



Social Network Analysis

School of Information Studies
Syracuse University

| Social Network Analysis (SNA)

“Study of human relationships using graph theory”

Used to analyze people or collections of people

Origins in social sciences and network analysis

Unique perspective on how society functions

Focuses on relations between individuals, groups, and social institutions

| Why and When to Use SNA

When studying a social network

Want to visualize data to uncover patterns

Want to follow paths that information follows

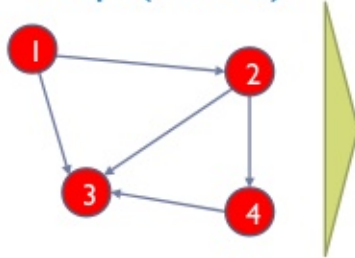
When doing quantitative research

Analyzing SNSs, OCs, and social media

Directed Graph of Networks

Entering data on a directed graph

Graph (directed)



Edge list

Vertex	Vertex
1	2
1	3
2	3
2	4
3	4

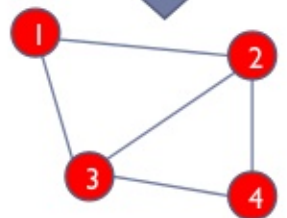
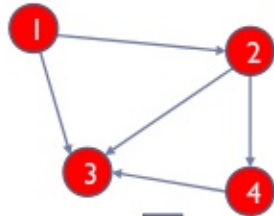
Adjacency matrix

Vertex	1	2	3	4
1	-	1	1	0
2	0	-	1	1
3	0	0	-	0
4	0	0	1	-

Undirected Graph

Representing an undirected graph

Directed
(who contacts whom)



Undirected
(who knows whom)

Edge list remains the same

Vertex	Vertex
1	2
1	3
2	3
2	4
3	4

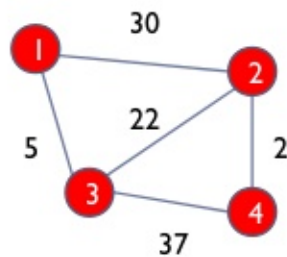
But interpretation is different now

Adjacency matrix becomes symmetric

Vertex	1	2	3	4
1	-	1	1	0
2	1	-	1	1
3	1	1	-	1
4	0	1	1	-

Strong/Weak Ties in Network

Adding weights to edges (*directed or undirected*)



Weights could be:

- Frequency of interaction in period of observation
- Number of items exchanged in period
- Individual perceptions of strength of relationship
- Costs in communication or exchange, e.g. distance
- Combinations of these

Edge list: add column of weights

Vertex	Vertex	Weight
1	2	30
1	3	5
2	3	22
2	4	2
3	4	37

Adjacency matrix: add weights instead of 1

Vertex	1	2	3	4
1	-	30	5	0
2	30	-	22	2
3	5	22	-	37
4	0	2	37	-

Key Players

Centrality Measure	Possible Interpretation
*Degree	Music Collaborations: How Many People?
*Betweenness	Spies: Who Collects the Most Information?
*Closeness	Sexual: How Fast Will an STD Spread?
*Eigenvector	Paper Citations: Which Author is Most Cited?

Network's Structure: Cohesion

Density

- Number of edges over total possible edges
- Measure of well-connected

Clustering

- Closed nodes as compared with all triplets
- Identify communities within networks

Preferential Attachment

Popularity

- Associated with popular people, ideas, and so on

Quality

- Objective criteria
- Higher quality attracts attention

Mixed model

- Nodes of similar attributes
- Reach “critical” mass first—STARS



Showing Geolocations on Maps

School of Information Studies
Syracuse University

Geolocations in Twitter

User has location in profile

User has “geo-enabled” True

- Location appears in tweet
- Given as longitude and latitude location
- Could be called “coordinates”

Twitter: given in order [longitude, latitude]

Other apps: given in order [latitude, longitude]

Reverse Geocoding

May want to summarize locations

Getting address information back

- Google Maps API
- Foursquare package
- Python package geopy
 - Geopy – <https://geopy.readthedocs.io/en/1.10.0/>
 - To install: `pip install geopy`
 - Geocoding and reverse geocoding

| Reverse Geocoding - Example

Given Latitude and Longitude for Syracuse University Hinds Hall

- Latitude: 43.038658 Longitude: -76.133498
- Use this to retrieve the location from the google api
- Parse the data to retrieve city and state from the 'short_name'
- Use the address to get the latitude and longitude
- Do the reverse (use latitude and longitude) to get address

Creating Maps

KML maps (Keyhole Markup Language)

- Used by Google Maps and Google Earth
- Used user location and how many tweets
- Put PlaceMarks on the map

KML maps (grouping data)

- User locations reverse geocoded
- Saved by city names and frequencies
- Key shows city names and frequencies

Google Chart API

- Region chart for states in the United States
- Can use satellite imagery
- Shows state locations with frequencies of tweets