http://arcs.di.unito.it

http://www.di.unito.it/~ruffo

giancarlo.ruffo@unito.it

@giaruffo

Giancarlo Ruffo *- Università degli Studi di Torino (Italy)*

The Science of (Fighting) Fake News

Using network science to model, analyze, and mitigate misinformation diffusion in social media

June 14-18th, 2021

What I do (and don't….)

- Academic and industrial research
- Data and network analysis
- Models of diffusion processes
- Social media and data as a resource
  - the interplay between 'segregation' and 'polarization'
  - rational motivations

- I don't debunk, I am not a journalist
- I don't look for automatic identification of true and false news
- I do not target social media as evil
  - I don't believe in censorship or freedom of speech limitations
  - I don't look for simple explanations to complex problems (e.g., gullible people is also stupid!)

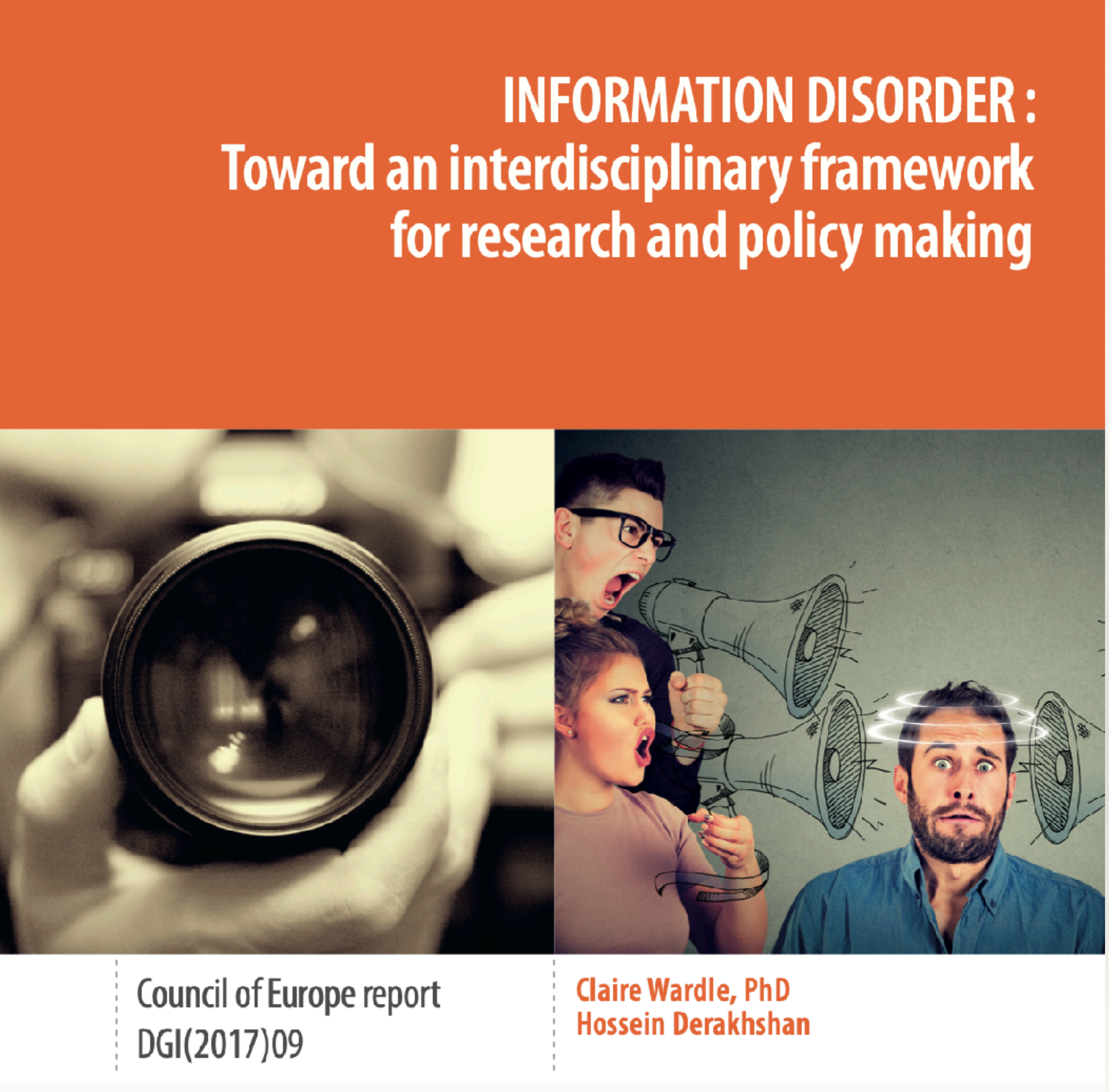
Course overview

- June 15th:
  - Problem definition and basic terminology
  - Introduction to Network Science
  - Understanding the structure of an information/misinformation network
- June 16th:
  - Introduction to dynamical processes on Networks
  - Social influence, the emergence of echo chambers and the interplay between segregation and polarization
  - Studying the impact of fact-checking
- June 17th:
  - The role of social bots
  - Open Problems and Trends

Introduction and Terminology

Terminology

- Misinformation
- Malinformation
- Fake-News
- Disinformation
- Unverified Information
- Propaganda
- Conspiracy Theories
- Urban Legend
- Rumors
- Astroturf
- Spam
- Troll
- Hate Speech
- Cyberbullying



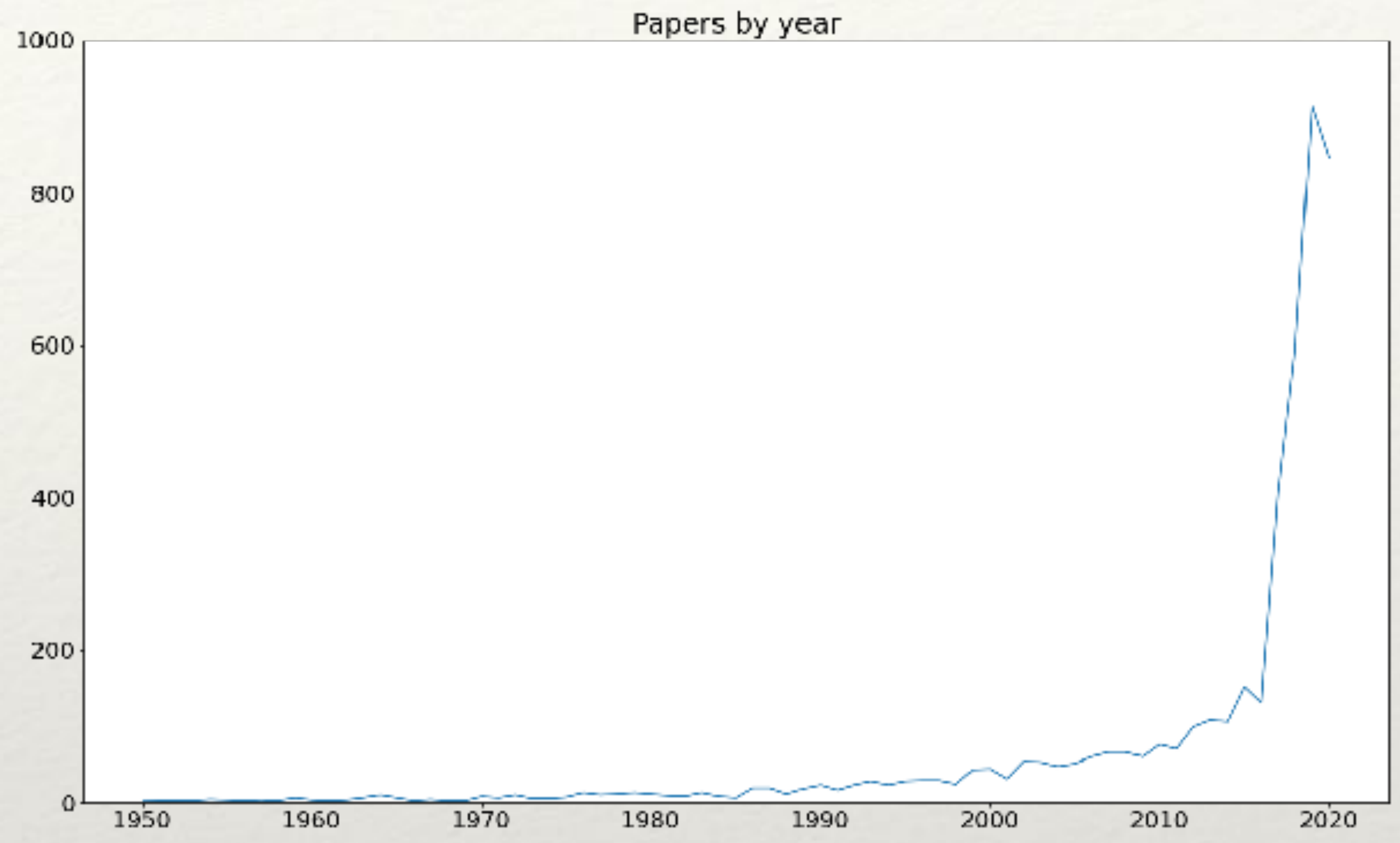
Open fronts:

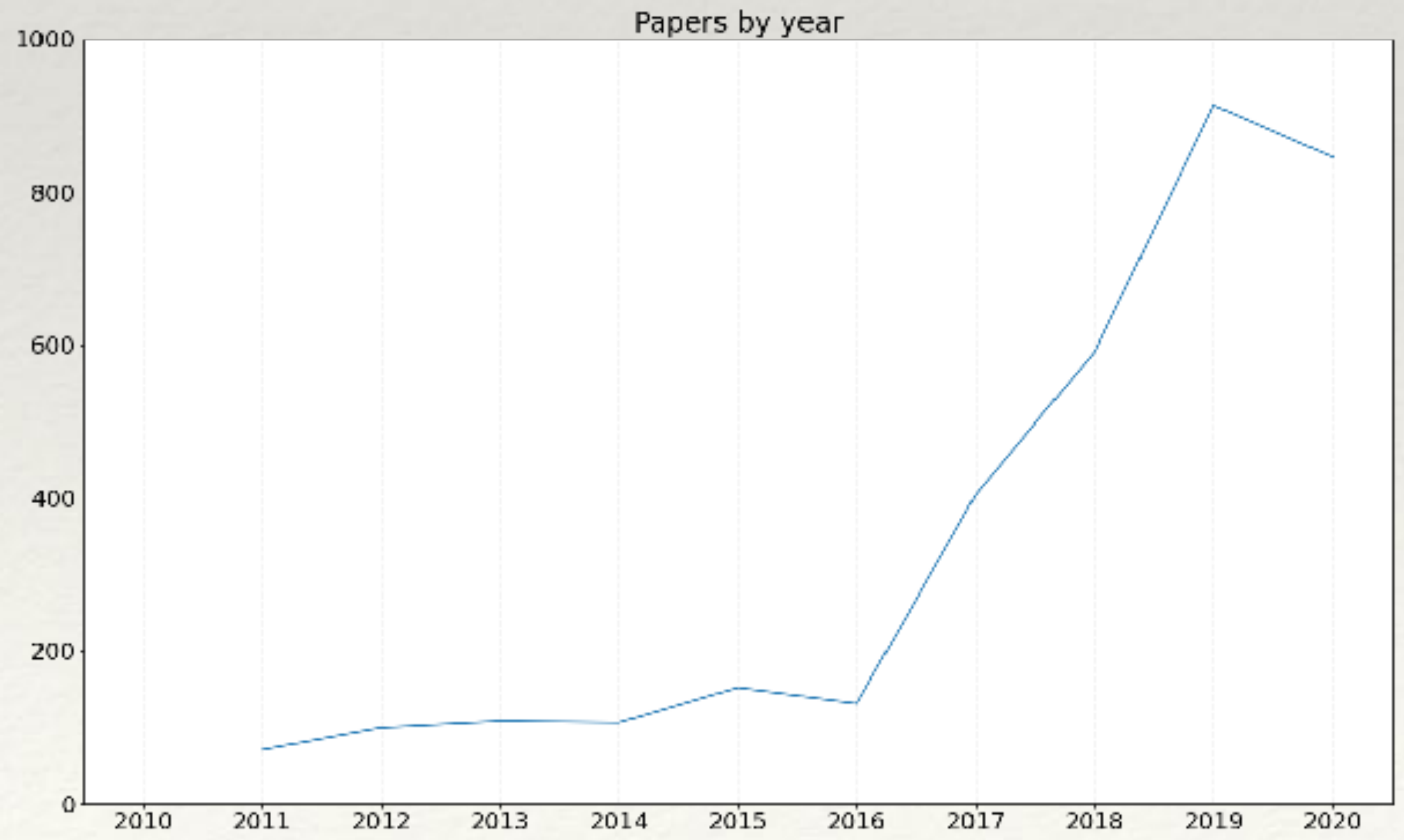
- defining a language to capture the complexity of the phenomenon
- implications for democracy?
- role of television?
- implications of weakened local media?
- micro-targeting
- computational amplification
- filter bubbles and echo chambers
- declining trust in evidence

https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

Scientific papers

- How many papers have been published with "fake news" (or related) in the title?
- We built a dataset from Microsoft Academic, and followed citations
- **40,971** papers (and still counting…)
- Explosive growth after 2016

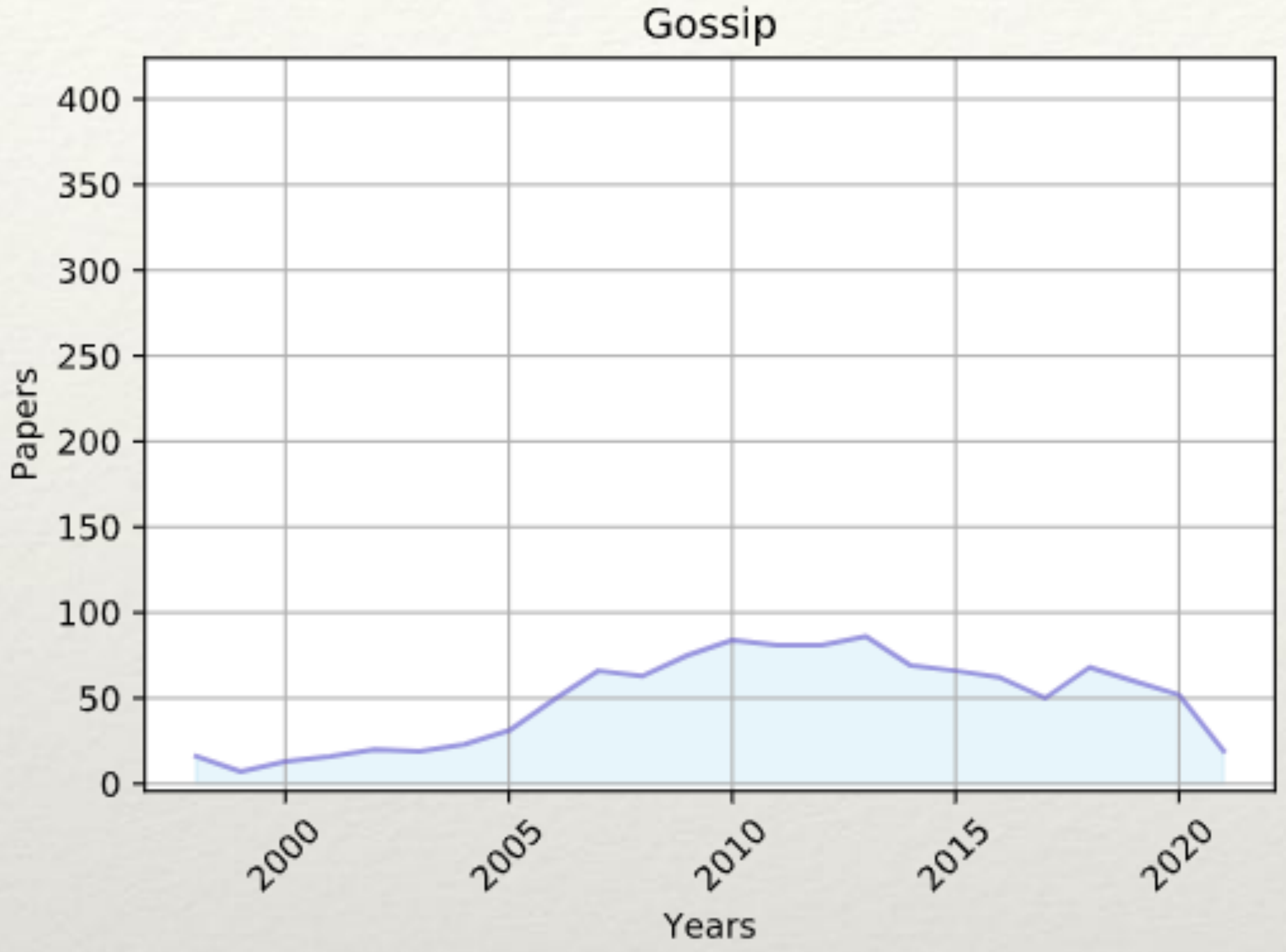


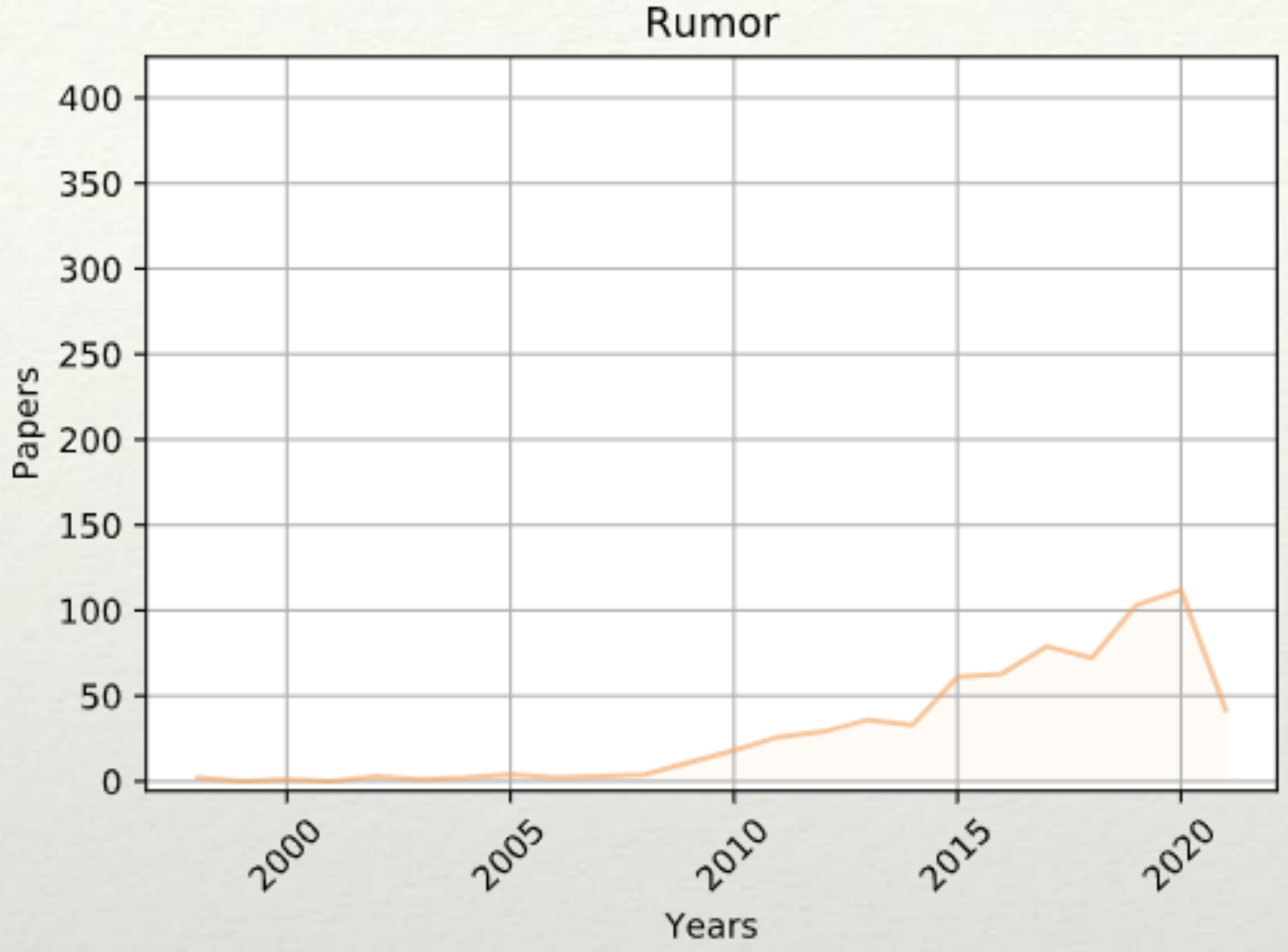


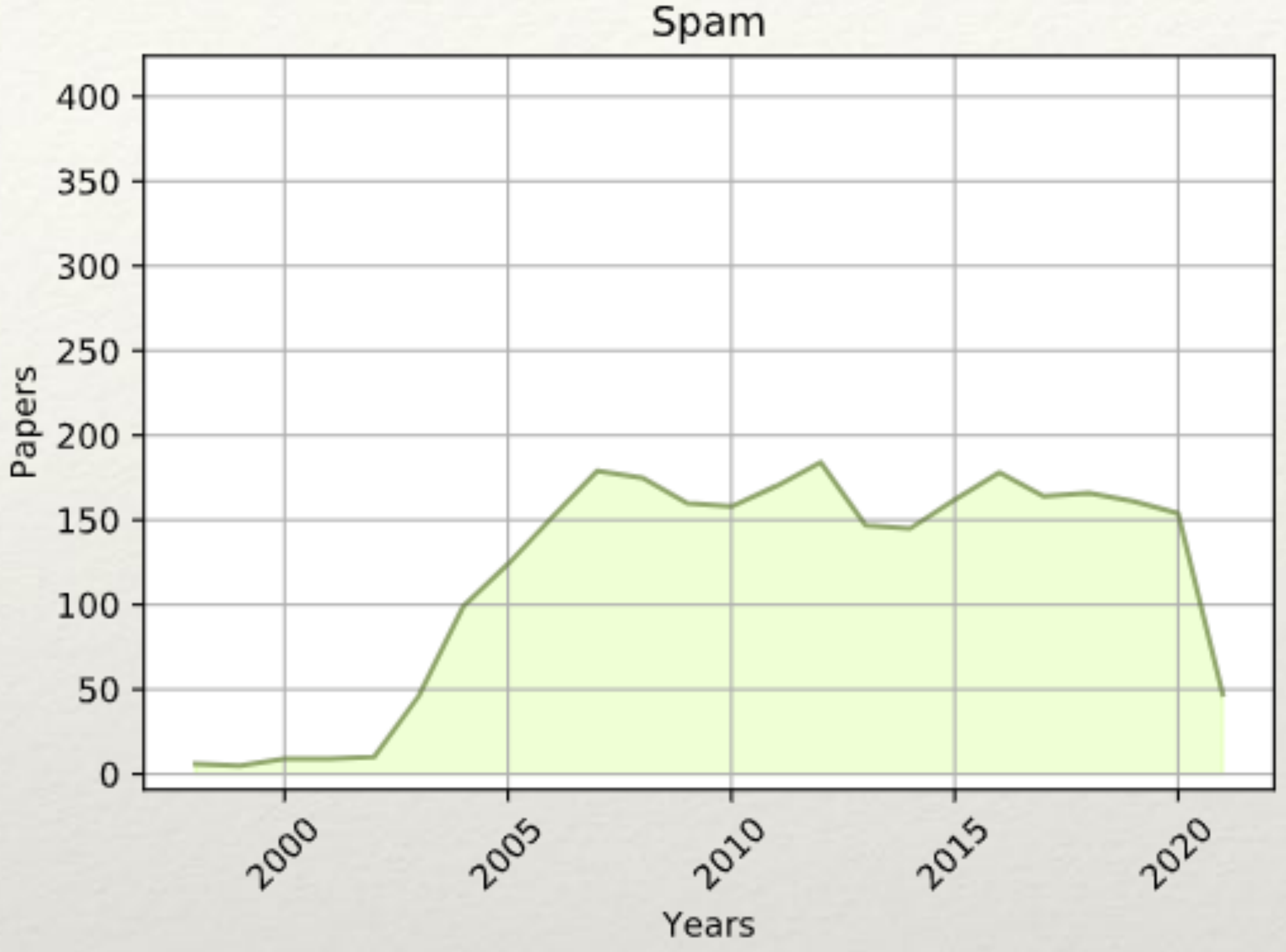
Google Trends

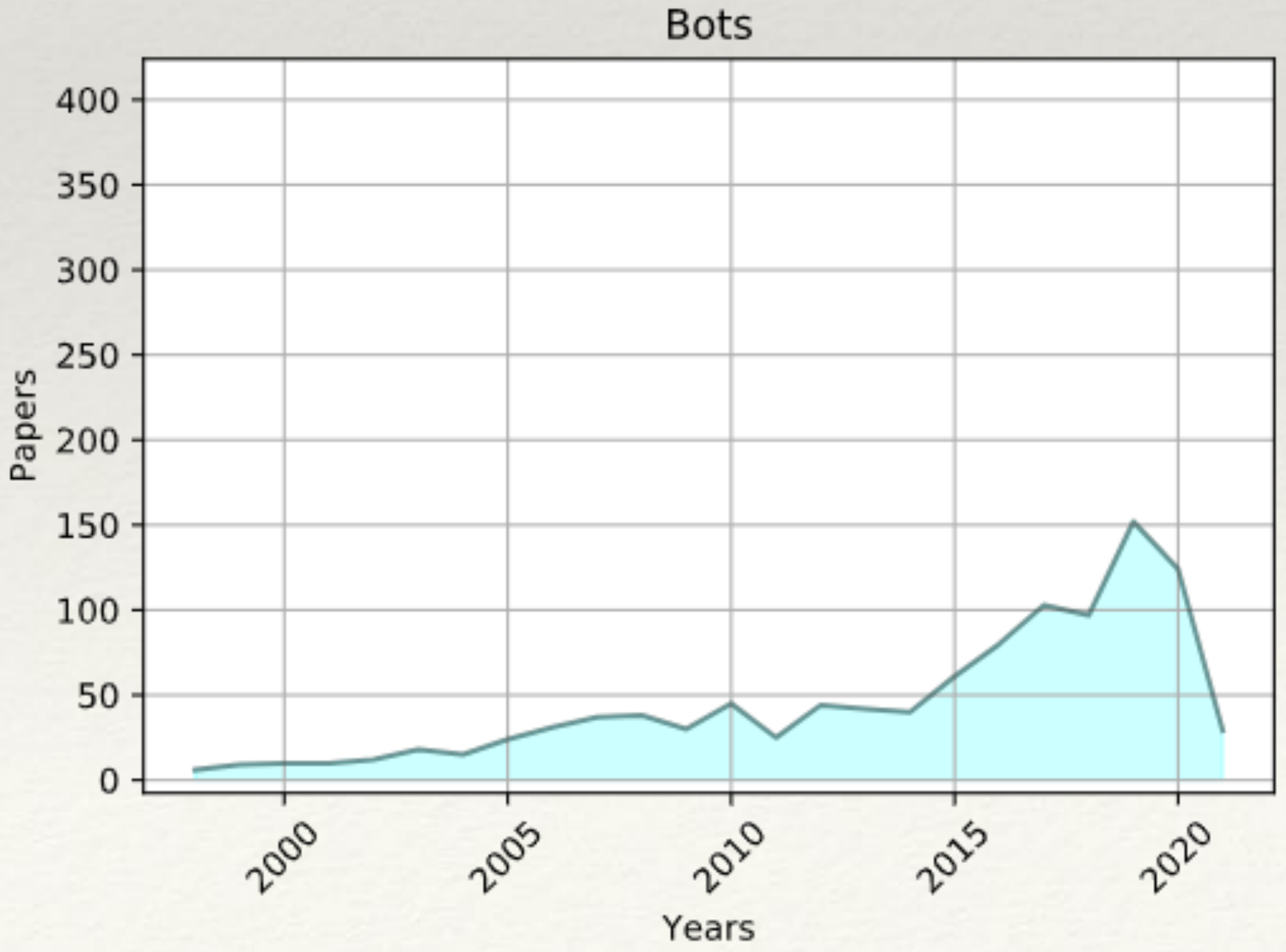
Updated: 2021, Jun. 13th

Focusing on DBLP only



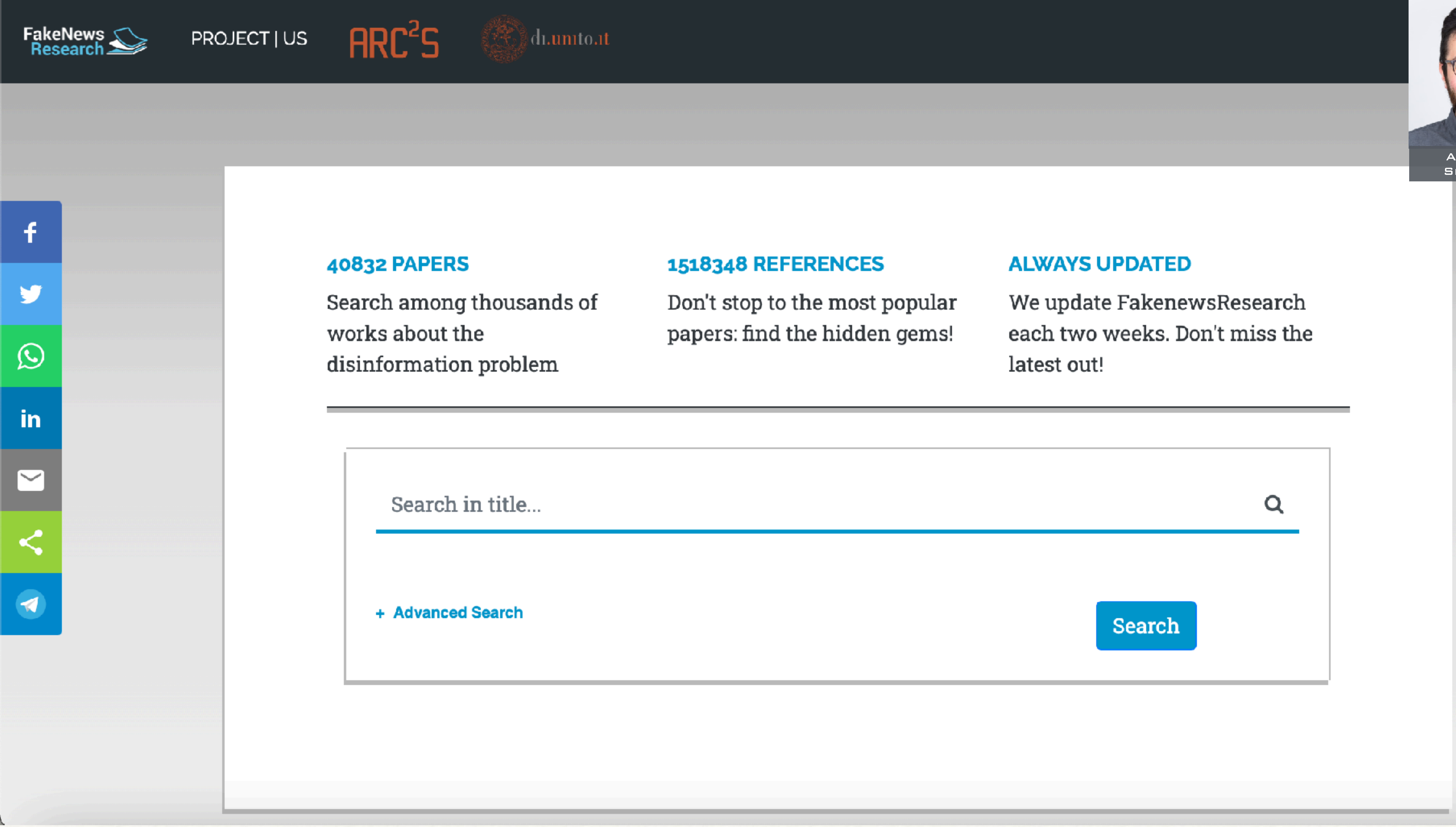


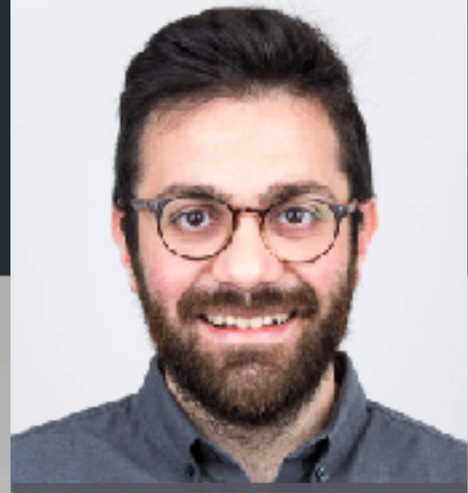




Fast growing literature problem

- Although the problem is considered "new", the literature is huge (and very multidisciplinary)
- Difficult to find an objective and general point of view
- This introductory course is necessarily subjective; however we tried to 'discover hidden gems' with a partially automatic search of relevant and potentially influential scientific contributions





ALFONSO SEMERARO

http://fakenewsresearch.net

2018 Manifesto

POLICY FORUM | SOCIAL SCIENCE

The science of fake news

David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts, Jonathan L. Zittrain

The list of author affiliations is provided in the supplementary materials.

Email: d.lazer@northeastern.edu

– Hide authors and affiliations

Science 09 Mar 2018:
Vol. 359, Issue 6380, pp. 1094-1096
DOI: 10.1126/science.aao2998

"… much remains unknown regarding the vulnerabilities of individuals, institutions, and society to manipulations by malicious actors."

Prevalence and impact

- ***How common is fake news, and what is its impact on individuals?***
- On average, an American encountered from 1 to 3 stories from fake news publishers before the 2016 elections - H. Allcott, M. Gentzkow, J. Econ. Perspect. 31, 211 (2017)
- False information on Twitter is typically **retweeted by many more people**, and far more rapidly, than true information, especially when the topic is politics - S. Vosoughi et al., Science 359, 1146 (2018)
- By liking, sharing, and searching for information, **social bots** can magnify the spread of fake news by orders of magnitude
    - Identification of bots is a moving target and will therefore remain major ongoing research challenge
- Evaluations of the **medium-to-long–run impact** on political behavior of exposure to fake news are essentially nonexistent in the literature.

Potential interventions

- *How can we empower individuals?*
  - **fact-checking**, whose efficacy is disputed
  - **education**, to improve individual evaluation of the quality of information
- *Ho can we prevent individuals' exposure to fake news?*
  - **adjusting social media business models** to increase emphasis on quality information
  - reducing **personalization** and '**echo-chambers**' effects
  - removing accounts associated to **bots**, when they are found
- Content curation decisions are subject to many **ethical considerations**

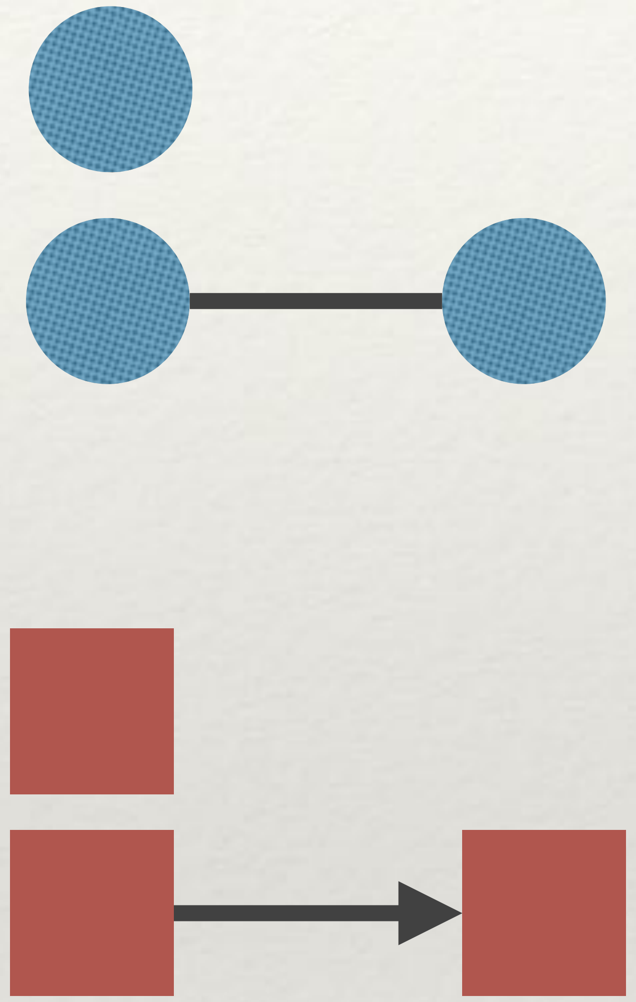
Main questions

- Can we find a language and a framework that is able to capture the complexity of the phenomenon?
- Which are the basic mechanisms that lead to the formation of echo-chambers?
- To which extent an account controlled by a human is vulnerable and manipulable by malicious actors and bots?
- How can we mitigate information spreading?

Introduction to Network Science

Networks are "everywhere"

- Social Networks
  - actors (individuals, also agents)
  - social ties
- Information systems
  - book, web page
  - citation, link, retweet

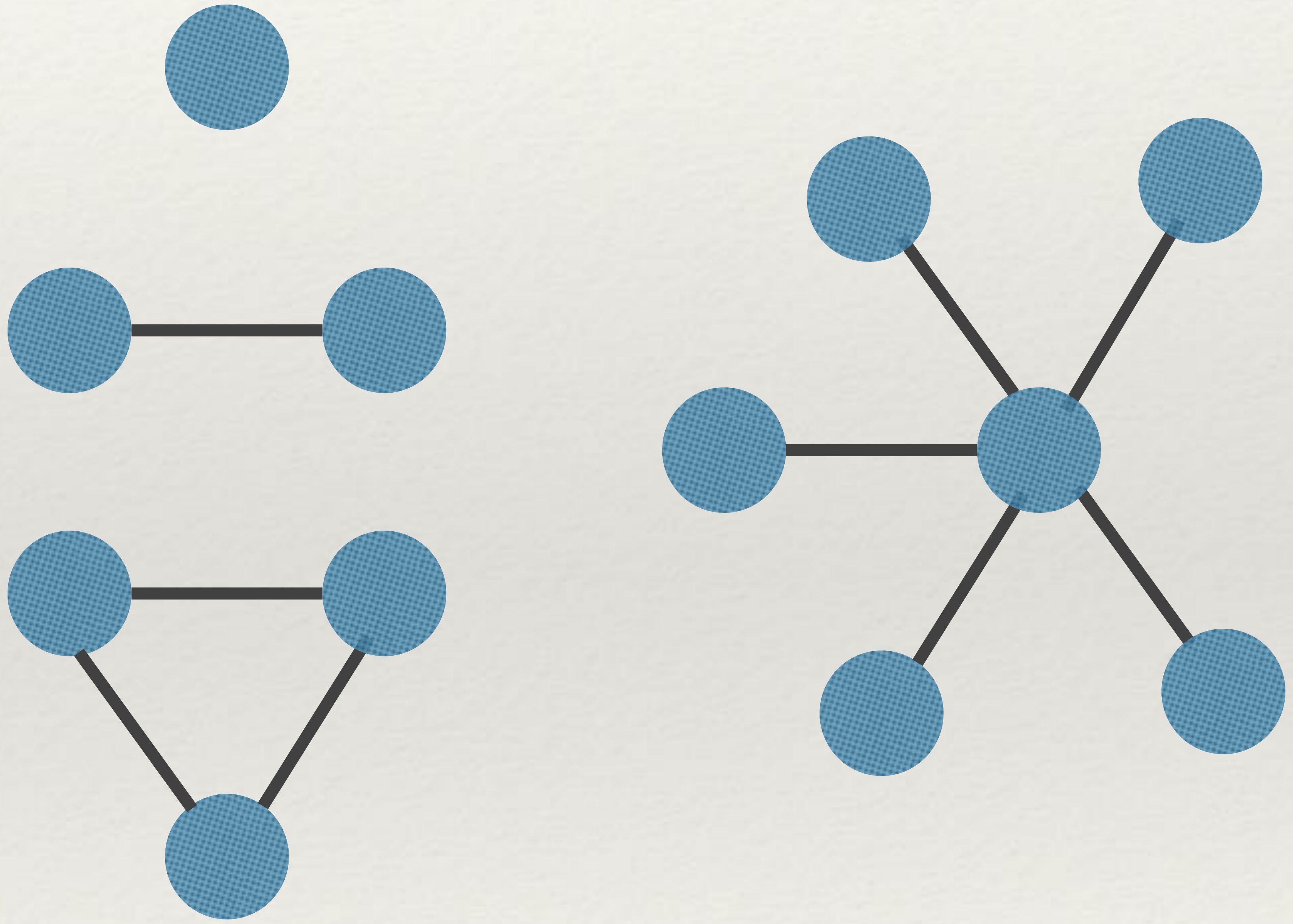


Complex Systems

- Complex != Complicated
- composed by many **interacting elements**
- they give rise to emergent **collective phenomena**
- **emergence**: not directly related to individual phenomena
- linearity vs **non linearity**
- **heterogeneous** vs homogeneous

from a local

- from the interconnection of small units



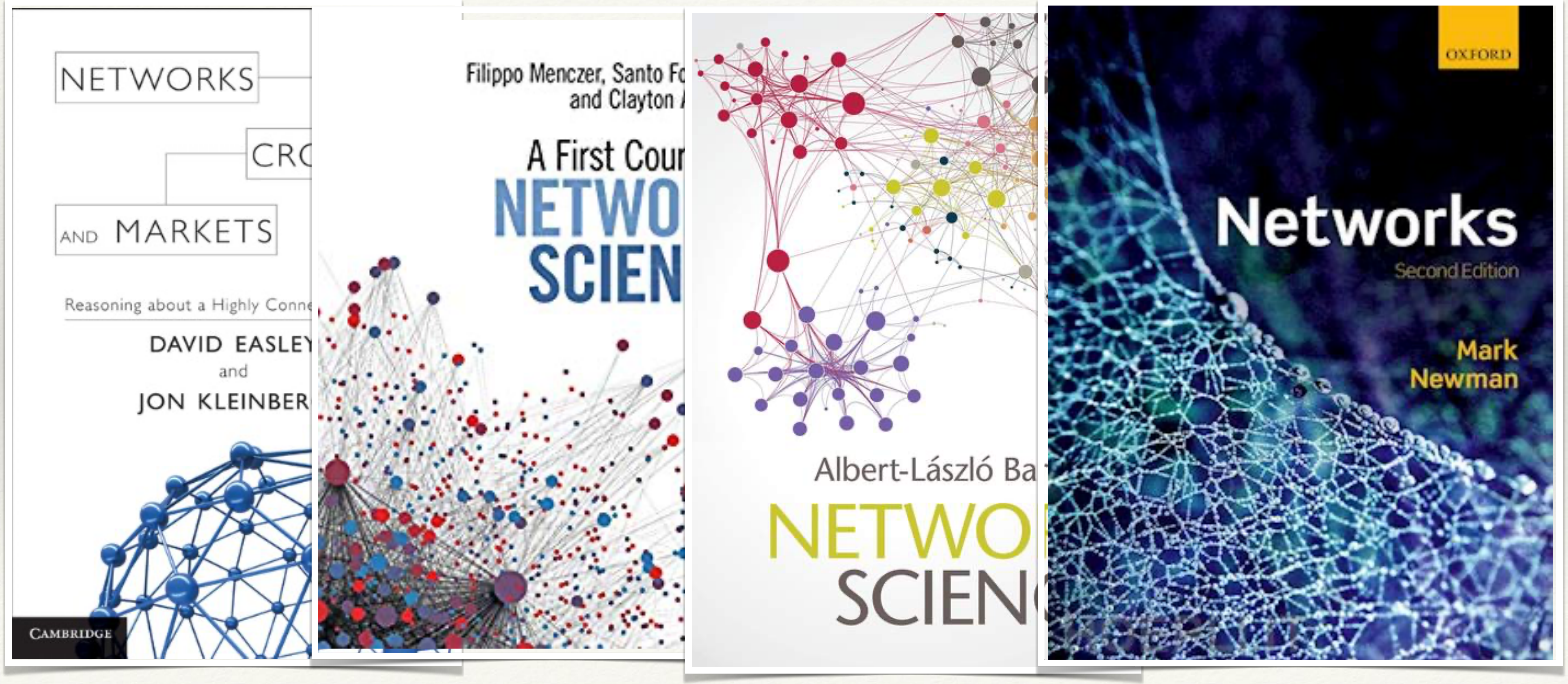
to global level phenomena

Complex Network Analysis

Epidemics

Social Contagion

Textbooks


NETWORKS
CRO
AND MARKETS
Reasoning about a Highly Conne
DAVID EASLEY
and
JON KLEINBER
CAMBRIDGE
Filippo Menczer, Santo Fo
and Clayton A
A First Cour
NETWO
SCIEN
Albert-László Ba
NETWO
SCIEN
OXFORD
Networks
Second Edition
Mark
Newman

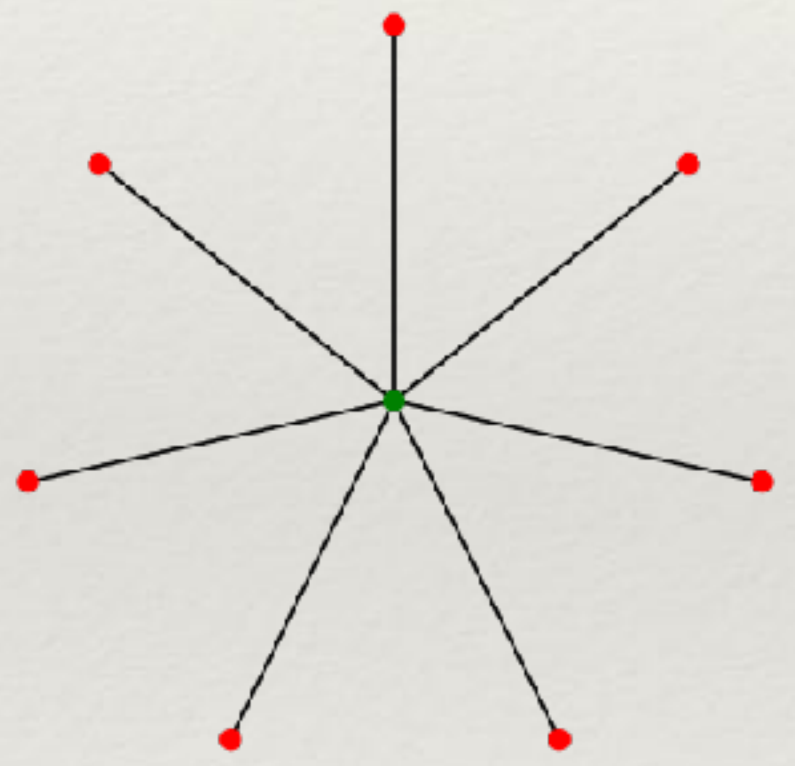
Basics of CNA: Complex Network Analysis

Overview of CNA

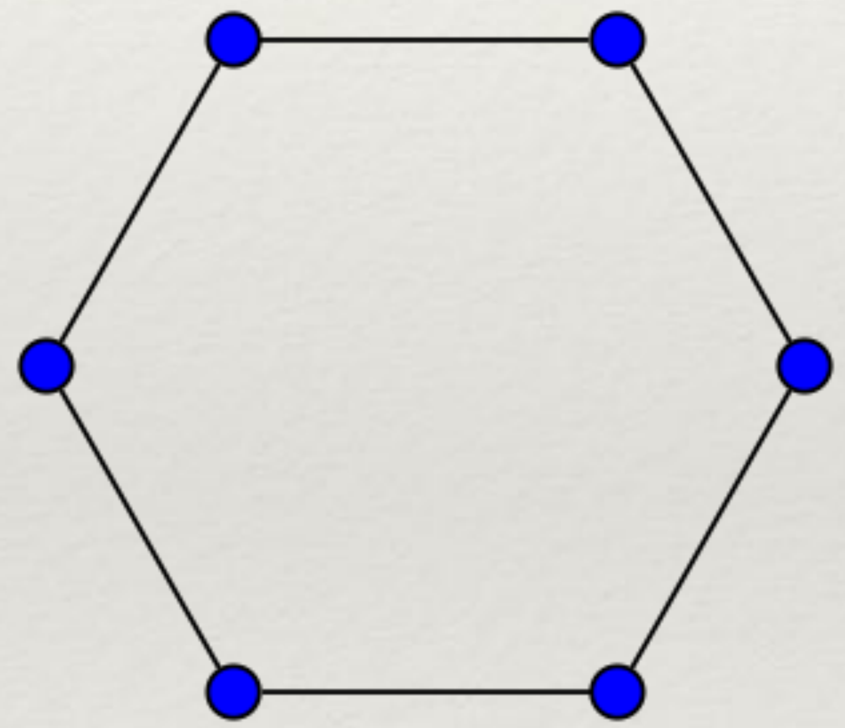
- different roles: hubs, weak ties, bridges, betweenness
- network heterogeneity
- robustness and immunization
- weighted and directed networks
- communities
- homophily
- the emergence of social clusters and segregation
- information and misinformation

Networks structural aspects

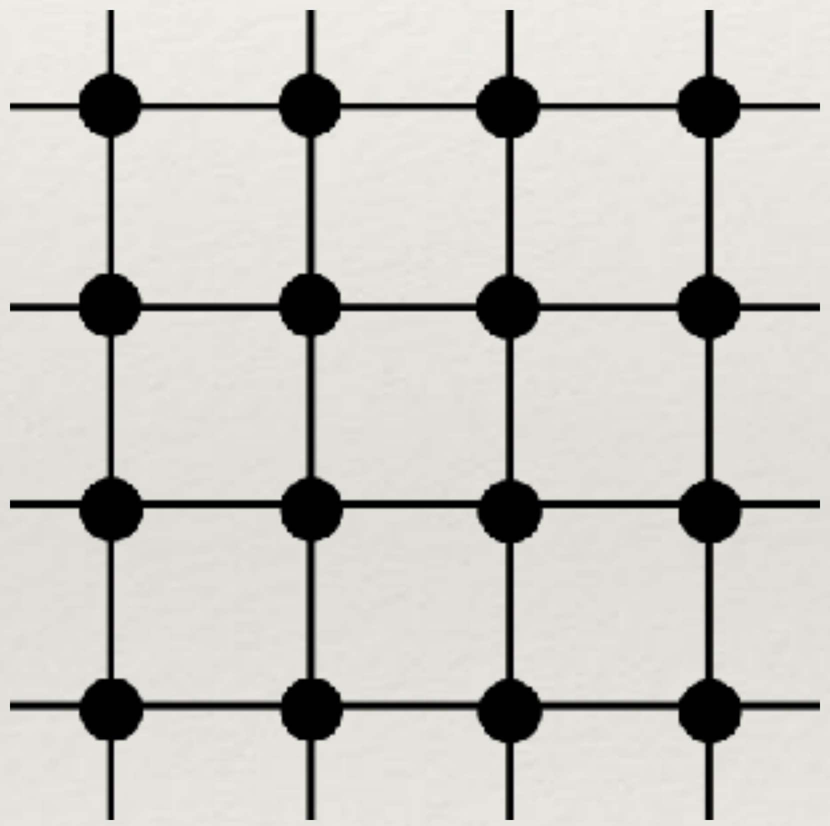
- "trivial" **representation** of a complex system
- **Simple** networks: few characteristics describe the network



Star



Ring



Grid

- We need a **language** and a **framework** to describe complex networks

Basic Definitions

- A **graph** (or a network) is made of nodes and links
- **nodes** (or vertices)
- **links** (or edges, or arcs)
- Graphs can be **directed** or **undirected**
- Graphs can be **weighted** or **unweighted**

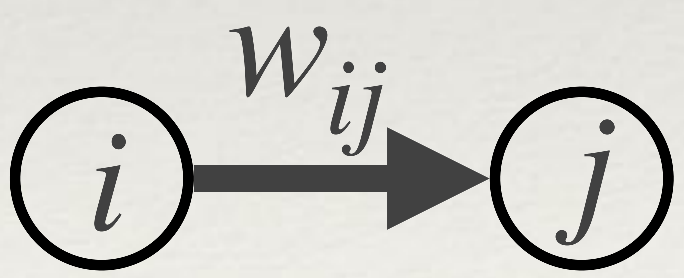
$$G = (N, L)$$

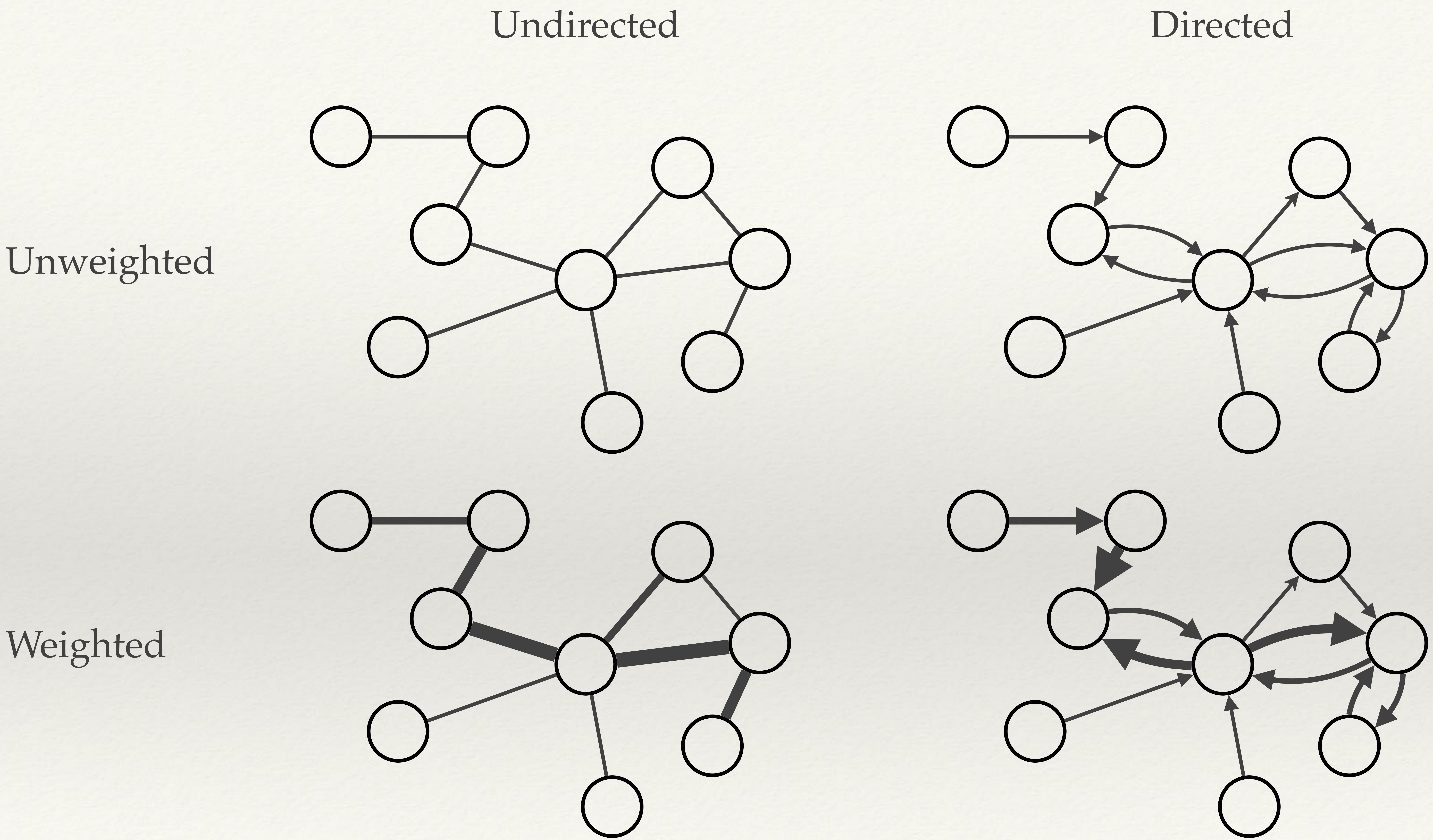
$$N = \{n_1, n_2, \ldots, n_l\} = \{1,2,\ldots,l\}$$

$$L = \{(i,j) : i,j \in N\}$$

(i,j)

w_{ij} $\quad (i,j,w)$




Undirected
Directed
Unweighted
Weighted

Degree

Number of links (or neighbors)

$i \to N_i$ $\quad k_i = |N_i|$ degree

Singleton: a node whose degree is zero

$N_i = \{\}, k_i = 0$

In directed networks

$k_i^{in} = |P_i|$ in-degree

$k_i^{out} = |S_i|$ out-degree

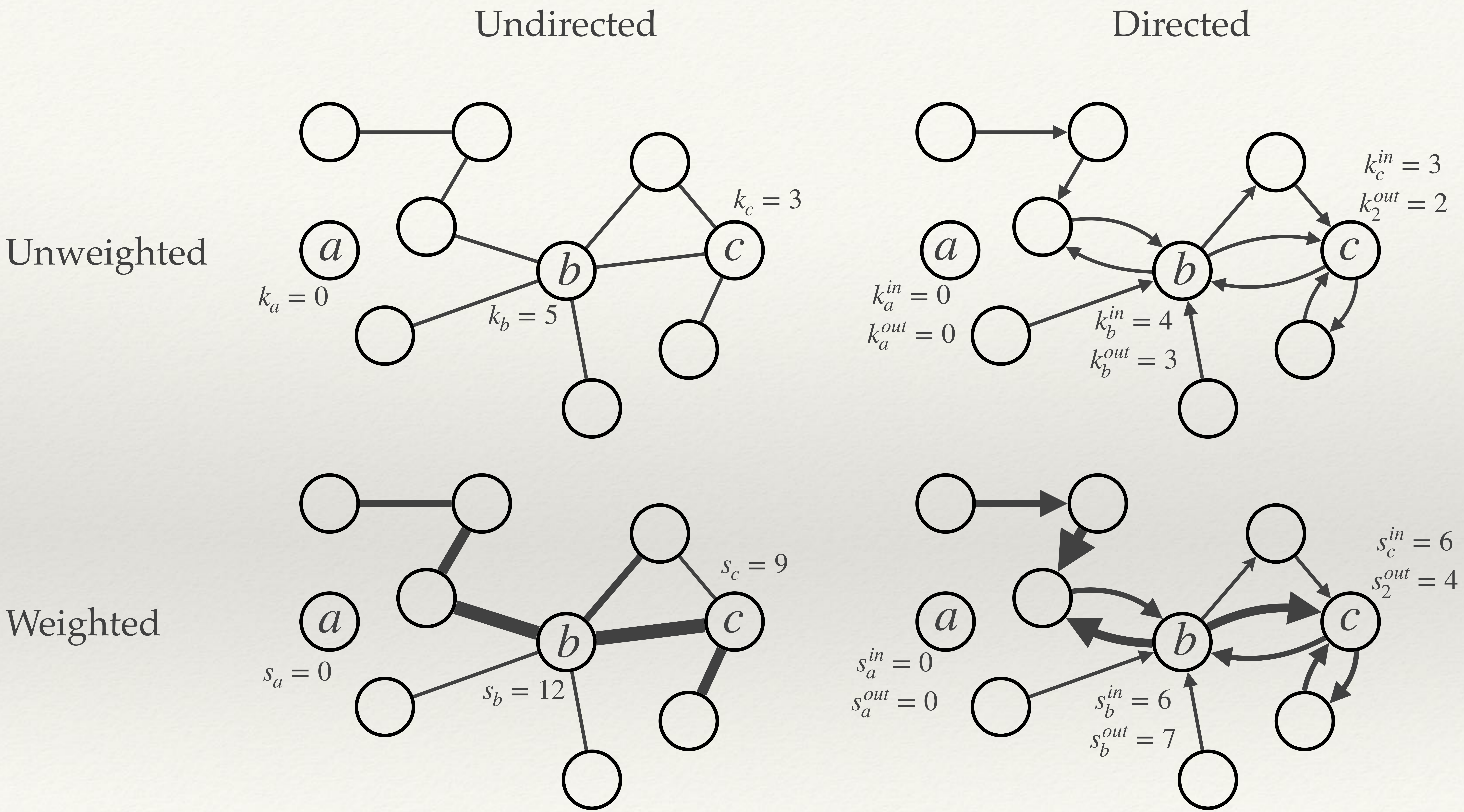
$$k_i = k_i^{in} + k_i^{out}$$

Strength

Strength: Weighted degree $s_i = \sum_{j \in N_i} w_{ij}$

in-strength $s_i^{in} = \sum_{j \in P_i} w_{ji}$

out-strength $s_i^{out} = \sum_{j \in S_i} w_{ij}$


Undirected
Directed
Unweighted
Weighted
a
b
c
$k_a = 0$
$k_b = 5$
$k_c = 3$
$k_a^{in} = 0$
$k_a^{out} = 0$
$k_b^{in} = 4$
$k_b^{out} = 3$
$k_c^{in} = 3$
$k_2^{out} = 2$
$s_a = 0$
$s_b = 12$
$s_c = 9$
$s_a^{in} = 0$
$s_a^{out} = 0$
$s_b^{in} = 6$
$s_b^{out} = 7$
$s_c^{in} = 6$
$s_2^{out} = 4$

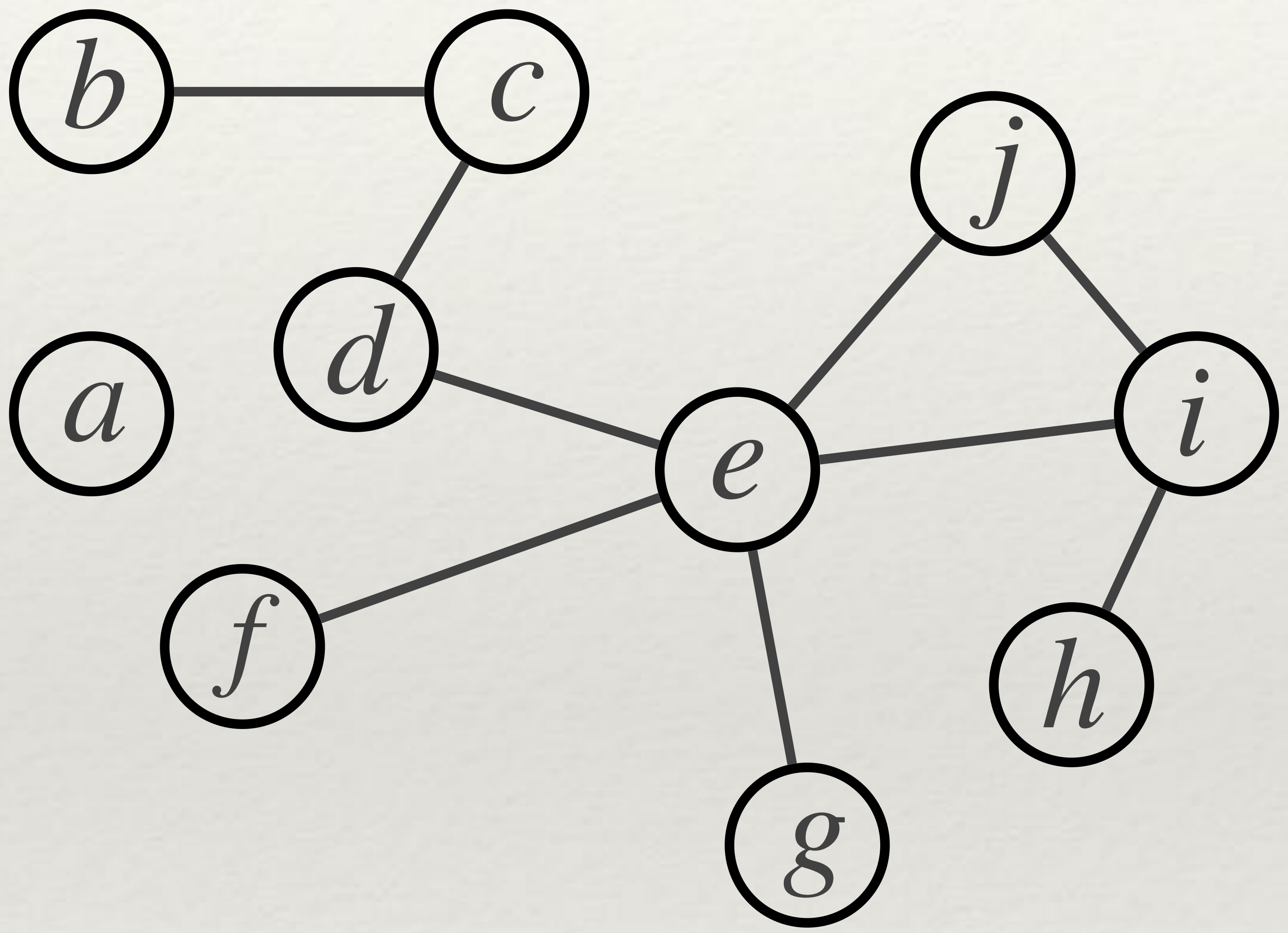
Network representations

Adjacency Matrix

NxN matrix

$$a_{ij} = \begin{cases} 0 & \text{no edge} \\ 1 & (i,j) \in L \end{cases}$$

Undirected network: $a_{ij} = a_{ji}$



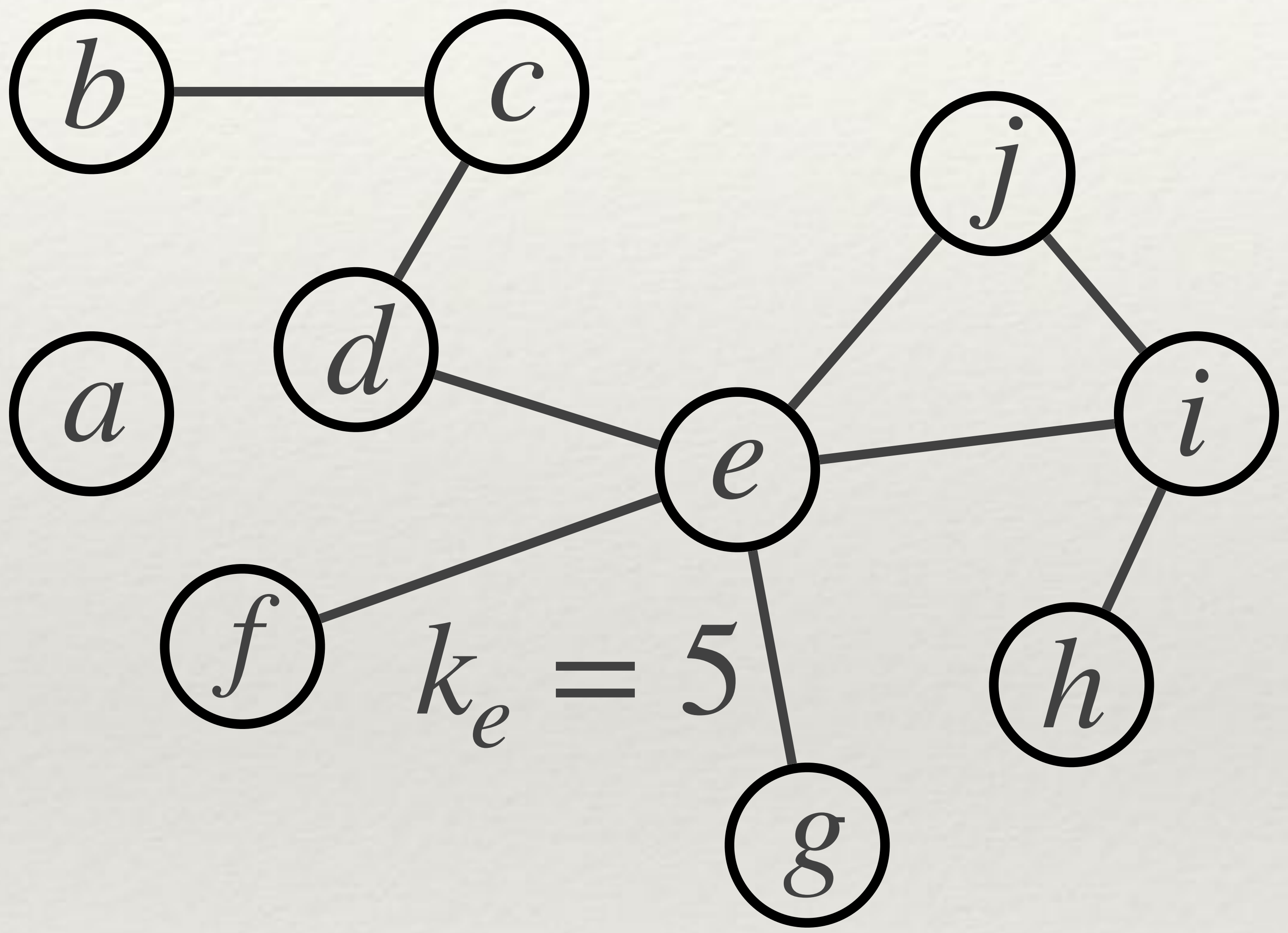
$$\begin{array}{cc} & \begin{array}{cccccccccc} \mathrm{a} & \mathrm{b} & \mathrm{c} & \mathrm{d} & \mathrm{e} & \mathrm{f} & \mathrm{g} & \mathrm{h} & \mathrm{i} & \mathrm{j} \end{array} \\ \begin{array}{c} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \\ \mathrm{d} \\ \mathrm{e} \\ \mathrm{f} \\ \mathrm{g} \\ \mathrm{h} \\ \mathrm{i} \\ \mathrm{j} \end{array} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Network representations

Adjacency Matrix

degree

$$k_i = \sum_j a_{ij} = \sum_j a_{ji}$$



$$
\begin{array}{c}
 \\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc} a & b & c & d & e & f & g & h & i & j \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}
$$

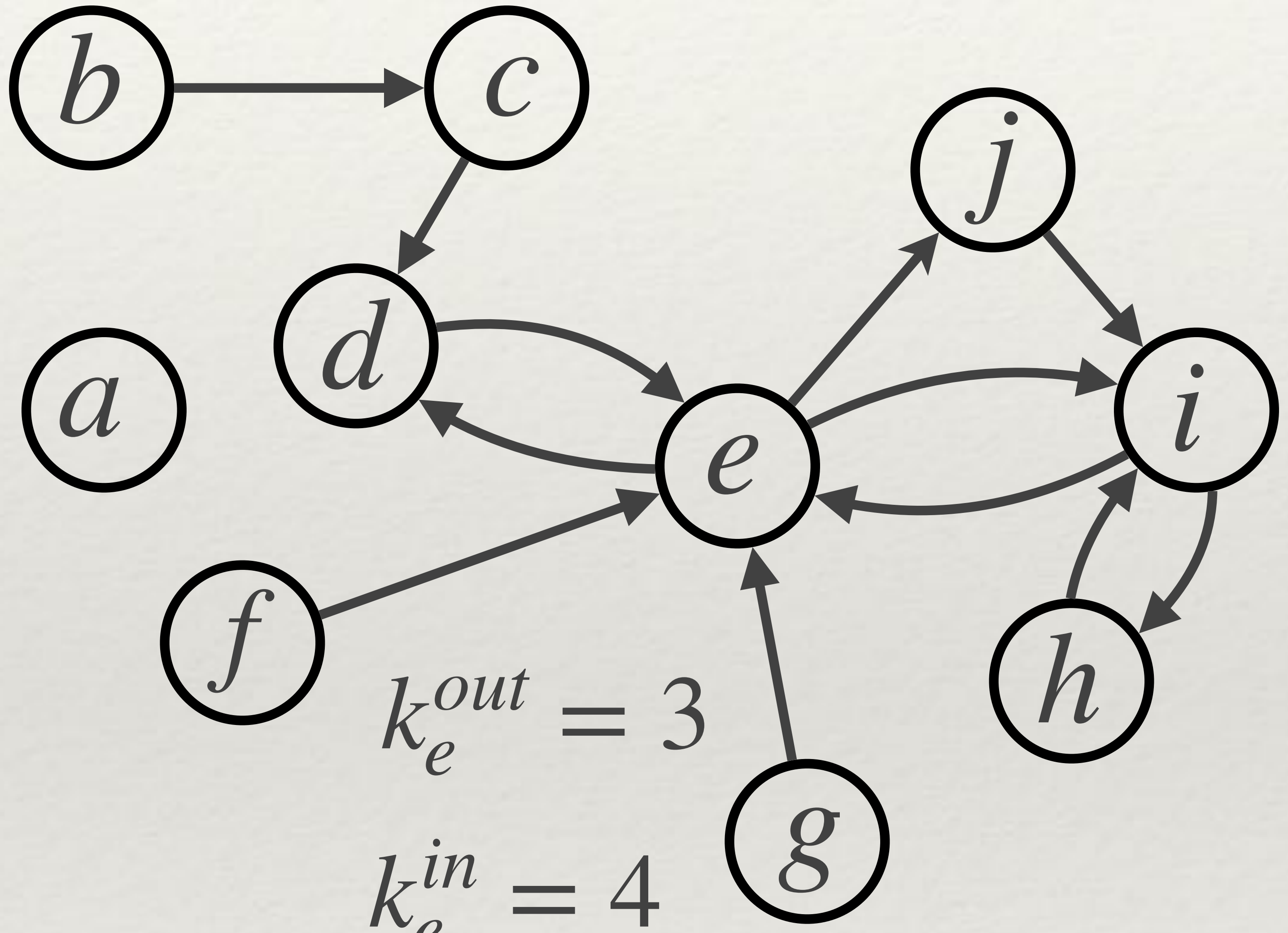
Network representations

Adjacency Matrix
directed

Matrix is not symmetric

$$k_i^{out} = \sum_j a_{ij}$$

$$k_i^{in} = \sum_j a_{ji}$$



$$k_e^{out} = 3$$

$$k_e^{in} = 4$$

$$\begin{array}{c@{\;}c}
 & \begin{array}{cccccccccc} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} & \mathtt{f} & \mathtt{g} & \mathtt{h} & \mathtt{i} & \mathtt{j} \end{array} \\
\begin{array}{c} \mathtt{a} \\ \mathtt{b} \\ \mathtt{c} \\ \mathtt{d} \\ \mathtt{e} \\ \mathtt{f} \\ \mathtt{g} \\ \mathtt{h} \\ \mathtt{i} \\ \mathtt{j} \end{array} &
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
\end{array}$$

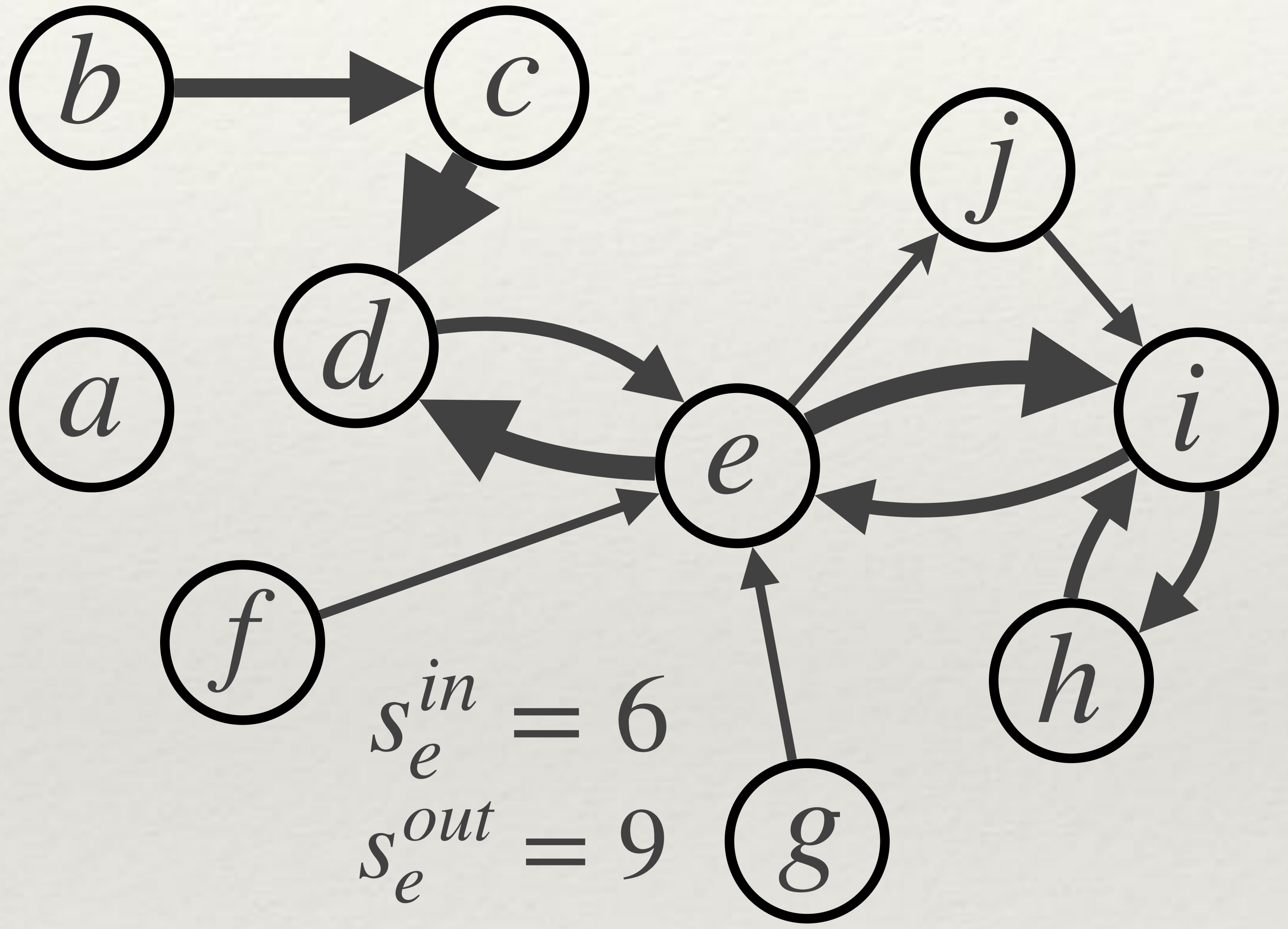
Network representations

Adjacency Matrix
weighted

$$w_{ij}$$

$$s_i^{out} = \sum_j w_{ij}$$

$$s_i^{in} = \sum_j w_{ji}$$

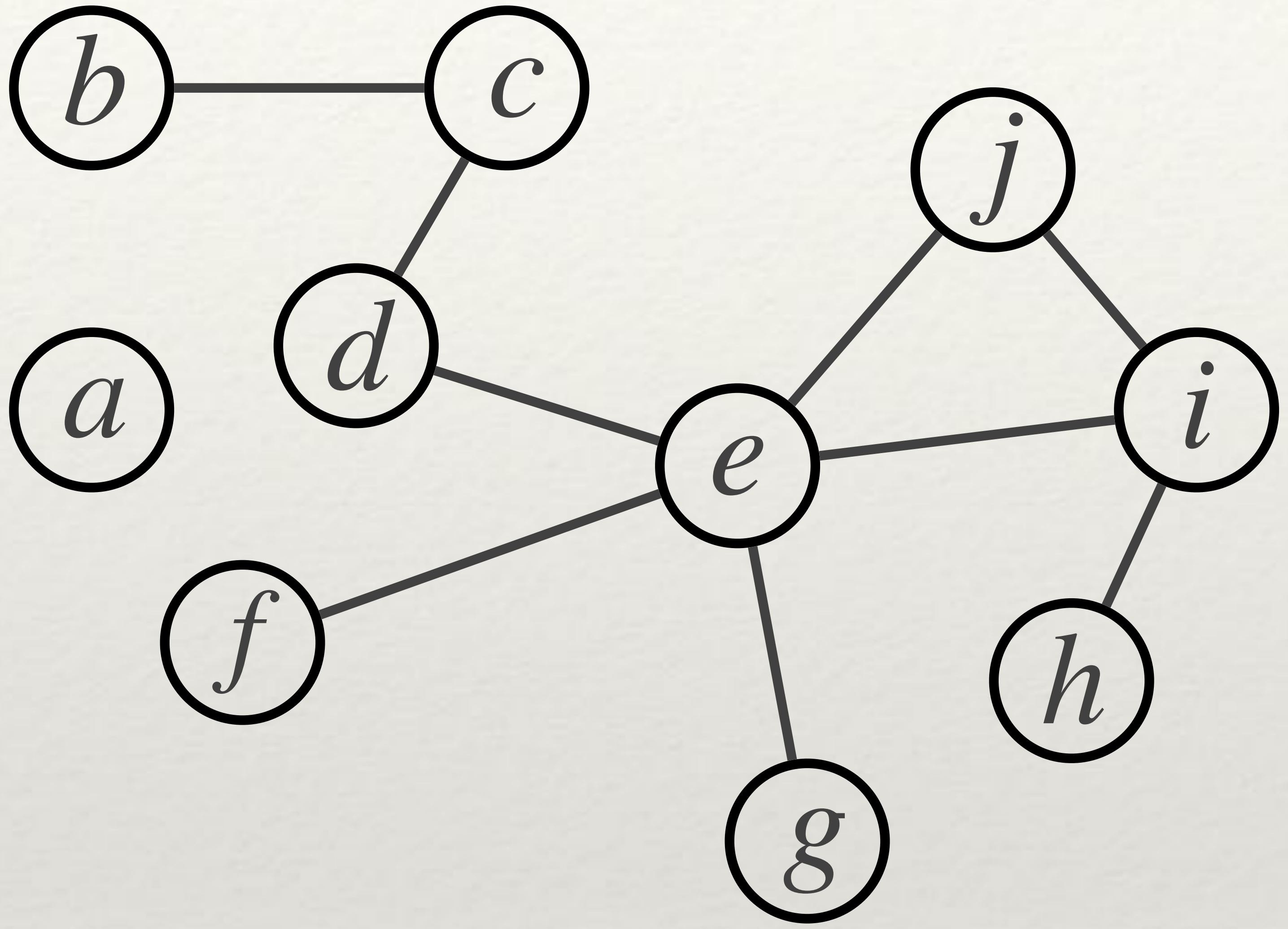


$$\begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \end{array}
\begin{array}{c} \begin{array}{cccccccccc} a & b & c & d & e & f & g & h & i & j \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 4 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0
\end{pmatrix} \end{array}$$

