Constructing Geographic and Long-term Temporal Graph for Traffic Forecasting

Yiwen Sun[1], Yulu Wang[1], Kun Fu[2], Zheng Wang[2], Changshui Zhang[1], Jieping Ye[2]

[1]Institute for Artificial Intelligence, Tsinghua University (THUAI),
State Key Lab of Intelligent Technologies and Systems,
Beijing National Research Center for Information Science and Technology (BNRist),
Department of Automation, Tsinghua University, Beijing, P.R.China

[2]DiDi AI Labs, Beijing, P.R.China

Background of Traffic Forecasting & GLT

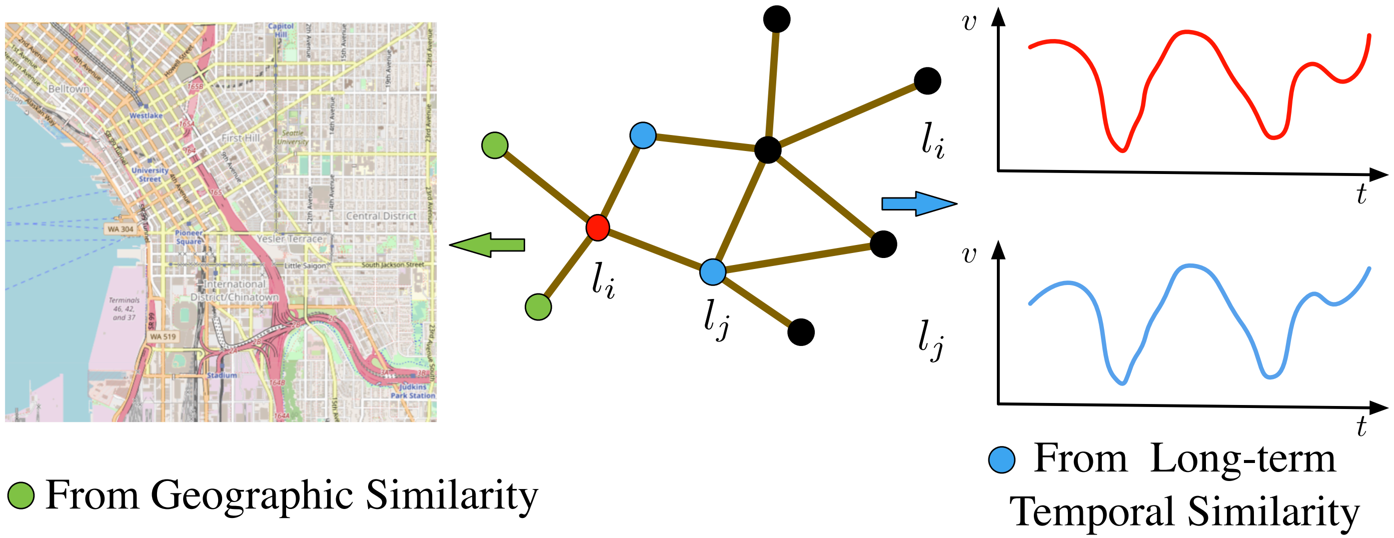

Fig.1. The conceptual demonstration of the construction process of Geographic and Long-term Temporal (GLT) Graph.

- Traffic forecasting: predicting future traffic states of road segments given sequential historical traffic states and the road network.
- Recent SOTA: GCN → spatial; RNN →temporal.
- The graph only based on the geographic information.
- Geographic and Long-term Temporal (GLT) Graph.
- GLT-GCRNN: the first deep learning framework which considers both the geographic and long-term temporal information.

