Network Reconstruction - Bayesian Sequential Inference of Sparse Network Connectivity

Jun Kataoka, Jayanth Sivakumar, Eric Peña, Yiming Che, Changqing Cheng*

State University of New York at Binghamton

BINGHAMTON UNIVERSITY | THOMAS J. WATSON SCHOOL OF ENGINEERING AND APPLIED SCIENCE

Abstract

Most existing work assumes the network structure is known or is readily available. However, the network topology could be radically altered due to an adversarial attack or a power outage. In this work, we propose a novel Bayesian sequential learning algorithm to adaptively reconstruct network connectivity. A sophisticated method of sequentially selecting the nodes is implemented using the between-node expected improvement. This algorithm has been applied to real network data: IEEE118-Bus System, and Barabasi-Albert network for $m = 1$ and $m = 2$. The performance is measured and compared against randomly selecting the nodes. Our algorithm is an improvement over traditional reconstruction efforts when faced with limited data and is robust to varying noise levels.

Objective

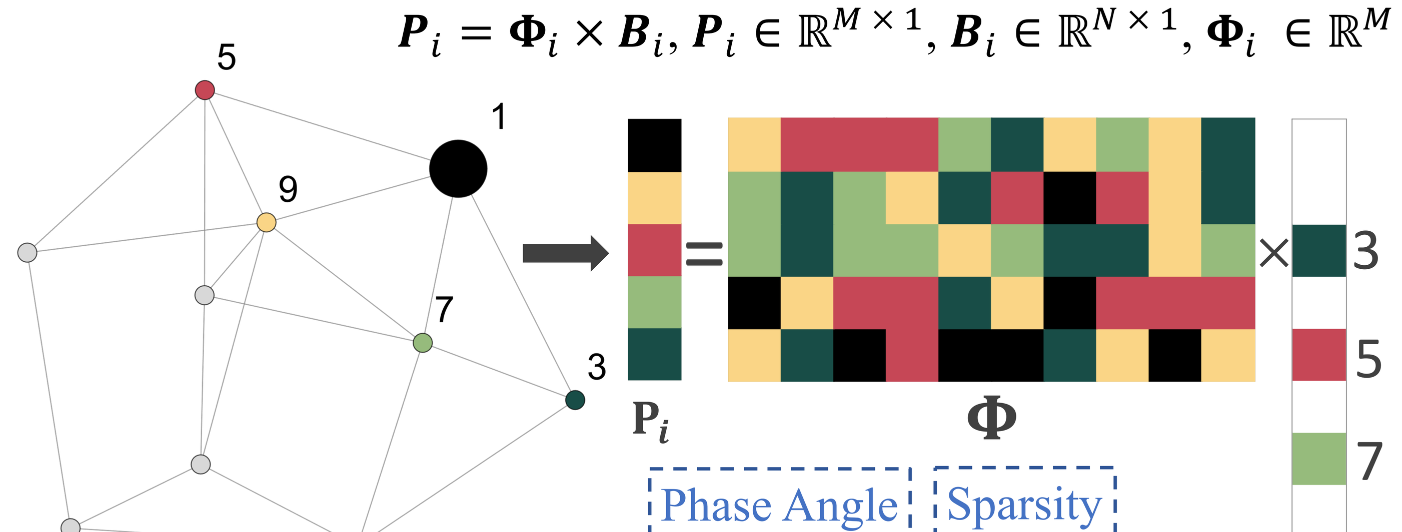
Given an undirected sparse network structure over time:

- Reconstruct the network node by node
- Find the order in which the nodes are to be reconstructed

Background

Lasso Convex Optimization [1]

$$\boldsymbol{P}_i = \boldsymbol{\Phi}_i \times \boldsymbol{B}_i,\ \boldsymbol{P}_i \in \mathbb{R}^{M\times 1},\ \boldsymbol{B}_i \in \mathbb{R}^{N\times 1},\ \boldsymbol{\Phi}_i \in \mathbb{R}^{M\times N}$$



$$argmin_{\boldsymbol{B}_i} \|\boldsymbol{P}_i - \boldsymbol{\phi}\boldsymbol{B}_i\|_2^2 + \lambda\|\boldsymbol{B}_i\|_0$$

