Spatio-Temporal Self-Supervised Learning for Traffic Flow Prediction

Jiahao Ji[1], Jingyuan Wang[1,2,3*], Chao Huang[4], Junjie Wu[3], Boren Xu[1], Zhenhe Wu[1], Junbo Zhang[5,6], Yu Zheng[5,6]

[1]School of Computer Science & Engineering, Beihang University, China
[2]Peng Cheng Laboratory, China [3]School of Economics & Management, Beihang University, China
[4]Department of Computer Science, Musketeers Foundation Institute of Data Science, University of Hong Kong, China
[5]JD Intelligent Cities Research, Beijing, China [6]JD iCity, JD Technology, Beijing, China

Robust Traffic Flow Prediction

- Importance: Crucial for advancing Intelligent Transportation System (ITS)

Mitigate tragedies caused by sudden flow spike

Enable effective traffic controls in time

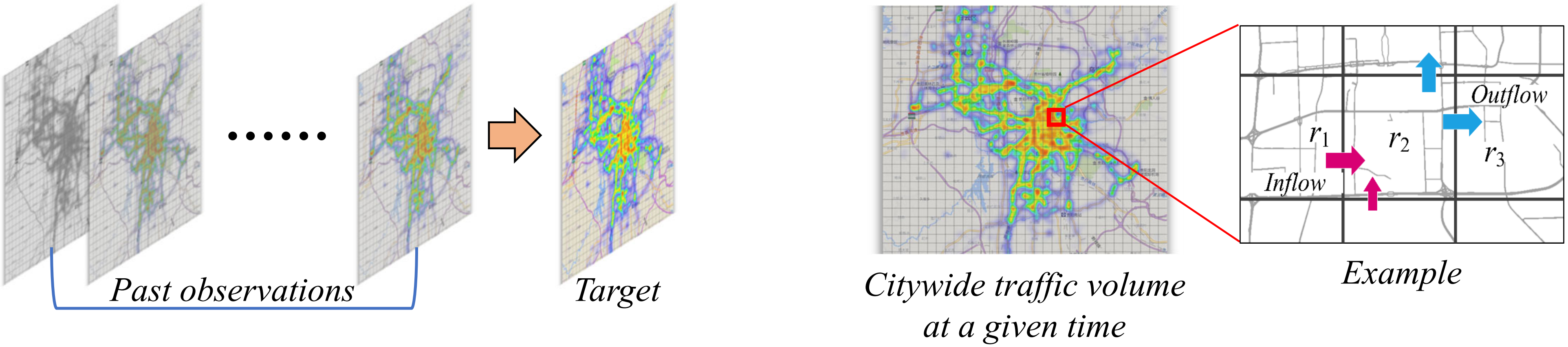
Robust Traffic Flow Prediction

- Importance: Crucial for advancing Intelligent Transportation System

Mitigate tragedies caused by sudden flow spike

Enable effective traffic controls in time

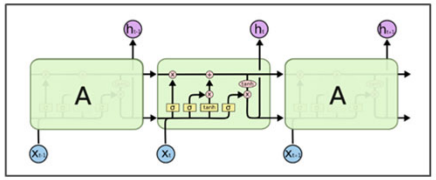
- Traffic flow prediction
  - Forecasting the future traffic volume from past traffic observations



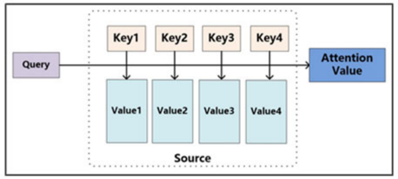
Challenges

- Existing methods focus on modeling spatio-temporal (ST) correlations

- Temporal modeling (closeness, period, trend)

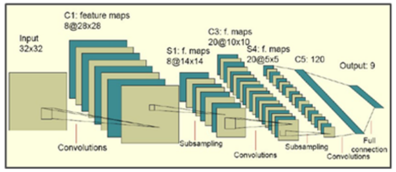


RNN

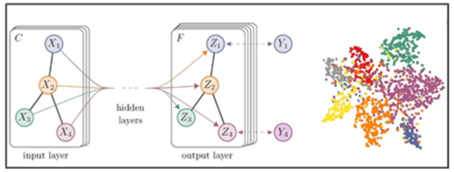


Attention Mechanism

- Spatial modeling



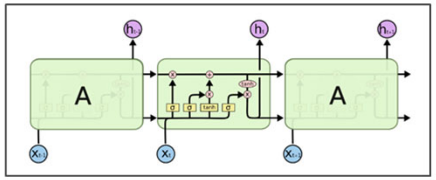
CNN



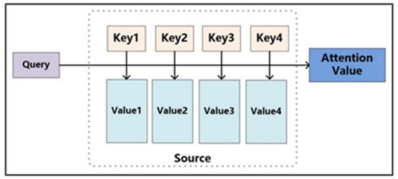
GNN

Challenges

- Existing methods focus on modeling spatio-temporal (ST) correlations
  - Temporal modeling (closeness, period, trend)

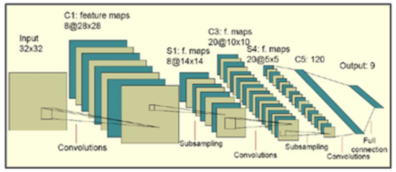


RNN

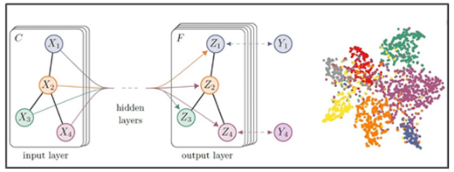


Attention Mechanism

  - Spatial modeling



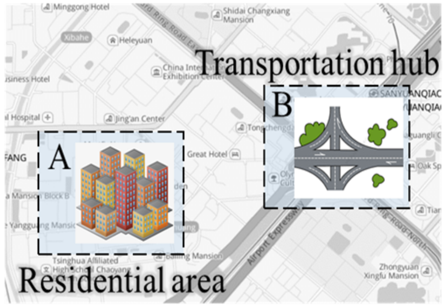
CNN



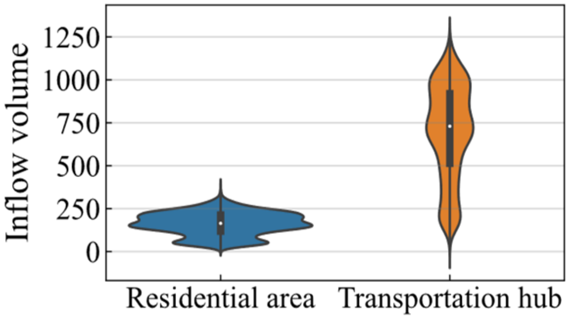
GNN

- Two main limitations:

Spatial heterogeneity



(a) Regions with different functions

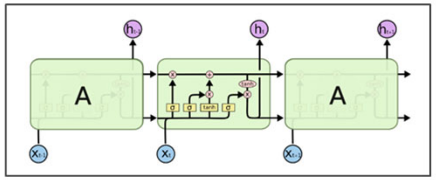


(b) Spatial heterogeneity

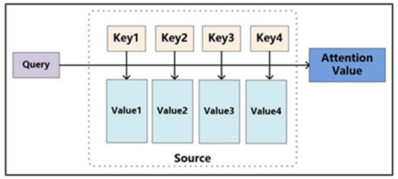
Ignorance of spatial heterogeneity

Challenges

- Existing methods focus on modeling spatio-temporal (ST) correlations
  - Temporal modeling (closeness, period, trend)

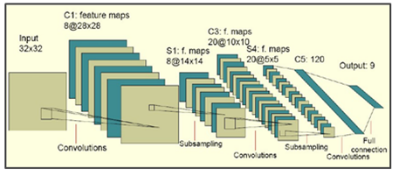


RNN

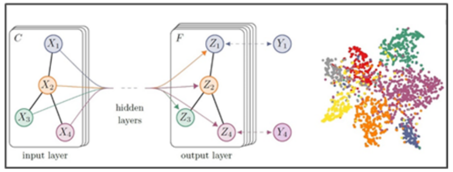


Attention Mechanism

  - Spatial modeling



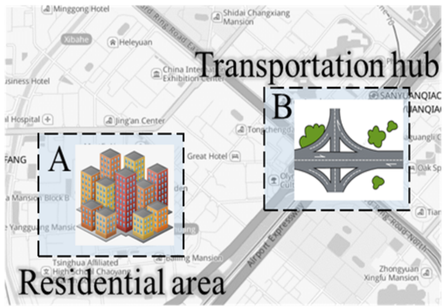
CNN



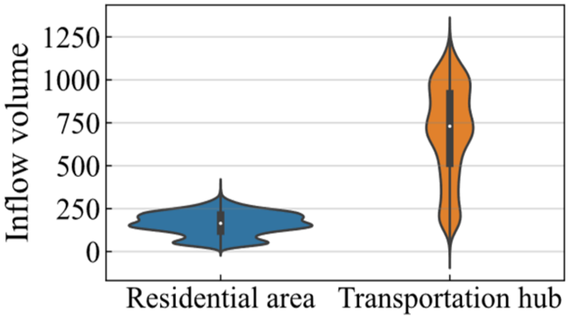
GNN

- Two main limitations

Spatial heterogeneity



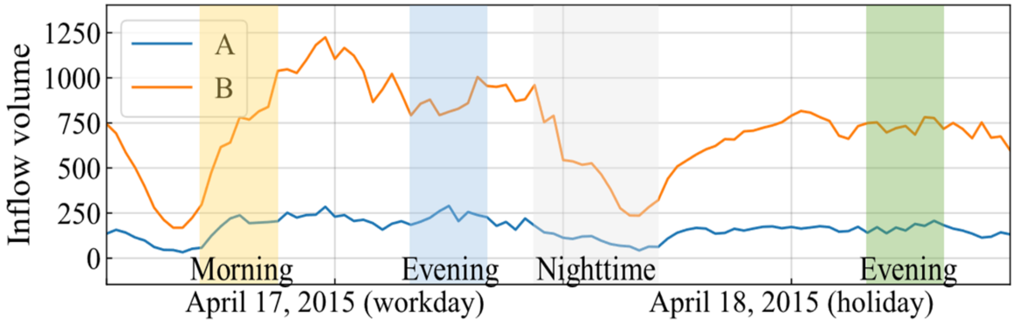
(a) Regions with different functions



(b) Spatial heterogeneity

Ignorance of spatial heterogeneity

Temporal heterogeneity

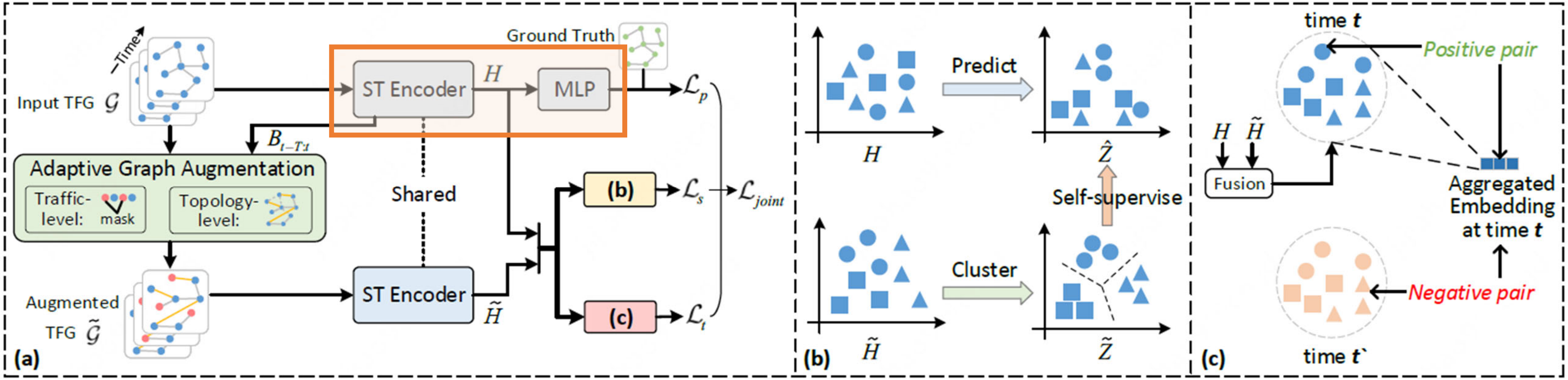


(c) Temporal heterogeneity

Using a shared parameter space for all time periods

Spatio-Temporal Self-Supervised Learning

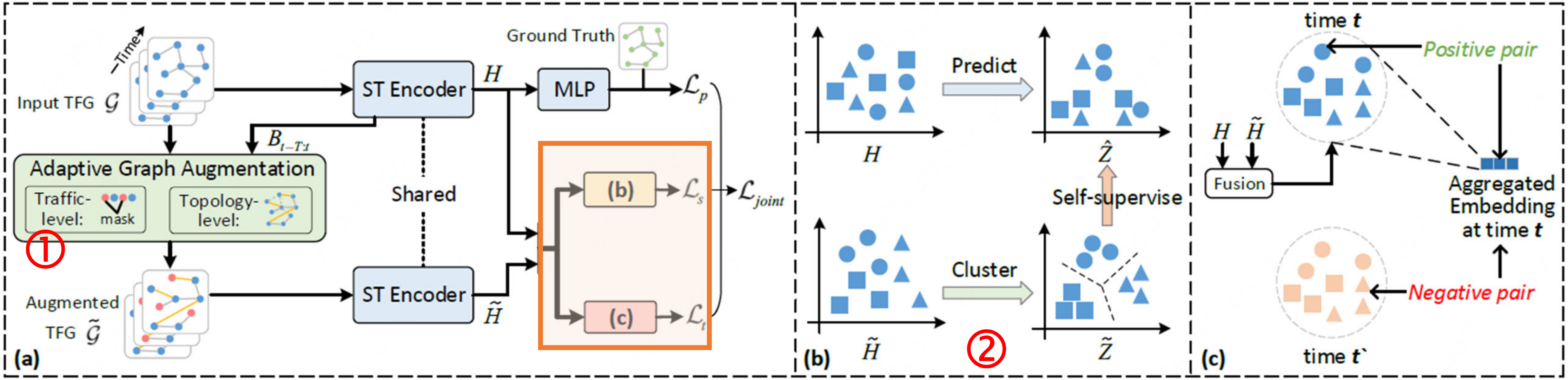
Spatio-Temporal Self-Supervised Learning (**ST-SSL**) for robust traffic prediction



- ST Encoder: encoding spatial-temporal traffic patterns into embeddings ***H***

Spatio-Temporal Self-Supervised Learning

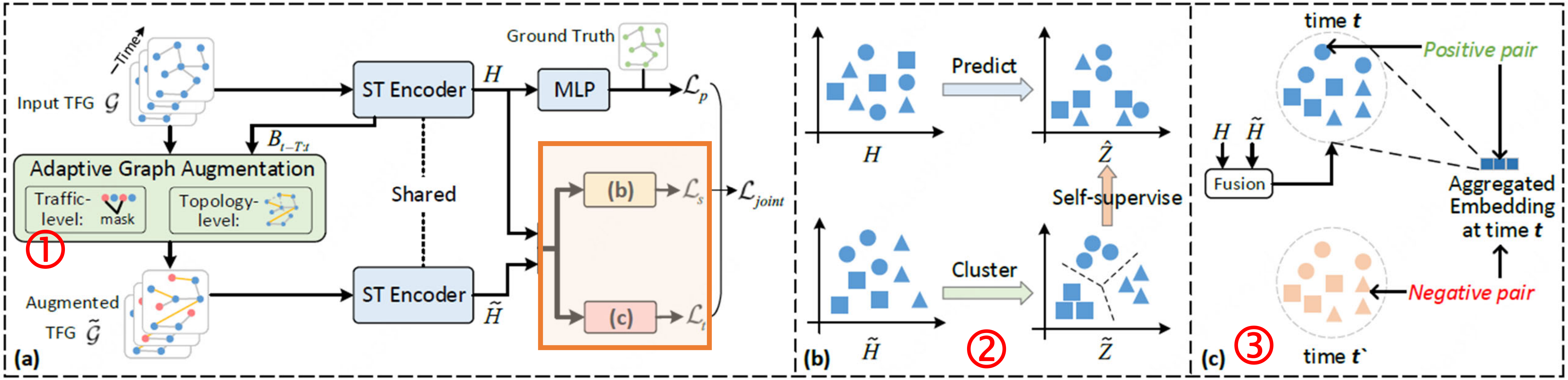
Spatio-Temporal Self-Supervised Learning (**ST-SSL**) for robust traffic prediction



