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

giancarlo.ruffo@unito.it

@giaruffo

di.unito.it

*Giancarlo **Ruffo***
*Alfonso **Semeraro***
Università degli Studi di Torino (Italy)

Research trends on misinformation

Computational approaches to study **fake news**, and **misinformation diffusion**

IEEE | Cosponsors ACM/ASA/CCF

DSAA 2020

The 7th IEEE International Conference on Data Science and Advanced Analytics

http://arcs.di.unito.it/2020/10/08/dsaatutorial/

6-9 October 2020
Sydney, Australia

Agenda

- Introduction and motivations
- Background and methods
- Problems and (some) answers
- Practical Lab (A. Semeraro)

Introduction and motivations

Prologue

ASTMH THE AMERICAN JOURNAL OF TROPICAL MEDICINE AND HYGIENE

ISSN: 0002-9637
E-ISSN: 1476-1645

OA COVID-19–Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis

Download

Md Saiful Is
Kamal, S. M.
Chowdhury,

Publisher: T
Source: The
Available or
Document T
DOI: https:/

Abstract Full-Text Ref

Infodemics, often including rumor
Monitoring social media data has
dispel misinformation and reduce
conspiracy theories in real time ar
conspiracy theories circulating on
newspapers, and their impacts on
descriptively analyzed. We perforn
other sources. We identified 2,311
Claims were related to illness, trar
disease including the origin (15%),
available, 1,856 claims were false (
serious implications on the individ
misinformation associated with th
debunk misinformation.

'Hundreds dead' because of Covid-19 misinformation

By Alistair Coleman
BBC Monitoring

At least 800 people may have died around the world because of coronavirus-related misinformation in the first three months of this year, researchers say.

Many died from drinking methanol or alcohol-based cleaning products.

They wrongly believing the products to be a cure for the virus.

However, the actual figure may never been known, as data from Iran - where many of the supposed methanol poisoning deaths occurred - is difficult to verify.

The World Health Organization (WHO) has previously said that the "infodemic" surrounding Covid-19 spread just as quickly as the virus itself, with conspiracy theories, rumours and cultural stigma all contributing to deaths and injuries.

What we 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

- We are not debunker, nor journalists
- We do not target social media as evil
  - We don't believe in censorship or freedom of speech limitations
  - We don't look for simple explanations to complex problems (e.g., *"gullible people are also stupid!"*)

The problem is worth studying

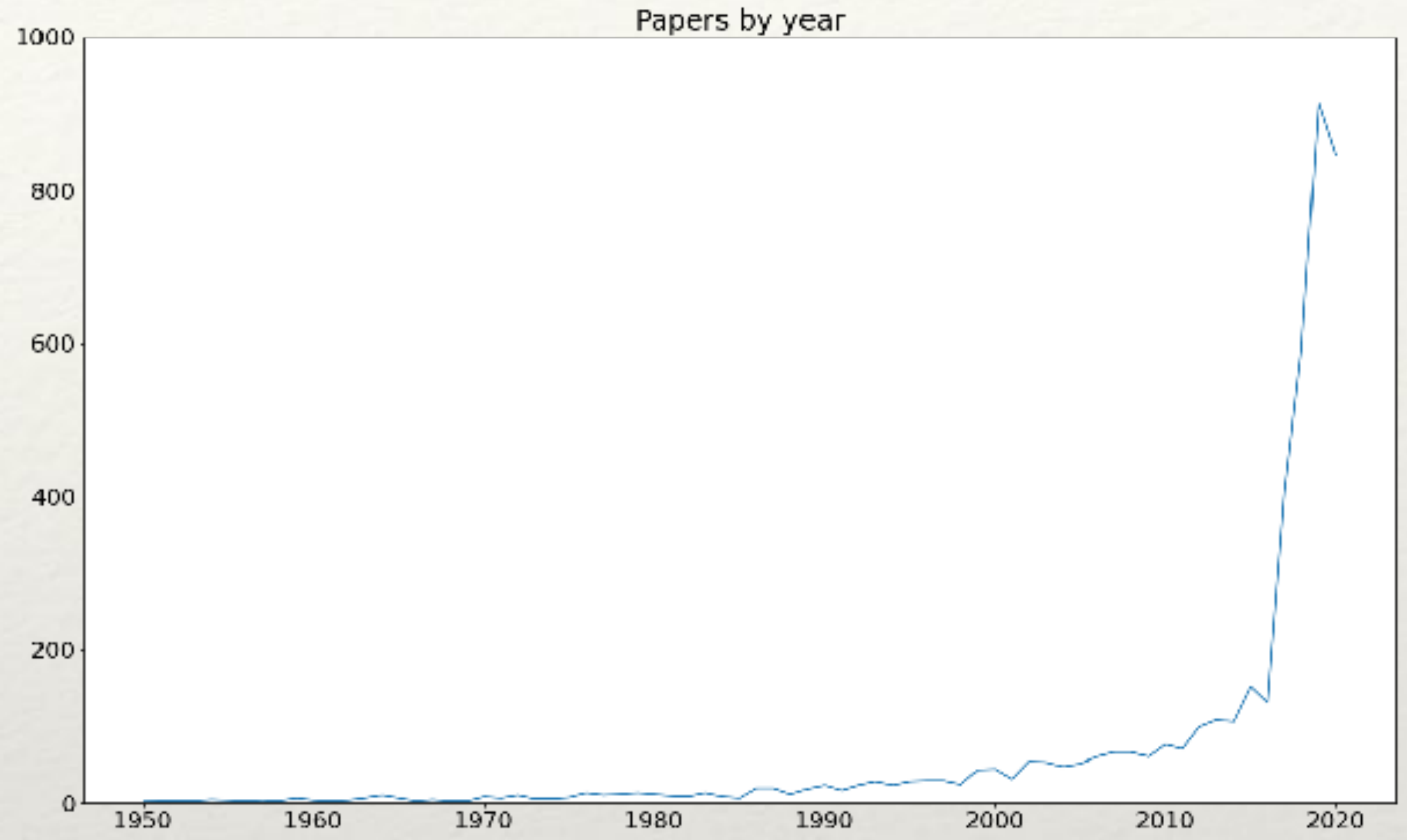
Prologue for Scholars

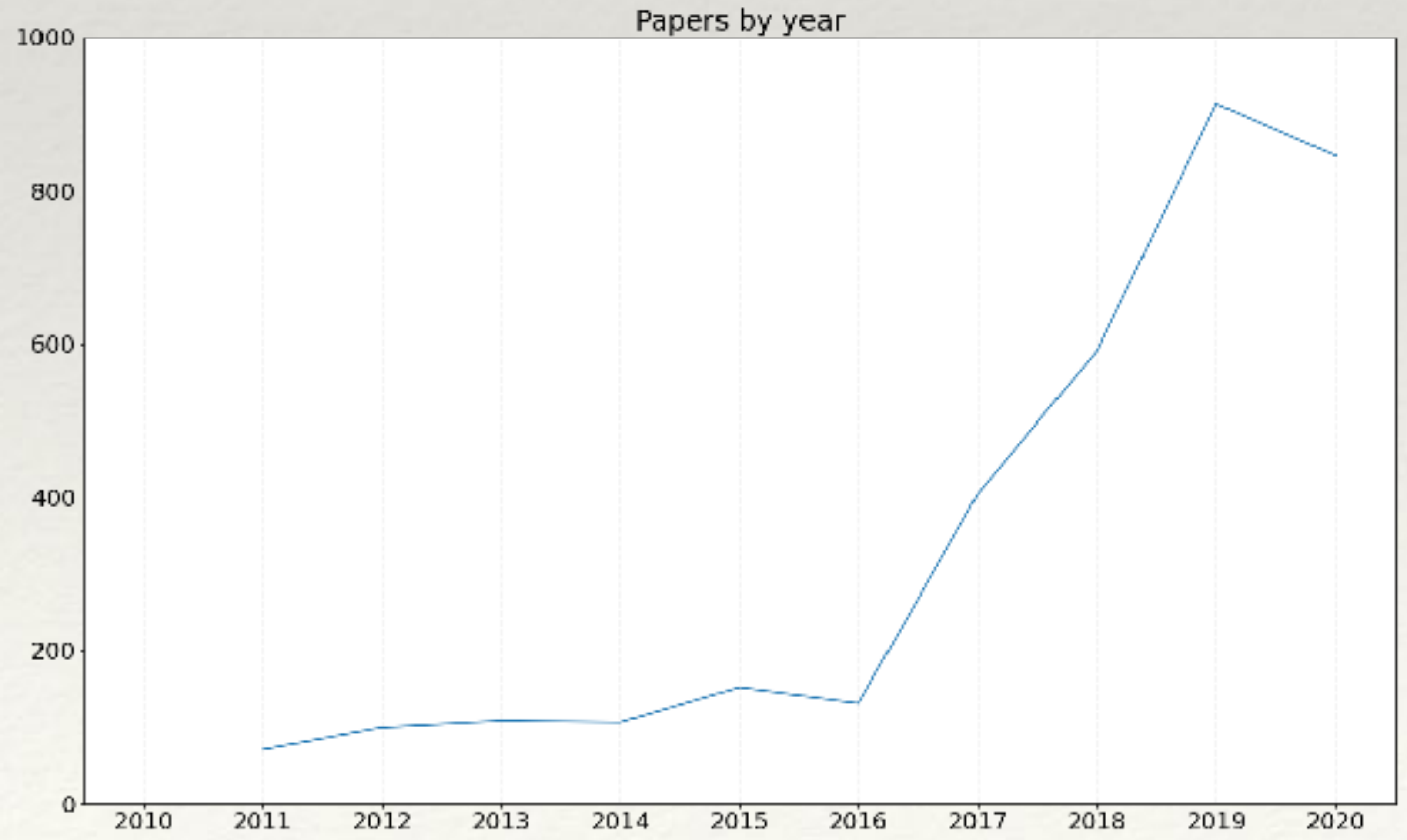
("Fake News" is a brand new field of study I wish to undertake)

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
- **4,258** papers (and counting)
- Explosive growth after 2016

