Tensor Graph Decomposition for Temporal Networks

Bishwadeep Das, Elvin Isufi

TU Delft

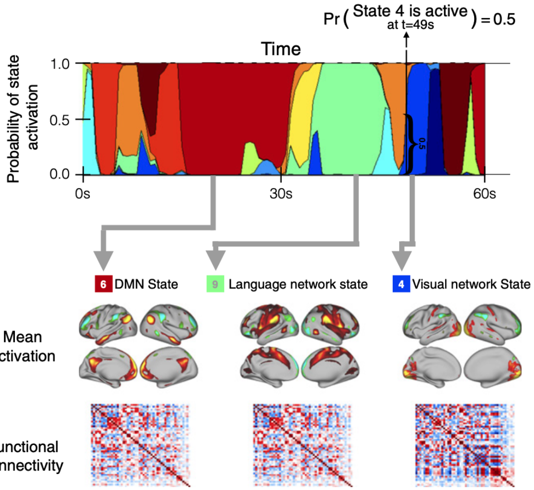
B.Das@tudelft.nl

Motivation

- Dynamic Networks represent a sequence of changing topology Holme and Saramäki (2012); Casteigts et al. (2012)
- Can be caused by a myriad of factors
- Studied under network models Zhang et al. (2017) and inferred from spatio-temporal data Natali et al. (2022); Kalofolias et al. (2017)
- Sometimes unobserved due to privacy or constraints

Where do they occur?

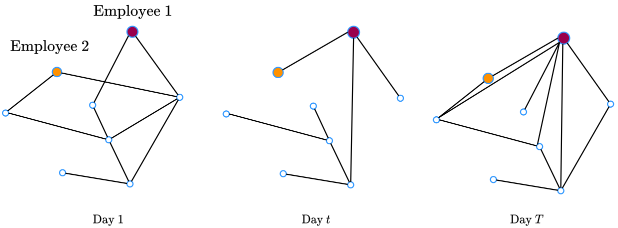
Time varying topology switching over a set of connectivity patterns [1]



[1] Figure taken from Vidaurre et al. (2017)

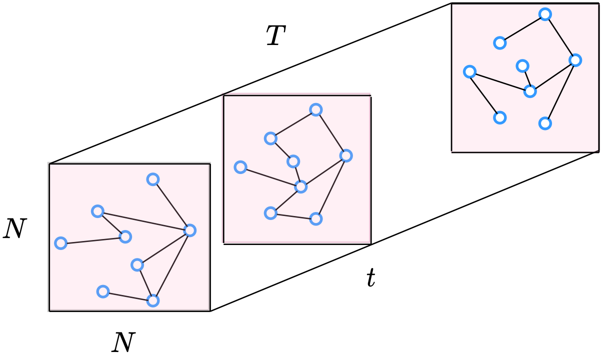
Where do they occur?

Communication pattern of a set of agents over a time period.



Representation through Tensors

- We can use three-dimensonal Tensors to represent the dynamic network



- Tensors allows the use of tools like Tensor Decompositions Kolda and Bader (2009); De Lathauwer et al. (2000)

Remarks

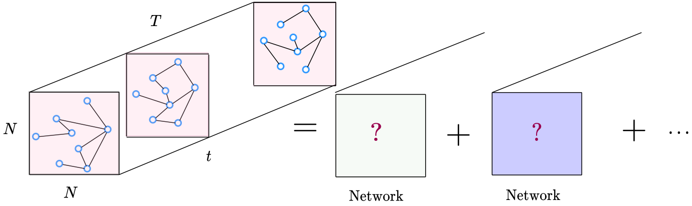
- Useful for downstream tasks dependent on embeddings Gujral et al. (2020)
- These decompositions are Low-Rank
- Low rank not always interpretable as a graph structure
- Data often not available

What we propose

- A two-way decomposition of this tensor
- Break down tensor as a sum of tensors
- Each component tensor corresponds to a mode or a graph scaled over the temporal dimension

Express the evolution of the structure as a sum of networks which can collectively capture this evolution

Illustration



Problem Formulation

- Consider a temporal graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$

Problem Formulation

- Consider a temporal graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$
- Adjacency tensor $\underline{\mathbf{X}} \in \mathbb{R}^{N \times N \times T}$

Problem Formulation

- Consider a temporal graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$
- Adjacency tensor $\underline{\mathbf{X}} \in \mathbb{R}^{N \times N \times T}$
- Represent $\mathcal{G}_t$ as a linear combination of $R \ll T$ *modal* graphs with adjacency matrices $\mathbf{A}_1, \ldots, \mathbf{A}_R \in \mathbb{R}^{N \times N}$.

