

# Exploiting Spatiotemporal Correlations of Arrive-Stay-Leave Behaviors for Private Car Flow Prediction

Chenxi Liu, Zhu Xiao, *Senior Member, IEEE*, Dong Wang, *Member, IEEE*, Lei Wang, Hongbo Jiang, *Senior Member, IEEE*, Hongyang Chen, *Senior Member, IEEE*, Jiangxia Yu

**Abstract**—Accurate prediction of private car flows in urban regions is strategically vital for constructing smart cities. Private car flows are essentially reflected in the *arrive-stay-leave* (ASL) behaviors of car users. Specifically, to fulfill daily travel needs, private car users usually *arrive* at specified locations, *stay* for a certain period participating in their activities, then *leave* to the next destination. In this paper, we strive to explore global spatiotemporal correlations of ASL behaviors to predict private car flow. Therefore, we formulate the dynamic distribution of ASL behaviors in urban regions through multiple graphs and propose the novel *multigraph dense convolutional network* (MGDCN) to represent spatiotemporal correlations of ASL behaviors for private car flow prediction. The proposed MGDCN is an end-to-end framework consisting of three modules: *i*) multigraph dense convolutions, in which we introduce densely connected blocks to aggregate global spatial correlations effectively, *ii*) convolutional gated recurrent units to capture sequential temporal correlations of ASL behaviors in all regions, and *iii*) attention networks for learning the stay duration correlations in each region. Experiments based on large-scale private car trajectory data and region-of-interest (ROI) datasets verify that the proposed method achieves superior performance than the baselines in all five metrics.

**Index Terms**—Private Car, Arrive-Stay-Leave, Region-of-Interest, Graph Convolutional Network.

## I. INTRODUCTION

Recent years have witnessed the continuous increase in private cars [1]. Taking China as an example, according to the Statistical Yearbook of China<sup>1</sup>, the average annual growth in private cars was 19.96 million from 2014 to 2019. By the end of 2019, the number of small passenger cars reached 220 million, 94% of which were private cars. In Europe, according to statistics from the Directorate-General for Energy and Transport of the European Commission [2], an average family can own one or two private cars, and it is estimated that the proportion of private car traffic activities will remain above 70% from 1990 to 2030; private cars occupy the main body of transportation activities in Europe. As seen in the

C. Liu, Z. Xiao, D. Wang, L. Wang, and H. Jiang are with the College of Computer Science and Electronic Engineering, Hunan University, Changsha, 410082, China. (e-mail: {cxliu, zhxiao, wangd}@hnu.edu.cn, wangl-hnu@gmail.com, hongbojiang2004@gmail.com, Corresponding authors: Zhu Xiao, Dong Wang, and Hongyang Chen.)

H. Chen is with the Research Center for Big Data Intelligence, Zhejiang Lab, Hangzhou, 311100, China. (e-mail: dr.h.chen@ieee.org)

J. Yu is with the an associate professor with the School of Economics & Management, Xidian University, Shaanxi, 710126, China. (e-mail: yujiangxia@126.com)

<sup>1</sup><https://www.mps.gov.cn/n2254314/n6409334/c6852472/content.html>

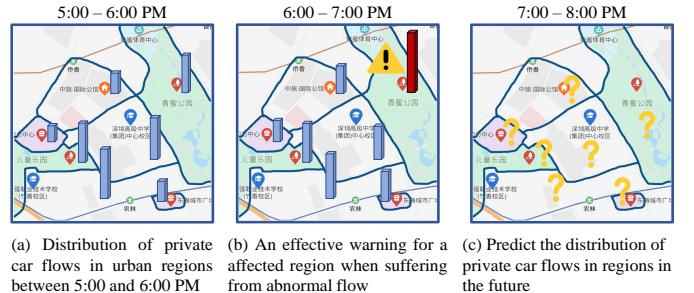


Fig. 1: An Example of Private Car Flow Prediction.

report<sup>2</sup>, vehicle ownership per person in the U.S. was 0.756 in 2015 and increased to 0.766 by 2016. In this context, private cars constitute a significant part of urban transportation [3].

The rapid growth in private cars has created many private car flows since increasingly more people drive private cars to travel among urban regions to perform daily commutes and meet travel needs [4], [5]. Accurate prediction of private car flows in urban regions is strategically vital for the construction of smart cities, such as smart parking [6], risk assessment [7], and public transportation planning [8]. Figure 1(a) depicts the distribution of private car flows in the early evening. Figure 1(b) shows that predicting private car flows can send a helpful warning for a possibly affected region when suffering from abnormal flows. In real life, foreseeing urban regions with abnormal flows helps the government implement traffic control and prevent tremendous risks to public safety. In Figure 1(c), predicting private car flows in all regions can reflect basic and flexible parking needs and provide an authoritative reference for the planning and construction of parking lots, thereby improving the utilization rate of transportation facilities.

Private car flows are primarily reflected in the *arrive-stay-leave* (ASL) behaviors of car users. For instance, to fulfill daily travel needs, private car users usually *arrive* at specified locations, *stay* for a certain period participating in their activities, then *leave* to the next destination. Private car flows exhibit ASL behaviors and thus are distinctly different from other types of car flows (e.g., taxi demands or bicycle flows), which only consider arrival (inflow) or leaving (outflow) and ignore the original stay volumes of cars. Particularly, the stay duration in the ASL behavior provides crucial information for

<sup>2</sup><http://www.umich.edu/~umtriswt>

understanding user travel behavior and its connection with the urban aggregation effect [9]. When private cars arrive at a specified region and spend a longer stay duration, the region has large attractiveness and a strong aggregation effect [5]. In other words, the longer the average stay duration in a region, the more attractive the region is. Therefore, apart from the spatial correlations and sequence temporal correlations (i.e., the location and time) of ASL behaviors, capturing the stay duration correlations is essential in private car flow prediction. In recent years, extensive studies have been conducted to predict car flows by exploring the spatiotemporal correlations of traffic data. Researchers usually model data as spatiotemporal graphs, use graph neural networks (GNNs) to extract the spatial correlation of geographic locations, and use recurrent neural networks (RNNs) to extract the temporal correlation of traffic time series [10], [11], [12], [13]. However, these methods have difficulty extracting global spatiotemporal correlations. For spatial correlations, most of these studies mainly focus on static undirected graphs and overlook semantic information; that is, the vertices of the graphs lack semantic information, and the edges connecting vertices are usually static spatial attributes. Although it is necessary to consider the static spatial attributes, the car flow is tightly related to the vertical functions and dynamic multidirectional interactions between the vertices. For temporal correlations, these works mainly focus on sequential temporal correlations. Sequence time is only a part of temporal correlations, and stay duration is the essential temporal attribute of private car travel, reflecting the urban agglomeration effect and should not be ignored in private car flow prediction tasks.

Despite these efforts, it is not straightforward to apply recent advances in deep learning to private car flow prediction since the following two challenges remain. *i)* Global spatial correlations & region functions. First, each region's function (e.g., region categories) will affect the ASL behaviors of private cars. In addition, during driving, private cars perform ASL behaviors, reflecting the dynamic interactions between regions. As such, how to formulate the global spatial correlations of ASL behaviors in different categories of functional regions and design an effective model for aggregating these correlations. *ii)* Global temporal correlations of ASL behaviors. The existing works apply sequence methods to extract the sequential temporal correlations [10], [12]. Nevertheless, in practical applications, in addition to sequential temporal correlations, stay duration correlations provide a crucial attribute reflecting the regularity of ASL behaviors. For instance, given regions with the same private car arrival and leaving flow, a considerable stay duration indicates substantial attractiveness. Existing works treat the regions as a whole and model sequential time series, while the stay duration correlation of ASL behaviors in each region is often overlooked. Therefore, their average performance for all regions may be excellent, but there is less explanation for the prediction effect, especially for individual regions. Hence, how to effectively capture the global temporal correlations of ASL behaviors remains unsolved.

To solve the abovementioned challenges, we first propose a topic clustering algorithm based on urban regions of interest (ROIs) to identify irregular urban regions with semantic

functions, which we call urban functional regions (UFRs). Second, we formulate the dynamic distribution of private car ASL behaviors through four collaborative graphs with various topologies, including distance, function, flow similarity, and transition graphs. Finally, we propose a novel **MultiGraph Dense Convolutional Network (MGDCN)** to explore spatiotemporal correlations of ASL behaviors for private car flow prediction. This end-to-end framework combines multigraph dense connected convolutions, convolutional gated recurrent networks, and attention networks to capture the complex correlations between regions from both spatial and temporal aspects. The main contributions of this paper are summarized as follows:

- We propose a multigraph dense convolutional network (MGDCN) to predict private car flows in UFRs through exploiting spatiotemporal correlations of ASL behaviors. Private car ASL behaviors are formulated by multiple graphs with various topologies. Notably, we propose a topic clustering algorithm to identify UFRs.
- The proposed MGDCN, on the one hand, expands the densely connected blocks to multigraph convolutional networks to model the global spatial correlations; on the other hand, MGDCN leverages convolutional gated recurrent units to capture the sequential temporal correlations. In addition, an attention network is proposed to learn the stay duration correlation of ASL behaviors in each UFR.
- We conduct experiments based on real-world private car trajectory data and ROI datasets. The experimental results demonstrate that the proposed method achieves better performance than the baselines in all five metrics. We also make our private car trajectory dataset publicly available<sup>3</sup>.

The remainder of this paper is organized as follows. Section 2 summarizes related work. Section 3 presents the preliminaries and the methodological framework of MGDCN. Section 4 details the three main components of MGDCN in detail. Experiments and further analysis are presented in Section 5. Finally, we conclude this paper in Section 6.

## II. RELATED WORK

Over the past decade, a plentiful number of works have addressed car flow prediction, including predicting traffic demand [14], mobility events [15], and OD (origin-destination) flows [16]. The task construct is based on different types of trajectory data, e.g., taxi OD demand, bus, or smart card data. Yu et al. proposed an adversarial architecture to capture the spatiotemporal distribution of taxi hotspots [17]. In [18], the authors present a deep learning architecture to forecast the metro inbound/outbound passenger flow. Qi et al. predicted regional mobility patterns of bus travelers using smartcard data [19]. However, little focus is on the vast quantity of private car data generated every day. Unlike travel with floating cars (taxis or buses) [20], people driving private cars in the city exhibit unique travel patterns, and they have apparent ASL behaviors [21]. Other types of transportation do not have sufficient information about ASL behavior and stay duration.

