Beijing Jiaotong University
Institute of Network Science and Intelligent System

Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting

Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, HuaiyuWan*

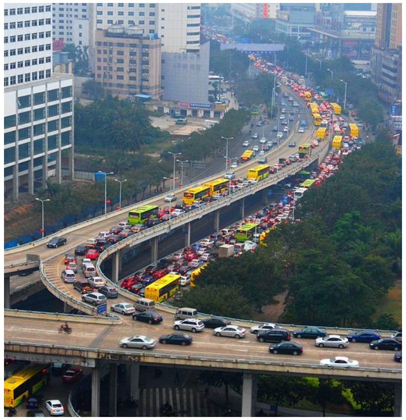
Introduction

Background

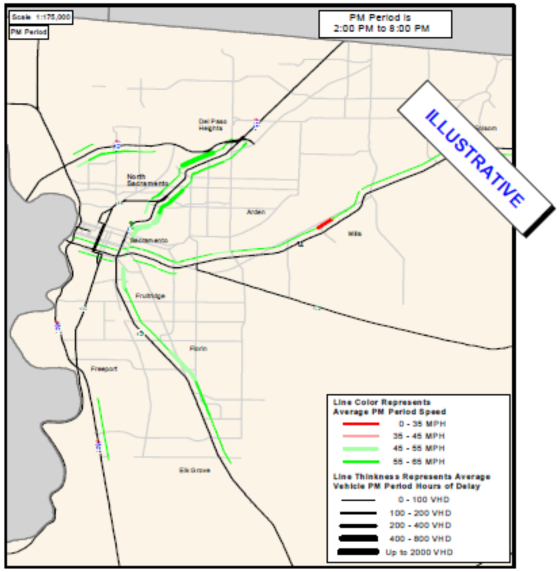
- Many countries are committed to developing Intelligent Transportation System (ITS)
- Traffic forecasting is an indispensable part of ITS
- Serious economic loss caused by traffic congestion

Value

- Traffic management & Traffic capacity
- Risk assessment
- Public safety



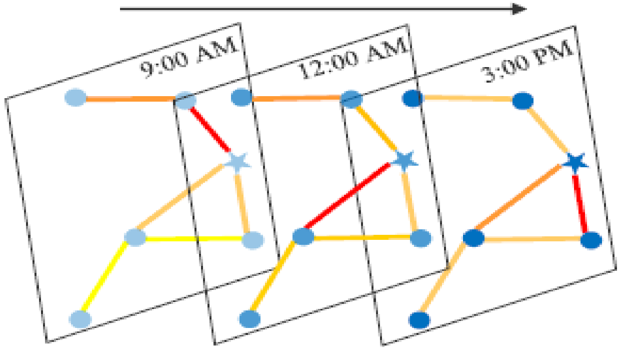
District 3 Congestion Monitoring Example



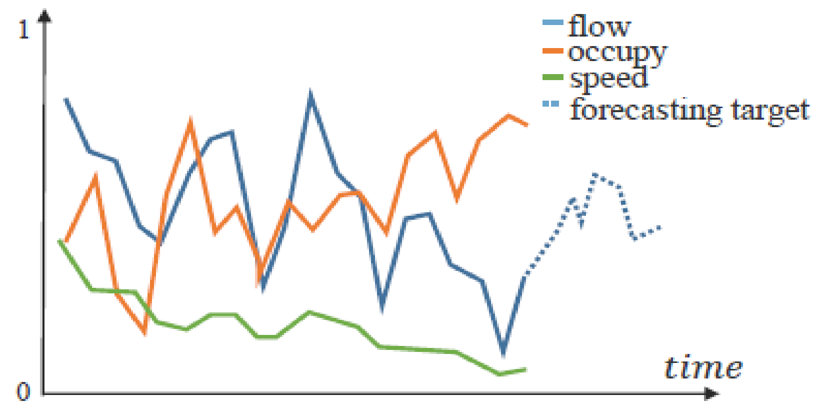
Preliminaries

Traffic Networks

- **Spatial**
  - A traffic network: $G = (V, E, A)$
- **Temporal**
  - Each node on the traffic network G detects F measurements with the same sampling frequency



(a)



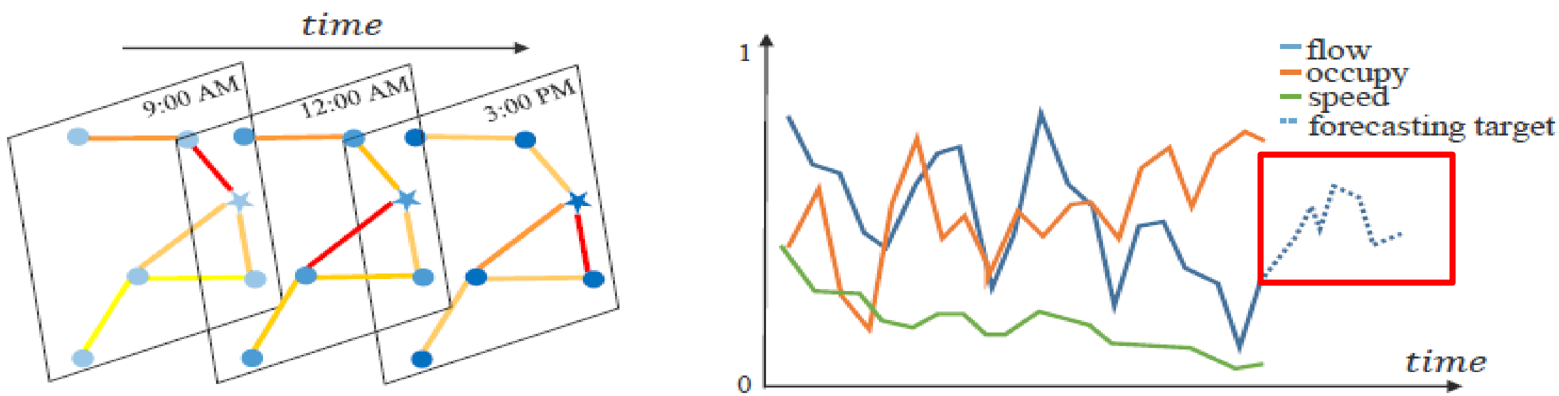
(b)

$$\mathcal{X} = (X_1, X_2, \ldots, X_\tau)^T \in \mathbb{R}^{N \times F \times \tau}$$

Preliminaries

Traffic Flow Forecasting

- Given all kinds of the historical measurements of all the nodes on the traffic network over past time slices, predict future traffic flow sequences of all the nodes over the following time slices.