Sparse network representations

- The memory/disk storage needed by an adjacency matrix is proportional to N^2
- In sparse networks (most real-world networks), this is terribly inefficient: most of the space is wasted storing zeros (non-links); for very large networks, adjacency matrices are unfeasible
- It is much more efficient, often necessary, to store only the actual links, and assume that if a link is not listed it means it is not present
- There are two commonly used sparse networks representations:
  - **Adjacency list**
  - **Edge list**

Adjacency list

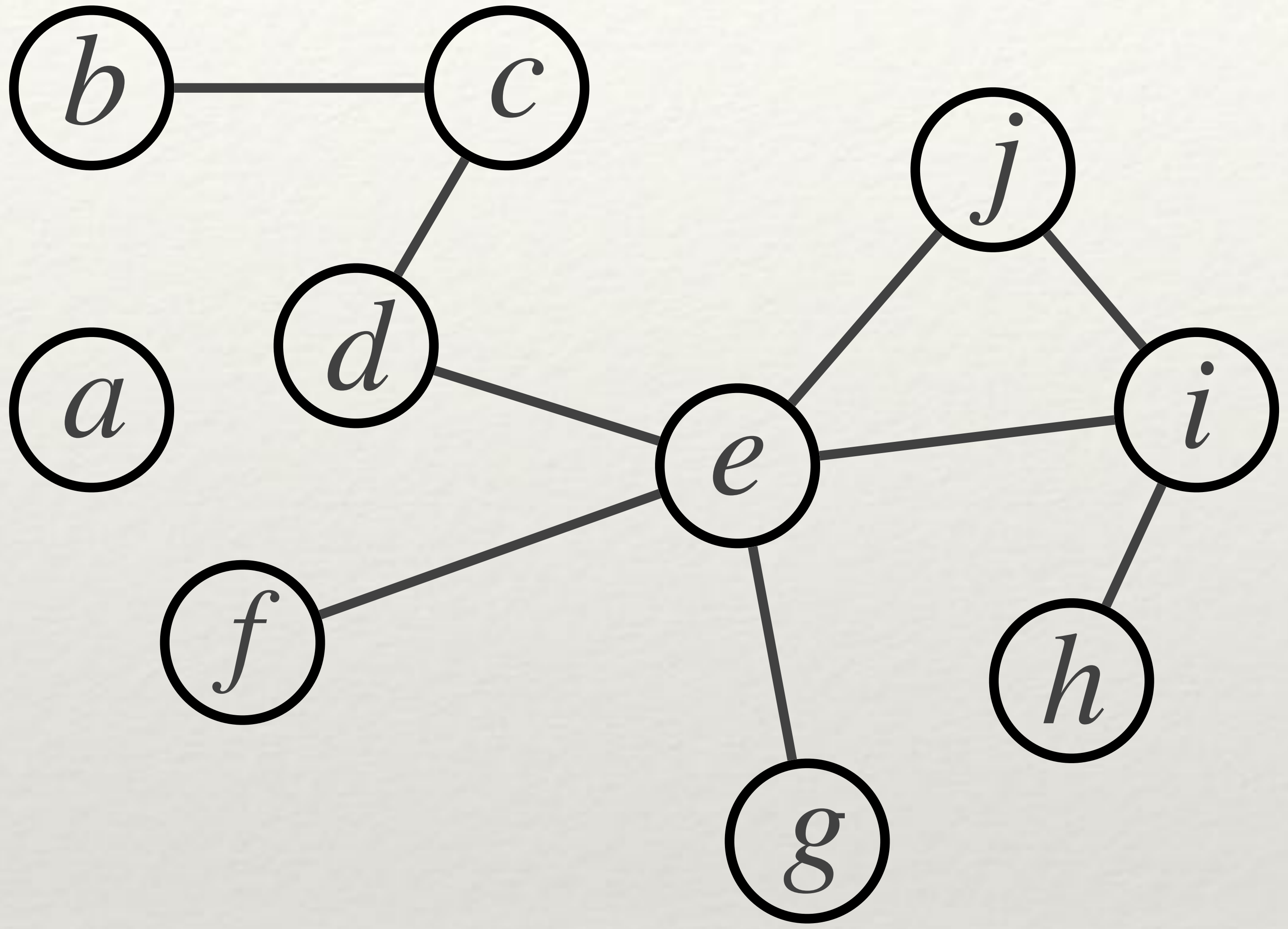


a	
b	c
c	b d
d	c e
e	d f i j g
f	e
g	e
h	i
i	h j
j	e i

Undirected network: list each link twice

Directed network: list only existing links

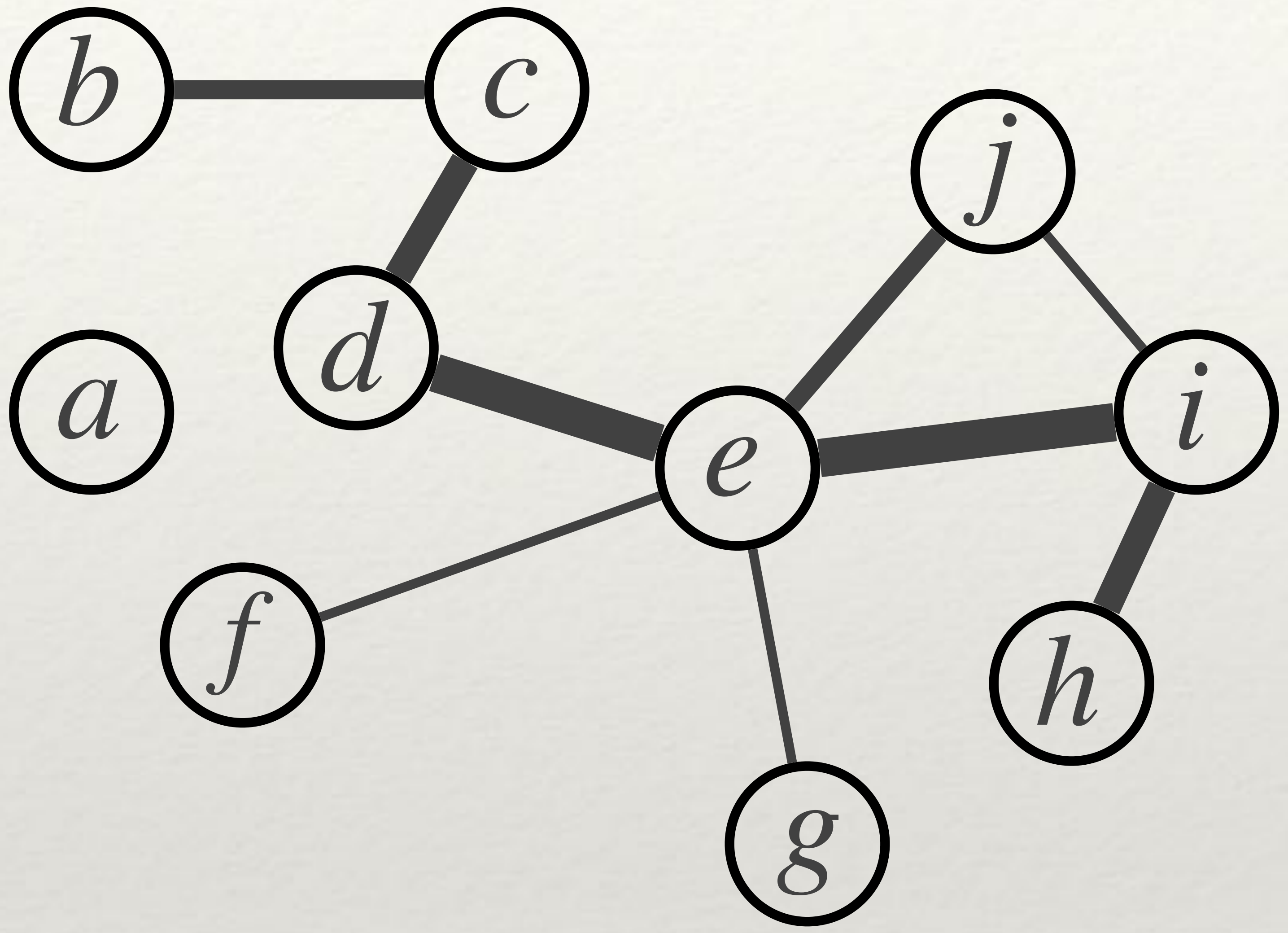
Edge list



b	c
c	d
d	e
e	f
e	g
e	i
e	j
h	i
i	j

L

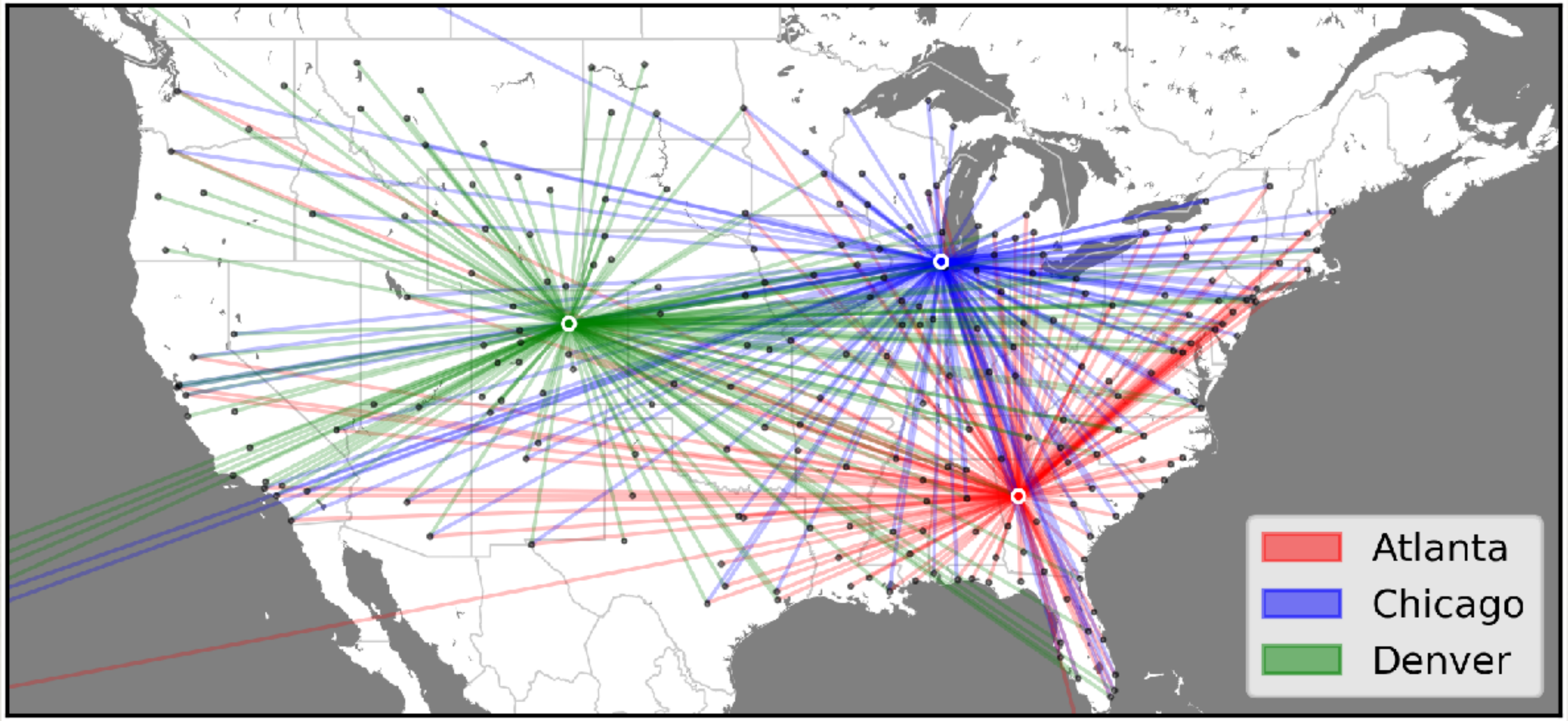
Edge (weighted) list



b	c	2
c	d	3
d	e	4
e	f	4
e	g	1
e	i	1
e	j	2
h	i	3
i	j	1

L

Real networks are heterogeneous



Some nodes (and links) are much more important (**central**) than others!

Centrality measures

- **Centrality**: measure of importance of a node
- **Measures**:

1. Degree
2. Closeness
3. Betweenness

Degree

- **Degree of a node:** number of neighbors of the node

$$k_i = \text{number of neighbors of node } i$$

- High-degree nodes are called **hubs**

- **Average degree of the network:**

$$\langle k \rangle = \frac{\sum_i k_i}{N} = \frac{2L}{N}$$

```
G.degree(2) # returns the degree of node 2
G.degree()  # dict with the degree of all nodes of G
```

Closeness

Idea: a node is the more central the *closer* it is to the other nodes, on average

$$g_i = \frac{1}{\sum_{j \neq i} \ell_{ij}}$$

where ℓ_{ij} is the distance between nodes i and j

```
nx.closeness_centrality(G, node) # closeness centrality
                                 # of node
```

Betweenness

Idea: a node is the more central the *more often it is crossed by paths*

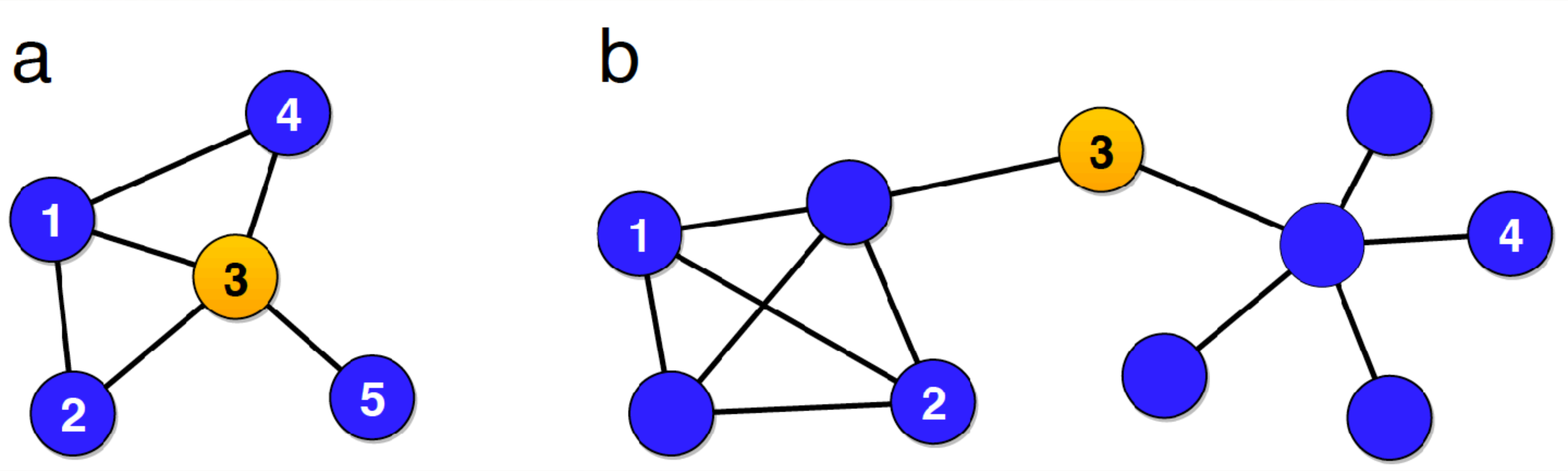
$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

σ_{hj} = **number of shortest paths from h to j**

$\sigma_{hj}(i)$ = **number of shortest paths from h to j running through i**

Betweenness

Hubs usually have high betweenness, but there can be nodes with high betweenness **that are not hubs**



A First Course in Network Science by F. Menczer, S. Fortunato & C.A. Davis. Cambridge University Press, 2020

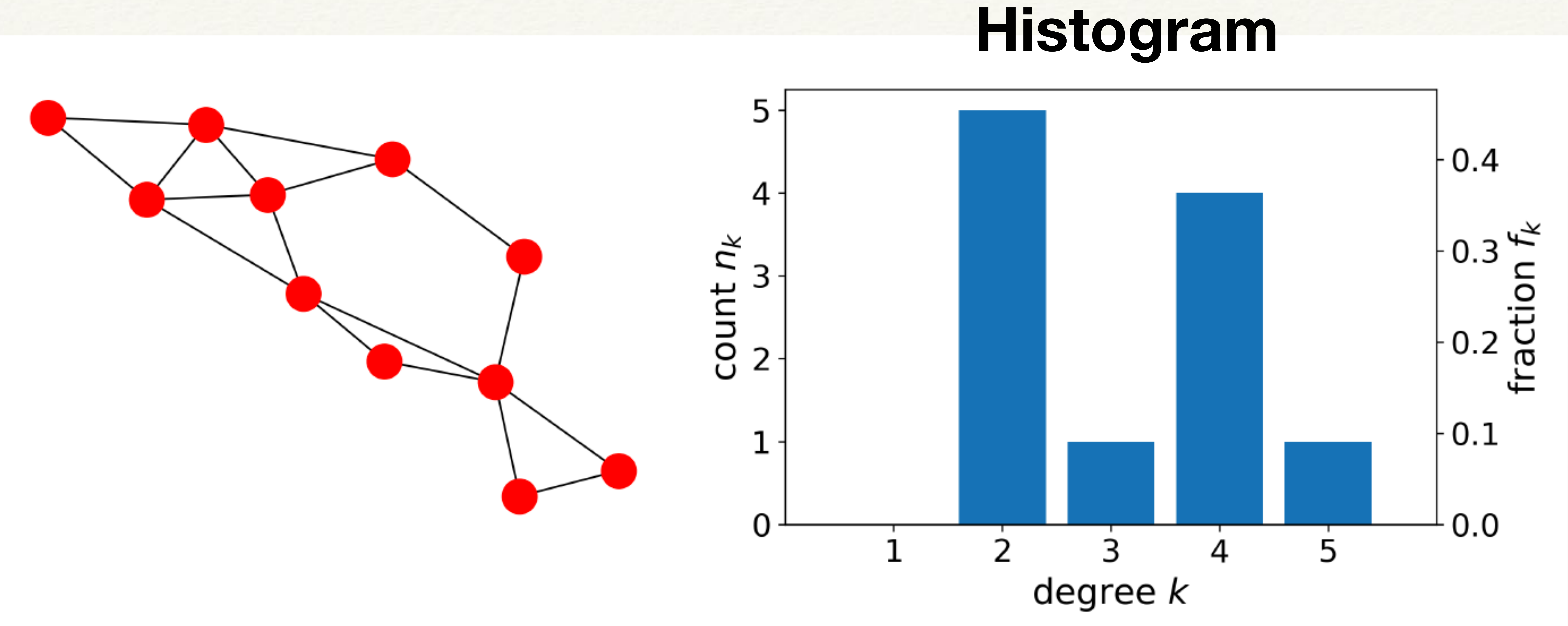
Betweenness

- Betweenness can be easily extended to links
- **Link betweenness:** fraction of shortest paths among all possible node pairs that pass through the link

Centrality distributions

- On small networks it makes sense to ask which nodes or links are most important
- On large networks **it does not**
- **Solution:** statistical approach
- Instead of focusing on individual nodes and links, we consider **classes** of nodes and links with similar properties

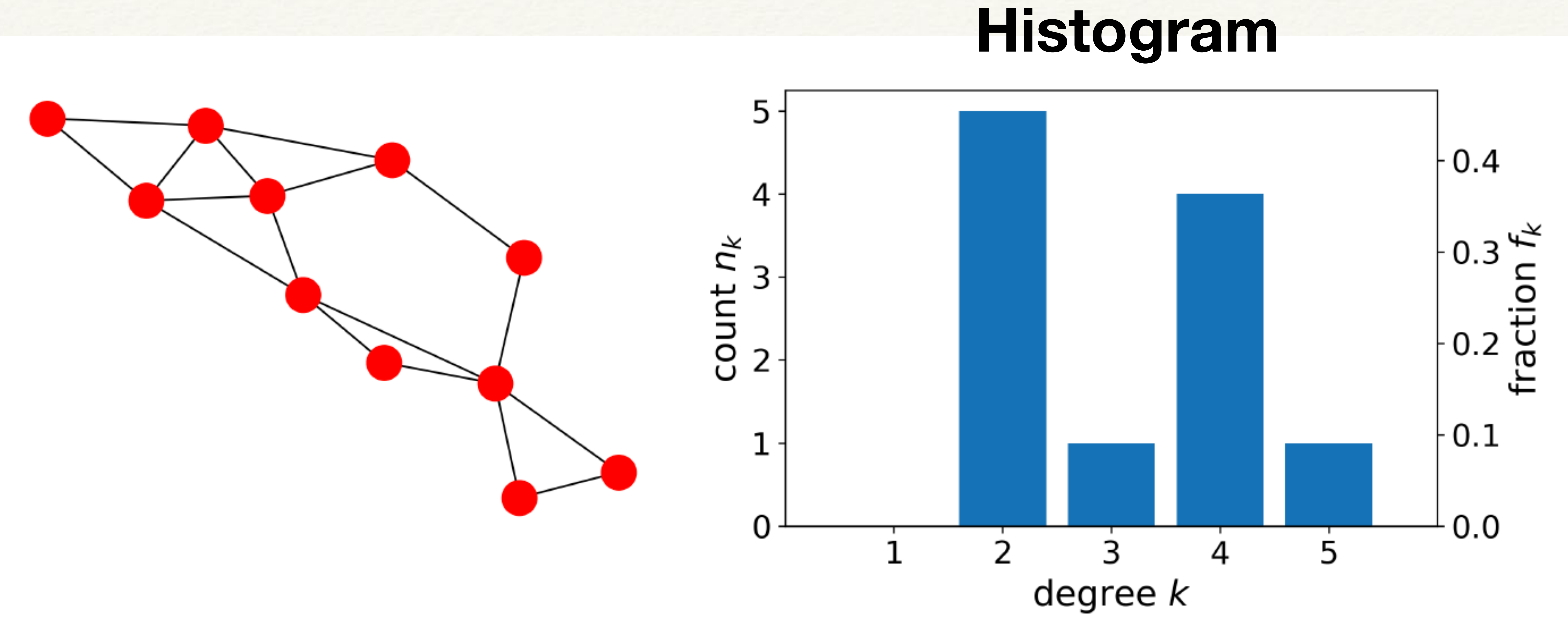
Centrality distributions



n_k = **number of nodes with degree** k

$f_k = \frac{n_k}{N}$ = **frequency of degree** k

Centrality distributions



- For large N, the frequency f_k becomes the **probability** p_k of having degree k
- **Probability distribution:** plot of probability p_k versus k

Cumulative distributions

- If the variable is *not integer* (e.g., betweenness), the range of the variable is divided into intervals (bins) and we count how many values fall in each interval

- **Cumulative distribution *P(x)*:** probability that the variable takes values *larger* than *x* as a function of *x*

- **How to compute it:** by summing the frequencies of the variable inside the intervals to the right of *x*

$$P(x) = \sum_{v \geq x} f_v$$

Logarithmic scale

- **Question:** how to plot a probability distribution if the variable spans a large range of values, from small to (very) large?

- **Answer:** use the **logarithmic** scale

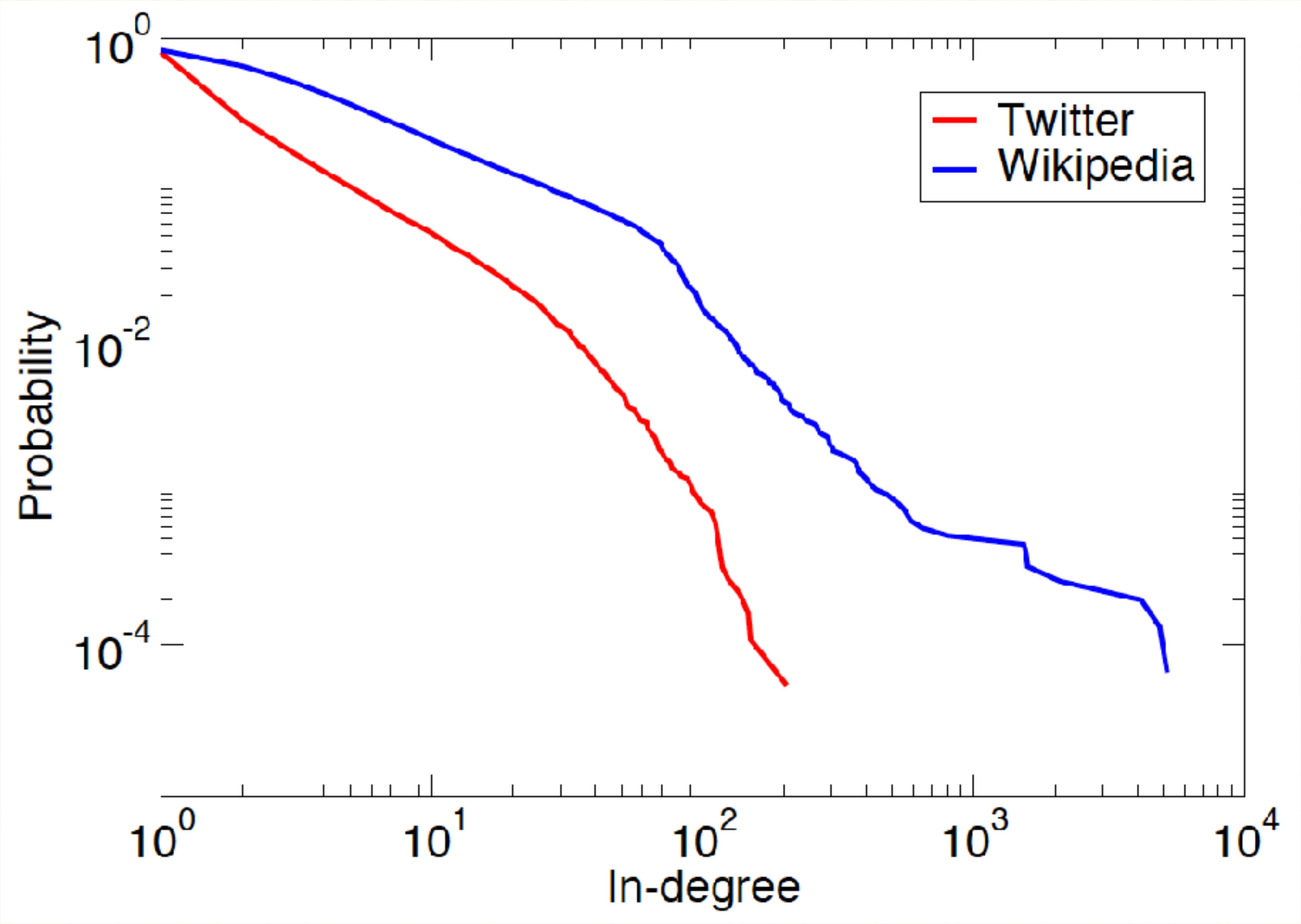
- **How to do it:** report the logarithms of the values on the x- and y-axes

$$\log_{10} 10 = 1$$

$$\log_{10} 1{,}000 = \log_{10} 10^3 = 3$$

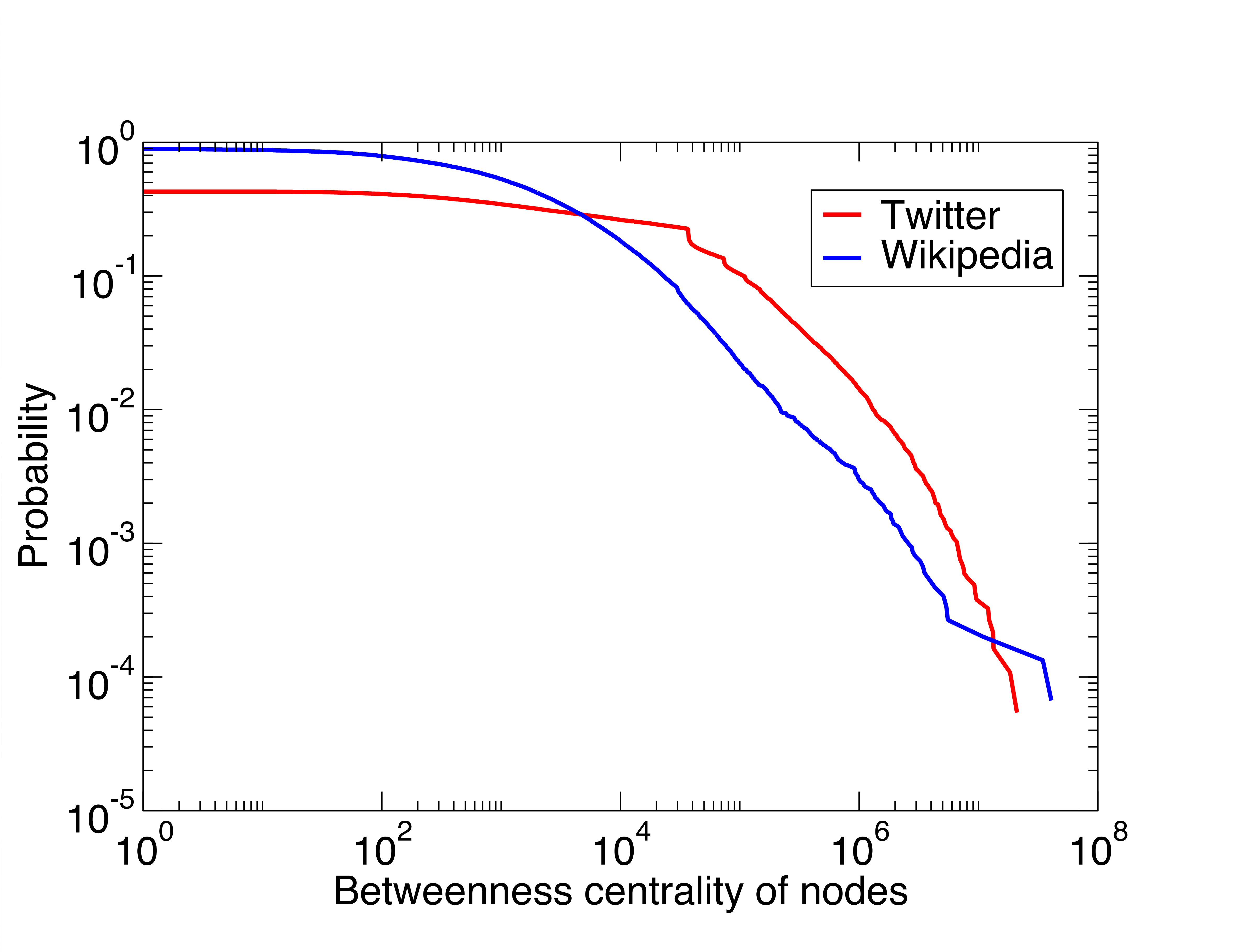
$$\log_{10} 1{,}000{,}000 = \log_{10} 10^6 = 6$$

Degree distributions



Heavy-tail distributions: the variable goes from small to large values

Betweenness distributions



Heavy-tail distribution: the variable goes from small to large values

A First Course in Network Science by F. Menczer, S. Fortunato & C.A. Davis. Cambridge University Press, 2020

Robustness

- A system is **robust** if the failure of some of its components does not affect its function

- **Question:** how can we define the robustness of a network?

- **Answer:** we remove nodes and/or links and see what happens to its structure

- **Key point:** *connectedness*

- If the Internet were not connected, it would be impossible to transmit signals (e.g., emails) between routers in different components

Robustness

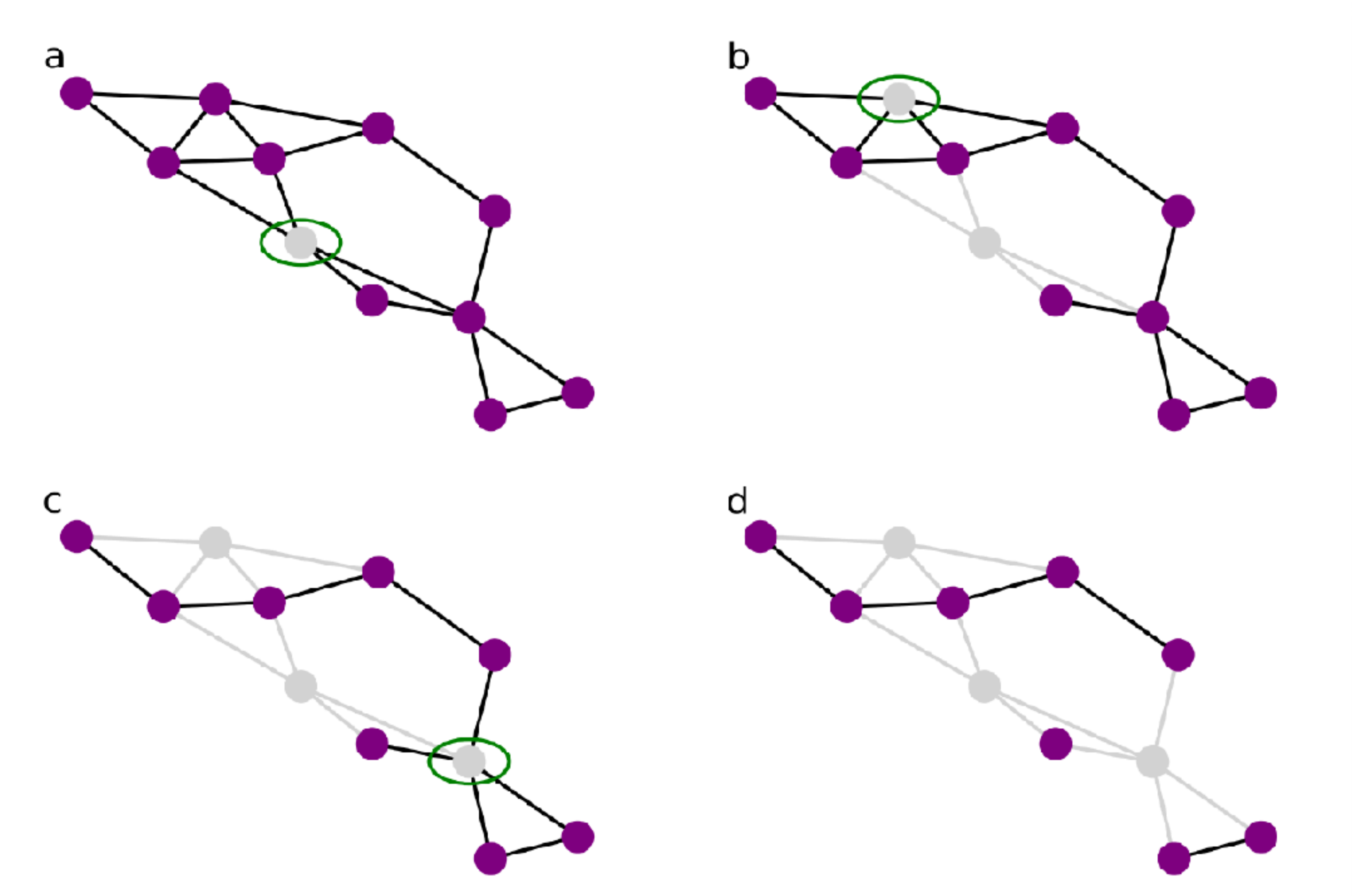
- **Robustness test:** checking how the connectedness of the network is affected as more and more nodes are removed

- **How to do it:** plot the relative size S of the largest connected component as a function of the fraction of removed nodes

- We suppose that the network is initially connected: there is only one component and $S = 1$

- As more and more nodes (and their links) are removed, the network is progressively broken up into components and S goes down

Robustness



A First Course in Network Science by F. Menczer, S. Fortunato & C.A. Davis. Cambridge University Press, 2020

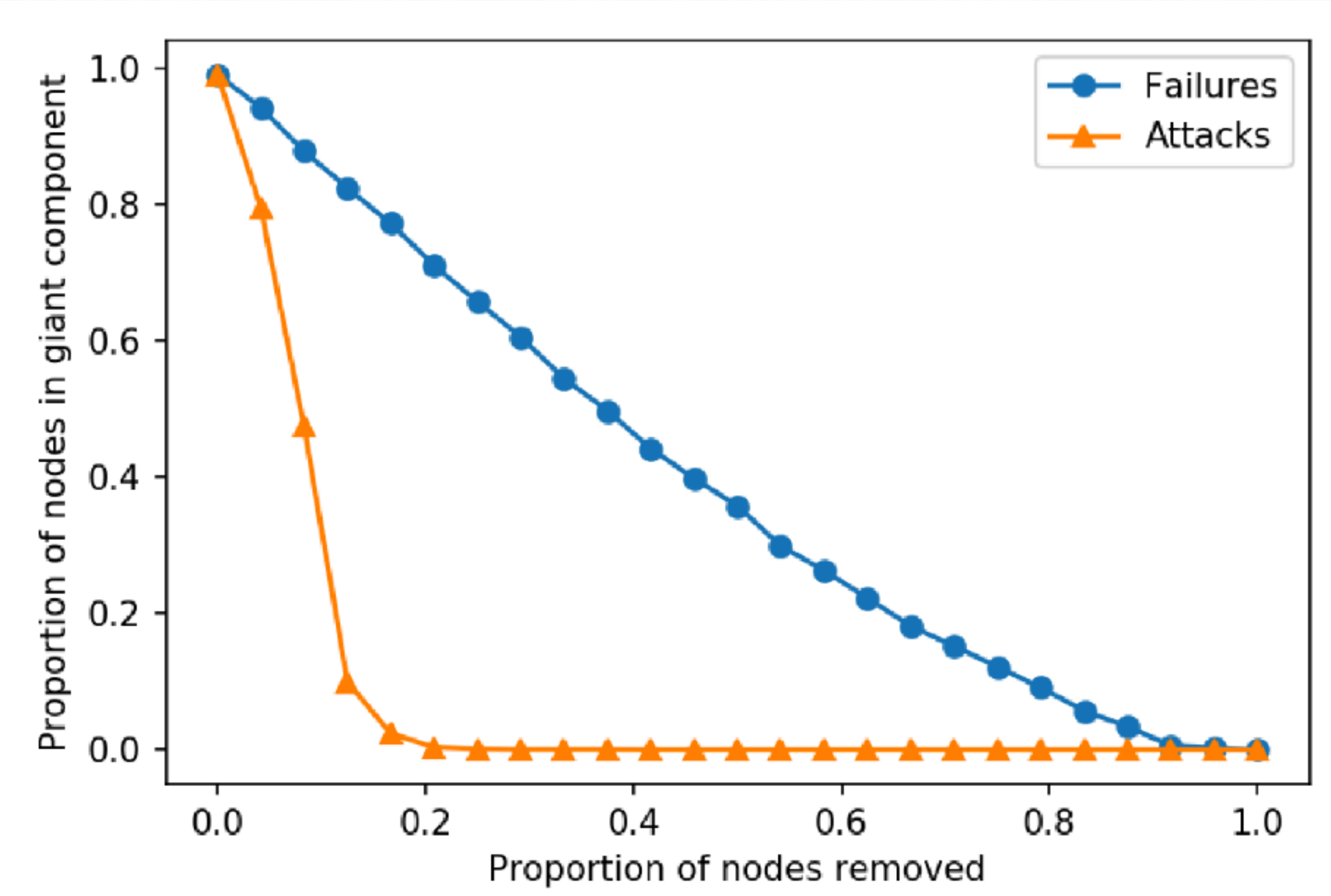
Robustness

- **Two strategies:**
  1. **Random failures:** nodes break down randomly, so they are all chosen with the **same probability**
  2. **Attacks:** hubs are deliberately targeted — the larger the **degree**, the higher the probability of removing the node

- In the first approach, we remove a fraction *f* of nodes, chosen at random

- In the second approach, we remove the fraction *f* of nodes with largest degree, from the one with largest degree downwards

Robustness



Conclusion: real networks are robust against random failures but fragile against targeted attacks!

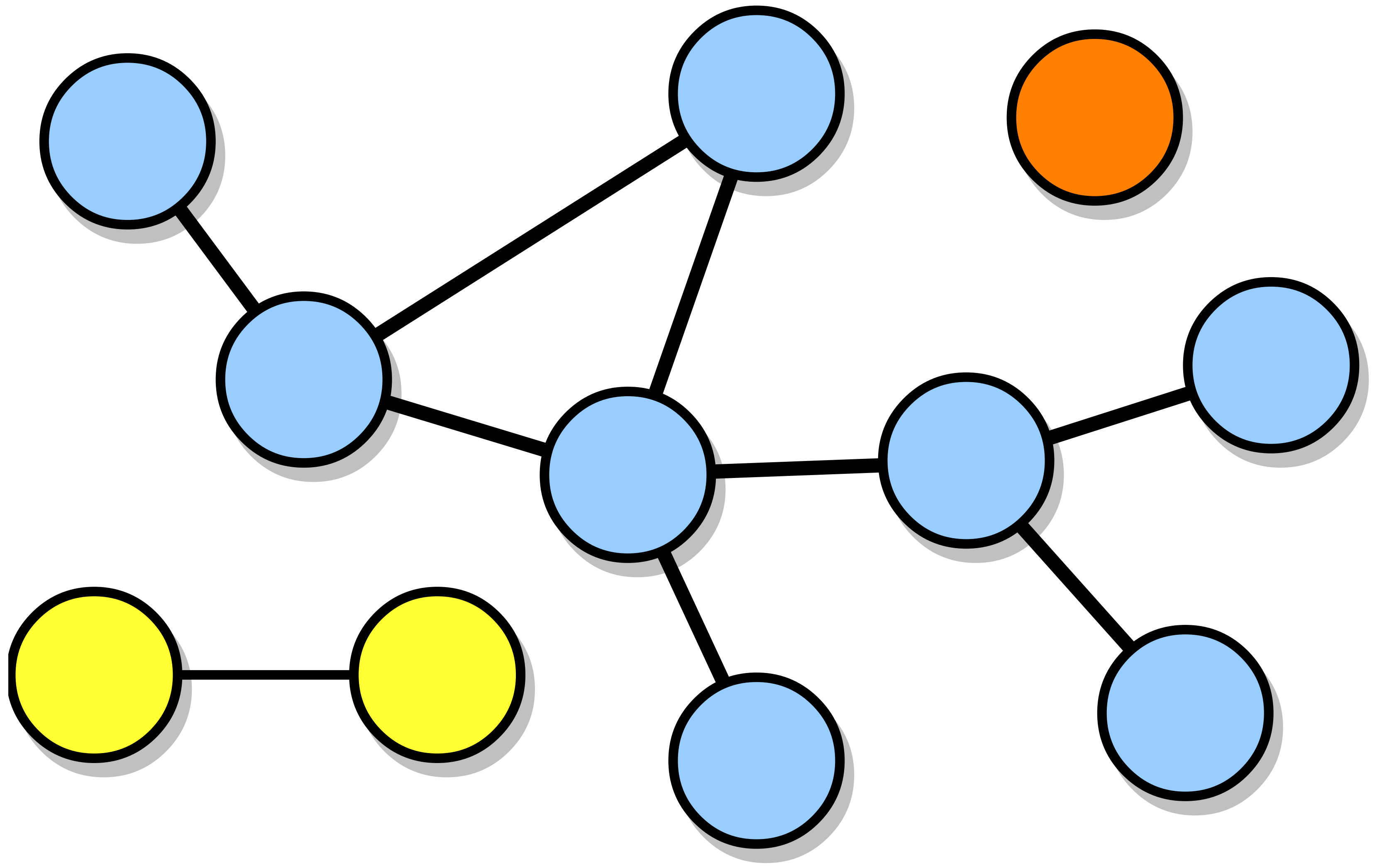
A First Course in Network Science by F. Menczer, S. Fortunato & C.A. Davis. Cambridge University Press, 2020

Pointer to epidemic modeling

- Studying network robustness is a great framework for comparing different immunization (vaccination) strategies
  - this very simple idea has been applied also for mitigating the diffusion of computer viruses
- Problem: real contact network is not usually available…

Connectedness and components

- A network is **connected** if there is a path between any two nodes
- If a network is not connected, it is **disconnected** and has multiple connected components
- A **connected component** is a connected subnetwork
  - The largest one is called **giant component**; it often includes a substantial portion of the network
  - A **singleton** is the smallest-possible connected component



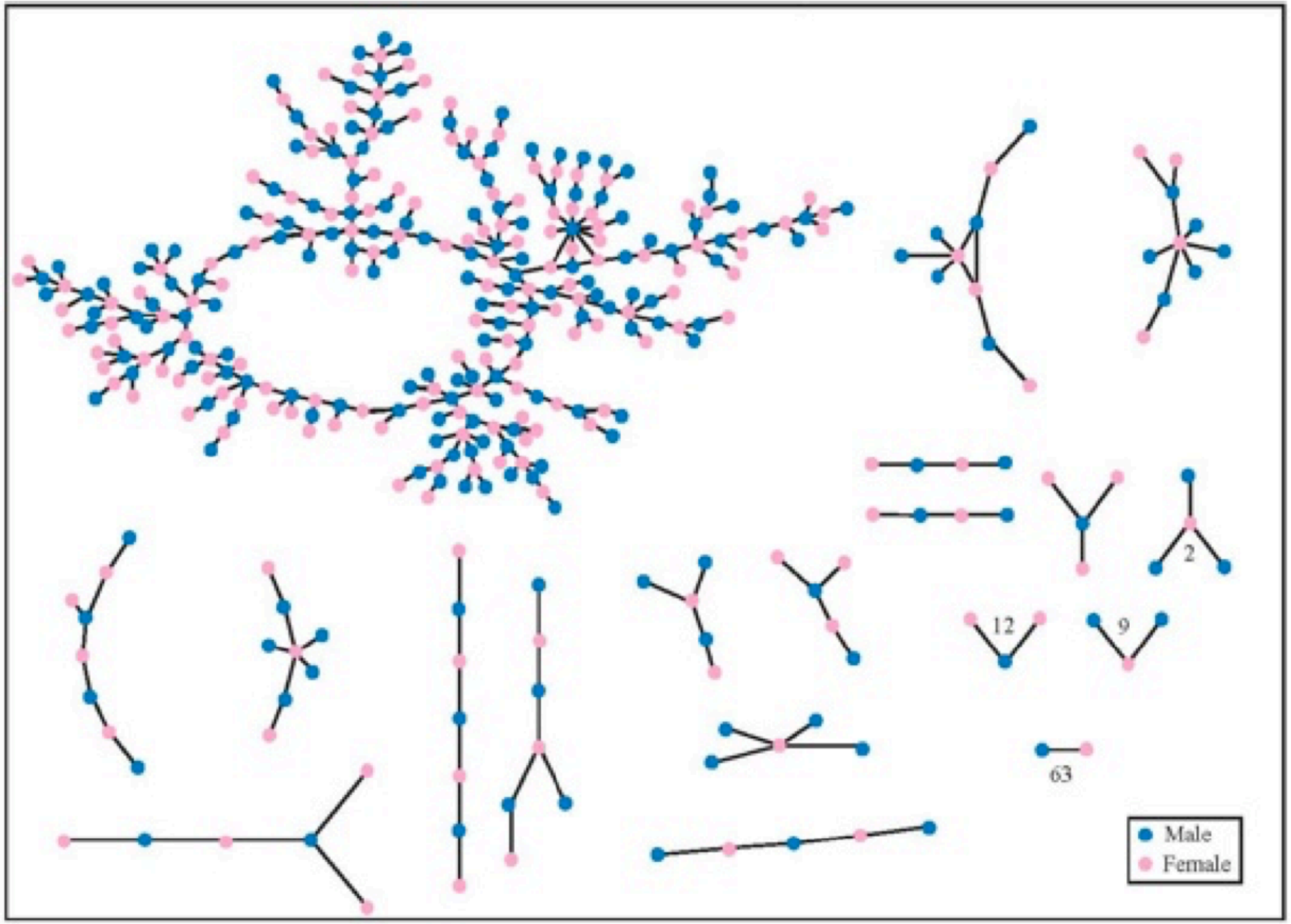
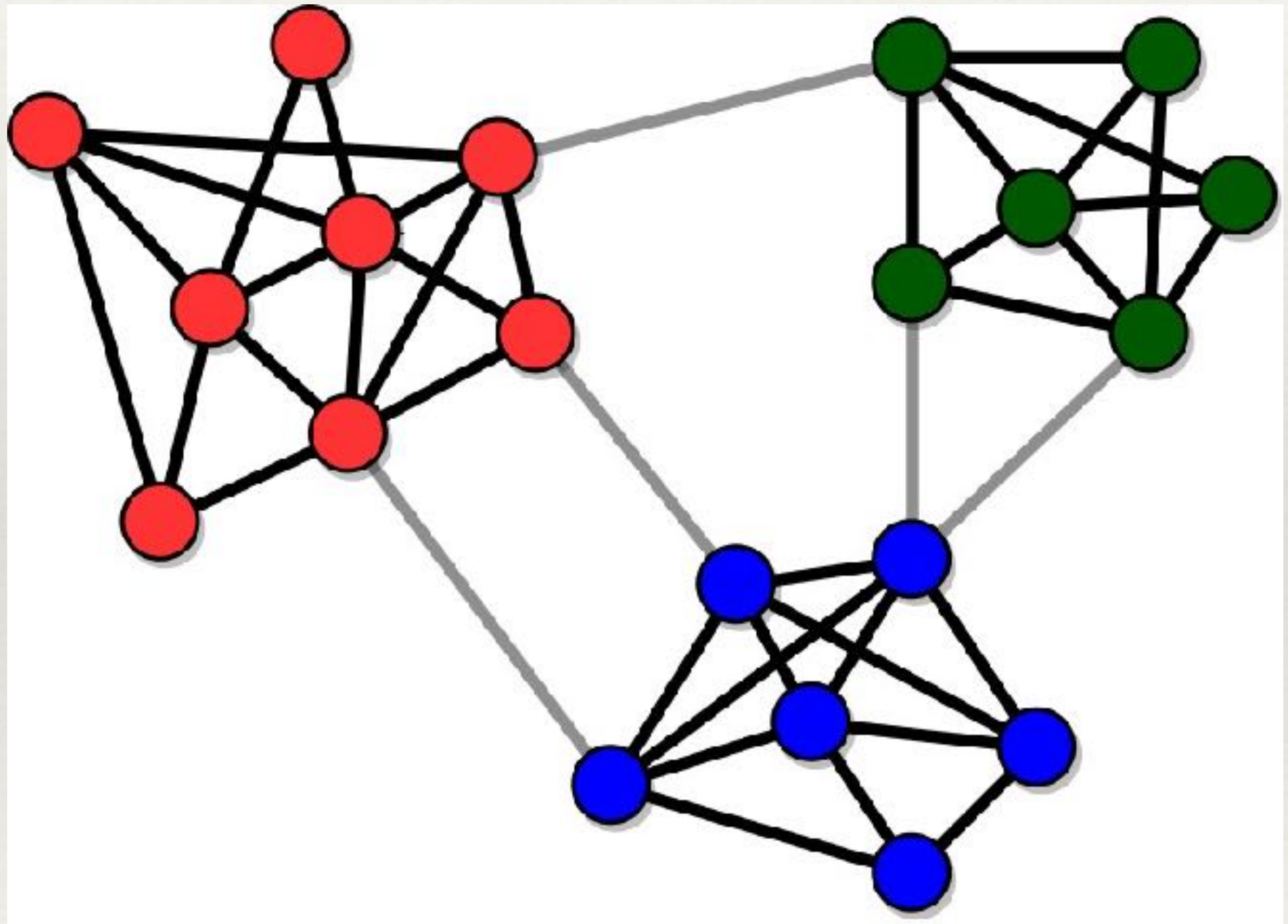


Figure 2.7: A network in which the nodes are students in a large American high school, and an edge joins two who had a romantic relationship at some point during the 18-month period in which the study was conducted [49].

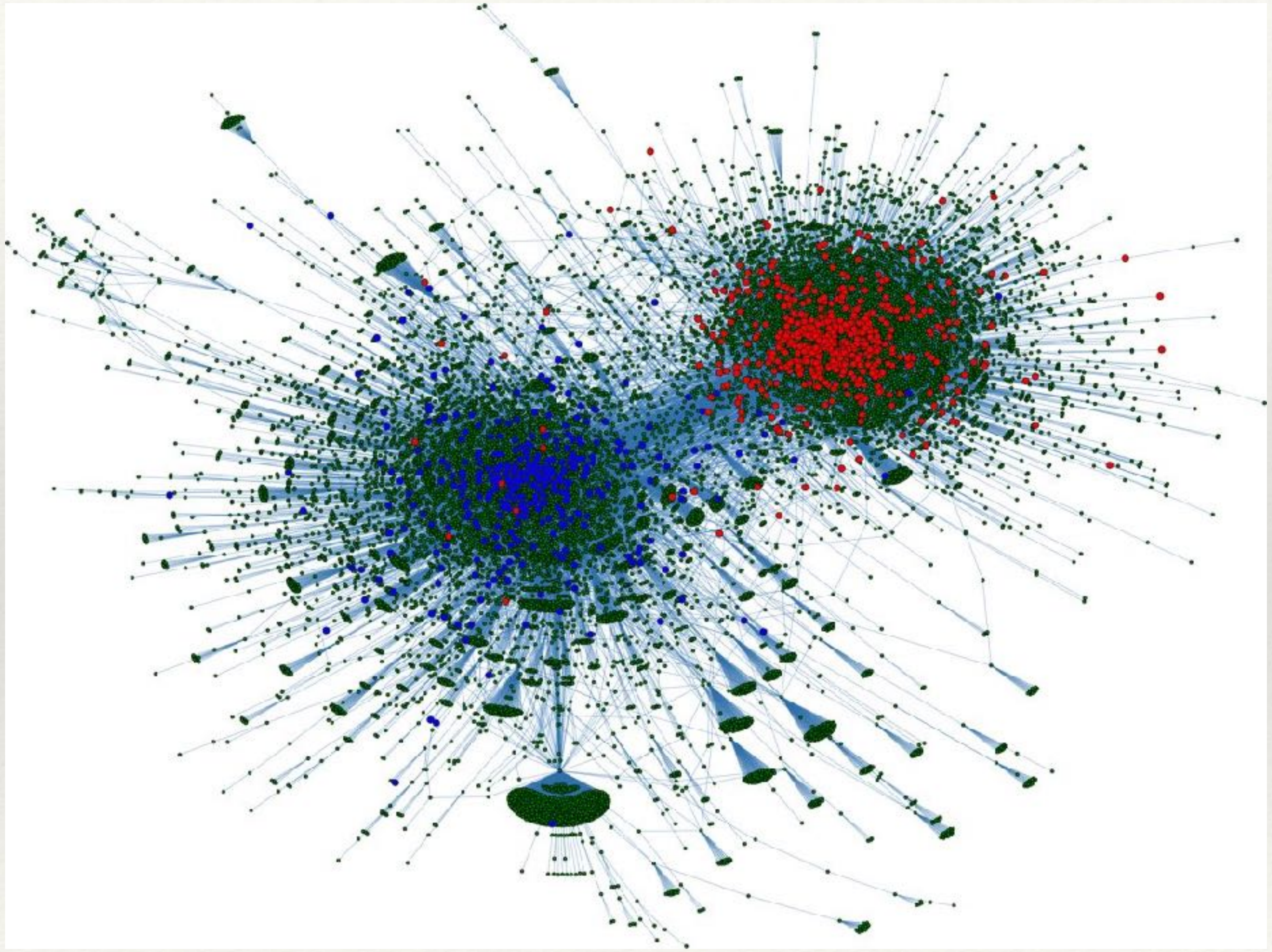
Community structure

Communities (or clusters): sets of tightly connected nodes



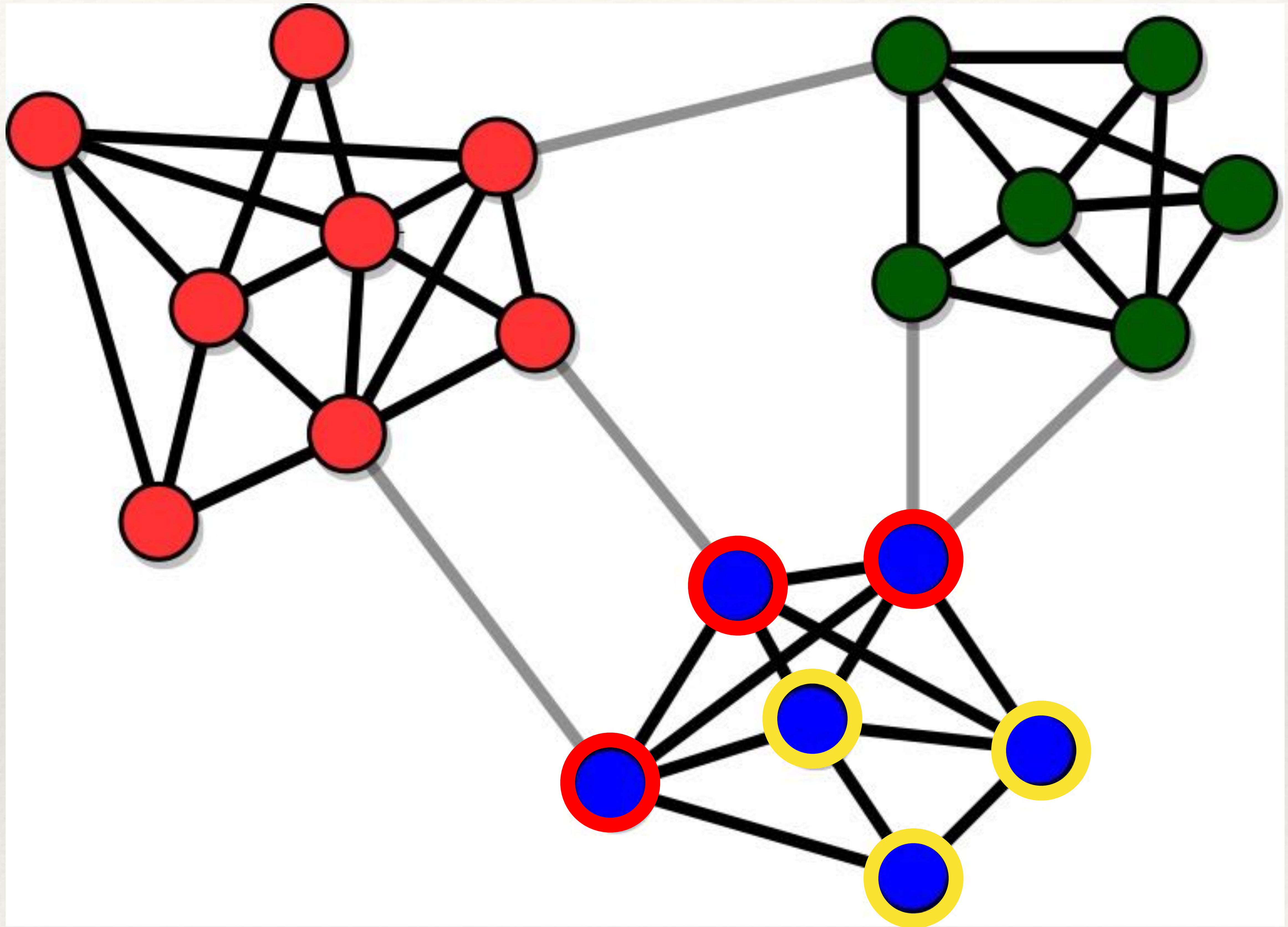
Community structure

- **Example:** Twitter users with strong political preferences tend to follow those aligned with them and not to follow users with different political orientation
- **Other examples:** social circles in social networks, functional modules in protein interaction networks, groups of pages about the same topic on the Web, etc.



Why study communities?

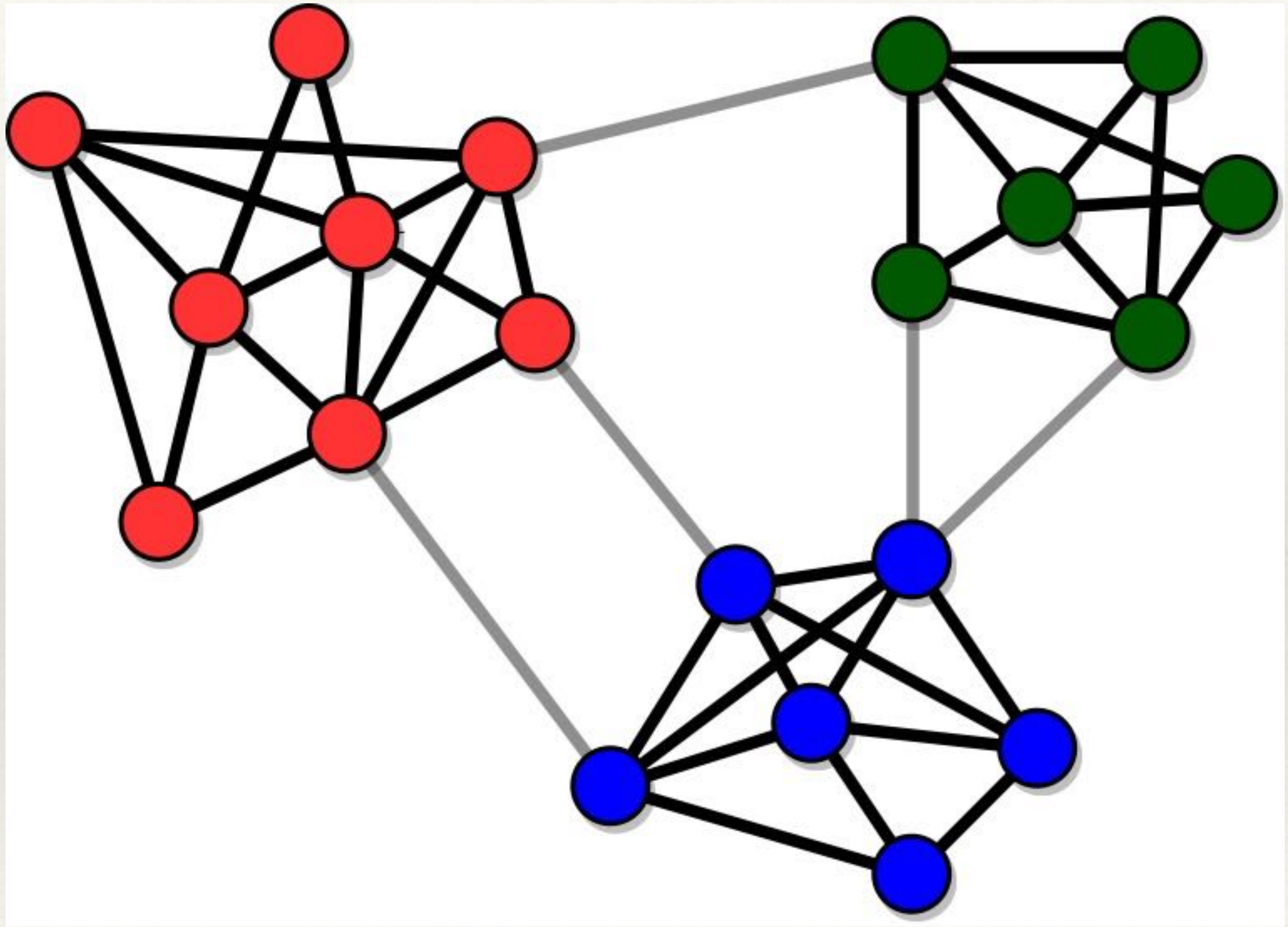
- Uncover the organization of the network
- Identify features of the nodes
- Classify the nodes based on their position in the clusters
- Find missing links



Basic definitions: community

Two main features:

- **High cohesion:** communities have many internal links, so their nodes stick together
- **High separation:** communities are connected to each other by few links



A First Course in Network Science by F. Menczer, S. Fortunato & C.A. Davis. Cambridge University Press, 2020

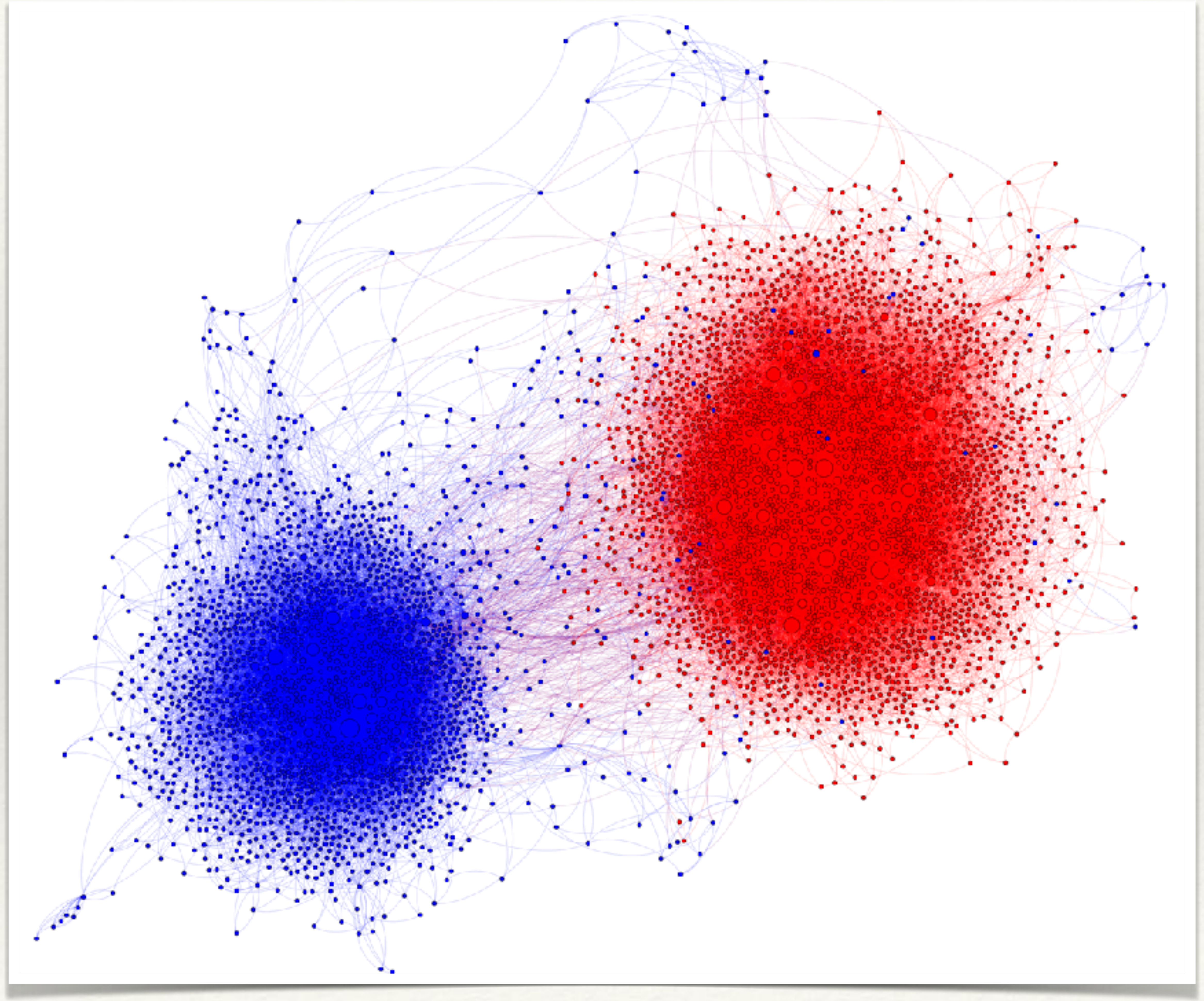
Partitions

- The number of partitions of n objects is the **Bell number *B_n***
- The Bell number **grows faster than exponentially with *n***
- **Conclusion:** it makes no sense to look for interesting community structures by exploring the whole space of partitions! A smart exploration of the partition space must be performed.

n	B_n
1	1
2	2
3	5
4	15
5	52
6	203
7	877
8	4140
9	21147
10	115975
11	678570
12	4213597
13	27644437
14	190899322
15	1382958545

example: retweet networks

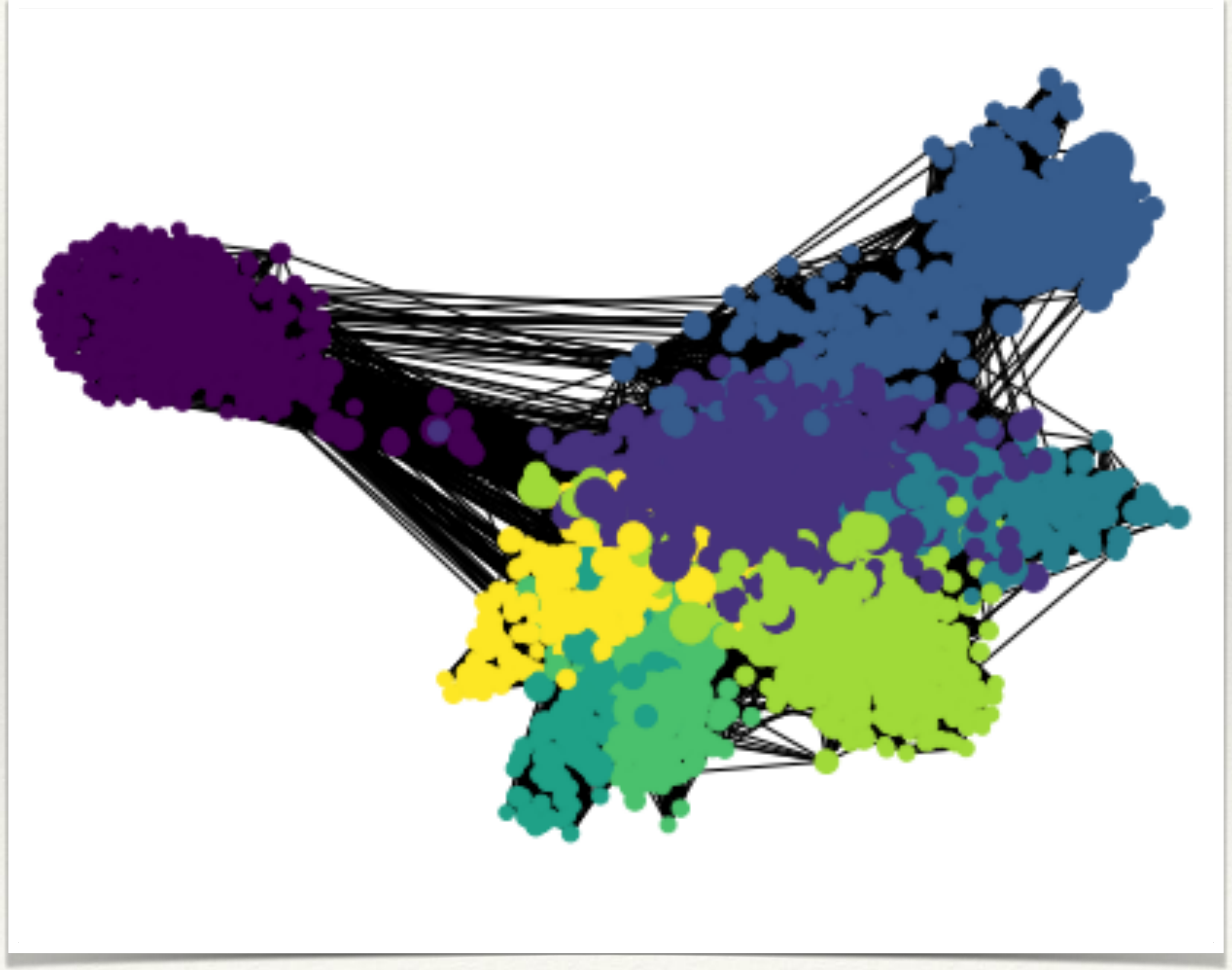
- goal: detecting communities or clusters
- "**echo chambers**"
- **homophily**: tendency of individuals to link with similar ones
- warning: no trivial linear relationships but **interplay**



Picture from: F. Menczer, S. Fortunato, C. A. Davis, A First Course in Network Science, Cambridge University Press, 2020

Drawing networks

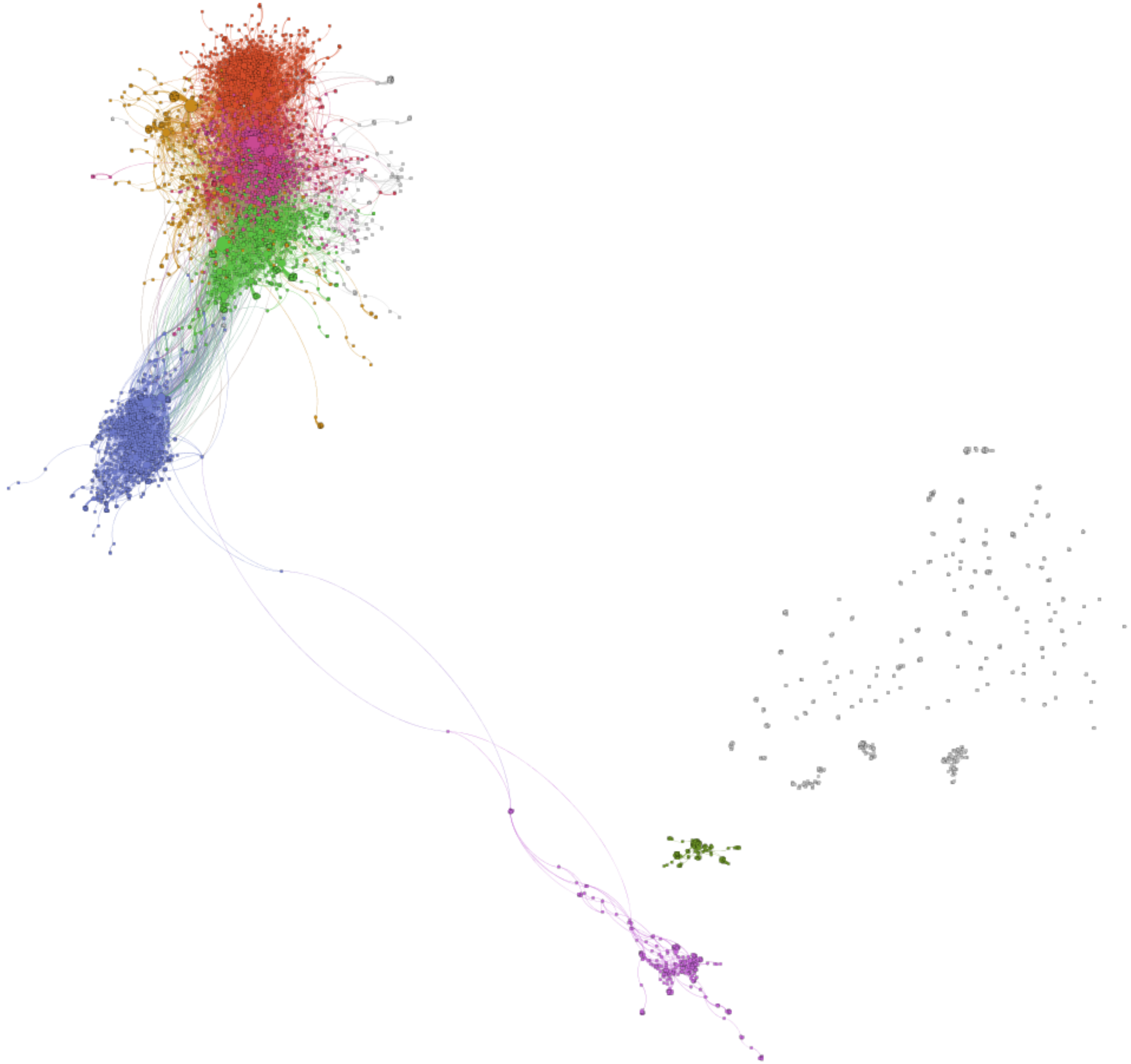
- A **network layout algorithm** places nodes on a plane to visualize the structure of the network
- There are many layout algorithms; the most commonly used are **force-directed layout** (a.k.a. **spring layout**) algorithms:
    - Connected nodes are placed near each other
    - Links have similar length
    - Link crossings are minimized
- This is done by simulating a physical systems where adjacent nodes are connected by springs and otherwise repel each other
- The community structure of the network can be revealed this way if the network is not too dense or too large



