated CNN and LSTM in a spatio-temporal recurrent convolutional network to predict the network-wide traffic speed. Zhang et al. (2017) fused several separate residual neural networks to predict the citywide crowd flows. Ke et al. (2017) presented a deep learning structure, which stacked ConvLSTM layers, LSTM layers, and CNN layers to better capture spatio-temporal correlation. Zhou et al. (2019) utilized a multi-level attention Network based on ConvLSTM units to forecast multi-step citywide passenger demand.

However, the distribution of passenger demand at the regional level is non-Euclidean structure due to the irregularity of regional relations. The methods including CNN and ConvLSTM are no longer suitable for passenger demand prediction at the regional level because they are commonly used for Euclidean data such as images, regular grids, and so on. Facing this issue and challenge, related works on graph neural networks provide us new light and possible solutions. Graph convolutional neural networks (GCN) (Defferrard et al., 2016) can effectively extract spatial features on graph-structured data by performing fast localized convolutional filters, and it has been successfully applied in image classification (Yi et al., 2017) and the segmentation of point cloud (Wang et al., 2019; Zhou and Lin, p, 2016.). Furthermore, several recent studies on traffic speed/volume prediction show that GCN can be employed to improve prediction performance by incorporating prior knowledge of traffic topology (Yu et al., 2018a, 2018b; Zhao et al., 2019). In addition, for the passenger demand prediction, Lin et al. (2018) combined GCN with a data-driven filter to predict hourly demand at bike-sharing stations. Geng et al. (2019) proposed a novel multi-graph convolutional neural network to reflect heterogeneous relationships between regions for ride-hailing demand forecasting.

In this context, considering the heterogeneous spatial correlation and non-Euclidean structure of passenger demand at the regional level, we propose an MC_STGCN architecture that combines GCN and GRU to capture spatio-temporal correlations of demand in different regions, and finally improve prediction performance by adopting community detection algorithm.

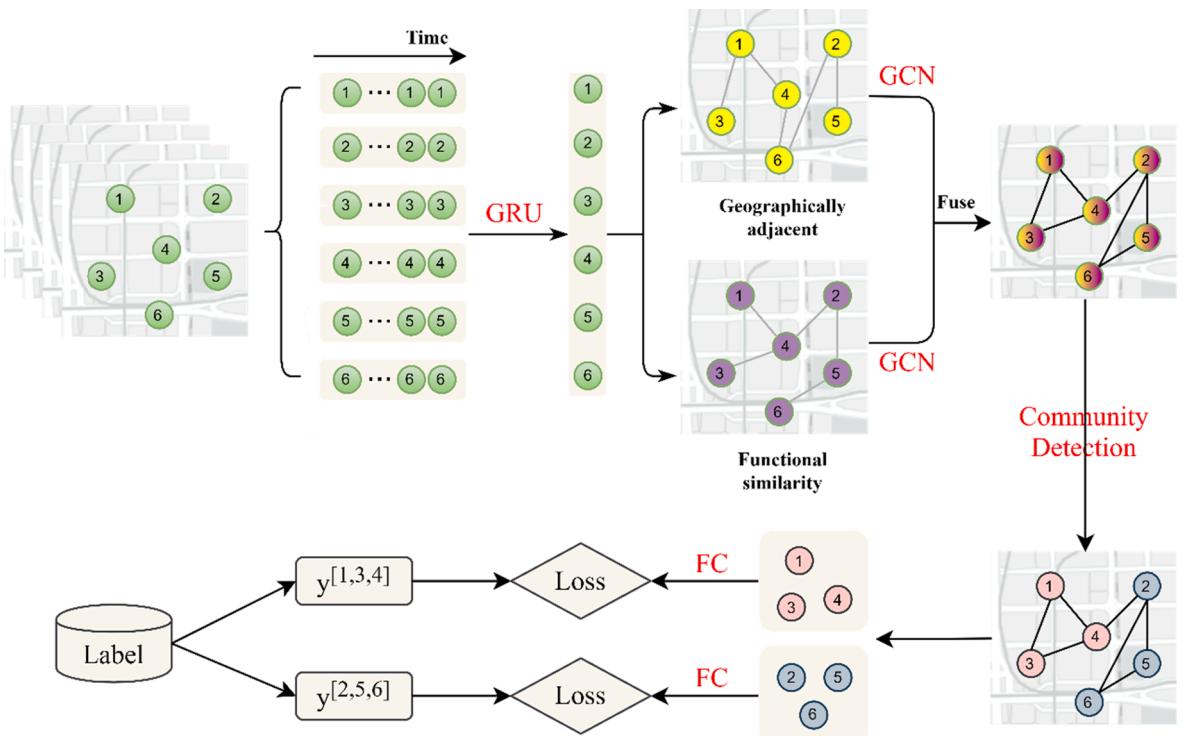


Fig. 1. Passenger demand prediction architecture at region level based on MC_STGCN.

3. Methodology

3.1. Problem definition

In this section, we first define several key parameters and then formalize the learning problem of the prediction of passenger demand at the regional level.

Definition 1. ((demand intensity):) *The study area is divided into many irregular regions according to the road network and then the number of travels is aggregated in each region at a specific time interval. The demand intensity is defined as the aggregated result. The demand intensity of the m th region at the t th time slot is denoted by d_t^m .*

Definition 2. ((region graph):) *An unweighted directed graph $G = (V, A, X)$ is used to combine all the irregular regions into a data structure. Each region is treated as a node, and the existence of edges is determined based on the heterogeneous spatial correlations between the regions. In the graph, V is the set of nodes, the adjacency matrix A indicates whether the nodes are connected, and X is the feature matrix of nodes. $X \in R^{N \times P}$, N is the number of regions and P is the length of the historical time series. $X^m = [d_{t-P}^m, \dots, d_{t-1}^m]$ represents the previous observations of the m th region, $X_t = [d_t^1, \dots, d_t^N]$ represents the demand intensity in all regions at the t th time slot.*

Thus, the forecasting problem of passenger demand at region-level can be formulated as follows:

Problem: *Given a region graph G over the previous P time intervals, the purpose of this study is to predict passenger demand intensity for all regions at the next time interval.*

3.2. Overview proposed model

In this section, we introduce the proposed MC_STGCN model in detail. Fig. 1 shows the architecture of the method and it consists of three parts:

- (1) Temporal correlation modeling: taking the region as a node in the graph, we first encode temporal correlation of passenger demand in different regions by extracting the keyframes from historical observations, and then the encoded non-spatial time-series of demand are taken as the input of the GRU layers.
- (2) Spatial correlation modeling: First, two graphs are constructed to encode heterogeneous spatial correlation based on the output of the GRU layers. Second, GCN layers are employed to characterize the spatio-temporal relations for each graph. Finally, the two graphs are fused from the node feature level considering the functional similarity and the topological structure level considering adjacent relationship.
- (3) Prediction module based on the Louvain algorithm: All regions are divided into several parts by applying the Louvain algorithm on the fused graph. The final output is obtained by establishing a fully connected layer (FC layer) for each region set. Then, the loss function is applied to evaluate prediction performance.

The details of the structure and calculation process for each model or algorithm will be introduced in the next sub-sections.

3.2.1. GRU structure

As an improved RNN, GRU (Cho et al., 2014) overcomes the defects of gradient disappearance through a gated mechanism. Compared with LSTM, GRU uses fewer parameters to achieve similar or better prediction performance (Chung et al., 2014). As shown in Fig. 2, each GRU unit obtains the current output (\hat{y}_t) and the hidden status at time t (h_t) by taking the current information (X_t) and the hidden status at time $t-1$ (h_{t-1}) as inputs. A typical GRU cell consists of two gates: reset gate and r_t and update gate z_t . The reset gate controls the extent of ignoring status information from the previous time step, and the update gate determines how much of the

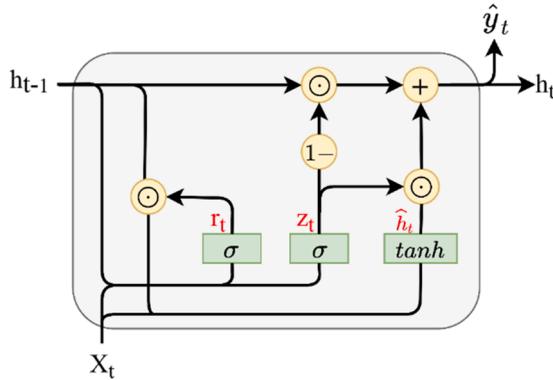


Fig. 2. Structure of the GRU unit.

historical information needs to be passed in current status. Both two gates are calculated by applying an activation function to the result of the weighted sum of X_t and h_{t-1} . Then, a current memory content \hat{h}_t is introduced to store the relevant information by using the reset gate from historical data. At last, the update gate is used to compute the current hidden status h_t . The calculation details are presented by the following equations:

$$r_t = \sigma(W^r X_t + U^r h_{t-1} + b^r) \quad (1)$$

$$z_t = \sigma(W^z X_t + U^z h_{t-1} + b^z) \quad (2)$$

$$\hat{h}_t = \tanh(W^h X_t + U^h (r_t \odot h_{t-1}) + b^h) \quad (3)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \hat{h}_t \quad (4)$$

where σ stands for the sigmoid function, which maps the input between 0 and 1. W and b denote the weights and bias respectively. \odot stands for the scalar product of two matrices.

