Analysis on employees skill awareness network in a Manufacturing Company

Noemi Benci
Federico Pirona

May 28, 2020

University of Florence

Table of contents

Introduction

Skill awareness (R&D team); R. Cross and A. Parker, ”The Hidden Power of Social Networks.” Harvard Business School Press (2004).

Introduction

Topic: ”awareness of each others” knowledge and skills

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

Introduction

Topic: ”awareness of each others” knowledge and skills
”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

- Directed graph

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

- Directed graph
- 77 employees

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

- Directed graph
- 77 employees
- Originally: 2326 edges; valued dyadic variable

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

- Directed graph
- 77 employees
- Originally: 2326 edges; valued dyadic variable
- New interpretation: 1842 edges; binary dyadic variable

Introduction

Topic: ”awareness of each others” knowledge and skills

”I understand this person’s knowledge and skills. This does not necessarily mean that I have these skills or am knowledgeable in these domains but that I understand what skills this person has and domains they are knowledgeable in”.

General informations:

- Directed graph
- 77 employees
- Originally: 2326 edges; valued dyadic variable
- New interpretation: 1842 edges; binary dyadic variable
- Three nodal categorical attributes, everyone with four levels

Introduction

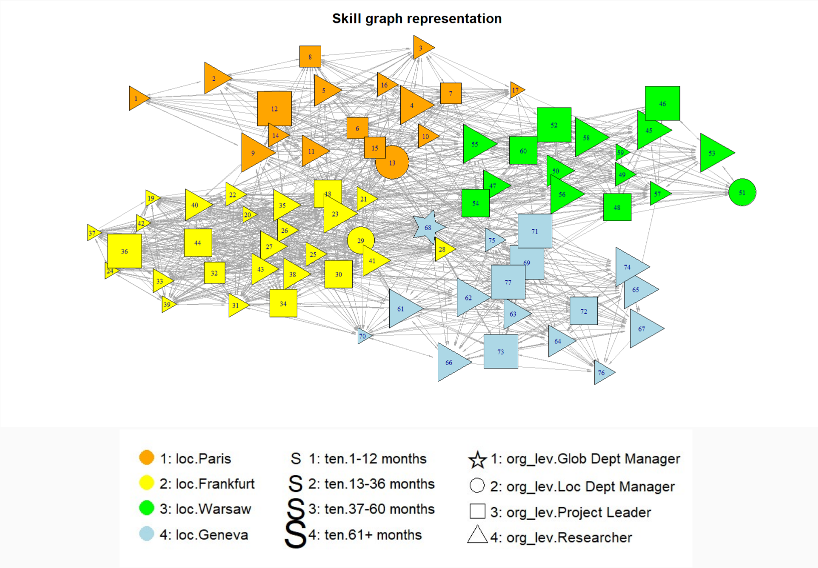


Figure 1: Graph representation of the data.

Network Statistics

Values to summarize information from the composition of the network

Network Statistics I

Density: In a probabilistic way, it can be seen as an estimate of the probability which two nodes are tied.

$$\rho = \frac{\sum_i \sum_j Y_{ij}}{n(n-1)} = \frac{m}{n(n-1)} = 0.315$$

Network Statistics I

Density: In a probabilistic way, it can be seen as an estimate of the probability which two nodes are tied.

$$\rho = \frac{\sum_i \sum_j Y_{ij}}{n(n-1)} = \frac{m}{n(n-1)} = 0.315$$

Reciprocity: it is the proportion of reciprocated ties:

$$R = \frac{\sum_{ij} Y_{ij} Y_{ji}}{m} = 0.819$$

Network Statistics I

Density: In a probabilistic way, it can be seen as an estimate of the probability which two nodes are tied.

$$\rho = \frac{\sum_i \sum_j Y_{ij}}{n(n-1)} = \frac{m}{n(n-1)} = 0.315$$

Reciprocity: it is the proportion of reciprocated ties:

$$R = \frac{\sum_{ij} Y_{ij} Y_{ji}}{m} = 0.819$$

Transitivity: In a probabilistic way, it can be seen as the probability which two nodes are tied given that they are tied to a third one.

$$C = \frac{\#\text{triangles}}{\#\text{triangles} + \#\text{two stars}} = 0.648$$

