Latent Space Models for Social Networks

Mark S. Handcock

Department of Statistics
University of Washington

Joint work with

UW Networks Working Group
Peter D. Hoff, University of Washington
Adrian E. Raftery, University of Washington
Jeremy S. Tantrum, Microsoft AdCenter Labs

Supported by NIDA Grant DA012831, NICHD Grant HD041877, and
ONR award N00014-08-1-1015.

MURI Kickoff Meeting, November 18
Example of Social Relationships between Monks

- Expressed “liking” between 18 monks within an isolated monastery
  ⇒ Sampson (1969)
- A directed relationship aggregated over a 12 month period before the breakup of the cloister.
Features of Many Social Networks

- Yutuality of ties: individual propensity to form ties varies by actor attributes.
- Homophily by actor attributes: higher propensity to form ties between actors with similar attributes (e.g., age, gender, geography).
- Transitivity of relationships: friends of friends have a higher propensity to be friends.
Features of Many Social Networks

- Mutuality of ties
Features of Many Social Networks

- Mutuality of ties
- Individual propensity to form ties varies by actor attributes
- Homophily by actor attributes
  ⇒ Lazarsfeld and Merton, 1954; Freeman, 1996; McPherson et al., 2001
  - higher propensity to form ties between actors with similar attributes
    e.g., age, gender, geography
Features of Many Social Networks

- Mutuality of ties
- Individual propensity to form ties varies by actor attributes
- Homophily by actor attributes
  \[ \Rightarrow \text{ Lazarsfeld and Merton, 1954; Freeman, 1996; McPherson et al., 2001} \]
  - higher propensity to form ties between actors with similar attributes
    - e.g., age, gender, geography
- Transitivity of relationships
  - friends of friends have a higher propensity to be friends
Clustering and Social Networks

Three types of clustering in social networks:
1. Transitivity of relationships
2. Homophily of actors with similar observed characteristics
3. Further clustering that could be due to homophily on unobserved attributes or "self-organization" into groups

Drawing conclusions about clustering of social actors is often a focus of interest in social network analysis. Most methods don't address it directly, instead conclusions about clustering are often drawn by informally eyeballing results from other methods.

We present a statistical model of social networks that incorporates clustering and allows formal inference about whether or not there is clustering beyond transitivity, if so how many groups there are, who is in what group, and uncertainty about group memberships.
Three types of clustering in social networks:

- transitivity of relationships
- homophily of actors with similar observed characteristics
- further clustering that could be due to:
  - homophily on unobserved attributes, or
  - “self-organization” into groups

We present a statistical model of social networks that incorporates clustering and allows formal inference about whether or not there is clustering beyond transitivity, if so how many groups there are, who is in what group, and uncertainty about group memberships.
Clustering and Social Networks

- Three types of clustering in social networks:
  - transitivity of relationships
  - homophily of actors with similar *observed* characteristics
  - further clustering that could be due to:
    - homophily on unobserved attributes, or
    - “self-organization” into groups

- Drawing conclusions about clustering of social actors is often a focus of interest in social network analysis
Clustering and Social Networks

- Three types of clustering in social networks:
  - transitivity of relationships
  - homophily of actors with similar observed characteristics
  - further clustering that could be due to:
    - homophily on unobserved attributes, or
    - “self-organization” into groups

- Drawing conclusions about clustering of social actors is often a focus of interest in social network analysis

- But most methods don’t address it directly
Three types of clustering in social networks:
- transitivity of relationships
- homophily of actors with similar *observed* characteristics
- further clustering that could be due to:
  - homophily on unobserved attributes, or
  - “self-organization” into groups

Drawing conclusions about clustering of social actors is often a focus of interest in social network analysis

But most methods don’t address it directly

Instead conclusions about clustering are often drawn by informally eyeballing results from other methods
Clustering and Social Networks

- Three types of clustering in social networks:
  - transitivity of relationships
  - homophily of actors with similar observed characteristics
  - further clustering that could be due to:
    - homophily on unobserved attributes, or
    - “self-organization” into groups

- Drawing conclusions about clustering of social actors is often a focus of interest in social network analysis
- But most methods don’t address it directly
- Instead conclusions about clustering are often drawn by informally eyeballing results from other methods
- We present a statistical model of social networks that incorporates clustering and allows formal inference about:
  - whether or not there is clustering (beyond transitivity)
  - if so, how many groups there are
  - who is in what group
  - uncertainty about group memberships
Statistical Models for Social Networks

*Notation*

A *social network* is defined as a set of \( n \) social “actors” and a social relationship between each pair of actors.
Statistical Models for Social Networks

**Notation**
A *social network* is defined as a set of $n$ social “actors” and a social relationship between each pair of actors.

$$Y_{ij} = \begin{cases} 1 & \text{relationship from actor } i \text{ to actor } j \\ 0 & \text{otherwise} \end{cases}$$
Statistical Models for Social Networks

Notation
A social network is defined as a set of \( n \) social “actors” and a social relationship between each pair of actors.

\[
Y_{ij} = \begin{cases} 
1 & \text{relationship from actor } i \text{ to actor } j \\
0 & \text{otherwise}
\end{cases}
\]

- call \( Y \equiv [Y_{ij}]_{n \times n} \) a sociomatrix
- a \( N = n(n - 1) \) binary array
Statistical Models for Social Networks

**Notation**

A *social network* is defined as a set of \( n \) social “actors” and a social relationship between each pair of actors.

\[
Y_{ij} = \begin{cases} 
1 & \text{relationship from actor } i \text{ to actor } j \\
0 & \text{otherwise} 
\end{cases}
\]

- call \( Y \equiv [Y_{ij}]_{n \times n} \) a *sociomatrix*
- a \( N = n(n - 1) \) binary array

The basic problem of stochastic modeling is to specify a distribution for \( Y \) i.e., \( P(Y = y) \)
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    Hoff (2003, 2004,...)
model an underlying latent “social space” of actors

Latent space models: Hoff, Raftery and Handcock (2002)
Hoff (2003, 2004 ,...)