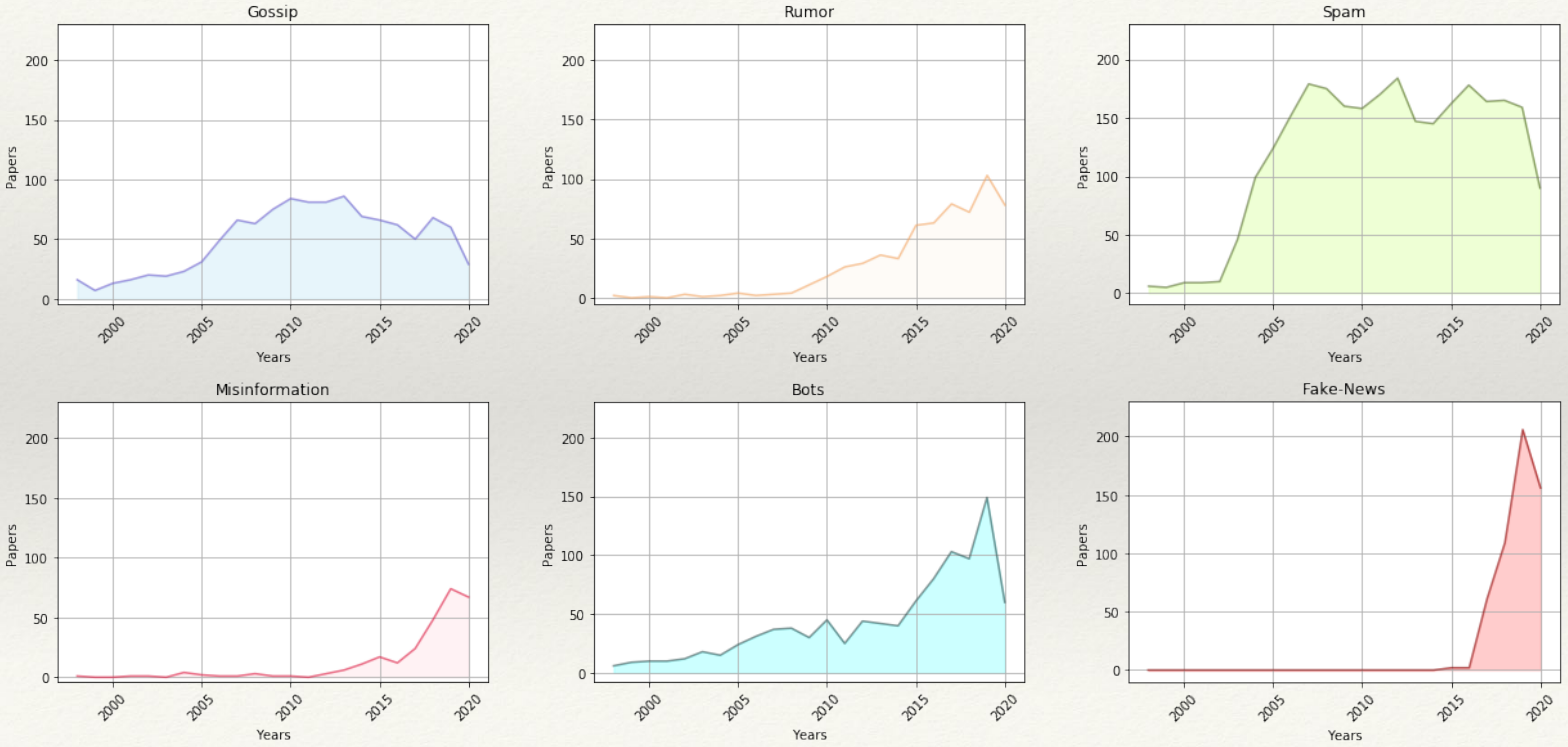
Google Trends

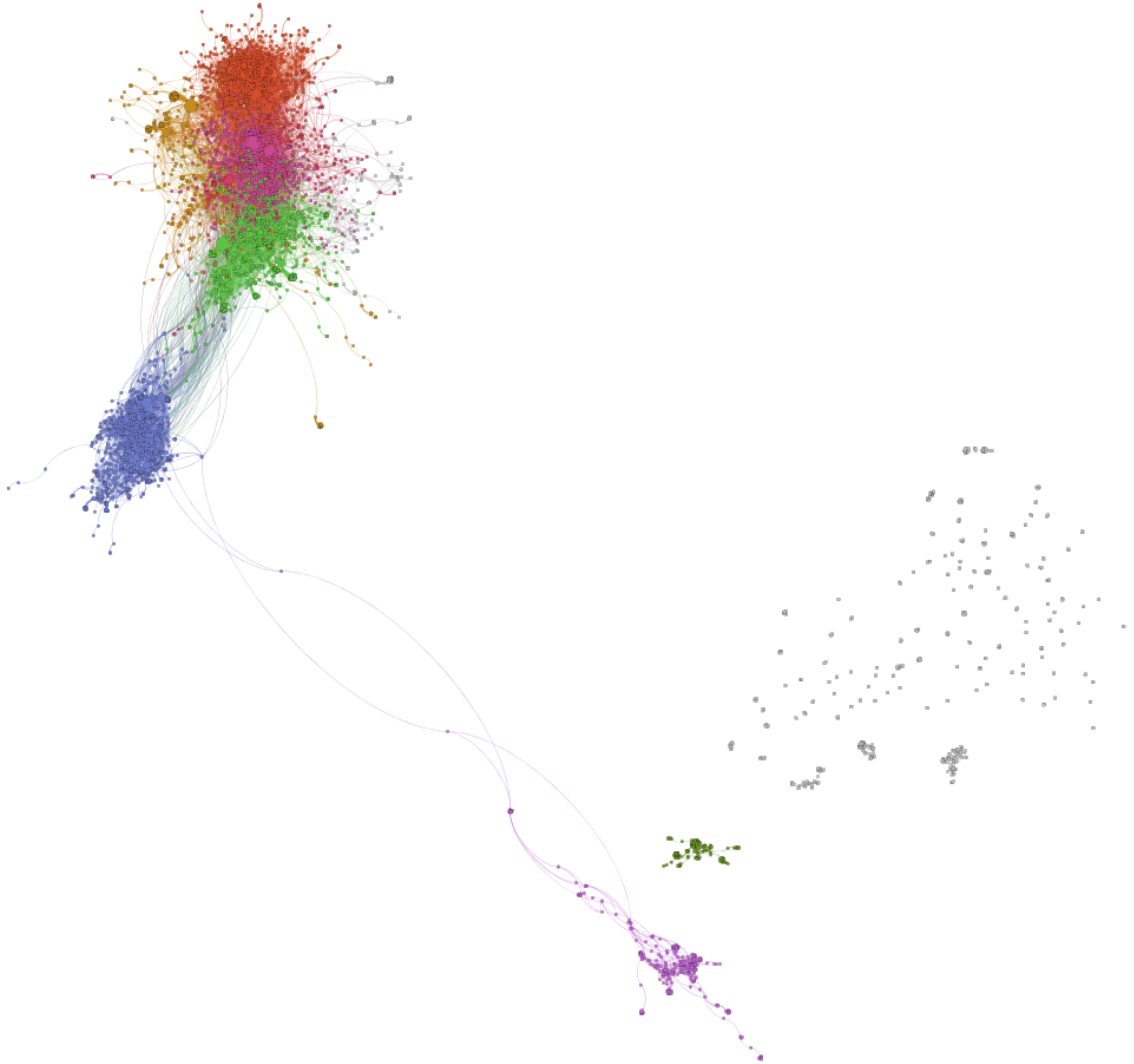




Updated: 2020, Oct. 6th

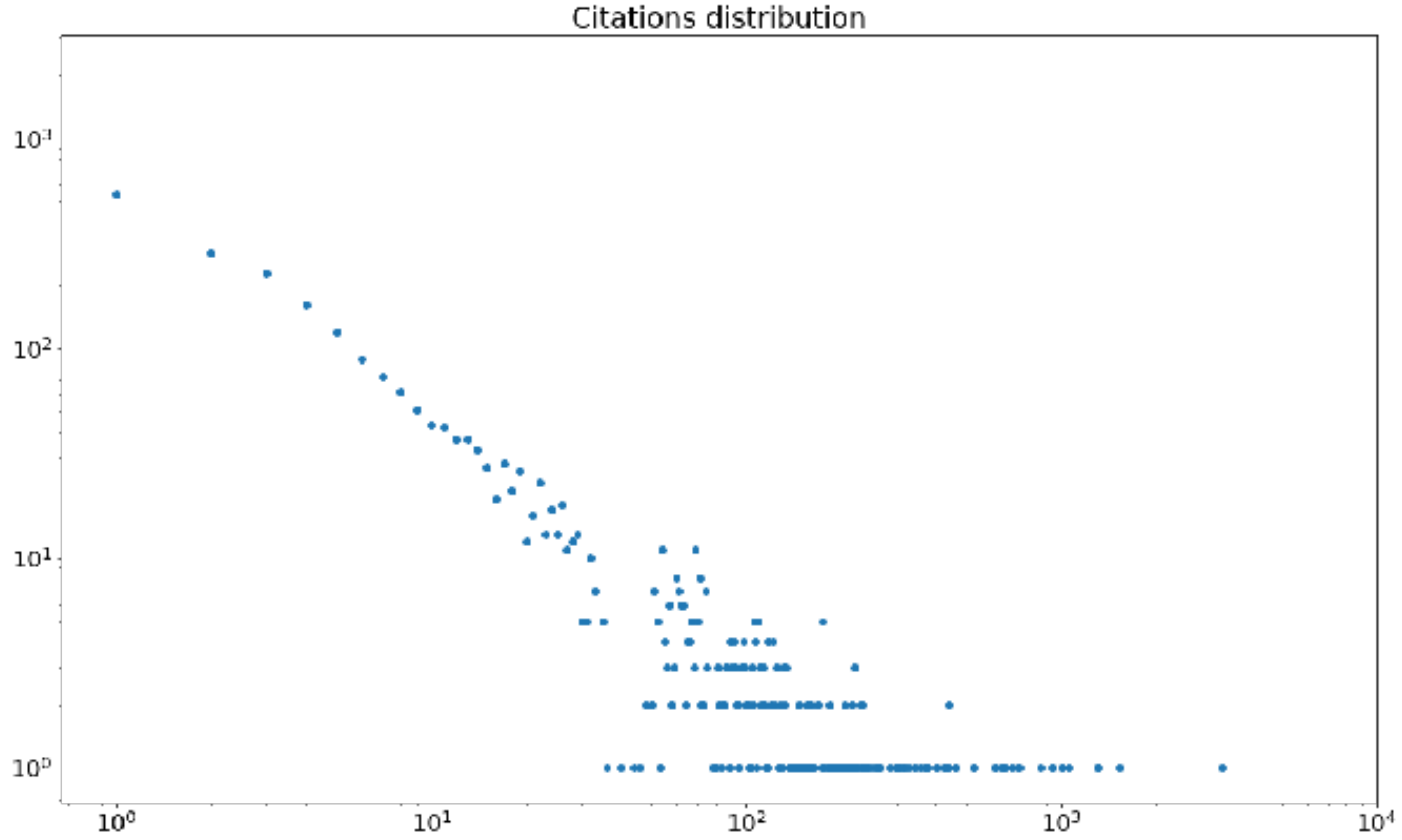
Focusing on DBLP only

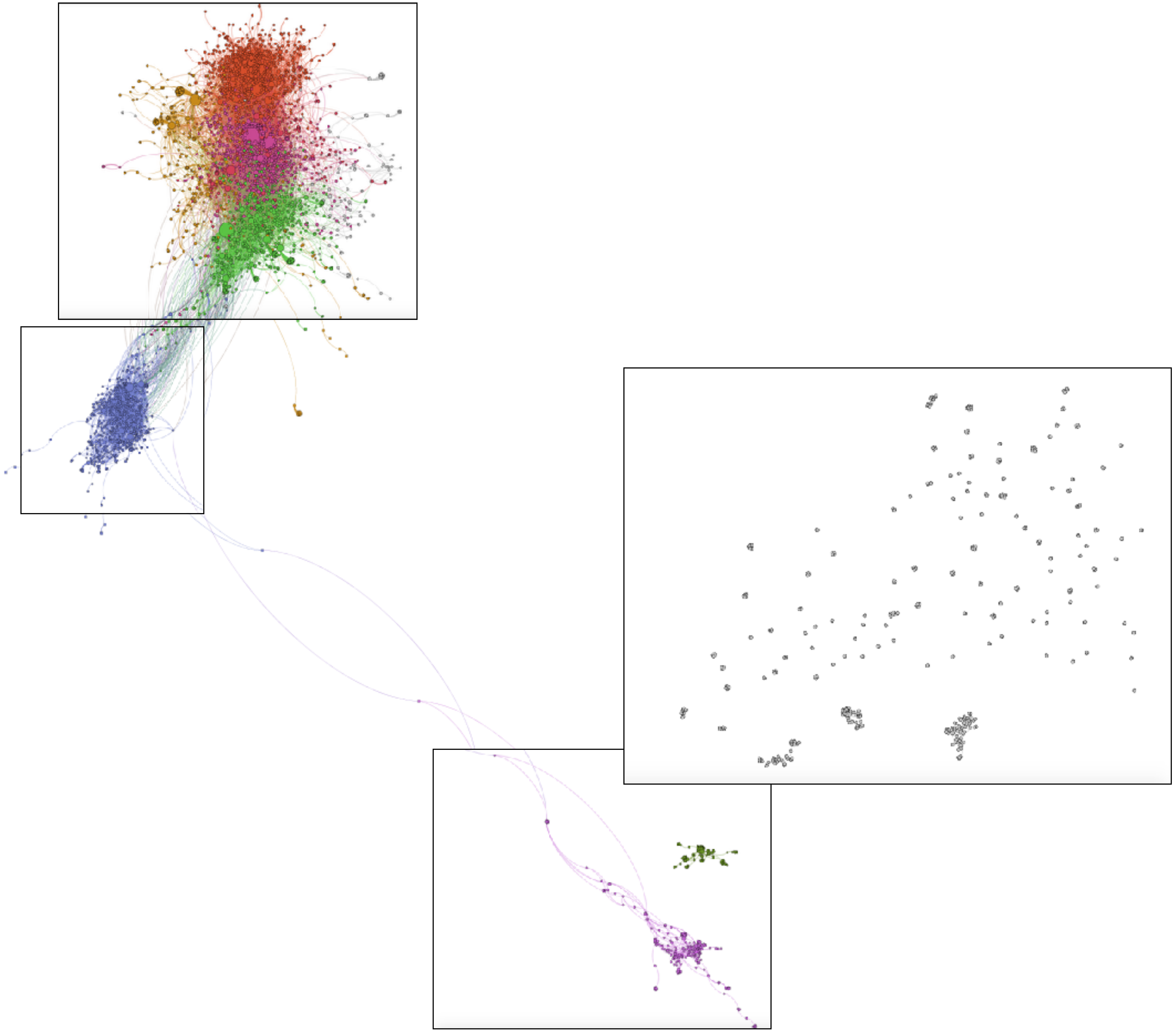


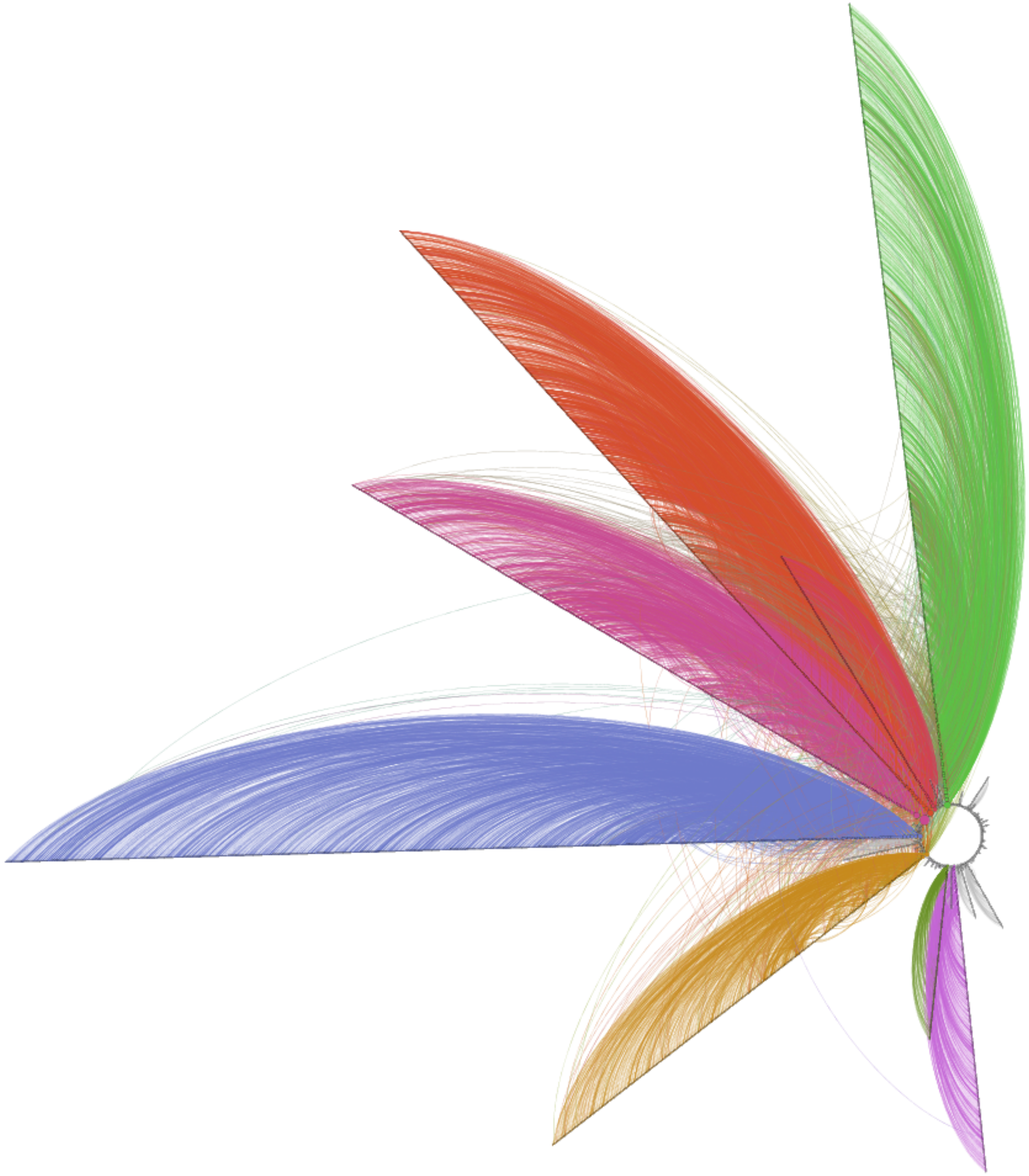


"fake news" citation network

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

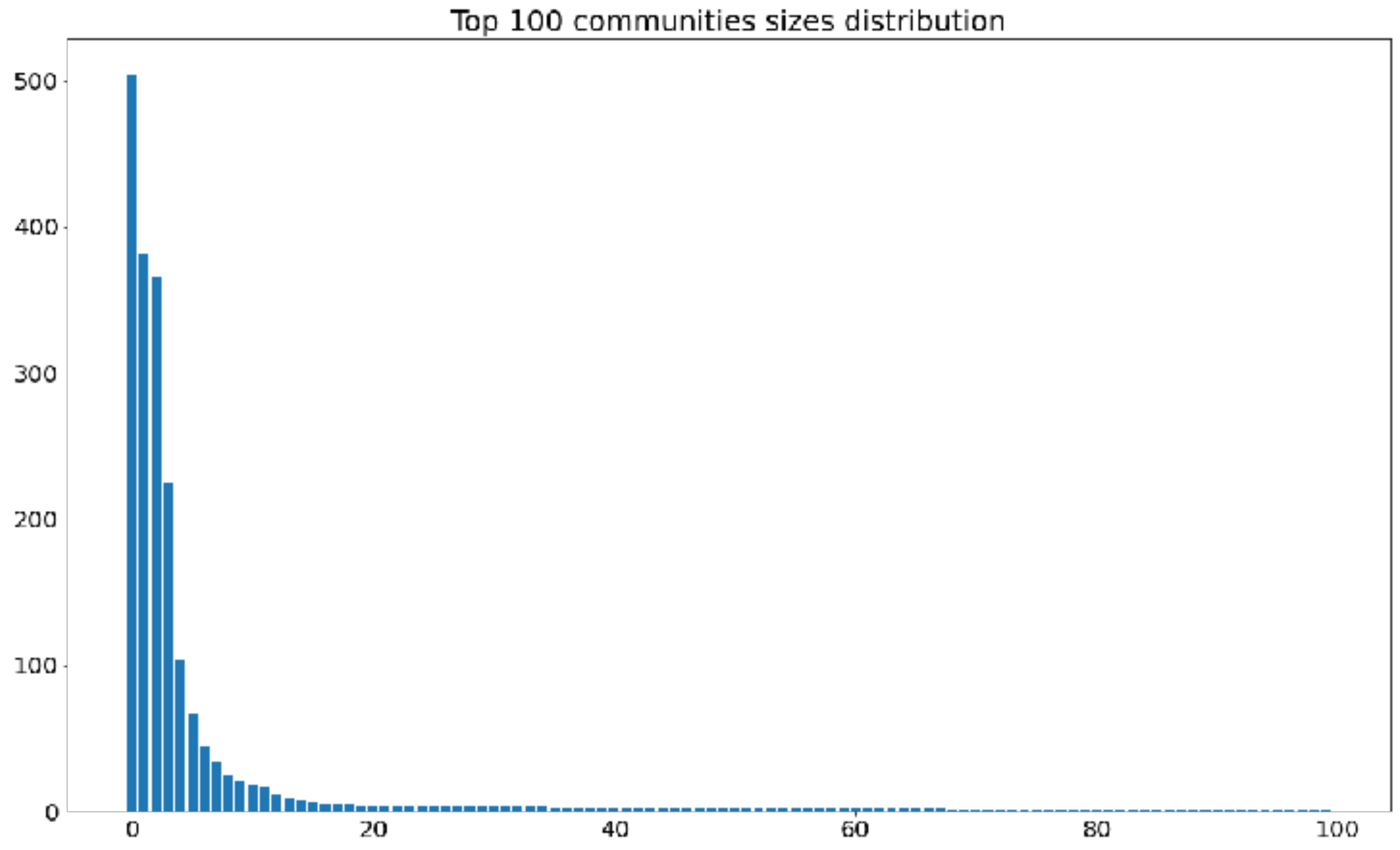






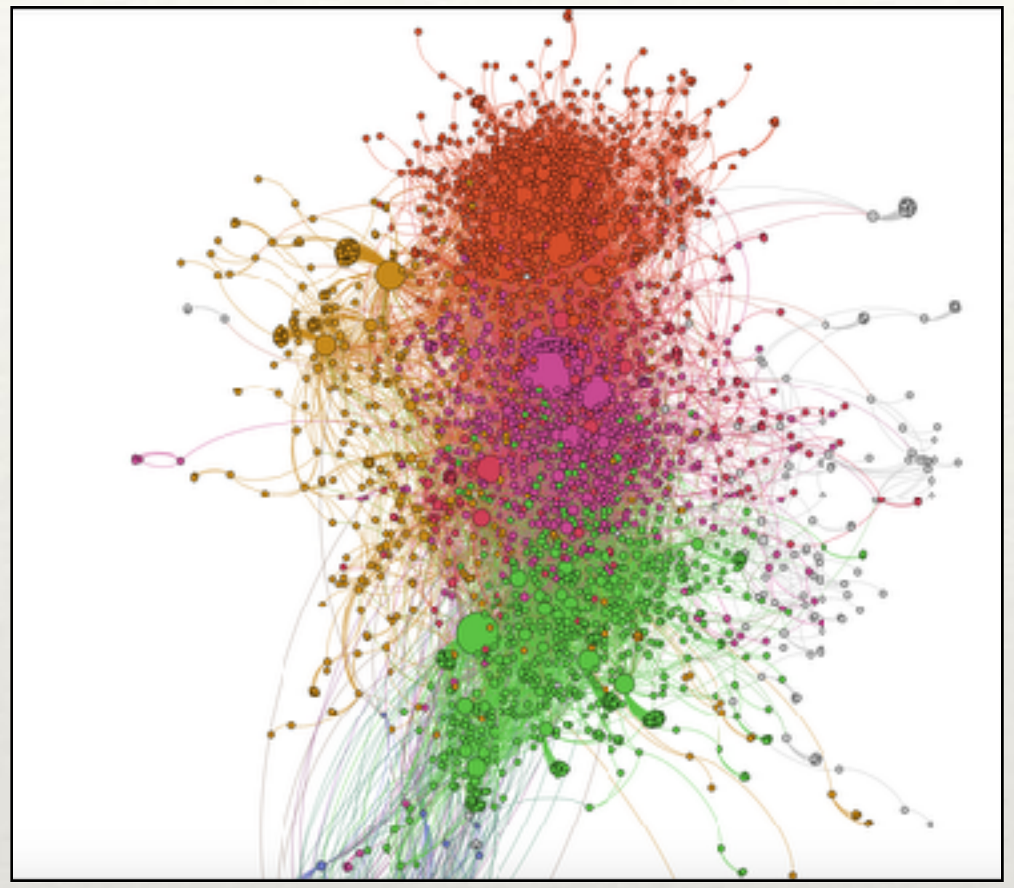
"fake news" citation network

- 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

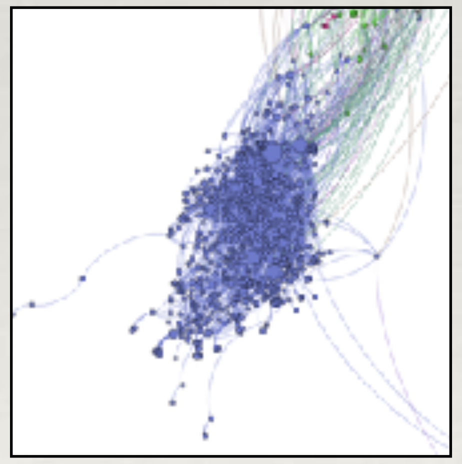


Data Mining
Neural Networks

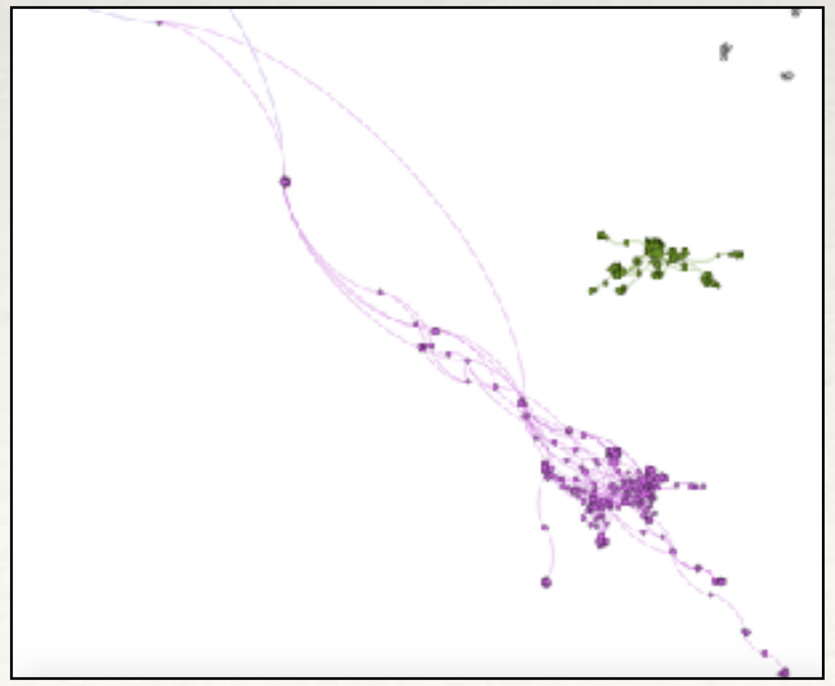
Identification
Rumors
Algorithms
Spreading

Spreading
Fact-checking
Networks
Social media data
Elections

Psychology