<sup>3</sup><https://github.com/HunanUniversityZhuXiao/PrivateCarTrajectoryData>

For example, taxi destinations are determined by different passengers [22], and taxi trajectories address only the travel behaviors of taxis rather than the owner's intention. Likewise, in the case of buses, their departure time and travel routes are planned. Our work attempts to bring a new perspective for studying car flow prediction based on large-scale private car trajectory data.

In early studies [23], [24], researchers usually modeled vehicle trajectory data into a time-series vector and then applied traditional methods to predict future traffic flow. The main methods of this type of study include gradient boosting regression tree (GBRT) [25], support vector regression (SVR) [26], autoregressive integrated moving average (ARIMA) [23], kernel density estimation (KDE) [5], and Kalman filtering [24]. These models rely on the ideal stationary assumption; in fact, real spatiotemporal data rarely meet this assumption. Thus, they fail to capture more complex spatiotemporal correlations of data. Later, deep learning methods such as RNNs (e.g., LSTM or GRU) and CNNs provided impressive performance in modeling spatiotemporal data [27], [28]. However, RNNs are incapable of parallel computing and overlooking spatial correlation modeling. CNNs restrict the model to processing tensor structures. Therefore, the non-Euclidean correlations in the real world are not considered. In addition, these studies divided the city into Euclidean tensors based on the grid division method and then straightforwardly applied convolutional networks for automatic representation learning [29]. Due to the restrictions of the road network, urban regions are often irregular, and their data structure is non-Euclidean [30]. Even though dividing the city into grids is a feasible solution, it is very sparse and difficult to maintain the correlation between two distant regions. Until recently, car flow prediction research at the urban irregular region level was still at an early stage.

However, the city is composed of various kinds of functional regions, where people work, study, entertain, and live [31]. Irregular regions with semantic information in cities, which we call urban functional regions [32], [33], are basic units of city and as a carrier of social and economic activities [34]. In this study, we focus on prediction at the level of urban functional regions to understand human mobility more intuitively.

Recently, emerging graph convolutional networks have been proven to be effective for non-Euclidean data embedding [35]. In this line, scholars model the data as spatiotemporal graphs for different tasks. Chen et al. constructed a heterogeneous graph to predict context-aware mobile application usage [36]. In [14], the authors proposed a transition graph for traffic demand prediction. In these studies, graphs are directly constructed based on a single attribute (e.g., physical topology). However, the relationship between the vertices of the graph in real life is complicated and dynamic. How to build graphs is a crucial step, and the flexible design of graph topology for specific tasks is the key to solving the problem. Our study formulated the dynamic distribution of private car ASL behaviors in urban regions through multiple graphs with various topologies, which can effectively capture the spatiotemporal correlations of ASL behaviors in UFRs.

Some studies are more relevant to us through multigraph

modeling. Chai et al. proposed a multigraph convolutional network to predict station-level traffic flows, with LSTM as the encoder-decoder [37]. However, they could not capture the semantic correlation between bicycle stations. Some scholars added semantic information to multigraph modeling. Geng et al. proposed multigraph convolution and contextual gated RNN to predict ride-hailing demand [38]. In [39], an attention multigraph convolutional sequence-to-sequence model was proposed for station-level flow prediction. Nevertheless, they ignore multidirectional interactions between regions and the stay duration correlation of each region. In contrast, we consider the regions' functional information by constructing a function graph and a transition graph to model the dynamic multidirectional interactions between these UFRs. In addition, our proposed MGDCN method learns the stay duration correlations of ASL behaviors in each UFR.

### III. PRELIMINARIES

In this section, we first introduce the real-world datasets in this study. Then, we present the essential definitions and problem statement. Finally, we give an overview of the proposed MGDCN.

#### A. Data Description

In this subsection, we introduce real-world private car trajectories and ROI datasets. Through collaboration with industry, we collect large-scale, encrypted private car trajectory data from the MapGoo intelligent cloud. This cloud service platform helps developers quickly access various terminal devices (such as GPS devices, OBD devices, and cloud recorders). The data collection process is ongoing, and the total size of the collected trajectory dataset has reached several TB.

All sensitive information about private cars is deleted from the original trajectory. After obtaining desensitized private car trajectory data, we extract private car ASL behavior data from the trajectory data [40], [5], [4]. An example of the ASL behavior data of private car users is illustrated in Table I.

ROI data are obtained from the AMAP. Unlike POIs (points of interest), ROIs refer to integrating irregular urban regions with specific functions. To generate the aggregated urban regions' main functions, we divide ROIs again into nine subcategories, as presented in Table II.

#### B. Definitions

In this subsection, we introduce essential definitions, which will be used throughout this study. Furthermore, we summarize the main notations used in this article in Table III.

**Definition 2.1 (Arrive-Stay-Leave Behavior).** Given a set of private cars  $\mathcal{P} = \{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{|\mathcal{P}|}\}$ . Each  $\mathbf{P}_{id} = (id, t_s, t_r, lon, lat, st)$  denotes a private car's ID, stop timestamp, restart timestamp, longitude, latitude, and stay duration. Specifically, the stop timestamp represents the time when the private car starts to park, and the restart timestamp represents the next time when the private car is started. The timestamp is Beijing time, and the coordinate type is WGS84ll.

**Definition 2.2 (Private Car Flow).** In our case, private car flows are the total number of private cars staying in regions

TABLE I: An Example of Private Car Arrive-Stay Behaviors.

Private Car ID	Stop Timestamp	Restart Timestamp	Longitude	Latitude	Stay Duration
97811	2018/09/01 09:52:45	2018/09/01 16:48:39	113.7885289	22.69536539	06:55:54
97811	2018/09/01 17:32:33	2018/09/01 19:00:08	113.7884202	22.69562659	01:27:35
116857	2018/09/01 00:59:52	2018/09/01 02:11:55	113.8658922	22.57464066	01:12:03
116857	2018/09/01 03:18:01	2018/09/01 19:33:20	113.8678928	22.57641296	16:14:19
116857	2018/09/01 20:58:45	2018/09/01 21:08:00	113.8679796	22.57648784	00:09:15
176160	2018/09/01 15:53:55	2018/09/01 16:24:11	113.9101283	22.55945648	00:30:16

TABLE II: Main Functions of Urban Regions.

Categories	Big Category in AMAP	Code
Life Service	Shopping, Food, Daily Life	1
Park	Tourist Attraction, Pass Facilities	2
Profession	Medical Service, Governmental Organization, Finance	3
Vehicle Service	Auto Service, Auto Dealers, Auto Repair, Motorcycle Service	4
Education	Science and Educational Service	5
Enterprise	Enterprises	6
Residence	Accommodation, Commercial House	7
Transportation	Transportation Service, Place Name, Road Furniture	8
Sport	Sport and Recreation	9

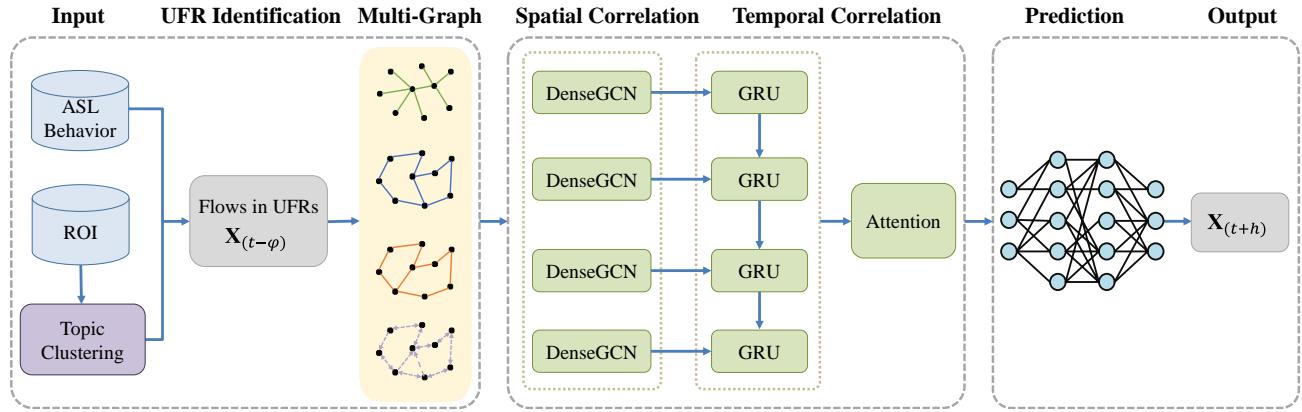


Fig. 2: Overview of the Proposed MGDCN.

TABLE III: Main Notations and Descriptions

Notation	Description
$\mathcal{P}$	Private car set
$\mathcal{U}$	Urban regions set
$\mathcal{R}$	ROI set
$\mathbf{w}_{f_i}$	Function set
$w_{f_{ij}}$	Word terms
$\mathcal{G}$	Spatiotemporal graph
$\mathbf{x}_t$	Private car flows in $[t, t + \tau]$
$\mathbf{x}_t^a$	Arrival flows from other regions in $[t, t + \tau]$
$\mathbf{x}_t^s$	Original private car stay volume in $[t, t + \tau]$
$\mathbf{x}_t^l$	Leaving flows to other regions in $[t, t + \tau]$
$\mathbf{T}_{ij,t}$	Region-to-region-transition
$W_{i,st}$	Average stay duration in region $i$
$\mathbf{A}_D$	Adjacency matrix of distance graph
$\mathbf{A}_S$	Adjacency matrix of flow similarity graph
$\mathbf{A}_F$	Adjacency matrix of function graph
$\mathbf{A}_T$	Adjacency matrix of transition similarity graph
$C^{l+1}$	Dense connectivity
$\mathbf{A}_m$	Aggregation matrix of multigraphs
$a_i$	Attention score
$\hat{\mathbf{X}}_{t+h}$	Predicted flows
$\mathbf{X}_{t+h}$	Real flows

during a period. It is expressed as  $\mathbf{X} \in \mathbb{R}^{N \times T}$ , where  $N$  represents the number of regions, and  $T$  represents the time

intervals. Specifically,  $\mathbf{x}_t$  denotes the private car flows in urban regions with time interval  $[t, t + \tau]$ , which consists of three parts. The first part is the arrival flows  $\mathbf{x}_t^a$ , which refer to private cars arriving from other regions. The second part is the original stay volume  $\mathbf{x}_t^s$ , which refers to the number of private cars staying in these regions before the current time. Leaving flows  $\mathbf{x}_t^l$  are the third part, which indicate that they leave for other regions.

$$\mathbf{x}_t = \mathbf{x}_t^a + \mathbf{x}_t^s - \mathbf{x}_t^l, \quad (1)$$

**Definition 2.3 (Urban Functional Regions).** Given UFR set  $\mathcal{U} = \{u_1, u_2, \dots, u_{|\mathcal{U}|}\}$  and ROI set  $\mathcal{R} = \{r_1, r_2, \dots, r_{|\mathcal{R}|}\}$ , each UFR  $u_i$  is a tuple  $(id, name, location, area, function)$ , and it is composed of different ROIs, where each  $r_i$  denotes an ROI.  $\mathbf{r}_i$  is a tuple  $(id, name, location, area, category)$ , and the location denotes the longitude and latitude of the center point of the ROI.