Hierarchical model for the network:

Actors $i$ and $j$ are an unknown distance apart in social space
Conditional on the distances the ties are independent
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    Hoff (2003, 2004 ,...)

- Hierarchical model for the network:
  - Actors $i$ and $j$ are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:
  - $\{\delta_i\}$ individual propensity of the actors to form ties
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    Hoff (2003, 2004,...)

- Hierarchical model for the network:
  - Actors \( i \) and \( j \) are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:
- \( \{\delta_i\} \) individual propensity of the actors to form ties
- \( \{\gamma_i\} \) individual propensity of the actors to receive ties
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    - Hoff (2003, 2004, ...)

- Hierarchical model for the network:
  - Actors $i$ and $j$ are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:
- $\{\delta_i\}$ individual propensity of the actors to form ties
- $\{\gamma_i\}$ individual propensity of the actors to receive ties
- $\{z_i\}$ be the positions of the actors in the social space $\mathbb{R}^k$
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    Hoff (2003, 2004, ...)

- Hierarchical model for the network:
  - Actors $i$ and $j$ are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:

- $\{\delta_i\}$ individual propensity of the actors to form ties
- $\{\gamma_i\}$ individual propensity of the actors to receive ties
- $\{z_i\}$ be the positions of the actors in the social space $\mathbb{R}^k$
- $\{x_{i,j}\}$ denote observed characteristics that may be dyad-specific and vector-valued
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    Hoff (2003, 2004,...)

- Hierarchical model for the network:
  - Actors $i$ and $j$ are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:

- $\{\delta_i\}$ individual propensity of the actors to form ties
- $\{\gamma_i\}$ individual propensity of the actors to receive ties
- $\{z_i\}$ be the positions of the actors in the social space $\mathbb{R}^k$
- $\{x_{i,j}\}$ denote observed characteristics that may be dyad-specific and vector-valued
Positing Latent Social Structure via Random Effects

- model an underlying latent “social space” of actors
  - Latent space models: Hoff, Raftery and Handcock (2002)
    - Hoff (2003, 2004,...)

- Hierarchical model for the network:
  - Actors $i$ and $j$ are an unknown distance apart in social space
  - Conditional on the distances the ties are independent

Let:
  - $\{\delta_i\}$ individual propensity of the actors to form ties
  - $\{\gamma_i\}$ individual propensity of the actors to receive ties
  - $\{z_i\}$ be the positions of the actors in the social space $\mathbb{R}^k$
  - $\{x_{i,j}\}$ denote observed characteristics that may be dyad-specific and vector-valued

Specifically:

$$\log \text{odds}(Y_{ij} = 1|z_i, z_j, x_{ij}, \beta) = \beta^T x_{ij} - |z_i - z_j| + \delta_i + \gamma_j$$

where $\beta$ denotes parameters to be estimated.
Model-based Clustering of Social Networks

...
Model-based Clustering of Social Networks

Model the latent positions as clustered into $G$ groups:

$$z_i \sim i.i.d. \sum_{g=1}^{G} \lambda_g \text{MVN}_d(\mu_g, \sigma^2_g I_d)$$
Model-based Clustering of Social Networks

- Model the latent positions as clustered into $G$ groups:
  \[ z_i \sim \sum_{g=1}^{G} \lambda_g \text{MVN}_d(\mu_g, \sigma_g^2 I_d) \]

- Spherical covariance motivated by invariance
Model-based Clustering of Social Networks

- Model the latent positions as clustered into $G$ groups:

$$z_i \overset{i.i.d.}{\sim} \sum_{g=1}^{G} \lambda_g \text{MVN}_d(\mu_g, \sigma_g^2 I_d)$$

- Spherical covariance motivated by invariance

- captures position, transitivity, homophily on attributes, and clustering
Model-based Clustering of Social Networks

- Model the latent positions as clustered into $G$ groups:

$$z_i^{i.i.d.} \sim \sum_{g=1}^{G} \lambda_g \text{MVN}_d(\mu_g, \sigma_g^2 I_d)$$

- Spherical covariance motivated by invariance
- captures position, transitivity, homophily on attributes, and clustering
- captures individual propensities to form and receive ties

$$\delta_i^{i.i.d.} \sim N(0, \sigma_\delta^2) \quad i = 1, \ldots, n,$$

$$\gamma_i^{i.i.d.} \sim N(0, \sigma_\gamma^2) \quad i = 1, \ldots, n,$$
Graphical Structure of the Model

Sampling Type
- Constant
- Gibbs
- Metropolis-Hastings
Structure of the algorithm

Bayesian inference implemented via Markov Chain Monte Carlo (MCMC)

Let $K_i$ be the cluster of actor $i$.