$\boldsymbol{B}_i$

Power | Connectivity

Data Simulation

$$P_{ij} = \frac{|V_i V_j|}{X_{ij}}\sin(\varphi_i - \varphi_j) \approx X_{ij}^{-1}(\varphi_i - \varphi_j)$$

Adding Noise:
$\boldsymbol{P}_i + (\mathcal{N}(0,\ \boldsymbol{\Sigma}_{\boldsymbol{P}_i})\cdot\sigma)$

$$B_{ij} = \begin{cases} \sum_{k=1}^{N} X_{ik}^{-1} & \text{for } i = j \\ -X_{ij}^{-1} & \text{for } i \neq j \end{cases}$$

M = Time instances for simulated power flow
N = Total Nodes in the network

$\varphi_i = (\omega + \Delta\omega_i)t,\ \omega = 2\pi\times 50,\ \Delta\omega_i \sim \mathcal{N}(0, 20)$, at time stamps $t_1, t_2, \ldots, t_M$, with $M \ll N$

Methodology

Spike and Slab Priors [2, 3]

- A sparse Spike-and-Slab prior distribution is placed on all edges
- The connectivity learned from the reconstructed nodes will be incorporated as a prior to select the next node and will then update the prior knowledge

$$p(\boldsymbol{B}_i|\boldsymbol{z}_i) = \prod_{j=1}^{N} p(B_{ij}|z_{ij}) = \prod_{j=1}^{N}\left[(1 - z_{ij})\delta(B_{ij}) + z_{ij}\mathcal{N}(B_{ij}|0, \tau_0)\right]$$

Point probability mass centered at spike ($\delta(B_{ij})$); Latent binary variable (z_{ij}); Controls the width of the slab (τ_0)

- $z_{ij} \sim Bernoulli(z_{ij}\,|\,\gamma_{ij})$
- γ_{ij} is a hyperparameter that controls B_{ij}
- If $\gamma_{ij} = 1$ then $z_{ij} = 1$ and $B_{ij} \neq 0$

Expected Improvement [4, 5]

Improvement by adding the node s is given by:

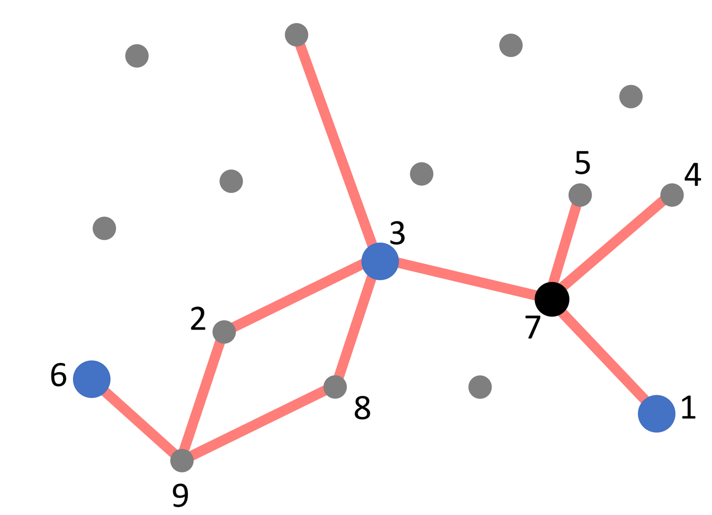
$$EI(s) = E_{Y\sim\mathcal{N}(\mu_s,\sigma_s^2)}[I(s)] = E_{\eta\sim\mathcal{N}(0,1)}[I(s)]$$

$$EI(s) = (\mu_s - f(\theta_k))\Gamma\left(\frac{\mu_s - f(\theta_k)}{\sigma_s}\right) + \sigma_s\gamma\left(\frac{\mu_s - f(\theta_k)}{\sigma_s}\right)$$

$$\mu_s = f(\theta_k \cup s) = \sum_{i=1}^{N}\left\|\mathrm{P}_i - \phi\hat{\mathrm{B}}_i\right\|_2 - MSE(\theta_k \cup s)$$

Next node to select (max EI): $s_{k+1} = \underset{s\in V\setminus\theta_k}{\operatorname{argmax}} EI(s)$