Influence
Media



Memory
Cognitive Biases



Health
Autism
Disorders
Vaccines

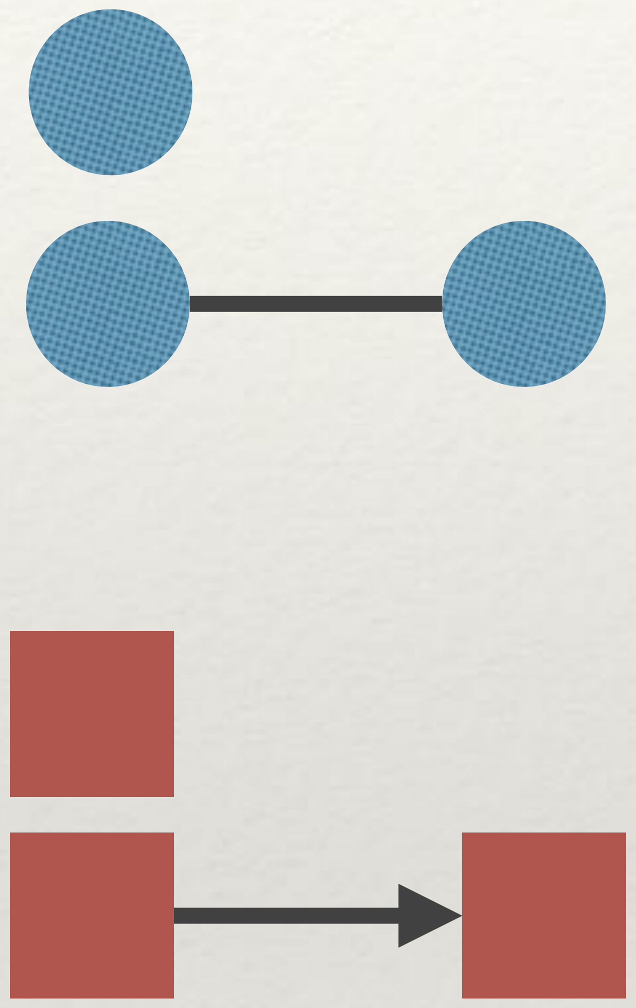
Huge literature (and very multidisciplinary)

Background and methods

Network Science

Networks are "everywhere"

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



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

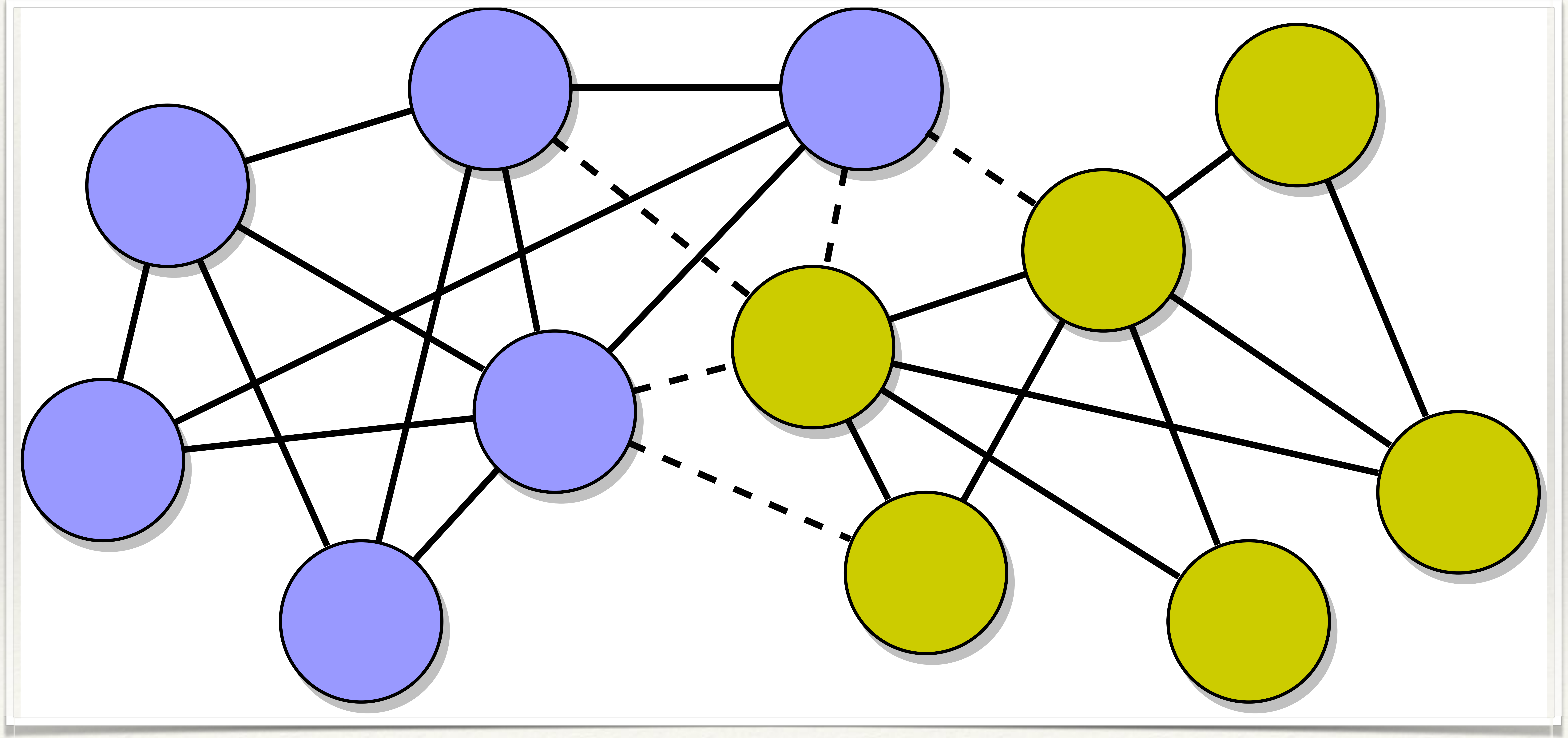
Epidemics

Social Contagion

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 charateristics
  - mutable characteristics

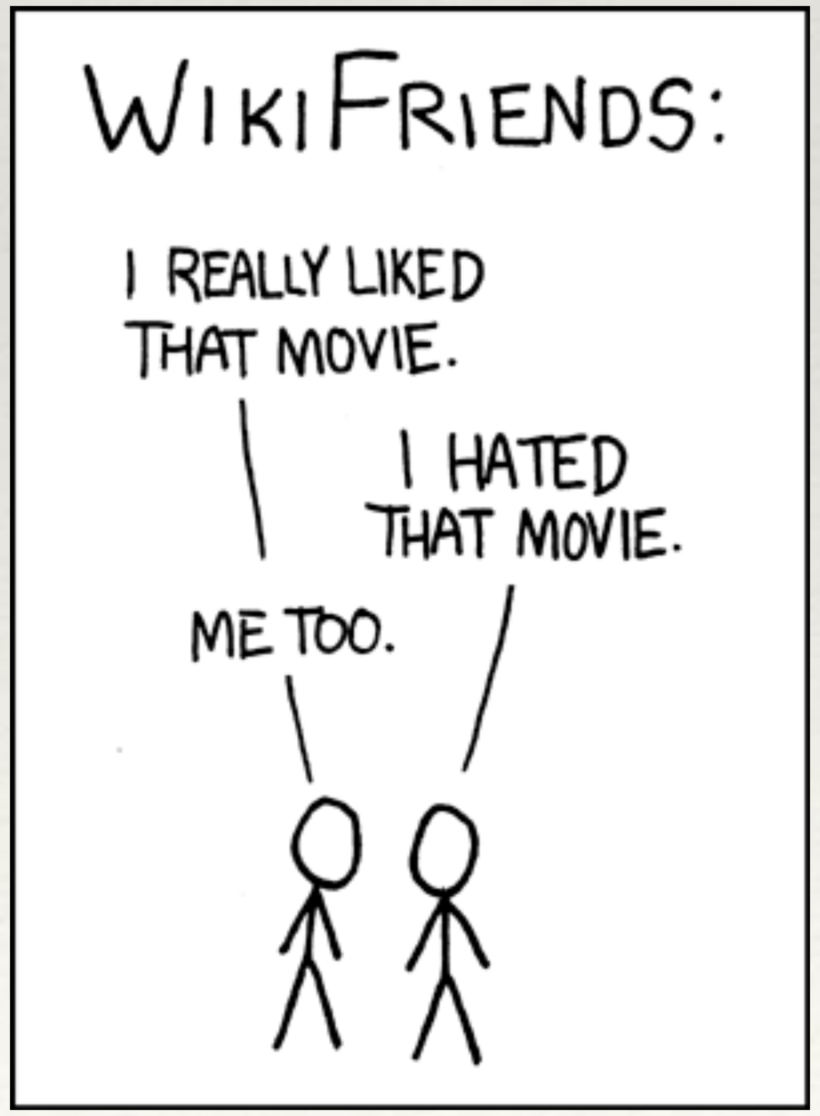
"Birds of a feather flock together"



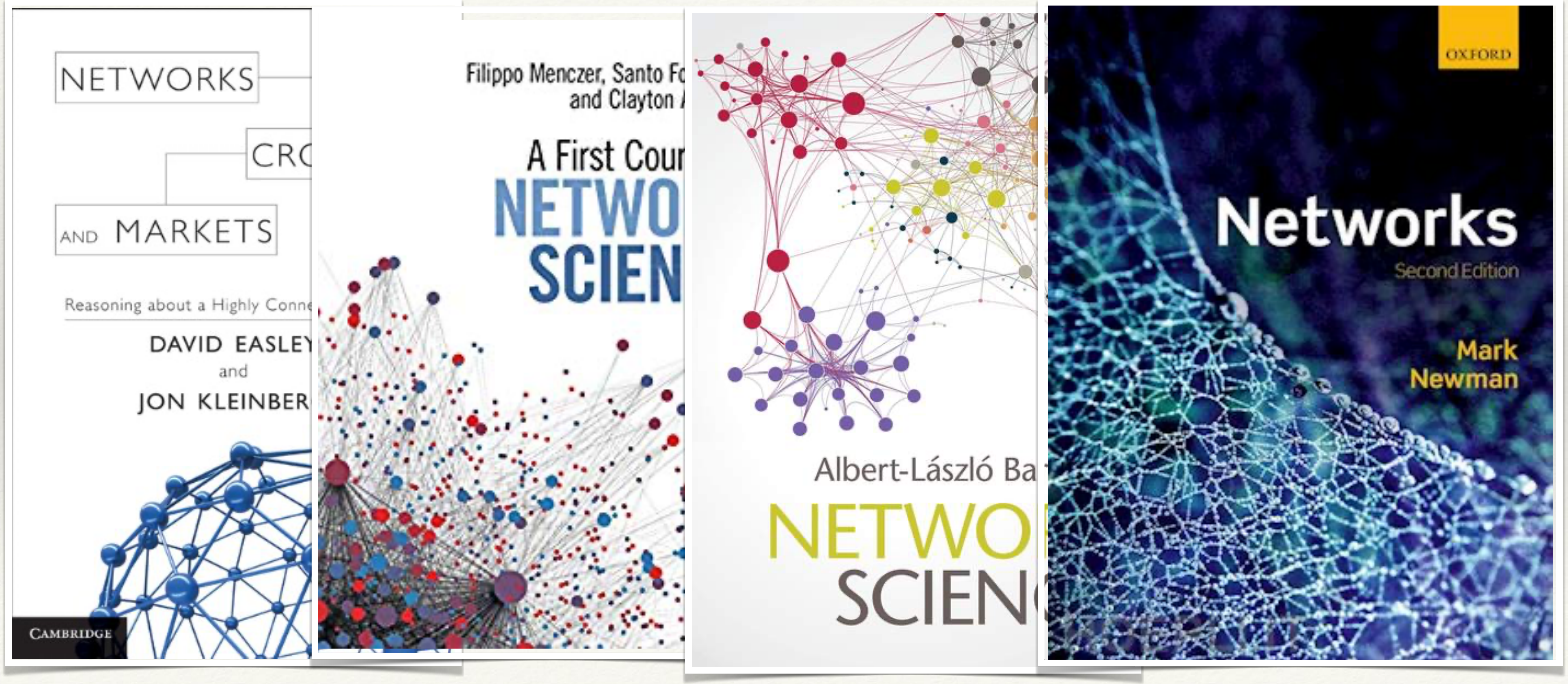
Picture from: F. Menczer, S. Fortunato, C. A. Davis, A First Course in Network Science, Cambridge University Press, 2020