Some full conditional posterior distributions are available:

\[
\sigma^2_\delta | \delta, \ldots \sim \left( \alpha_\delta \sigma^2_{0,\delta} + \sum_{i=1}^{n} \delta^2_i \right) \text{Inv} \chi^2_{\alpha_\delta + n},
\]

\[
\sigma^2_\gamma | \gamma, \ldots \sim \left( \alpha_\gamma \sigma^2_{0,\gamma} + \sum_{i=1}^{n} \gamma^2_i \right) \text{Inv} \chi^2_{\alpha_\gamma + n},
\]

\[
\mu_g | Z, K, \sigma^2_g, \ldots \overset{\text{ind}}{\sim} \text{MVN}_d \left( \frac{n_g \bar{Z}_g}{n_g + \sigma^2_g / \omega^2}, \frac{\sigma^2_g}{n_g + \sigma^2_g / \omega^2} \right),
\]

\[
\sigma^2_g | Z, K, \mu_g, \ldots \overset{\text{ind}}{\sim} \left( \alpha_Z \sigma^2_{Z,0} + SS_{Zg} \right) \text{Inv} \chi^2_{\alpha_Z + n_g d},
\]

\[
\lambda | K, \ldots \sim \text{Dirichlet} \left( \nu_1 + n_1, \ldots, \nu_G + n_G \right),
\]

\[
\Pr \left( K_i = g | \lambda, Z, \mu_g, \sigma^2_g, \ldots \right) \overset{\text{ind}}{=} \frac{\lambda_g f_{\text{MVN}_d}(\mu_g, \sigma^2_g I_d)(Z_i)}{\sum_{k=1}^{G} \lambda_k f_{\text{MVN}_d}(\mu_k, \sigma^2_k I_d)(Z_i)} \quad i = 1, \ldots, n,
\]
where

\[ SS_{Zg} = \sum_{i=1}^{n} 1_{K_i=g} (Z_i - \mu_g)^T (Z_i - \mu_g), \]

\[ n_g = \sum_{i=1}^{n} 1_{K_i=g} \]

and \( \phi_d(\cdot; \mu, \Sigma) \) is the d-dimensional multivariate normal density.
Algorithmic details

Our algorithm is then as follows:

1. Use Metropolis–Hastings to sample $Z_{t+1}$, updating each actor in random order:
   1. Propose $Z_i^* \sim \text{MVN}_d(Z_{ti}, \delta_Z^2 I_d)$.
   2. With probability equal to

   $$\frac{P(Y|Z^*, X, \beta_t)\phi_d(Z_i^*; \mu_{K_i}, \sigma_{K_i}^2 I_d)}{P(Y|Z_t, X, \beta_t)\phi_d(Z_{it}; \mu_{K_i}, \sigma_{K_i}^2 I_d)},$$

   set the $i$th element of $Z_{t+1}$ to $Z_i^*$. Otherwise set it to $Z_{it}$.

2. Use Metropolis–Hastings to sample $\beta_{t+1}$:
   1. Propose $\beta^* \sim \text{MVN}_d(\beta_t, \delta_\beta^2 I_p)$.
   2. With probability equal to

   $$\frac{P(Y|Z_{t+1}, X, \beta^*)\phi_p(\beta^*; \xi, \Psi)}{P(Y|Z_{t+1}, X, \beta_t)\phi_p(\beta_t; \xi, \Psi)},$$

   set $\beta_{t+1} = \beta^*$. Otherwise set $\beta_{t+1} = \beta_t$.

3. Update, $K_i, \mu_g, \sigma_g^2$ and $\lambda_g$ from (3), (4), (5) and (6).
Estimation
Estimation

Two-Stage Maximum Likelihood Estimation
Two-Stage Maximum Likelihood Estimation

1. Find the MLE of the latent positions via (unclustered) latent space model

   ⇒  Hoff, Raftery, Handcock (2002)
Two-Stage Maximum Likelihood Estimation

1. Find the MLE of the latent positions via (unclustered) latent space model
   \[ \Rightarrow \] Hoff, Raftery, Handcock (2002)

2. Apply model-based clustering conditional on estimated latent positions
Estimation

Two-Stage Maximum Likelihood Estimation

1. Find the MLE of the latent positions via (unclustered) latent space model
   \[ \Rightarrow \text{Hoff, Raftery, Handcock (2002)} \]

2. Apply model-based clustering conditional on estimated latent positions
   - Use EM as in the R package mclust (Fraley and Raftery 1998)
Two-Stage Maximum Likelihood Estimation

1. Find the MLE of the latent positions via (unclustered) latent space model
   - Hoff, Raftery, Handcock (2002)

2. Apply model-based clustering conditional on estimated latent positions
   - Use EM as in the R package mclust (Fraley and Raftery 1998)

- fast and simple
- cluster structure not used to estimate positions
Two-Stage Maximum Likelihood Estimation

1. Find the MLE of the latent positions via (unclustered) latent space model
   ⇒ Hoff, Raftery, Handcock (2002)

2. Apply model-based clustering conditional on estimated latent positions
   - Use EM as in the R package mclust (Fraley and Raftery 1998)