- ST Encoder: encoding spatial-temporal traffic patterns into embeddings ***H***
- SSL for Spatial heterogeneity modeling (b):
  - Adaptive graph augmentation on traffic flow graph ①
  - Soft clustering-based *predictive* SSL task ②
-

Spatio-Temporal Self-Supervised Learning

Spatio-Temporal Self-Supervised Learning (**ST-SSL**) for robust traffic prediction



- ST Encoder: encoding spatial-temporal traffic patterns into embeddings ***H***
- SSL for Spatial heterogeneity modeling (b):
  - Adaptive graph augmentation on traffic flow graph ①
  - Soft clustering-based *predictive* SSL task ②
- SSL for Temporal heterogeneity modeling (c): time-aware *contrastive* SSL task ③

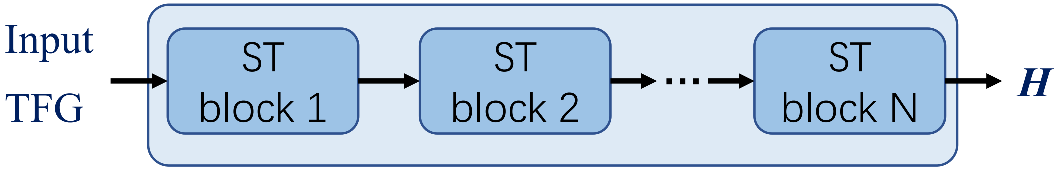
Spatio-Temporal Encoder

- Goal: encoding spatial-temporal traffic patterns into the embedding $\boldsymbol{H}$
  - It can be any spatio-temporal prediction model

Spatio-Temporal Encoder

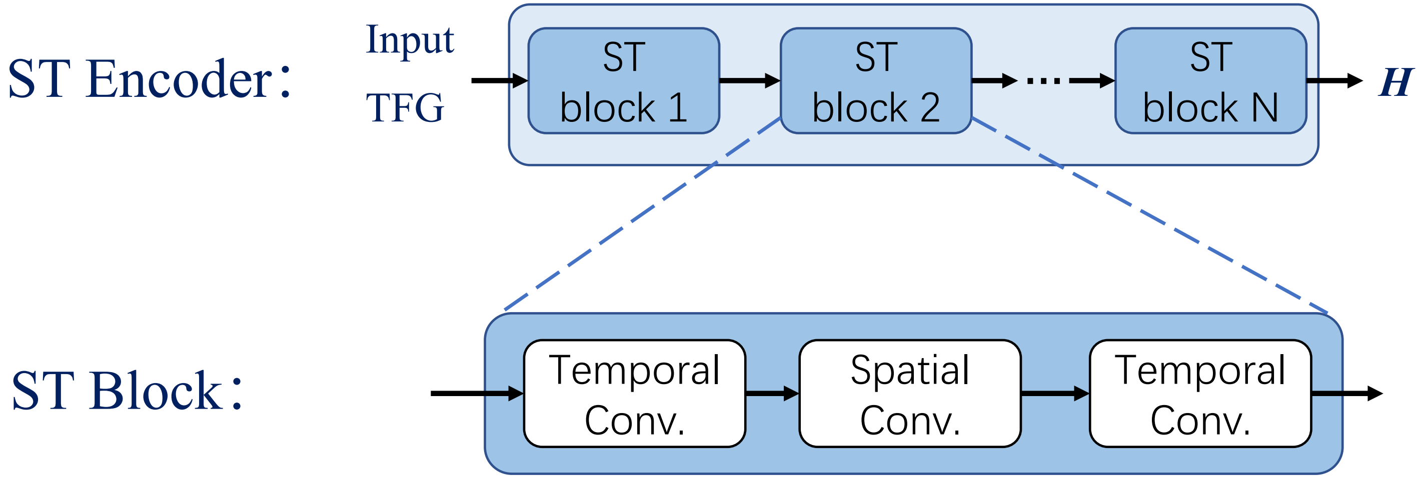
- Goal: encoding spatial-temporal traffic patterns into the embedding ***H***
  - It can be any spatio-temporal prediction model
- We choose the effective STGCN-like structure as our ST encoder:

ST Encoder：



Spatio-Temporal Encoder

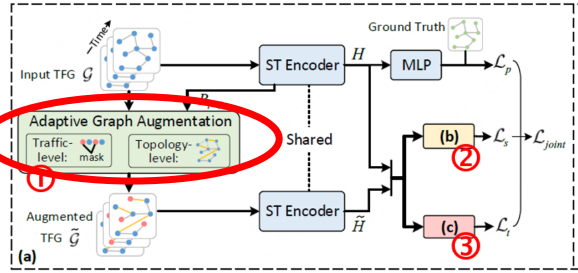
- Goal: encoding spatial-temporal traffic patterns into the embedding ***H***
  - It can be any spatio-temporal prediction model
- We choose the effective STGCN-like structure as our ST encoder:



Sandwich structure

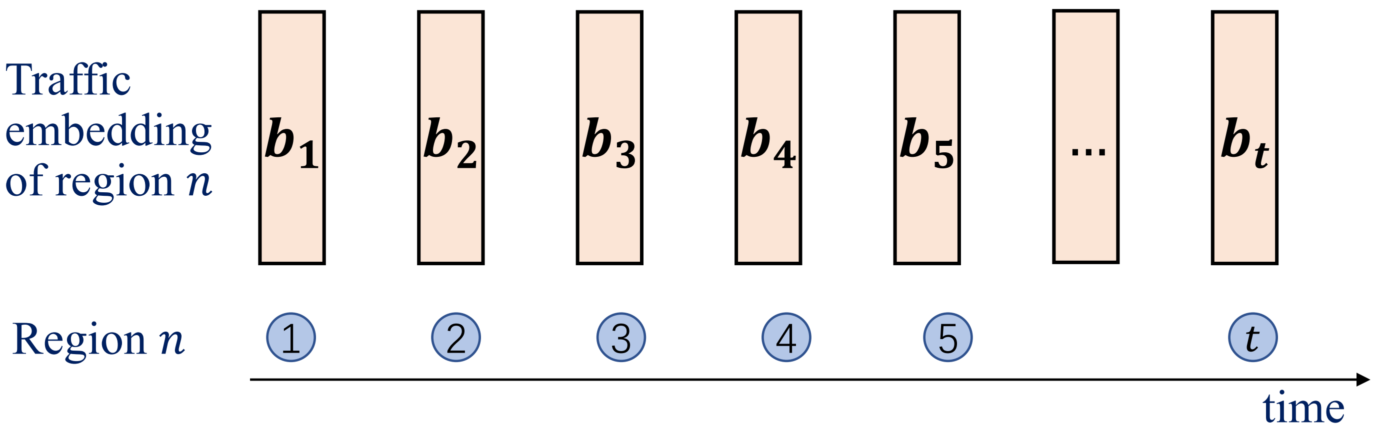
Adaptive Graph Augmentation on Traffic Flow Graph (TFG)

- Region-wise Heterogeneity Measurement



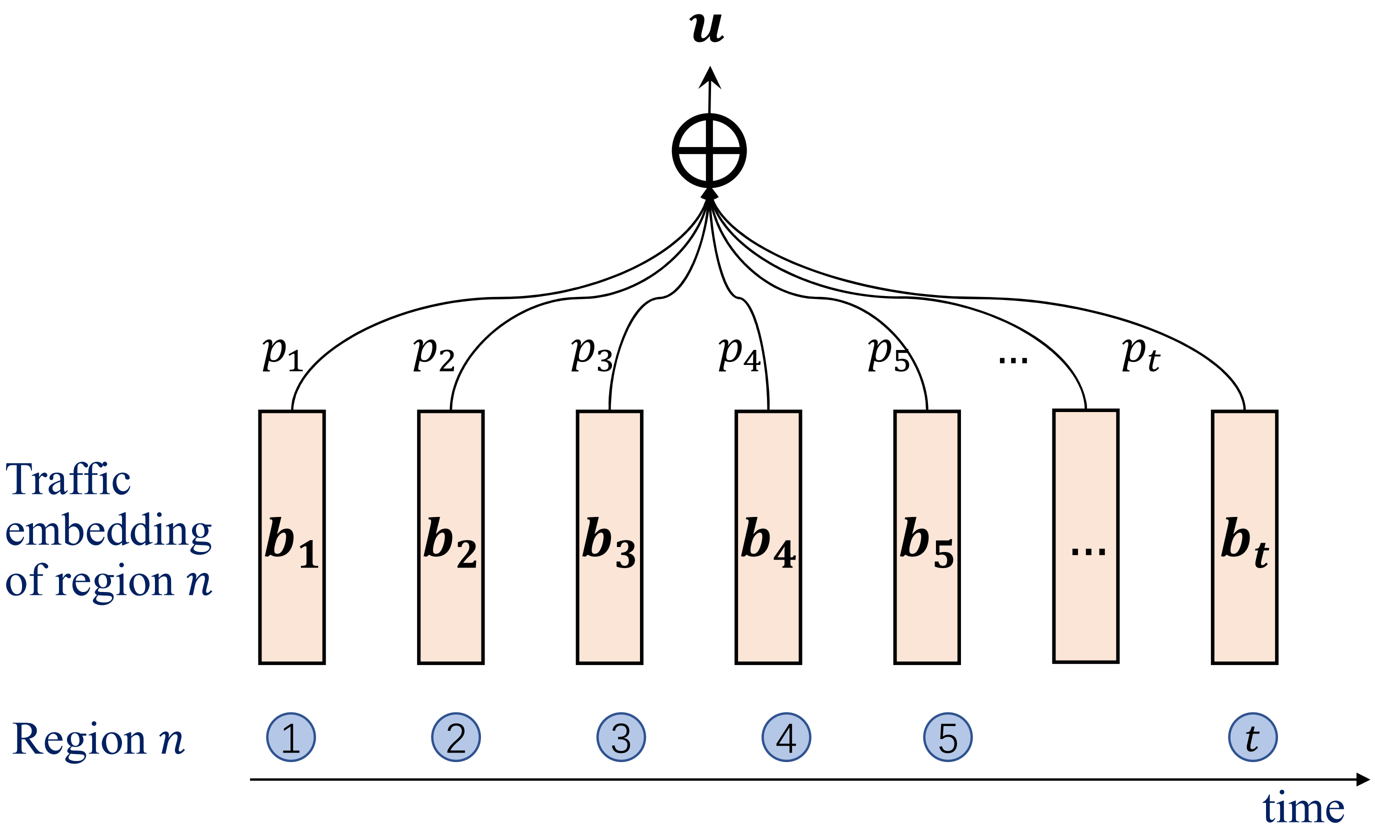
Adaptive Graph Augmentation on Traffic Flow Graph (TFG)

- Region-wise Heterogeneity Measurement



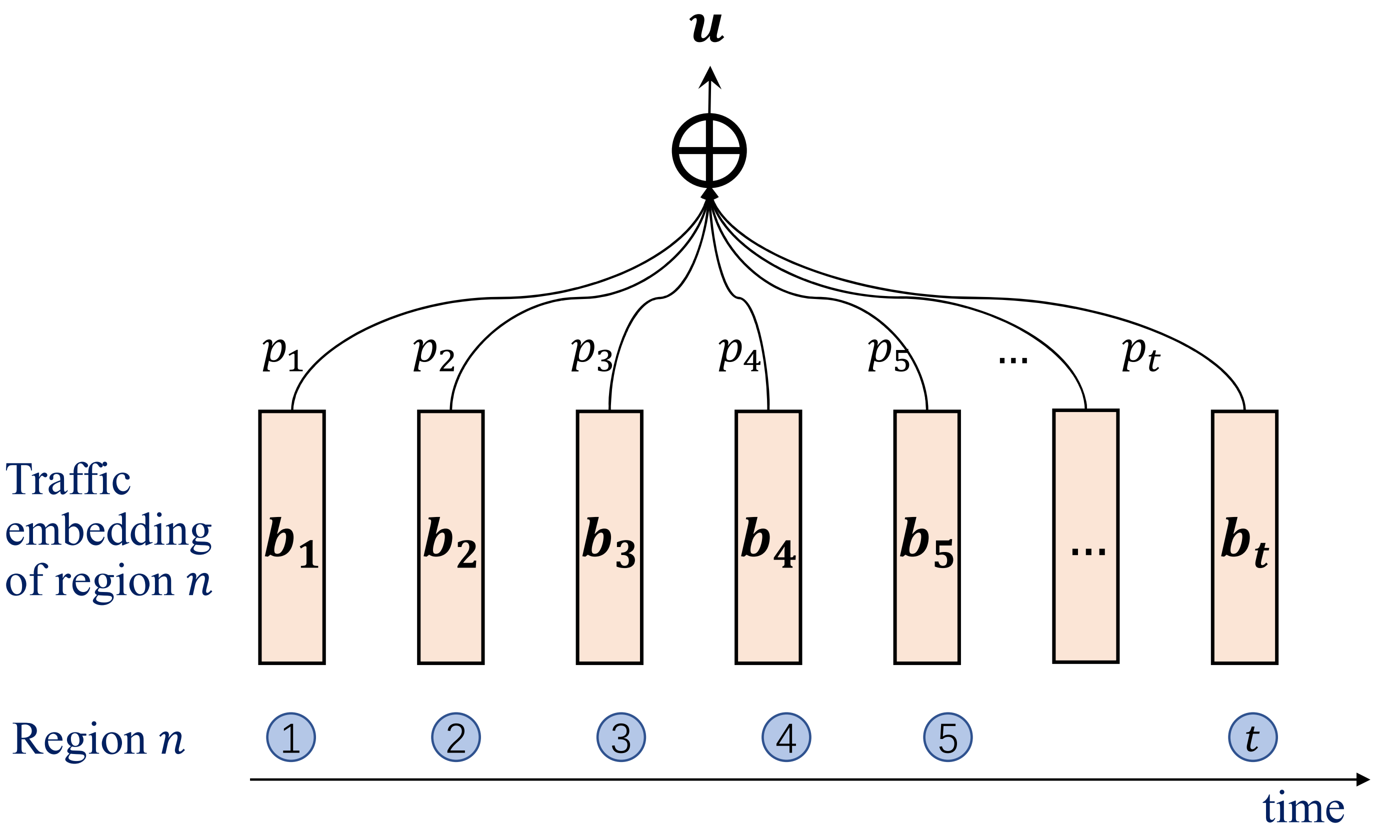
Adaptive Graph Augmentation on Traffic Flow Graph (TFG)

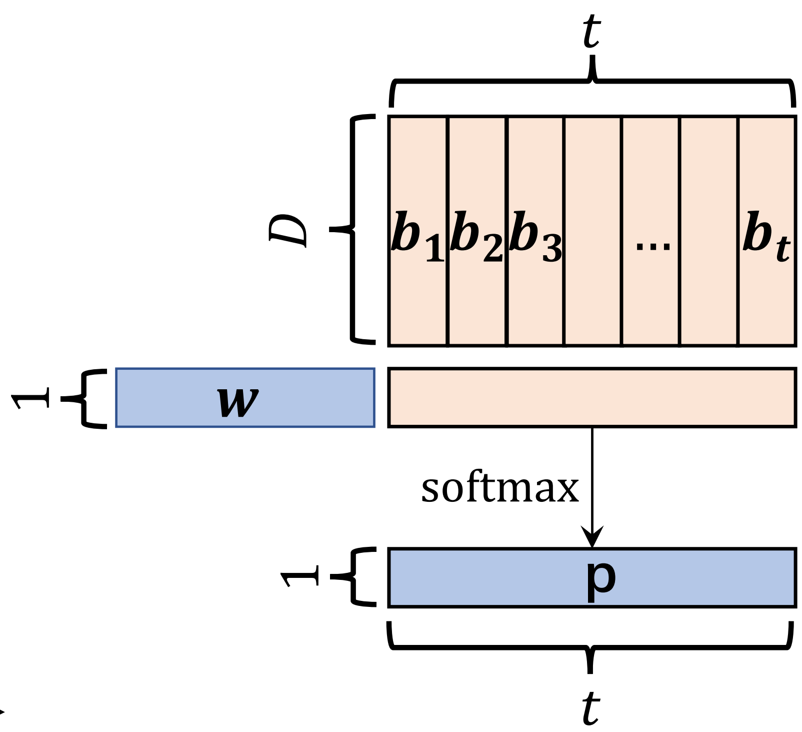
- Region-wise Heterogeneity Measurement



Adaptive Graph Augmentation on Traffic Flow Graph (TFG)