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 similar
- It can also be a bad thing. For example "**echo chambers**" and "**groupthink**" are situations where your friends are like you, diversity is killed, and you are only exposed to opinions that reinforce your pre-existing beliefs…



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

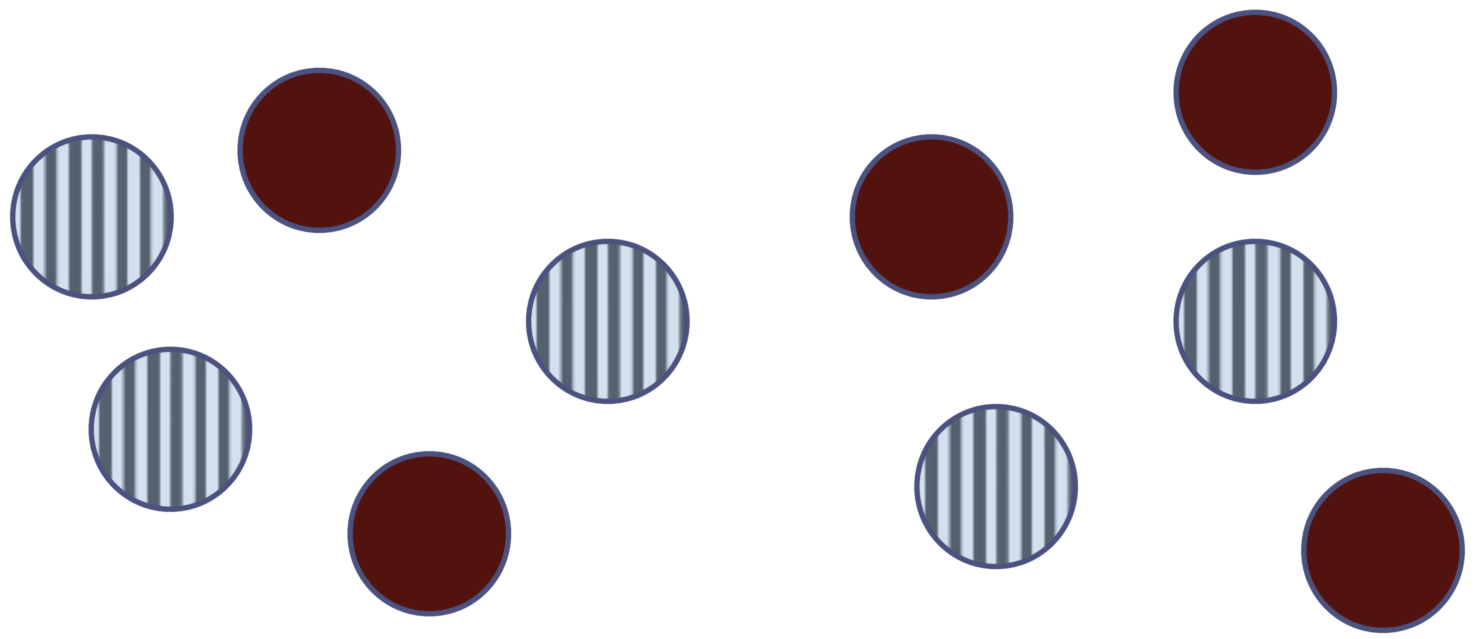
Segregation and Polarization

Segregation

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

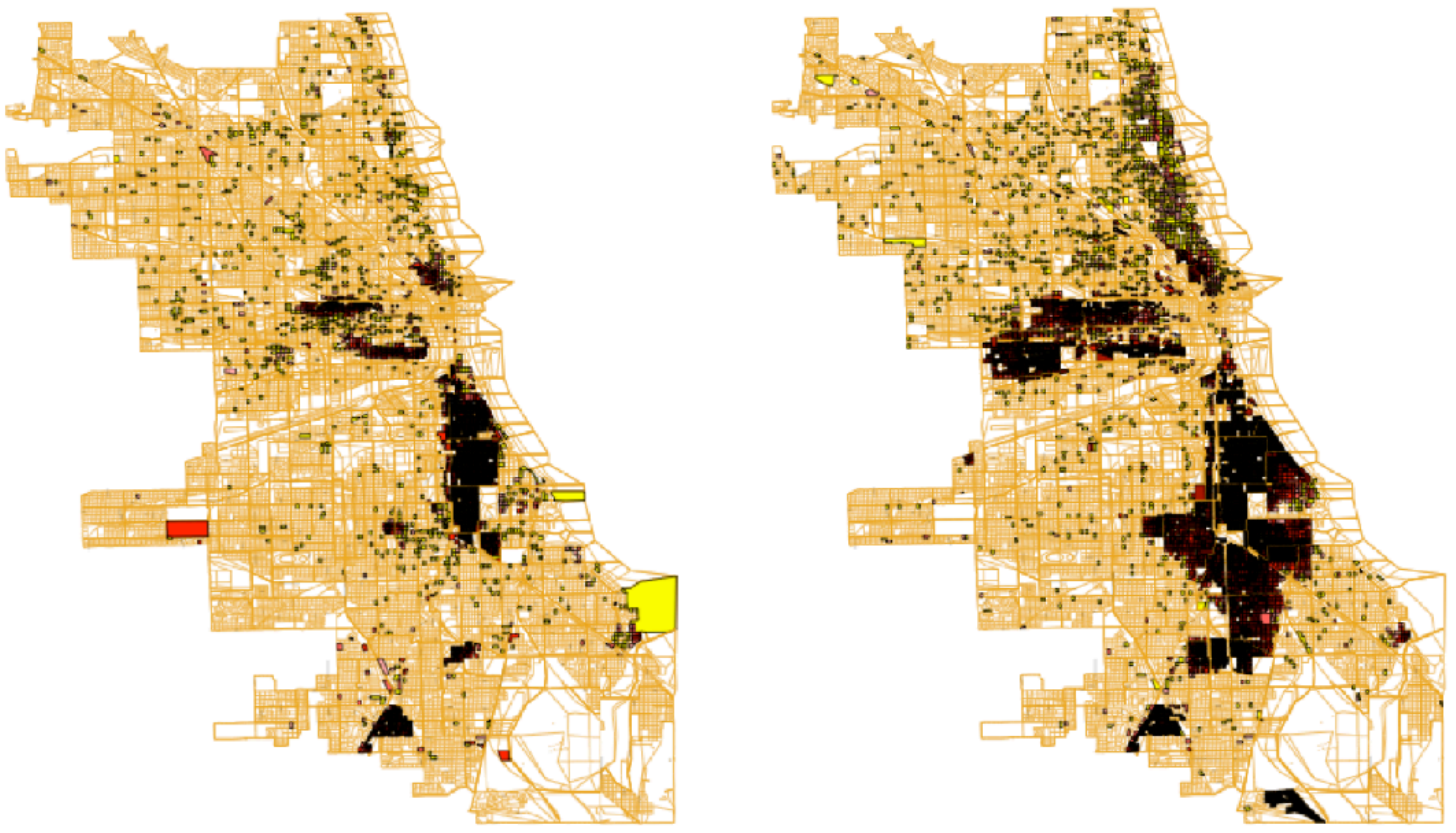
Segregation

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



Natural spatial "signature" in cities

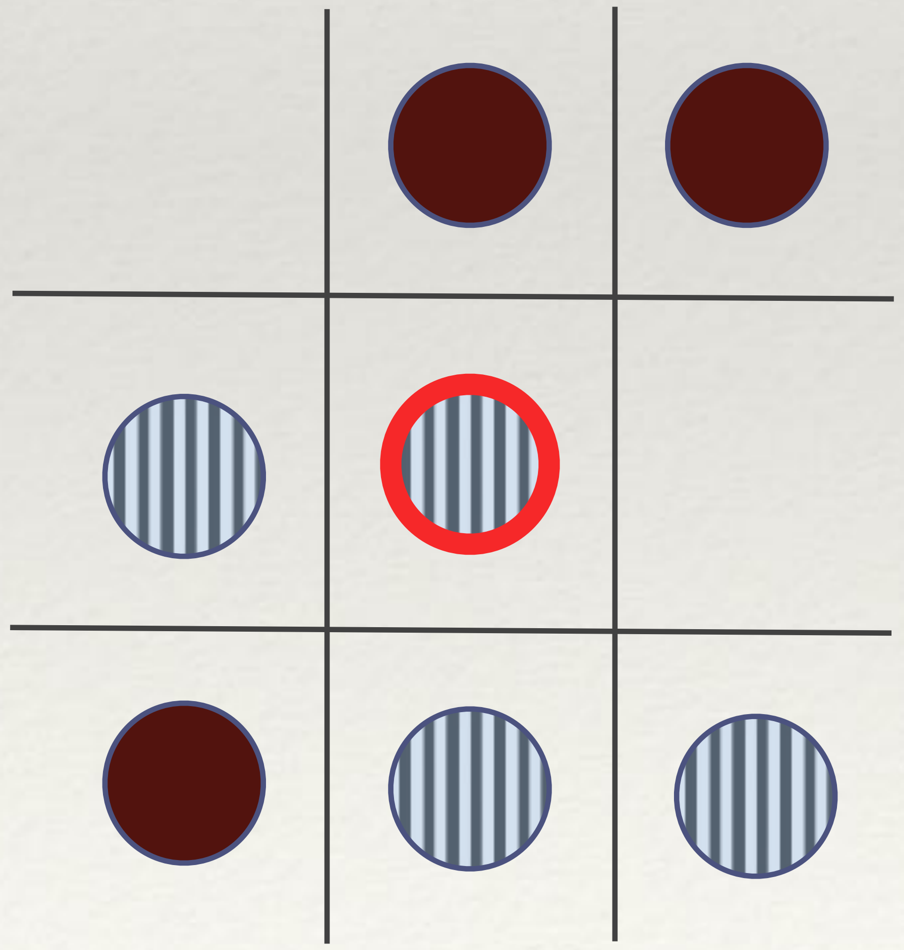
- Formation of homogeneous (according to some "type" or "class") neighbors in cities
- Which are the causes of "ghettization"?