Case studies

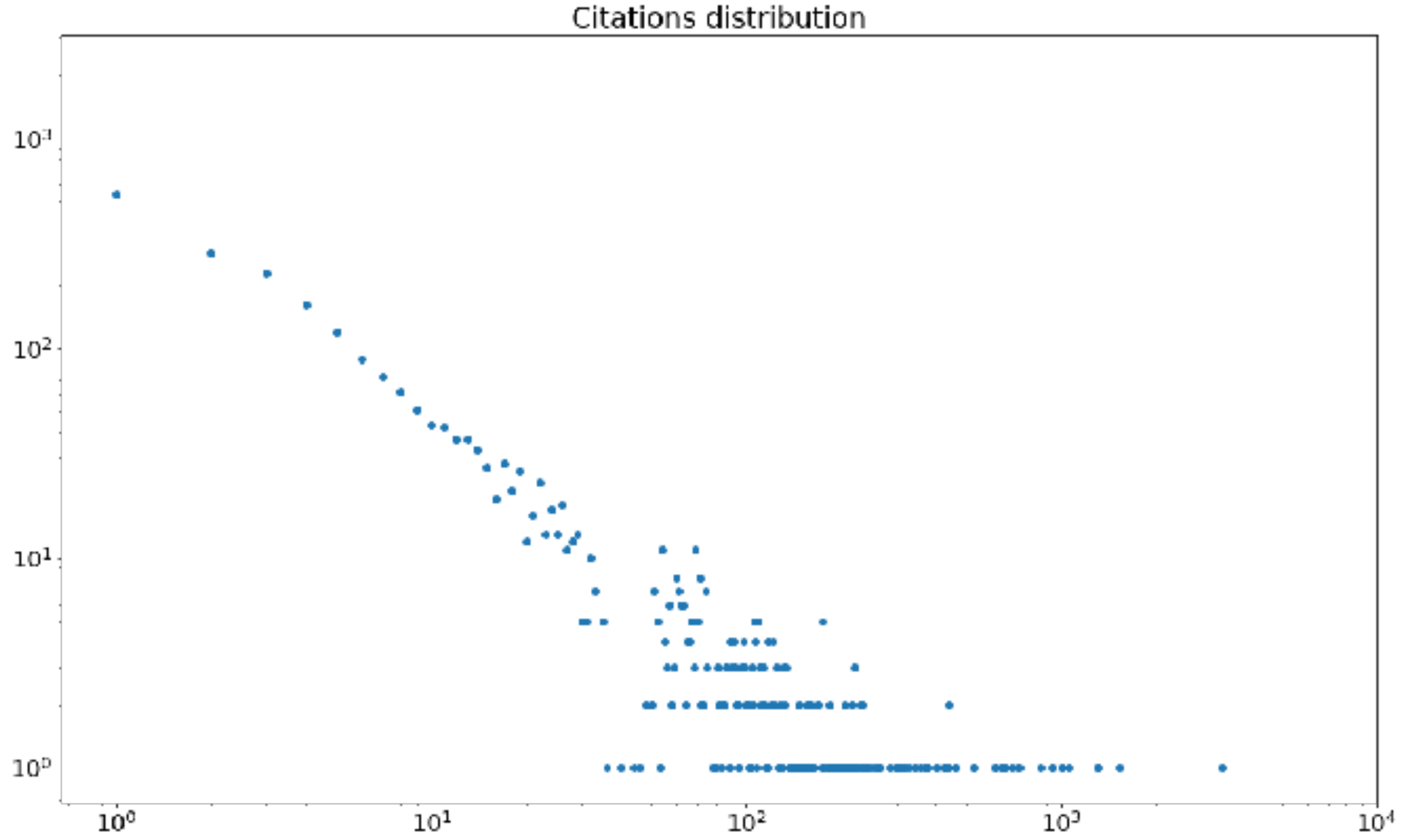
Case study: citation network

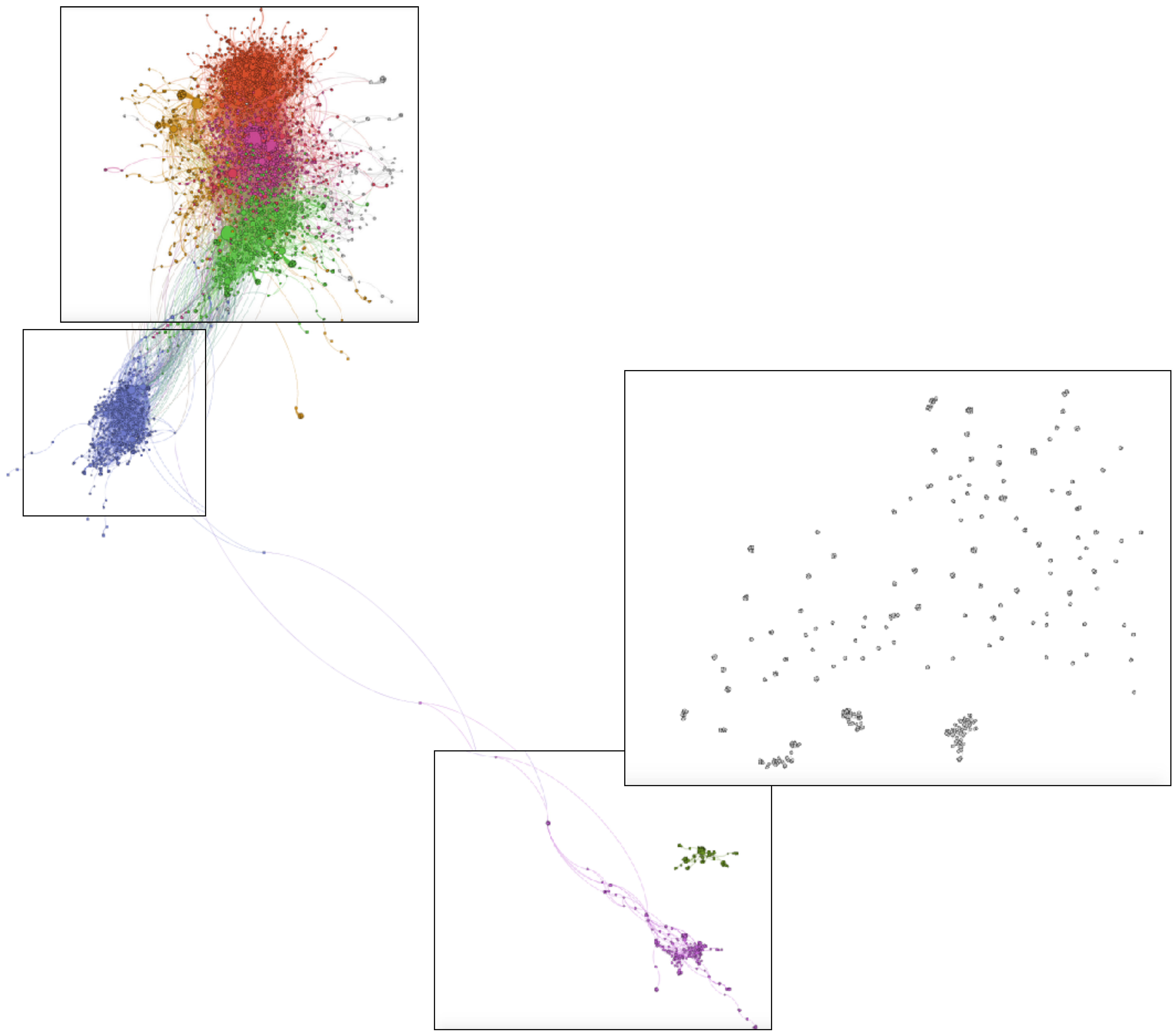
- Let's consider again our collection of scientific papers on fake news
- a citation is a directed link
- we can build a network and analyze its structure

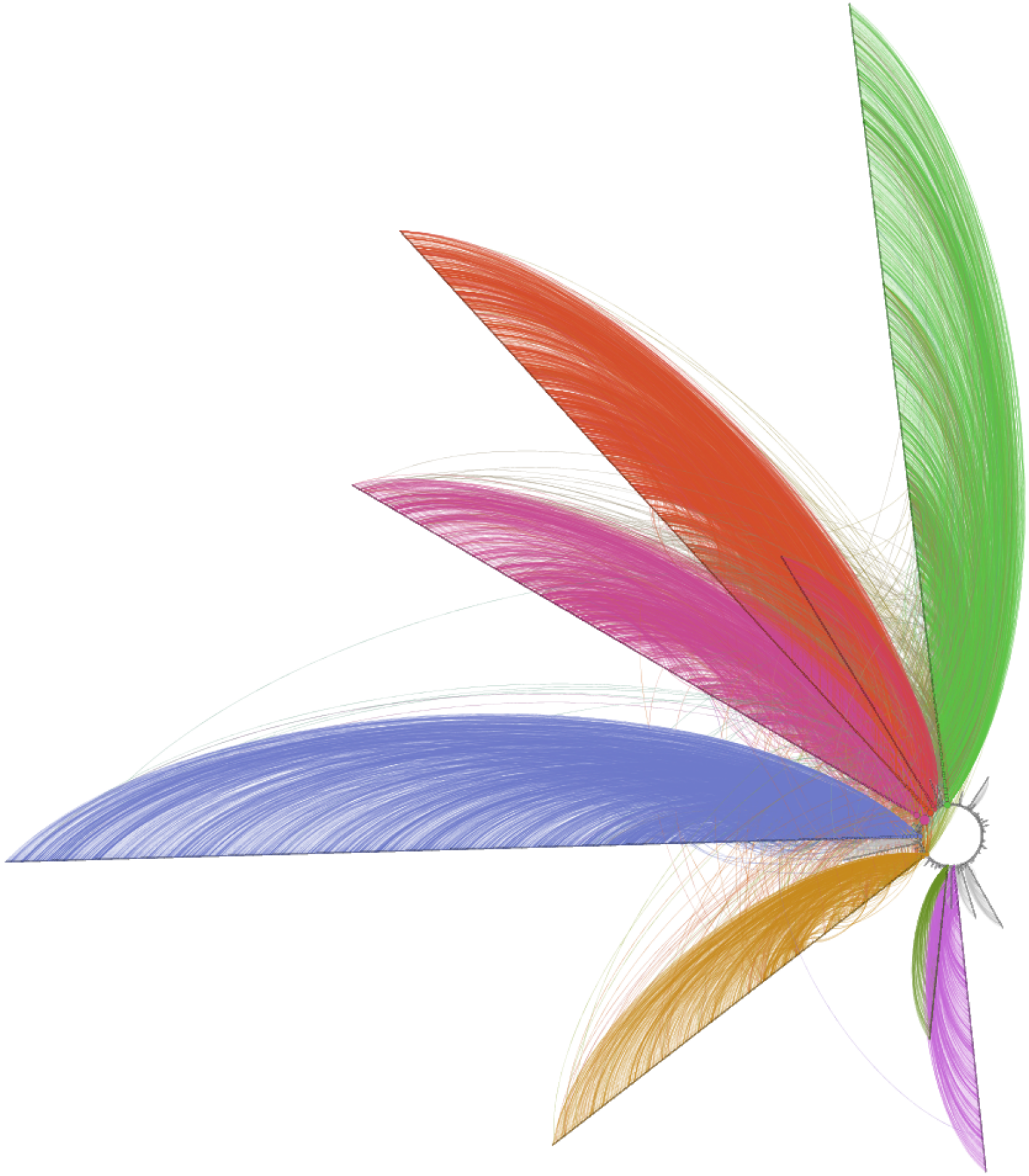


"fake news" citation network

- Force Atlas 2 Layout
- 8 biggest clusters
- Highly heterogeneous (both in terms of in-degree than of disciplines / venues)

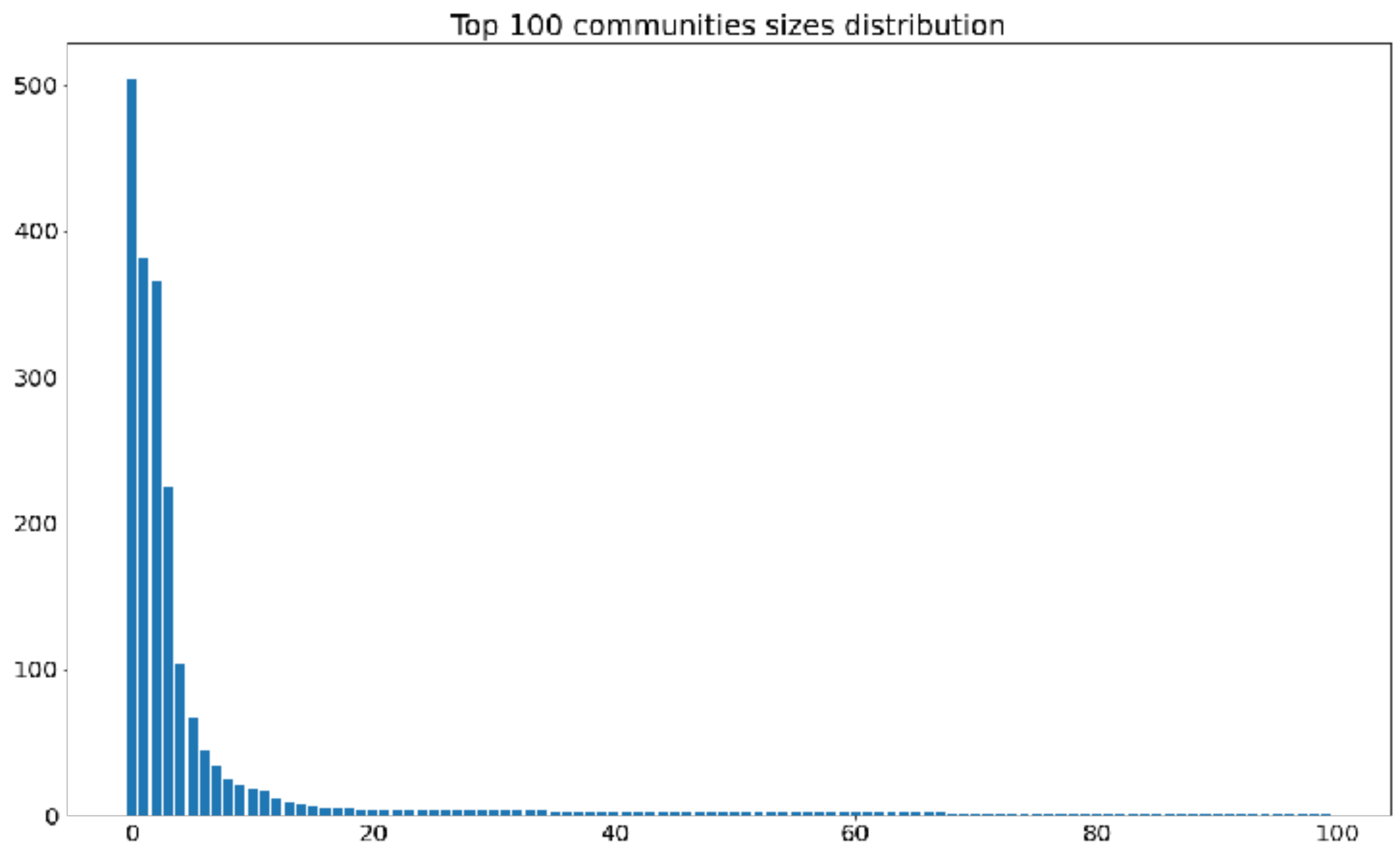






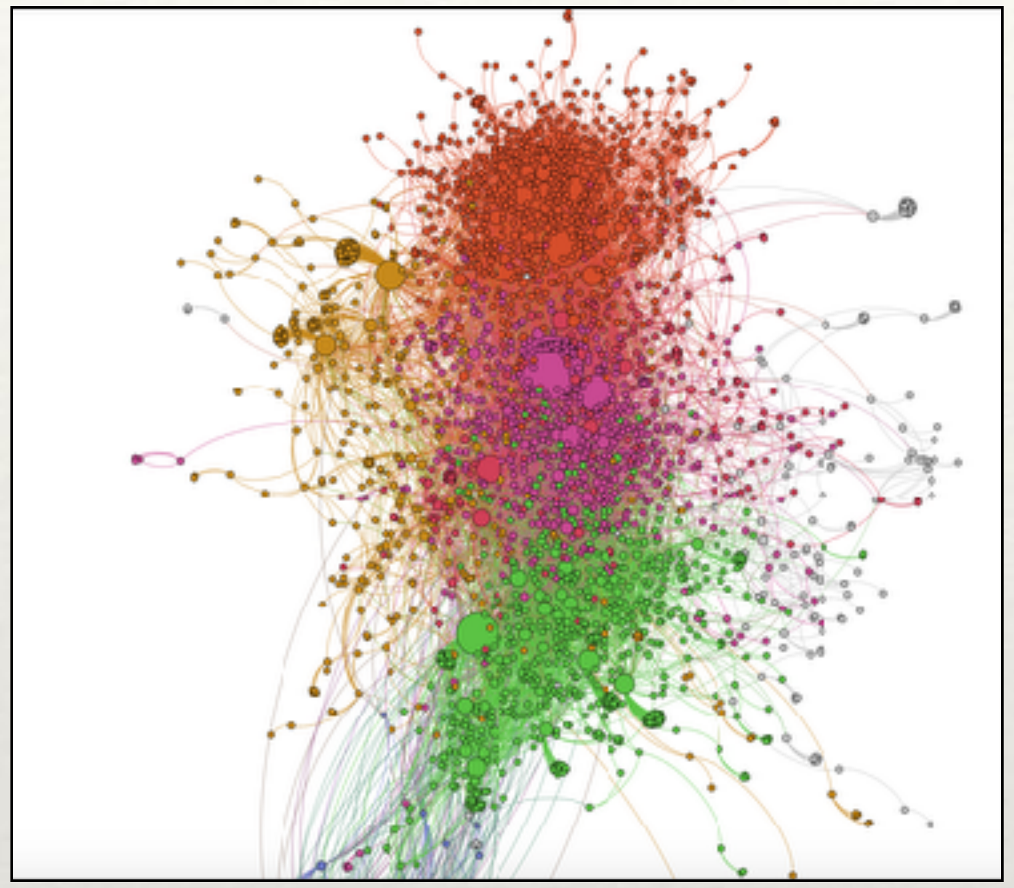
"fake news" citation network

- Axis Radial Layout
- intra vs inter clusters connections
- significantly different sizes



Largest clusters analysis

- Cluster 2: 365
- Cluster 3: 225
- Cluster 4: 103
- Cluster 6: 45
- Cluster 1: 382
- Cluster 0: 504
- Cluster 5: 67
- Cluster 7: 33



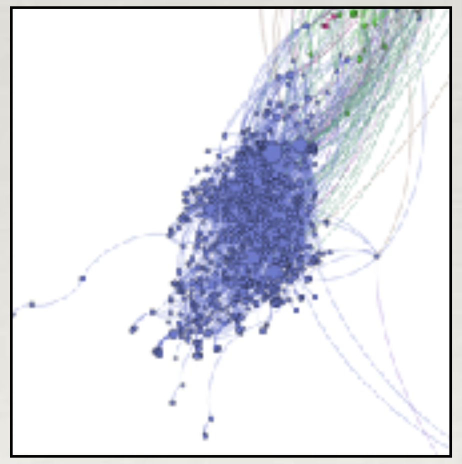
Data Mining
Neural Networks

Identification
Rumors
Algorithms

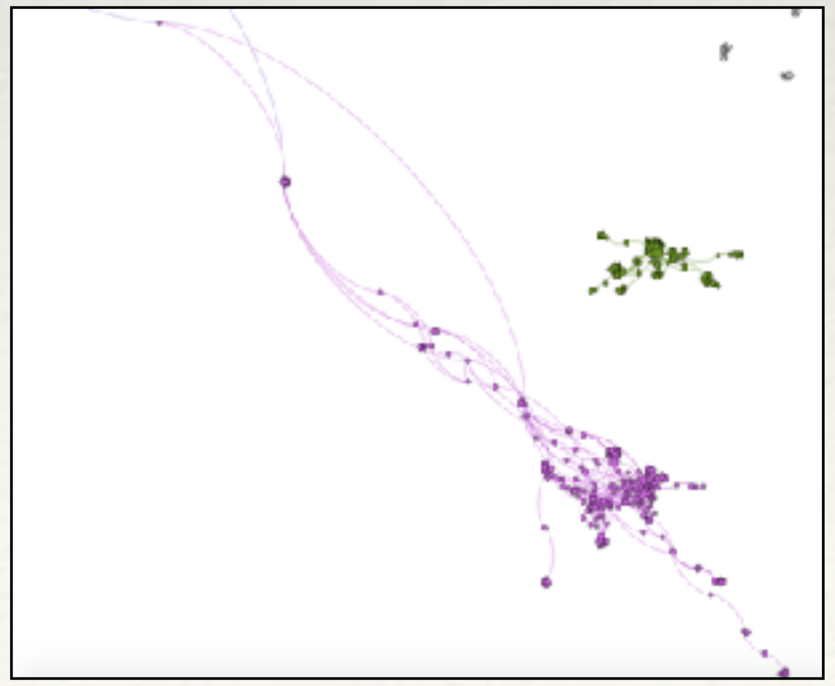
Spreading

Spreading
Fact-checking
Networks
Social media data
Elections

Psychology
Influence
Media



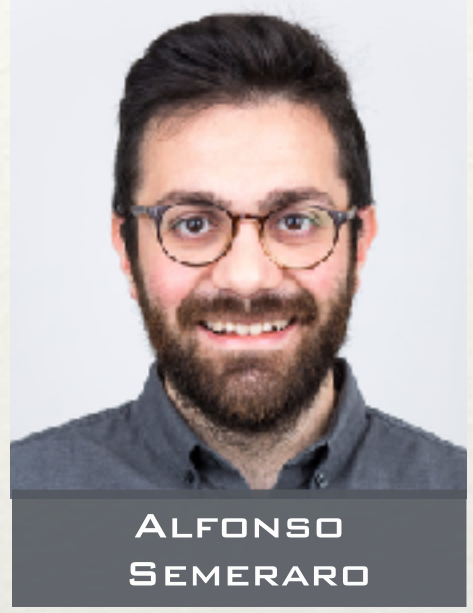
Memory
Cognitive Biases



Health
Autism
Disorders
Vaccines

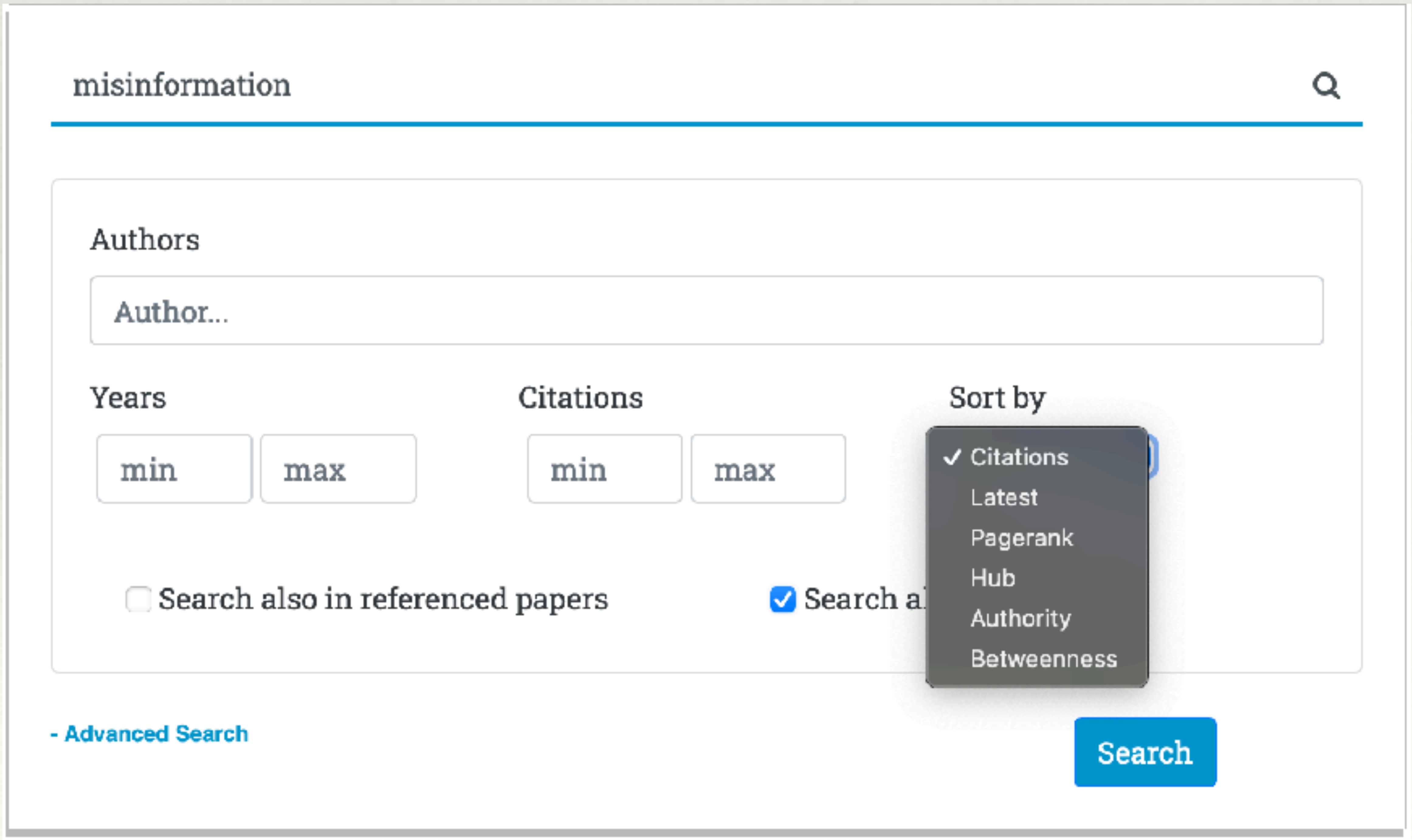
Search and research

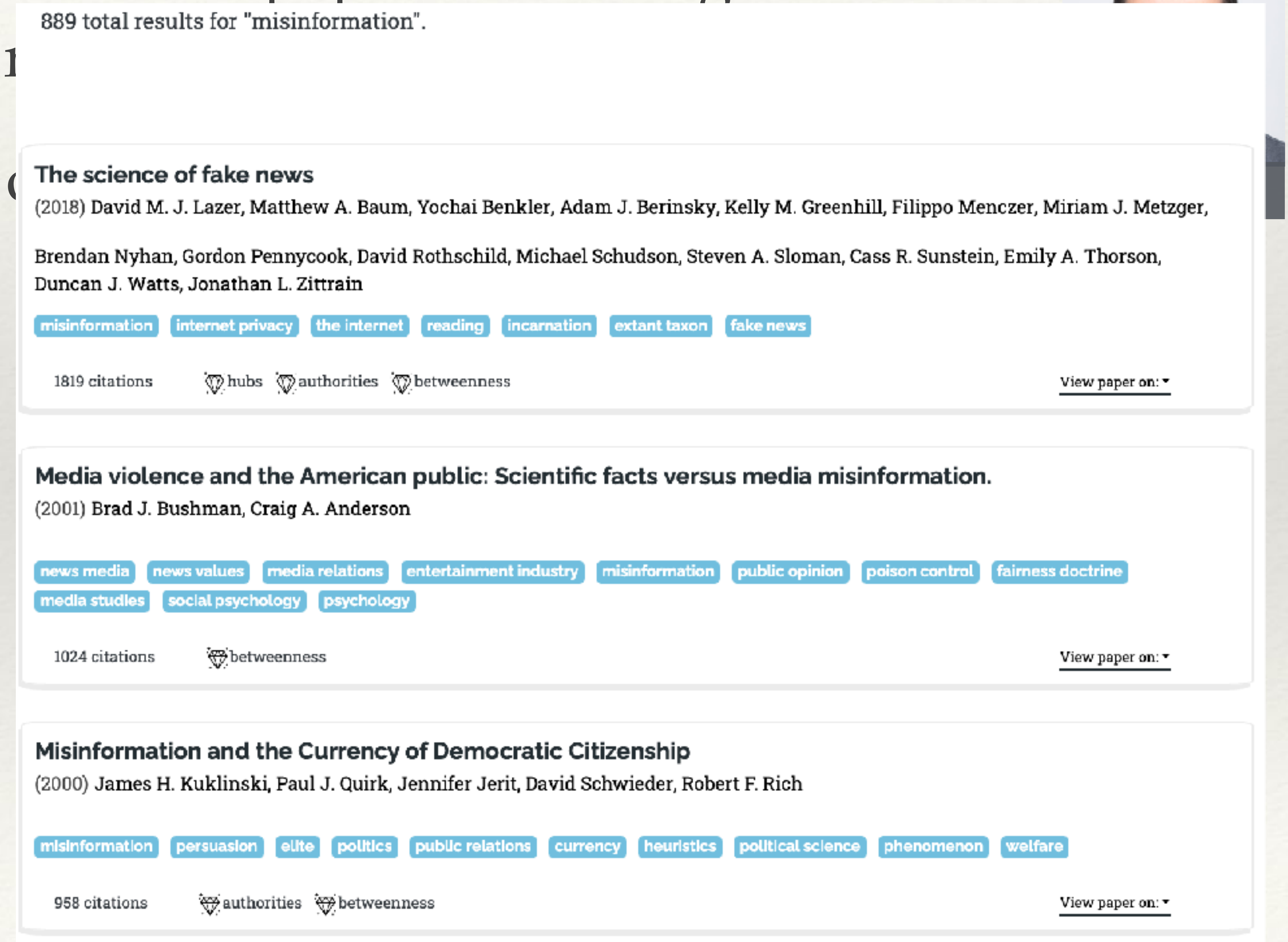
- Citation analysis allows us to identify 'relevant' papers according different metrics: in-degree, betweenness, page rank, hub/authority score
- We embedded these sorting criteria in our 'fakenewsresearch' search engine



Search and research

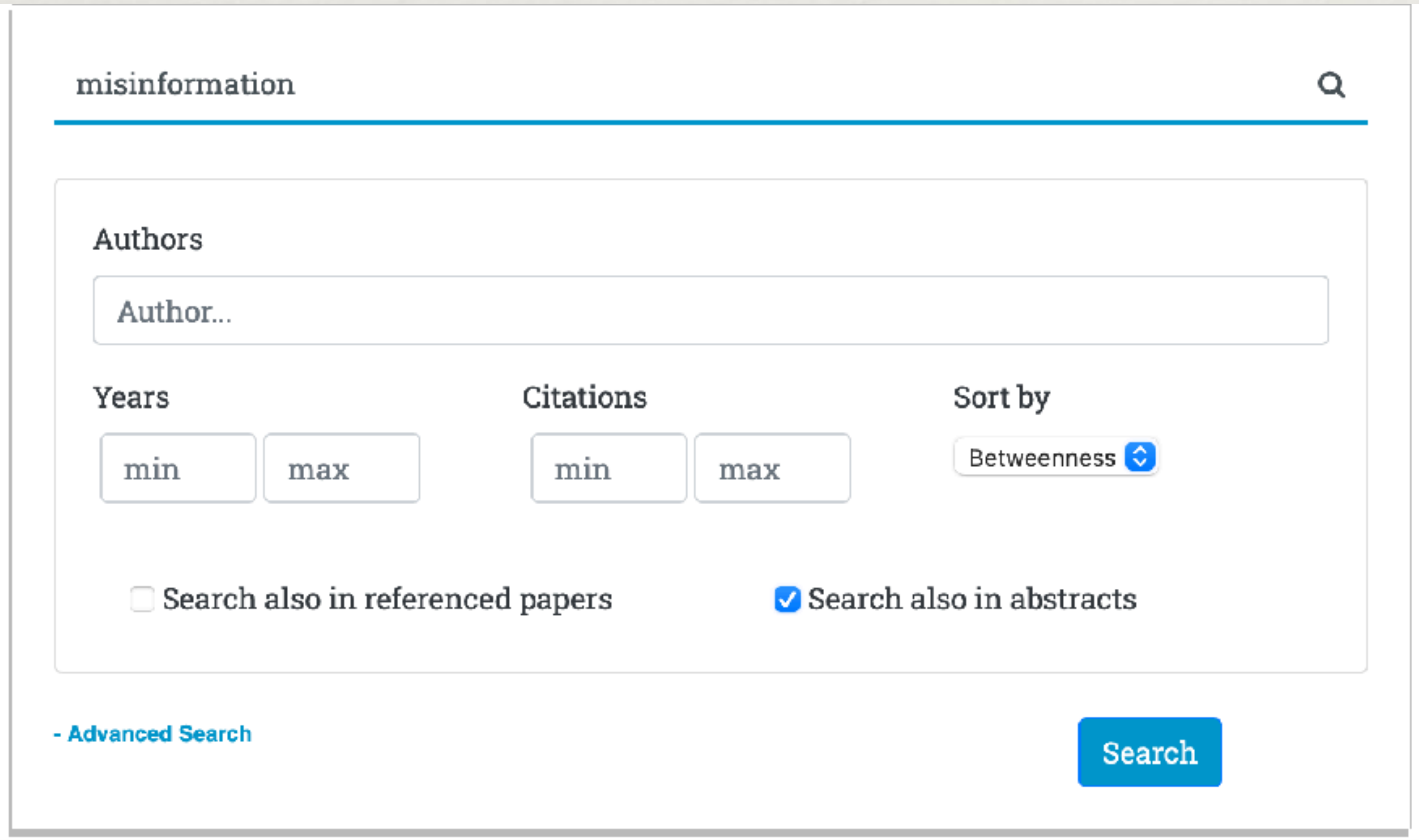
- Citation analysis allows us to identify 'relevant' papers according different metrics: in-degree, betweenness, page r
- We embedded these sorting criteria in c

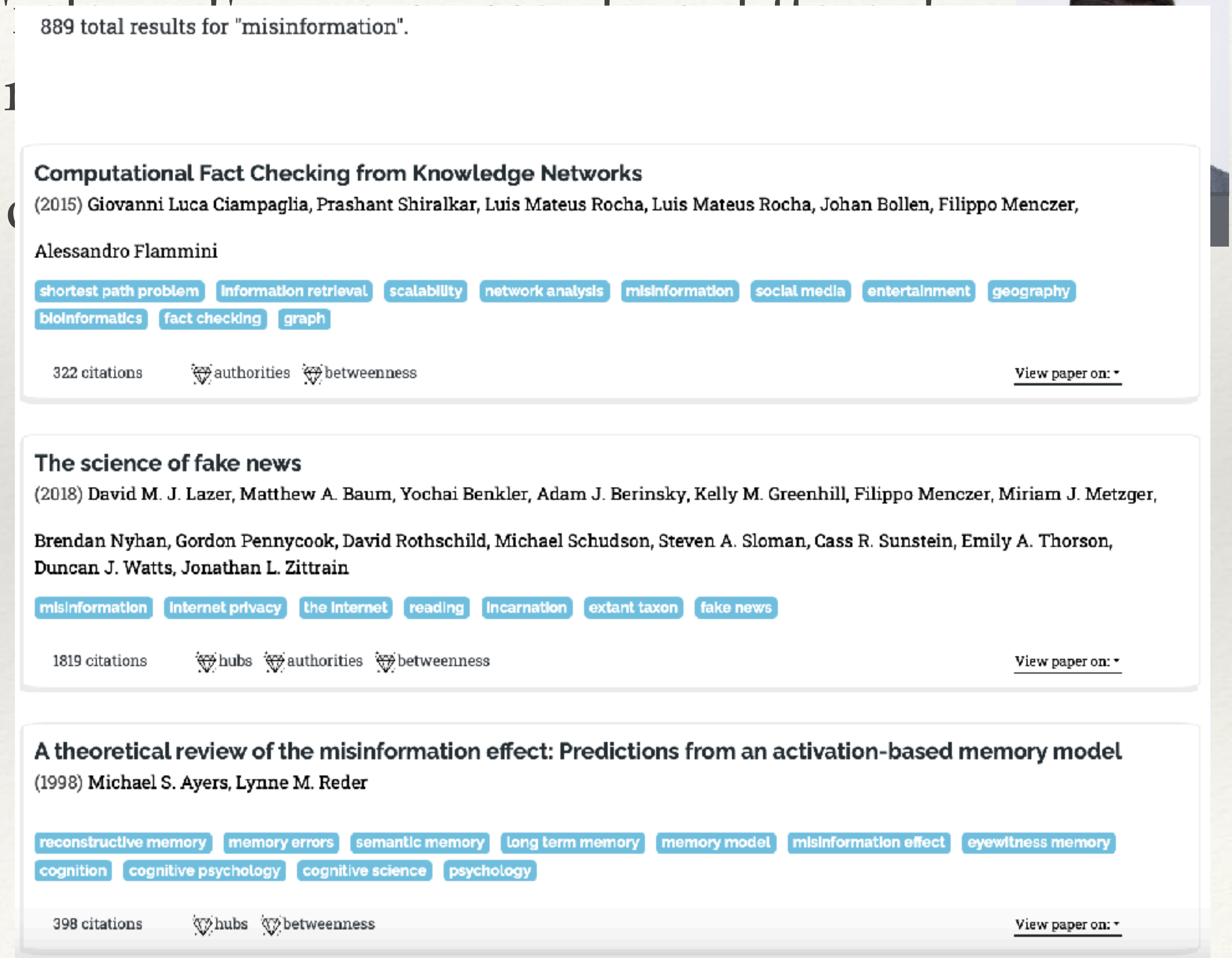




Search and research

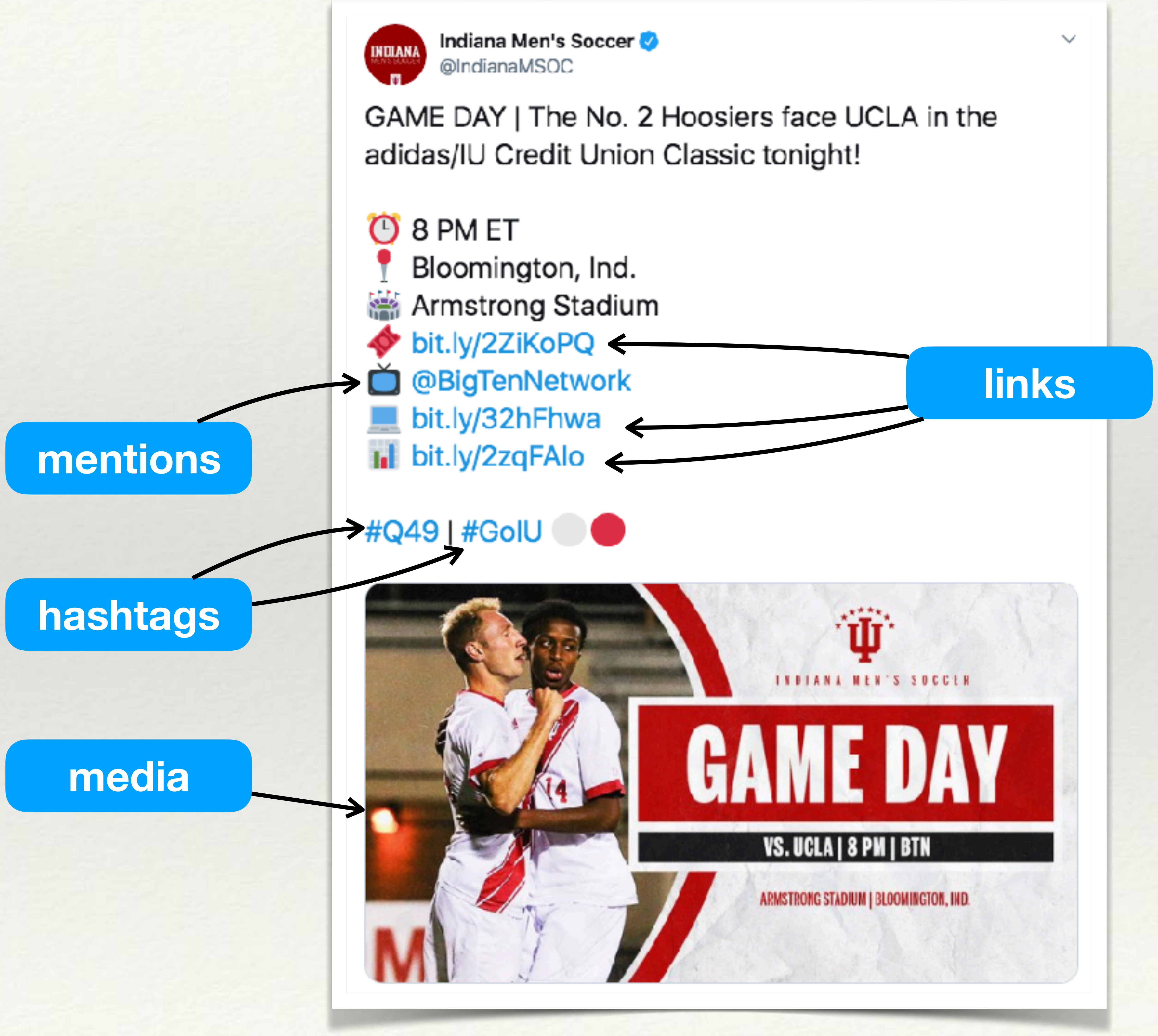
- Citation analysis allows us to identify '
metrics: in-degree, betweenness, page r
- We embedded these sorting criteria in c





Case Study: Information diffusion networks

- **Memes**: transmissible units of information, such as ideas, behaviors, news links, hashtags, and, yes, also images with captions (**image macros**)
- The definition of meme is due to Richard Dawkins, in analogy to genes transmitted from parent to offspring
- Like genes, memes can mutate and have fitness
- A tweet can carry several memes



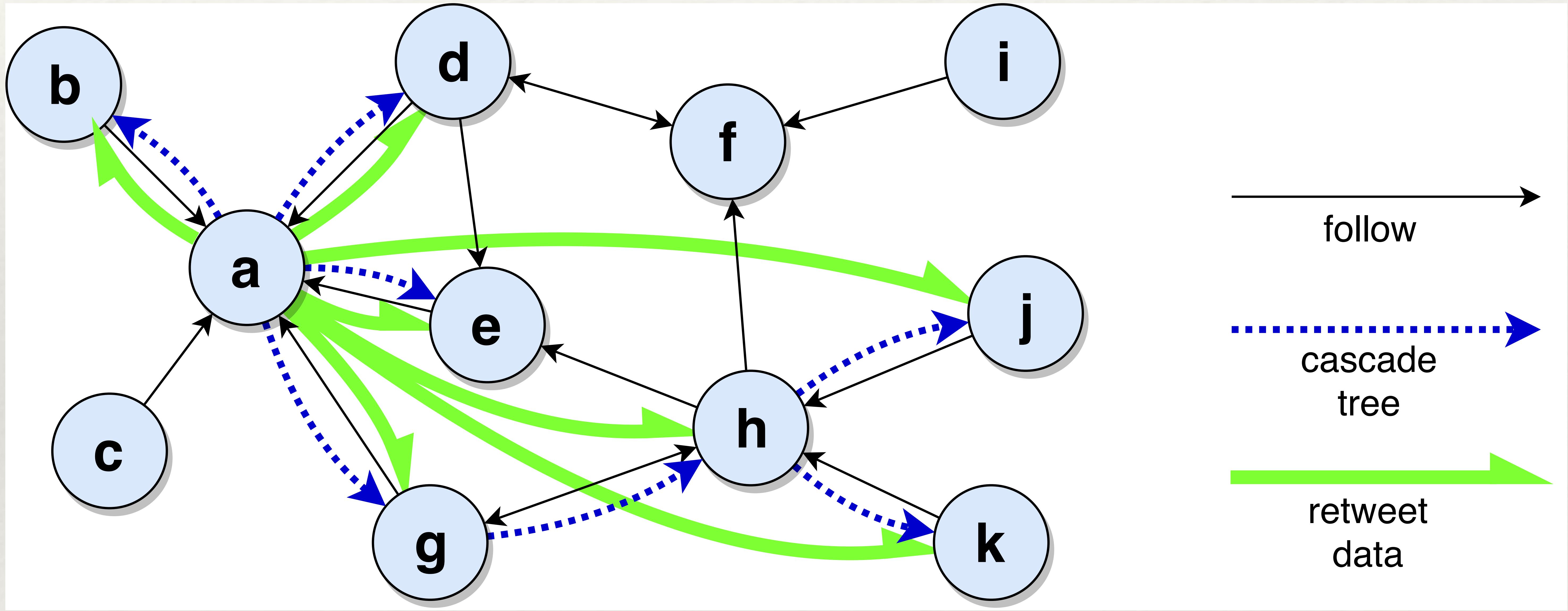
Networks from Twitter

- We can track, map, and analyze the spread of memes on **Twitter**
  - **Retweet network**: link from retweeted user to retweeter user
  - **Mention/reply network**: link to user who replies or who is mentioned
- Tweets are time-stamped; we can aggregate the temporal networks
- Can focus on a particular meme (eg, a hashtag) or multiple ones (eg, a set of accounts or links to a news source)

Play with the interactive diffusion network tools at osome.iuni.iu.edu

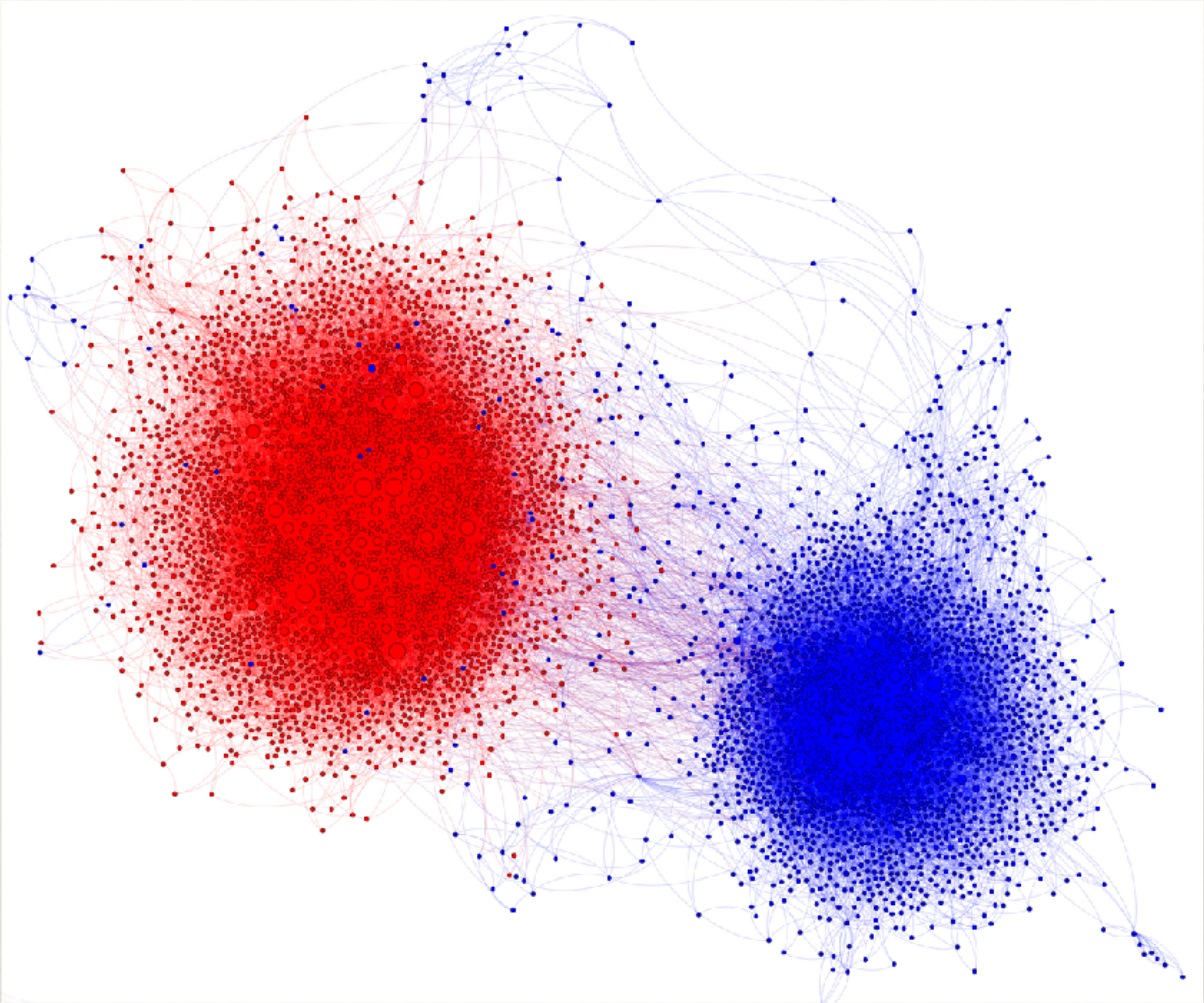
Retweet networks

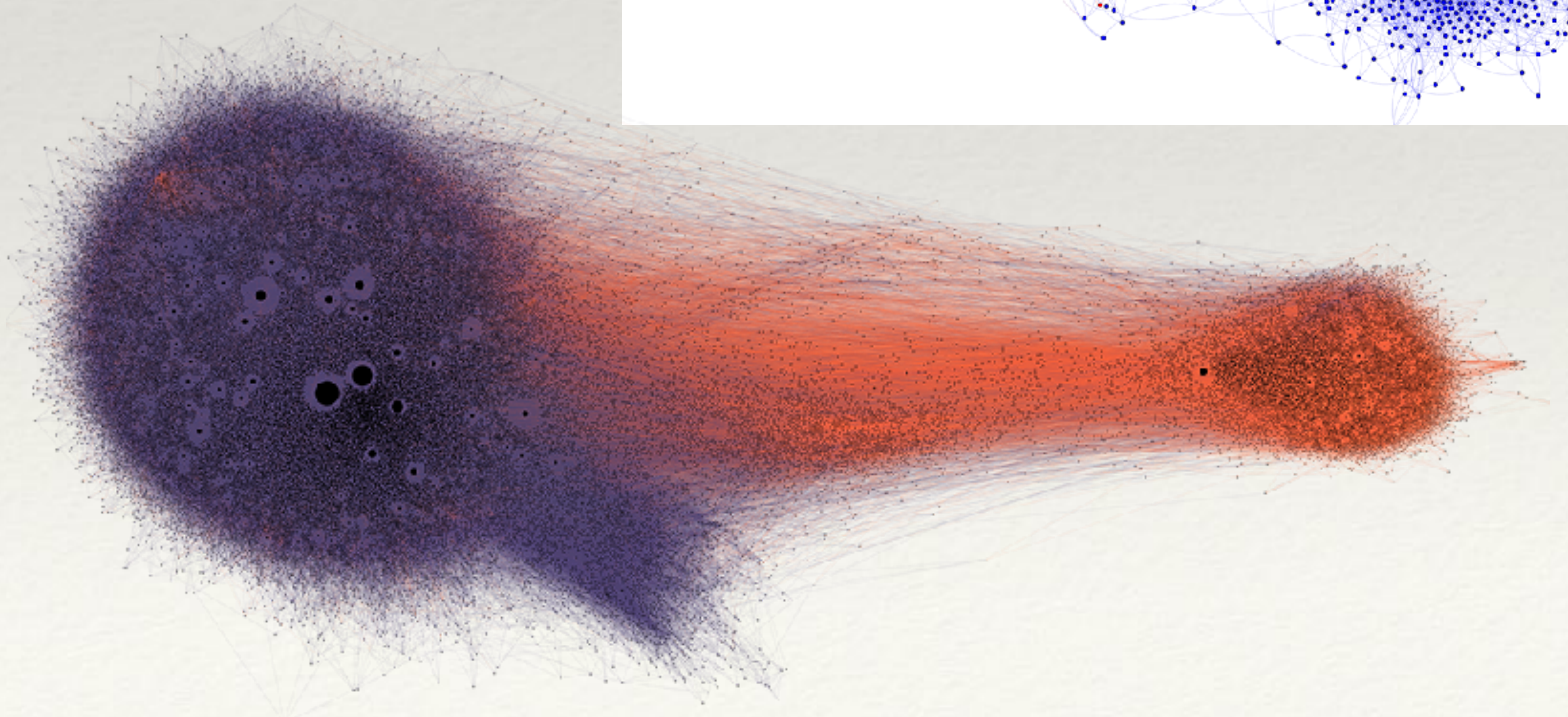
- In the data, each retweet cascade network is a star (all retweets point to original tweet)
- The actual cascade tree is difficult to reconstruct, but we can make some guesses based on the follower network and timestamps



Echo chambers

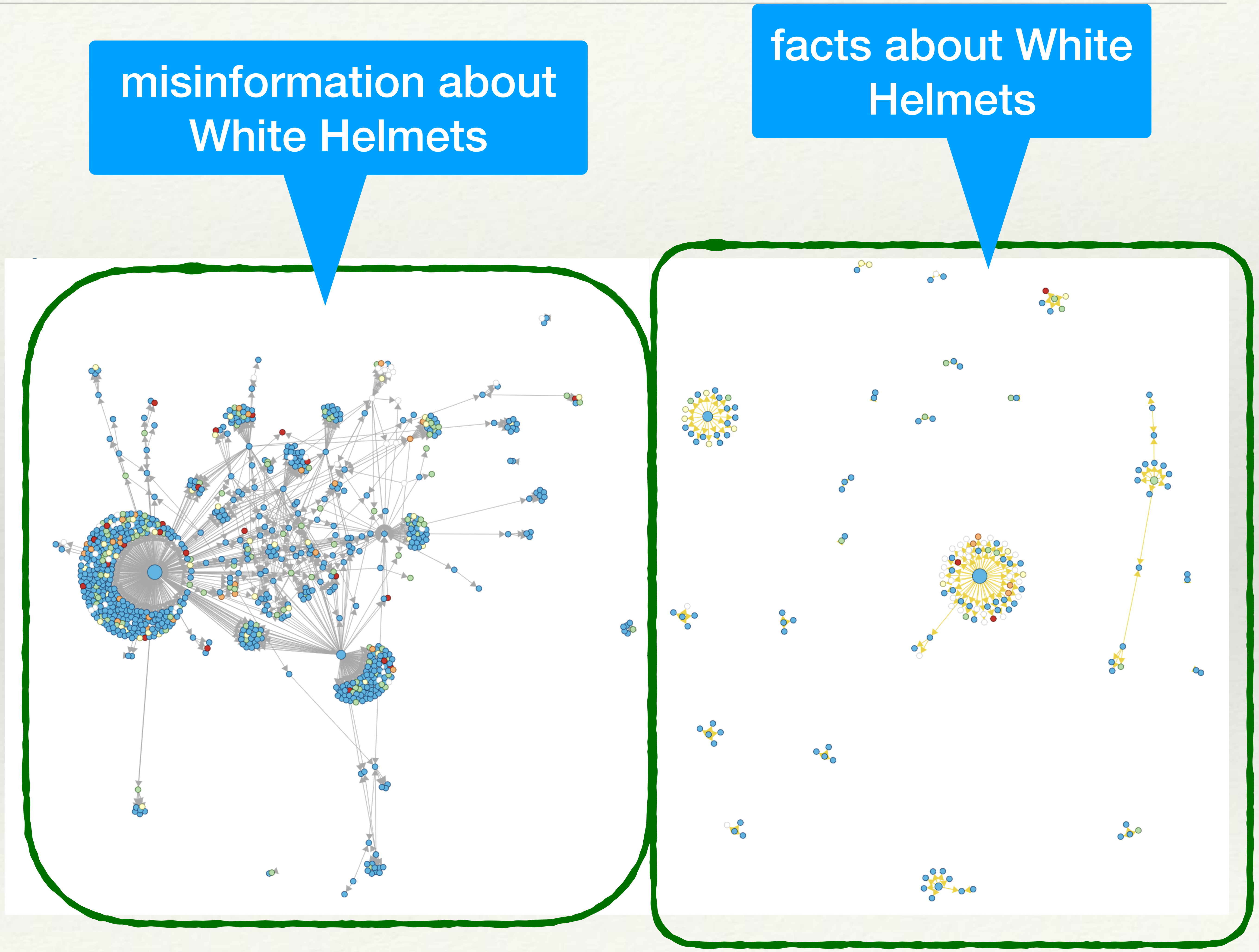
- Examples:
  - Retweets of tweets with progressive (blue) and conservative (red) political hashtags during 2010 US election (*k*=3 core)
  - Retweets of tweets with links to low-credibility (purple) and fact-checking (orange) sources during 2016 US election (*k*=5 core)





Virality

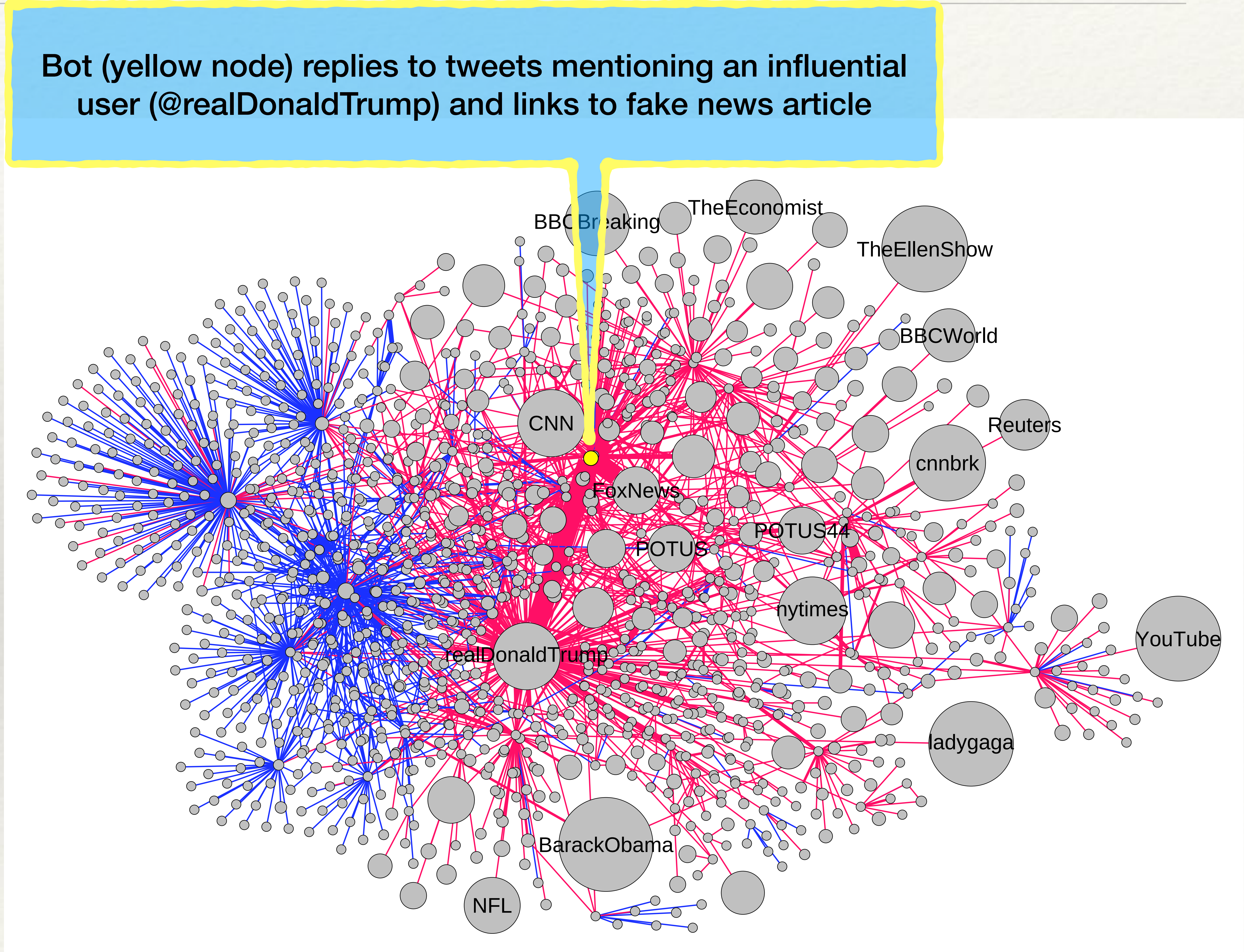
- Multiple ways to measure the virality of a meme:
  - Number of users exposed
  - Depth of diffusion tree
  - Fraction of users who retweet to users who are exposed
- Misinformation is often more viral than actual news reports



Source: hoaxy.iuni.iu.edu

Influence

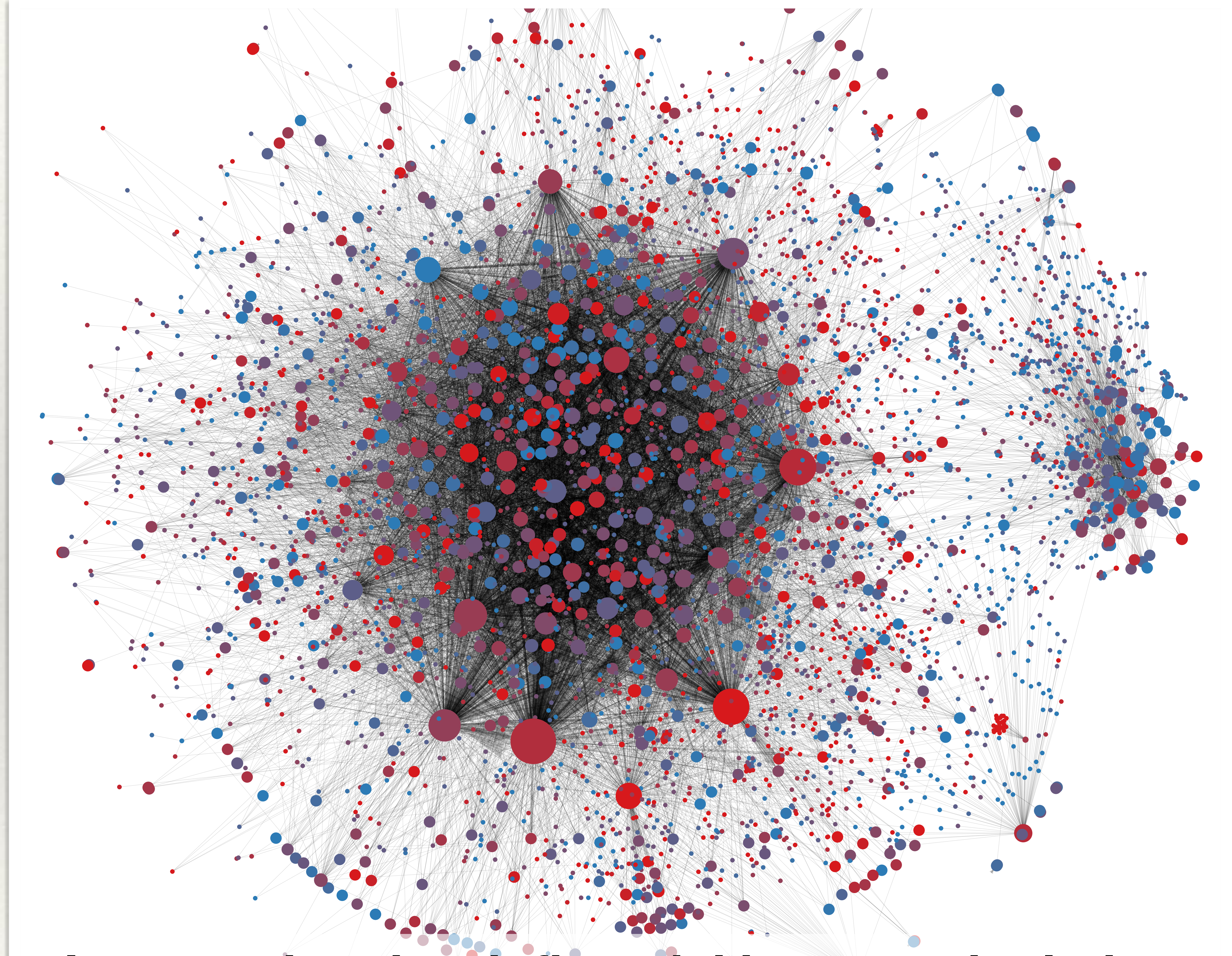
- Multiple ways to measure the influence of an account:
  - Number of followers (in-degree in follower network)
  - Number of users exposed (out-degree in retweet network)
  - Number of retweets (out-strength in retweet network)
  - Fraction of retweets to followers
- Social bots can target influential accounts hoping for retweet



Blue links: retweets and quotes. Red links: mentions and replies. Node size: number of followers.

Social bots

- Accounts controlled by an entity via software
- Malicious social bots can impersonate humans, deceive, and manipulate diffusion networks:
  - Fake followers
  - **Amplification**: fake retweets
  - **Astroturf**: appearance of organic virality
- All social media platforms and users are vulnerable



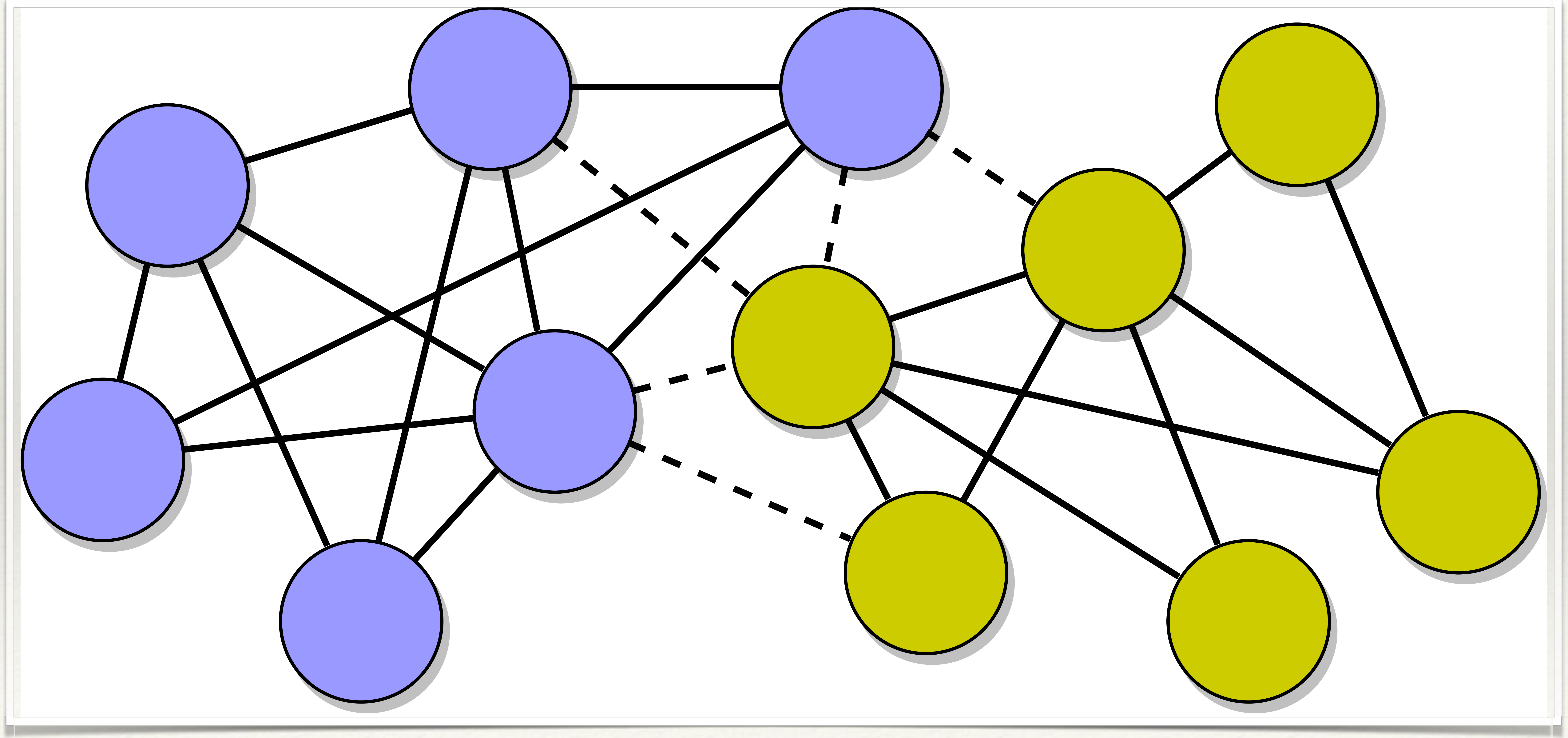
Large red nodes: influential bots manipulating online debate about vaccination policy

Homophily and Segregation

Homophily

- The principle that we tend to be similar to our friends
- This makes your friends not statistically significant as a random sample of the population
- Similarities:
  - immutable characteristics
  - mutable characteristics

"Birds of a feather flock together"



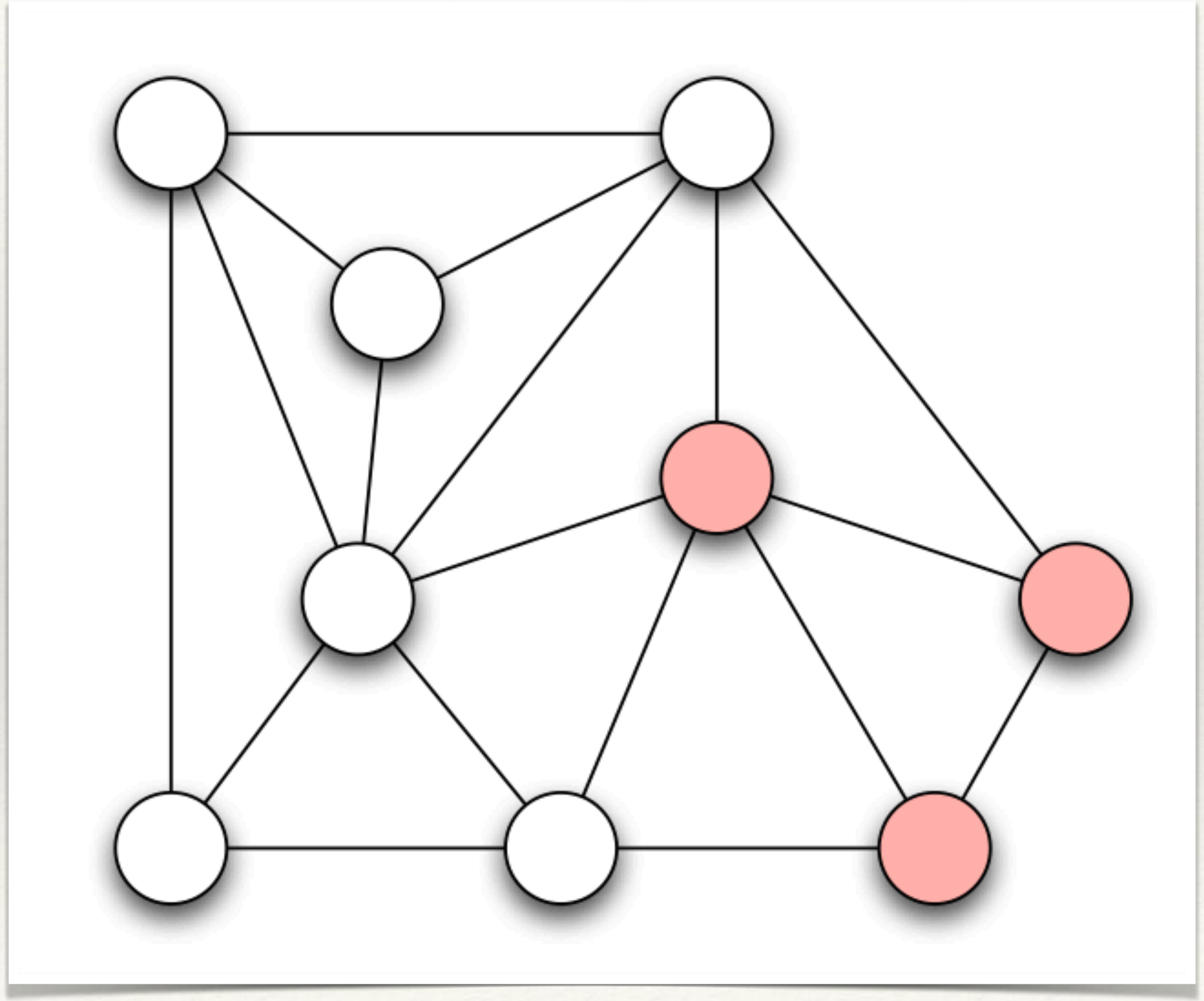
Picture from: F. Menczer, S. Fortunato, C. A. Davis, A First Course in Network Science, Cambridge University Press, 2020

Measuring homophily

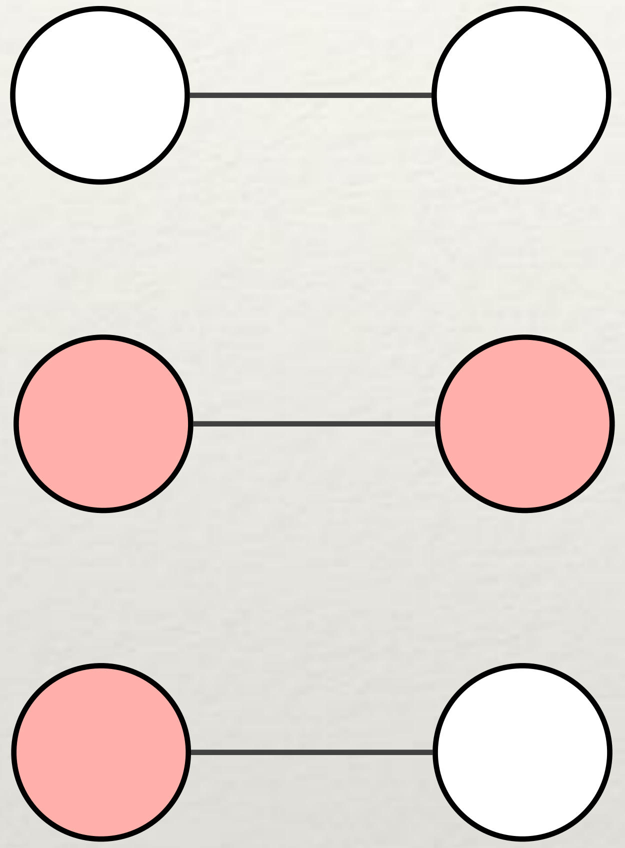
- Simple test:
  1. let's assign randomly a color to each node
  2. count number of cross-colors edges
  3. compare numbers with actual network

$p = \frac{6}{9} = \frac{2}{3}$

$q = \frac{3}{9} = \frac{1}{3}$



fraction of white nodes: $p = \frac{2}{3}$ fraction of pink nodes: $q = \frac{1}{3}$

