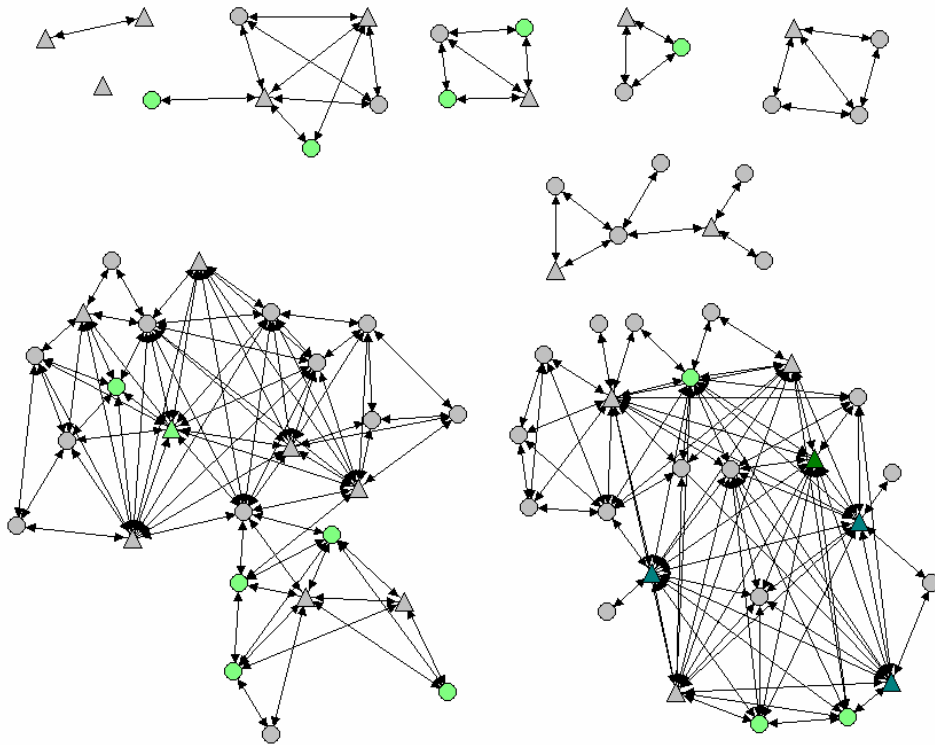


An Introduction to Applying Social Network Analysis to Behavioral Research on HIV/AIDS

Eric Rice, Ph.D.

Assistant Research Sociologist
Center for Community Health, and
Semel Institute for Neuroscience and Human
Behavior
University of California, Los Angeles



Outline of Talk

- Types of Data
- Basic Measures and Concepts
- Example of and application to HIV prevention
- Suggested further readings & Software

Standard Statistical Methods

- Standard statistical methods assume the world is an independent set of actors who have clusters of attributes
- On a basic level, measure individual behaviors and attributes and look for correlations

The social network perspective

- Society is a collection of **actors** connected by **relationships**
- Actors = nodes = individuals
 - Examples: persons, organizations, nations
- Relationships = ties = connections
 - Examples: friendship, social support, shared group membership, treaty members
- A specific network is a specific group of actors connected by a specific relationship
- Inherently violates the assumption of independence!

Opportunities and Constraints

- Networks serve as conduits of information and resources
 - Much applied work on influence and support in networks
- An individual's position in a network determines a person's opportunities and constraints
 - Much theoretical work on power and status, less applied

