

Introduction to Social Network Analysis and its Applications

Mark S. Handcock

`handcock@ucla.edu`

October 1, 2024

Focus on Relations

Relationship: an irreducible property of two or more entities

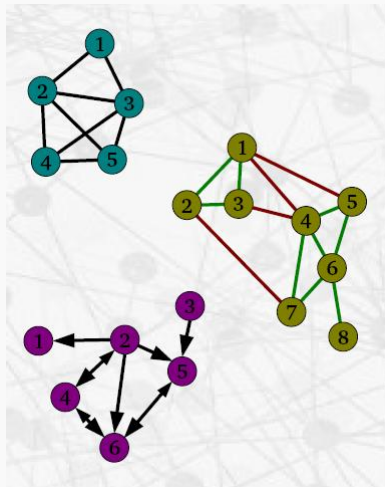
- Compare with properties of entities alone (“attributes”)

Focus: the properties and consequences of relations
(rather than individual properties)

- Entities can be persons, non-human animals, groups, locations, organizations, regions, etc.
- Relationships can be communication, acquaintanceship, sexual contact, proximity, migration rate, alliance/conflict, etc.
- Social network analysis: the study of relational data arising from social systems

Some Vocabulary

- *Network*: a collection of entities, together with a set of relations on those entities
 - Entities: nodes, vertices, actors
 - Relations: edges, ties, arcs, links.
 - Relations can be:
 - directed or undirected
 - signed or valued
 - multiplex



Using Mathematical Network to Represent Relations

- Networks are widely used to represent data on relations between interacting actors or nodes.
- The study of social networks is multi-disciplinary
 - plethora of terminologies
 - varied objectives, multitude of frameworks
- Understanding the structure of social relations has been the focus of the social sciences
 - *social structure*: a system of social relations tying distinct social entities to one another
 - Interest in understanding how social structures form and evolve

Using Mathematical Network to Represent Relations

- Attempt to represent the structure in social relations via networks
- The data are of at least three forms:
 - individual-level information on the social entities
 - relational data on pairs of entities
 - population-level data

Deep literatures available

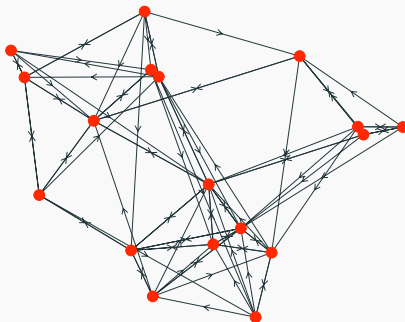
- Social networks community (Heider 1946; Frank 1972; Holland Leinhardt 1981)
 - social theory, description, survey design
- Statistical Networks Community (Fienberg 81, Frank&Strauss 86; Snijders 97)
 - statistical modeling, network sampling, design-based inference)
- Spatial Statistics Community (Besag 1974)
 - statistical models for spatial lattices, dependence
- Graphical Modeling Community (Lauritzen and Spiegelhalter 1988, ...)
 - multivariate analysis
 - nodes are variables, relations define dependency between variables
- Machine Learning Community (Jordan, Jensen, Xing,)
 - clustering, prediction, computation
- Physics and Applied Math (Newman, Watts, ...)
 - generative and agent-based models, emergent features

Core areas of statistical network analysis

- *Statistical modeling*: specification, evaluation and fitting of network models.
 - Specification: creating probability models for phenomena.
 - Testing: evaluation of competing theories of network formation.
 - Estimation: evaluation of parameters in a presumed network model.
 - Description: summaries of main network patterns.
 - Prediction: prediction of missing or future network relations.
- *Design-based inference*: Inferring a network from sampled data.
 - Design: survey and data-gathering procedures.
 - Inference: generalization of sample data to the full network.

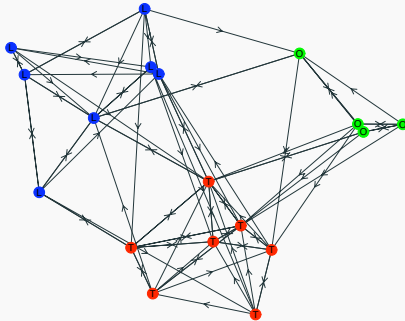
Example of Social Relationships between Monks

- Expressed “liking” between 18 monks within an isolated monastery \Rightarrow Sampson (1969)
 - A directed relationship aggregated over a 12 month period before the breakup of the cloister.