3.2.2. GCN structure

Many important datasets in real life do not have a regular spatial structure but are presented in the form of graphs, such as social networks, molecular structures, and so on. It is a challenge for how to generalize neural network models on such structured datasets. Each node in the graph structure has its feature and structure information. GCN (Kipf and Welling, 2016) is an extension of CNN on the graph domain and can perform end-to-end learning on node features and structure information simultaneously. GCN takes a graph $G = (V, A, X)$ as input, where V is the set of all nodes, X is the feature matrix, and A is an adjacency matrix that describes the graph structure. By calculating the spectral decomposition of the graph Laplacian matrix, GCN defines convolutional operation in the Fourier domain. Generally, GCN includes two parts: the propagation module and the output module. The propagation module consists of aggregator and updater. For each node in the graph, the aggregator learns the potential representations by aggregating information from its neighbors. The updater implements the status update of each node. The output module completes the corresponding tasks based on the potential representation of the nodes. Each GCN layer produces a node-level output H (an $N \times F$ feature matrix, where N is the number of nodes and F is the dimensions of the potential representation of each node). The propagation rule is as follow:

$$\hat{A} = A + I \quad (5)$$

$$H^{l+1} = \sigma\left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^l W^l\right) \quad (6)$$

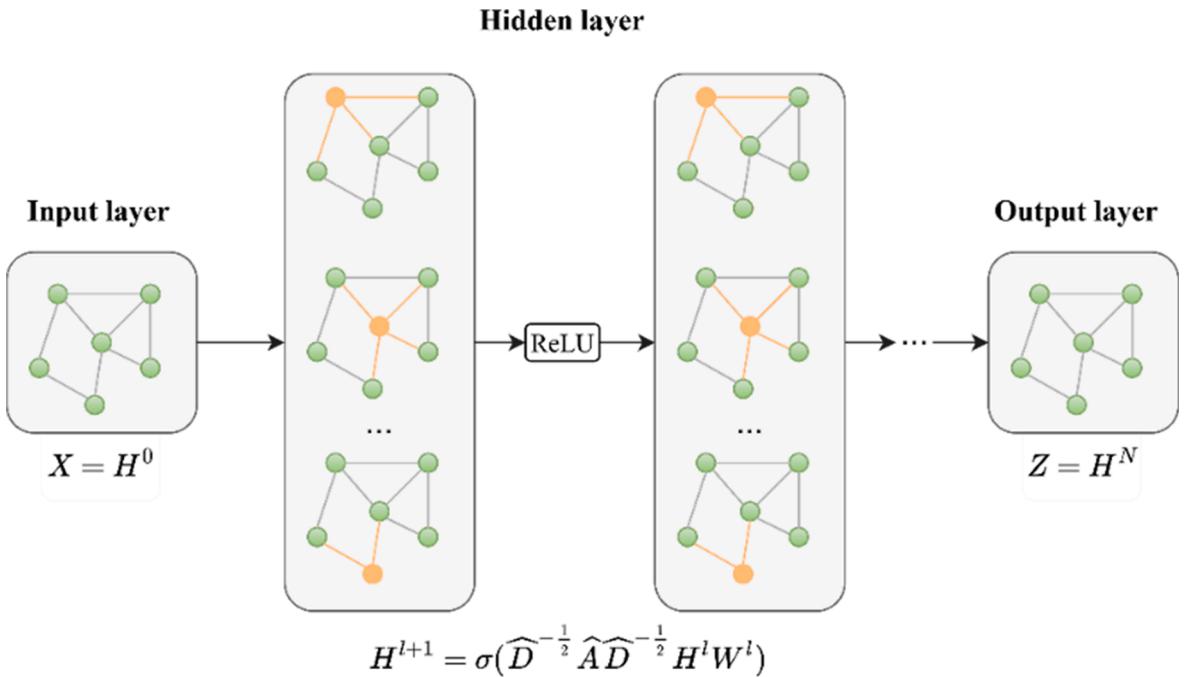


Fig. 3. The framework of a multi-layer GCN.

where I is the identity matrix, \hat{D} is the diagonal node degree matrix of \hat{A} , and W is the weight matrix for the l th neural network layer. Fig. 3 shows the framework of the multi-layer GCN model. The number of GCN layers is usually set to 2 or 3, because as the number of layers increases, the larger the receptive field of each node, the more features are aggregated, which may finally cause overfitting.

3.2.3. Louvain algorithm

Louvain is a widely used community detection algorithm, proposed by Blondel et al. (2008). Community structure refers to the occurrence of nodes groups in a network that is connected internally. Since a community structure acts as meta-nodes in the network, identifying the underlying community structures can provide insight into how the network topology affects the function of the community. There are often high similarities among nodes belonging to the same community, such as proteins with similar functions in biological cells, and literature with the same topic in a citation network. On the contrary, nodes belonging to different communities often have different properties. Louvain algorithm measures the closeness of a community through modularity (Q). It can be calculated as follows:

$$Q = \frac{1}{2m} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (7)$$

$$\delta(u, v) = \begin{cases} 1 & \text{when } u == v \\ 0 & \text{else} \end{cases} \quad (8)$$

where m is the sum of all of the edge weights in the graph, A_{ij} denotes the edge weight between node i and node j , K_i (K_j) denotes the sum of weights of the edges connected to node i (j), and c_i (c_j) represents the community to which node i (j) belongs. The Louvain algorithm includes two phases, which are repeatedly performed until the modularity of the graph no longer changes or the calculation reaches the maximum of iterations.

In phase one, each node in the graph is viewed as an independent community. A node is selected as the current node, and the change of modularity is calculated caused by adding the current node to the community where each of its neighbor nodes is located. Then, the current node is added into the community with the maximum modularity increment. This process is applied repeatedly and sequentially to all nodes until the community ownership of each node no longer changes. In phase two, the nodes belonging to the same community in phase one are combined into a super-node to reconstruct a graph. There is a self-loop on the super-node, whose weight is twice the sum of the weights of all connected edges in the community. The weight of the edge between the super-nodes is the sum of weights for the edges between the nodes in the two communities. The constructed graph is then reapplied to phase one. Fig. 4 shows a simple network with ten nodes as an example to express the calculation of phase one and two.

3.3. Spatial and temporal correlation analysis

3.3.1. Temporal correlation modeling

For a single region in the urban city, the distribution of passenger demand generally expresses obvious temporal correlation. On the one hand, the passenger demand is affected by the trend of recent historical demand; on the other hand, it has a strong periodicity, which includes daily and weekly cycles. For instance, passenger demand is expected to be high during rush hours and the variation pattern between working days and weekends is different. There are many historical observations for a particular forecasting period, but the length of the input variable in the prediction is limited. We extract the keyframes from the historical observations to encode the temporal correlation. Specifically, suppose the current time is t and the aggregating time interval is q . The demand intensity d_t to be predicted is obtained by aggregating the number of orders in the time interval $[t, t + q]$. We take multiple time series fragments along the time axis to represent hourly trends, daily cycle, and weekly cycle, respectively. Fig. 5 shows an example of keyframe extraction.

- (1) **Hourly trends component:** $D_H = [d_{t-a^*h}, d_{t-a^*h+q}, \dots, d_{t-q}]$. D_H includes all demand intensity over the past hours, corresponding to the red part of Fig. 5, used to describe the influence of the recent historical demand trends.
- (2) **Daily cycle component:** $D_D = [d_{t-b^*day}, \dots, d_{t-2^*day}, d_{t-1^*day}]$. D_D is composed of demand intensity at the same time period as the forecast interval in the last b days, which corresponds to the green part of Fig. 5 and reflects the daily periodicity.

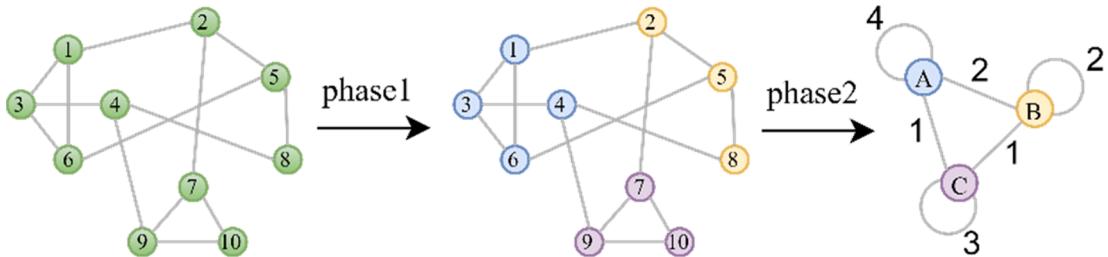


Fig. 4. Two phases in the Louvain algorithm (the weight of each edge is set as 1).

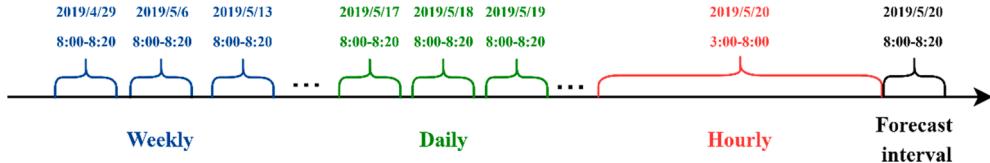


Fig. 5. An example of keyframe extraction ($t = 8:00$, $q = 20$ min).

- (3) **Weekly cycle component:** $D_W = [d_{t-c^*week}, \dots, d_{t-2^*week}, d_{t-1^*week}]$. D_W consists of demand intensity at the same time period during the last few weeks, which corresponds to the blue part of Fig. 5 and reflects the weekly periodicity.
- (4) **Long-range temporal correlation:**

$[D_H, D_D, D_W] = [d_{t-c^*week}, \dots, d_{t-1^*week}, d_{t-b^*day}, \dots, d_{t-1^*day}, d_{t-a^*h}, \dots, d_{t-q}]$. The three fragments D_H , D_D , and D_W are spliced into a feature vector to encode the long-range temporal correlation. This feature vector corresponds to the X^m defined in the problem definition.

