Spatio-temporal statistics

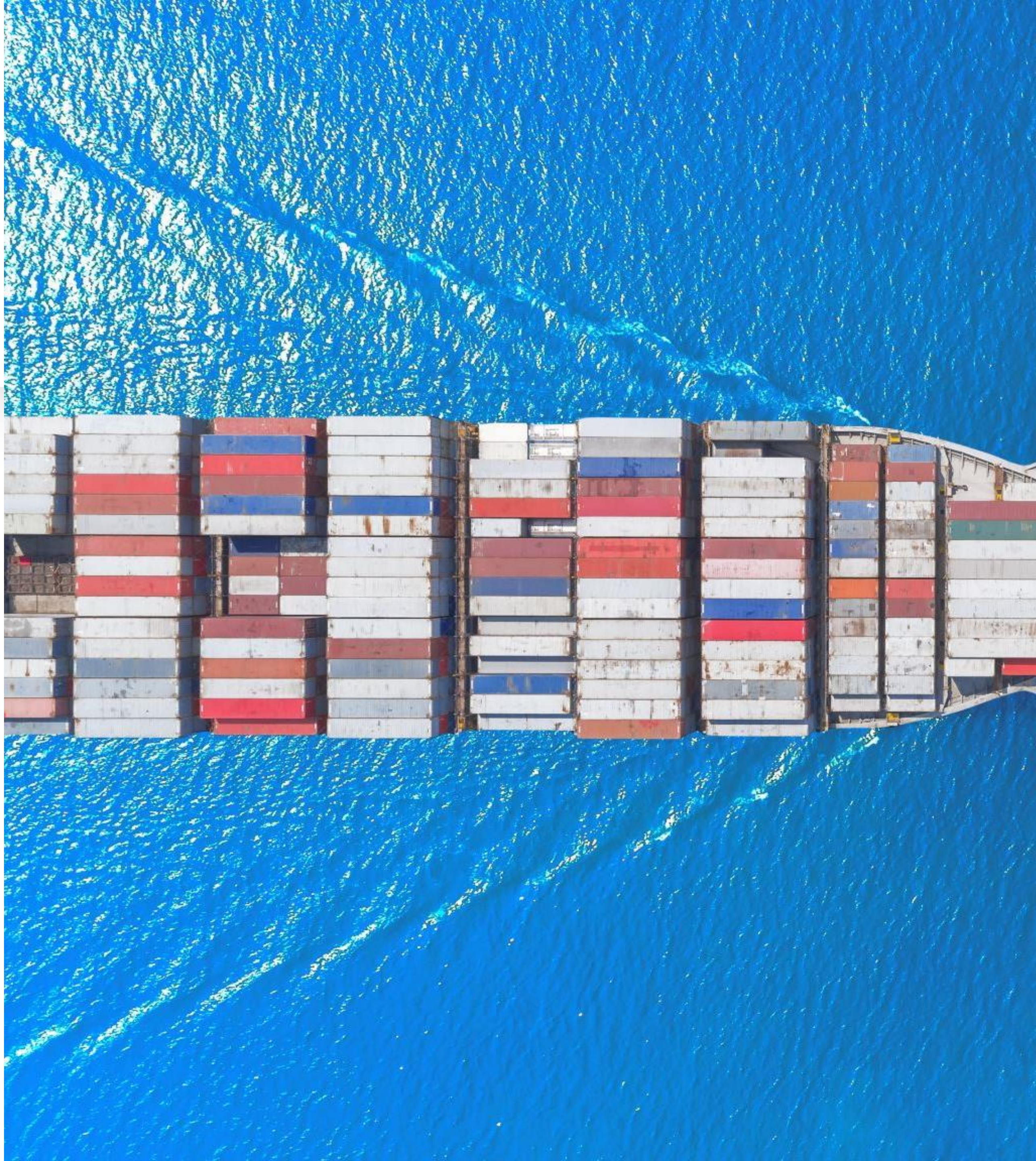
Anupam Sobti

Bibliography

- Shi, Zhicheng, and Lilian SC Pun-Cheng. "Spatiotemporal data clustering: A survey of methods." *ISPRS international journal of geo-information* 8, no. 3 (2019): 112.
- And more papers inline

Sources of spatio-temporal data

- Transportation
- Crime
- Social media
- Government censuses
  - Population data
  - Human movement
  - Economic characteristics of different periods

