Social Network Analysis: Theory, Visualization, and Modeling

Narisong Huhe

University of Strathclyde,
narisong.huhe@strath.ac.uk

Outline

An interdisciplinary enterprise

- Euler's solution laid the foundations of **graph theory** and **topology**.

An interdisciplinary enterprise

- Euler's solution laid the foundations of **graph theory** and **topology**.
- The roots of network science are particularly strong in **social psychology**, **sociology**, and **anthropology** (e.g., Granovetter, 1973).

An interdisciplinary enterprise

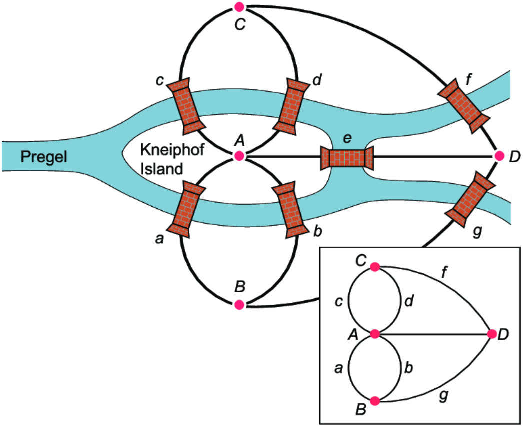
- Euler's solution laid the foundations of **graph theory** and **topology**.
- The roots of network science are particularly strong in **social psychology**, **sociology**, and **anthropology** (e.g., Granovetter, 1973).
- An increasingly number of **economists** also start to investigate the impacts of relational structure on economic activities (e.g., Jackson and Zenou, 2015)

An interdisciplinary enterprise

- Euler's solution laid the foundations of **graph theory** and **topology**.
- The roots of network science are particularly strong in **social psychology**, **sociology**, and **anthropology** (e.g., Granovetter, 1973).
- An increasingly number of **economists** also start to investigate the impacts of relational structure on economic activities (e.g., Jackson and Zenou, 2015)
- However, the recent increased visibility of network analysis is owed mainly to **statistical physicists** (e.g., Barabási and Albert, 1999).

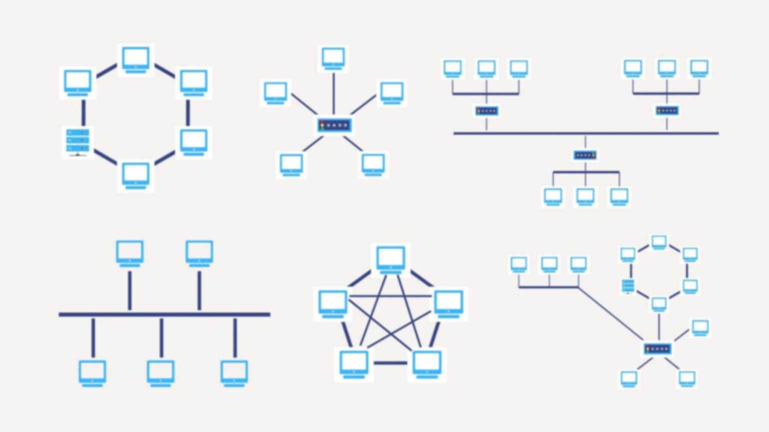
Euler and seven bridges of Königsberg

- Problem: Can we devise a walk through the city that would cross each of those bridges once and only once?
- A **negative** resolution was provided by Euler in 1736.



