Introduction to Social Network Analysis

Weihua An

Indiana University Bloomington
Departments of Statistics and Sociology
Presenation at the SSRC Workshop in Methods

April 18, 2014

Outline

- The Wide Use of SNA

Outline

- The Wide Use of SNA
- What is SNA

Outline

- The Wide Use of SNA
- What is SNA
    - Four Elements

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches

Outline

- The Wide Use of SNA
- What is SNA
    - Four Elements
    - Five Major Approaches
        - Descriptive analysis

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis
    - Causal analysis

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis
    - Causal analysis
    - Predictive analysis

Outline

- The Wide Use of SNA
- What is SNA
    - Four Elements
    - Five Major Approaches
        - Descriptive analysis
        - Formal analysis
        - Causal analysis
        - Predictive analysis
        - Intervention analysis

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis
    - Causal analysis
    - Predictive analysis
    - Intervention analysis
- More resources

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis
    - Causal analysis
    - Predictive analysis
    - Intervention analysis
- More resources
  - Books and Readings

Outline

- The Wide Use of SNA
- What is SNA
  - Four Elements
  - Five Major Approaches
    - Descriptive analysis
    - Formal analysis
    - Causal analysis
    - Predictive analysis
    - Intervention analysis
- More resources
  - Books and Readings
  - Courses

Figure 1. Map of Sciences and Social Sciences

Source: http://www.eigenfactor.org/map/images/Sci2004.pdf

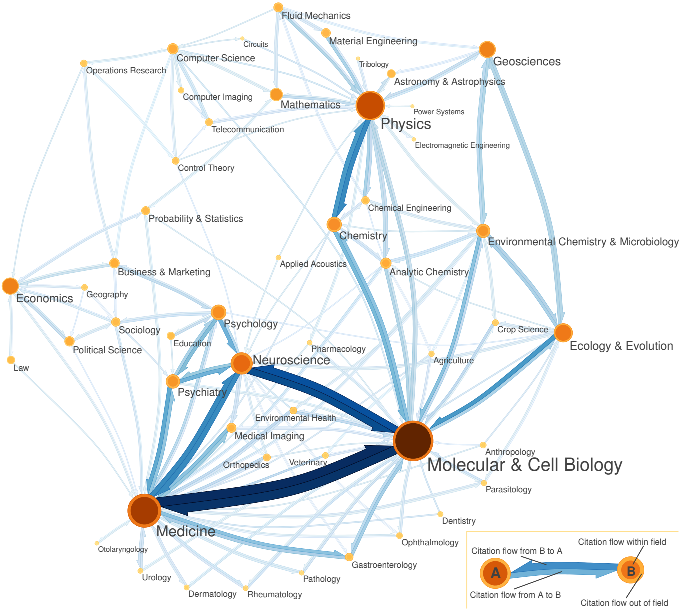


Figure 1b. Map of Social Sciences

Source: http://www.eigenfactor.org/map/images/SocSci2004.pdf

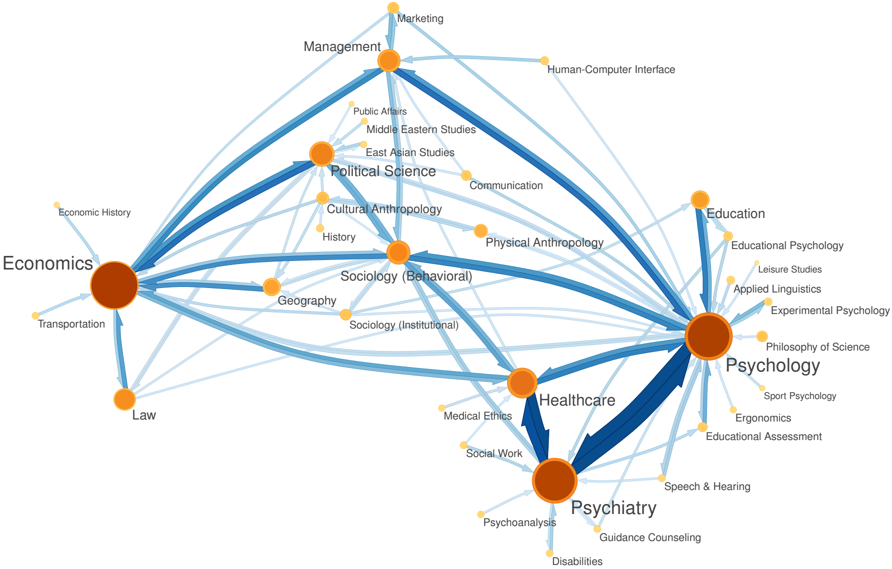


Figure 2. Friendships in a High School Colored by Grade and Excluding Isolates
Data Source: Goodreau et al. (2008)

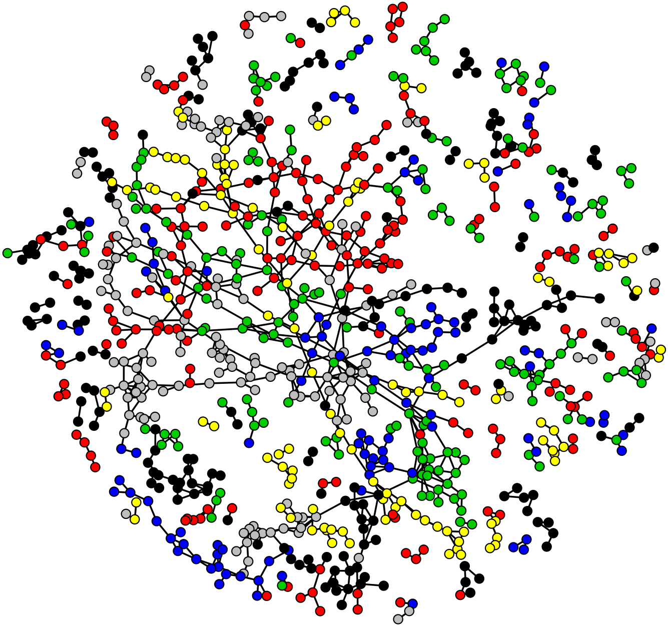


Figure 3. Friendships in a Middle School in China
Source: An (2011)

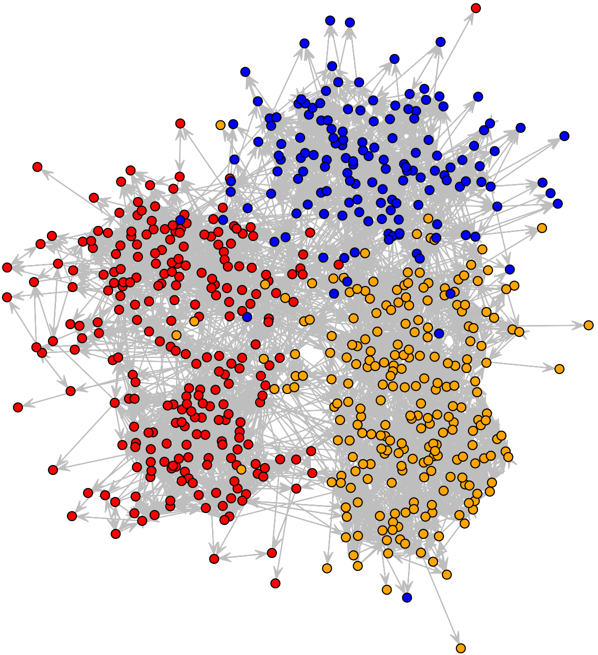


Figure 2. Friendships in a High School Colored by Grade and Excluding Isolates
Data Source: Goodreau et al. (2008)

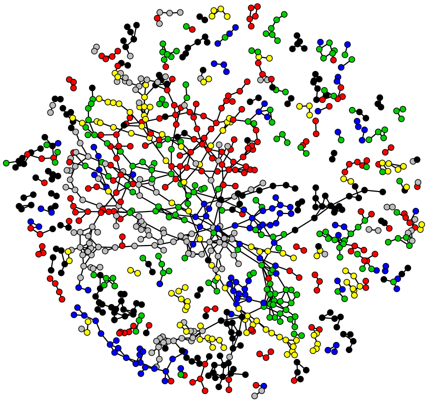


Figure 3. Friendships in a Middle School in China
Source: An (2011)