- fast and simple
- cluster structure not used to estimate positions
Bayesian Estimation

We use the following prior distributions:

\[
\beta \sim \text{MVN}(\mu, \Sigma), \\
\lambda \sim \text{dirichlet}(\nu), \\
\sigma_t \sim \text{univ}(\alpha, \beta), \\
\sigma_g \sim \text{univ}(\alpha, \gamma). \\
\]

where \(\mu, \Sigma, \nu, \alpha, \beta, \gamma\) are hyperparameters.
Bayesian Estimation

We use the following prior distributions:

\[
\begin{align*}
\beta & \sim \text{MVN}_p(\xi, \Psi), \\
\lambda & \sim \text{Dirichlet}(\nu), \\
\sigma^2_\delta & \sim \alpha_\delta \sigma^2_{0,\delta} \text{inv}\chi^2_{\alpha_\delta}, \\
\sigma^2_\gamma & \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{inv}\chi^2_{\alpha_\gamma}, \\
\sigma^2_g & \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{inv}\chi^2_{\alpha_Z}, \quad g = 1, \ldots, G, \\
\mu_g & \overset{\text{i.i.d.}}{\sim} \text{MVN}_d \left(0, \omega^2 I_d\right), \quad g = 1 \ldots G,
\end{align*}
\]

where \(\xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma,\) and \(\omega^2\) are hyperparameters.
Bayesian Estimation

We use the following prior distributions:

\[
\begin{align*}
\beta & \sim \text{MVN}_p(\xi, \Psi), \\
\lambda & \sim \text{Dirichlet}(\nu), \\
\sigma^2_\delta & \sim \alpha_\delta \sigma^2_{0,\delta} \text{Inv}\chi^2_{\alpha_\delta}, \\
\sigma^2_\gamma & \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{Inv}\chi^2_{\alpha_\gamma}, \\
\sigma^2_g & \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{Inv}\chi^2_{\alpha_Z} \quad g = 1, \ldots, G, \\
\mu_g & \overset{\text{i.i.d.}}{\sim} \text{MVN}_d(0, \omega^2 I_d) \quad g = 1 \ldots G,
\end{align*}
\]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma \), and \( \omega^2 \) are hyperparameters.

- We set:
Bayesian Estimation

We use the following prior distributions:

\[
\begin{align*}
\beta & \sim \text{MVN}_p(\xi, \Psi), \\
\lambda & \sim \text{Dirichlet}(\nu), \\
\sigma^2_\delta & \sim \alpha_\delta \sigma^2_{0,\delta} \text{ inv } \chi^2_{\alpha_\delta}, \\
\sigma^2_\gamma & \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{ inv } \chi^2_{\alpha_\gamma}, \\
\sigma^2_g & \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{ inv } \chi^2_{\alpha_Z} \quad g = 1, \ldots, G, \\
\mu_g & \overset{\text{i.i.d.}}{\sim} \text{MVN}_d \left(0, \omega^2 I_d\right) \quad g = 1 \ldots G,
\end{align*}
\]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_z, \alpha_z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma, \) and \( \omega^2 \) are hyperparameters.

- We set:
  - \( \nu_g = 3 \) \hspace{1cm} (low probability of small groups)
Bayesian Estimation

We use the following prior distributions:

\[ \beta \sim \text{MVN}_p(\xi, \Psi), \]
\[ \lambda \sim \text{Dirichlet}(\nu), \]
\[ \sigma^2_\delta \sim \alpha_\delta \sigma^2_{0,\delta} \text{ln}\chi^2_{\alpha_\delta}, \]
\[ \sigma^2_\gamma \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{ln}\chi^2_{\alpha_\gamma}, \]
\[ \sigma^2_g \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{ln}\chi^2_{\alpha_Z} \quad g = 1, \ldots, G, \]
\[ \mu_g \overset{\text{i.i.d.}}{\sim} \text{MVN}_d \left(0, \omega^2 I_d\right), \quad g = 1 \ldots G, \]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma, \) and \( \omega^2 \) are hyperparameters.

- We set:
  - \( \nu_g = 3 \) (low probability of small groups)
  - \( \xi = 0 \) and \( \Psi = 2I \) (allows a wide range of values of \( \beta \))
Bayesian Estimation

We use the following prior distributions:

\[
\begin{align*}
\beta & \sim \text{MVN}_p(\xi, \Psi), \\
\lambda & \sim \text{Dirichlet}(\nu), \\
\sigma^2_\delta & \sim \alpha_\delta \sigma^2_{0,\delta} \text{InvChi}^2_{\alpha_\delta}, \\
\sigma^2_\gamma & \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{InvChi}^2_{\alpha_\gamma}, \\
\sigma^2_g & \sim_{\text{i.i.d.}} \alpha_Z \sigma^2_{0,Z} \text{InvChi}^2_{\alpha_Z} \quad g = 1, \ldots, G, \\
\mu_g & \sim_{\text{i.i.d.}} \text{MVN}_d(0, \omega^2 I_d), \quad g = 1 \ldots G,
\end{align*}
\]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma, \) and \( \omega^2 \) are hyperparameters.

