A Spatial-Temporal Attention Multi-Graph Convolution Network for Ride-Hailing Demand Prediction Based on Periodicity with Offset

Dong Xing, Chenguang Zhao, Gang Wang
School of Electronics and Information Engineering, Beihang University
{xingdong, zchenguang, gwang}@buaa.edu.cn

Abstract

Ride-hailing service is becoming a leading part in urban transportation. To improve the efficiency of ride-hailing service, accurate prediction of transportation demand is a fundamental challenge. In this paper, we tackle this problem from both aspects of network structure and data-set formulation. For network design, we propose a spatial-temporal attention multi-graph convolution network (STA-MGCN). A spatial-temporal layer in STA-MGCN is developed to capture the temporal correlations by temporal attention mechanism and temporal gate convolution, and the spatial correlations by multi-graph convolution. A feature cluster layer is introduced to learn latent regional functions and to reduce the computation burden. For the data-set formulation, we develop a novel approach which considers the transportation feature of periodicity with offset. Instead of only using history data during the same time period, the history order demand in forward and backward neighboring time periods from yesterday and last week are also included. Extensive experiments on the three real-world data-sets of New-York, Chicago and Chengdu show that the proposed algorithm achieves the state-of-the-art performance for ride-hailing demand prediction.

1 Introduction

Recent years have witnessed a surge in the ride-hailing service market. Some services providers are rapidly developing all around the world, such as Uber, Left, Via, and Didi. Its ability to meet door-to-door transportation demand anytime anywhere has made it almost ubiquitous in the daily life. To improve the efficiency for ride-hailing, accurate prediction of the demand plays a vital role in guiding various operations, such as vehicle dispatching, vacant vehicle re-allocation, and surge pricing.

Some earliest prediction is realized by statistical methods [Ahmed and Cook, 1979; Moreira-Matias et al., 2013]. Those methods are constructed based on some assumptions, which may not be met in real-world traffic scenarios. The machine learning-based algorithms, especially neural network structured approaches, have grabbed considerable attention for demand prediction, thanks to its ability to discover complicated correlations from data directly [Wang et al., 2017; Zhao et al., 2018; Ke et al., 2017; Liu et al., 2019; Qiu et al., 2019; Ye et al., 2019]. However, the original CNN works only for regular Euclidean domain. The traffic network, on the other side, takes a natural graphical structure. To keep the topology relationship in the graph, graph convolution network (GCN) has been proposed, which can capture the spatial correlations in a graph network [Kipf and Welling, 2016]. Some work has been focusing on ride-hailing service demand prediction with GCN. [Geng et al., 2019; Feng et al., 2021].

Despite all of these efforts, the feature of periodicity with offset, one fundamental characteristics of transportation demand, has been ignored and under poor investigation. To understand this feature, consider the case in Figure 1. Suppose that we want to predict the demand during 8 o’clock and 9 o’clock in Dec. 15, 2018 (yellow), the first approach, (a), takes the most recent history demand as input data, such as during the last three hours 5 o’clock to 8 o’clock (blue). The second approach, (b), considers the periodicity by including the history data during the same period from yesterday and last week, i.e., 8 o’clock to 9 o’clock in Dec. 14 (red) and Dec. 8 (green). However, there may be offset between history and current demand, since the demand for each passenger has some extent of flexibility. For example, a person called for ride-hailing at 7:30 yesterday. But today it is 8:00. Therefore,
some useful information is ignored in (b), which will degrade the demand prediction. The periodicity has been studied in previous research [Guo et al., 2019]. The periodicity with offset, on the other side, is still under poor investigation.

In this paper, we consider the periodicity with offset by including history demand during forward and backward neighboring time slots. As in Figure 1 (c), when predicting demand during 08:00-09:00 Dec. 15, three sources of history demand will be incorporated, recent, 05:00-08:00 Dec. 15, daily, 07:00-10:00 Dec. 14, and weekly, 07:00-10:00 Dec. 8. The enrichment on the input data will bring some challenges to existing algorithms. Although there is similarity between the to-be-predicted time slot and the multiple history order demand, the effect of each history data on the result can be different. Even more, the effect can be dynamic and dependent on the specific data. Therefore, we further propose a novel spatial-temporal attention multi-graph convolution network (STA-MGCN) to discover the temporal and spatial correlations hidden in history order demand. The main contributions of this work are as follows.