Network Statistics II

Modularity: it is defined as the difference between the fraction of ties observed between nodes in the same category and the fraction of expected ties (based on the nodes degree) between nodes in the same category. A normalized measure:

$$r = \frac{\sum_{ij}(Y_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)}{2m - \sum_{ij} \frac{k_i k_j}{2m}\delta(c_i, c_j)}; \qquad \text{where:}$$

$$k_i = \deg(i), \qquad \delta(c_i, c_j) = \begin{cases} 1 & \text{if } c_i = c_j \\ 0 & \text{otherwise} \end{cases}$$

Network Statistics II

Modularity: it is defined as the difference between the fraction of ties observed between nodes in the same category and the fraction of expected ties (based on the nodes degree) between nodes in the same category. A normalized measure:

$$r = \frac{\sum_{ij}(Y_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)}{2m - \sum_{ij} \frac{k_i k_j}{2m}\delta(c_i, c_j)}; \qquad \text{where:}$$

$$k_i = \deg(i), \qquad \delta(c_i, c_j) = \begin{cases} 1 & \text{if } c_i = c_j \\ 0 & \text{otherwise} \end{cases}$$

$$r_{loc} = 0.527$$

Network Statistics II

Modularity: it is defined as the difference between the fraction of ties observed between nodes in the same category and the fraction of expected ties (based on the nodes degree) between nodes in the same category. A normalized measure:

$$r = \frac{\sum_{ij}(Y_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)}{2m - \sum_{ij} \frac{k_i k_j}{2m}\delta(c_i, c_j)}; \qquad \text{where:}$$

$$k_i = \deg(i), \qquad \delta(c_i, c_j) = \begin{cases} 1 & \text{if } c_i = c_j \\ 0 & \text{otherwise} \end{cases}$$

$$r_{loc} = 0.527 \qquad r_{org\_lev} = -0.032$$

Network Statistics II

Modularity: it is defined as the difference between the fraction of ties observed between nodes in the same category and the fraction of expected ties (based on the nodes degree) between nodes in the same category. A normalized measure:

$$r = \frac{\sum_{ij}(Y_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)}{2m - \sum_{ij} \frac{k_i k_j}{2m}\delta(c_i, c_j)}; \qquad \text{where:}$$

$$k_i = \deg(i), \qquad \delta(c_i, c_j) = \begin{cases} 1 & \text{if } c_i = c_j \\ 0 & \text{otherwise} \end{cases}$$

$$r_{loc} = 0.527 \qquad r_{org\_lev} = -0.032 \qquad r_{tenure} = 0.125$$

Nodal Statistics

Values to summarize nodal information

Nodal Statistics - Measures of Centrality I

In- and Out-Degree Centrality: The measure for node i is the simple count of observed ties incoming in it or outgoing from it.

$$\tilde{\zeta}_i^{in-d} = \frac{\zeta_i^{in-d}}{n-1} \qquad \tilde{\zeta}_i^{out-d} = \frac{\zeta_i^{out-d}}{n-1}$$

Nodal Statistics - Measures of Centrality I

In- and Out-Degree Centrality: The measure for node i is the simple count of observed ties incoming in it or outgoing from it.

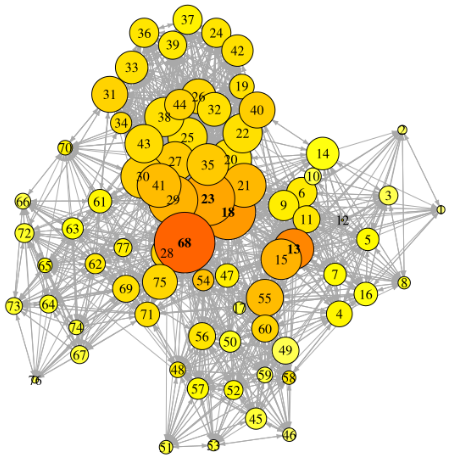
$$\tilde{\zeta}_i^{in-d} = \frac{\zeta_i^{in-d}}{n-1} \qquad \tilde{\zeta}_i^{out-d} = \frac{\zeta_i^{out-d}}{n-1}$$

	In-Degree	Out-Degree
Min.	0.1316	0.0263
1st Qu.	0.2237	0.2105
Median	0.3158	0.3026
Mean	0.3148	0.3148
3rd Qu.	0.3684	0.4079
Max.	0.7105	0.7105