- We set:
  - \( \nu_g = 3 \) (low probability of small groups)
  - \( \xi = 0 \) and \( \Psi = 2I \) (allows a wide range of values of \( \beta \))
  - \( \alpha = 2 \) and \( \sigma^2_0 = 0.103 \) (within-group variation can be small)
Bayesian Estimation

We use the following prior distributions:

\[
\begin{align*}
\beta & \sim \text{MVN}_p(\xi, \Psi), \\
\lambda & \sim \text{Dirichlet}(\nu), \\
\sigma^2_{\delta} & \sim \alpha_{\delta} \sigma^2_{0,\delta} \text{Inv}\chi^2_{\alpha_{\delta}}, \\
\sigma^2_{\gamma} & \sim \alpha_{\gamma} \sigma^2_{0,\gamma} \text{Inv}\chi^2_{\alpha_{\gamma}}, \\
\sigma^2_g & \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{Inv}\chi^2_{\alpha_Z}, \quad g = 1, \ldots, G, \\
\mu_g & \overset{\text{i.i.d.}}{\sim} \text{MVN}_d(0, \omega^2 I_d), \quad g = 1 \ldots G,
\end{align*}
\]

where $\xi$, $\Psi$, $\nu = (\nu_1, \ldots, \nu_G)$, $\sigma^2_{0,Z}$, $\alpha_Z$, $\sigma^2_{0,\delta}$, $\alpha_{\delta}$, $\sigma^2_{0,\gamma}$, $\alpha_{\gamma}$, and $\omega^2$ are hyperparameters.

- We set:
  - $\nu_g = 3$ (low probability of small groups)
  - $\xi = 0$ and $\Psi = 2I$ (allows a wide range of values of $\beta$)
  - $\alpha = 2$ and $\sigma^2_0 = 0.103$ (within-group variation can be small)
  - $\omega^2 = 2$ (prior density of the means is relatively flat)
Bayesian Estimation

We use the following prior distributions:

\[ \beta \sim \text{MVN}_p(\xi, \Psi), \]
\[ \lambda \sim \text{Dirichlet}(\nu), \]
\[ \sigma^2_\delta \sim \alpha_\delta \sigma^2_{0,\delta} \text{Inv}\chi^2_{\alpha_\delta}, \]
\[ \sigma^2_\gamma \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{Inv}\chi^2_{\alpha_\gamma}, \]
\[ \sigma^2_g \overset{\text{i.i.d.}}{\sim} \alpha_Z \sigma^2_{0,Z} \text{Inv}\chi^2_{\alpha_Z} \quad g = 1, \ldots, G, \]
\[ \mu_g \overset{\text{i.i.d.}}{\sim} \text{MVN}_d(0, \omega^2 I_d), \quad g = 1 \ldots G, \]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma, \) and \( \omega^2 \) are hyperparameters.

- We set:
  - \( \nu_g = 3 \) (low probability of small groups)
  - \( \xi = 0 \) and \( \Psi = 2I \) (allows a wide range of values of \( \beta \))
  - \( \alpha = 2 \) and \( \sigma^2_0 = 0.103 \) (within-group variation can be small)
  - \( \omega^2 = 2 \) (prior density of the means is relatively flat)
Bayesian Estimation

We use the following prior distributions:

\[ \beta \sim \text{MVN}_p(\xi, \Psi), \]
\[ \lambda \sim \text{Dirichlet}(\nu), \]
\[ \sigma^2_\delta \sim \alpha_\delta \sigma^2_{0,\delta} \text{Inv}\chi^2_{\alpha_\delta}, \]
\[ \sigma^2_\gamma \sim \alpha_\gamma \sigma^2_{0,\gamma} \text{Inv}\chi^2_{\alpha_\gamma}, \]
\[ \sigma^2_g \text{i.i.d.} \sim \alpha_Z \sigma^2_{0,Z} \text{Inv}\chi^2_{\alpha_Z}, \quad g = 1, \ldots, G, \]
\[ \mu_g \text{i.i.d.} \sim \text{MVN}_d(0, \omega^2 I_d), \quad g = 1 \ldots G, \]

where \( \xi, \Psi, \nu = (\nu_1, \ldots, \nu_G), \sigma^2_{0,Z}, \alpha_Z, \sigma^2_{0,\delta}, \alpha_\delta, \sigma^2_{0,\gamma}, \alpha_\gamma, \) and \( \omega^2 \) are hyperparameters.

- We set:
  - \( \nu_g = 3 \) (low probability of small groups)
  - \( \xi = 0 \) and \( \Psi = 2I \) (allows a wide range of values of \( \beta \))
  - \( \alpha = 2 \) and \( \sigma^2_0 = 0.103 \) (within-group variation can be small)
  - \( \omega^2 = 2 \) (prior density of the means is relatively flat)

- Posterior distribution approximated by Markov chain Monte Carlo
Identifiability of Positions and Cluster Labels

The likelihood is a function of the latent positions only through their distances. The likelihood is also invariant to relabelling of the clusters. Resolve nonidentifiabilities by postprocessing the output. First, perform a Procrustes transform to align the actor positions and the cluster means and covariances. Idea: choose the configuration that minimizes the Wasserstein divergence from the true distributions. Find the minimum Wasserstein positions of the actors relative to $P_M$, $Y | Z, X$, $\beta$. Find the minimum Wasserstein cluster membership probabilities over all label permutations.
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances.
- The likelihood is also invariant to relabelling of the clusters.
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances.
- The likelihood is also invariant to relabelling of the clusters.