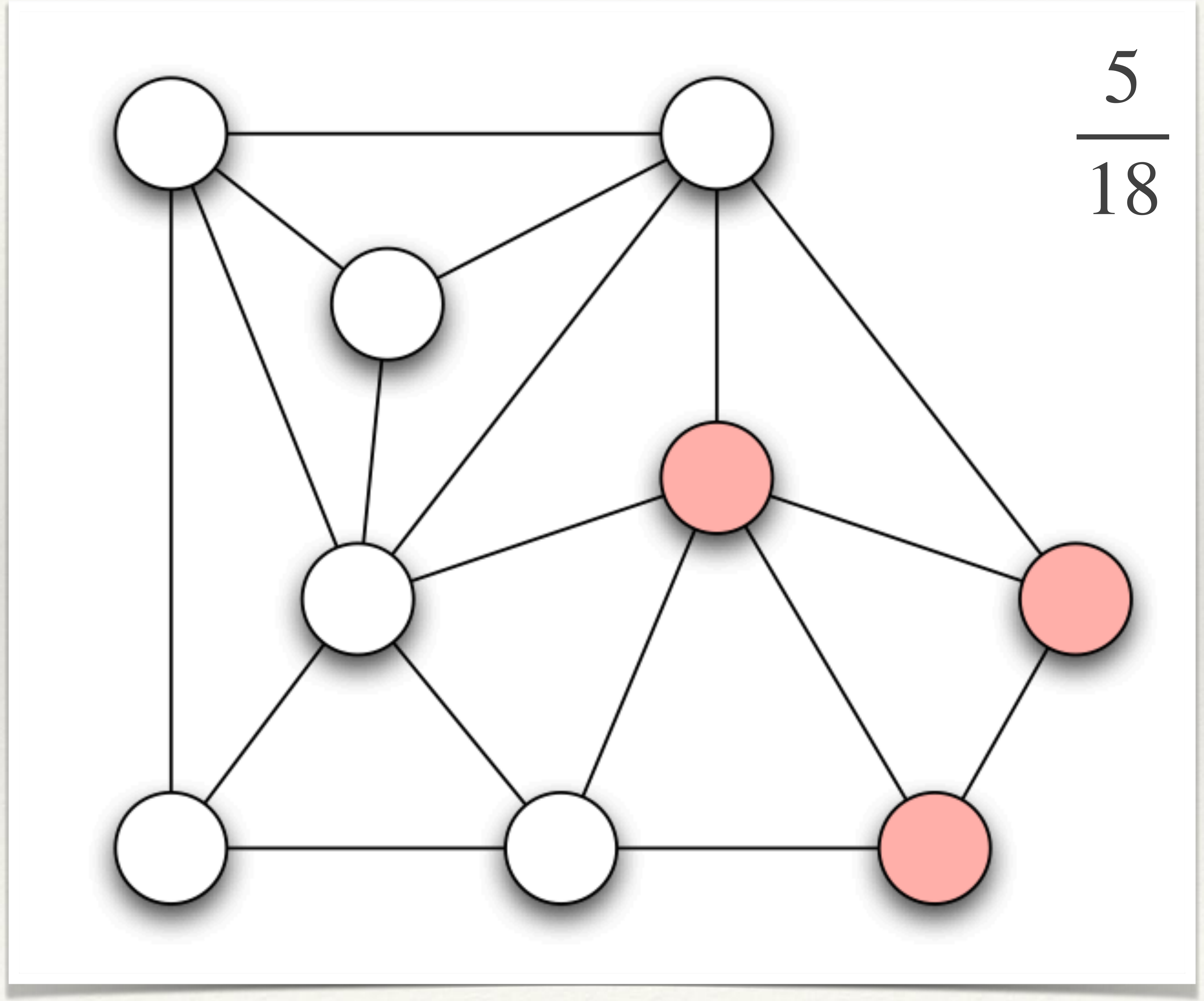


$$p \cdot p = p^2 = \frac{4}{9}$$

$$q \cdot q = q^2 = \frac{1}{9}$$

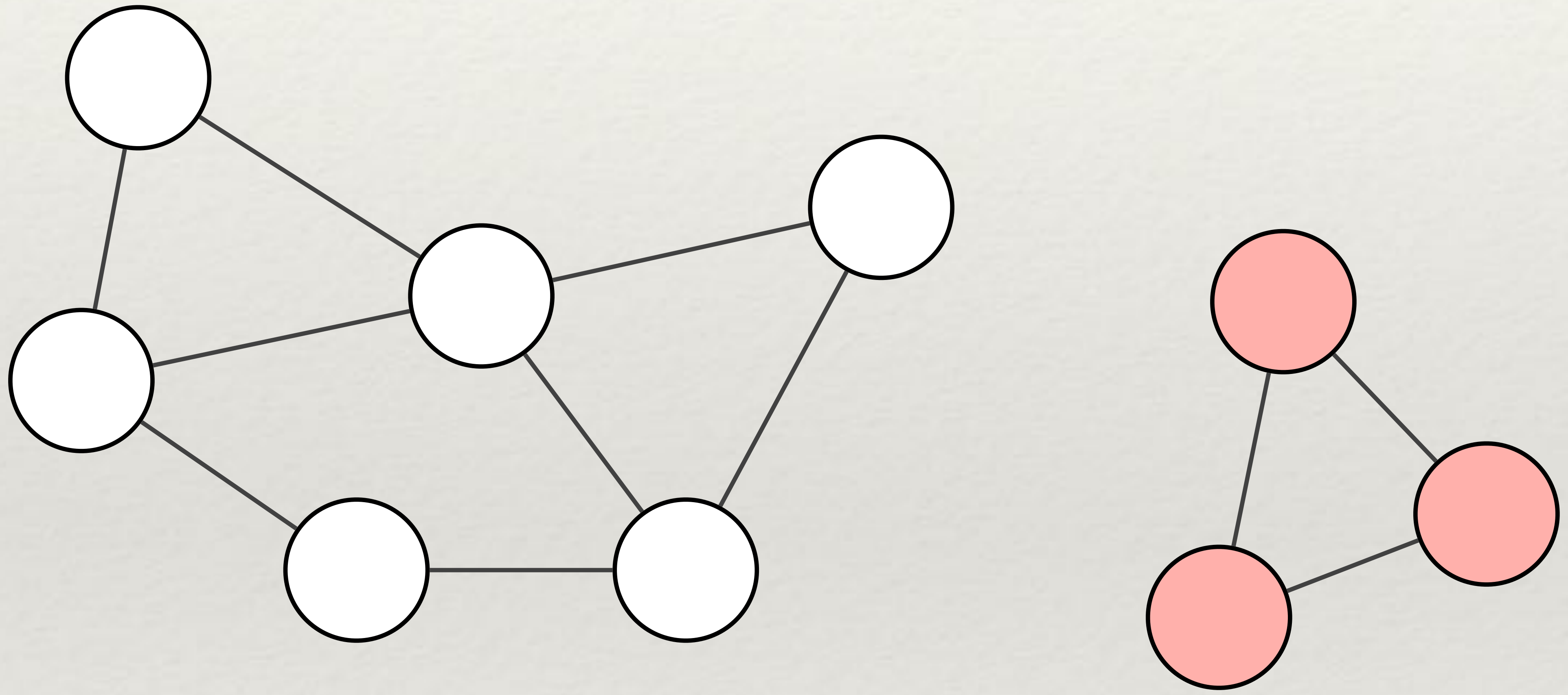
$$2 \cdot p \cdot q = 2\frac{2}{3}\frac{1}{3} = \frac{4}{9}$$

$$\frac{5}{18} < \frac{4}{9}$$



homophily test: check if #*actual cross groups edges* < *2pq*

$$\frac{5}{18} < \frac{4}{9} \quad \Rightarrow \quad \textbf{homophily!}$$



perfect **homophily:** $0 < \frac{4}{9}$

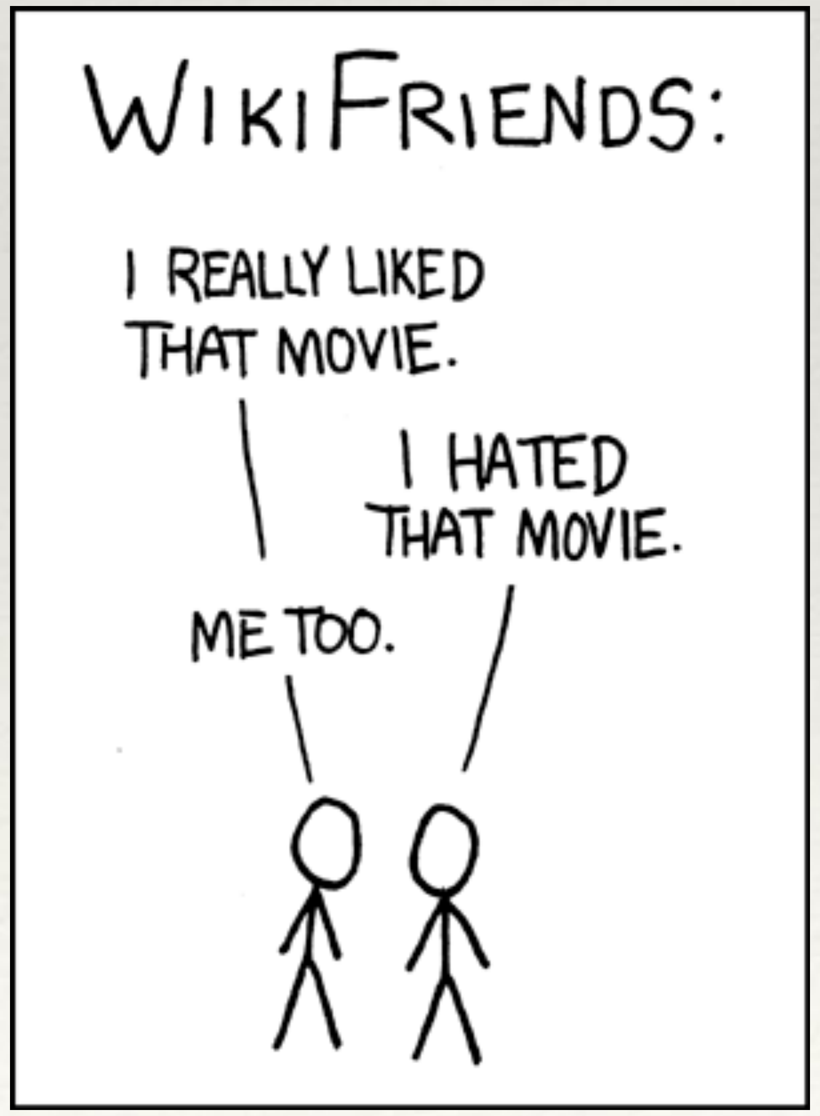
More precisely

homophily test: if the fraction of cross-types edges is *significantly less* then *2pq*, then there is a *signal* of homophily

Underlying mechanisms of homophily

- Two possible mechanisms by which homophily (also: assortativity) emerges naturally:

1. **Selection**: similar nodes become connected
2. (Social) **influence**: connected nodes become more



The interplay of selection and social influence

- longitudinal methodology:
  - observe a network for a long period of time
  - observe both factors in action
  - how do we quantify the impact?
- example: obesity as a social contagion phenomenon

Nicholas A. Christakis, and James H. Fowler, The Spread of Obesity in a Large Social Network over 32 Years, July 26, 2007, N Engl J Med 2007; 357:370-379, DOI: 10.1056/NEJMsa066082

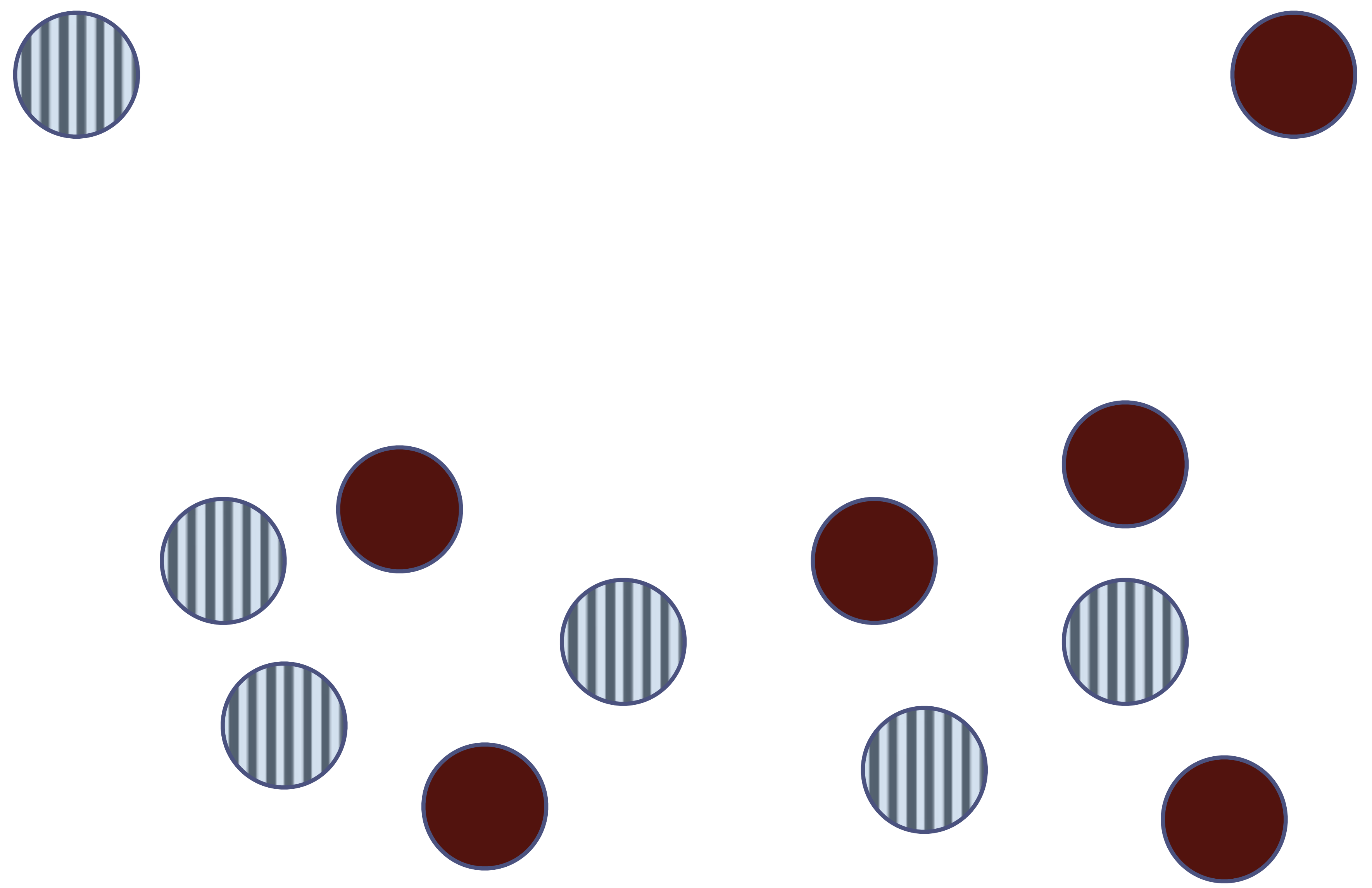
obesity "contagion"

- dataset: 12,000 people
- obesity status
- social network structure
- obese vs non obese: there is a tendency toward clustering
- homophily test: passed
- why?
  - selection?
  - homophily that correlates with something else?
  - social influence? —> contagion!

https://www.ted.com/talks/nicholas_christakis_the_hidden_influence_of_social_networks

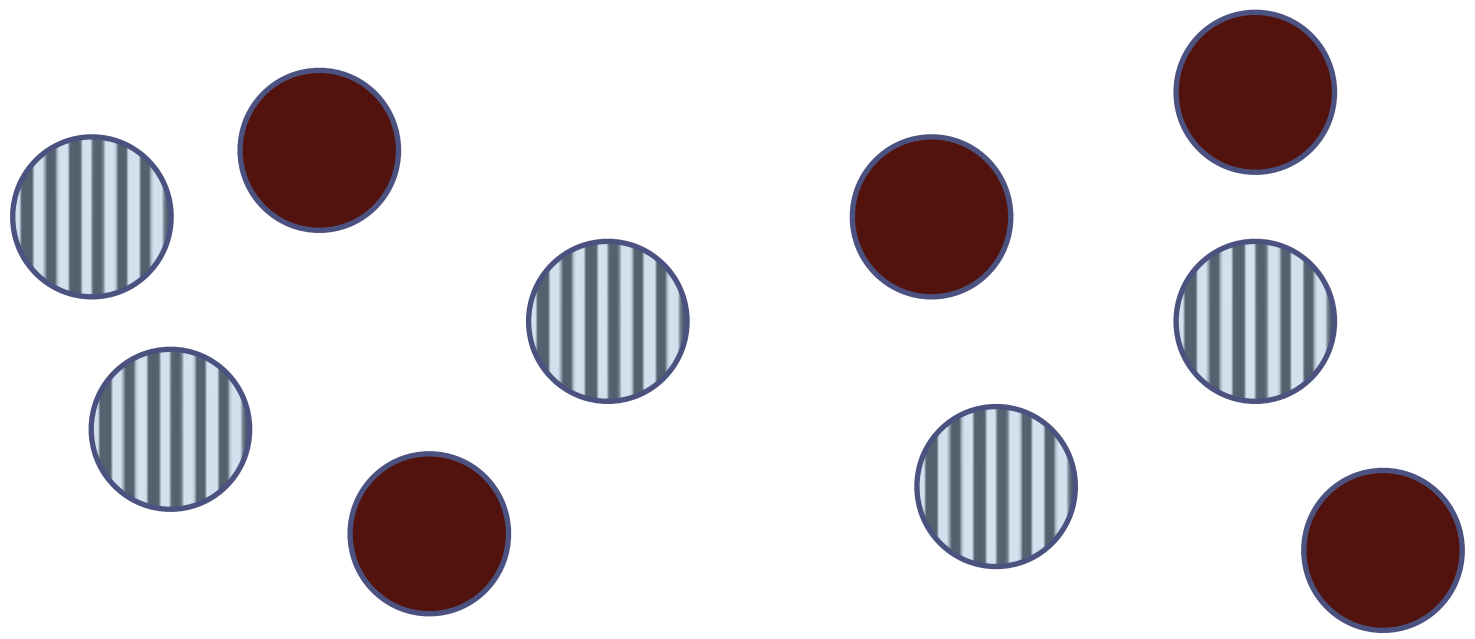
The emergence of segregation

- Society's structure is shaped in function of **immutable characteristics** of individuals
  - ethnic group
  - age
  - religious belief
  - …



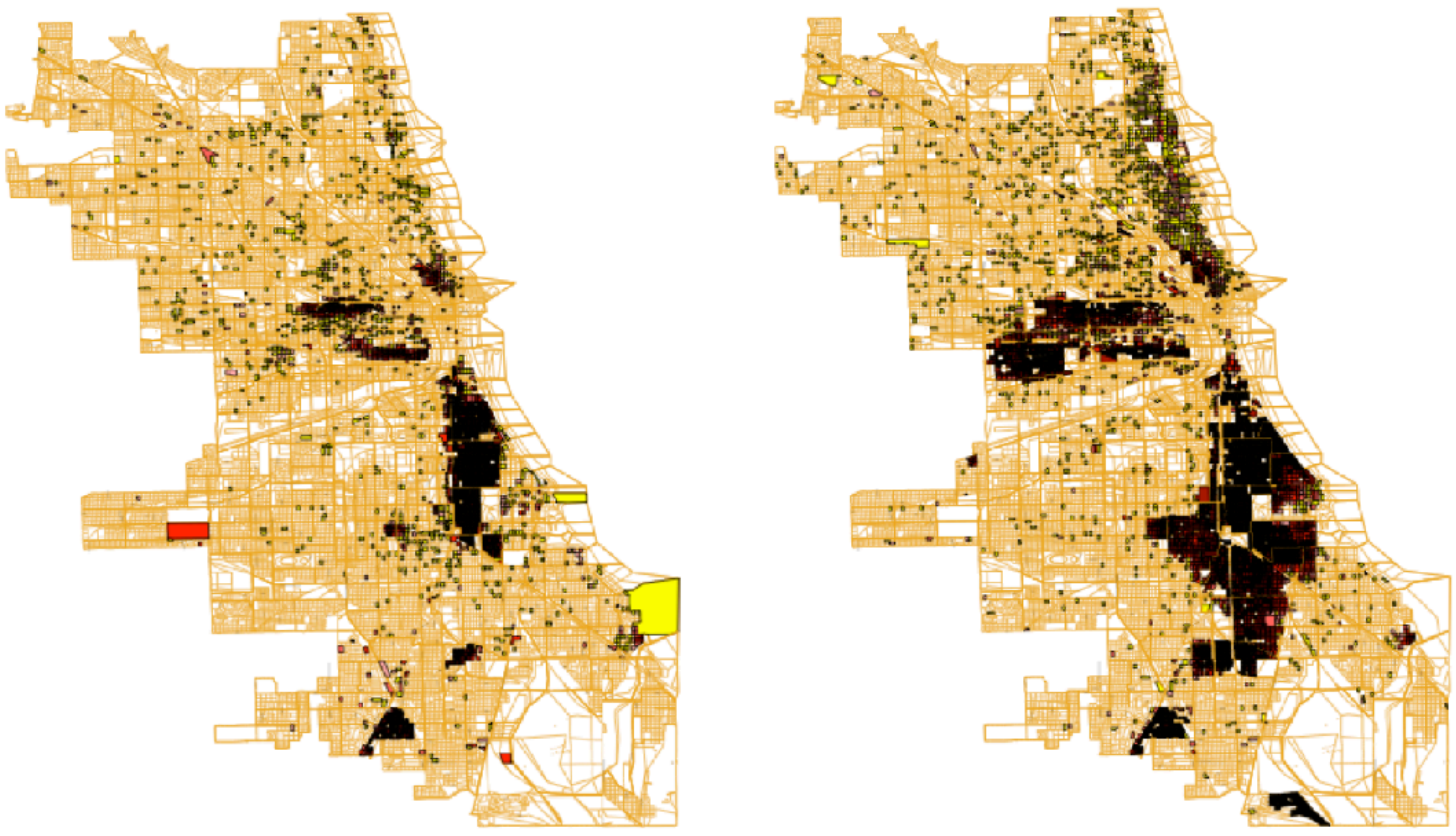
Segregation

- Society's structure is shaped in function of **immutable characteristics** of individuals
  - ethnic group
  - age
  - religious belief
  - …



Natural spatial "signature" in cities

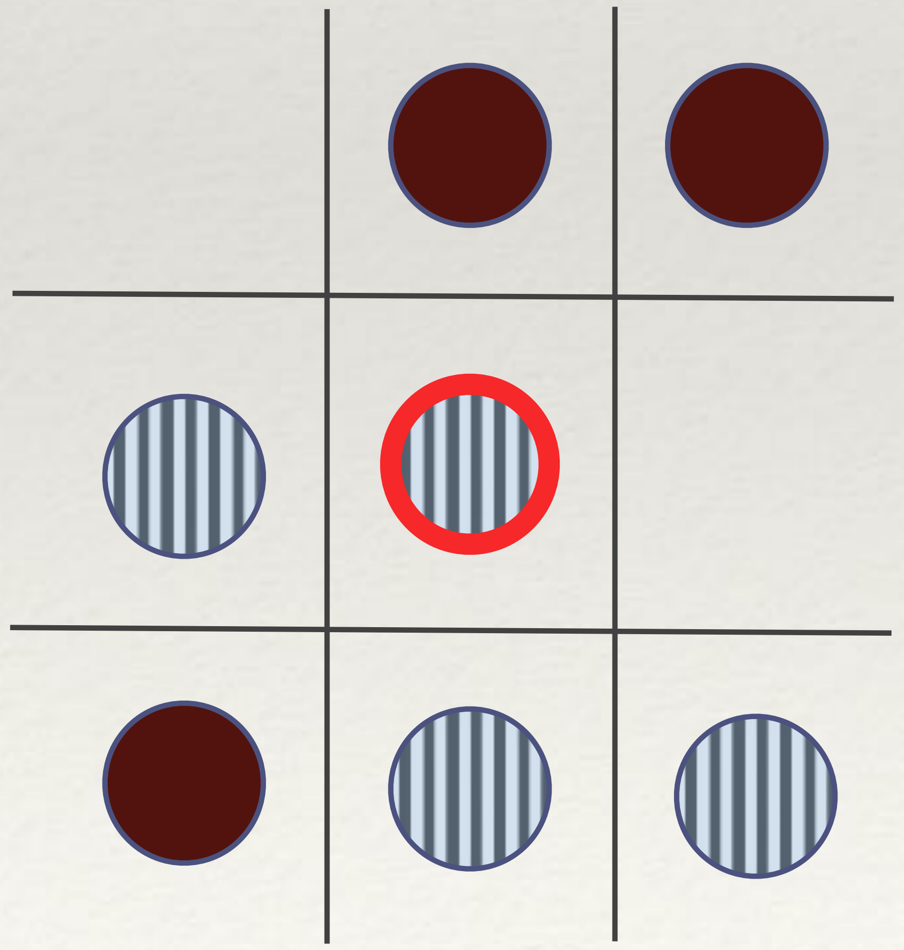
- Formation of homogeneous (according to some "type" or "class") neighbors in cities
- Which are the causes of "ghettization"?

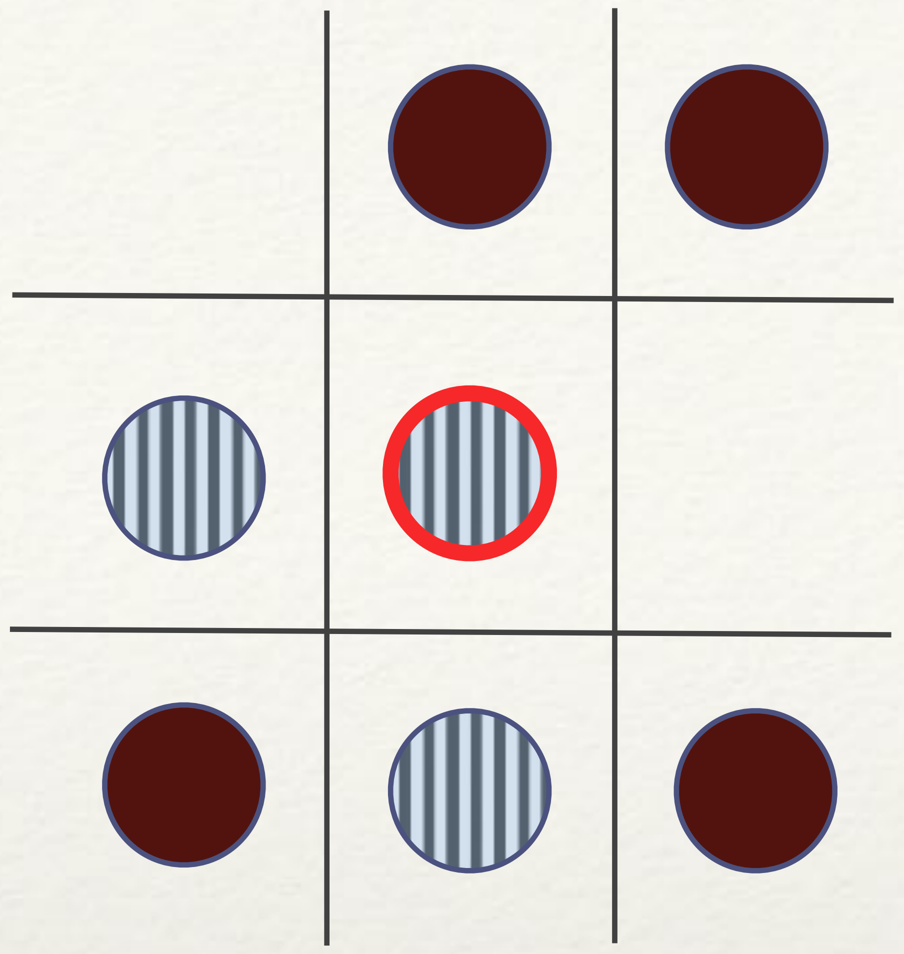


(a) *Chicago, 1940* (b) *Chicago, 1960*

The Schelling model

- Can spatial segregation arise from the effect of homophily operating at a local level?
- Assumption: no individual want segregation explicitly
- Agents:
  - two types: 
  - immutable characteristics
- Agents reside in a cell of a grid
  - some cells contain agents
  - some other cells are unpopulated
- Neighbors: 8 other cells "touching" an agent



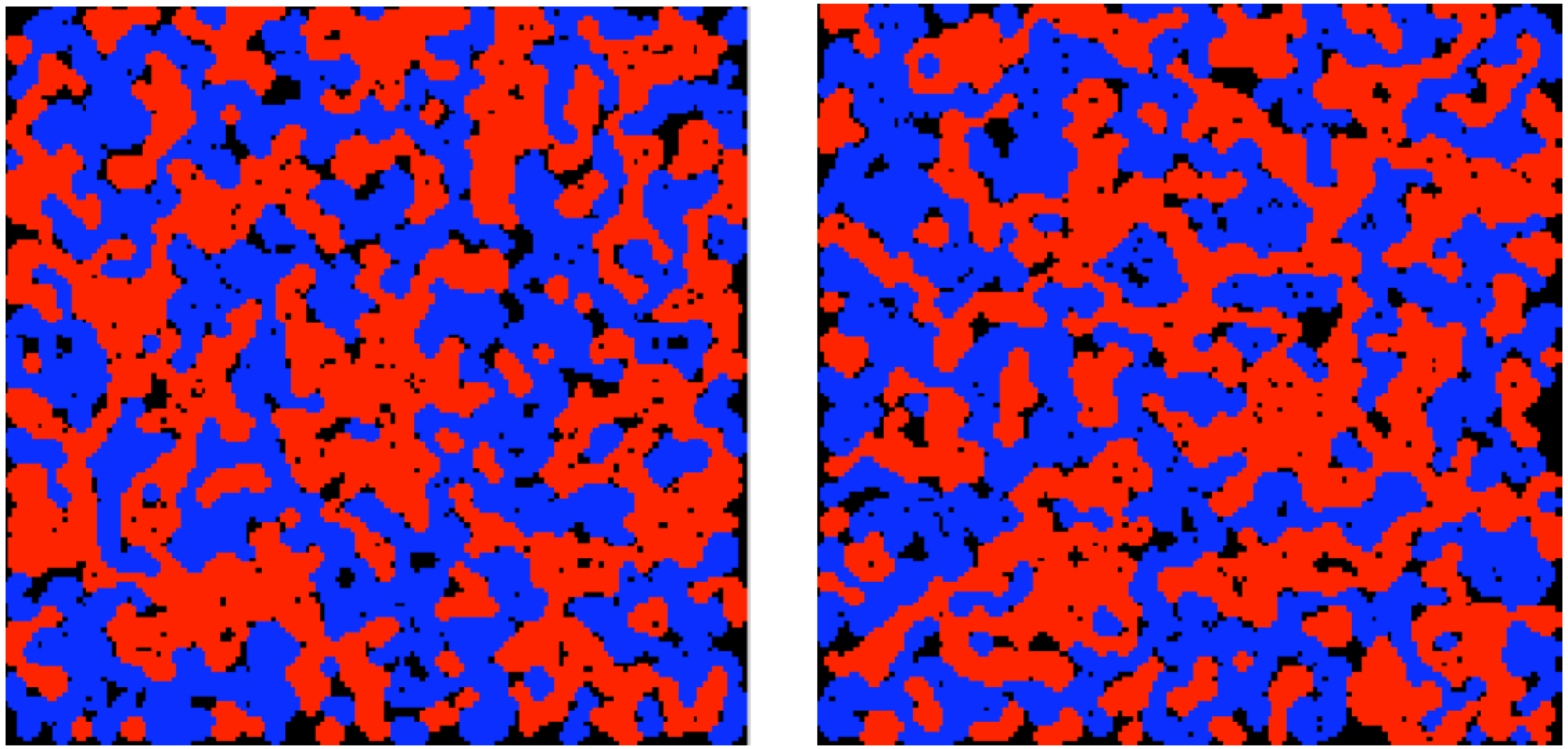


$t = 3 \Rightarrow$:-(

- Each agent wants to have at least ***t*** neighbors of their own type
- If an agent find < ***t*** neighbors of the same type, then they are **unsatisfied**
- If unsatisfied, they want to **move**

Larger examples

- Computer simulations to look for patterns at larger scale
- We want to run different simulations and make some comparisons
  => integrated pattern?
- on the right: two runs of a simulations of the Schelling model with a threshold *t* of *3*
  - 150x150 grid
  - 10,000 agents

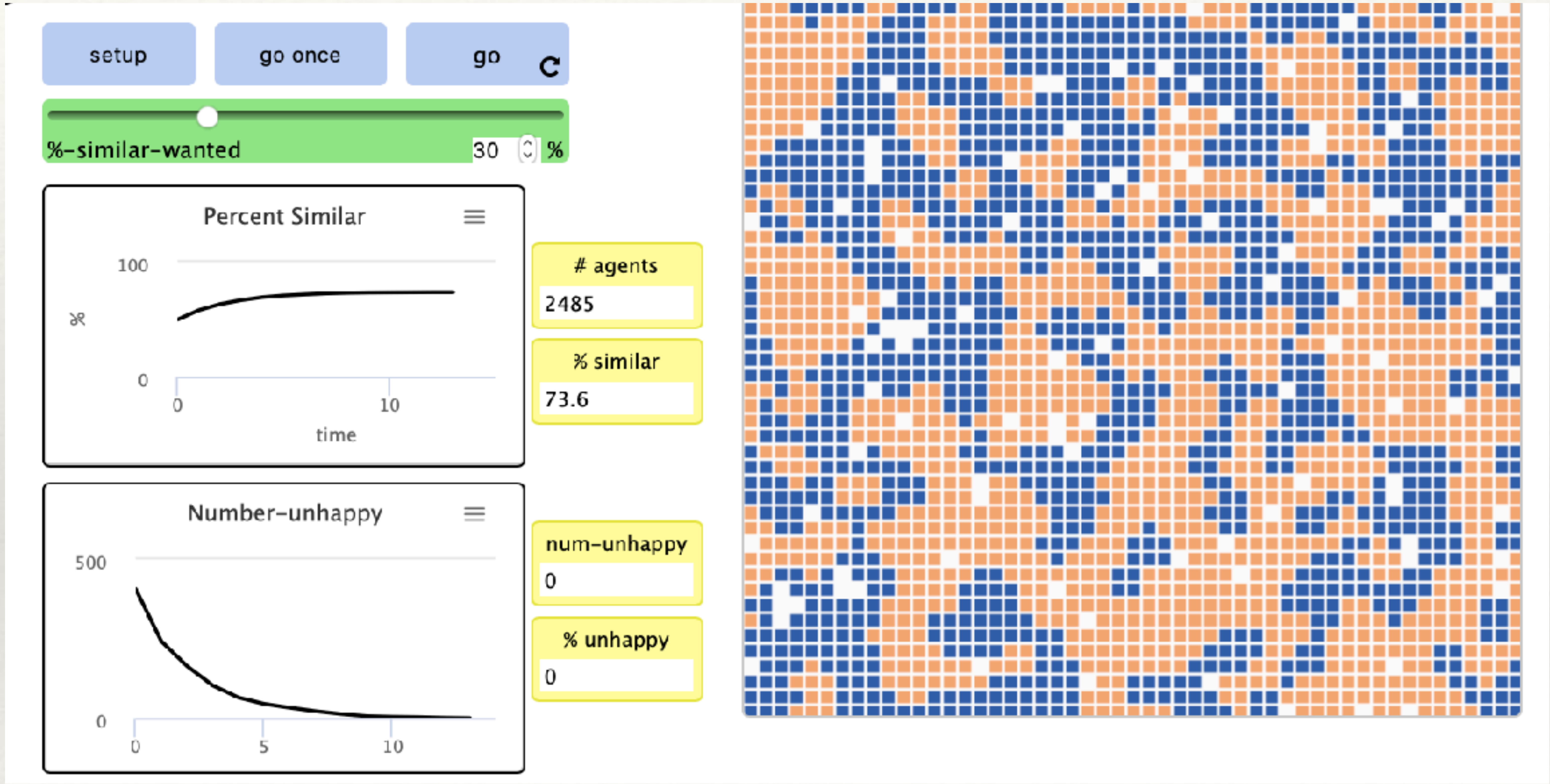


(a) *A simulation with threshold 3.* (b) *Another simulation with threshold 3.*

Segregation emerges even when agents accept to be a minority!

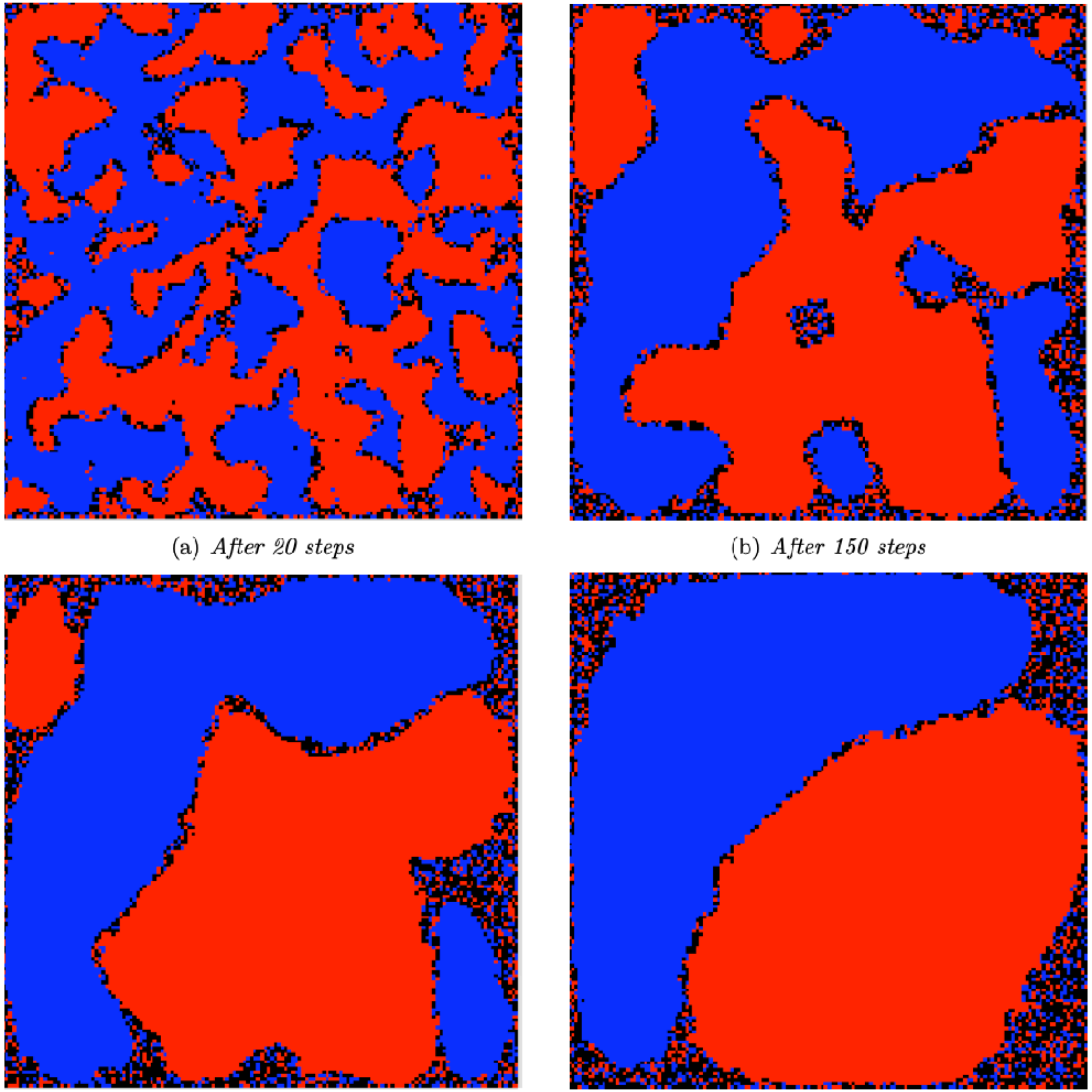
Example with NetLogo

Agent based simulations



http://www.netlogoweb.org/launch#http://www.netlogoweb.org/assets/modelslib/Sample%20Models/Social%20Science/Segregation.nlogo

$t > 3$ =>



(a) *After 20 steps* (b) *After 150 steps*

(c) *After 350 steps* (d) *After 800 steps*

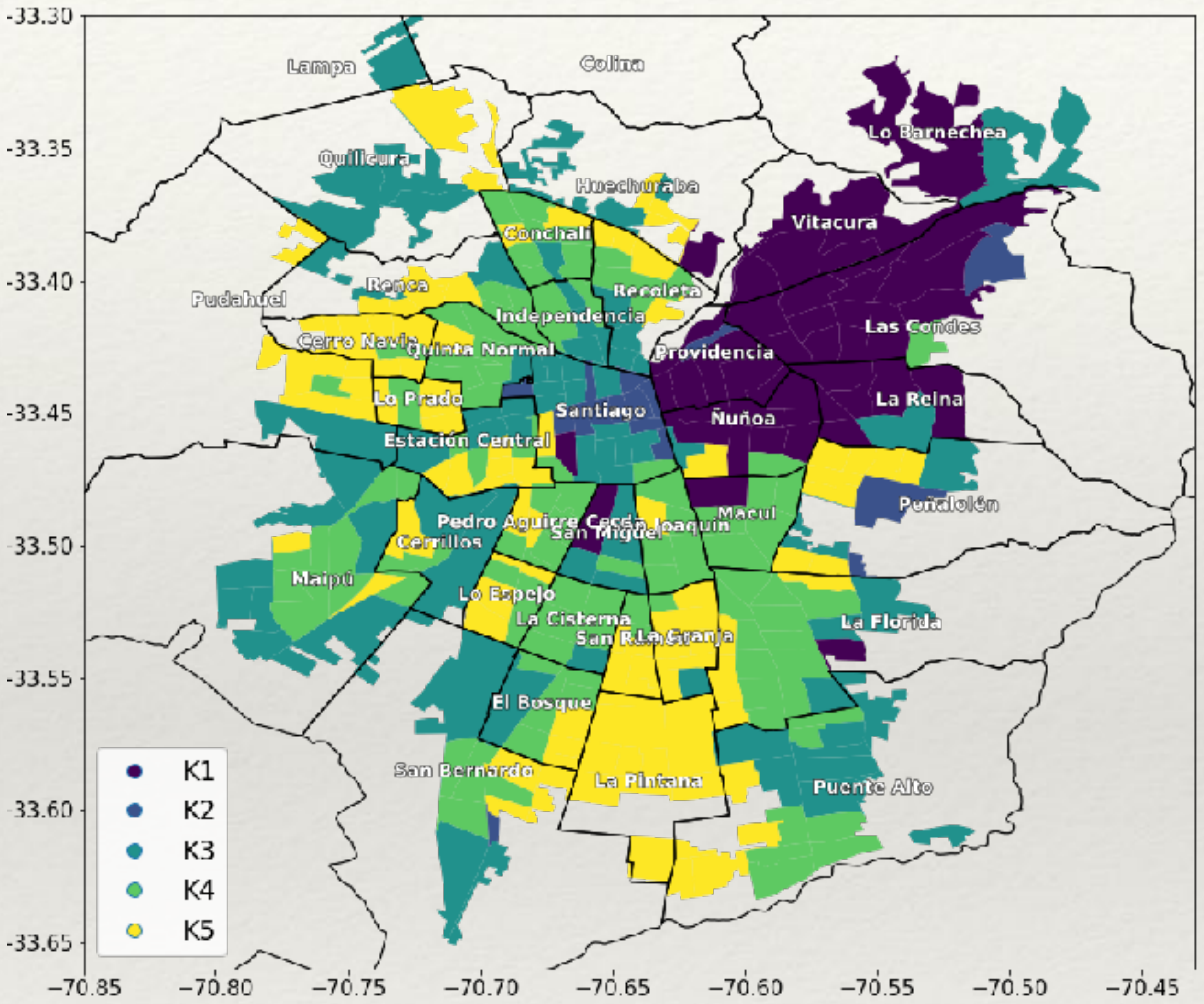
Segregation is (trivially) amplified in an intolerant society

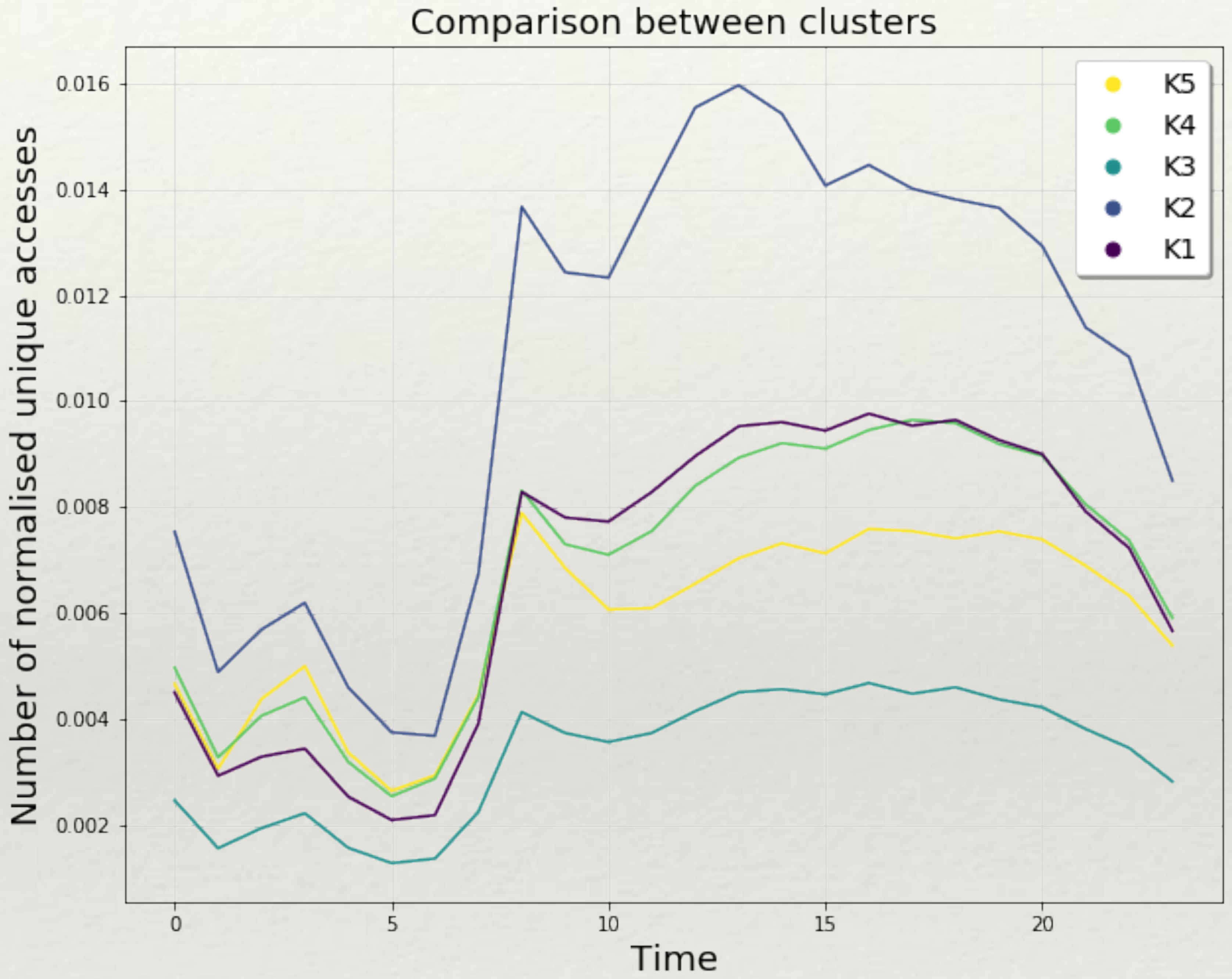
Impacts of segregation

- Let's accept that segregation emerges naturally even in the most tolerant society (unless we do not design our 'societies' properly)
- Segregation has consequences (not necessarily bad…)
- Examples:
  - on news consumption
  - on outbreaks diffusion

Segregation vs information consumption

Study of geo-located accesses to websites of **news media** revealed strong differences between different “classes” of the population of SCL.






Salvatore Vilella


Leo Ferres


Daniela Paolotti

Vilella, S., Paolotti, D., Ruffo, G. and Ferres, L.. News and the city: understanding online press consumption patterns through mobile data. EPJ Data Sci. 9, 10 (2020)

Segregation by age and virus transmission

Crowds take in the the cherry blossoms a
visitors from holding *sakura*-viewing par

COMMENTARY / JAPAN

Why is Japan still a cor

BY OSCAR BOYD
STAFF WRITER

At the time of writing, Japan has just
coronavirus. That's 900 cases record
first person — a man who had travel
have the disease while in a Japanese

In Italy, the first case was recorded t
23. Shortly after, 50,000 people were quarantined in a handful of towns in

CLICK TO ENLARGE

thought: that Japan is
pread in the way it has
ns: relatively less social
to wear masks when
us, already high
e voluntary self-
at Japan is flattening

Dynamical Processes in Information Networks

Overview of network dynamics

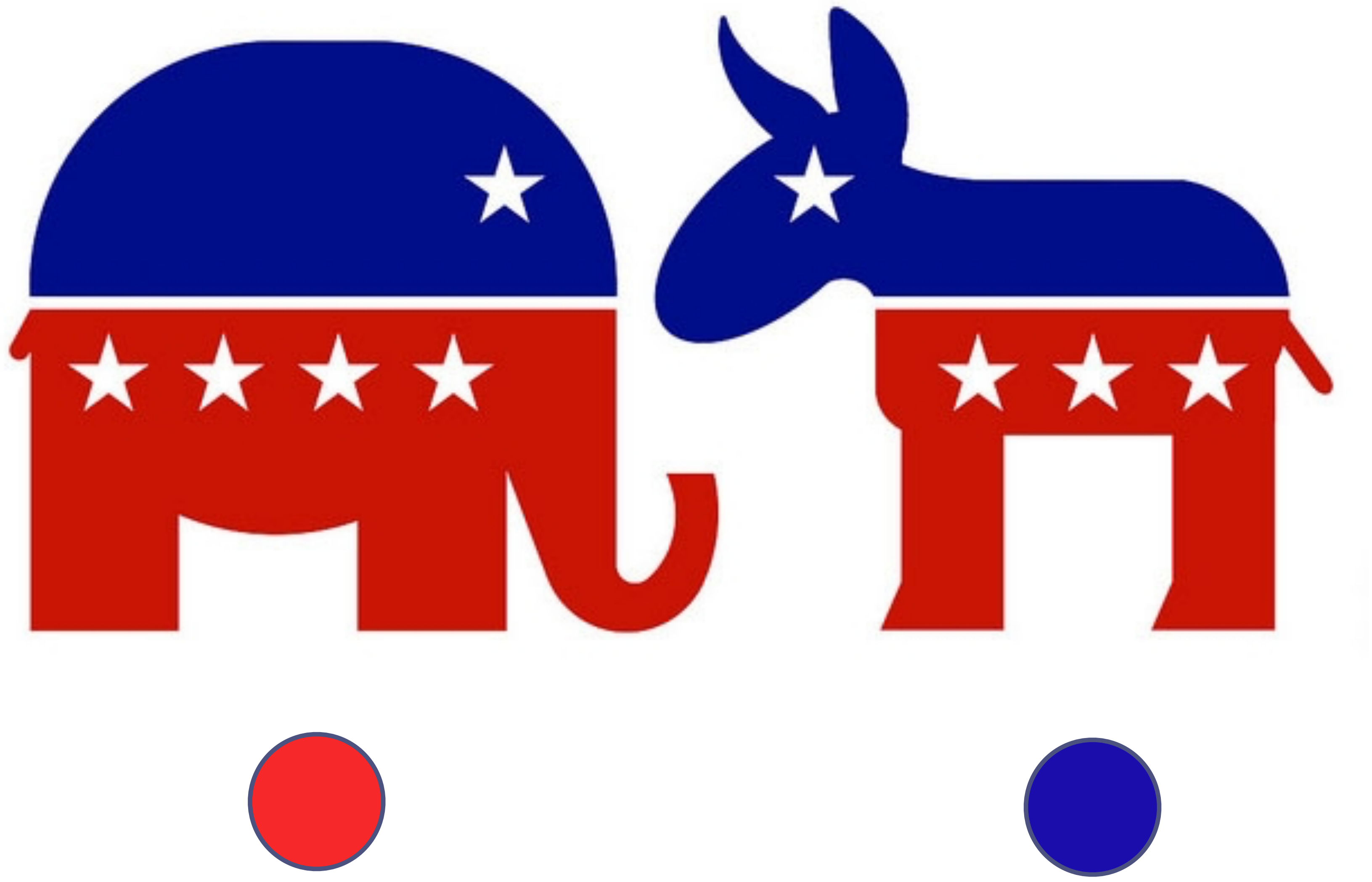
- Social contagion
- Emergence of polarization
- Consequences: confirmation biases, echo chambers
- Intro to epidemic spreading
- Impact of verification and fact checking: SBFC model and what-if analysis

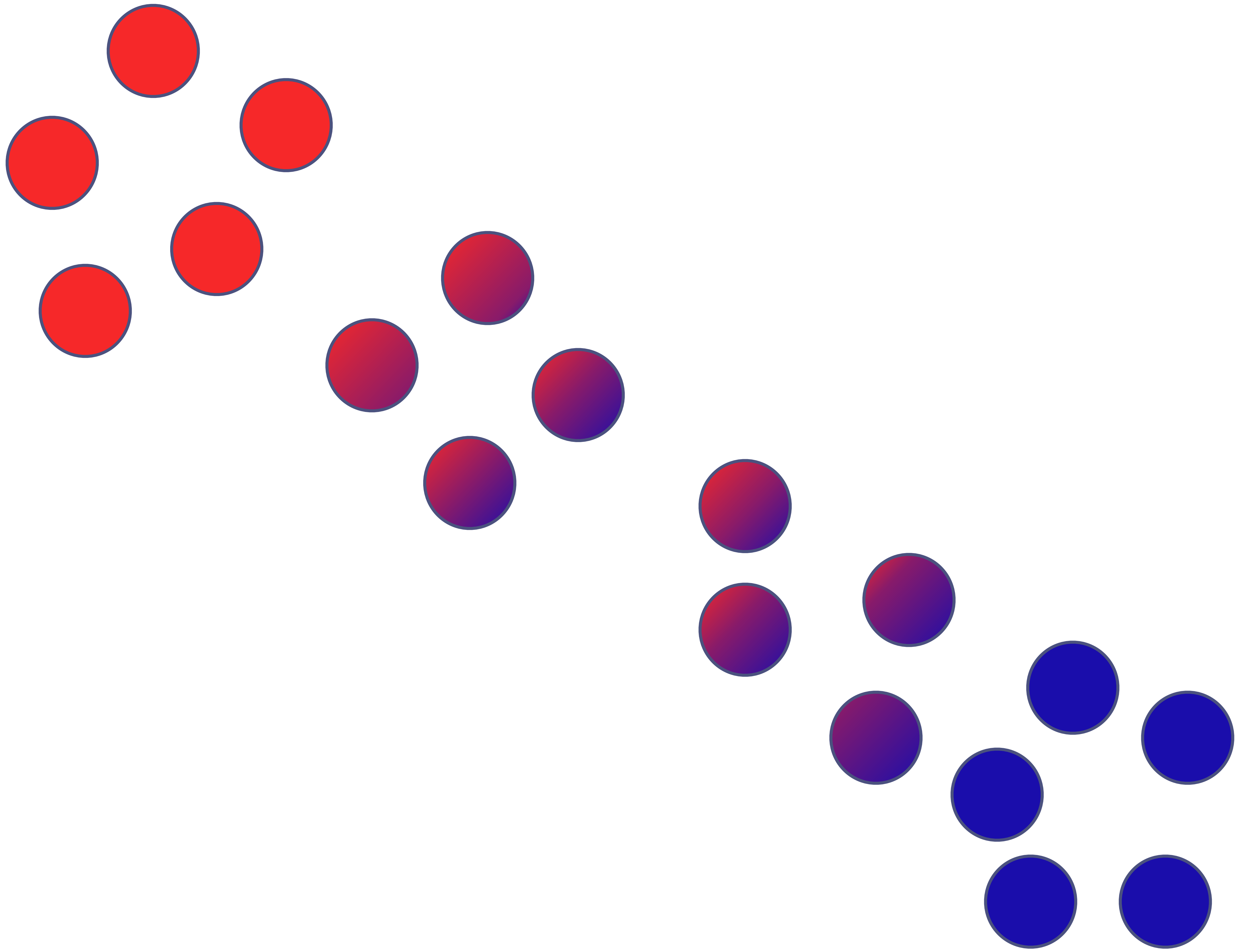
Emergence of polarization

*“**Polarization** is both a state and a process. Polarization as a **state** refers to the extent to which opinions on an issue are opposed in relation to some theoretical maximum. Polarization as a **process** refers to the increase in such opposition over time.”*

– DiMaggio et. al, American Journal of Sociology, 1996

Polarization





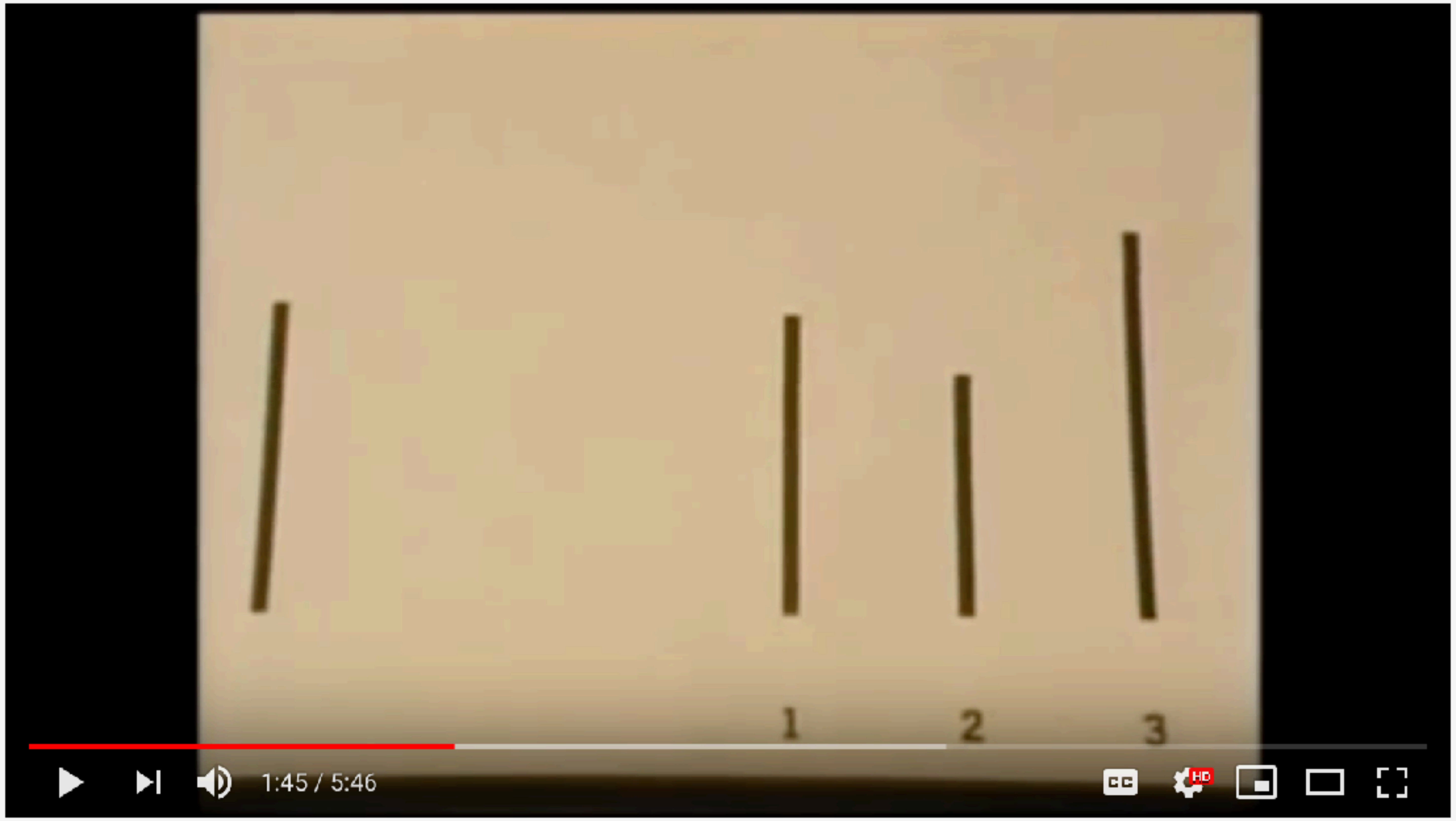
Polarization

Issues with studying polarization

- **State**: difficult to detect
  - e.g., NLP based techniques as "*stance detection*" are great, but errors prone
- **Process**: difficult to observe
  - e.g., opinions can mitigate or polarize over time, but people do not necessarily express them
- Polarization by **selection** and by **influence**
  - do I get along with people that share my opinion, or I am influenced by people with whom I get along? or both processes are at interplay?
- "**Social contagion**" is more rational than we may think…

Social contagion

Conformity experiment and group influence



Asch Conformity Experiment

https://www.youtube.com/watch?v=NyDDyT1lDhA

Different kinds of contagion

- **Epidemics**: a pathogen is transmitted by infected individuals
- **Social Contagion**: diffusion and adoption of ideas, opinions, innovations, behaviors, …

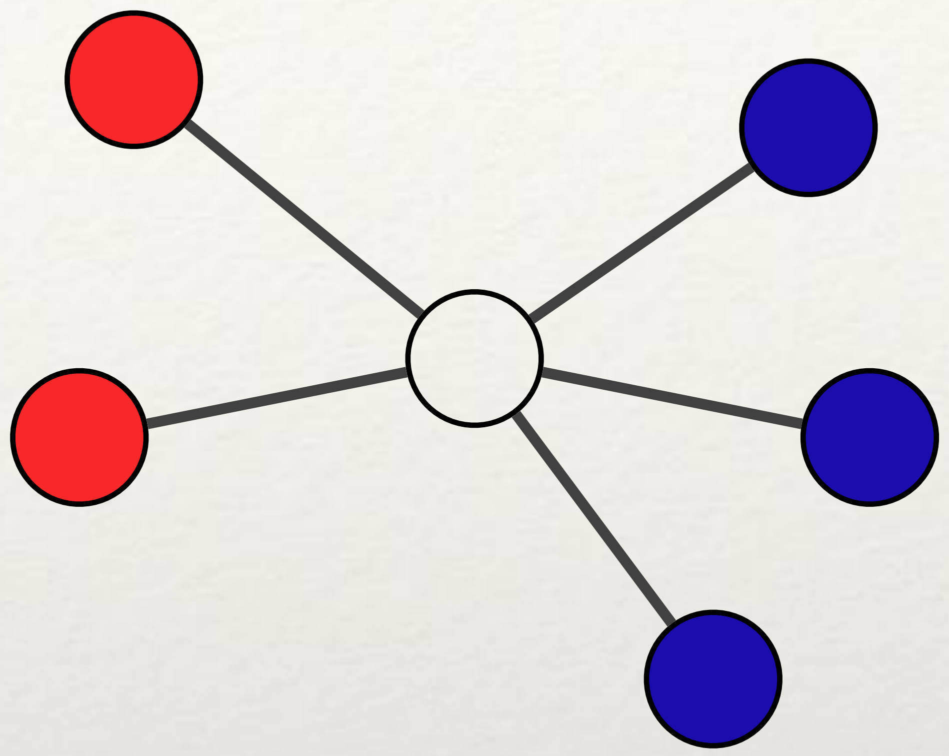
A diffusion of a new behavior

- Assumption: individuals make *decisions based* on the choices of *their neighbors*
  - focus on links
- Natural model introduced by Stephen Morris in 2000

Stephen Morris. Contagion. Review of Economic Studies, 67:57–78, 2000.

A simple (linear) threshold model

- It is natural to use a **coordination game**
  - each node has a choice between two possible behaviors, A and B
  - players have an incentive to adopt the same behavior



		w	
		A	B
v	A	a, a	0, 0
	B	0, 0	b, b

p fraction of neighbors adopting A

$1\text{-}p$ fraction of neighbors adopting B

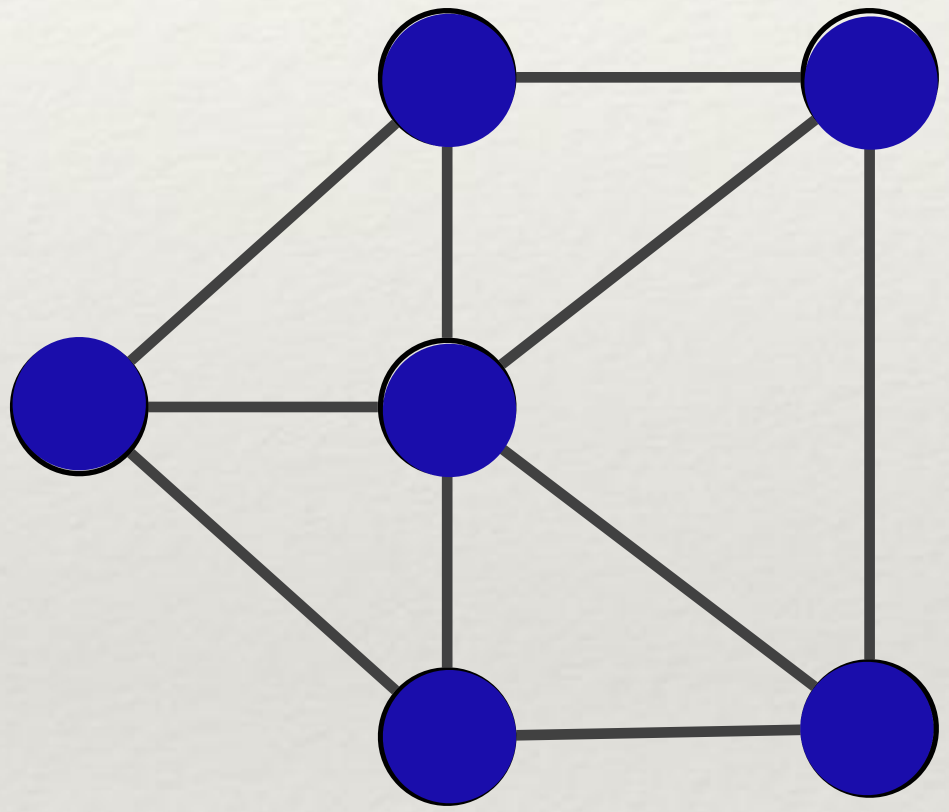
d is the number of neighbors

the node chooses A if $pda \geq (1-p)db$

$$\Rightarrow p \geq \frac{b}{a+b} = q$$

Example

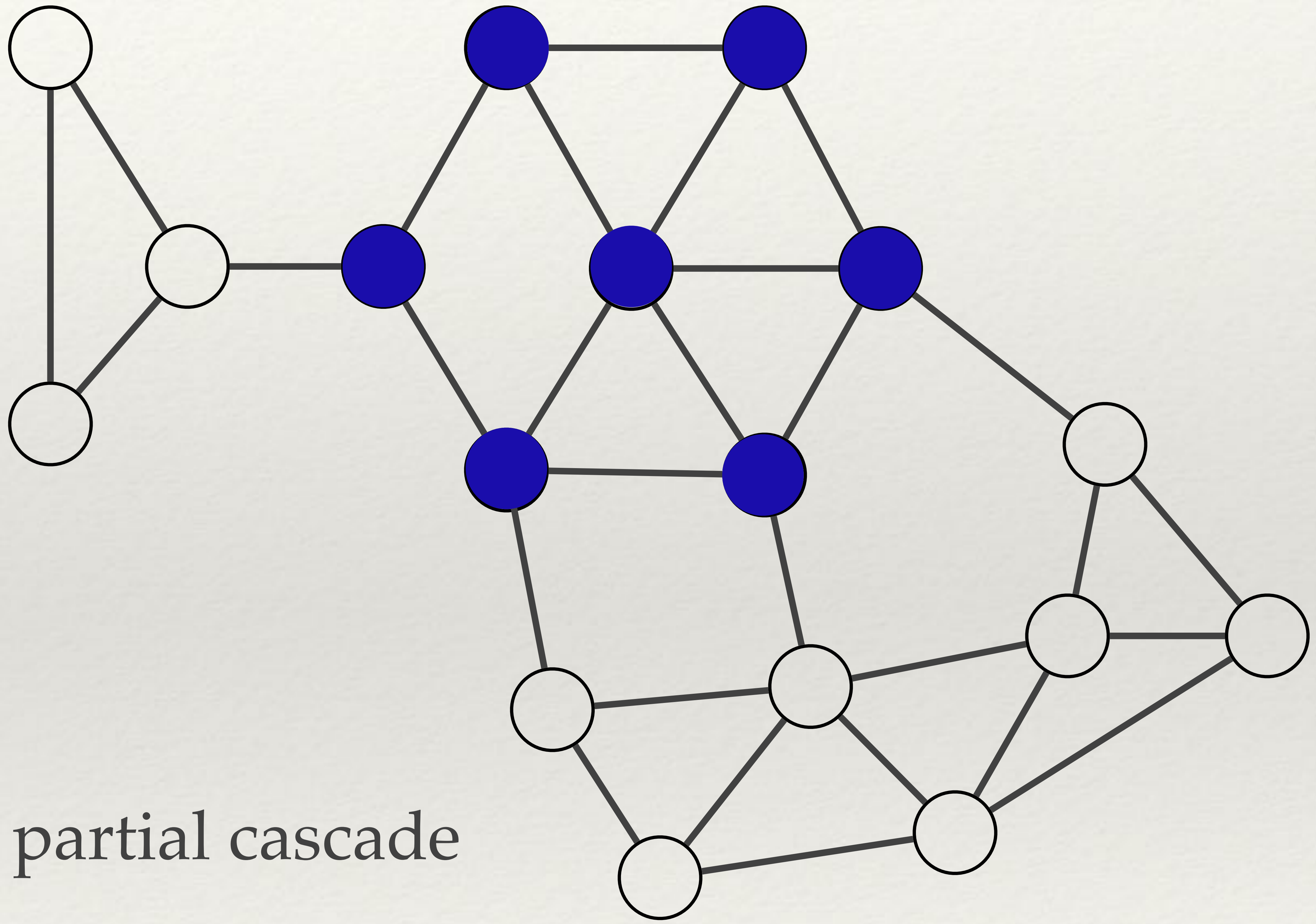
- $q = \dfrac{2}{5}$
- $S = \{u, v\}$



Chain reaction: complete cascade

Another example

- $q = \dfrac{2}{5}$
- $S = \{u, v\}$

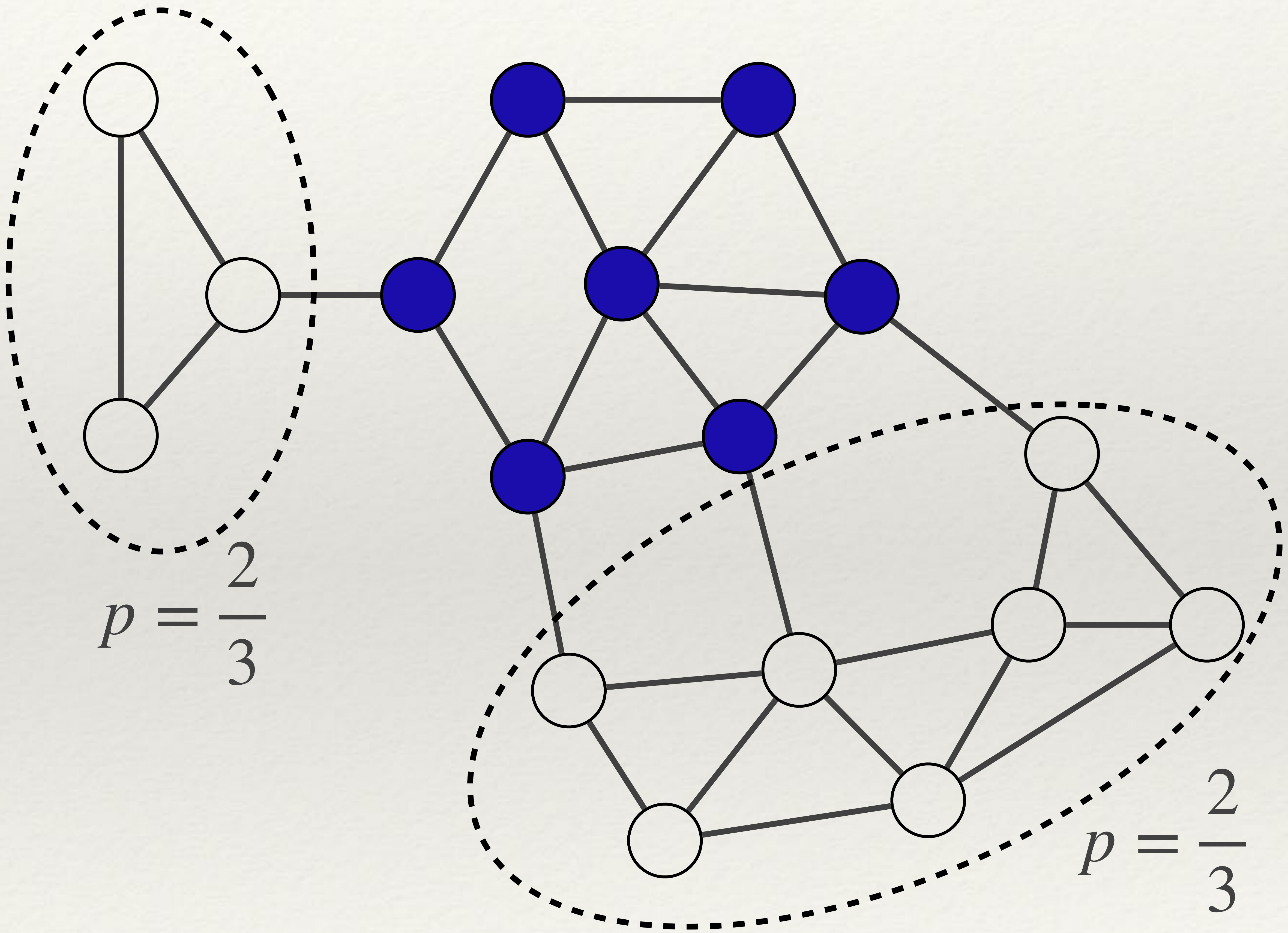


The diffusion of A stops here: partial cascade

Clusters are **barriers** to diffusion!

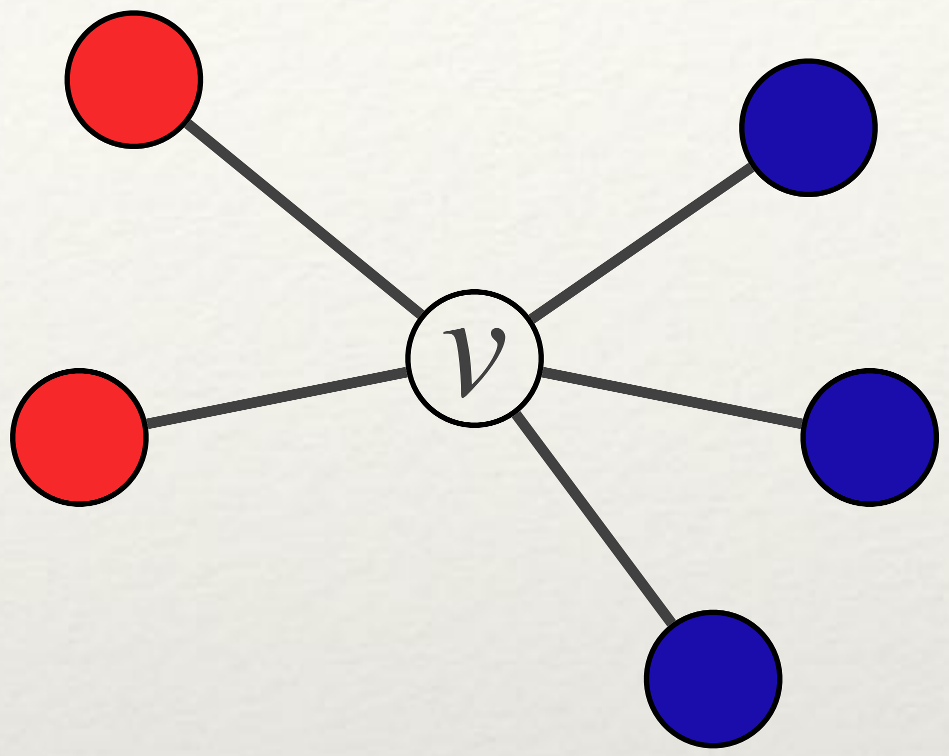
Stopping cascades

- What prevents cascades from spreading?
  - *Homophily* can serve as a barrier to diffusion: it is hard for innovation to arrive from outside densely connected communities
- Let's try to quantify this intuition:
  - <u>def</u>. *cluster of density p is a set of nodes C where **each node** in the set has at least p fraction of edges in C*



Heterogeneous thresholds

- Let's suppose each person gives values to *A* and *B* subjectively



w

v		**A**	**B**
	A	a_v, a_w	0, 0
	B	0, 0	b_v, b_w

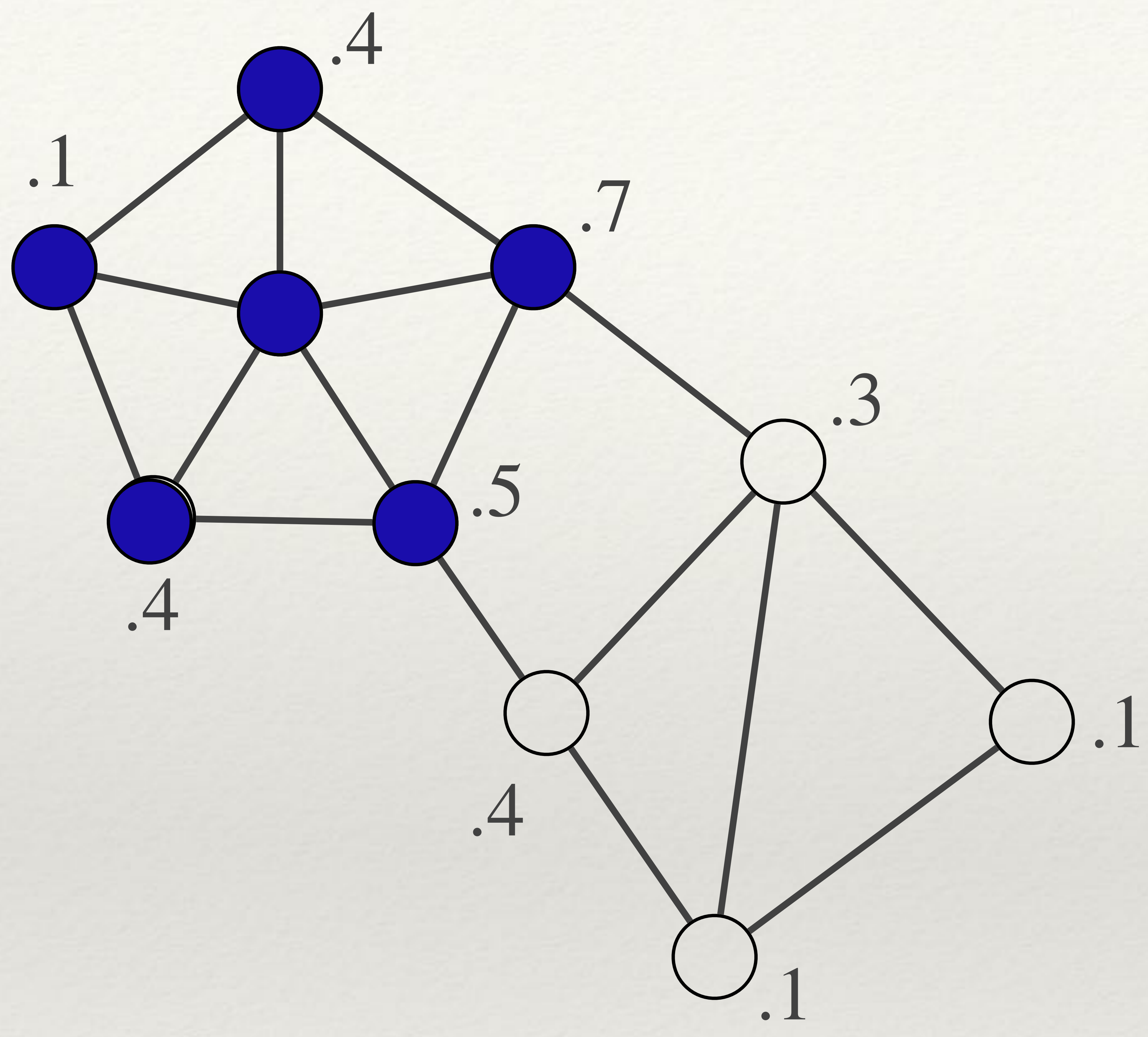
p fraction of neighbors adopting A

$1-p$ fraction of neighbors adopting B

d is the number of neighbors

the node chooses A if $pda_v \geq (1-p)db_v$

$$\Rightarrow p \geq \frac{b_v}{a_v + b_v} = q_v$$



Watts and Dodds: we need to take into account not just the power of influential nodes, but also the extent to which these influential nodes have access to easily **influenceable** people.

Reformulating the notion of **blocking clusters**: set of nodes for which each node v has a fraction $> (1 - q_v)$ of its friends inside the set.

The notion of density becomes **heterogeneous** as well: each node has a different requirement for the fraction of friends it needs to have in the cluster.