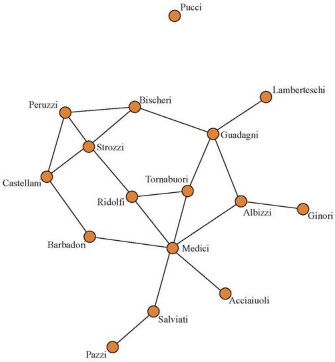
Network typology and Internet

- Computer scientists and the development of Internet



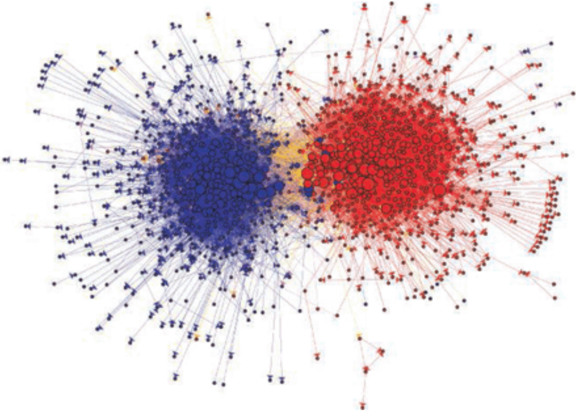
From 2-player games to multi-player games

- Strategies in a networked setting (e.g., Jackson, 2010)
  - e.g., 15th Century Florentine marriage networks (Padgett and McLean, 2006)



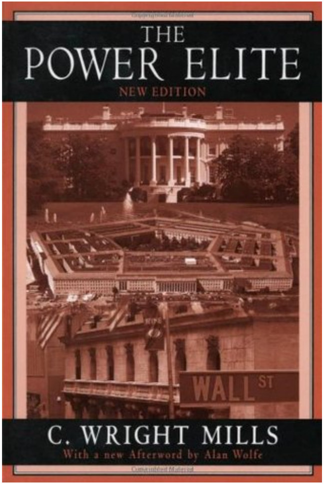
Computational social science

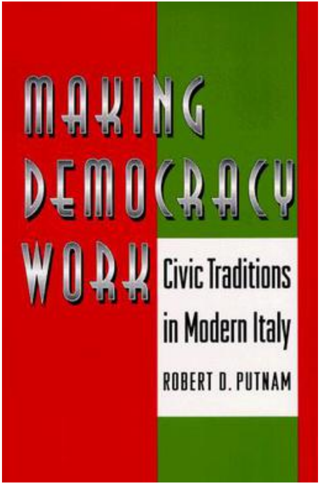
- Stronger computational capacity and bigger data (e.g., Lazer et al., 2009)



Outline

Network theory in PoliSci





Embeddedness and interdependencies

- Targets: Atomic **agents** and rigid **structures**
  - e.g., the agent-structure debate in IR
  - the dominant statistical paradigm also reinforces the primacy of atomic agents

Embeddedness and interdependencies

- Targets: Atomic **agents** and rigid **structures**
  - e.g., the agent-structure debate in IR
  - the dominant statistical paradigm also reinforces the primacy of atomic agents
- A network perspective
  - Agents are not atomic; they are **embedded** in complex networks. (No iid)
  - Power and influence do not derive solely from agents' **attributes**; they can be results of agents' **positions** in a network. (Levels of analysis)
  - Structures are not rigid; they can be molded by **changing interdependencies** among agents. (Causal loops)

Embeddedness and interdependencies

- Targets: Atomic **agents** and rigid **structures**
  - e.g., the agent-structure debate in IR
  - the dominant statistical paradigm also reinforces the primacy of atomic agents
- A network perspective
  - Agents are not atomic; they are **embedded** in complex networks. (No iid)
  - Power and influence do not derive solely from agents' **attributes**; they can be results of agents' **positions** in a network. (Levels of analysis)
  - Structures are not rigid; they can be molded by **changing interdependencies** among agents. (Causal loops)
- The nature of networks leads to **dependence between actors** and also to **dependence between network ties**.

Basic rationales: From local to global

- The ideas of embeddedness and interdependence are attractive, but how to study **complex networks**?

Basic rationales: From local to global

- The ideas of embeddedness and interdependence are attractive, but how to study **complex networks**?
- From **local dynamics** to **global patterns**
  - social network theories focus on local dynamics and how they generate complex global networks

From local competition to global evolution

From local rational persons to global equilibrium

Outline

Rating restaurants

- You are landed in a foreign city, and you have no travel guides or access to the Internet.
- You are not desperately hungry.