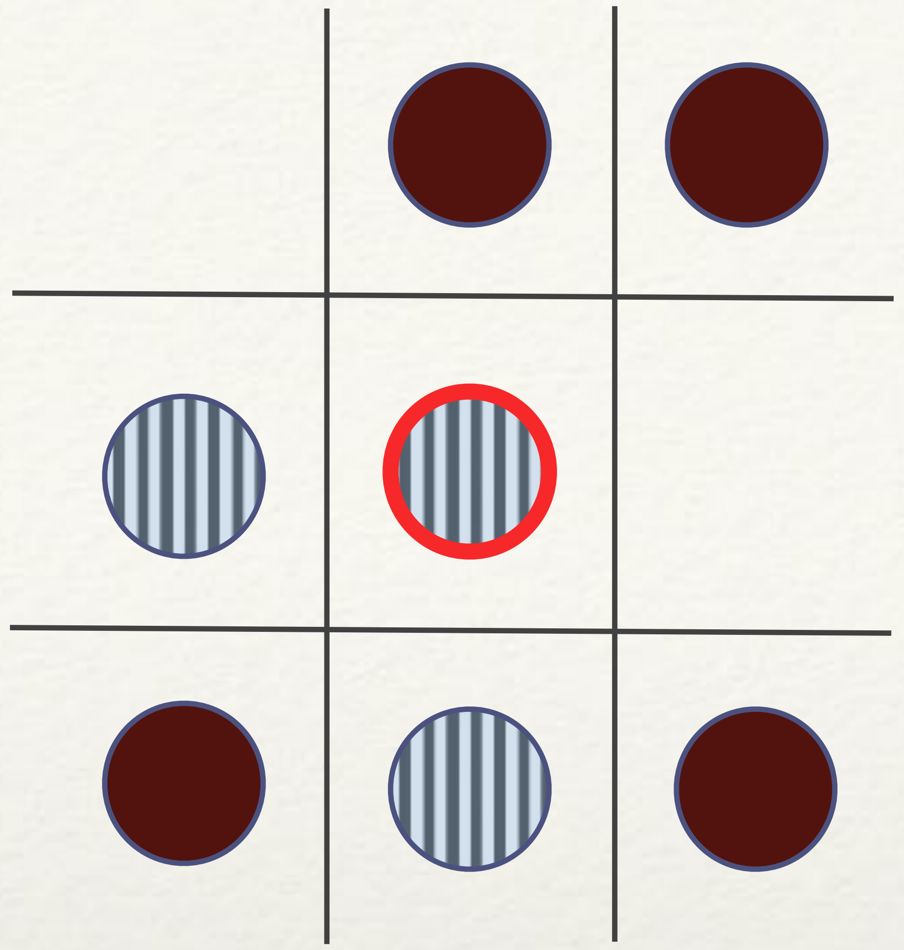


(a) *Chicago, 1940* (b) *Chicago, 1960*

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



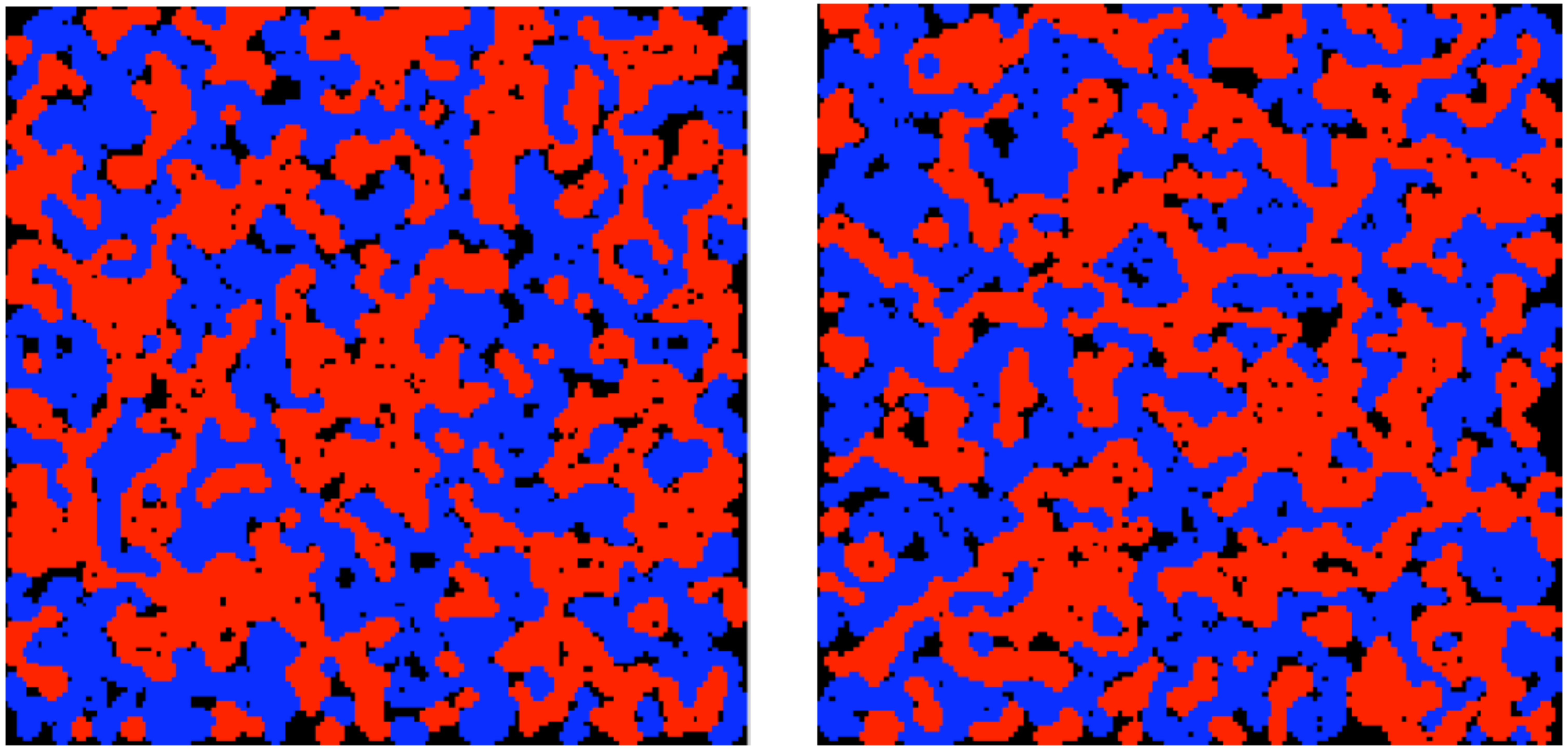


t = 3 => :-(

- 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

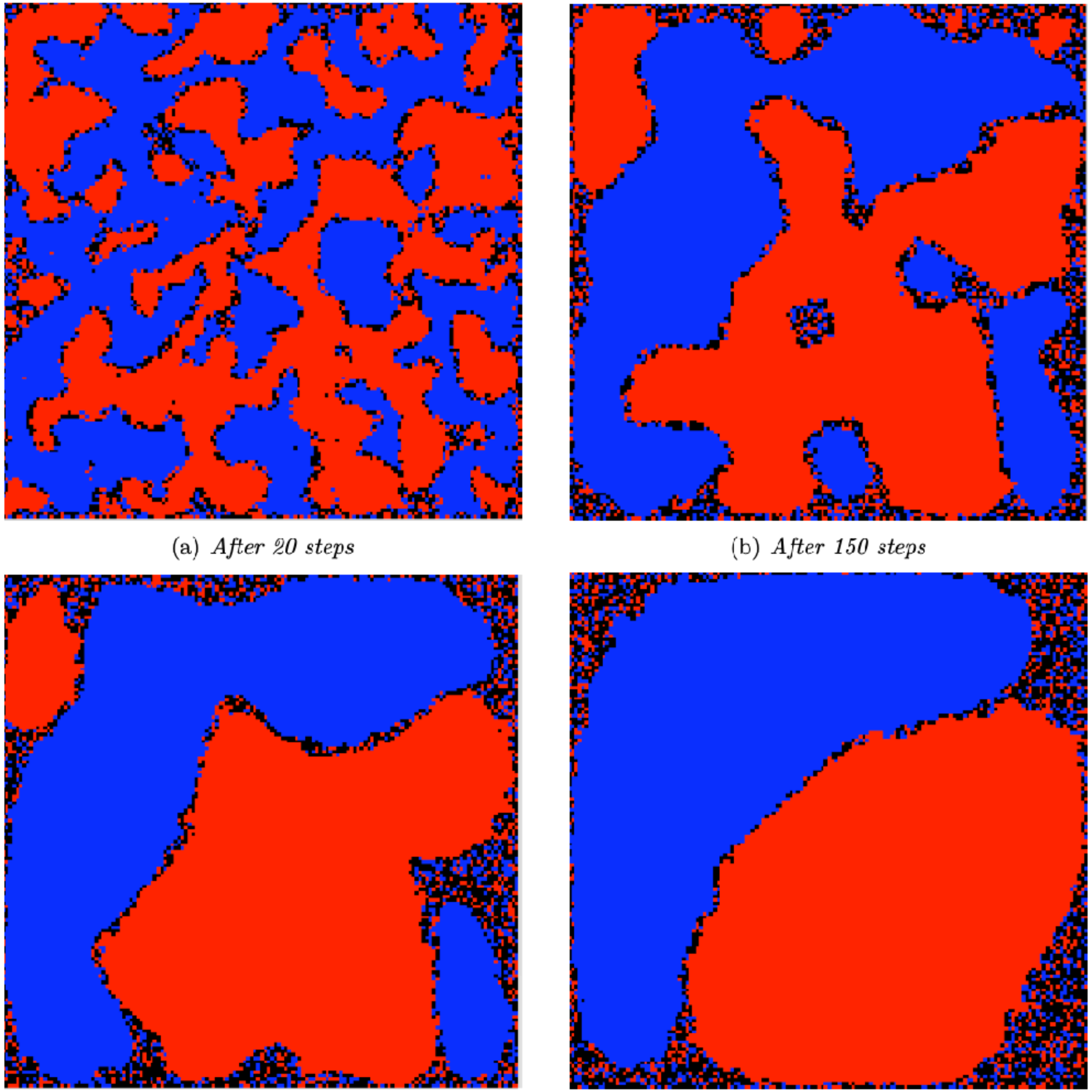
- Computer simulations to look for patterns at larger scale
- We want to run different simulations and make some comparisons
  => integrated pattern?
- on the right: two runs of a simulations of the Schelling model with a threshold *t* of *3*
  - 150x150 grid
  - 10,000 agents



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

Segregation emerges even when agents accept to be a minority!

$t > 3 =>$

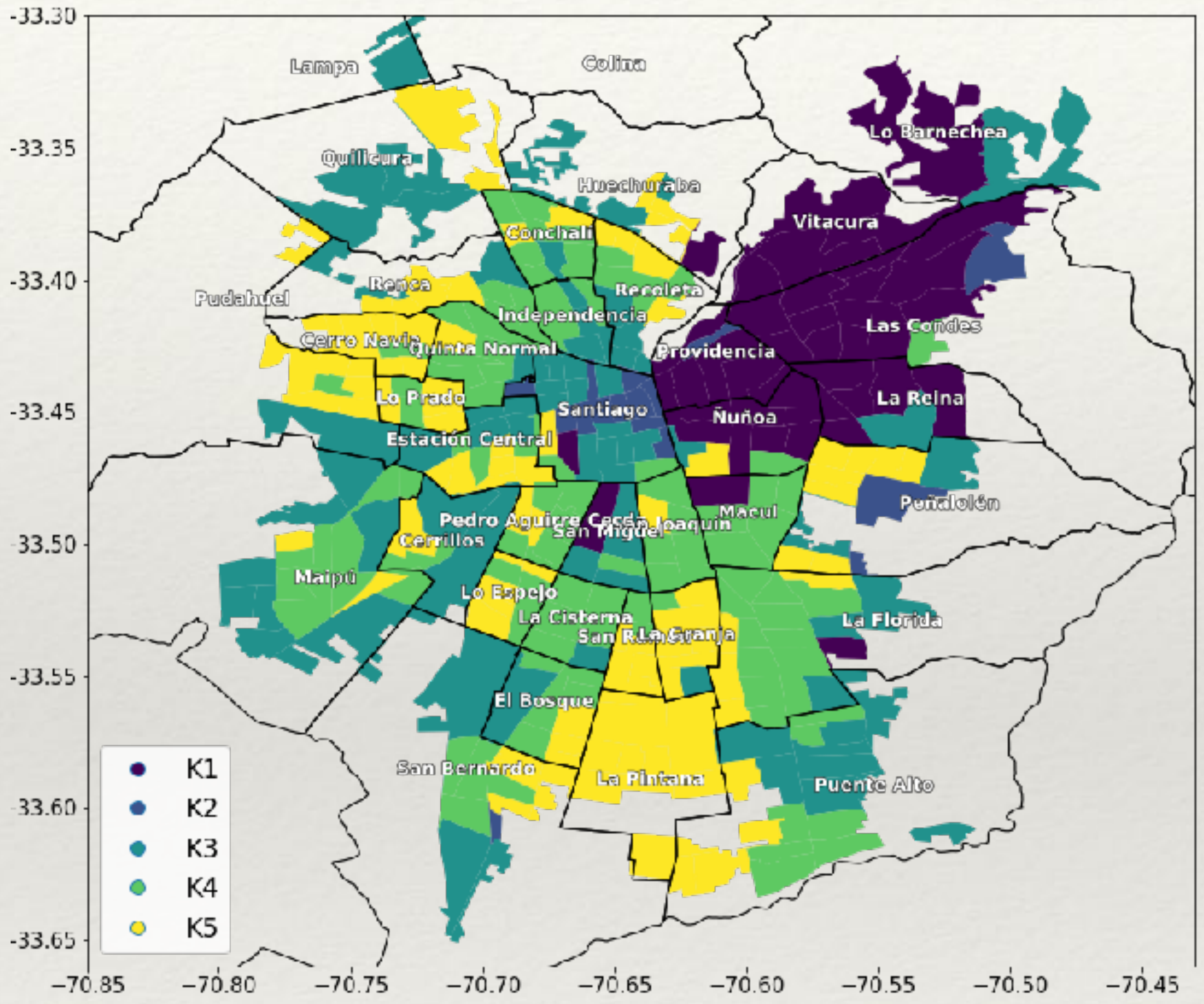


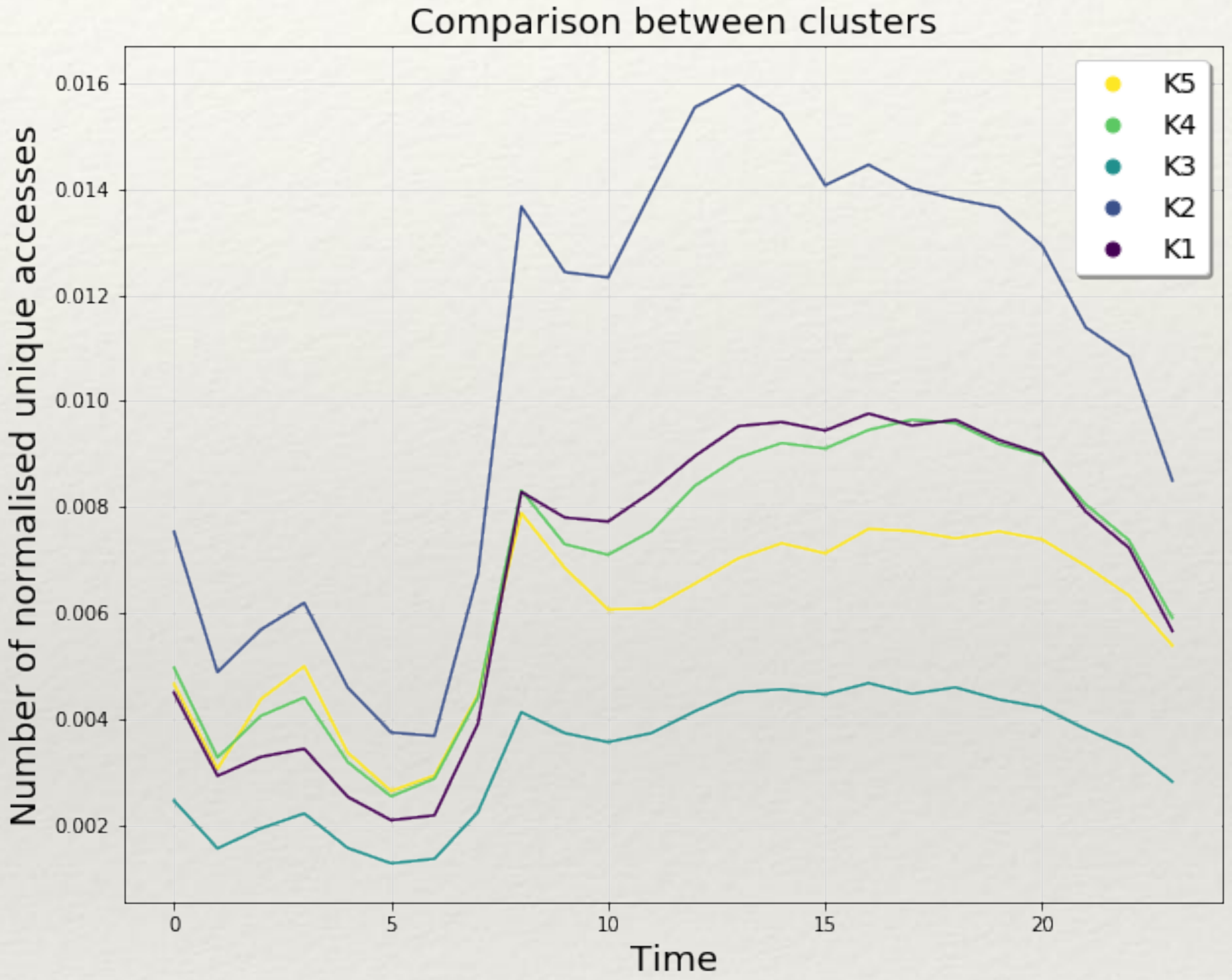
(a) *After 20 steps* (b) *After 150 steps* (c) *After 350 steps* (d) *After 800 steps*

Segregation is (trivially) amplified in an intolerant society

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.





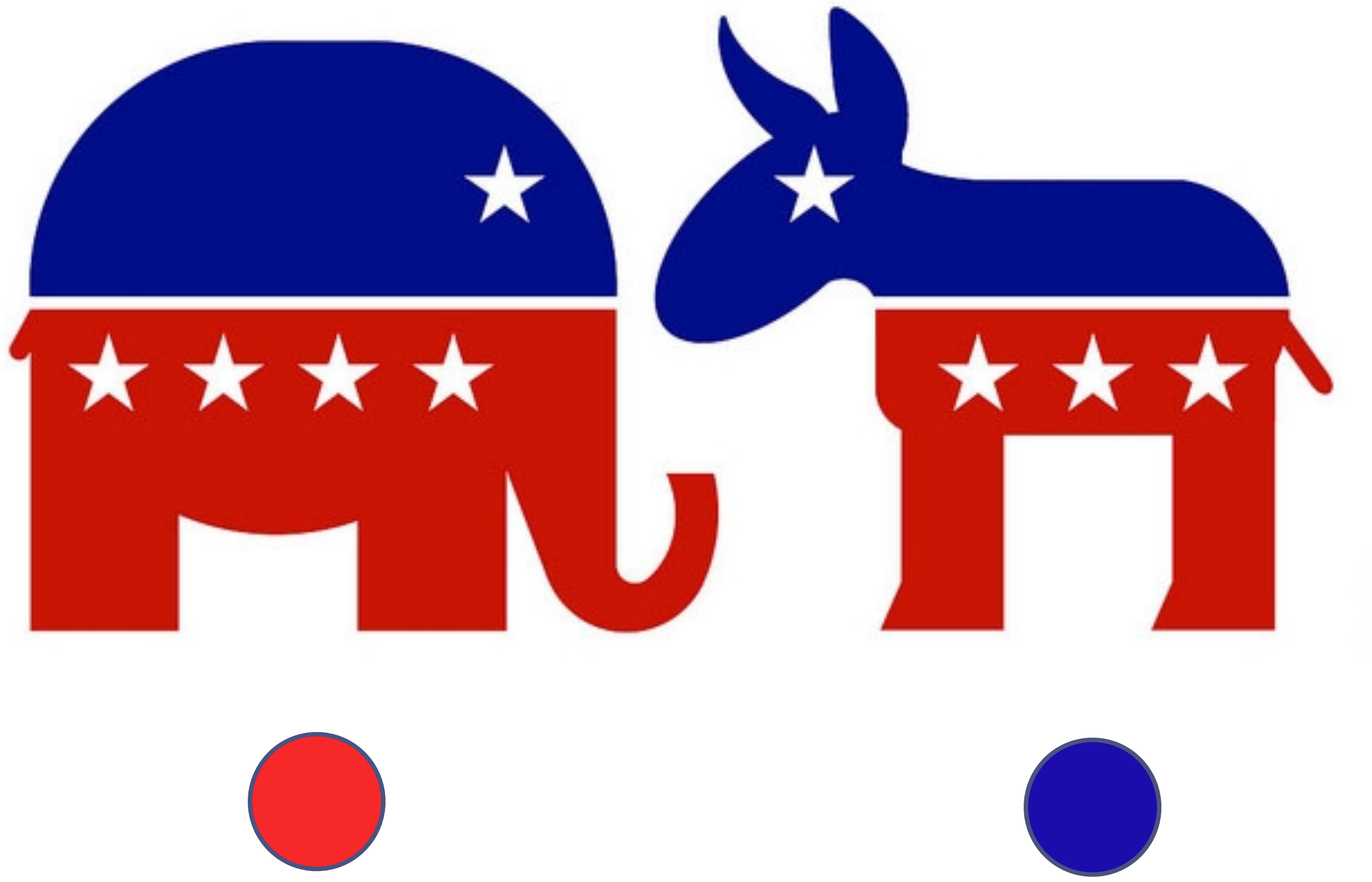
S Vilella, D Paolotti, G Ruffo, L Ferres, News and the city: understanding online press consumption patterns through mobile data, EPJ Data Sci. 9, 10 (2020). https://doi.org/10.1140/epjds/s13688-020-00228-9