Two Types of Network Data

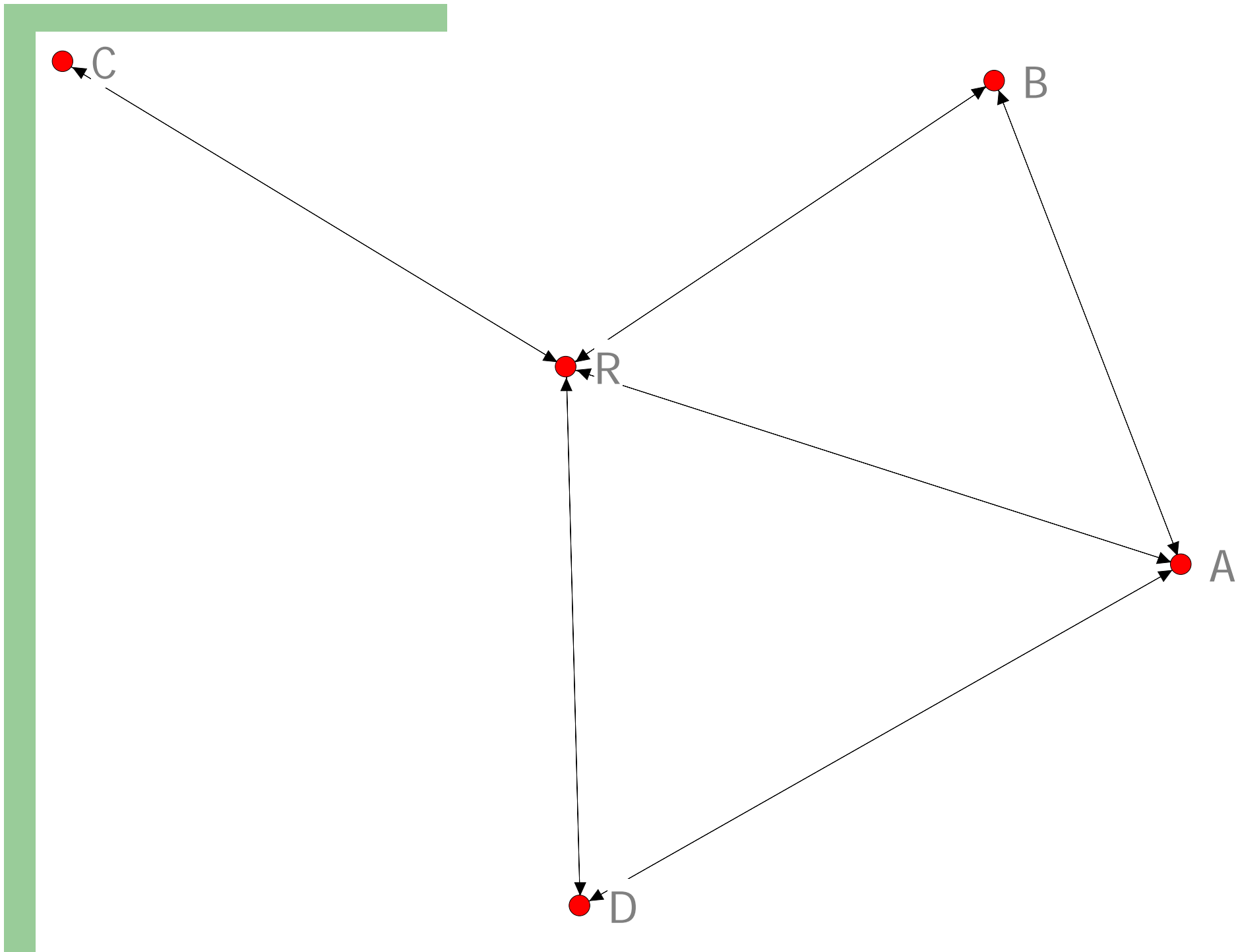
- Ego-Centric
 - Network from the perspective of an individual respondent
 - Their set of immediate ties and the relations among those ties
- Sociometric
 - Population view of a network
 - All relevant ties in among a group of actors

Ego Centric Network Data

- Easy to collect with standard survey techniques
- Doesn't go beyond the scope of the individual being interviewed
- Can get around violation of assumption of independence if your sample is of independent individuals – the network can become a set of attributes

Ego-centric HIV risk taking networks

- From a pilot study of mine:
 - List up to five friends you spent at least 1 hour with this week
 - Provide connections among them
 - Provide attributes of the actors (e.g. who uses methamphetamine?)

