Spatiotemporal Adaptive Gated Graph Convolution Network for Urban Traffic Flow Forecasting

Bin Lu[1], Xiaoying Gan[1], Haiming Jin[1], Luoyi Fu[1], Haisong Zhang[2]

[1]Shanghai Jiao Tong University, [2]Tencent AI Lab

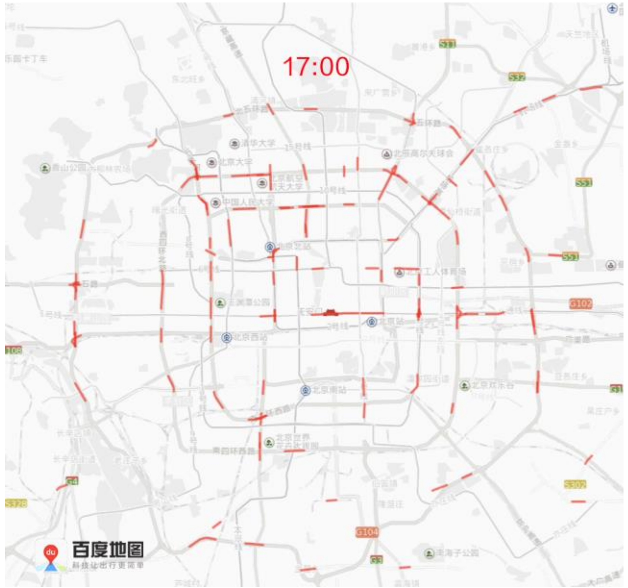
{robinlu1209, ganxiaoying, jinhaiming, yiluofu}@sjtu.edu.cn,

hansonzhang@tencent.com

Introduction

- Accurate and timely traffic flow forecasting plays a vital role in SMART CITY.

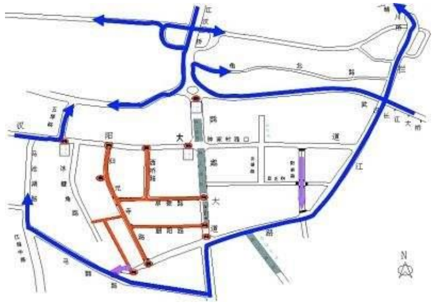
Road condition analysis and travel planning



Urban Traffic Flow Forecasting

Intelligent Transportation System(ITS)