Problem Formulation

- Consider a temporal graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$
- Adjacency tensor $\underline{\mathbf{X}} \in \mathbb{R}^{N \times N \times T}$
- Represent $\mathcal{G}_t$ as a linear combination of $R \ll T$ *modal* graphs with adjacency matrices $\mathbf{A}_1, \dots, \mathbf{A}_R \in \mathbb{R}^{N \times N}$.
- $\hat{\underline{\mathbf{X}}} = \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r := [[\mathbf{A}, \mathbf{C}]]$

Problem Formulation

- Consider a temporal graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$
- Adjacency tensor $\underline{\mathbf{X}} \in \mathbb{R}^{N \times N \times T}$
- Represent $\mathcal{G}_t$ as a linear combination of $R \ll T$ *modal* graphs with adjacency matrices $\mathbf{A}_1, \ldots, \mathbf{A}_R \in \mathbb{R}^{N \times N}$.
- $\underline{\hat{\mathbf{X}}} = \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r := [[\mathbf{A}, \mathbf{C}]]$
- Adjacency matrix at time t $\underline{\mathbf{X}}_{:,:,t} = \underline{\hat{\mathbf{X}}}_{:,:,t} + \underline{\mathbf{E}}_{:,:,t} = \sum_{r=1}^{R} c_{rt} \mathbf{A}_r + \underline{\mathbf{E}}_{:,:,t}$.

General Problem

$$\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C})$$
$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R] \tag{1}$$

Problem Statement

Given $\underline{\mathbf{X}}$, estimate $\{\mathbf{A}_r\}$ and $\{\mathbf{c}_r\}$ for a known R

Specifics

$$\begin{aligned}\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad & l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C}) \\ \text{subject to } & \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]\end{aligned}$$

- $l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) = ||\underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r||_F^2$

Specifics

$$\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C})$$
$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- $l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) = ||\underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r||_F^2$
- $\mathcal{S}$: Set of Adjacency Matrices without self loops

Specifics

$$\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C})$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S}, \;\; \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R], \;\; \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- $l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) = ||\underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r||_F^2$
- $\mathcal{S}$: Set of Adjacency Matrices without self loops
- $g(\mathbf{A}_r) = ||\mathbf{A}_r||_1$ for sparsity.

Specifics

$$\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C})$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- $l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) = ||\underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r||_F^2$
- $\mathcal{S}$: Set of Adjacency Matrices without self loops
- $g(\mathbf{A}_r) = ||\mathbf{A}_r||_1$ for sparsity.
- $h(\mathbf{A})$ penalises shared support across the modes.

Specifics

$$\underset{\mathbf{A},\mathbf{C}}{\text{minimize}} \quad l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) + \sum_{r=1}^{R} g(\mathbf{A}_r) + h(\mathbf{A}) + i(\mathbf{C})$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- $l(\underline{\mathbf{X}}, [[\mathbf{A}, \mathbf{C}]]) = ||\underline{\mathbf{X}} - \sum_{r=1}^{R} \mathbf{A}_r \circ \mathbf{c}_r||_F^2$
- $\mathcal{S}$: Set of Adjacency Matrices without self loops
- $g(\mathbf{A}_r) = ||\mathbf{A}_r||_1$ for sparsity.
- $h(\mathbf{A})$ penalises shared support across the modes.
- $i(\mathbf{C}) = \sum_{r=1}^{R} ||\mathbf{D}\mathbf{c}_r||_2^2$