- We explicitly consider the feature of periodicity with offset in transportation. Besides recent order demand, history data collected from forward and backward neighboring time slots in yesterday and last week is also included as input to the prediction module.
- We design a novel STA-MGCN to capture the spatial and temporal correlations in the history data. Spatial-Temporal Blocks (ST-Blocks) are introduced to decide the dependency in data, which consist of temporal attention module, multi-graph convolution, and temporal gate convolution.
- The proposed scheme is exhaustively tested through simulation on three real-world data-sets of New-York, Chicago, and Chengdu. For both short-term and long-term prediction, the proposed STA-MGCN outperforms state-of-the-art algorithms under multiple evaluation metrics. We also test some variants of the proposed algorithm to demonstrate the role of some modules in the proposed structure.

The remainder of this paper is structured as follows. We first summarize some related work on the topic of order demand prediction in Section 2. Then we formulate the prediction problem in Section 3. In Section 4, we introduce the proposed STA-MGCN algorithm, which is tested on three real-world data-sets in Section 5.

2 Related Work

The order demand prediction for ride-hailing service falls into the general category of temporal-spatial traffic prediction, which can trace back to 1980s [Ahmed and Cook, 1979]. In its early period, the main prediction methods are some statistical approaches. Among them, auto-regressive integrated moving average (ARIMA) has attracted particular attention [Moreira-Matias et al., 2013]. Despite the mathematical guarantee, it is constructed based on some assumptions, which may not be satisfied by the complicated real-world traffic data.

The development of machine learning (ML) has shed new light on traffic prediction. Multiple research has been carried out to realize ML’s ability to discover complicated mapping relationship from a large volume of data. Some earliest efforts are made from various approaches, such as KNN [Van Lint and Van Hinsbergen, 2012; Tak et al., 2014], SVM [Jeong et al., 2013], random forest [Leshem and Ritov, 2007; Yang and Qian, 2018], Bayesian network [Zhu et al., 2016] and so on. The breakthrough in deep leaning has further promoted the research. Some deep neural networks have yielded considerable improvement on prediction accuracy [Wang et al., 2017; Zhao et al., 2018; Qiu et al., 2019; Ke et al., 2017; Ye et al., 2019; Liu et al., 2019].

However, the CNN-based structure can work only for regular Euclidean domain, such as 2D or 3D grids. In a transportation network, however, the regions is formulated by a graph structure. To incorporate graph structure into regular CNN, the above work first transfers the traffic data into image-type data, which destroys the topology structure in the traffic network. To recognize the graphical structure, graph convolution has been proposed, with two main streams, spatial [Niepert et al., 2016] and spectral [Niepert et al., 2016]. The GCN-based demand prediction has been realized in some related work [Yu et al., 2017; Geng et al., 2019; Guo et al., 2019; Pan et al., 2019; Wu et al., 2019; Feng et al., 2021; Guo et al., 2021]. However, the periodicity with offset, which is one fundamental feature in traffic, is still under poor investigation.
generate the predicted demand. The details for each module will be introduced as follows.

4.1 Input Data

In this paper, we will consider explicitly the periodicity with offset in transportation. History demand from three different sources will be utilized, recent, yesterday, and last week. We adopt $\varsigma$ as subscript to different them. The three sources of input are then $X_{\varsigma}$ with $\varsigma \in \{r, d, w\}$.

For each input data $X_\varsigma$, suppose that history demand from $T_\varsigma$ slots are included, then we have that $X_\varsigma \in \mathbb{R}^{N \times F \times T_\varsigma}$.