**Definition 2.4 (Multiple Graphs).** A graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{A})$ ,  $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$  is the set of all vertices,  $|\mathcal{V}| = N$ .  $(v_i, v_j) \in \mathcal{E}$  describes the edges between vertices.  $\mathbf{A} \in \mathbb{R}^{N \times N}$  is a weighted adjacency matrix that associates each edge  $(v_i, v_j)$  with its element  $\mathbf{A}_{ij}$ . We formulate multiple correlations in UFRs as multiple graphs (in Section 4.2).

**Definition 2.5 (Region-to-Region Transition).** A region-to-region transition denoted as  $T_{ij,t}$ , which contains the following three items: origin region  $u_o$ , destination region  $u_d$ , and transition time  $t$ .

**Problem Formulation** The private car flows observed on  $\mathcal{G}$  at the  $t$ -th interval are denoted as  $\mathbf{x}_t \in \mathbb{R}^{N \times i}$ . Given historical  $\varphi$  interval private car flows  $\mathbf{X}_{t-\varphi} = \{\mathbf{x}_{t-\varphi}, \dots, \mathbf{x}_t\}$ , private car flow prediction aims to learn a function  $f$  that can predict flows  $\mathbf{X}_{t+h} = \{\mathbf{x}_{t+1}, \dots, \mathbf{x}_{t+h}\}$  in the next  $h$  interval:

$$[\mathbf{X}_{t-\varphi}; \mathcal{G}] \xrightarrow{f} [\mathbf{X}_{t+h}], \quad (2)$$

### C. Framework Overview

Figure 2 depicts an overview of the proposed MGDCN, which consists of the following essential tasks. (1) Identifying UFRs and combining them with the ASL behaviors to represent the private car flows in UFRs. (2) Using multiple graphs to formulate different types of correlations among regions. (3) Developing multigraph densely connected convolutions for capturing the global spatial correlations of multiple graphs. (4) Combining multigraph dense convolutions with gated recurrent units to capture the temporal correlations. (5) Using attention networks to learn the importance of each UFR. (6) Applying a multilayer feed-forward network to predict private car flows.

## IV. MGDCN: MULTIGRAPH DENSE CONVOLUTIONAL NETWORKS

In this section, we detail the various components of the multigraph dense convolutional networks. The pipeline of the proposed MGDCN is depicted in Algorithm 1.

### A. Urban Functional Region Identification

Urban functional regions reflect city's spatial structures [41], in this study, we propose a topic clustering algorithm to divide the city into UFRs with different topics based on urban ROIs. The inputs are the ROI set  $\mathcal{R}$ , the number of clusters  $m$ , and the maximum number of iterations, which we set to 100. The output is UFR set  $\mathcal{U}$ .

As shown in Algorithm 2, ROIs with similar distances are divided into the same subregion by the k-means++ clustering algorithm [42] with Haversine distance. Then, the latent Dirichlet allocation algorithm is applied to mine the topic function of each region. Specifically, we extract the terms that belong to the same functionality in all the functional topics. As shown in Table II, it is recorded as a set  $\mathbf{w}_{f_{ij}} = \{w_{f_{i1}}, w_{f_{i2}}, \dots, w_{f_{ik}}\}$ , and the strength of the function  $i$  in the functional region is calculated.

$$\mathbf{f}_i = \sum_{j=1}^n c \times p(w_{f_{ij}}) \times q(w_{f_{ij}}) \times \mathbf{w}(i), \quad (3)$$

where  $p(w_{f_{ij}})$  indicates the probability that the topic of the term  $w_{f_{ij}}$  belongs to the UFR document,  $q(w_{f_{ij}})$  indicates the probability of the term  $w_{f_{ij}}$  for the topic it belongs to,  $c$  is a constant set to prevent floating-point underflow during the calculation, and  $\mathbf{w}(i)$  is a weight value set for function  $i$ . We set the UFR with the highest  $\mathbf{f}_i$  to be the topic function.

---

### Algorithm 1: Multigraph Dense Convolutional Networks

---

**Input:**  $\mathcal{P} = \{\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_{|\mathcal{P}|}\}, \mathcal{U} = \{u_1, u_2, \dots, u_{|\mathcal{U}|}\}$ ;  
**Output:** Private car flow in the next  $h$  time interval.

- 1 Construct  $\mathbf{A}_D$  distance graph by geographic distance;
- 2 Construct  $\mathbf{A}_S$  similarity graph by historical pattern;
- 3 Construct  $\mathbf{A}_F$  function graph by ROI category;
- 4 Construct  $\mathbf{A}_T$  transition graph by transition flow;
- 5 **for**  $graph$  in  $[\mathbf{A}_D, \mathbf{A}_S, \mathbf{A}_F, \mathbf{A}_T]$  **do**
- 6     Train the densely connected GCN model with a graph;
- 7     Use DGRU to extract sequence temporal correlations  $\mathbf{H}_t[i, :]$ ;
- 8     **for** each  $u_i \in \mathcal{U}$  **do**
- 9         average stay duration  $W_{i,st}$  as the attention weight  
 $\mathbf{e}_i = (W_{i,st} \mathbf{H}_t)^T \mathbf{u};$   
 $\mathbf{a}_i = SoftMax(\mathbf{e}_i) = \frac{\exp(\mathbf{e}_i)}{\sum_{i=1}^k \exp(\mathbf{e}_i)};$
- 10     **endfor**
- 11     Predict private car flows by the above three components.
- 12     **endfor**
- 13     Aggregate the results of multiple graphs.

---

### B. Multigraph Construction and Refinement

In this subsection, we identify the non-Euclidean correlations between UFRs and model these correlations into multiple graphs, including distance, flow similarity, transition, and function graphs. As shown in Figure 3, we take nine regions in a city as an example to illustrate the construction strategy of multiple graphs.

**Distance Graph.** Inspired by the first law of geography [43], which states that *everything is related to everything else, but nearby things are more related than distant things*, private car users may go to the nearest region to complete activities. We define the distance graph based on the geographical distance between the region center points. The weight between the two regions is recorded by the distance's reciprocal, and the closer region pair has a higher weight. We denote the adjacency matrix of the distance graph as  $\mathbf{A}_D$ , and each element  $\mathbf{A}_{D,ij}$  in the matrix is defined as:

$$\mathbf{A}_{D,ij} = dist(v_i, v_j)^{-1} \in [0, 1], \quad (4)$$

$$dist(v_i, v_j) = 2 \times e \times \arcsin \sqrt{\varepsilon}, \quad (5)$$

$$\varepsilon = \sin^2 \left( \frac{\text{lat}_i - \text{lat}_j}{2} \right) + \cos(\text{lat}_i) \cos(\text{lat}_j) \sin^2 \left( \frac{\text{lon}_i - \text{lon}_j}{2} \right), \quad (6)$$

where  $dist(v_i, v_j)$  denotes the Haversine distance of two regions,  $v_i = (\text{lon}_i, \text{lat}_i)$  and  $v_j = (\text{lon}_j, \text{lat}_j)$  represent the region center points  $u_i$  and  $u_j$ , respectively, and  $e = 6371.004$  is the approximate Earth's radius.

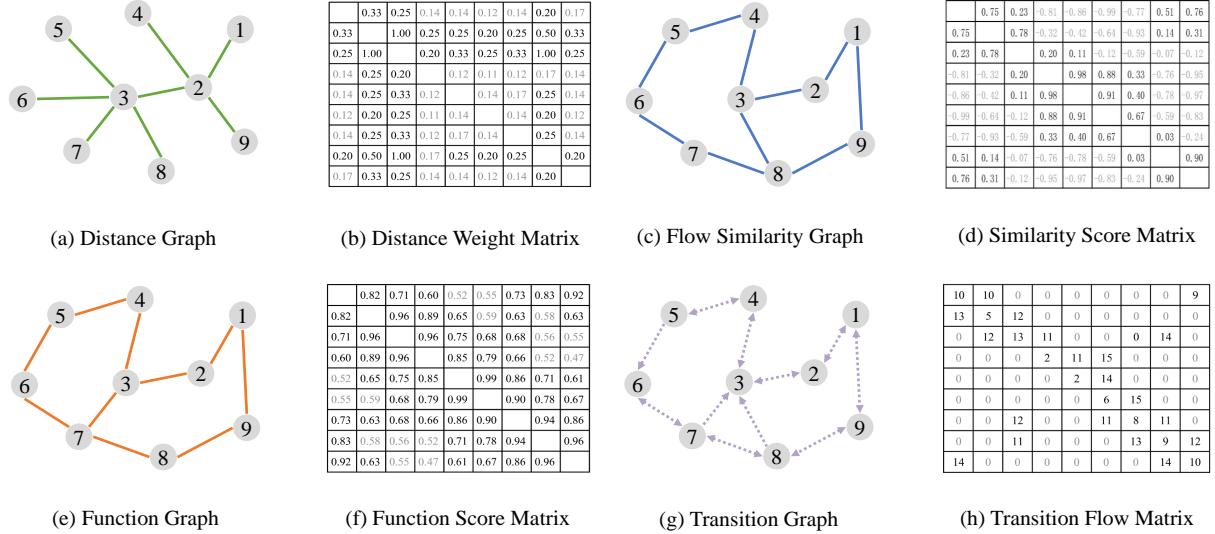


Fig. 3: An Example of Multiple Graphs and Weight Matrix.

