Introduction to Network Analysis Workshop

Basak Taraktas

Bogazici University

March 4, 2024

Northwestern University

Graph theory

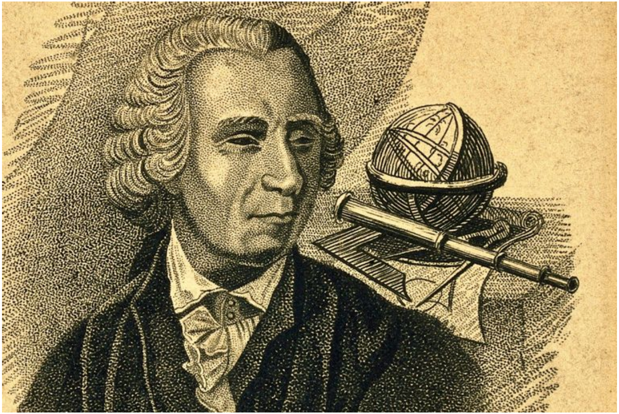


Figure: Euler

Seven Bridges of Konigsberg Problem

Figure: Bridges of Konigsberg

https://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg

Why care about the bridge problem in social sciences?

- Structure of relationships (power, hierarchy, institutional structure...)
- Devising a strategy to navigate the institutional structure
- Network effects (homophily, transitivity, diffusion, cascade...)

Dynamics of political polarization (2007) [1]

The first paradox is the simultaneous absence and presence of attitude polarization; the fact that global attitude polarization is relatively rare, even though pundits describe it as common.

The second paradox is the simultaneous presence and absence of social polarization; the fact that while individuals experience attitude homogeneity in their interpersonal networks, their networks are characterized by attitude heterogeneity.

[1] Delia Baldassarri and Peter Bearman, Dynamics of Political Polarization, American Sociological Review 72, no. 5 (2007): 784–811.

Dynamics of Political Polarization

What is the underlying mechanism?

“In general, people interact with people who are similar to them. There are strong pressures toward *homophily* in social relations.”

“At the same time, it happens that (...) Democrats know Republicans and rich people fall in love with not-so-wealthy ones. (...) People are also *in touch with people different from them.*” (*heterophily* (Baldassarri & Bearman, 2007))

The Snowball effect

https://one.npr.org/i/972648402:972648404

Modeling in social sciences

Statistical analysis

- Large samples
- Unit homogeneity
- Statistical independence

Game theory

- Interdependent decision making
- A few or n individuals

Modeling in social sciences

Statistical analysis

- Large samples
- Unit homogeneity
- Statistical independence

Game theory

- Interdependent decision making
- A few or n individuals

What about network analysis?

Network analysis studies *interdependencies* –relational data capturing relationships and flows between entities, which can be people, groups, organizations, etc.

Network analysis encompasses both empirical studies based on real-world data and theoretical inquiries with abstract network models.

Assumptions of network analysis

- **Embeddedness:** the environment structures behavior, preferences, beliefs, and perceptions. The environment, in turn, is structured by actors and patterns of relationships
- **Interdependence:** one's behavior is affected by the behavior of those who are in one's entourage (e.g., taking up smoking, logic of appropriateness...)

$\Rightarrow$ Network analysis studies *dynamic behavioral data*!

What are networks?

An informal definition

"A network is **a collection of points joined together** in pairs by lines." (Newman, 2010).

The formal definition

$$G = (V, E)$$

where G represents a graph (network), V a set of vertices and E edges between vertices with

$$V = \{v_1, v_2, \ldots, v_n\}$$

What are networks?

An informal definition

"A network is **a collection of points joined together** in pairs by lines." (Newman, 2010).

The formal definition

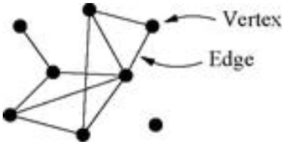
$$G = (V, E)$$

where G represents a graph (network), V a set of vertices and E edges between vertices with

$$V = \{v_1, v_2, \ldots, v_n\}$$

Network elements

- Nodes (vertex): entities
- Edges (ties): relationships
    - Directed
    - Undirected



