

POIR 613: Computational Social Science

Pablo Barberá

School of International Relations
University of Southern California
pablobarbera.com

Course website:
pablobarbera.com/POIR613/

Today

1. Announcements
 - ▶ Next milestone for project: 5-page summary that includes some data analysis by Nov 4th
 - ▶ No lecture next week (Nov 6)
2. Social network analysis
 - ▶ Motivation
 - ▶ Applications
 - ▶ Key concepts
 - ▶ Centrality
3. Solutions to challenge 8
4. Network analysis using R
5. Network visualization using Gephi

Introduction to social network analysis

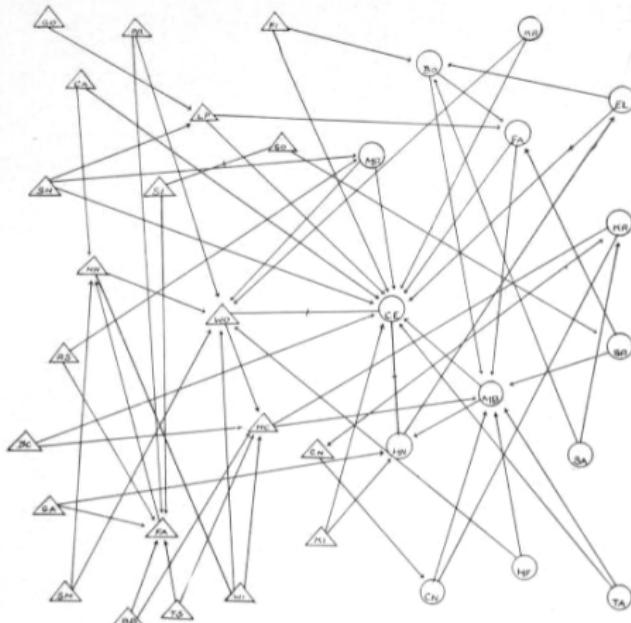


Human behavior is characterized by **connections to others**



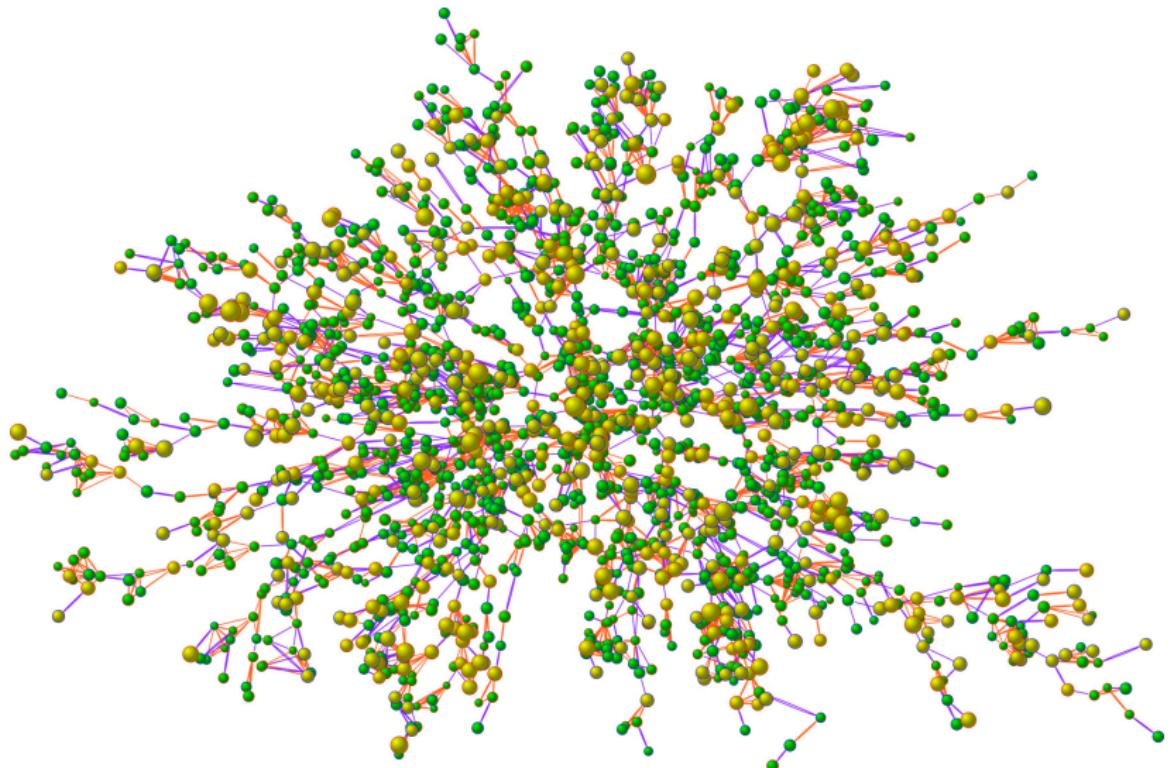
Digital technologies have led to an explosion in the availability of networked data

