Social Network Analysis: When Social Relationship is the Dependent Variable

Anabel Quan-Haase
Faculty of Information and Media Studies
Sociology

Overview of Presentation

- General overview of the social network approach
- Key terminology
- Uniqueness of collecting and analyzing social network data
- Recoding and importing data into STATA
- Example of analysis
- Understanding limitations and problems
- Conclusions

What is SNA?

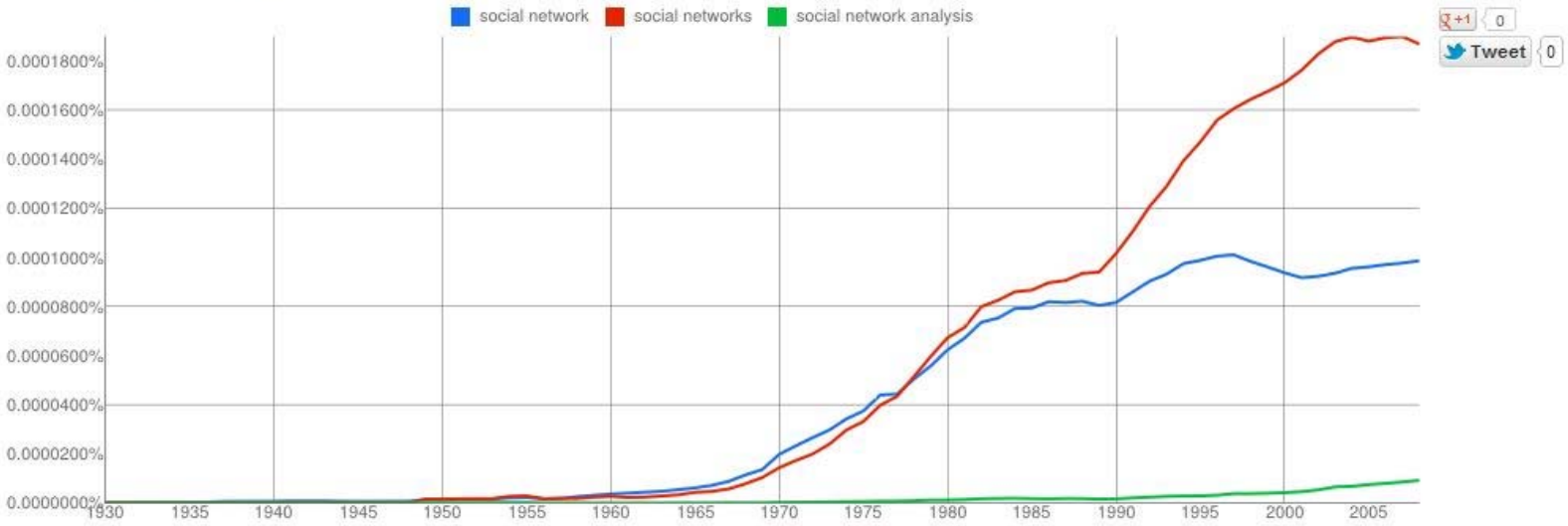
Social network analysis is focused on uncovering the patterning of people's interaction....Network analysts believe that how an individual lives depends in large part on how that individual is tied into the larger web of social connections. Many believe, moreover, that the success or failure of societies and organizations often depends on the patterning of their internal structure (Freeman, 1998, November 11).

Google books Ngram Viewer

Graph these ***case-sensitive*** *comma-separated phrases:* social network,social networks,social network analysis

between 1930 *and* 2008 *from the corpus* English *with smoothing of* 3 .

Search lots of books



Search in Google Books:

1930 - 1976	1977 - 1999	2000 - 2002	2003 - 2005	2006 - 2008	social network (English)
1930 - 1982	1983 - 2005	2006	2007	2008	social network analysis (English)
1930 - 1980	1981 - 2002	2003 - 2004	2005 - 2006	2007 - 2008	social networks (English)

Run your own experiment! Raw data is available for download here.

© 2010 Google - About Google - About Google Books - About Google Books NGram Viewer

What is Unique about SNA?

Social science research and theory tends to focus on social actors':

- attributes
- attitudes
- opinions
- behavior

Focus is on individual level of analysis, less on network-structural level.

