Online Network Topology Inference with Partial Connectivity Information

Rasoul Shafipour and **Gonzalo Mateos**

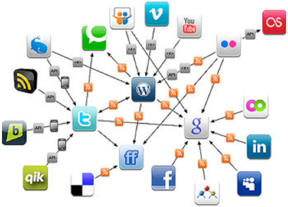
Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
`http://www.ece.rochester.edu/~gmateosb/`

Acknowledgment: NSF Awards CCF-1750428, ECCS-1809356, CCF-1934962

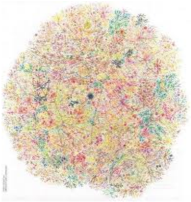
Guadeloupe, French West Indies, December 17, 2019

Network Science analytics

Online social media



Internet

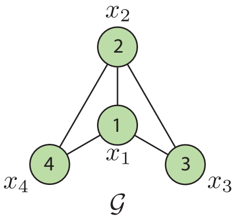


Clean energy and grid analytics

- Network as graph $G = (\mathcal{V}, \mathcal{E})$: encode pairwise relationships
- Desiderata: Process, analyze and learn from network data [Kolaczyk'09]
  $\Rightarrow$ Use G to study graph signals, data associated with nodes in $\mathcal{V}$
- Ex: Opinion profile, buffer congestion levels, neural activity, epidemic

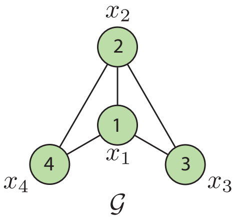
Graph signal processing (GSP)

- Undirected $\mathcal{G}$ with adjacency matrix $\mathbf{A}$
  - $\Rightarrow A_{ij}$ = Proximity between i and j
- Define a signal $\mathbf{x}$ on top of the graph
  - $\Rightarrow x_i$ = Signal value at node i



Graph signal processing (GSP)

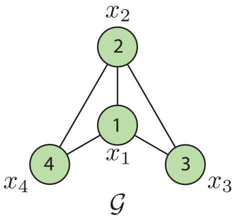
- Undirected $\mathcal{G}$ with adjacency matrix $\mathbf{A}$
  - $\Rightarrow A_{ij}$ = Proximity between i and j
- Define a signal $\mathbf{x}$ on top of the graph
  - $\Rightarrow x_i$ = Signal value at node i



- Associated with $\mathcal{G}$ is the graph-shift operator (GSO) $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$
  - $\Rightarrow S_{ij} = 0$ for $i \neq j$ and $(i, j) \notin \mathcal{E}$ (local structure in $\mathcal{G}$)
  - $\Rightarrow$ Ex: $\mathbf{A}$, degree $\mathbf{D}$ and Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$ matrices

Graph signal processing (GSP)

- Undirected $\mathcal{G}$ with adjacency matrix $\mathbf{A}$
  - $\Rightarrow A_{ij}$ = Proximity between i and j
- Define a signal $\mathbf{x}$ on top of the graph
  - $\Rightarrow x_i$ = Signal value at node i



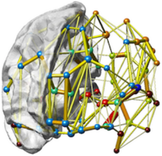
- Associated with $\mathcal{G}$ is the graph-shift operator (GSO) $\mathbf{S} = \mathbf{V\Lambda V}^T$
  - $\Rightarrow S_{ij} = 0$ for $i \neq j$ and $(i, j) \notin \mathcal{E}$ (local structure in $\mathcal{G}$)
  - $\Rightarrow$ Ex: $\mathbf{A}$, degree $\mathbf{D}$ and Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$ matrices
- Graph Signal Processing $\rightarrow$ Exploit structure encoded in $\mathbf{S}$ to process $\mathbf{x}$
  - $\Rightarrow$ Use GSP to learn the underlying $\mathcal{G}$ or a meaningful network model

Topology inference: Motivation and context

- Network topology inference from nodal observations [Kolaczyk'09]
  - Partial correlations and conditional dependence [Dempster'74]
  - Sparsity [Friedman et al'07] and consistency [Meinshausen-Buhlmann'06]

- Key in neuroscience [Sporns'10]