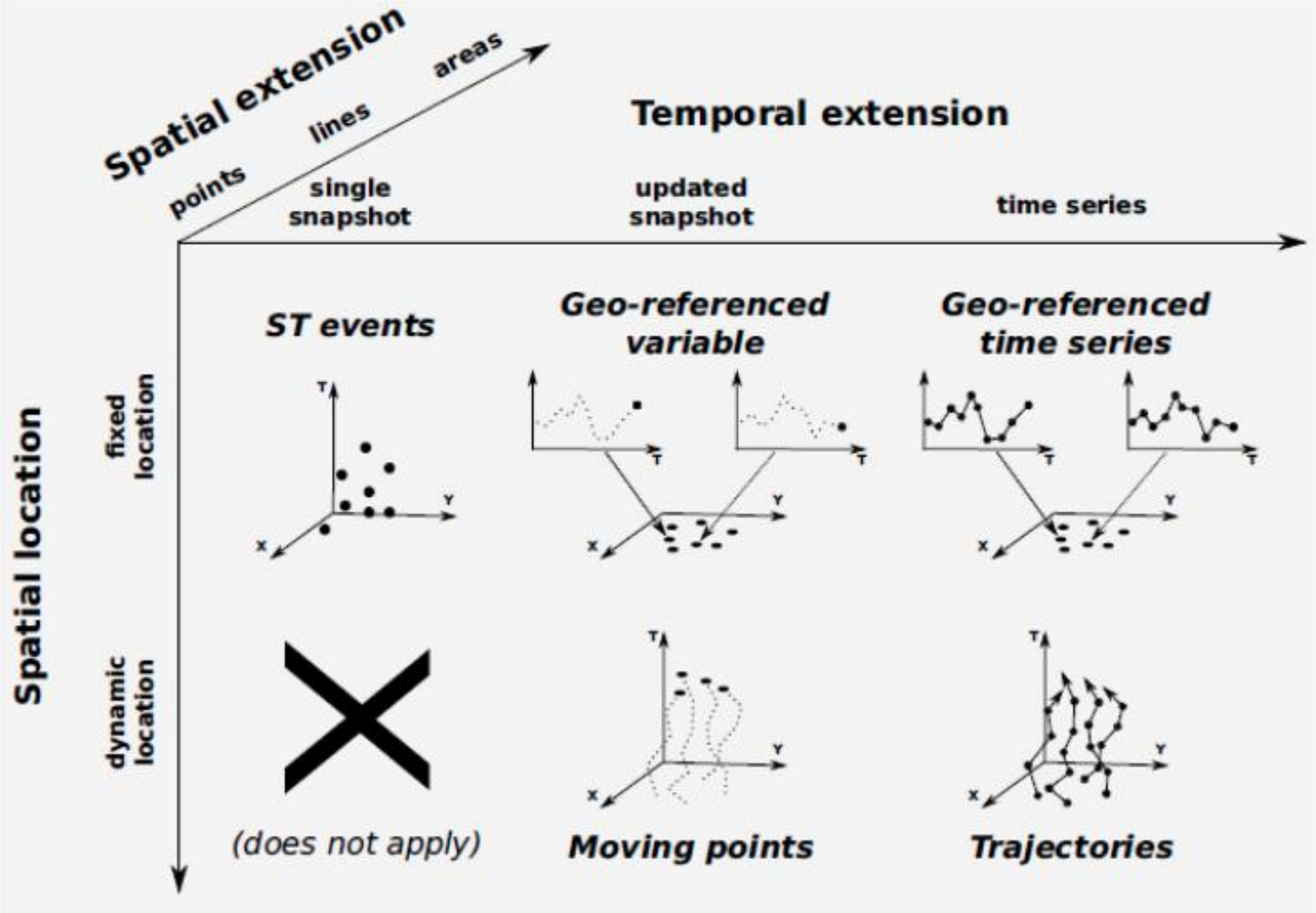


Types of spatio-temporal data

Lightning strike recorded at a weather station

Daily air temperature at weather station

10 year rainfall data at a certain location



GPS location of delivery trucks at one time

Food delivery updated snapshots in a city

Flight paths of airplanes (continuous time)

Understanding data

- Clustering
  - Distance functions – Euclidian, Manhattan, etc.?
- Prediction
- Change detection
- Frequent pattern mining
- Anomaly detection
- Relationship mining

The time dimension

- Is time just another dimension?
- Would clusters change when time scale changes?
  - Relative scale of distance and time changes?
  - Do you want time or another feature to dominate clustering?

Reminder: Spatial clustering

- K-means
- Hierarchial clustering methods
  - Ward’s hierarchical method
- DB-SCAN (for differently shaped clusters)

Including spatial weights for spatial constraints

Types of spatio-temporal clustering methods

- Hypothesis testing based
- Partition based

Hypothesis testing

Knox’s test

Alpha and beta are set manually (empirically)

1. **Spatial adjacency**:

$$d_{ij} = \begin{cases} 1, & \text{if distance between } i \text{ and } j < \alpha \\ 0, & \text{otherwise} \end{cases}$$

---

2. **Temporal adjacency**:

$$t_{ij} = \begin{cases} 1, & \text{if time difference between } i \text{ and } j < \beta \\ 0, & \text{otherwise} \end{cases}$$

---

3. **Total number of spatiotemporal adjacent pairs**:

$$K = \sum_{i=1}^{N} \sum_{j=1}^{i-1} d_{ij} t_{ij}$$

Knox, E.; Bartlett, M. The detection of space-time interactions. J. R. Stat. Soc. Ser. C (Appl. Stat.) 1964, 13, 25–30. [CrossRef]

Mantel's test

- Avoid the hyperparameters alpha and beta

$$Z = \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij}^{s} d_{ij}^{t}$$

$$M = \frac{1}{(N^2 - N - 1)} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\left(d_{ij}^{s} - \overline{d^{s}}\right)}{s_s} \frac{\left(d_{ij}^{t} - \overline{d^{t}}\right)}{s_t}$$

s_s and s_t are the standard deviations of data in space and time, respectively.