Nodal Statistics - Measures of Centrality I

In- and Out-Degree Centrality: The measure for node i is the simple count of observed ties incoming in it or outgoing from it.

$$\tilde{\zeta}_i^{in-d} = \frac{\zeta_i^{in-d}}{n-1} \qquad \tilde{\zeta}_i^{out-d} = \frac{\zeta_i^{out-d}}{n-1}$$



	In-Degree	Out-Degree
Min.	0.1316	0.0263
1st Qu.	0.2237	0.2105
Median	0.3158	0.3026
Mean	0.3148	0.3148
3rd Qu.	0.3684	0.4079
Max.	0.7105	0.7105

Nodal Statistics - Measures of Centrality II

In- and Out-Closeness Centrality: define central a node if it is close, in terms of geodesic distance, to many others.

$$\tilde{\zeta}_i^{in-c} = (\zeta_i^{in-c})(n-1) \qquad \tilde{\zeta}_i^{out-c} = (\zeta_i^{out-c})(n-1)$$

Nodal Statistics - Measures of Centrality II

In- and Out-Closeness Centrality: define central a node if it is close, in terms of geodesic distance, to many others.

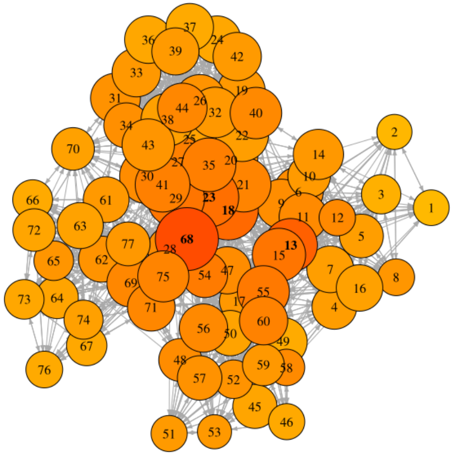
$$\tilde{\zeta}_i^{in-c} = (\zeta_i^{in-c})(n-1) \qquad \tilde{\zeta}_i^{out-c} = (\zeta_i^{out-c})(n-1)$$

	In-Closeness	Out-Closeness
Min.	0.4578	0.4199
1st Qu.	0.5278	0.5241
Median	0.5547	0.5801
Mean	0.5668	0.5724
3rd Qu.	0.5984	0.6281
Max.	0.7755	0.7755

Nodal Statistics - Measures of Centrality II

In- and Out-Closeness Centrality: define central a node if it is close, in terms of geodesic distance, to many others.

$$\tilde{\zeta}_i^{in-c} = (\zeta_i^{in-c})(n-1) \qquad \tilde{\zeta}_i^{out-c} = (\zeta_i^{out-c})(n-1)$$



	In-Closeness	Out-Closeness
Min.	0.4578	0.4199
1st Qu.	0.5278	0.5241
Median	0.5547	0.5801
Mean	0.5668	0.5724
3rd Qu.	0.5984	0.6281
Max.	0.7755	0.7755

Nodal Statistics - Measures of Centrality III

Betweenness Centrality: define central a node if it is located between many other nodes. $\tilde{\zeta}_i^b = \frac{\zeta_i^b}{(n-1)^2(n-2)}$

Eigenvector Centrality: defines central a node that is connected to other central nodes.

Nodal Statistics - Measures of Centrality III

Betweenness Centrality: define central a node if it is located between many other nodes. $\tilde{\zeta}_i^b = \frac{\zeta_i^b}{(n-1)^2(n-2)}$

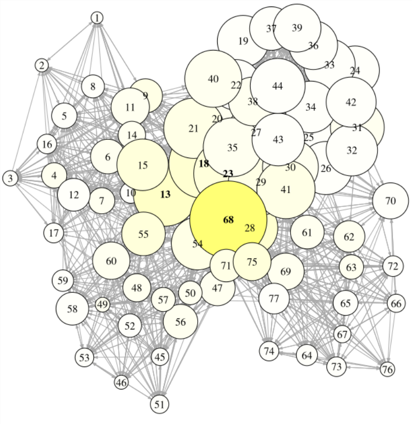
Eigenvector Centrality: defines central a node that is connected to other central nodes.