Sequential Network Reconstruction



- When node 7 is first selected and evaluated, $p(x_{7,3})$ and $p(x_{7,6})$ are obtained via Spike and Slab
- In the undirected network setting, $p(x_{3,7})$ and $p(x_{6,7})$ also becomes available
- In this case, owing to significant $p(x_{7,3})$ and negligible $p(x_{7,6})$, addition of node 7 diminishes the utility to select node 3
- Thus, node 6 is favored in the set through expected improvement

Results

Reconstruction Error:

$$\text{Error}_{\hat{\mathbf{A}}} = \frac{\|\boldsymbol{A} - \hat{\boldsymbol{A}}\|_2^2}{\|\boldsymbol{A}\|_2^2}$$

Estimated connectivity ($\hat{\boldsymbol{A}}$); Connectivity ($\boldsymbol{A}$)

- $N = 118, M = 60$ for both networks
- Conducted 30 experiments, average plotted

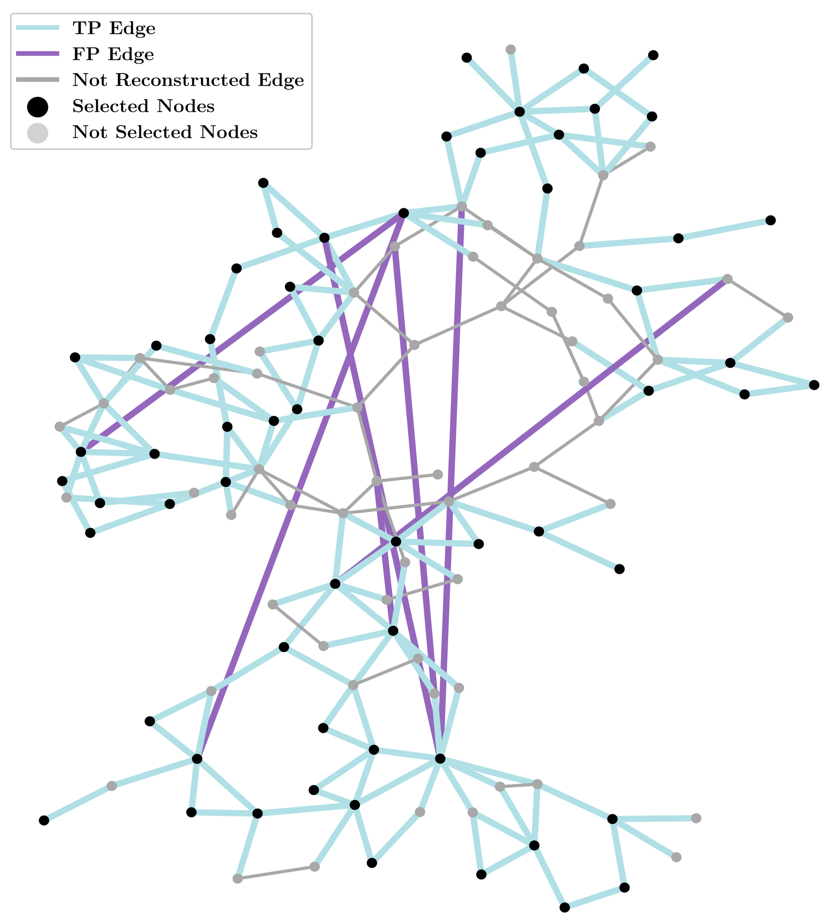


Fig. 1: Reconstruction of IEEE-118 for subset of 60 nodes with noise $\sigma = 0.03$