Tidal Lane



Road Planning

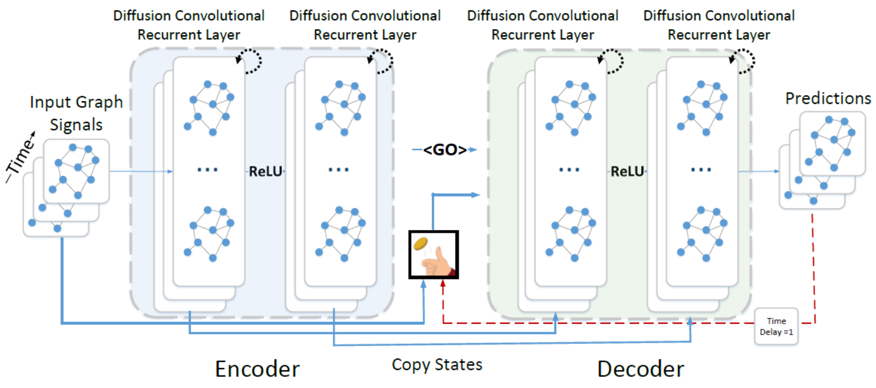
Smart Traffic Lights

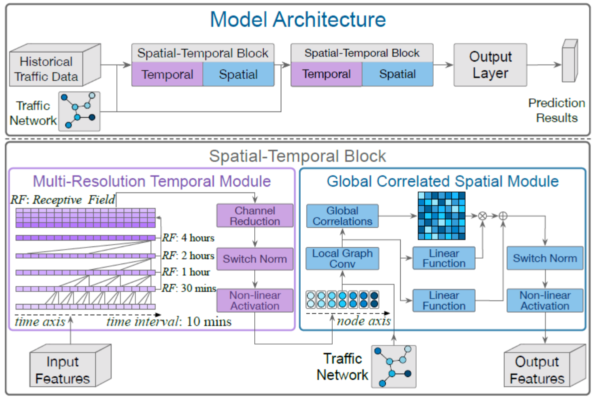
Traffic Management

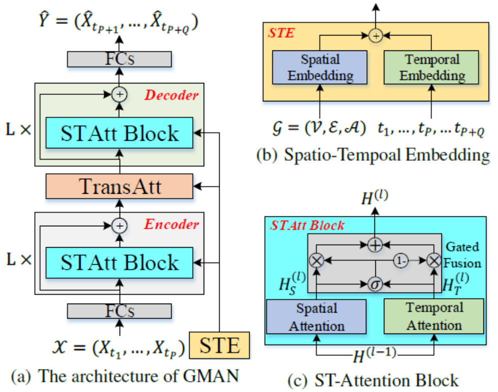
Introduction

■ **Traffic flow forecasting is a widely studied problem in recent years.**

- DCRNN [ICLR-2018]
- GSTNet [IJCAI-2019]
- GMAN [AAAI-2020]







How to further improve the accuracy of the prediction?

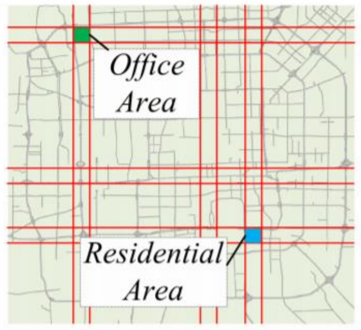
How to capture the dynamic and complex spatiotemporal correlations of urban traffic?

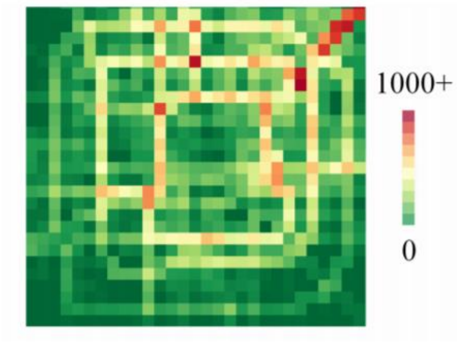
Problem Formulation

■ A dynamic weighted graph is proposed to model the spatial dependency.

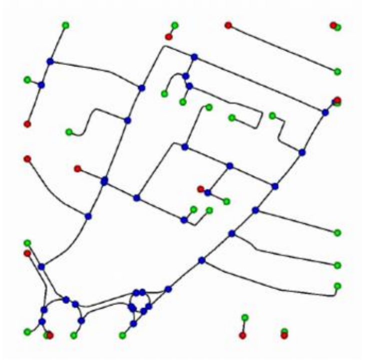
We consider both the local and contextual spatial information, and define the spatial neighbors and semantic neighbors of the road nodes.

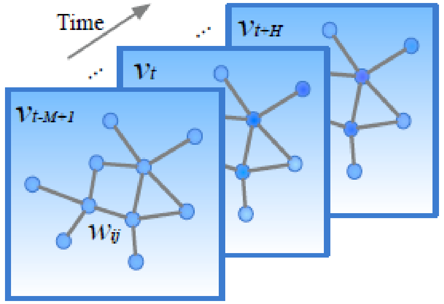
Grid-based map
(only consider local spatial proximity)