EVOLUTION OF GROUPS

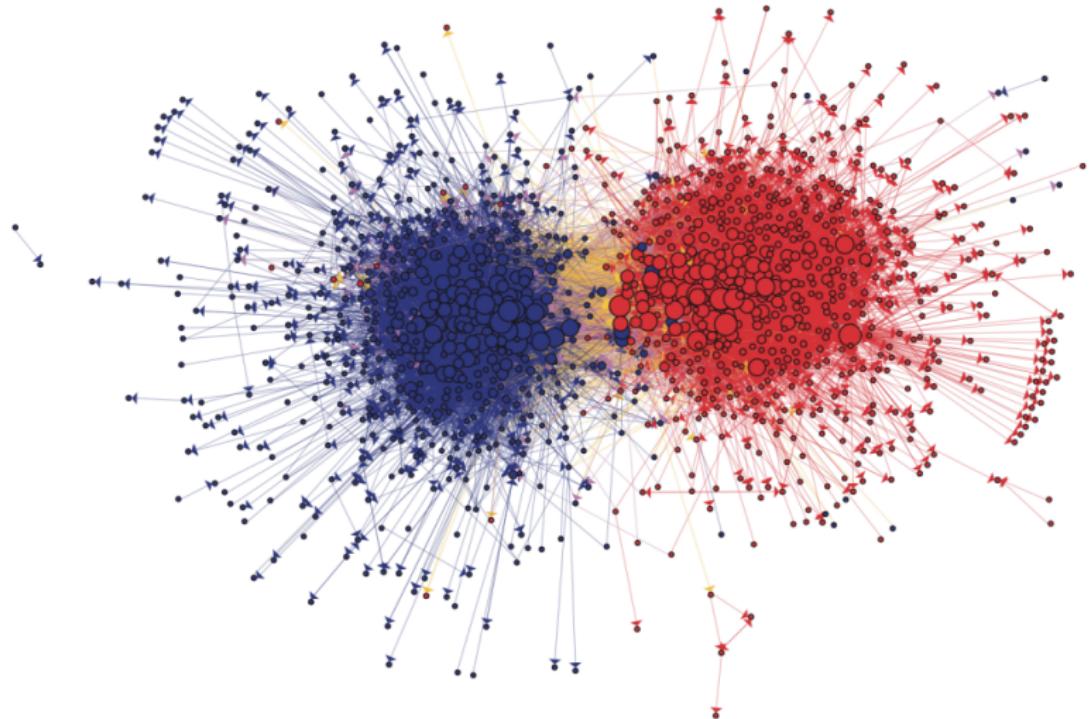


CLASS STRUCTURE, 1ST GRADE

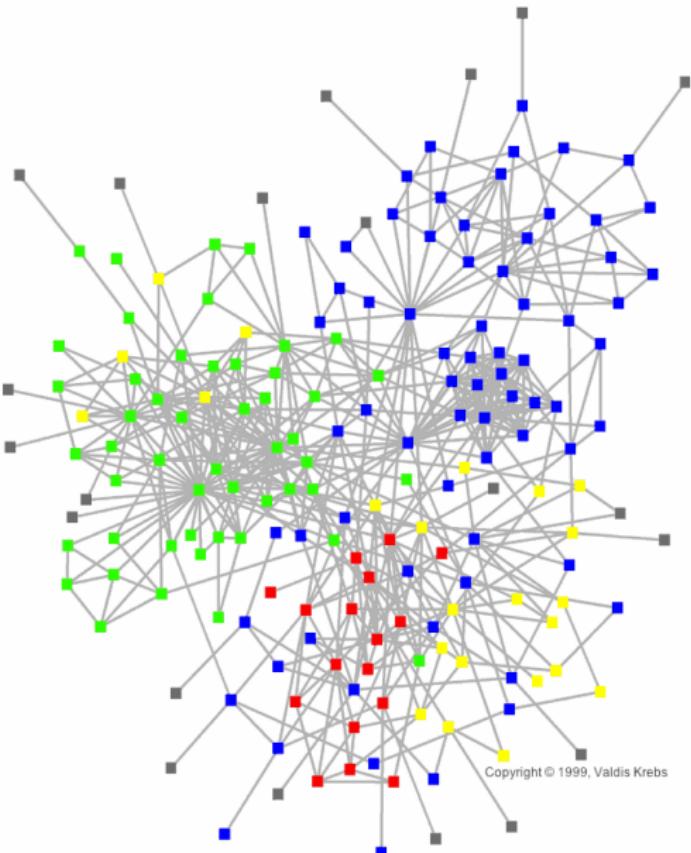
21 boys and 14 girls. *Unchosen*, 18, GO, PR, CA, SH, FI, RS, DC, GA, SM, BB, TS, WI, KI, TA, HF, SA, SR, KR; *Pairs*, 3, EI-GO, WO-CE, CE-HN; *Stars*, 5, CE, WO, HC, FA, MB; *Chains*, 0; *Triangles*, 0; *Inter-sexual Attractions*, 22.