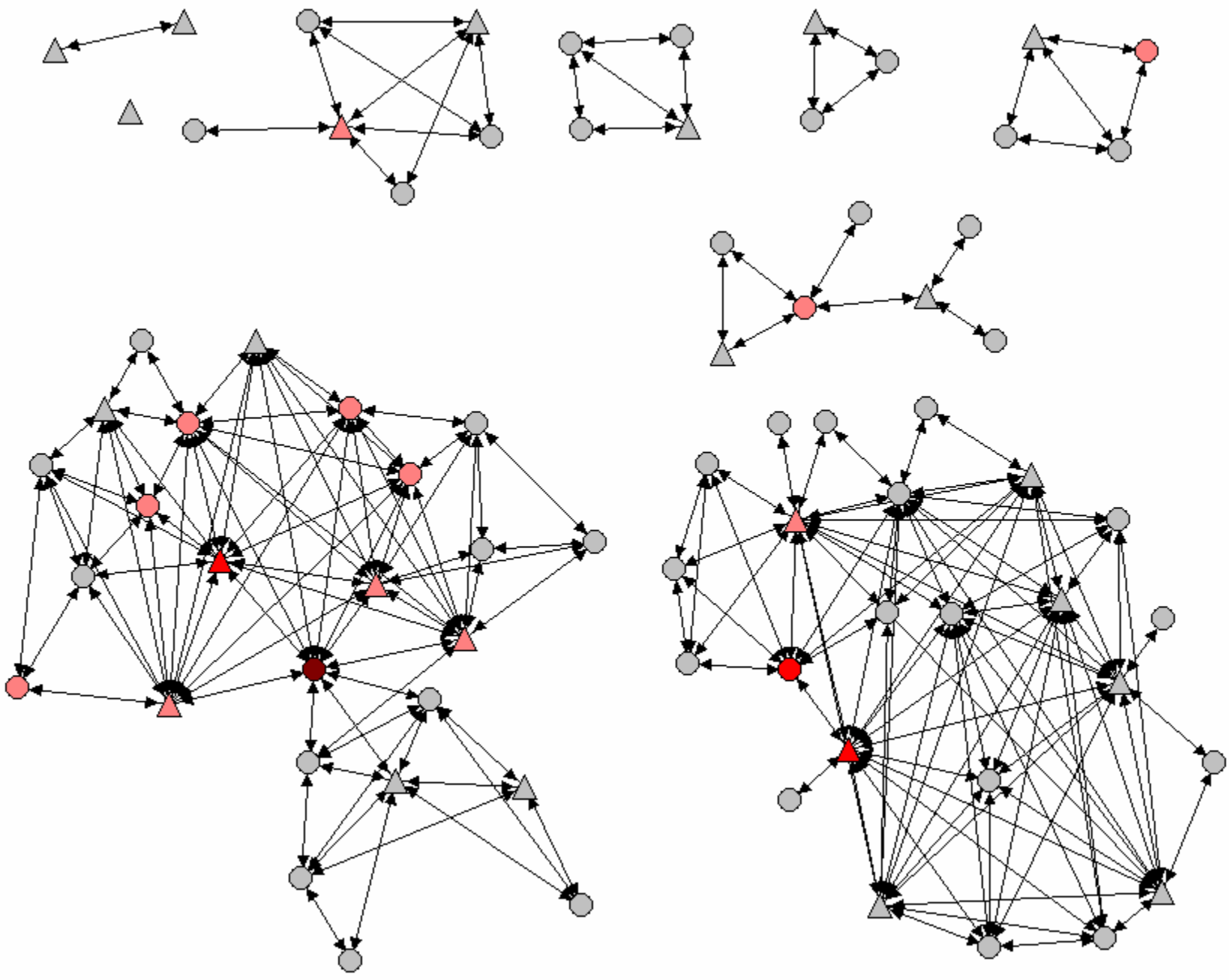
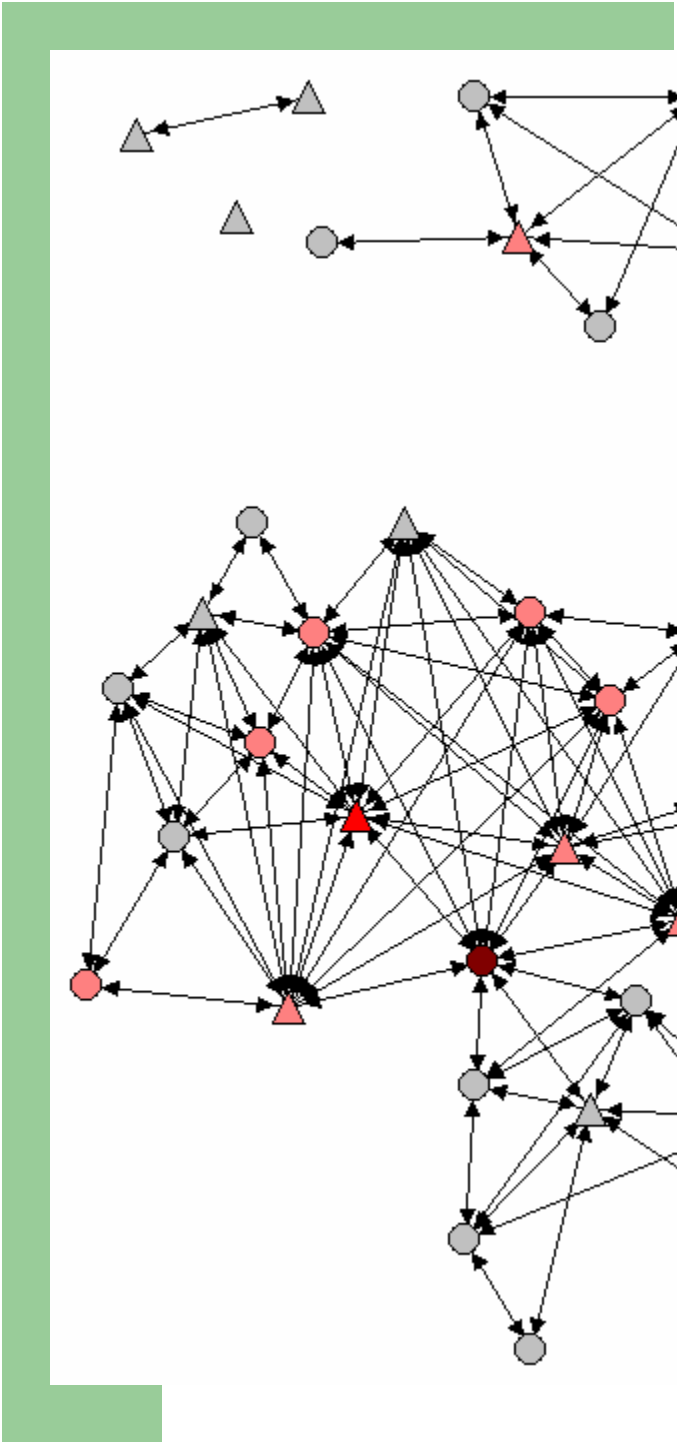


Sociometric Data

- Bird's eye view of the network
- Whole population of interest
- Get information on relevant actors and their ties
 - Can come from archival data
 - Membership rosters
 - Key informants