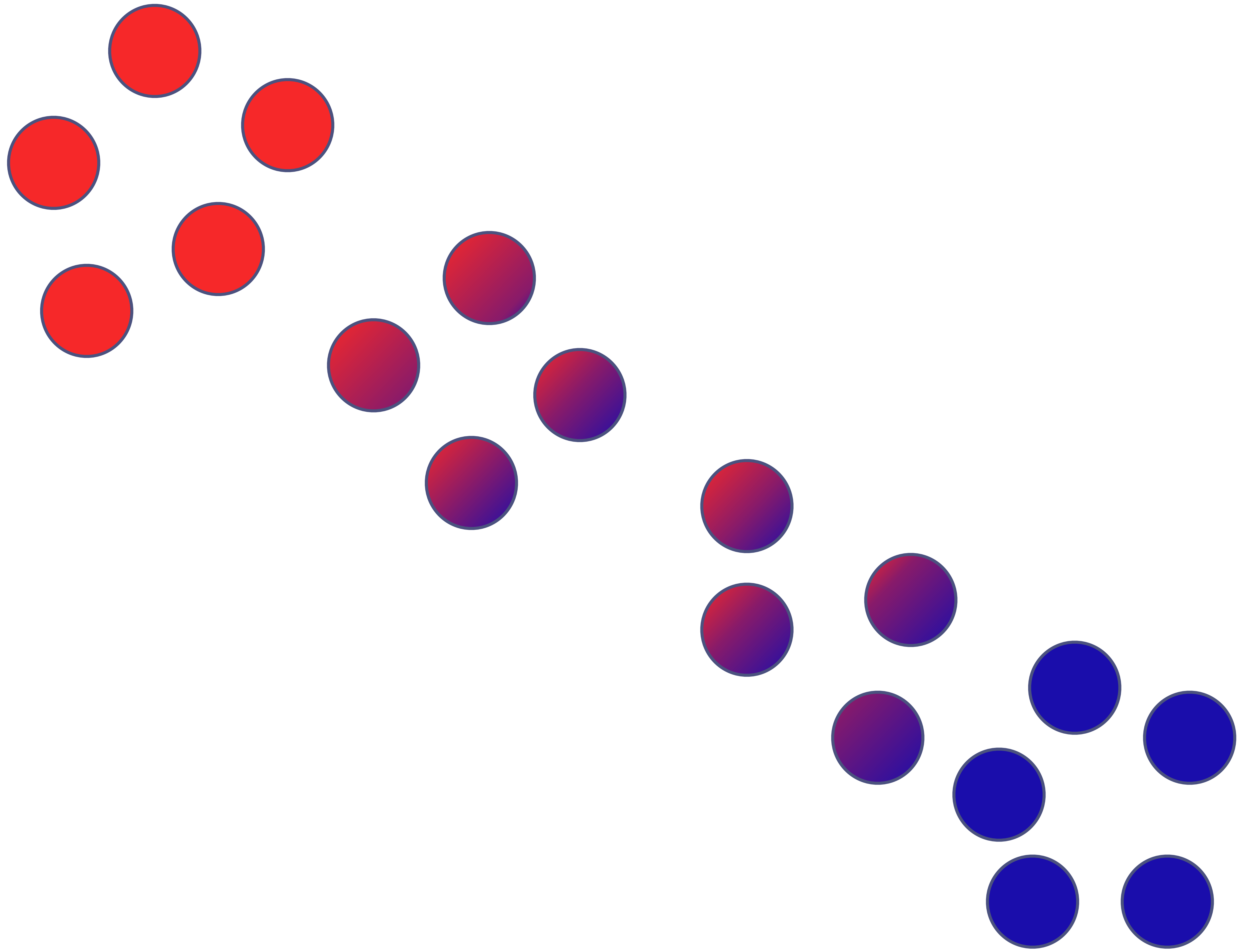
We live in a naturally segregated (or clustered) society - if social media amplifies or reduces such segregation is still a research question

*"**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





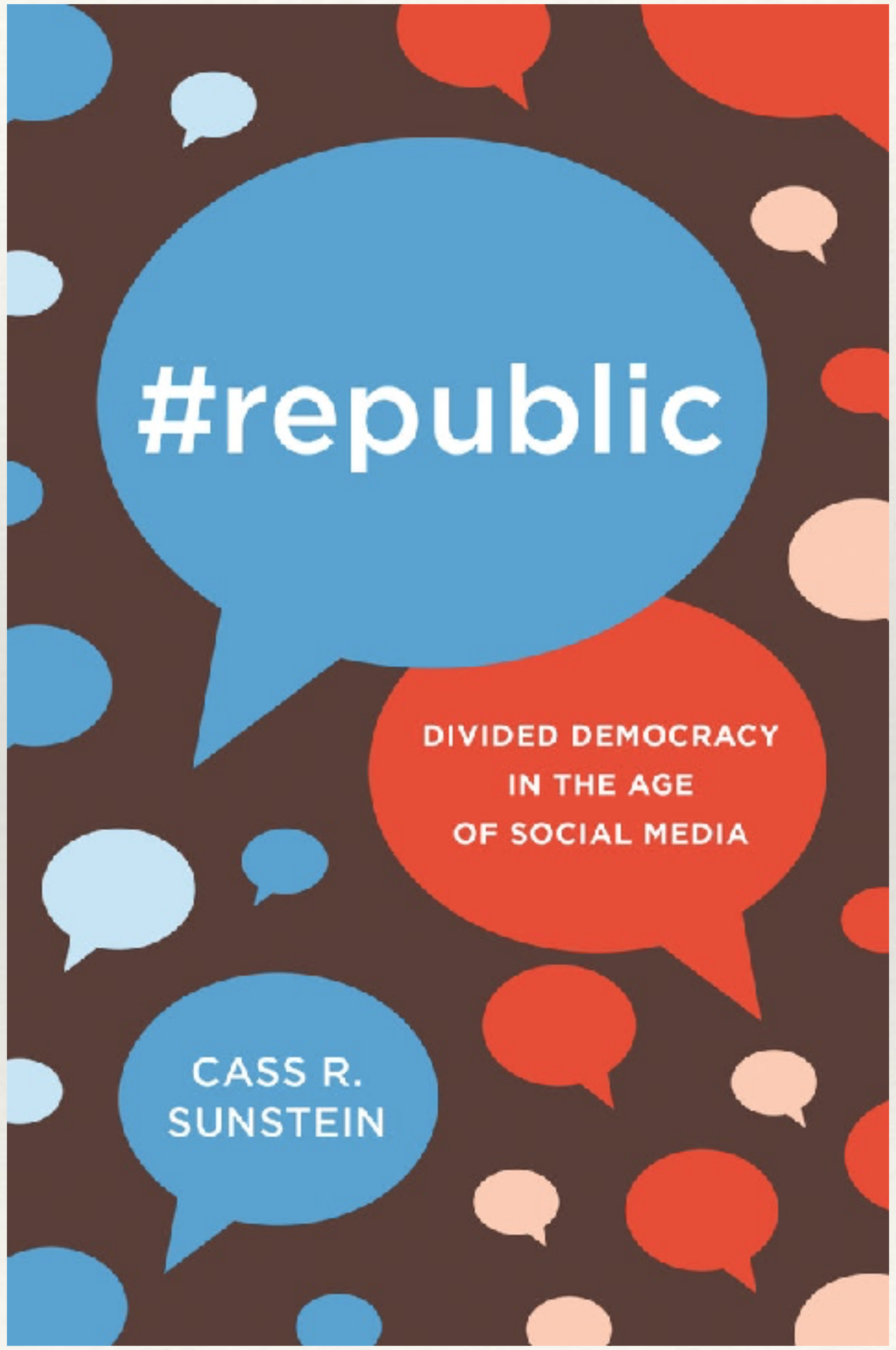
Polarization

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…

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

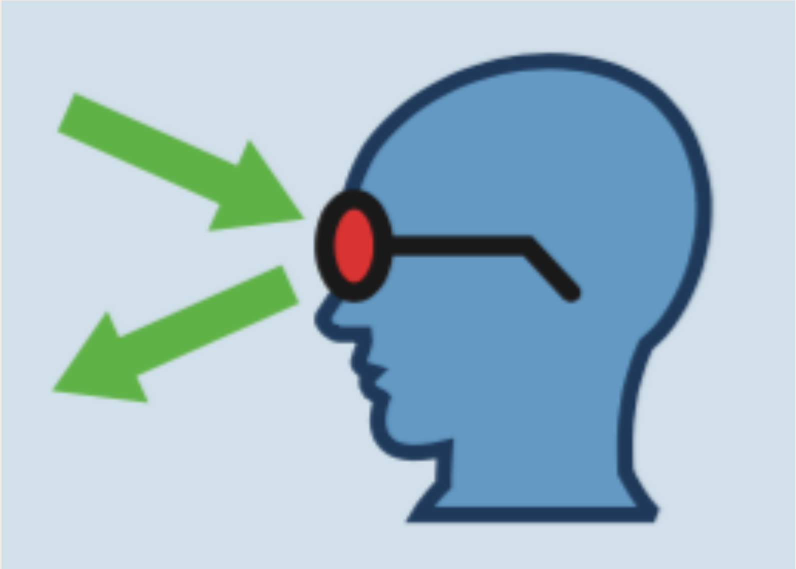


Psychological issues

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.

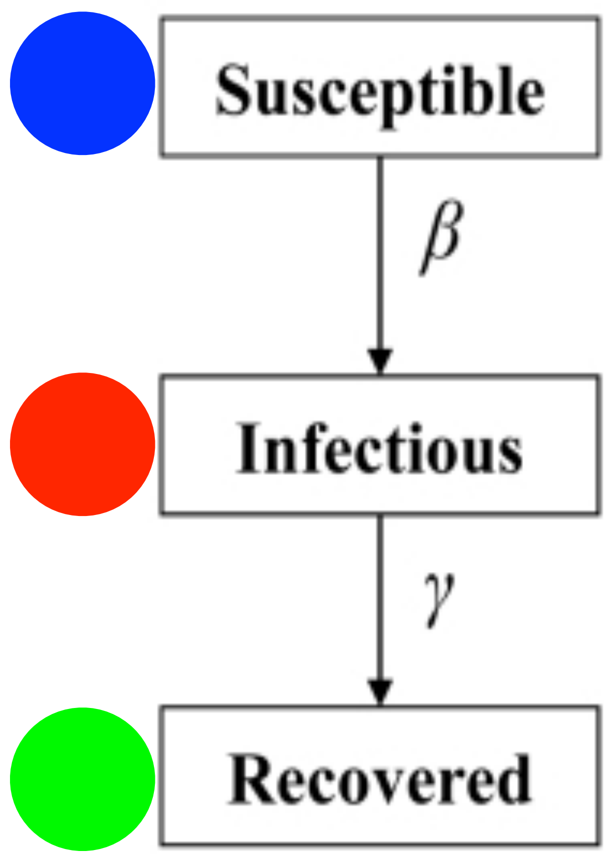
Polarization emerges from radicalized segregation, but not necessarily a segregated network is also polarized.
However, some topics are strongly divisive (echo-chambers), others are not.