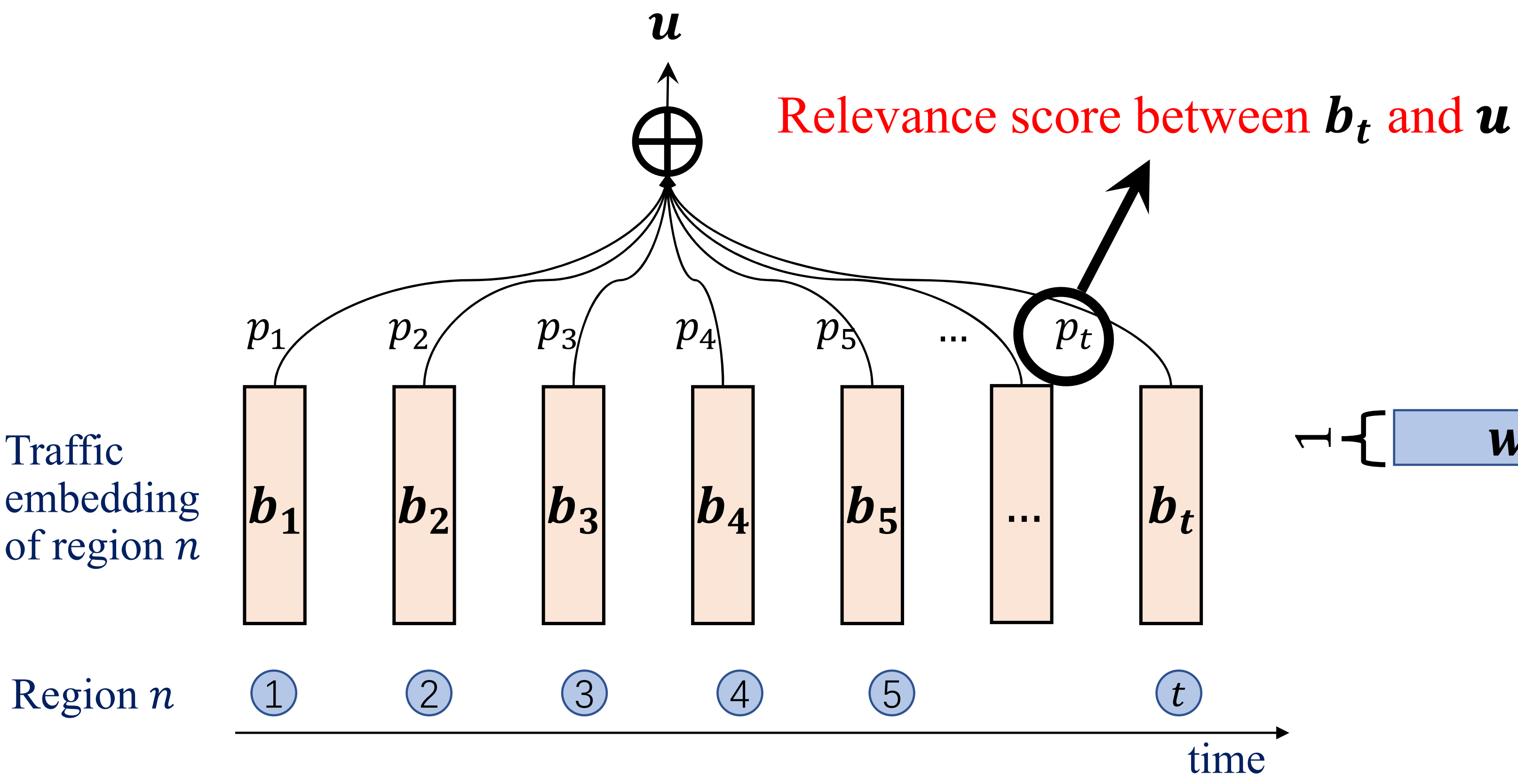
- Region-wise Heterogeneity Measurement





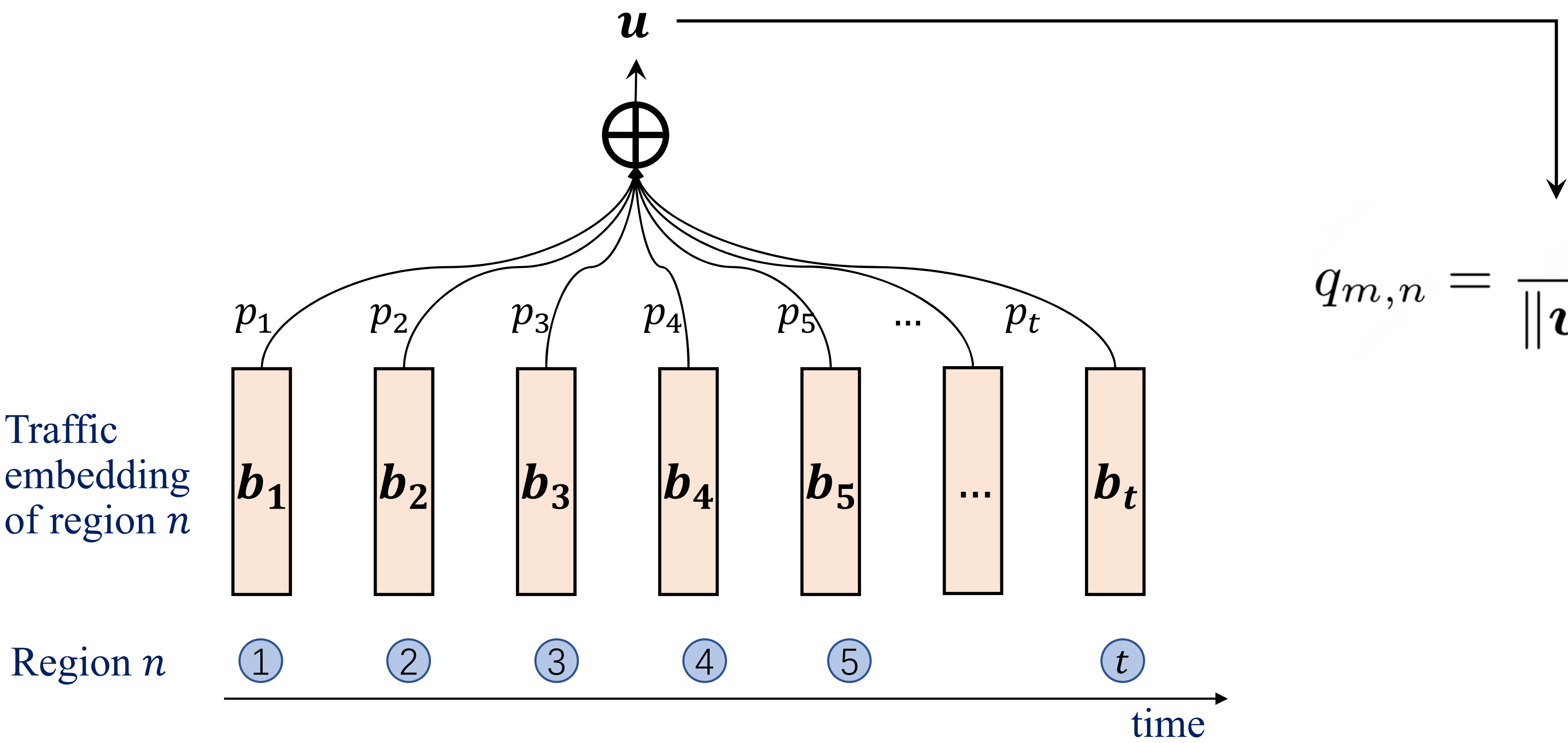
Adaptive Graph Augmentation on Traffic Flow Graph (TFG)

- Region-wise Heterogeneity Measurement



Adaptive Graph Augmentation on TFG

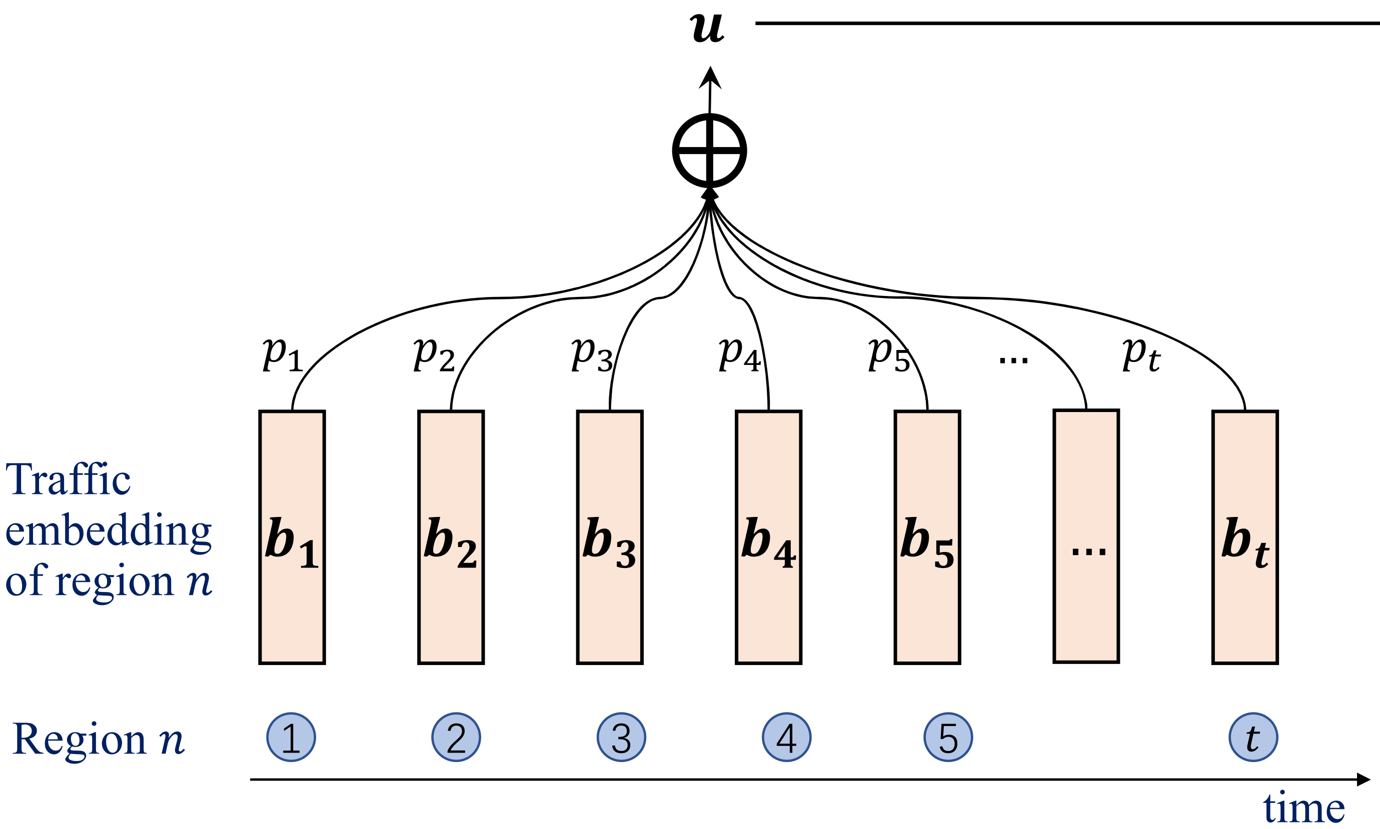
- Region-wise Heterogeneity Measurement



$$q_{m,n} = \frac{\boldsymbol{u}_m^\top \boldsymbol{u}_n}{\|\boldsymbol{u}_m\| \|\boldsymbol{u}_n\|}$$

Adaptive Graph Augmentation on TFG

- Region-wise Heterogeneity Measurement

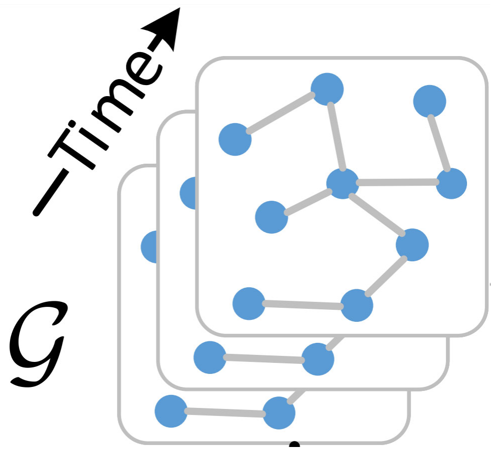


$$q_{m,n} = \frac{\boldsymbol{u}_m^\top \boldsymbol{u}_n}{\|\boldsymbol{u}_m\| \|\boldsymbol{u}_n\|}$$