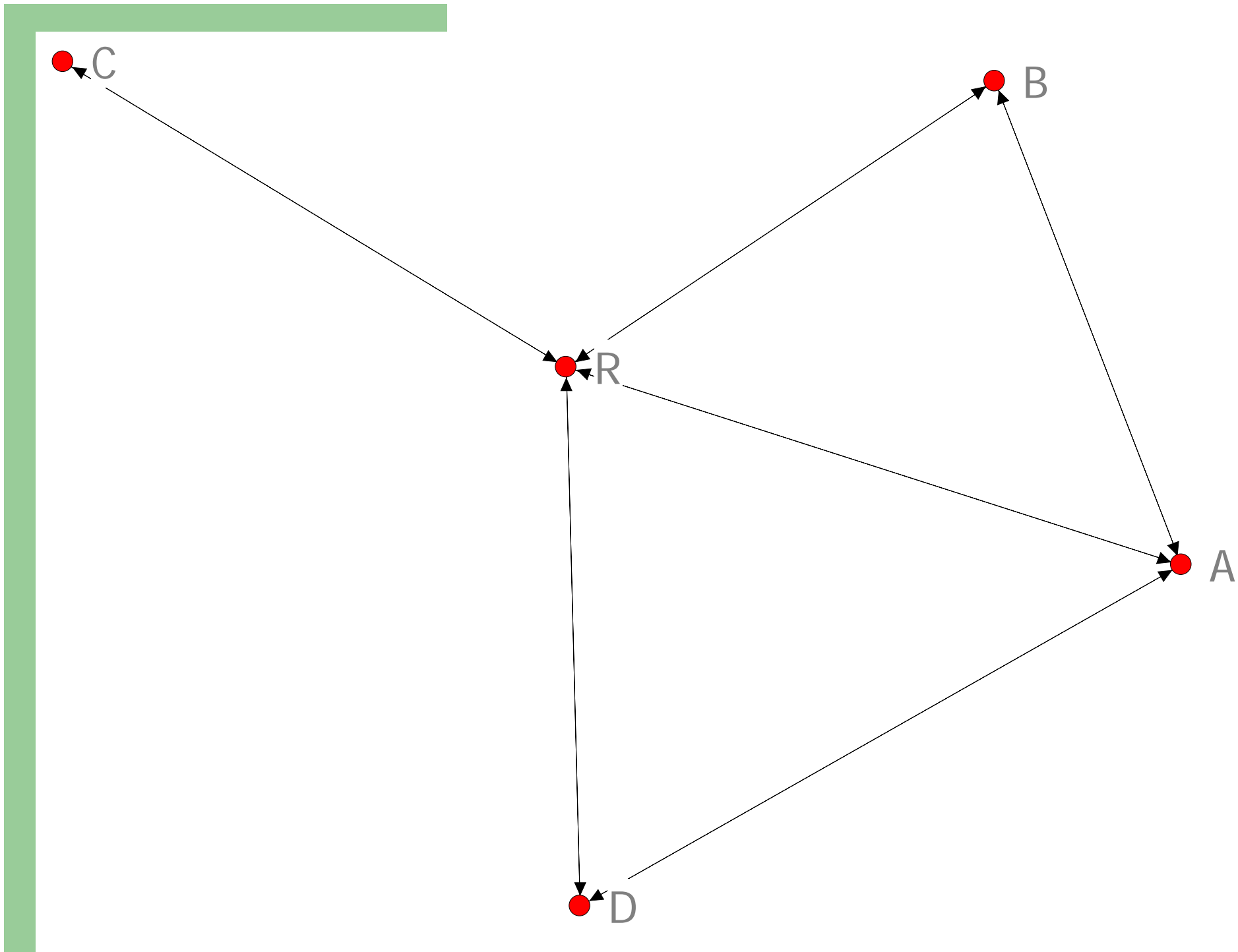
Sociometric applied data example

- Social ties: who hangs out with and/or talks to who at a particular drop in agency
- among a group of 76 youth as reported to me by 26 key informants
- Can be linked to key attributes: who gets you into trouble?



What does the data look like?

- Regardless of ego-centric or sociometric data collection: the data comes a matrix
- Rows and Columns represent actors, cells represent ties
- Start with an example from an ego-centric data set, because it is smaller



The Resulting Matrix

	R	A	B	C	D
R	0	1	1	1	1
A	1	0	1	0	1
B	1	1	0	0	0
C	1	0	0	0	0
D	1	1	0	0	0

For Ego-Centric Data Sets

- The data set is a series of independent matrices
- One per respondent
- Typically, perform calculations on each network one at a time, merge with larger data set of attributes for subsequent analysis

Some Basic Network Properties that Affect Behavior

- Size
- Density
- Behavior Frequency
- Centrality

Behavior Frequency

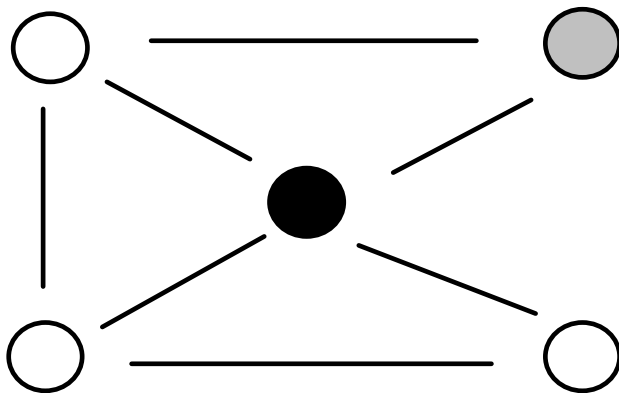
- Measure of the propensity of a behavior or attribute in a network
- Number of nodes with a given attribute divided by the number of nodes in the network
- Substantively, the more nodes with an attribute, the more likely a particular node will display that attribute
 - Could be the result of selection – people create groups based on shared attributes and tastes
 - Could be the result of social learning – people adopt behaviors of those close to them in network space