4.2 Temporal Attention Mechanism

The temporal attention layer is involved to reveal the correlations between different time slots. Suppose that the input is $X^{in} \in \mathbb{R}^{N \times F \times T}$, the attention matrix $V \in \mathbb{R}^{T \times T}$ is calculated as [Guo et al., 2019]

$$V = W \cdot \sigma \left( \left( (X^{ir})^T \cdot U_1 \right) \cdot \left( U_2 \cdot (U_3 \cdot X^{ir}) + b \right) \right)$$

where $X^{ir} \in \mathbb{R}^{c_{in} \times T}$ is reshaped from $X^{in}$ with $c_{in} = N \times F_{in}$, $U_1 \in \mathbb{R}^{c_{in} \times c_1}$, $U_2 \in \mathbb{R}^{c_1 \times c_2}$, $U_3 \in \mathbb{R}^{c_2 \times c_{in}}$, $W \in \mathbb{R}^{T \times T}$, and $b \in \mathbb{R}^{T \times T}$ are trainable parameters. We take the activation function $\sigma$ as ReLu. The value $v_{i,j}$ represents the dependency of time slot $i$ on $j$. To incorporate the inter-dependency between time slots, the input history traffic demand is weighted by the regularized temporal attention matrix.

$$X_1 = XV'$$

with

$$v'_{i,j} = \frac{\exp(v_{i,j})}{\sum_{j} \exp(v_{i,j})}.$$  

The output is then $X^T \in \mathbb{R}^{N \times F_{in} \times T}$, which is reshaped from $X_1 \in \mathbb{R}^{c_{in} \times T}$.

4.3 Multi-Graph Convolution

The multi-graph convolution aims to capture the spatial correlations among regions [Geng et al., 2019]. Suppose that we have a distance metric $\varrho$, which can calculate a specific kind of distance $\varrho_{i,j}$, for two node $i$ and $j$. The adjacency matrix $A_\varrho \in \mathbb{R}^{N \times N}$ is then defined as

$$[A_{\varrho}]_{i,j} = \begin{cases} 1, & \text{if } j \in A_{\varrho_{i}}, \\ 0, & \text{Otherwise} \end{cases}$$

where the set $A_{\varrho_{i}}$ contains the top 10% nodes closest to node $i$ with respect to the meaning of distance $\varrho$.

In this paper, we consider three type of adjacency, geographical, functional, and route distance.

- geographical distance $A_D$. The distance is defined as the geographical distance between node $i$ and node $j$ in the traffic network. It is intuitive that when two nodes are close, they may enjoy a similar traffic pattern.

- functional distance $A_F$. We first assign a function characteristic vector $f_{i}$ to each node $i$. The distance is then calculated based on this function vector as $\varrho_{i,j} = \|f_{i} - f_{j}\|$. For nodes that are remote in geography, their traffic patterns may still have approximate trend. For example, ride-hailing demand around residence areas may be similar, even if they are not geographically adjacent.

- mobility distance $A_M$. For two nodes $i$ and $j$, the distance is the inverse of number of history orders between them (If there is no order, the distance is then 0).

For the output data $X^T$ from the temporal attention layer, the multi-graph convolution calculates for each time slot $t$ as

$$X^M_t = \sigma \left( \sum_\varrho f(A_\varrho)X^T_t M_\varrho \right),$$

where $f(A_\varrho) \in \mathbb{R}^{N \times N}$ is a spectral filter on graphs constructed from the adjacency matrix [Kipf and Welling, 2016; Geng et al., 2019], and $M \in \mathbb{R}^{F_{in} \times F_{in}}$ is the trainable feature transformation matrix. The activation function $\sigma$ is taken as ReLu. $X^M \in \mathbb{R}^{N \times F_{in} \times T}$ is the output by the multi-graph convolution layer.

4.4 Temporal Gate Convolution

The temporal gate convolution is introduced to further extract the temporal correlations in the data [Guo et al., 2021]. In this layer, we adopt the dilated causal convolution (DCC) proposed by [Wu et al., 2019] to capture correlations for time series data during a long period.

$$X^D = DCC \left( X^M \right),$$

where $f$ is the element-wise Hadamard product. The two activation functions $\sigma_1$ and $\sigma_2$ are taken as tanh and sigmoid respectively.
Figure 3: Zone partitions for the three data-sets

Table 1: Detailed information of the three data-sets

|          | New-York | Chicago | Chengdu |
|----------|----------|---------|---------|
| Start Date | Jan. 1, 2018 | Jan. 1, 2018 | Nov. 1, 2016 |
| End Date  | Dec. 31, 2018 | Dec. 31, 2018 | Nov. 30, 2016 |
| Regions   | 63       | 77      | 90      |
| Orders    | 18,981k  | 42,092k | 5,901k  |

4.5 Matrix Factorization
Denote the output from the last ST-Block as $X^{ST} \in \mathbb{R}^{N \times F}$. While the spatial correlations has been captured in the ST-Blocks by graph convolution, the shared weight matrix in the neural network can disregard the uniqueness for each node. To accentuate distinction among regions, we adopt the matrix factorization technology to capture the individual feature for each regions [Pan et al., 2019]. The output is $X^{F} \in \mathbb{R}^{N \times F}$ where $F$ is the number of features in the traffic graph $G$.