Jacques's test

- Spatiotemporal k-nearest neighbors test

N: Number of cases.

d_{ij}: Spatial measure, when $d_{ij} = 1$ case j is a k nearest neighbor of case i in space, otherwise equal to 0.

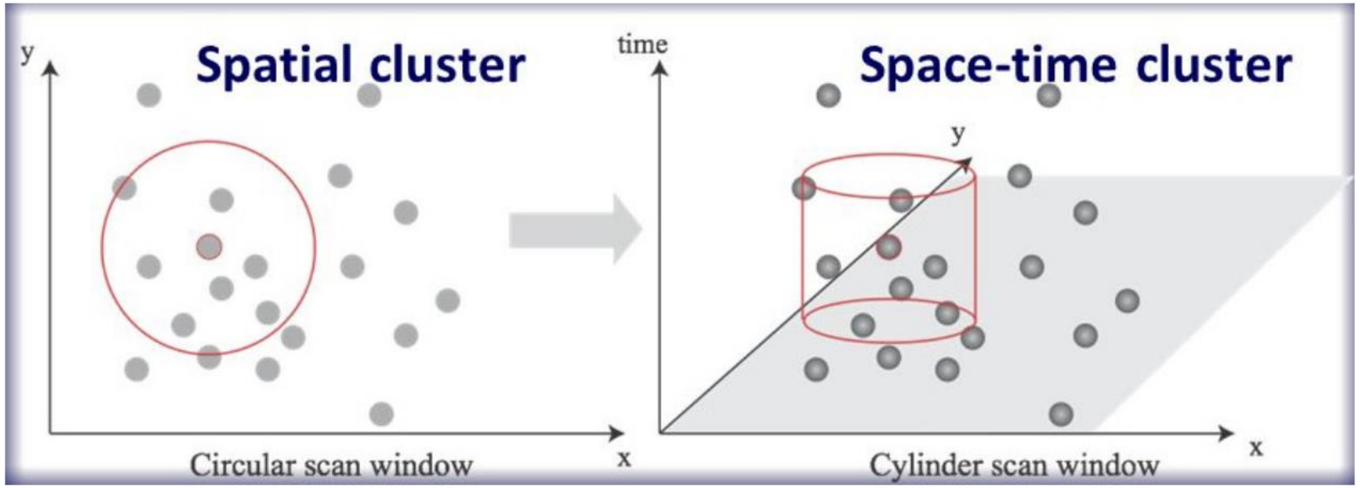
t_{ij}: Spatial measure, when $t_{ij} = 1$ case j is a k nearest neighbor of case i in time, otherwise equal to 0.

D_k: Is a cumulative test statistic, where $D_k = \sum_{i=1}^{N} \sum_{j=1}^{N} d_{ij} t_{ij}$.

ΔD_k: Is k– specific test statistic, where $\Delta D_k = D_k - D_{k-1}$.