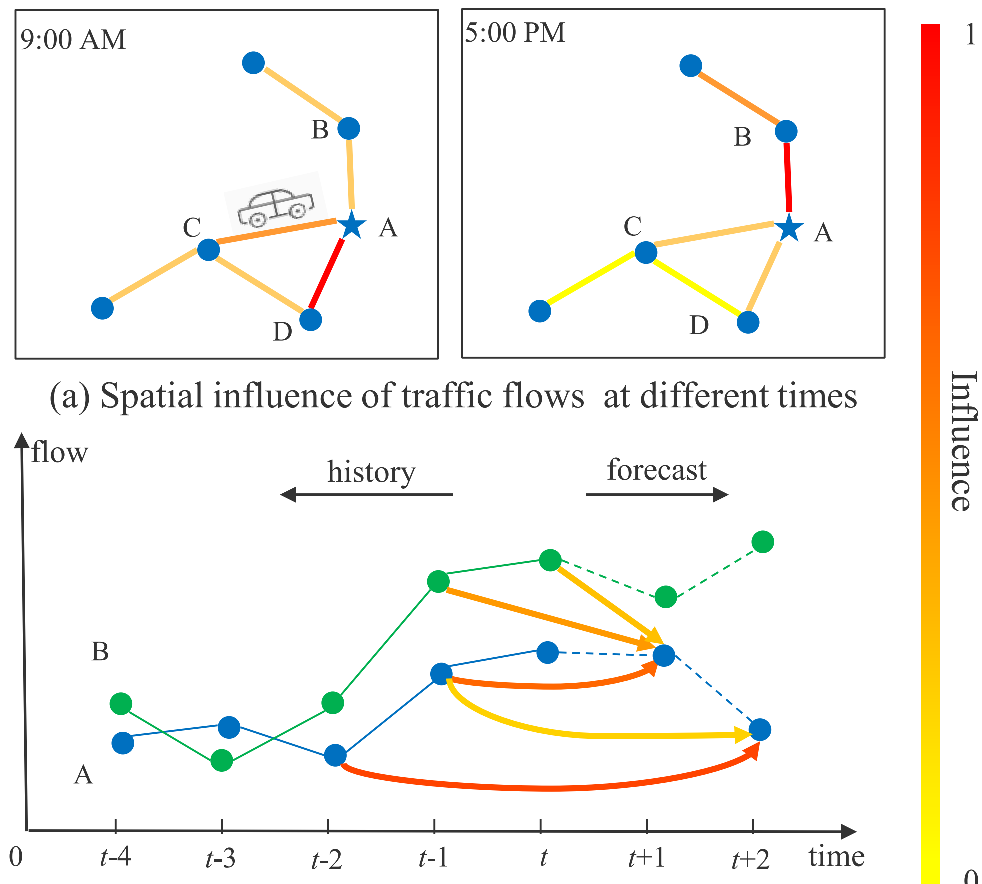


(a) (b)

Challenges

Dynamical spatial-temporal correlations

- Spatial
- Temporal



(a) Spatial influence of traffic flows at different times

(b) Temporal influence between traffic flows

Related Work

- **Traffic forecasting**
  - **statistical models**
    - HA
    - ARIMA, VAR
  - **traditional machine learning models**
    - KNN, SVR
  - **deep learning models**
    - ST-ResNet [1], DMVST-Net [2], GeoMAN [3]

[1] Zhang, J.; Zheng, Y.; Qi, D.; Li, R.; Yi, X.; and Li, T. 2018. Predicting citywide crowd flows using deep spatio-temporal residual networks. Artificial Intelligence 259:147-166.

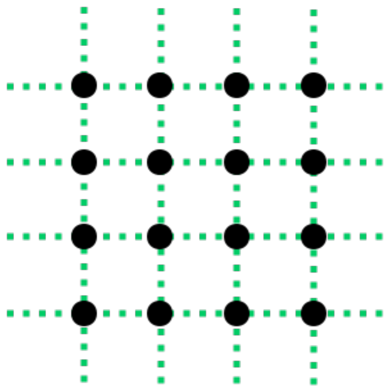
[2] Yao, H.; Wu, F.; Ke, J.; Tang, X.; Jia, Y.; Lu, S.; Gong, P.; and Ye, J. 2018b. Deep multi-view spatial-temporal network for taxi demand prediction. In AAAI Conference on Artificial Intelligence, 2588-2595.

[3] Liang, Y.; Ke, S.; Zhang, J.; Yi, X.; and Zheng, Y. 2018. GeoMAN: Multi-level Attention Networks for Geo-sensory Time Series Prediction. In International Joint Conference on Artificial Intelligence, 3428-3434.