4.6 Output Predictions
Finally, the prediction $X_{\varsigma}^{F} \in \mathbb{R}^{N \times F}$ from multiple history data $X^{F}$ is fused to generalize the prediction $\hat{X}$. To achieve adaptive balance of the multiple history data, the final prediction is calculated as a weighted sum of the multiple predictions

$$\hat{X} = \sum_{\varsigma} W_{\varsigma} \odot X^{F}_{\varsigma},$$

where $W_{\varsigma} \in \mathbb{R}^{N \times F}$ are trainable weighting matrices.

5 Experiments
In this section, we aim to test the proposed STA-MGCN and analyze the results.

5.1 Data-Sets
We run simulation experiments on three real-world data-sets, New-York1, Chengdu2, and Chicago3. The original data records can be downloaded from the website. In Table 1, we list more detailed information for the three data-sets, including the start date, end date, number of regions, and number of collected orders.

1https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page
2https://gaia.didichuxing.com
3https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Table 2: New-York 60min Inflow

| Model            | MAE   | RMSE  | MAPE  |
|------------------|-------|-------|-------|
| TCN              | 6.629 | 9.973 | 32.86 |
| ASTGCN           | 6.267 | 10.305| 33.20 |
| STMGCN           | 5.632 | 8.214 | 28.98 |
| MT-MFGCN         | 5.861 | 8.577 | 30.42 |
| STA-MGCN(no)     | 5.683 | 8.388 | 30.91 |
| STA-MGCN(fc)     | 5.440 | 7.923 | 29.15 |
| STA-MGCN         | 5.415 | 7.900 | 28.14 |

Table 3: New-York 60min Outflow

| Model            | MAE   | RMSE  | MAPE  |
|------------------|-------|-------|-------|
| TCN              | 6.602 | 9.843 | 32.61 |
| ASTGCN           | 6.553 | 10.475| 34.46 |
| STMGCN           | 5.789 | 8.573 | 29.64 |
| MT-MFGCN         | 6.089 | 9.093 | 31.52 |
| STA-MGCN(no)     | 5.878 | 8.737 | 31.43 |
| STA-MGCN(fc)     | 5.709 | 8.568 | 30.74 |
| STA-MGCN         | 5.712 | 8.515 | 29.22 |

In Figure 3, we plot the partition of regions for the three data-sets. For the New-York data-set, the partition follows the TLC zone id given in its website. For the Chicago data-set, the partition is based on the community area. And for the Chengdu data-set, the regions are partitioned by the main roads.

For all the three data-sets, we consider two types of demand, inflow, and outflow, which refers to the order terminated and originated at a node respectively. So we have $F = 2$.

5.2 Baselines
We compare the proposed STA-MGCN with several state-of-the-art algorithms for ride-hailing demand prediction.

- TCN (Temporal Convolution Network) [Wu et al., 2019]. A structure designed to capture the temporal correlations for long sequences.
- ASTGCN(Attention based spatial-temporal GCN) [Guo et al., 2019]. The periodicity of the transportation is considered but the offset is not. Only the history data from recent, and from the same time-slot in yesterday last week in included.
- STMGCN (Spatial temporal multi-graph GCN) [Geng et al., 2019]. The contextual information of the traffic network is incorporated, and the multi-graph convolution is adopted to capture the correlations.
- MT-MFGCN (Multi-Task Matrix Factorized GCN) [Feng et al., 2021]. A multi-task matrix factorization GCN for simultaneous zone-based and OD-based demand prediction.

We also test the performance for some variants of the proposed STA-MGCN.
Figure 4: Predicted and observed demand under New-York data-set.

- STA-MGCN(no). The input data ignores the feature periodicity with offset, and only considers recent demand, demand in the same time slot from yesterday and last week.
- STA-MGCN(fc). The matrix factorization layer is replaced by a fully connected layer.