Heterogeneity degree between two regions

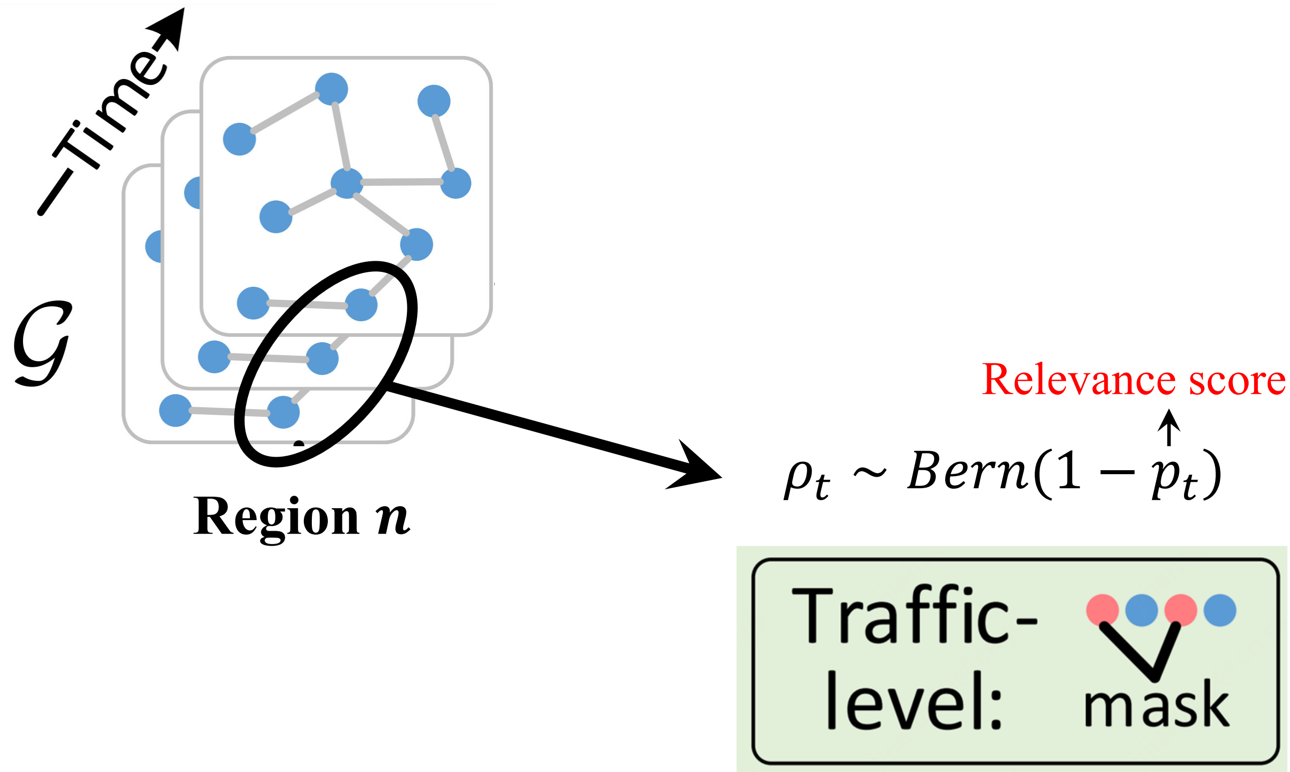
Adaptive Graph Augmentation on TFG

- Heterogeneity-guided Data Augmentation



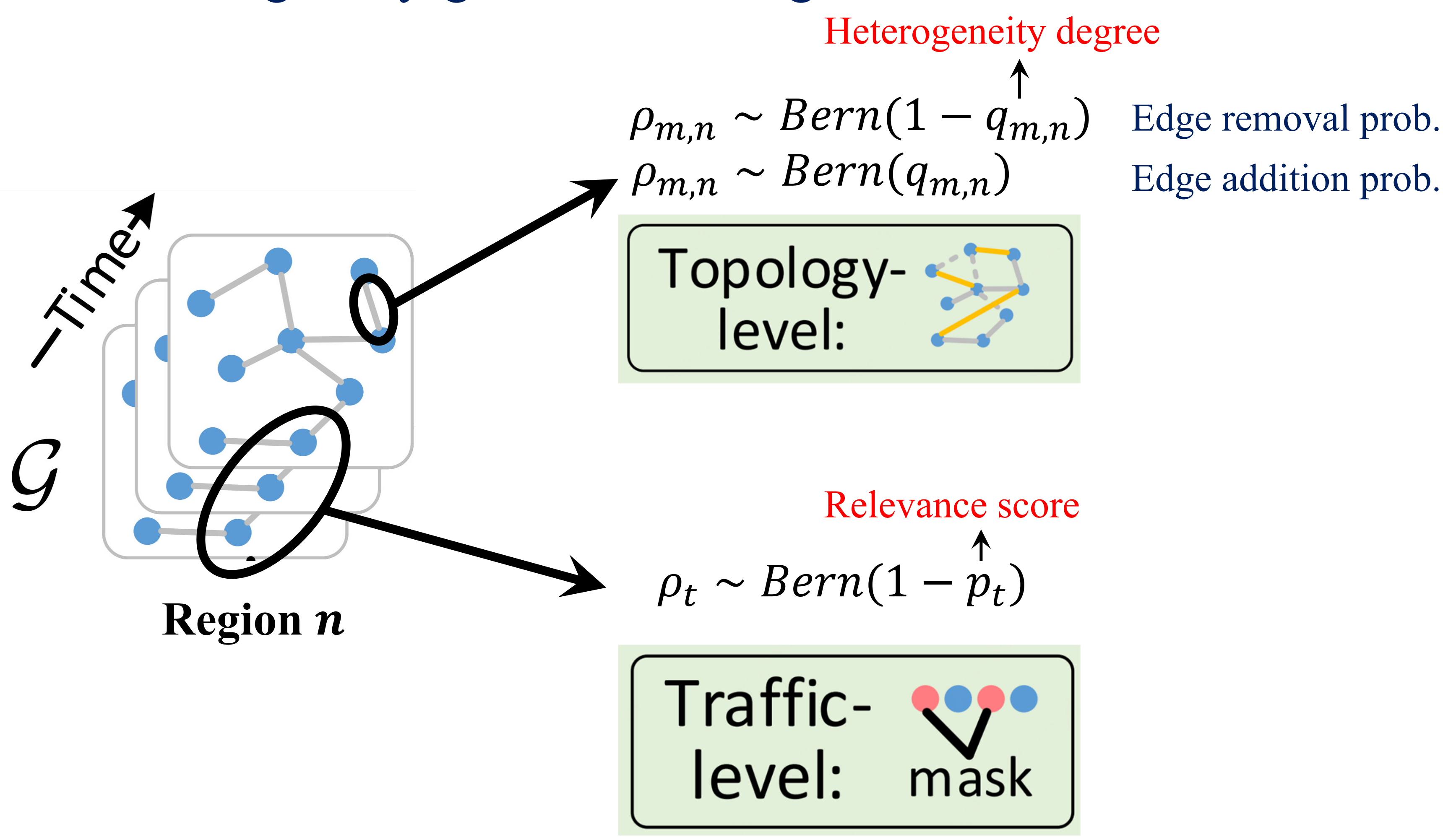
Adaptive Graph Augmentation on TFG

- Heterogeneity-guided Data Augmentation



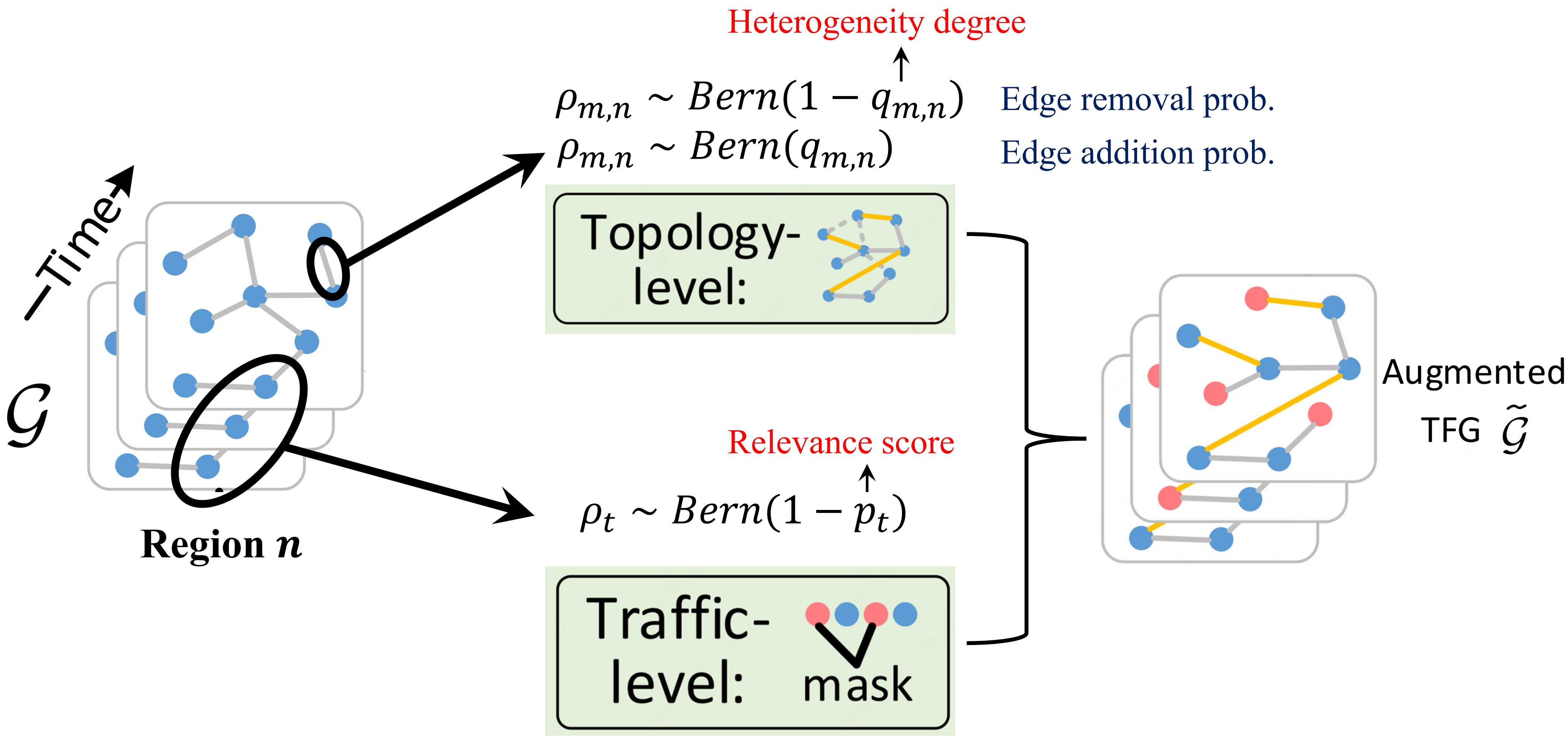
Adaptive Graph Augmentation on TFG

- Heterogeneity-guided Data Augmentation



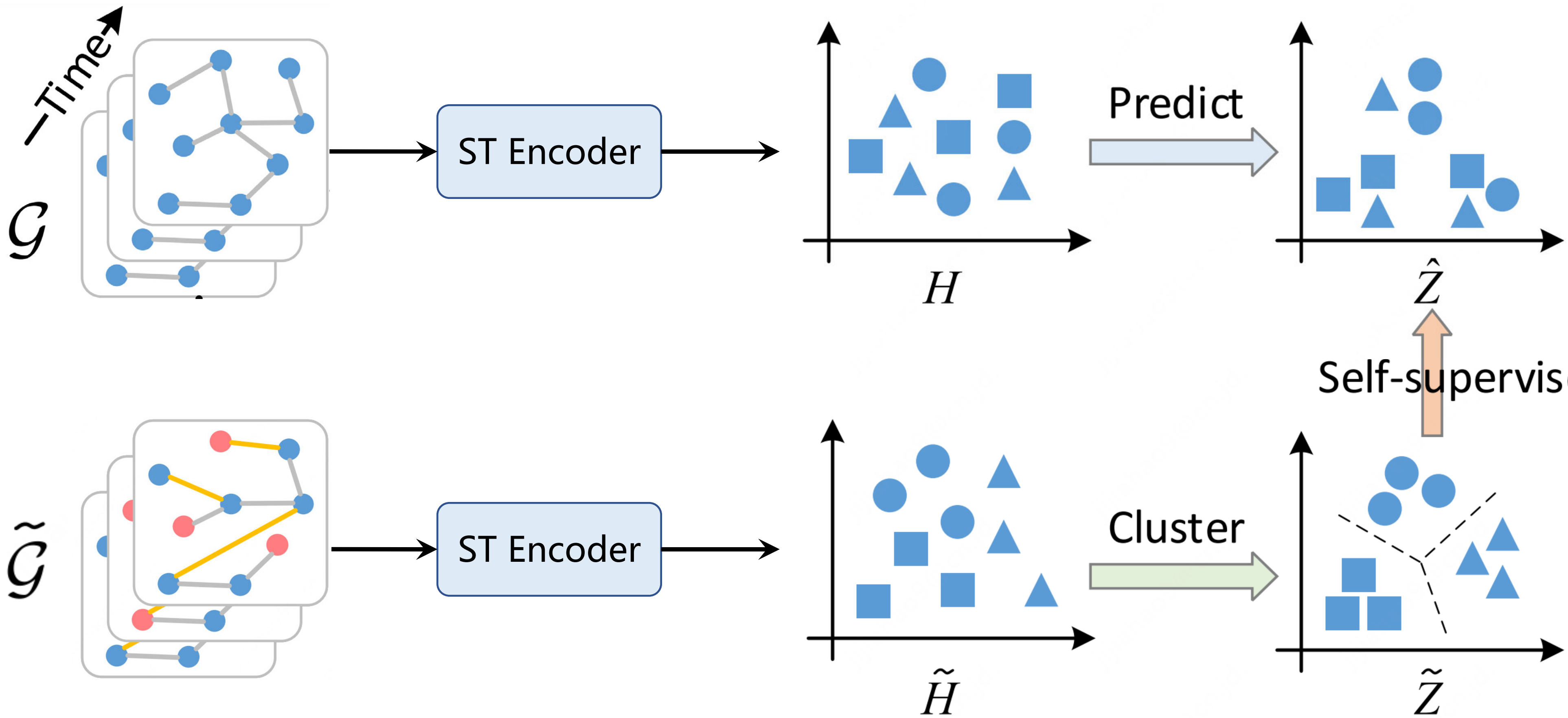
Adaptive Graph Augmentation on TFG

- Heterogeneity-guided Data Augmentation



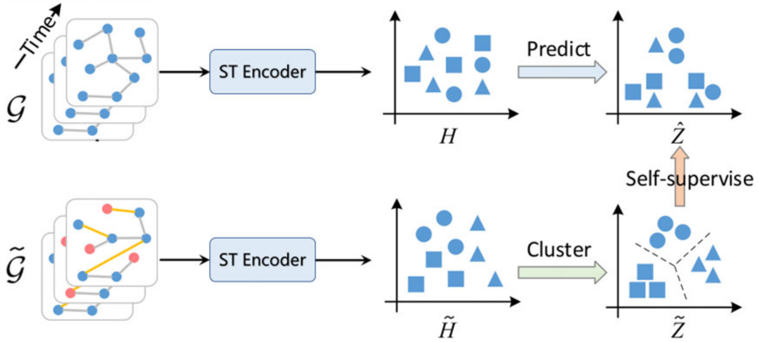
SSL for Spatial Heterogeneity Modeling

- Soft-clustering-based *predictive* SSL task



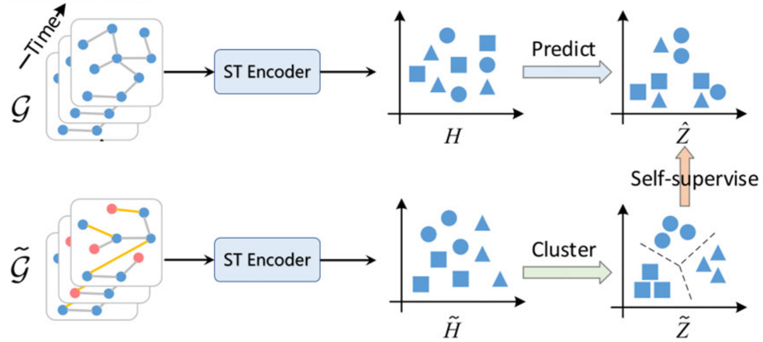
SSL for Spatial Heterogeneity Modeling

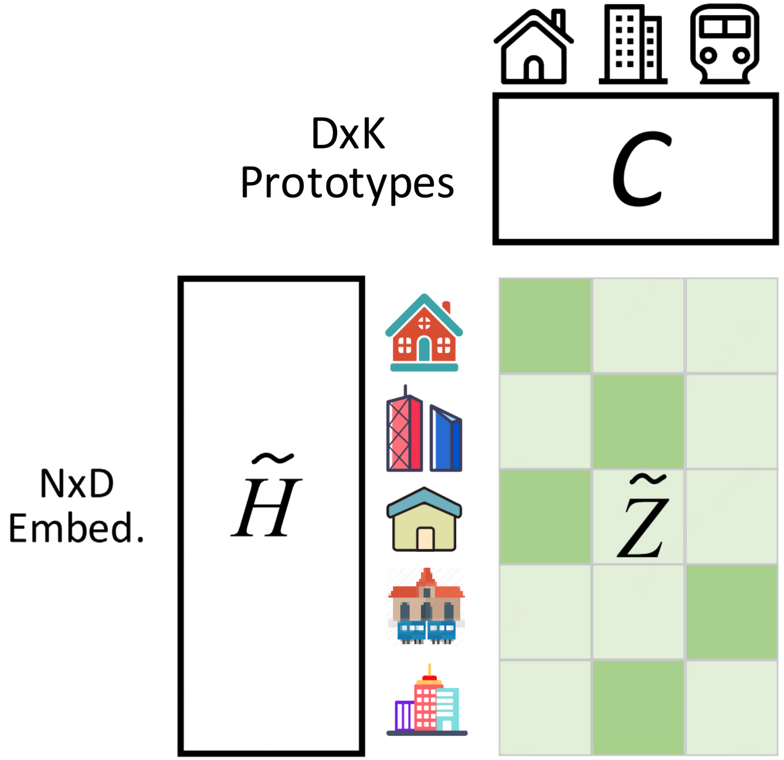
- Soft-clustering principal:
  - Generate *K* cluster embeddings (learnable)



SSL for Spatial Heterogeneity Modeling

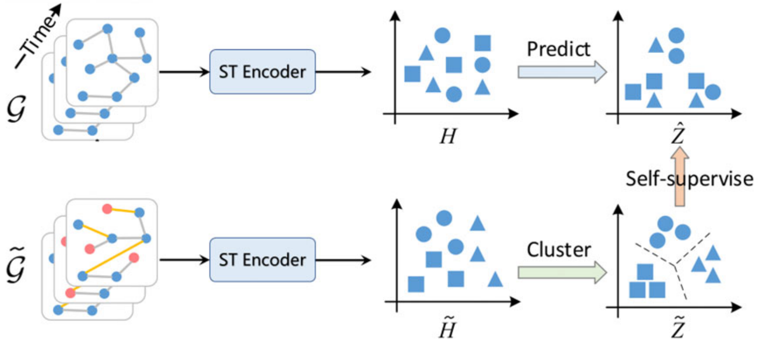
- Soft-clustering principal:
  - Generate K cluster embeddings (learnable)
  - Make cluster assignments using region embeddings

