

Introduction to Ego Network Analysis

Rich DeJordy



Dan Halgin

Boston College
and the
Winston Center for Leadership & Ethics

Goals for Today

1. Introduce the network perspective
 - How is ego-centric analysis different from socio-centric analysis?
 - When and why ego network analysis?
 - What theories are ego-centric?
2. Research design and data collection
3. Data analysis
4. Review and demo of software tools
 - Egonet, E-Net

What is Unique about Social Network Analysis?

- Phenomenon studied
 - Distinctive type of data,
 - It's about relations & structure
- How we study it
 - Distinctive tool
 - Typical statistical methods may not apply
- How we understand it
 - One “network perspective”
 - Based on multiple theories (Simmel, Blau)

Mainstream Logical Data Structure

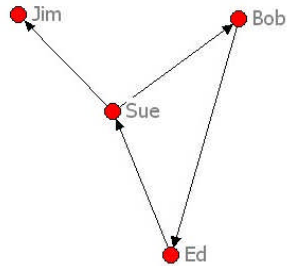
- 2-mode rectangular matrix in which rows (cases) are entities or objects and columns (variables) are attributes of the cases
- Analysis consists of correlating columns
 - Emphasis on explaining one variable

ID	Age	Education	Salary
1			
2			
3			
4			

Network Logical Data Structures

Friendship

	Ed	Sue	Jim	Bob
Ed	-	1	0	0
Sue	0	-	1	1
Jim	0	0	-	0
Bob	1	0	0	-



Email Communication

	Ed	Sue	Jim	Bob
Ed	-	4	0	2
Sue	0	-	5	1
Jim	0	0	-	0
Bob	3	0	4	-

- Individual characteristics only half the story...**RELATIONS MATTER!**
- People influence each other, ideas & material flow
- Values are assigned to pairs of actors
- Hypotheses can be phrased in terms of correlations between relations

Relational Data & Attribute Data

	Ed	Sue	Jim	Bob
Ed	-	1	0	0
Sue	0	-	1	1
Jim	0	0	-	0
Bob	1	0	0	-

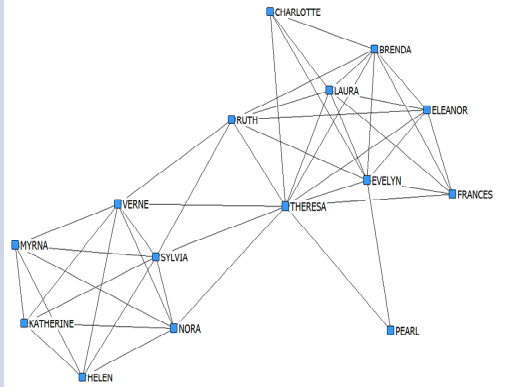
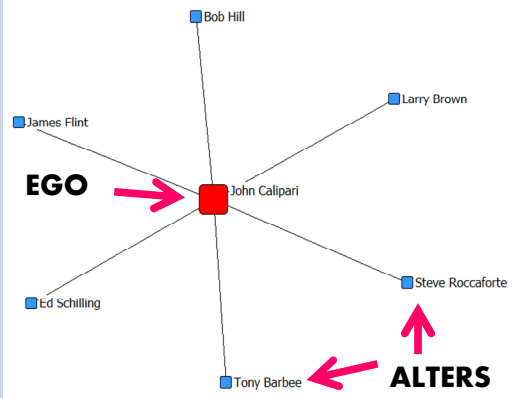
Relational Data

	Gender	Education	Salary
Ed	0	14	50000
Sue	1	15	99000
Jim	0	12	65000
Bob	0	8	15000