  $\Rightarrow$ Functional network from BOLD signal

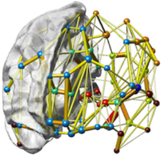


Topology inference: Motivation and context

- Network topology inference from nodal observations [Kolaczyk'09]
  - Partial correlations and conditional dependence [Dempster'74]
  - Sparsity [Friedman et al'07] and consistency [Meinshausen-Buhlmann'06]

- Key in neuroscience [Sporns'10]

  $\Rightarrow$ Functional network from BOLD signal



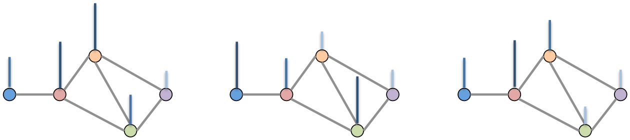
- Noteworthy GSP-based approaches
  - Graphical models [Egilmez et al'16], [Rabbat'17], [Kumar et al'19] ...
  - Smooth signals [Dong et al'15], [Kalofolias'16], [Sardellitti et al'17], ...
  - Stationary signals [Pasdeloup et al'15], [Segarra et al'16], ...
  - Directed graphs [Bazerque et al'13], [Mei-Moura'15], [Shen et al'16], ...
  - Streaming data [Baingana et al'14], [Vlaski et al'18], ...

- **Our contribution:** topology inference in the graph spectral domain
  - Here online topology inference with partial connectivity information

Problem formulation

Setup

- Undirected network $\mathcal{G}$ with unknown graph shift $\mathbf{S}$. Observe
  - $\Rightarrow$ Streaming signals $\{\mathbf{y}_i\}_{i=1}^{P}$ defined on $\mathbf{S}$
  - $\Rightarrow$ Edge status for $(i,j) \in \Omega \subset \mathcal{V} \times \mathcal{V}$

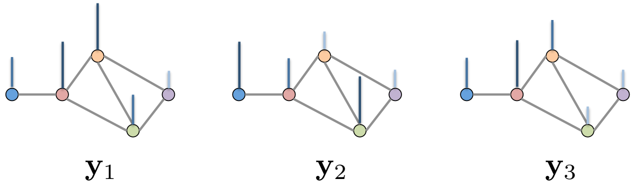


$\mathbf{y}_1$ $\mathbf{y}_2$ $\mathbf{y}_3$

Problem formulation

Setup

- Undirected network $\mathcal{G}$ with unknown graph shift $\mathbf{S}$. Observe
  - $\Rightarrow$ Streaming signals $\{\mathbf{y}_i\}_{i=1}^{P}$ defined on $\mathbf{S}$
  - $\Rightarrow$ Edge status for $(i,j) \in \Omega \subset \mathcal{V} \times \mathcal{V}$



Problem statement

> Given observations $\{\mathbf{y}_i\}_{i=1}^{P}$ and edge status in Ω, determine the network $\mathbf{S}$ knowing that $\{\mathbf{y}_i\}_{i=1}^{P}$ are generated via diffusion on $\mathbf{S}$.

Generating structure of a diffusion process

- Signal $\mathbf{y}_i$ is the response of a linear diffusion process to input $\mathbf{x}_i$

$$\mathbf{y}_i = \alpha_0 \prod_{l=1}^{\infty} (\mathbf{I} - \alpha_l \mathbf{S}) \mathbf{x}_i = \sum_{l=0}^{\infty} \beta_l \mathbf{S}^l \mathbf{x}_i, \quad i = 1, \ldots, P$$

$\Rightarrow$ Common generative model, e.g., heat diffusion, consensus

Generating structure of a diffusion process

- Signal $\mathbf{y}_i$ is the response of a linear diffusion process to input $\mathbf{x}_i$

$$\mathbf{y}_i \;=\; \alpha_0 \prod_{l=1}^{\infty} (\mathbf{I} - \alpha_l \mathbf{S}) \mathbf{x}_i \;=\; \sum_{l=0}^{\infty} \beta_l \mathbf{S}^l \mathbf{x}_i, \quad i = 1, \ldots, P$$