We extract such feature vector for each region and then stack all the feature vectors to obtain a feature matrix X (defined in the problem definition). The obtained feature matrix is then sent to the GRU layer to extract temporal features for each region.

3.3.2. Spatial correlation modeling

As we mentioned before, passenger demand in different regions expresses heterogeneous spatial correlations. The distribution of passenger demand in a specific region is affected by that in the neighboring regions and regions with similar functions. In order to fully model the spatial correlation of passenger demand in regions, we construct two graphs with the different topological structure to encode adjacent and functional correlation.

- (1) **Geographically adjacent graph (GAG)** $G = (V, A_{ga}, X)$: We construct the graph by connecting two geographically adjacent regions so that information can flow through the edges. The formulation of the edge definition is given as follows.

$$A_{ga}^{[i,j]} = \begin{cases} 1, & \text{region } i \text{ and region } j \text{ are adjacent} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

- (2) **Functional similarity graph (FSG)** $G = (V, A_{fs}, X)$: Two regions that perform a similar function may have similar mobility patterns even if they are far apart in space. For instance, residential regions generate high passenger demands during a specific time period on the working days due to the commuting, and regions including shopping malls and entertainment spots are expected to attract higher passenger demand on weekends. As the POI data distributed in a region describes the number of entities with different functions. Therefore, we use POI similarity to quantify the functional similarity between different regions in this study. We calculate the POI similarity between any two regions and decide whether there is an edge between the two regions based on the values of the POI similarity. The threshold of the similarity is set as 0.8.

$$A_{fs}^{[i,j]} = \begin{cases} 1, & \text{sim}(P_i, P_j) > 0.8 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where P_i, P_j are the POI vector of region i and region j , and $\text{sim}(\cdot)$ is the calculation function of the Pearson coefficient.

After constructing the above two graphs, the GCN is then applied to extract spatial correlation. The spatial features can be captured in each region from related regions in the geographically adjacent graph and functional similarity graph, shown as follows.

$$G = (V, A_{ga}, X) \xrightarrow{\text{GCN}} G = (V, A_{ga}, Z_{ga}) \quad (11)$$

$$G = (V, A_{fs}, X) \xrightarrow{\text{GCN}} G = (V, A_{fs}, Z_{fs}) \quad (12)$$

where Z_{ga} , Z_{fs} represent the extracted feature matrix of the two graphs. We further fuse the spatial features of two graphs from the node feature level and the topological structure level, see Eq. (15). At the node level, the feature vector of each corresponding node in the two graphs is updated as a new vector of double size, see Eq. (13). At the topological structure level, the union of the edge sets in two graphs is performed in Eq. (14).

$$Z = (Z_{ga}, Z_{fs}) \quad (13)$$

$$A^{[i,j]} = A_{ga}^{[i,j]} \text{or} A_{fs}^{[i,j]} \quad (14)$$

$$\begin{aligned} G &= (V, A_{ga}, Z_{ga}) \xrightarrow{\text{fuse}} G = (V, A, Z) \\ G &= (V, A_{fs}, Z_{fs}) \end{aligned} \quad (15)$$

where Z and A indicate the new feature matrix and adjacent matrix in the fused graph, or is a logical operation indicating that the operation result is true if and only if one or more elements are true.

3.4. Final prediction module

In this section, we will discuss how to finish the final prediction based on the Louvain algorithm. An important issue faced in passenger demand forecasting at the regional level is how to predict demand patterns in multi-regions. In general, there are three scenarios: (a) Construct a unique model for each region and only predict passenger demand in one region each time. (b) Divide all regions into several sets, and implement the learning process for all regions, then finish the demand forecasting task of several region sets. (c) Forecast all regions at once. Fig. 6 shows the comparison of prediction structure between the input and output of the above three schemes.

The Scenario (a) ignores the association between regions and does not consider spatial correlation. As for scenario (c), it is difficult to predict all regions at once because of the great difference in spatial location and functional features among regions. Furthermore, a large number of outputs could make the loss function be defined in a very high dimensional space, which leads to the optimization algorithm easily falling into a local optimum. The Scenarios (b) is selected to finish final prediction in this study due to the following two reasons: (1) Decomposing prediction into multiple subtasks can simplify the complex problem shown in scenario (b). (2) Different subtasks utilize the domain knowledge through parameter sharing in the shallow layer of the model, thereby improving the generalization effect.

Follow the idea in scenario (b), the Louvain algorithm (introduced in section 3.2.3) is applied to divide all regions into several communities in the hybrid graph. Passenger demand prediction of different communities can be considered as a single task, and then the prediction in all communities is regarded as a multi-task learning problem. For each task, a fully connected neural network is used to transform the node 0 features into the prediction process.

$$\text{pred} = \sigma(W_{fc}Z^i + b_{fc}) \quad (16)$$

where W_{fc} , b_{fc} are both parameters, Z^i is the feature matrix of i th community, and σ denotes an activation function.

3.5. Multi-community spatio-temporal graph convolutional network

In this section, we illustrate the elaborate description of the proposed MC_STGCN method based on the modules introduced earlier. Supposed the number of samples is M . Each sample contains the historical demand intensity of all regions X and the target demand intensity of all regions Y . X is the feature matrix defined in Section 3.1, and each row of this matrix denotes the historical demand intensity of a single region. Moreover, the adjacency matrix of geographically adjacent A_{ga} and functional similarity A_f are used as the input of GCN. Table 1 shows the input, output, and specific process of the MC_STGCN.

4. Data source description and results analysis

4.1. Datasets and evaluation measurements

In this section, we evaluate the proposed model on two real-world datasets: TaxiSZ (Taxi order records collected from Shenzhen City) and TaxiNY (Taxi order records collected from New York City). The details of the two datasets are described as follows:

(1) TaxiSZ: It contains 19,544,075 taxi order records from 17th February to 22nd April in 2019. Each order record mainly consists of the pickup datetime, pickup longitude, pickup latitude, and other information. The selected study area is located in the Futian

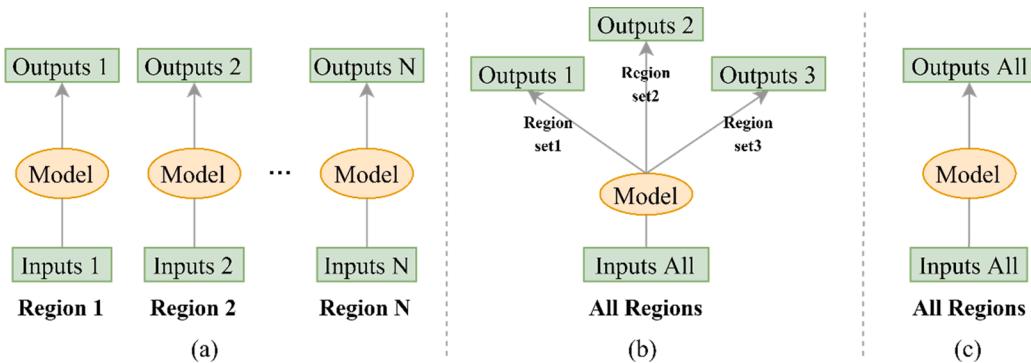
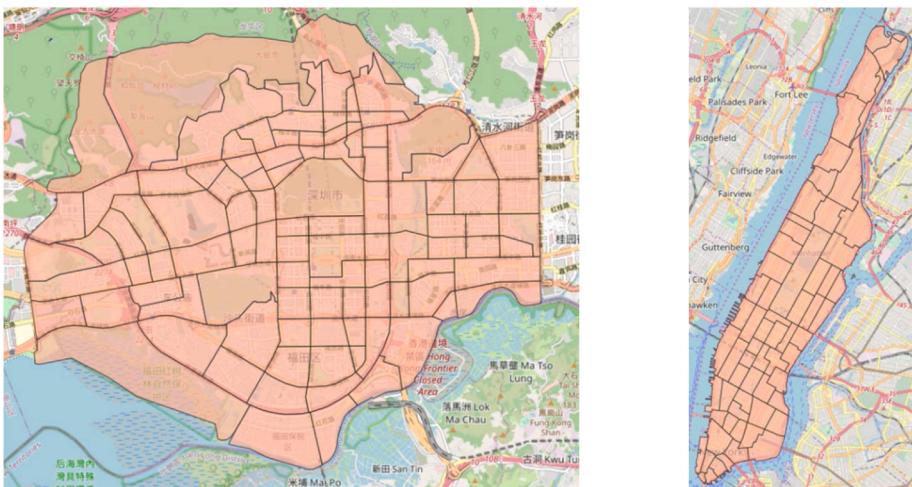


Fig. 6. Three scenarios in passenger demand forecasting (a) single input and single output; (b) all input and region set output; (c) all input and all output.

Table 1
Algorithm of MC_STGCN.