Behavior Frequency

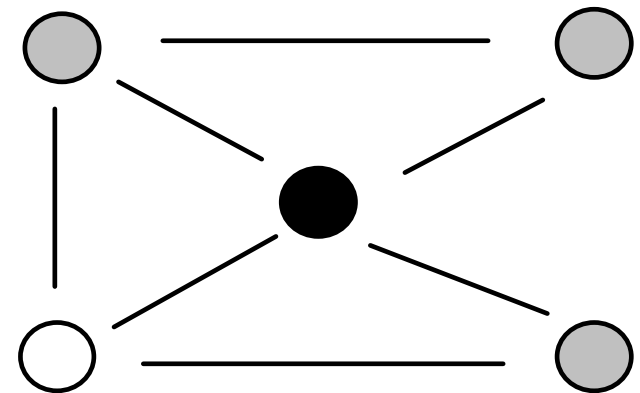
Frequency

Lower

Higher



$$\text{Freq} = 1/4 = .25$$



$$\text{Freq} = 3/4 = .75$$

Density

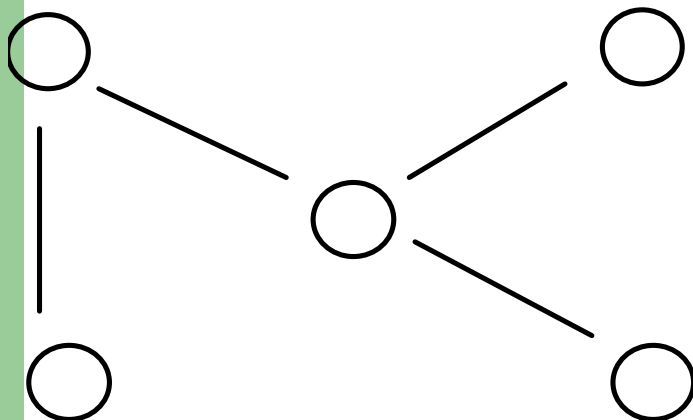
- Measure of the connectivity of the network overall.
- The number of ties divided by the number of possible ties.
- Substantively, denser networks spread information faster, are more cohesive, more homogeneous with respect to behavior

5 actors = 10 possible ties

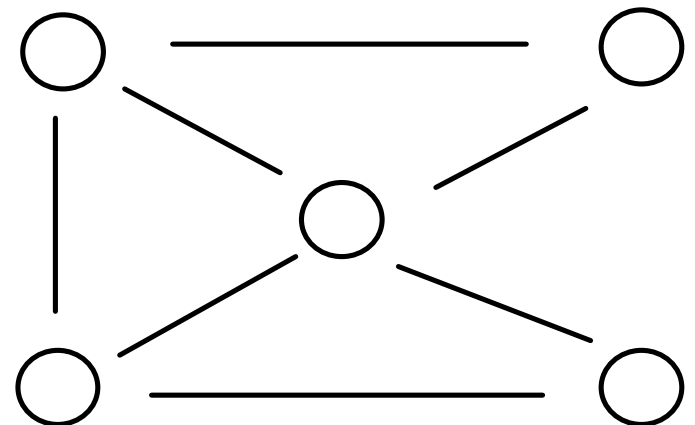
Density

Lower

Higher



density = $4/10 = 0.4$



density = $7/10 = 0.7$

Size

- How many actors are in the network?
- If the question forces the respondent to always list 10 or 5, this is meaningless
- People who can't list or won't list many others are likely isolated – could mean more alienated and self-destructive or..
- If in an anti-social environment, social-isolation may be protective – reduced peer pressure toward negative behaviors

Ego-centric HIV risk taking networks

- From a pilot study of mine:
 - List up to five friends you spent at least 1 hour with this week
 - Provide connections among them
 - Provide attributes of the actors (e.g. who uses methamphetamine?)

Ego-centric data can be used in standard linear modeling techniques

Table 5: Logistic Regression of the Effects of Network Structure on Methamphetamine Use

	Beta	Odds Ratio	95 % Confidence Interval	
Background Characteristics				
Age (years)	0.36	1.43	0.68	3.00
Gender (Male=1)	-0.44	0.64	0.06	7.14
Housing Status (Stable=1)	0.67	1.95	0.07	56.03
Network Properties				
Density	-3.72	0.02	0.00	28.14
Size	-1.57 a	0.21	0.04	1.16
# of Peers Using Meth	2.79 **	16.30	2.04	130.52
Intercept	-4.53			