A heuristic shortcut

Scale-free network (Barabási and Albert, 1999)

- Two assumptions:
  - a network is growing gradually
  - new comers have slightly higher chances to connect with the popular nodes (i.e., preferential attachment)

Scale-free network (Barabási and Albert, 1999)

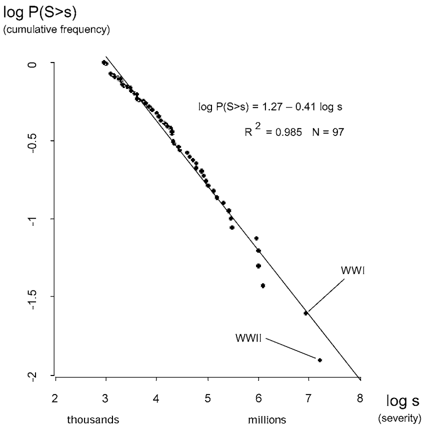
- Two assumptions:
  - a network is growing gradually
  - new comers have slightly higher chances to connect with the popular nodes (i.e., preferential attachment)
- Result: A hub-and-spoke network and it is scale free.

Scale-free network (Barabási and Albert, 1999)

- Two assumptions:
  - a network is growing gradually
  - new comers have slightly higher chances to connect with the popular nodes (i.e., preferential attachment)
- Result: A hub-and-spoke network and it is scale free.
- Evidence:
  - WWW, Facebook, metropolitan cities, ...
  - Cederman (2003): Sizes of war
  - Baumgartner, Box-Steffensmeier and Campbell (2018): Executions in US

Cederman (2003)

FIGURE 1. Cumulative Frequency Distribution of Severity of Interstate Wars, 1820–1997



Source: COW data.

Simulation time!

- Why **inequality** persists in human society (e.g., capitalism)?
- Is it **fair** (e.g., meritocracy)?

Simulation time!

- Why **inequality** persists in human society (e.g., capitalism)?
- Is it **fair** (e.g., meritocracy)?
- Simulation 1: Matthew effects, "the rich get richer and the poor get poorer"
- Simulation 2: Talent vs luck

Outline

Hunting flats

Strong ties, weak ties, and a small world

- My story: Richard told me that Neil had a mortgage friend good at handling cases like mine.
- Two kinds ties
  - type 1: me <–> Richard, Richard <–> Neil, me <–> Neil
  - type 2: Neil –> the mortgage guy

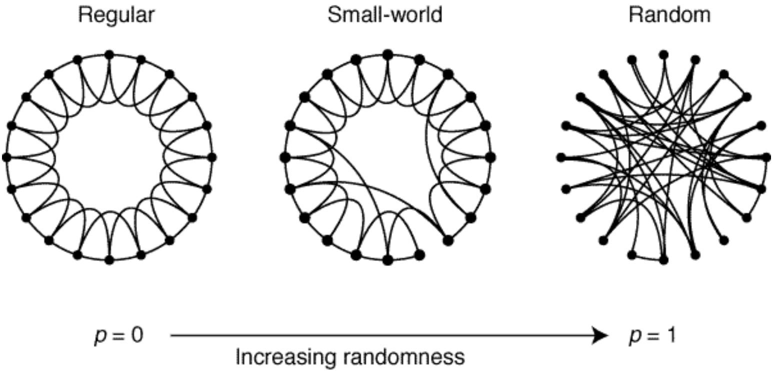
Strong ties, weak ties, and a small world

- My story: Richard told me that Neil had a mortgage friend good at handling cases like mine.
- Two kinds ties
  - type 1: me <–> Richard, Richard <–> Neil, me <–> Neil
  - type 2: Neil –> the mortgage guy
- Granovetter (1973): the "strength of weak ties"
  - While strong ties bond, weak ties bridge.