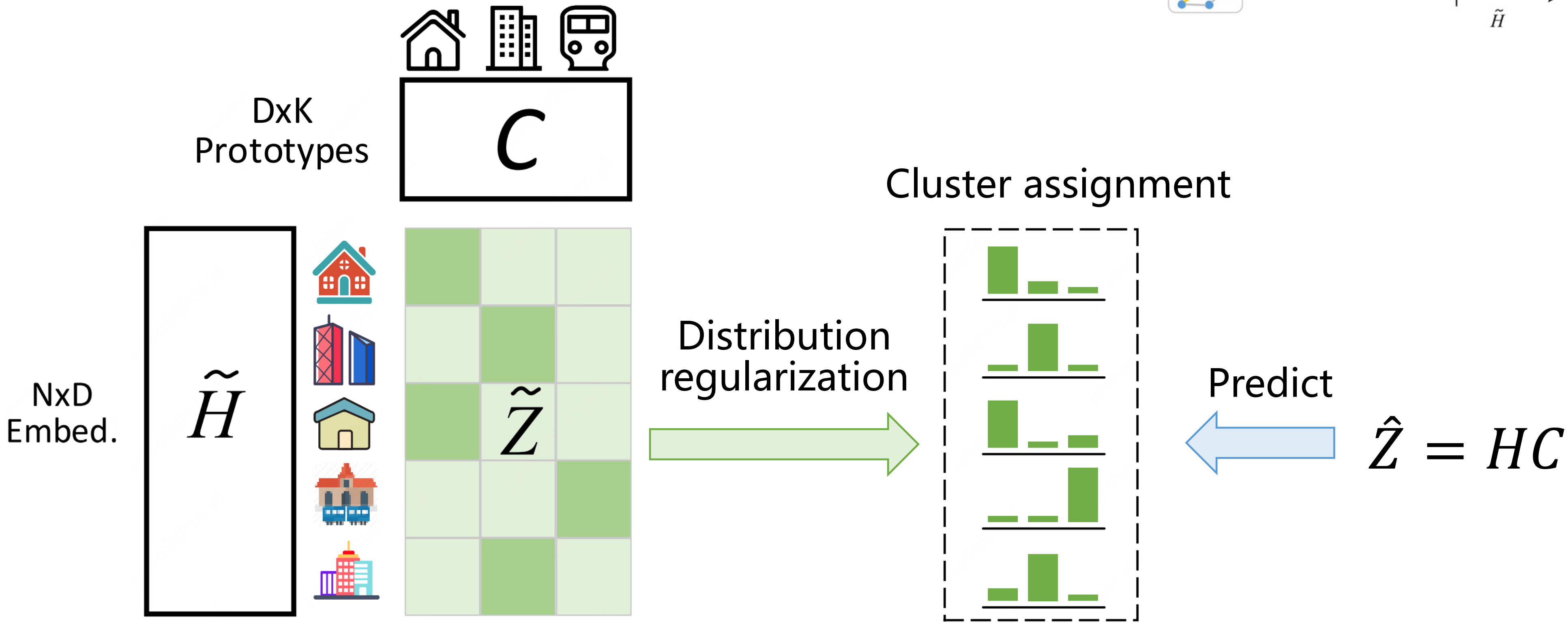




SSL for Spatial Heterogeneity Modeling

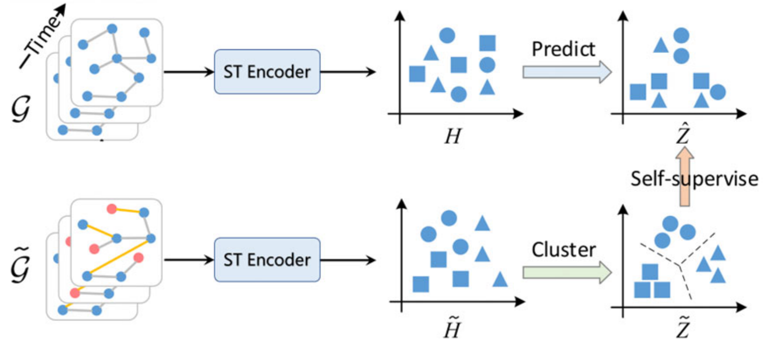
- Soft-clustering principal:
  - Generate *K* cluster embeddings (learnable)
  - Make cluster assignments using region embeddings
  - Predict cluster assignment score of each region

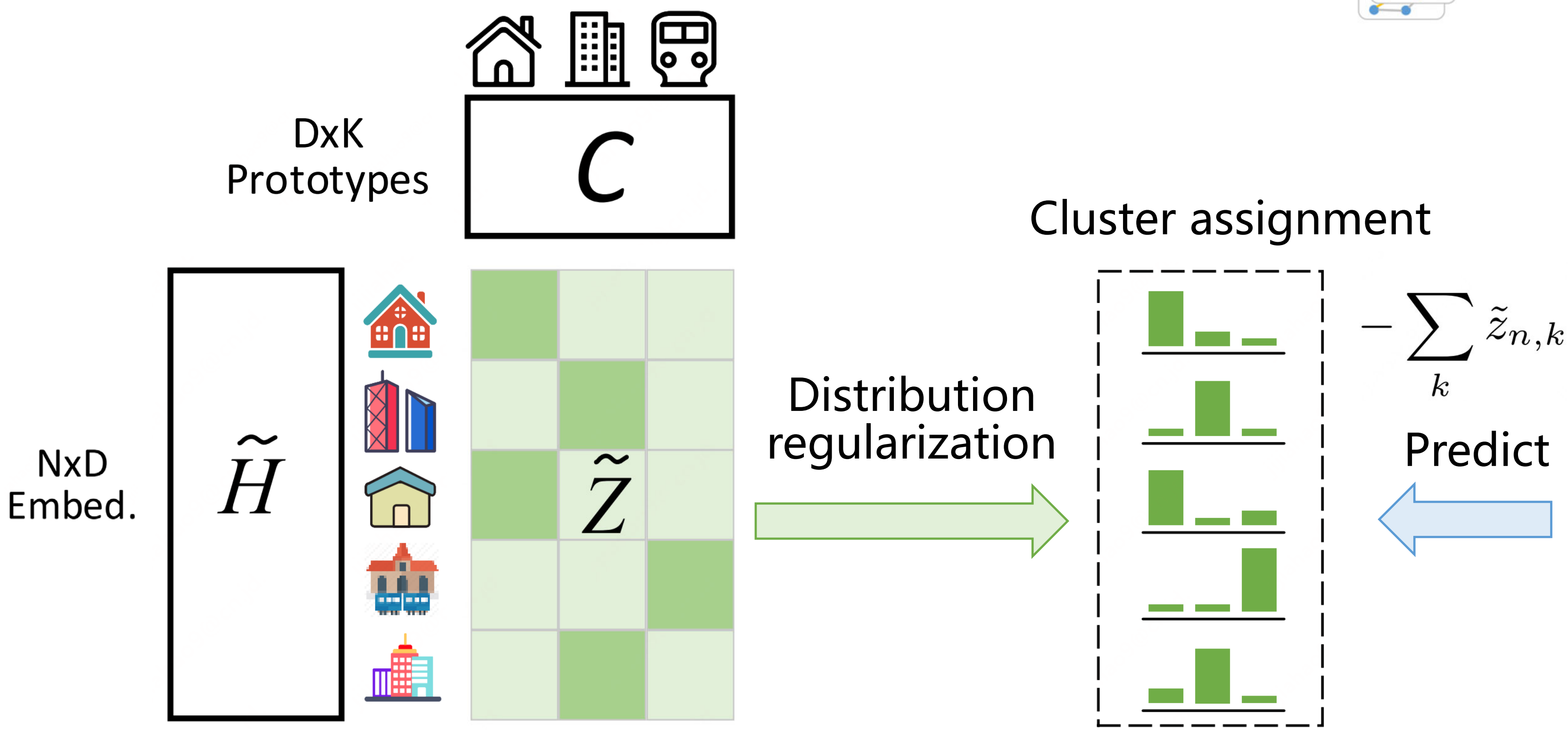




SSL for Spatial Heterogeneity Modeling

- Soft-clustering principal:
  - Generate *K* cluster embeddings (learnable)
  - Make cluster assignments using region embeddings
  - Predict cluster assignment score of each region





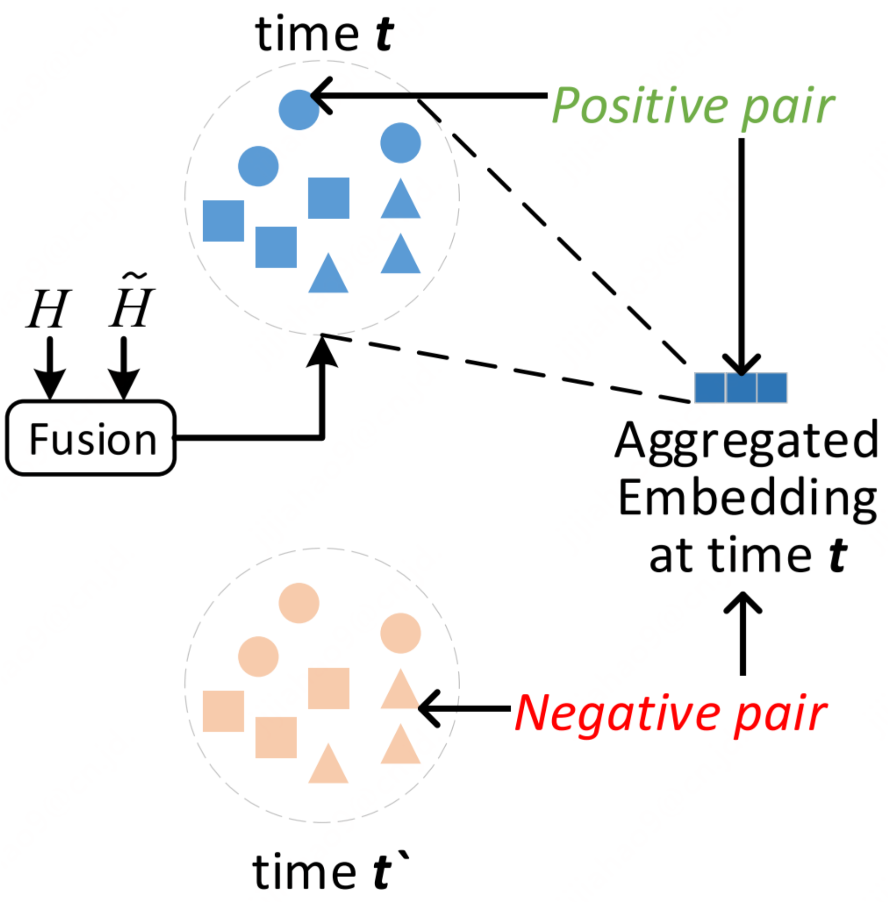
$$-\sum_{k} \tilde{z}_{n,k} \log \frac{\exp\left(\hat{z}_{n,k}/\gamma\right)}{\sum_{j} \exp\left(\hat{z}_{n,j}/\gamma\right)}$$

$$\hat{Z} = H\boxed{C}$$

Learnable

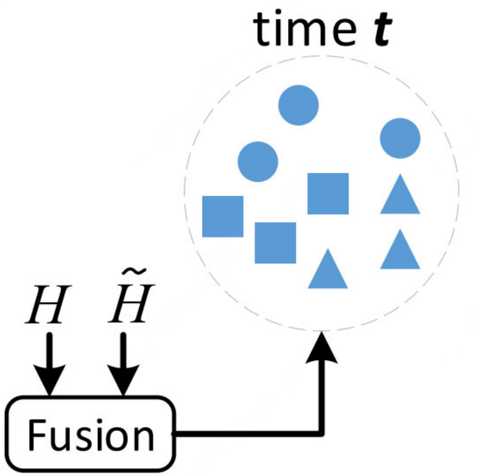
SSL for Temporal Heterogeneity Modeling

- Time-aware *contrastive* SSL task



SSL for Temporal Heterogeneity Modeling

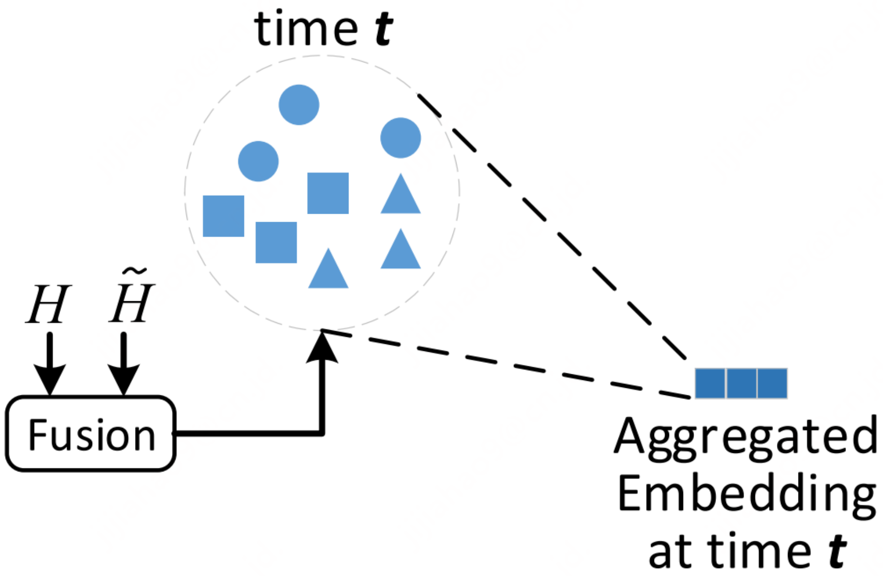
- Time-aware *contrastive* SSL task



Fusion: $\boldsymbol{v}_{t,n} = \boldsymbol{w_1} \odot \boldsymbol{h}_{t,n} + \boldsymbol{w_2} \odot \tilde{\boldsymbol{h}}_{t,n}$

SSL for Temporal Heterogeneity Modeling

- Time-aware *contrastive* SSL task

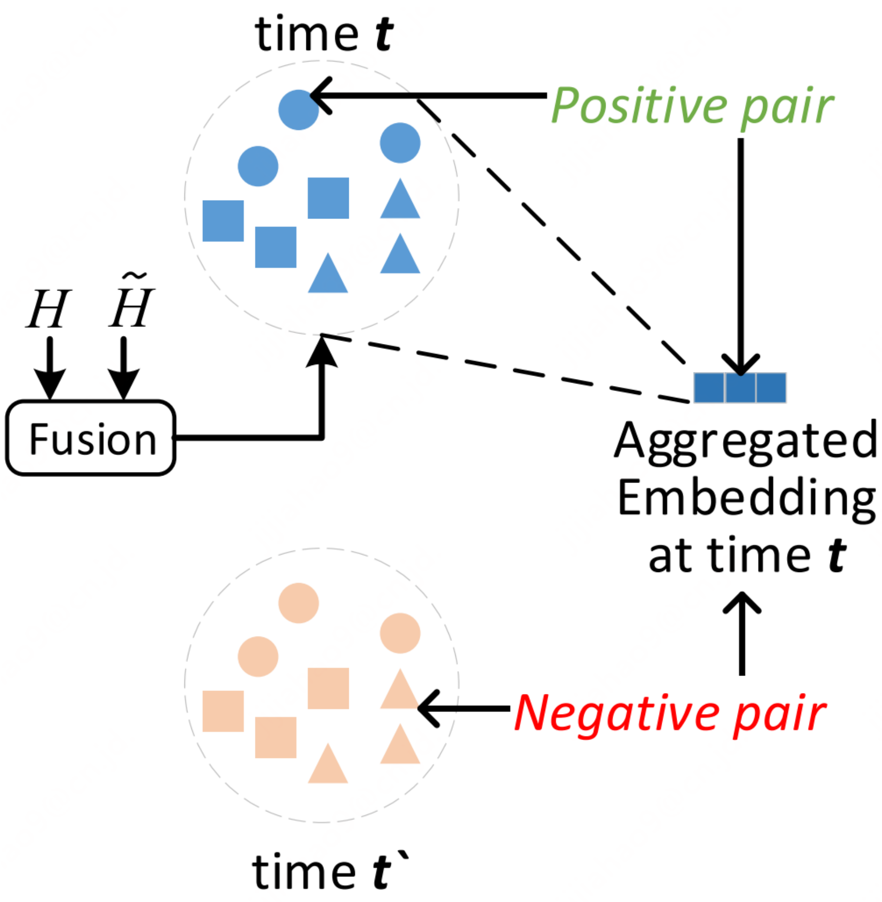


Fusion: $\boldsymbol{v}_{t,n} = \boldsymbol{w_1} \odot \boldsymbol{h}_{t,n} + \boldsymbol{w_2} \odot \tilde{\boldsymbol{h}}_{t,n}$

Aggregation: $\boldsymbol{s}_t = \sigma\left(\frac{1}{N}\sum_{n=1}^{N} \boldsymbol{v}_{t,n}\right)$

