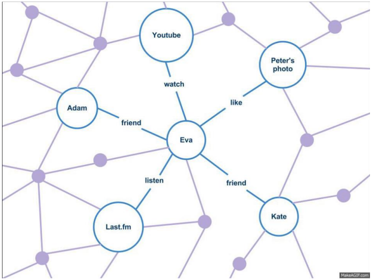
Estimating Centrality Blindly from Low-pass Filtered Graph Signals

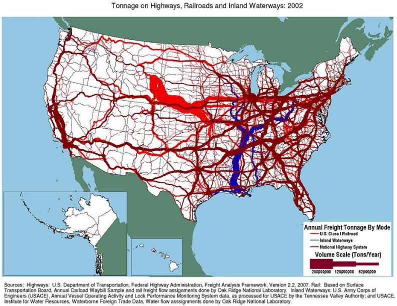
Yiran He, Hoi-To Wai

Department of Systems Engineering & Engineering Management, CUHK

April 16, 2020 ICASSP 2020

Motivation

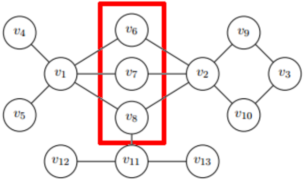


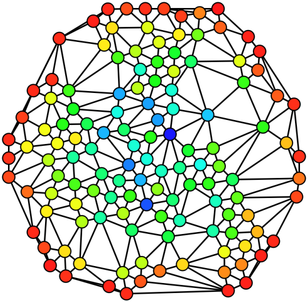


[Source: Wikipedia]

- Graphs are useful for describing the **geometric** structures of data from numerous fields, including social, energy, transportation, and neuronal networks.

Motivation





[Source: Wikipedia]

- Importance of nodes in the graph $\Longrightarrow$ **node centrality**
- E.g., social network – most influential individuals, transport network – cities with largest population mobility.

Research Question: *Can we learn node centrality <u>from data</u>?*

Prior works

- Graph learning: recover the complete topology
  - *Statistical/physical models* – GMRF [Friedman et al., 2008], dynamical systems / causation [Shen et al., 2017].
  - *Graph signal processing (GSP) models* – smoothness of graph signals [Dong et al., 2016], inference with structural constraints [Egilmez et al., 2017], spectral template [Segarra et al., 2017].
  - and many others...
- **This work**: learning graph features *without the graph*
  - *Community inference* – blind community detection [Wai et al., 2018], Bayesian learning [Hoffmann et al., 2018], recovery from multi-graph [Roddenberry et al., 2020].
  - *Centrality learning* – centrality ranking inference [Roddenberry and Segarra, 2019], and this work.

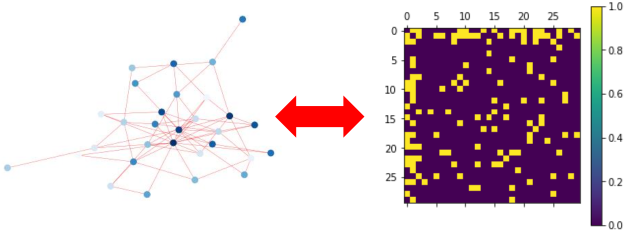
Contribution

- We show that the folklore heuristic based on PCA **works** if data is generated from a ‘strong’ low-pass filter.
- For data generated from a ‘weak’ low-pass filter, we propose a **boosting** method with provably better estimation quality.
- Numerical experiments on synthetic and stocks data.

Graph Model and Centrality Measure

- **Undirected** graph $G = (V, E, \mathbf{A})$ with $V = \{1, \ldots, N\}$, symmetrix adjacency matrix $\mathbf{A} \in \Re_{+}^{N \times N}$
- The **adjacency matrix** admits an eigenvalue decomposition (EVD) as $\mathbf{A} = \mathbf{V} \boldsymbol{\Lambda} \mathbf{V}^{\top}$ s.t. $\mathbf{V} =$ orthogonal, $\boldsymbol{\Lambda} = \mathrm{Diag}(\lambda_1, ..., \lambda_N)$.
- Centrality is given by the eigen-centrality