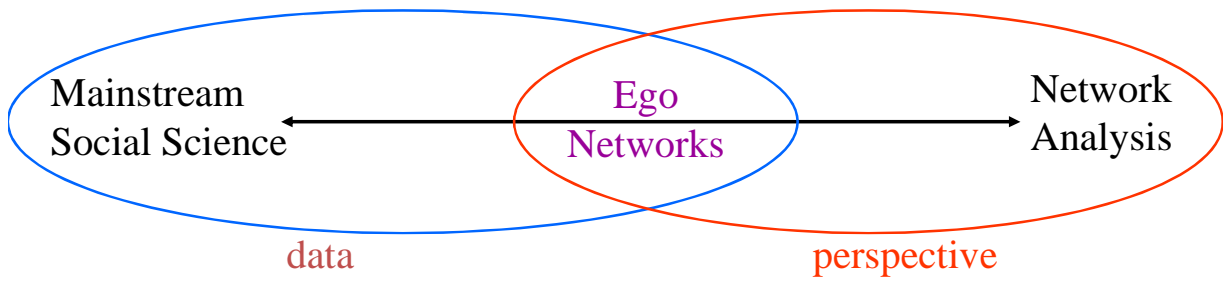
Attribute Data

SNA provides the ability to combine relational data with attribute data (e.g., homophily, heterogeneity, etc)

Introduction to Ego Network Analysis

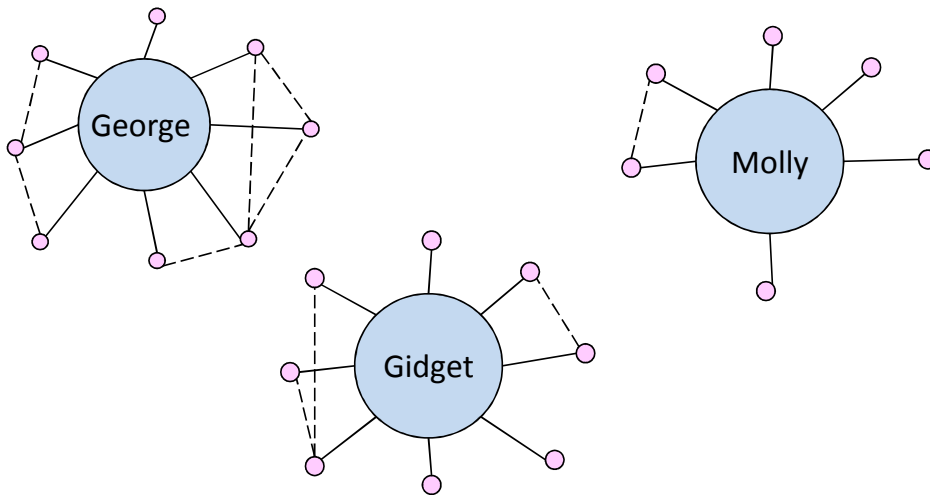
Socio-centric (Whole/ Complete network)	Ego-centric (Ego/Personal network)
	
<ul style="list-style-type: none">• Focus on the whole group<ul style="list-style-type: none">○ Global structure• Patterns of interaction used to explain:<ul style="list-style-type: none">○ Concentration of power○ Flow of information or resources○ Status structures• Cases are complete networks<ul style="list-style-type: none">○ Generalized to other networks	<ul style="list-style-type: none">• Focus on individual ego networks<ul style="list-style-type: none">○ Structure○ Composition○ Shape• Cases are individual ego networks<ul style="list-style-type: none">○ Generalized to other ego networks

Ego Network Analysis



- Combines the perspective of network analysis with the data of mainstream social science

Each Ego Network is Treated as its Own World



Or in more typical language, each ego network is treated as a separate case

Why Study Ego Networks?

Ego's network is a source of:

- Information
- Social support
- Access to resources
- Sense-making
- Normative pressures
- Influence
- etc.

All of which can influence Ego's behavior

When to use Ego Network Analysis

- If your research question is about phenomena of or affecting individual entities across different settings (networks) use the ego-centric approach
 - Individual people, organizations, nations, etc.
- If your research question is about different patterns of interaction within defined groups (networks), use the socio-centric approach
 - E.g., who are the key players in a group? How do ideas diffuse through a group?

Which Theories are Ego-centric?