Strong ties, weak ties, and a small world

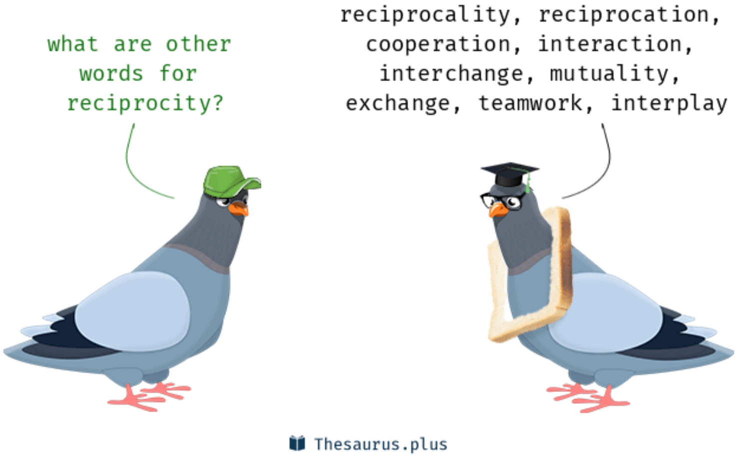
- My story: Richard told me that Neil had a mortgage friend good at handling cases like mine.
- Two kinds ties
  - type 1: me <–> Richard, Richard <–> Neil, me <–> Neil
  - type 2: Neil –> the mortgage guy
- Granovetter (1973): the "strength of weak ties"
  - While strong ties bond, weak ties bridge.
- Watts and Strogatz (1998) formalize the idea (also see "Six Degrees of Separation")
  - the network is numerically large
  - the network is sparse
  - the network is decentralized
  - the network is highly clustered

A small-world network

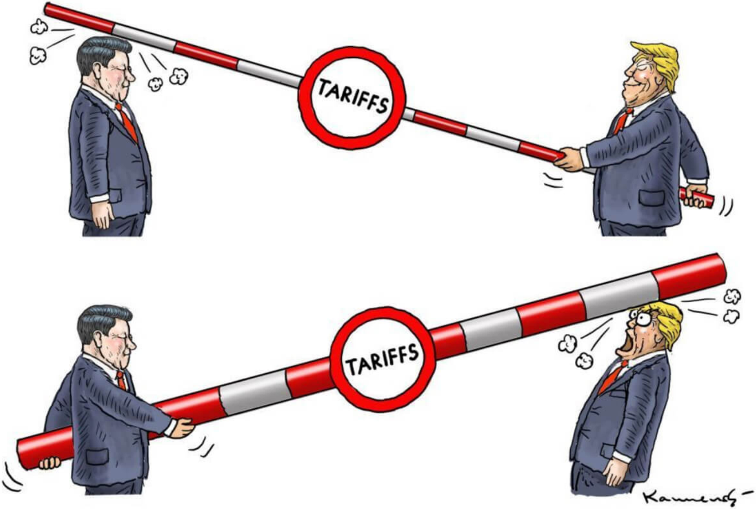


Outline

How to (un)make friends



How to (un)make friends: Axelrod (1984)



Confucius and reciprocity

Analects 15.24 *Zigong asked: "Is there a single teaching that can be practiced to the end of one's life?" Confucius replied: "It is reciprocity! What you don't desire for yourself, do not desire for others."*

Network dependencies

- **Degree differentials** refer to the tendency that some actors are highly connected and others have few connections.

Network dependencies

- **Degree differentials** refer to the tendency that some actors are highly connected and others have few connections.
- **Transitivity** of ties is expressed by the saying "friends of my friends are my friends."

Network dependencies

- **Degree differentials** refer to the tendency that some actors are highly connected and others have few connections.
- **Transitivity** of ties is expressed by the saying "friends of my friends are my friends."
- **Reciprocation of directed ties** is a basic feature of social networks.

