中国科学技术大学
University of Science and Technology of China

KDD2024
BARCELONA, SPAIN

Kill both Spatial and Temporal shifts with one *STONE*

Binwu Wang, Jiaming Ma, Pengkun Wang, Xu Wang, Yudong Zhang, Zhengyang Zhou, Yang Wang

July 3, 2024

Spatio-temporal OOD prediction problem

◊ Challenges

- ❖ Temporal shift.
- ❖ Spatial shift.

◊ Objective

- ❖ Default prediction:

$$\min_{\mathcal{F}} \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim P(\boldsymbol{x},\boldsymbol{y}|e)}[\mathcal{L}(\mathcal{F}(\mathbf{x},\mathcal{G}),\mathbf{y})]$$

- ❖ OOD prediction:

$$\min_{\mathcal{F}} \max_{e^*\in E} \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim P(\boldsymbol{x},\boldsymbol{y}|e^*)}[\mathcal{L}(\mathcal{F}(\mathbf{x},\mathcal{G}),\mathbf{y})]$$

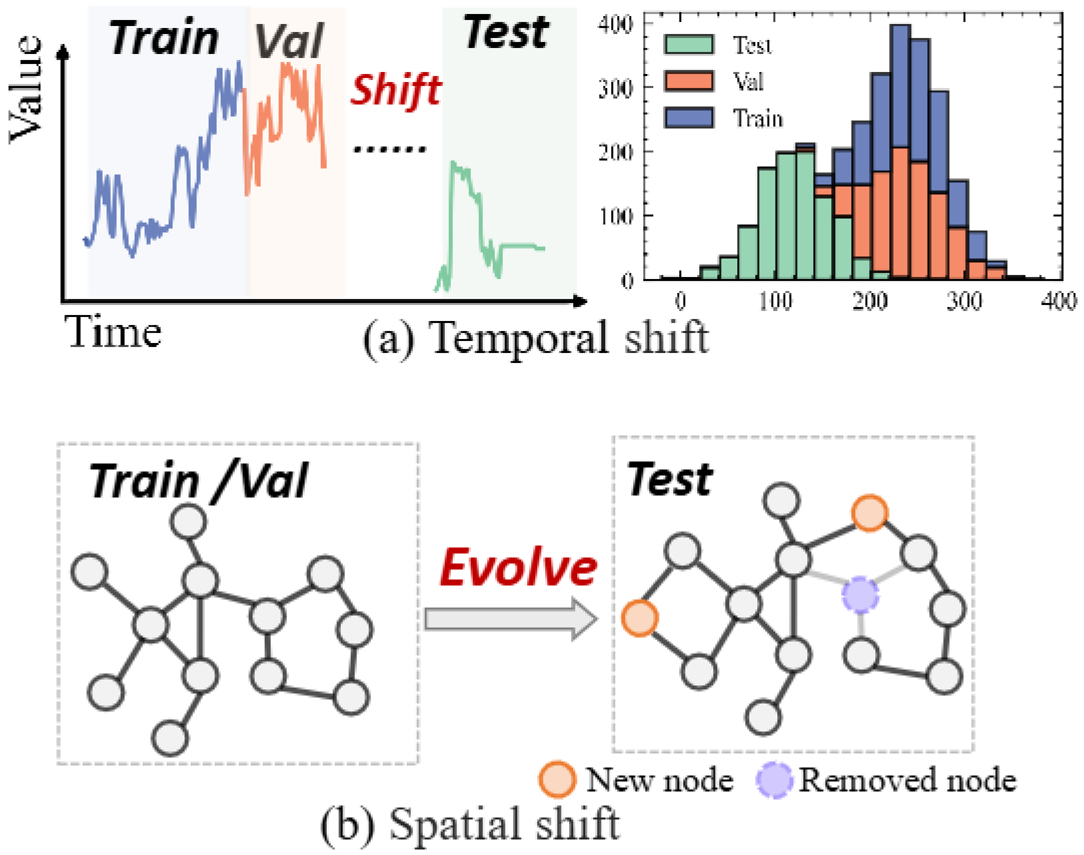


Figure 1. Visualization of temporal (a) and spatial (b) shifts.