Algorithm: MC_STGCN	
Input:	$[X_1, X_2, \dots, X_M]$;
	$[Y_1, Y_2, \dots, Y_M]$;
	A_{ga} and A_{fs} ;
	$A^{[i,j]} = A_{ga}^{[i,j]}$ or $A_{fs}^{[i,j]}$ (described in section 3.2.2)
	Number of regions K (X_m^k denotes the feature vector of region k in sample m).
Process:	
1.	for $m = 1, \dots, M$:
2.	for $k = 1, \dots, K$:
3.	$h^k = \text{GRU}(X_m^k)$
4.	end
5.	$Z_{ga} = \text{GCN}([h^1, \dots, h^K]; A_{ga})$
6.	$Z_{fs} = \text{GCN}([h^1, \dots, h^K]; A_{fs})$
7.	$Z = (Z_{ga}, Z_{fs})$ (described in section 3.2.2)
8.	[Community ₁ , ..., Community _N] = Louvain ($A^{[i,j]}$)
9.	for $n = 1, \dots, N$:
10.	$\text{pred}_n = \text{FC}(Z_i, i \in \text{community}_n)$
11.	end
12.	end
Output:	$[\text{pred}_1, \dots, \text{pred}_N]$ (Forecasted demand intensity of all regions)

district, Shenzhen City, which is the administrative center and the central business district of the city. The investigated area is partitioned into 101 regions, as shown in Fig. 7a. It is should be noted that the region's division in the study area is based on the census tract areas (CTAs) identified by Urban Planning land and Resources commission in Shenzhen city. Data collected from 17th February to 8th April are used as the training set, data collected from 9th April to 15 April are used as a validation set, and data collected from 16th April to 22nd April are used as the test set.



(a) the division of 101 regions in Futian district,
Shenzhen City

(b) the division of 63 regions in
Manhattan district, New York City

Fig. 7. Regions division of two selected study area in two cities.

- (2) **TaxiNY:** It is an open-source dataset provided by The New York City Taxi and Limousine Commission ([TLC](#)). TaxiNY consists of 19,024,073 taxi order records in Manhattan district, New York City from 1st January to 30th June in 2019. The research area located in Manhattan district is partitioned into 63 regions or zones based on the NYC Department of City Planning's Neighborhood Tabulation Areas (NTAs), shown in [Fig. 7b](#). We use the data collected from 3rd June to 16th June as a validation set, and the data collected from 17th June to 30th June as a test set. The remaining data are used as a training set.

In addition, the time interval to aggregate order records as demand intensity is set as 20 min. We also collect POI data to match two study areas and quantify the functional similarity between regions for both TaxiSZ and TaxiNY datasets. Each region is assigned a 12-dimensional vector corresponding to 12 categories of entities: Education Facility, Recreational Facility, Transportation Facility, Government Facility, Medical facilities, Automobile Service, Financial Service, Building, Parking, Hotel, Food, Shop. For each region, the specific values in the vector represent the number of entities belonging to each category.

Three common indicators are applied to evaluate the prediction performance of passenger demand in model comparison: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), shown as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (17)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (18)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i} \quad (19)$$

where y_i and \hat{y}_i are the i^{th} ground truth and the predicted value of demand intensity. Since MAPE is greatly affected by the very small values, we only calculate MAPE for samples with demand intensity higher or equal to 10.

4.2. Prediction results comparison

In the model comparison, we compare the proposed MC-STGCN model with several benchmark methods on TaxiSZ and TaxiNY dataset to demonstrate its effectiveness. The details of the candidate models are described as follows:

- (1) **HA:** the History Average model predicts the passenger demand in a specific time interval by averaging the historical observations at the same time interval.
- (2) **ARIMA:** the Auto-Regressive Moving Average is a traditional time-series prediction method and is widely used in passenger demand prediction ([Li et al., 2012](#); [Moreira-Matias et al., 2013](#)).
- (3) **XGBoost:** it is a powerful ensemble learning algorithm with excellent performance and efficiency, and it has been applied in many traffic prediction tasks for its high learning efficiency, such as travel time prediction ([Kankamage et al., 2019](#)) and traffic flow prediction ([Dong et al., 2018](#)). The historical demand data of each region are fed to the XGBoost for training and testing.
- (4) **MLP:** Multiple Layer Perceptron is a classical feedforward artificial neural network. It can solve the nonlinear and high-dimensional problems well due to its strong self-adaptation and self-learning ability. It is also widely used in traffic flow prediction and modeling. For each region, we build an MLP with 100 ReLU neurons.
- (5) **GRU:** Gated Recurrent Unit is well-studied to tackle the problem of traffic state prediction since its excellent performance in sequence modeling. We establish a GRU for each region, and its settings are consistent with the GRU settings in MC_STGCN.
- (6) **LSTM:** Long Short-Term Memory is a commonly used neural network in time-series modeling, and it has many successful applications in passenger demand prediction ([Chen et al., 2017](#); [Fu et al., 2016](#)). The settings of LSTM are the same as GRU.
- (7) **GCN:** We build two two-layer graph neural networks to encode two graphs separately, and use the fully connected layer to predict the demand of all regions at once after fusing the node features of the encoded two graphs.
- (8) **STGCN:** A variant of MC_STGCN, which forecasts demand intensity in all regions at once. The only difference between STGCN and MC_STGCN is whether community detection is used. STGCN corresponds to the third scenario in [Fig. 6](#).
- (9) **T-GCN, Graph WaveNet:** Both T-GCN [[Zhao et al., 2019](#)] and Graph WaveNet [[Wu et al., 2019](#)] are advanced spatiotemporal graph convolutional neural network for traffic prediction. Both models use GCN to capture non-Euclidean relationships between road networks. Graph WaveNet uses a Temporal Convolutional Network to capture temporal dependence, while T-GCN uses GRU to model temporal correlation. We use the union of the adjacency matrices of the geographic neighboring graph and the functionally similar graph as the input of GCN in the two methods.

As introduced in [Section 3.3.1](#), through keyframe extraction, at each time interval t , we extract a feature vector to describe long-range temporal correlation for each region. For MC_STGCN, GCN, STGCN, T-GCN, and Graph WaveNet, the input received each time includes the adjacency matrix used to describe the spatial dependence between regions, and the feature matrix composed of feature vectors of all regions in the same time interval. For XGBoost, MLP, LSTM, and GRU, the input that the model accepts each time is a feature vector of a single region in a time interval. The HA model takes the average of historical observations to predict. In this paper,

we treat the extracted feature vectors as historical observations, so the input accepted by HA is the same as XGBoost. For the ARIMA model, the input is the original demand intensity sequence. For each region, we first make the demand sequence stationary, and then use the ARIMA model to predict.

Furthermore, in order to study the effectiveness of temporal and spatial correlation in demand prediction, we further consider four corresponding models related to MC_STGCN in the comparison. MC_STGCN encodes time correlations into vector consisting of three time-series fragments. For the temporal correlation, two corresponding models considering hourly trends component and daily cycle component are selected in the comparison. The model considering only the weekly cycle component is not selected because its effects are long-term and relatively weak on the demand prediction. As for spatial correlation, two corresponding models considering different graphs are selected as candidates.

- (10) MC_STGCN^{hourly}: It takes only historical passenger demand intensity from the past 6 h for each region as inputs dataset, and it has the same structure as the MC_STGCN model.
- (11) MC_STGCN^{daily}: It takes only the historical passenger demand intensity for the same time interval in the past 21 days as inputs dataset, and it also has the same structure as the MC_STGCN model.
- (12) MC_STGCN^{GA}: It has the same inputs dataset as the MC_STGCN model but only keeps the geographically adjacent (GA) graph in the spatial correlation modeling.
- (13) MC_STGCN^{FS}: It has the same inputs dataset as the MC_STGCN model but only keeps the functional similarity (FS) graph in the spatial correlation modeling.

For all above models in comparison, in order to make a fair evaluation, we use the same dataset for model training, validating, and testing. Furthermore, in this section, we provide the implementation details and experimental settings for the proposed MC_STGCN model. (1) In the temporal correlation modeling part, it consists of 2-layer GRU for each region. Both two layers in GRU contain 20 hidden states. The value of a, b, and c defined in section 3.3.1 is set to 5, 3, and 3, respectively. (2) The modeling results of the temporal correlation are then organized into two graphs according to the spatial correlation described above. In the spatial correlation modeling part, it contains a 2-layer GCN model. The number of hidden units is set to 32 for the lower layer and 20 for the higher layer. (3) The study areas are divided into several communities by applying the Louvain algorithm. As shown in Table 2, using the Louvain algorithm in community detection, 101 regions in TaxiSZ are divided into 8 communities, and 63 regions in TaxiNY are divided into 6 communities. (4) The fully connected neural network in the prediction module for each region set contains 2 layers: a layer with 100 ReLU neurons and a linear layer for output.

We implement all the methods in Python and use PyTorch for the neural network-based approaches on a server with NVIDIA GTX 2080Ti. The MC_STGCN is trained using Adam optimizer with a learning rate of 0.001 for TaxiSZ and 0.0005 for TaxiNY. The batch size of MC_STGCN on both TaxiSZ and TaxiNY is set to 72. We use the L2 loss to train the proposed model and use early stopping on the validation set. A performance comparison of different models is performed on the same test set.

Table 3 shows the mean average results of all regions on the same test set at a 20-minute aggregation interval under ten experimental tests. It can be found that the proposed MC_STGCN model achieves the best prediction results for both two datasets under three measurements. Besides, corresponding models related to MC_STGCN only considering the temporal or spatial correlation between different regions also produce better prediction performance than conventional methods, which demonstrates the effectiveness and superiority of the proposed framework considering spatio-temporal correlations of passenger demand. More specifically, we can observe that the MAPEs of HA and ARIMA model are significantly higher than that of other models, and the MAEs and RMSEs of these two traditional models are also relatively high, which reflects the limited learning ability of these two time-series methods in passenger demand prediction with long-term and non-stationary features at the region level.