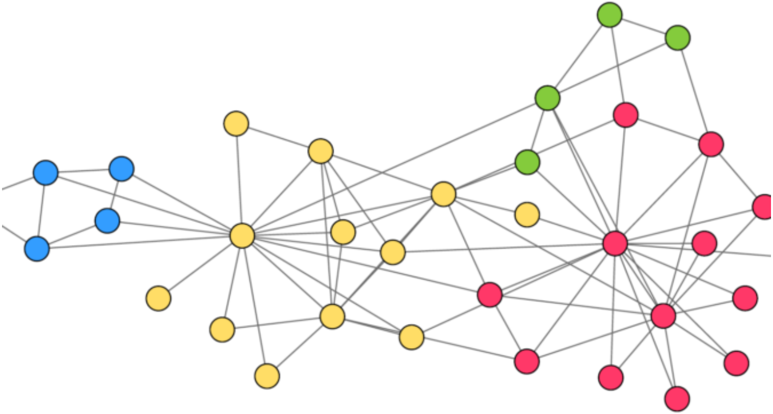
Network dependencies

- **Degree differentials** refer to the tendency that some actors are highly connected and others have few connections.
- **Transitivity** of ties is expressed by the saying “friends of my friends are my friends.”
- **Reciprocation of directed ties** is a basic feature of social networks.
- **Homophily** denotes the tendency of similar actors to relate to each other.

Network dependencies

- **Degree differentials** refer to the tendency that some actors are highly connected and others have few connections.
- **Transitivity** of ties is expressed by the saying "friends of my friends are my friends."
- **Reciprocation of directed ties** is a basic feature of social networks.
- **Homophily** denotes the tendency of similar actors to relate to each other.
- There are many other important types of dependencies.

Bringing back agent attributes: McPherson, Smith-Lovin and Cook (2001)



Outline

How to model network structures

- Incorporating network structure via **covariates**

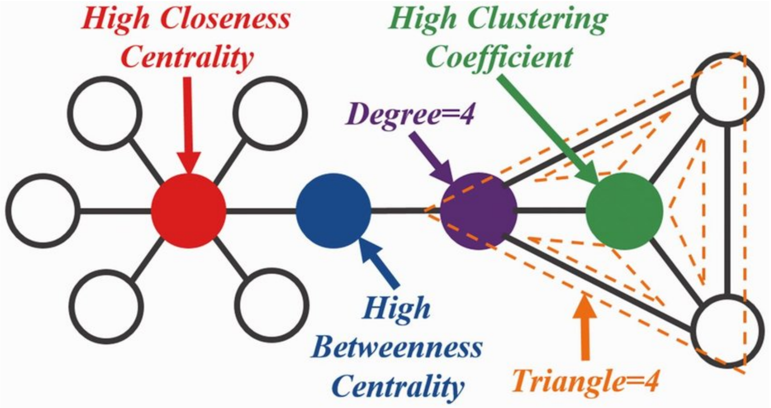
How to model network structures

- Incorporating network structure via **covariates**
- Controlling for certain aspects of network dependencies while not explicitly modeling them (e.g., the quadratic assignment procedure, QAP)

How to model network structures

- Incorporating network structure via **covariates**
- Controlling for certain aspects of network dependencies while not explicitly modeling them (e.g., the quadratic assignment procedure, QAP)
- Modeling network structure

Centrality measures



Outline

Some common models

- Conditionally uniform models

Some common models

- Conditionally uniform models
- Gravity models

Some common models

- Conditionally uniform models
- Gravity models
- Latent space models
  - discrete space
  - distance model
  - sender and receiver effects

Some common models

- Conditionally uniform models
- Gravity models
- Latent space models
  - discrete space
  - distance model
  - sender and receiver effects
- **Exponential random graph models (ERGMs)**

Outline

ERGMs (a.k.a. p^* models)

- ERGMs assume Markov dependence for distributions on network.