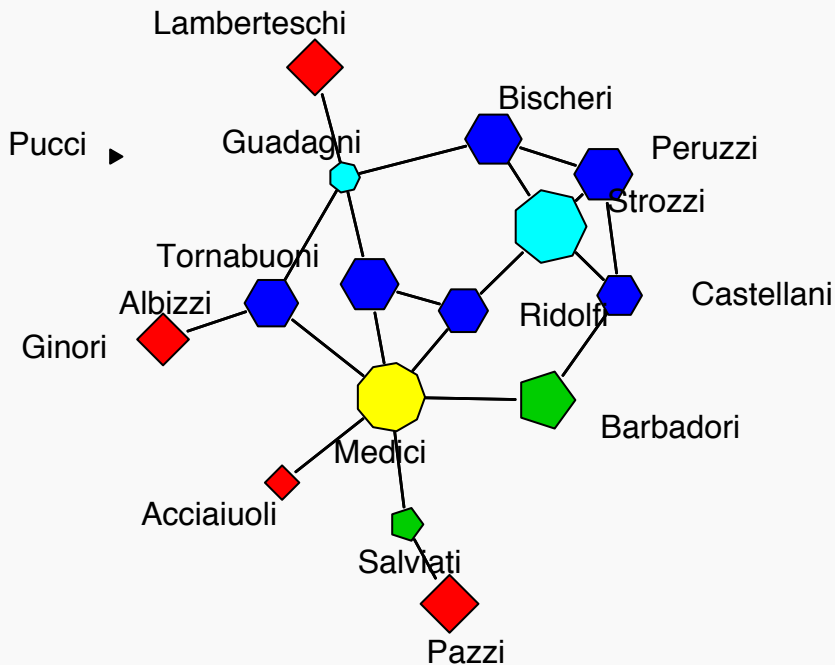
Example of Social Relationships between Monks

- Expressed “liking” between 18 monks within an isolated monastery \Rightarrow Sampson (1969)
 - A directed relationship aggregated over a 12 month period before the breakup of the cloister.
- Sampson identified three groups plus:
(T)urks, (L)oyal Opposition, (O)utcasts and (W)averers



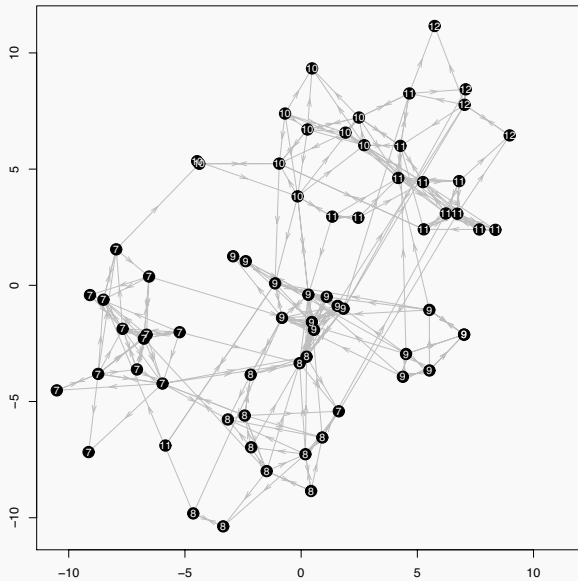
Marriage ties among Renaissance Florentine families

- Collected from historical documents.
 - ⇒ Padgett, John F. 1994. "Marriage and Elite Structure in Renaissance Florence, 1282-1500".
- the data include families who were locked in a struggle for political control of the city of Florence around 1430. Two factions were dominant in this struggle:
 - one revolved around the infamous Medicis,
 - the other around the powerful Strozzi.
- There is also a business relations network and covariates on each family (such as wealth and political representation)