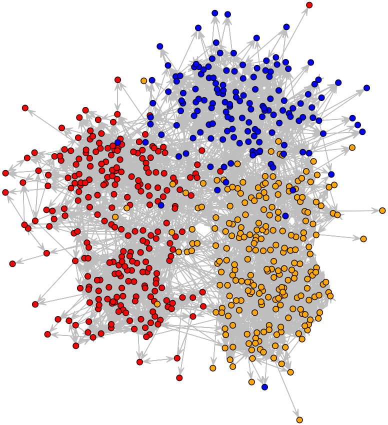


Figure 2b. Friendships in a High School Colored by Sex and Excluding Isolates
Data Source: Goodreau et al. (2008)

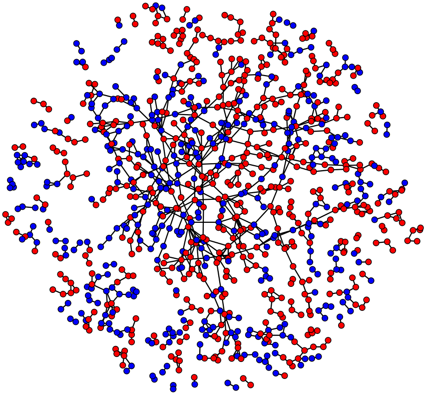


Figure 3b. Friendships in a Middle School in China Colored by Sex
Source: An (2011)

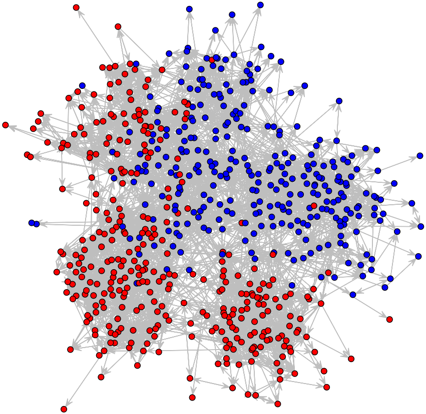


Figure 4. Friendship and Lunchroom Seating Networks in an Elementrary School

Source: Calarco, An, and McConnell (2013)

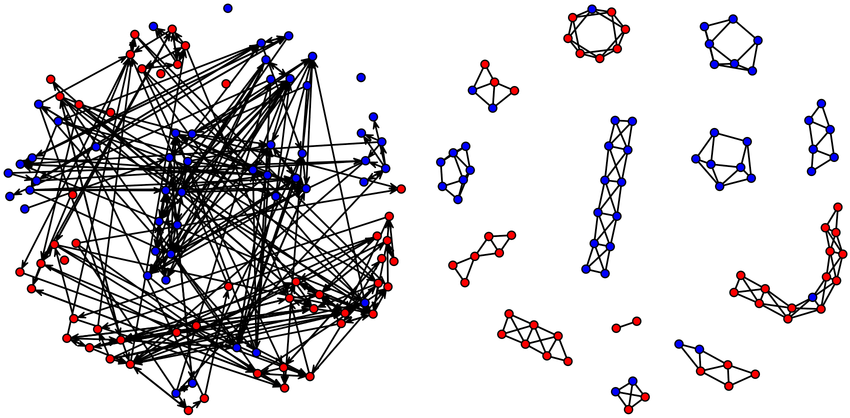


Figure 5. Chains of Affection: Romantic Relationships in Jefferson High

Source: Bearman et al. (2004)

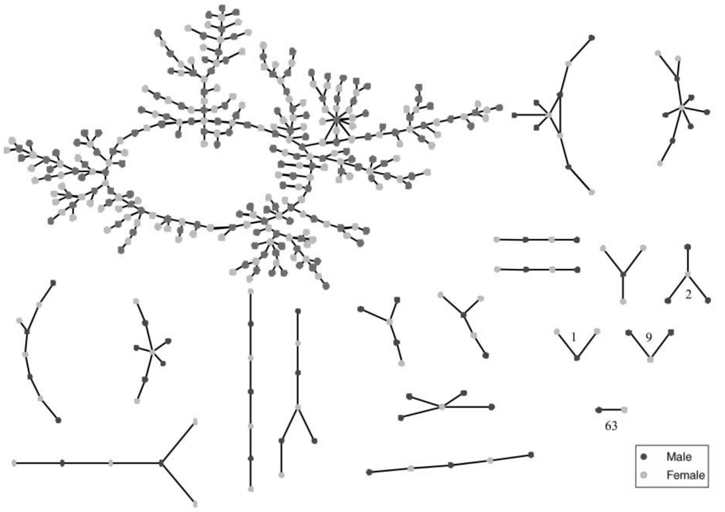


Figure 6. Marriage and Business Networks of the Florentien Notable Families

Source: Padgett (1994)

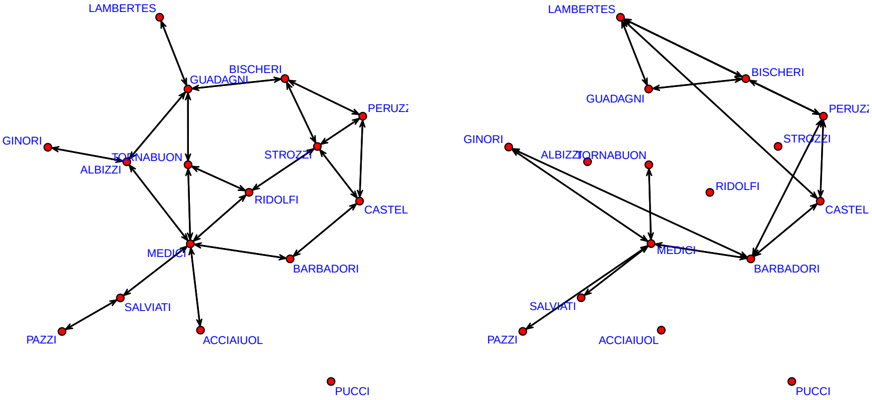


Figure 7. Inter-organizational Network in Response to Hurricane Katrina

Source: Kapucu et al. (2010)

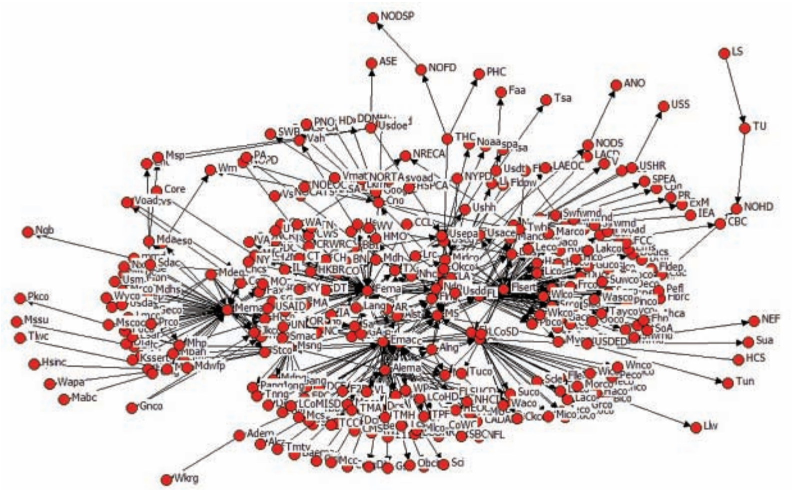


Figure 8. Policy Network of Elected Officials in the Orlando Metropolitan Area

Source: Feiock et al. (2010)