Furthermore, the models which simultaneously consider both temporal and spatial correlations of passenger demand, including MC_STGCN and its corresponding models, achieve better prediction performance than traditional neural network approaches such as MLP and LSTM. This result further confirms that integrating spatio-temporal patterns into GCN mode ls plays a critical role in the improvement of passenger demand prediction. GCN ignores the impact of temporal dependence and produces large errors. Models that complete predictions for all regions at once, such as GCN, STGCN, T-GCN, Graph WaveNet, perform poorly compared to MC_STGCN model and its variants, which indicates that predicting the demand for all regions at the same time could results in the decline of

Table 2

. Results of community detection on TaxiNY and TaxiSZ dataset.

Community	TaxiSZ (101 regions)	TaxiNY (63 regions)
	8 region sets	6 region sets
No.1	[1, 2, 3, 4, 5, 7, 10, 64, 70, 78]	[1, 6, 8, 11, 12, 13, 14, 33, 34, 50, 55]
No.2	[6, 8, 9, 11, 23, 26, 46, 47, 48, 49, 50, 51, 52, 59, 67, 72, 82]	[2, 3, 5, 7, 22, 23, 25, 26, 44, 49, 57, 58, 60, 61, 62, 63]
No.3	[12, 22, 25, 30, 31, 36, 38, 40, 41, 42, 43, 44, 54, 93, 97]	[4, 9, 10, 16, 17, 29, 30, 31, 35, 56]
No.4	[13, 14, 16, 17, 24, 35, 37, 55, 56, 73, 79, 80, 98]	[15, 18, 32, 37, 38, 39, 40, 42, 47, 48, 51, 53, 54]
No.5	[15, 18, 19, 20, 21, 39, 45, 85, 86, 87, 88, 89, 90, 91, 92, 94, 95, 96, 99, 101]	[19, 27, 43, 46, 52]
No.6	[27, 28, 29, 32, 33, 34]	[19, 27, 43, 46, 52]
No.7	[53, 57, 58, 60, 61, 62, 63, 65, 66]	-
No.8	[68, 69, 71, 74, 75, 76, 77, 81, 83, 84, 100]	-

Table 3

. Comparison of prediction performance.

Methods	TaxiSZ			TaxiNY		
	MAE	RMSE	MAPE ₁₀ (%)	MAE	RMSE	MAPE ₁₀ (%)
HA	10.894	18.257	36.149	7.968	17.996	43.553
ARIMA	12.576	20.634	113.569	5.328	11.620	101.042
XGBoost	7.552	11.898	26.074	4.464	8.904	23.435
MLP	10.224	17.405	29.266	6.150	10.554	24.249
GRU	7.711	12.851	27.780	4.585	9.345	26.375
LSTM	8.369	13.623	30.128	5.328	10.435	29.118
GCN	14.263	23.164	44.704	11.662	25.418	56.455
STGCN	10.180	15.334	32.655	5.241	9.778	23.489
T-GCN	7.737	11.740	24.290	5.585	9.155	21.370
Graph WaveNet	9.735	13.920	23.610	4.641	9.229	21.617
MC_STGCN ^{daily}	8.078	13.699	25.675	5.690	13.190	25.470
MC_STGCN ^{hourly}	7.200	12.054	23.383	4.035	8.638	19.817
MC_STGCN ^{GA}	6.764	11.279	21.944	4.609	10.345	21.944
MC_STGCN ^{FS}	6.753	11.365	22.136	4.133	9.235	20.310
MC_STGCN	6.555	11.005	21.312	3.727	7.863	19.241

prediction accuracy. From the Table 3, we can see that two state-of-the-art methods, T-GCN and Graph WaveNet, express lower prediction accuracy than the proposed MC_STGCN model and its variant. The reason is that these two models cannot accurately extract heterogeneous spatial dependence of travel demand in multi-graph fusion process.

In Fig. 8, we further show the comparison between observed and predicted passenger demand based on the MC_STGCN model for one week. For TaxiSZ and TaxiNY dataset, we both select three specific regions in different communities. As we can see, the distributions of passenger demand in these regions express obvious different patterns with different demand intensity and varying shapes. The comparison also proves that the proposed model can effectively follow the fluctuation of real data in different regions.

4.3. Prediction results discussion

(1) Effect of temporal and spatial correlations

In the prediction comparison, Table 3 shows the test errors of MC_STGCN and its corresponding models in two datasets, while this part will further show the validation errors of these models as the number of training epoch increases, shown in Fig. 9. Compared with

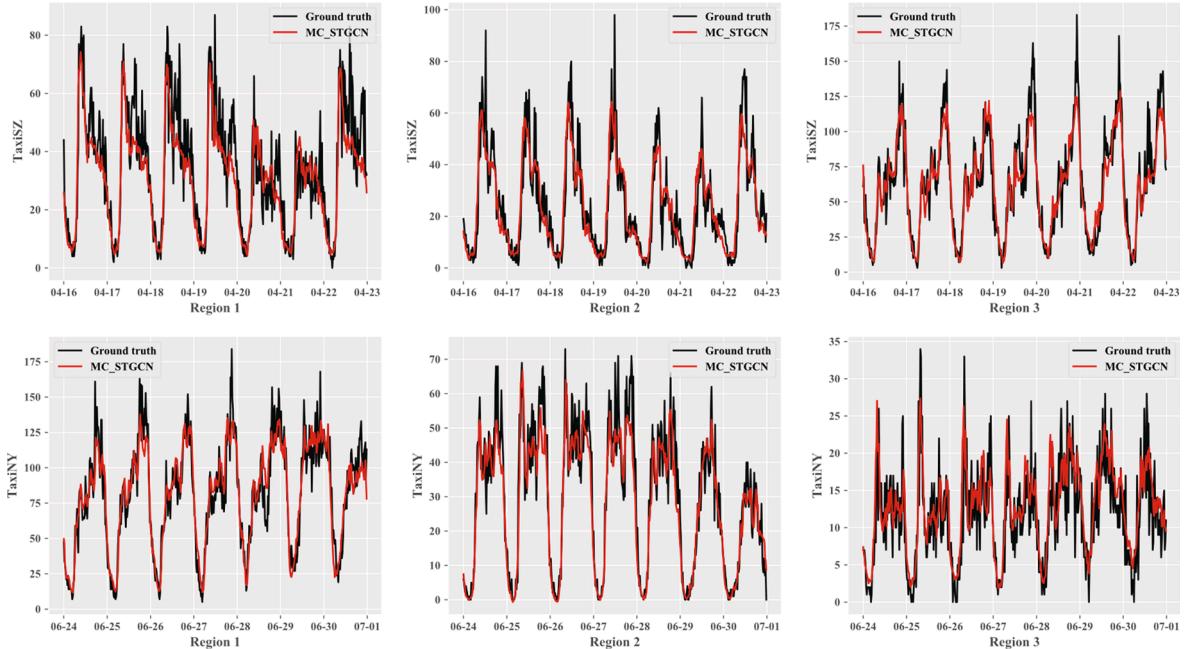


Fig. 8. Comparison of prediction results and ground truth based on MC_STGCN model during one week.

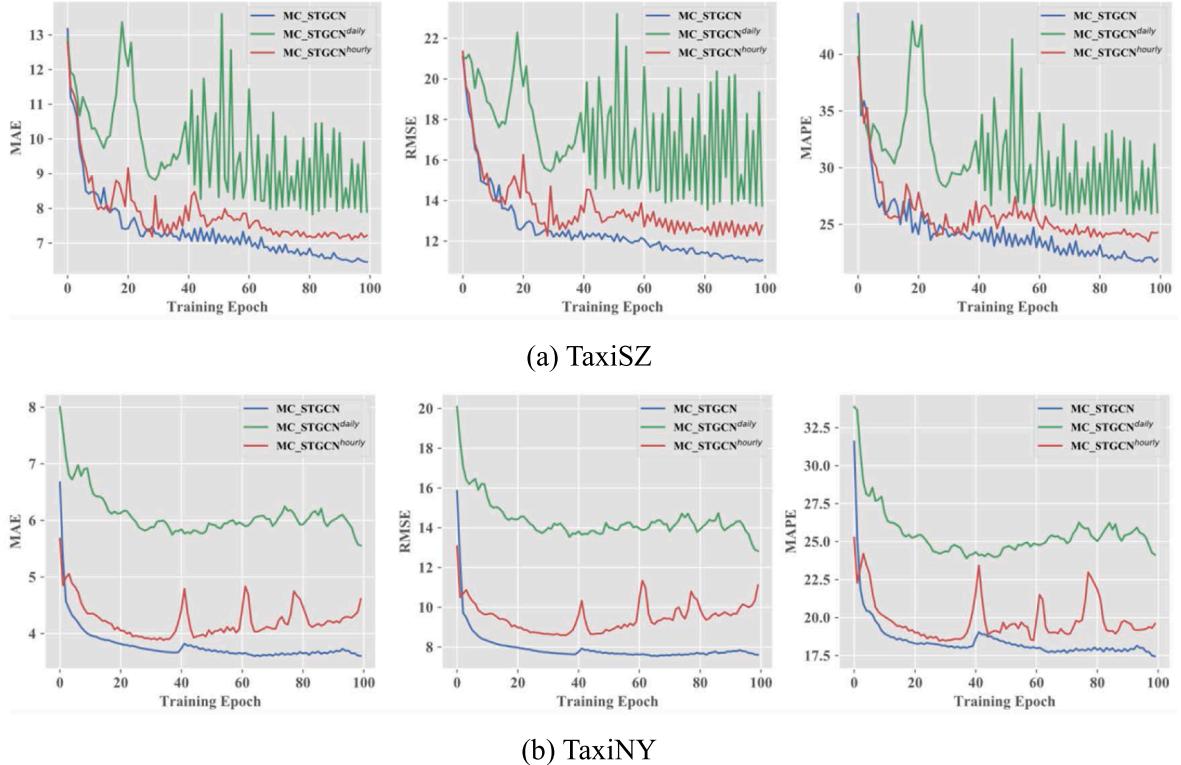


Fig. 9. Performance comparison of the MC_STGCN model and its corresponding models at different epochs on the validation set.