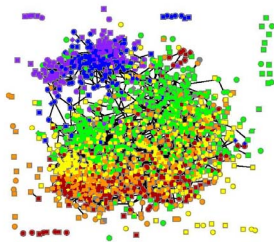


Examples of Friendship Relationships

- The National Longitudinal Study of Adolescent Health
 - ⇒ www.cpc.unc.edu/projects/addhealth
 - “Add Health” is a school-based study of the health-related behaviors of adolescents in grades 7 to 12.
- Each nominated up to 5 boys and 5 girls as their friends
- 160 schools: Smallest has 69 adolescents in grades 7–12



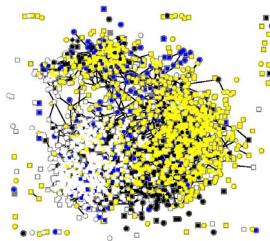
School Community Stratum 44
mutual friendships by Grade



2209 Students

- Grade 7
- Grade 8
- Grade 9
- Grade 10
- Grade 11
- Grade 12
- Grade NA

School Community Stratum 44
mutual friendships by Race



2209 Students

- White (non-Hispanic)
- Black (non-Hispanic)
- Hispanic (of any race)
- Asian / Native Am / Other (non-Hispanic)
- Race NA

Features of Many Social Networks

- *Density* of ties
- *Individual heterogeneity* in the propensity to form or receive ties
 - degree heterogeneity
 - sociability
 - popularity
 - attractiveness
- *Reciprocity* of ties (in a directed network)
mutuality



Features of Many Social Networks



- *Homophily* by actor attributes

⇒ Lazarsfeld and Merton, 1954; Freeman, 1996; McPherson et al., 2001

- higher propensity to form ties between actors with similar attributes
e.g., age, gender, geography, major, social-economic status
- attributes may be observed or unobserved

Features of Many Social Networks

- *Transitivity* of relationships



- friends of friends have a higher propensity to be friends
- *Balance* of relationships \Rightarrow Heider (1946)
 - people establish and maintain balance in their relationships
- *Structural Equivalence* of nodes
 - some nodes may have identical or similar patterns of relationships
- *Context* is important \Rightarrow Simmel (1908)
 - triad, not the dyad, is the fundamental social unit

Defining Features of Social Network Analysis

- Motivated by a relational, structural perspective
- Grounded in systematic empirical data
- Draws heavily on visualization
- Relies on mathematical and statistical models

(Linton Freeman, *Development of Social Network Analysis*)

Social Network Perspective

Simmel (1917/1950, p. 10) went so far as to define society simply as

“...the name for a number of individuals connected by interaction.”

For Simmel, sociology was no more and no less than the study of the patterning of interaction.

(Linton Freeman, *Development of Social Network Analysis*)

Social Network Perspective

Bruce Mayhew 1980 “Structuralism vs. Individualism”

“...the individual is never the unit of analysis in either research or theory construction. *Rather, in this structuralist conception of social life, sociologists are studying a communication network mapped on some human population.* That network, the interaction which proceeds through it, and the social structures which emerge in it are the subject matter of sociology.”

Social Network Perspective

- Actors and their actions are interdependent
- Ties between actors are channels for flow of resources
- Network ties both enable and constrain actions
- “Structure” is seen as enduring patterns of relations

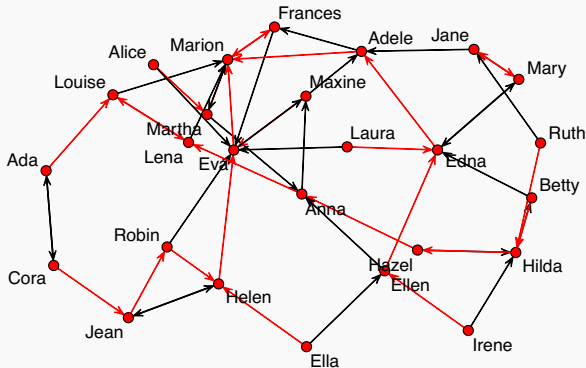
Network Topics

- Diffusion of innovation
- Migration
- Exchange
- Organizations
- Political alliances
- Community
- Epidemiology
- Communication
- Resource flows
- Social Support
- ...