Christakis & Fowler, NEJM, 2007

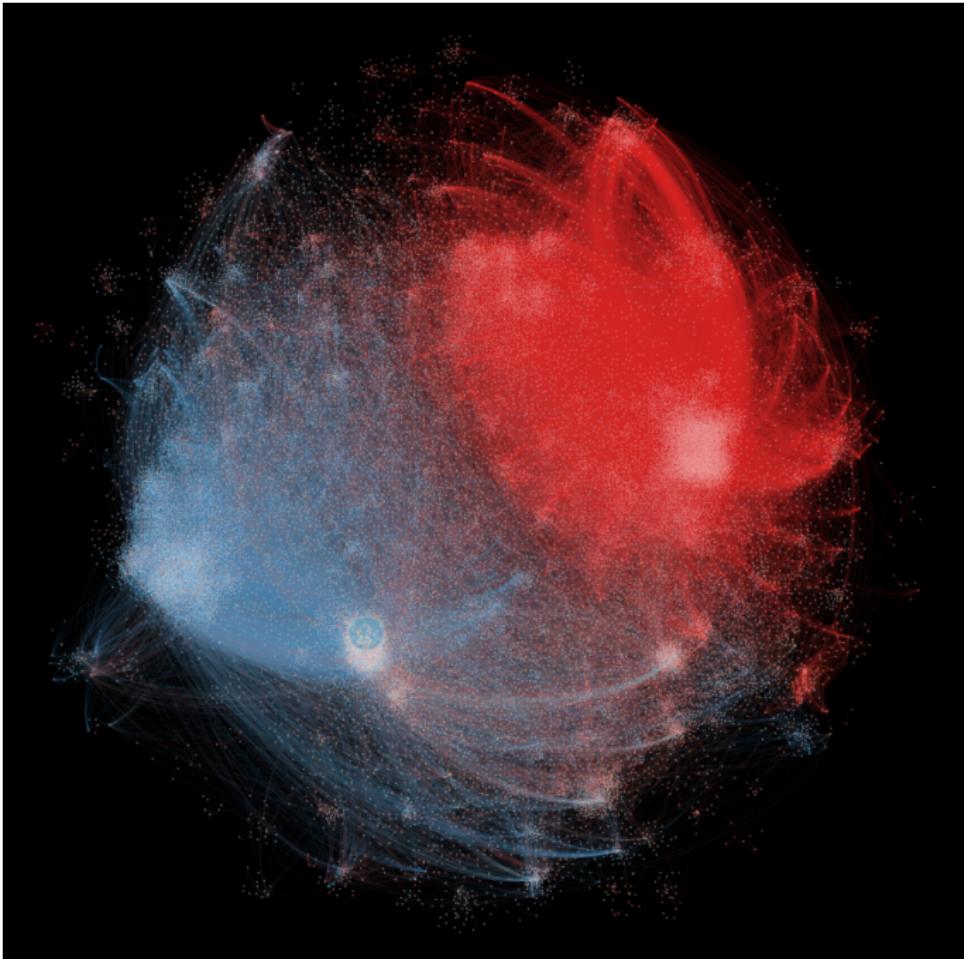


Adamic & Glance, 2004, IWLD

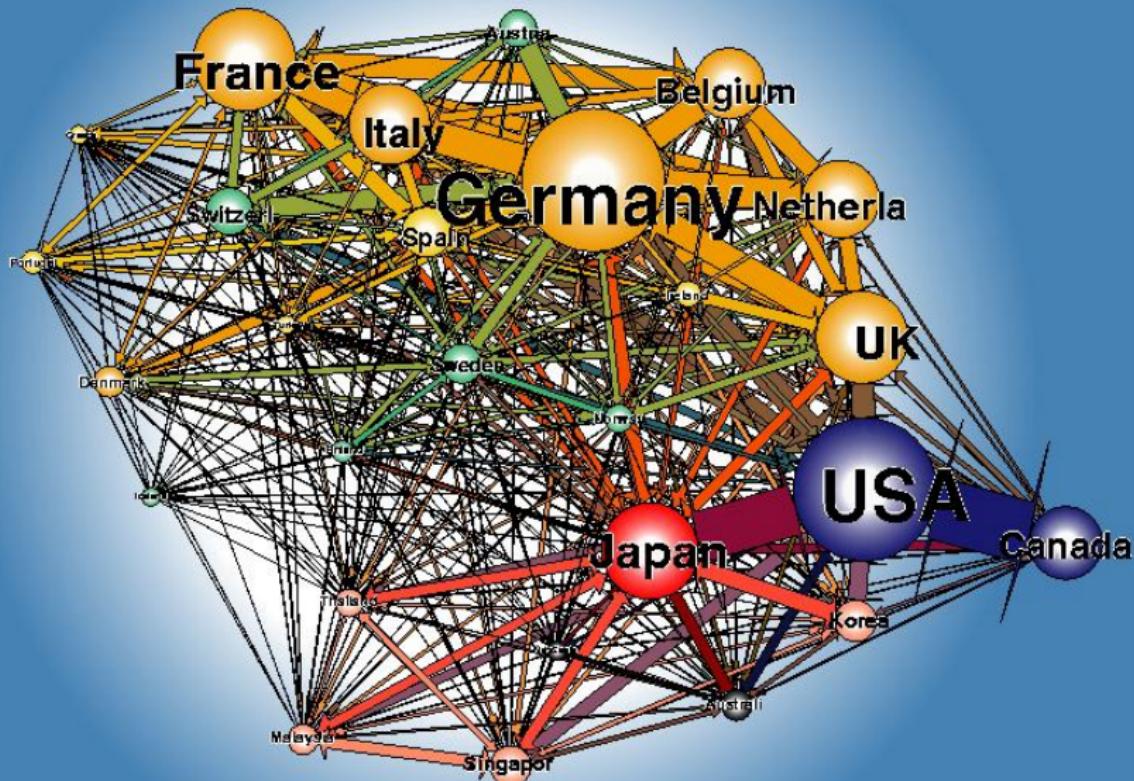


Copyright © 1999, Valdis Krebs

Email network of a company



Barbera et al, 2015, Psychological Science

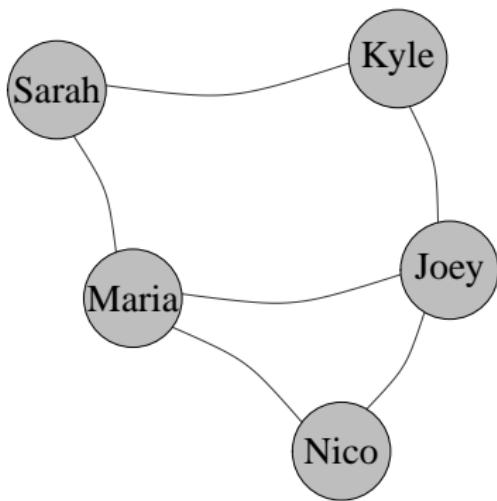


Basic concepts

- ▶ **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
 - ▶ Unweighted: all edges have same strength
 - ▶ Weighted: some edges have more strength than others
- ▶ A **network** consists of a set of nodes and edges
i.e. a set of actors and their relationships

Basic concepts

Network Visualization