Resolve nonidentifiabilities by postprocessing the MCMC output.

Idea: Choose the configuration that minimizes the Hellinger divergence from the "true" distributions. Find the minimum Hellinger positions of the actors relative to \( P, M, Y | Z, X, \beta, N \). Find the minimum Hellinger cluster membership probabilities over all label permutations.
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances.
- The likelihood is also invariant to relabelling of the clusters.

Resolve nonidentifiabilities by postprocessing the MCMC output.

- First, Procrustes transform the actor positions and the cluster means and covariances.
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances.
- The likelihood is also invariant to relabelling of the clusters.

Resolve nonidentifiabilities by postprocessing the MCMC output.

- First, Procrustes transform the actor positions and the cluster means and covariances.
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances
- The likelihood is also invariant to relabelling of the clusters

Resolve nonidentifiabilities by postprocessing the MCMC output.

- First, Procrustes transform the actor positions and the cluster means and covariances.

**Idea**

Choose the configuration that minimizes the Kullback-Leibler divergence from the “true” distributions.
The likelihood is a function of the latent positions only through their distances.
The likelihood is also invariant to relabelling of the clusters.

Resolve nonidentifiabilities by postprocessing the MCMC output:

First, Procrustes transform the actor positions and the cluster means and covariances.

Idea

Choose the configuration that minimizes the Kullback-Leibler divergence from the “true” distributions.

Find the minimum Kullback-Leibler positions of the actors relative to $P(Y|Z, X, \beta)$. 
Identifiability of Positions and Cluster Labels

- The likelihood is a function of the latent positions only through their distances.
- The likelihood is also invariant to relabelling of the clusters.

Resolve nonidentifiabilities by postprocessing the MCMC output.
- First, Procrustes transform the actor positions and the cluster means and covariances.

Idea

Choose the configuration that minimizes the Kullback-Leibler divergence from the “true” distributions.

1. Find the minimum Kullback-Leibler positions of the actors relative to $P(Y|Z, X, \beta)$.
2. Find the minimum Kullback-Leibler cluster membership probabilities over all label permutations.
Choosing the Number of Groups

We recast the choice of number of groups as a model selection problem, where each number of groups is viewed as a different statistical model. A Bayesian model selection approach is approximated by a version of the $p$-hierarchical model, which determines the number of groups. If the preferred number of groups is $p$, there is no evidence for clustering.

We use conditional posterior model probabilities, conditioning on an estimate of the latent positions $\hat{Z}$, and integrating over the other parameters. This method is motivated by the specific latent positions that will be used. Worked well in a similar model-based clustering with dissimilarities $\hat{O}$. After $p$, the method simplifies calculations.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups
- If the preferred number of groups is 1, there is no evidence for clustering
- We use *conditional* posterior model probabilities:
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use conditional posterior model probabilities:
  - conditioning on an estimate of the latent positions, \( \hat{Z} \)
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use *conditional* posterior model probabilities:
  - conditioning on an estimate of the latent positions, \( \hat{Z} \)
  - integrating over the other parameters.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use *conditional* posterior model probabilities:
  - conditioning on an estimate of the latent positions, $\hat{Z}$
  - integrating over the other parameters
- **Justification:**
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use conditional posterior model probabilities:
  - conditioning on an estimate of the latent positions, \( \hat{Z} \)
  - integrating over the other parameters
- Justification:
  - Evaluates the specific latent positions that will be used.
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use conditional posterior model probabilities:
  - conditioning on an estimate of the latent positions, $\hat{Z}$
  - integrating over the other parameters
- Justification:
  - Evaluates the specific latent positions that will be used
  - Worked well in a similar model: model-based clustering with dissimilarities (Oh & Raftery 2003)
Choosing the Number of Groups

- We recast the choice of number of groups as a model selection problem: Each number of groups is viewed as a different statistical model.
- Bayesian model selection (approximated by a version of BIC) determines the number of groups.
- If the preferred number of groups is 1, there is no evidence for clustering.
- We use *conditional* posterior model probabilities:
  - conditioning on an estimate of the latent positions, \( \hat{Z} \)
  - integrating over the other parameters.
- Justification:
  - Evaluates the specific latent positions that will be used.
  - Worked well in a similar model: model-based clustering with dissimilarities (Oh & Raftery 2003).
  - Simplifies calculations.
Approximate Posterior Model Probabilities
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[ P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*) \]
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[ P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \]  

(*)
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[ P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*) \]

\[ = \text{logistic regression factor} \times \text{model-based clustering factor}, \]

where \( \theta = (\mu_g, \lambda_g, \sigma^2_g)^G_{g=1} \)
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[ P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*) \]

= logistic regression factor × model-based clustering factor,

where \( \theta = (\mu_g, \lambda_g, \sigma^2_g)^G \)

We approximate both factors in (*) using the BIC approximation (in generic form):

\[ P(W) = \int P(W|\phi)p(\phi)d\phi \approx P(W|\hat{\phi})m^{-p/2}, \]

where \( m = \text{dim}(W) \) and \( p = \text{dim}(\phi) \).
Approximate Posterior Model Probabilities

- Then the integrated likelihood is:

\[
P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*)
\]

= logistic regression factor \times model-based clustering factor,

where \( \theta = (\mu_g, \lambda_g, \sigma^2_g)_{g=1}^G \)

- We approximate both factors in (*) using the BIC approximation (in generic form):

\[
P(W) = \int P(W|\phi)p(\phi)d\phi \approx P(W|\hat{\phi})m^{-p/2},
\]

where \( m = \text{dim}(W) \) and \( p = \text{dim}(\phi) \).