a whole is not simply the sum of its parts

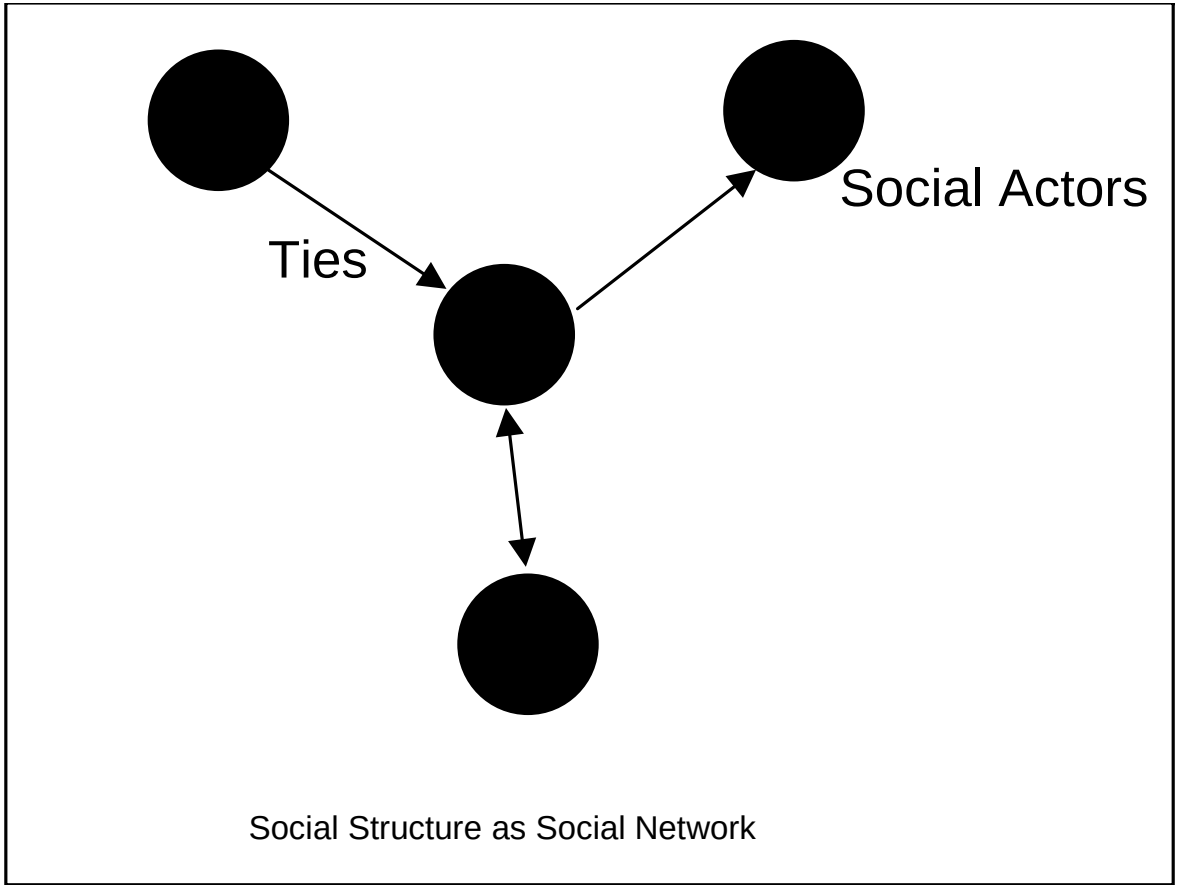
Key Terminology

- 1. Social structure
- 2. Social network
- 3. Social actors
- 4. Social ties or relations
- 5. Additional terms of relevance:
  - Nodes & edges
  - Directed graphs vs. undirected graphs
  - Ego
  - Alter
  - Homophily

1. Social Structure

- Sociological inquiry consists of understanding the constraining influence of social structure on social action
- BUT; how do we study social structure?

2. Social Network



Social Structure as Social Network

3. Social Actors

- The actors considered in a social network are exclusively social (alternatively referred to as agents, nodes, or social entities).
- These include individuals, organizations, institutions, nations, or groups (Wasserman & Faust, 1994).

Blurred Social Actors

- Social actors can therefore be distinguished from non-social actors – e.g., neurons comprising a neural network.
- On occasion, the distinction between a social and a non-social actor is not absolute. For example, computer networks represent a hybrid type of network.

One-mode vs. Two-mode

- Most social network analysis methods allow only one type of social actor (for instance, individuals or corporations) in their analysis; these are referred to as one-mode networks (Wasserman & Faust, 1994).
- However, methods exist which allow two different types of social actors in their analysis; these are referred to as two-mode networks. For instance, a study may simultaneously analyze corporations and their directors.
- Two-mode networks may also include social actors from distinct networks, for example, a network comprised of adults and a network comprised of children.
- Two-mode networks allow for comparison between different types and sets of social actors.

4. Ties

- Ties are links that connect social actors, and are the main focus of social network analysis. Ties are seen as “channels for transfer or “flow” of resources (either material or nonmaterial)” (Wasserman & Faust, 1994, p. 4).