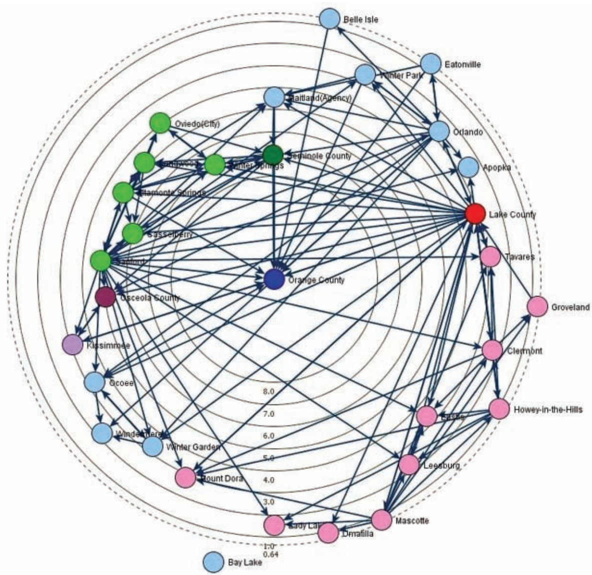
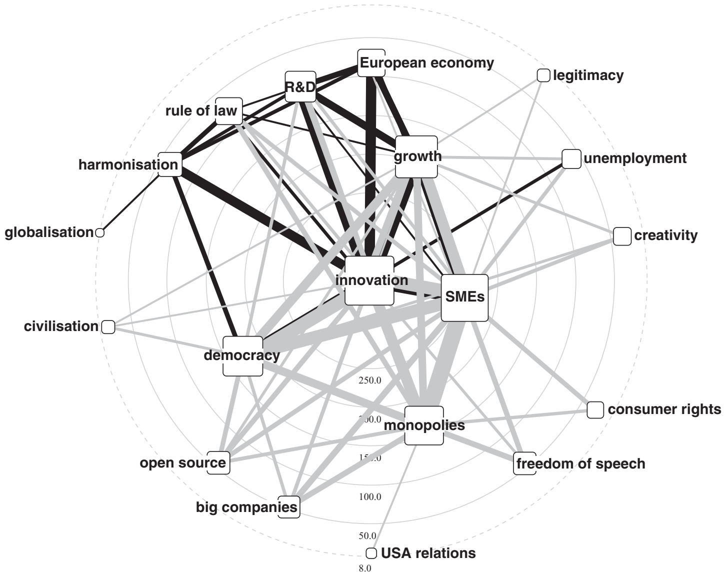


Figure 9. Concept Network in Discourse Analysis

Source: Leifeld and Haunss (2012)



Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks
- Biology: cell interactions, brain activities, system biology

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks
- Biology: cell interactions, brain activities, system biology
- Computer science and informatics: computer networks, social media (e.g., Facebook, Twitter)

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks
- Biology: cell interactions, brain activities, system biology
- Computer science and informatics: computer networks, social media (e.g., Facebook, Twitter)
- Statistics: random network models

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks
- Biology: cell interactions, brain activities, system biology
- Computer science and informatics: computer networks, social media (e.g., Facebook, Twitter)
- Statistics: random network models
- Math: graph theory

Bursts of Social Network Analysis (SNA)

- Sociology: career attainment and mobility, friendships, advising relationships, gift exchange, holiday visits, board interlocking, diffusion of innovations, contagion of health and criminal behaviors and outcomes
- Political science: inter-governmental cooperation, international relations, networking and networks in bureaucracy, bill sponsorship, voting and election influence, social movement and collective action
- Economics: suppliers, international trade, shareholders network, spillover of productivity
- Communication: social marketing, information diffusion, citation networks
- Biology: cell interactions, brain activities, system biology
- Computer science and informatics: computer networks, social media (e.g., Facebook, Twitter)
- Statistics: random network models
- Math: graph theory
- Literature: conversation networks, co-play networks

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions
  - What roles do motivation, percetion, and cognition play?

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions
  - What roles do motivation, percetion, and cognition play?
- Relational data: Not only between people but also between organizations or objects (e.g., words, books, concepts, topics) that can co-occur.

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions
  - What roles do motivation, percetion, and cognition play?
- Relational data: Not only between people but also between organizations or objects (e.g., words, books, concepts, topics) that can co-occur.
- Graphic display

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions
  - What roles do motivation, percetion, and cognition play?
- Relational data: Not only between people but also between organizations or objects (e.g., words, books, concepts, topics) that can co-occur.
- Graphic display
- Quantitative analysis

What is SNA

Freeman (2004) defined four essential elements of SNA

- Structural perspective: Patterns of relationships and interactions
  - What roles do motivation, percetion, and cognition play?
- Relational data: Not only between people but also between organizations or objects (e.g., words, books, concepts, topics) that can co-occur.
- Graphic display
- Quantitative analysis
  - The revival of qualitative approaches (e.g., interviews, ethnographic observations) to SNA

Five Major Approaches

- Descriptive analysis: Describe the features of social connections (Wasserman and Faust 1994)

Five Major Approaches

- Descriptive analysis: Describe the features of social connections (Wasserman and Faust 1994)
- Formal analysis: Use statistical or mathematical models to characterize the network formation process (Jackson 2008; Kolaczyk 2009)

Five Major Approaches

- Descriptive analysis: Describe the features of social connections (Wasserman and Faust 1994)
- Formal analysis: Use statistical or mathematical models to characterize the network formation process (Jackson 2008; Kolaczyk 2009)
- Causal analysis: Identify and quantify the effects of social connections and networks

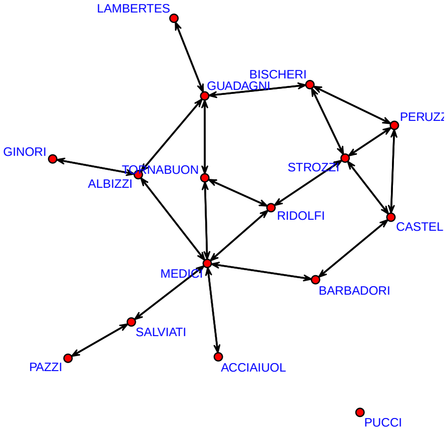
Five Major Approaches

- Descriptive analysis: Describe the features of social connections (Wasserman and Faust 1994)
- Formal analysis: Use statistical or mathematical models to characterize the network formation process (Jackson 2008; Kolaczyk 2009)
- Causal analysis: Identify and quantify the effects of social connections and networks
- Predictive analysis: Use principles found in social network analysis to predict connections or behaviors

Five Major Approaches

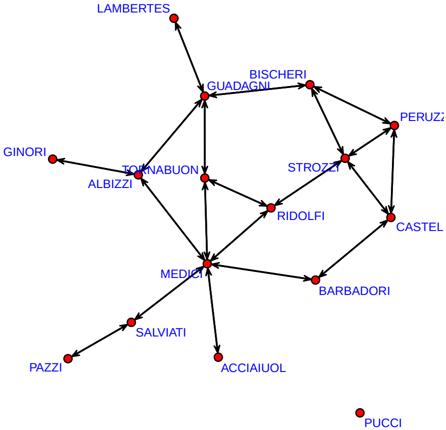
- Descriptive analysis: Describe the features of social connections (Wasserman and Faust 1994)
- Formal analysis: Use statistical or mathematical models to characterize the network formation process (Jackson 2008; Kolaczyk 2009)
- Causal analysis: Identify and quantify the effects of social connections and networks
- Predictive analysis: Use principles found in social network analysis to predict connections or behaviors
- Intervention analysis: Utilize the features of social networks to design more effective policy programs

1. Descriptive Analysis



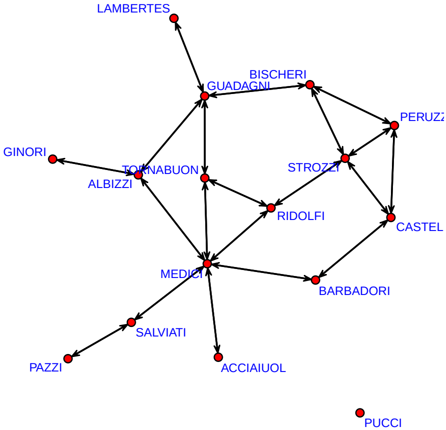
- Node

1. Descriptive Analysis



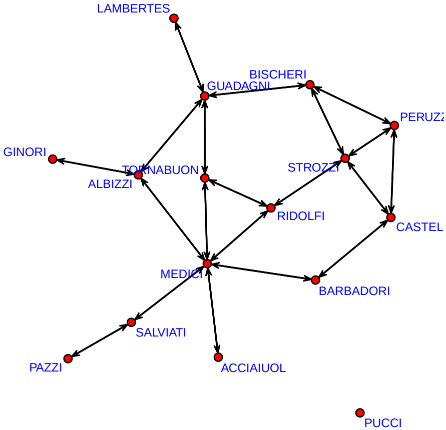
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector