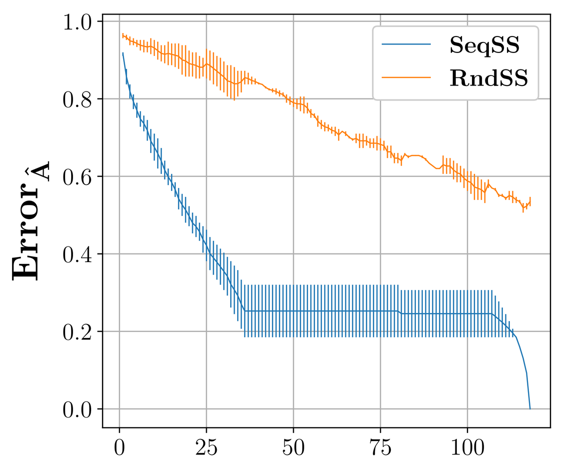


Fig. 2: Recon. Error for BA $m = 1$ with noise $\sigma = 0.05$

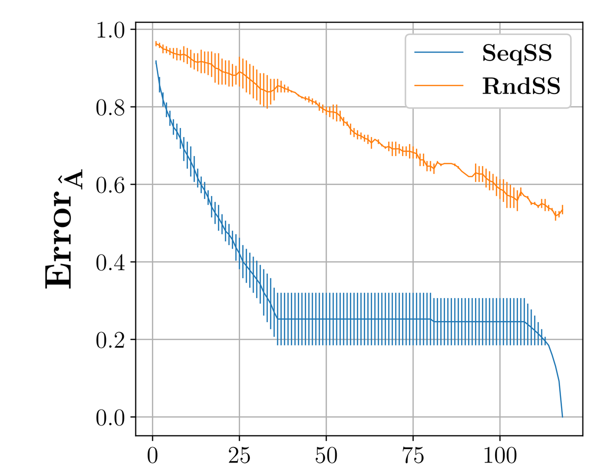


Fig. 3: Recon. Error for BA $m = 2$ with noise $\sigma = 0.05$

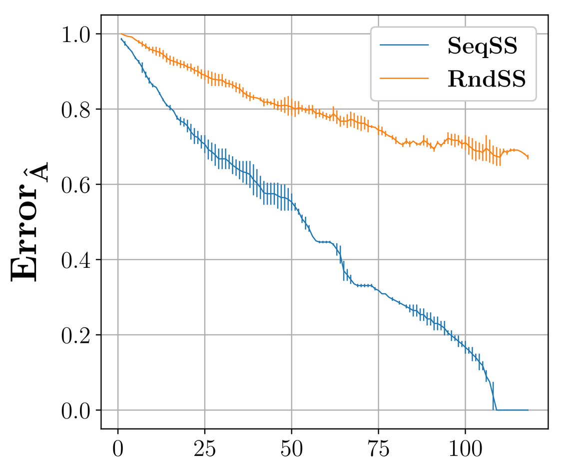


Fig. 4: Recon. Error for IEEE-118 with noise $\sigma = 0.05$

Conclusions

- The proposed method outperforms the reconstruction done by random node selection in terms of accuracy and robustness
- This study has the potential to significantly scale up the reconstruction radically transform the operation of various realistic networked systems including power grid, transportation, and communication networks particularly in a host of military operation scenarios
- This research finds its critical application in the host of military operation scenarios

Contact

Changqing Cheng
State University of New York at Binghamton
4400 Vestal Parkway East, Vestal, NY 13902
ccheng@binghanmton.edu

References

1. X. Han, Z. Shen, W.-X. Wang, and Z. Di, "Robust reconstruction of complex networks from sparse data," *Physical Review Letters*, vol. 114, no. 2, p. 028701, 2015, doi: 10.1103/PhysRevLett.114.028701.
2. D. Hernández-Lobato, J. M. Hernández-Lobato, and P. Dupont, "Generalized spike-and-slab priors for bayesian group feature selection using expectation propagation," *Journal of Machine Learning Research*, vol. 14, no. 23, pp. 1891–1945, 2013.
3. J. M. Hernández-Lobato, D. Hernández-Lobato, and A. Suárez, "Expectation propagation in linear regression models with spike-and-slab priors," *Machine Learning*, vol. 99, no. 3, pp. 437–487, 2015, doi: 10.1007/s10994-014-5475-7.
4. J. Wilson, F. Hutter, and M. Deisenroth, "Maximizing acquisition functions for Bayesian optimization," in *32nd Conference on Neural Information Processing Systems, NeurIPS 2018*, Montréal, Canada, Dec. 2018, pp. 9884–9895.
5. D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, vol. 13, no. 4, pp. 455–492, Dec. 1998, doi: 10.1023/A:1008306431147.