Centrality

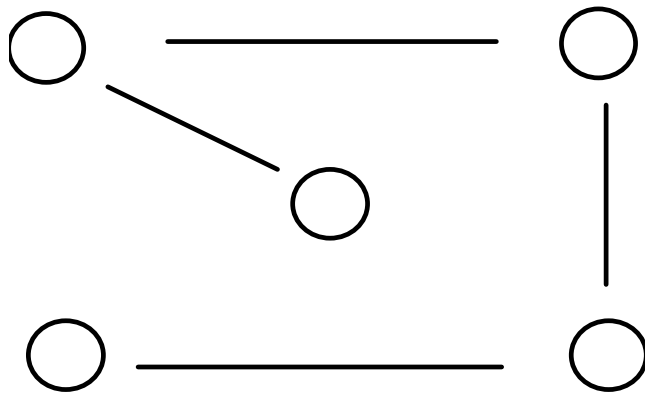
- Substantively, Centrality is a structural level measure of social status
- There are a host of ways to measure centrality
- Easiest is in-degree, a count of how many ties an actor has
- Can then be standardized by dividing by the number actors in the network – 1 (you don't count yourself)

Centrality of the Middle Node

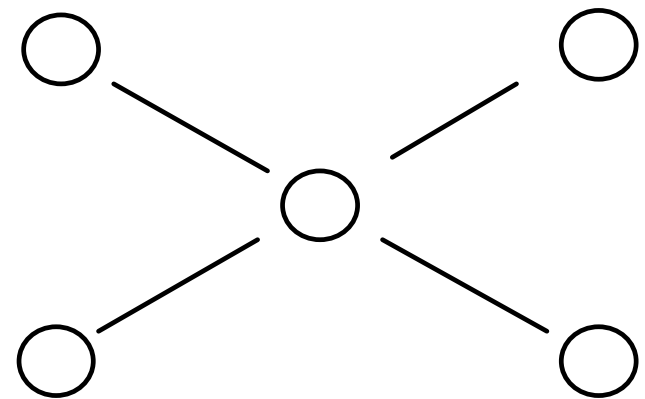
Centrality

Lower

Higher



$$\text{Cent} = 1/4 = 0.25$$



$$\text{Cent} = 4/4 = 1.0$$

But...

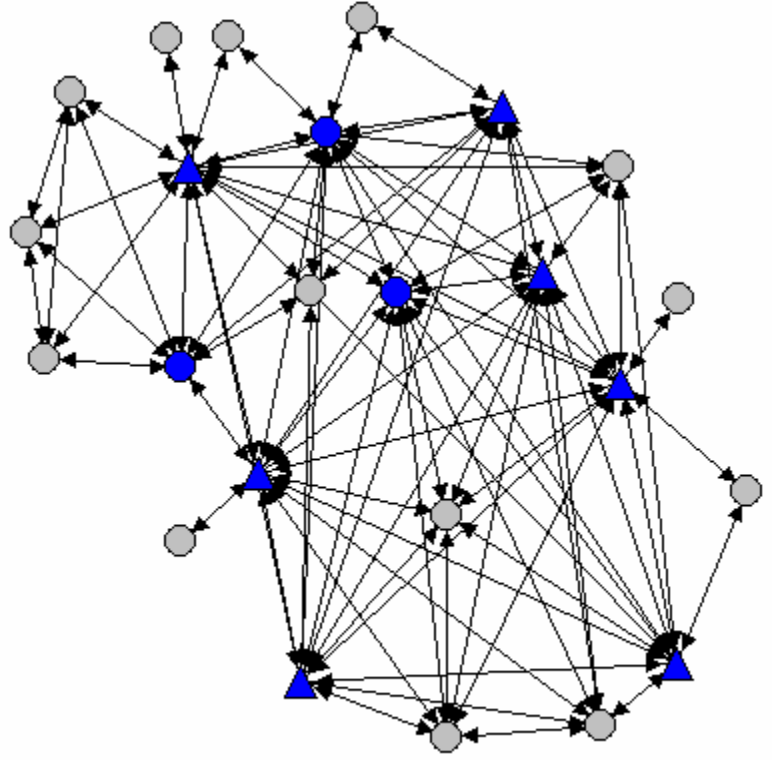