1. Descriptive Analysis



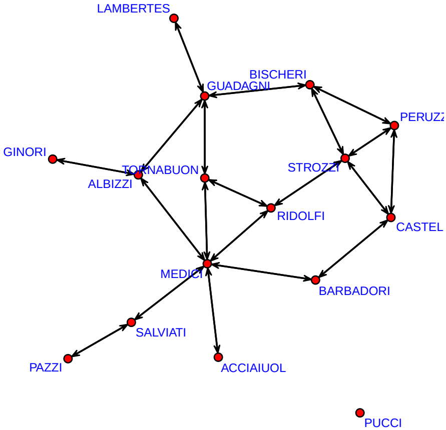
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad

1. Descriptive Analysis



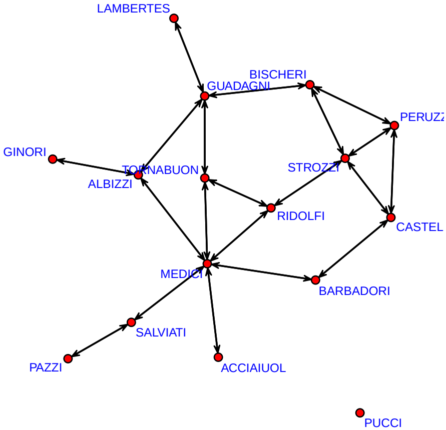
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence

1. Descriptive Analysis



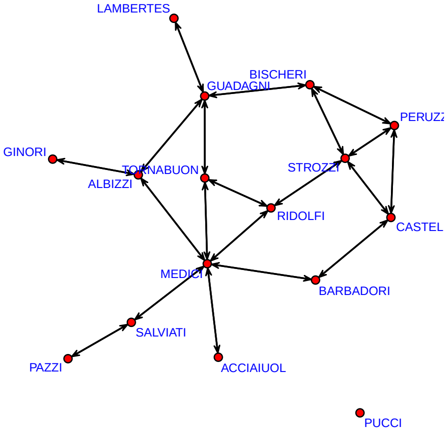
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group

1. Descriptive Analysis



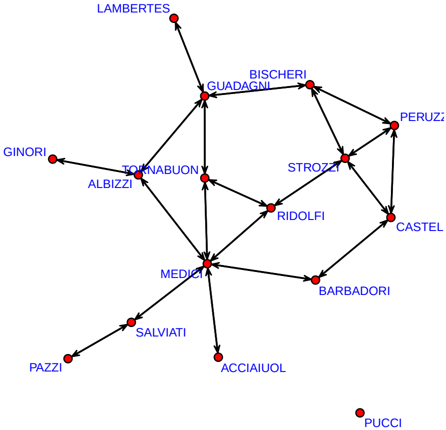
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group
  - Triad, cliques, component

1. Descriptive Analysis



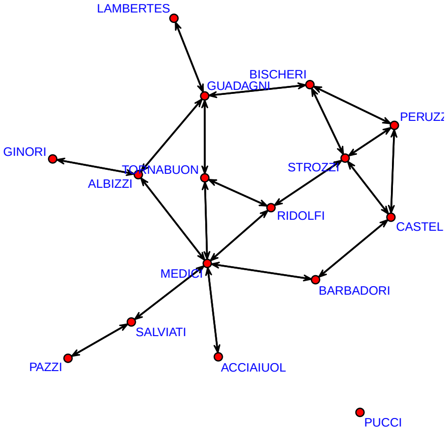
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group
  - Triad, cliques, component
  - Hierarchical clustering

1. Descriptive Analysis



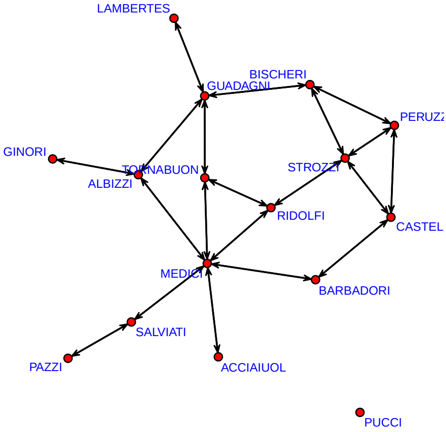
- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group
  - Triad, cliques, component
  - Hierarchical clustering
  - Core and periphery

1. Descriptive Analysis



- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group
  - Triad, cliques, component
  - Hierarchical clustering
  - Core and periphery
- Network

1. Descriptive Analysis



- Node
  - Centrality: indegree, outdegree, betweenness, closeness, and eigenvector
- Dyad
  - Distance, structural equivalence
- Group
  - Triad, cliques, component
  - Hierarchical clustering
  - Core and periphery
- Network
  - Density, centralization, transitivity, clustering coefficient

Matrix Presentation of the Florentine Marriage Network

	ACCIAIUOL	ALBIZZI	BARBADOR	BISCHERI	CASTELLAN	GINORI	GUADAGNI	LAMBERTE	MEDICI	PAZZI	PERUZZI	PUCCI	RIDOLFI	SALVIATI	STROZZI	TORNABUC
ACCIAIUOL	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
ALBIZZI	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0
BARBADOR	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
BISCHERI	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
CASTELLAN	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0
GINORI	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GUADAGNI	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
LAMBERTE	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
MEDICI	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1
PAZZI	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
PERUZZI	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0
PUCCI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RIDOLFI	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
SALVIATI	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
STROZZI	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0
TORNABUC	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0

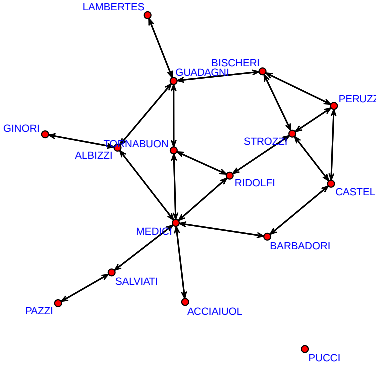


Table 1. Centrality Measures

	Degree	Closeness	Betweenness	Eigenvector
MEDICI	6	0.63	95.00	0.43
GUADAGNI	4	0.54	46.33	0.29
STROZZI	4	0.52	18.67	0.36
ALBIZZI	3	0.52	38.67	0.24
BISCHERI	3	0.48	19.00	0.28
CASTELLAN	3	0.46	10.00	0.26
PERUZZI	3	0.45	4.00	0.28
RIDOLFI	3	0.53	20.67	0.34
TORNABUON	3	0.52	16.67	0.33
BARBADORI	2	0.47	17.00	0.21
SALVIATI	2	0.44	26.00	0.15
ACCIAIUOL	1	0.39	0.00	0.13
GINORI	1	0.36	0.00	0.07
LAMBERTES	1	0.36	0.00	0.09
PAZZI	1	0.32	0.00	0.04
PUCCI	0	0.00	0.00	0.00

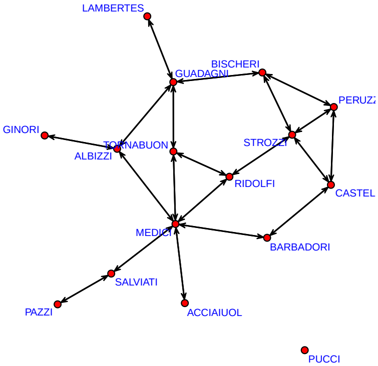
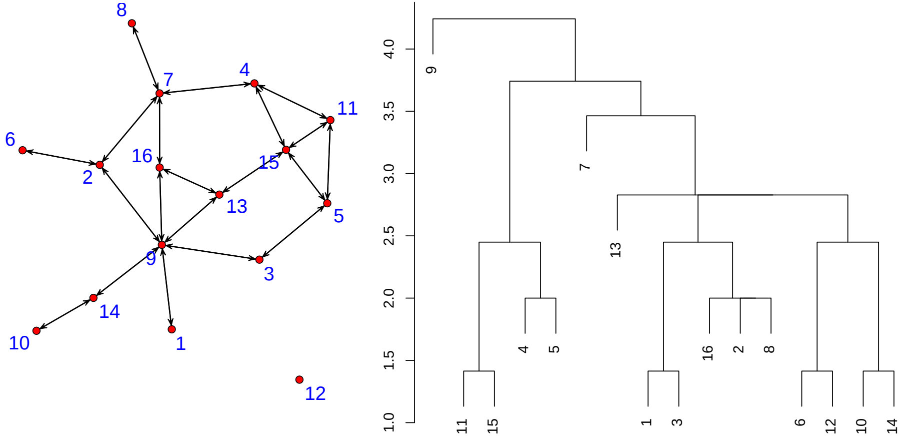


