Correlated Stochastic Block Models: Exact Graph Matching with Applications to Recovering Communities

Miklós Z. Rácz
Princeton University, ORFE Department

Anirudh Sridhar
Princeton University, Electrical and Computer Engineering Department

Q: How can information from multiple correlated networks improve community recovery algorithms?

Example: Correlated Social Networks

Users have similar but not identical information across networks

Correlated Networks

Mathematical model: Correlated Stochastic Block Models

Graph Matching / Network Alignment

Step 1: Graph Matching / Network Alignment

Step 2: Network Synthesis

Step 3: Find communities in the synthesized network

Recovering communities from correlated networks

Objectives

Exact Community Recovery: When can we exactly recover the community labels \( s_i \) from \( (G_1, G_2) \)?

Prior work: Phase transition in the log-degree regime \([1,2]\)

Exact recovery is possible

Exact recovery is information-theoretically impossible

Exact graph matching: When can we exactly recover the alignment \( \pi \) from \( (G_1, G_2) \)?

Prior work: Phase transition in the log-degree regime \([3]\)

Exact recovery is possible

Exact recovery is information-theoretically impossible

Theoretical analysis:

If \( \pi = \pi^* \), synthesized network is \( SBM(n, p, q') \)

Then

\[ p' = a \left( 1 - \frac{a}{2} \right) \frac{np}{n} \]

\[ q' = a \left( 1 - \frac{a}{2} \right) \frac{np}{n} \]

Exact community recovery possible from \( (G_1, G_2) \) but not from \( G_1 \) or \( G_2 \) if

\[ \sqrt{\frac{2}{1 - \frac{a}{2}}} > 1 \]

\[ \sqrt{\frac{2}{1 - \frac{a}{2}}} < 1 \]

Citations


This work was supported in part by NSF Grant DMS 1811724.