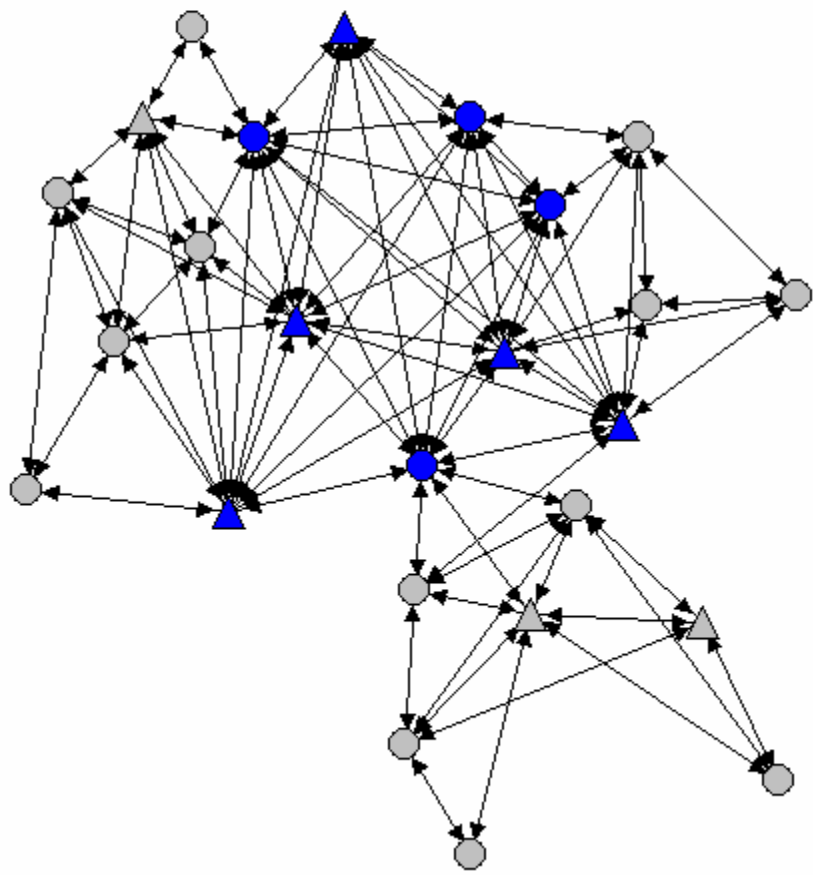
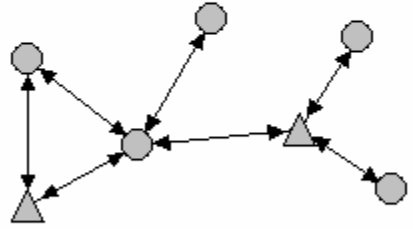
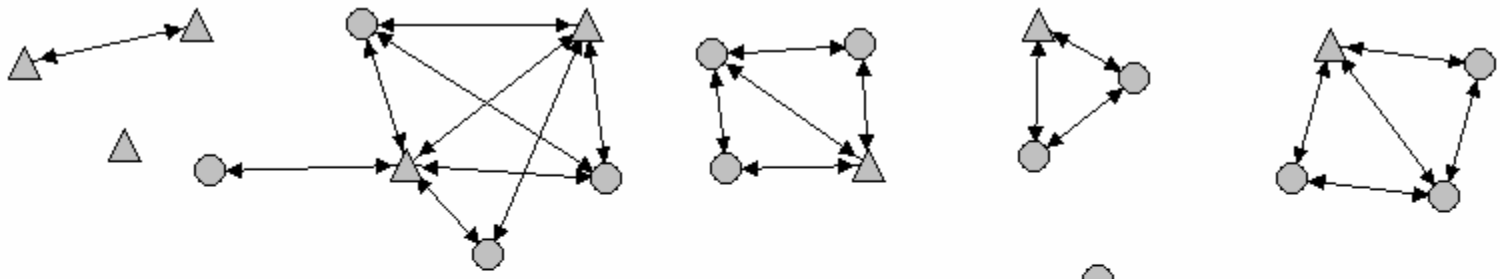
- Centrality can be calculated for each node in the network and then one can make comparisons
- Substantively, central actors are important conduits of information and dissemination
- Less immediately obvious how to incorporate into standard regression, but one idea is:
 - Score centrality of whole network from ego-net data
 - Determine if a behavior in question is associated with most central actor, is respondent's behavior more associated with the most central actor