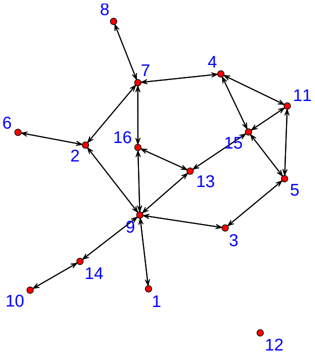
Table 2. Summary Statistics of the Network

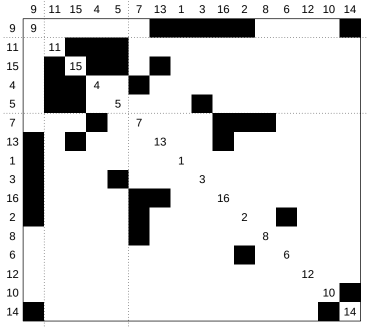
Statistics	Frequence
Dyad	
Mutual	20
Asymmetric	0
Null	100
Triangle	3
Clique	
3	3
2	12
1	1
Component	
15	1
1	1
Network	**Coefficient**
Density	0.17
Centralization	0.27
Transitivity	0.19

Hierarchical Clustering Based on Structural Equivalence

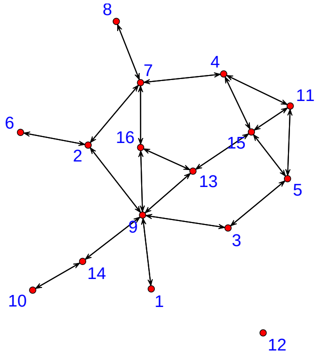


Blockmodeling





Blockmodeling



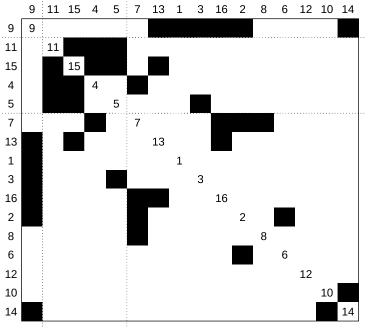
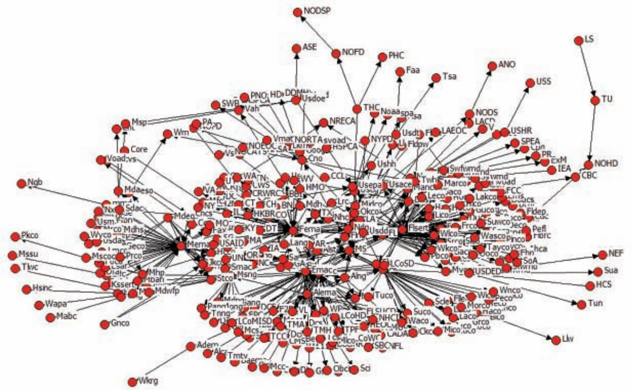


Table 3. Inter-Block Relationships

	Block 1	Block 2	Block 3
Block 1	0.10	0.07	0.55
Block 2	0.07	0.63	0.00
Block 3	0.55	0.00	0.00

Figure 7. Inter-organizational Network in Response to Hurricane Katrina

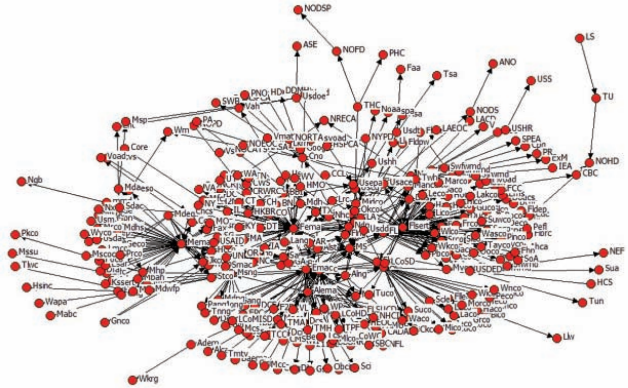
Source: Kapucu et al. (2010)



- Nine of the central players are state-level agencies.

Figure 7. Inter-organizational Network in Response to Hurricane Katrina

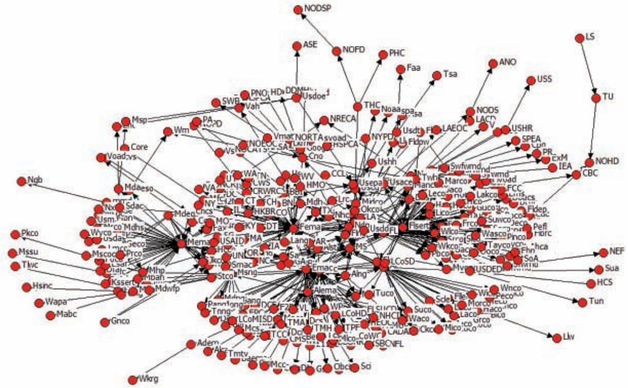
Source: Kapucu et al. (2010)



- Nine of the central players are state-level agencies.
- Large distance between actors.

Figure 7. Inter-organizational Network in Response to Hurricane Katrina

Source: Kapucu et al. (2010)



- Nine of the central players are state-level agencies.
- Large distance between actors.
- A great heterogeneity in the betweenness power of the actors.

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)
- Strength of weak ties (Granovetter 1973)

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)
- Strength of weak ties (Granovetter 1973)
- Small world (Kochen and Pool 1978; Watts 1999)

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)
- Strength of weak ties (Granovetter 1973)
- Small world (Kochen and Pool 1978; Watts 1999)
- Preferential attachment (Barabsi 1999)

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)
- Strength of weak ties (Granovetter 1973)
- Small world (Kochen and Pool 1978; Watts 1999)
- Preferential attachment (Barabsi 1999)
- Biases in cognitive networks: surplus of balancing relationships, overestimation of self-centrality

Important Findings in Descriptive Network Analysis

- Chains of opportunity (White 1970)
- Strength of weak ties (Granovetter 1973)
- Small world (Kochen and Pool 1978; Watts 1999)
- Preferential attachment (Barabsi 1999)
- Biases in cognitive networks: surplus of balancing relationships, overestimation of self-centrality
- Measurement error: forgetting friends

2. Formal Analysis

- Exponential random graph models (ERGMs)

2. Formal Analysis

- Exponential random graph models (ERGMs)
- Mathematical models of networks

ERGMs

Researchers have developed ERGMs to study the patterns of connections in an observed network in a more quantitative way (Handcock et al. 2003; Robins et al. 2007). Briefly speaking, in an ERGM the probability of observing a network, w, is assumed to be

$$\text{Prob}(W = w|X) = \frac{exp(\theta^T g(w, X))}{K},$$

where W is a random network, w represents the observed network, X the covariates, $g(w, X)$ is a function of the covariates and some network formation processes of interest (e.g., mutuality, transitivity), a vector of coefficients measuring their effects, and K a normalizing constant which ensures the probability sum to 1.

ERGMs

Prior research (Hunter et al. 2008) has shown that the ERGM is somewhat equivalent to an extended logit model:

$$\text{logit}(w_{ij} = 1 | w^r, X) = \theta^T \delta^{ij}(w, X),$$

where the log odds of actor i sending a tie to j (i.e., $w_{ij} = 1$), conditioning on the covariates X and the rest of the network w^r, is dependent on the change statistics $\delta^{ij}(w, X)$ (i.e., the changes in the covariates values and network features when w_{ij} flips from 0 to 1) and their effects as measured by the coefficient vector θ. Hence, the estimated coefficients from the ERGM can be interpreted as the logged odds ratio.