- Most theories under the rubric of social capital are ego-centric
- Topological
 - Structural holes / Brokerage
 - Embeddedness
- Compositional
 - Size
 - Alter attributes

Steps to a SNA study

1. Identify the population
 - Sampling, gaining access

2. Determine the data sources
 - Surveys, interviews, observations, archival

3. Collect the data
 - Instrument design

Step 1. Identify the Population

- Sampling Criteria
 - Determined by research question
 - High tech entrepreneurs
 - Alumni of defunct organizations
 - Basketball coaches
 - First time mothers returning to the workforce
 - Baseball Hall of Fame inductees
 - Contingent workers
 - People with invisible stigmatized identities

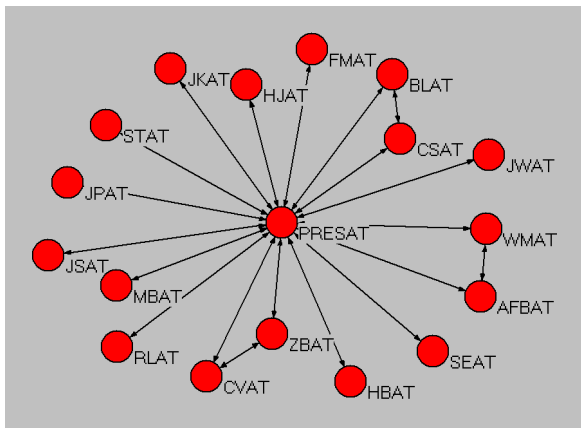
Step 1. Identify the Population

- Gaining Access
 - Same concerns as other research
 - It depends on the sensitivity of the questions that you are asking
 - Length of interview can be daunting
 - Depends on the number of alters

Step 2: Determine Data Sources

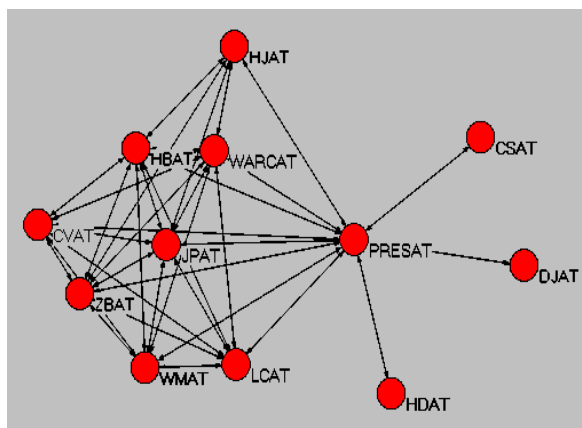
- Surveys
- Interviews
- Observations
- Archival data

White House Diary Data, Carter Presidency



Year 1

Data courtesy of Michael Link



Year 4

Step 3: Collect the Data

- What data should you collect?
 - What questions need to be answered?
- How to format your data collection instrument (e.g., a survey, spreadsheet, database, etc.)?

What Questions to Ask?

- **IT DEPENDS!!!**
 - Ego’s relations to alters form variables. Size of ego’s social support network is to ego network analysis what “attitude toward gun-control” is to traditional case based research.
- It is the researcher who defines the relations of interest. What’s relevant for the phenomena in question?
 - What influences an employee’s turn-over intention?
 - What influences one’s likelihood of adoption of a new technology?

How to ask: Tick or Rate?

- Record yes/no decisions or quantitative assessment?
 - Yes/no are cognitively easier to determine (therefore reliable, believable)
 - Yes/no **much** faster to administer
 - But yes/no provides no discrimination among levels
- One quantitative rating can replace a series of binaries
 - “How often do you see each person?”
 - 1 = once a year; 2 = once a month; 3 = once a week; etc.
 - Instead of three questions:
 - Who do you see at least once a year?
 - Who do you see at least once a month?
 - Who do you see at least once a week?
 - However, if categories are too similar it may be difficult to differentiate

Question Wording Issues

- “Friendship” does not mean the same thing to everyone
 - Especially across national cultures
- Some helpful practices
 - Use one word label plus two or three sentence description, plus have full paragraph detailed explanation available
 - Use homogeneous samples

Ethnographic Sandwich

- Ethnography at front end helps to ...
 - Select the right questions to ask
 - Word the questions appropriately
 - Create enough trust to get the questions answered

- Ethnography at the back end helps to ...
 - Interpret the results
 - Can sometimes use respondents as collaborators

Instrument Design: Paper or Plastic?