As shown in Figure 3(a), the distance graph is a weighted undirected graph, and the distance between regions is calculated in kilometers. Figure 3(b) shows the distance weight between the regions; the weight range is between 0 and 1. The larger the weight, the shorter the distance between the regions, and the higher the correlation. Notably, the weight in gray indicates that the correlation between the regions is less than a given threshold.

**Flow Similarity Graph.** Historical flow is used to capture the similarity correlation between a region pair. If the historical private car flows of the two regions are similar, they also share similar crowd mobility patterns. We denote the adjacency matrix of the flow similarity graph as  $\mathbf{A}_S$  and use the Pearson correlation coefficient between time-series flows of a region pair to quantify its correlations. Each element  $\mathbf{A}_{S,ij}$  in the matrix can be calculated by:

$$\mathbf{A}_{S,ij} = \frac{\sum_{m=1}^n (v_{i,m} - \bar{v}_i)(v_{j,m} - \bar{v}_j)}{\sqrt{\sum_{m=1}^n (v_{i,m} - \bar{v}_i)^2} \sqrt{\sum_{m=1}^n (v_{j,m} - \bar{v}_j)^2}} \in [-1, 1], \quad (7)$$

where  $\mathbf{v}_i$  and  $\mathbf{v}_j$  are time series that represent private car flows of two regions  $v_i$  and  $v_j$ .  $v_{i,m}$  and  $v_{j,m}$  represent the  $m$ -th element in  $\mathbf{v}_i$  and  $\mathbf{v}_j$ .  $\bar{v}_i$  and  $\bar{v}_j$  are the means of  $\mathbf{v}_i$  and  $\mathbf{v}_j$ .

Figure 3(c) shows that the flow similarity graph is an undirected weighted graph. The similarity score between the regions is shown in Figure 3(d), and its range is between -1 and 1. The closer the score is to 1, the stronger the correlation between the two regions.

**Function Graph.** Urban regions with the same functional property have strong connections [44], and their attractiveness to private cars is also similar. The functional similarity of the two regions depends on the distribution of the regions between different ROI categories, as illustrated in Table II. We denote the adjacency matrix of the function graph as  $\mathbf{A}_F$ , and each

element  $\mathbf{A}_{F,ij}$  in the matrix is defined as:

$$\mathbf{A}_{F,ij} = \cos(\mathbf{p}_{v_i}, \mathbf{p}_{v_j}) \in [0, 1], \quad (8)$$

where  $\mathbf{p}_{v_i}$  and  $\mathbf{p}_{v_j}$  are the AOI category vectors of regions  $v_i$  and  $v_j$ , the dimension equals the number of AOI categories, and each entry represents the number of a specific AOI category in the region.

As shown in Figure 3(e), the function graph is an undirected weighted graph. The function score between the regions is shown in Figure 3(f), and its range is between 0 and 1. The larger the score, the higher the correlation between the regions.

**Transition Graph.** The transition graph differs from the other types of graphs; it is a dynamic graph of time evolution, as depicted in Figure 3. We use region-to-region transition flows to quantify such correlations, which are the count of private cars traveling from one region to another. We denote the adjacency matrix of the distance graph as  $\mathbf{A}_T$ , and each element  $\mathbf{A}_{T,ij}$  at time slot  $t$  is the sum of transition flows with time interval  $[t, t + \tau]$ :

$$\mathbf{A}_{T,ij} = \text{sum}(\mathbf{T}_{ij,t}) \in [0, \infty), \quad (9)$$

As shown in Figure 3(g), the transition graph is a directed weighted graph, and it retains changes over time. The flow transferred between regions is shown in Figure 3(h), and its value is greater than or equal to 0. The larger the score, the more dynamic the interactions between the regions and the higher their correlations.

### C. Spatial Correlation Modeling

Acquiring complex spatial correlations is a crucial step. We leverage GCN to transform and propagate information for graph  $\mathcal{G}$ :

$$\mathbf{X}^{l+1} = \sigma \left( \tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2} \mathbf{X}^l \mathbf{W}^l \right), \quad (10)$$

**Algorithm 2:** Urban Functional Region Identification

---

**Input:**  $\mathcal{R} = \{r_1, r_2, \dots, r_{|\mathcal{R}|}\}, m, ItrMax$ ;  
**Output:**  $\mathcal{U} = \{u_1, u_2, \dots, u_n\}$ .

- 1 Select  $m$  ROIs as the initial center point from  $\mathcal{R}$ :  
 $s = \{s_1, s_2, \dots, s_m\}$ .  $Itr = 0$ ;
- 2 **while**  $Itr < ItrMax$  **do**
- 3      $Itr = Itr + 1$ ;
- 4     **for**  $i \in [1, m]$  **do**
- 5         **for**  $j \in [1, n]$  **do**
- 6              $d_{i,j} \leftarrow \text{HarvenSin}(r_j, s_i)$ ;
- 7              $\lambda_j \leftarrow \arg\min_{j \in \{1, 2, \dots, n\}} d_{ji}$ ;
- 8              $U_{\lambda_j} \leftarrow U_{\lambda_j} \cup r_j$ ;
- 9         **endfor**
- 10     **endfor**
- 11     **for**  $i \in [1, m]$  **do**
- 12          $s_i' \leftarrow \frac{1}{|U_i|} \sum_{r \in U_i} r$ ;
- 13         **if**  $s_i' \neq s_i$  **then**
- 14              $s_i \leftarrow s_i'$ ;
- 15         **end**
- 16     **endfor**
- 17     Extract the categories that belong to the same function from  $\mathcal{U}$  and record them as sets  
 $\mathbf{W}_{f_{ij}} = \{\mathbf{w}_{f_{i1}}, \mathbf{w}_{f_{i2}}, \dots, \mathbf{w}_{f_{ik}}\}$ ;
- 18     **for** function  $f_i$  belonging to each region  $u_i$  **do**
- 19         **for**  $i \in [1, n]$  **do**
- 20             **for**  $j \in [1, k]$  **do**
- 21                  $\mathbf{f}_{ij} = \sum_{j=1}^k c \times p(\mathbf{W}_{f_{ij}}) \times q(\mathbf{W}_{f_{ik}}) \times \mathbf{w}(i)$ ;
- 22             **endfor**
- 23              $\mathbf{f}_i = \sum_{j=1}^n \mathbf{f}_{ij}$ ;
- 24         **endfor**
- 25     **endfor**
- 26 **end**

---

where  $\mathbf{X}^l$  and  $\mathbf{X}^{l+1}$  are the correlation vectors of  $N$  regions in layers  $l$  and  $l+1$ , respectively.  $\sigma$  is the activation function  $\text{ReLU}$ .  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$  is the adjacency matrix with self-connection.  $\tilde{\mathbf{D}}$  is the diagonal degree matrix of  $\tilde{\mathbf{A}}$ .  $\mathbf{W}^l$  is a learnable weighted matrix shared over all regions. To simplify notations, we summarize the operation in Equation (10) as follows:

$$\mathbf{X}^{l+1} = f_g(\mathbf{A}, \mathbf{X}^l), \quad (11)$$

Unlike the traditional GCN, to fully exploit the global spatial correlations of graphs, we propose graph convolutional networks with densely connected blocks (DenseGCN) to capture the spatial correlation. Inspired by DenseNets, we introduce densely connected blocks in the GCN since convolutional networks can make training more accurate and effective if they contain shorter connections between layers close to the input and those close to the output [45]. As depicted in Figure 4, a dense graph convolutional network is primarily composed of two subblocks with different layers: the first is  $n$  layers, and the second is  $m$  layers. With densely connected blocks, the input of the vertex  $v$  in the  $l$  layer comes

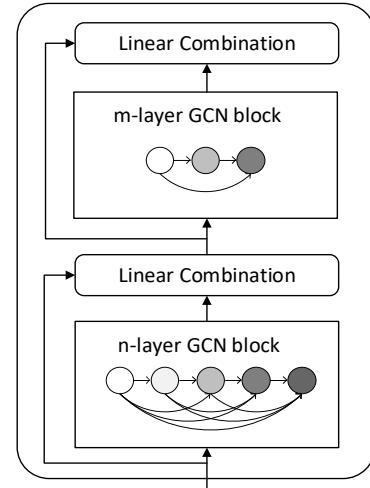


Fig. 4: Graph Convolutional Networks with Densely Connected Blocks.

not only from the correlation representation of the previous layer  $\mathbf{X}^l$  but also from the output from all previous layers. The  $l$ th layer  $\mathbf{X}^{l+1}$  receives the correlation maps of all preceding layers as input:

$$\mathbf{C}^{l+1} = H_c[\mathbf{X}^0; \mathbf{X}^1; \dots; \mathbf{X}^l], \quad (12)$$

where  $H_c(\cdot)$  is a linear combination layer between multilayer convolutional networks [46].

We expand the densely connected blocks to multigraph convolutional networks to model the global spatial correlations:

$$\mathbf{X}^{l+1} = f_g(\mathcal{A}_m, \mathbf{C}^{l+1}), \quad (13)$$

$$\mathcal{A}_m = f_a(\mathbf{A}; \theta_i), \quad (14)$$

where  $\mathbf{A}$  denotes a type of graph in multiple graphs,  $\mathbf{A} \in \{\mathbf{A}_D, \mathbf{A}_F, \mathbf{A}_S, \mathbf{A}_T\}$ .  $\mathcal{A}_m \in \mathbb{R}^{N \times N \times 4}$  represents the aggregation matrix of different samples based on the graph  $\mathbf{A}$  parameterized by  $\theta_i$ .  $f_g$  is the spatial feature extraction function.