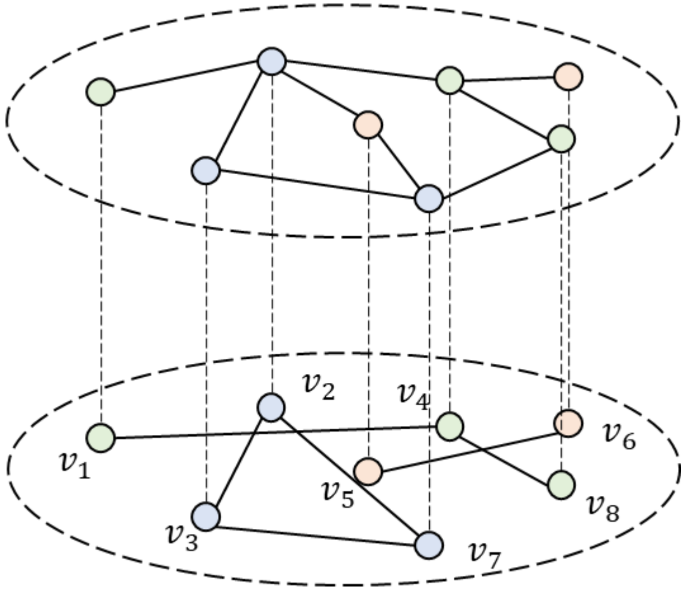


Fixed graph
(only consider local spatial proximity)





Dynamic weighted graph
(consider both spatial and semantic correlations)

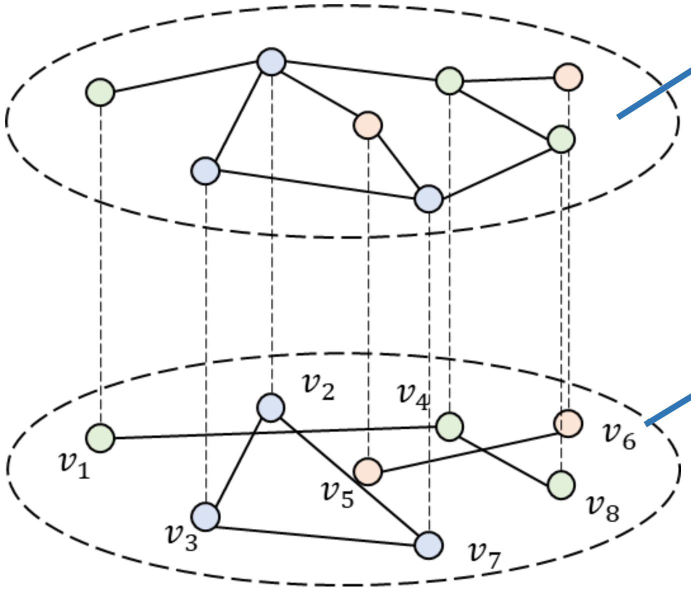


Problem Formulation

■ **A dynamic weighted graph is proposed to model the spatial dependency.**

We consider both the local and contextual spatial information, and define the spatial neighbors and semantic neighbors of the road nodes.

➢ **Dynamic weighted graph** (consider both spatial and semantic correlations)

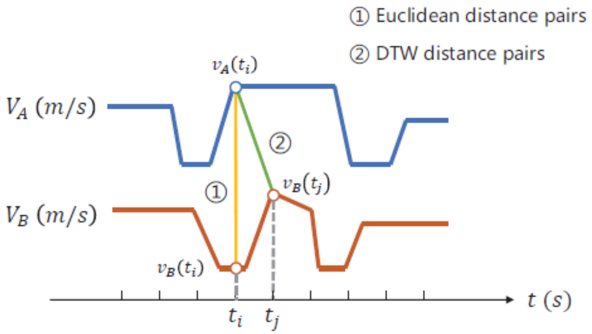


➢ Spatial Neighbor Subgraph Adjacency Matrix A_{ij}^{sp}

$$A_{ij}^{sp} = \begin{cases} 1, & v_i \text{ and } v_j \text{ share the same intersection,} \\ 0, & \text{otherwise.} \end{cases}$$

➢ Semantic Neighbor Subgraph Adjacency Matrix A_{ij}^{se}

$$A_{ij}^{se} = \begin{cases} 1, & \mathrm{DTW}(v_i, v_j) > \epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