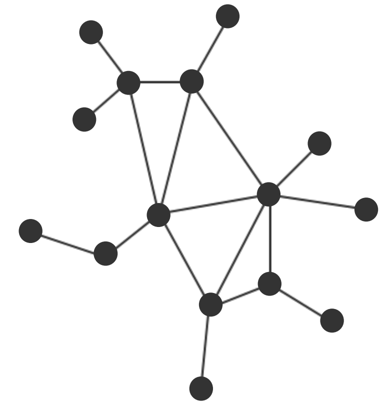
Related Work

■ Convolutions on graphs

- Standard grid data



- Data of graph structure



- GLU-STGCN [4]

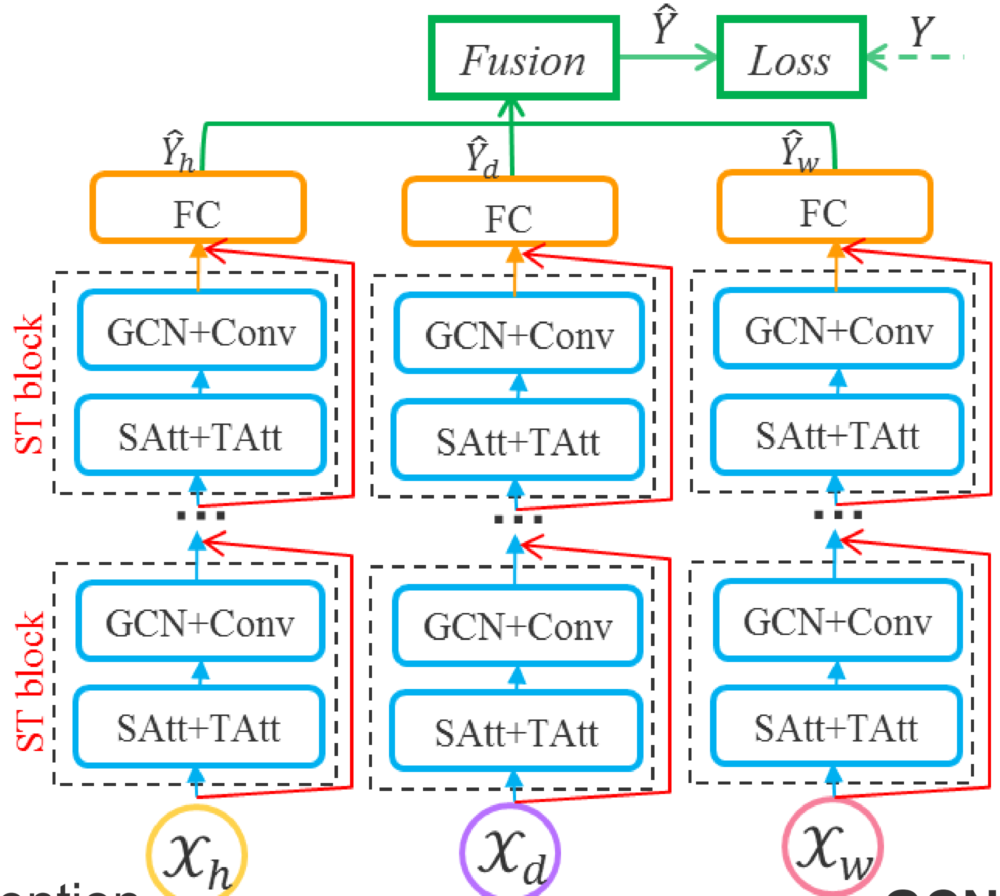
■ Attention mechanism

- Natural language processing
- Image caption
- Speech recognition

[4] Yu, B.; Yin, H.; and Zhu, Z. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In International Joint Conference on Artificial Intelligence, 3634-3640.

ASTGCN Architecture

- *Attention based Spatial-Temporal Graph Convolutional Network*

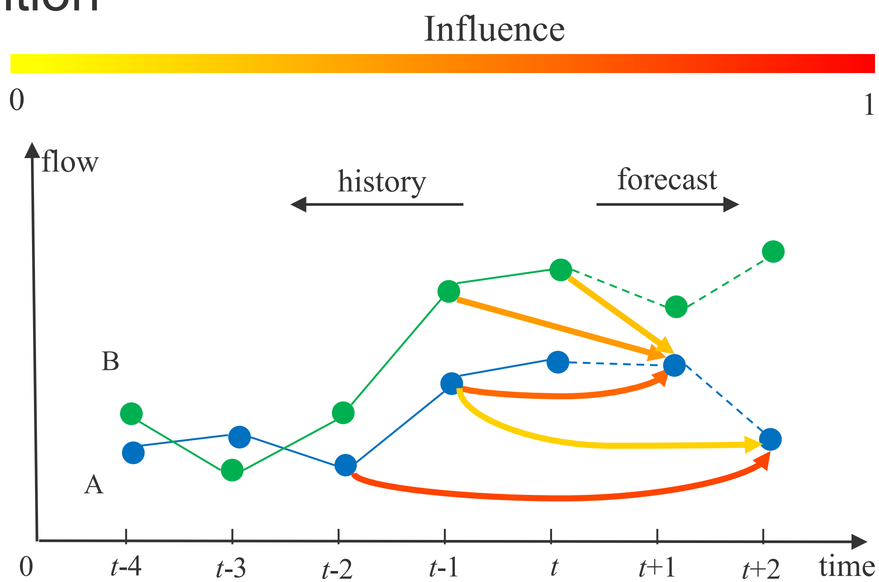


- **SAtt** Spatial Attention
- **TAtt** Temporal Attention
- **GCN** Graph Convolution
- **Conv** Convolution

ASTGCN Architecture

Spatial-Temporal Attention

- Temporal attention



(b) Temporal influence between traffic flows

$$\mathbf{E} = \mathbf{V}_e \cdot \sigma(((\mathcal{X}_h^{(r-1)})^T \mathbf{U}_1)\mathbf{U}_2(\mathbf{U}_3 \mathcal{X}_h^{(r-1)}) + \mathbf{b}_e)$$