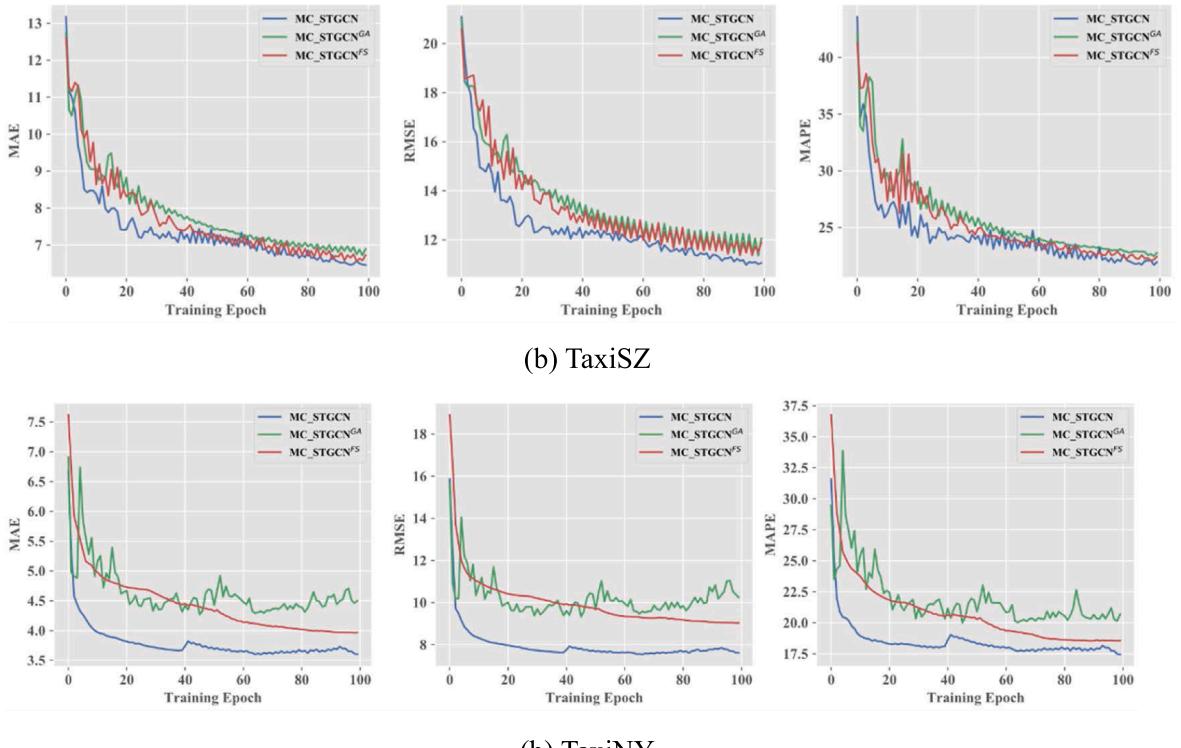


Fig. 10. Performance comparison of the MC_STGCN model and its corresponding models at different epochs on the validation set.

$\text{MC_STGCN}^{\text{hourly}}$ and $\text{MC_STGCN}^{\text{daily}}$, which only considers a single temporal fragment, the MC_STGCN model expresses an obvious improvement by combining the hourly trend, daily cycle, and weekly cycle. In addition, $\text{MC_STGCN}^{\text{hourly}}$ displays better prediction performance than $\text{MC_STGCN}^{\text{daily}}$ on both TaxiSZ and TaxiNY dataset. Compared with the $\text{MC_STGCN}^{\text{hourly}}$ model, the $\text{MC_STGCN}^{\text{daily}}$ model not only produces higher MAE, RMSE, and MAPE, but also generates an obvious shake in the training process on TaxiSZ dataset. The reason for this phenomenon may be that the passenger demand on the daily cycle is more unstable and the correlation of historical observations is relatively weak.

Furthermore, in the model comparison, we find that the MC_STGCN model outperforms the $\text{MC_STGCN}^{\text{GA}}$ and $\text{MC_STGCN}^{\text{FS}}$ models, which indicates that spatial correlation can effectively enhance the accuracy of passenger demand prediction. Fig. 10 further shows the validation errors of the MC_STGCN model and its corresponding models considering spatial correlation with the number of training epoch increases. It can be observed that the MC_STGCN fusing two spatial graphs are superior to two corresponding models during validating processes.

(2) Effect of community detection

In this part, we further discuss how the community division in the prediction module affects the performance. Two community division strategies are compared: the Louvain algorithm based partition and the random partition. More specifically, we adopt the method of random sampling without replacement to divide all regions into 4, 6, 8, and 10 communities. The number of regions in each community is approximately equal. Each sampling is repeated three times to reduce the uncertainty caused by a random partition. We evaluate the prediction results under 13 community divisions on both TaxiSZ and TaxiNY dataset under a 20-minute aggregation interval using the same training, validation, and test set as model comparison introduced before. Table 4 shows the MAE, RMSE, and MAPE results on the test set for the MC_STGCN model compared with different community division strategies.

From the observation of prediction results in the table, we can obtain several findings:

- (1) The number of communities expresses an obvious impact on the prediction performance of passenger demand. As the number of communities grows, the MAE, RMSE, and MAPE of MC_STGCN vary greatly on both TaxiSZ and TaxiNY datasets. We also notice that the performance under 4 and 8 communities are better than that under 6 and 10 communities in TaxiSZ dataset, while the errors under 4 and 6 communities are lower than that under 8 and 10 communities in TaxiNY dataset. This result indicates, in the random division, a larger number of communities may weaken the correlation of passenger demand in different regions, and this then causes the decline of prediction performance.
- (2) The results for different sampling in rounds under the same number of communities are different, which indicates that the way of random matching the regions into a specific community could affect the prediction performance of passenger demand. For instance, in TaxiNY dataset, the errors of the first sampling round under 6 communities are significantly larger than that of the second and third samplings round.
- (3) The Louvain algorithm divides all regions into several communities based on the spatial correlations between regions, which results in an obvious reduction of MAE, RMSE, and MAPE. This further shows that the application of the Louvain algorithm can provide effective improvement for predicting passenger demand in multi-region simultaneously.

(3) Effect of region division

In the proposed model, passenger demand forecasting is performed at the regional level instead of the usual grid level. Thus, in this part, we explore the impact of different region division results on demand forecasting and compare the prediction results of the MC_STGCN model with two widely used grid-based methods on TaxiSZ dataset, including a method using ConvLSTM as the core structure and a method integrating CNN and LSTM models.

The definition of the region in this study combines the division criterion from Urban Planning agency and the structure of the actual road network. In order to explore the possibility of using a finer grid to replace the region, we further analyze the impact of the division of region or grid under different scales on the demand prediction performance. We firstly subdivided the original 101 regions into 198 regions and 370 regions according to the levels of the road in the local network, for example, arterial road, secondary road, or branch road (see Fig. 11). The road network is used to subdivide regions because the road network structure undertakes urban travel demand and is the main division criteria for region division. Finally, urban areas are naturally divided into different irregular regions by complex road network. Moreover, to explore demand prediction performance under the different scale of grid, we divide the study area into grids of different granularity corresponding to the size of the region: grids with the size equal to 1 km (total 104 valid grids); grids with the size equal to 0.7 km (total 194 valid grids); grids with the size equal to 0.5 km (total 379 valid grids). A valid grid means that there is sufficient overlap between the grid and the study area. Fig. 11 shows the result of the region division and grid division.

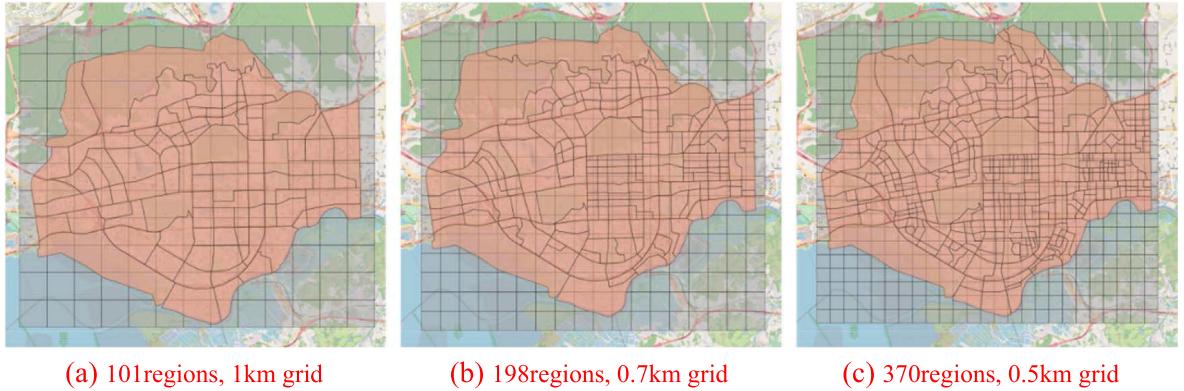
The passenger demand intensity in a grid during a time interval can be obtained by accumulating the order records. The time interval is set as 20 min to be consistent with the MC_STGCN model. Even if the time interval is the same, the passenger demand intensity collected from two kinds of models is still different because of the different divisions (region-based or grid-based) in the study area. For two kinds of prediction models, we also use the same training set, validation set, and test set in performance comparison.