Simple Ties

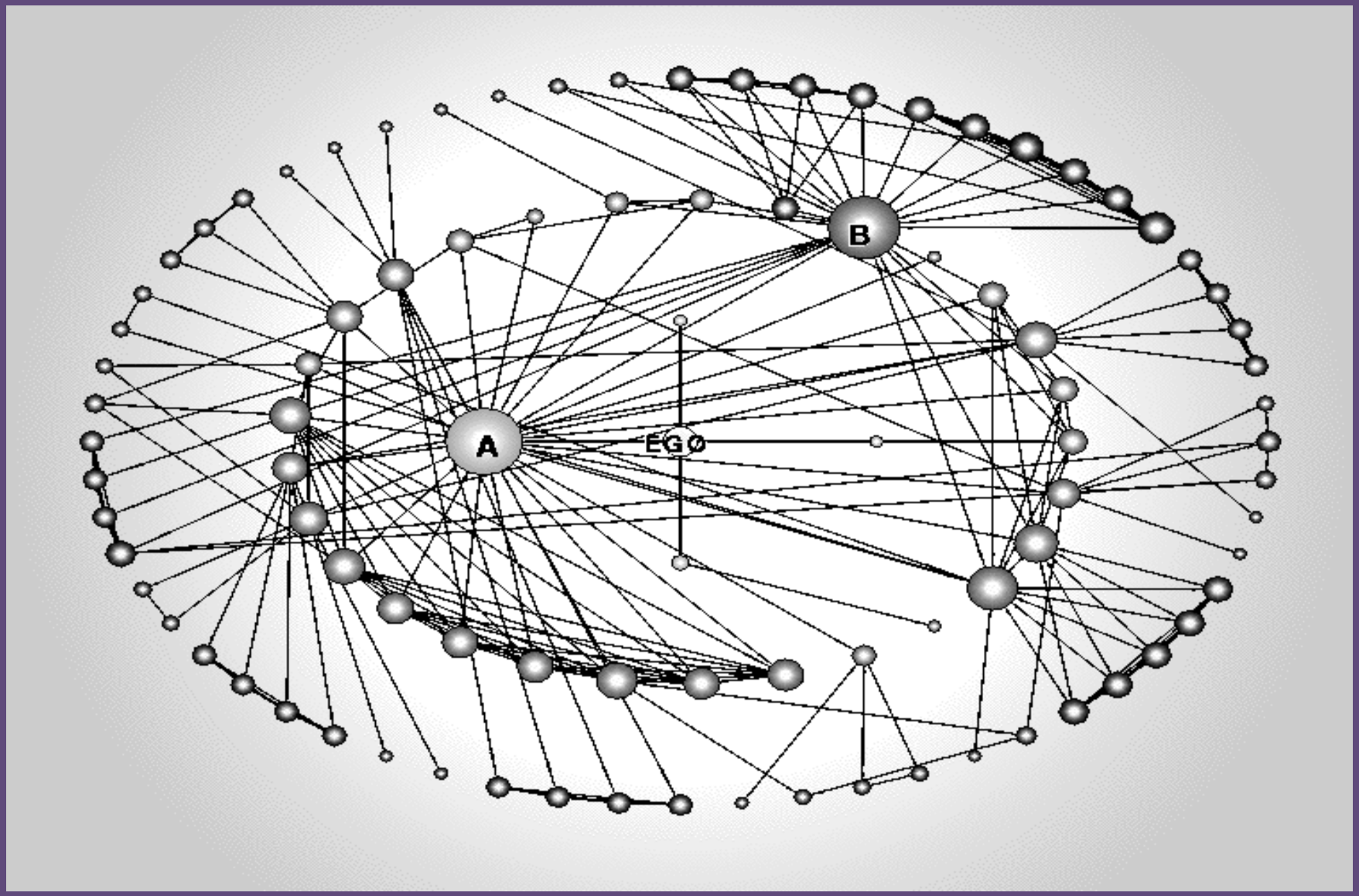
- Naturally occurring ties among social actors are inherently complex and consist of numerous different interaction activities.
- However, unlike ethnographers network analysts do not focus on the complexity of interactions among individuals (Burt, 1983).
- Instead, social network analysts focus more on the pattern of relations amongst individuals and to do so simplify the inherent complexity of social relationships by categorizing interactions into different broad types. The types can be manifold. For example, a pair of social actors may have friendship, working, cooperation, or citation ties.

5. Additional Terms

- Nodes & edges
- Directed graphs vs. undirected graphs
- Ego
- Alter
- Homophily