Duncan J. Watts and Peter S. Dodds. Networks, influence, and public opinion formation. Journal of Consumer Research, 34(4):441–458, 2007.

Independent cascade models

- Principle of threshold models: **peer pressure**, the more people try to persuade you, the more likely they will succeed

- **Remark:** social influence often works **one-to-one**, we may be persuaded by a single passionate individual

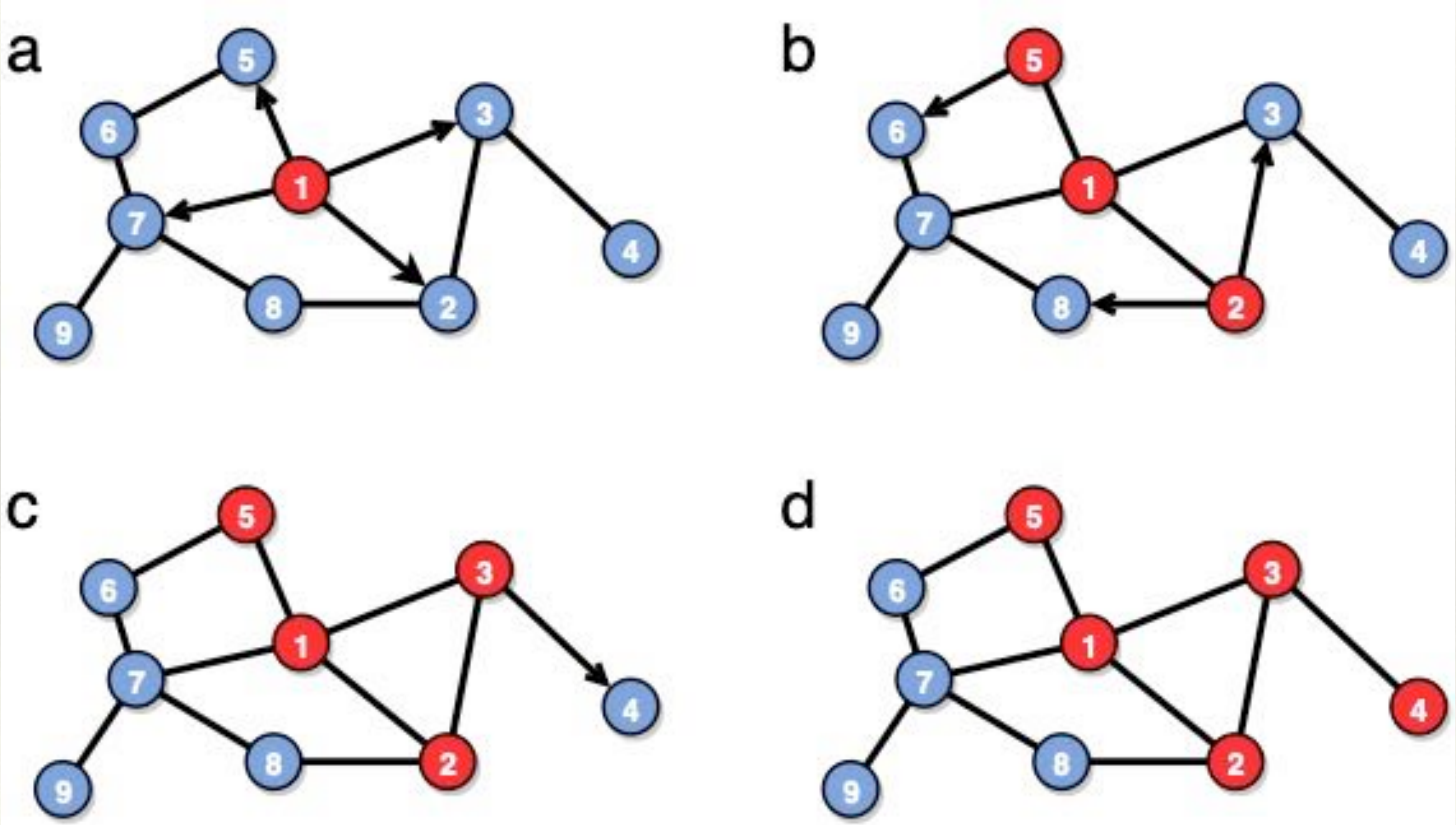
- **Alternative principle:** each of our contacts has their own influence

- **Independent cascade models** are based on node-node interactions!

Independent cascade models

- **Model dynamics:**
  - An active node *i* has a probability p_{ij} to convince its inactive neighbor *j* ($p_{ij} \neq p_{ji}$, in general)
  - All active nodes are considered in sequence: the inactive neighbor *j* of the active node *i* is activated with probability p_{ij}. All inactive neighbors of *i* have one chance to be persuaded by *i*
  - If a node *j* is activated, it has only one chance to activate its inactive neighbors

Independent cascade models



Independent cascade models

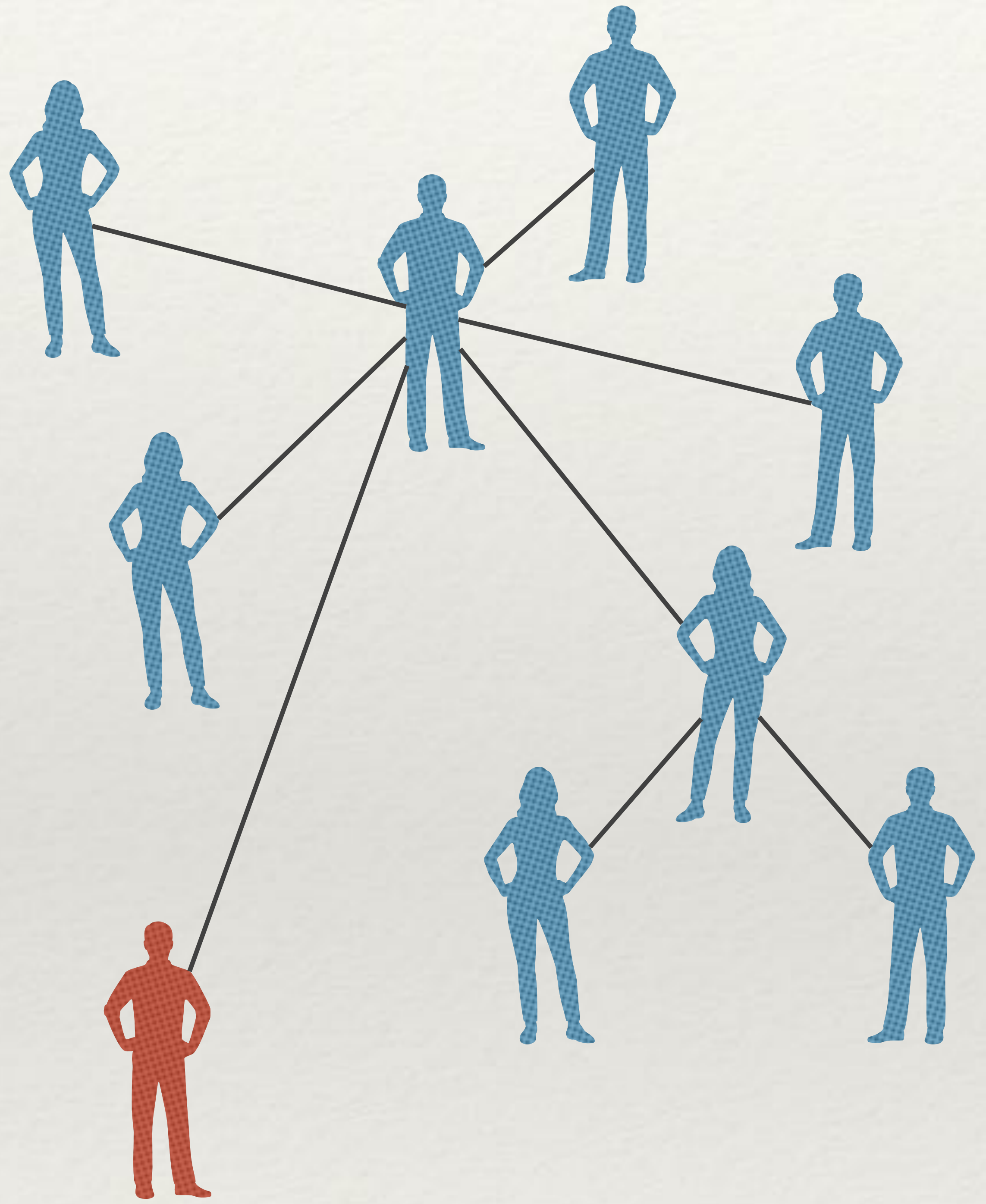
- **Remark**: the more active neighbors, the more likely a node will be activated
- **Independent cascade versus threshold models**:
  - Threshold models focus on the inactive nodes, independent cascade models on the active ones
  - Threshold models are (usually) **deterministic**: the dynamics depends on whether the threshold condition is satisfied or not
  - Independent cascade models are **probabilistic**: nodes are activated with a given probability —> it is more **difficult to control a cascade!**

Information diffusion

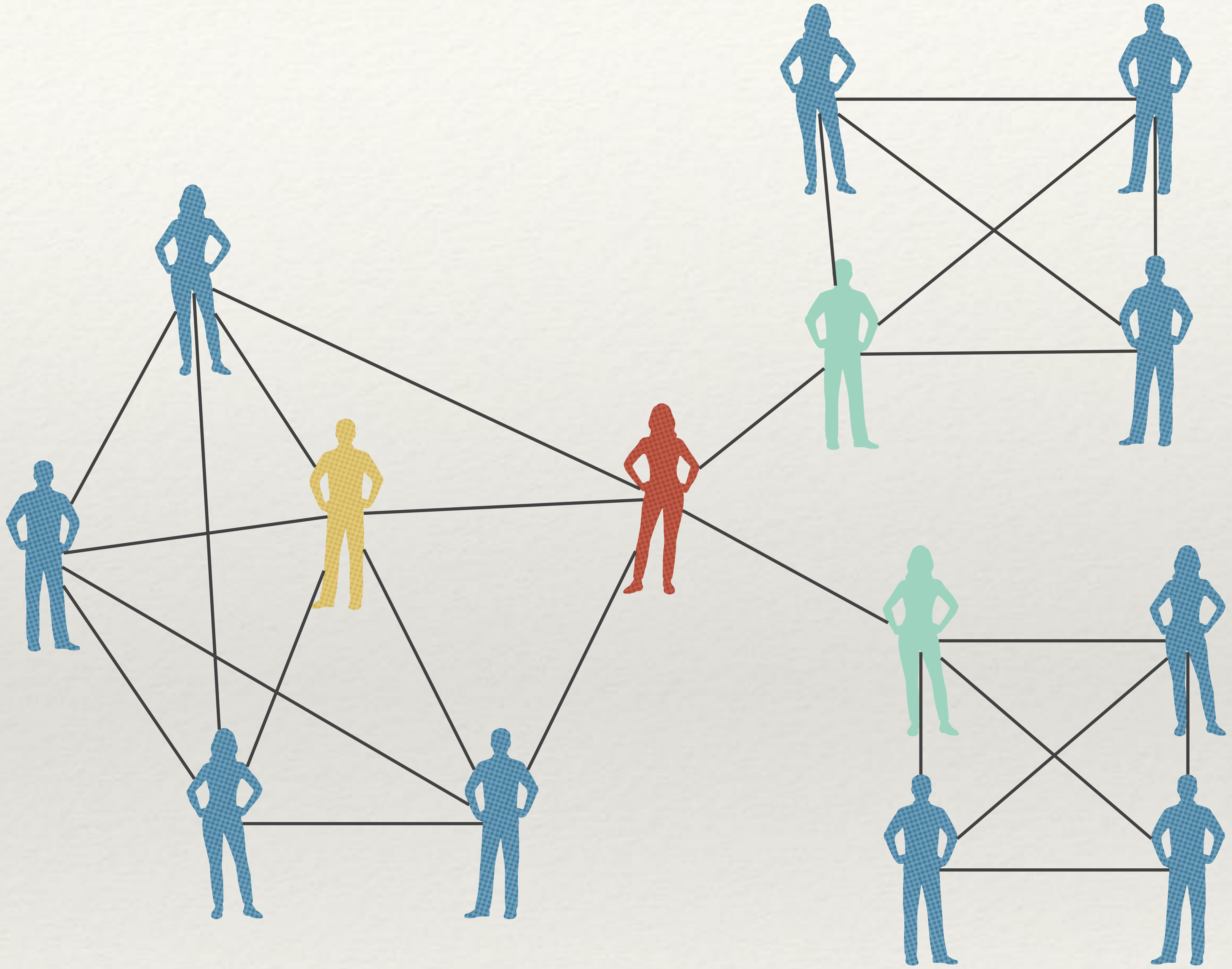
- **Problem**: models are too simple to be realistic
- **Solution**: more sophisticated variants!
- **Example**:
  - Probabilistic version of threshold model, in which the chance of being activated grows with the number of active neighbors (instead of the usual yes/no dynamics)
  - Similar to independent cascade model, except that the active neighbors **do not exert influence independently of each other!**
- **Complex contagion:** each new person exposing us to a new idea or product has greater influence than the previous ones!

Adrien Guille, Hakim Hacid, Cecile Favre, and Djamel A. Zighed. 2013. Information diffusion in online social networks: a survey. SIGMOD Rec. 42, 2 (May 2013), 17–28. DOI:https://doi.org/10.1145/2503792.2503797

Recall: real networks are heterogeneous



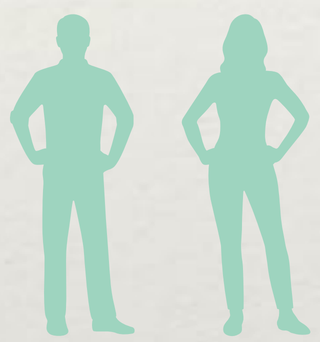
Rich-get-richer dynamics
(aka preferential attachment)



weak/strong ties, betweenness,
homophily, clusters

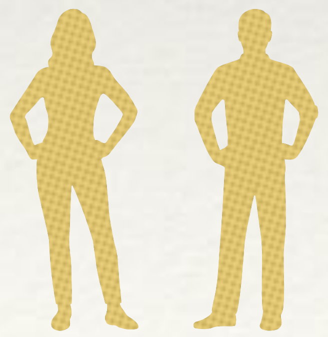
The role of weak ties

Threshold models highlight some important implications of 'the strength of weak ties' theory

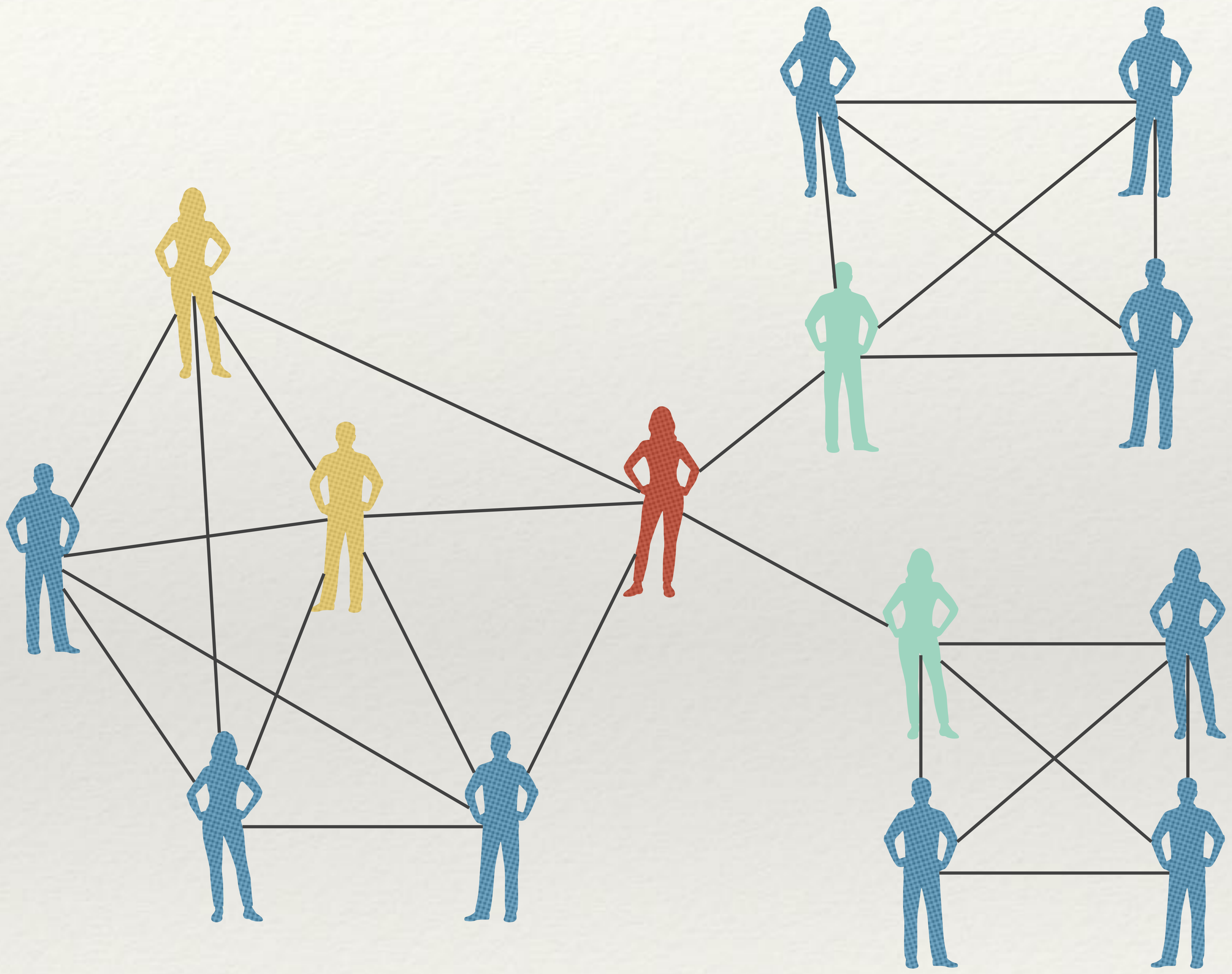


They receive very **fresh ideas** from other communities; not enough for adoption and spread (try threshold model with with $q = \dfrac{1}{2}$)

Bridges and weak ties are great for **spreading rumors** or jokes across the network, but **not for diffusion of innovation or social mobilization**



Strong ties can have more significant role for others in the community to take actions



Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. American Journal of Sociology, 113:702–734, 2007.

Complex contagion

Simple contagion: a single contact with an "infected" individual is usually sufficient to transmit the behavior.

Complex contagion: when behaviors require **social reinforcement**, a network with more clustering may be more advantageous, even if the network has a larger diameter.

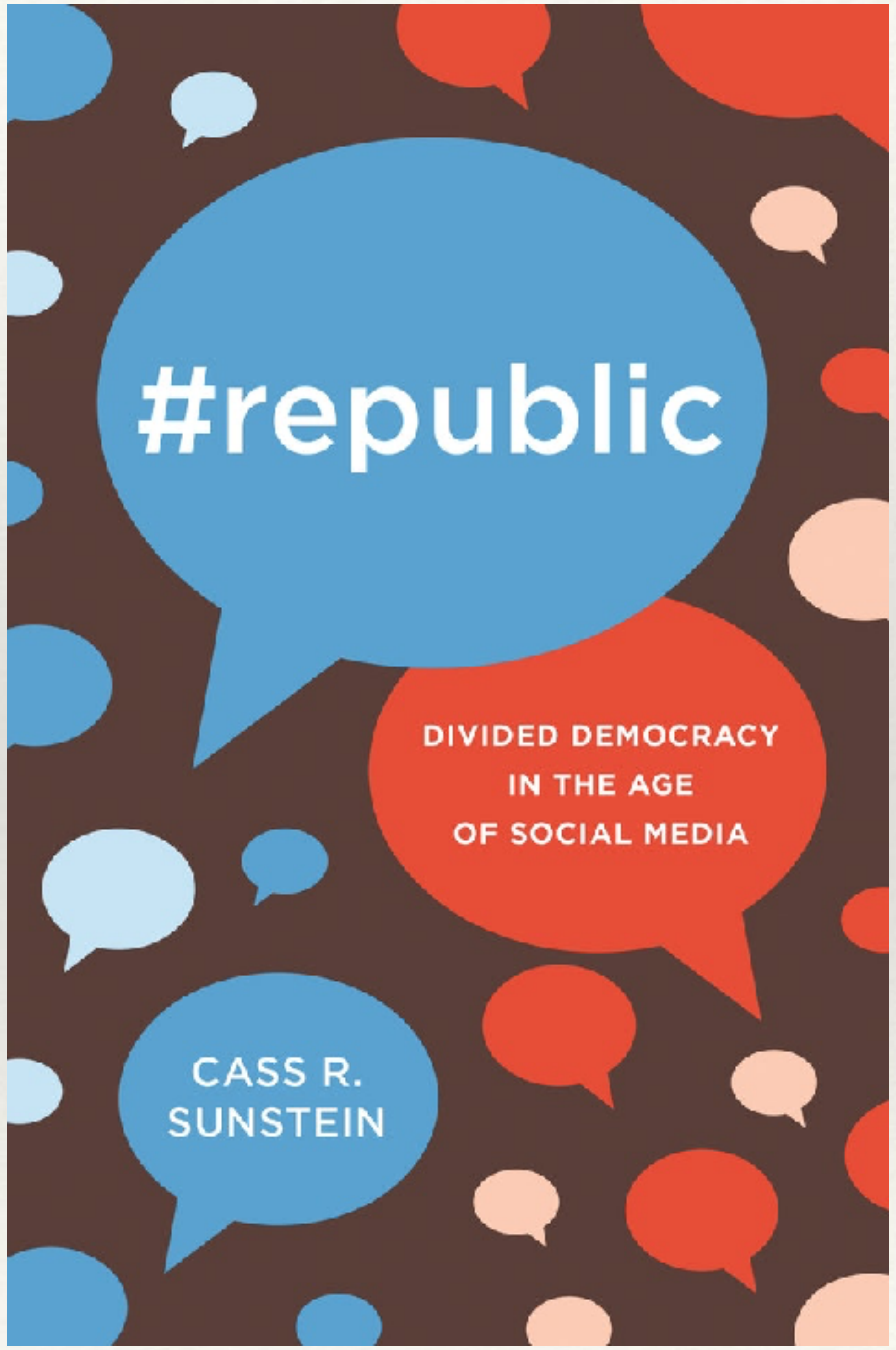
Centola investigated the effects of network structure on diffusion by studying t*he spread of health behavior through artificially structured online communities*

D. Centola, The Spread of Behavior in an Online Social Network Experiment, Science 03 Sep 2010: 1194-1197

Echo-chambers

Echo-chambers

- "Echo-chambers" metaphor superbly explained by Cass Sunstein
- Group of like-minded people amplifies their's members view
- Many factors:
  - Homophily (selection & influence)
  - Confirmation bias
  - Back-fire effect
  - Hypercorrection effect
  - Bandwagon effect

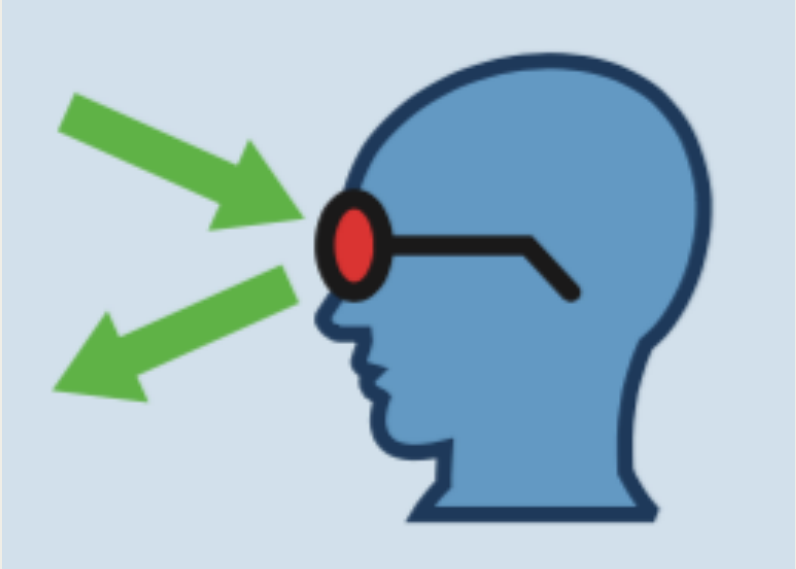


Psychological issues

Confirmation Bias

Hypercorrection Effect

Backfire effect



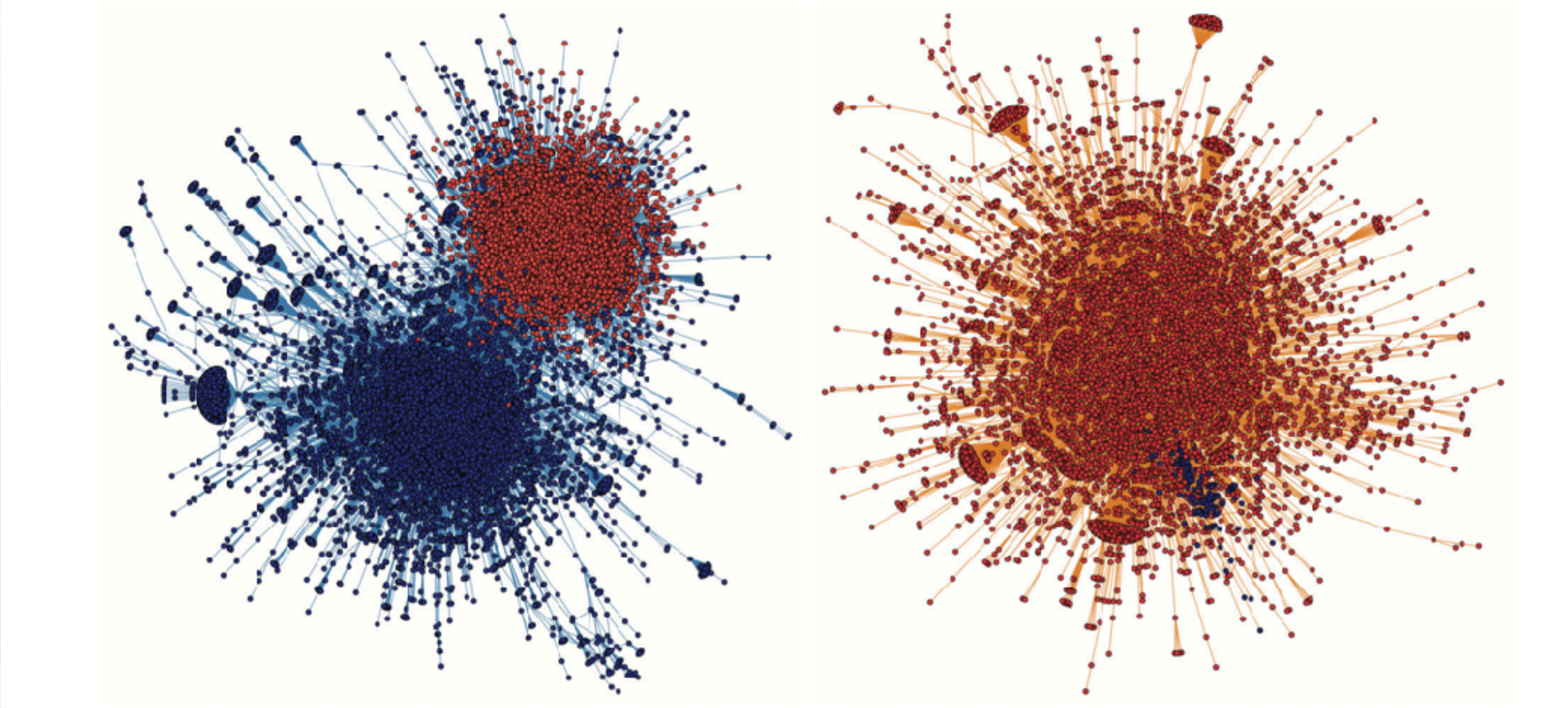
Bandwagon effect

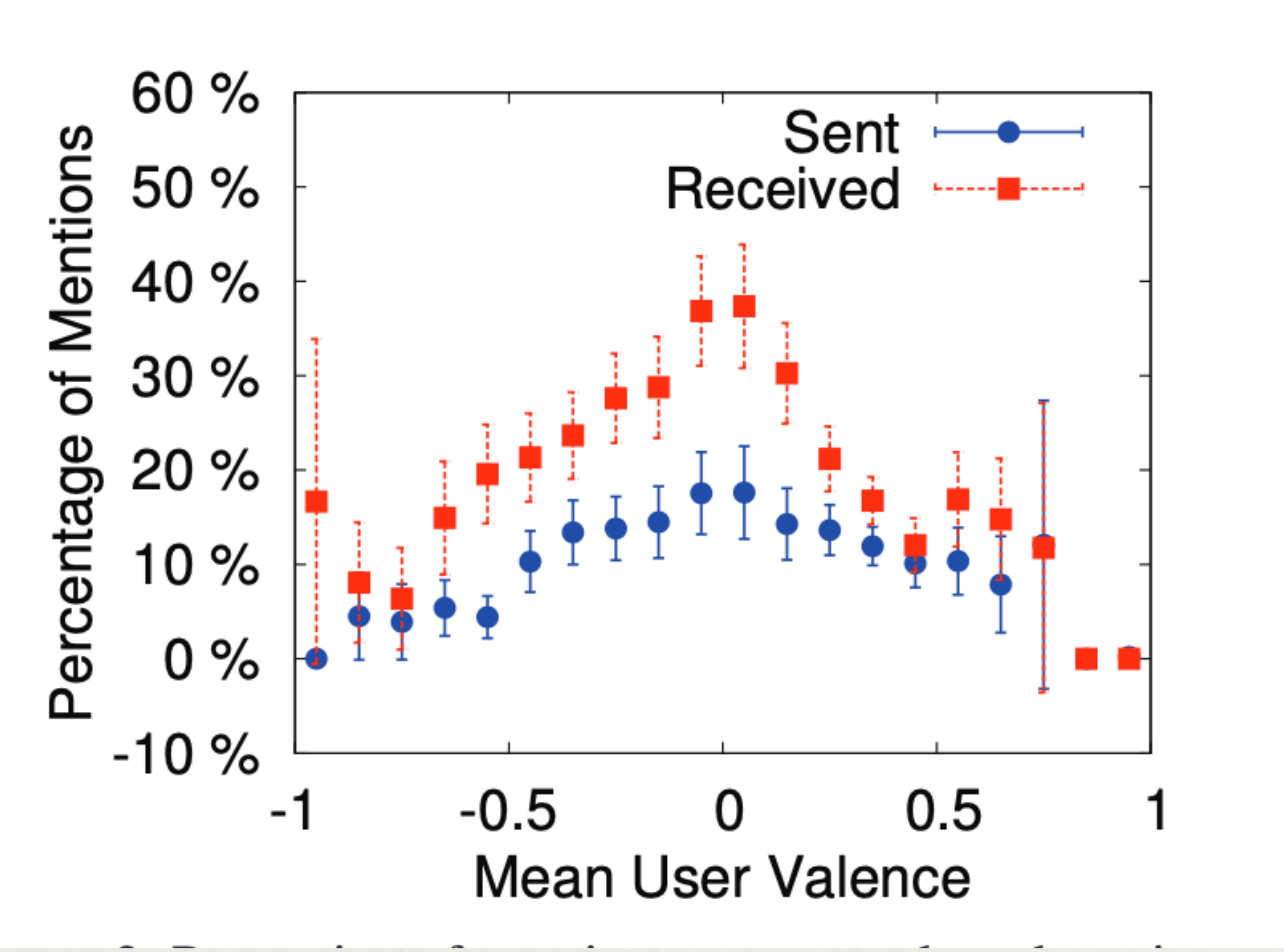
Butler AC, Fazio LK, Marsh EJ. The hypercorrection effect persists over a week, but high-confidence errors return. Psychon Bull Rev. 2011 Dec;18(6):1238-44. doi: 10.3758/s13423-011-0173-y. PMID: 21989771.

Lewandowsky, S. et al. (2012) Misinformation and Its Correction: Continued Influence and Successful Debiasing, Psychological Science in the Public Interest, 13(3), pp. 106–131. doi: 10.1177/1529100612451018.

Polarization emerges from radicalized segregation, but not necessarily a segregated network is also polarized.
However, some topics are strongly divisive (echo-chambers), others are not.

Political polarization on Twitter

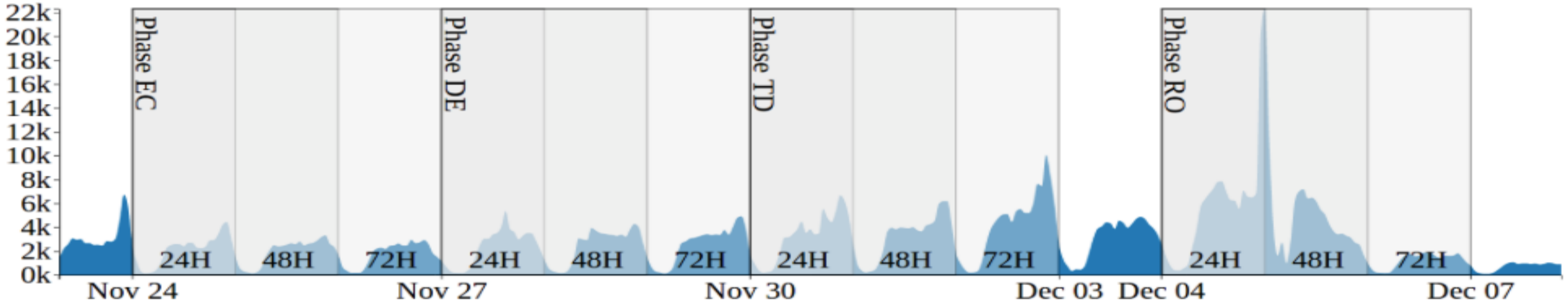




Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011, July). Political polarization on twitter. In *Proc. of the Intern. AAAI Conference on Web and Social Media* (Vol. 5, No. 1) - ICWSM 2011.

Italian 2016 Constitutional Referendum

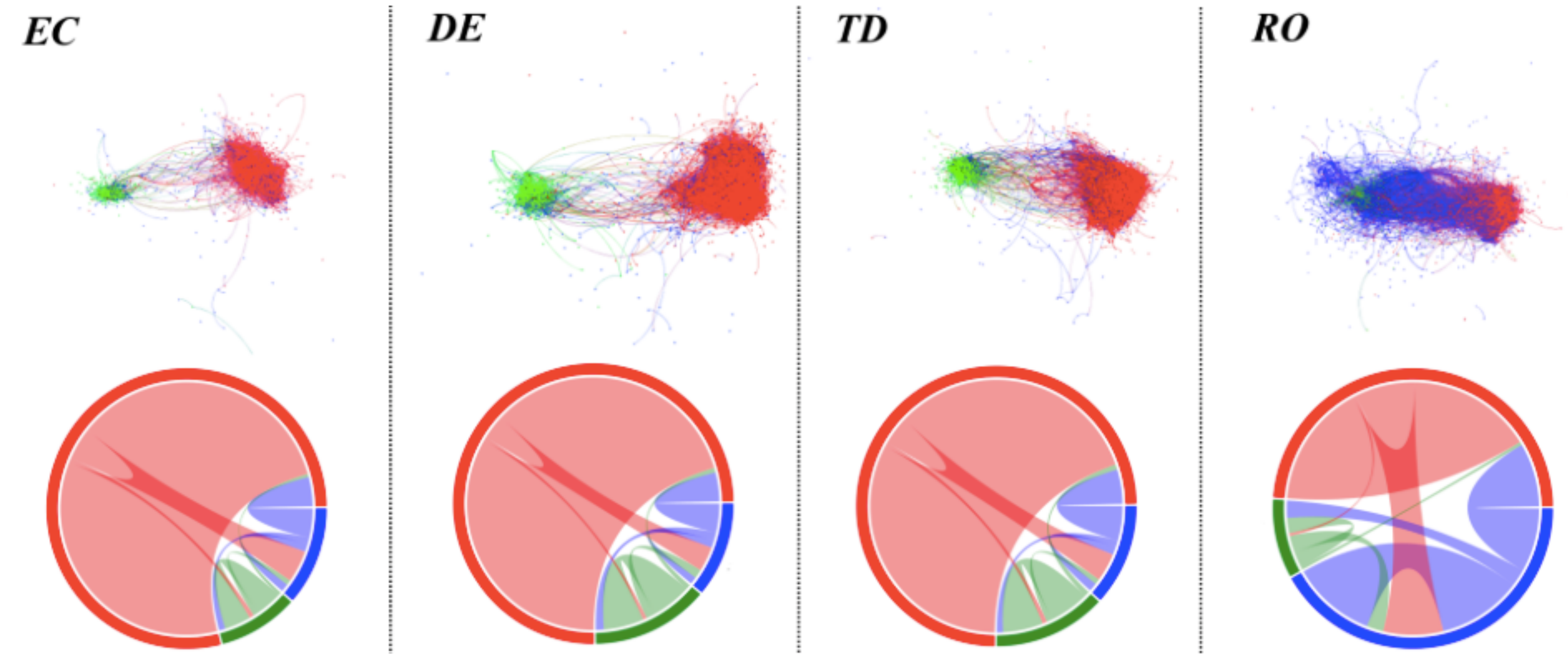
Collected Tweets



stance detected as **AGAINST**

stance detected as **IN FAVOR**

stance detected as **NONE**



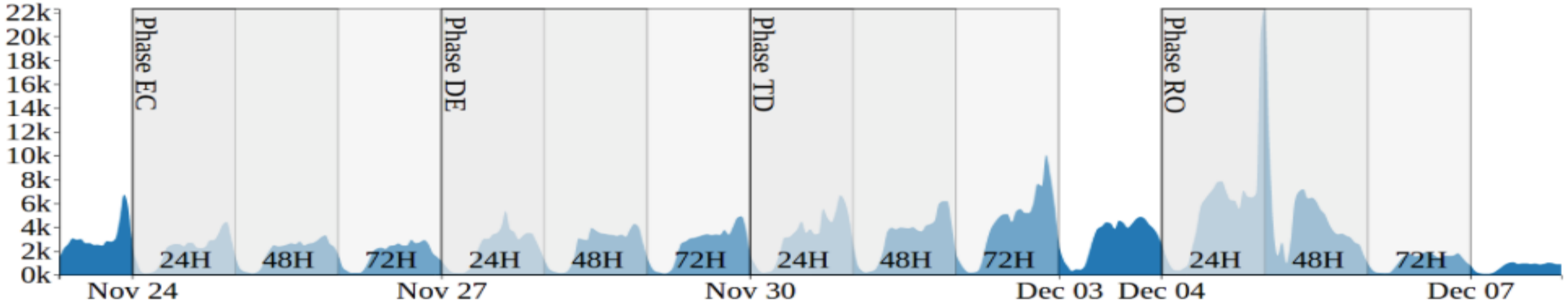
Retweet Network

strong signal of **homophily**

M Lai, M Tambuscio, V Patti, P Rosso, G. Ruffo, Stance Polarity in Political Debates: a Diachronic Perspective of Network Homophily and Conversations on Twitter, Data & Knowledge Engineering Journal, online: September 2019

Italian 2016 Constitutional Referendum

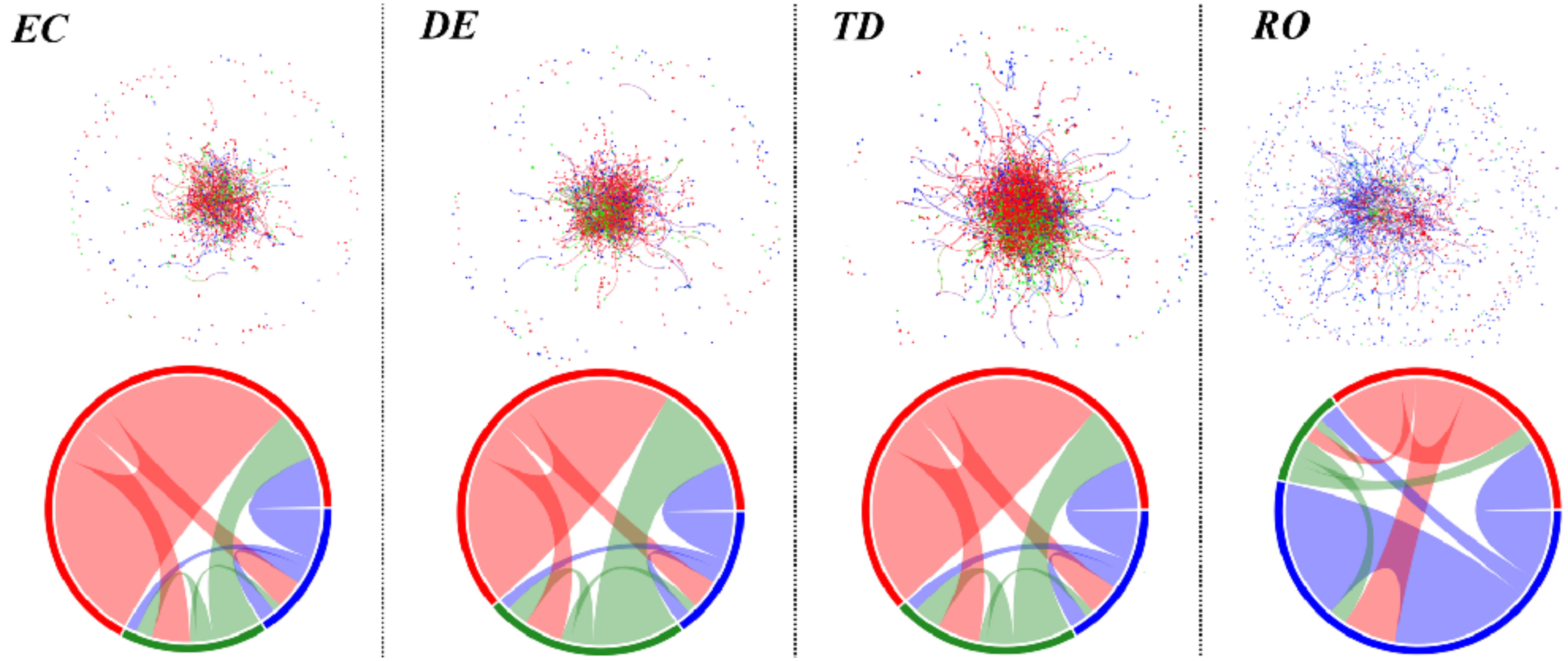
Collected Tweets



stance detected as **AGAINST**

stance detected as **IN FAVOR**

stance detected as **NONE**

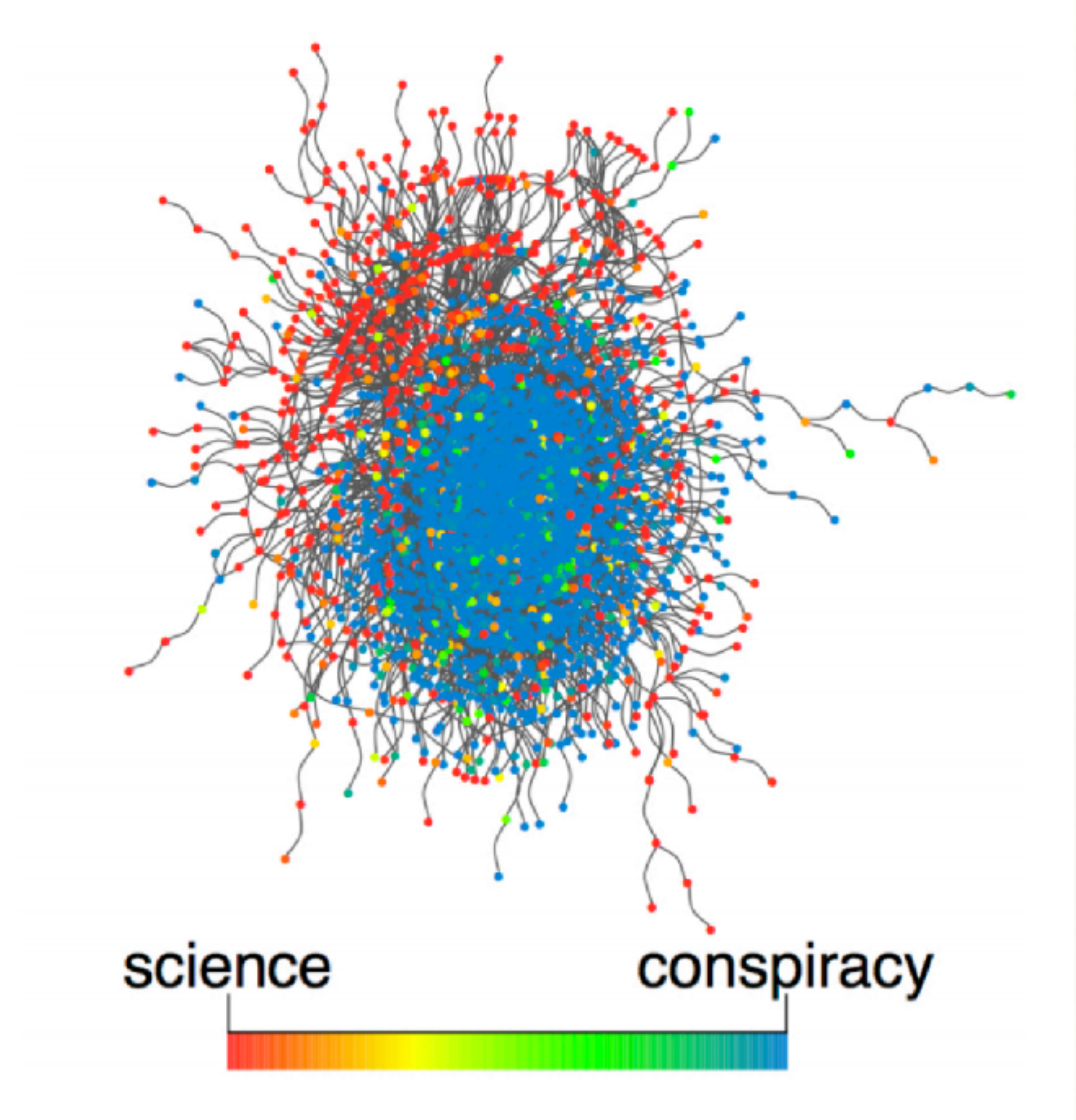


Mention Network

signal of **inverse homophily**

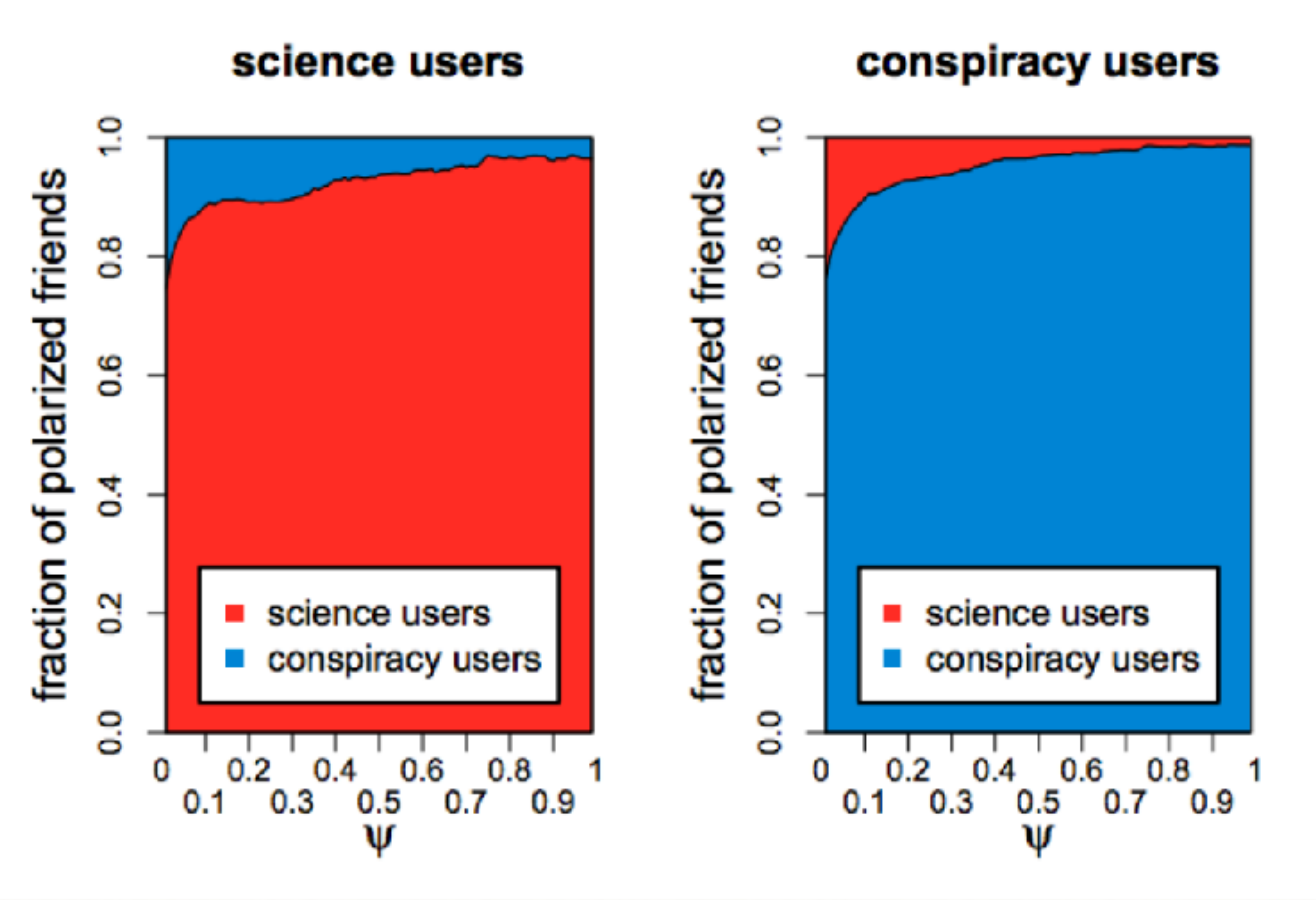
M Lai, M Tambuscio, V Patti, P Rosso, G. Ruffo, Stance Polarity in Political Debates: a Diachronic Perspective of Network Homophily and Conversations on Twitter, Data & Knowledge Engineering Journal, online: September 2019

Misinformation tends to polarize



Users engagement correlates with the number of friends having similar consumption patterns

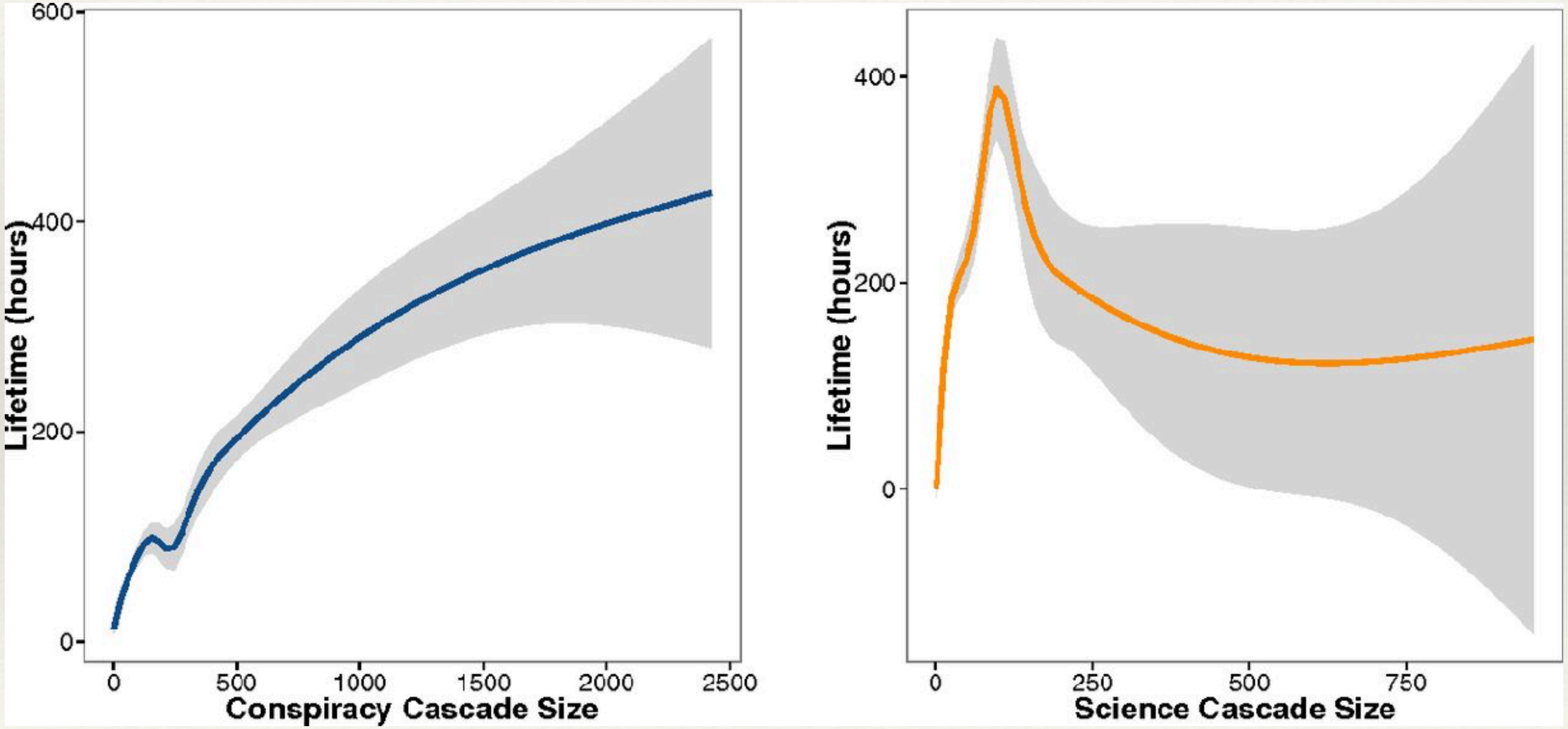
homophily!



A. Bessi, …., G. Caldarelli, W. Quattrociocchi, Viral Misinformation: The Role of Homophily and Polarization, WWW 2015 Companion, May 18–22, 2015, Florence, Italy.

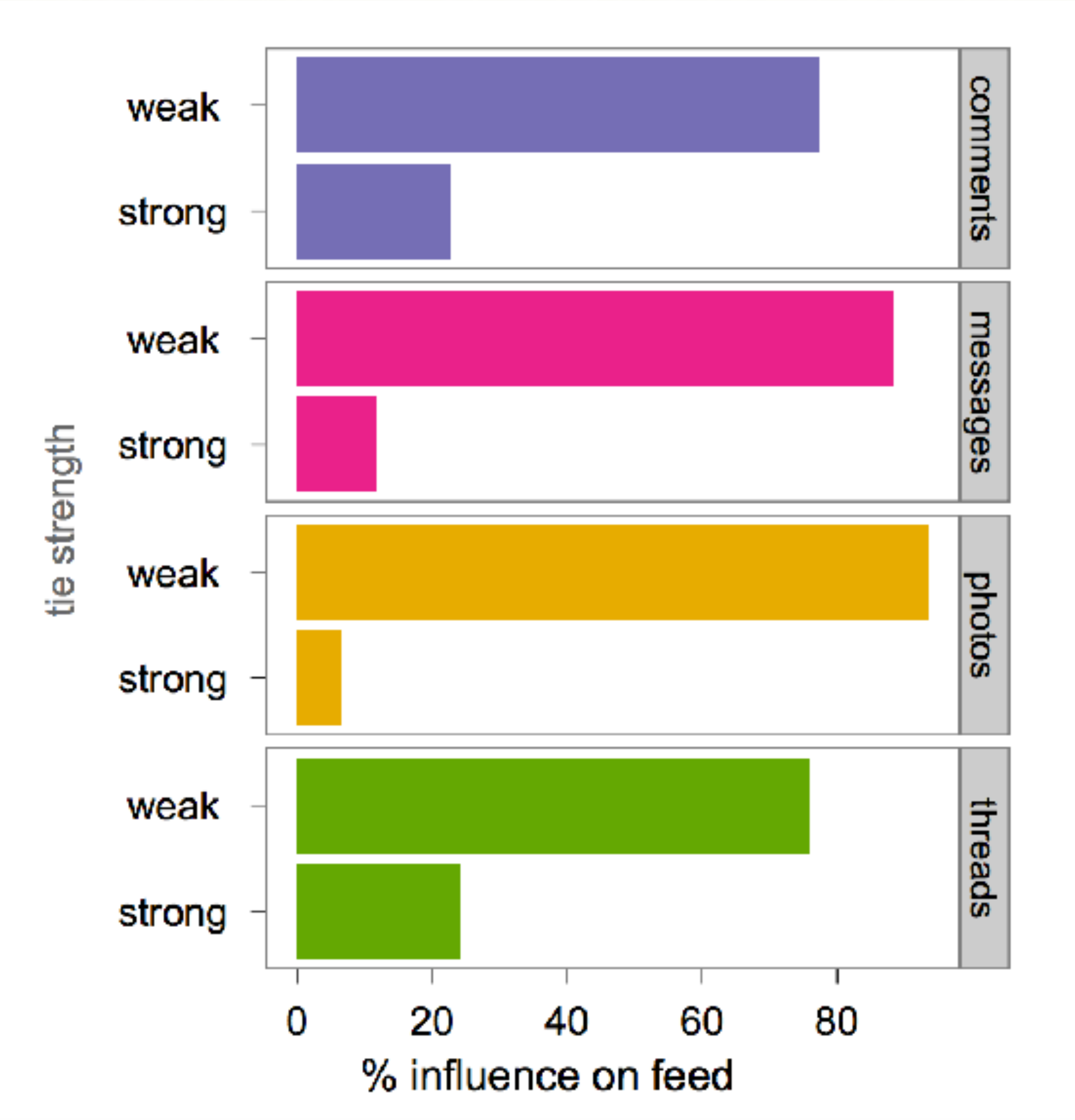
… and polarization fuels misinformation spread

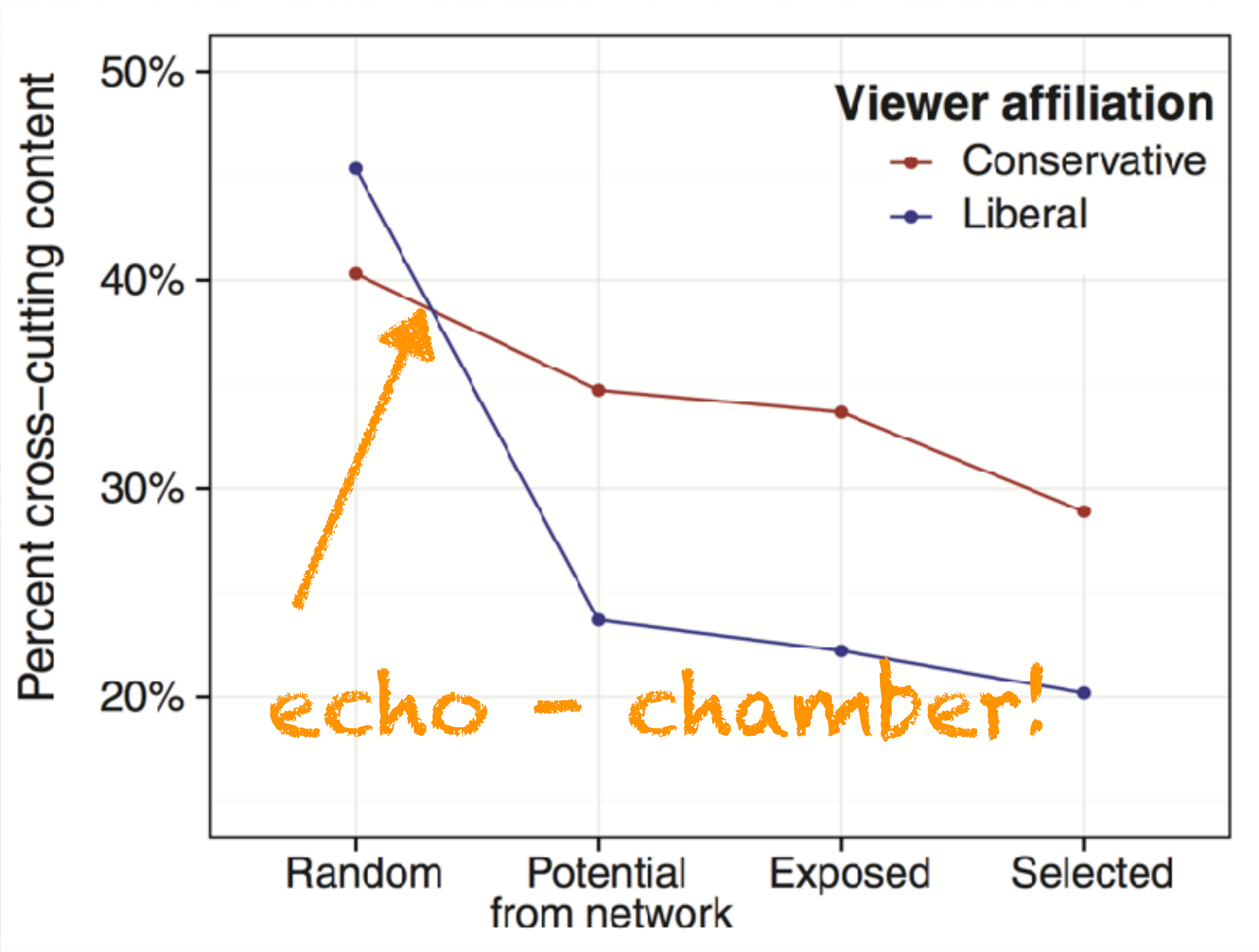
A data-driven percolation model of rumor spreading that demonstrates that homogeneity and polarization are the main determinants for predicting cascades' size



M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, W. Quattrociocchi, Echo chambers in the age of misinformation, PNAS, Jan 2016, 113 (3) 554-559; DOI: 10.1073/pnas.1517441113

"Weak ties" are important, too





E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. 2012. The role of social networks in information diffusion. In Proc of the 21st Int. Conf. on World Wide Web (WWW '12). ACM, New York, NY, USA, 519–528. DOI:https://doi.org/10.1145/2187836.2187907

E. Bakshy, S. Messing, L. Adamic, Exposure to ideologically diverse news and opinion on Facebook, Science 05 Jun 2015: Vol. 348, Issue 6239, p. 1130-1132, DOI: 10.1126/science.aaa1160(Bakshy et al. 2015)

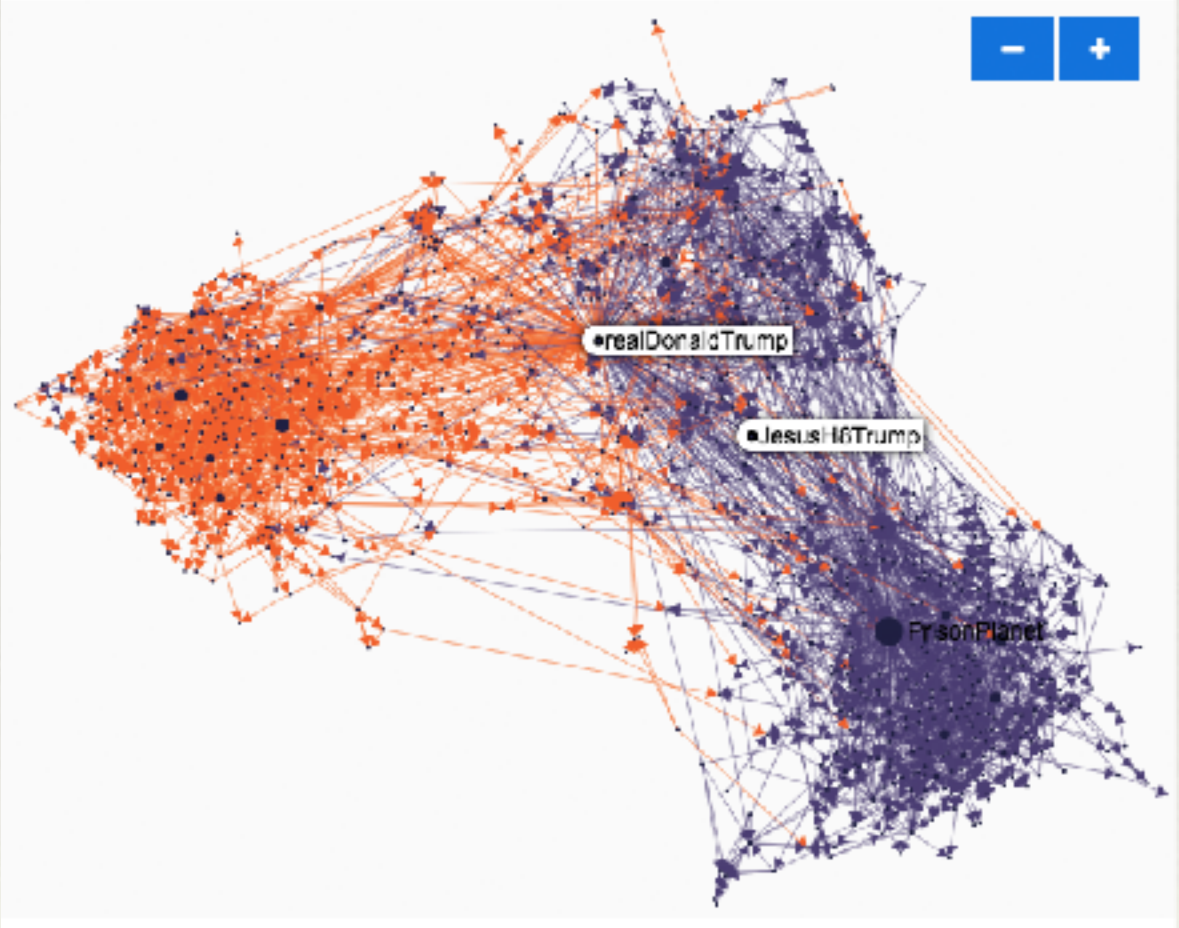
Analyzing the structure of a misinformation network

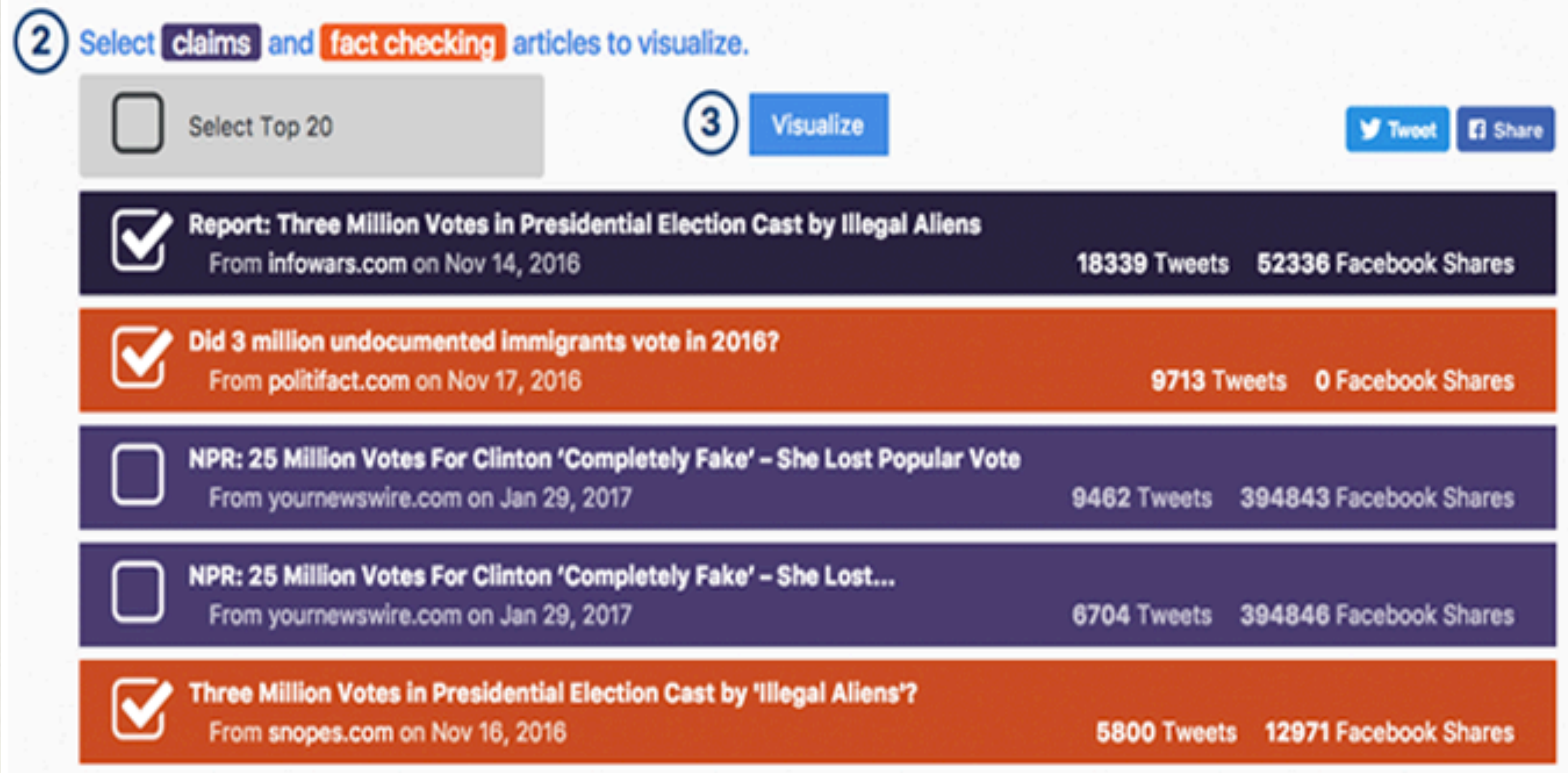
- *What are the structural and dynamic characteristics of the core of the misinformation diffusion network, and who are its main purveyors?*
- "As we move from the periphery to the core of the network, fact-checking nearly disappears, while social bots proliferate."

HOAXY beta ®

Visualize the spread of claims and fact checking.

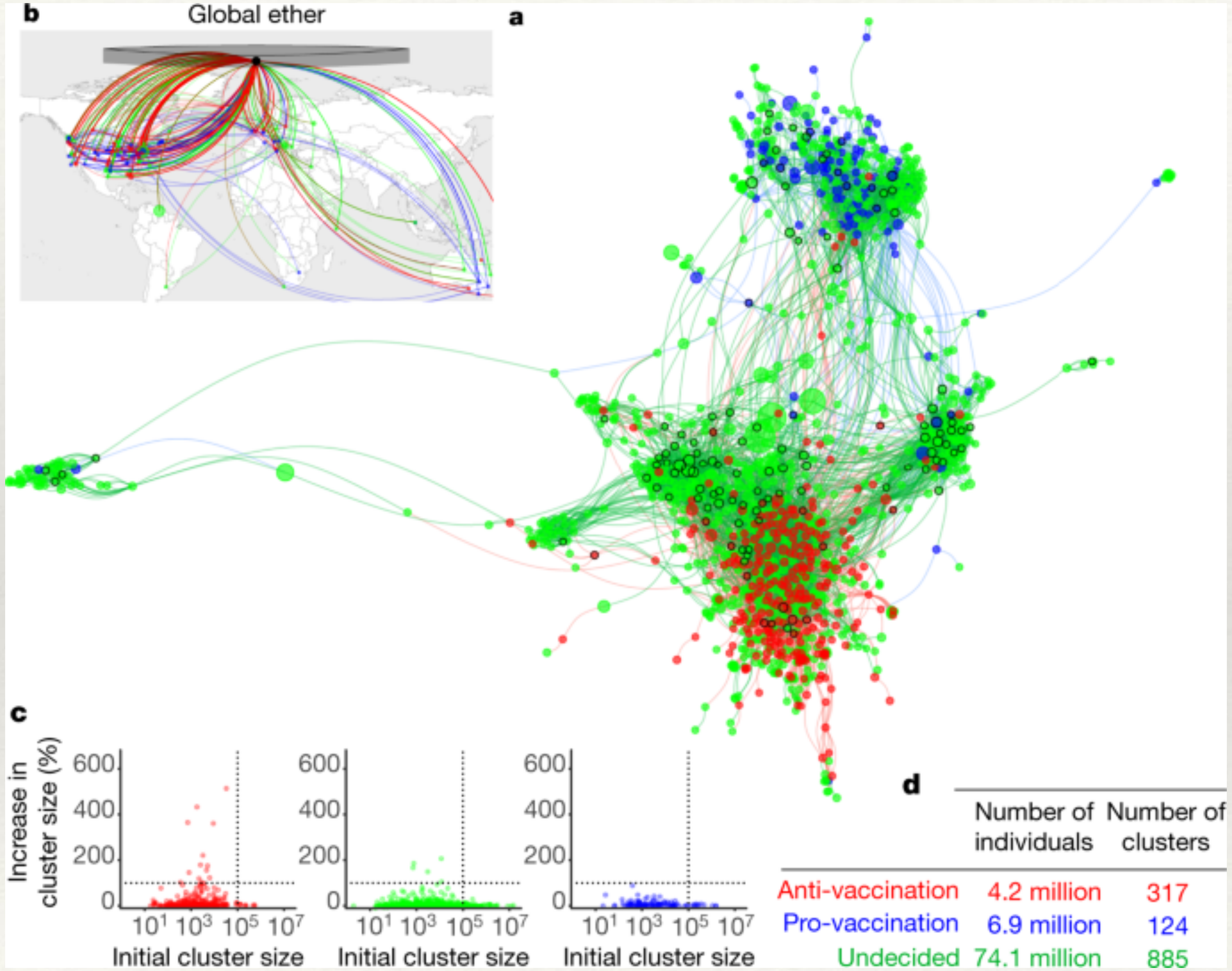
https://hoaxy.iuni.iu.edu



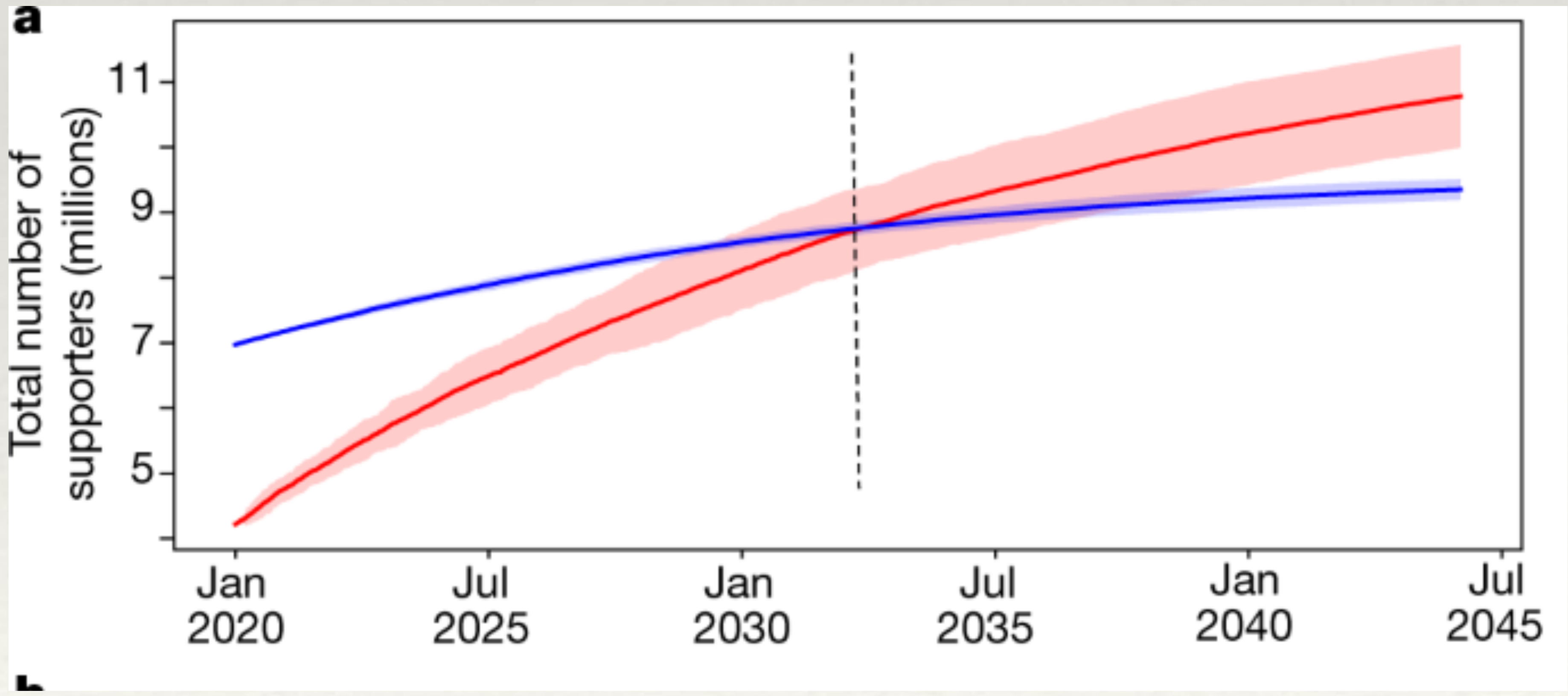


Shao C, Hui P-M, Wang L, Jiang X, Flammini A, Menczer F, et al. (2018) Anatomy of an online misinformation network. PLoS ONE 13(4):e0196087. https://doi.org/10.1371/journal.pone.0196087

The role of the undecided



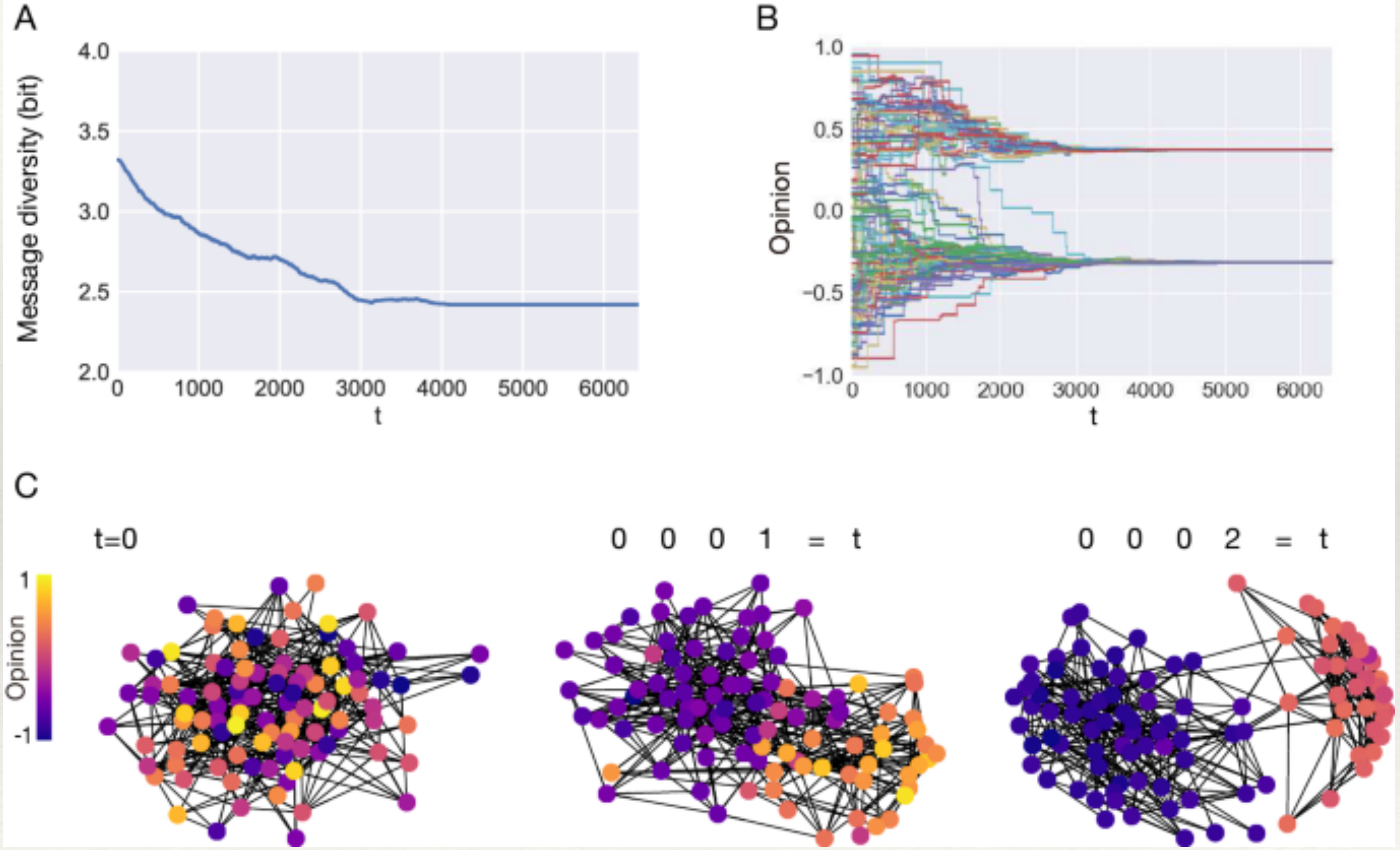
- Theoretical prediction for the future total size of anti-vaccination and pro-vaccination support
- Under the present conditions, it predicts that total anti-vaccination support reaches dominance in around 10 years



Johnson, N.F., Velásquez, N., Restrepo, N.J. *et al.* The online competition between pro- and anti-vaccination views. *Nature* **582,** 230–233 (2020).