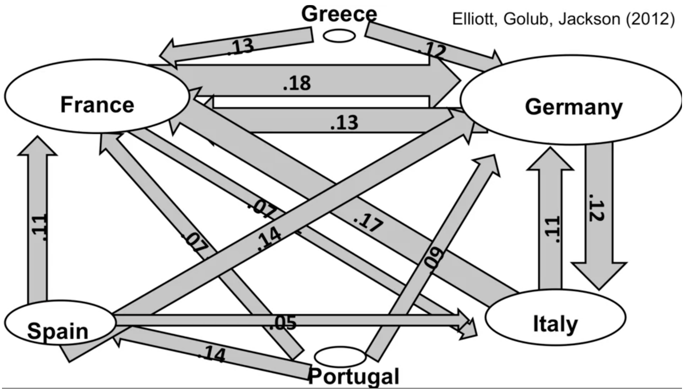
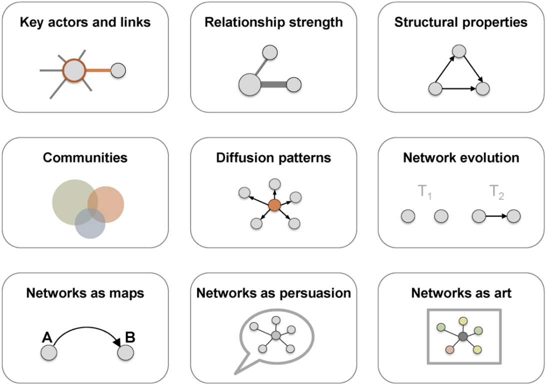


Figure: (Elliott et al., 2014)

Focus of your research



Questions of network analysis

Network as DV

How does a particular *structure of relationship* come about and evolve? (e.g., segregation, cooperation, coalition formation)

Network as IV

What *role* do networks play in the observed outcome? (e.g., socialization, quitting smoking, finding jobs)

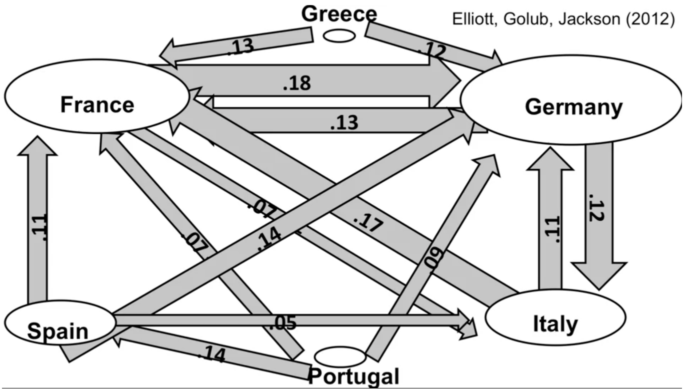
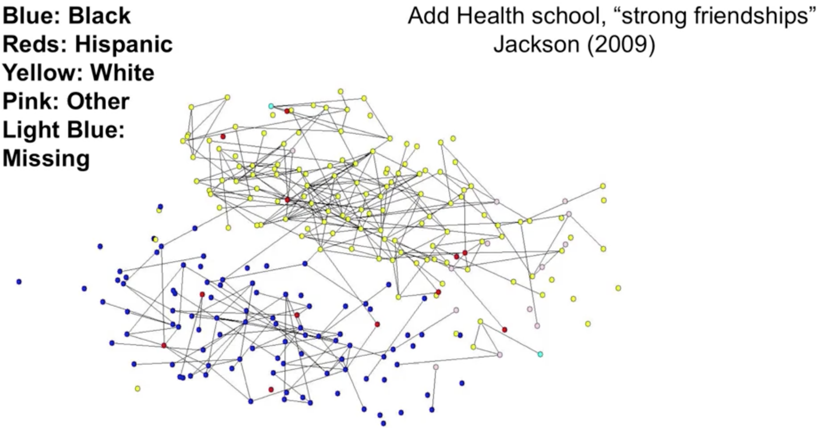


Figure: (Elliott et al., 2014)



(Jackson, 2009)

Data

- **Unit of analysis:** dyads, triads, subgroups of networks, or an entire network
- **Data:** observational data, archives...
- **Data size:** moderate (e.g., classroom), large (e.g., CTA stops in Chicago), big data (e.g., social media)
- **Data format:** adjacency matrices

Data

- **Unit of analysis:** dyads, triads, subgroups of networks, or an entire network
- **Data:** observational data, archives...
- **Data size:** moderate (e.g., classroom), large (e.g., CTA stops in Chicago), big data (e.g., social media)
- **Data format:** adjacency matrices

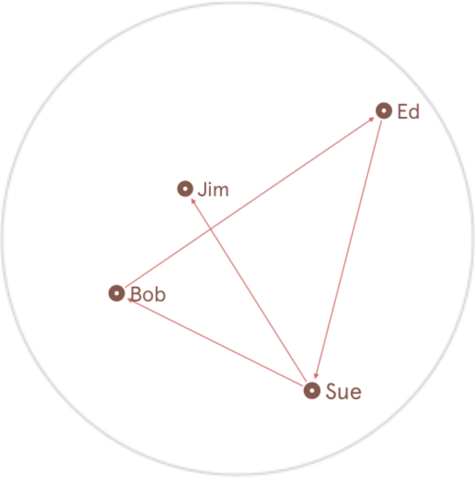
Data

- **Unit of analysis:** dyads, triads, subgroups of networks, or an entire network
- **Data:** observational data, archives...
- **Data size:** moderate (e.g., classroom), large (e.g., CTA stops in Chicago), big data (e.g., social media)
- **Data format:** adjacency matrices

Data

- **Unit of analysis:** dyads, triads, subgroups of networks, or an entire network
- **Data:** observational data, archives...
- **Data size:** moderate (e.g., classroom), large (e.g., CTA stops in Chicago), big data (e.g., social media)
- **Data format:** adjacency matrices

What does relational data look like?