5.3 Evaluation Metric
To evaluate the performance of the demand prediction, we adopt three commonly used evaluation metrics. Denote the set of all time slots as $\mathcal{T}$. The three evaluation metrics are defined as

- Mean Absolute Error (MAE)
  $$\frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \| \hat{X}_t - X_t \|.$$

- Root Mean Square Error (RMSE)
  $$\sqrt{\frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \| \hat{X}_t - X_t \|^2}.$$

- Mean Absolute Percentage Error (MAPE)
  $$\frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \frac{\| \hat{X}_t - X_t \|}{\| X_t \|} \times 100\%.$$

5.4 Result Analysis
We consider both short-term prediction and long-term prediction. For the New-York and Chicago data-sets, duration of each time-slot for short-term and long-term are $\tau = 30$ and 60 minutes respectively. For the Chengdu data-set, they are 10 minutes and 20 minutes.

In Table 2 to 5, we list the prediction results in the New-York data-set, for inflow and outflow, long-term and short-term. The results for Chicago and Chengdu are listed in Table 6 to 9, and Table 10 to 13 respectively. Compared with the stat-of-the-arts, it can be seen that the proposed STA-MGCN generates more accurate prediction.

5.5 Case Study
In Figure 4, we plot the predicted demand and ground truth in the zone 243 in the New-York data-set from Dec. 10 to Dec. 16. It can be seen that the proposed STA-MGCN approximately predicts order demand.

To provide a more understanding of the results, we visualize the both long-term and short-term demand the New-York data-set. We pick three time slots in the day Dec. 14. The true and predicted order demand for long-term and short-term are presented in Figure 5 and 6 respectively.

6 Conclusion
In this paper, we propose to explicitly exploit the feature of periodicity with offset in transportation. Besides history data in the same time slot, order demand during neighboring periods is also included as input. To capture the temporal and spatial correlations hidden in the history data, we design a novel STA-MGCN structure, which contains three main modules, ST-Block, matrix factorization, and history fusion. We have run simulations on three real-world data-sets for both short-term and long-term prediction. Compared with several state-of-the-arts, remarkable improvement on the prediction accuracy can be achieved by the proposed STA-MGCN, which demonstrates its superiority.
### Table 4: New-York 30min Inflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 4.392| 6.410| 39.86 |
| ASTGCN         | 3.679| 5.293| 35.61 |
| STMGCN         | 3.610| 5.115| 33.79 |
| MT-MFGCN       | 3.680| 5.282| 40.22 |
| STA-MGCN(no)   | 3.617| 5.076| 37.94 |
| STA-MGCN(fc)   | 3.655| 5.138| 35.74 |
| STA-MGCN       | **3.550**| **4.992**| **35.17**|

### Table 5: New-York 30min Outflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 4.512| 6.560| 41.23 |
| ASTGCN         | 3.963| 5.846| 38.04 |
| STMGCN         | 3.692| 5.284| 35.35 |
| MT-MFGCN       | 3.795| 5.487| 41.35 |
| STA-MGCN(no)   | 3.730| 5.300| 38.32 |
| STA-MGCN(fc)   | 3.978| 6.075| 37.62 |
| STA-MGCN       | **3.685**| **5.236**| **35.21**|

### Table 6: Chicago 60min Inflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 6.634| 12.710| 30.35 |
| ASTGCN         | 6.563| 15.315| 27.31 |
| STMGCN         | 6.371| 13.849| 27.09 |
| MT-MFGCN       | 6.367| 13.470| 27.73 |
| STA-MGCN(no)   | 5.823| 11.731| 26.56 |
| STA-MGCN(fc)   | 5.728| 11.014| 27.33 |
| STA-MGCN       | **5.657**| **10.638**| **27.25**|

### Table 7: Chicago 60min Outflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 6.781| 12.936| 30.17 |
| ASTGCN         | 6.478| 14.618| 26.59 |
| STMGCN         | 6.522| 14.050| 26.47 |
| MT-MFGCN       | 6.407| 13.465| 27.50 |
| STA-MGCN(no)   | 5.889| 11.375| 26.38 |
| STA-MGCN(fc)   | 5.823| 11.122| 26.30 |
| STA-MGCN       | **5.684**| **10.843**| **26.38**|

