Discrete Temporal Models of Social Networks

Steve Hanneke & Eric Xing

MACHINE LEARNING DEPARTMENT
Carnegie Mellon

Outline

- The Setting
- Exponential Random Graphs
- Extending ERGMs for Evolving Networks
- Estimation
- Conclusions and Future Work

Evolving Networks

- We observe the network at discrete, evenly spaced time points t = 1,2,…,T,
- The observed network at time t: $N^{(t)}$.
- We want a statistical model of the network's evolution.

Markov Assumption

- To simplify things, assume the network observed at time t is independent of the rest of history, given knowledge of the network at time t-1 (paper relaxes this).

$$P(N^{(T)},N^{(T-1)},\ldots,N^{(2)},N^{(1)}) =$$

$$P(N^{(T)}|N^{(T-1)})\cdot P(N^{(T-1)}|N^{(T-2)})\cdots P(N^{(2)}|N^{(1)})\cdot P(N^{(1)})$$

- What should the conditional look like?

Exponential Random Graphs

- Very general families for modeling a single static network observation.

$$P(N) = \exp\{\theta \cdot u(N) - \ln Z(\theta)\}$$

- Can estimate the θ parameters by MCMC MLE

ERGM Example

- Classic example: (Frank & Strauss 1986)

$u_1(N)$ = # edges in N

$u_2(N)$ = # 2-stars in N

$u_3(N)$ = # triangles in N

$$P(N) \propto \exp\{\theta_1 u_1(N) + \theta_2 u_2(N) + \theta_3 u_3(N)\}$$

Temporal Extension of ERGMs

- Can we build on all the work on ERGMs when designing a temporal model?

$$P\left(N^{(t)} \middle| N^{(t-1)}\right) = \exp\left\{\theta \cdot \Psi\left(N^{(t)}, N^{(t-1)}\right) - \ln Z\left(\theta, N^{(t-1)}\right)\right\}$$

An Example

$$P\left(N^{(t)} \middle| N^{(t-1)}\right) = \exp\left\{\theta \cdot \Psi\left(N^{(t)}, N^{(t-1)}\right) - \ln Z\left(\theta, N^{(t-1)}\right)\right\}$$

- Say the network has a single relation, and its value is either 0 or 1 (e.g., “friends” or “not friends”).
- Let $A_{ij}^{(t)}$ equal the value of the relation between i^{th} actor and j^{th} actor.

An Example (continued)

$$P\left(N^{(t)} \middle| N^{(t-1)}\right) = \exp\left\{\theta \cdot \Psi\left(N^{(t)}, N^{(t-1)}\right) - \ln Z\left(\theta, N^{(t-1)}\right)\right\}$$

- "Continuity": $\Psi_1\left(N^{(t)}, N^{(t-1)}\right) = \sum_{ij} A_{ij}^{(t)} A_{ij}^{(t-1)} + \left(1 - A_{ij}^{(t)}\right)\left(1 - A_{ij}^{(t-1)}\right)$
- "Reciprocity": $\Psi_2\left(N^{(t)}, N^{(t-1)}\right) = \sum_{ij} A_{ij}^{(t)} A_{ji}^{(t-1)}$
- "Transitivity": $\Psi_3\left(N^{(t)}, N^{(t-1)}\right) = \sum_{ijk} A_{ik}^{(t)} A_{ij}^{(t-1)} A_{jk}^{(t-1)}$
- "Density": $\Psi_4\left(N^{(t)}, N^{(t-1)}\right) = \sum_{ij} A_{ij}^{(t)}$

An Example (continued)

$$P\left(N^{(t)} \middle| N^{(t-1)}\right) \propto \exp\{$$

$$\theta_1 \sum_{ij} \left[A_{ij}^{(t)} A_{ij}^{(t-1)} + (1 - A_{ij}^{(t)})(1 - A_{ij}^{(t-1)})\right]$$

$$+ \theta_2 \sum_{ij} A_{ij}^{(t)} A_{ji}^{(t-1)}$$

$$+ \theta_3 \sum_{ijk} A_{ik}^{(t)} A_{ij}^{(t-1)} A_{jk}^{(t-1)}$$

$$+ \theta_4 \sum_{ij} A_{ij}^{(t)} \}$$

Maximum Likelihood Estimation

- Approximate MLE by MCMC (Z intractable)