$$c_{\text{eig}} := \text{TopEV}(\mathbf{A}) = \mathbf{v}_1$$

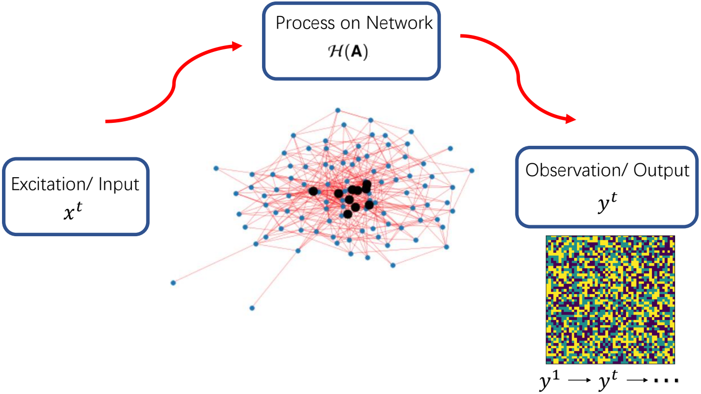


Network Adjacency Matrix

Graph Signal Model

- Observed data y^t is produced by an excitation x^t to be 'processed' by a **graph filter** $\mathcal{H}(\mathbf{A})$

$$y^t = \mathcal{H}(\mathbf{A})x^t$$



Graph Signal Model (cont'd)

Graph Filter

- The graph filter $\mathcal{H}(\mathbf{A})$ is a matrix polynomial:

$$\mathcal{H}(\mathbf{A}) = \sum_{t=0}^{T-1} h_t \mathbf{A}^t$$

Set $h(\lambda) := \sum_{t=0}^{T-1} h_t \lambda^t$.

- Assume **1-low pass** $\mathcal{H}(\mathbf{A})$:

$$\max_{j=2,\ldots,N} |h(\lambda_j)|/|h(\lambda_1)| =: \eta < 1$$

- $\eta \ll 1 \implies$ **strong** low-pass.
  $\eta \approx 1 \implies$ **weak** low-pass.
- E.g., diffusion, op. dynamics.

Excitation Signal

- The input x^t is controlled by an external source z^t:

$$x^t = \mathbf{B} z^t$$

- Assume a **sparse** influence matrix $\mathbf{B} \in \Re^{N \times k}$, $(k < N)$.
- E.g., influence from external source z^t are localized to specific nodes on graph.

Blind centrality estimation

- Idea: apply PCA on the covariance of filtered graph signals, use the principal eigenvector as an estimate for c_{eig}

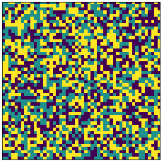
$$C_y = \frac{1}{m}\sum_{t=1}^{m} y^t (y^t)^T$$

Sample Covariance

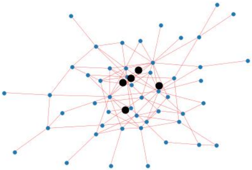
PCA

$$\hat{v}_1 := \text{TopEV}(C_y)$$

Centrality Estimation



$Y = [\, y_1 \cdots y_m]$
Observation



Detected K possible central nodes

Blind centrality estimation

$$C_y = \mathcal{H}(\mathbf{A})\mathbf{B}\mathbf{B}^\top(\mathcal{H}(\mathbf{A}))^\top = \mathbf{V}\begin{bmatrix} h(\lambda_1) & & \\ & \ddots & \\ & & h(\lambda_N) \end{bmatrix}\mathbf{V}^\top\mathbf{B}\mathbf{B}^\top\mathbf{V}\begin{bmatrix} h(\lambda_1) & & \\ & \ddots & \\ & & h(\lambda_N) \end{bmatrix}\mathbf{V}^\top$$

(for 'strong' low pass filter) $\approx \mathrm{const} \cdot \mathbf{v}_1\mathbf{v}_1^\top$

Lemma

Suppose $h(\lambda_1) > max_{j=2,\dots,n} h(\lambda_j)$. Then it holds that

$$||c_{eig} - \hat{v}_1||_2 = \mathcal{O}\left(\frac{\max_{j=2,\dots,n} |h(\lambda_j)|}{|h(\lambda_1)|}\right) = \mathcal{O}(\eta)$$

- Centrality estimation may be **inaccurate** for 'weak' low pass filter (i.e., $\eta \approx 1$).