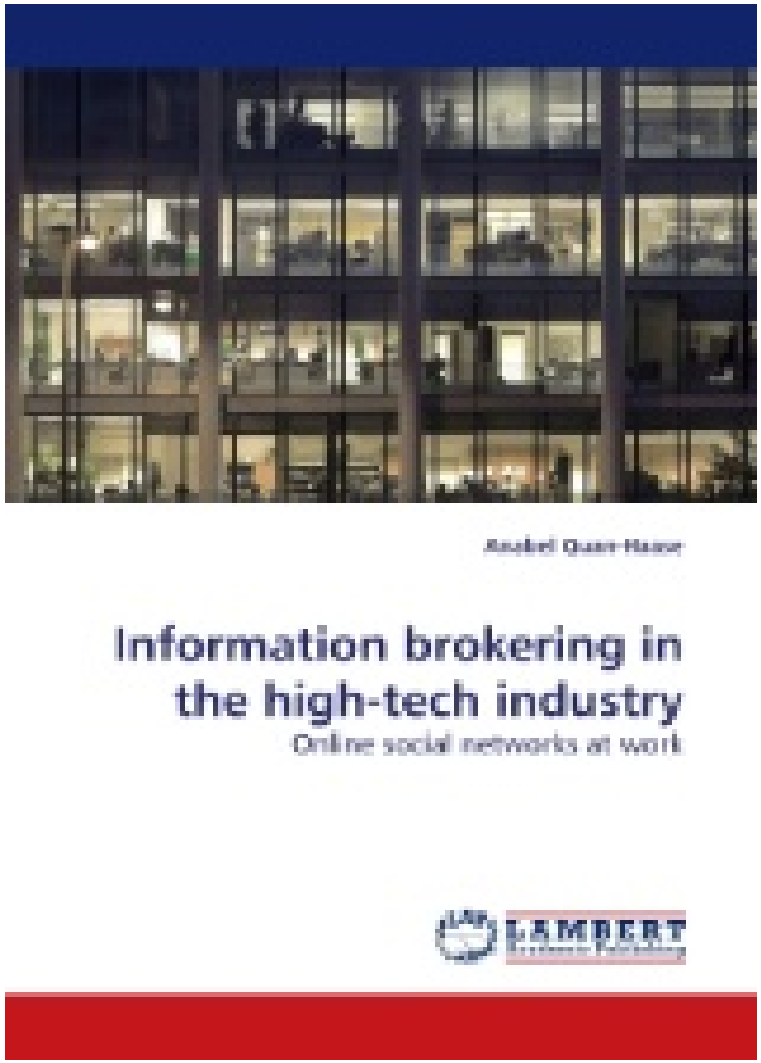
Types of Network Analysis

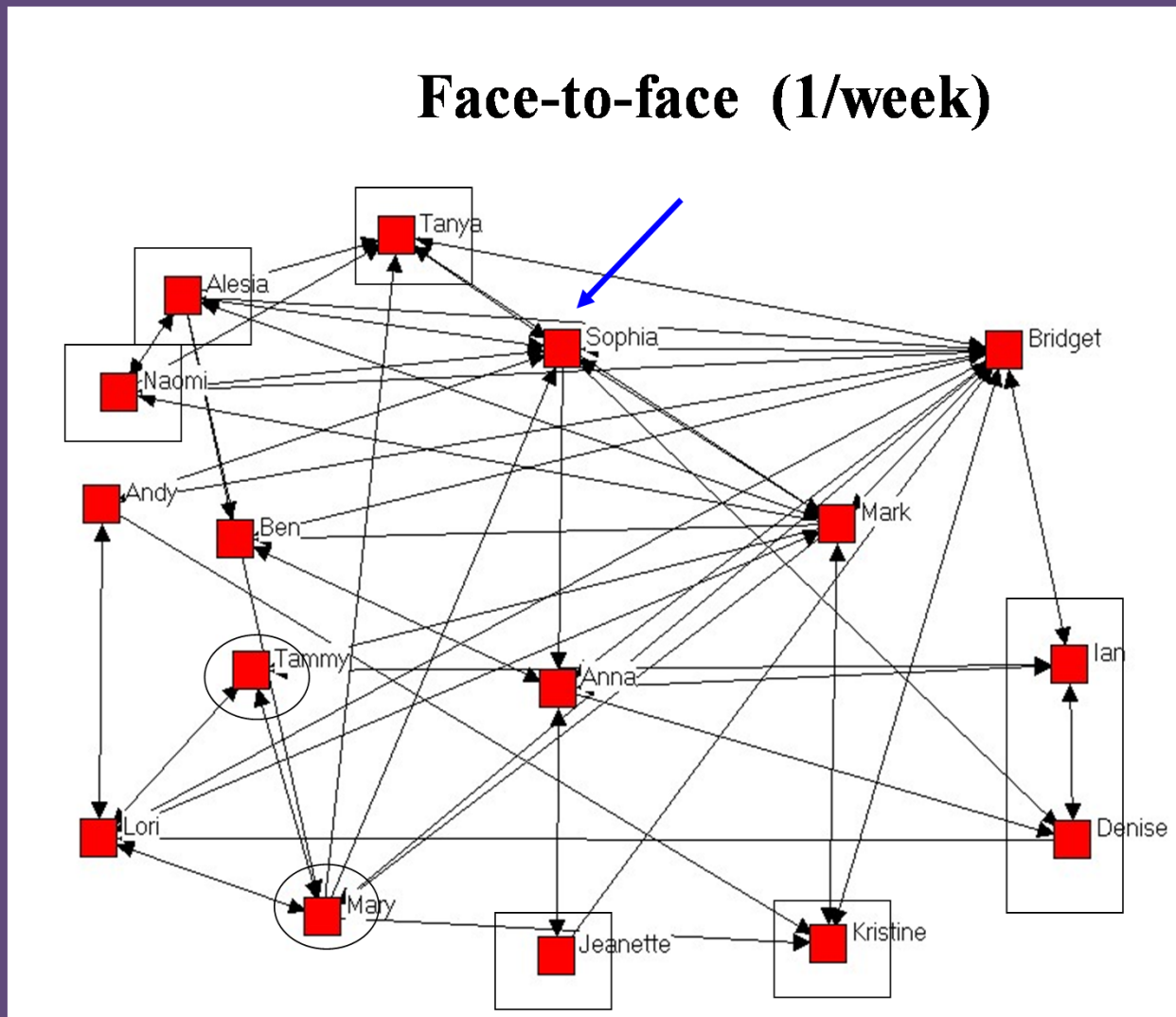
- *Ego-centered/Socio-centered Social Networks*
- *Community-centered social networks*



Ego-centered/Socio-centered Social Networks

Quan-Haase, A. (2009). Information Brokering in the High-Tech Industry: Online Social Networks at Work. Berlin: LAP Publishing.