$$\ln P\left(N^{(T)}, N^{(T-1)}, \ldots, N^{(2)} \middle| N^{(1)}\right) = \theta \cdot \sum_{t=2}^{T} \Psi\left(N^{(t)}, N^{(t-1)}\right) - \sum_{t=2}^{T} \ln Z\left(\theta, N^{(t-1)}\right)$$

$$\nabla \ln P\left(N^{(T)}, N^{(T-1)}, \ldots, N^{(2)} \middle| N^{(1)}\right) = \sum_{t=2}^{T} \Psi\left(N^{(t)}, N^{(t-1)}\right) - \sum_{t=2}^{T} E_{\theta}\left[\Psi\left(N, N^{(t-1)}\right) \middle| N^{(t-1)}\right]$$

- Use gradient ascent, using MCMC to estimate the expectation on each iteration (as in ERGM).

Estimation Toy Example

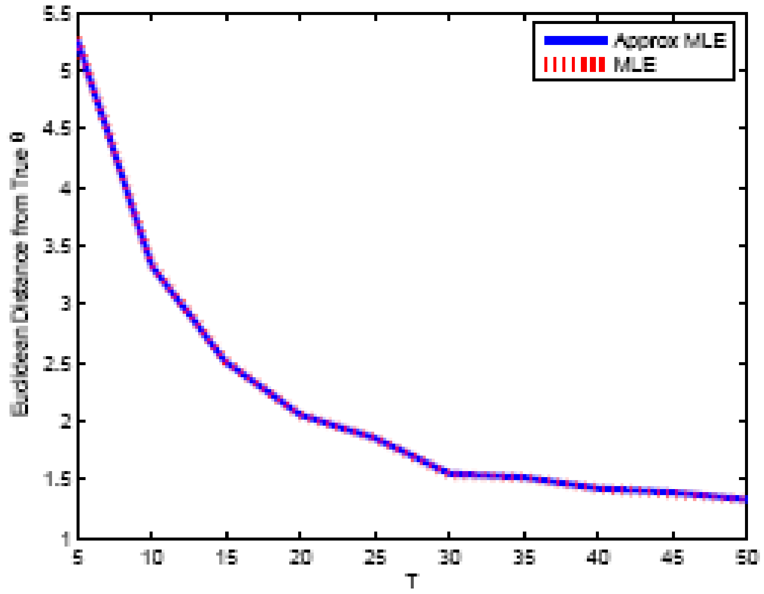
- Generate a series of 10 networks from the example model
- True model has $\theta_1=0$, $\theta_2=5$, $\theta_3=0$, $\theta_4=-20$ (ie, reciprocity and density only)

- Estimated parameters:

$\theta_1=-0.5$, $\theta_2=4.2$, $\theta_3=-0.08$, $\theta_4=-20.2$

Simulation

- Uses the example model
- True parameters random in [0,10)
- 100 actors



What's it good for?

- Hypothesis Testing
- Data Exploration
- Foundation for Learning

An idea for specifying a model

- A network might be decomposable into different types of “motifs” (e.g., “hub & spokes”, “k-clique”, “triangle”,…).
- Write the potential functions to encode your understanding about how each motif evolves.
- It’s nice because we can “plug in” our intuition about the data.

Conclusions

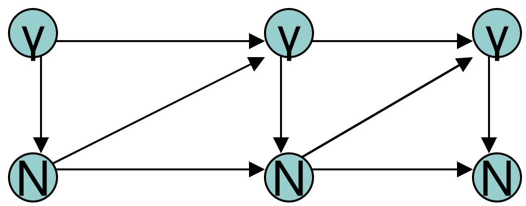
- Pretty much anything you can do with ERGMs can be adapted for this temporal model.

Future Work

- This type of model converges to an ERGM stationary distribution; can we give a general characterization of that distribution? Can we characterize the set of temporal models that give rise to a particular ERGM stationary distribution?
- Continuous time Markov chain

Future Work (continued)

- Latent variables to explain the behavior in a simple way (e.g., groups).



- We would like to preserve the model generality and retain the ability to “plug in” our knowledge of the data, while still allowing for generic inference algorithms.