

Network Analysis

An Introduction

Bruce A. Desmarais
University of Massachusetts Amherst

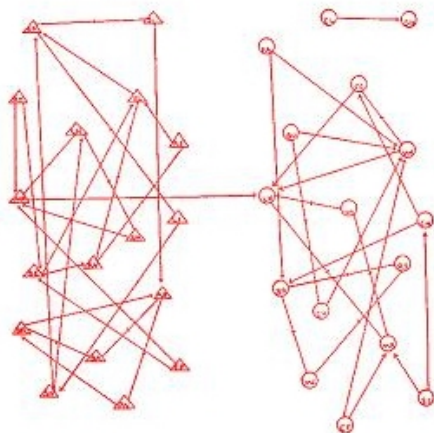
Jacob Moreno's Sociogram

EMOTIONS MAPPED BY NEW GEOGRAPHY

Charts Seek to Portray the
Psychological Currents of
Human Relationships.

New York Times

April 3, 1933

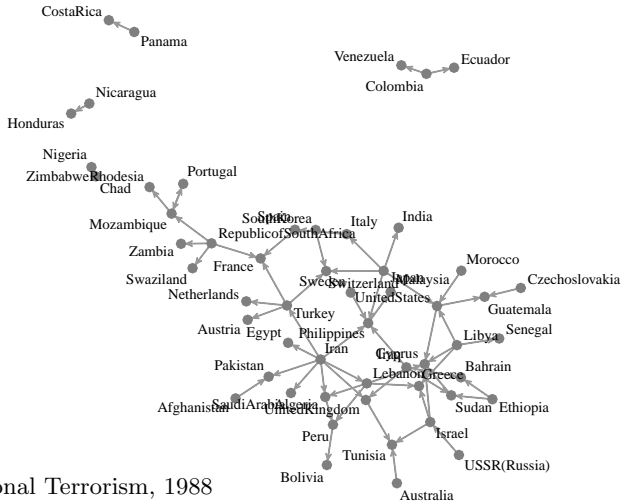


Quoting Moreno on the Analysis of Networks

“Such an invisible structure underlies society and has its influence in determining the conduct of society as a whole. Deep psychological evolutions have been evident throughout the world in the last few years in clashes between groups within nations, and between nations themselves. Until we have at least determined the nature of these fundamental structures which form the networks, we are working blindly in a hit-or-miss effort to solve problems which are caused by group attraction, repulsion and indifference.”

What is a Network (i.e., a graph)?

Set of **nodes** and **relation(s)** defined on them



Transnational Terrorism, 1988

Nodes

Units, the connections among which are to be studied.

Synonyms:

- ▶ **Actor** – from sociometry, common terminology in sociology and psychology.
- ▶ **Vertex** – from graph theory (i.e., math), common terminology in mathematics and physics.

Term **node** is common in statistics and applied sciences outside of soc and psych.

Let's brainstorm a list of nodes...

Relation

Set of **relational ties** or **ties** that span **dyads**

Synonyms:

- ▶ **Link** – common in computer science (e.g., huge lit on “Link Prediction”) and social sciences.
- ▶ **Edge** – graph theoretic terminology common in physics and math, but also elsewhere.

Edge Characteristics

- ▶ Weight
- ▶ Qualitative attributes
- ▶ Direction

e.g., we meet for 2.5hrs for me to teach you network analysis.

Let's brainstorm tie types to go with our nodes...

Subgraphs

- ▶ **Dyad** – Pair of nodes and relation(s) on them.
- ▶ **Triad** – Triple of nodes and relation(s) on them.
- ▶ **Subgroup** – Flexible; can refer to subnetwork of any size.

To qualify as a node in a network (even an isolate), some sort of group membership is often required.

Were groups implicitly used to
define our nodes?

Why network science?

How does network analysis depart from
conventional quantitative social science?

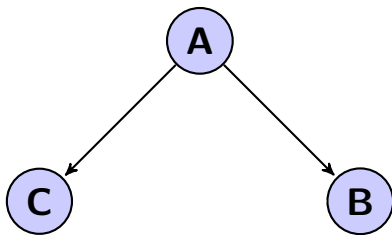
In conventional quantitative social science...

The assumption that units are independent is
leveraged to define quantities of interest
(e.g., a mean, a regression coefficient)

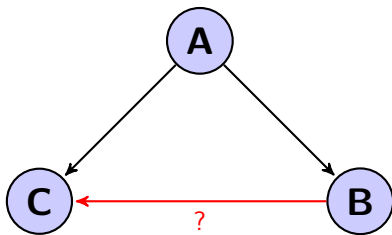
In the network analytic approach

The analyst is interested in the ways in which nodes and relations are interdependent.