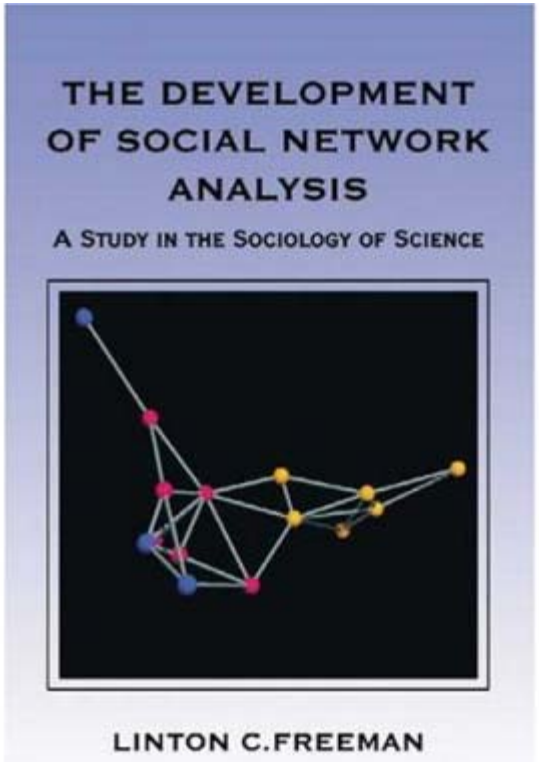
Community-Centered Social Networks

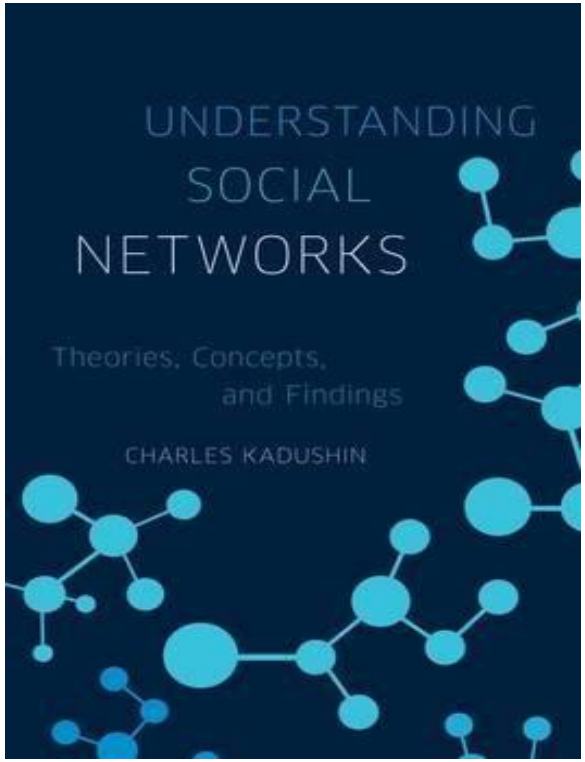
Uniqueness of Collecting and Analyzing Social Network Data

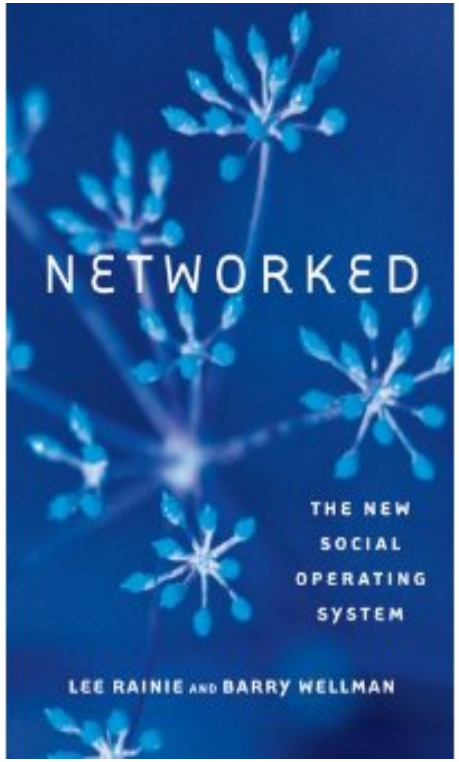
- Relational data
- Boundary specification and sampling
- Interdependence of data points
- Complexity of data collection
  - Behavioral
  - Self-report
- Key Sources
  - Marsden, P. V., & Campbell, K. E. (1984). Measuring tie strength. *Social Forces, 62(2), 482-501.*
  - Marsden, P. V. (1987). Core discussion networks of Americans. *American Sociological Review, 52, 122-131.*
  - McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in America: Changes in core discussion networks over two decades. *American Sociological Review, 71(3), 353-375.*
- Visualization
  - Sylva (CulturePlex)