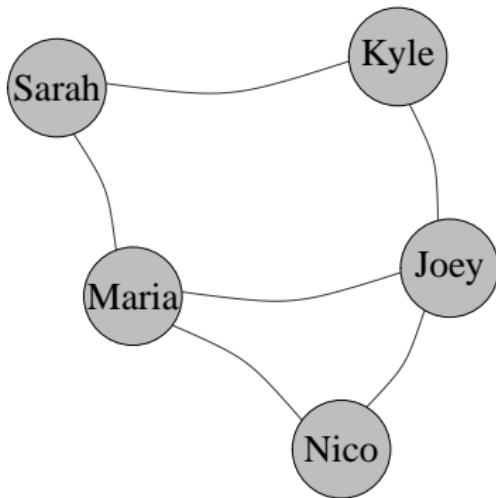


Adjacency Matrix

	K	J	S	M	N
K	0	1	1	0	0
J	1	0	0	1	1
S	1	0	0	1	0
M	0	1	1	0	1
N	0	1	0	1	0

Basic concepts

Network Visualization



Edgelist

	Node1	Node2
1	Kyle	Joey
2	Kyle	Sarah
3	Joey	Maria
4	Joey	Nico
5	Maria	Nico
6	Sarah	Maria

Networks everywhere

- ▶ Academic literature: papers / citations
- ▶ Internet: websites / hyperlinks
- ▶ Twitter: users / retweets
- ▶ Power grid: plants, transformers / cables
- ▶ Biology: neurons / connections
- ▶ Text: documents / cosine similarity
- ▶ Hollywood: actors / playing in same movie

Intellectual and societal impact of networks

Why do networks matter?

- ▶ **Economic impact:** most successful companies in 21st century base their technology and business model on networks
- ▶ **Health:** importance of networks in molecular biology, spread of human diseases, pharmacology
- ▶ **Fighting terrorism:** network-centric warfare can disrupt the financial networks of terrorist organizations and map adversarial networks
- ▶ **Epidemics:** role of transportation networks in the spread of viruses
- ▶ **Neuroscience:** the human brain, one of the least-understood networks
- ▶ **Political behavior:** voting is contagious; opinion formation as a social process