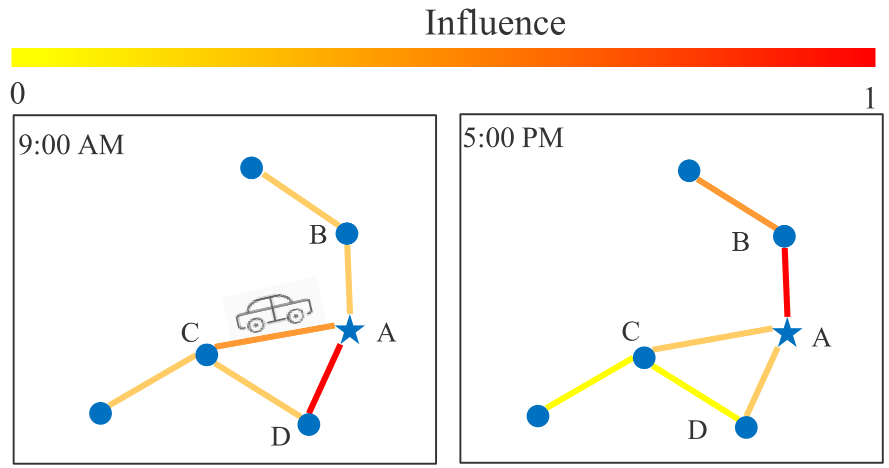
$$\mathbf{E}'_{i,j} = \frac{\exp(\mathbf{E}_{i,j})}{\sum_{j=1}^{T_{r-1}} \exp(\mathbf{E}_{i,j})}$$

where $\mathbf{V}_e, \mathbf{b}_e \in \mathbb{R}^{T_{r-1} \times T_{r-1}}$, $\mathbf{U}_1 \in \mathbb{R}^N$, $\mathbf{U}_2 \in \mathbb{R}^{C_{r-1} \times N}$, $\mathbf{U}_3 \in \mathbb{R}^{C_{r-1}}$

$$\begin{aligned}\hat{\mathcal{X}}_h^{(r-1)} &= (\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, ..., \hat{\mathbf{X}}_{T_{r-1}}) \\ &= (\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_{T_{r-1}})\mathbf{E}'\end{aligned}$$

■ Spatial-Temporal Attention

● Spatial attention



(a) Spatial influence of traffic flows at different times

$$\mathbf{S} = \mathbf{V}_s \cdot \sigma((\boldsymbol{\mathcal{X}}_h^{(r-1)}\mathbf{W}_1)\mathbf{W}_2(\mathbf{W}_3\boldsymbol{\mathcal{X}}_h^{(r-1)})^T + \mathbf{b}_s)$$

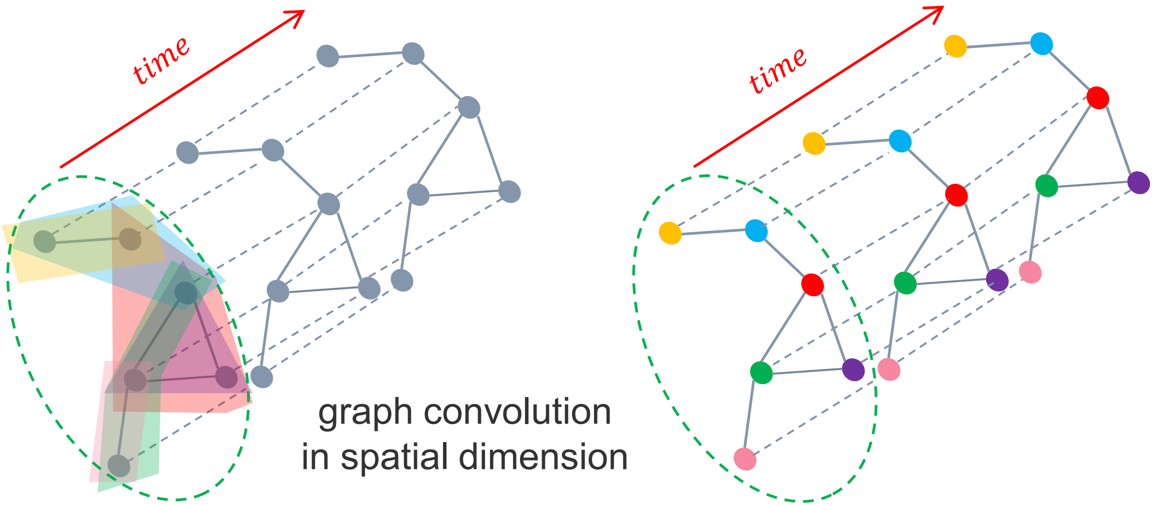
$$\mathbf{S}'_{i,j} = \frac{\exp(\mathbf{S}_{i,j})}{\sum_{j=1}^{N}\exp(\mathbf{S}_{i,j})}$$

$$\mathbf{V}_s, \mathbf{b}_s \in \mathbb{R}^{N \times N}, \mathbf{W}_1 \in \mathbb{R}^{T_{r-1}}, \mathbf{W}_2 \in \mathbb{R}^{C_{r-1} \times T_{r-1}}, \mathbf{W}_3 \in \mathbb{R}^{C_{r-1}}$$

ASTGCN Architecture

Spatial-Temporal Convolution

- Spatial graph convolution



In spectral graph analysis, the properties of the graph structure can be obtained by analyzing Laplacian matrix and its eigenvalues.

$$g_\theta *_G x = g_\theta(\mathbf{L})x$$

$$= \sum_{k=0}^{K-1} \theta_k T_k(\tilde{\mathbf{L}})x$$ [5]