$\Rightarrow$ Common generative model, e.g., heat diffusion, consensus

- Cayley-Hamilton asserts we can write diffusion as ($L \leq N$)

$$\mathbf{y}_i = \left(\sum_{l=0}^{L-1} h_l \mathbf{S}^l \right) \mathbf{x}_i := \mathbf{H} \mathbf{x}_i, \quad i = 1, \ldots, P$$

$\Rightarrow$ Graph filter $\mathbf{H}$ is shift invariant [Sandryhaila-Moura'13]

$\Rightarrow$ $\mathbf{H}$ diagonalized by the eigenvectors $\mathbf{V}$ of the shift operator

Generating structure of a diffusion process

- Signal $\mathbf{y}_i$ is the response of a linear diffusion process to input $\mathbf{x}_i$

$$\mathbf{y}_i \;=\; \alpha_0 \prod_{l=1}^{\infty} (\mathbf{I} - \alpha_l \mathbf{S})\mathbf{x}_i \;=\; \sum_{l=0}^{\infty} \beta_l \mathbf{S}^l \mathbf{x}_i, \quad i = 1, \ldots, P$$

$\Rightarrow$ Common generative model, e.g., heat diffusion, consensus

- Cayley-Hamilton asserts we can write diffusion as ($L \leq N$)

$$\mathbf{y}_i = \left(\sum_{l=0}^{L-1} h_l \mathbf{S}^l \right) \mathbf{x}_i := \mathbf{H}\mathbf{x}_i, \quad i = 1, \ldots, P$$

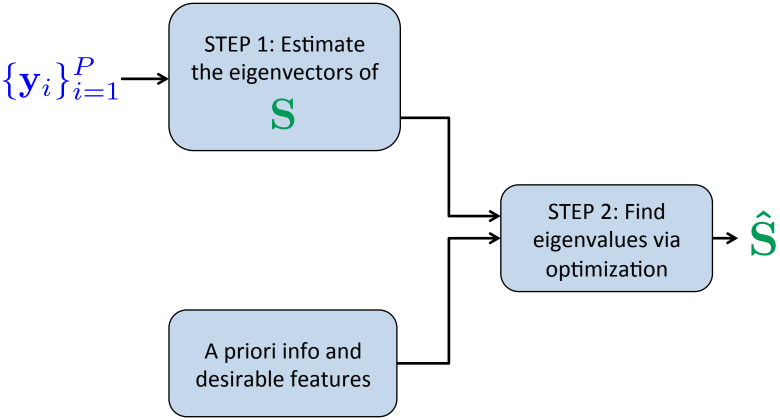
$\Rightarrow$ Graph filter $\mathbf{H}$ is shift invariant [Sandryhaila-Moura'13]

$\Rightarrow$ $\mathbf{H}$ diagonalized by the eigenvectors $\mathbf{V}$ of the shift operator

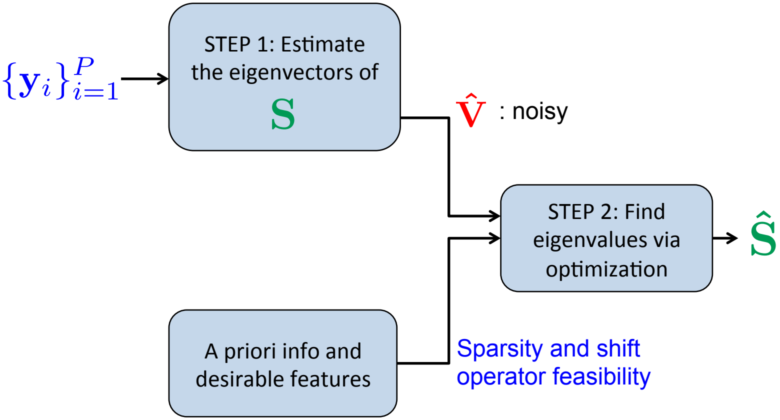
- **Goal**: estimate undirected network $\mathbf{S}$ online from signals $\{\mathbf{y}_i\}_{i=1}^P$

$\Rightarrow$ Unknowns: filter order L, coefficients $\{h_l\}_{l=1}^{L-1}$, inputs $\{\mathbf{x}_i\}_{i=1}^P$

Blueprint of batch solution



Blueprint of batch solution



Two-step approach

- Suppose that the input is white, i.e., $\mathbf{C}_x = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right] = \mathbf{I}$