#### D. Temporal Correlation Modeling

We combine multigraph dense convolutions with gated recurrent units to capture the spatiotemporal correlations. We refer to this structure as the multigraph dense GRU. For convenience,  $\tilde{\mathbf{X}}_t[i, :]$  represents multigraph dense convolutions, and  $\tilde{\mathbf{H}}_{t-1}[i, :]$  hidden is the hidden representations from the previous step:

For each node  $v_i$  at time step  $t$ , the process of multigraph dense GRU extracting temporal correlations can be expressed

as follows:

$$\begin{aligned}\mathbf{u}_t &= \sigma_u \left( \mathbf{W}_u \tilde{\mathbf{X}}_t[i, :] + \mathbf{U}_u \tilde{\mathbf{H}}_{t-1}[i, :] + \mathbf{b}_u \right), \\ \mathbf{r}_t &= \sigma_r \left( \mathbf{W}_r \tilde{\mathbf{X}}_t[i, :] + \mathbf{U}_r \tilde{\mathbf{H}}_{t-1}[i, :] + \mathbf{b}_r \right), \\ \mathbf{c}_t &= \tanh \left( \mathbf{W}_c \tilde{\mathbf{X}}_t[i, :] + \mathbf{U}_c \left( \mathbf{r}_t \odot \mathbf{U}_c \tilde{\mathbf{H}}_{t-1}[i, :] \right) + \mathbf{b}_c \right), \\ \mathbf{H}_t[i, :] &= (1 - \mathbf{u}_t) \tilde{\mathbf{H}}_{t-1}[i, :] + \mathbf{u}_t \odot \mathbf{c}_t,\end{aligned}\quad (15)$$

where  $\tilde{\mathbf{X}}_t[i, :]$ ,  $\mathbf{H}_t[i, :]$  are the input and output at time step  $t$ .  $\odot$  denotes the elementwise multiplication.

From the perspective of private cars, the longer the private car stays in a region, the more attractive the region is. We take the average stay duration of private cars in each region as the weight of regional importance. We propose attention networks based on the private cars' stay duration. Combined with the stay duration correlations, and assign different weights to different vertices nearby to learn the importance.

For node  $v_i$ , the output sequence  $(\mathbf{H}_1[i, :], \dots, \mathbf{H}_T[i, :])$  from multigraph dense GRU is the input for the attention. We first organize this input according to the sequential order by stacking them row-wise as a matrix  $\mathbf{H}_i \in \mathbb{R}^{T \times d}$ , where  $i$  indicates the vertex  $v_i$ . Then, a shared linear transformation is parameterized by a weight vector  $\mathbf{W}_{i,st}$ . Specifically, each  $\mathbf{W}_{i,st}$  represents the average stay duration of a region in the previous  $\varphi$  intervals. The attention is formulated as:

$$\begin{aligned}\mathbf{e}_i &= (\mathbf{W}_{i,st} \mathbf{H}_i)^T \mathbf{u}, \\ \mathbf{a}_i &= \text{SoftMax}(\mathbf{e}_i) = \frac{\exp(\mathbf{e}_i)}{\sum_{l=1}^k \exp(\mathbf{e}_i)},\end{aligned}\quad (16)$$

where  $\mathbf{u}$  is a weight vector, which is initialized as a random vector and jointly learned during the training process.

After obtaining the attention coefficient, we calculate the linear combination represented for each vertex as follows:

$$\mathbf{H}_{out}^i = \sum_{l=1}^k \mathbf{a}_i \mathbf{H}_i, \quad (17)$$

Finally, we generate the output of the attention layer, which can be denoted as  $\mathbf{H}_{out}^i \in \mathbb{R}^{T \times d}$ .

### E. Prediction Network

After the attention layer, we take  $\{\mathbf{H}_{out}^i | i \in \mathcal{V}\}$  as input and apply a multilayer feed-forward network to predict private car flows of future periods. In the training process, our goal is to minimize the error between the real private car flows  $\mathbf{X}_{t+h}$  and predicted private car flows  $\hat{\mathbf{X}}_{t+h}$ . The loss function of our framework is shown in Equation (18).

$$\text{loss} = \|\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h}\| + \lambda L_{reg}. \quad (18)$$

where the first term is used to minimize the error. The second term is the L2 regularization term, which helps avoid overfitting problems.

## V. EXPERIMENTS

In this section, we evaluate the performance of private car flow prediction based on real-world datasets.

TABLE IV: Data Description.

Data Source	Yuhua-Changsha	Baoan-Shenzhen
Latitude	27.92-28.21	22.51-22.83
Longitude	112.97-113.18	113.75-113.99
Private car trajectories	1258620	4249660
ROIs	1407	3995
Training Timespan	2018/8/1-2018/9/21	
Validation Timespan	2018/9/22-2018/9/24	
Testing Timespan	2018/9/25-2018/9/30	

### A. Datasets

Table IV details the private car trajectory and ROI data for two months in Baoan District, Shenzhen, and Yuhua District, Changsha, China. To protect the privacy of private car owners, all sensitive information was removed from the raw trajectories. All researchers are subject to a strict nondisclosure license.

### B. Baselines

We compare the proposed MGDCN with baselines. The description and main parameter settings of the baselines are as follows.

- **HA.** The historical average simply predicts the historical average in a particular region. We choose the first 70% as the training set of the HA model, 10% as the validation set, and the remaining 20% as the test set.
- **ARIMA<sub>kal</sub>.** The autoregressive integrated moving average model with a Kalman filter is a classical time-series prediction model [47]. We input 1,464 time intervals, and the length of each interval is 1 hour.
- **BO-SVR.** The support vector regression approach based on Bayesian optimization performs well for time-series and nonlinear prediction [48]. We set three tuning parameters  $(C; \varepsilon; \sigma)$  as  $(1; 0.1; 1)$ .
- **GBRT.** The gradient boost regression tree is a nonparametric method of solving regression problems that produce a prediction model [49]. We set the estimator range from 10 to 100.
- **Multi-LSTM.** Multidimensional LSTM can be directly used to deal with private car flow sequences. We set the number of dimensions to 5 and the number of hidden units to 32.
- **DCRNN.** Diffusion convolutional recurrent neural network models spatial and temporal correlations by integrating graph convolution and GRU [12]. The number of diffusion steps is set as 2, and the size of graph embedding is set as 32. We set the number of hidden units to 64 in the Shenzhen dataset and 32 in the Changsha dataset.
- **T-GCN.** The temporal GCN combines the GCN and the GRU to predict traffic flow [10]. We set the number of hidden units to 100 in the Shenzhen dataset and 64 in the Changsha dataset.
- **GraphWaveNet.** Graph WaveNet is a model that integrates GCN and WaveNet [50]. We set the number of graph WaveNet layers to 4, the size of graph embedding to 32, and the dimension of TCN layers to 32.

- **GMAN.** The graph multiattention network adapts an encoder-decoder architecture and relies on a self-attention mechanism to model spatial-temporal correlations for traffic speed and volume prediction [51]. In our case, the ST-Attention is set to 3. The attention head is set to 8, and its dimension is set to 8.

### C. Settings and Metrics

We choose the first 70% as the training set, 10% as the validation set, and the remaining 20% as the test set. Our framework is trained using the Adam optimizer. We set the GRU hidden units to 64. The time threshold  $\tau = 1$  h denotes dividing the day into 24 equal-length time slices. The initial learning rate  $lr = 0.0015$  decreased to 9/10 of itself every 1,000 iterations, the drop rate  $d = 0.5$ , and the batch size is 64. This code is carefully designed in the early stopping strategy; if the performance does not improve, the training process is terminated.

In this study, private car flows in urban regions are extracted from real-world private car trajectory and ROI datasets. We identify this as the ground truth in the experiments. To evaluate the performance of each method, we employ five standard metrics: mean absolute error (MAE), root mean squared error (RMSE), weighted mean absolute percentage error (wMAPE), coefficient of determination ( $R^2$ ), and variance regression score (Var).

$$\text{MAE} = \frac{\sum_{i=1}^N |\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h}|}{N}, \quad (19)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h})^2}{N}}, \quad (20)$$

$$\text{wMAPE} = \frac{\sum_{i=1}^N |\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h}|}{\sum_{i=1}^N \mathbf{X}_{t+h}}, \quad (21)$$

$$R^2 = 1 - \frac{\sum_{j=1}^M \sum_{i=1}^N (\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h})^2}{\sum_{j=1}^M \sum_{i=1}^N (\mathbf{X}_{t+h} - \bar{\mathbf{X}}_{t+h})^2}, \quad (22)$$

$$\text{Var} = 1 - \frac{\text{var}\{\hat{\mathbf{X}}_{t+h} - \mathbf{X}_{t+h}\}}{\text{var}\{\mathbf{X}_{t+h}\}}, \quad (23)$$

where  $M$  denotes the number of time samples and  $\bar{\mathbf{X}}_{t+h}$  represents the average of  $\mathbf{X}_{t+h}$ .

### D. Results and Discussion

1) *Urban Functional Regions Identification:* We first select the silhouette coefficient (SC) as the evaluation metrics to calculate the  $m$ . The calculation of SC is shown in Equation (24):

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \quad (24)$$

where  $a(i)$  is the average of the distance from  $r_i$  to all other points in the cluster it belongs to, and  $b(i)$  is  $r_i$  to the nearest adjacent cluster, the minimum value of the average distance

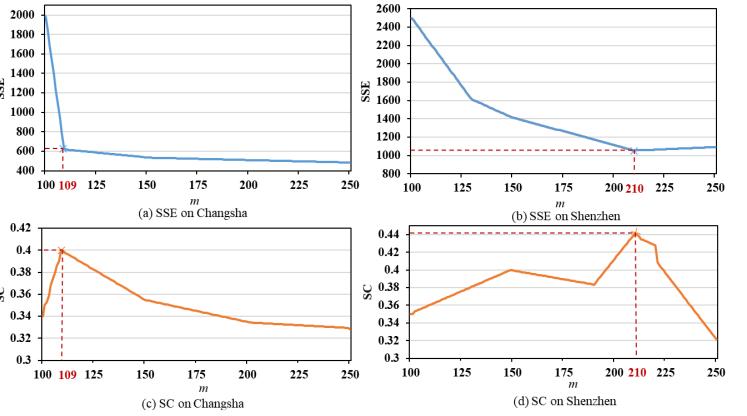
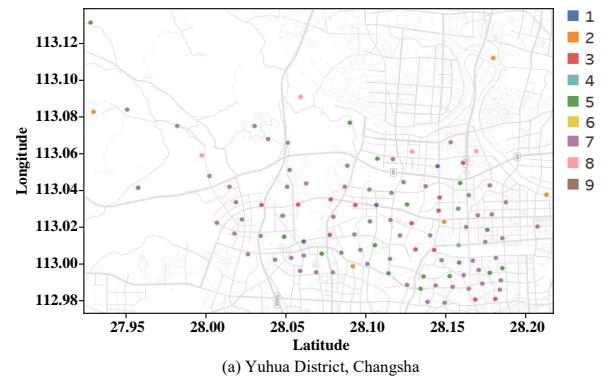
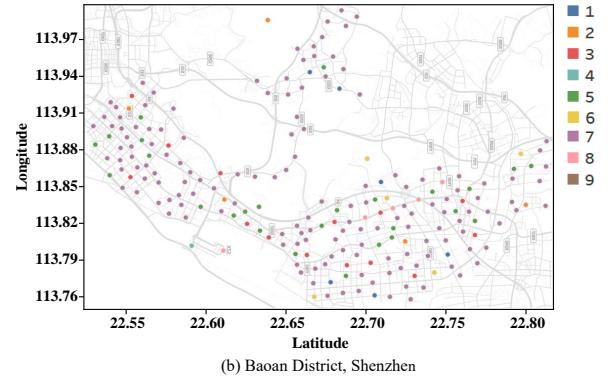


Fig. 5: Comparison of UFR Identification under Different  $m$ .