Constructing the GLT Graph

- In the geographic aspect,

$$S_{G}{}_{i,j}^{k} = \min\left((A+I)_{i,j}^{k}, 1\right)$$

- In the long-term temporal aspect, the difference matrix Q.

$$Q_{ij} = Q_{ji} = \|\hat{\boldsymbol{v}}(i) - \hat{\boldsymbol{v}}(j)\|_2$$

- The similar matrix in the long-term temporal aspect.

$$S_{LT}{}_{i,j} = \begin{cases} 1, & Q_{i,j} \in Q_{i,:} \text{ top } \gamma \text{ small elements} \\ 0, & \text{otherwise} \end{cases}$$

- The k-hop geographic and long-term temporal similar matrix:

$$\boldsymbol{S_{GLT}}^{k} = \boldsymbol{S_{G}}^{k} + \boldsymbol{S_{LT}}$$

Constructing the GLT Graph

- Two representative groups of long-term temporal similar links,

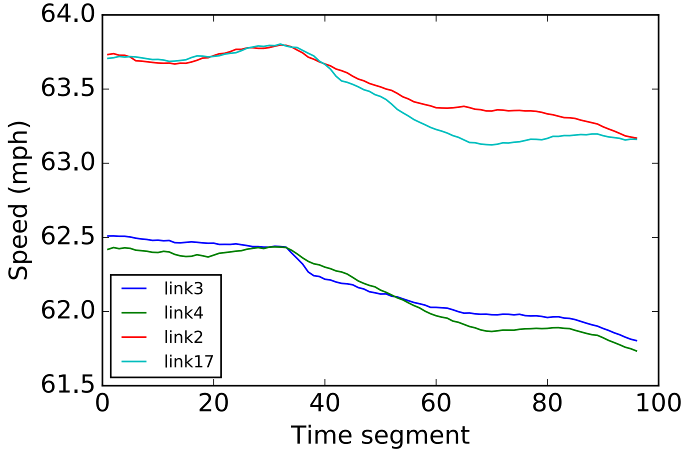

Fig.2. Two representative groups of long term temporal similar links.

Link 2 is the most LT similar link of link 17, nevertheless, they are not geographic similar which is ignored by S_G.

S_{LT} → let them interact.

- The free-flow reachable matrix is also adopted:

$$S_{F\,i,j} = \begin{cases} 1, & V_{i,j} m \Delta t \geq D_{i,j} \\ 0, & \text{otherwise} \end{cases}$$

- The k-hop ultimate similar matrix :

$$\boldsymbol{S_U}^k = \boldsymbol{S_{GLT}}^k \odot \boldsymbol{S_F}$$

GLT Graph Convolution and Modified LSTM

- GLT graph convolution operation

$$\boldsymbol{g}_t^k = \left(\boldsymbol{W_g}^k \odot \boldsymbol{S_U}^k\right) \boldsymbol{x}_t$$

- g_t^k extracted according to various hops are concatenated:

$$\boldsymbol{G}_t^K = \left[\boldsymbol{g}_t^1, \boldsymbol{g}_t^2, \ldots, \boldsymbol{g}_t^K\right]$$

- The cell state gate is also modified:

$$C_{t-1}^* = \boldsymbol{W_C} \odot \boldsymbol{S_U}^K \cdot C_{t-1}$$

Overall Framework of GLT-GCRNN

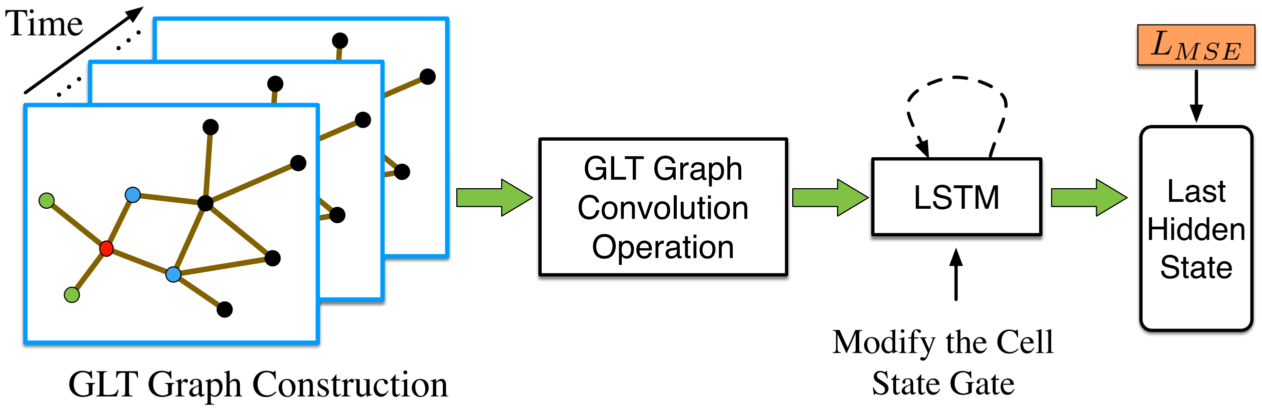

Fig.3. The overall framework of GLT-GCRNN.

(1) The GLT graph construction.

(2) GLT graph convolution operation for mining the spatial correlation.

(3) Modified LSTM to capture the time series dependency.

The objective function:

$$L_{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2$$

Dataset & Evaluation Metrics

TABLE I
STATISTICS OF THE REAL-WORLD TRAFFIC FORECASTING DATASET

location	time span	# sensor	interval
Greater Seattle Area	1 year (2015)	323	5-minute

Rooted Mean Square Error:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2}$$

Mean Absolute Percentage Error:

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - y_i'|}{y_i}$$

Mean Absolute Error:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - y_i'|$$

Experimental results

TABLE II
TRAFFIC FORECASTING RESULT COMPARISON ON THE NETWORK-SCALE DATASET

	RMSE (mph)	MAPE (%)	MAE (mph)
ARIMA	10.65	13.85	6.10
SVR	11.12	14.39	6.85
FNN	7.83	10.19	4.45
LSTM	4.97	6.83	2.70
DiffGRU	8.22	11.18	4.64
Conv+LSTM	5.02	6.79	2.71
SGC+LSTM	4.80	6.52	2.64
LSGC+LSTM	6.18	7.51	3.16
TGC-LSTM	4.63	6.01	2.57
GLT-GCRNN (ours)	**3.59**	**5.90**	**2.45**

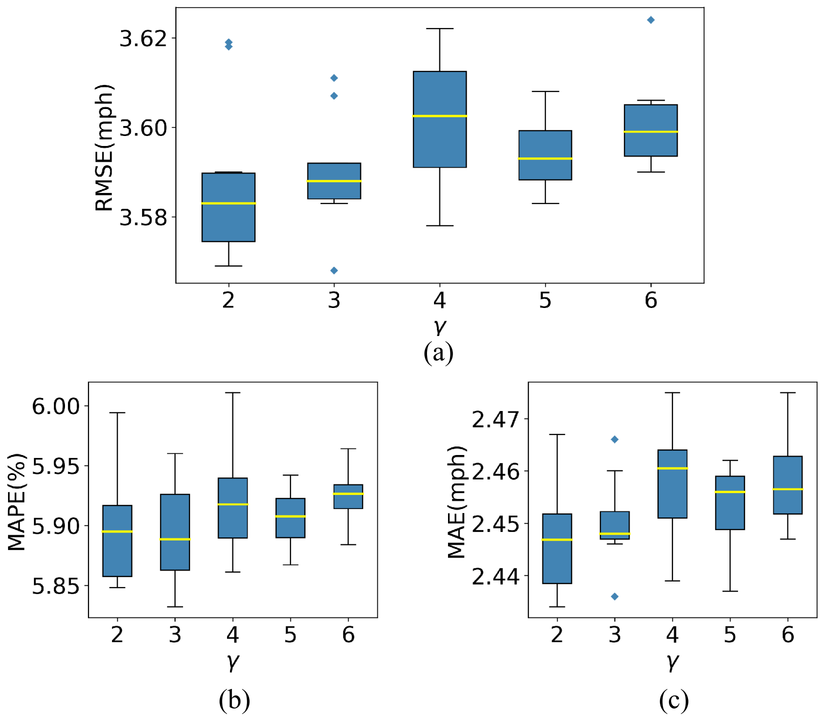

Fig.4. The influence of hyper-parameter γ: (a) Regarding to RMSE. (b) Regarding to MAPE. (c) Regarding to MAE.