The role of unfollowing

- The **model dynamics** show that even with minimal amounts of **influence** and **unfriending**, the social network rapidly devolves into polarized communities
- Predictions are consistent with **empirical data** from Twitter



Sasahara, K., Chen, W., Peng, H. Ciampaglia, G. L., Flammini, A., Menczer, F. Social influence and unfollowing accelerate the emergence of echo chambers. *J Comput Soc Sc* (2020). https://doi.org/10.1007/s42001-020-00084-7

Modeling epidemics on networks

Epidemic spreading

the Black Death

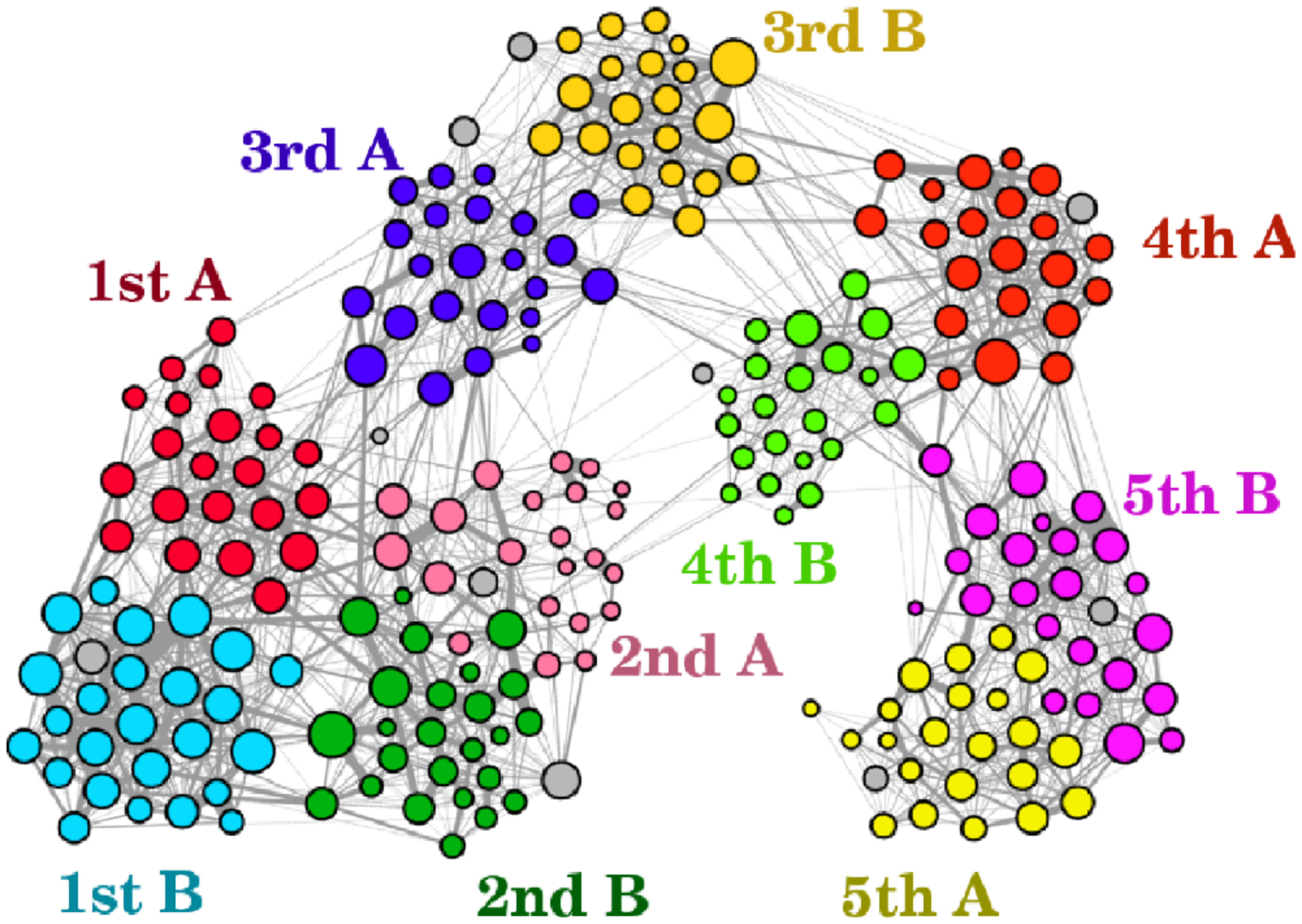
Probably originated in Central Asia, it spread throughout all of Europe between 1346 and 1353. The Black Death is estimated to have killed 30-60% of Europe’s population

Epidemic spreading

- Problems:
  - Nowadays the speed of epidemic spreading has increased enormously due to advances in transportation: someone contracting Ebola in Africa can travel to Europe, America and Asia and spread the disease before being aware of it
  - Technology has created new types of epidemics: computer viruses & malware spread over the Internet. Mobile phone viruses spread via Bluetooth or MMS. Misinformation spreads through social media, etc.

Contact networks

- Epidemics spread on contact networks, such as networks of physical contacts, transportation, the Internet, email, online social networks, and mobile phone communication



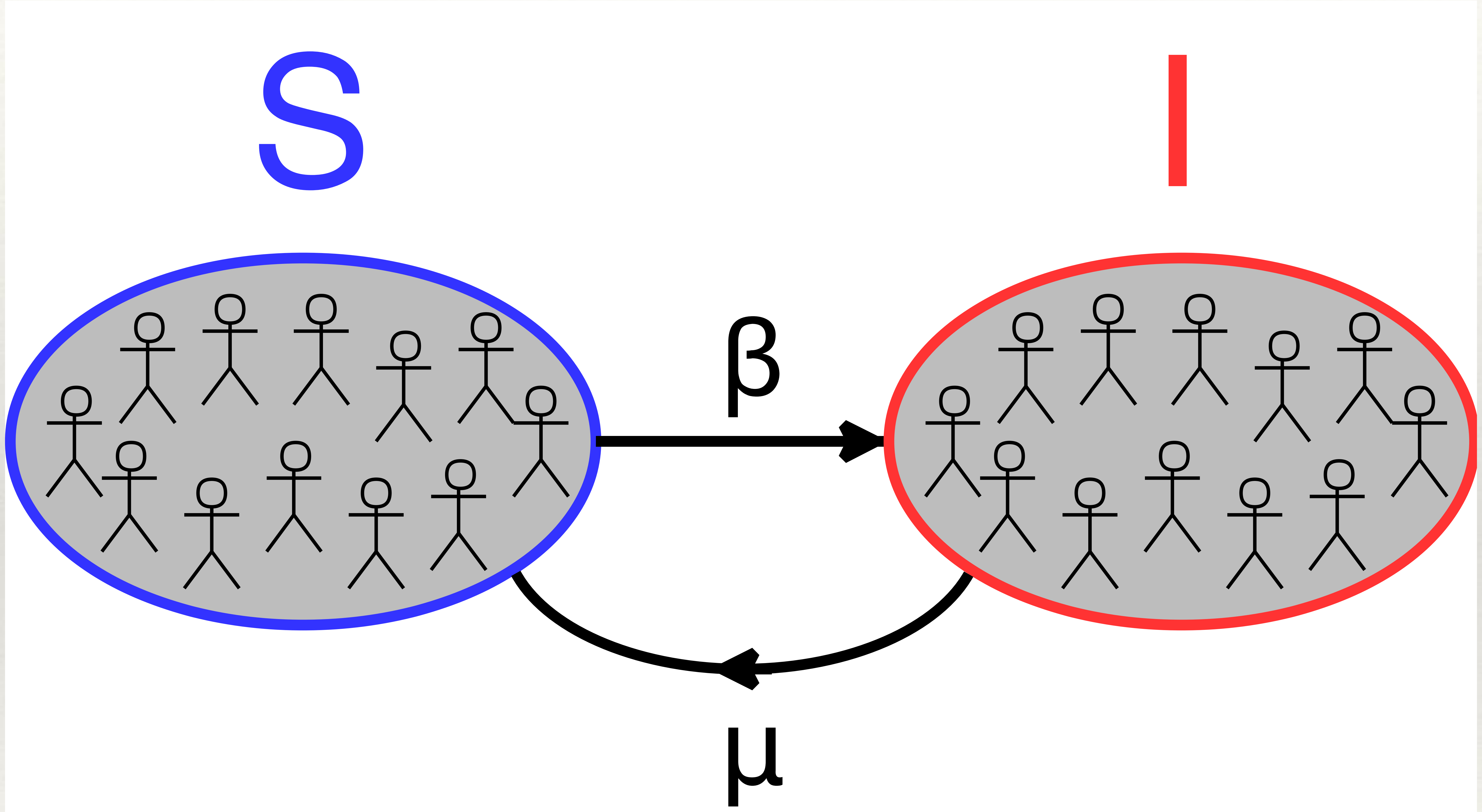
Epidemic models

- Classic epidemic models divide the population into **compartments**, corresponding to different stages of the disease
  - **Key compartments**:
    - **Susceptible (S)**: individuals who can contract the disease
    - **Infected (I)**: individuals who have contracted the disease and can transmit it to susceptible individuals
    - **Recovered (R)**: individuals who recovered from the disease and cannot be infected anymore

The SIS model

- Just two compartments: **Susceptible (S)** and **Infected (I)**
- Dynamics:
  - A susceptible individual gets infected with a probability β (**infection rate**)
  - An infected individual recovers and becomes susceptible again with a probability μ (**recovery rate**)
  - The model applies to diseases that do not confer long-lasting immunity (e.g., common cold)

The SIS model

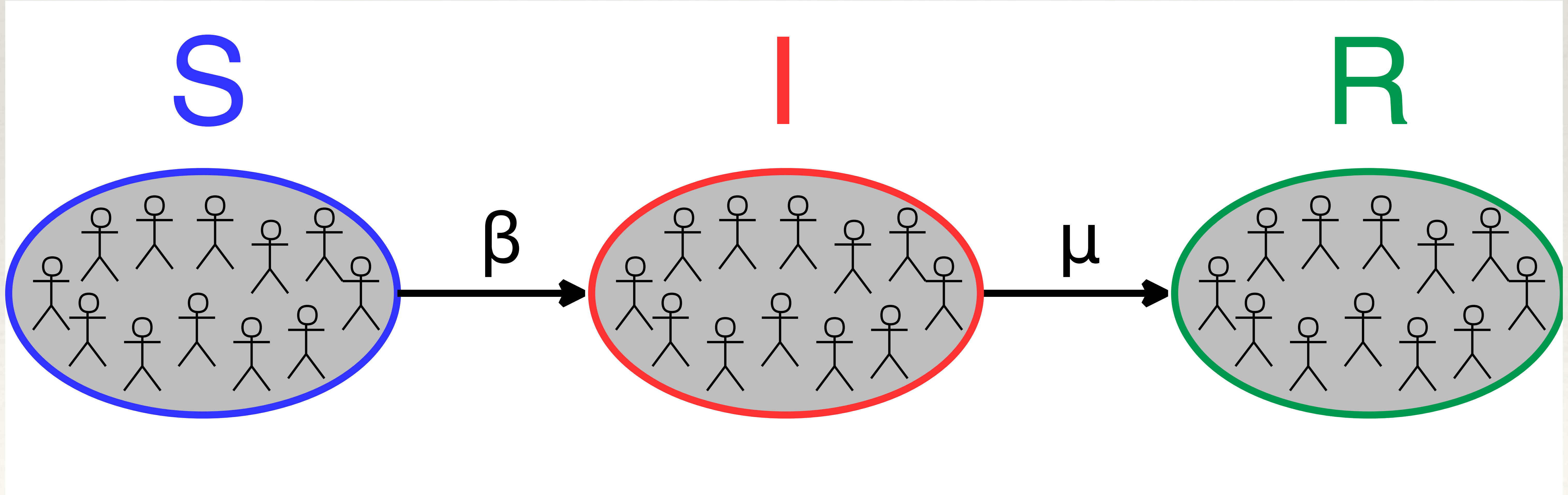


The SIS model

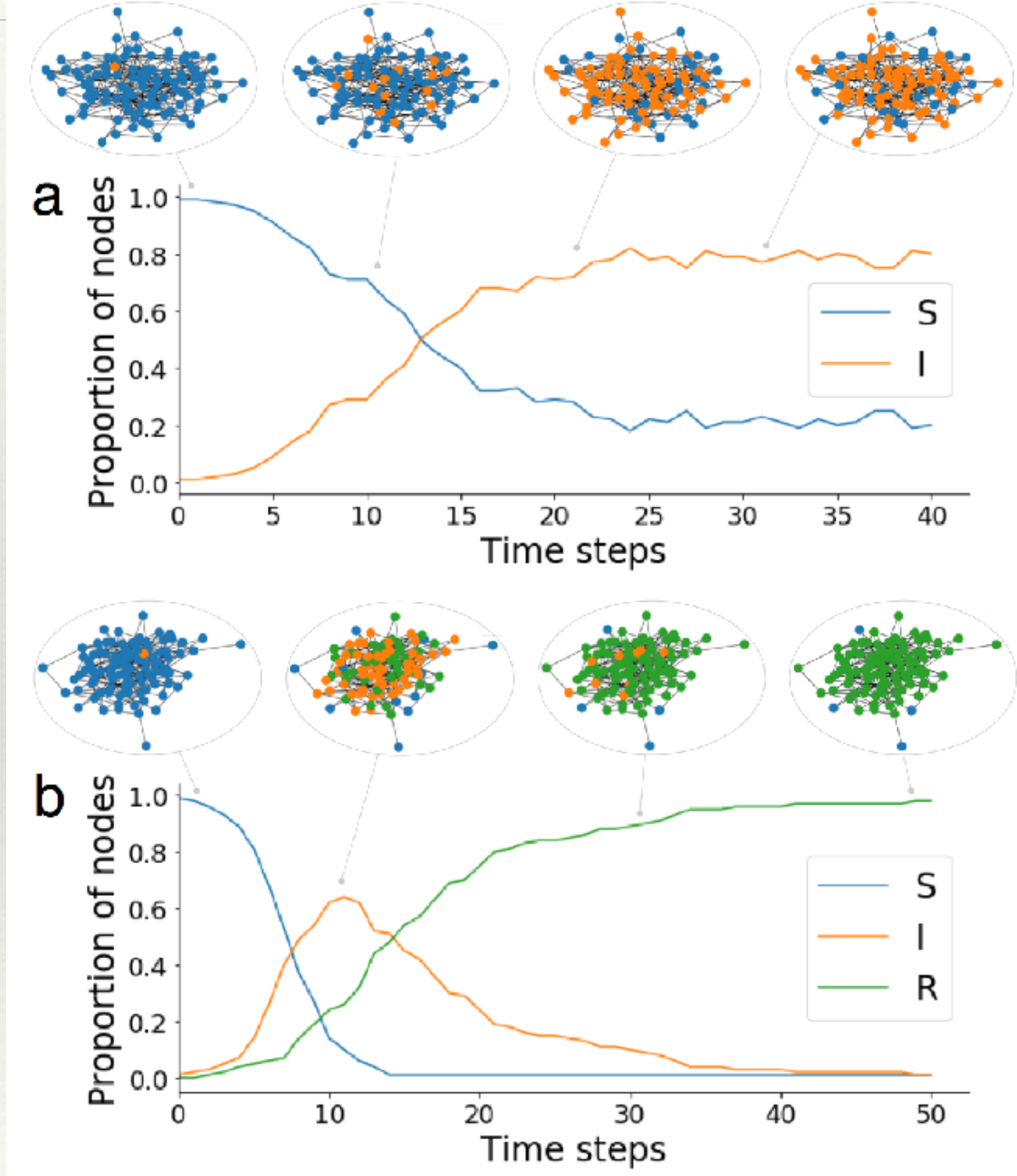
- Simulation of SIS dynamics on networks:
  - Take a network (e.g., a random network or a real contact network)
  - A number (fraction) of the nodes are infected (e.g., at random), all others are susceptible
  - All nodes are visited in sequence
  - For each node i:
    - If i is susceptible, loop over its neighbors: for each infected neighbor, i becomes infected with probability β
    - If i is infected, it becomes susceptible with probability μ

The SIR model

- **Difference from SIS model**: when infected individuals recover, they do not become susceptible again, but they are moved to the compartment R and play no further role in the dynamics
- The model applies to diseases that confer long-lasting immunity (e.g., measles, mumps, rubella, etc.)



Epidemic spreading



- Three characteristic stages of the **dynamics**:
  - **Initial stage**: just a few people are infected, and the diffusion of the epidemic is irregular and slow
  - **Ramp-up phase of exponential growth**, that can quickly affect a large number of people
  - **Stationary state**, in which the disease is either endemic, i.e. it affects a stable fraction of the population over time, or eradicated

Homogeneous mixing

- **Hypothesis**: every individual is in contact with every other
- **Consequence**: all individuals in the same compartment have identical behavior and only the relative proportions of people in the various compartments matter for the model dynamics
- Justified for a small population, e.g., the inhabitants of a little village where all people are in touch with each other.
- In real large-scale epidemics, individuals can only be infected by the people they come in contact with. In this case it is **necessary to reconstruct the actual network of contacts**

SIS & SIR models on networks

- **Start**: homogeneous contact network, with all nodes having degree approximately equal to $<k>$
- **Early stage**: few people are infected, so we can assume that every infected individual is in contact with mostly susceptible individuals
- Each infected individual can transmit the disease to about $<k>$ people at each iteration —> the expected number of people infected by a single person after one iteration is $\beta<k>$
- If there are I infected individuals, we expect to have $I_{sec} = \beta<k>I$ new infected people after one iteration and $I_{rec} = \mu I$ recovered people

SIS & SIR models on networks

- **Threshold condition** for epidemic spreading: $I_{sec} > I_{rec}$

$$\beta\langle k\rangle I > \mu I \implies R_0 = \frac{\beta}{\mu}\langle k\rangle > 1$$

- $R_0 = \beta\langle k\rangle/\mu$ is the **basic reproduction number**
- If $R_0 < 1$, the initial outbreak **dies out in a short time**, affecting only a few individuals
- If $R_0 > 1$, the epidemic keeps spreading

SIS & SIR models on networks

- **Problem**: real contact networks are not homogeneous
- **Hubs drastically change the scenario.** On contact networks with hubs there is effectively no epidemic threshold —> even diseases with low infection rate and/or high recovery rate may end up affecting a sizable fraction of the population!
- **Reason**: even if the infection rate is low, the process is likely to eventually infect a hub, via one of its many contacts; the hub can in turn infect a large number of susceptible individuals, including possibly other hubs, and so on
- Effective disease containment strategies should aim at isolating/vaccinating individuals with many contacts. The latter can be identified by picking the endpoints of randomly selected links, as this increases the chance to bump into hubs. So, don't vaccinate a random sample of the population: **vaccinate their friends**!

Modeling the spread of misinformation

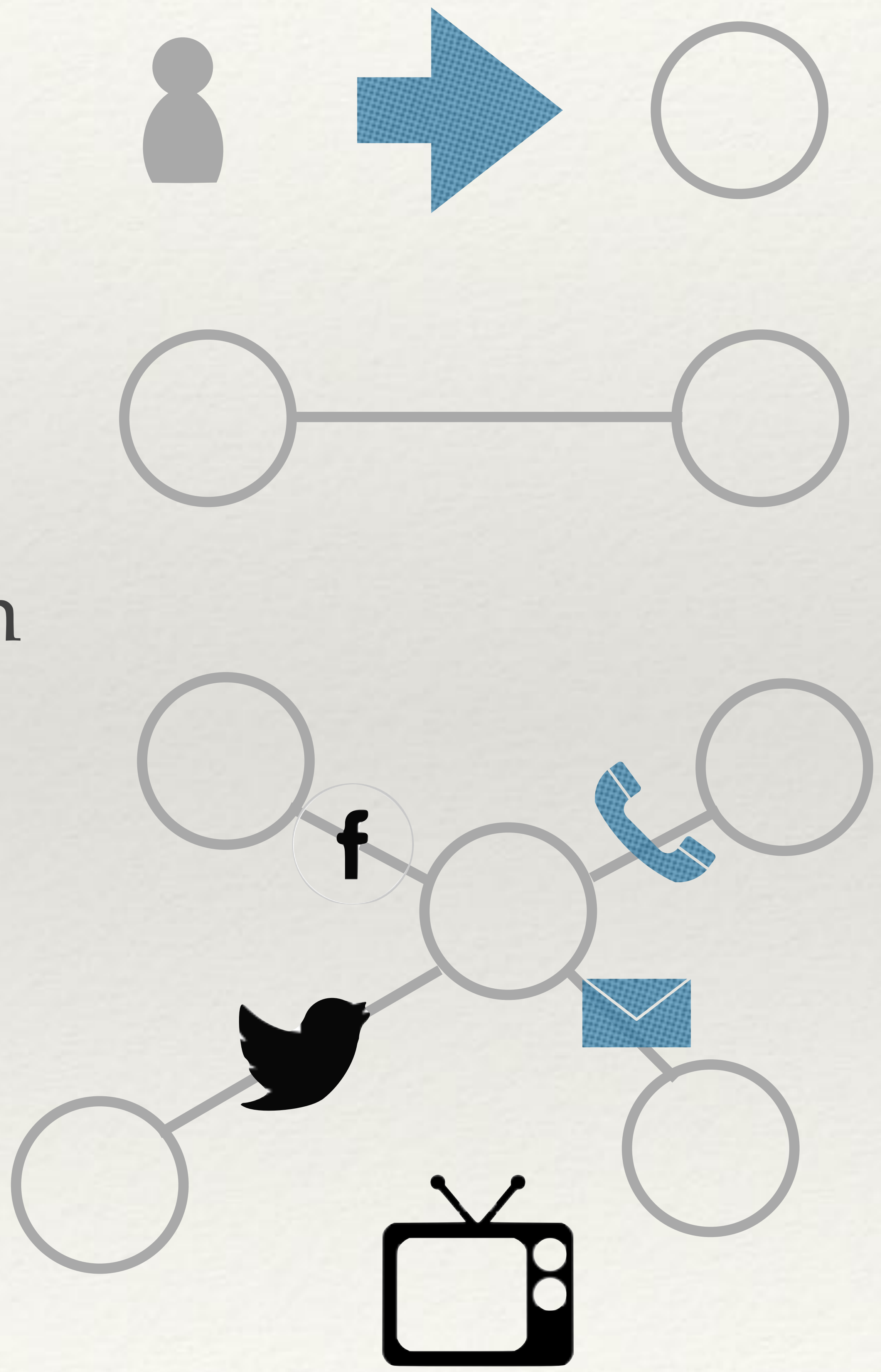
Questions

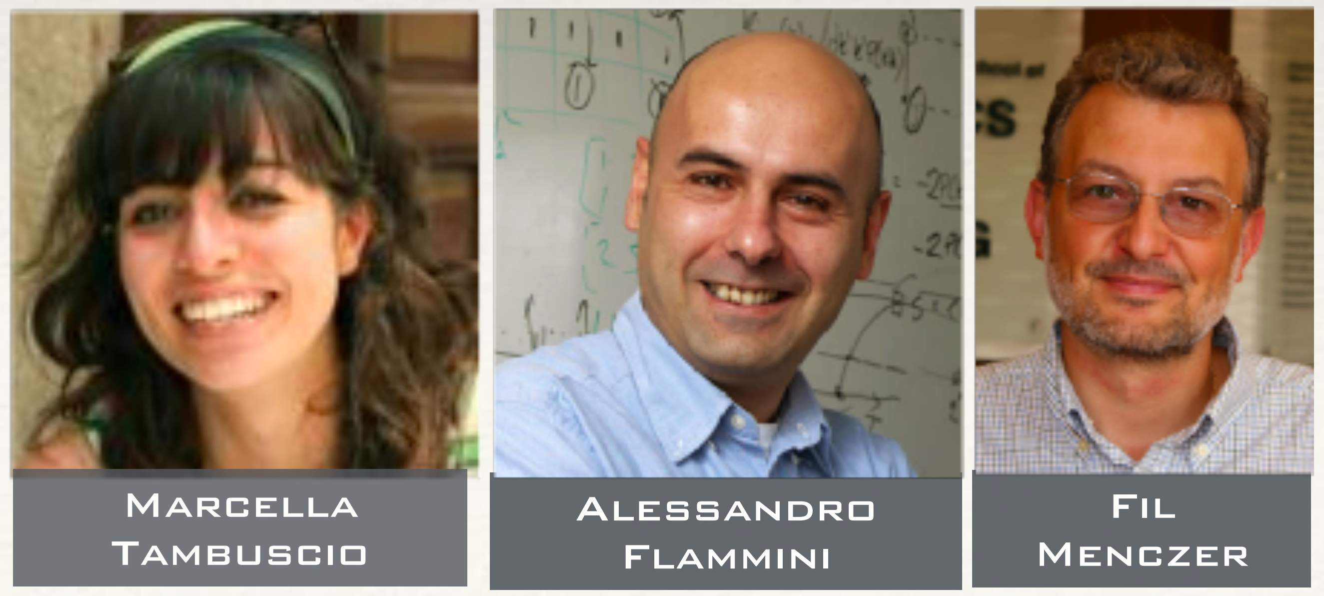
- Is fact-checking effective against the diffusion of fake-news?

- Do "echo-chambers" play a role as inhibitors or facilitators of fake-news spreading?

Networks and their context

- nodes are **actors** involved in a **generic** social network (no assumption is given)
- links are **social relationships**
- nodes can be exposed to news from both **internal and external sources** and via different communication devices

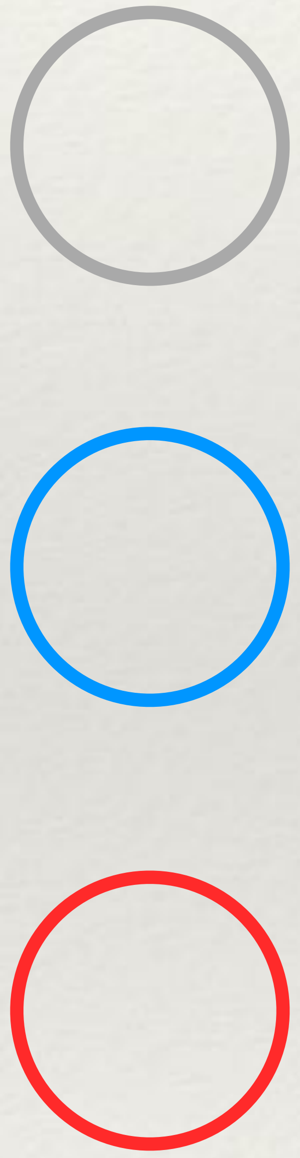


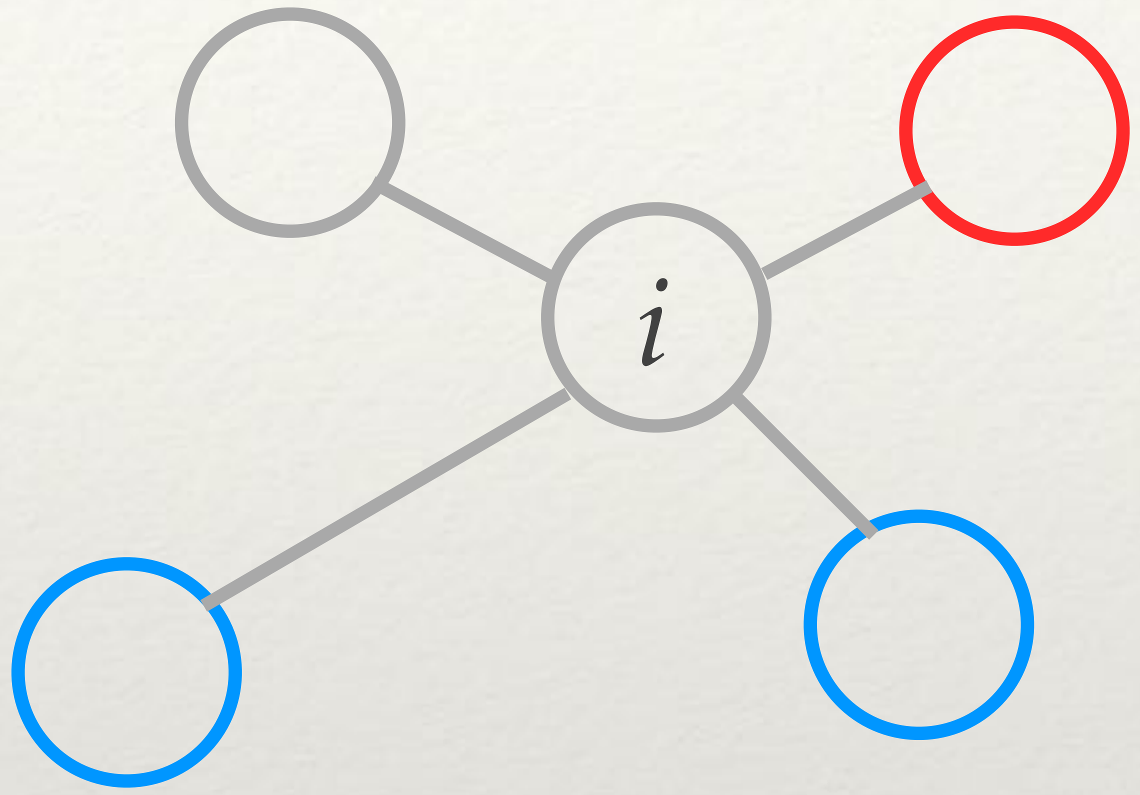


- **network topologies** can be created artificially or built from real data
- The **news is factually false** (can be debunked or someone else has already debunked it)
- We need a **model** for predictions and what-if analysis; data for validation and tuning only

Node states in the SBFC model

- Susceptible
- Believer
- Fact-Checker



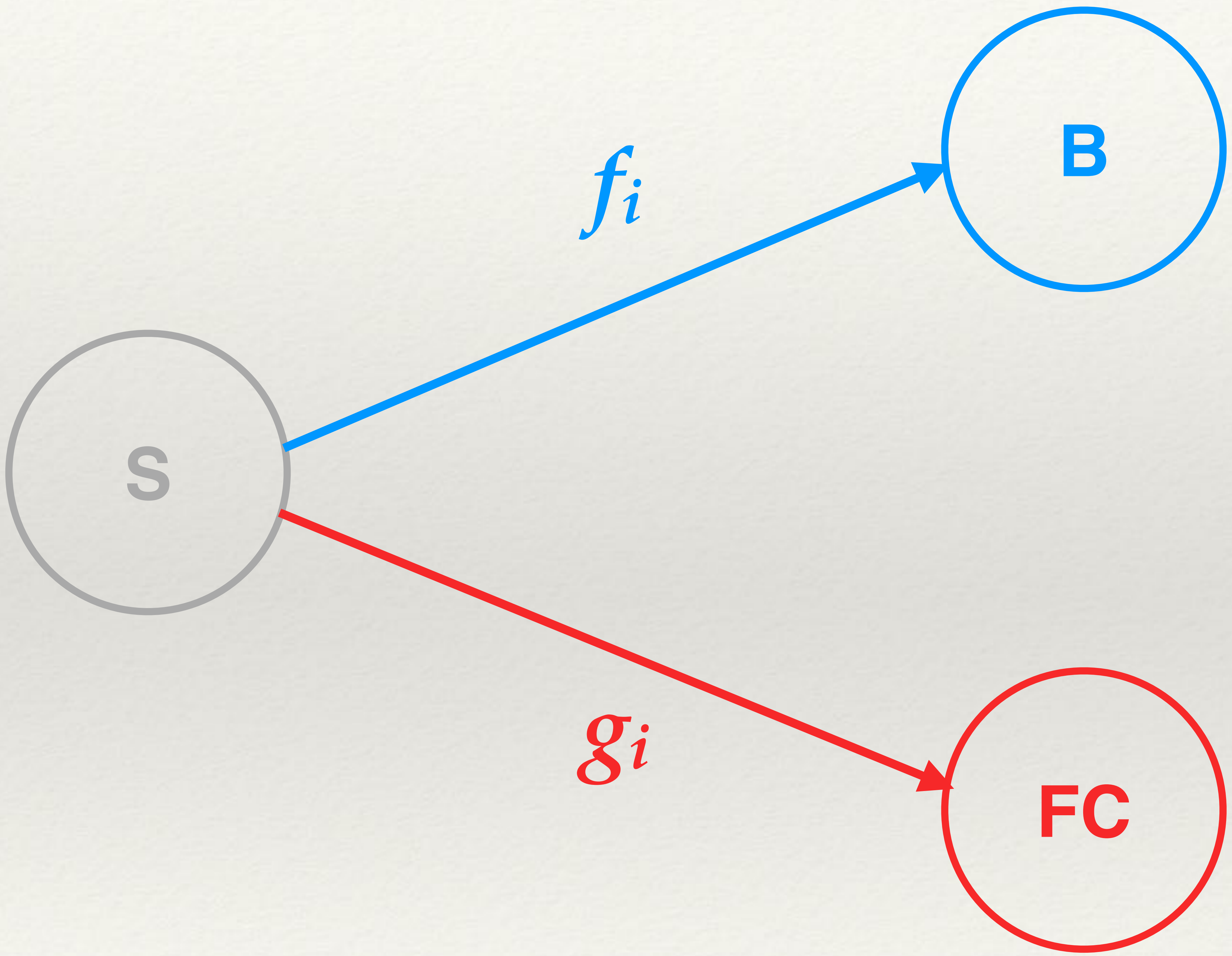


neighbors of i: $\boldsymbol{n_i}$

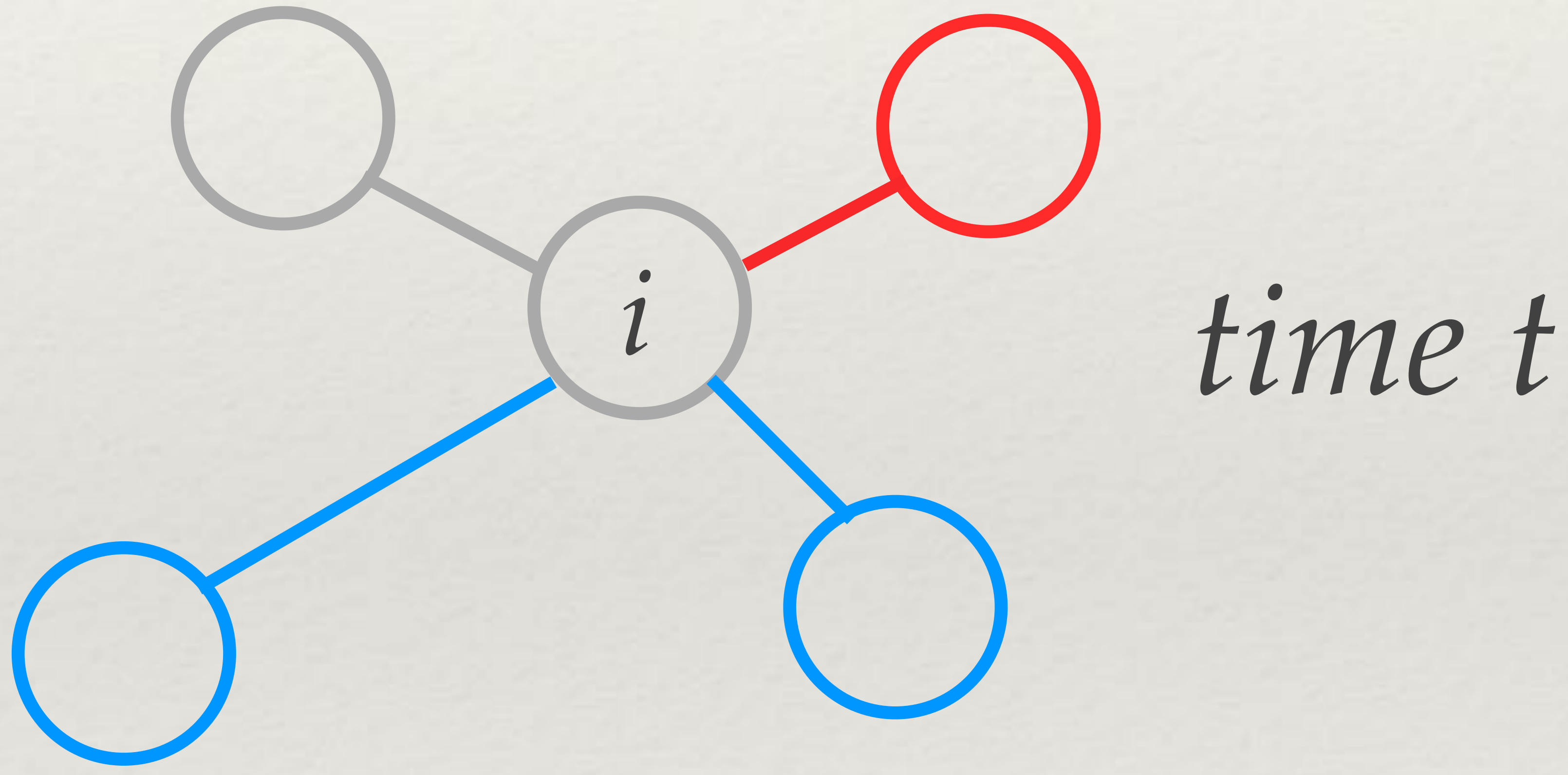
credibility of the hoax: $\boldsymbol{\alpha}$

spreading rate: $\boldsymbol{\beta}$

From Susceptible to Believer/Fact-Checker

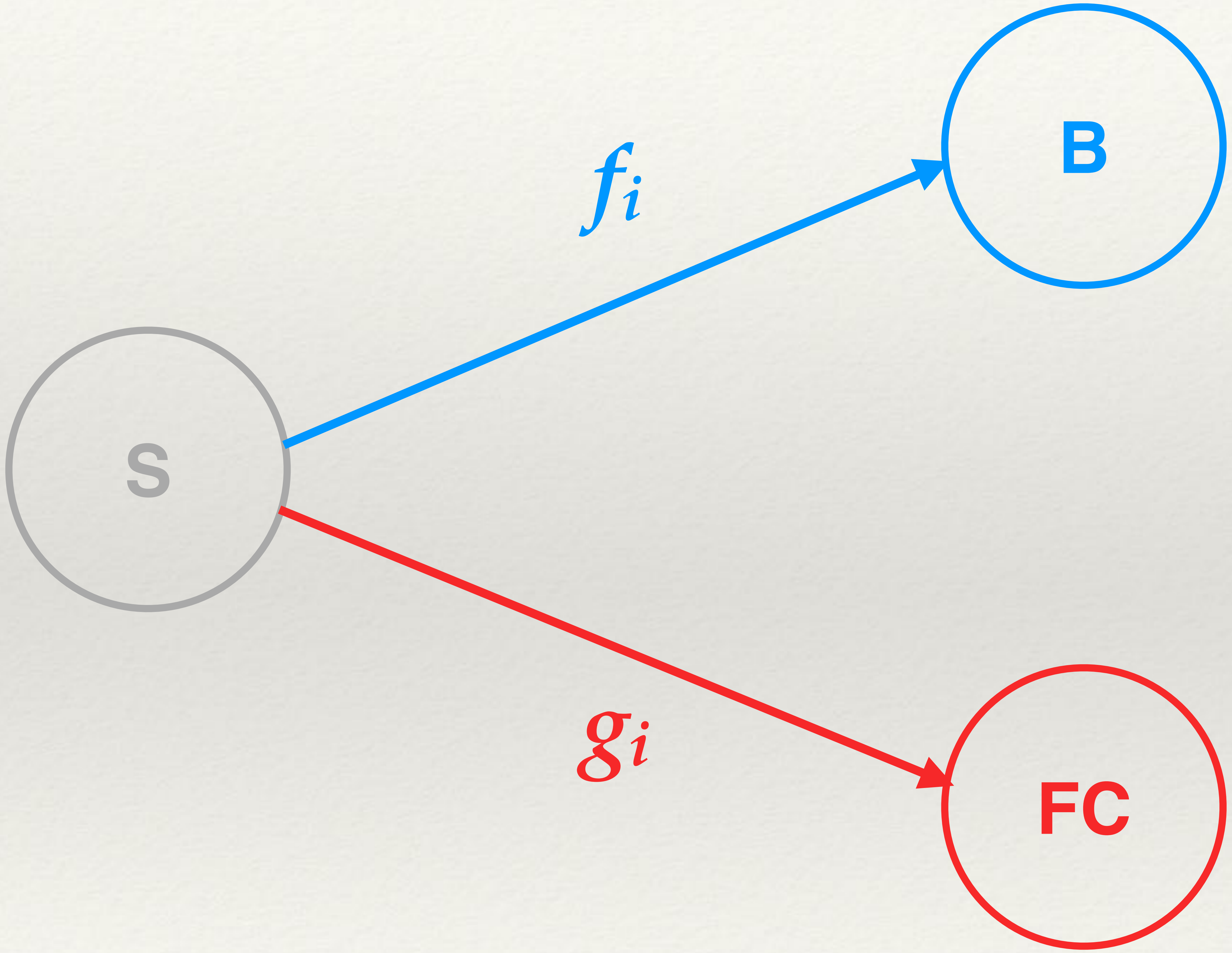


$$f_i(t) = \beta \frac{n_i^B(t)(1+\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$

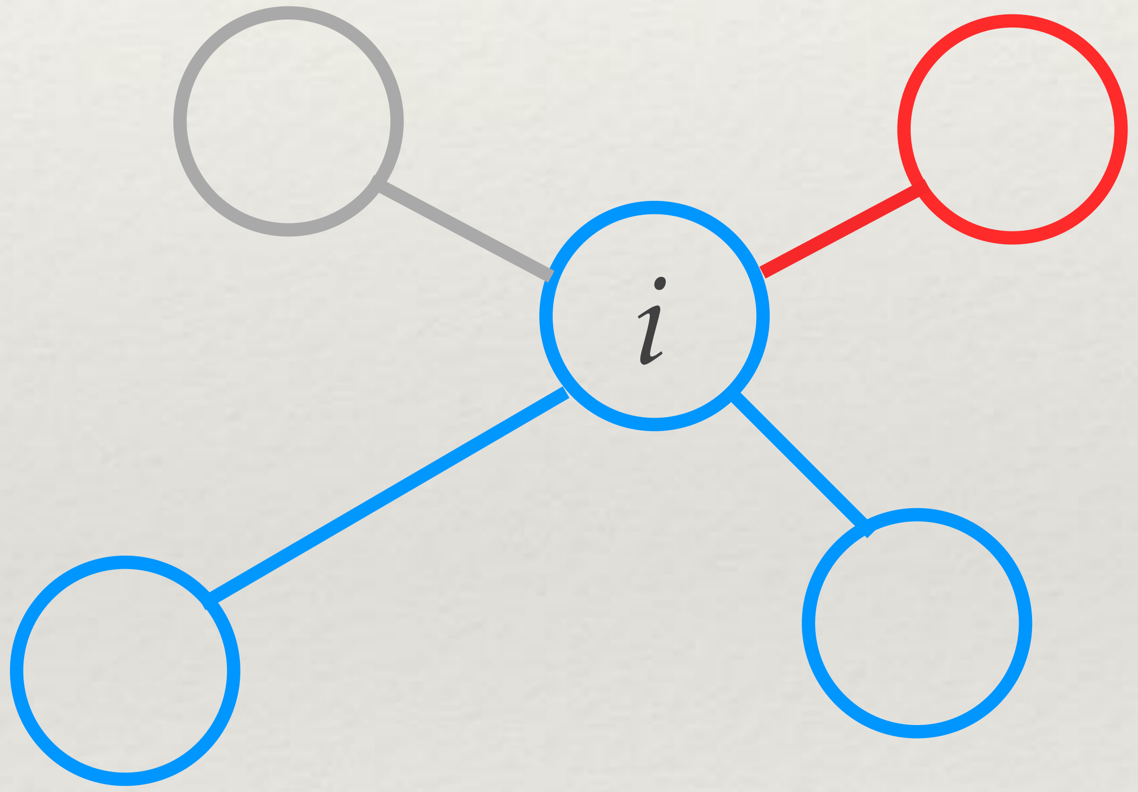


$$g_i(t) = \beta \frac{n_i^F(t)(1-\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$

From Susceptible to Believer/Fact-Checker



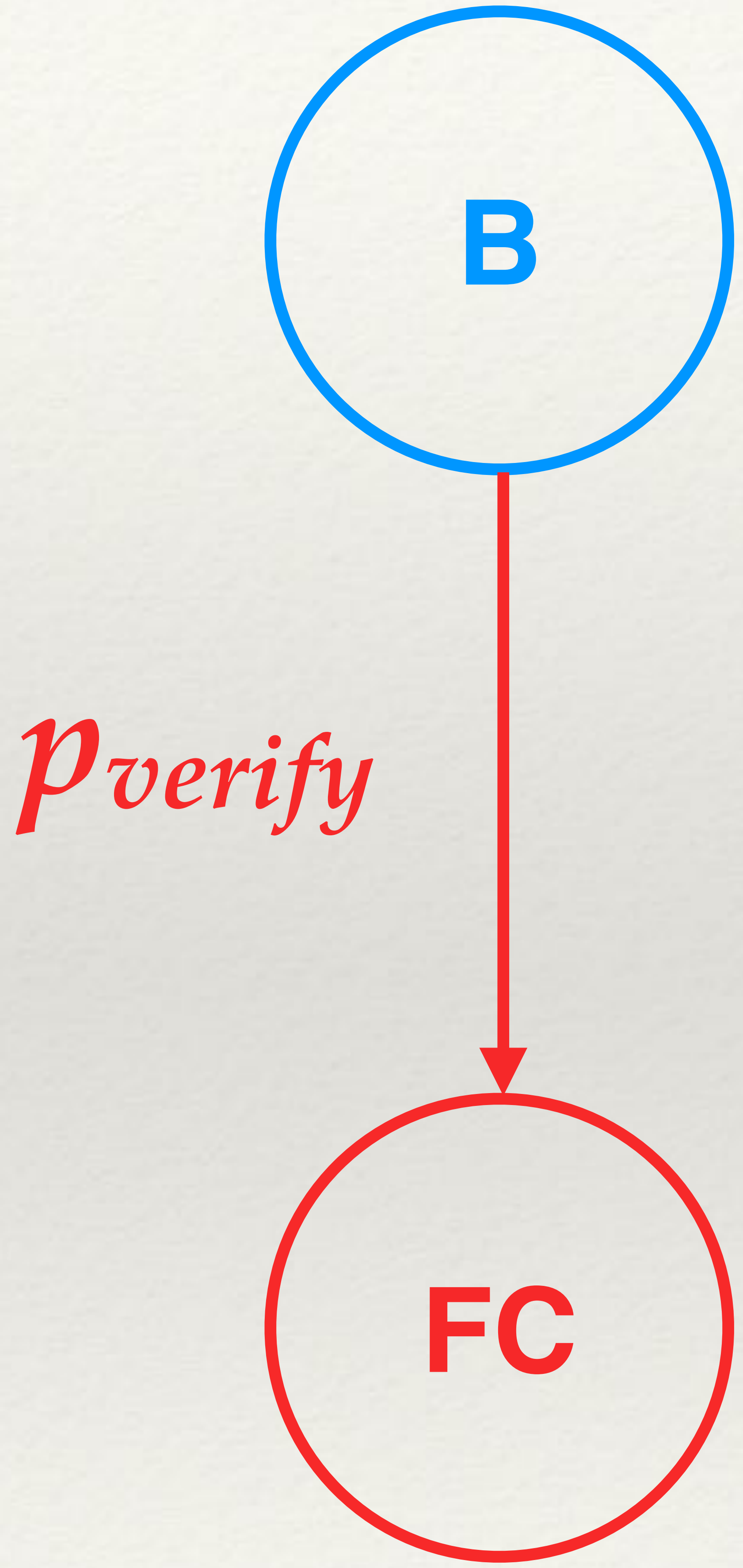
$$f_i(t) = \beta \frac{n_i^B(t)(1+\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$



time t+1

$$g_i(t) = \beta \frac{n_i^F(t)(1-\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$

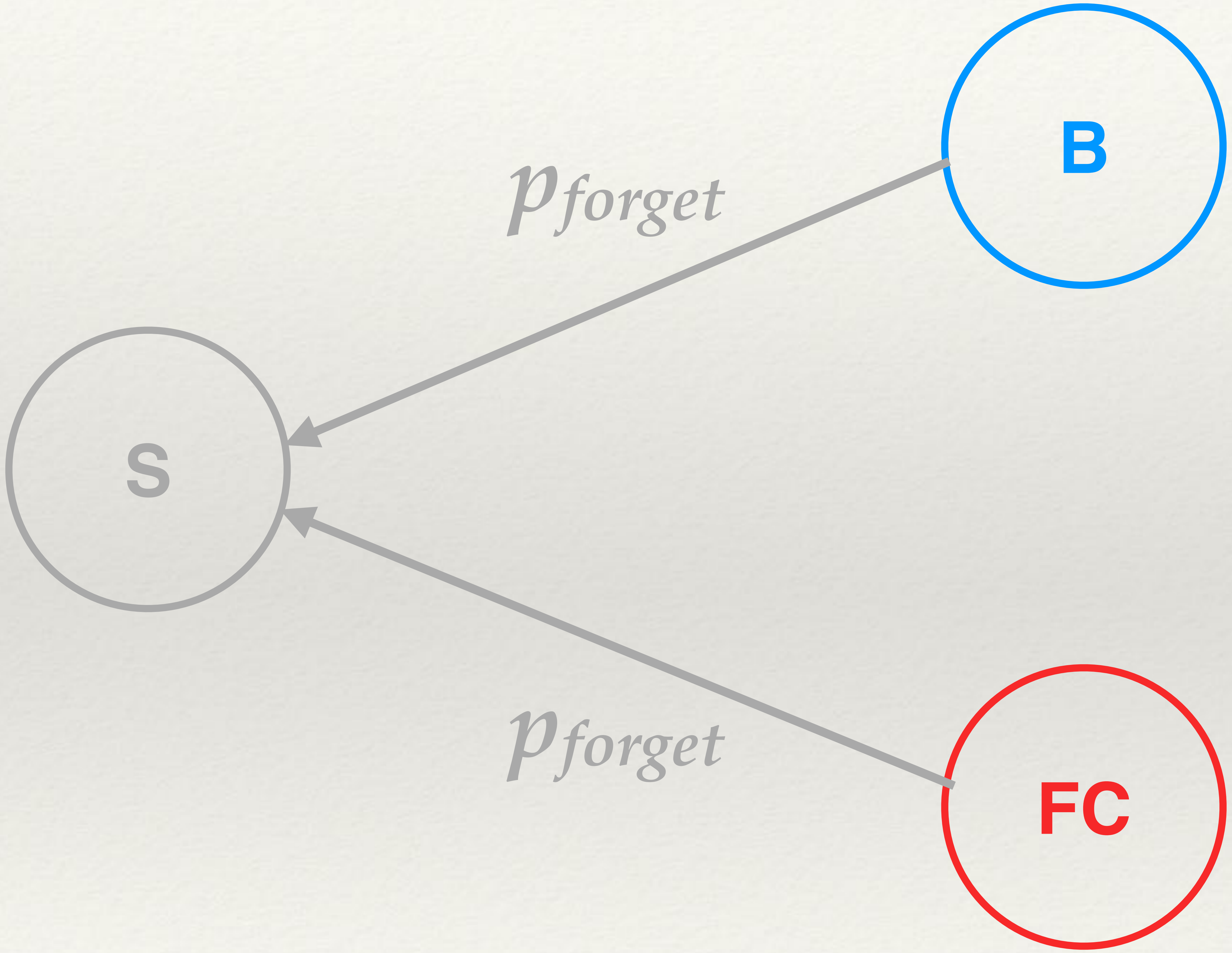
From Believer to Fact-Checker



VERIFYING

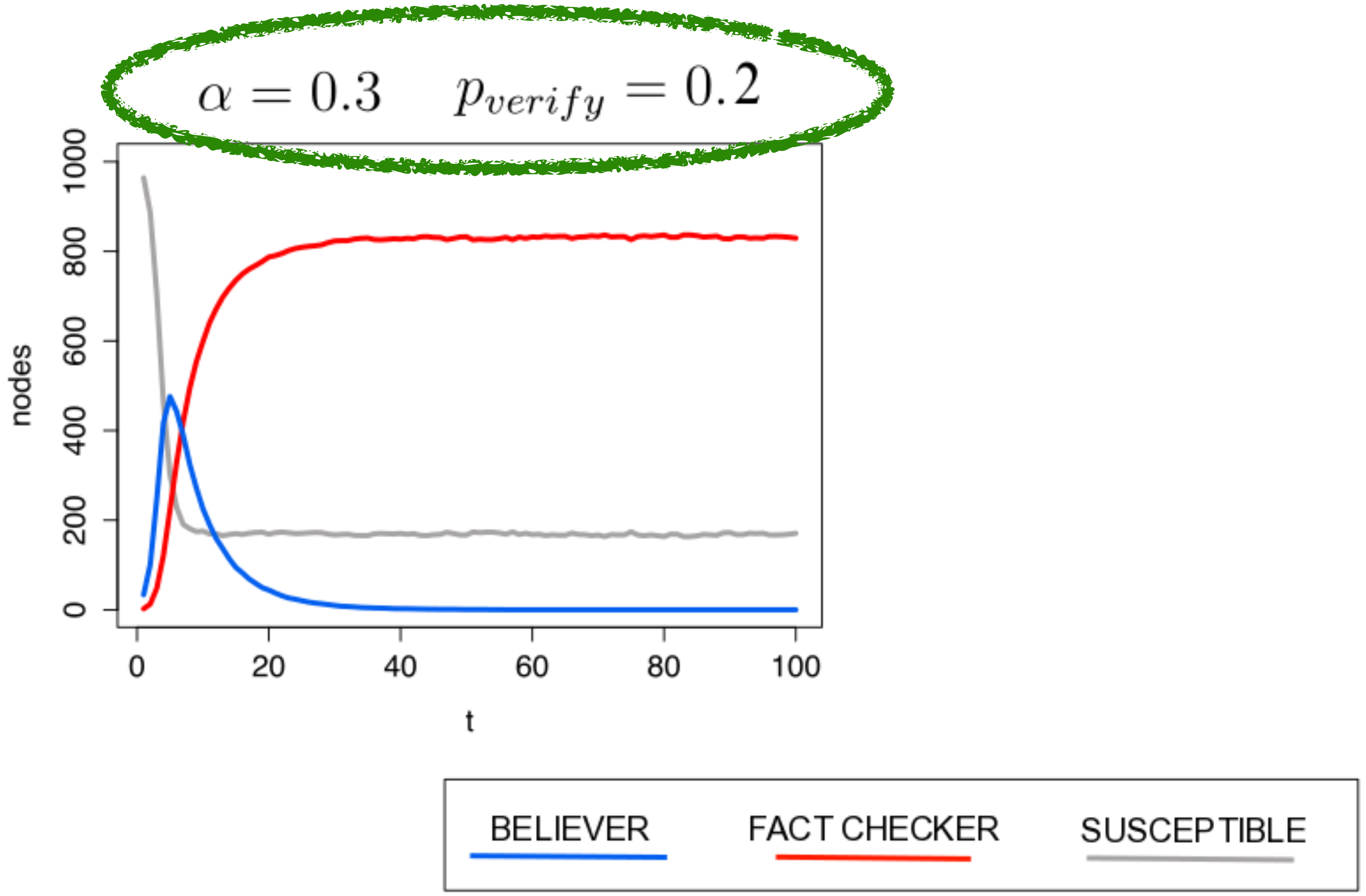
probability of fact-checking (or just deciding not to believe)

From Believer/Fact-Checker to Susceptible

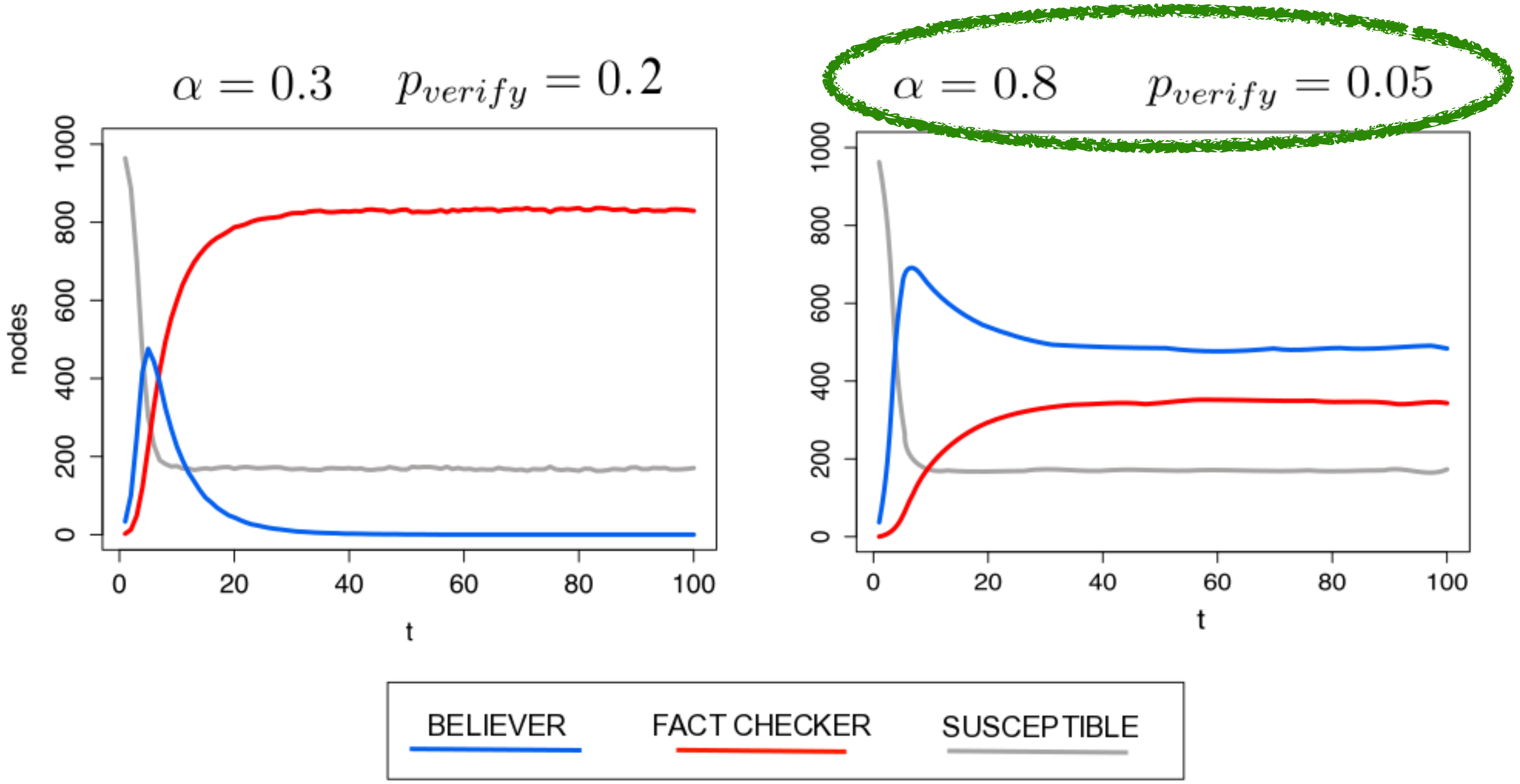


FORGETTING

Dynamics (agent-based simulations)



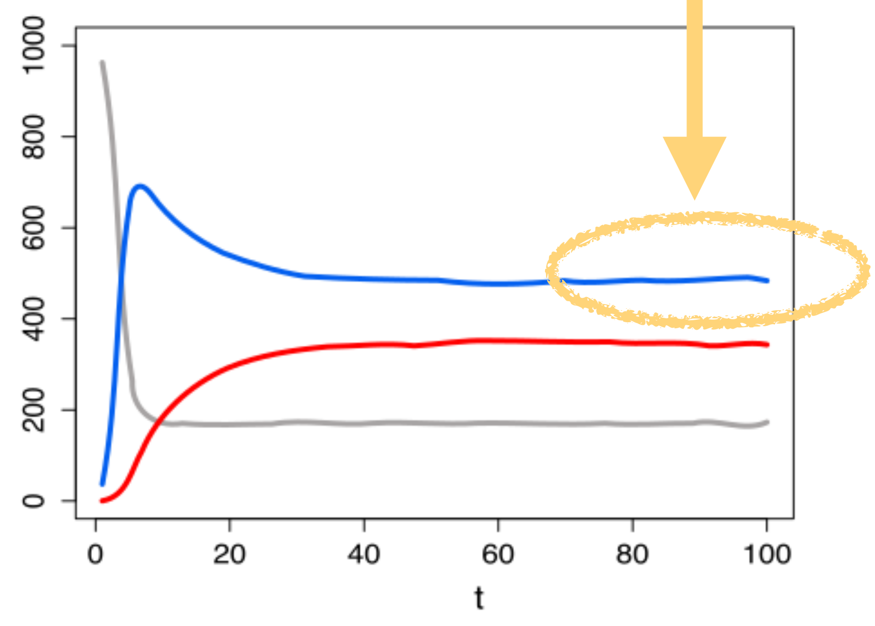
Dynamics (agent-based simulations)

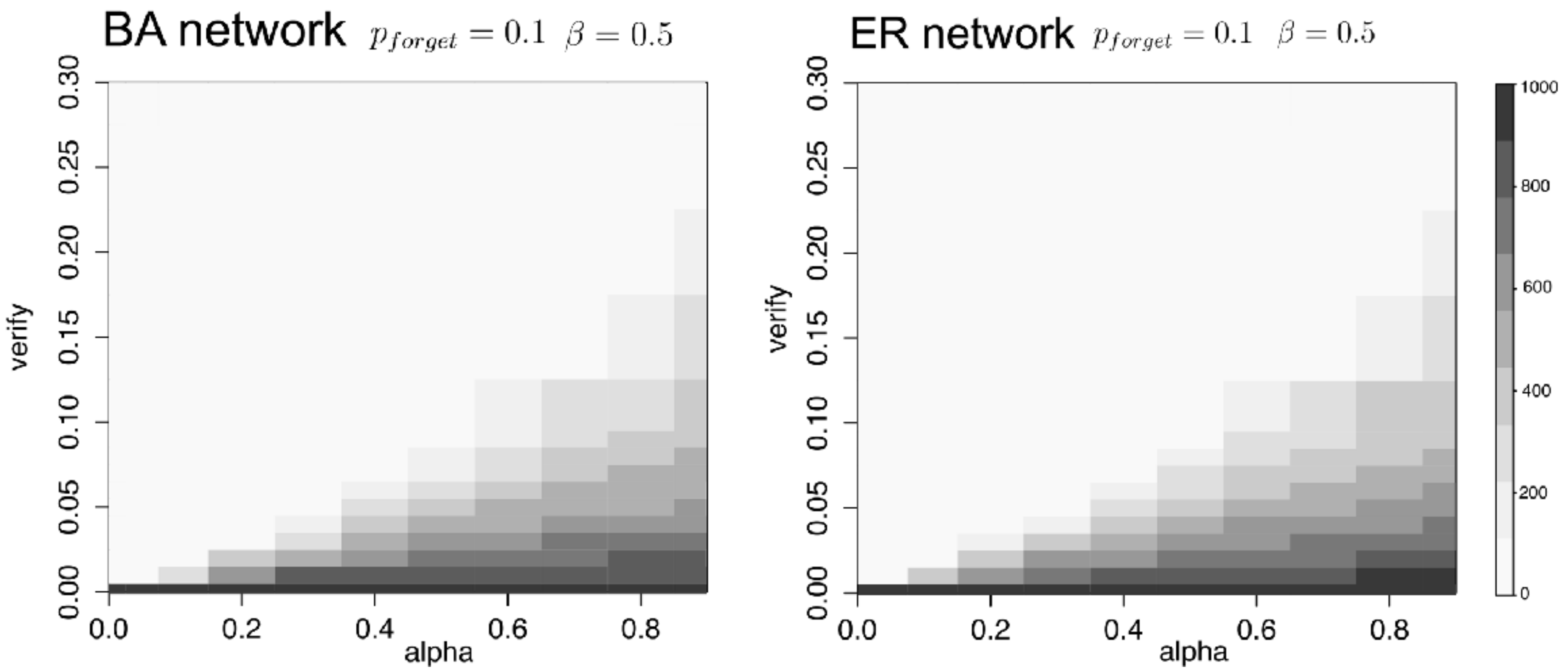


hoax **credibility** and **fact-checking probability** rule hoax persistence in the network

Dynamics (agent-based simulations)

number of 'believers' at the equilibrium





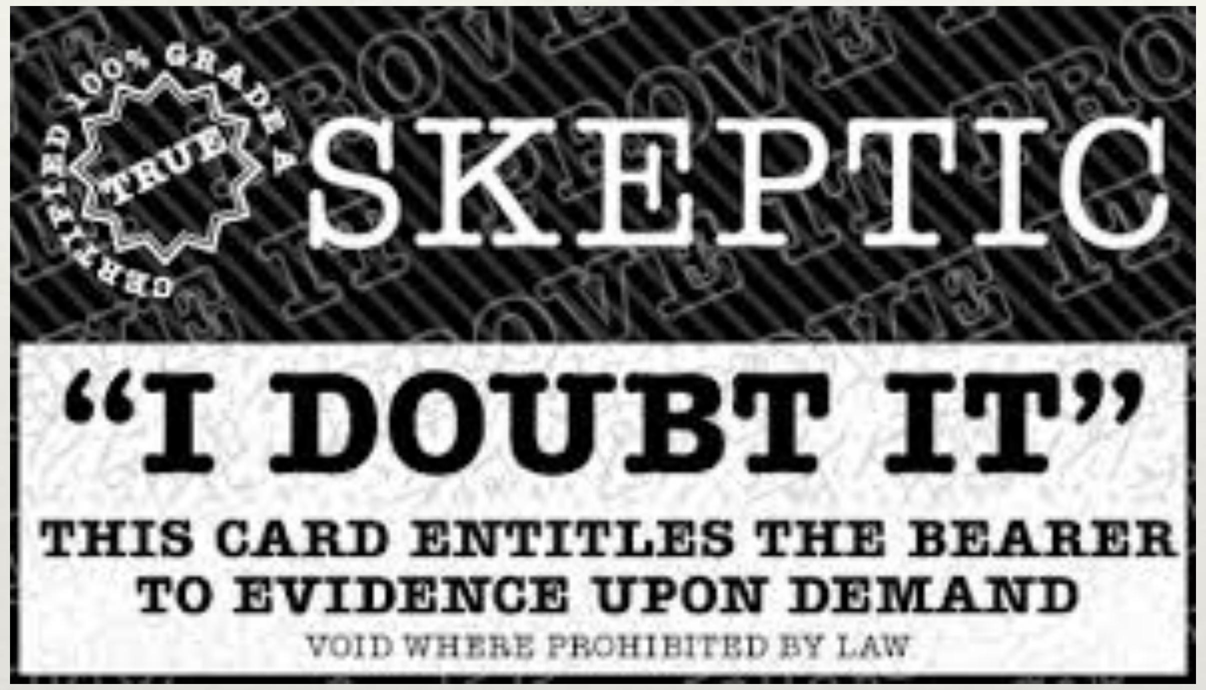
First step toward "good practices" understanding

threshold on verifying probability: our model provides an idea of how many believers we need to convince to guarantee the removal of the hoax

M Tambuscio, G Ruffo, A Flammini, and F Menczer. 2015. Fact-checking Effect on Viral Hoaxes: A Model of Misinformation Spread in Social Networks. In Proc. of the 24th Int. Conf. on World Wide Web (WWW '15 Companion)

The role of segregation

Skeptical and gullible agents



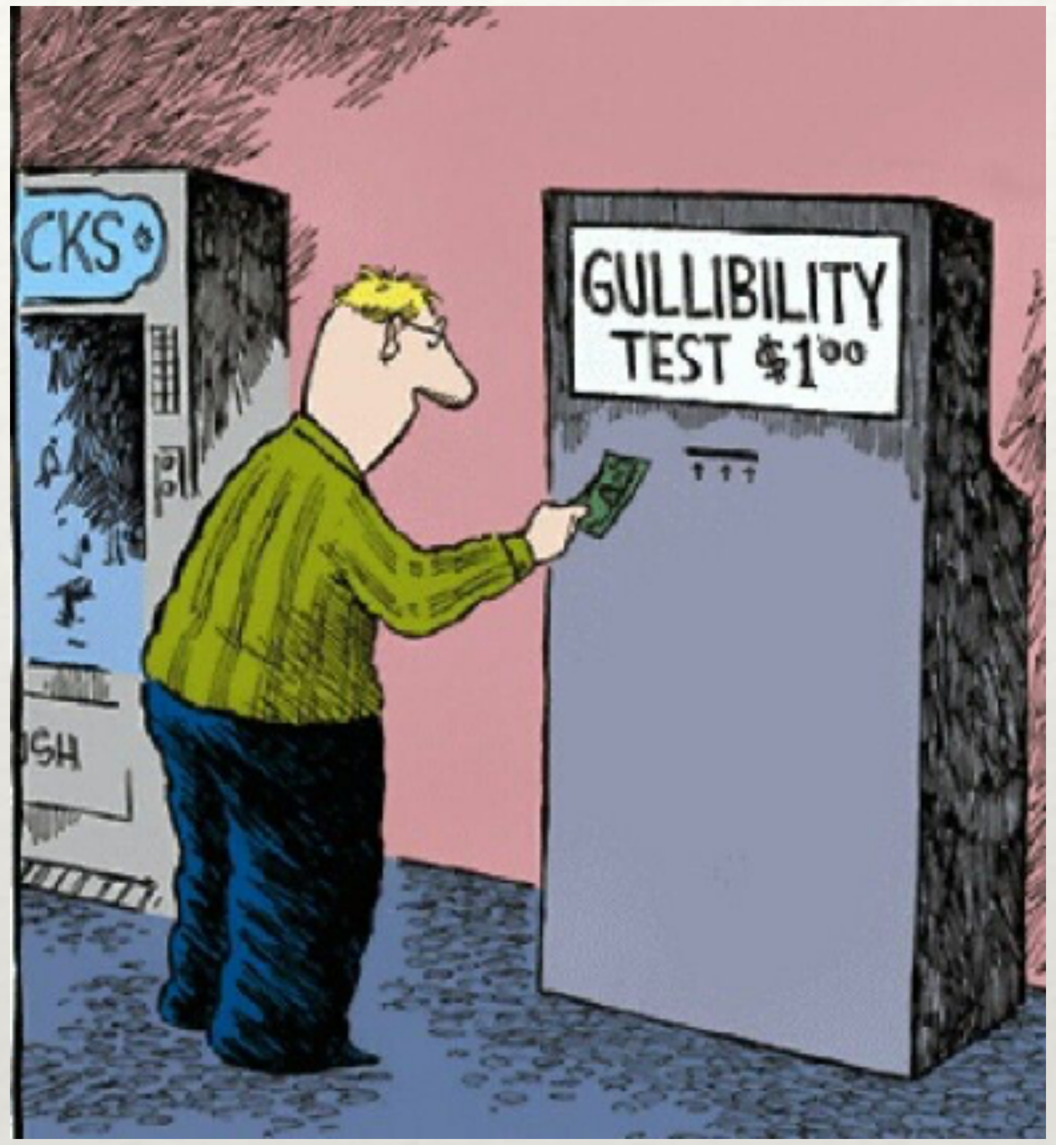
let's tune **credibility** accordingly

α

less credible — *more credible*

0 ⟷ 1

*more **skeptical*** — *more **gullible***



the propensity to believe is also a property of the node (**gullibility**)

What does it happen when skeptics and gullible agents are segregated?