	Betweeness	Eigenvector
Min.	0.000007	0.1550
1st Qu.	0.0016	0.2985
Median	0.0048	0.4489
Mean	0.0104	0.4856
3rd Qu.	0.0119	0.6662
Max.	0.1305	1.0000

Nodal Statistics - Measures of Centrality III

Betweenness Centrality: define central a node if it is located between many other nodes. $\tilde{\zeta}_i^b = \frac{\zeta_i^b}{(n-1)^2(n-2)}$

Eigenvector Centrality: defines central a node that is connected to other central nodes.



	Betweeness	Eigenvector
Min.	0.000007	0.1550
1st Qu.	0.0016	0.2985
Median	0.0048	0.4489
Mean	0.0104	0.4856
3rd Qu.	0.0119	0.6662
Max.	0.1305	1.0000

Nodal Statistics

Centralization indexes:

Index	Formula	Value
CI^{in-d}	$\frac{\sum_i(\zeta_{max}^{in-d} - \zeta_i^{in-d})}{n-1}$	0.401
CI^{out-d}	$\frac{\sum_i(\zeta_{max}^{out-d} - \zeta_i^{out-d})}{n-1}$	0.401
CI^{in-c}	$\left[\sum_i \zeta_{max}^{in-d} - \zeta_i^{in-d}\right](n-1)$	0.214
CI^{out-c}	$\left[\sum_i \zeta_{max}^{in-d} - \zeta_i^{in-d}\right](n-1)$	0.208
CI^{b}	$\frac{\sum_i \zeta_{max}^{in-d} - \zeta_i^{in-d}}{(n-1)^2(n-2)}$	0.122

Models

Comparing different data generating process for our network

Simple Random Graph

Assumption:

$$Y_{ij} \overset{iid}{\sim} Bern(p)$$

Is number of edges ($\sum_{ij} y_{ij}$) fixed?

- No: $\implies$ Binomial Random Graph (BRG) $\in$ ERG family:

$\kappa(\mu)^{-1} = \prod_{ij} 1 + \exp\{\mu\}$.

- Yes: $\implies$ Uniform Random Graph (URG)

Simple Random Graph

Assumption:

$$Y_{ij} \overset{iid}{\sim} Bern(p)$$

Is number of edges ($\sum_{ij} y_{ij}$) fixed?

- No: $\implies$ Binomial Random Graph (BRG) $\in$ ERG family:

$\kappa(\mu)^{-1} = \prod_{ij} 1 + \exp\{\mu\}.$

- Yes: $\implies$ Uniform Random Graph (URG)

Results:

Simple Random Graph

Assumption:

$$Y_{ij} \stackrel{iid}{\sim} Bern(p)$$

Is number of edges ($\sum_{ij} y_{ij}$) fixed?

- No: $\implies$ Binomial Random Graph (BRG) $\in$ ERG family:

$\kappa(\mu)^{-1} = \prod_{ij} 1 + \exp\{\mu\}$.

- Yes: $\implies$ Uniform Random Graph (URG)

Results:

$$\hat{\mu} = -0.7779;$$

Simple Random Graph

Assumption:

$$Y_{ij} \overset{iid}{\sim} Bern(p)$$

Is number of edges ($\sum_{ij} y_{ij}$) fixed?

- No: $\Longrightarrow$ Binomial Random Graph (BRG) $\in$ ERG family:

$\kappa(\mu)^{-1} = \prod_{ij} 1 + \exp\{\mu\}.$

- Yes: $\Longrightarrow$ Uniform Random Graph (URG)

Results:

$$\hat{\mu} = -0.7779; \qquad \frac{\widehat{P(Y_{ij} = 1)}}{1 - \widehat{P(Y_{ij} = 1)}} = \exp\{\hat{\mu}\} = 0.4593;$$

Simple Random Graph

Assumption:

$$Y_{ij} \overset{iid}{\sim} Bern(p)$$

Is number of edges ($\sum_{ij} y_{ij}$) fixed?

- No: $\implies$ Binomial Random Graph (BRG) $\in$ ERG family:

$\kappa(\mu)^{-1} = \prod_{ij} 1 + \exp\{\mu\}$.