  $\Rightarrow$ The covariance matrix of $\mathbf{y} = \mathbf{H}\mathbf{x}$ shares $\mathbf{V}$ with $\mathbf{S}$

$$\mathbf{C}_y = \mathbb{E}\left[\mathbf{H}\mathbf{x}\left(\mathbf{H}\mathbf{x}\right)^T\right] = \mathbf{H}^2 = h_0^2\mathbf{I} + 2h_0h_1\mathbf{S} + h_1^2\mathbf{S}^2 + \ldots$$

- Form sample covariance $\hat{\mathbf{C}}_y$ using $\{\mathbf{y}_i\}_{i=1}^P$ $\Rightarrow$ Diagonalize $\Rightarrow$ Obtain $\hat{\mathbf{V}}$

Two-step approach

- Suppose that the input is white, i.e., $\mathbf{C}_x = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right] = \mathbf{I}$

  $\Rightarrow$ The covariance matrix of $\mathbf{y} = \mathbf{H}\mathbf{x}$ shares $\mathbf{V}$ with $\mathbf{S}$

$$\mathbf{C}_y = \mathbb{E}\left[\mathbf{H}\mathbf{x}\left(\mathbf{H}\mathbf{x}\right)^T\right] = \mathbf{H}^2 = h_0^2\mathbf{I} + 2h_0h_1\mathbf{S} + h_1^2\mathbf{S}^2 + \ldots$$

- Form sample covariance $\hat{\mathbf{C}}_y$ using $\{\mathbf{y}_i\}_{i=1}^P$ $\Rightarrow$ Diagonalize $\Rightarrow$ Obtain $\hat{\mathbf{V}}$

- Of all graphs, given $\hat{\mathbf{V}}$ select one that is sparse and feasible

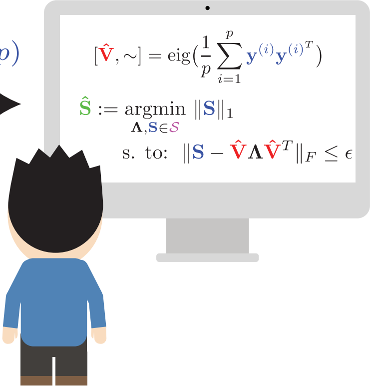
$$\hat{\mathbf{S}} := \underset{\mathbf{S},\mathbf{\Lambda}}{\text{argmin}} \ \|\mathbf{S}\|_1 \quad \text{subject to:} \quad \|\mathbf{S} - \hat{\mathbf{V}}\mathbf{\Lambda}\hat{\mathbf{V}}^T\|_F \leq \epsilon, \ \ \mathbf{S} \in \mathcal{S}$$

- Set $\mathcal{S}$ contains all admissible scaled adjacency matrices

$$\mathcal{S} := \{\mathbf{S} \mid S_{ij} \geq 0, \mathbf{S}^T = \mathbf{S}, S_{ii} = 0, S_{ij} = s_{ij}, \ (i,j) \in \Omega\}$$

Online network topology inference

$$\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(p)}$$

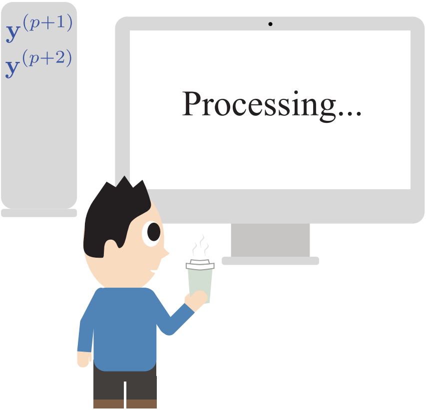


$$[\hat{\mathbf{V}}, \sim] = \mathrm{eig}\Big(\frac{1}{p}\sum_{i=1}^{p} \mathbf{y}^{(i)} {\mathbf{y}^{(i)}}^T\Big)$$

$$\hat{\mathbf{S}} := \underset{\mathbf{\Lambda}, \mathbf{S} \in \mathcal{S}}{\mathrm{argmin}} \; \|\mathbf{S}\|_1$$

$$\text{s. to:} \;\; \|\mathbf{S} - \hat{\mathbf{V}} \mathbf{\Lambda} \hat{\mathbf{V}}^T\|_F \leq \epsilon$$