SSL for Temporal Heterogeneity Modeling

- Time-aware *contrastive* SSL task



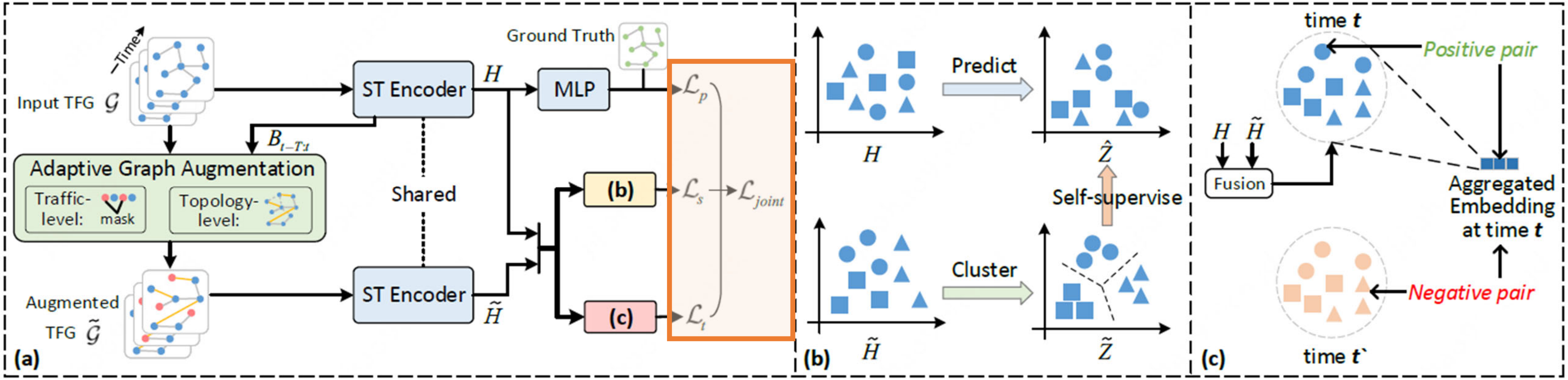
Fusion: $\boldsymbol{v}_{t,n} = \boldsymbol{w_1} \odot \boldsymbol{h}_{t,n} + \boldsymbol{w_2} \odot \tilde{\boldsymbol{h}}_{t,n}$

Aggregation: $\boldsymbol{s}_t = \sigma\left(\frac{1}{N}\sum_{n=1}^{N} \boldsymbol{v}_{t,n}\right)$

Contrastive loss: $\mathcal{L}_t = -\left(\overbrace{\sum_{n=1}^{N} \log g\left(\boldsymbol{v}_{t,n}, \boldsymbol{s}_t\right)}^{\text{Positive}} + \underbrace{\sum_{n=1}^{N} \log\left(1 - g\left(\boldsymbol{v}_{t',n}, \boldsymbol{s}_t\right)\right)}_{\text{Negative}}\right)$

Model Training

Spatio-Temporal Self-Supervised Learning (**ST-SSL**) for robust traffic prediction



- Loss of traffic prediction branch: $\mathcal{L}_p$
- Loss of spatial heterogeneity modeling branch: $\mathcal{L}_s$
- Loss of temporal heterogeneity modeling branch: $\mathcal{L}_t$

$$\left.\begin{matrix}\mathcal{L}_p\\ \mathcal{L}_s\\ \mathcal{L}_t\end{matrix}\right\}\mathcal{L}_{joint}$$

Experiments: Setup

- Datasets
  - Four public datasets[1, 2] belonging to two types of real-world traffic mode

Data type	Bike rental		Taxi GPS	
Dataset	NYCBike1	NYCBike2	NYCTaxi	BJTaxi
Time interval	1 hour	30 min	30 min	30 min
# regions	16×8	10×20	10×20	32×32
# taxis/bikes	6.8k+	2.6m+	22m+	34k+

[1] Deep spatio-temporal residual networks for citywide crowd flows prediction. AAAI'17.
[2] Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction. AAAI'19.

Experiments: Setup

- Baseline methods
  - Time series prediction approaches
    - Autoregressive Integrated Moving Average Model (ARIMA)
    - Support Vector Regression (SVR)
  - Spatio-temporal prediction methods
    - Spatio-Temporal Residual Networks (ST-ResNet) [Zhang, Zheng and Qi 2017]
    - Spatio-Temporal Graph Convolutional Network (STGCN) [Yu, Yin and Zhu 2018]
    - Graph Multi-Attention Network (GMAN) [Zheng et al. 2020]
  - Spatial-temporal methods considering heterogeneity
    - Adaptive Graph Convolutional Recurrent Network (AGCRN) [Bai et al. 2020]
    - Spatial-Temporal Synchronous Graph Convolutional Networks (STSGCN) [Song et al. 2020]
    - Spatial-Temporal Fusion Graph Neural Networks (STFGNN) [Li and Zhu 2021]

Experiments: Overall results

Dataset	Metric	Type	ARIMA	SVR	ST-ResNet	STGCN	GMAN	AGCRN	STSGCN	STFGNN	ST-SSL
NYCBike1	MAE	In	10.66	7.27	5.53±0.06	5.33±0.02	6.77±3.42	5.17±0.03	5.81±0.04	6.53±0.10	**4.94±0.02**
		Out	11.33	7.98	5.74±0.07	5.59±0.03	7.17±3.61	5.47±0.03	6.10±0.04	6.79±0.08	**5.26±0.02**
	MAPE	In	33.05	25.39	25.46±0.20	26.92±0.08	31.72±12.29	25.59±0.22	26.51±0.32	32.14±0.23	**23.69±0.11**
		Out	35.03	27.42	26.36±0.50	27.69±0.14	34.74±17.04	26.63±0.30	27.56±0.39	32.88±0.19	**24.60±0.27**
NYCBike2	MAE	In	8.91	12.82	5.63±0.14	5.21±0.02	5.24±0.13	5.18±0.03	5.25±0.03	5.80±0.10	**5.04±0.03**
		Out	8.70	11.48	5.26±0.08	4.92±0.02	4.97±0.14	4.79±0.04	4.94±0.05	5.51±0.11	**4.71±0.02**
	MAPE	In	28.86	46.52	32.17±0.85	27.73±0.16	27.38±1.13	27.14±0.14	29.26±0.13	30.73±0.49	**22.54±0.10**
		Out	28.22	41.91	30.48±0.86	26.83±0.21	26.75±1.14	26.17±0.22	28.02±0.23	29.98±0.46	**21.17±0.13**
NYCTaxi	MAE	In	20.86	52.16	13.48±0.14	13.12±0.04	15.09±0.61	12.13±0.11	13.69±0.11	16.25±0.38	**11.99±0.12**
		Out	16.80	41.71	10.78±0.25	10.35±0.03	12.06±0.39	9.87±0.04	10.75±0.17	12.47±0.25	**9.78±0.09**
	MAPE	In	21.49	65.10	24.83±0.55	21.01±0.18	22.73±1.20	18.78±0.04	22.91±0.44	24.01±0.30	**16.38±0.10**
		Out	21.23	64.06	24.42±0.52	20.78±0.16	21.97±0.86	18.41±0.21	22.37±0.16	23.28±0.47	**16.86±0.23**
BJTaxi	MAE	In	21.48	52.77	12.12±0.11	12.34±0.09	13.13±0.43	12.30±0.06	12.72±0.03	13.83±0.04	**11.31±0.03**
		Out	21.60	52.74	12.16±0.12	12.41±0.08	13.20±0.43	12.38±0.06	12.79±0.03	13.89±0.04	**11.40±0.02**
	MAPE	In	23.12	65.51	15.50±0.26	16.66±0.21	18.67±0.99	15.61±0.15	17.22±0.17	19.29±0.07	**15.03±0.13**
		Out	20.67	65.51	15.57±0.26	16.76±0.22	18.84±1.04	15.75±0.15	17.35±0.17	19.41±0.07	**15.19±0.15**

Table 2: Model comparison on four datasets in terms of MAE and MAPE (%). In and Out represent the inflow and outflow.

Experiments: Overall results

Dataset	Metric	Type	ARIMA	SVR	ST-ResNet	STGCN	GMAN	AGCRN	STSGCN	STFGNN	ST-SSL
NYCBike1	MAE	In	10.66	7.27	5.53±0.06	5.33±0.02	6.77±3.42	5.17±0.03	5.81±0.04	6.53±0.10	**4.94±0.02**
		Out	11.33	7.98	5.74±0.07	5.59±0.03	7.17±3.61	5.47±0.03	6.10±0.04	6.79±0.08	**5.26±0.02**
	MAPE	In	33.05	25.39	25.46±0.20	26.92±0.08	31.72±12.29	25.59±0.22	26.51±0.32	32.14±0.23	**23.69±0.11**
		Out	35.03	27.42	26.36±0.50	27.69±0.14	34.74±17.04	26.63±0.30	27.56±0.39	32.88±0.19	**24.60±0.27**
NYCBike2	MAE	In	8.91	12.82	5.63±0.14	5.21±0.02	5.24±0.13	5.18±0.03	5.25±0.03	5.80±0.10	**5.04±0.03**
		Out	8.70	11.48	5.26±0.08	4.92±0.02	4.97±0.14	4.79±0.04	4.94±0.05	5.51±0.11	**4.71±0.02**
	MAPE	In	28.86	46.52	32.17±0.85	27.73±0.16	27.38±1.13	27.14±0.14	29.26±0.13	30.73±0.49	**22.54±0.10**
		Out	28.22	41.91	30.48±0.86	26.83±0.21	26.75±1.14	26.17±0.22	28.02±0.23	29.98±0.46	**21.17±0.13**
NYCTaxi	MAE	In	20.86	52.16	13.48±0.14	13.12±0.04	15.09±0.61	12.13±0.11	13.69±0.11	16.25±0.38	**11.99±0.12**
		Out	16.80	41.71	10.78±0.25	10.35±0.03	12.06±0.39	9.87±0.04	10.75±0.17	12.47±0.25	**9.78±0.09**
	MAPE	In	21.49	65.10	24.83±0.55	21.01±0.18	22.73±1.20	18.78±0.04	22.91±0.44	24.01±0.30	**16.38±0.10**
		Out	21.23	64.06	24.42±0.52	20.78±0.16	21.97±0.86	18.41±0.21	22.37±0.16	23.28±0.47	**16.86±0.23**
BJTaxi	MAE	In	21.48	52.77	12.12±0.11	12.34±0.09	13.13±0.43	12.30±0.06	12.72±0.03	13.83±0.04	**11.31±0.03**
		Out	21.60	52.74	12.16±0.12	12.41±0.08	13.20±0.43	12.38±0.06	12.79±0.03	13.89±0.04	**11.40±0.02**
	MAPE	In	23.12	65.51	15.50±0.26	16.66±0.21	18.67±0.99	15.61±0.15	17.22±0.17	19.29±0.07	**15.03±0.13**
		Out	20.67	65.51	15.57±0.26	16.76±0.22	18.84±1.04	15.75±0.15	17.35±0.17	19.41±0.07	**15.19±0.15**

Table 2: Model comparison on four datasets in terms of MAE and MAPE (%). In and Out represent the inflow and outflow.

- ST methods outperform time series approaches: necessity to capture spatial dependencies

Experiments: Overall results

Dataset	Metric	Type	ARIMA	SVR	ST-ResNet	STGCN	GMAN	AGCRN	STSGCN	STFGNN	ST-SSL
NYCBike1	MAE	In	10.66	7.27	5.53±0.06	5.33±0.02	6.77±3.42	5.17±0.03	5.81±0.04	6.53±0.10	**4.94±0.02**
		Out	11.33	7.98	5.74±0.07	5.59±0.03	7.17±3.61	5.47±0.03	6.10±0.04	6.79±0.08	**5.26±0.02**
	MAPE	In	33.05	25.39	25.46±0.20	26.92±0.08	31.72±12.29	25.59±0.22	26.51±0.32	32.14±0.23	**23.69±0.11**
		Out	35.03	27.42	26.36±0.50	27.69±0.14	34.74±17.04	26.63±0.30	27.56±0.39	32.88±0.19	**24.60±0.27**
NYCBike2	MAE	In	8.91	12.82	5.63±0.14	5.21±0.02	5.24±0.13	5.18±0.03	5.25±0.03	5.80±0.10	**5.04±0.03**
		Out	8.70	11.48	5.26±0.08	4.92±0.02	4.97±0.14	4.79±0.04	4.94±0.05	5.51±0.11	**4.71±0.02**
	MAPE	In	28.86	46.52	32.17±0.85	27.73±0.16	27.38±1.13	27.14±0.14	29.26±0.13	30.73±0.49	**22.54±0.10**
		Out	28.22	41.91	30.48±0.86	26.83±0.21	26.75±1.14	26.17±0.22	28.02±0.23	29.98±0.46	**21.17±0.13**
NYCTaxi	MAE	In	20.86	52.16	13.48±0.14	13.12±0.04	15.09±0.61	12.13±0.11	13.69±0.11	16.25±0.38	**11.99±0.12**
		Out	16.80	41.71	10.78±0.25	10.35±0.03	12.06±0.39	9.87±0.04	10.75±0.17	12.47±0.25	**9.78±0.09**
	MAPE	In	21.49	65.10	24.83±0.55	21.01±0.18	22.73±1.20	18.78±0.04	22.91±0.44	24.01±0.30	**16.38±0.10**
		Out	21.23	64.06	24.42±0.52	20.78±0.16	21.97±0.86	18.41±0.21	22.37±0.16	23.28±0.47	**16.86±0.23**
BJTaxi	MAE	In	21.48	52.77	12.12±0.11	12.34±0.09	13.13±0.43	12.30±0.06	12.72±0.03	13.83±0.04	**11.31±0.03**
		Out	21.60	52.74	12.16±0.12	12.41±0.08	13.20±0.43	12.38±0.06	12.79±0.03	13.89±0.04	**11.40±0.02**
	MAPE	In	23.12	65.51	15.50±0.26	16.66±0.21	18.67±0.99	15.61±0.15	17.22±0.17	19.29±0.07	**15.03±0.13**
		Out	20.67	65.51	15.57±0.26	16.76±0.22	18.84±1.04	15.75±0.15	17.35±0.17	19.41±0.07	**15.19±0.15**

Table 2: Model comparison on four datasets in terms of MAE and MAPE (%). In and Out represent the inflow and outflow.

- ST methods outperform time series approaches: necessity to capture spatial dependencies
- Methods considering heterogeneity perform better: rationality of learning spatial and temporal heterogeneity

Experiments: Overall results

Dataset	Metric	Type	ARIMA	SVR	ST-ResNet	STGCN	GMAN	AGCRN	STSGCN	STFGNN	ST-SSL
NYCBike1	MAE	In	10.66	7.27	5.53±0.06	5.33±0.02	6.77±3.42	5.17±0.03	5.81±0.04	6.53±0.10	**4.94±0.02**
		Out	11.33	7.98	5.74±0.07	5.59±0.03	7.17±3.61	5.47±0.03	6.10±0.04	6.79±0.08	**5.26±0.02**
	MAPE	In	33.05	25.39	25.46±0.20	26.92±0.08	31.72±12.29	25.59±0.22	26.51±0.32	32.14±0.23	**23.69±0.11**
		Out	35.03	27.42	26.36±0.50	27.69±0.14	34.74±17.04	26.63±0.30	27.56±0.39	32.88±0.19	**24.60±0.27**
NYCBike2	MAE	In	8.91	12.82	5.63±0.14	5.21±0.02	5.24±0.13	5.18±0.03	5.25±0.03	5.80±0.10	**5.04±0.03**
		Out	8.70	11.48	5.26±0.08	4.92±0.02	4.97±0.14	4.79±0.04	4.94±0.05	5.51±0.11	**4.71±0.02**
	MAPE	In	28.86	46.52	32.17±0.85	27.73±0.16	27.38±1.13	27.14±0.14	29.26±0.13	30.73±0.49	**22.54±0.10**
		Out	28.22	41.91	30.48±0.86	26.83±0.21	26.75±1.14	26.17±0.22	28.02±0.23	29.98±0.46	**21.17±0.13**
NYCTaxi	MAE	In	20.86	52.16	13.48±0.14	13.12±0.04	15.09±0.61	12.13±0.11	13.69±0.11	16.25±0.38	**11.99±0.12**
		Out	16.80	41.71	10.78±0.25	10.35±0.03	12.06±0.39	9.87±0.04	10.75±0.17	12.47±0.25	**9.78±0.09**
	MAPE	In	21.49	65.10	24.83±0.55	21.01±0.18	22.73±1.20	18.78±0.04	22.91±0.44	24.01±0.30	**16.38±0.10**
		Out	21.23	64.06	24.42±0.52	20.78±0.16	21.97±0.86	18.41±0.21	22.37±0.16	23.28±0.47	**16.86±0.23**
BJTaxi	MAE	In	21.48	52.77	12.12±0.11	12.34±0.09	13.13±0.43	12.30±0.06	12.72±0.03	13.83±0.04	**11.31±0.03**
		Out	21.60	52.74	12.16±0.12	12.41±0.08	13.20±0.43	12.38±0.06	12.79±0.03	13.89±0.04	**11.40±0.02**
	MAPE	In	23.12	65.51	15.50±0.26	16.66±0.21	18.67±0.99	15.61±0.15	17.22±0.17	19.29±0.07	**15.03±0.13**
		Out	20.67	65.51	15.57±0.26	16.76±0.22	18.84±1.04	15.75±0.15	17.35±0.17	19.41±0.07	**15.19±0.15**

Table 2: Model comparison on four datasets in terms of MAE and MAPE (%). In and Out represent the inflow and outflow.