Table 4. Covariates

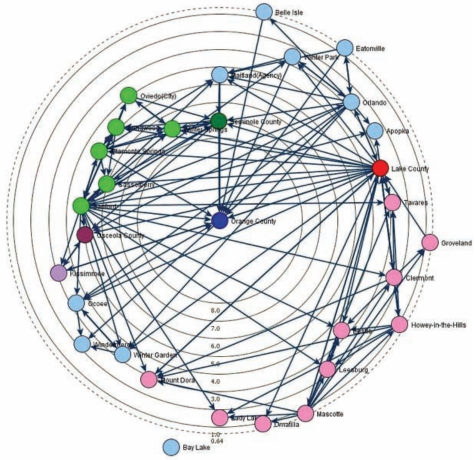
ID	Family	Wealth	Seats	Ties
1	ACCIAIUOL	10	53	1
2	ALBIZZI	36	65	0
3	BARBADORI	55	0	12
4	BISCHERI	44	12	6
5	CASTELLAN	20	22	15
6	GINORI	32	0	8
7	GUADAGNI	8	21	10
8	LAMBERTES	42	0	13
9	MEDICI	103	53	48
10	PAZZI	48	0	6
11	PERUZZI	49	42	29
12	PUCCI	3	0	1
13	RIDOLFI	27	38	1
14	SALVIATI	10	35	3
15	STROZZI	146	74	25
16	TORNABUON	48	0	4

Table 5. ERGM Results

	Model I			Model II		
	Coef.	SE	P	Coef.	SE	P
Constant	-3.15	0.50	0.00	-3.17	0.64	0.00
Main Effect						
Wealth	0.00	0.01	0.99	0.00	0.01	0.96
Seats in city coucil	0.02	0.01	0.09	0.02	0.01	0.11
Ties with other families	-0.01	0.02	0.44	-0.01	0.02	0.45
Homophily						
Abs. difference in wealth	0.02	0.01	0.02	0.02	0.01	0.02
Abs. difference in seats	-0.01	0.01	0.37	-0.01	0.01	0.38
Abs. difference in other ties	0.01	0.02	0.53	0.01	0.02	0.53
Other Network Tie						
Business tie	2.70	0.52	0.00	2.69	0.52	0.00
Structural Effect						
Tirangles (gwesp)				0.08	0.29	0.79
Twopaths (gwdsp)				-0.02	0.16	0.92
AIC	185.90			189.80		

Figure 8. Policy Network of Elected Officials in the Orlando Metropolitan Area

Source: Feiock et al. (2010)



- Build clustered local networks with high reciprocity and transitivity to enhance trustworthiness and resolve cooperative problems.

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.
- Examples:

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.
- Examples:
  - Utilitarian networks: If people form links purely due to utilitarian considerations, the structure will be composed of simple stars, etc.

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.
- Examples:
  - Utilitarian networks: If people form links purely due to utilitarian considerations, the structure will be composed of simple stars, etc.
  - Games in social network: the effects of network size and the efficiency of networks

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.
- Examples:
  - Utilitarian networks: If people form links purely due to utilitarian considerations, the structure will be composed of simple stars, etc.
  - Games in social network: the effects of network size and the efficiency of networks
  - Transmission of infectious diseases: how much immunization is sufficient to prevent the outbreaks of epidemics depends on the structure of social networks, especially the level of heterogeneity in degree. The higher, the faster.

Mathematical models of networks

- Main goals: Use mathematical models to describe or simulate the generation, development, and structural features of social networks.
- Examples:
  - Utilitarian networks: If people form links purely due to utilitarian considerations, the structure will be composed of simple stars, etc.
  - Games in social network: the effects of network size and the efficiency of networks
  - Transmission of infectious diseases: how much immunization is sufficient to prevent the outbreaks of epidemics depends on the structure of social networks, especially the level of heterogeneity in degree. The higher, the faster.
  - Phase transition: When $P = 1/2$, a big component will arise almost surely.

3. Causal Network Analysis

Three types of network effects:

- Relational effects

3. Causal Network Analysis

Three types of network effects:

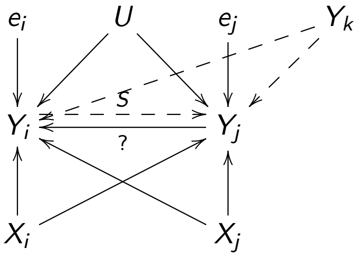
- Relational effects
- Positional effects: structural holes, structural equivalence

3. Causal Network Analysis

Three types of network effects:

- Relational effects
- Positional effects: structural holes, structural equivalence
- Structural effects: density, cohesion, structure

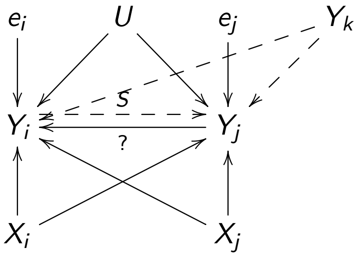
The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding

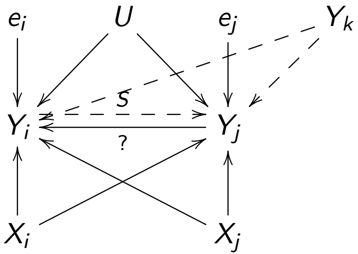
The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding
- Peer selection (homophily)

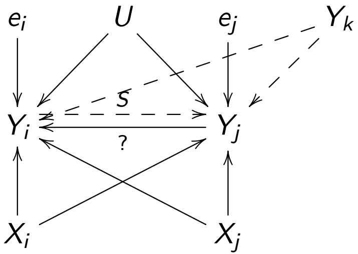
The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding
- Peer selection (homophily)
- Simultaneity

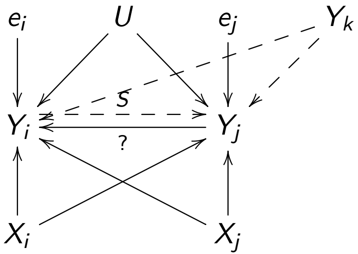
The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding
- Peer selection (homophily)
- Simultaneity
- Measurement error

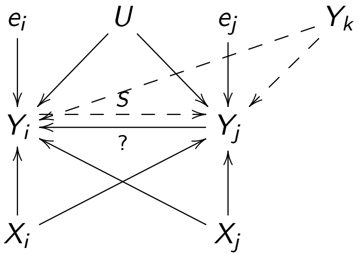
The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding
- Peer selection (homophily)
- Simultaneity
- Measurement error

The Challenges



However it turns out to be very difficult to estimate causal peer effects due to

- Contextual confounding
- Peer selection (homophily)
- Simultaneity
- Measurement error

A heated debate has been going on in the field for a while.

Possible solutions

An (2011) and VanderWeele and An (2013) discuss some possible solutions:

- Experiments

Possible solutions

An (2011) and VanderWeele and An (2013) discuss some possible solutions:

- Experiments
- Instrument variable methods

Possible solutions

An (2011) and VanderWeele and An (2013) discuss some possible solutions:

- Experiments
- Instrument variable methods
- Dynamic network models (Snijders 2001; Steglich and Snijders 2010)

3a. Experiments

There are two types of experiments that are useful to provide causal estimates of peer effects.

- Type I: random assignment of contacts

3a. Experiments

There are two types of experiments that are useful to provide causal estimates of peer effects.

- Type I: random assignment of contacts
- Type II: partial treatment design

Type I Experiment

The type I experiment is random assignment of contacts. This is meant to eliminate the selection problem.

- Sacerdote (2001) found that randomly assigned roommates and dormmates had significant impact on the grade point average (GPA) of students in a college and their decisions to join social groups such as fraternities.

Type I Experiment

The type I experiment is random assignment of contacts. This is meant to eliminate the selection problem.

- Sacerdote (2001) found that randomly assigned roommates and dormmates had significant impact on the grade point average (GPA) of students in a college and their decisions to join social groups such as fraternities.
- Boisjoly et al. (2006) found that students randomly assigned with African-American roommates were more likely to endorse affirmative action.

Type II Experiment

However, sometimes it might be infeasible or unethical to randomly assign contacts to subjects. In this study, I propose a second type of experiment which is particularly useful in such situations.

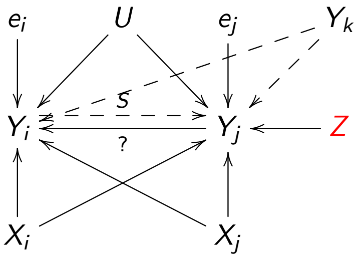
Type II Experiment

However, sometimes it might be infeasible or unethical to randomly assign contacts to subjects. In this study, I propose a second type of experiment which is particularly useful in such situations.