- Yes: $\implies$ Uniform Random Graph (URG)

Results:

$$\hat{\mu} = -0.7779; \qquad \frac{\widehat{P(Y_{ij} = 1)}}{1 - \widehat{P(Y_{ij} = 1)}} = \exp\{\hat{\mu}\} = 0.4593;$$

$$\widehat{P(Y_{ij} = 1)} = \frac{\exp\{\hat{\mu}\}}{1 + \exp\{\hat{\mu}\}} = 0.3148.$$

Simple Random Graph

Goodness of fit:

- Residual Deviance: 7290; (Null Deviance: 8113)
- MLE.LogLikelihood: -3644.975 (Null.LogLikelihood: -4056.297)
- Number of parameters: 1
- BIC: 7299

Simple Random Graph

Goodness of fit:

- Residual Deviance: 7290; (Null Deviance: 8113)
- MLE.LogLikelihood: -3644.975 (Null.LogLikelihood: -4056.297)
- Number of parameters: 1
- BIC: 7299

Good simulations only for density ρ and modularity *organization level* r_{org}.

Non Homogeneous Binomial Random Graph

Relaxing Homogeneity Assumption:

Non Homogeneous Binomial Random Graph

Relaxing Homogeneity Assumption:

$$Y_{ij} \sim Bern(p_{ij})$$

According the model:

$$Y_{ij} = \mu + a_i + b_j + \epsilon_{ij}; \quad \Longrightarrow \quad E[Y_{ij}] = \mu + a_i + b_j = \mu_{ij}$$

ERGM form:

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \kappa(\mu, \boldsymbol{a}, \boldsymbol{b})^{-1} \exp\left\{\mu y_{..} + \sum_i a_i y_{i.} + \sum_j b_j y_{.j}\right\}$$

$$\kappa(\mu, \boldsymbol{a}, \boldsymbol{b})^{-1} = \prod_{ij} 1 + \exp\{\mu + a_i + b_j\}.$$

Non Homogeneous Binomial Random Graph

Relaxing Homogeneity Assumption:

$$Y_{ij} \sim Bern(p_{ij})$$

According the model:

$$Y_{ij} = \mu + a_i + b_j + \epsilon_{ij}; \quad \implies \quad E[Y_{ij}] = \mu + a_i + b_j = \mu_{ij}$$

ERGM form:

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \kappa(\mu, \boldsymbol{a}, \boldsymbol{b})^{-1} \exp\left\{\mu y_{..} + \sum_i a_i y_{i.} + \sum_j b_j y_{.j}\right\}$$

$$\kappa(\mu, \boldsymbol{a}, \boldsymbol{b})^{-1} = \prod_{ij} 1 + \exp\{\mu + a_i + b_j\}.$$

Results:

Non Homogeneous Binomial Random Graph

Relaxing Homogeneity Assumption:

$$Y_{ij} \sim Bern(p_{ij})$$

According the model:

$$Y_{ij} = \mu + a_i + b_j + \epsilon_{ij}; \quad \Longrightarrow \quad E[Y_{ij}] = \mu + a_i + b_j = \mu_{ij}$$

ERGM form:

$$P(\mathbf{Y} = \mathbf{y}) = \kappa(\mu, \mathbf{a}, \mathbf{b})^{-1} \exp\left\{\mu y_{..} + \sum_i a_i y_{i.} + \sum_j b_j y_{.j}\right\}$$

$$\kappa(\mu, \mathbf{a}, \mathbf{b})^{-1} = \prod_{ij} 1 + \exp\{\mu + a_i + b_j\}.$$

Results:

- $\hat{\mu}$ = -0.7779;
- Greater values of a_i and b_j for the Global Department Manager (vertex 68).

Non Homogeneous Binomial Random Graph

Goodness of fit:

- Residual Deviance: 6399; (BRG Residual Deviance: 7290)
- MLE.LogLikelihood: -3199.748 (MLE.LogLikelihood BRG: -3644.975)
- Number of parameters: $2(n-1)+1=153$ (BRG: 1)
- BIC: 7727 (BRG BIC: 7299)

Compared to BRG, good simulations also for centralization measures in-degree CI^{in-d}, out-degree CI^{out-d}, in-closeness CI^{in-c} and out-closeness CI^{out-c}.