- ST methods outperform time series approaches: necessity to capture spatial dependencies
- Methods considering heterogeneity perform better: rationality of learning spatial and temporal heterogeneity
- Our ST-SSL performs best over all datasets: effectiveness of *jointly* modeling the spatial and temporal heterogeneity in a *self-supervised* manner

Experiments: Ablation study

- Ablation study on sub-modules, including
  - Adaptive augmentation: graph topology-level and traffic-level
  - Spatial heterogeneity modeling and temporal heterogeneity modeling

➢ ST-SSL-sa: replaces heterogeneity-guided structure augmentation on graph topology with random augmentations

➢ ST-SSL-ta: replaces heterogeneity-guided traffic-level augmentation with random masking

Experiments: Ablation study

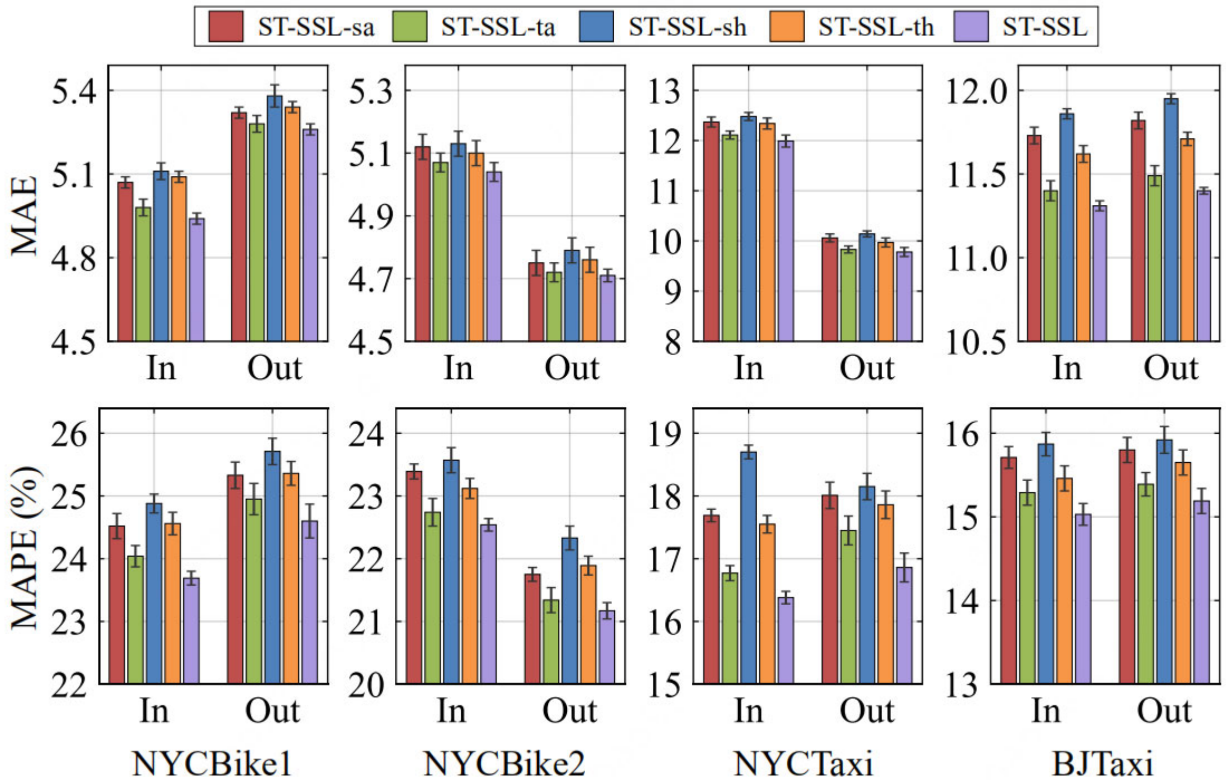
- Ablation study on sub-modules, including
  - Adaptive augmentation: graph topology-level and traffic-level
  - Spatial heterogeneity modeling and temporal heterogeneity modeling

- ST-SSL-sa: replaces heterogeneity-guided structure augmentation on graph topology with random augmentations
- ST-SSL-ta: replaces heterogeneity-guided traffic-level augmentation with random masking
- ST-SSL-sh: removes spatial heterogeneity modeling
- ST-SSL-th: removes temporal heterogeneity modeling

Experiments: Ablation study

- Ablation study on sub-modules, including
  - Adaptive augmentation: graph topology-level and traffic-level
  - Spatial heterogeneity modeling and temporal heterogeneity modeling

- ST-SSL-sa: replaces heterogeneity-guided structure augmentation on graph topology with random augmentations
- ST-SSL-ta: replaces heterogeneity-guided traffic-level augmentation with random masking
- ST-SSL-sh: removes spatial heterogeneity modeling
- ST-SSL-th: removes temporal heterogeneity modeling

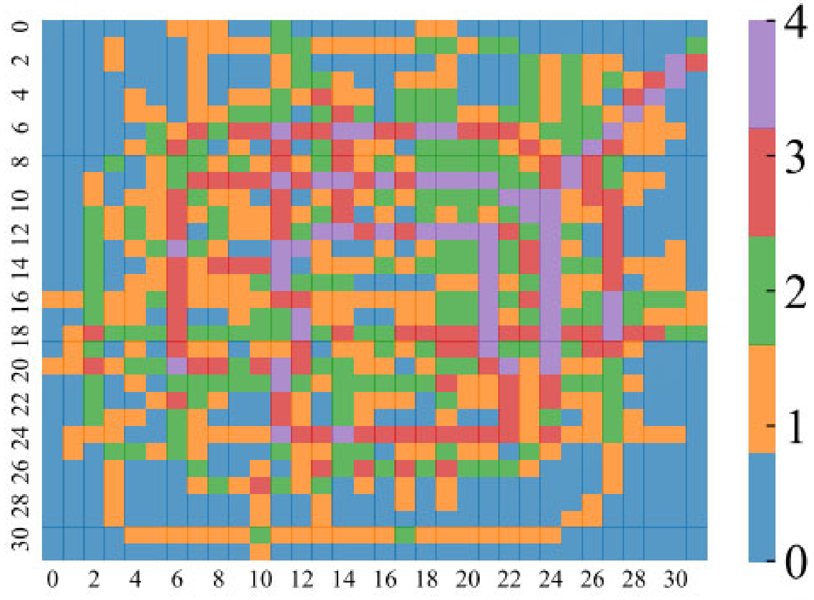


Experiments: Robustness Analysis (1/2)

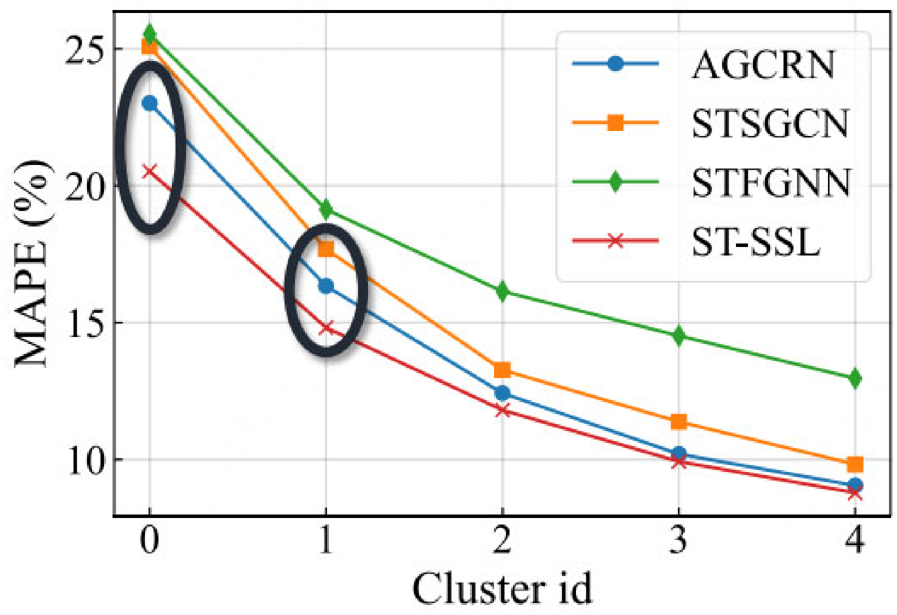
- Traffic prediction for *spatial* regions with heterogeneous data distributions

Experiments: Robustness Analysis (1/2)

- Traffic prediction for *spatial* regions with heterogeneous data distributions



(a) Spatial clusters

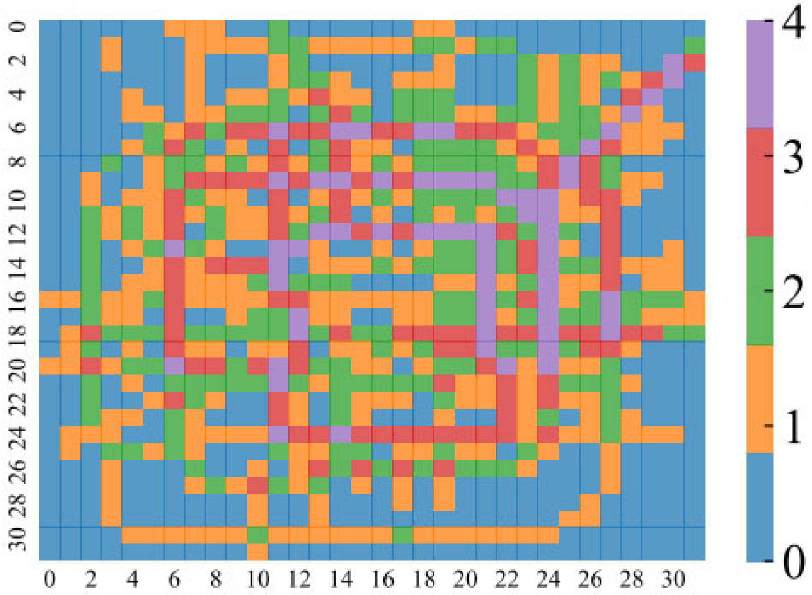


(b) Spatial performance

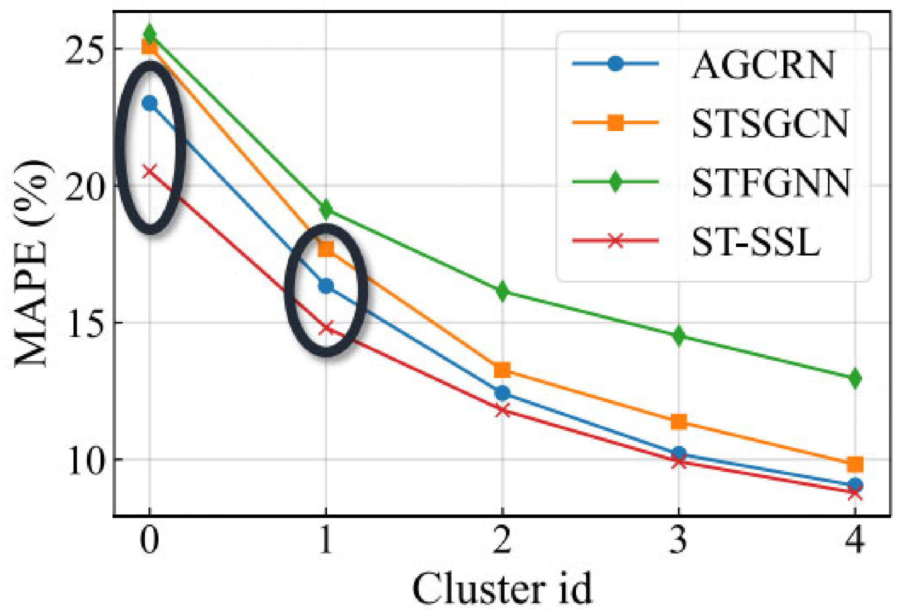
Cluster by: (mean, median, standard deviation)

Experiments: Robustness Analysis (1/2)

- Traffic prediction for *spatial* regions with heterogeneous data distributions
  - ST-SSL surpasses other baselines in different types of spatial regions
  - Particularly for less popular regions (with smaller cluster id)



(a) Spatial clusters



(b) Spatial performance

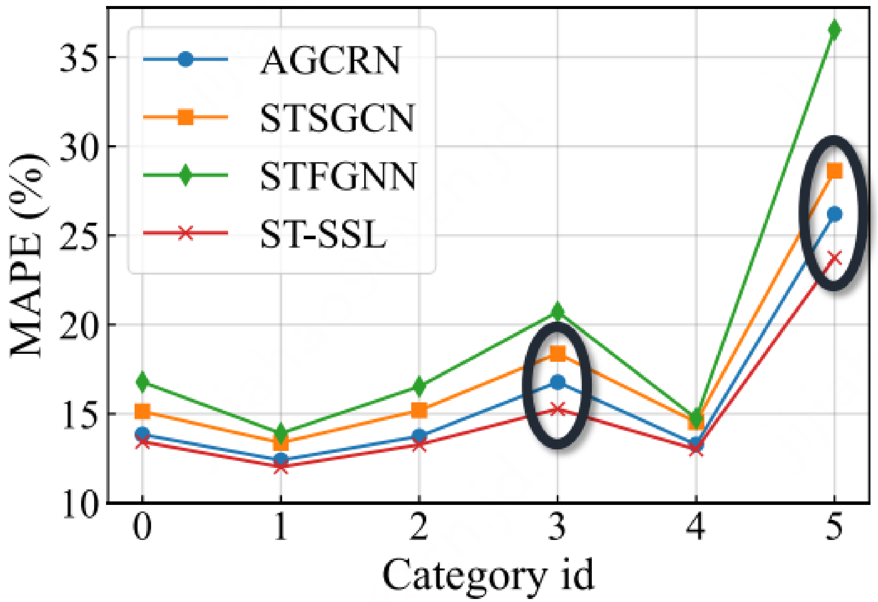
Cluster by: (mean, median, standard deviation)

Experiments: Robustness Analysis (2/2)

- Traffic prediction for *time* periods with different traffic patterns

Day type	Time period	Category (id)
Workday	7:00-10:00	Morning (0)
	10:00-17:00	Regular (1)
	17:00-20:00	Evening (2)
	20:00-7:00	Night (3)
Holiday	9:00-22:00	Day (4)
	22:00-9:00	Night (5)

(c) Temporal categories



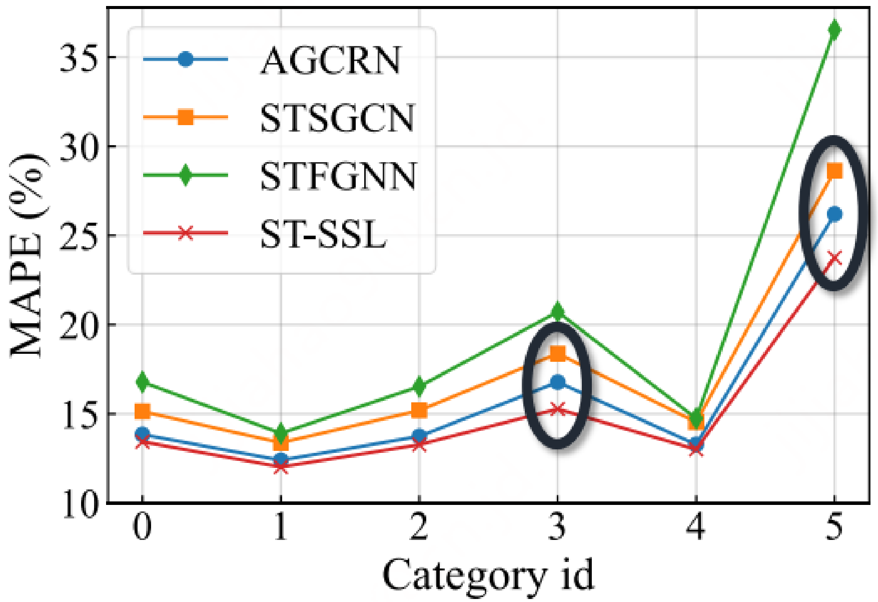
(d) Temporal performance

Experiments: Robustness Analysis (2/2)

- Traffic prediction for *time* periods with different traffic patterns
  - ST-SSL beats the baselines in terms of every temporal category, verifying its robustness
  - ST-SSL shows a significant improvement in categories 3 and 5, during which times traffic flow data are typically sparse