According to the comparing results shown in Table 5, the prediction performance of the MC_STGCN model is generally lower than the two grid-based methods in the comparison. This demonstrates that the proposed deep learning framework can reach a similar or even better level to the commonly used grid-based methods in capturing spatio-temporal correlation of passenger demand from region dimension. Moreover, as the scale of grid/region becomes smaller, the prediction accuracy becomes higher.

Table 4

Performance comparison of the Louvain algorithm and the random partition.

Division strategies	TaxiSZ			TaxiNY		
	MAE	RMSE	MAPE ₁₀ (%)	MAE	RMSE	MAPE ₁₀ (%)
4	First sampling	9.309	15.922	29.460	4.593	9.938
	Second sampling	8.530	14.687	26.763	4.920	10.556
	Third sampling	8.232	14.059	25.803	4.655	10.007
6	First sampling	13.229	26.589	39.301	9.431	30.384
	Second sampling	13.569	24.949	43.733	5.241	15.818
	Third sampling	14.505	31.211	39.801	5.815	14.426
8	First sampling	8.802	15.181	27.614	12.444	29.398
	Second sampling	8.948	15.272	27.806	12.225	33.733
	Third sampling	8.912	15.180	27.798	8.551	21.851
10	First sampling	19.631	37.538	61.095	11.029	28.104
	Second sampling	20.906	37.988	64.667	14.837	36.523
	Third sampling	21.699	41.477	60.899	15.738	40.386
Louvain method	6.555	11.005	21.312	3.727	7.863	19.241

**Fig. 11.** Region division and grids division of Futian District in Shenzhen City.**Table 5**

Comparison of prediction performance between region-based and grid-based models.

Methods	101 region & 1 km grid			198 region & 0.7 km grid			370 region & 0.5 km grid		
	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
CNN + LSTM	6.727	16.291	29.583	3.949	9.328	30.466	3.289	8.762	40.952
ConvLSTM	7.115	12.460	28.208	3.184	6.285	21.396	3.185	6.391	34.650
MC_STGCN	6.555	11.005	21.312	3.971	5.546	28.758	2.727	3.822	36.829

Besides, in order to make an in-depth comparison between the region-based method and the grid-based method, we use heat maps to analyze the spatial distribution of RMSE on different regions/grids, shown in Fig. 12. Under different granular region/grid divisions, the performance of the proposed MC_STGCN expresses stable prediction superiority and outperforms the two grid-based methods. In addition, it can be found from Fig. 12 that the region-based method expresses larger prediction errors in natural scenic areas with sparse population (e.g. Mountains, Parks), and in areas with high human activity intensity, the prediction errors are lower than the commonly used grid-based methods. As for the grid-based methods, the grids with higher prediction errors tend to be located in areas with higher human activity intensity. Moreover, multi-graphs used in the proposed MC_STGCN model can better encoder spatial dependence than uniform grids. Grid-based methods use convolution operations to capture spatial dependence. In this process, irrelevant grid information will inevitably be included in the prediction of the target grid. Therefore, the region-based prediction model has the advantage in spatio-temporal features exploring comparing to the grid-based methods. This can also provide practical insight that the region-based method is more suitable for the area with high intensity of human activities in the application of demand prediction.

5. Conclusion

In this study, we propose an end-to-end neural network-based method MC_STGCN for region-level passenger demand forecasting.

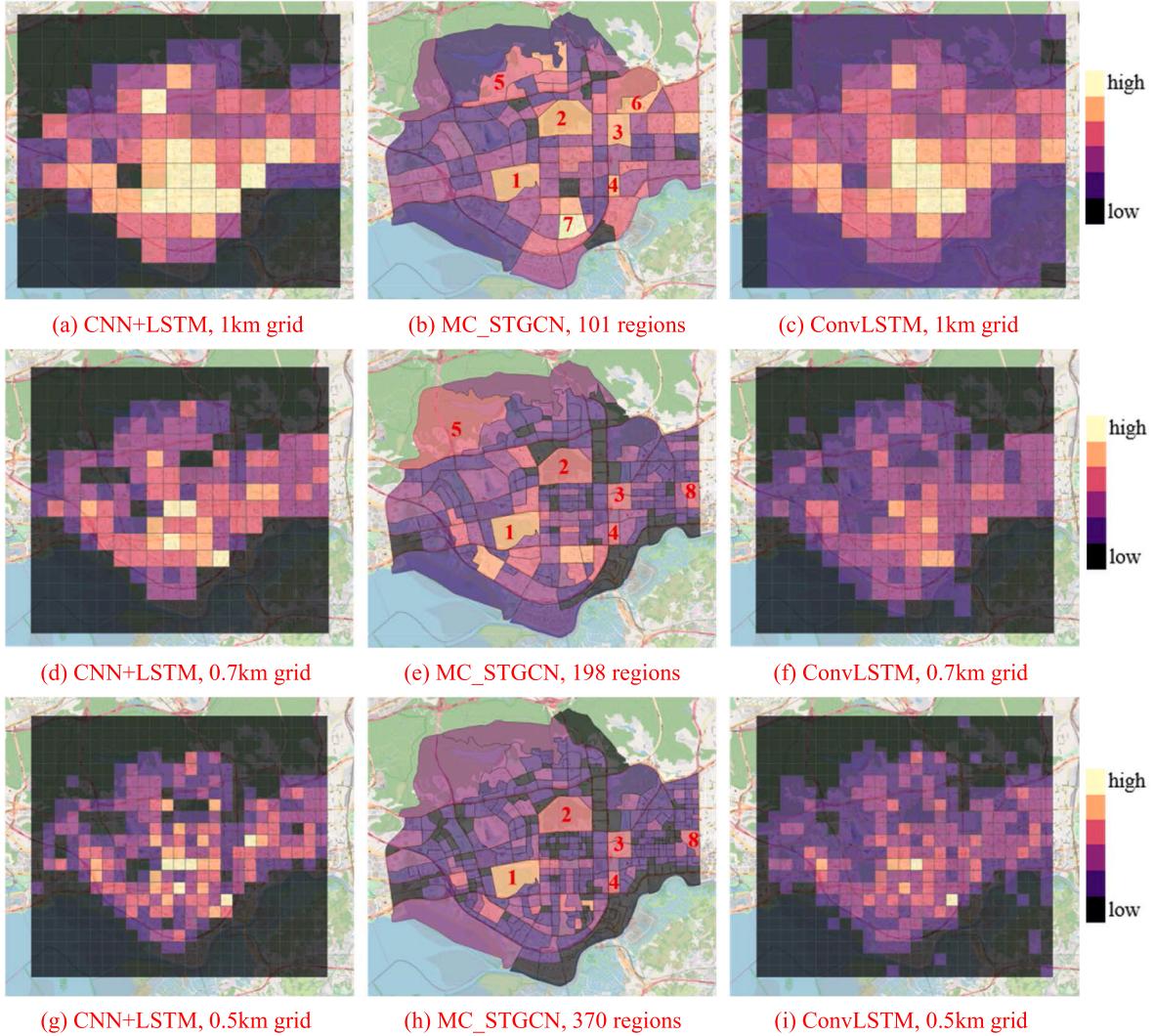


Fig. 12. Spatial distribution of RMSE for the MC_STGCN model and two grid-based models. (1: Shenzhen Golf Club; 2: Lianhuashan Mountain; 3,4: Shenzhen Central Park; 5: Meilin Mountain Park; 6: Bijia Moutain Park; 7: Huanggang Park; 8: Lychee Park).

We convert passenger demand states of all regions into graphs, in which the nodes represent regions, the edges represent the spatial correlations of demand in different regions, and the historical passenger demand observations in the regions are defined as the node attributes. On the one hand, the GAT and the GCN are applied to model temporal and spatial correlation, respectively; on the other hand, a prediction module based on the Louvain algorithm simplifies the problem for simultaneous prediction of multi-regions. We evaluate the proposed model on two large-scale taxi order datasets collected in Shenzhen City and New York City. The experiment results demonstrate that: (1) The MC_STGCN achieves better prediction results than other candidate methods in the comparison, indicating its good ability to capture the spatio-temporal correlation of multi-regions. (2) The MC_STGCN performs better predictions than its corresponding models, implying that keyframe extraction and graph construction can effectively encode the temporal and spatial correlation, respectively. (3) The Louvain algorithm can effectively improve the performance of multi-region predictions.

Although the MC_STGCN model shows high and stable passenger demand prediction accuracy, several limitations of the current study need further to be improved and theses are also the research directions in our future works. First, in this study, there is a sequential order for capturing spatiotemporal correlations (temporal correlation is first captured by GRU and spatial correlation is captured by GCN). It is our future work to develop more efficient algorithms to capture spatiotemporal dependencies at the same time. Second, the external information (e.g., weather and emergency) needs to be integrated into the framework to reduce the impact of fluctuations in demand caused by external events. Finally, spatial dependence (e.g., the connectivity between regions) needs to be further explored. For instance, two regions that are directly connected by main roads, even if they are far away and have different functions, will still express an obvious effect on each other through road networks. Encoding these dependencies more reasonably and feasibly is our future research direction. Therefore, further research plans will focus on how to integrate more information (e.g., complex spatial dependence and related external information) and develop more efficient algorithms to process different inputs.

CRediT authorship contribution statement

Jinjun Tang: Conceptualization, Methodology, Software, Writing - review & editing. **Jian Liang:** Data curation, Methodology, Writing - original draft. **Fang Liu:** Conceptualization, Writing - review & editing, Visualization. **Jingjing Hao:** Software, Validation. **Yinhai Wang:** Writing - review & editing.