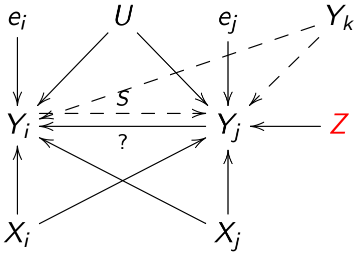
An (2011) proposed a type II experiment with a partial treatment design, in which only partial members of the treated groups are assigned to an intervention and how the effects of the intervention diffuse via social ties are examined.

$$\begin{tikzcd} & \text{Intervention} \arrow[d] \\ Y_i & Y_j \arrow[l, "?"'] \end{tikzcd}$$

3b. IV Methods



3b. IV Methods



An (2011) used six variables as IVs for peer smoking in order to study peer effects on smoking:

- Parental attitudes toward their childrens smoking
- Father's smoking status
- Siblings' smoking status
- Whether any relatives are sick due to smoking
- Whether cigarettes are stored at home year-round
- Distance from home to the nearest cigarette store

3c. Dynamic Network Models

Here I focus on the stochastic actor-oriented model (SAOM) (Snijders 2001, 2005; Snijders et al. 2009; Steglich et al. 2010).

3c. Dynamic Network Models

Here I focus on the stochastic actor-oriented model (SAOM) (Snijders 2001, 2005; Snijders et al. 2009; Steglich et al. 2010).

SAOM assumes changes in network and behavior follow two continuous Markov processes. The frequency of the two types of changes are determined by two rate functions: λ_N for network and λ_B for behavior. The waiting time for any change is assumed to follow an exponential distribution, $P(T > t) = e^{-(\lambda_N + \lambda_B)t}$. Subjects make changes according to two objective functions, which are assumed to be a linear summation of the effects of network structures and behavioral features.

$$f_i^N(w, w', z) = \sum_k \beta_k^N S_k^N(i, w, w', z, z'), \quad (1)$$

$$f_i^B(w, w', z) = \sum_k \beta_k^B S_k^B(i, w, w', z, z'). \quad (2)$$

w and w' represent the network statistics of subject i and its peers, and z and z' their covariates and behaviors.

Table 6. SAOM Results of Friendship Dynamics among Students

Friendship Dynamics	Explanations	Estimates	SE
smoking alter	Smokers tend to have more friends.	0.15	0.21
smoking ego	Smokers tend to nominate more friends.	0.36	0.22
same smoking	Smokers tend to be friends with other smokers.	-0.34	0.40
same smoking (break)	Smokers tend to break ties with other smokers.	1.07	0.72
eversmoking alter	Eversmokers tend to have more friends.	-0.02	0.08
eversmoking ego	Eversmokers tend to nominate more friends.	-0.25	0.07
same eversmoking	Eeversmokers tend to be friends with other eversmokers.	0.09	0.05
basic rate friendship	Basic rate of friendship changes.	18.23	0.81
outdegree (density)	Basic pattern of the network.	-3.05	0.19
reciprocity	Friendships tend to be reciprocated.	1.65	0.06
transitive ties	Friendships tend to form triangles.	1.28	0.05
indegree - popularity	Popular students tend to attract more friends.	0.00	0.01
outdegree - popularity	Active students tend to have more friends.	-0.07	0.02
boy alter	Boys tend to have more friends.	0.00	0.05
same boy	Friends tend to be same gender.	1.23	0.13
same boy (break)	Friendship ties with same gender tend to break.	-1.40	0.27
age alter	Older students tend to have more friends.	-0.01	0.03
age similarity	Students with similar age tend to be friends.	0.34	0.12
height alter	Taller students tend to have more friends.	0.00	0.00
height similarity	Students with similar height tend to be friends.	-0.27	0.13
weight alter	Heavier students tend to have more friends.	0.00	0.00
weight similarity	Students with similar weight tend to be friends.	0.22	0.19
ranking alter	Low ranked students tend to have more friends.	-0.05	0.02
ranking similarity	Similar ranked students tend to be friends.	0.13	0.08
paedu similarity	Students with similar family background tend to be friends.	0.10	0.12

Table 6 (Continued). SAOM Results of Smoking Dynamics among Students

Behavior Dynamics	Explanations	Estimates	SE
average alter	Students' smoking status is influenced by their friends.	-4.83	38.37
rate smoking period 1	Prevalence of smoking.	1.22	0.35
linear shape	Smoking trend in the long run.	-6.02	19.42
indegree	Popular students tend to smoke.	0.89	4.15
outdegree	Active students tend to smoke.	-1.52	7.72
treatment	Students in treatment groups tend to smoke.	-1.30	7.12
pasmoking	Students whose father smoke tend to smoke.	-1.92	8.85
sibsmoking	Students whose siblings smoke tend to smoke.	6.29	26.31
boy	Boys tend to smoke.	0.80	5.35
age	Older students tend to smoke.	2.46	9.88
height	Taller students tend to smoke.	-0.16	0.85
weight	Heavier students tend to smoke.	-0.16	0.77
ranking	Lower ranked students tend to smoke.	0.47	1.91
paedu	Students with better educated dad tend to smoke.	0.59	3.28

4. Network Predictions

- Relational Predictions

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity
- Behavioral Predictions

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity
- Behavioral Predictions
  - Nearest neighbor predicting

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity
- Behavioral Predictions
  - Nearest neighbor predicting
  - Network sensoring

4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity
- Behavioral Predictions
  - Nearest neighbor predicting
  - Network sensoring
  - Network surveillance

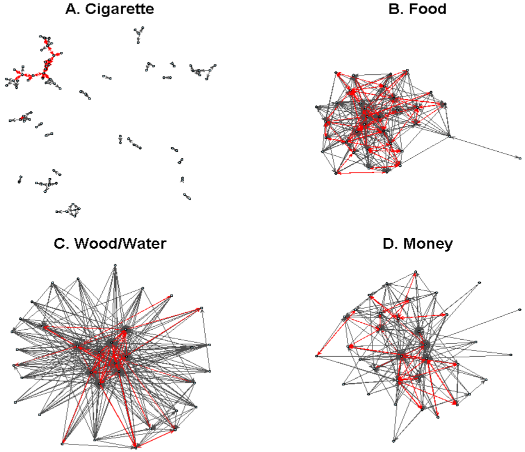
4. Network Predictions

- Relational Predictions
  - Model based. Training data $->$ Estimate parameters $->$ make predictions.
  - Quotation (text analysis), phone calls
  - Random walks: friends of friends are usually more central; the persons you meet are usually more active
  - Attributes-based homophily or complementarity
- Behavioral Predictions
  - Nearest neighbor predicting
  - Network sensoring
  - Network surveillance
  - Using network reports to correct self-reporting bias

One Example for Relational Predictions

An and Schramski (2013) proposed two methods for correcting contested reports in exchange networks.

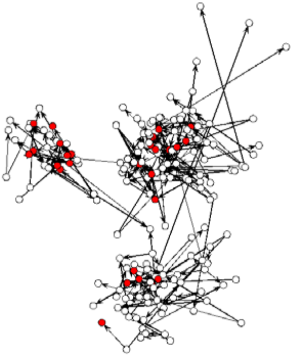
Figure 10. Four Exchange Networks



One Example for Behavioral Predictions

An and Doan (2013) proposed a network-based method to monitor health behaviors. They found that smokers, optimistic students, and popular students make better informants than their counterparts. Using three to four positive peer reports seem to uncover a good number of under-reported smokers while not producing excessive false positives.