- We thus approximate (twice the log) integrated likelihood by
Approximate Posterior Model Probabilities

- Then the integrated likelihood is:

\[
P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \tag{\star}
\]

\[
= \text{logistic regression factor} \times \text{model-based clustering factor},
\]

where \(\theta = (\mu_g, \lambda_g, \sigma_g^2)^G_{g=1}\)

- We approximate both factors in (\star) using the BIC approximation (in generic form):

\[
P(W) = \int P(W|\phi)p(\phi)d\phi \approx P(W|\hat{\phi})m^{-p/2},
\]

where \(m = \text{dim}(W)\) and \(p = \text{dim}(\phi)\).

- We thus approximate (twice the log) integrated likelihood by
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[ P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*) \]

\[ = \text{logistic regression factor } \times \text{model-based clustering factor}, \]

where \( \theta = (\mu_g, \lambda_g, \sigma^2_g)_{g=1}^G \)

We approximate both factors in (*) using the BIC approximation (in generic form):

\[ P(W) = \int P(W|\phi)p(\phi)d\phi \approx P(W|\hat{\phi})m^{-p/2}, \]

where \( m = \text{dim}(W) \) and \( p = \text{dim}(\phi). \)

We thus approximate (twice the log) integrated likelihood by

\[ \text{BIC} = \text{BIC}_{lr} + \text{BIC}_{mbc} \]
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

$$P(Y, \hat{Z}|G) = \int P(Y|\hat{Z}, X, \beta)p(\beta)d\beta \times \int P(\hat{Z}|\theta)p(\theta)d\theta \quad (*)$$

$$= \text{logistic regression factor} \times \text{model-based clustering factor},$$

where $\theta = (\mu_g, \lambda_g, \sigma_g^2)_{g=1}^G$

We approximate both factors in (*) using the BIC approximation (in generic form):

$$P(W) = \int P(W|\phi)p(\phi)d\phi \approx P(W|\hat{\phi})m^{-p/2},$$

where $m = \text{dim}(W)$ and $p = \text{dim}(\phi)$.

We thus approximate (twice the log) integrated likelihood by

$$\text{BIC} = \text{BIC}_{lr} + \text{BIC}_{mbc}$$

where

$$\text{BIC}_{lr} = 2\log P(Y|\hat{Z}, X, \hat{\beta}(\hat{Z})) - d_{\text{logit}} \log n_{\text{logit}},$$
Approximate Posterior Model Probabilities

Then the integrated likelihood is:

\[
P(Y, \hat{Z} | G) = \int P(Y | \hat{Z}, X, \beta) p(\beta) d\beta \times \int P(\hat{Z} | \theta) p(\theta) d\theta \quad (*)
\]

\[= \text{logistic regression factor } \times \text{model-based clustering factor},\]

where \( \theta = (\mu_g, \lambda_g, \sigma_g^2)_{g=1}^G \)

We approximate both factors in (*) using the BIC approximation (in generic form):

\[
P(W) = \int P(W | \phi) P(\phi) d\phi \approx P(W | \hat{\phi}) m^{-p/2},
\]

where \( m = \text{dim}(W) \) and \( p = \text{dim}(\phi) \).

We thus approximate (twice the log) integrated likelihood by

\[
\text{BIC} = \text{BIC}_{lr} + \text{BIC}_{mbc}
\]

where

\[
\text{BIC}_{lr} = 2 \log P \left( Y | \hat{Z}, X, \hat{\beta}(\hat{Z}) \right) - d_{\text{logit}} \log n_{\text{logit}},
\]

and

\[
\text{BIC}_{mbc} = 2 \log P \left( \hat{Z} | \hat{\theta}(\hat{Z}) \right) - d_{\text{mbc}} \log n
\]
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

Strong evidence for clustering because the one-group model has little support from the data. This is the number of groups “known” to be in the data.
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

Strong evidence for clustering because the one-group model has little support from the data. This is the number of groups "known" to be in the data.
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

- Strong evidence for clustering
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

- Strong evidence for clustering
  - because the one-group model has little support from the data
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

- Strong evidence for clustering
  - because the one-group model has little support from the data
- 3 groups are strongly supported:
Sampson’s Monks: Inference About Presence of Clustering and Number of Groups

- Strong evidence for clustering
  - because the one-group model has little support from the data
- 3 groups are strongly supported:
  - This is the number of groups “known” to be in the data
Sampson’s Monks: Estimated Positions
Sampson’s Monks: Estimated Positions

The “known” groups Mas defined by White et al W are shown by letters M based on much more information than we use here Neut. The groups identified by our method are the same as the “known” groups. The probability of assignment of each monk to each latent cluster is shown by a colored pie chart.
Sampson’s Monks: Estimated Positions

- Bayesian estimates of positions in latent social space shown
Bayesian estimates of positions in latent social space shown