- Paper medium
 - Reliable
 - Reassuring to respondents
 - Errors in data entry
 - Data entry is time-consuming
- Electronic
 - Span distances, time zones
 - Harder to lose
 - Fewer data handling errors
 - Lower response rate
 - Emailed documents vs. survey instruments

Data Collection in an Ego-centric Study

1. Attributes about Ego
 2. Name generator
 - Obtain a list of alters
 3. Name interpreter
 - Assess ego's relationships with generated list of alters?
 4. Alter Attributes
 - Collect data on the list of alters
 5. Alter – Alter Relationships
 - Determine whether the listed alters are connected
-
-
-
-
-
-
-
-
-
-
-

Attributes about Ego

- Typical variables for case based analysis
 - Age
 - Gender
 - Education
 - Profession
 - SES
 - Etc.

Sample Name Interpreter

- Questions that deal with ego's relationship with [or perception of] each alter
 - How close are you with <alter>?
 - How frequently do you interact with <alter>?
 - How long have you known <alter>?
- All of these questions will be asked for each alter named in the previous section

Sample Alter Attribute Questions

- As far as you know, what is <alter>' s highest level of education? (Adapted from Burt, 1984)
 - Age, occupation, race, gender, nationality, salary, drug use habits, etc

- Some approaches do not distinguish between name interpreters and alter attribute questions

Sample Alter-Alter Relationship Questions

- Think about the relationship between <alter1> and <alter2>. Would you say that they are strangers, just friends, or especially close?
(Adapted from Burt, 1998)
- Note: this question is asked for each unique alter-alter pair. E.g., if there are 20 alters, there are 190 alter-alter relationship questions!
 - Typically, we only ask one alter-alter relationship question

Why Ego-Centric Analysis

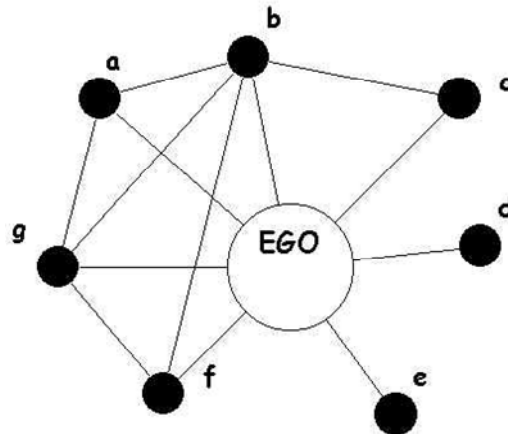
- Asks different questions than whole network analysis.
- In fact, many of the various approaches to “Social Capital” lend themselves particularly to the analysis of Ego-Centric or Personal networks

Kinds of Analyses

- In Ego-Centric Network analyses we are typically looking to use network-derived measures as variables in more traditional case-based analyses
 - E.g., instead of just age, education, and family SES to predict earning potential, we might also include heterogeneity of network or brokerage statistics
 - Many different kinds of network measures, the simplest is degree (size)

Size

- Degree = 7



Access to social support, resources, information

Composition: Content

- The attributes (resources) of others to whom I am connected affect my success or opportunities
 - Access to resources or information
 - Probability of exposure to/experience with
- Paris Hilton..Why is she a celebrity?