Spatio-temporal environments shifts

◊ Portray and perceive

❖ Semantic graph: A suitable metric is utilized to portray the temporal or spatial similarity between nodes at the current time.

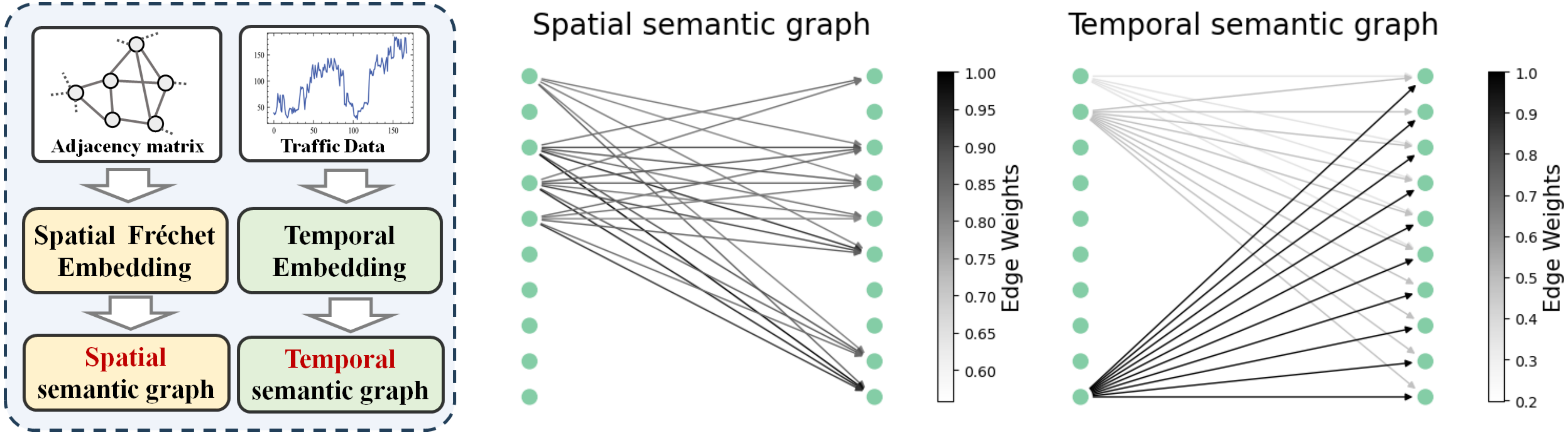


Figure 2. Spatial and Temporal semantic graphs.

Spatio-temporal environments shifts

◊ The case of traffic prediction

- ❖ Node #1 ~ #5 in SD.
- ❖ Similarity of DTW distances based on 12 time steps.