The “known” groups (as defined by White et al 1976) are shown by letters (based on much more information than we use here): Young Turks (T), Loyal opposition (L), Outcasts (O)
Bayesian estimates of positions in latent social space shown

The “known” groups (as defined by White et al 1976) are shown by letters (based on much more information than we use here): Young Turks (T), Loyal opposition (L), Outcasts (O)

The groups identified by our method are the same as the “known” groups
Bayesian estimates of positions in latent social space shown
The “known” groups (as defined by White et al 1976) are shown by letters (based on much more information than we use here): Young Turks (T), Loyal opposition (L), Outcasts (O)
The groups identified by our method are the same as the “known” groups
The probability of assignment of each monk to each latent cluster is shown by a colored pie chart
How important at the random propensities?

(a) without random effects  
(b) with receiver effects  

**Figure:** For panel (b), the area of the pie chart is proportional to the conditional odds ratio of a nomination for the monk due to his receiver effect and the pie chart represents the proportions of the MCMC draws assigning each monk to each cluster. The radii of the unfilled circles are equal to the cluster standard deviations, $\sigma_g$, conditional on the MKL point estimates.
Add Health Data
Add Health Data

- Friendship network (Bearman et al 1997)
Add Health Data

- Friendship network (Bearman et al 1997)
- 69 adolescents in grades 7–12 from one school
Add Health Data

- Friendship network (Bearman et al 1997)
- 69 adolescents in grades 7–12 from one school
- Each nominated up to 5 boys and 5 girls as their friends
Add Health Data

- Friendship network (Bearman et al 1997)
- 69 adolescents in grades 7–12 from one school
- Each nominated up to 5 boys and 5 girls as their friends
- Grade not taken into account in clustering
Add Health: Inference About Number of Groups
Add Health: Inference About Number of Groups

There is little evidence for one group model, The same number of groups as of grades
Clear evidence for clustering, because little evidence for one-group model
Clear evidence for clustering, because little evidence for one-group model

6 groups chosen by BIC
Clear evidence for clustering, because little evidence for one-group model

6 groups chosen by BIC

The same number of groups as of grades
Add Health: Estimated Clusters
Add Health: Estimated Clusters

- Estimated clusters shown by color
- Grades shown by number
- Good Mbut not total agreement between grades and groups
Add Health: Estimated Clusters

- Latent clusters shown by color
Latent clusters shown by color

Grades shown by number
Add Health: Estimated Clusters

- Latent clusters shown by color
- Grades shown by number
- Good (but not total) agreement between grades and groups
Add Health: Uncertainty about Cluster Memberships
Add Health: Uncertainty about Cluster Memberships
Add Health: Uncertainty about Errant 11th grader
Add Health: Unconstrained (with 2 isolates)

ah.lc ~ latent(d = 2, G = 6)
Summary

Yodel-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering but permits inference about whether there is clustering, how many groups there are, who is in what group, uncertainty about group memberships, and the actors' latent social positions. It gave reasonable results for two examples.

Software: The R package `latentnet` R available on CRAN.

Some future work: Account for degree distribution by including actor random effects, extend to nonebinary relations counting continuous duration, model cluster generating process, longitudinal and dynamic models, and generalize to hypergraphs.
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors' latent social positions

Some future work:
- Account for degree distribution by including actor random effects
- Extend to non-binary relations (count, continuous, duration)
- Model cluster generating process
- Longitudinal and dynamic models
- Generalize to hypergraphs

Software: The R package latentnet is available on CRAN
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering

Some future work:
- account for degree distribution by including actor random effects
- Extend to nonebinary relations, continuous, and duration
- Model cluster generating process
- Longitudinal and dynamic models
- Generalize to hypergraphs
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group

Some future work:
- Account for degree distribution by including actor random effects
- Extend to non-binary relations and continuous/duration data
- Model cluster generating process
- Longitudinal and dynamic models
- Generalize to hypergraphs
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships}

Some future work:
- Account for degree distribution by including actor random effects.
- Extend to non-binary relations (e.g., continuous, duration-based).
- Model cluster generating process.
- Longitudinal and dynamic models.
- Generalize to hypergraphs.
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.

- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions

- It gave reasonable results for two examples.
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
  - Account for degree distribution by including actor random effects.
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
  - Account for degree distribution by including actor random effects
  - Extend to non-binary relations (count, continuous, duration)
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
  - Account for degree distribution by including actor random effects
  - Extend to non-binary relations (count, continuous, duration)
  - Model cluster generating process
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
  - Account for degree distribution by including actor random effects
  - Extend to non-binary relations (count, continuous, duration)
  - Model cluster generating process
  - Longitudinal and dynamic models
Summary

- Model-based clustering of latent positions for social networks provides a formal model of social networks that incorporates clustering.
- It permits inference about:
  - whether there is clustering
  - how many groups there are
  - who is in what group
  - uncertainty about group memberships
  - the actors’ latent social positions
- It gave reasonable results for two examples.
- Software: The R package `latentnet`, available on CRAN.
- Some future work:
  - Account for degree distribution by including actor random effects
  - Extend to non-binary relations (count, continuous, duration)
  - Model cluster generating process
  - Longitudinal and dynamic models
  - Generalize to hypergraphs