Modeling two segregated communities

Skeptic

α small

size $(0 < \gamma < N)$

nodes in the gullible community

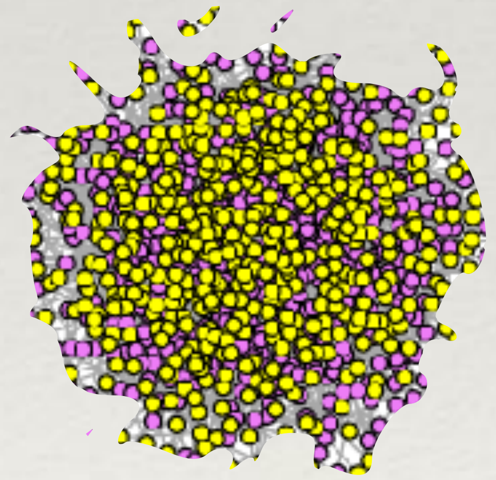
segregation (0.5 < **s** < 1)

fraction of edges within same community

[Gu-Gu, Sk-Sk]

Gullible

α large

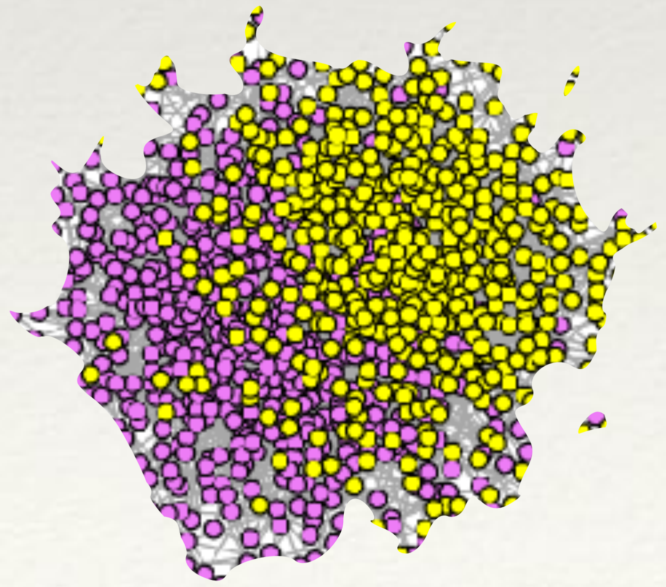


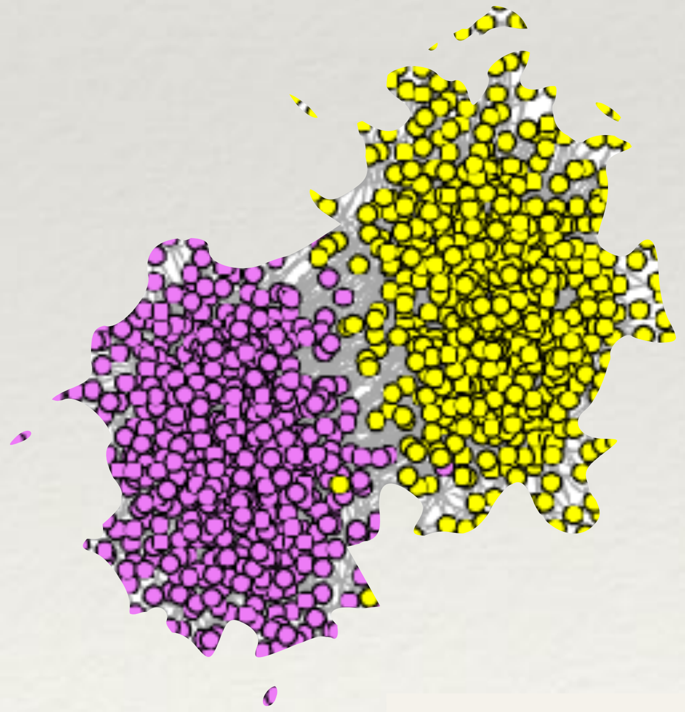
s=0.55

γ=500

s=0.8

γ=500

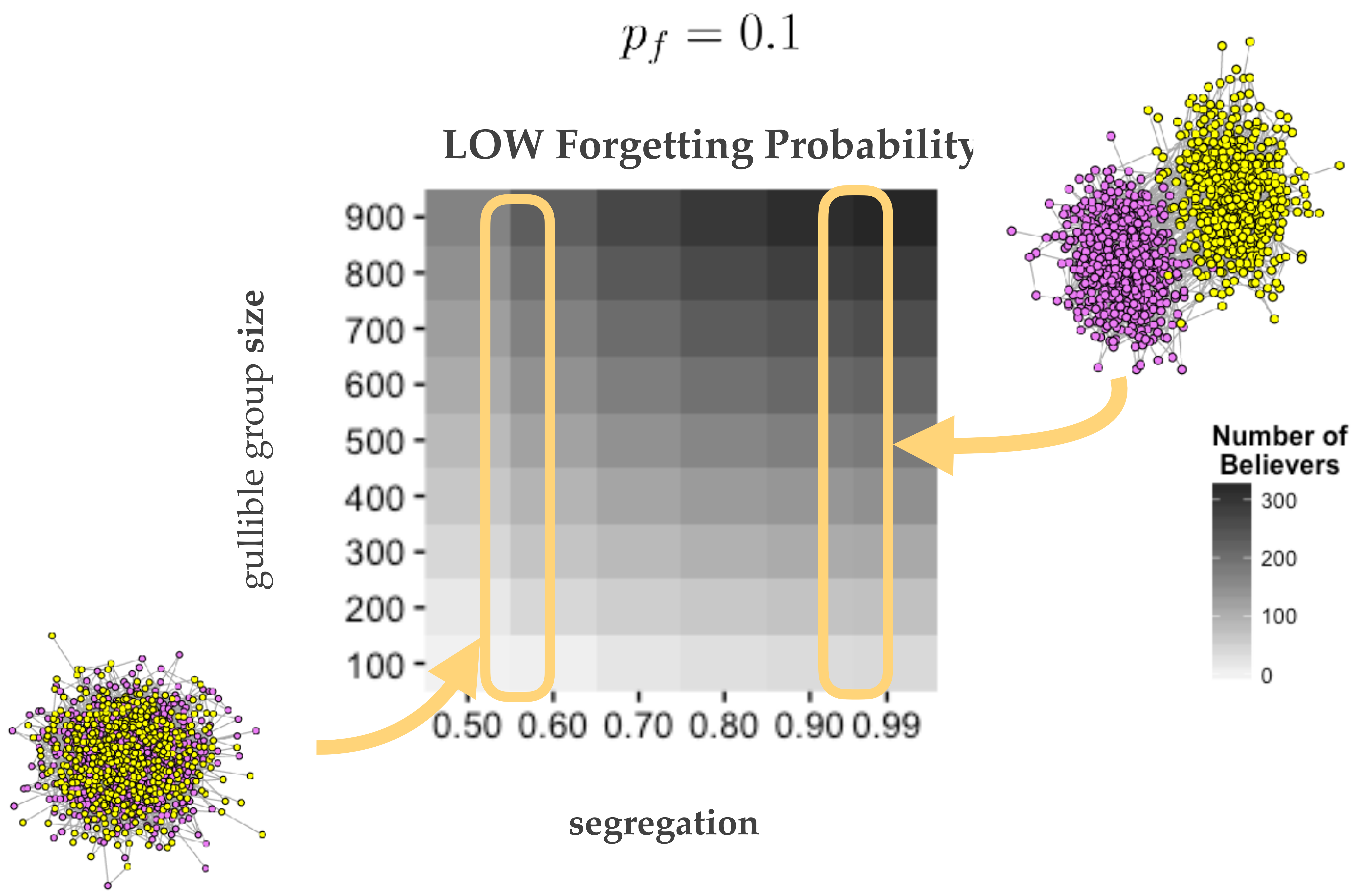




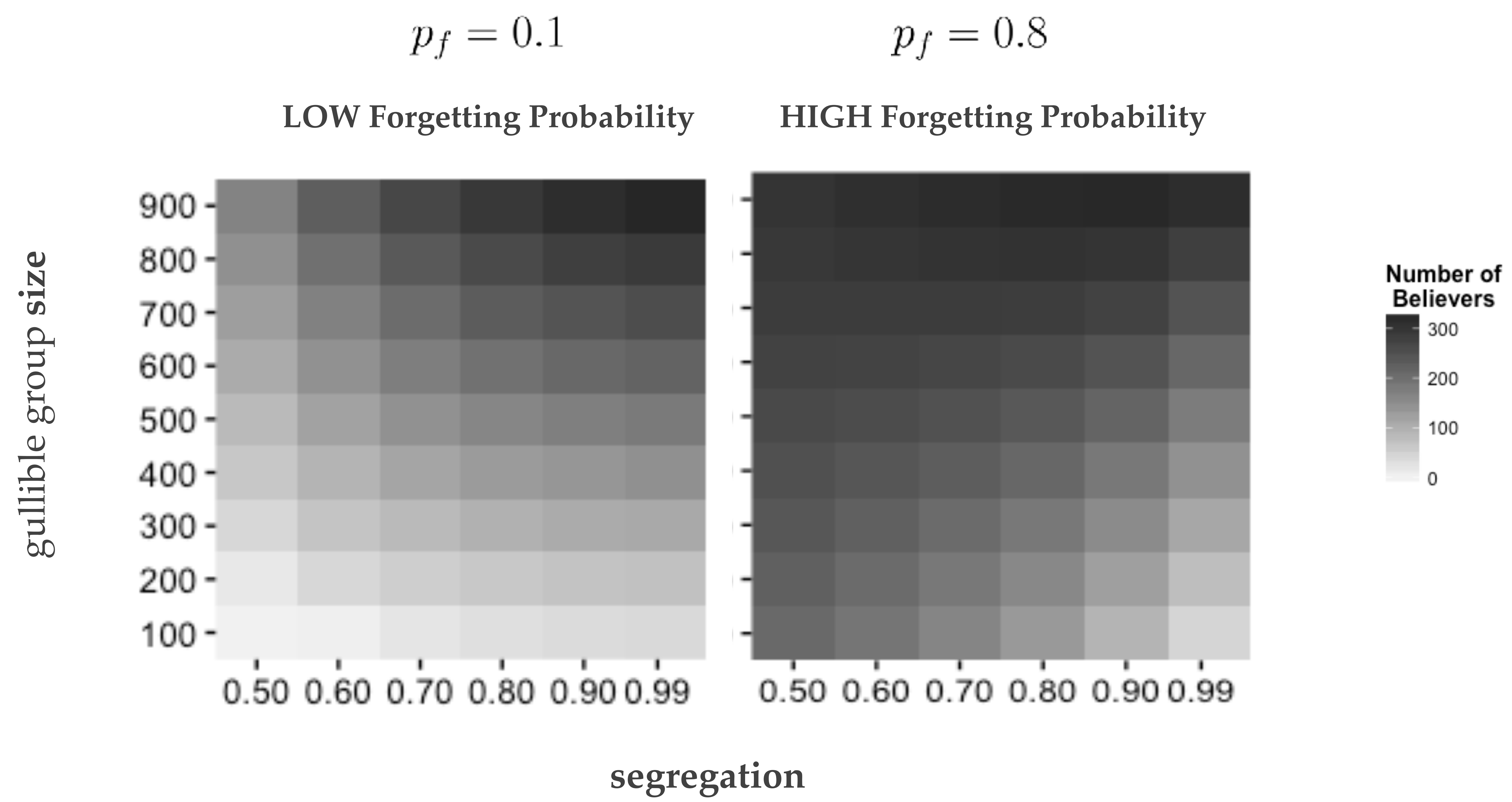
s=0.95

γ=500

Size vs segregation



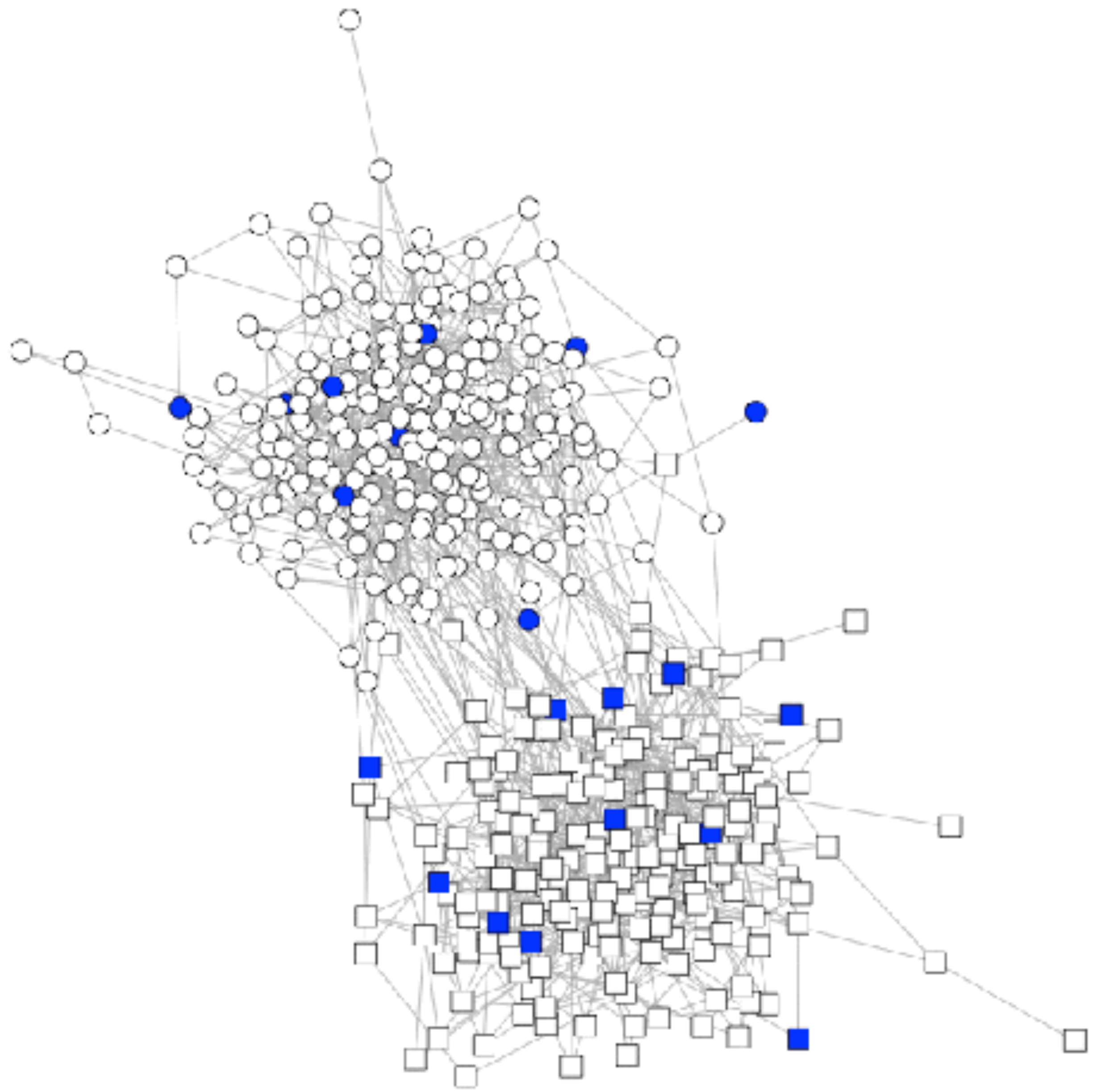
Size vs segregation



Role of forgetting

LOW Forgetting Rate

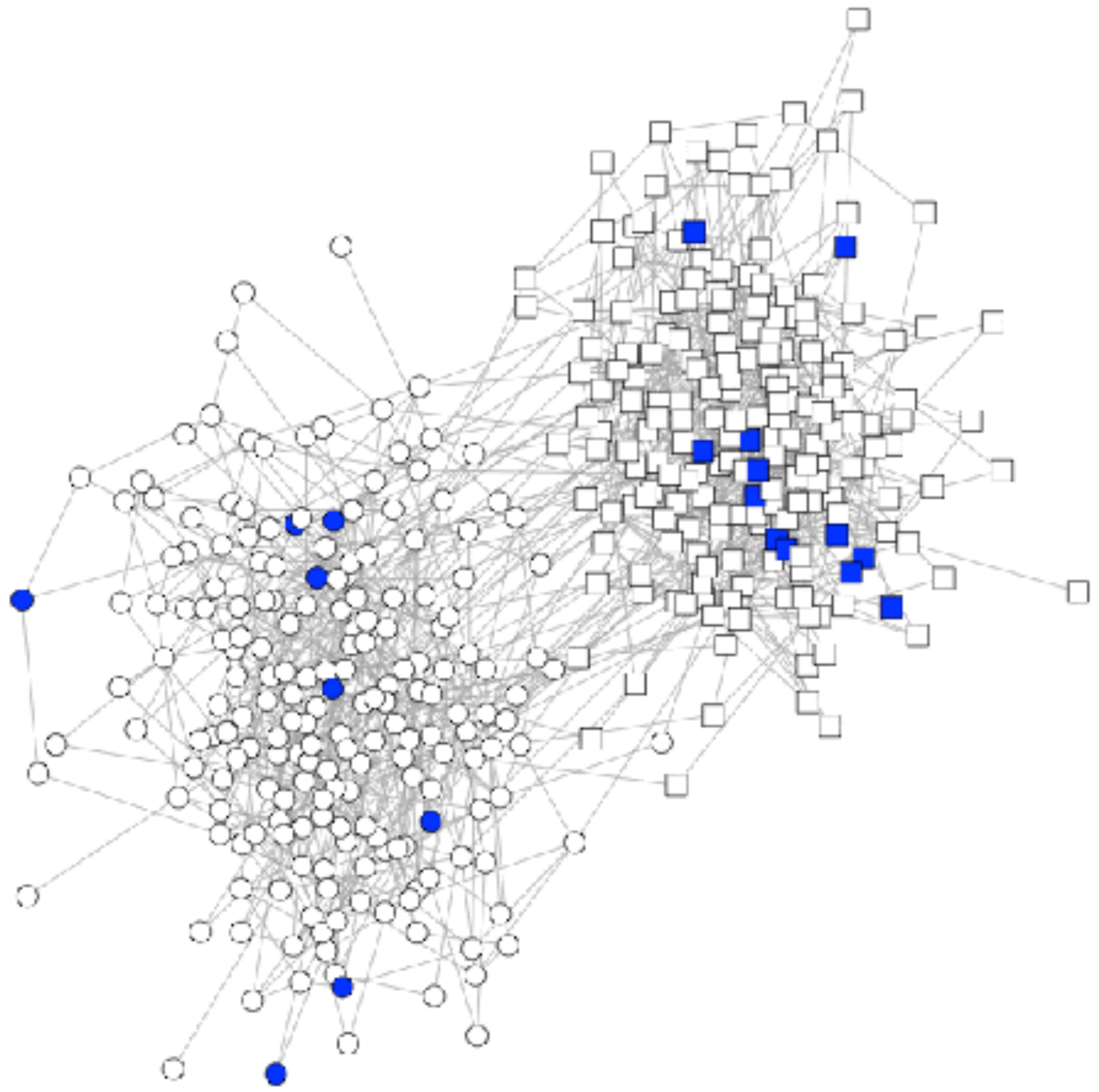
$p_f = 0.1$



Time = 1

HIGH Forgetting Rate

$p_f = 0.8$



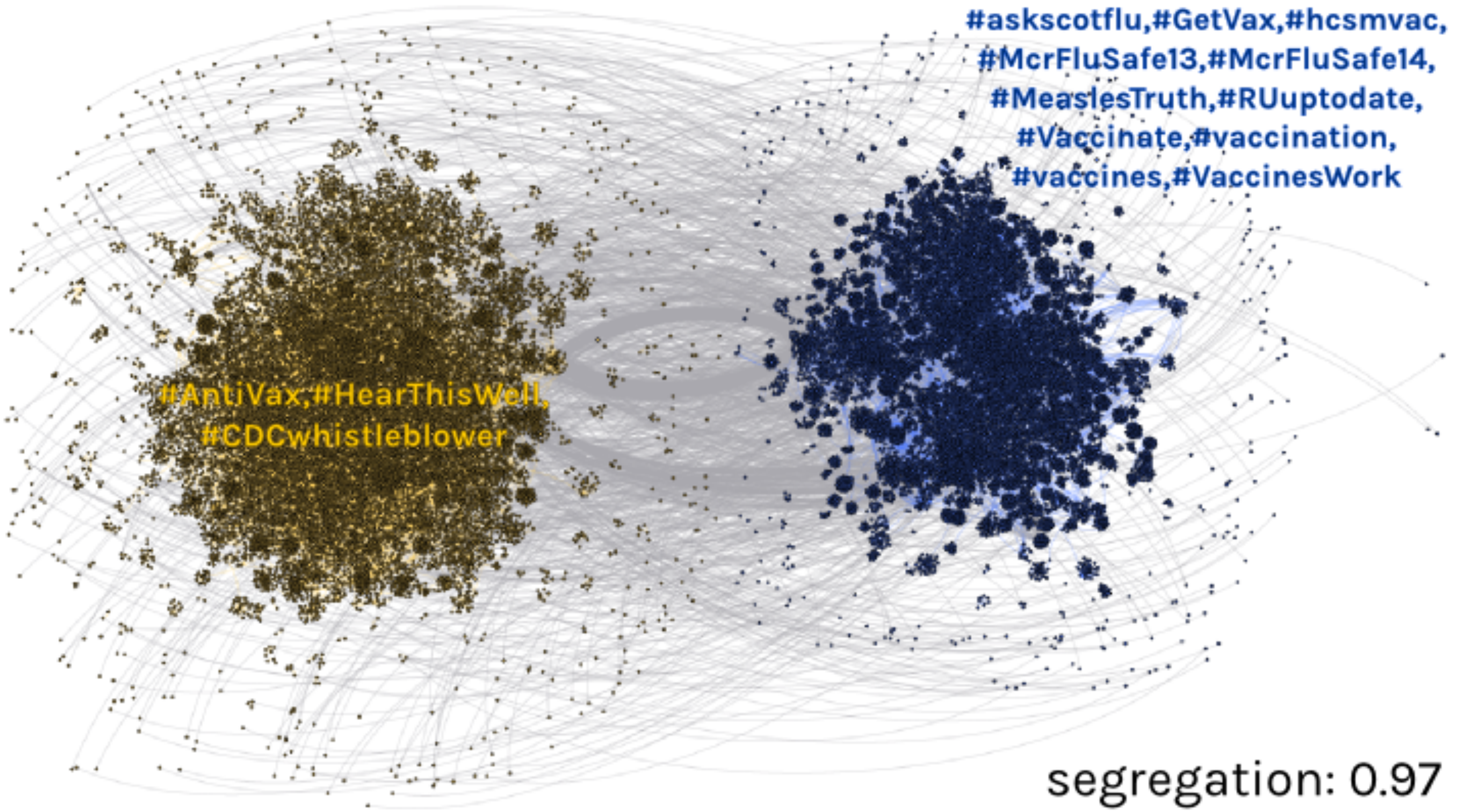
Time = 1

Lessons learned and observations

- We can use our model to study the fake-news diffusion process in **segregated community**
- **Complex contagion** is observed: interplay and not trivial outcomes
- **Forgetting** probability becomes relevant as well as the **level of segregation:**
  - **high forgetting** probability (e.g., just `normal' unfounded gossip) vanishes soon in **segregated communities**
  - **low forgetting** probability (e.g., conspiracy theories or partisanship beliefs) requires **low segregation**

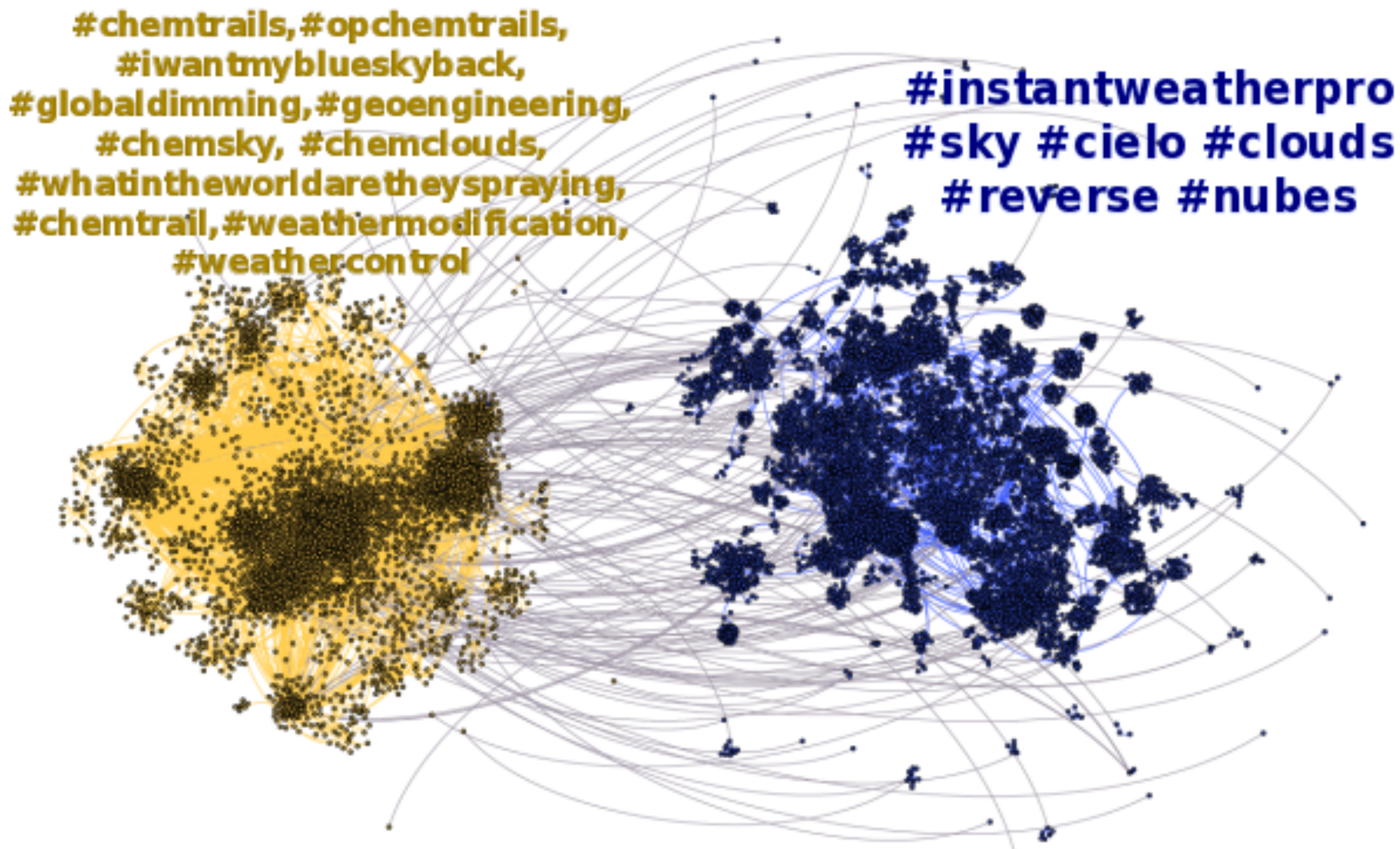
M Tambuscio, D F M Oliveira, G L Ciampaglia, G Ruffo, Network segregation in a model of misinformation and fact-checking, Journal of Computational Social Science (2018) 1: 261.

real data: vaccines



twitter data from IU https://osome.iuni.iu.edu

real data: chemtrails



twitter data from IU https://osome.iuni.iu.edu

segregation: 0.99

Evaluating debunking strategies

What-if analysis


MARCELLA TAMBUSCIO


GIOVANNI LUIGI CIAMPAGLIA

- We live in a **segregated** society: let's accept it!
- Misinformation can survive in the network for a long time: **low forgetting** probability
- **Computational epidemiology**: immunization works better if some node in the network (e.g., hubs, bridges) is vaccinated first
- **Where** to place fact-checkers?
- Stronger hypothesis: a believer do not verify ($\mathbf{p_{verify} = 0}$)
  - they can still forget
  - we can accept to leave half of the population in their own (false) beliefs, but we want at least to protect the skeptics!

Basic settings with no verification

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

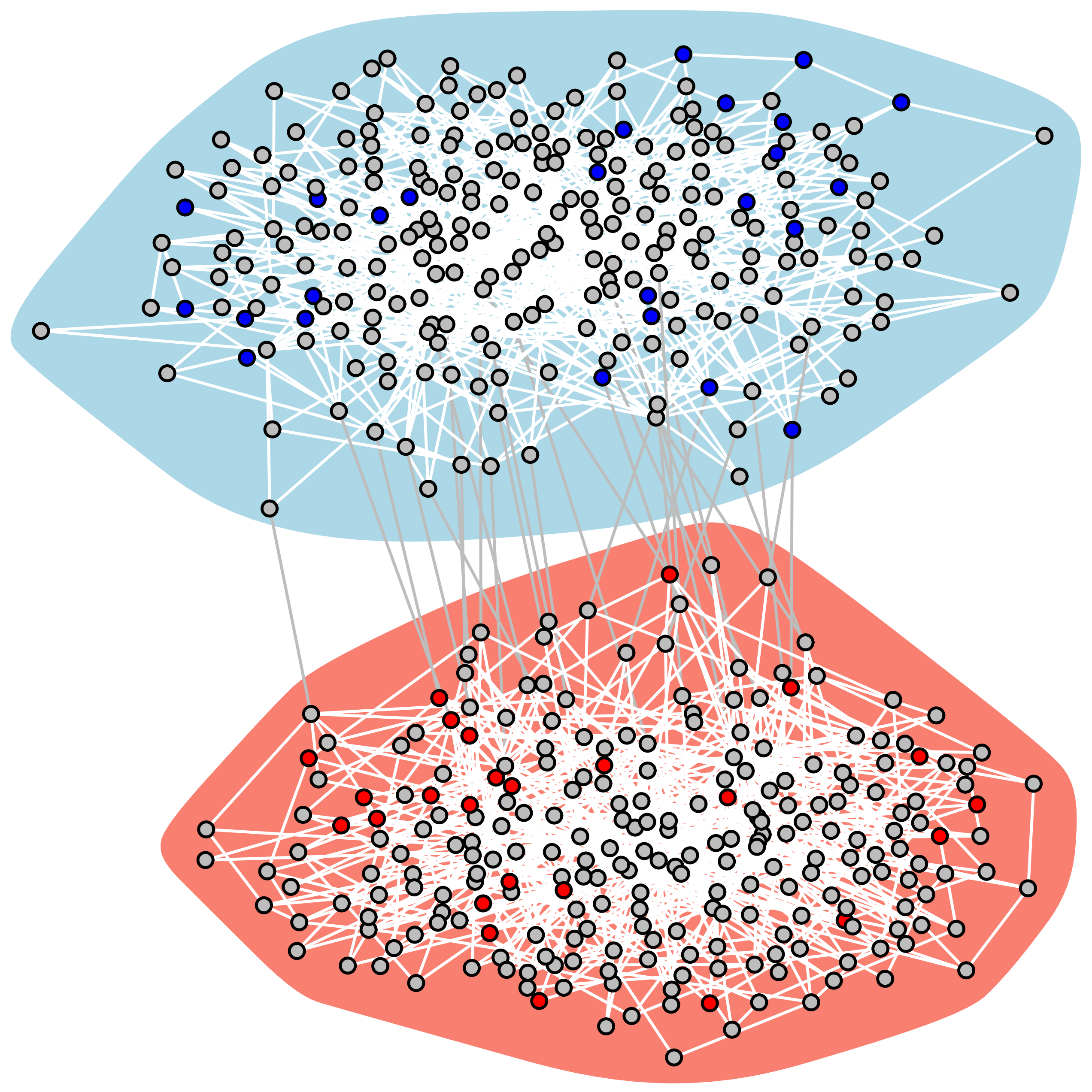
gullible group:

- α: 0.8
- seeders B: 10%

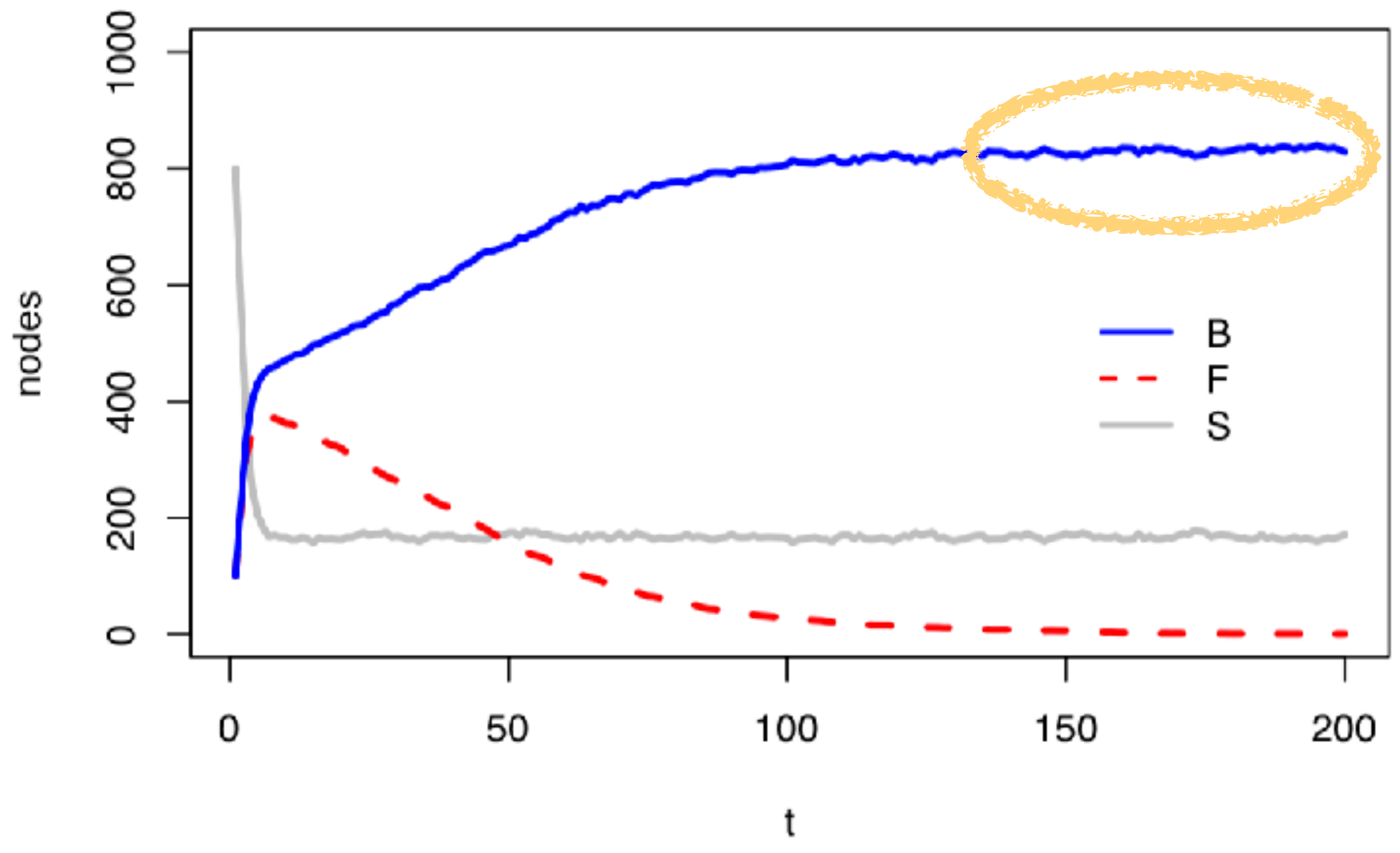
skeptical group:

- α: 0.3
- seeders FC: 10%

Simulation start



Simulation results



As expected: very **bad**!

Eternal fact-checkers placed at random

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

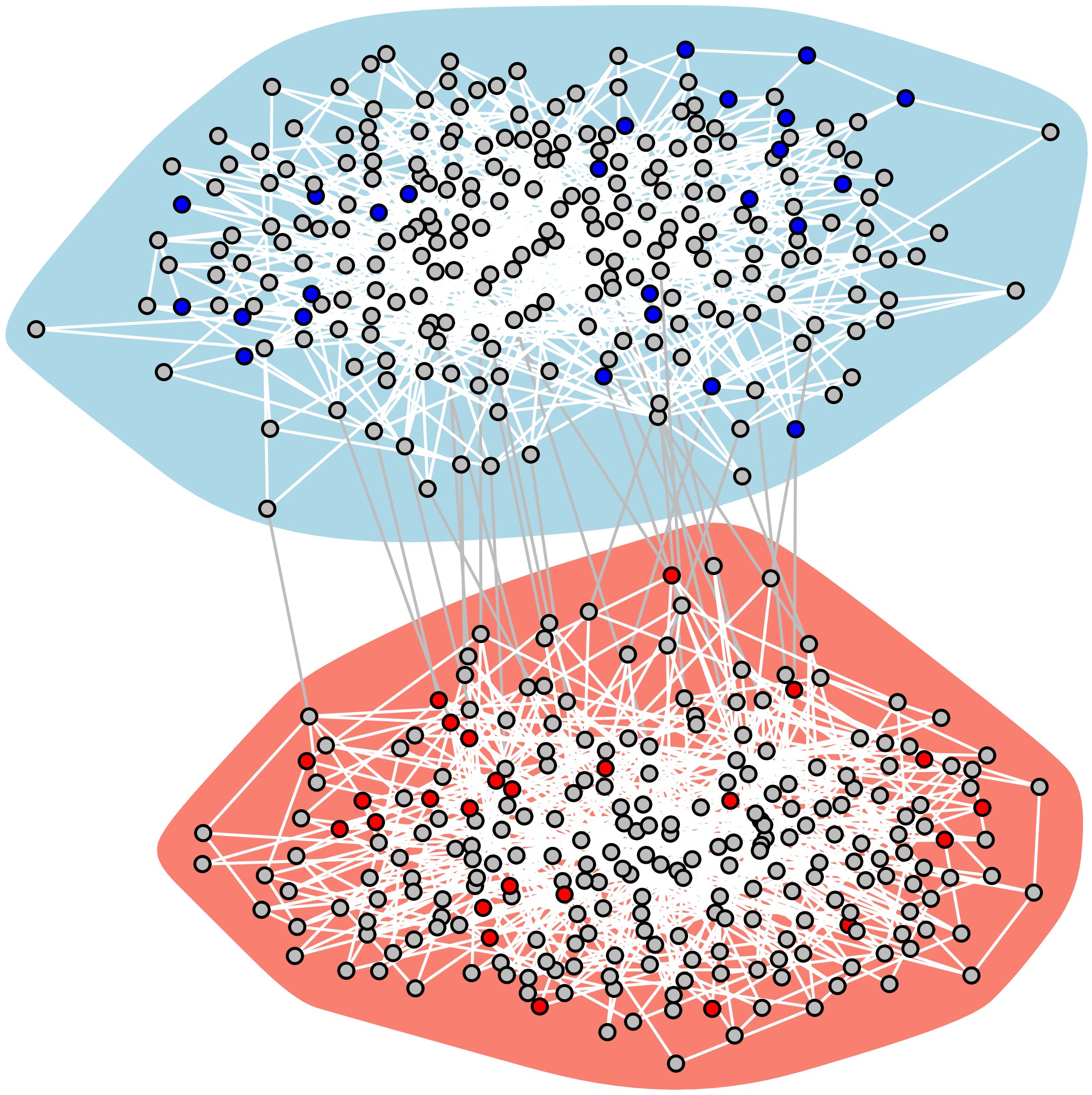
gullible group:

- α: 0.8
- seeders B: 10%

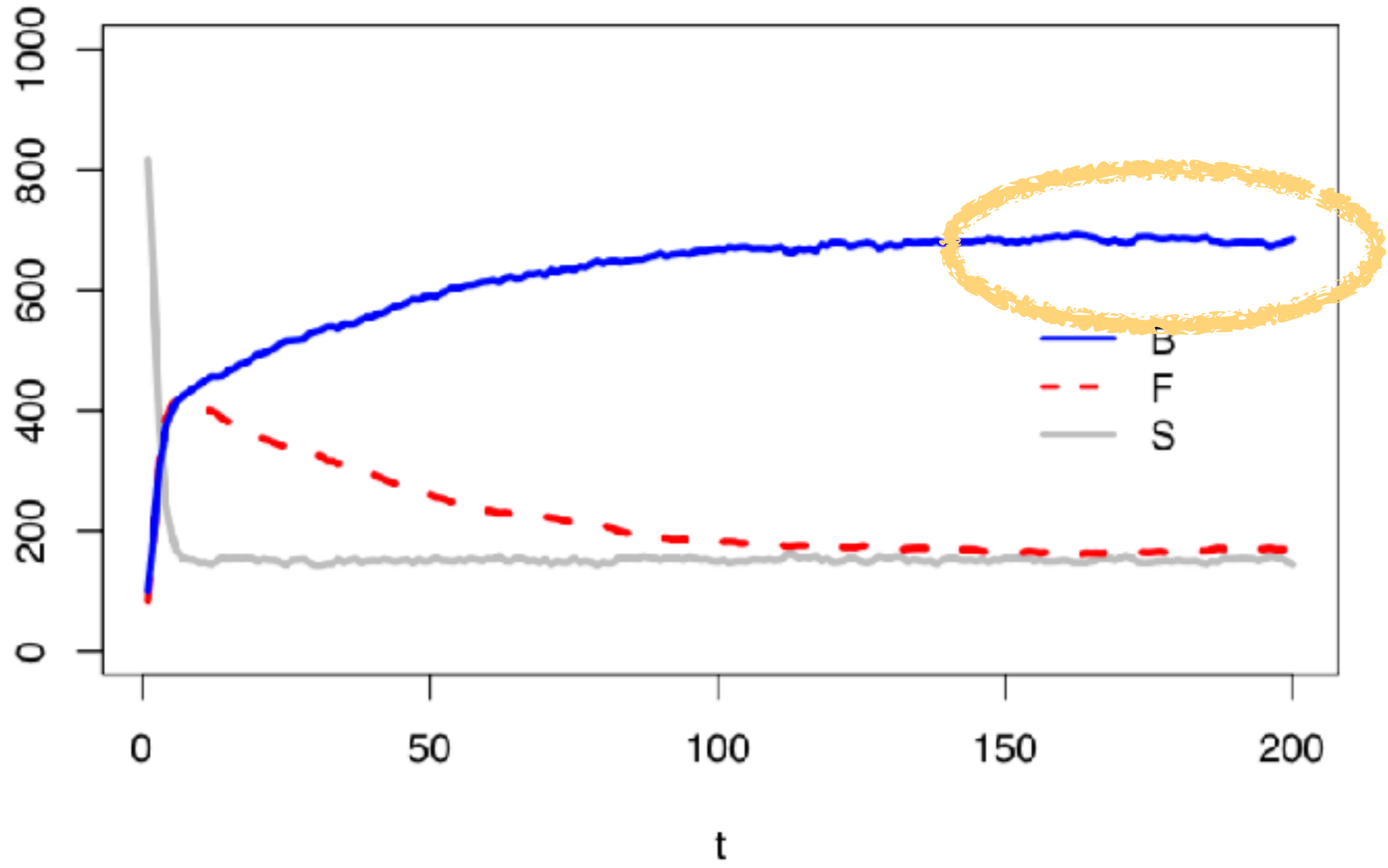
skeptical group:

- α: 0.3
- seeders FC: 10%
- **seeders are eFC**

Simulation start



Simulation results



better, but still…

Hubs as eternal fact-checkers

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

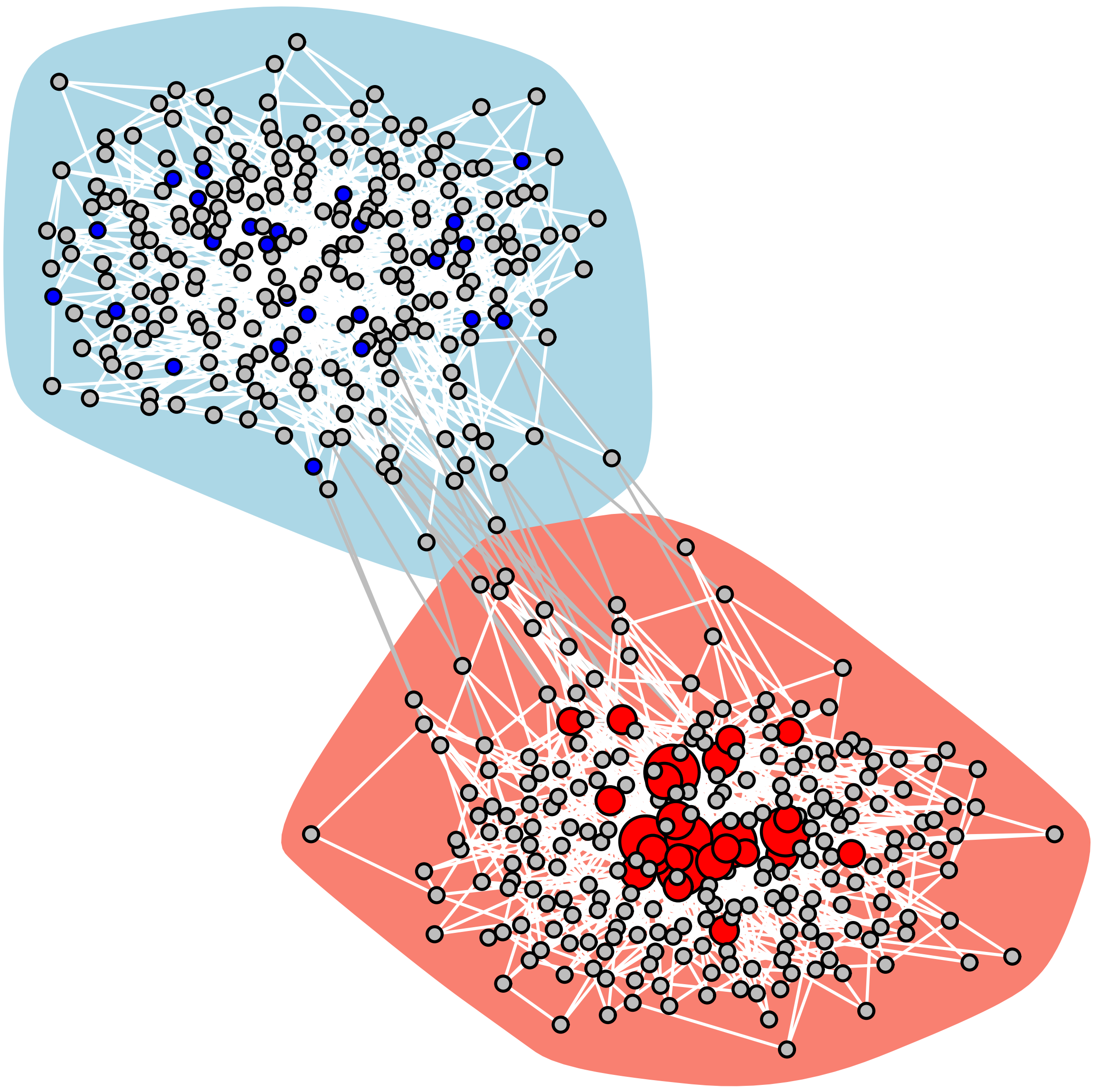
gullible group:

- α: 0.8
- seeders B: 10%

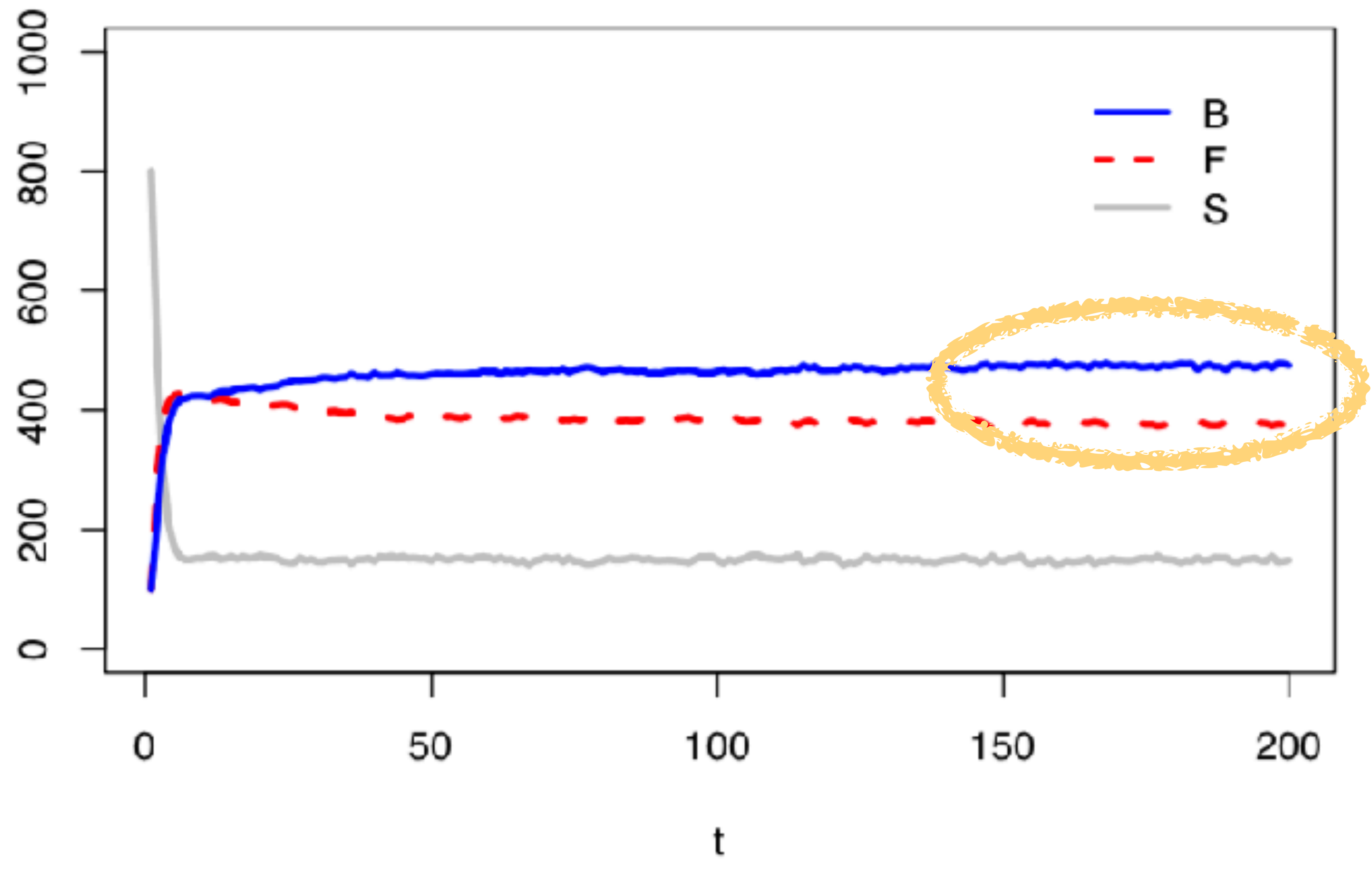
skeptical group:

- α: 0.3
- seeders FC: 10%
- **HUBS are eFC!**

Simulation start



Simulation results



better

Bridges as eternal fact-checker

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

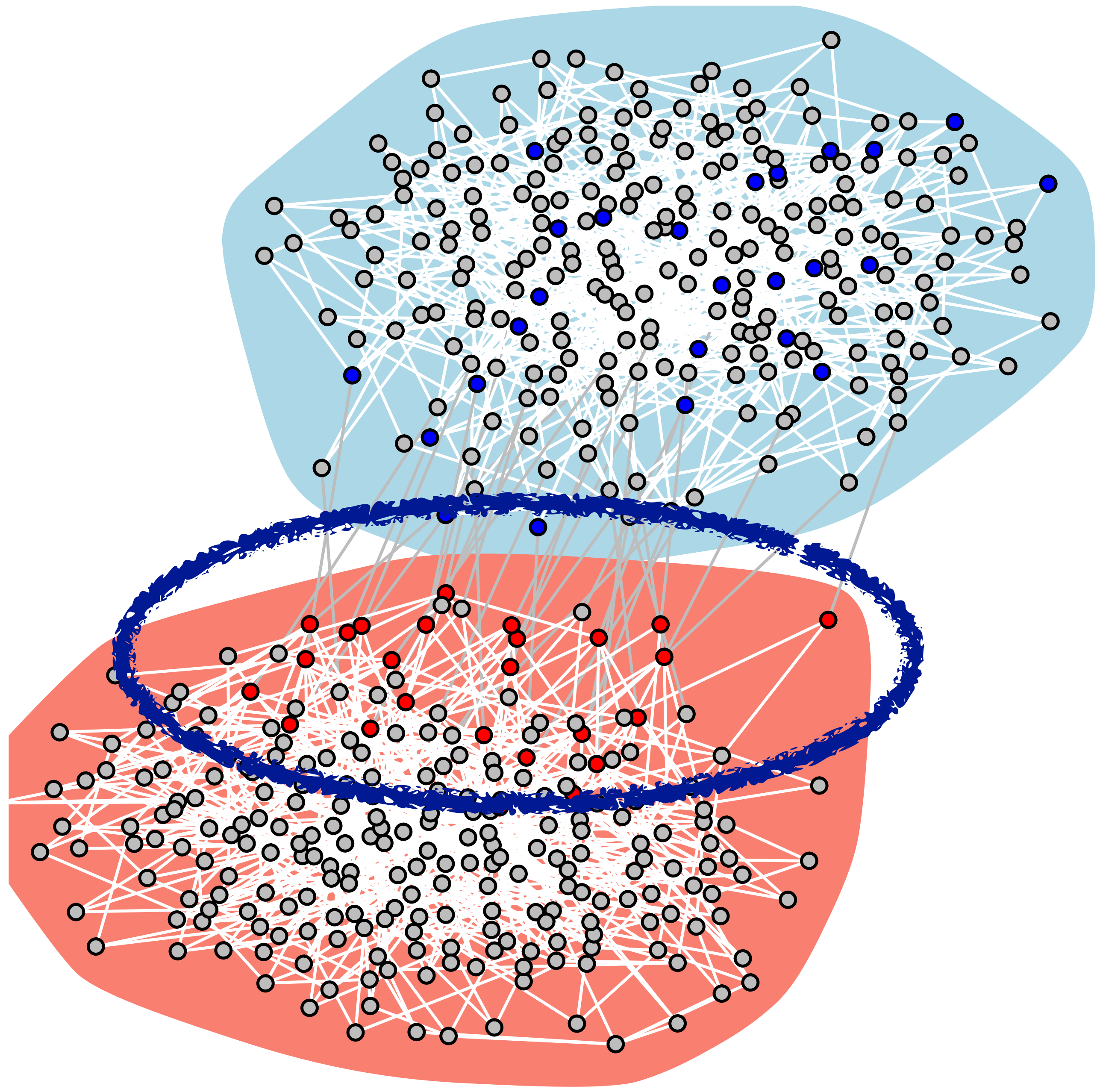
gullible group:

- α: 0.8
- seeders B: 10%

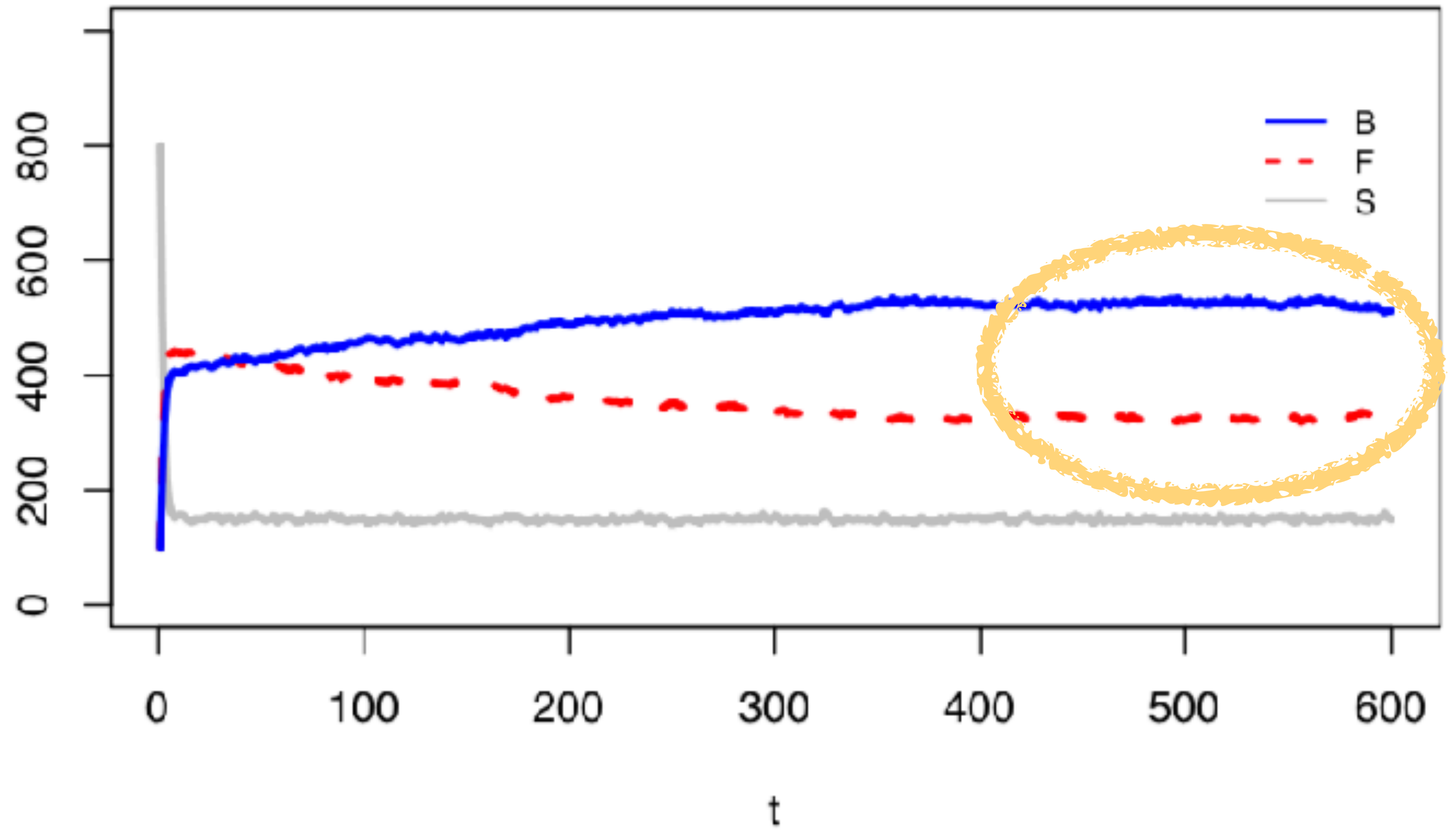
skeptical group:

- α: 0.3
- ~~seeders FC: 10%~~
- **BRIDGES are eFC!**

Simulation start



Simulation results



comparable, more realistic

Lessons learned and observations

- **Debunking activism** is often considered useless or **counterproductive**
- However, a world without fact-checking is harmless against fake-news circulation: **skeptics exposed to misinformation** will turn into **believers** because of **social influence**
- **Skeptics with links to gullible subjects** should be the first to be exposed to the fact-checking: misinformation will survive in the network, but their communities can be 'protected' by such **gatekeepers**
- Note: no socio-psychological assumption so far. Real world is much more complicated

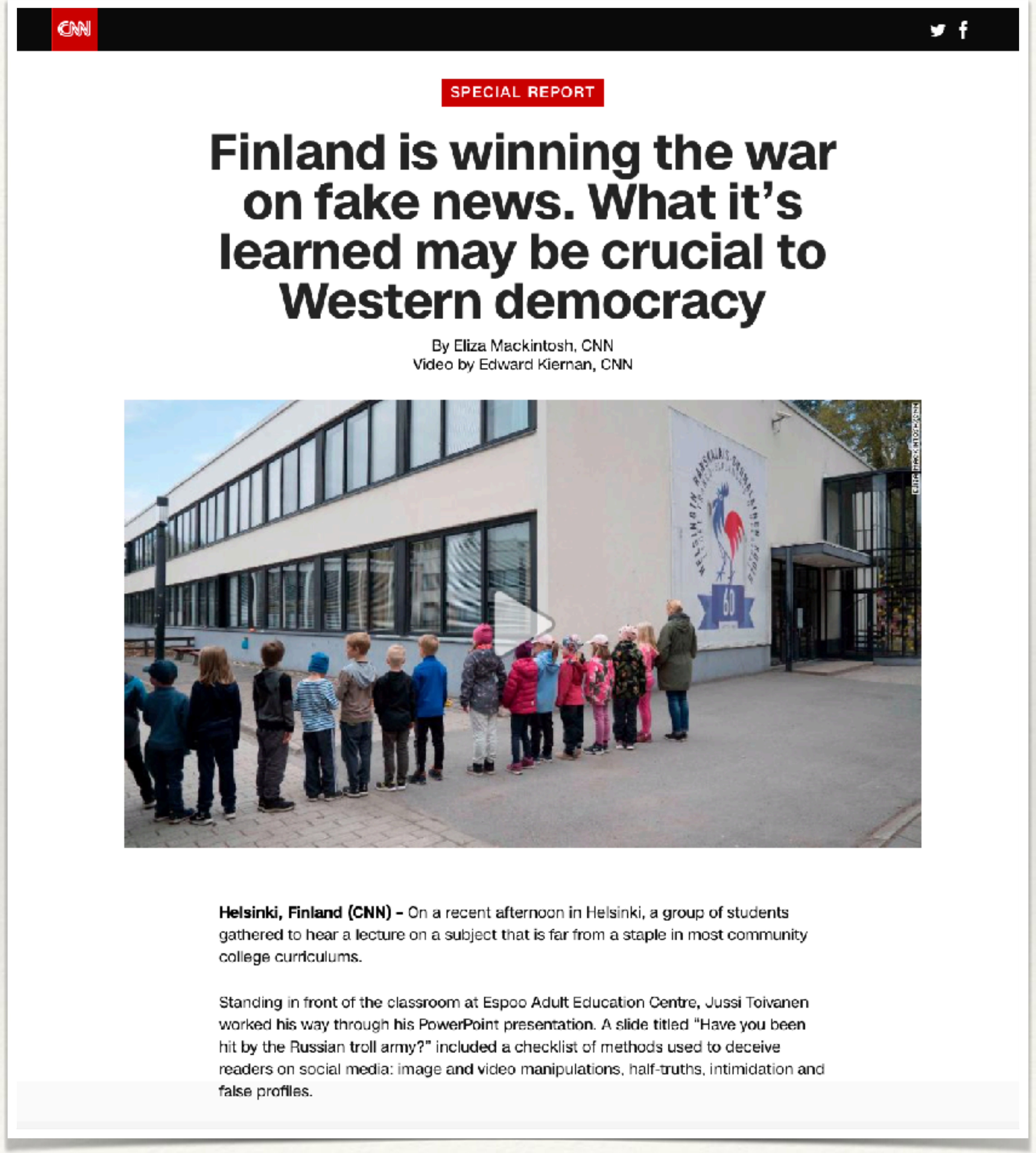
M Tambuscio, G. Ruffo, Fact-checking strategies to limit urban legends spreading in a segregated society, in Applied Network Science 4, 116 (2019), Springer, https://appliednetsci.springeropen.com/articles/10.1007/s41109-019-0233-1

protect the vulnerable, encourage skepticism

Who is the gatekeeper?

Finland is reported as winning the war against fake news in the classrooms: **education** first

Teachers and the education system have a great **responsibility**



CNN

SPECIAL REPORT

Finland is winning the war on fake news. What it's learned may be crucial to Western democracy

By Eliza Mackintosh, CNN
Video by Edward Kiernan, CNN

Helsinki, Finland (CNN) – On a recent afternoon in Helsinki, a group of students gathered to hear a lecture on a subject that is far from a staple in most community college curriculums.

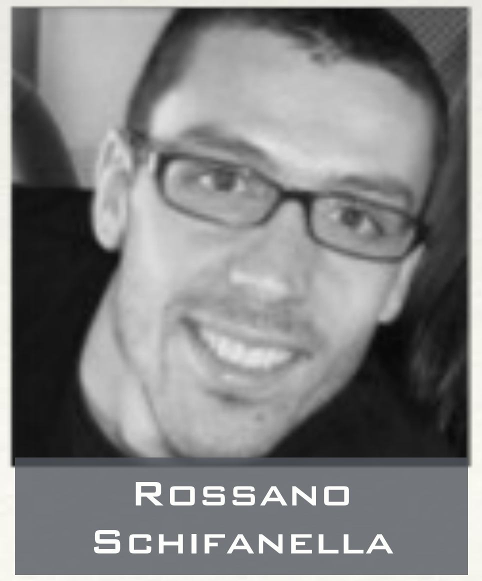
Standing in front of the classroom at Espoo Adult Education Centre, Jussi Toivanen worked his way through his PowerPoint presentation. A slide titled "Have you been hit by the Russian troll army?" included a checklist of methods used to deceive readers on social media: image and video manipulations, half-truths, intimidation and false profiles.

The Rise of Social Bots

Overview of the impact of bots

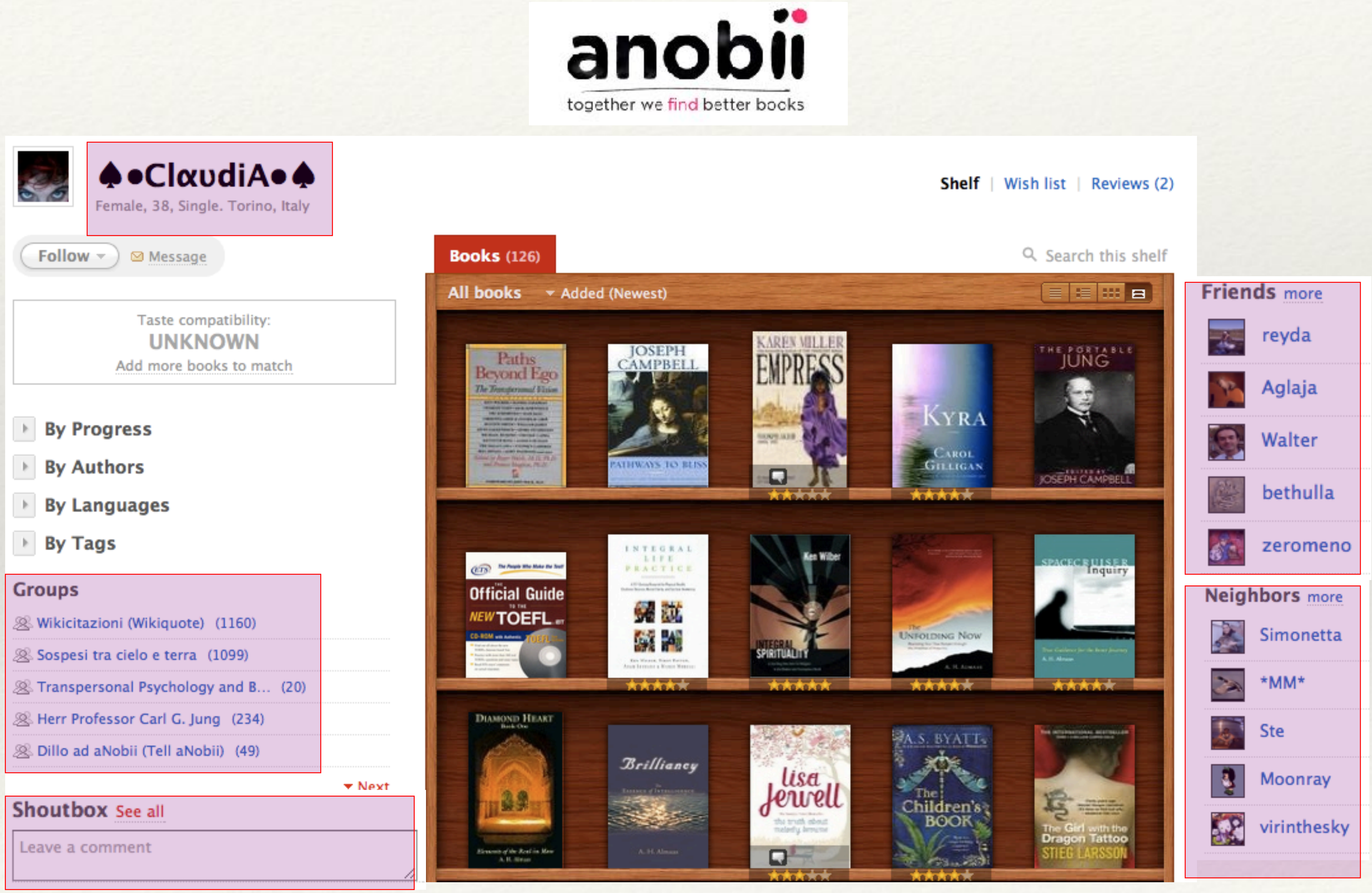
- The strange case of lajello
- The path to botometer
- The impact of bots on disinformation diffusion
- Case study: the interplay between bots and low quality information diffusion in the italian debate on immigration on twitter

The strange case of Lajello



Analyzing social network with a bot

- Anobii was a social networks for book lovers
- Scraping users' profiles from the Web was admitted
- Users' libraries and their links were collected periodically

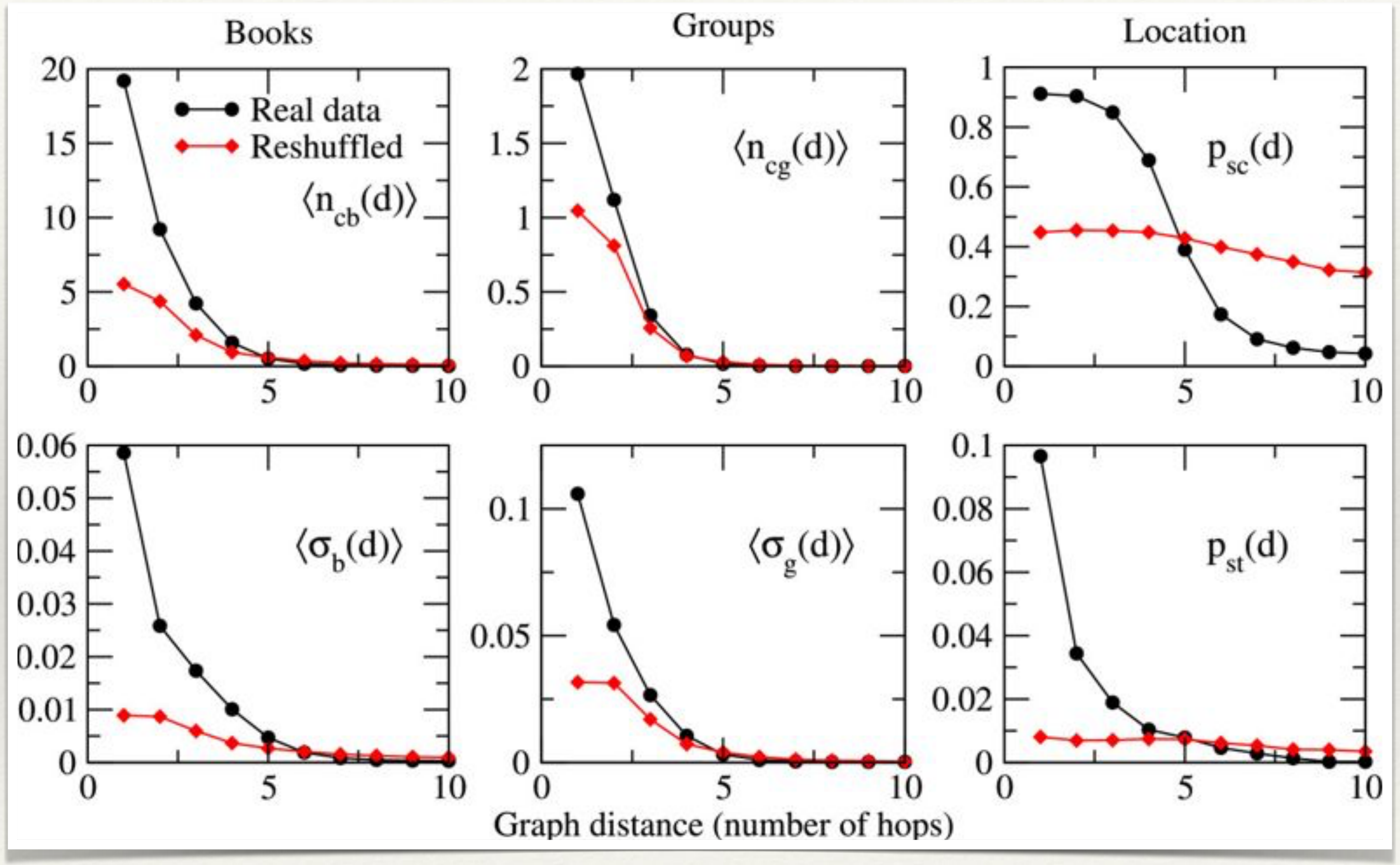


Analyzing social network with a bot

- Anobii was a social networks for book lovers
- Scraping users' profiles from the Web was admitted
- Users' libraries and their links were collected periodically
- **The bot "Lajello" used to silently navigate Anobii twice a month for one year**

Analyzing social network with a bot

- Anobii was a social networks for book lovers
- Scraping users' profiles from the Web was admitted
- Users' libraries and their links were collected periodically
- The bot "Lajello" used to silently navigate Anobii twice a month for one year
- **homophily by selection and by influence analysed**

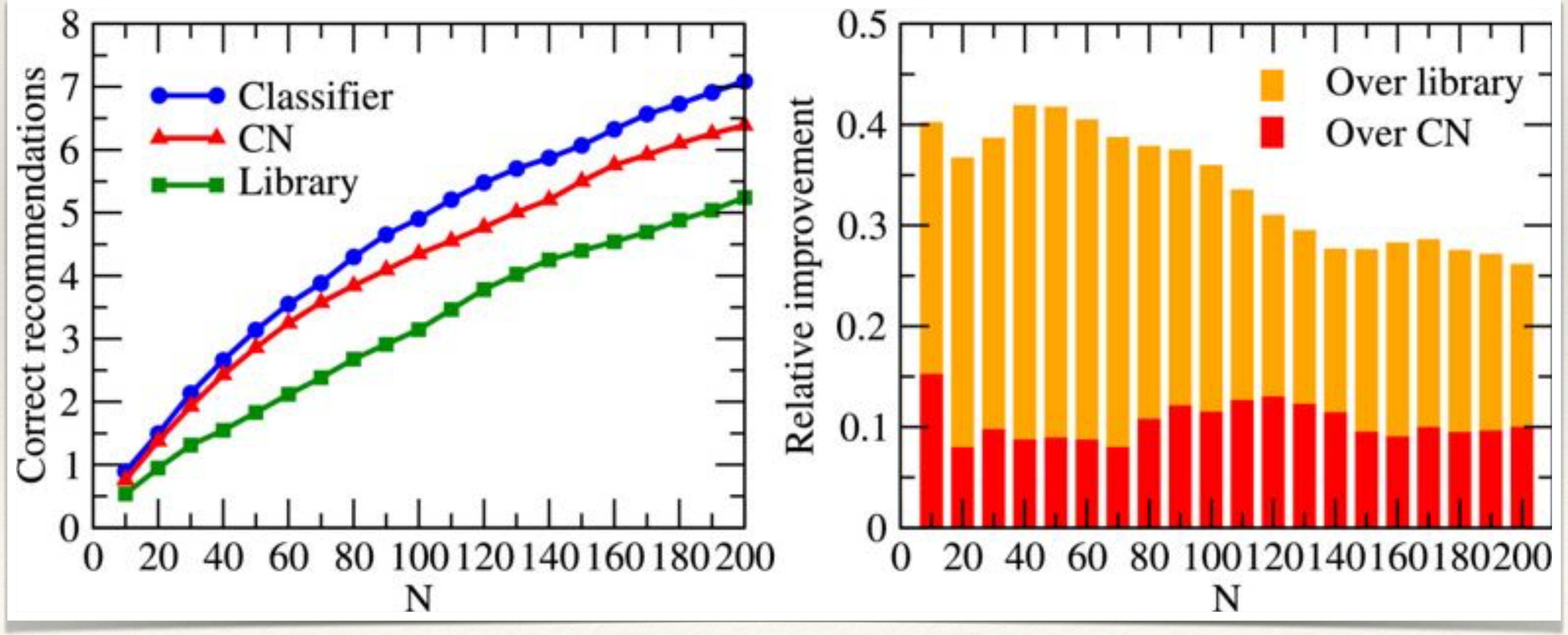


LM Aiello, A Barrat, C Cattuto, G Ruffo, R Schifanella, Link creation and profile alignment in the aNobii social network, 2010 IEEE 2nd Int.. Conf. on Social Computing, 249-256

LM Aiello, A Barrat, C Cattuto, G Ruffo, R Schifanella, Link creation and information spreading over social and communication ties in interest based online social network, EPJ Data Science 1 (1), 12

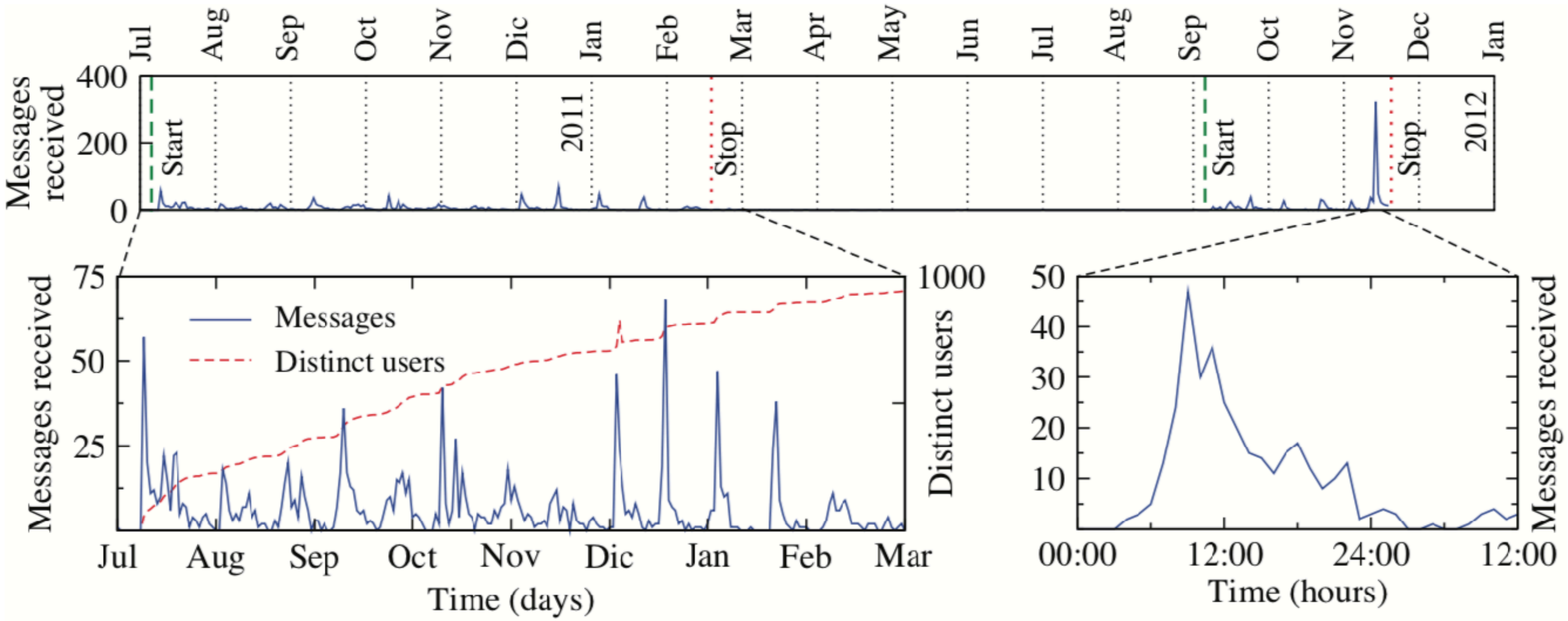
Application: a link recommendation algorithm

- A link recommendation algorithm based on prediction of profile similarities was proposed and tested
- Results showed an improvement w.r.t. the baselines



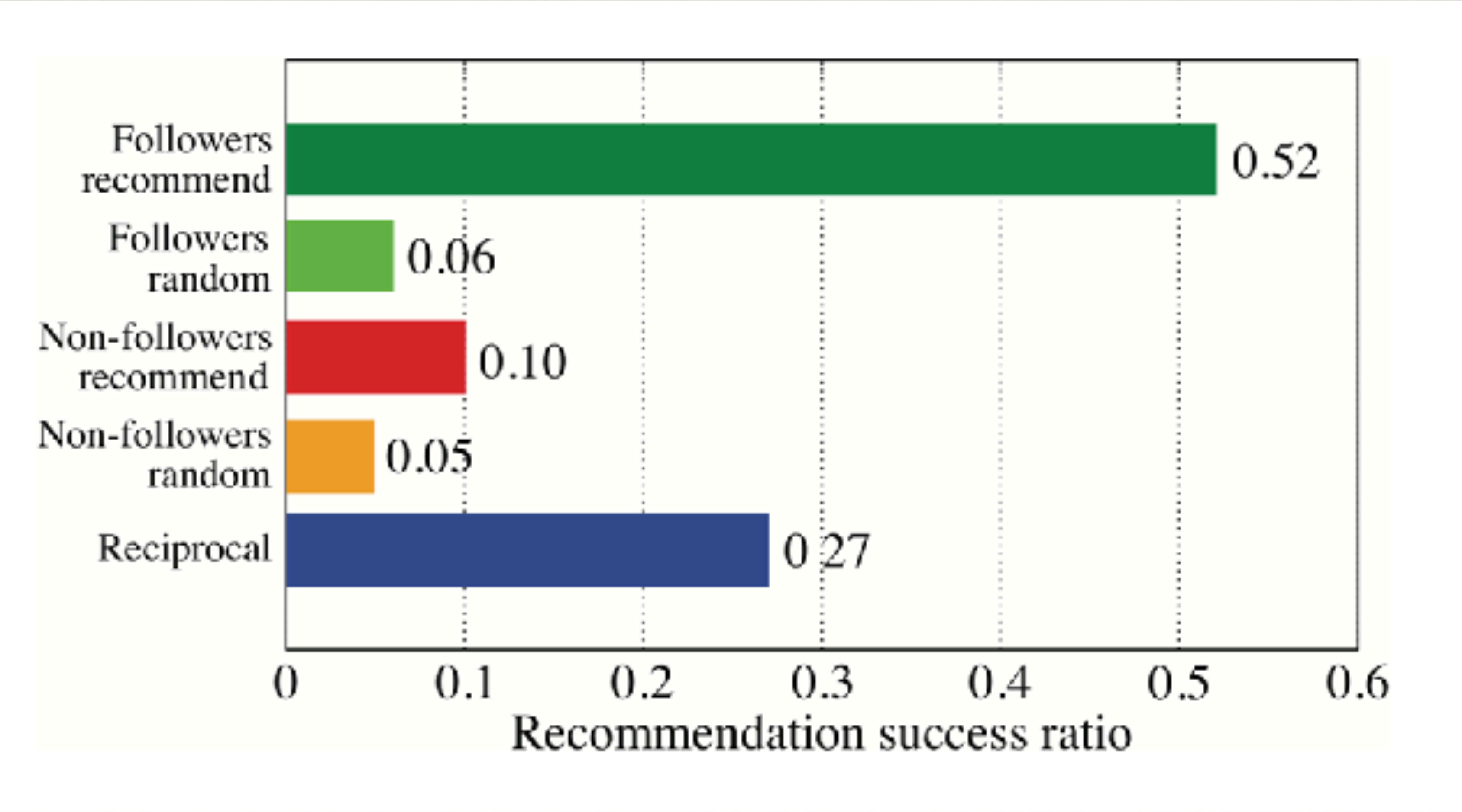
What happened to Lajello?