SAtt

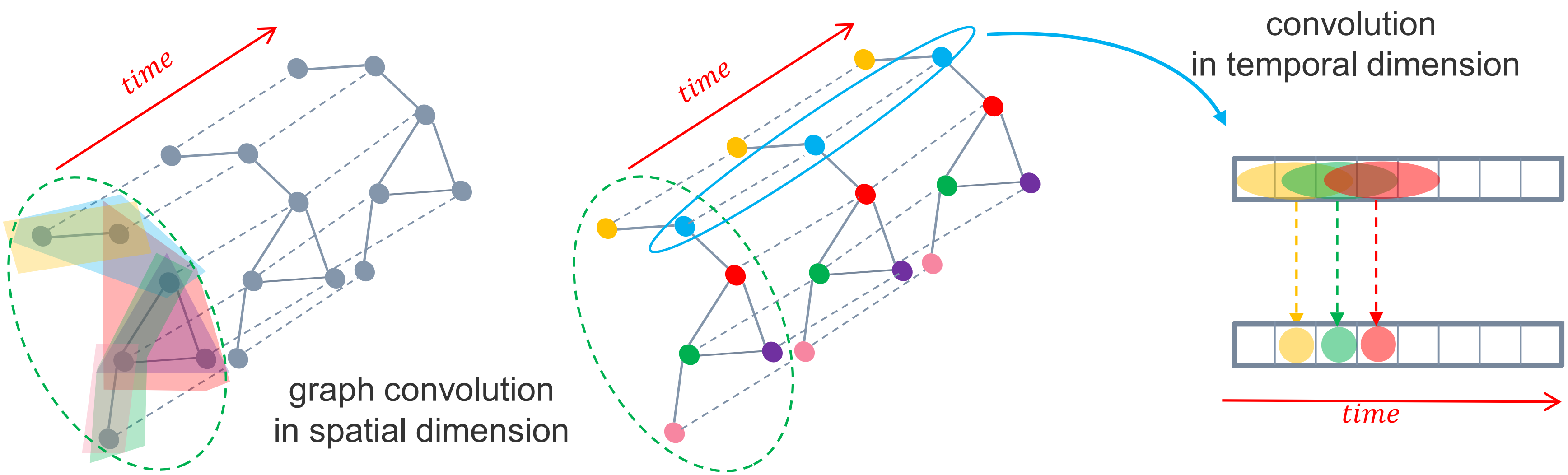
$$\sum_{k=0}^{K-1} \theta_k (T_k(\tilde{\mathbf{L}}) \odot \mathbf{S}')x$$

[5] Defferrard, M.; Bresson, X.; and Vandergheynst, P. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In Advances in Neural Information Processing Systems, 3844–3852.

ASTGCN Architecture

■ Spatial-Temporal Convolution

- Temporal convolution



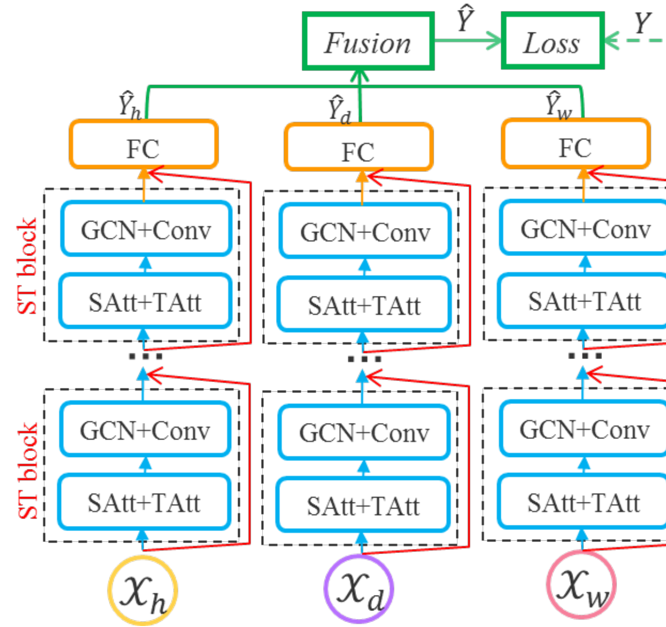
ASTGCN Architecture

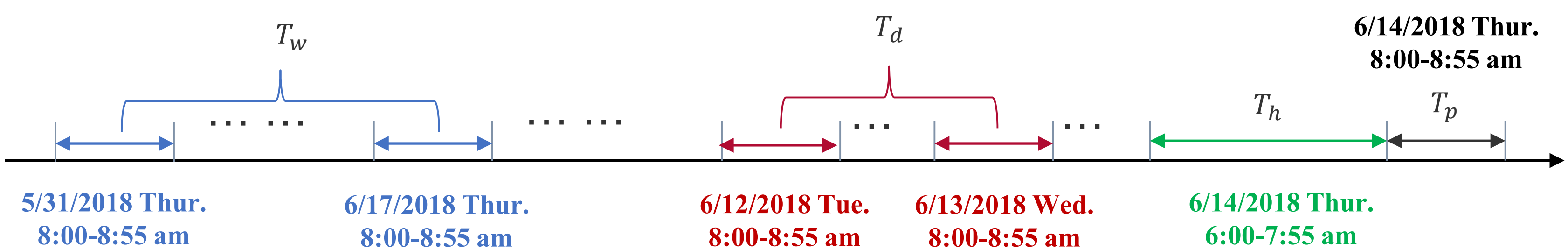
Multi-Component Fusion

- Intercept three time series segments of length T_h, T_d and T_w along the time axis as the input of three components respectively

- Weighted fused:

$$\hat{\mathbf{Y}} = \mathbf{W}_h \odot \hat{\mathbf{Y}}_h + \mathbf{W}_d \odot \hat{\mathbf{Y}}_d + \mathbf{W}_w \odot \hat{\mathbf{Y}}_w$$