Some idiosyncratic networks

- Moreno's dining table friendships
- Radio communication networks during WTC disaster
- Emergent multi-organizational networks (EMON) during WTC and Hurricane Katrina
- Friendship among graduate students
- Global migration patterns

The black lines are the most liked, the red lines are the second most liked.



WTC Radio Communication Networks for Specialists

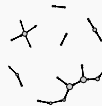
Channel lincoln.tunnel.police



Channel newark.ch23.ewr.command



Channel newark.ch25.ewr.TACI



Channel newark.ch26.CPD



Channel spen1.ch15



Channel spen2.ch16



Channel wtc.chw.police



Channel path.ch21.r2.trainmaster



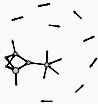
Channel wtc.chx.security



Communication networks amongst emergence units. The direction of the arrow indicated transmission to that unit

WTC Radio Communication Networks for Non-specialists

Channel path.ch27.r3.communications



Channel newark.ch36.ops.terminals



Channel newark.ch39.maintenance



Channel path.ch26.r1.trainmaster



Channel wtc.chy.ops



Channel wtc.chz.vertical.transp



Channel newark.ch38.facilities



Channel wtc.chb.mech.electric

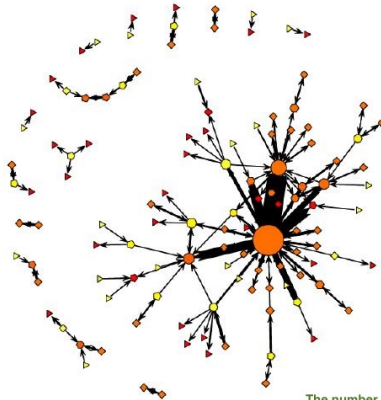


Communication networks amongst non-specialist emergence units.

The direction of the arrow indicated transmission to that unit

WTC Radio Communication Networks

WTC Data: Plot for Channel wtc.chy.ops

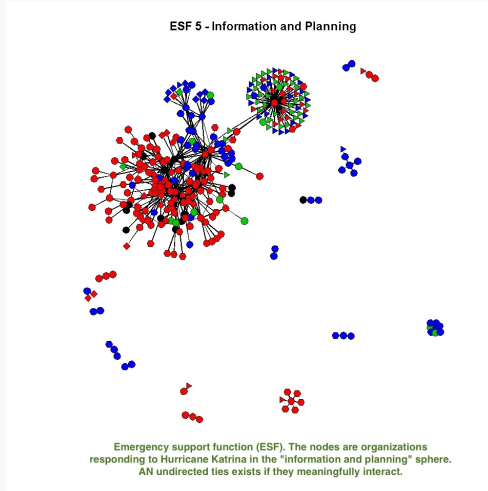


For the operations personnel from WTC who self-evacuated and reassembled at an ad-hoc collection point on a nearby street

The number of sides of the node increases with its degree.

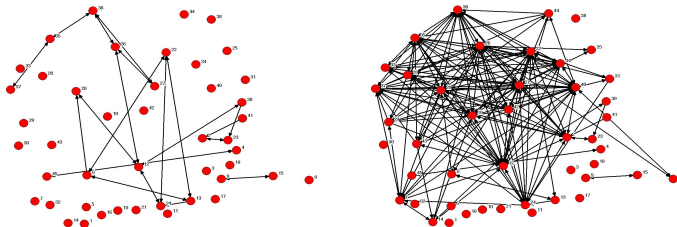
(C.T. Butts and M. Petrescu-Prahova, 2005)

Emergency Support Function 5 - Information and Planning

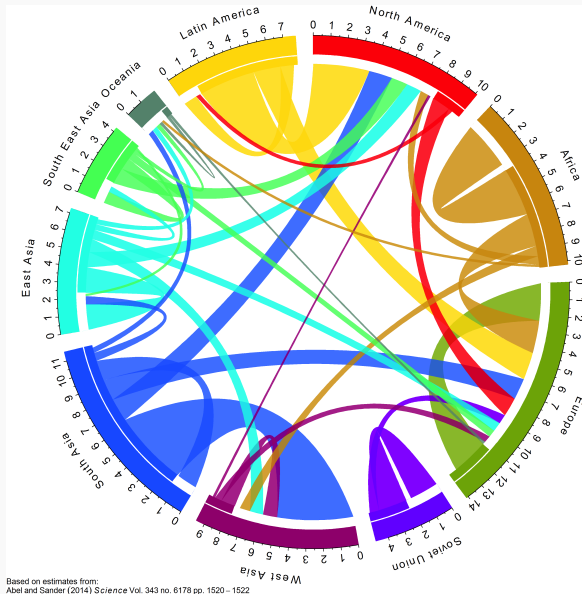


Friendship among Graduate Students

Two networks?

