Network Models

Eric Xing

Lecture 22, November 18, 2013

Reading: See class webpage

© Eric Xing @ CMU, 2013
Network Data

- Disease Spread
- Food Web
- Internet
- Social Network

Electronic Circuit
Social networks
Information on a social network

- Social graph
  - Friendship networks
  - User-ads network ...
- Text
  - News feed
  - Messages
  - Ads text ...
- Images
  - Album
  - Random posts
  - Ads figures ...
- Demographics
  - Age, occupation ...

- Very high-dimensional
- Non-independent
- Insufficient training data (this is true even we use the whole web)
- Hard to optimize and interpret
Network analysis -- visualization
Visualization, cont’d
Global topological measures

- Indicate the gross topological structure of the network

Connectivity (Degree)  
Path length  
Clustering coefficient

© Eric Xing @ CMU, 2013
Local network motifs profiles

- Regulatory modules within the network

SIM

MIM

FBL

FFL

© Eric Xing @ CMU, 2013
Models for Macroscopic and Descriptive Network Analysis

A. Random Networks [Erdos and Rényi (1959, 1960)]

\[ P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!} \]

Mean path length \( \sim \ln(k) \)

Phase transition:

Connected if: \( p \geq \ln(k) / k \)

B. Scale Free [Price, 1965 & Barabasi, 1999]

\[ P(k) \sim k^{-\gamma}, \quad k \gg 1, \quad 2 < \gamma \]

Mean path length \( \sim \ln\ln(k) \)

Preferential attachment. Add proportionally to connectedness

C. Hierarchical

Copy smaller graphs and let them keep their connections.
Summary

- Impressive graphs!
- Seemingly impressive statements about network properties
  --- scale free, small world, ...
- Can even fit data with some models
- But … What about details?
  - Can we infer …
    - e.g., the role of every node, the meaning of every edge?
    - the network topology itself?
  - Can we predict …
  - Can we simulate …

- Most current analyses tell **BIG** stories, or obvious stories, but not so useful to serious detail-hunters
“We can express a role through a relation (or set of relations) and thus a social system by the inventory of roles. If roles equate to positions in an exchange system, then we need only identify particular aspects of a position. But what aspect?”

Structural Equivalence:

Two actors are structurally equivalent if they have the same types of ties to the same people.
Structural Equivalence
Structural Equivalence

Graph reduced to positions
Blockmodeling is the process of identifying these types of positions. A **block** is a section of the adjacency matrix - a “group” of structurally equivalent ACTORS.
Cohesive Subgroups

\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 \\
2 & . & 0 & 0 & 1 & 1 & 0 \\
3 & 1 & 0 & . & 1 & 0 & 0 \\
4 & 1 & 0 & 1 & . & 0 & 0 \\
5 & 0 & 1 & 0 & 0 & 1 & 0 \\
6 & 0 & 1 & 0 & 0 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{ccccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 \\
2 & 1 & 0 & 0 & 1 & 0 & 0 \\
3 & 1 & 0 & 0 & 1 & 0 & 0 \\
4 & 0 & 1 & 0 & 1 & 0 & 1 \\
5 & 0 & 0 & 1 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{array}
\]
Stochastic Cohesive Subgroups
Regard each network tie as a \textit{random variable} (often binary)

\begin{align*}
X_{ij} & = 1 \text{ if there is a network link from person } i \text{ to person } j \\
& = 0 \text{ if there is no link,} \\
& \text{for } i, j \text{ members of some set of actors } N.
\end{align*}

A \textit{directed network}: $X_{ij}$ and $X_{ji}$ are distinct.
A \textit{non-directed network}: $X_{ij} = X_{ji}$

Formulate a hypothesis about interdependencies and construct a \textit{dependence graph}

The \textit{dependence graph} represents the contingencies among network variables $X_{ij}$ (e.g., defined on cliques), i.e., a set of "potential functions".
The Hammersley-Clifford Theorem

$$\Pr(X = x) = p^*(x) = \frac{1}{c} \exp \left\{ \sum_{\text{all cliques}} \lambda_A z_A \right\}$$

where:

- the summation is over all cliques $A$;
- $z_A = \prod_{x_{ij}\in A} x_{ij}$ is the *network statistic* corresponding to the clique $A$;
- $\lambda_A$ is the parameter corresponding to clique $A$;
- $c = \Sigma_X \exp\{\Sigma_A \lambda_A z_A(x)\}$ is a normalising constant

(Besag, 1974)
Bernoulli Blockmodels

- Suppose actors are either in block 1 or 2, and pairwise potentials

- Hammersley-Clifford:
  \[ \Pr(\mathbf{X} = \mathbf{x}) = (1/c) \exp\{\sum_{i,j} \lambda_{ij} x_{ij}\} \]

- Block homogeneity:
  \[ \lambda_{ij} = \theta_{11} \text{ if } i \text{ and } j \text{ both in block 1} \]
  \[ \lambda_{ij} = \theta_{12} \text{ if } i \text{ in block 1 and } j \text{ in block 2}, \text{ etc.} \]

- \[ \Pr(\mathbf{X} = \mathbf{x}) = (1/c) \exp\{\theta_{11} L_{11} + \theta_{12} L_{12} + \theta_{21} L_{21} + \theta_{22} L_{22}\} \]
  where \( L_{rs} \) is the number of edges from block \( r \) to block \( s \).