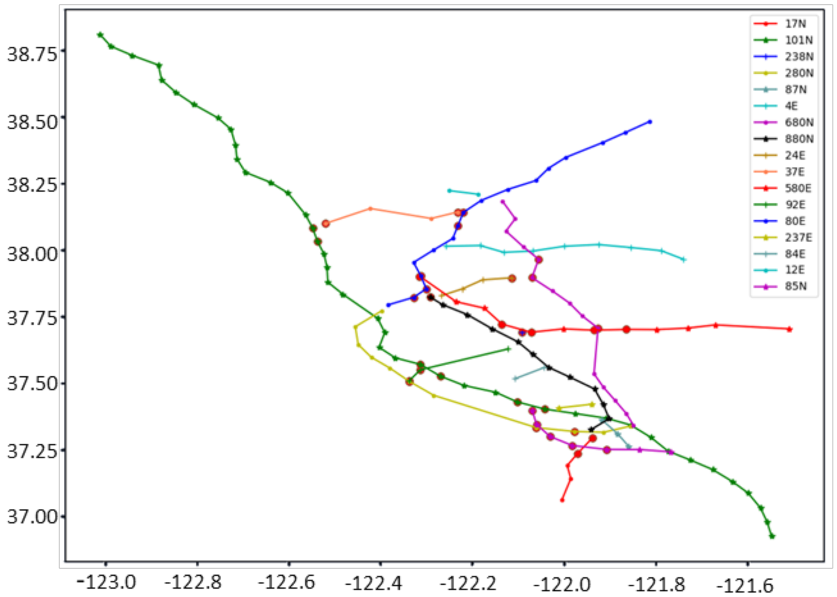


Experiments

Datasets

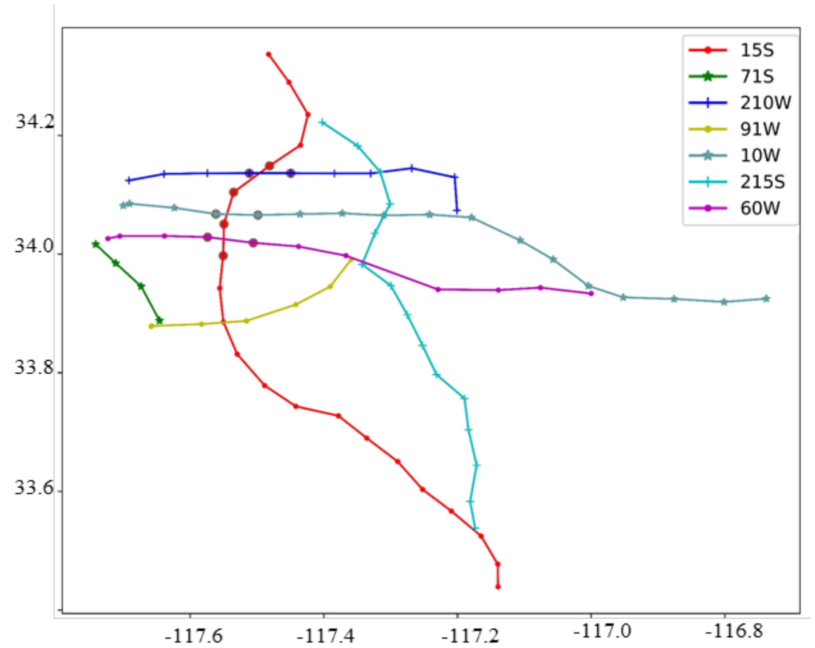
- **Three measurements (every 5min):** total flow, average speed, average occupancy
- **Goal:** predicting the traffic flow over one hour in the future.

- PeMSD4



- 307 detectors on 17 roads
- 2018.01 – 2018.02

- PeMSD8



- 170 detectors on 7 roads
- 2016.07 – 2016.08

Evaluation Metrics & Baselines

Evaluation Metrics

$$MAE = \frac{1}{n}\sum_{i}^{n}\left|x_i - \hat{x}_i\right| \qquad RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}\left(x_i - \hat{x}_i\right)^2}$$

Baselines

- **HA**： Historical Average method
- **ARIMA**： Auto-Regressive Integrated Moving Average method
- **VAR**： Vector Auto-Regressive
- **LSTM** [6]： Long Short-Term Memory network
- **GRU** [7]： Gated Recurrent Unit network
- **STGCN** [8]： A graph convolution model based on the spatial method
- **GLU-STGCN** [4]： A graph convolution network with a gating mechanism
- **GeoMAN** [3]： A multi-level attention-based recurrent neural network model
- **MSTGCN**： a degraded version of ASTGCN, without spatial-temporal attention

Result Analysis

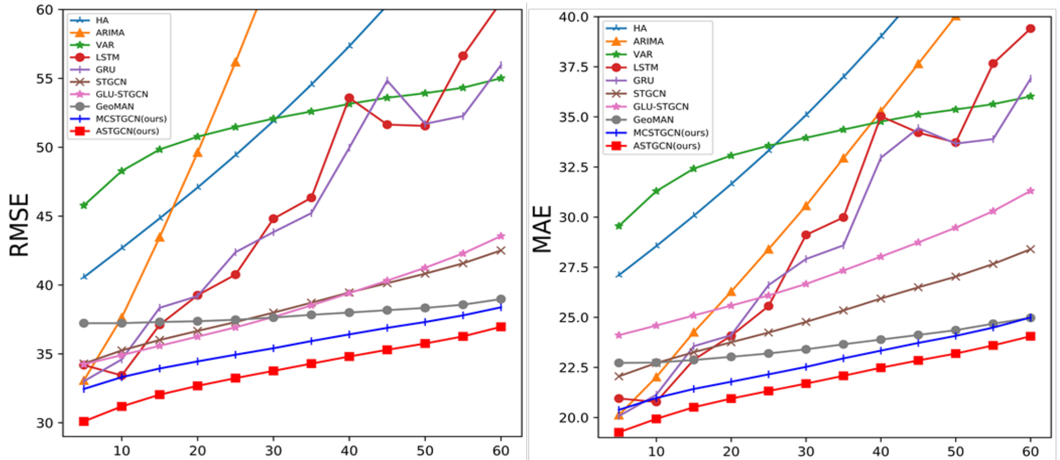
- **Average results of traffic flow prediction performance over the next one hour**