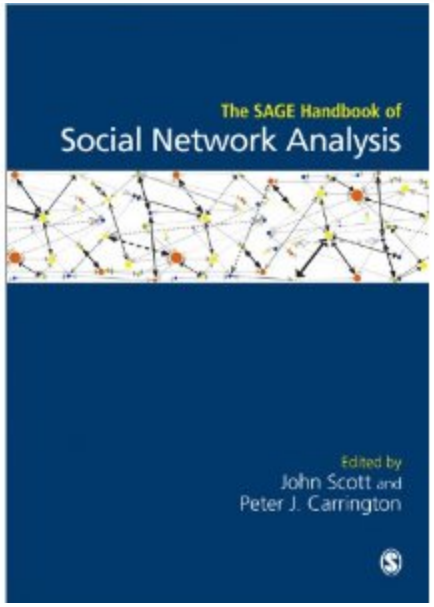
Internet Resources of Social Network Analysis

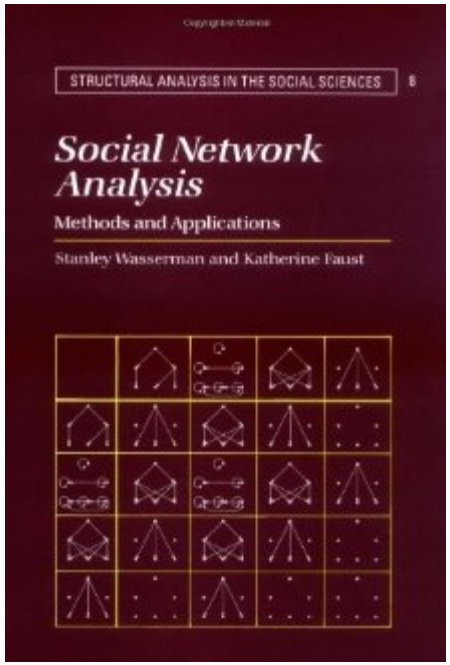
- **Center for the Study of Group Processes**
  http://lime.weeg.uiowa.edu/~grpproc/
- **INSNA International Network of Social Network Analysis**
  http://www.heinz.cmu.edu/project/INSNA/
- **Barry Wellman's Homepage**
  http://www.chass.utoronto.ca/~wellman/index.html
- **NodeXL**
  http://nodexl.codeplex.com/

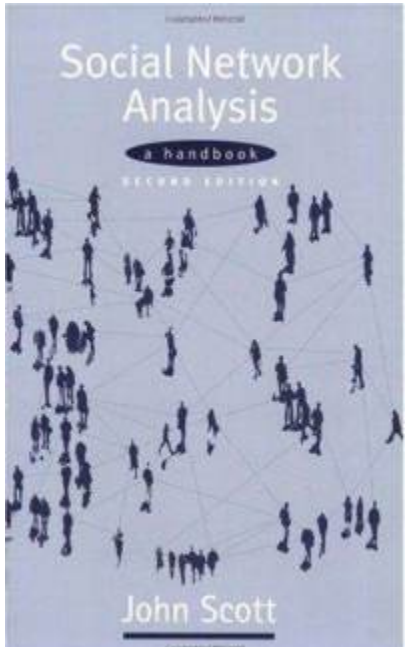

THE DEVELOPMENT OF SOCIAL NETWORK ANALYSIS
A STUDY IN THE SOCIOLOGY OF SCIENCE
LINTON C. FREEMAN

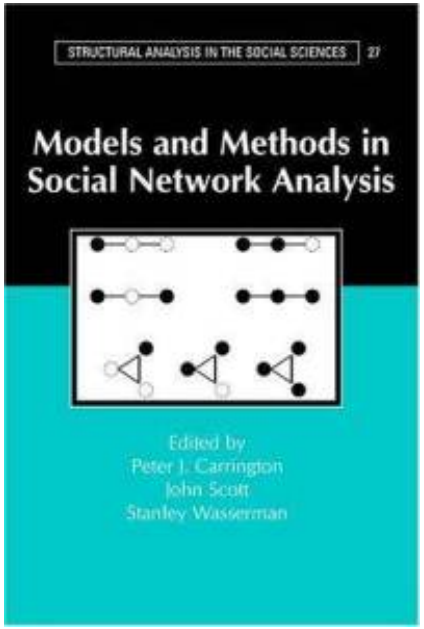

UNDERSTANDING SOCIAL NETWORKS
Theories, Concepts, and Findings
CHARLES KADUSHIN


NETWORKED
THE NEW SOCIAL OPERATING SYSTEM
LEE RAINIE AND BARRY WELLMAN


The SAGE Handbook of
Social Network Analysis
Edited by
John Scott and
Peter J. Carrington


STRUCTURAL ANALYSIS IN THE SOCIAL SCIENCES 8
Social Network Analysis
Methods and Applications
Stanley Wasserman and Katherine Faust


Social Network Analysis
a handbook
John Scott


STRUCTURAL ANALYSIS IN THE SOCIAL SCIENCES 27
Models and Methods in Social Network Analysis
Edited by
Peter J. Carrington
John Scott
Stanley Wasserman

Example

- Mesch, Gustavo, Talmud, Ilan, & Quan-Haase, A. (forthcoming 2012). **Instant messaging social networks: Individual, relational, and cultural characteristics.** *Journal of Social and Personal Relationships.*

- Collaborators: Department of Sociology and Anthropology, University of Haifa, Israel.