(a) Yuhua District, Changsha



(b) Baoan District, Shenzhen

Fig. 6: Distribution of Urban Regional Central Points in Changsha and Shenzhen.

of all points. We average the SC of all regions, when the SC is the largest, the corresponding  $m$  is the optimal number of UFRs.

The division of UFRs will be more refined with the increasing of  $m$ . In this case, the degree of aggregation of each region will gradually increase, and the SSE will gradually become smaller. When  $m$  is less than the ground truth, since the increase of  $m$  will greatly increase the degree of aggregation of each region, the SSE declines quickly. When  $m$  reaches the ground truth, the return of aggregation degree obtained by increasing  $m$  will rapidly decrease so that the SSE will decrease sharply, and then it will level off as  $m$  continues to increase. In short, the relationship between SSE and  $m$  is the shape of an elbow, and the value corresponding to the

inflection point of the elbow is the optimal candidate value of  $m$ . The calculation of SSE is shown in Equation (25):

$$SSE = \sum_{i=1}^n \sum_{l \in u_i} |l - m_i|^2. \quad (25)$$

where  $l$  is the location point in  $u_i$ .  $m_i$  is the centroid of  $u_i$ .

We first obtain the number of UFR  $m$ . As shown in Figure 5, on the Changsha data set, when the value of  $m$  is set to 109, the value of SSE is the inflection point, that is, the optimal value is closest to the ground truth. Similarly, the value of SC at this time reaches the maximum. On the Shenzhen data set, the optimal number of UFR can be obtained as 210 through the two metrics of SSE and SC. Then, we visualize the center points of the generated urban regions. Figure 6(a) presents the distribution of 109 center points in the Yuhua District of Changsha. Figure 6(b) depicts the distribution of 210 center points in the Baoan District of Shenzhen. The color of the center point corresponds to Table II, and different colors represent different URFs. Figure 6 verifies that the center point distribution in urban regions is relatively uniform.

2) *Parameter Sensitivity*: We report the impact of the input time step  $\varphi$  and the prediction time steps  $h$  using MAE. Specifically, we examine the performance by observing the variation in one parameter and setting the other parameters' default values. We test the input time step's impact and choose the number of  $\varphi$  from [6, 8, 10, 12, 14, 16, 18]. The results are reported in Figure 7(a). MGDCN achieves least errors when  $\varphi = 14$ . One possible reason is that too many short-term inputs cannot provide sufficient temporal correlations, and too long inputs will introduce more noise for temporal correlation modeling. To test the impact of the prediction step, we vary  $h$  from 1 to 6. The results are depicted in Figure 7(b). By increasing  $h$ , the MAE increases consistently; this makes sense because the temporal correlations between observed flows and future flows become lower when  $h$  increases.

The number of hidden units is an essential hyperparameter of the MGDCN model, as the different number of hidden units may substantially affect prediction precision. We experiment with different hidden units and select the optimal value by comparing the predictions to choose the best value. In the experiment, for the Changsha dataset, we choose the number of hidden units from [8, 16, 32, 64, 128] and analyze the change in prediction precision.

As shown in Figure 8, the horizontal axis represents the number of hidden units, and the vertical axis represents the

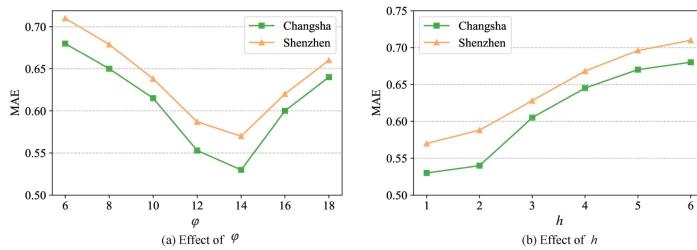


Fig. 7: Parameter Sensitivity on the Changsha and Shenzhen Datasets.

change in different metrics. Figure 8(a) depicts the RMSE, MAE, and wMAPE results for different hidden units. The error is the smallest when the number is 32. Figure 8(b) shows the variation in  $R^2$  and variance regression score for different hidden units. Similarly, the results reach a maximum when the number is 32.

In summary, the prediction results are better when the number is set to 32. When the number of hidden units increases, the prediction precision increases and then decreases because when the hidden unit is larger than a certain degree, the model complexity and computational difficulty greatly increase. As a result, overfitting of the training data occurs. Therefore, we set the number of hidden units to 32 in our experiments on the Changsha dataset. The results of the Shenzhen dataset are shown in Figure 8(c) and 8(d). When the number of hidden units is 64, the prediction precision is the highest, and the prediction error is the lowest.

3) *Overall Performance*: For the above urban regions, we predict private car flows in different periods. Tables V and VI present the metrics for all the methods in the two cities. “—” denotes that the values are small enough to be negligible. Compared with the other baselines, our proposed MGDCN achieves superior performance, followed by GMAN, GraphWaveNet, DCRNN, and T-GCN. Traditional machine learning methods fail to perform well because they fail to capture the spatiotemporal correlations synchronously. MGDCN outperforms the traditional regression-based methods and achieves better performance than deep-learning models such as Multi-LSTM. The potential reason is that Multi-LSTM typically applies temporal correlations but ignores the spatial correlations. T-GCN simply combines the GCN and GRU to predict but ignores the multiple correlations between regions. T-GCN cannot obtain better results than DCRNN. DCRNN also overlooks these multiple correlations and lacks personalized considerations for each region. The diffusion operation module of DCRNN still relies on the road network structure to propagate information, failing to represent the actual connections between regions truly. It is challenging to model a private car's ASL behavior using GraphWaveNet with adaptive dynamic graphs, especially the stay duration correlations. MGDCN outperforms GMAN since it is difficult for GMAN to capture the semantics, dynamic transfer, and stay duration correlations of private car ASL behaviors.

TABLE V: Prediction Errors and Goodness of Fit on Changsha.

City	Changsha					
	Metrics	MAE	RMSE	wMAPE	$R^2$ *	Var *
HA	3.39	4.91	44.56%	-	-	-
ARIMA <sub>kal</sub>	4.1	4.96	38.71%	-	-	-
GBRT	2.19	4.03	31.81%	0.54	0.61	
SO-SVR	1.5	3.32	21.9%	0.69	0.71	
Multi-LSTM	2.22	5.24	31.3%	0.21	0.28	
DCRNN	0.69	0.93	7.11%	0.81	0.80	
T-GCN	0.97	2.73	14.36%	0.79	0.80	
GraphWaveNet	0.68	0.91	7.96%	0.82	0.80	
GMAN	0.62	0.87	6.98%	0.87	0.86	
MGDCN	0.53	0.74	6.13%	0.92	0.93	

\* Higher is better.

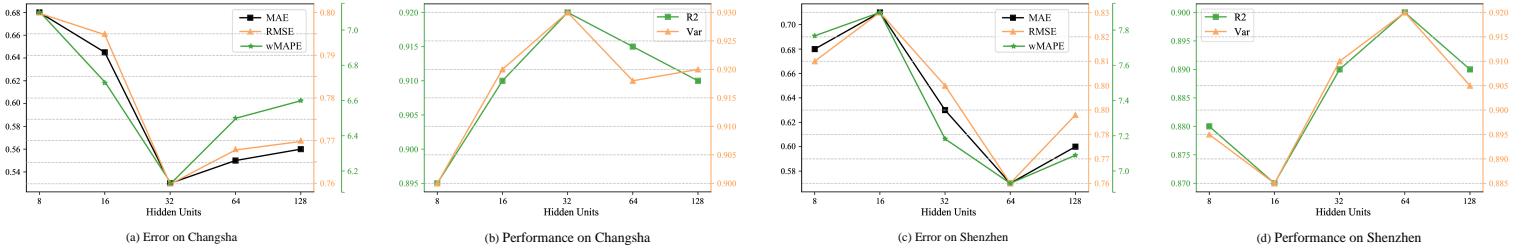


Fig. 8: Comparison of Predicted Performance under Different Hidden Units.

TABLE VI: Prediction Errors and Goodness of Fit on Shenzhen.

City	Shenzhen					
	Metrics	MAE	RMSE	wMAPE	R <sup>2</sup> *	Var *
HA	4.31	5.40	42.01%	-	-	-
ARIMA <sub>kal</sub>	4.51	5.15	36.37%	-	-	-
GBRT	2.21	5.22	31.3%	0.21	0.28	
SO-SVR	1.76	4.47	24.93%	0.42	0.46	
Multi-LSTM	2.96	5.01	43.2%	0.29	0.3	
DCRNN	0.69	0.97	7.23%	0.80	0.79	
T-GCN	1.09	4.29	15.38%	0.77	0.77	
GraphWaveNet	0.68	0.96	8.04%	0.81	0.80	
GMAN	0.64	0.93	7.62%	0.83	0.82	
MGDCN	0.57	0.76	6.77%	0.90	0.92	

\* Higher is better.