Tensor Graph Decomposition (TGD)

$$\min_{\mathbf{A},\mathbf{C}} \frac{1}{2}\left\|\underline{\mathbf{X}} - [[\mathbf{A}, \mathbf{C}]]\right\|_F^2 + \gamma \sum_{r=1}^{R} ||\mathbf{A}_r||_1 + \frac{\beta}{2} h(\mathbf{A}) + ||\mathbf{D}\mathbf{C}||_F^2$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

Problem Statement

Given $\underline{\mathbf{X}}$, estimate $\{\mathbf{A}_r\}$ and $\{\mathbf{c}_r\}$ for a known R

Solving the TGD

$$\min_{\mathbf{A},\mathbf{C}} \frac{1}{2}\left\|\underline{\mathbf{X}} - [[\mathbf{A},\mathbf{C}]]\right\|_F^2 + \gamma \sum_{r=1}^{R} ||\mathbf{A}_r||_1 + \frac{\beta}{2} h(\mathbf{A}) + ||\mathbf{DC}||_F^2$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- Alternate minimization between **A** and **C**

Solving the TGD

$$\min_{\mathbf{A},\mathbf{C}} \frac{1}{2}\left\|\underline{\mathbf{X}} - [[\mathbf{A},\mathbf{C}]]\right\|_F^2 + \gamma \sum_{r=1}^{R} ||\mathbf{A}_r||_1 + \frac{\beta}{2} h(\mathbf{A}) + ||\mathbf{D}\mathbf{C}||_F^2$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- Alternate minimization between **A** and **C**
- For **A**, alternate between each $\mathbf{A}_r$
- Solve for each $\mathbf{A}_r$ in the dual domain.

Solving the TGD

$$\min_{\mathbf{A},\mathbf{C}} \frac{1}{2}\left\|\underline{\mathbf{X}} - [[\mathbf{A},\mathbf{C}]]\right\|_F^2 + \gamma \sum_{r=1}^{R} ||\mathbf{A}_r||_1 + \frac{\beta}{2} h(\mathbf{A}) + ||\mathbf{D}\mathbf{C}||_F^2$$

$$\text{subject to } \mathbf{A}_r \in \mathcal{S},\ \mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_R],\ \mathbf{C} = [\mathbf{c}_1, \ldots, \mathbf{c}_R]$$

- Alternate minimization between $\mathbf{A}$ and $\mathbf{C}$
- For $\mathbf{A}$, alternate between each $\mathbf{A}_r$
- Solve for each $\mathbf{A}_r$ in the dual domain.
- Closed form solution for each $\mathbf{C}$ update

Errors

Dataset	**TGD** $\gamma = 0.01$ $\beta = 0.01$	**TGD** $\gamma = 9$ $\beta = 16$	**Uncon**	**BTD**	**SBTD**
SBM	0.15	0.24	0.15	0.37	0.44
Enron	0.74	0.82	0.71	0.76	0.98

- NMSE score for all methods.

Errors

Dataset	**TGD** $\gamma = 0.01$ $\beta = 0.01$	**TGD** $\gamma = 9$ $\beta = 16$	**Uncon**	**BTD**	**SBTD**
SBM	0.15	0.24	0.15	0.37	0.44
Enron	0.74	0.82	0.71	0.76	0.98
SBM	0.92	0.93	0.58	0.48	0.48
Enron	0.18	0.40	0.07	0.03	0.03

- F1 score evaluated on 20 percent removed observations

Errors

Dataset	**TGD** $\gamma = 0.01$ $\beta = 0.01$	**TGD** $\gamma = 9$ $\beta = 16$	**Uncon**	**BTD**	**SBTD**
SBM	0.15	0.24	0.15	0.37	0.44
Enron	0.74	0.82	0.71	0.76	0.98
SBM	0.92	0.93	0.58	0.48	0.48
Enron	0.18	0.40	0.07	0.03	0.03
SBM	0.90	0.90	0.52	0.43	0.43
Enron	0.17	0.19	0.06	0.03	0.0

- F1 score over a fixed set of unobserved nodes.
- Higher NMSE for better F1
- Constraints help recover structure

Properties of modes

Graph entropy (no. of connected components) for the recovered components.