Hypotheses

- *H1*: Perceived tie closeness is expected to be positively associated with the frequency of communication via IM and the multiplexity of communication topics.
- *H2:* Same-sex relationships are expected to communicate more frequently via IM and on a larger number of topics than mixed-sex relationships.

Miller-McPherson, J., Smith-Lovin, L. & Cook, J.M. (2002). Birds of a feather: Homophile in social networks. *Annual Review of Sociology, 27*, 415-444.

Methods

- The two dependent variables examined in this study are dyadic, with the unit of observation being relational.
- We defined *i* as ego, the focal individual in the study, and *j* as the network contact listed in the name generator: *j* = 1 or 2.

The Name Generator

- An efficient means to study ego-networks.
- Provides quick insight into ties.
- Allows to study several ties at one time in point.

Name Generator Approach

The next couple of questions examine your relationship with instant messaging contacts.

Consider all your communication instances of the past **WEEK** that took place via instant messaging. Considering every single day of the week, please write down the names of **TWO** people who you communicated with most frequently. This list can include friends on or off campus, acquaintances, professors and family. If you communicated with strangers, you can also write them down. All these names will remain confidential.

(Modified question from Burt, 1992)

1. _____
2. _____
….

Frequency of Communication

- The first dependent variable measured the frequency of communication via IM for each pair (*i, j*) on an eight-point scale from 1 = “never” to 8 = “several times a day”.

Topic Multiplexity

- The second dependent variable, IM topic multiplexity, measured the number of topics discussed by each pair (*i*, *j*), ranging from 0 to 11 topics. The measure of multiplexity was developed and tested in previous studies (Marsden & Campbell, 1984; Mesch & Talmud, 2006).
- Respondents were provided with a list of 11 topics of conversation from which to choose, such as school, family, and personal. Positive responses were coded 1, and a scale was created based on the sum of all the topics discussed among each pair *(i,j)*.

Independent Variables: Age and Sex of Alter

In this section, a series of questions will be asked about each of the people you have listed above. Please indicate which response applies best.

What is the approximate age of each person listed?

- Less than 10 years
- 10-14 years
- 15-18 years
- 18-25 years
- 25-30 years
- More than 30 years

What is the sex of each person listed?

- Female
- Male

Relational Variables: Closeness & Distance

Closeness

How **close** do you feel to each person listed?

- Very close
- Somewhat close
- Neutral
- Somewhat distant
- Distant

Distance

Where is each person listed located?

__ On campus

__ Off campus

- __ In London
- __ Within 3 hours driving distance
- __ Within 6 hours driving distance
- __ Within Canada, but outside Ontario
- __ Within North America, but outside Canada
- __ Outside North America

Are We More Connected?

- A key question in the literature evolves around our levels of connectivity and the role of technology.
- Are we less connected, or are these connections of lesser quality?

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in America: Changes in core discussion networks over two decades. *American Sociological Review, 71(3), 353-375. Wang, H., &*

Wang, H. & Wellman, B. (2010). Social connectivity in America: Changes in adult friendship network size from 2002 to 2007. *American Behavioral Scientist, 53*(8), 1148-1169.

Topic Multiplexity

Which topics do you discuss with each of the 6 friends in the IM list?

- Classes and/or professors in the university
- About relationship with parents
- About mutual friends
- Sports and movies
- About family problems
- About personal problems
- Favorite television programs
- Romantic relationships
- Music
- Religion/Spirituality
- Fashion and clothes
- Food and Diet
- Politics
- Places to go out
- Computers and technology

Recoding and importing data into STATA

Excel ➡ • 1. Reconfiguring data from standard social science data into relational data.

SPSS ➡ • 2. Recoding homophily variables.

STATA ➡ • 3. GLMS analysis.

Standard Data=>Relational Data

Ego=*i*

case	age_ego	sex_ego	age_alter1	age_alter2	sex_alter1	sex_alter2	closeness_alter1	closeness_alter2
1	26	1	22	30	0	1	4	3
2	19	0	20	23	1	1	1	5
3	23	1	21	17	1	0	4	2

	case	age_ego	age_alter	sex_ego	sex_alter	closeness	distance	topic multiplexity
(1,1)	1a	26	22	1	0	4	0	8
(1,2)	1b	26	30	1	1	3	1	3
(2,1)	2a	19	20	0	1	1	1	9
(2,2)	2b	19	23	0	1	5	1	6
(3,1)	3a	23	21	1	1	4	1	5
(3,2)	3b	23	17	1	0	2	0	4

Homophily

Ego=*i* Alter=*i, j*

	case	age_ego	age_alter	sex_ego	sex_alter	closeness	distance	topic multiplexity
(1)	1a	26	22	1	0	4	0	8
(1)	1b	26	30	1	1	3	1	3
(2)	2a	19	20	0	1	1	1	9
(2)	2b	19	23	0	1	5	1	6
(3)	3a	23	21	1	1	4	1	5
(3)	3b	23	17	1	0	2	0	4

Calculate new homophily variable for sex:
If ego and alter same sex than equal 1
If ego and alter different sex than equal 0

Baumgarte, R., & Nelson, D.W. (2009). Preference for same versus cross sex friendships. *Journal of Applied Social Psychology, 39,* 901-917.