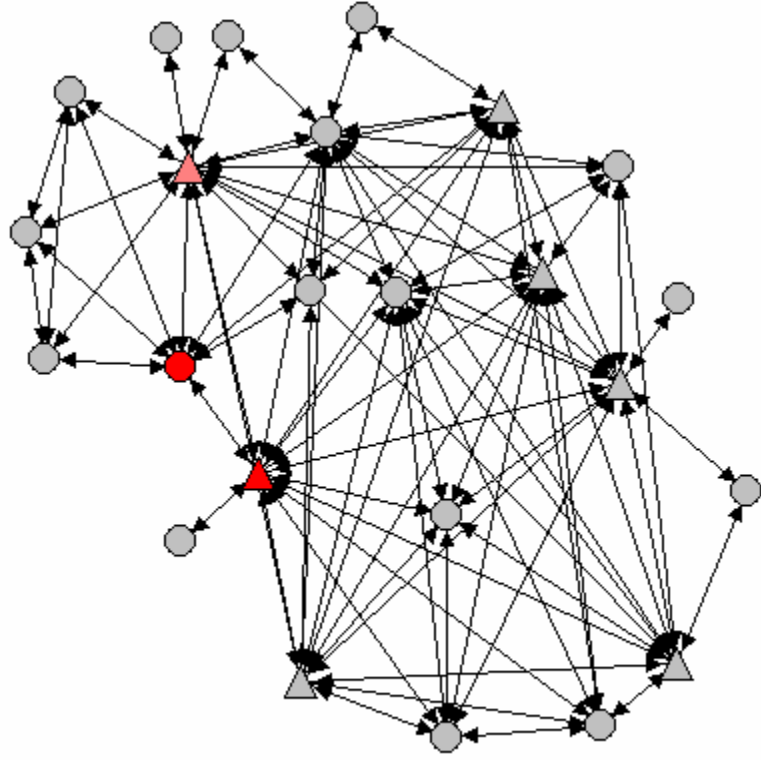
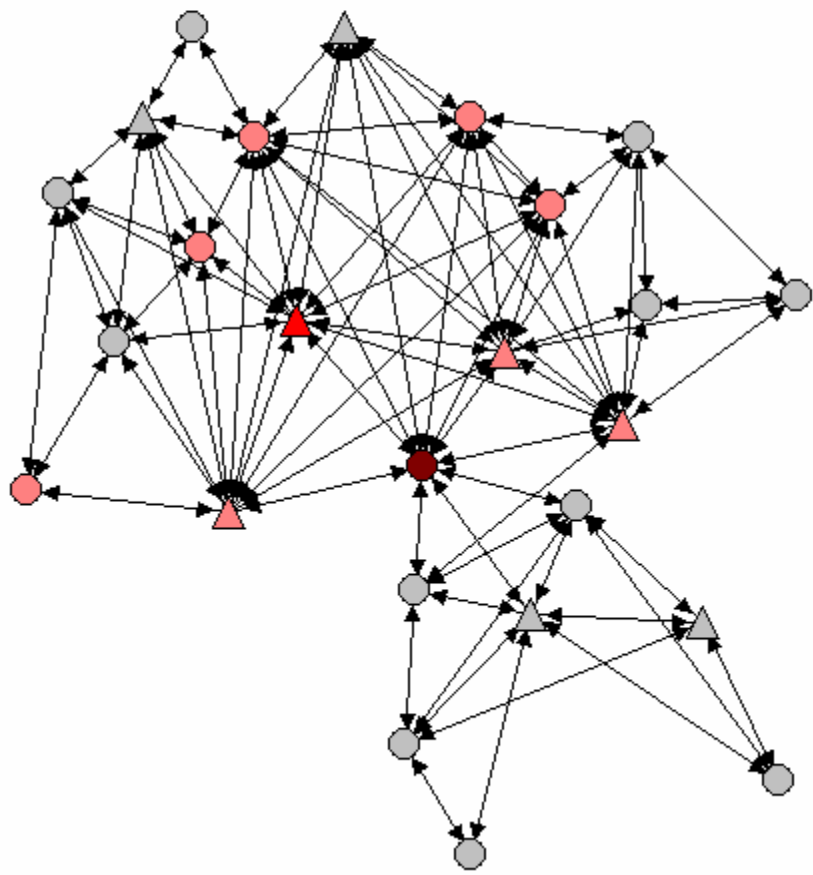
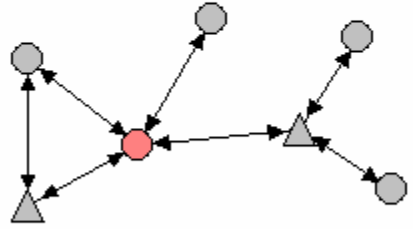
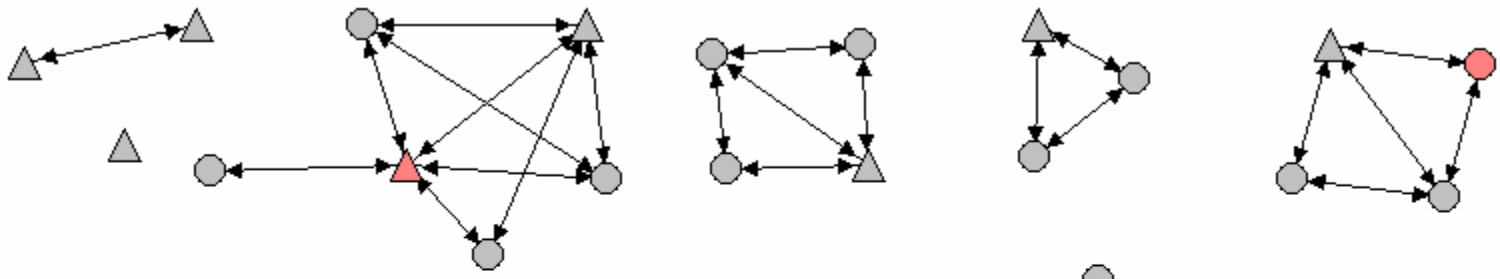
Sociometric Data

- Huge matrices
- Even simple computations, like in-degree centrality get cumbersome
- UCINET is a great software package

Centrality and Social Status

- A substantive use of centrality from my work: use it to identify key persons in the network who could be trained as popular opinion leaders





Suggested Further Readings

- 1. The best and most comprehensive text on network methods:
Wasserman S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press: Cambridge.
- 2. A basic overview of concepts and some analytic techniques:
Knoke, D. & Kuklinski, J. H. (1982). *Network Analysis*. Sage: Newbury Park, London, New Delhi.
- 3. A great text on field issues in data collection:
Schensul, J. J., LeCompte, M. D., Cromley, E. K., & Singer, M. (1999). *Mapping Social Networks, Spatial Data, & Hidden Populations*. Alta Mira Press: Walnut Creek, London, New Delhi.

Software

- <http://www.analytictech.com/>
- UCINET – matrix manipulation and extensive network computation capacity
- NetDraw – Great for graphical representations of networks, enter matrix, get a picture, assists conceptual work, can attach attributes

**Feel free to contact me if you want to
learn more or need consultation**

erice@mednet.ucla.edu