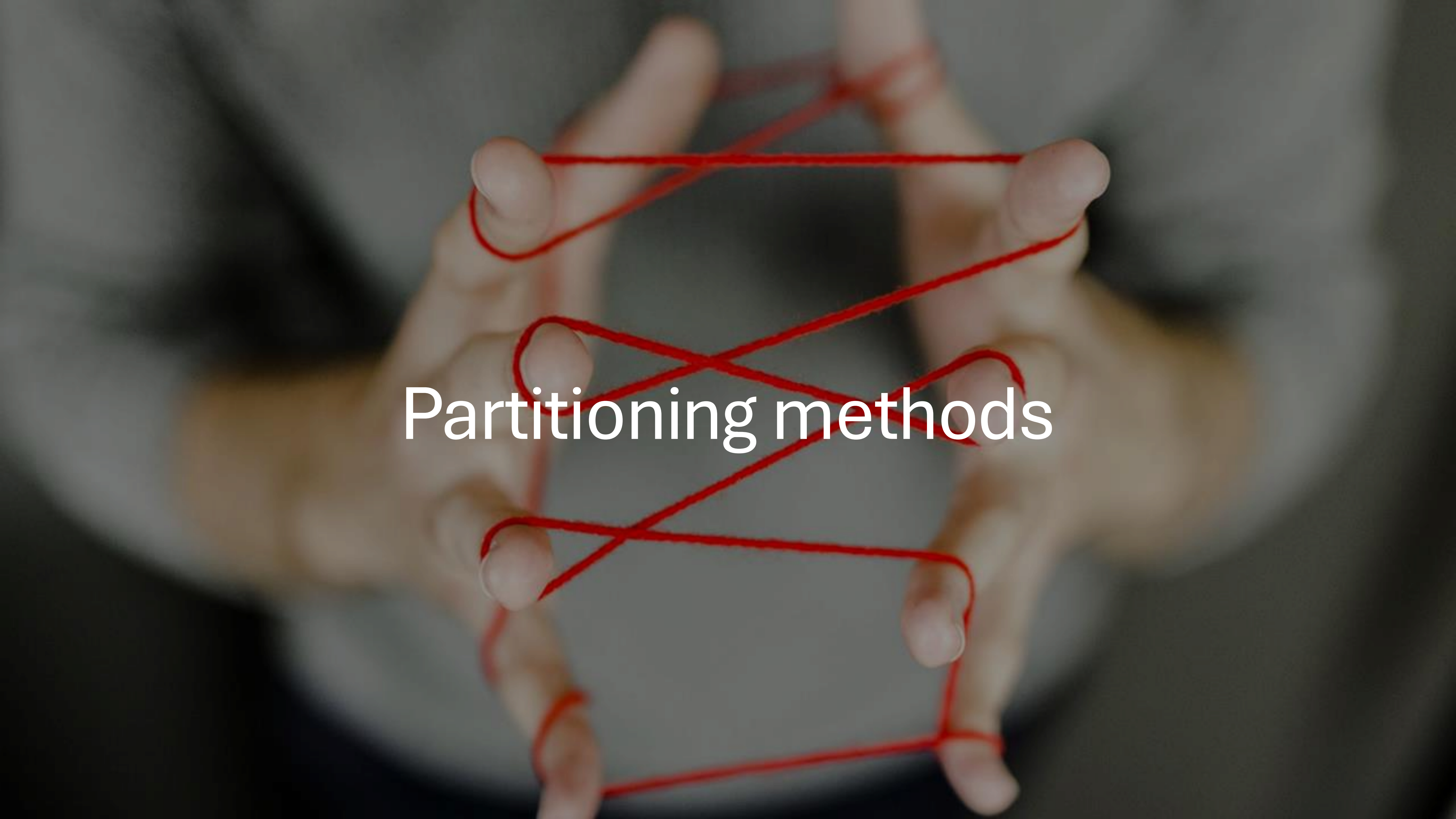
Scan Statistics

$$S = log\left(\frac{n_z}{u_z}\right)^{n_z}\left(\frac{N-n_z}{N-u_z}\right)^{(N-n_z)} I\left(\frac{n_z}{u_z} > \frac{N-n_z}{N-u_z}\right)$$



Nz = observed points in the cylinder
Uz = expected points in the cylinder

Reminder, null hypothesis is that points are distributed randomly. The log likelihood above shows a clustering in the data



Partitioning methods

ST-DBSCAN

- Added a temporal radius to DBSCAN
  - Used a k-dist graph to determine noisy points vs dense clusters

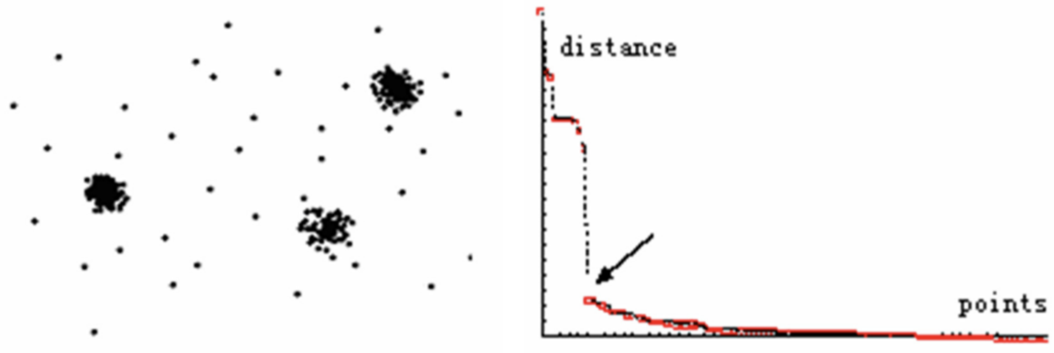
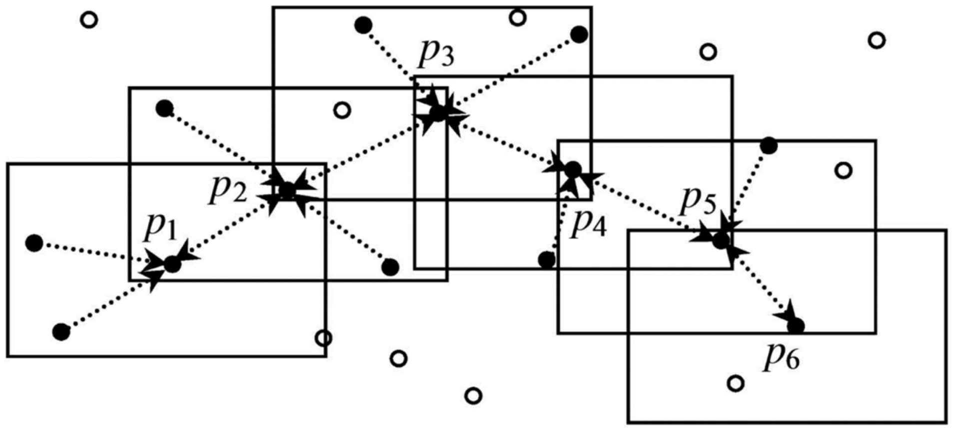


Figure 4. A point set with its sorted 4-dist graph (source: Wang, et al. [49]).