Contagion

Different kinds of contagion

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

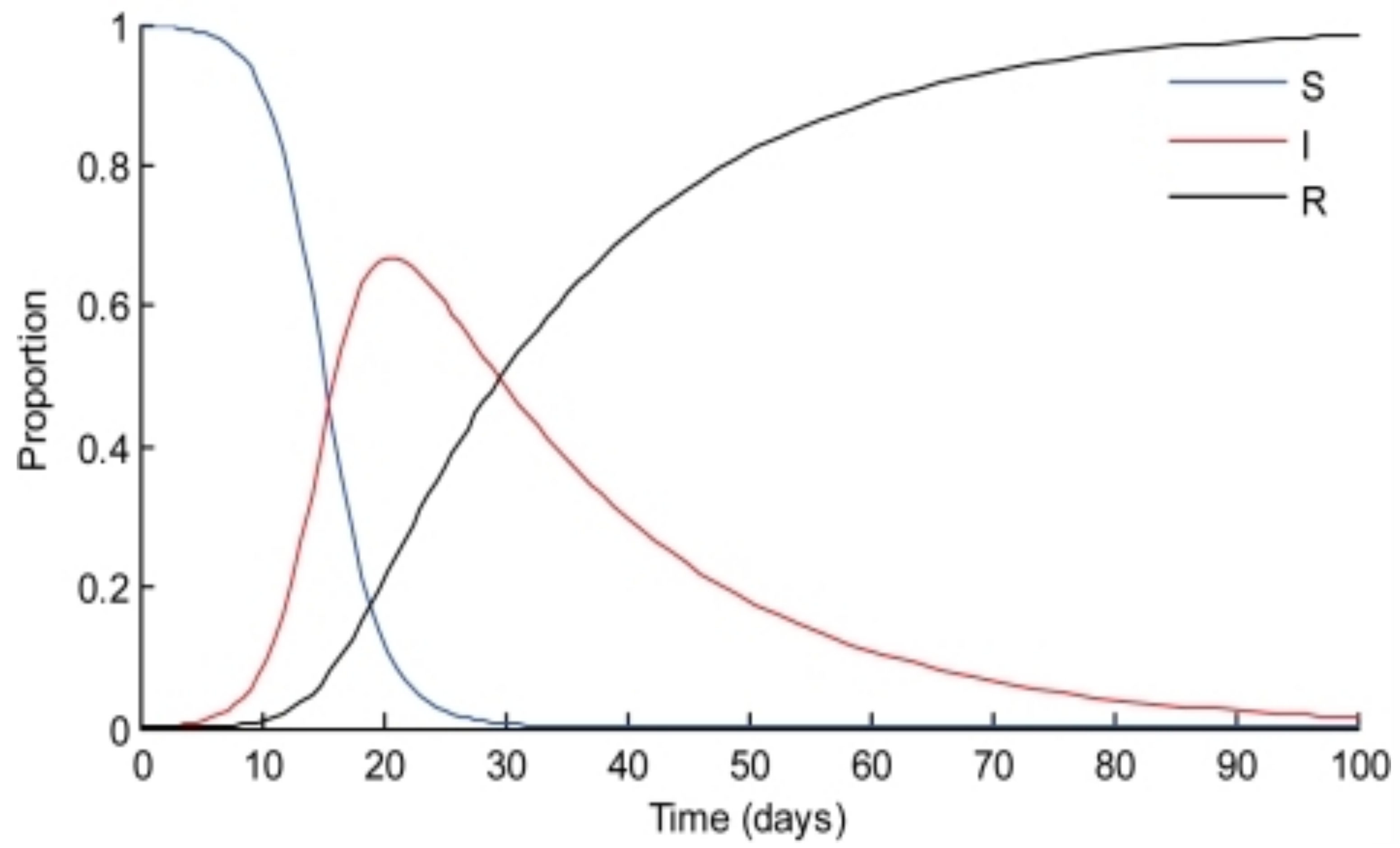
Epidemics

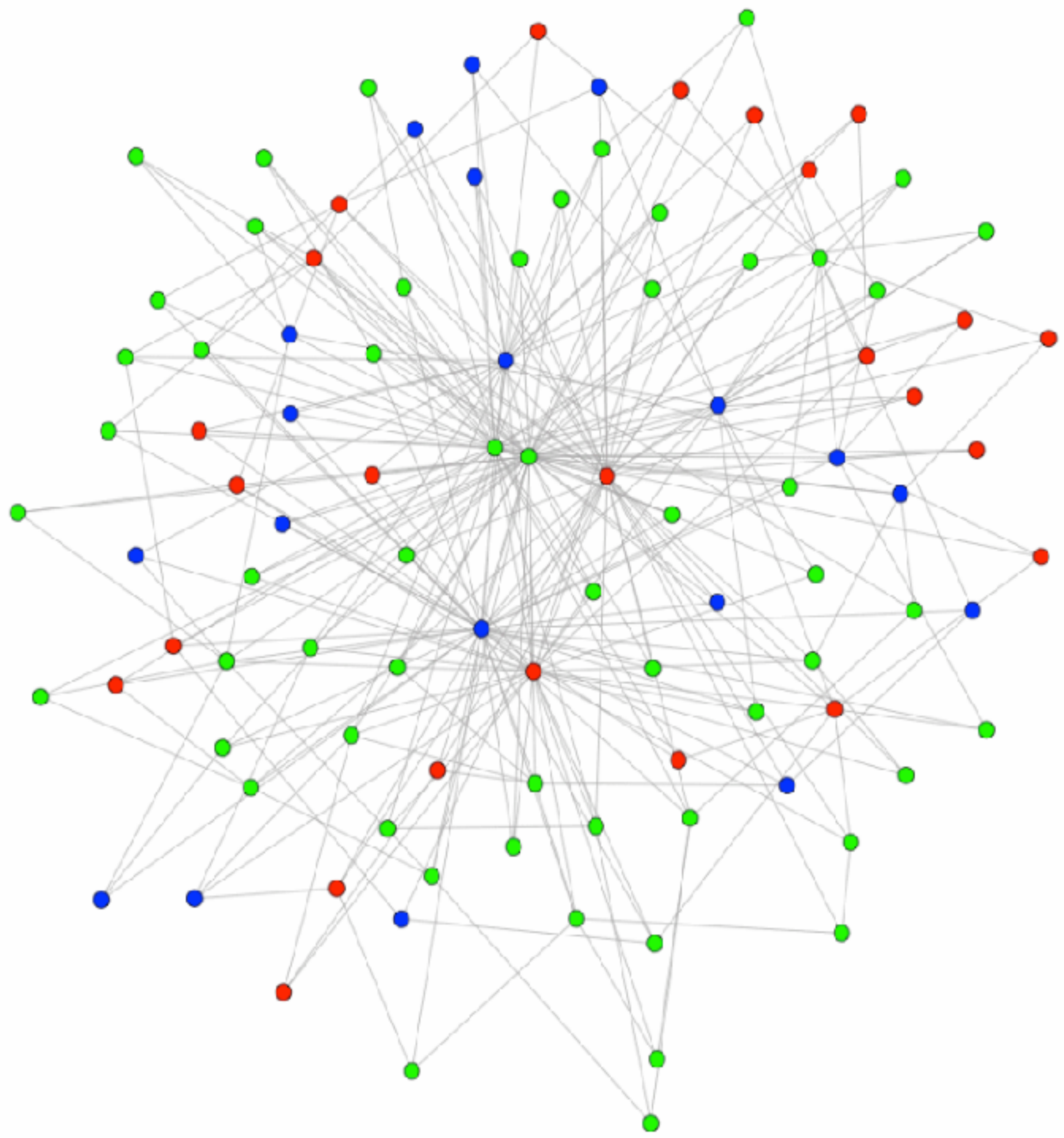


$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$





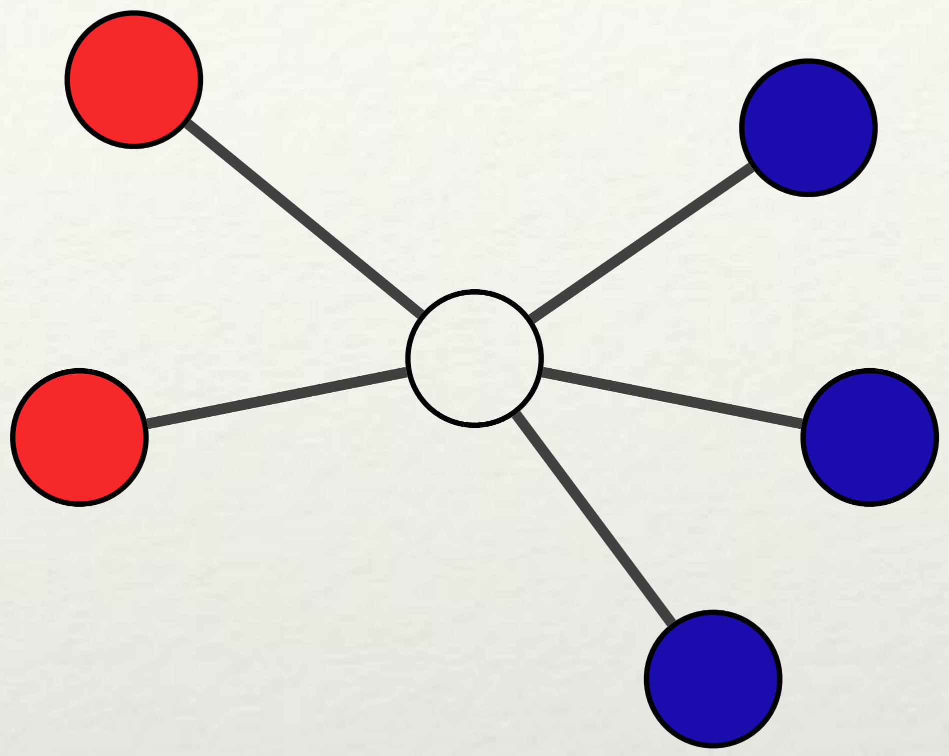
A diffusion of a new behavior

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

p fraction of neighbors adopting A

1-p fraction of neighbors adopting B

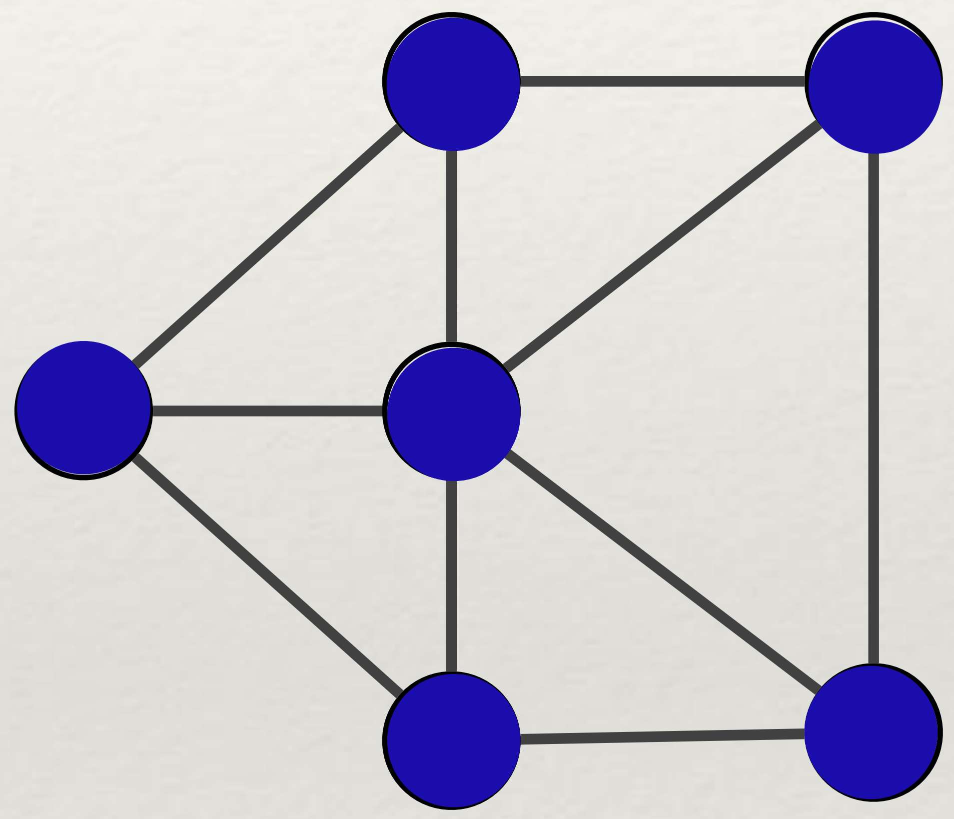
d is the number of neighbors

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

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

Example

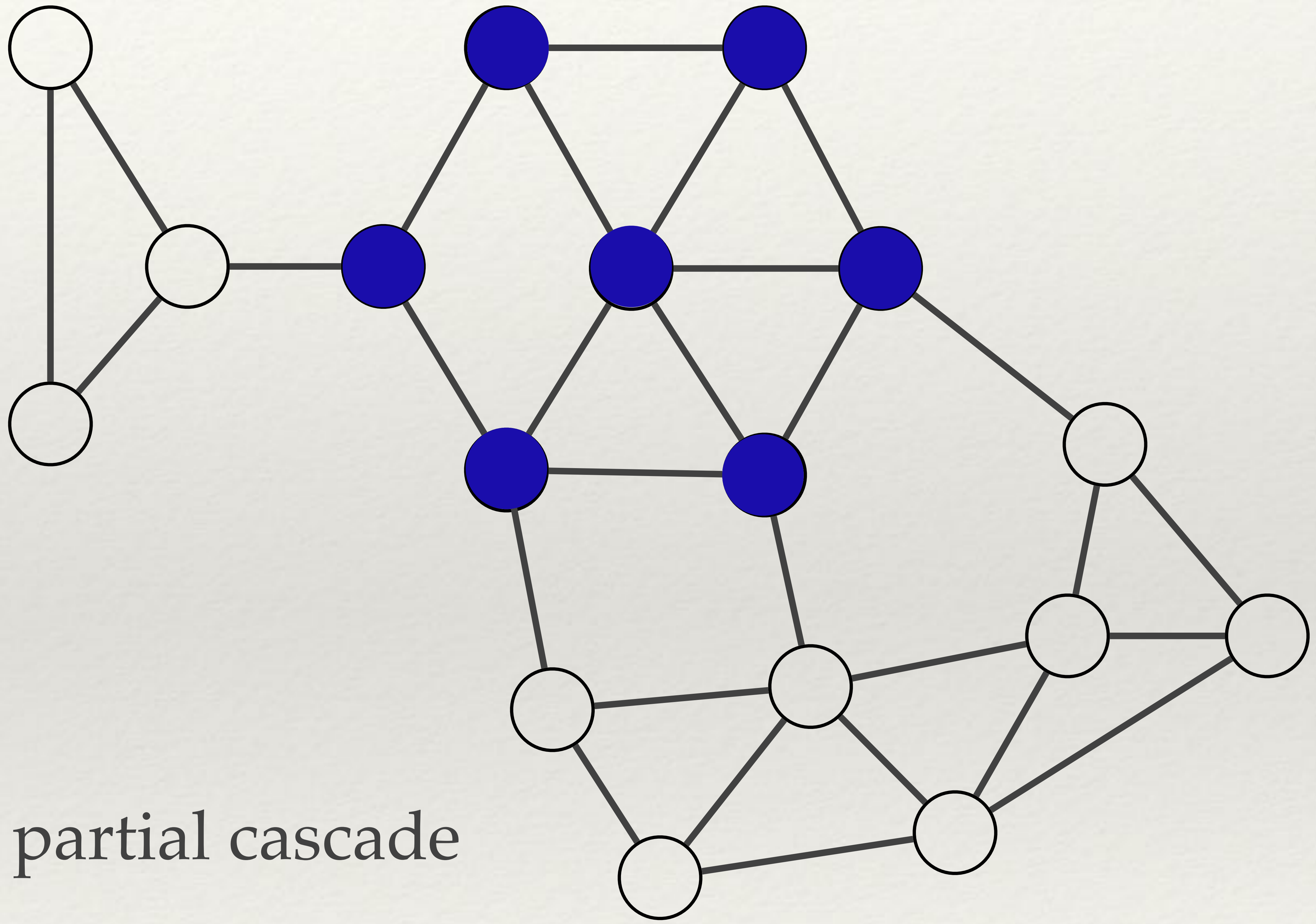
- $q = \frac{2}{5}$
- $S = \{u, v\}$



Chain reaction: complete cascade

Another example

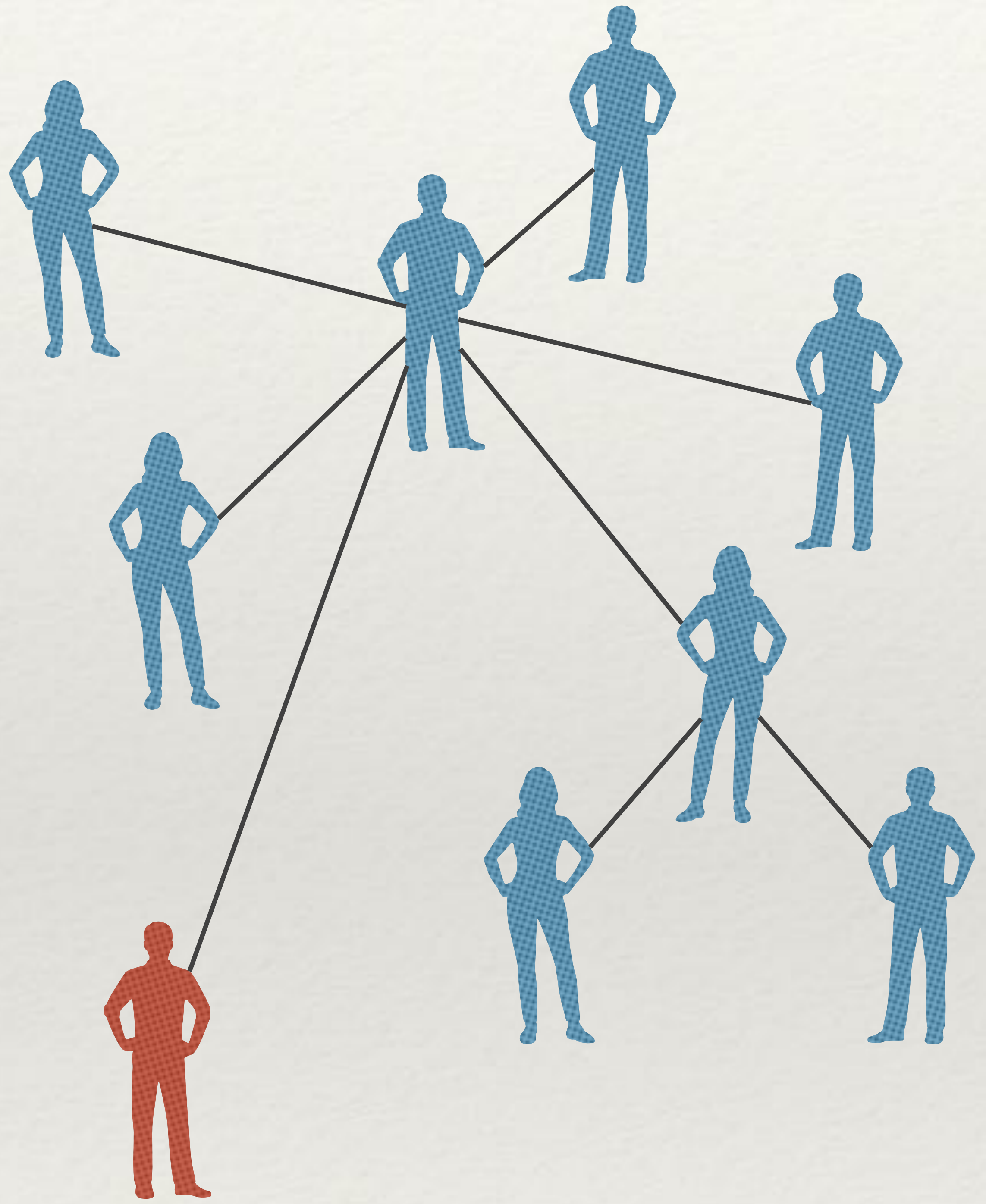
- $q = \dfrac{2}{5}$
- $S = \{u, v\}$



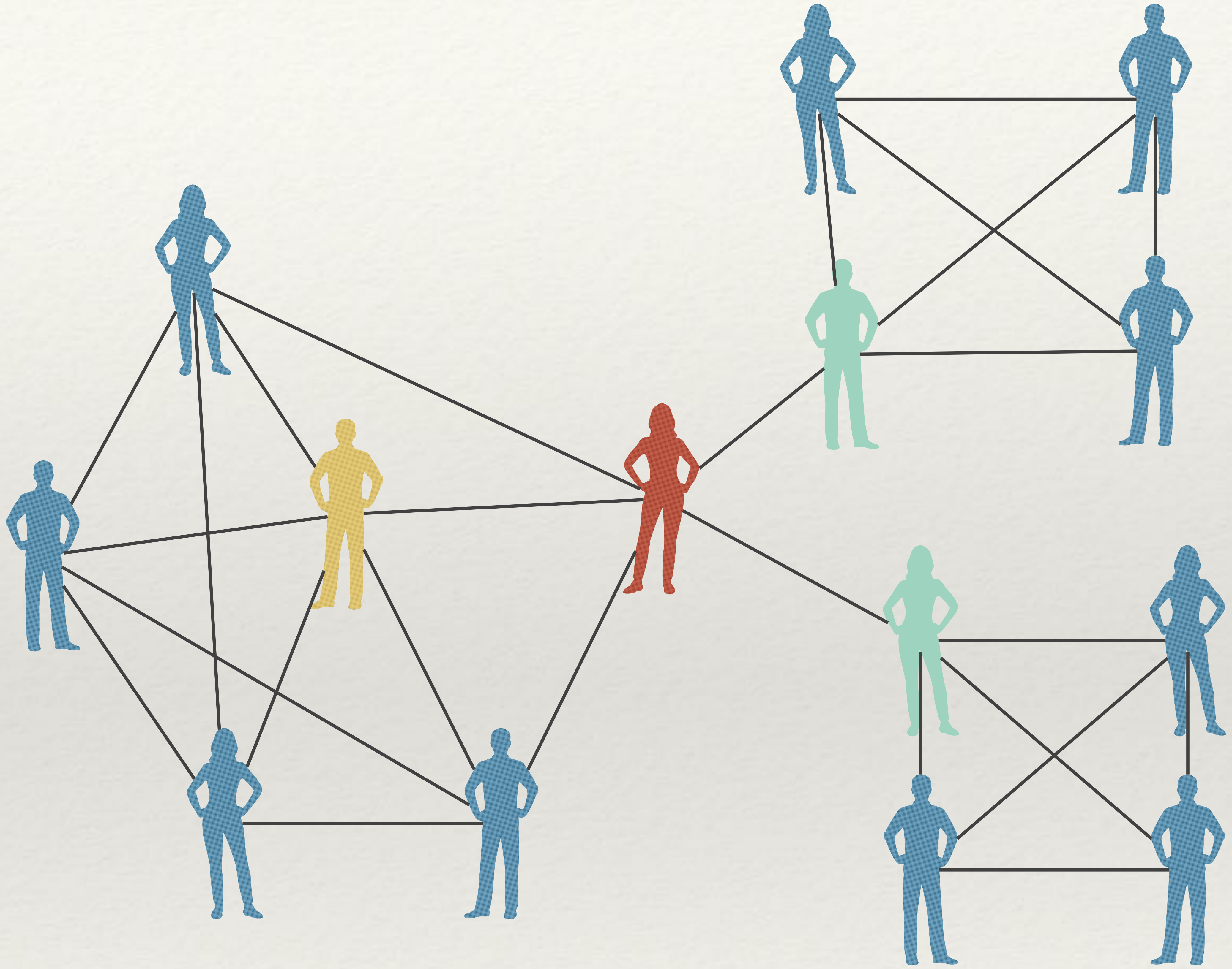
The diffusion of A stops here: partial cascade

Clusters are **barriers** to diffusion!