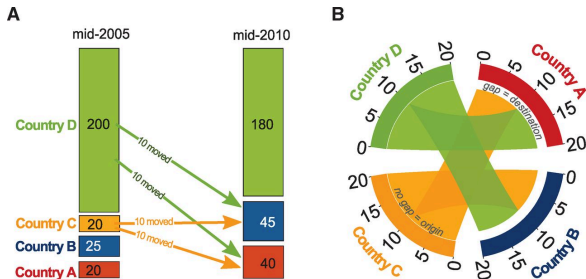


Global Migration Flowmap



Global Migration Flowmap

Fig. 1 Linking migrant flow to stock data and visualizing flows via circular plots. (A) The simplified example illustrates our method for estimating 5-year migration flows from changes in stock data between mid-2005 and mid-2010 (details are available in the supplementary materials).



G J Abel, and N Sander Science 2014;343:1520-1522

Published by AAAS



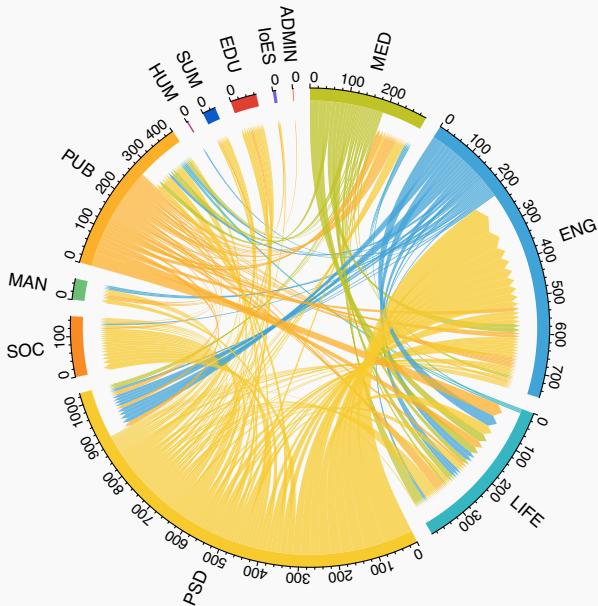
Global Migration Flowmap

- The circular plot visualizes the migrant flows estimated in the hypothetical example.
- The origins and destinations of migrants (Countries A to D) are each assigned a color and represented by the circle's segments.
- The direction of the flow is encoded by both the origin country's color and a gap between the flow and the destination country's segment.

Global Migration Flowmap

- The volume of movement is indicated by the width of the flow. Because the flow width is nonlinearly adapted to the curvature, it corresponds to the flow size only at the beginning and end points.
- Tick marks on the circle segments show the number of migrants (inflows and outflows).

Graduate Data Science classes at UCLA



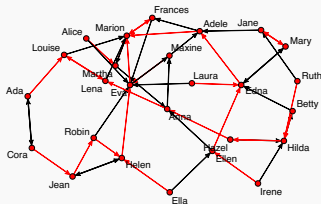
Think Formally

- Network is not just a metaphor
- It is a precise mathematical construct
- Represent, analyze, model relations

Network Methods

- Cannot simply use existing statistical methods
- Units are not independent
- Want to explicitly model interdependencies, relations, linkages, structural patterns
- Network formulation of social relational constructs

Network Analysis Goals



How can we quantify such network patterns?

1. How can we represent features of social relations?
 - reciprocity/sociability/popularity/transitivity :
descriptive statistics
2. Can we identify nodes with similar network roles?
 - stochastic equivalence, node partitioning
3. Can we relate the network to covariate information?
 - homophily, regression modeling