### Table 8: Chicago 30min Inflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 4.125| 7.123| 36.33 |
| ASTGCN         | 3.753| 6.687| 32.59 |
| STMGCN         | 3.752| 6.495| 34.06 |
| MT-MFGCN       | 3.716| 6.392| 33.09 |
| STA-MGCN(no)   | 3.658| 6.315| 32.72 |
| STA-MGCN(fc)   | 3.637| 6.128| 33.00 |
| STA-MGCN       | **3.607**| **6.086**| **32.78**|

### Table 9: Chicago 30min Outflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 4.310| 7.385| 36.83 |
| ASTGCN         | 3.824| 6.690| 32.96 |
| STMGCN         | 3.858| 7.673| 34.01 |
| MT-MFGCN       | 3.824| 6.627| 33.23 |
| STA-MGCN(no)   | 3.769| 6.494| 32.42 |
| STA-MGCN(fc)   | 3.776| 6.435| 33.49 |
| STA-MGCN       | **3.735**| **6.331**| **32.93**|

### Table 10: Chengdu 20min Inflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 5.370| 8.415| 31.04 |
| ASTGCN         | 4.966| 7.905| 28.95 |
| STMGCN         | 4.652| 7.096| 28.76 |
| MT-MFGCN       | 4.795| 7.458| 29.03 |
| STA-MGCN(no)   | 4.495| 6.805| 27.24 |
| STA-MGCN(fc)   | 4.542| 6.862| 26.88 |
| STA-MGCN       | **4.469**| **6.762**| **26.57**|

### Table 11: Chengdu 20min Outflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 5.605| 8.502| 33.96 |
| ASTGCN         | 4.883| 7.371| 30.30 |
| STMGCN         | 4.709| 7.007| 28.76 |
| MT-MFGCN       | 4.690| 7.050| 28.60 |
| STA-MGCN(no)   | 4.798| 7.745| 28.99 |
| STA-MGCN(fc)   | 4.603| 6.909| 27.49 |
| STA-MGCN       | **4.576**| **6.902**| **27.78**|

### Table 12: Chengdu 10min Inflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 3.456| 5.153| 35.57 |
| ASTGCN         | 3.254| 4.840| 33.74 |
| STMGCN         | 3.248| 4.788| 33.58 |
| MT-MFGCN       | 3.147| 4.650| 33.04 |
| STA-MGCN(no)   | 3.159| 4.665| 32.68 |
| STA-MGCN(fc)   | 3.157| 4.658| 33.37 |
| STA-MGCN       | **3.145**| **4.637**| **32.28**|

### Table 13: Chengdu 10min Outflow

| Model          | MAE  | RMSE | MAPE  |
|----------------|------|------|-------|
| TCN            | 3.584| 5.276| 37.91 |
| ASTGCN         | 3.337| 4.906| 35.07 |
| STMGCN         | 3.287| 4.808| 34.50 |
| MT-MFGCN       | 3.192| 4.668| 33.66 |
| STA-MGCN(no)   | 3.183| 4.651| 33.94 |
| STA-MGCN(fc)   | 3.186| 4.650| 33.66 |
| STA-MGCN       | **3.189**| **4.667**| **33.60**|
References

[Ahmed and Cook, 1979] Mohammed S Ahmed and Allen R Cook. Analysis of freeway traffic time-series data by using Box-Jenkins techniques. Number 722. 1979.

[Feng et al., 2021] Siyuan Feng, Jintao Ke, Hai Yang, and Jieping Ye. A multi-task matrix factorized graph neural network for co-prediction of zone-based and od-based ride-hailing demand. IEEE Transactions on Intelligent Transportation Systems, 2021.

[Geng et al., 2019] Xu Geng, Yaguang Li, Leye Wang, Lingyu Zhang, Qiang Yang, Jieping Ye, and Yan Liu. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting. In Proceedings of the AAAI conference on artificial intelligence, volume 33, pages 3656–3663, 2019.

[Guo et al., 2019] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatio-temporal graph convolutional networks for traffic flow forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, pages 922–929, 2019.

[Guo et al., 2021] Kan Guo, Yongli Hu, Yanfeng Sun, Sean Qian, Junbin Gao, and Baocai Yin. Hierarchical graph convolution networks for traffic forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, pages 151–159, 2021.