- Extendable to multiple blocks
Richer Network Tomography

- Multi-role of every node
- Context dependent role-instantiation
- Role dynamics
Mixed Membership of Actors

- Sampson’s Monastery (Sampson 1968)

- What are the factions?
- How do the monks involve themselves in these factions? Multi-role? Time evolution?
- How do factions relate to one another?
Motivation

- In many networks (e.g., biological network, citation networks), each node may be "multiple-class", i.e., has multiple functional/topical aspects.
- The interaction of a node (e.g., a protein) with different nodes (partners) may be under different function context.
- Prior knowledge of group interaction may be available.
A Generative Perspective

\[ \theta_j \]

\[ \theta_i \]

\[ z_i \rightarrow j \]

\[ e_{ij} \sim \beta_{z_i \rightarrow j, z_j \leftarrow i} \]
A Generative Perspective

- Mixed membership and peer-specific role realizations:

\[ \theta_i, \theta_j \]

MM vector of node \( i \)

MM vector of node \( j \)

MM indicators of node \( i \) as initiator

MM vector of node \( j \) as acceptor

\[ \{Z_{i \rightarrow 1}, Z_{i \rightarrow 2}, \ldots, Z_{i \rightarrow j}, \ldots, Z_{i \rightarrow n}\} \]

\[ \{Z_{j \rightarrow 1}, Z_{j \rightarrow 2}, \ldots, Z_{j \rightarrow i}, \ldots, Z_{j \rightarrow n}\} \]

\[ R_{ij} \]
Mixed Membership Stochastic Blockmodel [Airoldi, Blei, Fienberg and Xing, 2008]

1. \( \{\theta_i\}_{i=1}^{N} \sim p(\theta|\alpha) \equiv \text{Dirichlet}(\theta; \alpha) \)

   sample mixed membership vectors.

2. For each actor \( v_j \) that actor \( v_i \) possibly interacts with:
   - \( z_{i\rightarrow j} \sim \text{Multinomial}(z|\theta_i) \)
     sample an indicator for \( v_i \);
   - \( z_{i\leftarrow j} \sim \text{Multinomial}(z|\theta_j) \)
     sample an indicator for \( v_j \);
   - \( e_{ij} \sim \text{Bernoulli}(e|z_{i\rightarrow j}^TBz_{i\leftarrow j}) \)
     sample a link.
Modeling Ideas

- Hierarchical Bayes
  - Latent MM vector encodes degree of semantic aspects
  - Latent structures on semantic-dependent interactions

- Combination of 2 classes of models

1. Models of mixed membership (e.g., admixture model)
2. Network models (block models)

= Stochastic block models of mixed membership
The Hierarchical Model

Mixed membership Vectors (latent*)

Interactions (observed*)

Group-to-group patterns (latent*)

\[ \Pr (y_{ij} = 1 | \theta_i, \theta_j, \beta) = \theta_i^T \beta \theta_j \]

© Eric Xing @ CMU, 2013
Variational EM

The Complete likelihood:

\[ p(r, z, \theta, \gamma) \propto \prod_i \theta_i^{\sum_j z_{i\rightarrow j} + \alpha - 1} \times \gamma^{\sum_i r_{i,j} z^g_{i \rightarrow j} z^h_{i \rightarrow j} + \beta_1 - 1} (1 - \gamma_{g,h})^{\sum_i (1 - r_{i,j}) z^g_{i \rightarrow j} z^h_{i \rightarrow j} + \beta_2 - 1} \]

- The mean-field approximation

\[ q(r, z, \theta, \gamma | \alpha, \beta) = \left( \prod_{i=1}^N q(\theta_i | \mu_i) \right) \times \left( \prod_{g=1}^K q(\gamma_{g,h} | \nu_{g,h}) \right) \times \left( \prod_{i=1}^N q(z_{i \rightarrow j} | \phi_{i \rightarrow j}) q(z_{i \leftarrow j} | \phi_{i \leftarrow j}) q(r_{i,j} | \phi_{i,j}) \right) \]

- Parameters: \( K^*, \alpha, \beta \)

- Variational EM:
  - init \((\alpha, \beta)\)
  - while \((\approx \text{log-lik increases})\)
    - variational inference
      \[(v^*, \phi^*, ...) = \arg\max_{(v, \phi)} F(\alpha, \beta; v, \phi, ...)\]
    - update \((\alpha, \beta)\) with variational-MLEs given \((v^*, \phi^*, ...)\)
• Outcasts are an isolated faction
• Young Turks *like* members of the Loyal Opposition, although the sentiment is not reciprocal.
In the mixed-membership simplex  [Airoldi, Blei, Fienberg and Xing, 2008]
In the mixed-membership simplex

― Visualizing social space
Travel time tomography

- How to model dynamics in a simplex?