Model	PeMSD4		PeMSD8	
	RMSE	MAE	RMSE	MAE
HA	54.14	36.76	44.03	29.52
ARIMA	68.13	32.11	43.30	24.04
VAR	51.73	33.76	31.21	21.41
LSTM	45.82	29.45	36.96	23.18
GRU	45.11	28.65	35.95	22.20
STGCN	38.29	25.15	27.87	18.88
GLU-STGCN	38.41	27.28	30.78	20.99
GeoMAN	37.84	23.64	28.91	17.84
MSTGCN (ours)	**35.64**	**22.73**	**26.47**	**17.47**
ASTGCN (ours)	**32.82**	**21.80**	**25.27**	**16.63**

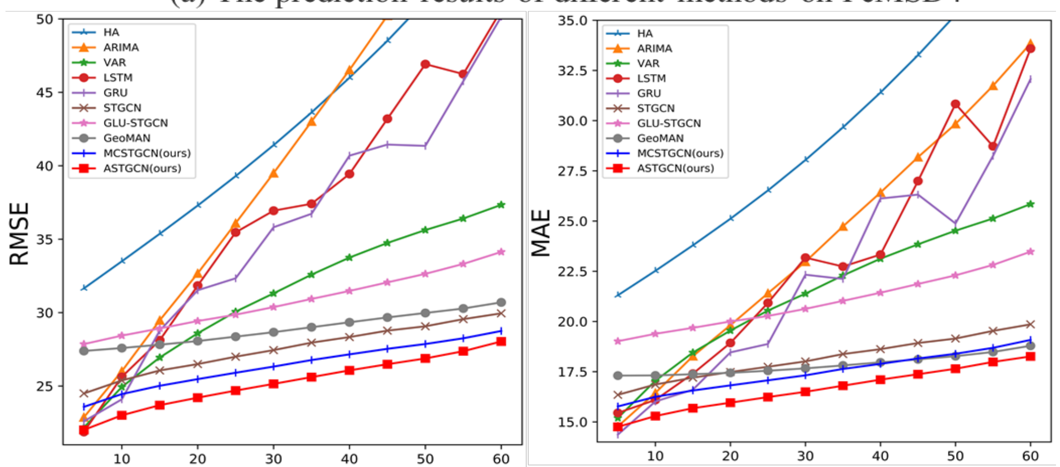
- PeMSD4
  - RMSE ↓13.27%
  - MAE ↓7.78%
- PeMSD8
  - RMSE ↓9.33%
  - MAE ↓6.78%

Result Analysis

■ Changes of prediction performance as the prediction interval increases



(a) The prediction results of different methods on PeMSD4

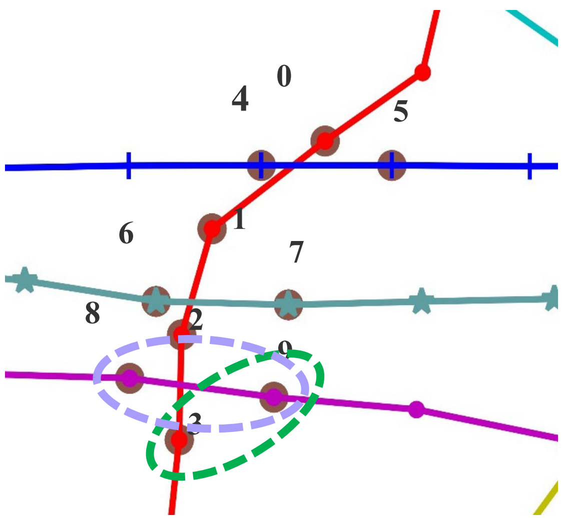


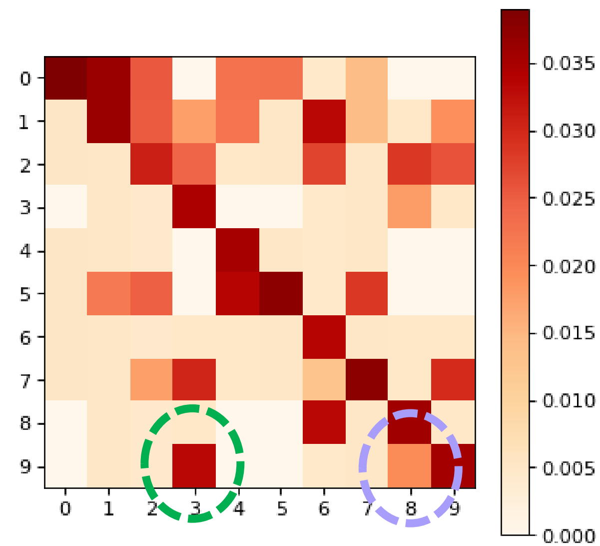
(b) The prediction results of different methods on PeMSD8

