Inferential Network Analysis with Exponential Random Graph Models

Bruce A. Desmarais

March 21, 2016

What does it mean to model the network?

Construct a probability distribution that

accurately approximates the network

Why build models?

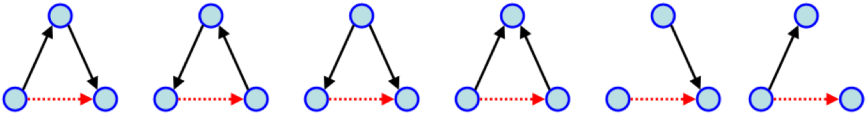
- Test hypotheses
  Example: Does the cosponsorship network exhibit reciprocity?
- Simulation for theoretical exploration
  Example: How should seats be assigned in a classroom to encourage cross-racial friendships?
- Tie prediction
  Example: Will Canada attack next year?

Modeling Interdependence

Two Classes of Questions: Covariate and Interdependence

1. Covariate
   - Do legislators in the same political party collaborate more frequently than those in opposite parties?
   - Do states with democratic governments have more alliances than those with autocratic regimes?
2. Interdependence
   - Are two states at war with the same third state less likely to be at war with each other?
   - Are there popularity effects in the choice of co-authors?

ERGM: integrate effects for any forms of (1) and (2) into a unified model of a network.



Parametric Probabilistic Modeling and the Likelihood Framework of Inference

We observe x, a draw of a random variable X..

$$X \sim f(X, \boldsymbol{\theta})$$

f is a family of probability distributions and $\boldsymbol{\theta}$ is unknown.
X could be

- The dependent variable in a regression model
- An adjacency matrix
- The text in a document

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}} [f(x, \boldsymbol{\theta})]$$

1. In many cases, $\hat{\boldsymbol{\theta}}_{MLE}$ is asymptotically normally distributed
2. If f is exponential family, $\ln [f(x, \boldsymbol{\theta})]$ globally concave in $\boldsymbol{\theta}$

The Exponential Random Graph Model (ERGM)

The probability (likelihood function) of observing network N is:

$$\mathcal{P}(N, \boldsymbol{\theta}) = \frac{\exp\{\boldsymbol{\theta}' \boldsymbol{h}(N)\}}{\sum_{N^* \in \mathcal{N}} \exp\{\boldsymbol{\theta}' \boldsymbol{h}(N^*)\}}$$

Decomposition:

$$\underbrace{\boldsymbol{h}(N)}_{Net\ Stats} \qquad \underbrace{\boldsymbol{\theta}}_{Effects} \qquad \underbrace{\exp\{\boldsymbol{\theta}' \boldsymbol{h}(N)\}}_{+\ Weight} \qquad \underbrace{\sum_{N^* \in \mathcal{N}} \exp\{\boldsymbol{\theta}' \boldsymbol{h}(N^*)\}}_{Normalizer}$$

Flexible: $\boldsymbol{h}$ can capture virtually any form of interdependence among the edges + covariates

Normalizing constant can make estimation difficult

Definining $\boldsymbol{h}$

How would we measure **reciprocity**?

A statistic we would expect to be high if ties were reciprocated a lot and low if they were not reciprocated.

Unpacking $\boldsymbol{h}$

- Dyadic Covariate

$$h_D(N, X) = \sum_{ij} N_{ij} X_{ij}$$

- Sender Covariate

$$h_{VS}(N, VS) = \sum_{i} VS_i \sum_{j \neq i} N_{ij}$$

- Receiver Covariate

$$h_{VR}(N, VR) = \sum_{i} VR_i \sum_{j \neq i} N_{ji}$$

- Reciprocity

$$h_R(N) = \sum_{i<j} N_{ij} N_{ji}$$

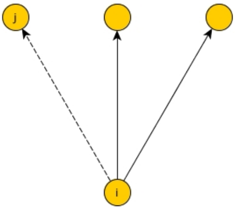
Unpacking $\boldsymbol{h}$

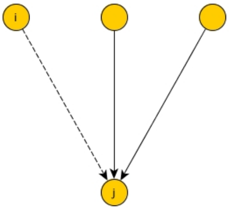
- Popularity

$$h_P(N) = \sum_{i,j,k} N_{ji}N_{ki} + N_{kj}N_{ij} + N_{ik}N_{jk}$$

- Sociality

$$h_S(N) = \sum_{i,j,k} N_{ij}N_{ik} + N_{jk}N_{ji} + N_{ki}N_{kj}$$

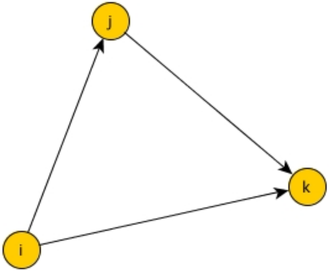




Unpacking $\boldsymbol{h}$

- Transitivity

$$h_T(N) = \sum_{i} \sum_{i \neq j,k} N_{ij} N_{ik} N_{jk}$$



- For detailed discussions on the selection of network statistics, see Snijders et al. (2006) and Goodreau (2007)

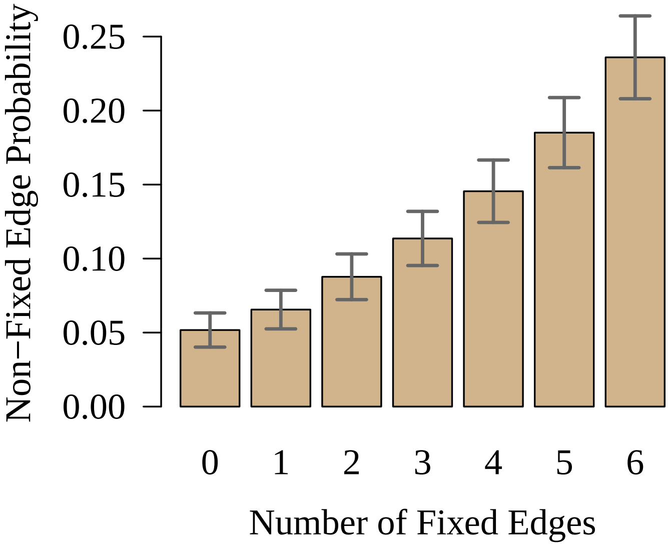
Interpretation of ERGM

ERGM offers an incredibly flexible model – it can be used to investigate individual, dyad, node and network level effects.

Two levels of interpretation

1. (**Network**) The relative likelihood of observing N^{j+} to observing N^j is $\exp(\theta_j)$, where
   - θ_j is the estimate of the parameter that corresponds to statistic j.
   - N^{j+} is one unit greater than N^j on statistic j (e.g., one more closed triangle, one more edge), *ceteris paribus*.

2. (**Edge**) $P(N_{ij} = 1 | N_{-ij}, \boldsymbol{\theta}) = \text{logit}^{-1}\left(\sum_{r=1}^{k} \theta_r \delta_r^{(ij)}(N)\right)$
   - N_{-ij} indicates the network excluding N_{ij}
   - $\delta_r^{(ij)}(N)$ is equal to the change in h_r when N_{ij} is changed from zero to one
   - $\text{logit}^{-1}(x) = 1/(1 + \exp(-x))$ (i.e., inverse logit function)


Non−Fixed Edge Probability
0.25
0.20
0.15
0.10
0.05
0.00
0
1
2
3
4
5
6
Number of Fixed Edges

Wrap-up

- ERGM
  - Test covariate effects
  - Test interdependence effects
  - Nothing like it in the literature
- Extensions
  - Weighted Ties
  - Network time series
  - Bipartite networks