Lajello, incidentally, became the second most popular user in Anobii in terms of messages from distinct users



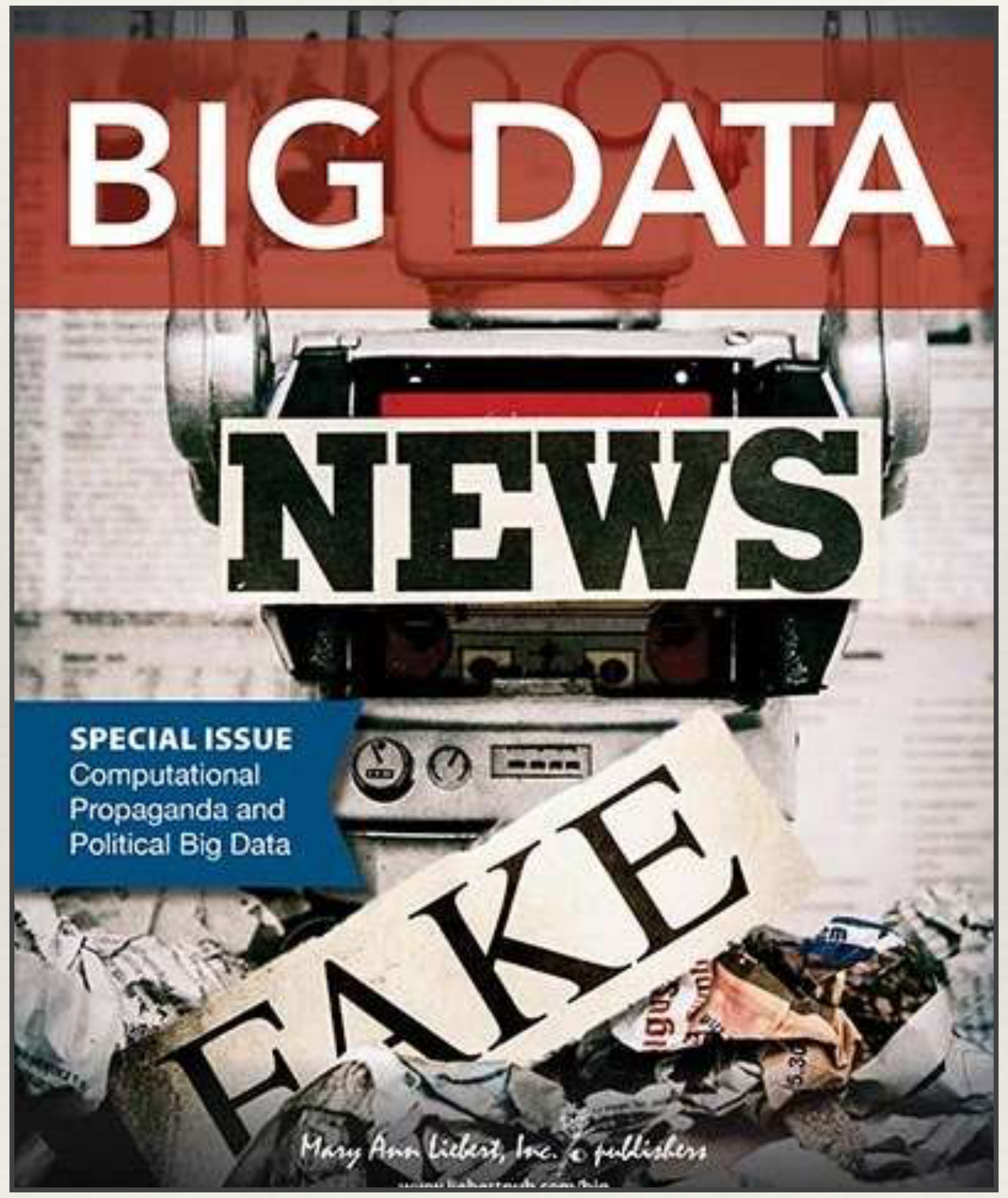
Exploiting Lajello popularity

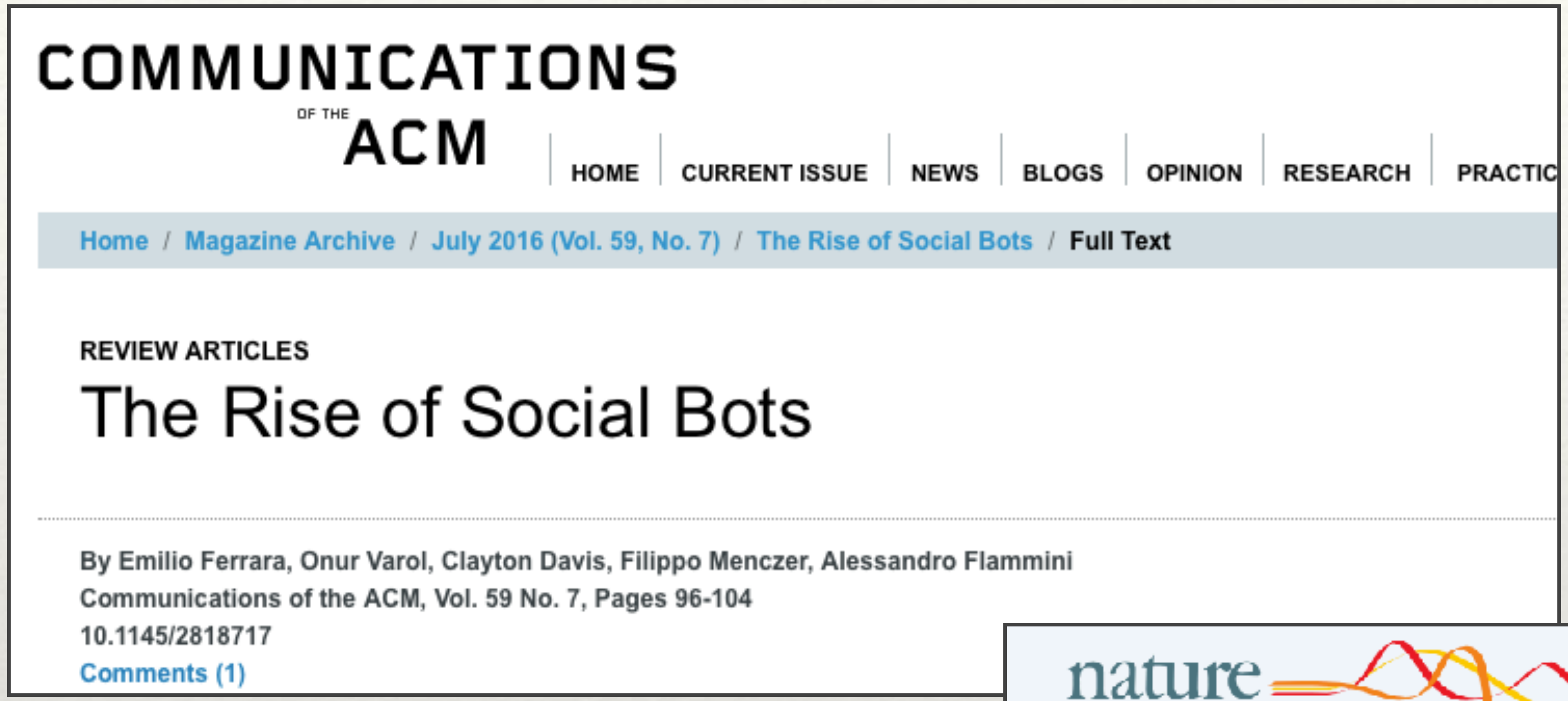
- Lajello started to introduce users to each other according our link recommendation algorithm
- First result: users acceptance of the recommendation skyrocketed if they previously wrote in Lajello's wall



LM Aiello, M. Deplano, R Schifanella, G Ruffo, People are Strange when you're a Stranger: Impact and Influence of Bots on Social Networks, in Proc. of the 6th Intern. AAAI Conf. on Weblogs and Social Media (ICWSM'12), Dublin, Ireland, 2012

Influence of bots




COMMUNICATIONS OF THE ACM

HOME | CURRENT ISSUE | NEWS | BLOGS | OPINION | RESEARCH | PRACTIC

Home / Magazine Archive / July 2016 (Vol. 59, No. 7) / The Rise of Social Bots / Full Text

REVIEW ARTICLES

The Rise of Social Bots

By Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, Alessandro Flammini
Communications of the ACM, Vol. 59 No. 7, Pages 96-104
10.1145/2818717
Comments (1)


nature COMMUNICATIONS

Article | Open Access | Published: 20 November 2018

The spread of low-credibility content by social bots

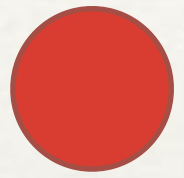
Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini & Filippo Menczer

Nature Communications **9**, Article number: 4787 (2018) | Download Citation

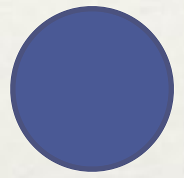
Incidentally, we created an "egg war"

- After our initial experiment, Lajello remained silent for one year and then he "talked". The recommendations changed the net structure and lajello account was banned after 24 hours. This ignited a "war"
- Two polarized opinions emerged: Anobii users created immediately two thematic groups: "the (not requested) suggestions of Lajello" and "Hands-off Lajello"
- A large portion of users that were contacted by Lajello joined to one of these groups
- We observed a strong interplay between the existing relationships in the social network and the opinion that emerged from the users at the end of the links: "**echo chamber**" effect?

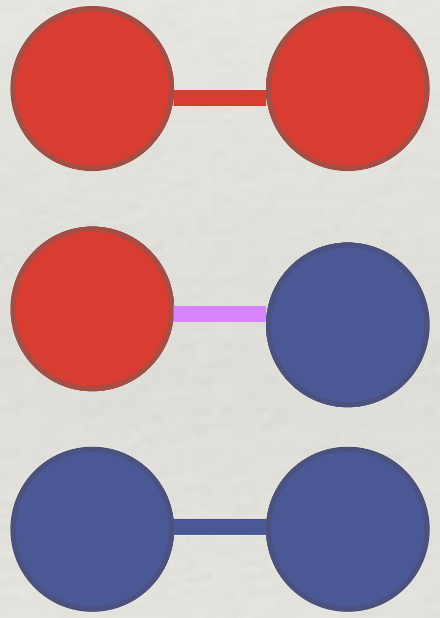
Social polarization and emotional reaction



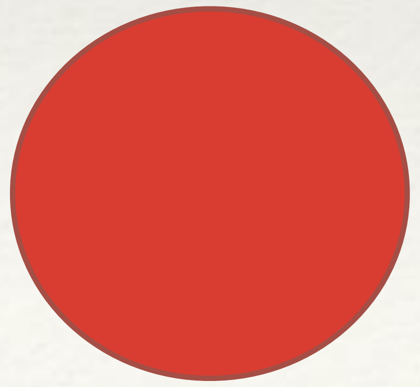
red dots are lajello **supporters**



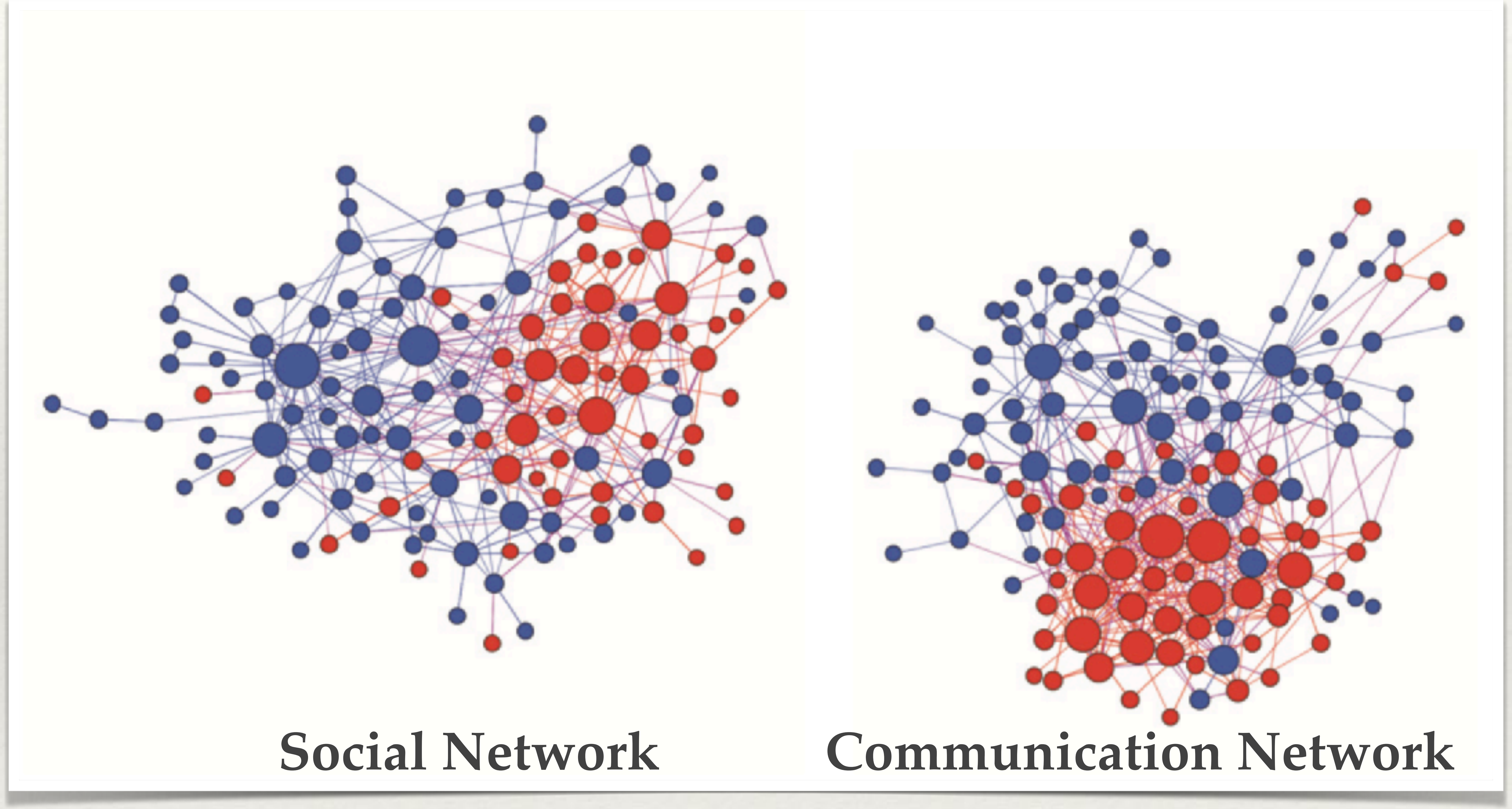
blu dots are lajello **haters**



links are existing
social connections
or **direct messages**
(graph is directed)



bigger dots are
users with more links



Automatic network-based community detection algorithm (OSLOM) accurately finds clusters (80% - Social network, 72% - Communication network), confirming a signal of **segregation** between the two groups before link recommendations

LAJELLO... HAI STUFATO..NON SE NE PUO' PIU'...STA ATTENTO/A CHE SONO CAPACE DI ASSOLDARE UN HACKER PER VEDERE CHI SEI..E PO' SONO C...TUOI

Tre settimane fa

chi sei?

già che mi ritrovo qui mi faccio pubblicità! Venite a vedere la mia libreria è la più bella -del mondo-. (l'ultima parte andava sottolineata..)

Due sett

Le tue visite cominciano ad essere inquietanti....

ahahahhahahaha tu sei un genio!!!!! sei davvero un genio!!! insomma ma quante visualizzazioni hai???? sei un grande!!!!! riesci a farti visitare e a farti scrivere pur non avendo libri!!! ti adoro sei grandissimo :P

Aug 13, 2010

un grande.
continua così. Grazie delle visite, si vede che ti sto simpatica....

P.S: propongo di aprire un gruppo the Lajellos fans...

3 giorni fa

chapeau!!

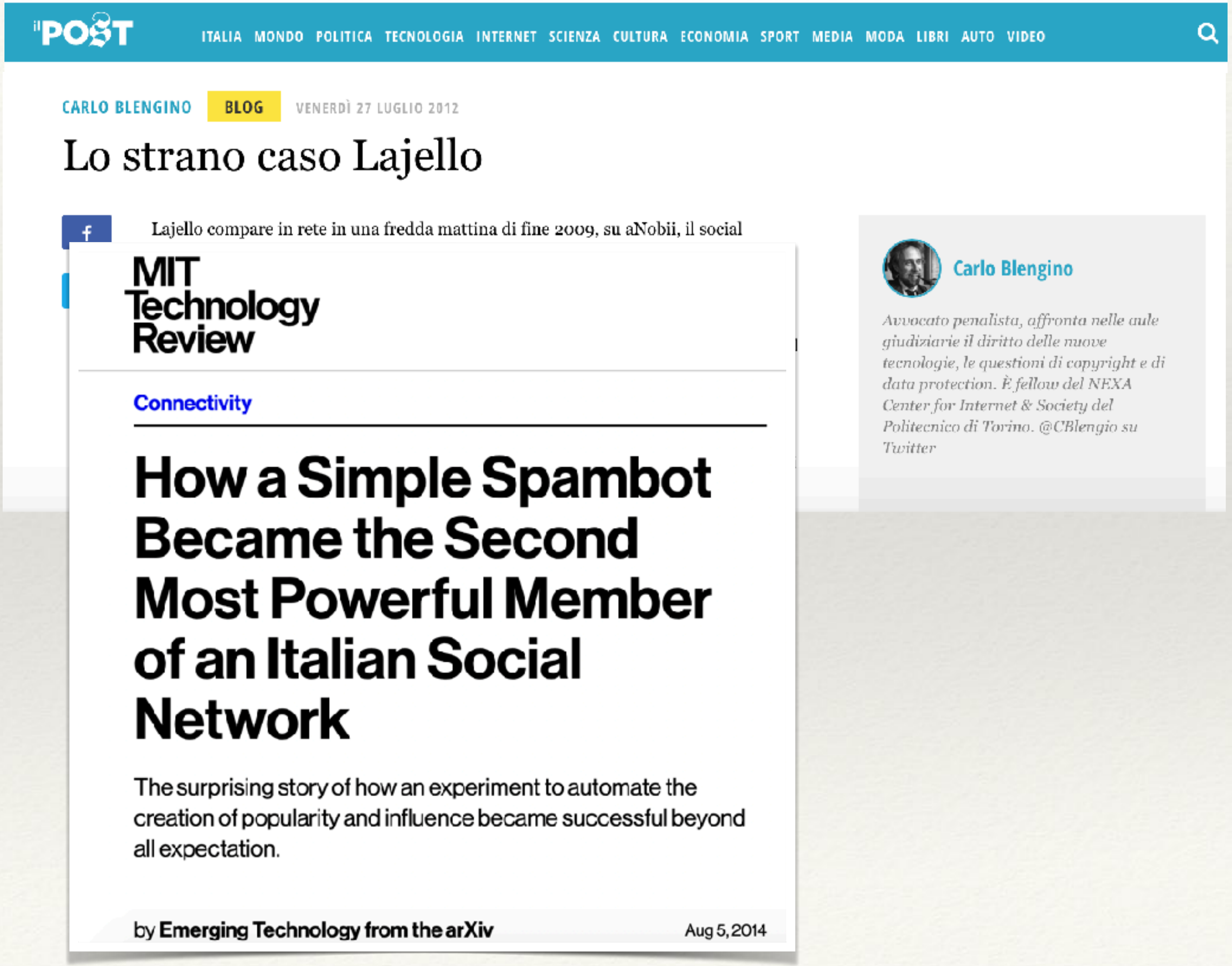
ahahahaahah tu sei un genio!!

Grazie Lajello, mi sono divertita un sacco a leggere i commenti degli altri anobiani. Sembra un esperimento di psicologia sociale, se non ti dispiace ti aggiungo come vicino! e resisti eh...non pubblicare un libro! ;)

Due settimane fa

Lessons learned and observations

- Handle **experiments in social media** with care :)
- A simple **spambot can take power** in a social network
- A seed of **polarization** found in pre-existing network **structure**
- … also the structure changed after our experiment was run!
- What if the real identity and motivations of Lajello were fact-checked?



Bots and the problem of prevalence

Prevalence

- Many observed that false stories in social media are more successful (in numbers and speed) than true stories
- *Which are the key factors?*
- *Who is to blame: bots or humans?*

Lies are faster than truth

- Dataset: ~126,000 stories tweeted by ~3 million people more than 4.5 million times.
- News classified as true or false using six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications.

S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

- Falsehood diffused significantly **farther**, **faster**, **deeper**, and **more broadly** than the truth in all categories of information

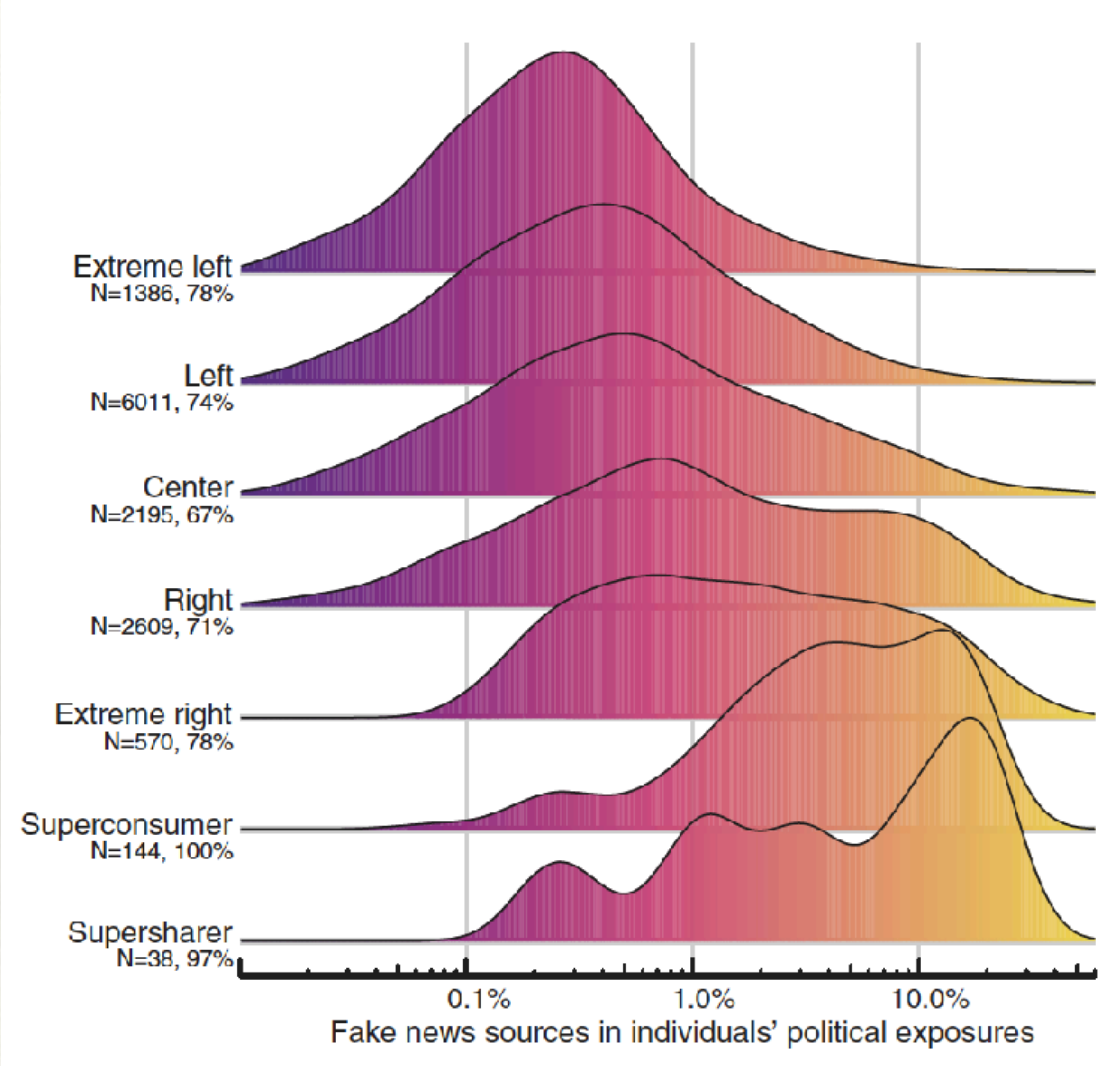
S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

- Effects were **more pronounced for false political news** than for false news about terrorism, natural disasters, science, urban legends, or financial information.

S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

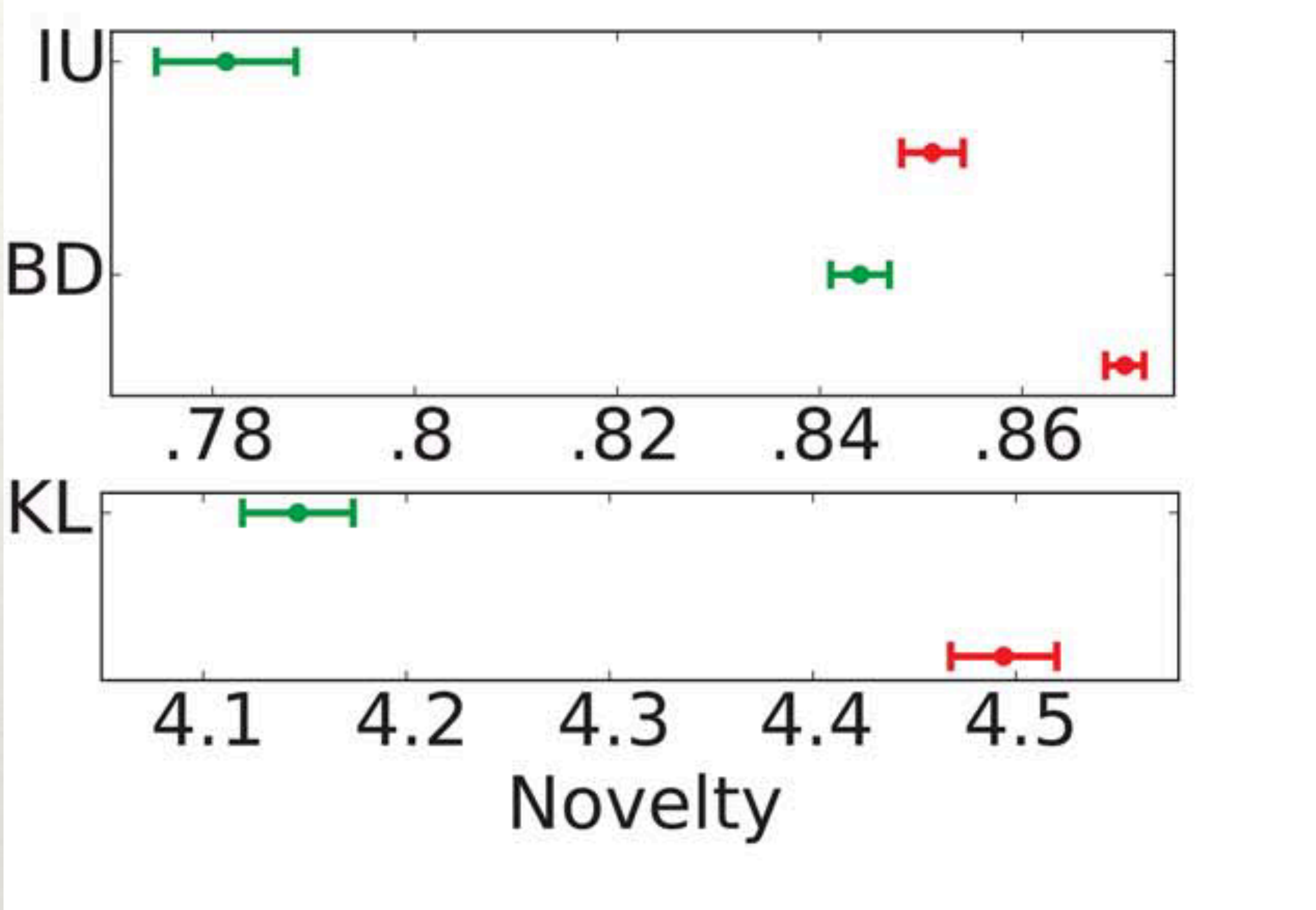
Fake-News and elections

- Engagement with fake news sources extremely concentrated in 2016 US presidential elections
- Only 1% of individuals accounted for 80% of fake news source exposures, and 0.1% accounted for nearly 80% of fake news sources shared.
- Individuals most likely to engage with fake news sources were conservative leaning, older, and highly engaged with political news.



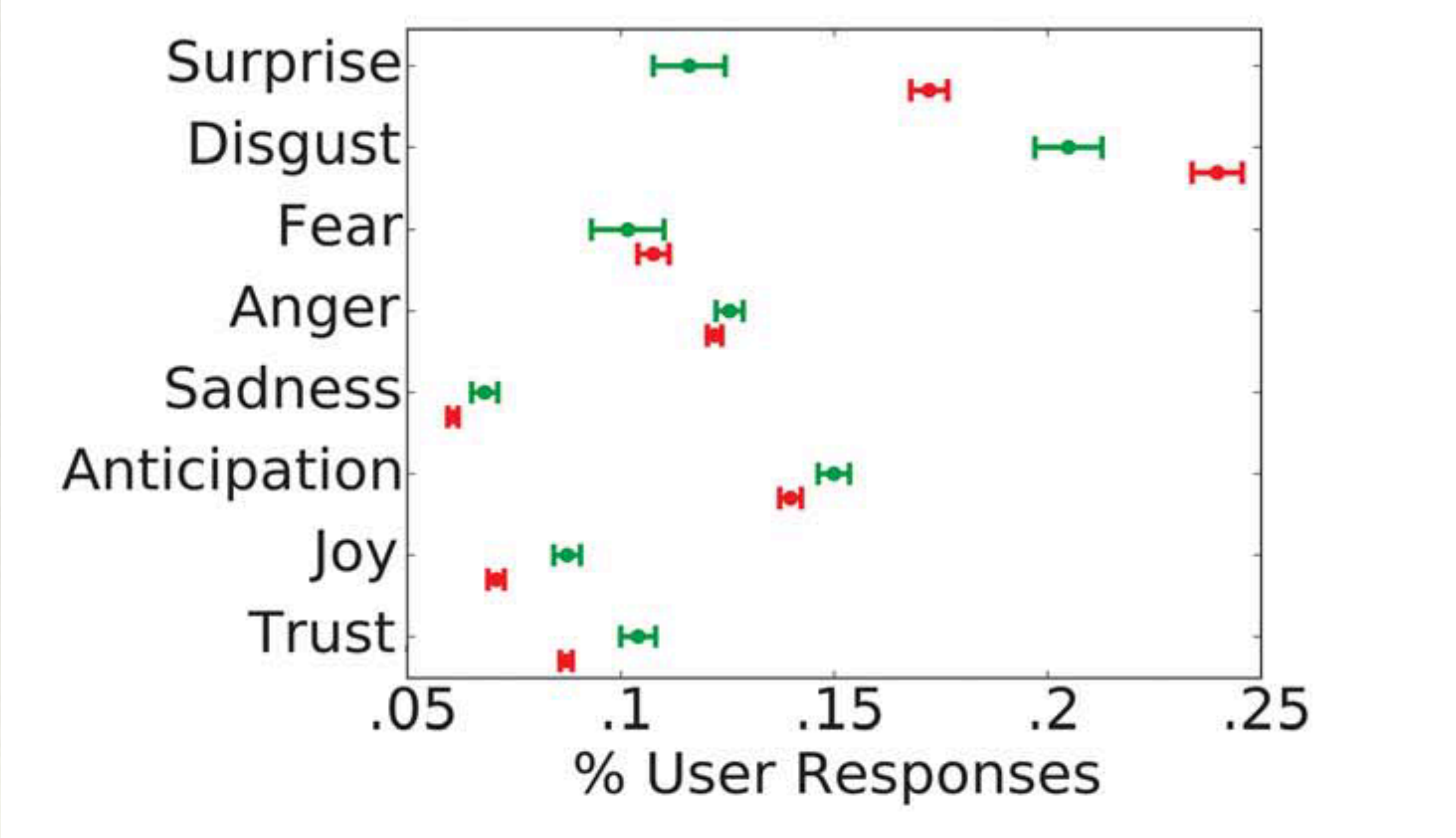
N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Fake news on Twitter during the 2016 U.S. presidential election, Science 25 Jan 2019:Vol. 363, Issue 6425, pp. 374-378 DOI: 10.1126/science.aau2706

Novelty and emotions



- False news **more novel** than true news, which suggests that people were more likely to share novel information

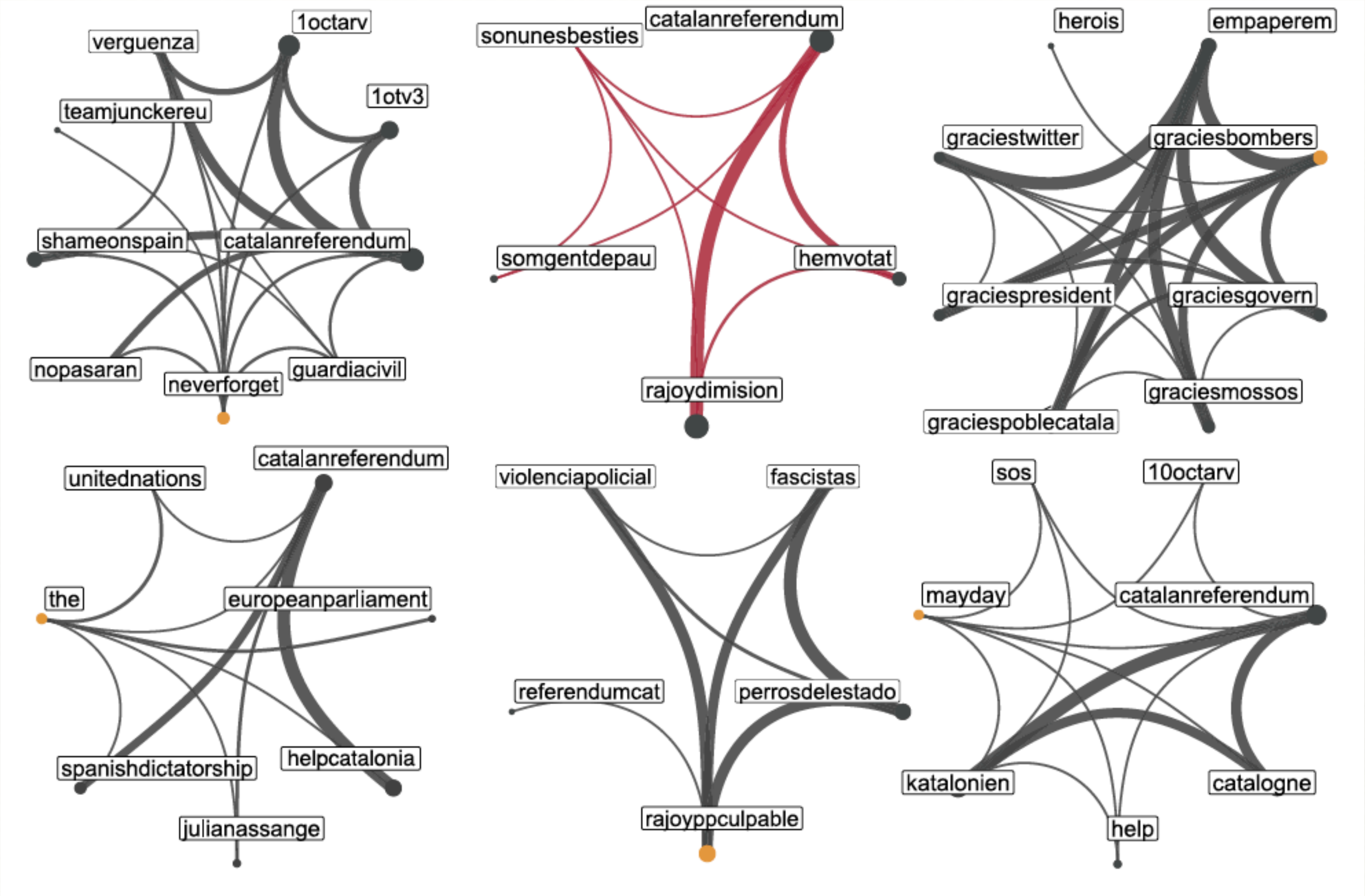
- False stories inspired **fear**, **disgust**, and **surprise** in replies, true stories inspired anticipation, sadness, joy, and trust.



S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

The role of emotions

- Large-scale social data collected during the **Catalan referendum for independence** on October 1, 2017, consisting of nearly 4 millions Twitter posts generated by almost 1 million users;
- Two polarized groups: **Independentists** vs **Constitutionalists**
- Structural and emotional roles played by **social bots**
    - Bots act from **peripheral areas** to target **influential humans** of both groups;
    - Bots bombard Independentists with **violent contents**, **increasing their exposure to negative and inflammatory narratives**, and exacerbating social conflict online.



M. Stella, E. Ferrara, M. Di Domenico, Bots increase exposure to negative and inflammatory content in online social systems, PNAS, Dec. 4, 2018, Vol. 115, no. 49, 12435–12440. www.pnas.org/cgi/doi/10.1073/pnas.1803470115

The role of social bots

- 14 million messages spreading 400 thousand articles on Twitter during ten months in 2016 and 2017
- Social bots played a disproportionate role in spreading articles from low-credibility sources.
- Bots amplify such content in the early spreading moments, before an article goes viral.
- They also target users with many followers through replies and mentions. Humans are vulnerable to this manipulation, resharing content posted by bots.

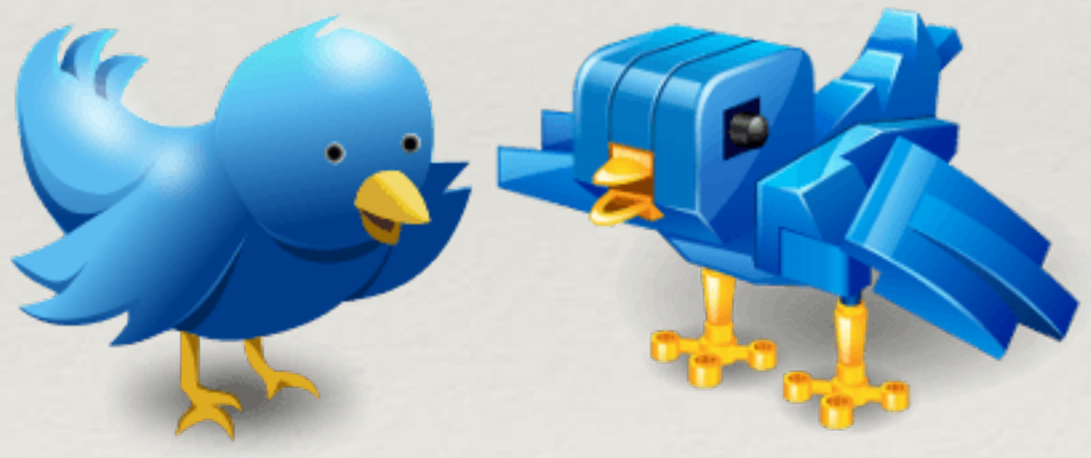
Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., The spread of low-credibility content by social bots. *Nat Commun* **9**, 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

BotSlayer and Botometer (IU)

- **BotSlayer**: it tracks and detect potential manipulation of information spreading on Twitter

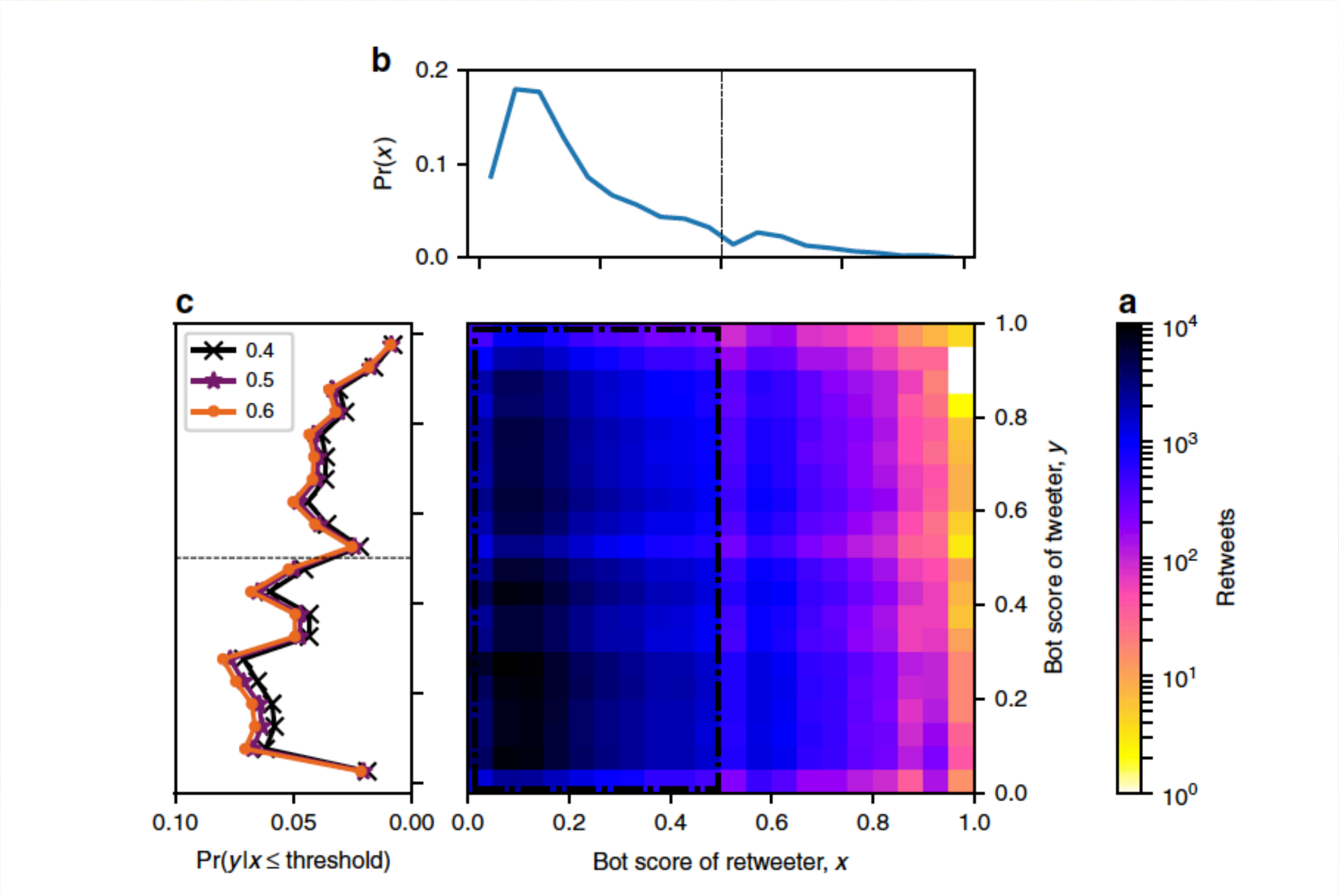
https://osome.iuni.iu.edu/tools/botslayer/

- **Botometer** (formerly known as BotOrNot) :checks the activity of a Twitter account and gives it a score. Higher scores mean more bot-like activity.



https://botometer.osome.iu.edu

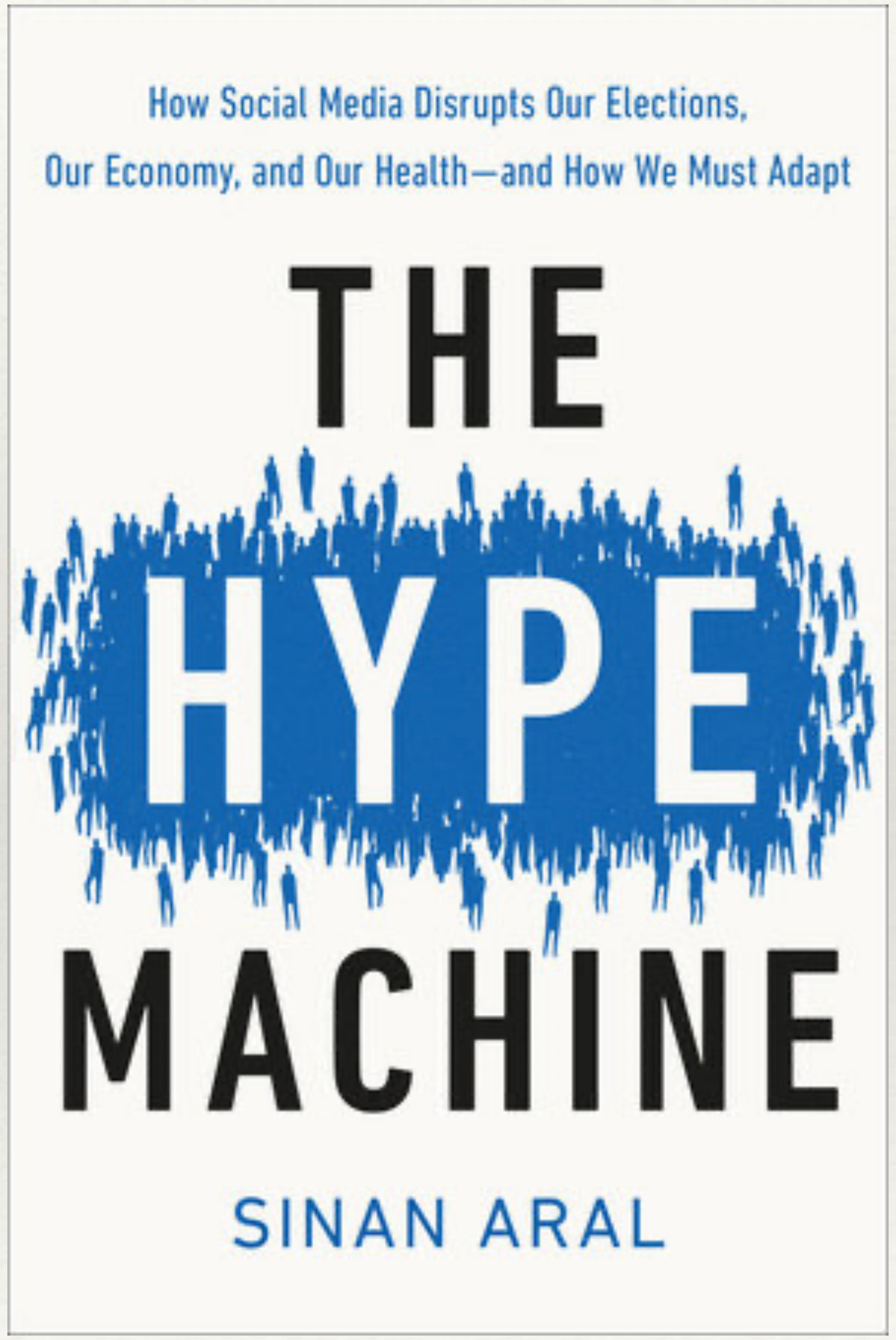
…but humans should be blamed the most



Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., The spread of low-credibility content by social bots. *Nat Commun* **9,** 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

The Hype Machine

- Prevalence of fake-news and role of social bots in spreading misinformation
- Bots share **novel** fake news and retweet it broadly
- Bots **mention influential humans** incessantly
- The strategy works when influential people are fooled into sharing the content.
- **Misleading humans is the ultimate goal of any misinformation campaign**



https://www.salon.com/2020/09/27/fake-news-bots-spreading-misinformation-2020-election-propaganda/

Open Problems and Trends

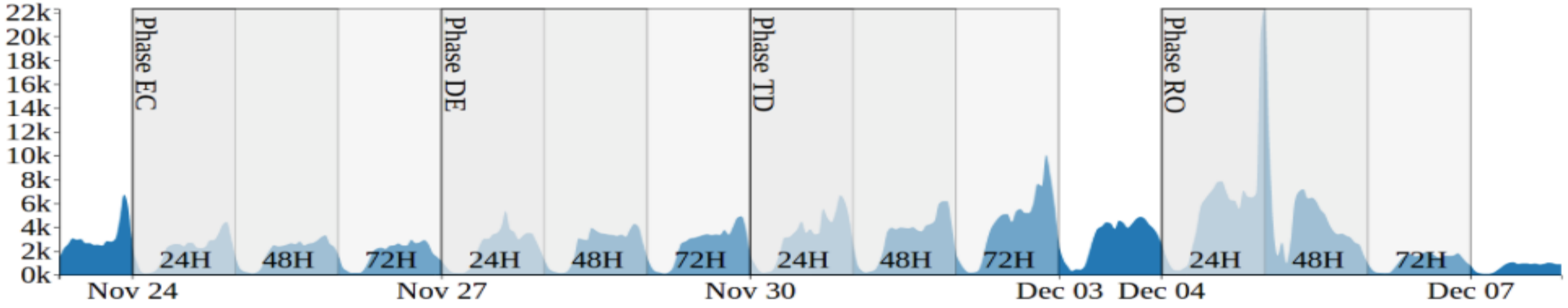
Language and network structure

Links to NLP

- Individual's opinions are often hidden
- Social Media provide much data for stance detection, emotion analysis, and so on
- Communication styles can be another trigger or just a reaction to news exposition and partisanships
- Relationships between structural segregation and opinion formation and polarization should be explored further by a joint effort between our scientific communities

Italian 2016 Constitutional Referendum

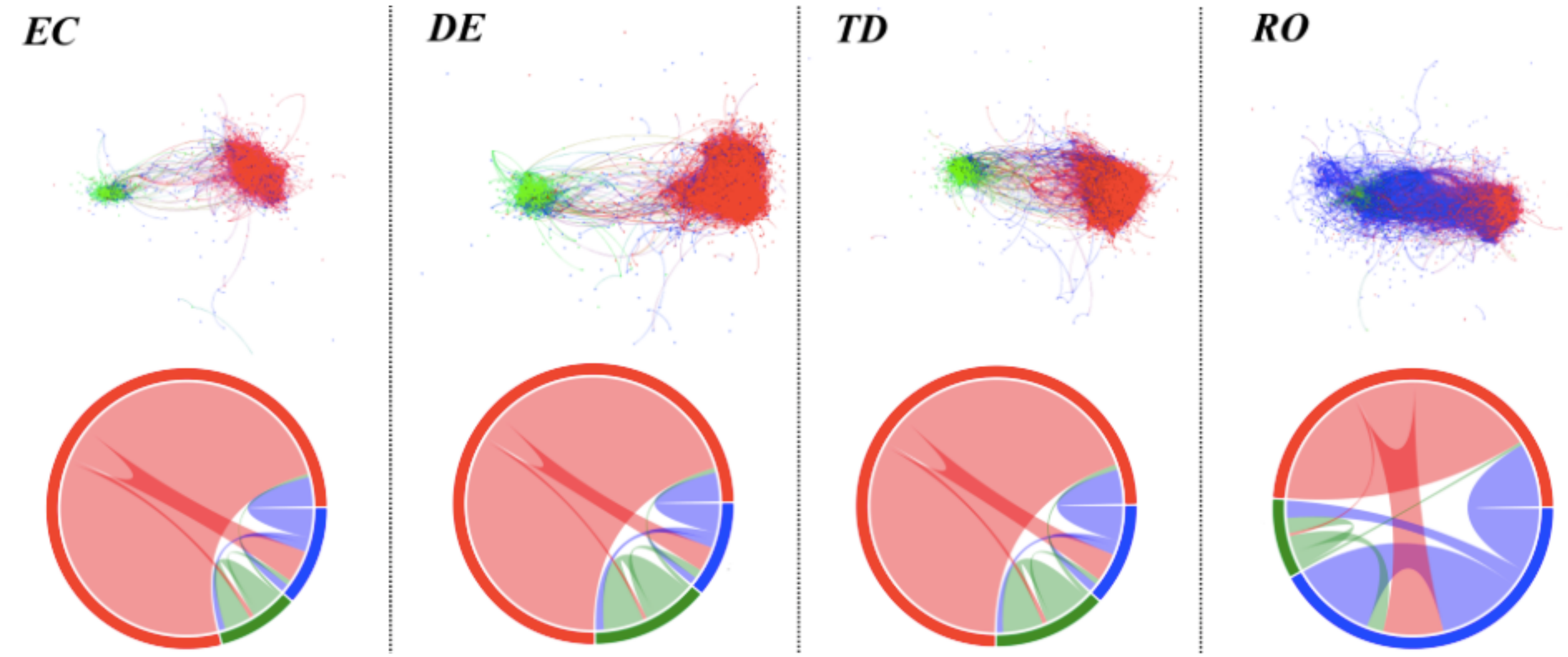
Collected Tweets



stance detected as AGAINST

stance detected as IN FAVOR

stance detected as NONE

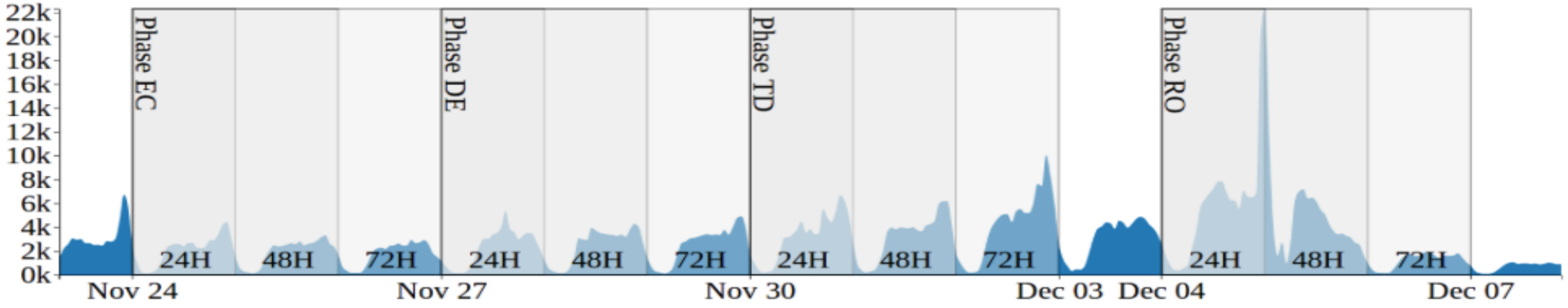


Retweet Network

strong signal of **homophily**

Italian 2016 Constitutional Referendum

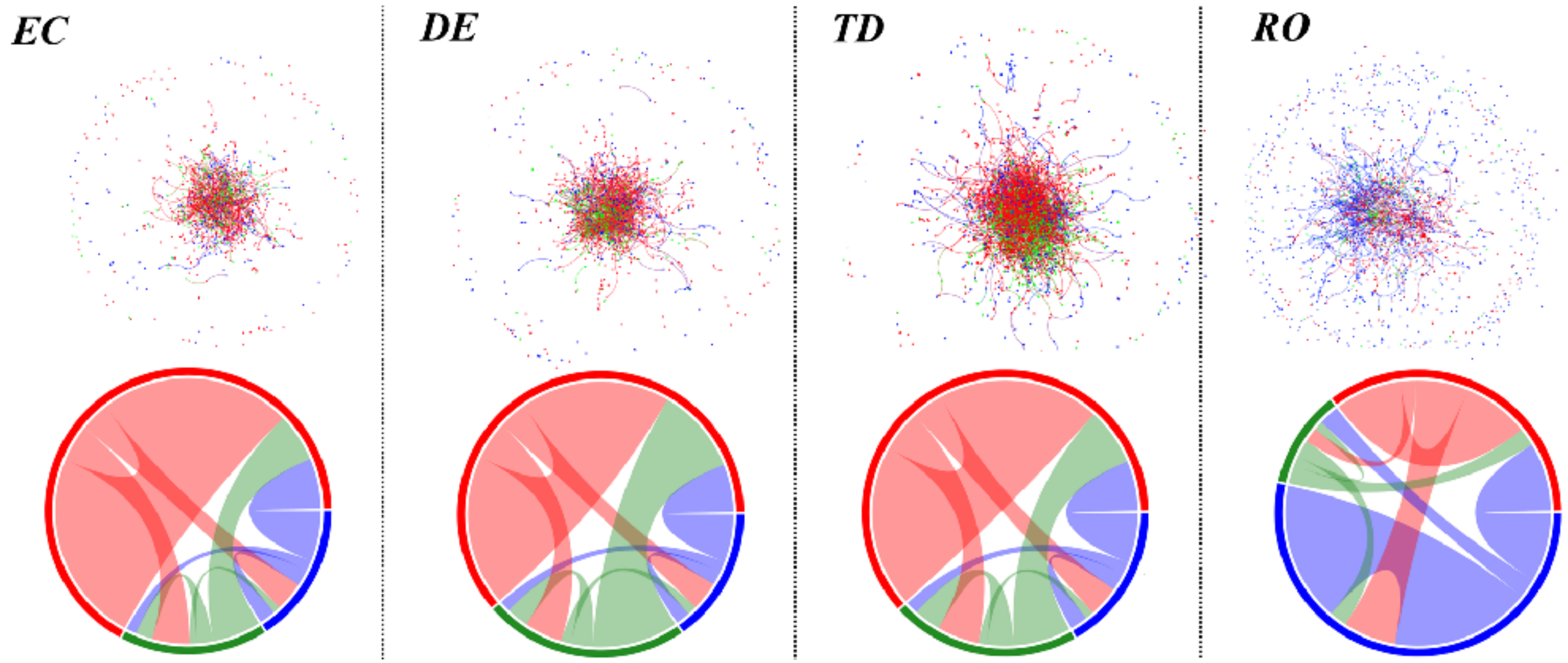
Collected Tweets



stance detected as **AGAINST**

stance detected as **IN FAVOR**

stance detected as **NONE**



Mention Network

signal of **inverse homophily**

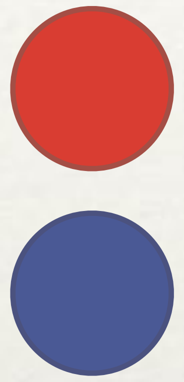
Stance detection and Network Homophily

- ML-based **stance detection** is a NLP tool extremely useful for computational social science analyses
- We need **approximation** of users' opinions
- Building networks that **evolve** when the polarizing debate takes place is an opportunity to study the **interplay between structure and opinions**
- Apparently in Twitter retweets and reply-to are used to respectively show agreement or disagreement. If you look for disputes, **dig mentions**

M Lai, M Tambuscio, V Patti, P Rosso, G. Ruffo, Stance Polarity in Political Debates: a Diachronic Perspective of Network Homophily and Conversations on Twitter, Data & Knowledge Engineering Journal, online: September 2019
https://www.sciencedirect.com/science/article/pii/S0169023X19300187

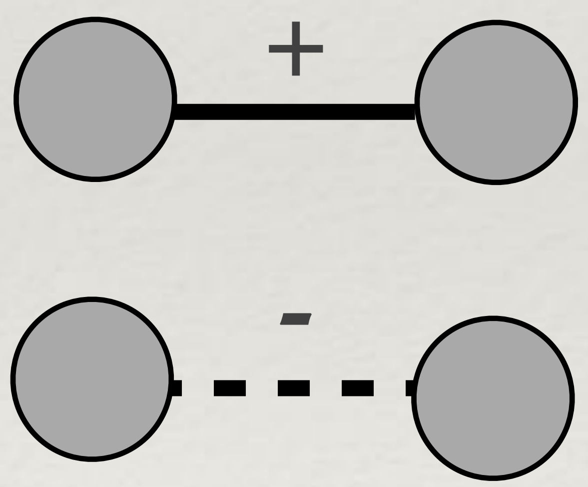
Balance in networks: algorithms and visualization

Signed nets

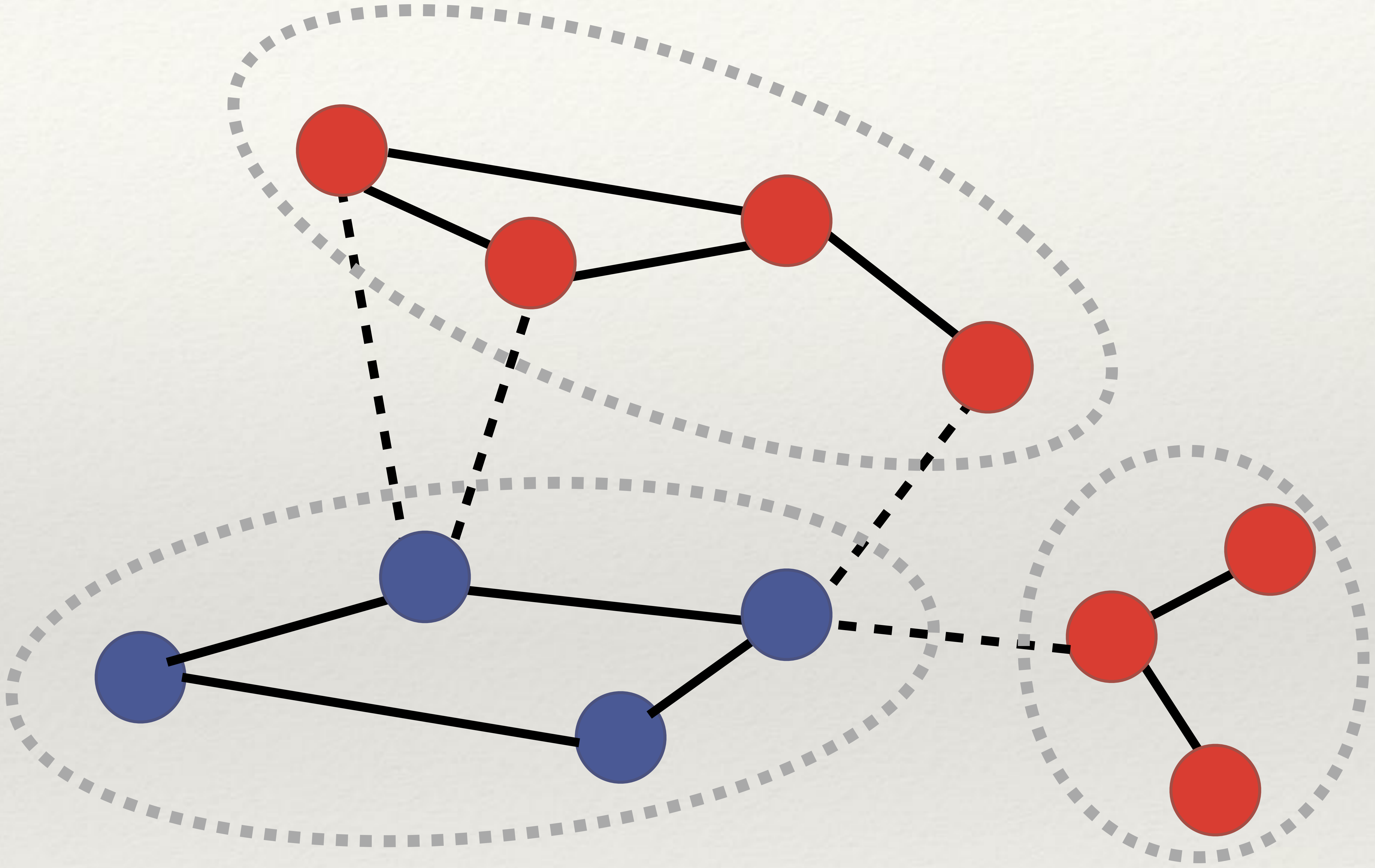


journalists

scientists

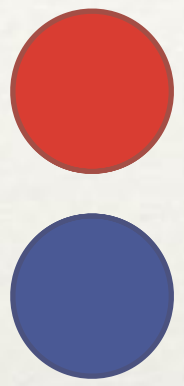


signs make explicit the type of the relationship



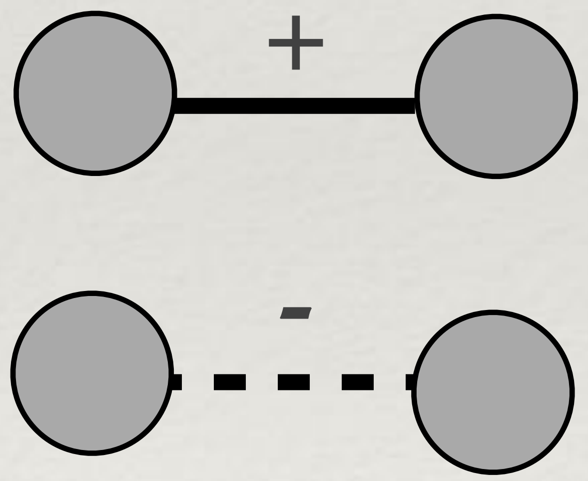
Balanced

Signed nets

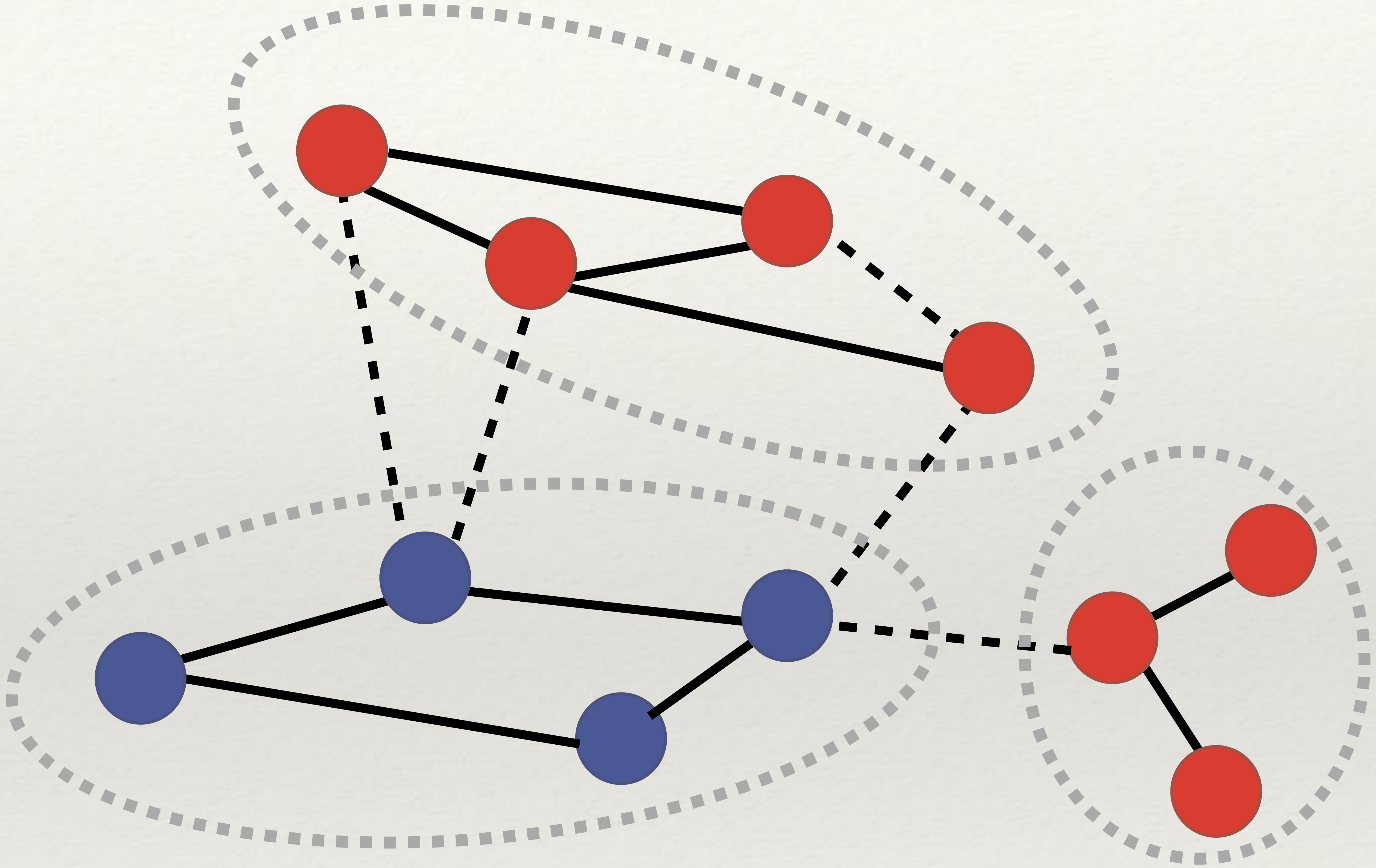


journalists

scientists



signs make explicit
the type of the
relationship



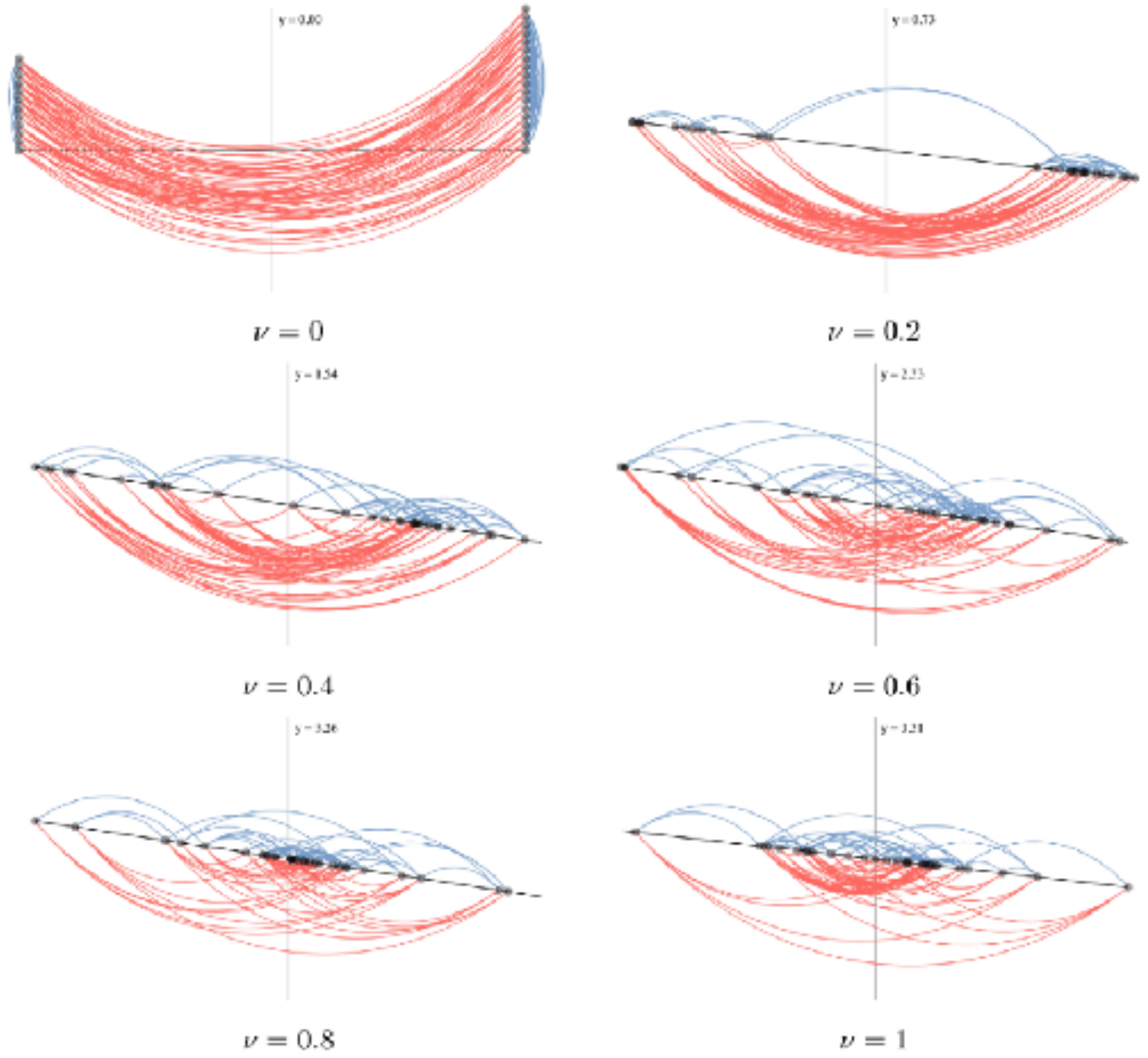
Not balanced

Balance in networks

- Balance is not always good: if journalists hate scientists and vice versa, we would live in a perfectly balanced world!
- There are different levels of balance when few negative edges cross boundaries
- Partial balance is a measure of polarization (or to predict a forthcoming egg war?) - *frustration index problem*
- Probably a great framework, not fully exploited so far, to better understand polarization and segregation dynamics in socio-political systems

Algorithms for communities detection and visualization

EDOARDO GALIMBERTI



2-Polarized-Communities: an algorithm based on spectral properties of the graph

Stuctural-balance-viz: spectral properties used to emphasize balance/unbalance

F Bonchi, E Galimberti, A Gionis, B Ordozgoiti and G Ruffo, Discovering polarized communities in signed networks, in Proc. of CIKM 2019 (Beijing, China)

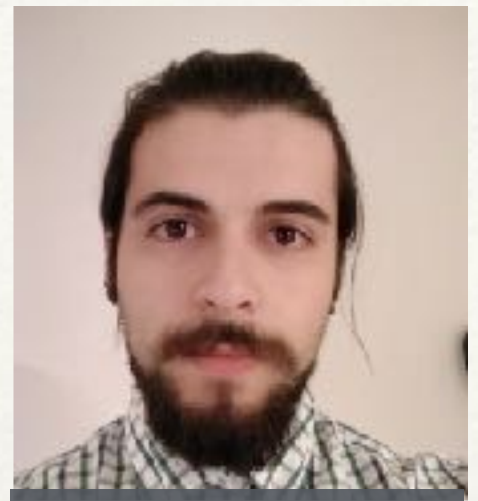
E Galimberti, C Madeddu, F Bonchi, and G Ruffo, Visualizing structural balance in signed networks, in Proc. of COMPLEX NETWORKS 2019 (Lisbon, Portugal)

Discussion and conclusions

Recap

- **Structural segregation** may be one of the main triggers of opinion **polarization**
- **Fake-news spreading**, especially when partisanship and antagonistic behavior reinforce the debate, is **facilitated** in segregated networks
- Fact-checking is needed and skeptics with links to more gullible (vulnerable) contacts can be recruited as **gatekeepers**
- **Network Analysis** and **NLP** are great tools for modeling and analyzing data in this domain
- **Balance theory** provides a so far neglected framework to study the interplay between opinion polarization and structural segregation: new **algorithms** and **visualizations tools** can be added to the analytical loop
- Beware of the **interplay**: segregation causes polarization and vice-versa

di.unito.it


Arthur Capozzi


Alfonso Semeraro


Alessandra Urbinati

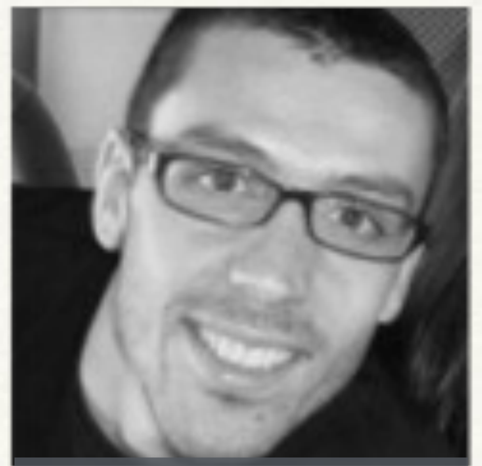

Salvatore Vilella


Mirko Lai


Marcella Tambuscio


André Panisson


Luca Aiello


Rossano Schifanella


Viviana Patti


Edoardo Galimberti


Emilio Sulis


Martina Deplano


Ciro Cattuto


Cristina Bosco

ARC^{2}S: Applied Research on Computational Complex Systems

Thanks!

Bookmark this:
(slides and annotated bibliography available soon)

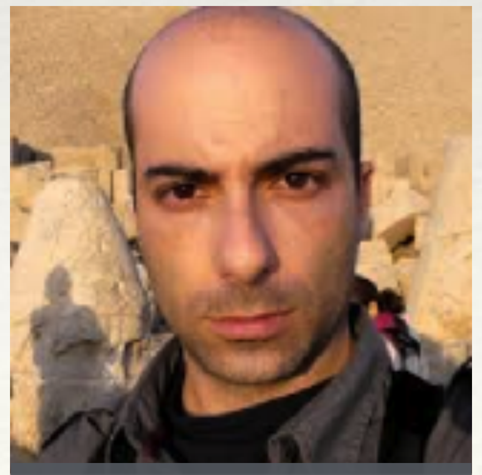
http://www.di.unito.it/~ruffo/talks/2021_COINS.pdf

UNIVERSITAT POLITÈCNICA DE VALÈNCIA


Paolo Rosso

Universidad del Desarrollo


Leo Ferres

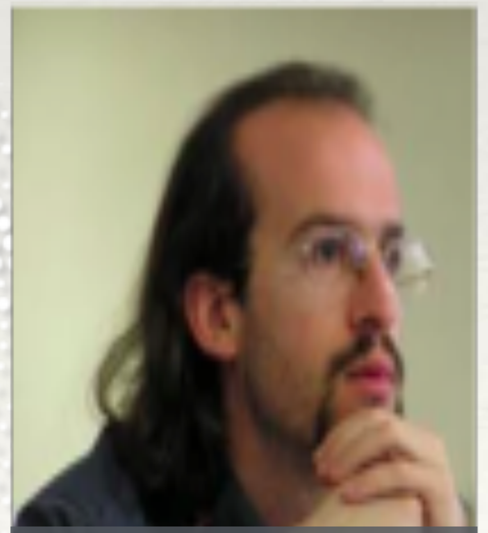

Giovanni Luigi Ciampaglia


Chengcheng Shao


Alessandro Flammini


Fil Menczer

CPT


Alain Barrat

Institute for Scientific Interchange Foundation


Daniela Paolotti