P1 Model

Dyad Assumption:

$$D_{ij} = (Y_{ij}, Y_{ji}) \overset{ind}{\sim} Multinomial(p_{ij})$$

$$P(D_{ij} = d_{ij}) = P\left[(Y_{ij}, Y_{ji}) = (y_{ij}, y_{ji})\right] = \frac{e^{\mu_{ij} y_{ij} + \mu_{ji} y_{ji} + \gamma y_{ij} y_{ji}}}{1 + e^{\mu_{ij}} + e^{\mu_{ji}} + e^{\gamma + \mu_{ij} + \mu_{ji}}}$$

where: γ reciprocity parameter constant across dyads

$\mu_{ij} = \mu + a_i + b_j$ depends on popularity and sociability

P1 Model

Dyad Assumption:

$$D_{ij} = (Y_{ij}, Y_{ji}) \overset{ind}{\sim} Multinomial(p_{ij})$$

$$P(D_{ij} = d_{ij}) = P\left[(Y_{ij}, Y_{ji}) = (y_{ij}, y_{ji})\right] = \frac{e^{\mu_{ij}y_{ij} + \mu_{ji}y_{ji} + \gamma y_{ij}y_{ji}}}{1 + e^{\mu_{ij}} + e^{\mu_{ji}} + e^{\gamma + \mu_{ij} + \mu_{ji}}}$$

where: γ reciprocity parameter constant across dyads

$\mu_{ij} = \mu + a_i + b_j$ depends on popularity and sociability

Results (after few tries):

P1 Model

Dyad Assumption:

$$D_{ij} = (Y_{ij}, Y_{ji}) \overset{ind}{\sim} Multinomial(p_{ij})$$

$$P(D_{ij} = d_{ij}) = P\left[(Y_{ij}, Y_{ji}) = (y_{ij}, y_{ji})\right] = \frac{e^{\mu_{ij}y_{ij} + \mu_{ji}y_{ji} + \gamma y_{ij}y_{ji}}}{1 + e^{\mu_{ij}} + e^{\mu_{ji}} + e^{\gamma + \mu_{ij} + \mu_{ji}}}$$

where: γ reciprocity parameter constant across dyads
$\mu_{ij} = \mu + a_i + b_j$ depends on popularity and sociability

Results (after few tries):

- $\hat{\mu} = -4.254$
- $\hat{\gamma} = 4.738$
- sender effects significant for numerous vertexes, mostly of Frankfurt
- receiver effects significant only for nodes 12, 13 and 54

P1 Model

Goodness of fit:

- Residual Deviance: 4965; (BRG Residual Deviance: 7290)
- MLE.LogLikelihood: -2482.381 (MLE.LogLikelihood BRG: -3644.975)
- Number of parameters: $2(n-1)+2=154$ (BRG: 1)
- BIC: 6301 (BRG BIC: 7299)

P1 Model

Goodness of fit:

- Residual Deviance: 4965; (BRG Residual Deviance: 7290)
- MLE.LogLikelihood: -2482.381 (MLE.LogLikelihood BRG: -3644.975)
- Number of parameters: $2(n-1)+2=154$ (BRG: 1)
- BIC: 6301 (BRG BIC: 7299)

Good simulations for μ, γ, r_{org}, In- and Out-Degree and In- and Out-Closeness.

Markov Model

Markov Assumption: two tie-variables are assumed to be independent unless they share a node.

$$Y_{ij} \perp Y_{kl} \quad \Longleftrightarrow \quad \{i, j\} \cap \{k, l\} = \emptyset.$$

Markov Model

Markov Assumption: two tie-variables are assumed to be independent unless they share a node.

$$Y_{ij} \perp Y_{kl} \quad \Longleftrightarrow \quad \{i, j\} \cap \{k, l\} = \emptyset.$$

Homogeneity Assumption: the effect of a configuration is the same regardless of vertexes involved.

Markov Model

Markov Assumption: two tie-variables are assumed to be independent unless they share a node.

$$Y_{ij} \perp Y_{kl} \iff \{i,j\} \cap \{k,l\} = \emptyset.$$

Homogeneity Assumption: the effect of a configuration is the same regardless of vertexes involved.

MCMC frequently did not converge: restricted choice of parameters.

Markov Model

Results: All parameters highly significant

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793
- $\hat{\sigma}_2^{out}$ = 0.248, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793
- $\hat{\sigma}_2^{out}$ = 0.248, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{nodefactor.location.2}$ = -0.294 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793
- $\hat{\sigma}_2^{out}$ = 0.248, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{nodefactor.location.2}$ = -0.294 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{nodefactor.org\_level.1}$ = 2.579 , (Global Manager)
  $\widehat{nodefactor.org\_level.2}$ = 0.457 (Local Manager);
  baseline: Project Leader and Researcher

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793
- $\hat{\sigma}_2^{out}$ = 0.248, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{\textit{nodefactor.location.2}}$ = -0.294 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{\textit{nodefactor.org\_level.1}}$ = 2.579 , (Global Manager)
  $\widehat{\textit{nodefactor.org\_level.2}}$ = 0.457 (Local Manager);
  baseline: Project Leader and Researcher
- $\widehat{\textit{nodefactor.tenure.1}}$ = -0.457 (1-12 months);
  baseline: 13-36, 37-60 and +61 months.