Independence of Observations

- Given that the dyadic relations between contact *i* and contact *j* are not independent observations from the relations between contacts *i* and *k*, we utilized a clustering procedure with Robust Standard Error Estimates (RSE) to control for the dependence of the observations. In this way, we analyzed both contacts in a single model, controlling for joint variance resulting from the clustered contacts (Cameron & Trivedi, 2010; Mok, Wellman, & Carrasco, 2010; Wellman & Frank, 2001).

Clustering Variable

Ego=*i*

	case	age_ego	age_alter	sex_ego	sex_alter	closeness	distance	topic multiplexity
(1)	1a	26	22	1	0	4	0	8
(1)	1b	26	30	1	1	3	1	3
(2)	2a	19	20	0	1	1	1	9
(2)	2b	19	23	0	1	5	1	6
(3)	3a	23	21	1	1	4	1	5
(3)	3b	23	17	1	0	2	0	4

Analysis in STATA

This approach allows us to control for the interdependence of cases.

glm DV IV1 IV2,
family(gaussian) link(identity) vce(cluster id)

Relationship Type

Table 1. Relationship type as a percentage of IM contacts for Israel and Canada

	Israel		Canada	
Relationship type	First IM contact	Second IM contact	First IM contact	Second IM contact
Family	11%	10%	16%	20%
Close friend	48%	44%	42%	49%
Distant friend	23%	35%	12%	16%
Romantic partner	14%	5%	21%	6%
Online tie	4%	9%	1%	1%

Note. Israel $N = 492$; Canada $N = 293$

Table 2. GLM results for the final models predicting IM topic multiplexity for Israel and Canada

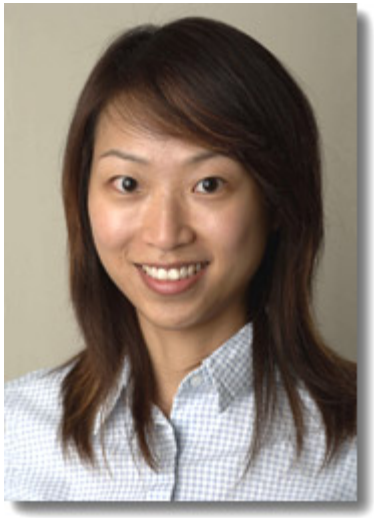
	Israel			Canada		
Predictors	*β*	RSE	*z*	*β*	RSE	*z*
Age	-.09*	.04	-2.25	-.05	.10	-.89
Gender (1=male)	-.11*	.25	-2.44	-.08	.42	-1.8
Relationship status (1=single)	.0	.25	.09	.01	.37	.21
Gender similarity (1=same)	-.09*	.21	-2.22	.05	.34	1.08
Propinquity (1=nearby)	.03	.22	.82	-.03	.33	-.74
Hours of daily IM use	.15**	.15	2.72	.14**	.24	2.65
Number of IM regular contacts	.12*	.14	2.47	.09*	.17	2.09
Online relationship	-.08	.68	-1.74	-.04	.71	-1.77
Family	-.18***	.35	-4.64	-.21***	.48	-4.31
Distant friend	-.14**	.27	-3.26	-.09	.52	-1.75
Partner	.01	.39	0.14	.19***	.57	3.80
Close friend	------	-----	------			
Duration of relationship	.04	.11	.92	.13*	.15	2.51
Closeness	.24***	.17	5.60	.22***	.29	4.78
Constant	2.0	1.23	1.62	6.33*	2.48	2.55
Adjusted R Square	.23***			.23***		

Note. *** $p < .001$, ** $p < .01$, * $p < .05$, two-tailed.

Israel $F(13, 633) = 16.17$; Canada $F(13, 513) = 13.81$.

Israel $N = 647$; Canada $N = 526$

Other Approaches



- Mok, D., Wellman, B., & Carrasco, J.A. (2010). Does distance still matter in the age of the Internet? *Urban Studies*, *47*(13), 2747-2783.
- Wellman, B., Frank, O., Espinoza, V., Lundquist, S., & Wilson, C. (1991). Integrating individual, relational and structural analysis. *Social Networks, 13,* 223-250.
- Wellman, B., & Frank, K. (2001). Network capital in a multi-level world: Getting support from personal communities. In N. Lin, K. Cook & R. Burt (Eds.), Social capital: Theory and research (pp. 233-273). Chicago: Aldine DeGruyter.
- Contractor, N. S., Wasserman, S., & Faust, K. (1999). Testing multi-level, multi-theoretical hypotheses about networks in 21st century organizational forms: An analytic framework and empirical example. Retrieved from http://www.spcomm.uiuc.edu/contractor/pstarpaper.html

Limitations of Social Network Analysis

- Boundary specification
- Data source
- Definition of social actors
- No distinct method

Conclusions

- Powerful for research questions that focus on the relationship and the flow of resources.
- Statistical software does not yet accommodate all forms of social network analysis.
- A number of considerations are necessary:
  - boundary specification;
  - interdependence;
  - specifications of other statistical methods.

Thank you!

“With enough effort and perseverance: Anything is possible”