Figure 11. A Smoking Detection Network



5. Network Interventions

- Change the context

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure
  - Physical segregation & relocation

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure
  - Physical segregation & relocation
  - Management. Mao's three strategies

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure
  - Physical segregation & relocation
  - Management. Mao's three strategies
- Change the process

5. Network Interventions

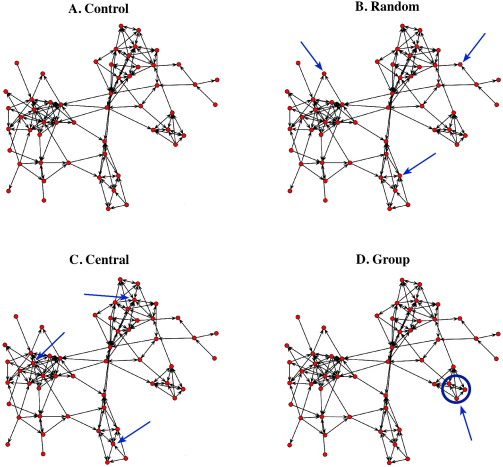
- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure
  - Physical segregation & relocation
  - Management. Mao's three strategies
- Change the process
  - Speeding up or halting diffusion

5. Network Interventions

- Change the context
  - How actors activate social ties to navigate through the uncertainties created by institutional reforms or leadership changes
  - How political and socioeconomic changes alter the culture of networking and the importance of network capital
- Change the structure
  - Physical segregation & relocation
  - Management. Mao's three strategies
- Change the process
  - Speeding up or halting diffusion
  - Synchronization

One Example

An (2011) assigned a smoking intervention to random, central students, and students with their best friends in selected classes, respectively.



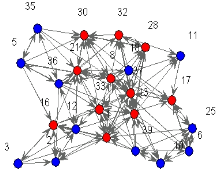
Uniqueness of This Study

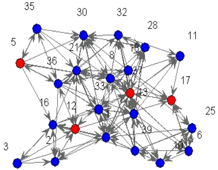
- Unlike previous interventions that assign intervention to all members in the treated groups, the partial treatment design assigns intervention to only partial members in the treated groups, which enables us to estimate several different kinds of causal peer effects.

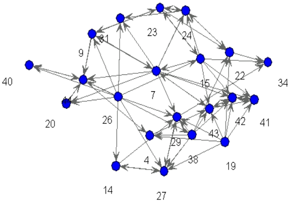
Uniqueness of This Study

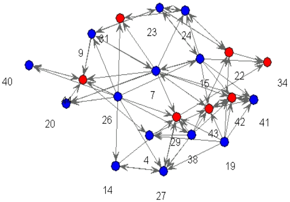
- Unlike previous interventions that assign intervention to all members in the treated groups, the partial treatment design assigns intervention to only partial members in the treated groups, which enables us to estimate several different kinds of causal peer effects.
- Unlike previous network interventions (e.g., Kelly et al. 1991; Latkin 1998; Campbell et al. 2008), this study includes a random intervention as an additional benchmark, which enables us to provide more proper evaluations of the effectiveness of network interventions.

Selecting Central Students

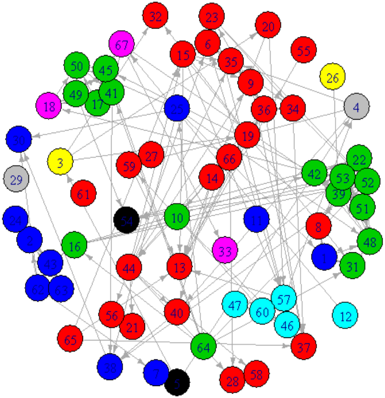


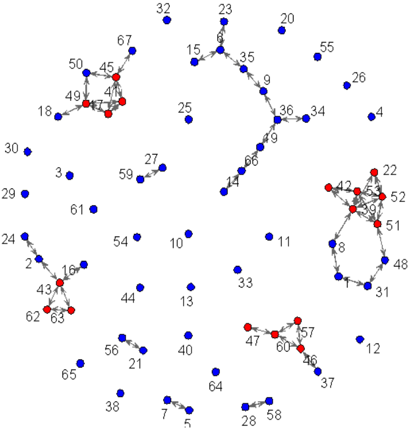




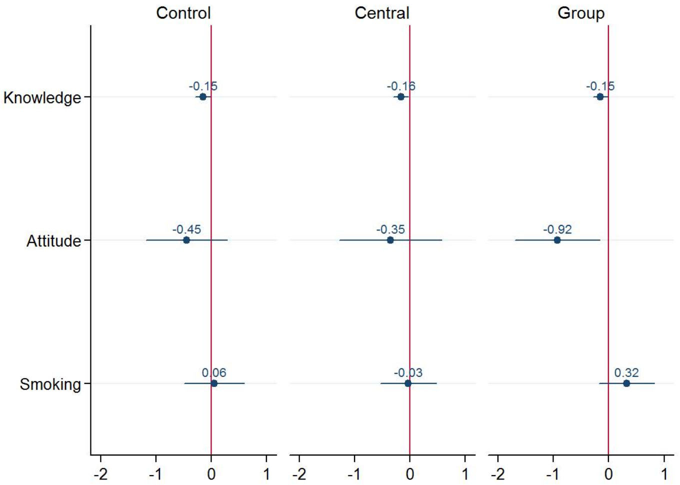


Selecting Student Groups



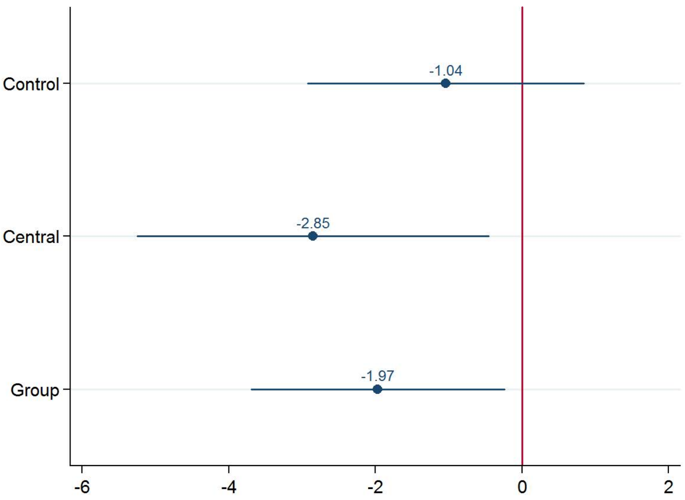


No Attididunal or Behavioral Effects



Also, no evidence for PEC, PEA, or PET.

Effects on Networks?!



Smokers are much more marginalized in the network interventions than in the random intervention.

Implications

1. The relative marginalization of smokers will restrict their influence on others, which may enable network interventions to outperform non-network interventions in the long run.

Implications

1. The relative marginalization of smokers will restrict their influence on others, which may enable network interventions to outperform non-network interventions in the long run.
2. The finding suggests that the strict separation between peer selection and peer influence as has been treated in the literature is inappropriate, because peer selection can act as a way to resist or exert peer influence.

Implications

1. The relative marginalization of smokers will restrict their influence on others, which may enable network interventions to outperform non-network interventions in the long run.
2. The finding suggests that the strict separation between peer selection and peer influence as has been treated in the literature is inappropriate, because peer selection can act as a way to resist or exert peer influence.
3. It also suggests that when evaluating interventions, we should put more attention to examining network outcomes, not just attitudinal or behavioral outcomes.

Books and Readings

1. Wasserman, Stanley and Katherine L. Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
2. Hanneman, Robert A. and Mark Riddle. 2005. *Introduction to Social Network Methods*. Riverside: University of California, Riverside (Available at `http://www.faculty.ucr.edu/~hanneman/nettext/`.
3. John Scott and Peter J. Carrington. 2011. *The SAGE Handbook of Social Network Analysis*. London: The Sage Publications.
4. Kolaczyk, Eric D. 2009. *Statistical Analysis of Network Data: Methods and Models*. New York: Springer.
5. Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Courses

- Title: Soc-S651: Topics in Quantitative Sociology: Social Network Analysis
- Instructor: Weihua An, Assistant Professor of Statistics and Sociology, weihuaan@indiana.edu
- Time: Thursdays 2:30PM - 5:00PM
- Location: Wells Library (LI) 851 (Subject to change)
- Description: This course covers the major approaches and methods to collect, represent, and analyze social network data. Students will learn hands-on skills to conduct their own network research using popular software such as UCINet and R.
- Prerequisites: This course requires a basic understanding of logistic regressions at the level of Statistics 503 or Sociology 650 (Categorical Data Analysis).
- A past syllabus can be found at `http://mypage.iu.edu/~weihuaan/Documents/Soc651_2012.pdf`.