Boosted centrality estimation

- A simple modification to *strengthen* the low-pass filter.
- Let $\rho > 0$. Consider

$$\widetilde{\mathcal{H}}(\mathbf{A}) := \mathcal{H}(\mathbf{A}) - \rho \mathbf{I}, \quad \tilde{h}_\rho(\lambda) := h(\lambda) - \rho.$$

- Let $\mu := \frac{\max_{j=2,\ldots,n} |\lambda_j|}{|\lambda_1|}$
- **Observation**: there exists $\rho > 0$ such that

$$\frac{\max_{j=2,\ldots,n} |\tilde{h}_\rho(\lambda_j)|}{|\tilde{h}_\rho(\lambda_1)|} = \mathcal{O}\left(\frac{\max_{j=2,\ldots,n} |\lambda_j|}{|\lambda_1|} \frac{\max_{j=2,\ldots,n} |h(\lambda_j)|}{|h(\lambda_1)|}\right) = \mathcal{O}(\mu\eta)$$

- $\widetilde{\mathcal{H}}(\mathbf{A})$ has a **better low-pass condition** than $\mathcal{H}(\mathbf{A})$.

Boosted centrality estimation

- Assume the external signals $\mathbf{Z} \in \Re^{k \times M} (k < M)$ are known, $\mathcal{H}(\mathbf{A})\mathbf{B} =: \widehat{\mathbf{HB}} = \mathbf{YZ}^\top(\mathbf{ZZ}^\top)^{-1}$
- $\mathcal{H}(\mathbf{A})\mathbf{B}$ admits a low-rank + sparse decomposition as:

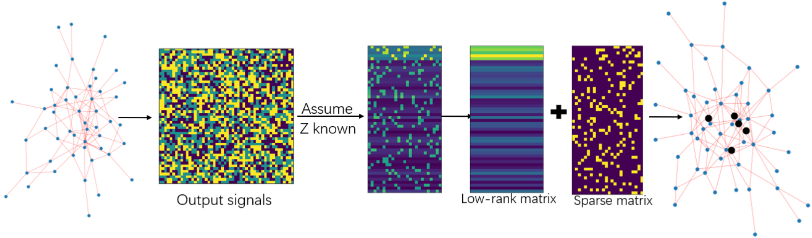
$$\mathcal{H}(\mathbf{A})\mathbf{B} = \widetilde{\mathcal{H}}(\mathbf{A})\mathbf{B} + \rho\mathbf{B} \equiv \mathbf{L} + \mathbf{S}$$

- **L** is a **low-rank** matrix and **S** is a **sparse** matrix.
- To obtain **L**, we solve the convex problem:

$$\min_{\hat{\mathbf{L}},\hat{\mathbf{S}}} \ \|\widehat{\mathbf{HB}} - \hat{\mathbf{L}} - \hat{\mathbf{S}}\|_F^2 + \lambda_L \|\hat{\mathbf{L}}\|_\star + \lambda_S \|\mathrm{vec}(\hat{\mathbf{S}})\|_1$$

where λ_L, λ_S – regularization for low-rankness, sparseness.

The whole process:



Corollary

Let $\tilde{v}_1$ be the top left *singular vector of* $\mathbf{L}$. *Under the same conditions as the previous Lemma. It holds*

$$\|c_{eig} - \tilde{v}_1\|_2 = \mathcal{O}\left(\frac{\max_{j=2,\ldots,n}|\lambda_j|}{|\lambda_1|}\frac{\max_{j=2,\ldots,n}|h(\lambda_j)|}{|h(\lambda_1)|}\right) = \mathcal{O}(\mu\eta)$$

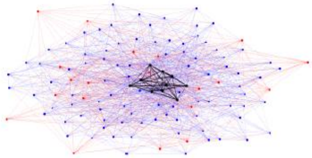
Numerical Results

- Graph G: Core periphery model with connectivity $p = 0.05$

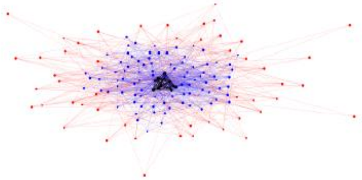
$$\begin{bmatrix} 1 & 4p \\ 4p & p \end{bmatrix}$$

- N = 100 nodes (10 core nodes), M = 10^5 observations
- Graph filter: $\mathcal{H}(\mathbf{A}) = (\mathbf{I} - 0.1\mathbf{A})^{-1}$, $\lambda_L = 0.1$, $\lambda_S = 0.2 + \frac{2}{\sqrt{k}}$.
- Three settings of **B** for different locations of external sources (**black** - central, **blue** - regular, **red** - external):