Day type	Time period	Category (id)
Workday	7:00-10:00	Morning (0)
	10:00-17:00	Regular (1)
	17:00-20:00	Evening (2)
	20:00-7:00	Night (3)
Holiday	9:00-22:00	Day (4)
	22:00-9:00	Night (5)

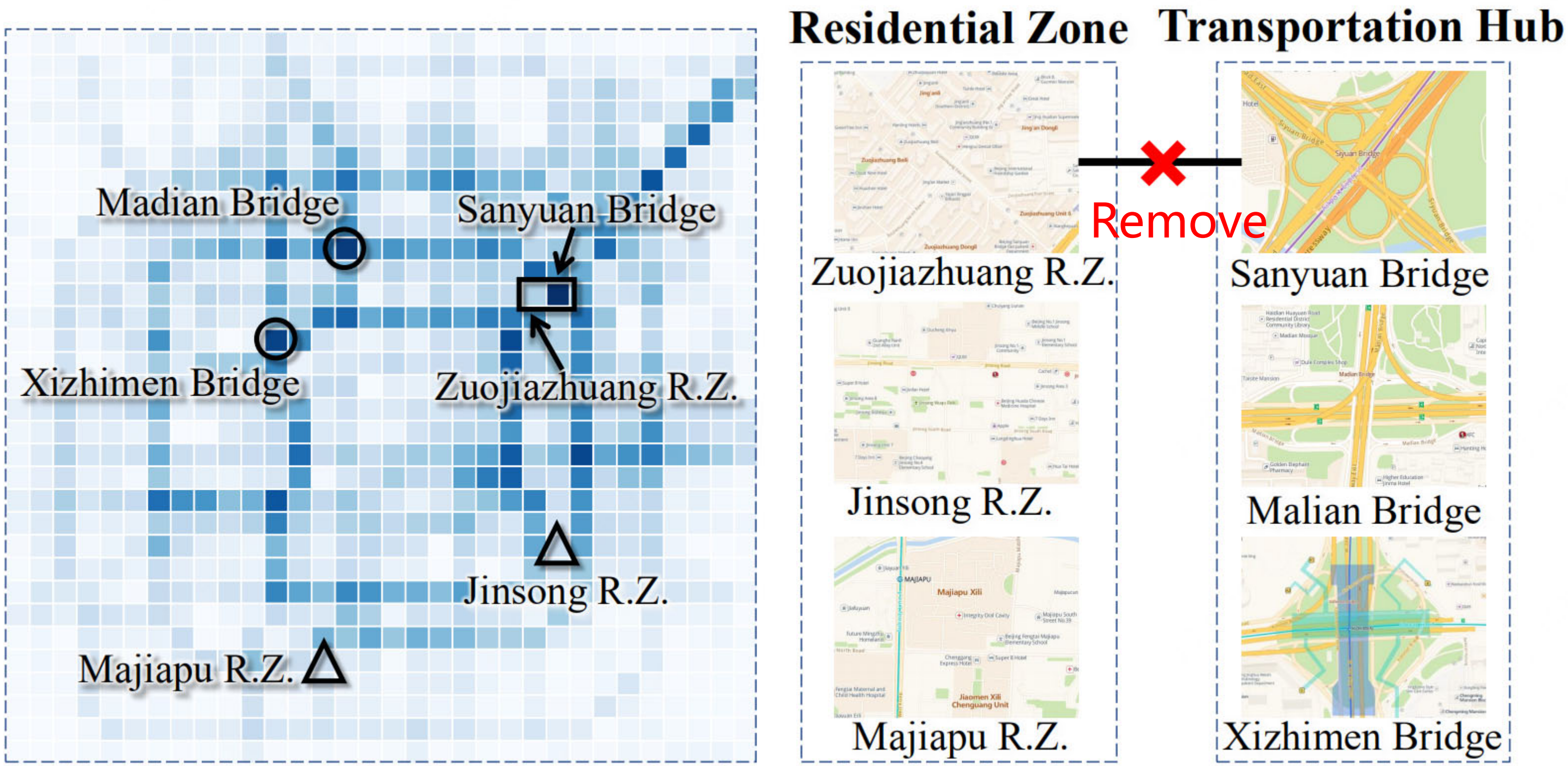
(c) Temporal categories



(d) Temporal performance

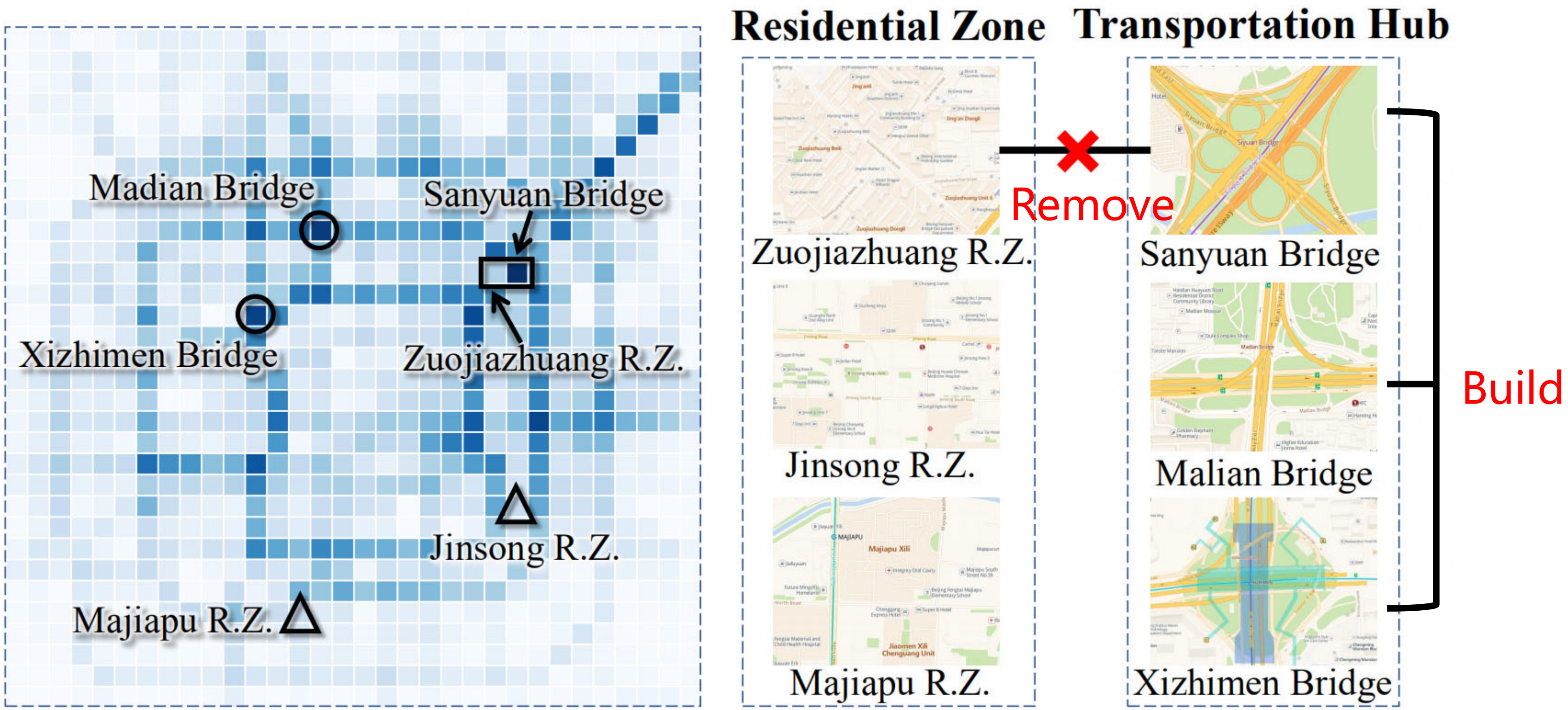
Experiments: Qualitative Study (1/2)

- Investigation on heterogeneity-guided graph topology-level augmentation
  - Remove connections between adjacent regions with heterogeneous traffic patterns



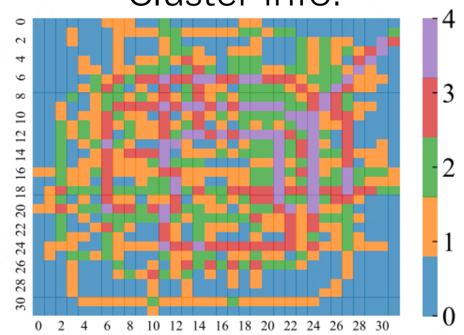
Experiments: Qualitative Study (1/2)

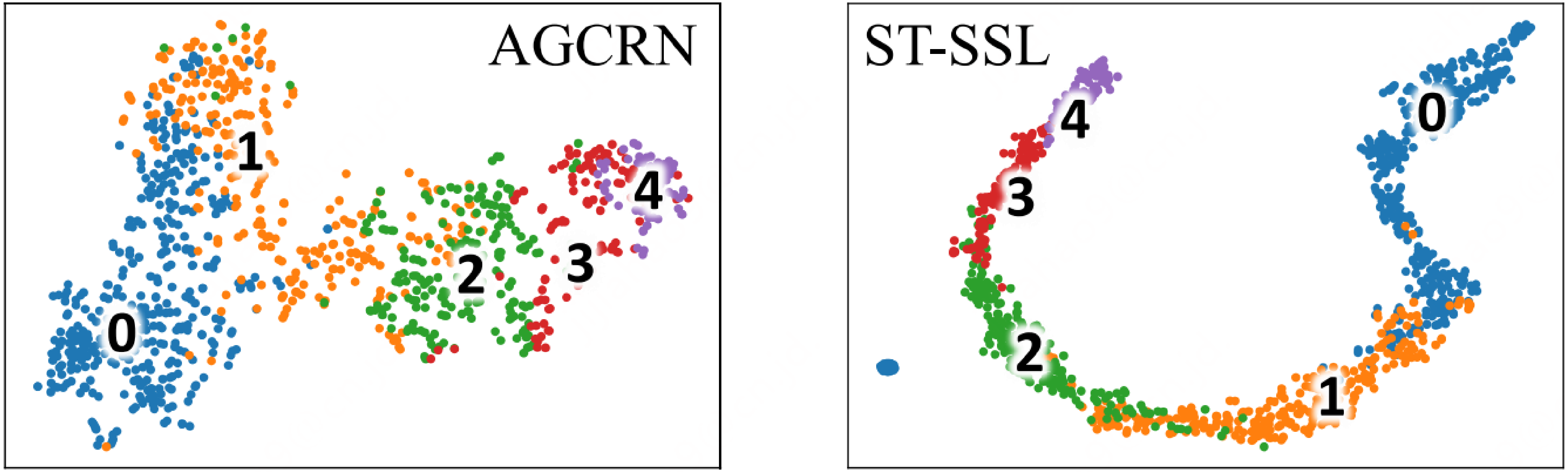
- Investigation on heterogeneity-guided graph topology-level augmentation
  - Remove connections between adjacent regions with heterogeneous traffic patterns
  - Build connections between distant regions with similar latent urban function



Experiments: Qualitative Study (2/2)

- How the learned embeddings benefit the model?



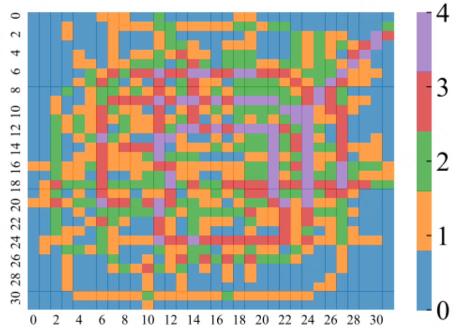


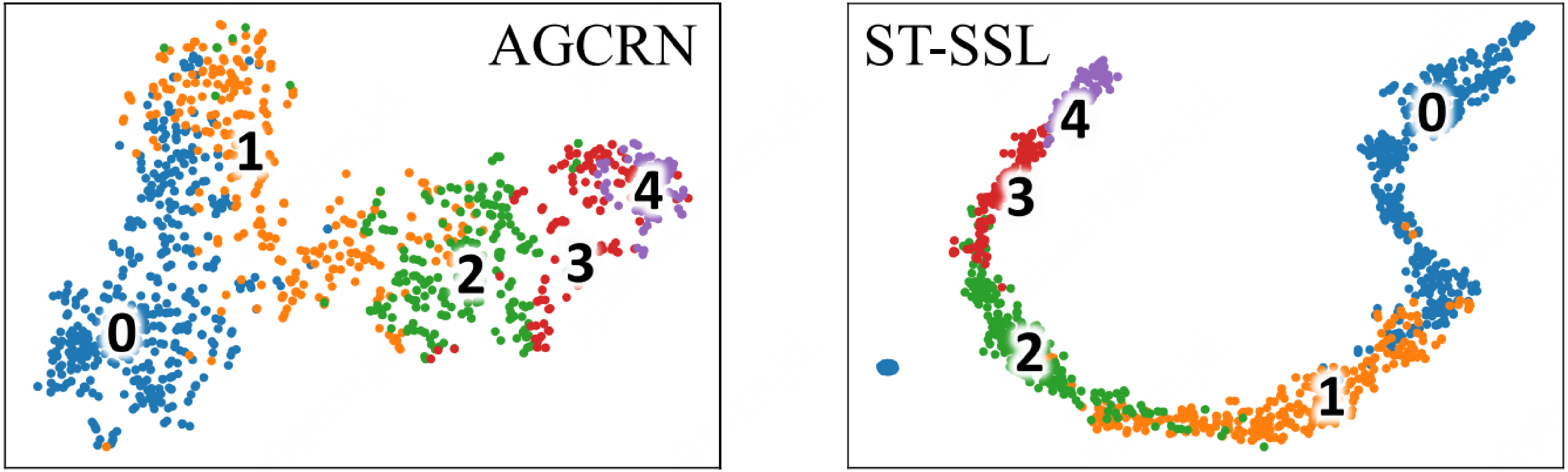
Embedding visualization using t-SNE

Experiments: Qualitative Study (2/2)

- How the learned embeddings benefit the model?
  - Samples are more compact and those of different classes are better separated for ST-SSL

Cluster info.



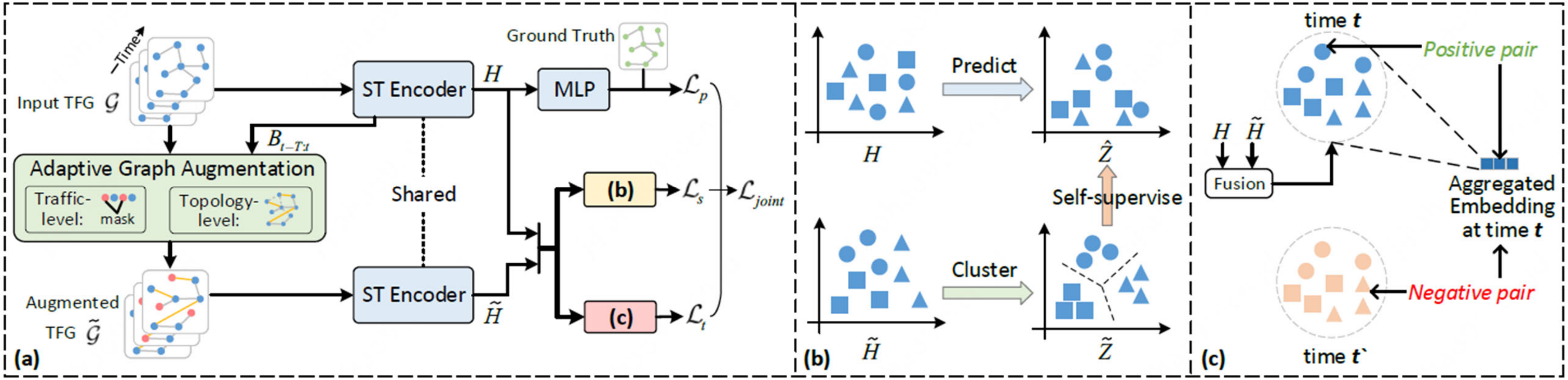


Embedding visualization using t-SNE

Broader Impact

Spatio-Temporal Self-Supervised Learning (**ST-SSL**) for robust traffic prediction

Code: https://github.com/Echo-Ji/ST-SSL



- Provide confidence for the **marriage of SSL and ST prediction**
- Cast light on **other ST applications**, such as air quality prediction
- Can be used as a **new paradigm** for ST prediction in **low-quality data** settings

Thank you!

Paper: https://arxiv.org/abs/2212.04475

Code: https://github.com/Echo-Ji/ST-SSL

Homepage: https://echo-ji.github.io/academicpages/