Online network topology inference



Online network topology inference

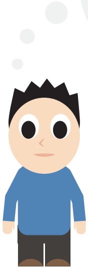
Online network topology inference

Online network topology inference

- Develop an iterative algorithm for the topology inference
- Upon sensing new diffused output signals

  $\Rightarrow$ - Update $\hat{\mathbf{V}}$ efficiently

  - Take one or a few steps of the iterative algorithm



Alternating minimization

$$\begin{aligned} \underset{\mathbf{\Lambda},\mathbf{S}}{\text{minimize}} \quad & \mu\|\mathbf{S}\|_1 + \frac{1}{2}\|\mathbf{S} - \hat{\mathbf{V}}\mathbf{\Lambda}\hat{\mathbf{V}}^T\|_F^2, \\ \text{subject to} \quad & \mathbf{\Lambda} \in \mathcal{D}, \quad \mathbf{S}^T = \mathbf{S}, \quad S_{ij} \geq 0, \quad S_{ii} = 0, \\ & S_{ij} = s_{ij}, \ (i,j) \in \Omega. \end{aligned}$$

Alternating minimization

$$\begin{aligned} \underset{\boldsymbol{\Lambda},\mathbf{S}}{\text{minimize}} \quad & \mu\|\mathbf{S}\|_1 + \frac{1}{2}\|\mathbf{S} - \hat{\mathbf{V}}\boldsymbol{\Lambda}\hat{\mathbf{V}}^T\|_F^2, \\ \text{subject to} \quad & \boldsymbol{\Lambda} \in \mathcal{D}, \quad \mathbf{S}^T = \mathbf{S}, \quad S_{ij} \geq 0, \quad S_{ii} = 0, \\ & S_{ij} = s_{ij}, \ (i,j) \in \Omega. \end{aligned}$$

- $\boldsymbol{\Lambda}$-update:

$$\boldsymbol{\Lambda}(k) := \underset{\boldsymbol{\Lambda}\in\mathcal{D}}{\text{argmin}} \ \|\mathbf{S}(k) - \hat{\mathbf{V}}\boldsymbol{\Lambda}\hat{\mathbf{V}}^\top\|_F^2 = \text{diag}(\hat{\mathbf{V}}^T\mathbf{S}(k)\hat{\mathbf{V}})$$

- $\mathbf{S}$-update: Let $\mathbf{B}(k) := \hat{\mathbf{V}}\boldsymbol{\Lambda}(k)\hat{\mathbf{V}}^T$, then

$$S_{ij}(k+1) = \begin{cases} 0, & i = j \\ s_{ij}, & (i,j) \in \Omega \\ \max(0, B_{ij}(k) - \mu), & \text{otherwise} \end{cases}$$

- Convex, thus block coordinate-descent converges to a global minimizer

Online eigenvector estimates

- **Q:** How can we efficiently update the sample covariance eigenvectors $\hat{\mathbf{V}}$?
- $\hat{\mathbf{C}}_{\mathbf{y}}^{(P)}$ is the sample covariance from P streaming observations

$$\hat{\mathbf{C}}_{\mathbf{y}}^{(P+1)} = \frac{1}{P+1}\left(P\hat{\mathbf{C}}_{\mathbf{y}}^{(P)} + \mathbf{y}^{(P+1)}\mathbf{y}^{(P+1)}\right)$$

Online eigenvector estimates

- **Q:** How can we efficiently update the sample covariance eigenvectors $\hat{\mathbf{V}}$?
- $\hat{\mathbf{C}}_{\mathbf{y}}{}^{(P)}$ is the sample covariance from P streaming observations

$$\hat{\mathbf{C}}_{\mathbf{y}}{}^{(P+1)} = \frac{1}{P+1}(P\hat{\mathbf{C}}_{\mathbf{y}}{}^{(P)} + \mathbf{y}^{(P+1)}\mathbf{y}^{(P+1)})$$