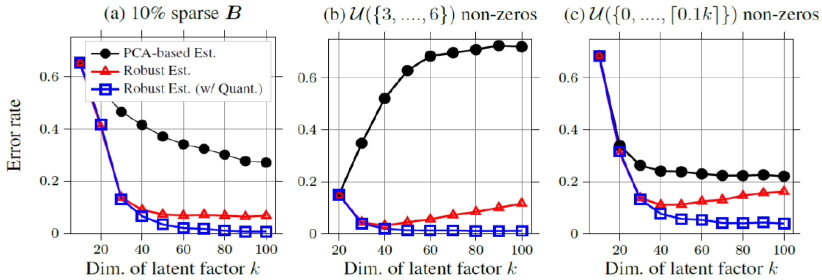
Setting (a)/(b)/(c), $k = 20$



Setting (b), $k = 50$



Numerical Results



- $\widehat{\mathbf{S}}^{\text{thres}} = \mathbf{1}(\hat{\mathbf{S}} \geq 0.1) \odot \hat{\mathbf{S}}$ and replace $\hat{\mathbf{L}}$ with $\widehat{\mathbf{HB}} - \widehat{\mathbf{S}}^{\text{thres}}$.
- Error rate = $\mathbb{E}\left[\frac{1}{10}|\{1, ..., 10\} \cap \tilde{v}_1|\right]$
- PCA suffers from a higher error rate than the robust methods.
- The error rate for the robust methods decreases with k.
- Results are consistent with our theoretical analysis

Real data

- Data: daily return data from S&P100 stocks in May 2018 to Aug 2019, consisting of $n = 99$ stocks and $m = 300$ samples, collected from `https://alphavantage.co`.
- External source: the latent input z^t on the relevant days estimated from the *interest level* on Google Trend (`https://trends.google.com`) on $k = 5$ key words: 'trade war', 'sales tax', 'Iran', 'oil crisis' and 'election'.
- Method: Robust Estimation with Quantization

Real data

- Estimated most influenced stocks:

PCA:

Robust PCA:

[Source: Wikipedia]

Real data

- Estimated most affected areas:

Trade war

pharmaceutical industry
(WBA, PEF, MDT)

Sales tax

technology
(INTC, ORCL, etc.)

Oil crisis

oil field (e.g., SLB)
and technology
(e.g., QCOM, WBA)

Election

technology
(e.g., GE, EMR, etc.)
and service (e.g., CVS,
SBUX, COST) stocks

Iran

food (KHC), finance
(UNH, BLK), technology
(LLY, ORCL), energy (EXC)
and others (GM, HD).

Summary

- PCA **works** if the related filter is 'strong' low-pass.
- With 'weak' low-pass filter, **boosting** method is applied.
- Numerical experiments on synthetic and stocks data.

Future Question: *Can we learn node centrality from data without knowing external sources?*

References

[Dong et al., 2016] Dong, X., Thanou, D., Frossard, P., and Vandergheynst, P. (2016).
Learning laplacian matrix in smooth graph signal representations.
IEEE Transactions on Signal Processing, 64(23):6160–6173.

[Egilmez et al., 2017] Egilmez, H. E., Pavez, E., and Ortega, A. (2017).
Graph learning from data under laplacian and structural constraints.
IEEE Journal of Selected Topics in Signal Processing, 11(6):825–841.

[Friedman et al., 2008] Friedman, J., Hastie, T., and Tibshirani, R. (2008).
Sparse inverse covariance estimation with the graphical lasso.
Biostatistics, 9(3):432–441.

[Hoffmann et al., 2018] Hoffmann, T., Peel, L., Lambiotte, R., and Jones, N. S. (2018).
Community detection in networks with unobserved edges.
arXiv preprint arXiv:1808.06079.

[Roddenberry et al., 2020] Roddenberry, T. M., Schaub, M. T., Wai, H.-T., and Segarra, S. (2020).
Exact blind community detection from signals on multiple graphs.
arXiv preprint arXiv:2001.10944.

[Roddenberry and Segarra, 2019] Roddenberry, T. M. and Segarra, S. (2019).
Blind inference of centrality rankings from graph signals.
arXiv preprint arXiv:1910.10846.

[Segarra et al., 2017] Segarra, S., Marques, A. G., Mateos, G., and Ribeiro, A. (2017).
Network topology inference from spectral templates.
IEEE Transactions on Signal and Information Processing over Networks, 3(3):467–483.

[Shen et al., 2017] Shen, Y., Baingana, B., and Giannakis, G. B. (2017).
Kernel-based structural equation models for topology identification of directed networks.
IEEE Transactions on Signal Processing, 65(10):2503–2516.

[Wai et al., 2018] Wai, H.-T., Segarra, S., Ozdaglar, A. E., Scaglione, A., and Jadbabaie, A. (2018).
Blind community detection from low-rank excitations of a graph filter.
arXiv:1809.01485v2.