Experimental results

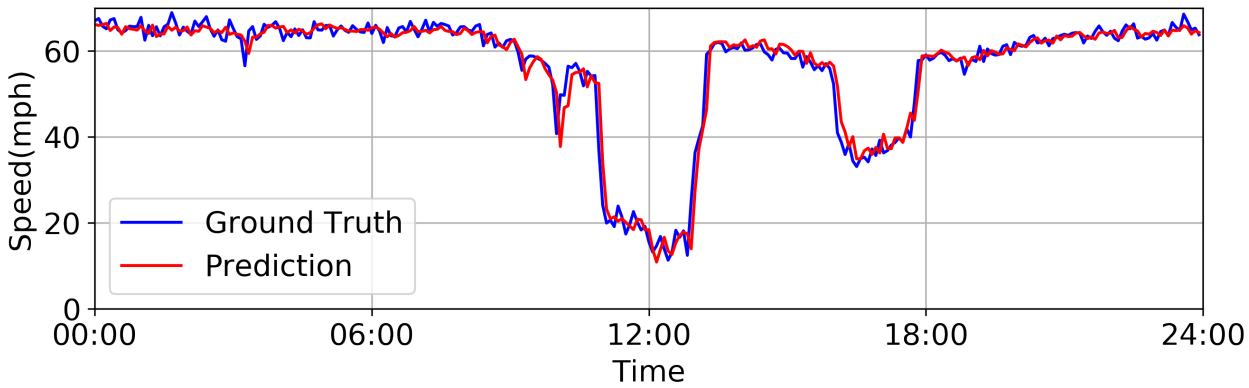

Fig.5. The traffic forecasting visualization of GLT-GCRNN. The link's ID is 190 and the date is 2015-12-27.

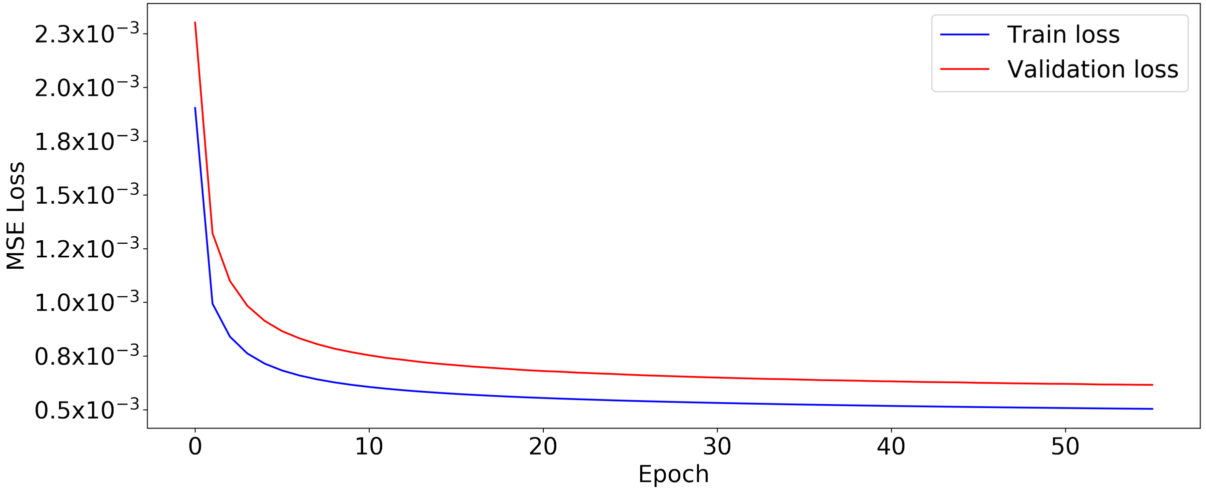

Fig.6. The training and validation process of GLT-GCRNN.

Thank you for listening.