Other methods

- Versions of
  - KDE
  - Windowed nearest neighbor

That add a time dimension



Other common statistical approaches

- STARIMA (Space-Time Arima)
  - Economics and epidemiology
- Gaussian process regression can be quantified in uncertainty using spatio-temporal covariance
  - Great for sparse data modeling, e.g., pollution
- Bayesian hierarchical models
  - Disease epidemiology where priors and interpretability is important
- Point process models
  - Models where probabilities of events nearby (in space and time) get increased based on an event, e.g., Hawkes process

ML/DL Algorithms

GCN → STGCN

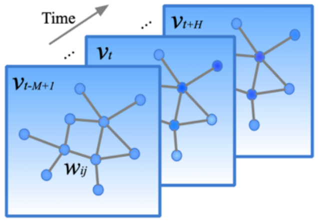


Figure 1: Graph-structured traffic data. Each v_t indicates a frame of current traffic status at time step t, which is recorded in a graph-structured data matrix.

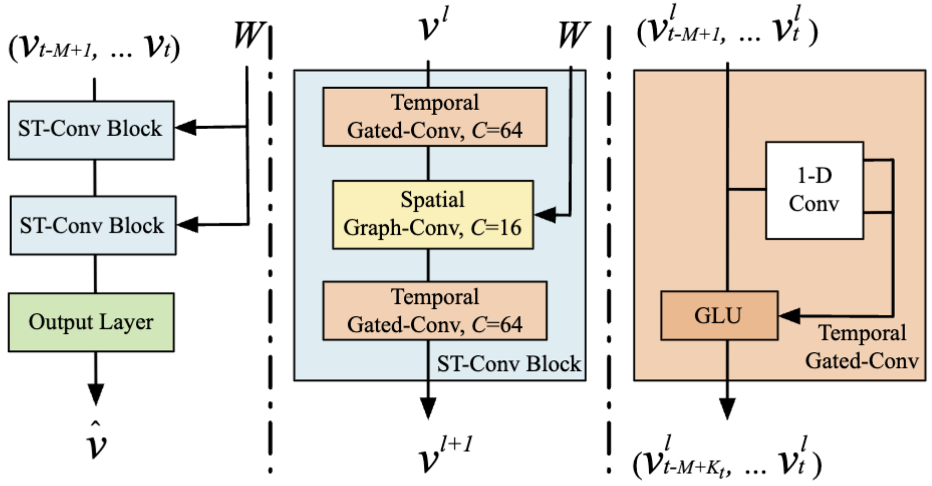
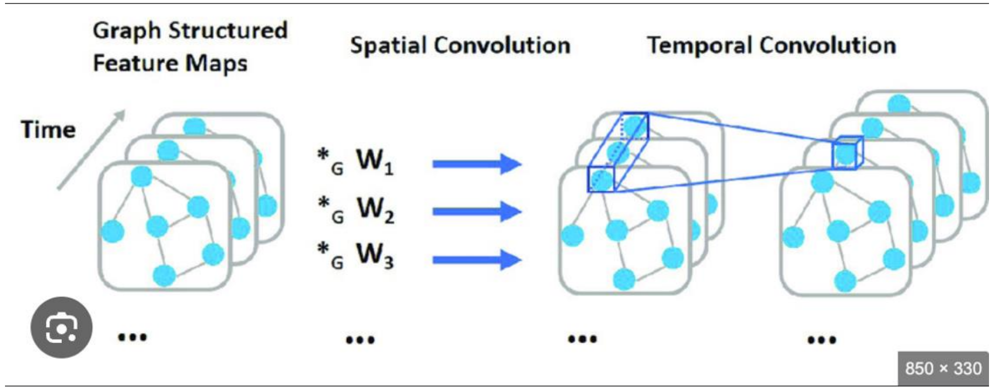


Figure 2: Architecture of spatio-temporal graph convolutional networks. The framework STGCN consists of two spatio-temporal convolutional blocks (ST-Conv blocks) and a fully-connected output layer in the end. Each ST-Conv block contains two temporal gated convolution layers and one spatial graph convolution layer in the middle. The residual connection and bottleneck strategy are applied inside each block. The input $v_{t-M+1}, \ldots, v_t$ is uniformly processed by ST-Conv blocks to explore spatial and temporal dependencies coherently. Comprehensive features are integrated by an output layer to generate the final prediction $\hat{v}$.