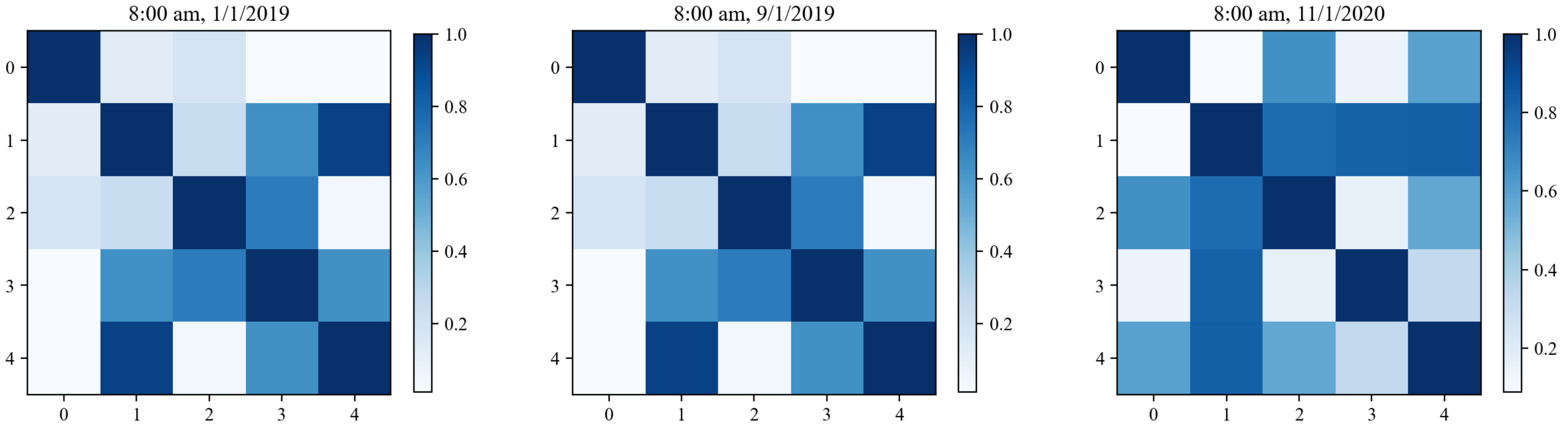


Figure 3. Case study of #1 to #5 sensors in SD datasets.

Spatial Side Information

◊ Bourgain's Theorem (1985)

❖ If (X, d) is an N-point metric space and f is an Fréchet embedding, then

$$\frac{1}{\mathcal{O}(\log N)} d(x,y) \leq \mathbb{E}_f \|f(x) - f(y)\| \leq d(x,y), \forall x, y \in X.$$

◊ Spatial Fréchet Embedding

❖ Dimension is not affected by the addition of new nodes.

❖ Solid Theory.

❖ Based on Hamming distance. Low complexity.

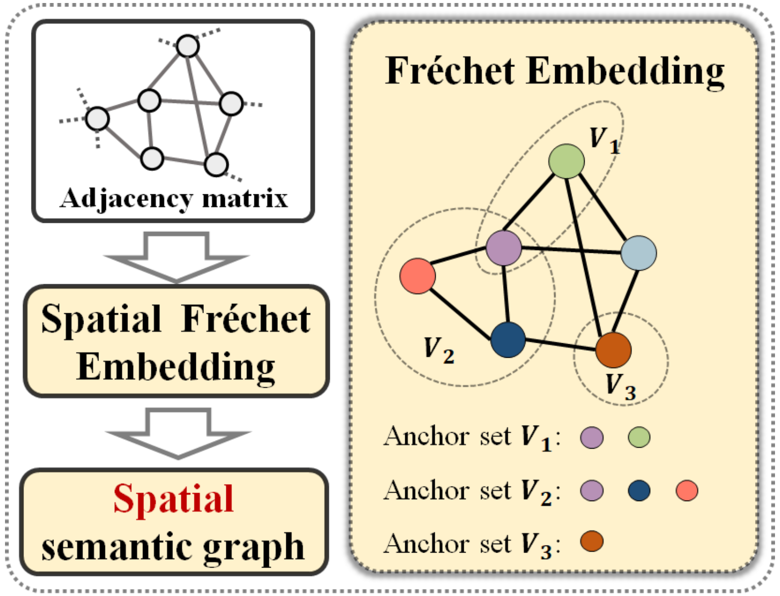


Figure 4. Fréchet Embedding

Environment Generators

◊ Optimized objects for multiple environments

❖ Mask operators $\mathbb{M}$ for disturbing spatio-temporal environment in training phase.

$$\min_{\Theta} \mathrm{Var}\{\mathcal{L}_{\Theta}(\mathcal{F}(\mathbf{x},\mathcal{G}),\mathbf{y}|\mathbb{M}^{*},\Theta)\} + \beta\mathbb{E}[\mathcal{L}_{\Theta}(\mathcal{F}(\mathbf{x},\mathcal{G}),\mathbf{y}|\mathbb{M}^{*})]\,,$$

$$\text{s.t.} \quad \mathbb{M}^{*} = \left(\mathbb{M}_1^{*},\mathbb{M}_2^{*},\ldots,\mathbb{M}_{K_M}^{*}\right) = \underset{\mathbb{M}_m \in \{0,1\}^{N\times N}}{\operatorname{argmax}} \mathrm{Var}\{\mathcal{L}_{\Theta}(\mathcal{F}(\mathbf{x},\mathcal{G}),\mathbf{y}|\mathbb{M}^{\boldsymbol{m}},\Theta)\}\,.$$