Social network analysis: key dimensions of analysis

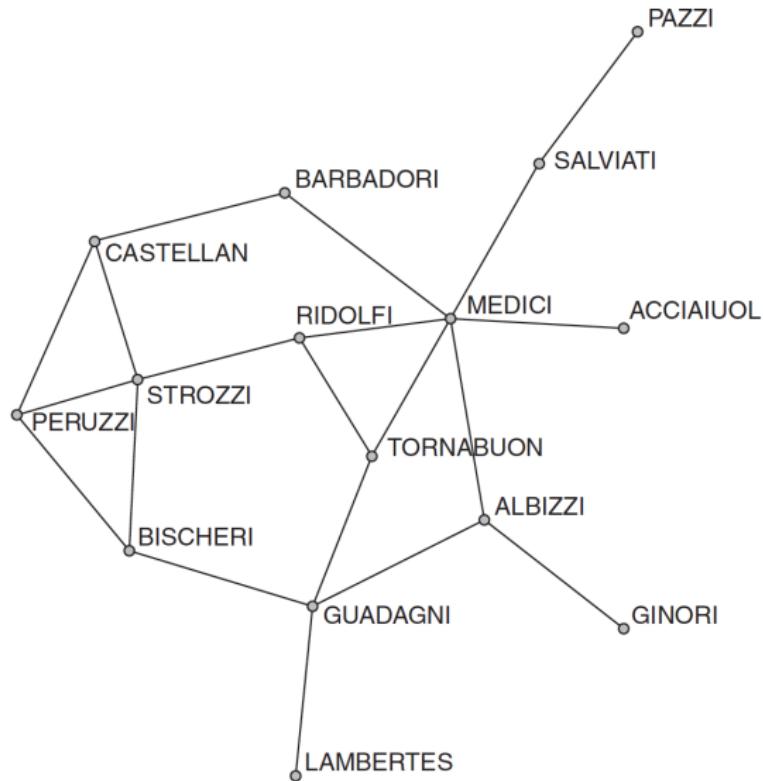
Node centrality

How to measure actor influence or importance in a network?

Two main conceptual definition of **centrality**:

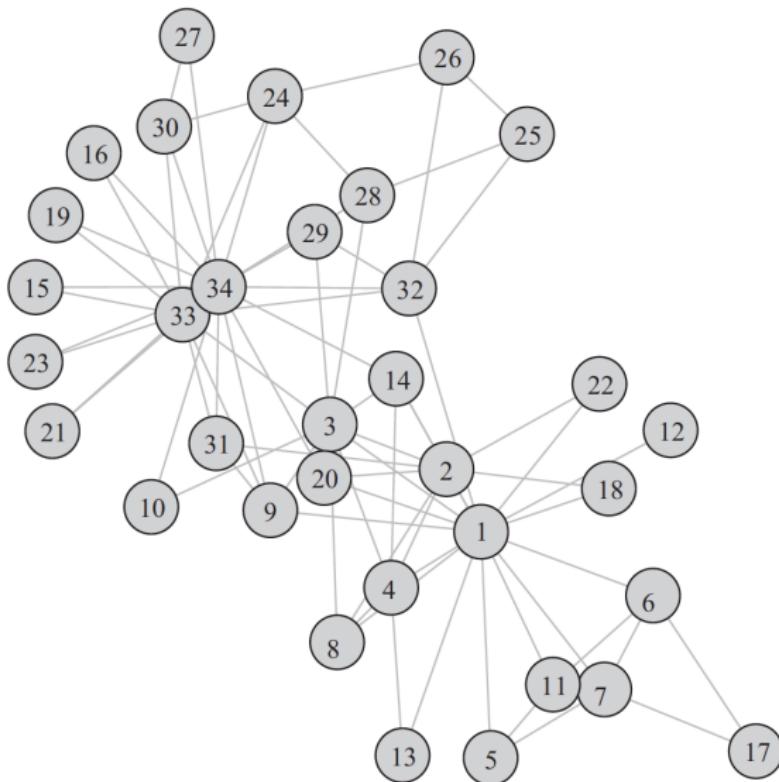
1. **Degree centrality**: number of connections for each node (potential for direct reach)
 - ▶ Indegree: incoming connections
 - ▶ Outdegree: outgoing connections
 2. **Betweenness centrality**: gatekeeping potential
 - ▶ How well a node connects different parts of the network
 - ▶ Fraction of shortest paths between any two nodes on which a particular node lies
- Other measures:
- ▶ **Closeness centrality**: broadcasting potential
 - ▶ **Eigenvector centrality and coreness**: centrality measured as being connected to other central neighbors