Composition: Similarity Between Ego & Alter

- Homophily
 - We may posit that a relationship exists between some phenomenon and whether or not ego and alters in a network share an attribute
 - Selection
 - Teens who smoke tend to choose friends who also smoke
 - Influence
 - Overtime, having a network dominated by people with particular views may lead to one taking on those views

Composition: Homophily

- A CFO who surrounds herself with all finance people
- A Politician who surrounds himself with all members of the same political party

Composition: Dissimilarity Between Ego & Alter

- Heterophily
 - We may posit that a relationship exists between some phenomenon and a difference between ego and alters along some attribute
 - Mentoring tends to be heterophilous with age

Composition: Homophily/Heterophily

- Krackhardt and Stern's E-I index

$$\frac{E - I}{E + I}$$

- E is number of ties to members in different groups (external), I is number of ties to members of same group (internal)
- Varies between -1 (homophily) and +1 (heterophily)

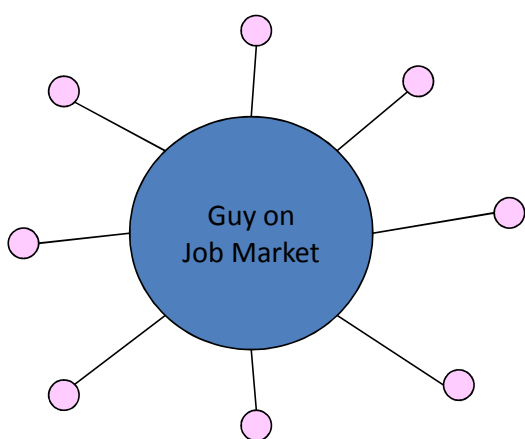
Composition: Heterogeneity

- Similar to homophily, but distinct in that it looks not at similarity to ego, but just among the alters
- Diversity on some attribute may provide access to different information, opinions, opportunities, etc.
 - My views about social welfare may be affected by the diversity in SES present in my personal network (irrespective of or in addition to my own SES)
- Blau's Heterogeneity Index

Structural Analyses

- Burt’s work is particularly and explicitly ego-network based in calculation
 - My opportunities are affected by the connections that exist (or are absent) between those to whom I am connected

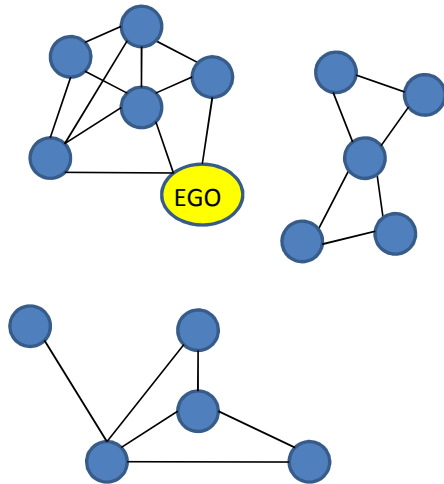
Structural Holes



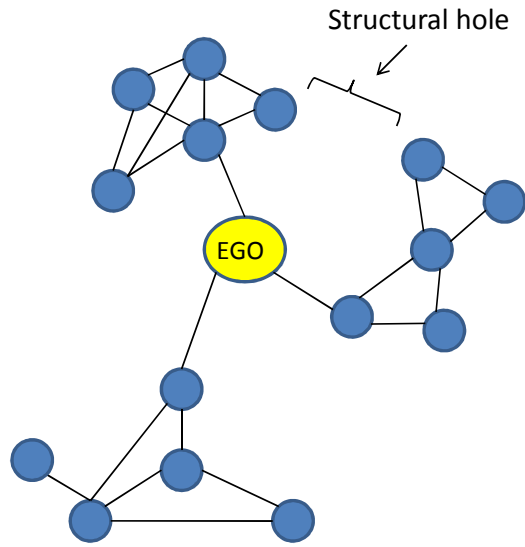
- Basic idea: Lack of ties among alters may benefit ego
- Benefits
 - Autonomy
 - Control
 - Information