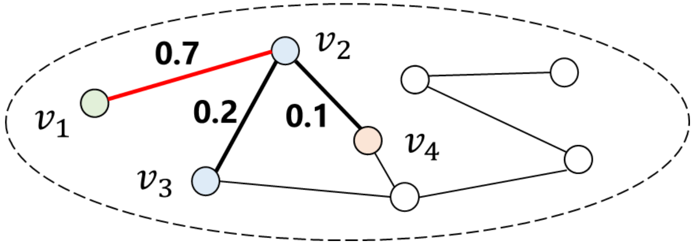


Problem Formulation

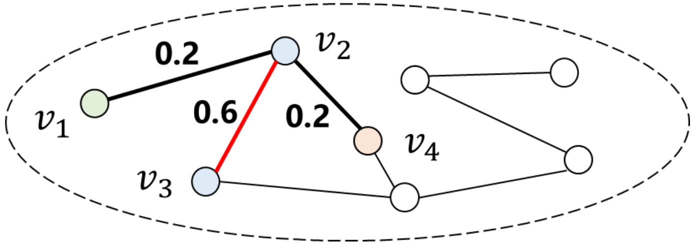
■ A dynamic weighted graph is proposed to model the spatial dependency.

Multi-head graph attention mechanism is utilized to model the dynamic road relationship.

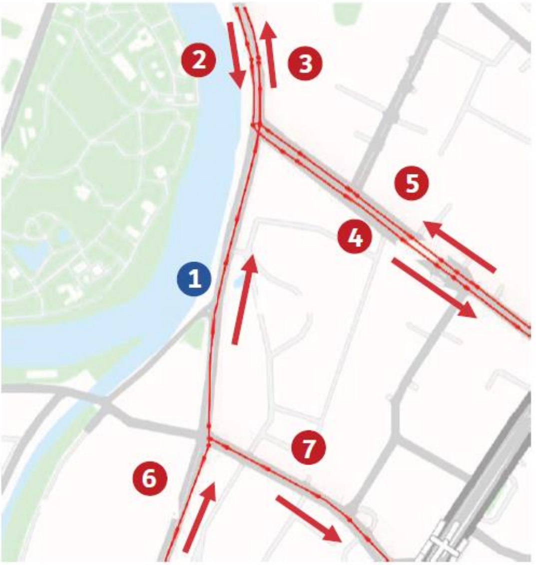
Attention score: $$\alpha_{ij} = \frac{exp\left(LeakyReLU\left(\theta^T\left[Wh_i \parallel Wh_j\right]\right)\right)}{\sum_{k \in N_i} exp\left(LeakyReLU\left(\theta^T\left[Wh_i \parallel Wh_k\right]\right)\right)}$$