Project an individual/stock in network into a "tomographic" space

Trajectory of an individual/stock in the "tomographic" space
Objective

Discover latent role space for a time-evolving network

From time-evolving network graphs...  ...to latent space trajectories

Time = 1

Latent role discovery

Role A

Role B

Role C

© Eric Xing @ CMU, 2013
Basic Model

The mixed-membership blockmodel for networks

Alice’s “hidden” roles

- Student
- Cyclist
- Democrat

Bob’s “hidden” roles

- Professor
- Guitarist
- Democrat

Role “compatibility” matrix

<table>
<thead>
<tr>
<th></th>
<th>Democrat</th>
<th>Professor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democrat</td>
<td>Compatible!</td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td></td>
<td>Incompatible!</td>
</tr>
</tbody>
</table>

Friends?

Friends!

Not friends...

© Eric Xing @ CMU, 2013
Time-varying Network Model

High level overview

Time evolution

Hidden role prior \( (t) \)

Actor hidden roles \( (t) \)

Observed interactions \( (t) \)

Role compatibility matrix \( (t) \)

Time evolution
Dynamic MMSB (dMMSB)
Dynamic MMSB (dMMSB)
Algorithm: Generalized Mean Field (xing et al. 2004)

Approximate the joint posterior
\[ p\left(\{\tilde{Z}^{(t)}, \tilde{\pi}^{(t)}, \tilde{\mu}^{(t)}, B^{(t)}\}_{t=1}^T | \Theta, \{G^{(t)}\}_{t=1}^T\right) \]
where \( \Theta \) denotes the model parameters, by a factored approximate distribution:

\[ q\left(\{\tilde{Z}^{(t)}, \tilde{\pi}^{(t)}, \tilde{\mu}^{(t)}, B^{(t)}\}_{t=1}^T\right) = q_1\left(\{\tilde{Z}^{(t)}, \tilde{\pi}^{(t)}\}_{t=1}^T\right) \times q_2\left(\{\tilde{\mu}^{(t)}\}_{t=1}^T\right) \times q_3\left(\{B^{(t)}\}_{t=1}^T\right), \]

- **Inference via variational EM**
  - Generalized mean field
  - Laplace approximation
  - Kalman filter & RTS smoother
dMMSB vs. MMSB

The graph compares the average error in L-2 norm for dMMSB and MMSB over time points. The error for dMMSB is represented by a blue line with markers, while MMSB is shown with a green dotted line with crosses. The error spikes at time points 2, 8, and 10 for dMMSB, while MMSB shows a different pattern with higher error at time points 2 and 10.
Case Study 1: Sampson’s Monk Network

- **Dataset Description**
  - 18 monks (junior members in a monastery)
  - Liking relations recorded
  - 3 time-points in one year period
  - Timing: before a major conflict outbreak

- **Recall static analysis:**

![Diagram of monk network]

© Eric Xing @ CMU, 2013
Sampson’s Monk Network: role trajectories

- The trajectories of the varying role-vectors over time
Sampson’s Monk Network: Dynamic Analysis

- Observations
  - Big changes in time 1 to time 2
  - From time 2 to time 3, role-vectors purifying
    - More isolated
    - Led to the separation
Case Study 2: The 109th congress

US senator voting records
100 senators, 109th Congress (Jan 2005 – Dec 2006) in 8 epochs
A closer look at the outliers

#28 (Democrat, New Jersey):
- Senator Jon Corzine from time points 1-4
  - Jon was extremely liberal – so much that his role vector is not “Democratic”
- His seat was taken over by Bob Menendez during time points 5-8
  - Bob was moderate-left, so his role vector is more “Democratic” than Jon’s

#75 (Democrat, Nebraska):
- Senator Ben Nelson
- Starts out centrist; neither Democrat nor Republican
- Ends up strongly Republican at the last time point (around Dec 2006)
  - About the same time as his re-election in Nebraska, in which he won a sizeable portion of the Republican vote.
109th US Congress senator vote data:
- 100 senators, 8 time points from Jan 2005 to Dec 2006
- Raw data provided in the form of yes/no votes for each senator
  - Applied the method of Kolar et al. (2008) to convert into a network
  - An edge between two senators indicates they have similar voting patterns

Results:
- BIC chooses 4 roles and 2 clusters
- Role 1 = Democrats, Role 2 = Republicans
- Roles 3 and 4 indicate unconventional behavior
- Most senators behave according to their party affiliation, but there are outliers:
  - #28 (Democrat, New Jersey)
  - #75 (Democrat, Nebraska)
Summary

- Stochastic Block Model for single role inference in networks

- Mixed-Membership Stochastic Block Model: A hierarchical Bayesian extension of stochastic block models, and applications to multi-role network semantic distillation

- Dynamic MMSB for dynamic network analysis and an example in the social network context