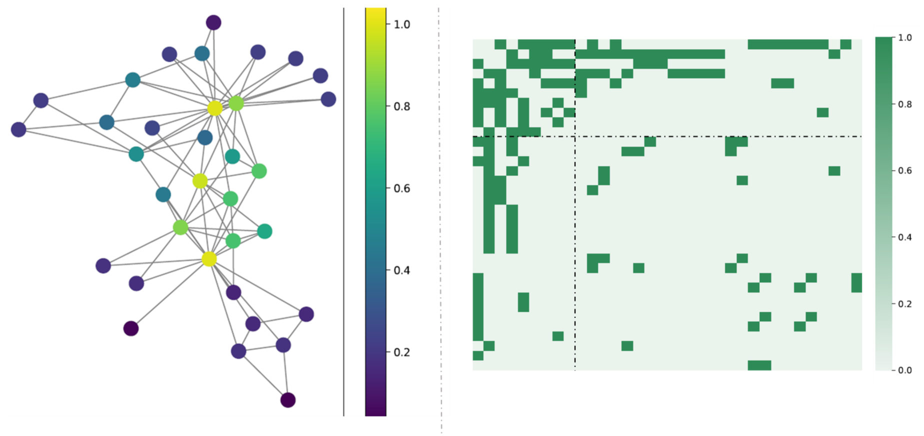
Learning Sparse Graphs with a Core-periphery Structure

Sravanthi Gurugubelli and Sundeep Prabhakar Chepuri
Indian Institute of Science, Bangalore, India

icassp 2022
Singapore, China, Online

Core-Periphery Structure

- Densely connected core vertices and sparsely connected periphery vertices



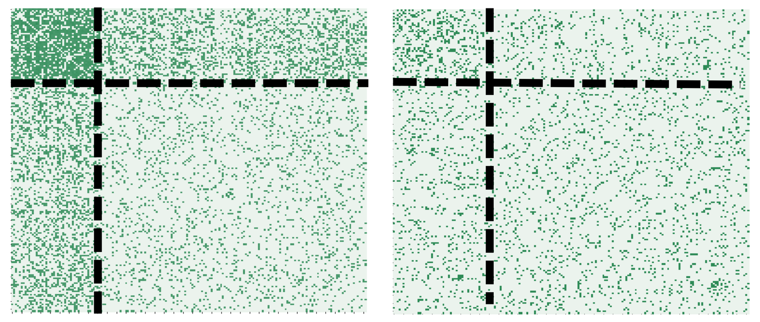
- Identifying the core and peripheral vertices helps in analyzing the central processes in complex networks

Prior Art

Existing algorithms estimate core scores given the network topology

- We often have access only to node attributes
- Underlying graph structure may not always be available

Conventional approaches to network topology inference do not readily incorporate a core-periphery structure



Adjacency matrices estimated from (a) the proposed method and (b) graphical lasso

Gaussian Graphical Model

Feature matrix of graph $\mathcal{G}$: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_d]$

$$\mathbf{x}_i \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Theta}^{-1}), \forall\, i = 1, \ldots, d$$

Encodes the conditional dependencies between the variables associated with the vertices of $\mathcal{G}$

Graphical lasso learns the sparsity pattern in $\boldsymbol{\Theta}$ by solving

$$\underset{\boldsymbol{\Theta} \succeq 0}{\text{maximize}} \quad \log\det \boldsymbol{\Theta} - \operatorname{tr}(\mathbf{S}\boldsymbol{\Theta}) - \lambda \|\boldsymbol{\Theta}\|_1$$

Does not readily incorporate core-periphery structure!

$\mathbf{S}$: empirical covariance matrix

$\lambda > 0$: regularization parameter that controls the sparsity in

Gaussian Graphical Model with a Core-periphery Structure

We model the dependence of the node attributes on the core scores through a latent graph structure

$\mathbf{c} \rightarrow \boldsymbol{\Theta} \rightarrow \mathbf{X}$

$\boldsymbol{\Theta} \rightarrow \mathbf{X}$:

$$p(\mathbf{X}|\boldsymbol{\Theta}) = \det \boldsymbol{\Theta} \, \exp(-\operatorname{tr}(\mathbf{S}\boldsymbol{\Theta}))$$

$$w_{ij} = 1 - c_i - c_j + e\log(d_{ij})$$

$\mathbf{c} \rightarrow \boldsymbol{\Theta}$:

$$p(\boldsymbol{\Theta}; \mathbf{c}) = \prod_{i,j=1}^{N} p(\Theta_{ij}; c_i, c_j) = Z \prod_{i,j=1}^{N} \exp\left(-\lambda w_{ij} |\Theta_{ij}|\right)$$

Proposed Learning Algorithm

$$\underset{\boldsymbol{\Theta} \succeq 0, \mathbf{c}}{\text{maximize}} \quad \log\det \boldsymbol{\Theta} - \operatorname{tr}(\mathbf{S}\boldsymbol{\Theta}) - \lambda \sum_{i,j=1}^{N} w_{ij} |\Theta_{ij}|$$

$$\text{s. to} \quad w_{ij} = 1 - c_i - c_j + e\log(d_{ij})$$

$$w_{ij} > 0, \quad i, j = 1, 2, \ldots, N$$

$$\sum_{i=1}^{N} c_i = M, \quad c_i \in [0, 1], \quad i = 1, 2, \ldots, N$$

To prevent the case where all the weights tend to zero

To fix the scale of the core scores

Updating the Graph

$$\underset{\boldsymbol{\Theta} \succeq 0}{\text{maximize}} \quad \log\det \boldsymbol{\Theta} - \operatorname{tr}(\mathbf{S}\boldsymbol{\Theta}) - \lambda \sum_{i,j=1}^{N} w_{ij} |\Theta_{ij}|$$

