Social Network Analysis

Basic concepts

- Node (vertex): each of the units in the network

Basic concepts

- Node (vertex): each of the units in the network
- Edge (tie): connection between nodes

Basic concepts

- Node (vertex): each of the units in the network
- Edge (tie): connection between nodes
  - Undirected: symmetric connection, represented by lines

Basic concepts

- Node (vertex): each of the units in the network
- Edge (tie): connection between nodes
  - Undirected: symmetric connection, represented by lines
  - Directed: imply direction, represented by arrows

Basic concepts

- Node (vertex): each of the units in the network
- Edge (tie): connection between nodes
  - Undirected: symmetric connection, represented by lines
  - Directed: imply direction, represented by arrows
- A network consists of a set of nodes and edges

Basic concepts

A few examples:

- Classroom: students / friendships

Basic concepts

A few examples:

- Classroom: students / friendships
- Twitter: users / retweets

Basic concepts

A few examples:

- Classroom: students / friendships
- Twitter: users / retweets
- Academic literature: papers / citations

Basic concepts

A few examples:

- Classroom: students / friendships
- Twitter: users / retweets
- Academic literature: papers / citations
- Internet: websites / hyperlinks

Basic concepts

A few examples:

- Classroom: students / friendships
- Twitter: users / retweets
- Academic literature: papers / citations
- Internet: websites / hyperlinks
- Trade: countries / trade flows

Basic concepts

A few examples:

- Classroom: students / friendships
- Twitter: users / retweets
- Academic literature: papers / citations
- Internet: websites / hyperlinks
- Trade: countries / trade flows
- Biology: neurons / connections

Basic concepts

Network Visualization

Adjacency Matrix

	J	J	E	W	T
J	0	1	1	0	0
J	1	0	0	1	1
E	1	0	0	1	0
W	0	1	1	0	1
T	0	1	0	1	0

Basic concepts

Network Visualization

Edgelist

	Node1	Node2
1	Jennifer	Josh
2	Jennifer	Evgeniia
3	Josh	Whitney
4	Josh	Tom
5	Whitney	Tom
6	Evgeniia	Whitney

Latent structure of social networks

Latent structure of social networks

The dreaded *hairball*

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)
  - Also hierarchy: core-periphery detection

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)
  - Also hierarchy: core-periphery detection
- Locating nodes on latent spaces

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)
  - Also hierarchy: core-periphery detection
- Locating nodes on latent spaces
  - **Latent space models of networks**

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)
  - Also hierarchy: core-periphery detection
- Locating nodes on latent spaces
  - **Latent space models of networks**
  - Proximity on latent space (ideology) predicts existence of edges

Discovery in large-scale networks

How to understand the structure of large-scale networks?

- Latent communities or clusters
  - **Community detection algorithms**
  - Finding groups of nodes that densely connected internally, more so than to the rest of the networks
  - Overlap with shared visible or latent similarities (homophily)
  - Also hierarchy: core-periphery detection
- Locating nodes on latent spaces
  - **Latent space models of networks**
  - Proximity on latent space (ideology) predicts existence of edges
  - Inference about latent positions based on multidimensional scaling of the adjacency matrix

Community detection

Community structure:

- Network nodes often cluster into tightly-knit groups with a high density of within-group edges and a lower density of between-group edges
- **Modularity score**: measures clustering of nodes compared to random network of same size
- Many different community detection algorithms based on different assumptions

Source: Newman (2012)

Network hierarchy

- **Intuition**

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals (not captured by simple topological measures)

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals (not captured by simple topological measures)
- **k-core decomposition**

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals (not captured by simple topological measures)
- **k-core decomposition**
  - Algorithm to partition a network in nested shells of connectivity

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals (not captured by simple topological measures)
- **k-core decomposition**
  - Algorithm to partition a network in nested shells of connectivity
  - The k-core of a graph is the maximal subgraph in which every node has at least degree k

Network hierarchy

- **Intuition**
  - Large-scale networks have hierarchical properties
- **Network core:**
  1. *Centrality*: high relative importance in network
  2. *Connectivity*: many possible distinct paths between individuals (not captured by simple topological measures)
- **k-core decomposition**
  - Algorithm to partition a network in nested shells of connectivity
  - The k-core of a graph is the maximal subgraph in which every node has at least degree k
  - Many applications; scales well to large networks: $\mathcal{O}(n + e)$

k-core decomposition

Source: Alvarez-Hamelin et al, 2005

k-core decomposition

Source: Alvarez-Hamelin et al, 2005

k-core decomposition of #OccupyGezi network

Relative importance of core and periphery

reach: aggregate size of participants' audience

activity: total number of protest messages published (not only RTs)