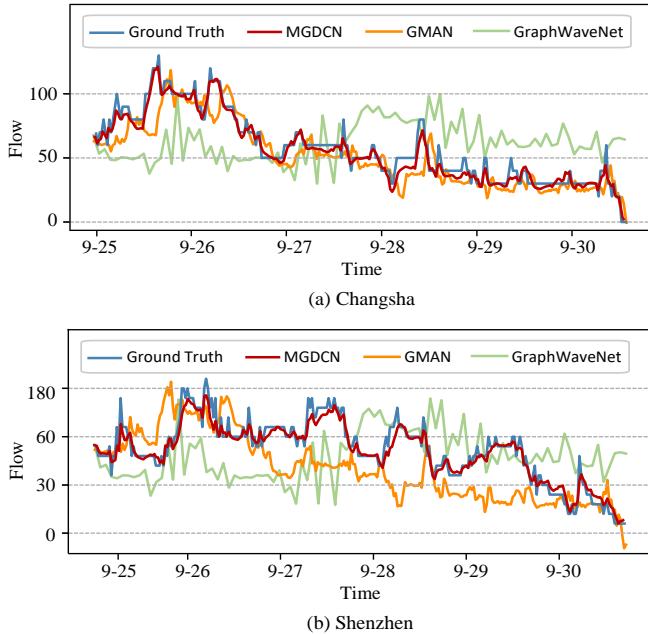


Fig. 9: Prediction Results in September 2018.

4) *Visualization of Prediction Results:* To further explain the superior effect of MGDCN, we visualized the prediction results of MGDCN, GMAN, and GraphWaveNet on a single UFR. Figure 9(a) visualizes the prediction results of the region of profession function in Changsha, which consists of 11 ROIs. Figure 9(b) depicts the prediction result of the region of enterprise function in Shenzhen, which consists of 20 ROIs. In the last two weekend days, we observe more private car

flows on weekdays in these two UFRs than on weekends. The changes in the ground truth curve are complicated, especially in the Shenzhen dataset. MGDCN can quickly learn and make reasonable predictions in these complex situations, which indicates that MGDCN can learn the complex dynamic changes in private car flows.

5) *Ablation Study:* To investigate the effect of spatial and temporal correlation modeling, we evaluated the following MGDCN variants by removing different components from the model: (1) flow similarity graph, (2) distance graph, (3) transition graph, (4) function graph, and (5) attention. The result is shown in Figure 10, "None" indicates that no component is removed. Removing any component causes a significant error increase, which justifies the importance of each component. The results demonstrate that removing the attention component has the most significant impact on the MGDCN because the attention component captures the stay duration feature in each region. Following the transition graph, the reason is that the transition graph can model dynamic multidirectional interactions between the regions. The flow similarity graph models the relationship between regions from the perspective of historical flow similarity. Compared with dynamic transition modeling, this method has less contribution to MGDCN. Due to the ability to fully use urban regions' semantic information, the functional graph also plays a significant role. The least impact is the distance graph.

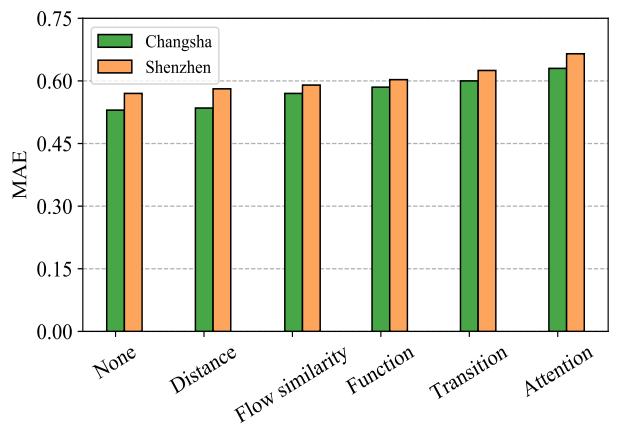


Fig. 10: Ablation Study.

### E. Discussion

The above experimental results confirm that our proposed MGDCN has good performance in predicting private car flows. Our results are helpful for groups such as urban traffic management departments and private car owners. In the following, we discuss three aspects in detail:

**Smart parking.** Smart parking guarantees to construct a smart city and an essential part of the smart city's realization [52]. The purpose of smart parking has three main aspects: maximizing the utilization of parking resources, maximizing parking lot profits, and optimizing parking services for car owners. Our prediction results are the total number of private cars staying in UFRs during a period, including the arrival flows, leave flows, and fixed stay flows. For parking space managers, accurate prediction results provide advanced decisions for parking lot spaces in the city, enabling them to manage parking spaces reasonably, maximizing parking space resource utilization and parking lot profits. For private car owners, accurate prediction results can recommend the best vacant parking lot nearby for them.

**Risk assessment.** For urban traffic managers, if a large number of private car flows in some regions are different than usual or exceed the original vehicle quota, more severe congestion or abnormal events will usually occur. If we can accurately predict this phenomenon, managers can take timely risk prevention and control measures. If this risk is predicted in advance, risk prevention and control measures can be implemented in time [53].

**Public transportation planning.** Through predicting private car flows, the routes with large car flows and a lack of public transport facilities can be determined and optimized. The public transport capacity can be reasonably arranged according to the car flow's elastic demand in different periods. Therefore, analyzing private car flows can effectively alleviate traffic congestion, facilitate public travel, provide data support for public transportation departments' decision-making, create economic value, and improve public transportation investment [54].

## VI. CONCLUSION

In this study, we attempted to predict private car flows in urban functional regions by exploiting spatiotemporal correlations of arrive-stay-leave (ASL) behaviors of private car users. It is noteworthy that we focus on private car flows in irregular regions with semantic information. In this context, we propose a novel end-to-end framework MGDCN. We evaluated the MGDCN based on two months of private car trajectory and ROI data in Changsha and Shenzhen, China. The robust performance of MGDCN proves the practical applicability of our framework, which is meaningful for risk assessment, smart parking, and urban planning. Additionally, the proposed MGDCN can be applied to other prediction tasks, such as travel demand, fuel consumption, and speed distribution prediction.

In the future, we will integrate multisource heterogeneous data, such as weather, holidays, important events, and city populations, to more comprehensively predict private car flows

or private car distributions. Moreover, it is further meaningful to promptly deploy our framework under a cloud/client architecture to collect data and visualize prediction results online promptly.

## VII. ACKNOWLEDGMENTS

This work was supported in part by the Humanities and Social Sciences Foundation of MOE under grant 21YJCZH183, the National Natural Science Foundation of China under grants 61772184 and U20A20181, in part by the Key Research and Development Project of Hunan Province of China under Grant 2021GK2020, the funding projects of Zhejiang Lab under grants 2021LC0AB05 and 2020LC0PI01, and the funding project of Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology under grant KF-J21012.

## REFERENCES

- [1] L. Zheng, D. Xia, L. Chen, and D. Sun, "Understanding citywide resident mobility using big data of electronic registration identification of vehicles," *IEEE Trans. Intell. Transp. Syst.*, pp. 4363–4377, Sep. 2019.
- [2] L. Mantzos, P. Capros, N. Kouvaritakis, and M. Zeka-Paschou, "European energy and transport-trends to 2030," 2009.
- [3] D. Kondor, H. Zhang, R. Tachet, P. Santi, and C. Ratti, "Estimating savings in parking demand using shared vehicles for home-work commuting," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 8, pp. 2903–2912, Oct. 2018.
- [4] Y. Huang, Z. Xiao, D. Wang, H. Jiang, and D. Wu, "Exploring individual travel patterns across private car trajectory data," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 12, pp. 1–15, Oct. 2019.
- [5] D. Wang, J. Fan, Z. Xiao, H. Jiang, H. Chen, F. Zeng, and K. Li, "Stop-and-wait: Discover aggregation effect based on private car trajectory data," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3623–3633, Nov. 2019.
- [6] W. Zhang, H. Liu, Y. Liu, J. Zhou, and H. Xiong, "Semi-supervised hierarchical recurrent graph neural network for city-wide parking availability prediction," in *AAAI*, vol. 34, no. 01, Feb. 2020, pp. 1186–1193.
- [7] Z. Yuan, X. Zhou, and T. Yang, "Hetero-convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data," in *ACM SIGKDD*, Aug. 2018, pp. 984–992.
- [8] F. Xia, J. Wang, X. Kong, Z. Wang, J. Li, and C. Liu, "Exploring human mobility patterns in urban scenarios: A trajectory data perspective," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 142–149, Mar. 2018.
- [9] H. Jiang, Y. Zhang, Z. Xiao, P. Zhao, and A. Iyengar, "An empirical study of travel behavior using private car trajectory data," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 1, pp. 53–64, Sep. 2021.
- [10] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, "T-gcn: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Aug. 2019.
- [11] H. Peng, H. Wang, B. Du, M. Z. A. Bhuiyan, H. Ma, J. Liu, L. Wang, Z. Yang, L. Du, S. Wang *et al.*, "Spatial temporal incidence dynamic graph neural networks for traffic flow forecasting," *Inf. Sci.*, vol. 521, pp. 277–290, Jun. 2020.
- [12] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *ICLR*. OpenReview.net, May. 2018, pp. 1–16.
- [13] H. Zhu, Y. Luo, Q. Liu, H. Fan, T. Song, C. W. Yu, and B. Du, "Multistep flow prediction on car-sharing systems: A multi-graph convolutional neural network with attention mechanism," *Int. J. Softw. Eng. Knowl. Eng.*, vol. 29, pp. 1727–1740, Nov. 2019.
- [14] B. Du, X. Hu, L. Sun, J. Liu, Y. Qiao, and W. Lv, "Traffic demand prediction based on dynamic transition convolutional neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 1237–1247, Jan. 2021.
- [15] B. Shen, X. Liang, Y. Ouyang, M. Liu, W. Zheng, and K. M. Carley, "Stepdeep: A novel spatial-temporal mobility event prediction framework based on deep neural network," in *ACM SIGKDD*, London, UK, Aug. 2018, pp. 724–733.