Result Analysis

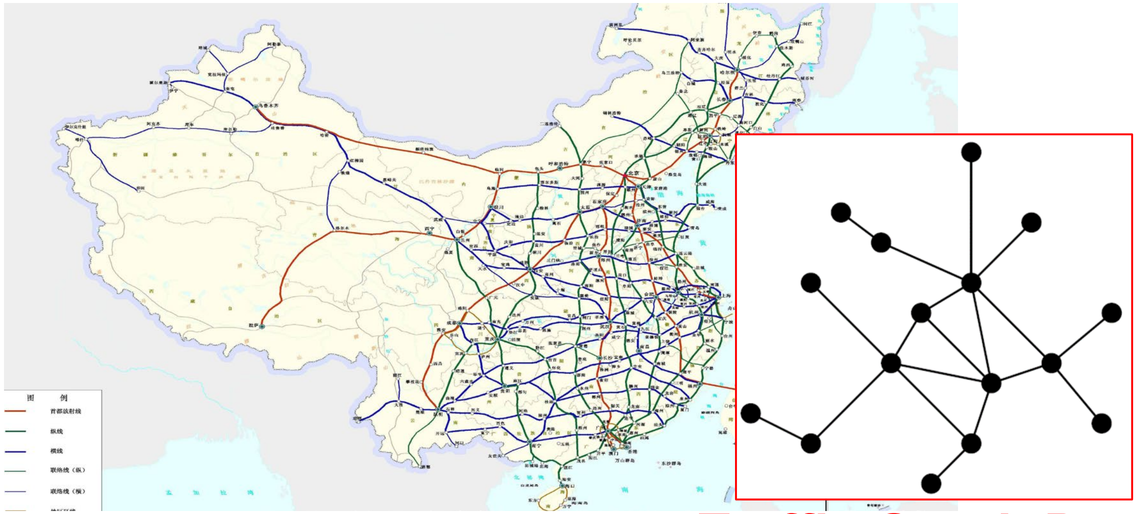
Case study

- A sub-graph with 10 detectors from PeMSD8
- The average spatial attention matrix among detectors in the training set

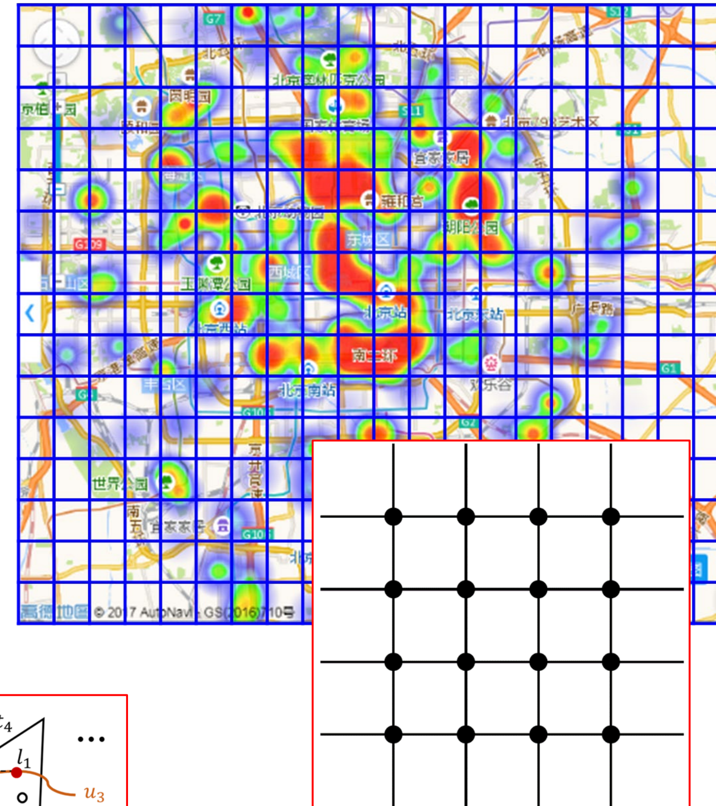




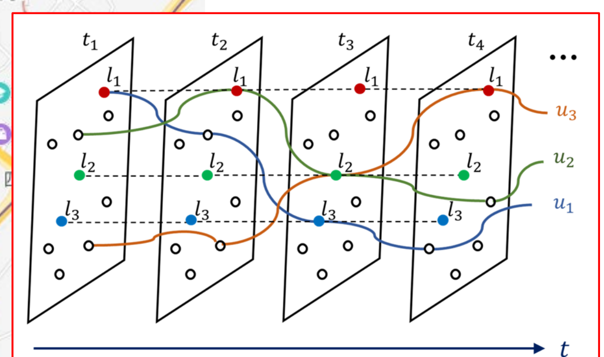
Related work



Traffic Graph Data



Traffic Raster Data



Trajectory Data

http://insis.bjtu.edu.cn/

Reference

[1] Zhang, J.; Zheng, Y.; Qi, D.; Li, R.; Yi, X.; and Li, T. 2018. Predicting citywide crowd flows using deep spatio-temporal residual networks. Artificial Intelligence 259:147-166.

[2] Yao, H.; Wu, F.; Ke, J.; Tang, X.; Jia, Y.; Lu, S.; Gong, P.; and Ye, J. 2018b. Deep multi-view spatial-temporal network for taxi demand prediction. In AAAI Conference on Artificial Intelligence, 2588-2595.

[3] Liang, Y.; Ke, S.; Zhang, J.; Yi, X.; and Zheng, Y. 2018. GeoMAN: Multi-level Attention Networks for Geo-sensory Time Series Prediction. In International Joint Conference on Artificial Intelligence, 3428-3434.

[4] Yu, B.; Yin, H.; and Zhu, Z. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In International Joint Conference on Artificial Intelligence, 3634-3640.

[5] Defferrard, M.; Bresson, X.; and Vandergheynst, P. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In Advances in Neural Information Processing Systems, 3844–3852.

[6] Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. Neural Computation 9(8):1735–1780.

[7] Chung, J.; Gulcehre, C.; Cho, K.; and Bengio, Y. 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. In NIPS 2014 Workshop on Deep Learning.

[8] Li, C.; Cui, Z.; Zheng, W.; Xu, C.; and Yang, J. 2018. Spatio-Temporal Graph Convolution for Skeleton Based Action Recognition. In AAAI Conference on Artificial Intelligence, 3482–3489.

Beijing Jiaotong University
Institute of Network Science and Intelligent System

Thanks!