- Let $\mathbf{z} = (\hat{\mathbf{V}}^{(p)})^\top \mathbf{y}^{(p+1)}$ and $\{d_j\}_{j=1}^N$ denote the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}{}^{(P)}$
  - $\Rightarrow$ Eigenvalues of rank-one modification of $\hat{\mathbf{C}}_{\mathbf{y}}{}^{(P)}$ are the roots (γ) of

$$1 + \sum_{j=1}^{N} \frac{z_j^2}{Pd_j - \gamma} = 0 \quad \text{[Bunch et al'78]}$$

  - $\Rightarrow$ Can be solved using the Newton method with $\mathcal{O}(N^2)$ complexity

Online eigenvector estimates

- **Q:** How can we efficiently update the sample covariance eigenvectors $\hat{\mathbf{V}}$?
- $\hat{\mathbf{C}}_{\mathbf{y}}^{(P)}$ is the sample covariance from P streaming observations

$$\hat{\mathbf{C}}_{\mathbf{y}}^{(P+1)} = \frac{1}{P+1}(P\hat{\mathbf{C}}_{\mathbf{y}}^{(P)} + \mathbf{y}^{(P+1)}\mathbf{y}^{(P+1)})$$

- Let $\mathbf{z} = (\hat{\mathbf{V}}^{(p)})^{\top}\mathbf{y}^{(p+1)}$ and $\{d_j\}_{j=1}^{N}$ denote the eigenvalues of $\hat{\mathbf{C}}_{\mathbf{y}}^{(P)}$

  $\Rightarrow$ Eigenvalues of rank-one modification of $\hat{\mathbf{C}}_{\mathbf{y}}^{(P)}$ are the roots (γ) of

$$1 + \sum_{j=1}^{N} \frac{z_j^2}{Pd_j - \gamma} = 0 \quad \text{[Bunch et al'78]}$$

  $\Rightarrow$ Can be solved using the Newton method with $\mathcal{O}(N^2)$ complexity

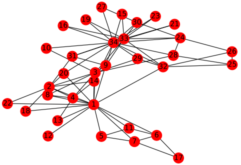
- For the updated eigenvalue γ_j, the corresponding eigenvector $\mathbf{v}_j$ is given by

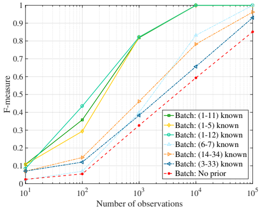
$$\mathbf{v}_j^{(p+1)} = \alpha_j \mathbf{y}^{(p+1)} \circ \mathbf{q}_j,$$

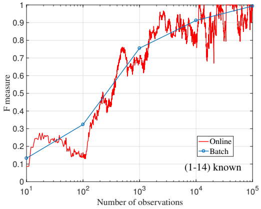
  where $\mathbf{q}_j = [1/(Pd_1 - \gamma_j), \cdots, 1/(Pd_N - \gamma_j)]$ and α_j is a normalizing factor

Recovery of Zachary's karate club

- Consider Zachary's karate club with $N = 34$ nodes
  - Diffusion filter $\mathbf{H} = \sum_{l=0}^{2} h_l \mathbf{A}^l$, $h_l \sim \mathcal{U}[0,1]$
  - Generate streaming signals $\{\mathbf{y}^{(1)}, \cdots, \mathbf{y}^{(p)}, \mathbf{y}^{(p+1)}, \cdots\}$ via $\mathbf{y}^{(i)} = \mathbf{H}\mathbf{x}^{(i)}$
  - Both batch and online inference for different Ω







- Least informative 3 edges in Ω for recovering $\mathbf{A}$
  - Among the top links in terms of the edge betweenness centrality

- The online scheme attains the performance of its batch counterpart

Closing remarks

- Online topology inference from streaming diffused graph signals
  - Graph shift $\mathbf{S}$ and covariance $\mathbf{C}_y$ are simultaneously diagonalizable
  - Promote desirable properties on $\mathbf{S}$ via convex criteria

- Developed an iterative algorithm for the topology inference
- Upon sensing new diffused output signals

$\Rightarrow$ - Updated $\hat{\mathbf{V}}$ efficiently

- Took one or a few steps of the iterative algorithm