ERGMs (a.k.a. p^* models)

- ERGMs assume Markov dependence for distributions on network.
- Conditioning on the other random variables, two random variables are **independent** unless they are tied.

ERGMs (a.k.a. p^* models)

- ERGMs assume Markov dependence for distributions on network.
- Conditioning on the other random variables, two random variables are **independent** unless they are tied.
- A stochastic graph, $X = (X_{ij})$, is a Markov graph if, for each set of four distinct nodes $\{i, j, h, k\}$, the random ties X_{ij} and X_{hk} are conditionally independent, given all the other random variables in X.

ERGMs

- A generalized Markov graph model incorporates any kind of dependence between tie variable (Cranmer and Desmarais, 2010),

$$\Pr(\mathrm{Y} = \mathrm{y}) = \frac{1}{\kappa} \exp\left\{\sum_{A} \eta_A g_A(\mathrm{y})\right\} \tag{1}$$

where,

1. the summation is over all configuration types A; different sets of configuration types represent different models;
2. η_A is the parameter corresponding to configuration of type A;
3. $g_A(\mathrm{y})$ is the *network statistic* corresponding to configuration A (for homogeneous Markov graph models, this is the number of configurations of type A observed in the network: for example, the number of triangles); and
4. κ is a normalizing quantity to ensure that (1) is a proper probability distribution.

Problems

- κ is difficult, if impossible, to be reliably calculated.
- The Markov model is not a reasonable representation for most empirically observed social networks when $N > 30$.
- Degeneracy problems

Problems

- κ is difficult, if impossible, to be reliably calculated.
- The Markov model is not a reasonable representation for most empirically observed social networks when $N > 30$.
- Degeneracy problems
- The complexity of dependencies in networks is so great that modeling large networks in a way that passes the high requirements of a good statistical fit seems intrinsically difficult to achieve (Snijders et al., 2006).

Longitudinal relational data

- Modeling network dynamics is less complicated than modeling single network observations.
  - Dependencies are spread out in time.

Longitudinal relational data

- Modeling network dynamics is less complicated than modeling single network observations.
  - Dependencies are spread out in time.
- Temporal ERGMs (TERGMs)

Longitudinal relational data

- Modeling network dynamics is less complicated than modeling single network observations.
    - Dependencies are spread out in time.
- Temporal ERGMs (TERGMs)
- Stochastic agent-oriented models (SAOMs)

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York, NY: Basic Books.

Barabási, Albert-László and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439):509–512.

Baumgartner, Frank R., Janet M. Box-Steffensmeier and Benjamin W. Campbell. 2018. "Event Dependence in U.S. Executions." *PLOS ONE* 13(1):e0190244.

Cederman, Lars-Erik. 2003. "Modeling the Size of Wars: From Billiard Balls to Sandpiles." *American Political Science Review* 97(1):135–150.

Cranmer, Skyler J. and Bruce A. Desmarais. 2010. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19:66–86.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6):1360–1380.

Jackson, Matthew O. 2010. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

Jackson, Matthew O. and Yves Zenou. 2015. Games on Networks. In *Handbook of Game Theory with Economic Applications*, ed. H. Peyton Young and Shmuel Zamir. Vol. 4 pp. 95–163.

Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy and Marshall Van Alstyne. 2009. "Computational Social Science." *Science* 323(5915):721–723.

McPherson, Miller, Lynn Smith-Lovin and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–444.

Padgett, John F. and Paul D. McLean. 2006. "Organizational Invention and Elite Transformation: The Birth of Partnership Systems in Renaissance Florence." *American Journal of Sociology* 111(5):1463–1568.

Snijders, Tom A. B., Philippa E. Pattison, Garry L. Robins and Mark S. Handcock. 2006. "New Specifications for Exponential Random Graph Models." *Sociological Methodology* 36:99–153.

Watts, Duncan J. and Steven H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393(6684):440–442.