[Jeong et al., 2013] Young-Seon Jeong, Young-Ji Byon, Manoel Mendonça Castro-Neto, and Said M Easa. Supervised weighting-online learning algorithm for short-term traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 14(4):1700–1707, 2013.

[Ke et al., 2017] Jintao Ke, Hongyu Zheng, Hai Yang, and Xiquan Michael Chen. Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. Transportation Research Part C: Emerging Technologies, 85:591–608, 2017.

[Kipf and Welling, 2016] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907, 2016.

[Leshem and Ritov, 2007] Guy Leshem and Yaacov Ritov. Traffic flow prediction using adaboost algorithm with random forests as a weak learner. In Proceedings of world academy of science, engineering and technology, volume 19, pages 193–198. Citeseer, 2007.

[Liu et al., 2019] Lingbo Liu, Zhihui Qu, Guanbin Li, Qing Wang, Wanli Ouyang, and Liang Lin. Contextualized spatial–temporal network for taxi origin-destination demand prediction. IEEE Transactions on Intelligent Transportation Systems, 20(10):3875–3887, 2019.

[Moreira-Matias et al., 2013] Luis Moreira-Matias, Joao Gama, Michel Ferreira, Joao Mendes-Moreira, and Luis Damas. Predicting taxi–passenger demand using streaming data. IEEE Transactions on Intelligent Transportation Systems, 14(3):1393–1402, 2013.

[Niepert et al., 2016] Mathias Niepert, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. In International conference on machine learning, pages 2014–2023. PMLR, 2016.

[Pan et al., 2019] Zheyi Pan, Zhaoyuan Wang, Weifeng Wang, Yong Yu, Junbo Zhang, and Yu Zheng. Matrix factorization for spatio-temporal neural networks with applications to urban flow prediction. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pages 2683–2691, 2019.

[Qiu et al., 2019] Zhihui Qu, Lingbo Liu, Guanbin Li, Qing Wang, Nong Xiao, and Liang Lin. Taxi origin-destination demand prediction with contextualized spatio-temporal network. In 2019 IEEE International Conference on Multimedia and Expo (ICME), pages 760–765. IEEE, 2019.

[Tak et al., 2014] Sehyun Tak, Sunghoon Kim, Kiate Jang, and Hwasoo Yeo. Real-time travel time prediction using multi-level k-nearest neighbor algorithm and data fusion method. In Computing in Civil and Building Engineering (2014), pages 1861–1868. 2014.

[Van Lint and Van Hinsbergen, 2012] JWC Van Lint and CPIJ Van Hinsbergen. Short-term traffic and travel time prediction models. Artificial Intelligence Applications to Critical Transportation Issues, 22(1):22–41, 2012.

[Wang et al., 2017] Dong Wang, Wei Cao, Jian Li, and Jieping Ye. Deepsd: Supply-demand prediction for online car-hailing services using deep neural networks. In 2017 IEEE 33rd international conference on data engineering (ICDE), pages 243–254. IEEE, 2017.

[Wu et al., 2019] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. Graph wavenet for deep spatio-temporal graph modeling. arXiv preprint arXiv:1906.00121, 2019.

[Yang and Qian, 2018] Shuguan Yang and Zhen Sean Qian. Understanding and predicting roadway travel time with spatio-temporal features of network traffic flow, weather conditions and incidents. Technical report, 2018.

[Ye et al., 2019] Junchen Ye, Leilei Sun, Bowen Du, Yanjie Fu, Xinran Tong, and Hui Xiong. Co-prediction of multiple transportation demands based on deep spatio-temporal neural network. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 305–313, 2019.

[Yu et al., 2017] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875, 2017.

[Zhao et al., 2018] Jing Zhao, Jiajie Xu, Rui Zhou, Pengpeng Zhao, Chengfei Liu, and Feng Zhu. On prediction of user destination by sub-trajectory understanding: A deep learning based approach. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pages 1413–1422, 2018.

[Zhu et al., 2016] Zheng Zhu, Bo Peng, Chenfeng Xiong, and Lei Zhang. Short-term traffic flow prediction with linear conditional gaussian bayesian network. Journal of Advanced Transportation, 50(6):1111–1123, 2016.