Introduction to Ego Network Analysis

FEW STRUCTURAL HOLES

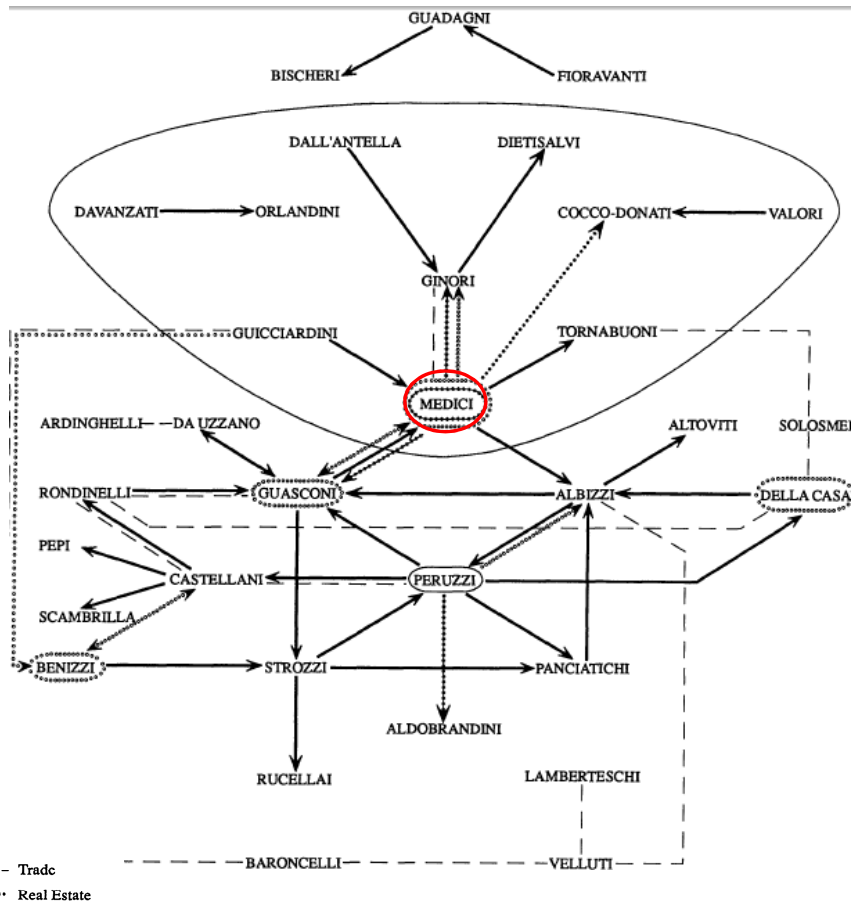


MANY STRUCTURAL HOLES



Structural Holes provide Ego with access to novel information, power, freedom

Power



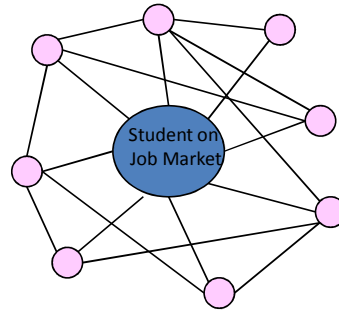
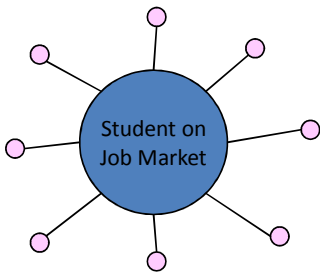
(Padgett & Ansell, 1993)

Types of Ties:

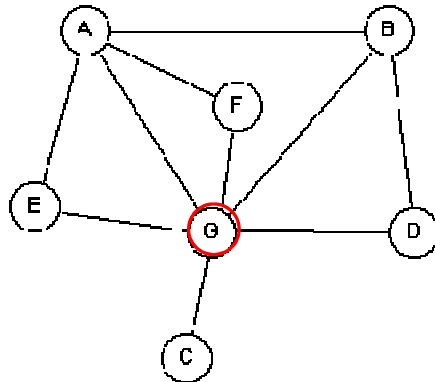
- ♀ → ♂ Marriage
- Partnership
- - - - Bank Employment
- - - - Trade
- Real Estate

Burt's Measures of Structural Holes

- Effective size
- Efficiency
- Constraint



Effective Size



Adapted from Burt (1995:56)

Node "G" is EGO	A	B	C	D	E	F	Total
Redundancy with EGO's other Alters:	3/6	2/6	0/6	1/6	1/6	1/6	1.33

Effective Size of G = Number of G's Alters – Sum of Redundancy of G's alters
 = 6 – 1.33 = 4.67