Acknowledgments

The research is funded by the National Natural Science Foundation of China (No.71701215), Natural Science Foundation of Hunan Province (No. 2020JJ4752), Innovation-Driven Project of Central South University (No.2020CX041), Foundation of Central South University (No.502045002), Science and Innovation Foundation of the Transportation Department in Hunan Province (No.201725), Postdoctoral Science Foundation of China (No.2018M630914 and 2019T120716).

References

- About TLC - TLC [WWW Document], n.d. URL <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page> (accessed 2.11.20).
- Ahmed, M.S., Cook, A.R., 1979. Analysis of freeway traffic time-series data by using bos-jenkins techniques. *Transp. Res. Rec.* 722.
- Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *J. Stat. Mech: Theory Exp.* 2008 (10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- Chai, D., Wang, L., Yang, Q., 2018. Bike flow prediction with multi-graph convolutional networks. In: *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 397–400.
- Chen, P.-C., Hsieh, H.-Y., Sigalingging, X.K., Chen, Y.-R., Leu, J.-S., 2017. Prediction of station level demand in a bike sharing system using recurrent neural networks. In: *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, pp. 1–5.
- Chen, Q., Song, X., Yamada, H., Shibasaki, R., 2016. In: Learning deep representation from big and heterogeneous data for traffic accident inference. AAAI Press, Phoenix, Arizona, pp. 338–344.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Chung, J., Gulcehre, C., Cho, K., Bengio, Y., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Davis, N., Raina, G., Jagannathan, K., 2016. A multi-level clustering approach for forecasting taxi passenger demand. In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. Presented at the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 223–228.
- Defferrard, M., Bresson, X., Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. *Adv. Neural Inform. Process. Syst.* 3844–3852.
- Dong, X., Lei, T., Jin, S., Hou, Z., 2018. Short-Term Traffic Flow Prediction Based on XGBoost. In: *2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS)*. Presented at the 2018 IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS), pp. 854–859.
- Fu, R., Zhang, Z., Li, L., 2016. Using LSTM and GRU neural network methods for traffic flow prediction, in: *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. Presented at the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), pp. 324–328.
- Geng, X., Li, Y., Wang, L., Zhang, L., Yang, Q., Ye, J., Liu, Y., 2019. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. *AAAI* 33, 3656–3663.
- Hamed Mohammad M., Al-Masaeid Hashem R., Said Zahi M. Bani, 1995. Short-term prediction of traffic volume in urban arterials. *Journal of Transportation Engineering* 121, 249–254.
- Kankanamge, K.D., Witharana, Y.R., Withanage, C.S., Hansini, M., Lakmal, D., Thayavaram, U., 2019. Taxi Trip Travel Time Prediction with Isolated XGBoost Regression. In: *2019 Moratuwa Engineering Research Conference (MERCon)*. Presented at the 2019 Moratuwa Engineering Research Conference (MERCon), pp. 54–59.
- Ke, J., Yang, H., Zheng, H., Chen, X., Jia, Y., Gong, P., Ye, J., 2018. Hexagon-based convolutional neural network for supply-demand forecasting of ride-sourcing services. *IEEE Trans. Intell. Transp. Syst.* 20, 4160–4173.
- Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2019). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. *arXiv preprint arXiv:1910.09103*.
- Ke, J., Zheng, H., Yang, H., Xiqun, Chen, 2017. Short-Term Forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach. *Transport. Res. Part C: Emerging Technol.* 85, 591–608.
- Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
- Li, X., Pan, G., Wu, Z., Qi, G., Li, S., Zhang, D., Zhang, W., Wang, Z., 2012. Prediction of urban human mobility using large-scale taxi traces and its applications. *Front. Comput. Sci.* 6 (1), 111–121.
- Lin, L., He, Z., Peeta, S., 2018. Predicting station-level hourly demand in a large-scale bike-sharing network: a graph convolutional neural network approach. *Transport. Res. C: Emerg. Technol.* 97, 258–276.
- Liu, L., Qiu, Z., Li, G., Wang, Q., Ouyang, W., Lin, L., 2019. Contextualized spatial-temporal network for taxi origin-destination demand prediction. *IEEE Trans. Intell. Transp. Syst.* 20, 3875–3887.
- Ma, X., Dai, Z., He, Z., Ma, J., Wang, Yong, Wang, Yunpeng, 2017. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction. *Sensors* 17, 818.
- Moreira-Matias, L., Gama, J., Ferreira, M., Damas, L., 2012. A predictive model for the passenger demand on a taxi network. In: *2012 15th International IEEE Conference on Intelligent Transportation Systems*. IEEE, pp. 1014–1019.
- Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., Damas, L., 2013. Predicting taxi-passenger demand using streaming data. *IEEE Trans. Intell. Transp. Syst.* 14, 1393–1402.
- Okutani, I., Stephanedes, Y.J., 1984. Dynamic prediction of traffic volume through Kalman filtering theory. *Transport. Res. B: Methodol.* 18, 1–11.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., Woo, W., 2015. Convolutional LSTM Network: a machine learning approach for precipitation nowcasting, in: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, NIPS'15*. MIT Press, Montreal, Canada, pp. 802–810.
- Sun, S., Zhang, C., Yu, G., 2006. A Bayesian network approach to traffic flow forecasting. *IEEE Trans. Intell. Transp. Syst.* 7, 124–132.
- Tang, J., Liu, F., Zou, Y., Zhang, W., Wang, Y., 2017. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic. *IEEE Trans. Intell. Transport. Syst.* 18 (9), 2340–2350.
- Tang, J., Zhang, G., Wang, Y., Wang, H., Liu, F., 2015. A hybrid approach to integrate fuzzy c-means based imputation method with genetic algorithm for missing traffic volume data estimation". *Transport. Res. C: Emerg. Technol.* 51, 29–40.
- Tobler, W.R., 1970. A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* 46, 234.
- Tong, Y., Chen, Y., Zhou, Z., Chen, L., Wang, J., Yang, Q., Ye, J., Lv, W., 2017. The simpler The better: a unified approach to predicting original taxi demands based on large-scale online platforms. In: *Halifax, N.S. (Ed.)*. Presented at the the 23rd ACM SIGKDD International Conference, ACM Press. Canada, pp. 1653–1662.
- Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M., 2019. Dynamic graph cnn for learning on point clouds. *ACM Trans. on Graphics (TOG)* 38, 1–12.

- Wu, Y., Tan, H., Qin, L., Ran, B., Jiang, Z., 2018. A hybrid deep learning based traffic flow prediction method and its understanding. *Transport. Res. C: Emerg. Technol.* 90, 166–180.
- Wu, Z., Pan, S., Long, G., Jiang, J., Zhang, C., 2019. Graph WaveNet for Deep Spatial-Temporal Graph Modeling. Twenty-Eighth International Joint Conference on Artificial Intelligence IJCAI-19.
- Xu, J., Rahmatizadeh, R., Boloni, L., Turgut, D., 2018. Real-time prediction of taxi demand using recurrent neural networks. *IEEE Trans. Intell. Transp. Syst.* 19, 2572–2581.
- Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., Li, Z., 2018. Deep multi-view spatial-temporal network for taxi demand prediction, Presented at the Thirty-Second AAAI Conference on Artificial Intelligence.
- Yi, L., Su, H., Guo, X., Guibas, L., 2017. SyncSpecCNN: Synchronized spectral CNN for 3D shape segmentation. In: Presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6584–6592.
- Yu, B., Yin, H., Zhu, Z., 2018a. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, pp. 3634–3640.
- Yu, H., Chen, X., Li, Z., Zhang, G., Liu, P., Yang, J., Yang, Y., 2019. Taxi-based mobility demand formulation and prediction using conditional generative adversarial network-driven learning approaches. *IEEE Trans. Intell. Transp. Syst.* 20, 3888–3899.
- Yu, H., Wu, Z., Wang, S., Wang, Y., Ma, X., 2017. Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks. *Sensors* 17, 1501.
- Yu, H.T., Jiang, C.J., Xiao, R.D., Liu, H.O., Lv, W., 2018b. Passenger flow prediction for new line using region dividing and fuzzy boundary processing. *IEEE Trans. Fuzzy Syst.* 27 (5), 994–1007.
- Zhang, J., Zheng, Y., Qi, D., 2017. Deep spatio-temporal residual networks for citywide crowd flows prediction, Presented at the Thirty-First AAAI Conference on Artificial Intelligence.
- Zhang, J., Zheng, Y., Qi, D., Li, R., Yi, X., 2016. DNN-based prediction model for spatio-temporal data. In: Press, A.C.M. (Ed.), Presented at the the 24th ACM SIGSPATIAL International Conference. Burlingame, California, pp. 1–4.
- Zhang, K., Liu, Z., Zheng, L., 2019. Short-term prediction of passenger demand in multi-zone level: temporal convolutional neural network with multi-task learning. *IEEE Trans. Intell. Transp. Syst.* 1–11 <https://doi.org/10.1109/TITS.2019.2909571>.
- Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., 2019. T-GCN: a temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* 1–11.
- Zhou, F., Lin, Y., 2016. Fine-grained image classification by exploring bipartite-graph labels. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1124–1133.
- Zhou, X., Shen, Y., Huang, L., Zang, T., Zhu, Y., 2019. Multi-level attention networks for multi-step citywide passenger demands prediction. *IEEE Trans. Knowl. Data Eng.* 1–11 <https://doi.org/10.1109/TKDE.2019.2948005>.