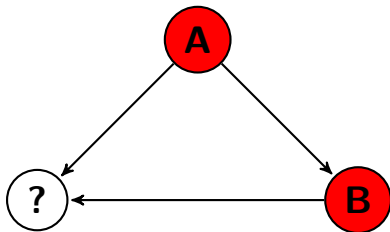
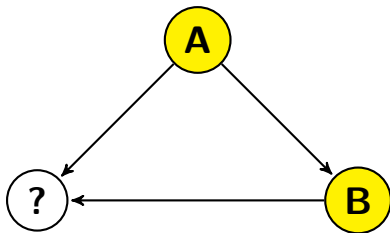
Interdependence: is a friend of a friend a friend?



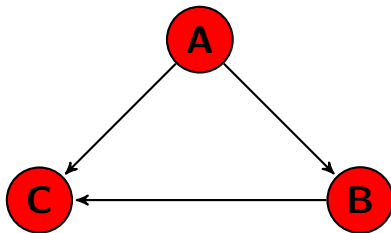
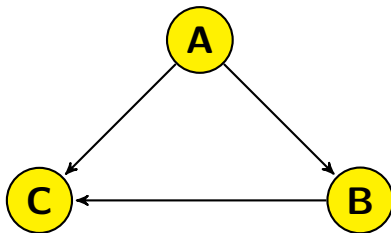
Interdependence: is a friend of a friend a friend?



Do birds of a feather flock together?



Do birds of a feather flock together?



What forms of interdependence
would we expect to see in our
example networks?

Network Theory

Precise statements about network structure and measurements thereof.

Transitivity Example:

- ▶ **Structure expectation:** If A is tied to B, and B is tied to C, then A is more likely to be tied to C than if B was not tied to C.
- ▶ **Measure expectation:** In a transitive network, we'd expect to see more closed triangles than in a network that is not transitive.

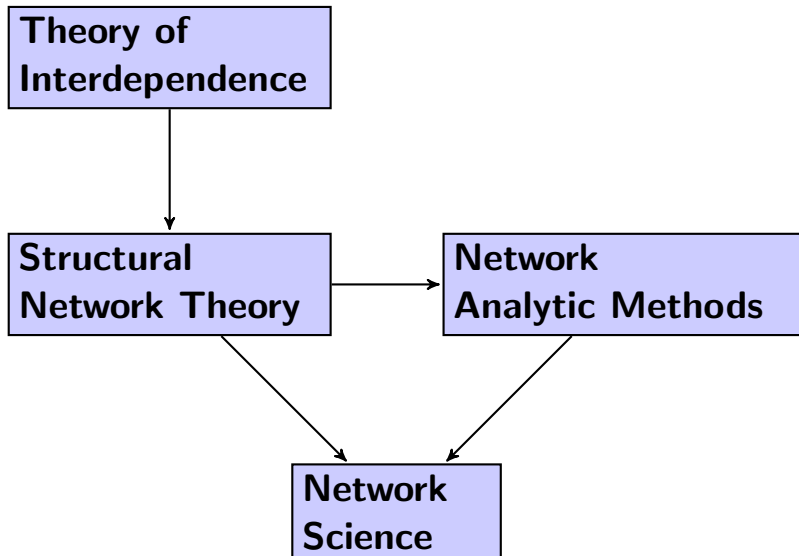
Network Methods

Descriptive, visual and statistical routines designed to suggest and/or assess network theory.

Transitivity Example:

- ▶ **Computation:** How can we count the number of closed triangles in the network?
- ▶ **Comparison:** How do we compare the triangle count to a hypothetical network without transitivity/clustering?

The big picture



Network Data: Structural Variables

No Cases by Variables...

	US	UK	FR	...
US	0	0	0	...
UK	0	0	1	...
FR	1	1	0	...
\vdots	\vdots	\vdots	\vdots	\ddots

- ▶ Called an **adjacency matrix**
- ▶ The matrix is square (row and column labels are the same)
- ▶ 1 in row A, column B indicates A sends a tie to B

Network Data: Node Attributes

Cases by Variables...

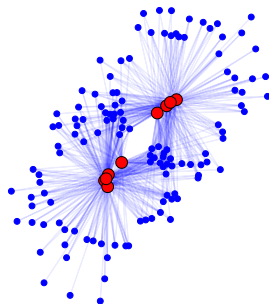
	Continent
US	North America
UK	Europe
FR	Europe
⋮	⋮

- ▶ Also called **compositional variables**
- ▶ Combined into conventional dataset with nodes as units of analysis
- ▶ Interval, ordinal, categorical

Can we think of some node attributes
that would be related to our example networks?

Node Attributes and Multi-mode Networks

- ▶ Node attributes may constrain relations
- ▶ **Modes:** Node types
- ▶ **Multi-mode network:** only connections between different node types



Measuring Network Data

Step 1: Defining the population

Measuring Network Data

Step 2: Sample or Census?

Sampling Approaches:

- ▶ Simple random sample (i.e., simple egonet)
- ▶ Snowball sampling

Measuring Network Data

Step 3: Collection Method

- ▶ Questionnaire/Interview (open-ended vs. list)
- ▶ Archival (e.g., web scraping)
- ▶ Observation
- ▶ Experiment (i.e., create the network)

Important Considerations (Measurement Error)

- ▶ Validity
- ▶ Reliability

How could we go about measuring
the example networks?