	$\beta = 0.01$	$\beta = 6$	$\beta = 16$
$\gamma = 0.01$	4.35(1)	4.29(1)	4.18(1.2)
$\gamma = 5$	4.2(1)	4.25(1.2)	4.12(1.26)
$\gamma = 9$	4.2(1.3)	4.19(1.4)	3.94(1.66)

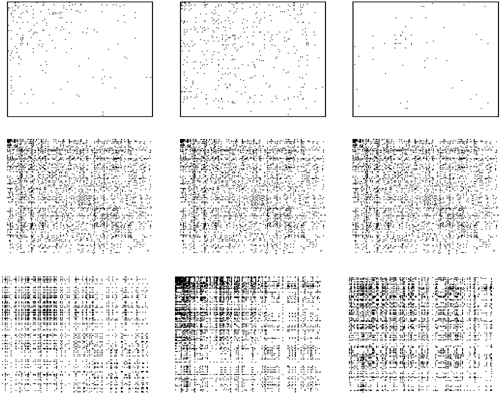
- Imposing graph structure tends to less diversity in degree
- Also leads to more connected components

Visualizing the modes

TGD

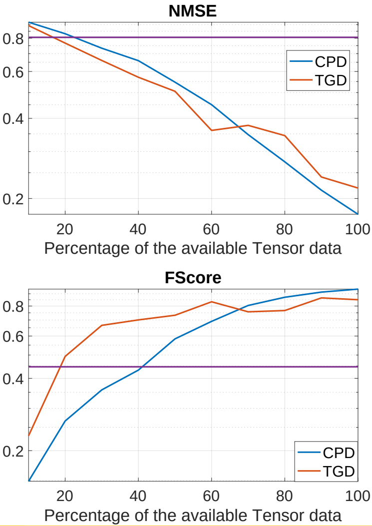
Uncon

SBTD

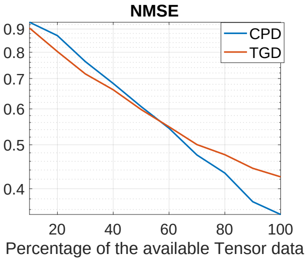


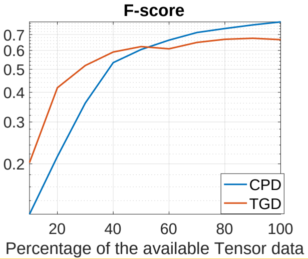
- TGD recovers different structures
- Unconstrained replicating same structure
- Symm BTD recovers typical low rank structure

SBM Graph with rewiring



Molene Temperature Sensor Graph





Summary

Conclusion

- A two-way decomposition on dynamic networks to recover structure.
- Helps recover structure, at the cost of reconstruction error.

Future Work

- Consider alternative priors on relationship between recovered modes.
- Conditions of Uniqueness, Identifiability and Recovery.

References

Casteigts, A., Flocchini, P., Quattrociocchi, W., and Santoro, N. (2012). Time-varying graphs and dynamic networks. *International Journal of Parallel, Emergent and Distributed Systems*, 27(5):387–408.

De Lathauwer, L., De Moor, B., and Vandewalle, J. (2000). A multilinear singular value decomposition. *SIAM journal on Matrix Analysis and Applications*, 21(4):1253–1278.

Gujral, E., Pasricha, R., and Papalexakis, E. (2020). Beyond rank-1: Discovering rich community structure in multi-aspect graphs. In *Proceedings of The Web Conference 2020*, pages 452–462.

Holme, P. and Saramäki, J. (2012). Temporal networks. *Physics reports*, 519(3):97–125.

Kalofolias, V., Loukas, A., Thanou, D., and Frossard, P. (2017). Learning time varying graphs. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2826–2830. Ieee.

Kolda, T. G. and Bader, B. W. (2009). Tensor decompositions and applications. *SIAM review*, 51(3):455–500.

Natali, A., Isufi, E., Coutino, M., and Leus, G. (2022). Learning time-varying graphs from online data. *IEEE Open Journal of Signal Processing*, 3:212–228.

Vidaurre, D., Smith, S. M., and Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832.

Zhang, X., Moore, C., and Newman, M. E. (2017). Random graph models for dynamic networks. *The European Physical Journal B*, 90:1–14.