Efficiency

- Efficiency = (Effective Size) / (Actual Size)

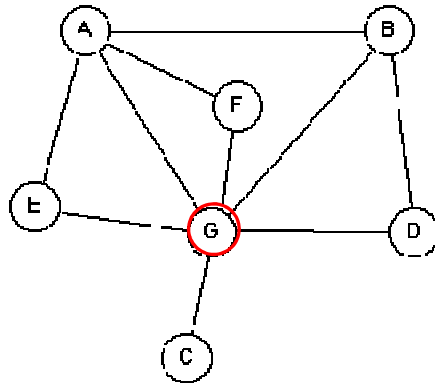


Figure 1. Adapted from Burt (1992:56)

Actual Size = 6

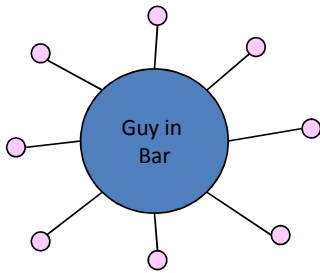
Effective Size of G = 4.67

Efficiency = $4.67/6 = \sim 0.78$

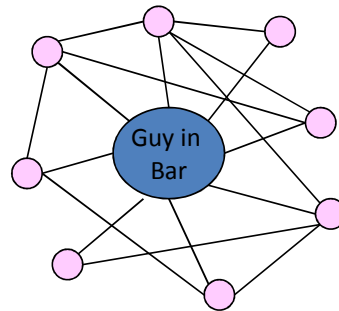
Constraint: The Basic Idea

- Constraint is a summary measure that taps the extent to which ego's connections are to others who are connected to one another.
- If ego's boyfriend bowls with her brother and father every Wednesday night, she may be constrained in terms of distancing herself from him, even if they break up.
- There's a normative bias in much of the literature that less constraint is good

Constraint



No constraint



More constraint

Ego-Centric Network Analysis

- When conducted across many, independent egos, presents different problems
- Many Social Network Analysis tools ill suited to the nature of such analyses
 - Really designed for “whole network” analysis
- Ego Network analyses require either:
 - joining into one large, sparse, blocked network, or
 - repetition of analysis of individual networks
 - Can be tedious if there’s no facility for batching them

Statistical Analyses

- Because Ego-Networks more readily meet the requirements of OLS models based on inferential statistics, the final analyses can be done using statistical packages like SPSS, Stata, SAS, etc.
- But getting the data into an appropriate format is complicated, and generating network statistics is cumbersome for even simple measures, and structural measures require extensions or very complex algorithms
 - Barry Wellman has a stream of articles on how to do compositional analysis on ego-centric networks in SAS (1985, 1992) and SPSS (Müller, Wellman, & Marin, 1999)
- Some tools (Egonet & E-Net) facilitate the process of performing these analyses and getting the data to statistical packages

Using the Programs

- Egonet
- E-Net

Egonet

- Tool available for free from www.sourceforge.net/projects/egonet
 - Written in java, runs on any java-enabled platform
- Tool facilitates design, collection, and analysis of ego-centric network data
 - Exports to other packages

Egonet: What it does

- Allows researcher to build and administer ego-net interview/survey questions
- Collects and summarizes data from respondents
- Allows for calculation of summary network metrics across all cases
- Visualization

E-NET

- Tool available for free from www.analytictech.com
 - Reads data in UCINET & ego-VNA format
 - Also reads EXCEL “column-wise” data
 - Runs on Windows/Intel platforms
- Tool designed specifically for analysis of Ego-Centric Network data
 - Built in function to export data to other packages
- Still in Beta

E-NET: What it does

- Allows for loading of “cases” of ego networks
- Allows for simultaneous calculation of network metrics across all cases, presently including:
 - Structural measures
 - Degree/Density, EffSize, Efficiency, Constraint, Hierarchy
 - Compositional Measures
 - Proportions for categorical
 - Mean, sum, min, max for continuous
 - Heterogeneity
 - Homophily
- Visualization

Demo