Florentine family marriages in the 15th century



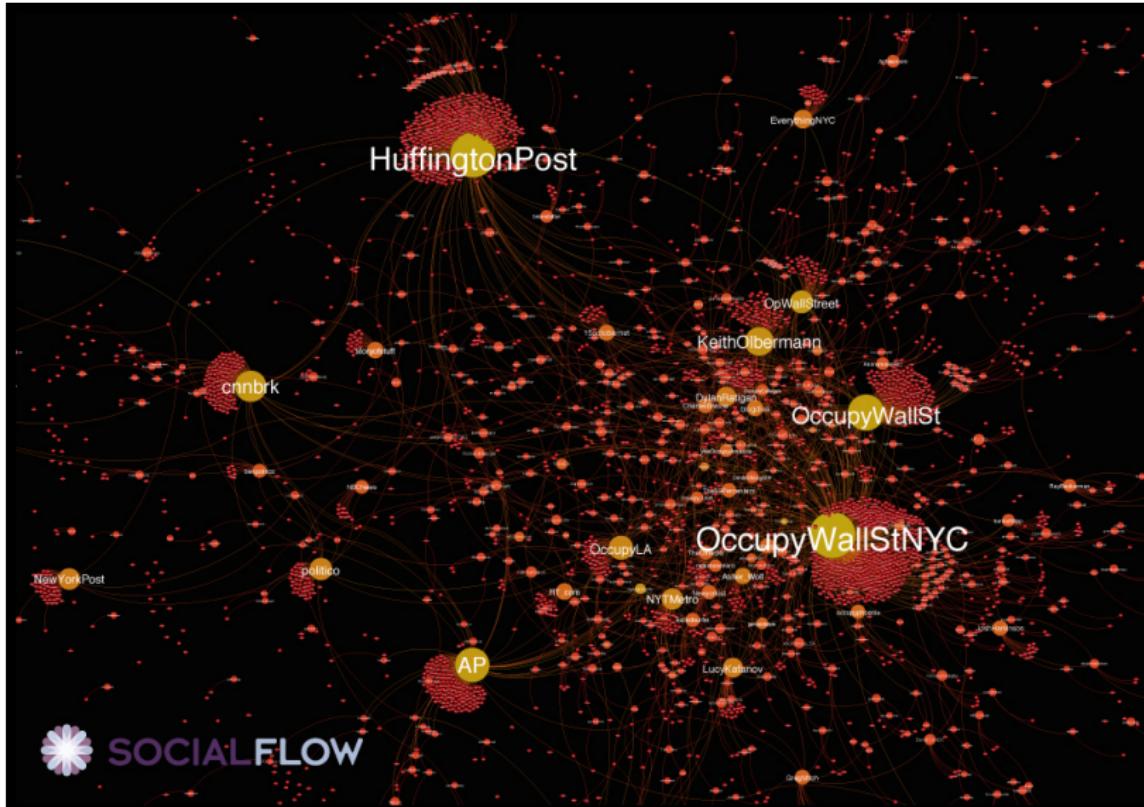
Source: Padgett (1993) and Sinclair (2016)

Social activities in a Karate club



Source: Zachary (1997) and Sinclair (2016)