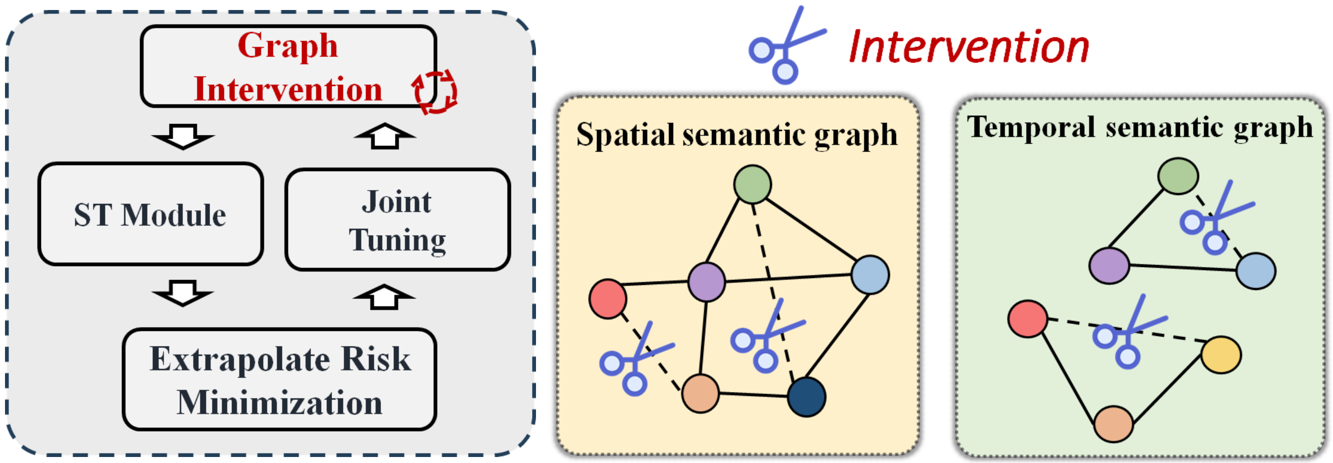
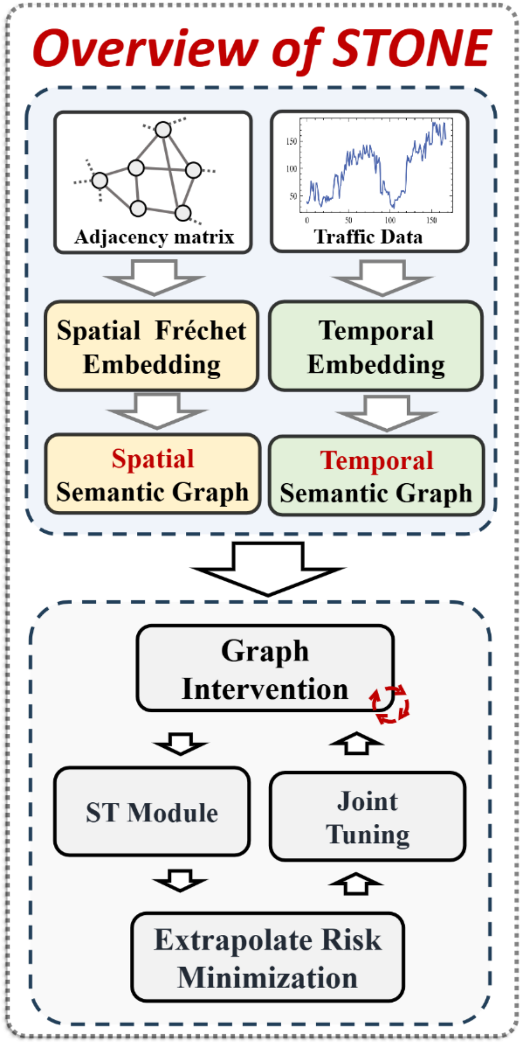
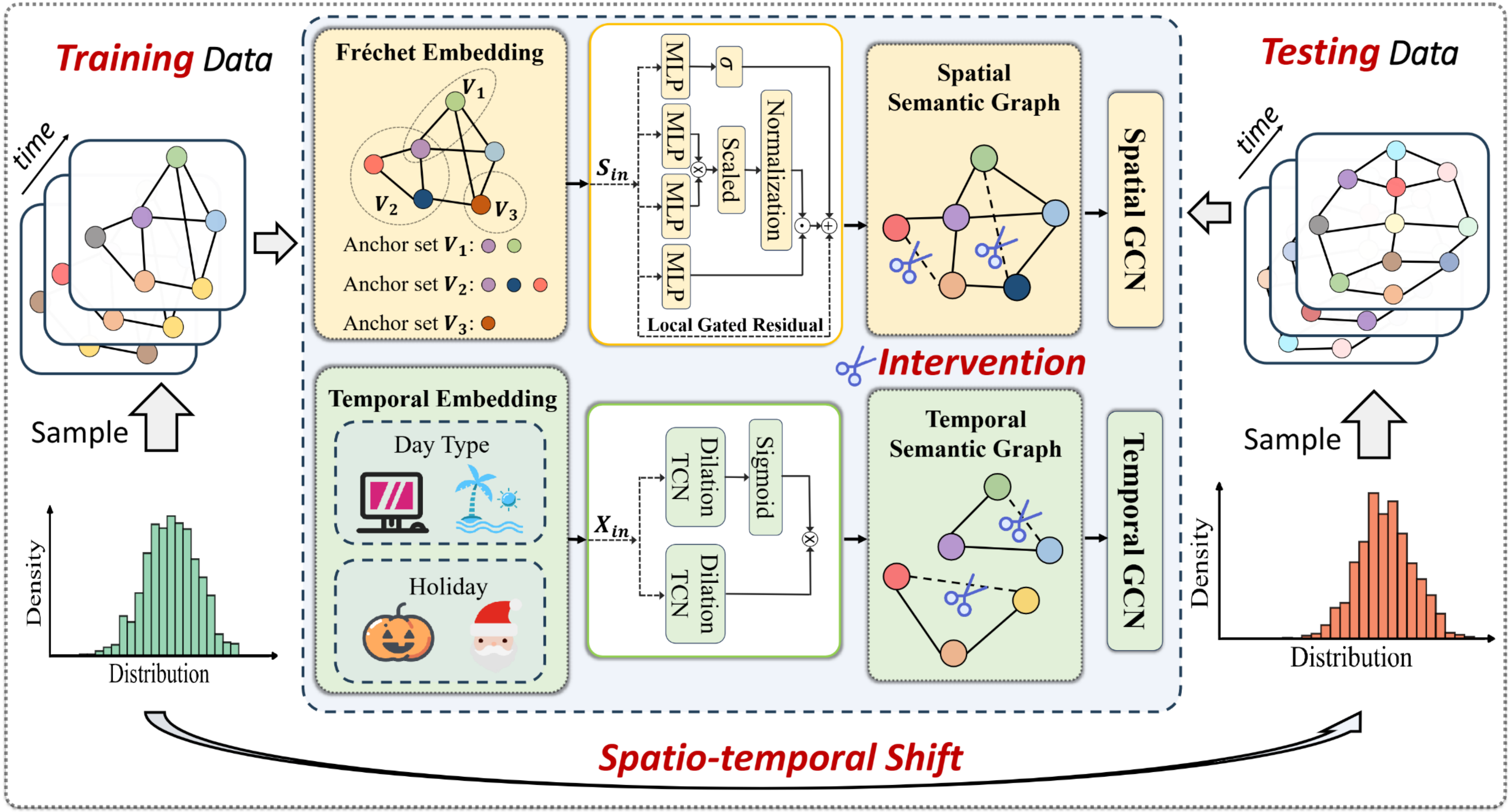


Figure 5. Mask operators and optimization.

STONE

❖ Spatio-Temporal OOD Graph Learning Networks with Fréchet Embedding





Experiments for STONE

◊ Datasets: SD and GBA in LargeST under OOD setting

- [1/1/2019, 8/31/2019] for training,
  [9/1/2019, 10/31/2019] for validation,
  [11/1/2020, 12/31/2020] for test, etc.
- Vertices increases by **5%/10%/15%** and decreases by **5**% in validation and test sets.
- STONE exhibits a relative improvement of up to nearly **20%**.
- Result details in paper.

中国科学技术大学
University of Science and Technology of China

KDD2024
BARCELONA, SPAIN

Thanks,
See you Barcelona!