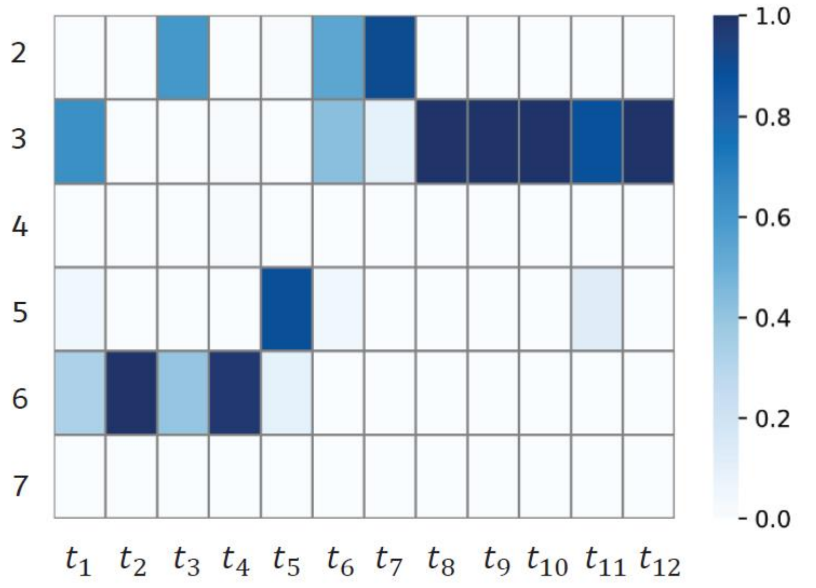


$t = t_1$

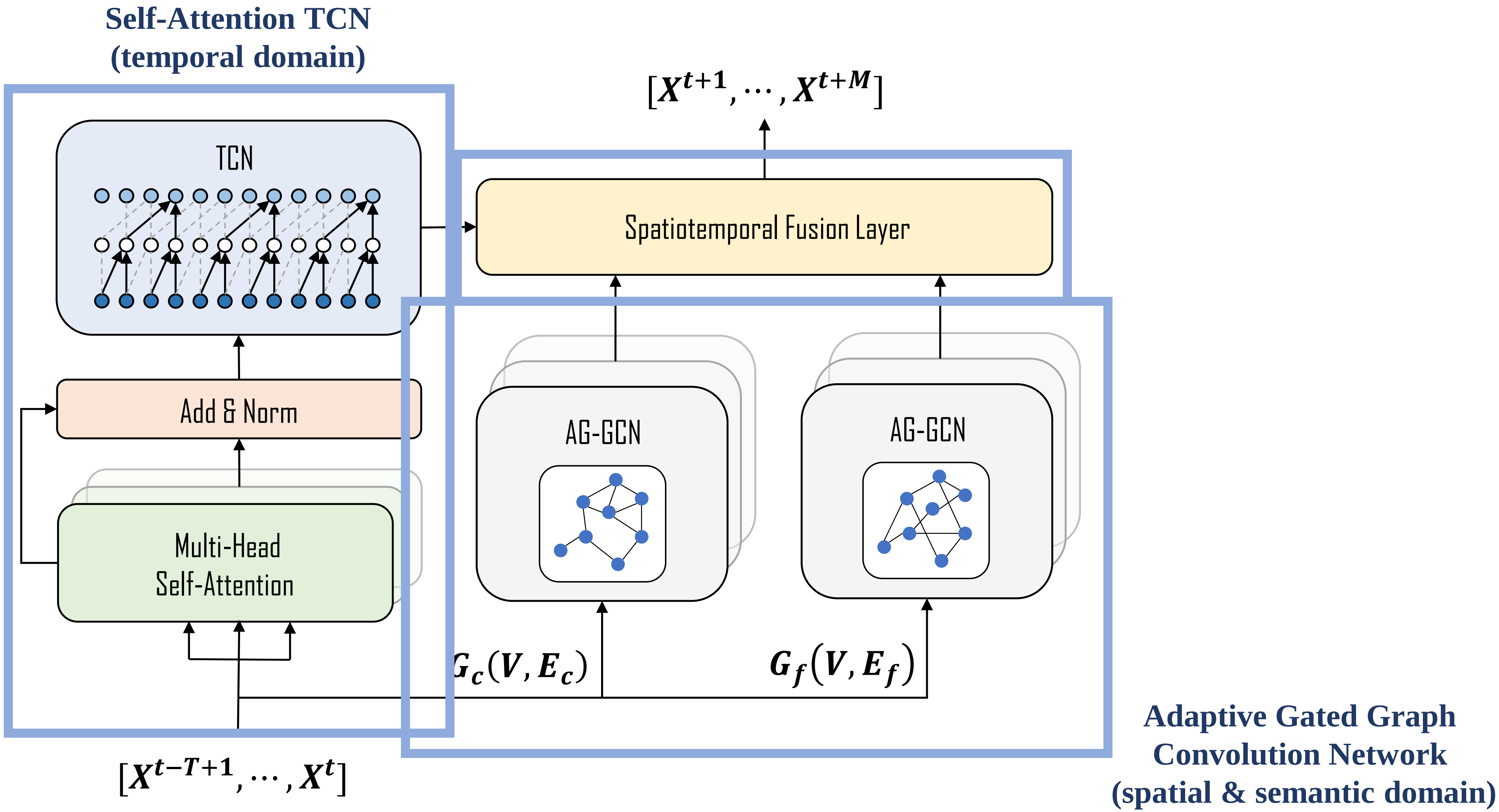


$t = t_2$



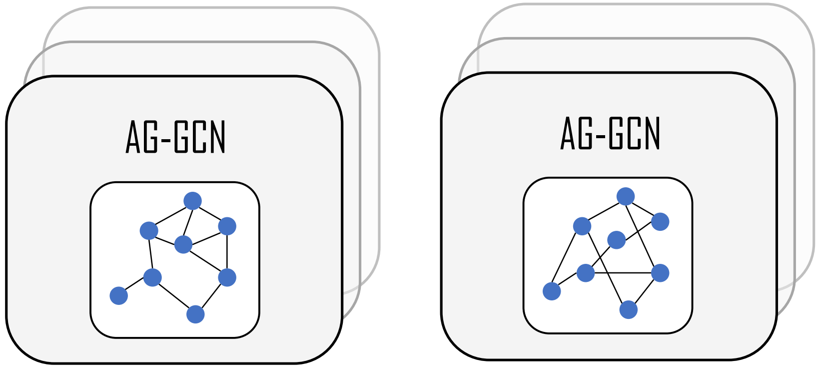


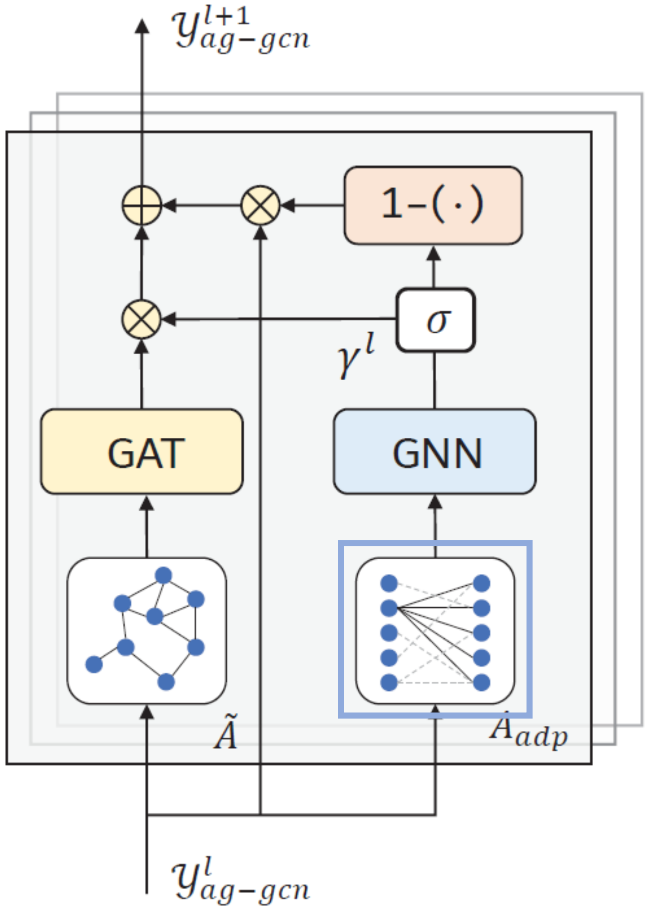
Methodology



SpatioTemporal **A**daptive **G**ated **G**raph **C**onvolution **N**etwork (**STAG-GCN**)

Methodology



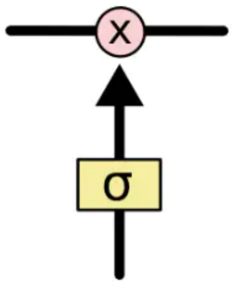


Adaptive Gated Graph Convolution Network (AG-GCN)

■ Multi-head Graph Attention Network.

$$h_i' = \sigma\left(\sum_{j \in N_i} \alpha_{ij} W h_j\right) \xrightarrow{\text{Multi-head}} h_i' = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j \in N_i} \alpha_{ij}^k W^k h_j\right)$$

■ Adaptive Graph Gating Mechanism.



Gating mechanism

- Adaptive adjacency matrix: $\tilde{A}_{adp} = softmax\left(ReLU(E_s E_t^T)\right)$
- Gated Value: $\gamma^l = \sigma\left(\tilde{A}_{adp} X^l \Theta\right)$
- How to update:

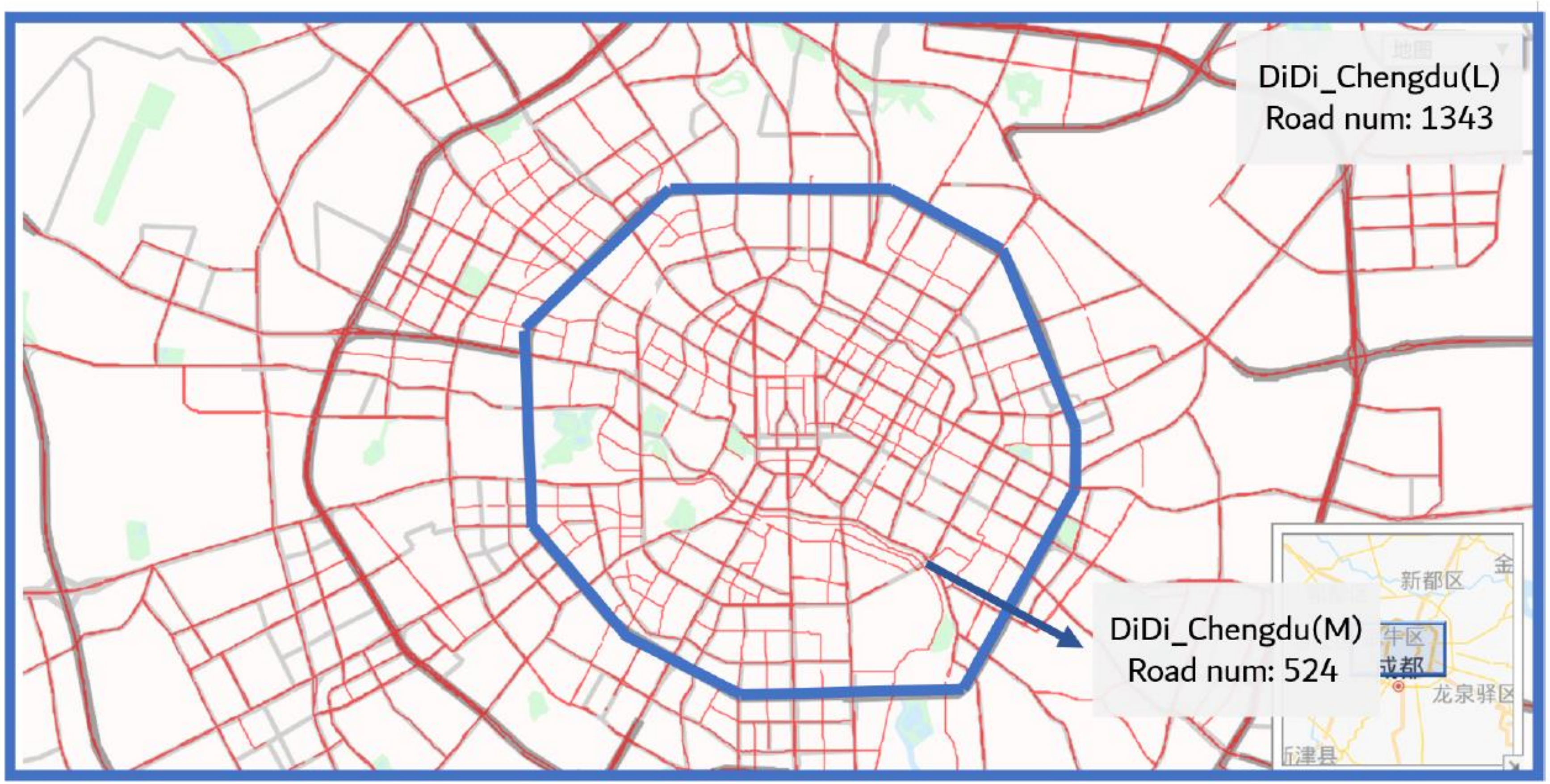
$$X^{l+1} = \gamma^l \odot g\left((\tilde{A} \odot M) X^l W\right) + (1 - \gamma^l) \odot X^l$$

$$\begin{cases} \tilde{A} = A_{sp} + I \\ \tilde{A} = A_{se} + I \end{cases}$$

Dynamic attention coefficients

Experiment

- **Dataset:** urban traffic index dataset of Chengdu, China, provided by Didi Chuxing GAIA Initiative



Experiment

CIKM 2020 19-23 OCTOBER Online

■ Performance Experiment

$$MAE(y,\hat{y}) = \frac{1}{N}\sum_{i=1}^{N} |\hat{y}_i - y_i|. \qquad MAPE(y,\hat{y}) = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right|. \qquad RMSE(y,\hat{y}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}.$$

Table 1: Performance comparison of STAG-GCN and other baseline models on DiDi_Chengdu datasets. STAG-GCN achieves the best performance with all three metrics for all forecasting horizons.