Occupy Wall Street Twitter networks

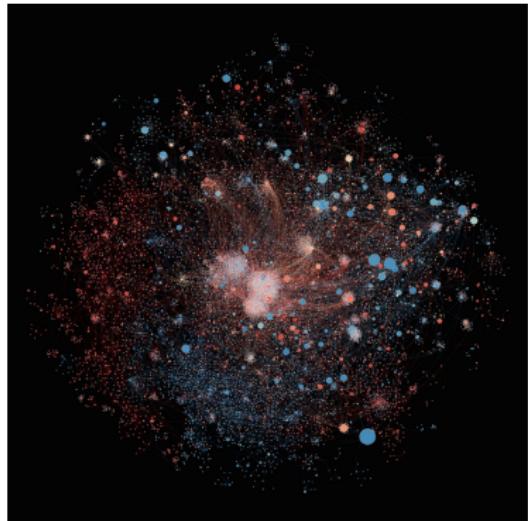


SOCIALFLOW

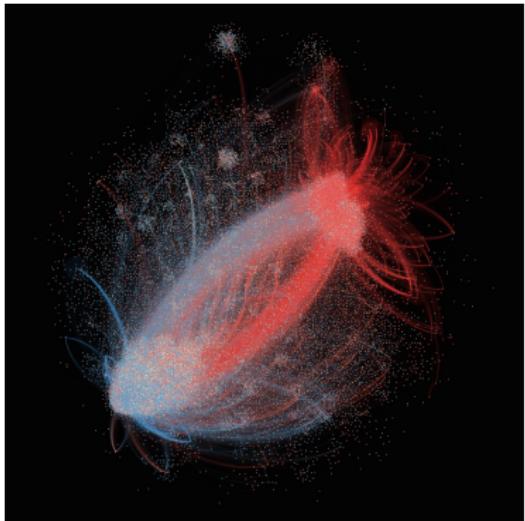
Source: Lotan (2011)

Social network analysis: network properties

Mechanisms explaining network structure



2013 SuperBowl



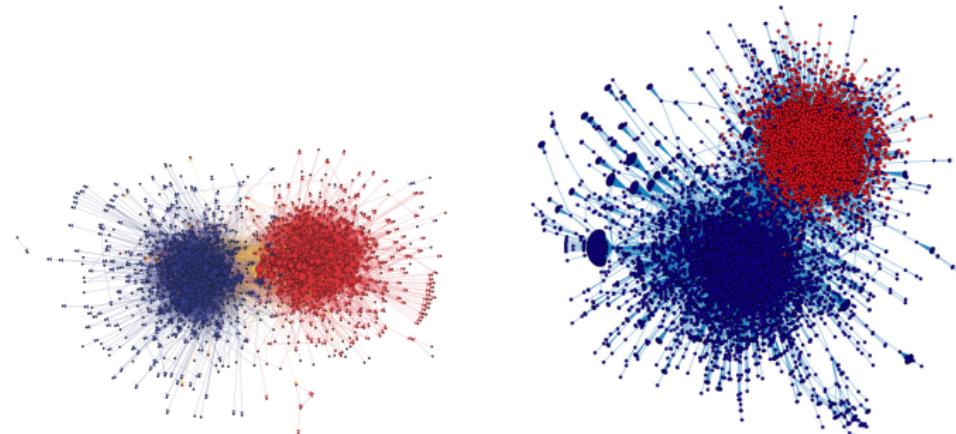
2012 Election

Barberá et al (2015) "Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber?" *Psychological Science*

Mechanisms explaining network structure

Three key mechanisms behind tie formation:

1. **Homophily**: propensity to form ties based on shared traits (age, gender, ideology, location...)



Adamic and Glance (2005)

Conover et al (2012)

2. **Reciprocity**: propensity of directed ties to be mutual
e.g. in **holiday gifts exchanges**, if you observe A giving a gift to B, then you're also likely to see B giving a gift back to A.
3. **Clustering** (aka *transitivity*): propensity of triads to close
e.g. friendship networks grow via introductions: if James is friends with Marcia, and Marcia is friends with José, then