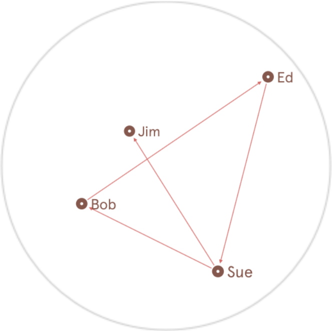


What does relational data look like?

Adjacency Matrix

Friendship

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



Adjacency matrix

The adjacency matrix is an algebraic representation of the graph structure.

Network of n nodes as an $n \times n$ matrix

$$A = [A_{ij}]$$

where

$$A_{ij} = \begin{cases} 1 & \text{if node } i \text{ is connected to node } j; \\ 0 & \text{otherwise} \end{cases}$$

Let's build a network

https://github.com/enaline/northwestern-network-analysis-workshop

Degree & average degree

Degree

The number of edges incident upon a node.

- Undirected network: degree = number of connections a node has
- Directed network: in-degree= number of incoming edges & out-degree= number of outgoing edges

.

Average degree

Sum of all degrees / Total number of vertices

$$\bar{k} = \frac{\sum_{i=1}^{n} k_i}{n}$$

with k representing degree

Density

Density

The ratio of observed edges to all possible edges. Often used to make an argument about power and patterns of connections. A density score close to 0 implies a sparse network.

Undirected graphs

$$\text{density} = \frac{2m}{n(n-1)}$$

Directed graphs

$$\text{density} = \frac{m}{n(n-1)}$$

Centrality

Degree centrality

Centrality is calculated based on the degree of a node.

$$\text{Degree centrality} = \frac{\text{Number of edges connected to the node}}{\text{Total number of nodes} - 1}$$

$\Rightarrow$ *Measure of prestige and prominence.*

Centrality (cont'd)

Eigenvector centrality

Centrality is defined by the number of important neighbors a node has. $\Rightarrow$ Great for undirected networks!

$$x_i = \frac{1}{\lambda} \sum_k a_{k,i}\, x_k$$

where λ is the eigenvector with the largest eigenvalue

Centrality

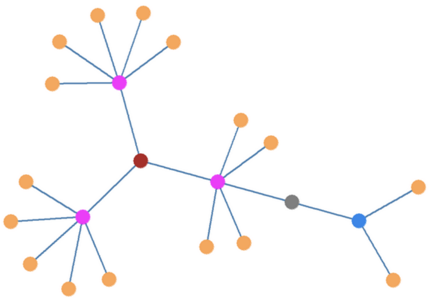


Figure 7.4: An undireced graph with degree centrality indicated by node color.

Centrality (cont'd)

degree centrality

Mag	Mag	Mag	Red	Blue	Gray	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan
6	6	6	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

eigenvector centrality

Mag	Mag	Mag	Red	Blue	Gray	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan	Tan
0.99	0.93	0.93	1	0.2	0.42	0.35	0.35	0.35	0.35	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.33	0.07	0.07

Centrality (cont'd)

Betweenness centrality

The extent to which a vertex lies on paths between other vertices. ⇔ fraction of the total number of geodesic paths between two nodes. Bridging (e.g., Brokers)!

$$\text{Betweenness centrality}(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where:

- v is a vertex in the graph,
- σ_{st} is the total number of shortest paths from node s to node t,
- $\sigma_{st}(v)$ is the number of those paths that pass through vertex v.

References I

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, *72*(5), 784–811. https://doi.org/10.1177/000312240707200507

Elliott, M., Golub, B., & Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, *104*(10), 3115–53. https://doi.org/10.1257/aer.104.10.3115

Jackson, M. (2009). Networks and economic behavior. *Annual Review of Economics*, *1*, 489–513. https://doi.org/10.1146/annurev.economics.050708.143238

Newman, M. E. J. (2010). *Networks: An introduction*. Oxford Univ. Press.

THANK YOU FOR LISTENING!

Contact information

Basak TARAKTAS

basak.taraktas@boun.edu.tr
basakt@sas.upenn.edu
https://basaktaraktas.com