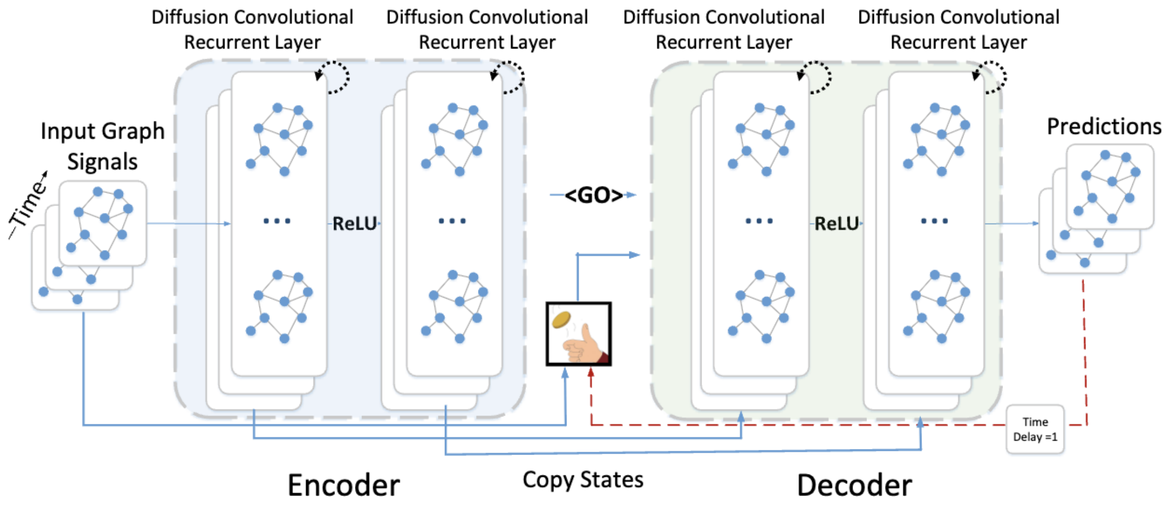
$$L(\hat{v}; W_\theta) = \sum_t ||\hat{v}(v_{t-M+1}, \ldots, v_t, W_\theta) - v_{t+1}||^2,$$

Loss function

Yu, Bing, Haoteng Yin, and Zhanxing Zhu. "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting." *arXiv preprint arXiv:1709.04875* (2017).

DCRNN – Diffusion Convolution Recurrent Neural network

Note, here Diffusion is diffusion of information on the graph



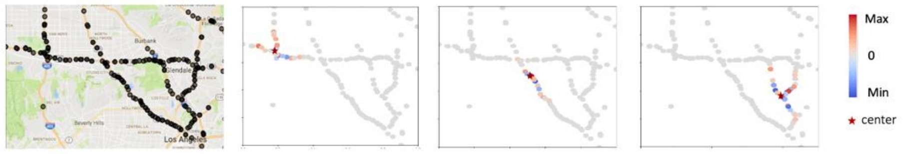


Figure 7: Visualization of learned localized filters centered at different nodes with $K = 3$ on the METR-LA dataset. The star denotes the center, and the colors represent the weights. We observe that weights are localized around the center, and diffuse alongside the road network.

Li, Yaguang, Rose Yu, Cyrus Shahabi, and Yan Liu. "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting." *arXiv preprint arXiv:1707.01926* (2017).