- [16] Y. Wang, H. Yin, H. Chen, T. Wo, J. Xu, and K. Zheng, "Origin-destination matrix prediction via graph convolution: a new perspective of passenger demand modeling," in *ACM SIGKDD*, Aug. 2019, pp. 1227–1235.
- [17] H. Yu, Z. Li, G. Zhang, P. Liu, and J. Wang, "Extracting and predicting taxi hotspots in spatiotemporal dimensions using conditional generative adversarial neural networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 3680–3692, Mar. 2020.
- [18] Y. Liu, Z. Liu, and R. Jia, "Deepf: A deep learning based architecture for metro passenger flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 18–34, Feb. 2019.
- [19] G. Qi, A. Huang, W. Guan, and L. Fan, "Analysis and prediction of regional mobility patterns of bus travellers using smart card data and points of interest data," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1197–1214, Apr. 2018.
- [20] P. Wang, L. T. Yang, Y. Peng, J. Li, and X. Xie, "M2t2: The multivariate multi-step transition tensor for user mobility pattern prediction," *IEEE Trans. Netw. Sci. Eng.*, pp. 907–917, Feb. 2019.
- [21] J. Chen, Z. Xiao, D. Wang, W. Long, and V. Havyarimana, "Stay of interest: A dynamic spatiotemporal stay behavior perception method for private car users," in *IEEE HPCC*, Aug. 2019, pp. 1526–1532.
- [22] H. Chen, B. Guo, Z. Yu, C. Zheng, and A. Wang, "The framework of increasing drivers' income on the online taxi platforms," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 4, pp. 2182–2191, Apr. 2020.
- [23] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal arima model with limited input data," *European Transport Research Review*, vol. 7, no. 3, pp. 1–9, Jun. 2015.
- [24] L. Cai, Z. Zhang, J. Yang, Y. Yu, T. Zhou, and J. Qin, "A noise-immune kalman filter for short-term traffic flow forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 536, p. 122601, Dec. 2019.
- [25] Y. Li, Y. Zheng, H. Zhang, and L. Chen, "Traffic prediction in a bike-sharing system," in *ACM SIGSPATIAL*, Nov. 2015, pp. 1–10.
- [26] J. Xiao, Z. Xiao, D. Wang, J. Bai, V. Havyarimana, and F. Zeng, "Short-term traffic volume prediction by ensemble learning in concept drifting environments," *Knowl. Based Syst.*, vol. 164, pp. 213 – 225, Nov. 2019.
- [27] Y. Liu, C. Lyu, A. Khadka, W. Zhang, and Z. Liu, "Spatio-temporal ensemble method for car-hailing demand prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 12, pp. 1–6, Oct. 2019.
- [28] J. Xu, R. Rahmatizadeh, L. Böloni, and D. Turgut, "Real-time prediction of taxi demand using recurrent neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2572–2581, Oct. 2017.
- [29] L. Liu, J. Chen, H. Wu, J. Zhen, G. Li, and L. Lin, "Physical-virtual collaboration modeling for intra-and inter-station metro ridership prediction," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–15, 2020.
- [30] J. Ye, J. Zhao, K. Ye, and C. Xu, "How to build a graph-based deep learning architecture in traffic domain: A survey," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–21, Nov. 2020.
- [31] M. Zhang, T. Li, Y. Li, and P. Hui, "Multi-view joint graph representation learning for urban region embedding," in *IJCAI*, Yokohama, Japan, Jan. 2021, pp. 4431–4437.
- [32] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong, "Discovering urban functional zones using latent activity trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 3, pp. 712–725, Aug. 2015.
- [33] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and pois," in *ACM SIGKDD*, Beijing, China, Aug. 2012, pp. 186–194.
- [34] X. Zhang, S. Du, and Z. Zheng, "Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with vhr images and poi data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 161, pp. 1–12, Mar. 2020.
- [35] B. Yu, Y. Lee, and K. Sohn, "Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (gcn)," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 189–204, Feb. 2020.
- [36] X. Chen, Y. Wang, J. He, S. Pan, Y. Li, and P. Zhang, "Cap: Context-aware app usage prediction with heterogeneous graph embedding," *ACM IMWUT*, vol. 3, no. 1, pp. 1–25, Mar. 2019.
- [37] D. Chai, L. Wang, and Q. Yang, "Bike flow prediction with multi-graph convolutional networks," in *ACM SIGSPATIAL*, pp. 397–400.
- [38] X. Geng, Y. Li, L. Wang, L. Zhang, Q. Yang, J. Ye, and Y. Liu, "Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting," in *AAAI*, vol. 33, Jan. 2019, pp. 3656–3663.
- [39] H. Shi, Q. Yao, Q. Guo, Y. Li, L. Zhang, J. Ye, Y. Li, and Y. Liu, "Predicting origin-destination flow via multi-perspective graph convolutional network," in *IEEE ICDE*, Apr. 2020, pp. 1818–1821.
- [40] Z. Xiao, D. Xiao, V. Havyarimana, H. Jiang, D. Liu, D. Wang, and F. Zeng, "Toward accurate vehicle state estimation under non-gaussian noises," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10 652–10 664, Sep. 2019.
- [41] W. Zhou, D. Ming, X. Lv, K. Zhou, H. Bao, and Z. Hong, "So-cnn based urban functional zone fine division with vhr remote sensing image," *Remote Sensing of Environment*, vol. 236, p. 111458, Jan. 2020.
- [42] D. Arthur and S. Vassilvitskii, "k-means++: the advantages of careful seeding," in *ACM SIAM, New Orleans, Louisiana, USA*. Stanford, Jan. 2007, pp. 1027–1035.
- [43] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic Geography*, vol. 46, pp. 234–240, Jan. 1970.
- [44] Y. Gong, Z. Li, J. Zhang, W. Liu, and Y. Zheng, "Online spatio-temporal crowd flow distribution prediction for complex metro system," *IEEE Trans. Knowl. Data Eng.*, pp. 1–16, Apr. 2020.
- [45] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *CVPR*, Jul. 2017, pp. 2261–2269.
- [46] Z. Guo, Y. Zhang, Z. Teng, and W. Lu, "Densely connected graph convolutional networks for graph-to-sequence learning," *Trans. Assoc. Comput. Linguistics*, vol. 7, pp. 297–312, Jun. 2019.
- [47] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, no. 50, pp. 159–175, Jan. 2003.
- [48] D. Wang, C. Wang, J. Xiao, Z. Xiao, W. Chen, and V. Havyarimana, "Bayesian optimization of support vector machine for regression prediction of short-term traffic flow," *Intell. Data Anal.*, vol. 23, no. 2, pp. 481–497, Apr. 2019.
- [49] Y. Wang, X. Lin, H. Wei, T. Wo, Z. Huang, Y. Zhang, and J. Xu, "A unified framework with multi-source data for predicting passenger demands of ride services," *ACM Trans. Knowl. Discov. Data*, vol. 13, no. 6, pp. 56:1–56:24, Dec. 2019.
- [50] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," in *IJCAI*, Aug. 2019, pp. 1907–1913.
- [51] C. Zheng, X. Fan, C. Wang, and J. Qi, "Gman: A graph multi-attention network for traffic prediction," in *AAAI*, vol. 34, no. 1, Feb. 2020, pp. 1234–1241.
- [52] T. Rajabioun and P. A. Ioannou, "On-street and off-street parking availability prediction using multivariate spatiotemporal models," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 5, pp. 2913–2924, May. 2015.
- [53] Y. Liang, K. Ouyang, J. Sun, Y. Wang, J. Zhang, Y. Zheng, D. S. Rosenblum, and R. Zimmermann, "Fine-grained urban flow prediction," in *WWW*, Apr. 2021, pp. 1833–1845.
- [54] E. Suryani, R. Hendrawan, P. Adipraja, and R. Indraswari, "System dynamics simulation model for urban transportation planning: a case study," *International Journal of Simulation Modelling*, vol. 19, no. 1, pp. 5–16, Mar. 2020.



**Chenxi Liu** She is currently pursuing a Ph.D. degree in computer science and technology at Hunan University, China. She is also currently conducting Ph.D. joint training at Nanyang Technological University. Her research interests include spatiotemporal data mining and urban computing.



**Zhu Xiao** received M.S. and Ph.D. degrees in communication and information systems from Xidian University, China, in 2007 and 2010, respectively. From 2010 to 2012, he was a research fellow at the Department of Computer Science and Technology, University of Bedfordshire, U.K. He is currently an associate professor at the College of Computer Science and Electronic Engineering, Hunan University, China. His research interests include mobile communications, wireless localization, the Internet of Vehicles, next-generation communications, and heterogeneous networks. He is a senior member of the IEEE. He is currently an Associate Editor for the *IEEE Transactions on Intelligent Transportation Systems*.



**Hongyang Chen** received the B.S. degree from Southwest Jiaotong University, Chengdu, China, in 2003, M.S. degree in Institute of Mobile Communications, Southwest Jiaotong University, Chengdu, China, in 2006, and Ph.D. degree from the Univ. of Tokyo, in 2011. From 2004 to 2006, he was a research assistant in the Institute of Computing Technology, Chinese Academy of Science. In 2009, he was a visiting researcher in the UCLA Adaptive Systems Laboratory at the University of California Los Angeles. From Apr. 2011 to June 2020, He worked as a researcher for Fujitsu Ltd, Japan. He is currently a Senior Research Expert with Zhejiang Lab, China. His research interests include IoT, data-driven intelligent networking and systems, machine learning, localization, location-based big data, B5G, and statistical signal processing.



**Dong Wang** received M.E. and Ph.D. degrees on computer science from Hunan University, China. He is a Ph.D. director and a director of overseas graduate students in the College of Computer Science and Electronics Engineering, Hunan University. His main research interests are computer networks and vehicular multimedia networks.



**Lei Wang** received the B.S. degree from Southwest Minzu University, China, in 2020. He is currently pursuing the M.S. degree with Hunan University, China. His research interests include traffic data mining and intelligent transportation system.



**Jiangxia Yu** received the B.S. degree, the M.S. degree and the Ph.D. degree from Chang'an University, Shaanxi, China. She is currently an associate professor with the School of Economics & Management, Xidian University, Shaanxi. Her research interests include traffic volume forecast, route choice, system evaluation and optimization.



**Hongbo Jiang** is currently a full professor in the College of Computer Science and Electronic Engineering, Hunan University. He was a professor at Huazhong University of Science and Technology. He received his Ph.D. from Case Western Reserve University in 2008. His research concerns computer networking, especially algorithms and protocols for wireless and mobile networks. He currently serves as the editor for IEEE/ACM Transactions on Networking, associate editor for IEEE Transactions on Mobile Computing, and associate technical editor for