Model	DiDi_Chengdu[M] (10/30/60 min)			DiDi_Chengdu[L] (10/30/60 min)		
	MAE	MAPE(%)	RMSE	MAE	MAPE(%)	RMSE
HA	3.89	13.42	5.61	3.02	13.24	4.56
ARIMA	2.32/ 3.26/ 4.15	9.80/ 14.21/ 18.23	3.45/ 4.90/ 6.22	2.44/ 3.41/ 4.36	9.00/ 12.92/ 16.67	3.73/ 5.19/ 6.65
KNN	2.31/ 2.89/ 3.32	10.39/ 13.50/ 15.69	3.33/ 4.21/ 4.83	2.42/ 3.01/ 3.45	9.40/ 12.07/ 14.08	3.56/ 4.45/ 5.14
RF	2.29/ 2.94/ 3.37	9.66/ 13.11/ 15.31	3.14/ 4.08/ 4.69	2.40/ 3.05/ 3.52	9.18/ 12.188/ 14.30	3.53/ 4.50/ 5.20
FNN	2.42/ 2.91/ 3.43	10.63/ 12.59/ 14.72	3.35/ 3.97/ 4.65	2.77/ 2.97/ 3.30	11.43/ 12.25/ 13.66	4.12/ 4.35/ 4.79
LSTM-FC	2.37/ 2.52/ 2.91	11.30/ 11.96/ 13.66	3.45/ 3.68/ 4.23	2.60/ 2.72/ 3.09	10.74/ 11.29/ 12.81	3.96/ 4.13/ 4.63
STGCN	2.22/ 2.67/ 3.05	9.94/ 12.67/ 14.65	3.18/ 3.90/ 4.46	2.50/ 2.81/ 3.14	10.07/ 11.56/ 13.02	3.61/ 4.12/ 4.61
DCRNN	2.04/ 2.65/ 3.16	9.00/ 12.34/ 14.83	2.99/ 3.92/ 4.61	2.27/ 2.61/ 2.78	9.22/ 10.97/ 11.78	3.49/ 4.04/ 4.32
ASTGCN	2.05/ 2.44/ 2.70	9.17/ 11.48/ 12.71	2.99/ 3.58/ 3.91	2.20/ 2.66/ 2.95	8.60/ 10.74/ 11.98	3.28/ 3.96/ 4.38
STAG-GCN	**1.98/ 2.36/ 2.54**	**8.84/ 11.05/ 11.90**	**2.89/ 3.46/ 3.69**	**2.08/ 2.52/ 2.79**	**8.07/ 10.04/ 11.02**	**3.12/ 3.75/ 4.11**

Conclusion

Contributions

- We consider both the **local and contextual spatial information**, and define the spatial neighbors and semantic neighbors of the road nodes. Multi-head graph attention mechanism is utilized to model the road relationship as a **dynamic weighted graph**.
- We propose a novel **adaptive graph gating mechanism** to selectively update and forget the high-order neighbor information of nodes within the multi-layer stacking. GNN based on **adaptive adjacency matrix** can identify deviations caused by artificially defined spatial relationships and characterize global spatial correlations.
- We conduct extensive experiments on real-world urban traffic datasets by using our proposed model, STAG-GCN, and outperform the performance compared to existing baselines. We also carry out rich experiments on the model itself, and explain the performance and design rationale in detail.

Future work

- How to capture the influence caused by external factors, like social events, extreme weather, traffic regulations, etc.
- The proposed AG-GCN module can be generalized into dynamical graph features learning in various applications.

Spatiotemporal Adaptive Gated Graph Convolution Network for Urban Traffic Flow Forecasting

Bin Lu[1], Xiaoying Gan[1], Haiming Jin[1], Luoyi Fu[1], Haisong Zhang[2]

[1]Shanghai Jiao Tong University, [2]Tencent AI Lab

{robinlu1209, ganxiaoying, jinhaiming, yiluofu}@sjtu.edu.cn,

hansonzhang@tencent.com