SOTA: STGFormer

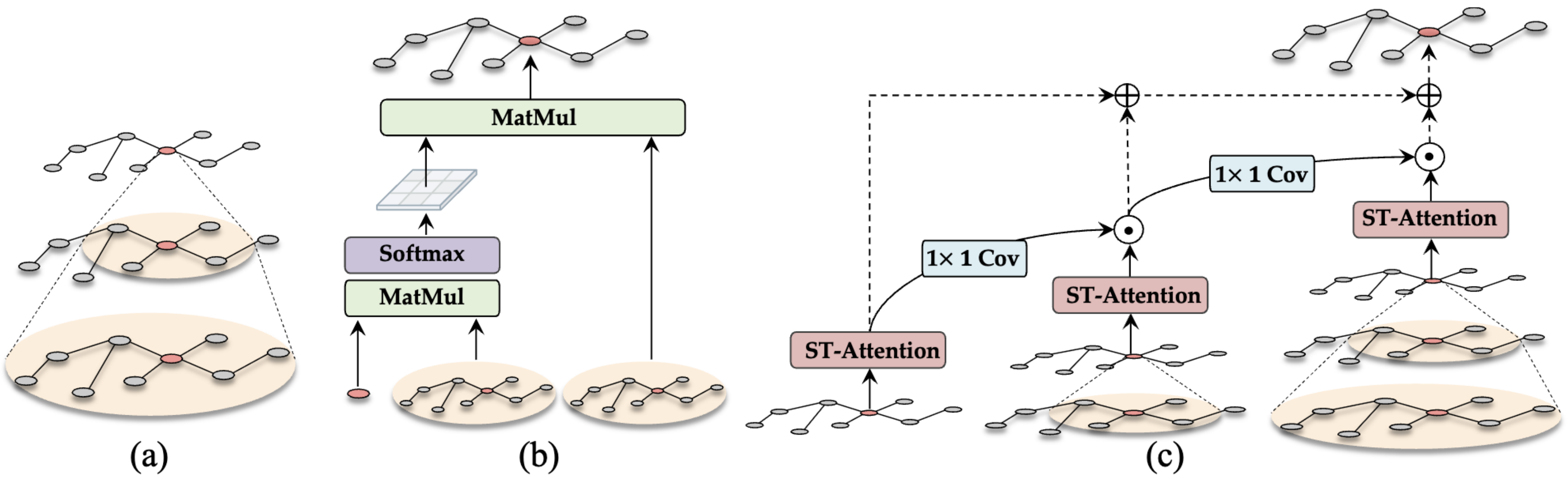


Fig. 3. **Illustration the main idea of STGformer.** We illustrate typical spatial modeling in graphs, highlighting various levels of interactions and receptive fields. (a) Conventional Graph Convolution [9], [11] explicitly handles arbitrary-order spatial interaction but within a limited receptive area. (b) Spatiotemporal Transformer [5], [6] performs interactions up to two orders through two consecutive matrix multiplications, covering a broad receptive area. (c) Our STGformer achieves high-order global spatial interactions with structure information by integrating graph convolutional networks with Transformer architecture.

Wang, Hongjun, Jiyuan Chen, Tong Pan, Zheng Dong, Lingyu Zhang, Renhe Jiang, and Xuan Song. "STGformer: Efficient Spatiotemporal Graph Transformer for Traffic Forecasting." *arXiv preprint arXiv:2410.00385* (2024).

Innovations in handling large scale data

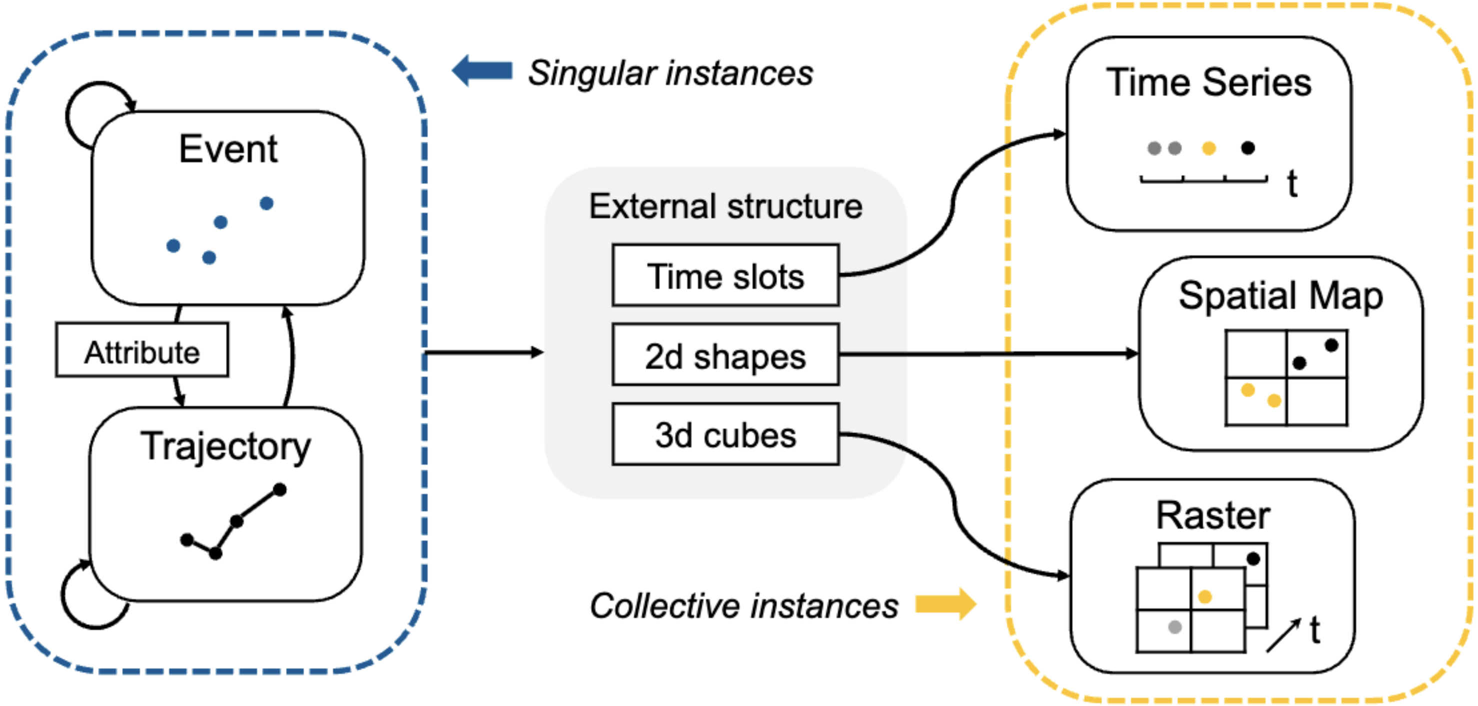
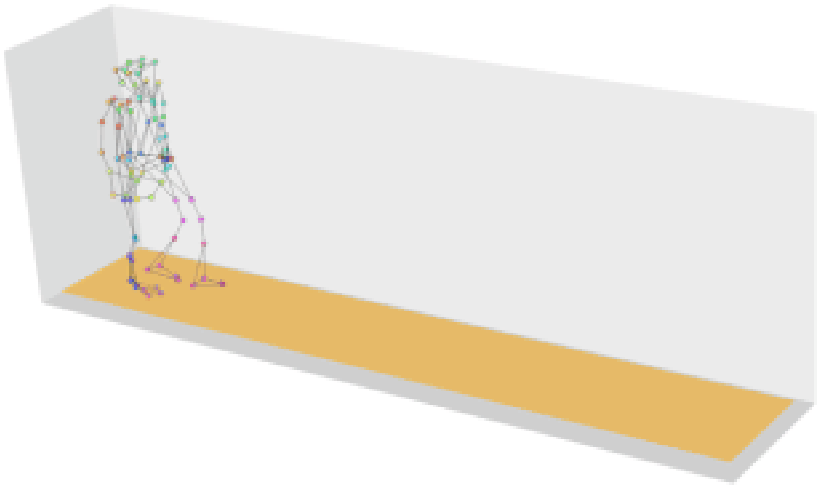