This is a convex program that can be solved using existing solvers, e.g., QUIC

Updating the Core Scores

$$\underset{c_1, \cdots, c_N}{\text{maximize}} \quad \sum_{i,j=1}^{N} |\Theta_{ij}|(c_i + c_j)$$

$$\text{s. to} \quad \sum_{i=1}^{N} c_i = M, \quad c_i \in [0, 1],$$

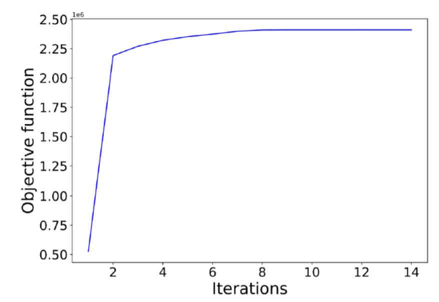
$$c_i + c_j < 1 + e\log(d_{ij}), \; i, j = 1, \cdots, N$$

This is a linear program and can be solved using standard off-the shelf solvers

Model Evaluation

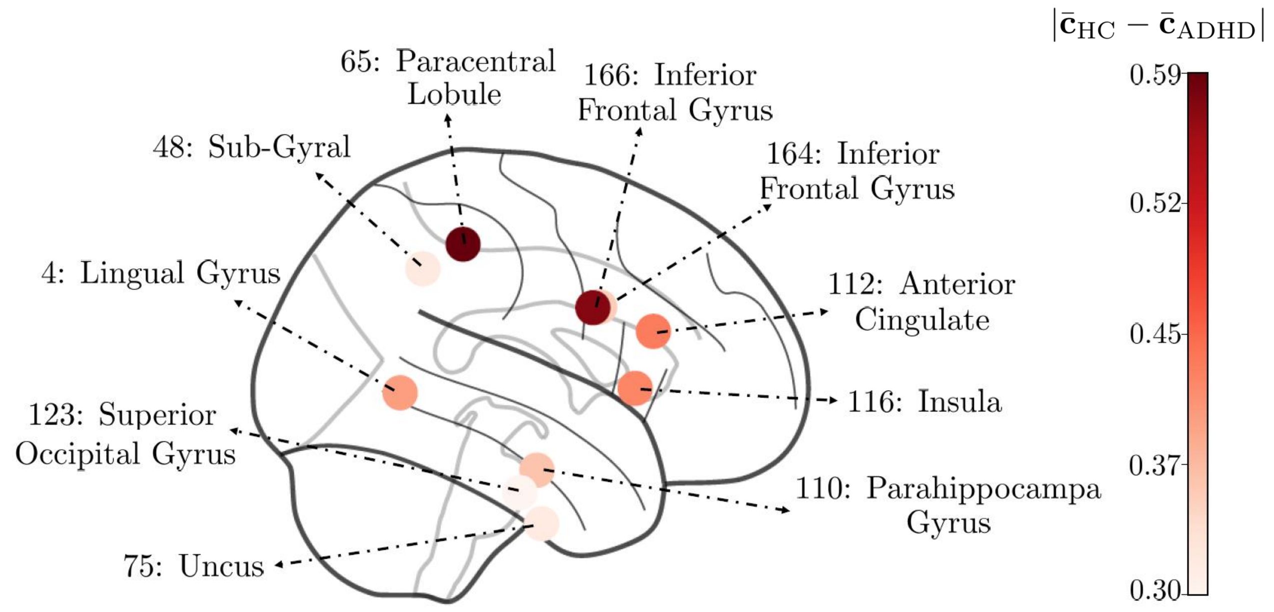
		Proposed	MINRES	Rombach	RandomWalk	k-cores
Celegans	$\|\boldsymbol{\Theta}_0 - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	41.940	41.821	39.076	40.877	39.051
	$\||\boldsymbol{\Theta}| - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	32.642	32.841	32.538	32.707	32.748
Cora	$\|\boldsymbol{\Theta}_0 - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	55.488	55.434	54.690	55.326	54.909
	$\||\boldsymbol{\Theta}| - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	47.626	55.722	47.884	47.983	47.625
London underground	$\|\boldsymbol{\Theta}_0 - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	79.216	79.249	78.7563	79.338	79.169
	$\||\boldsymbol{\Theta}| - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	78.818	78.905	78.811	78.858	78.856
Twitter	$\|\boldsymbol{\Theta}_0 - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	134.692	137.142	124.112	131.278	129.221
	$\||\boldsymbol{\Theta}| - \boldsymbol{\Theta}_{\text{ideal}}\|_F^2$	110.526	111.837	111.427	111.429	110.526

The core-periphery partitioning of the networks by the proposed method is similar to the others, in spite of not knowing the network directly!



The proposed algorithm converges in about 10 iterations

Brain Network Analysis



The regions with a large difference in the cores scores of the two groups coincide with the regions that have differences in activation for healthy individuals and patients with ADHD

sravanthig@iisc.ac.in,spchepuri@iisc.ac.in
https://ece.iisc.ac.in/~spchepuri//