Markov Model

Results: All parameters highly significant

- $\hat{\mu}$ = -1.793
- $\hat{\sigma}_2^{out}$ = 0.248, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{\textit{nodefactor.location.2}}$ = -0.294 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{\textit{nodefactor.org\_level.1}}$ = 2.579 , (Global Manager)
  $\widehat{\textit{nodefactor.org\_level.2}}$ = 0.457 (Local Manager);
  baseline: Project Leader and Researcher
- $\widehat{\textit{nodefactor.tenure.1}}$ = -0.457 (1-12 months);
  baseline: 13-36, 37-60 and +61 months.
- $\widehat{\textit{nodematch.location}}$ = 4.634;
  other two *nodematch* estimates not significant and removed.

Markov Model

Goodness of fit:

- Residual Deviance: 3765; (P1 Residual Deviance: 4965)
- MLE.LogLikelihood: -1882.63 (MLE.LogLikelihood P1: -2482.381)
- Number of parameters: 8 (P1: 154)
- BIC: 3835 (P1 BIC: 6301)

Markov Model

Goodness of fit:

- Residual Deviance: 3765; (P1 Residual Deviance: 4965)
- MLE.LogLikelihood: -1882.63 (MLE.LogLikelihood P1: -2482.381)
- Number of parameters: 8 (P1: 154)
- BIC: 3835 (P1 BIC: 6301)

Observations:

- Without nodal attributes as covariates BIC results equal to 7013.
- Passable simulations except for reciprocity, transitivity and eigenvector centrality.

Social Circuit Model

Extension of Markov Assumption:

$$Y_{ij} \not\perp Y_{kl} \quad \Longleftrightarrow \quad Y_{ik} = Y_{jl} = 1 \vee Y_{il} = Y_{jk} = 1$$

Social Circuit Model

Extension of Markov Assumption:

$$Y_{ij} \not\perp Y_{kl} \quad \Longleftrightarrow \quad Y_{ik} = Y_{jl} = 1 \vee Y_{il} = Y_{jk} = 1$$

MCMC frequently did not converge: models with alternating k-2 paths parameter σ_w not available

Social Circuit Model

Results: All parameters highly significant

- $\hat{\mu}$ = -8.042
- $\hat{\sigma}_2^{out}$ = 0.243, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{\textit{nodefactor.location.2}}$ = -0.287 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{\textit{nodefactor.org\_level.1}}$ = 2.484, (Global Manager)
  $\widehat{\textit{nodefactor.org\_level.2}}$ = 0.443 (Local Manager);
  baseline: Project Leader and Reasercher
- $\widehat{\textit{nodefactor.tenure.1}}$ = -0.446 (1-12 months);
  baseline: 13-36, 37-60 and +61 months.
- $\widehat{\textit{nodematch.location}}$ = 4.639;
  other two *nodematch* estimates not significant and removed.

Social Circuit Model

Results: All parameters highly significant

- $\hat{\mu}$ = -8.042
- $\hat{\sigma}_2^{out}$ = 0.243, $\hat{\sigma}_3^{out}$ = -0.006;
  (other Markov typical configurations not available).
- $\widehat{\textit{nodefactor.location.2}}$ = -0.287 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{\textit{nodefactor.org\_level.1}}$ = 2.484, (Global Manager)
  $\widehat{\textit{nodefactor.org\_level.2}}$ = 0.443 (Local Manager);
  baseline: Project Leader and Reasercher
- $\widehat{\textit{nodefactor.tenure.1}}$ = -0.446 (1-12 months);
  baseline: 13-36, 37-60 and +61 months.
- $\widehat{\textit{nodematch.location}}$ = 4.639;
  other two *nodematch* estimates not significant and removed.
- $\hat{\sigma}_v^2$ = 2.0196

Social Circuit Model

Goodness of fit:

- Residual Deviance: 3758 (Markov Model Residual Deviance: 3765)
- MLE.LogLikelihood:-1878.84 (MLE.LogLikelihood Markov Model: -1882.63)
- Number of parameters: 10 (Markov Model: 9)
- BIC: 3836 (Markov Model BIC: 3835)

P1 model with attributes

Results:

- $\hat{\mu}$ = -2.941;
- $\hat{\gamma}$ = 2.548;
- $\widehat{\textit{nodefactor.location.2}}$ = 0.446 (Frankfurt);
  baseline: Paris, Geneva and Warsaw.
- $\widehat{\textit{nodefactor.org\_level.1}}$ = 1.879, (Global Manager)
  $\widehat{\textit{nodefactor.org\_level.2}}$ = 0.6103 (Local Manager);
  baseline: Project Leader and Reasercher
- $\widehat{\textit{nodefactor.tenure.1}}$ = -0.578 (1-12 months);
  baseline: 13-36, 37-60 and +61 months.
- $\widehat{\textit{nodematch.location}}$ = 2.750;
  other two *nodematch* estimates not significant and removed.

P1 model with attributes

Goodness of fit:

- Residual Deviance: 3686 (Markov Model Residual Deviance: 3765)
- MLE.LogLikelihood: -1842.802 (MLE.LogLikelihood Markov Model: -1882.63)
- Number of parameters: 7 (Markov Model: 9)
- BIC: 3746 (Markov Model BIC: 3835)

Model Choice

Model	BIC
SRG	7299
NH-SRG	7727
P1	6301
Markov	3835
Social	3836
P1 with attr.	3746

Model Choice

Model	BIC
SRG	7299
NH-SRG	7727
P1	6301
Markov	3835
Social	3836
P1 with attr.	3746

- Social and Markov are very similar and both very good;
- The **model** with lowest BIC is **P1 with attributes**.

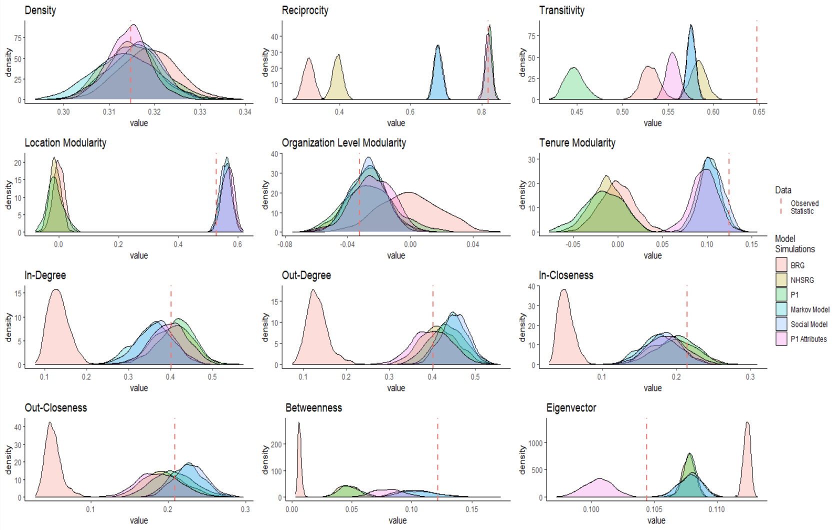
Model Choice

Model	BIC
SRG	7299
NH-SRG	7727
P1	6301
Markov	3835
Social	3836
P1 with attr.	3746

- Social and Markov are very similar and both very good;
- The **model** with lowest BIC is **P1 with attributes**.

MCMC diagnostics: All parameters converged.

All Model Simulations



Conclusion

Final observations

Conclusion

Statistics:

- Quite dense network;
- High values of transitivity and reciprocity $\rightarrow$ tendency to form clusters within same location (due to modularity values);
- Not very centralized network;
- Very important node: the only global department manager.

Conclusion

Statistics:

- Quite dense network;
- High values of transitivity and reciprocity $\rightarrow$ tendency to form clusters within same location (due to modularity values);
- Not very centralized network;
- Very important node: the only global department manager.

Models:

- Bad performance of simpler models (SRG and NH-SRG);
- Improvements with P1;
- Good performances with Social Circuit and Markov Model both with respect of BIC and fit in simulations;
- Best performance for P1 with attributes.

Thank you for the attention!