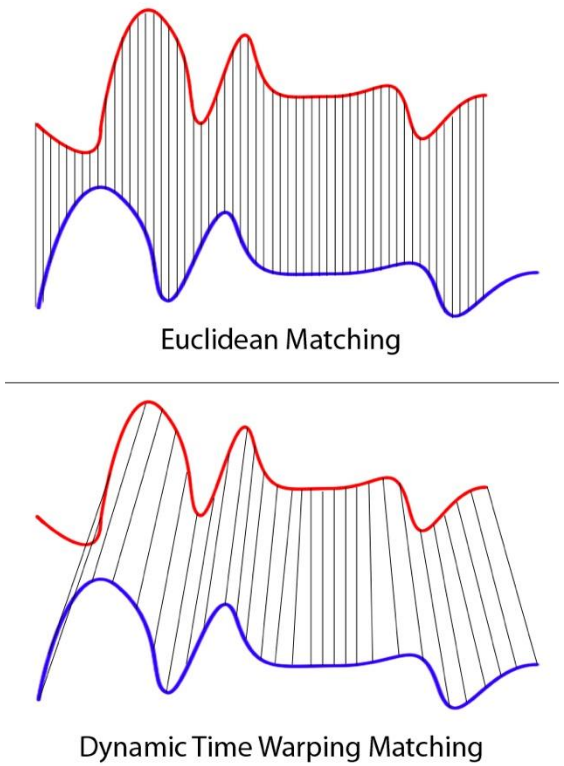
Fig. 3. Conversions between ST instances.

Liu, Kaiqi, Panrong Tong, Mo Li, Yue Wu, and Jianqiang Huang. "St4ml: Machine learning oriented spatio-temporal data processing at scale." *Proceedings of the ACM on Management of Data* 1, no. 1 (2023): 1-28.

DTWNet (Dynamic Time warping)

Something to try with 'not so large' data





Cai, Xingyu, Tingyang Xu, Jinfeng Yi, Junzhou Huang, and Sanguthevar Rajasekaran. "DTWNet: A dynamic time warping network." *Advances in neural information processing systems* 32 (2019).

Some example areas

Agriculture and Environment monitoring

- Drone data collected over the crop cycle for yield prediction
- Spatio-temporal anomalies
  - Crop stress and Drought monitoring
- Fourcastnet
  - Graphcast (available to run on your PC)

Yu, Manzhu, Qunying Huang, and Zhenlong Li. "Deep learning for spatiotemporal forecasting in Earth system science: a review." *International Journal of Digital Earth* 17, no. 1 (2024): 2391952.

Urban Planning

- Urban sprawl as spatio temporal regression
  - Gómez, Jairo A., Jorge E. Patiño, Juan C. Duque, and Santiago Passos. "Spatiotemporal modeling of urban growth using machine learning." *Remote Sensing* 12, no. 1 (2019): 109.
  - Population growth modeling
  - Urban land use modeling
- Urban mobility and infrastructure
  - Gariazzo, Claudio, Armando Pelliccioni, and Maria Paola Bogliolo. "Spatiotemporal analysis of urban mobility using **aggregate mobile phone derived presence** and demographic data: A case study in the city of rome, italy." *Data* 4, no. 1 (2019): 8.
  - Classify residential/commercial areas based on their diurnal population patterns
- Urban anomalies and events
  - Events in the city that could cause Traffic changes (e.g., Diljit Dosanjh concert)
  - Diagnosing a viral outbreak in the city

Urban planning contd

- ST models to predict city demand for better demand response modeling and grid adjustments
- Garbage collection or sewage cleaning routes
  - Xu, Shixiong, Sara Shirowzhan, and Samad Sepasgozar. "Spatiotemporal analysis and GIS-based dashboard development for urban household waste." *Smart and Sustainable Built Environment* (2025).
- Traffic flow and demand (DCRNN)
  - Public transit ridership
  - Uber eater/rider matching [link]
  - Identifying accidents and traffic with tweets [link]
  - Multimodal traffic integration with weather
    - Do you take the car when it rains? Increased uber ridership?

Some more

- COVID spreads
- Precipitation/Wind/Solar nowcasting
  - Google MetNet [link]
  - GraphCast