Real networks are heterogeneous



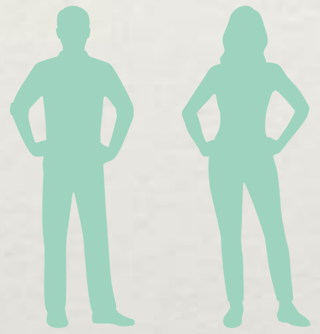
Rich-get-richer dynamics
(aka preferential attachment)



weak/strong ties, betweenness,
homophily, clusters

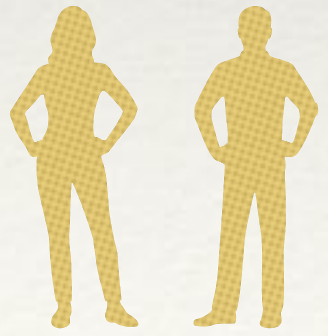
The role of weak ties

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

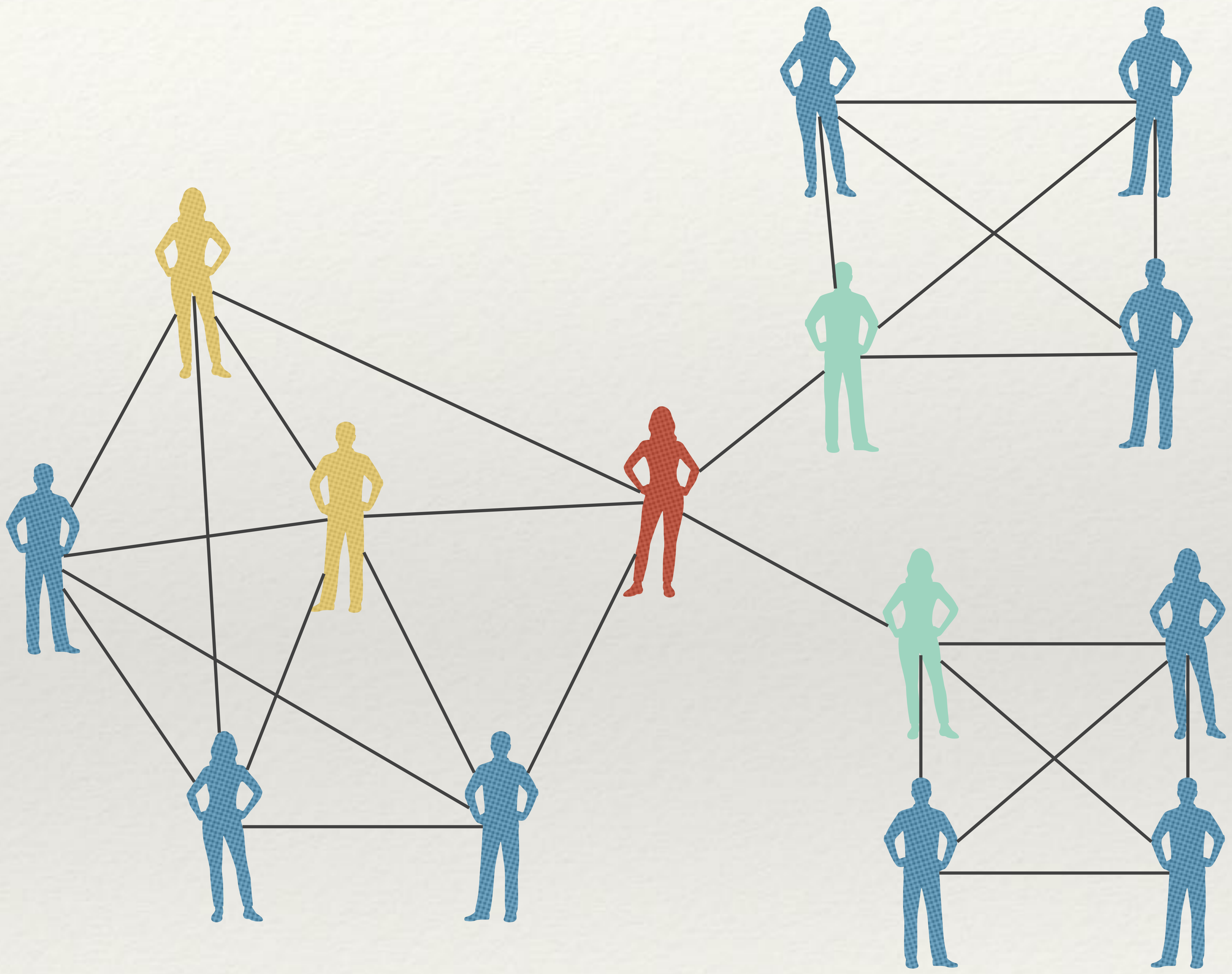


They receive very **fresh ideas** from other communities; not enough for adoption and spread (try with $q = \frac{1}{2}$)

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 *the 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

Problems and (some) answers

Terminology

Misinformation

Malinformation

Fake-News

Disinformation

Unverified Information

Propaganda

Conspiracy Theories

Urban Legend

Rumors

Astroturf

Spam

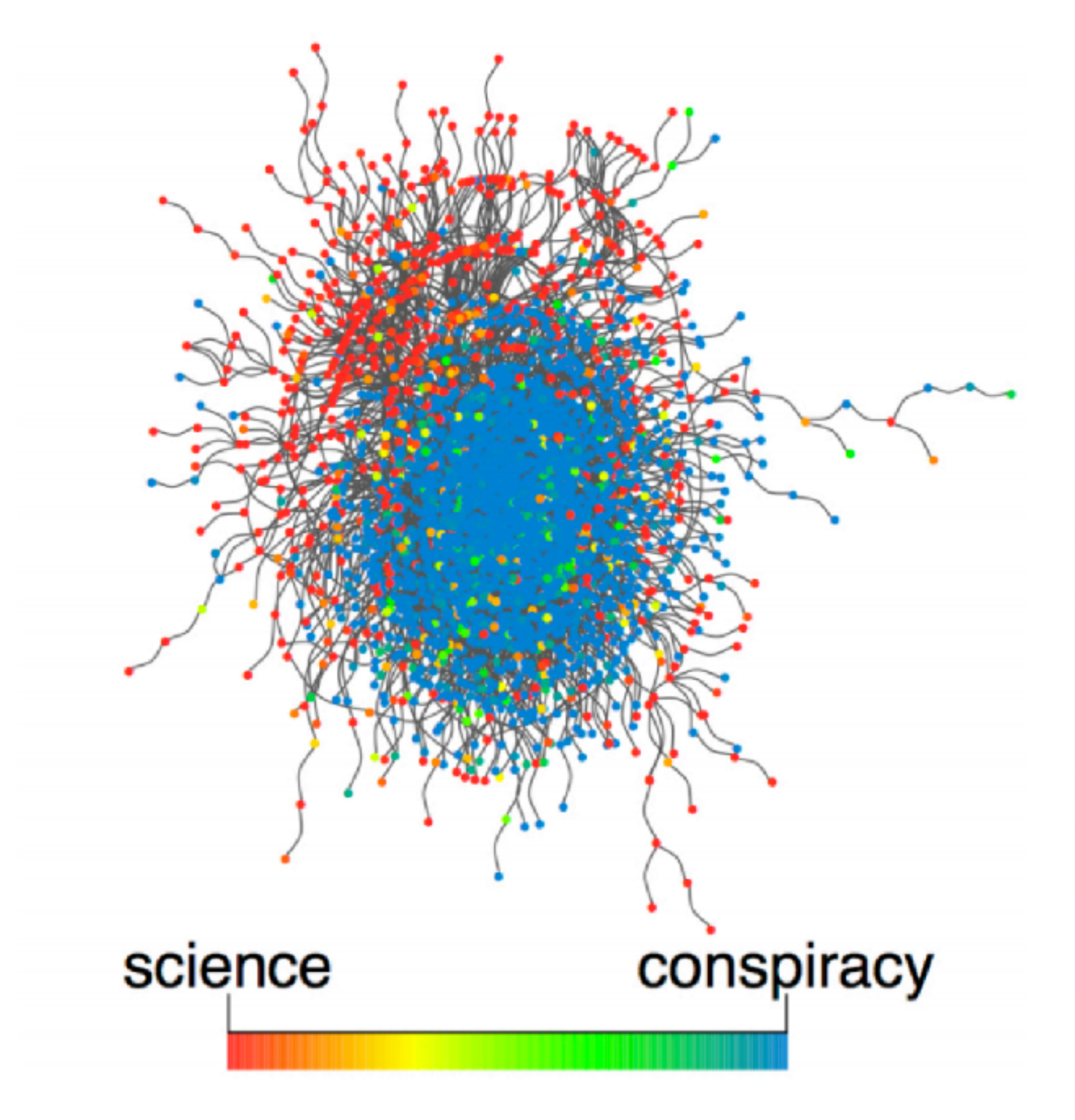
Troll

Hate Speech

Cyberbullying

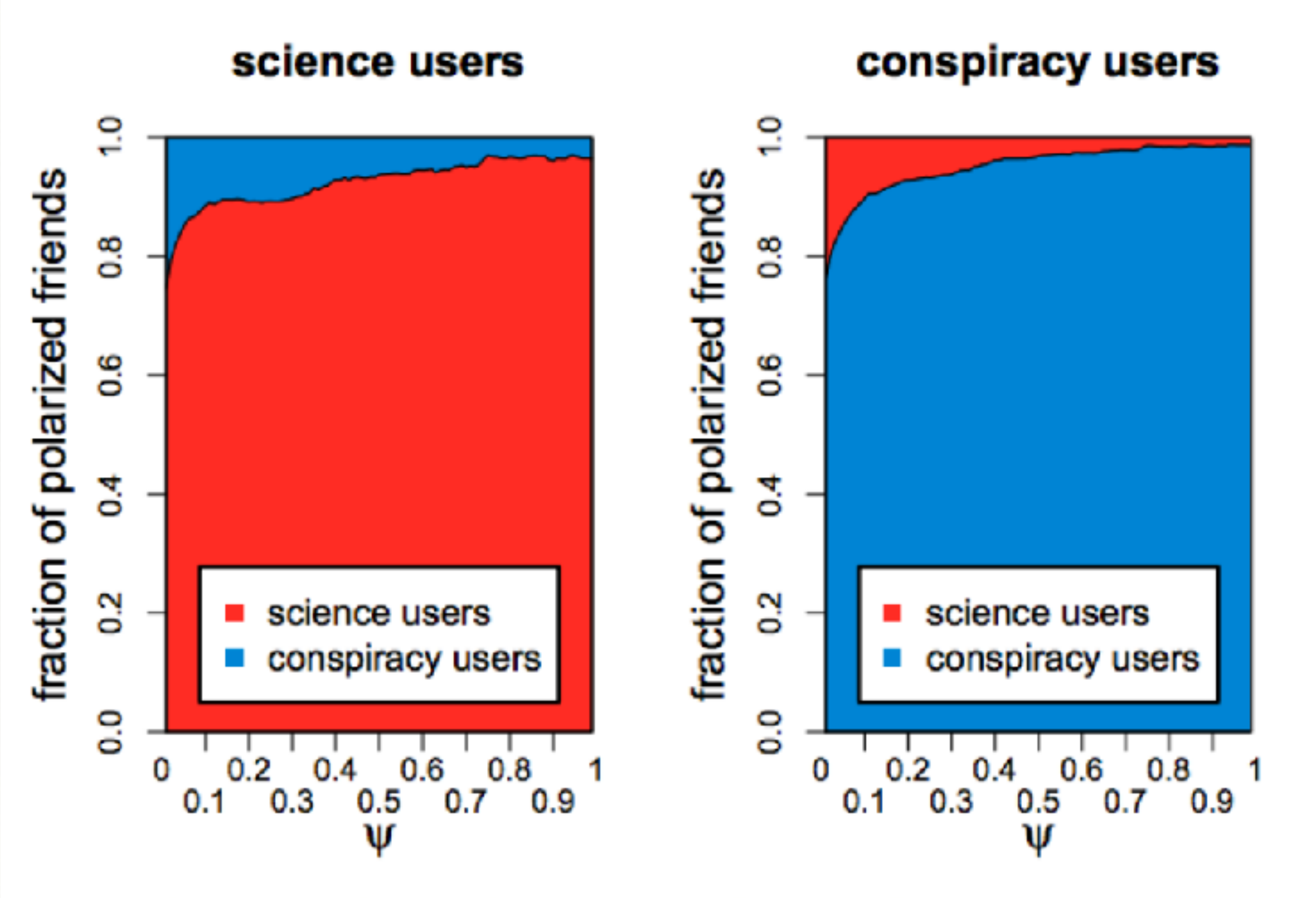
Echo-chambers

Misinformation tends to polarize



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

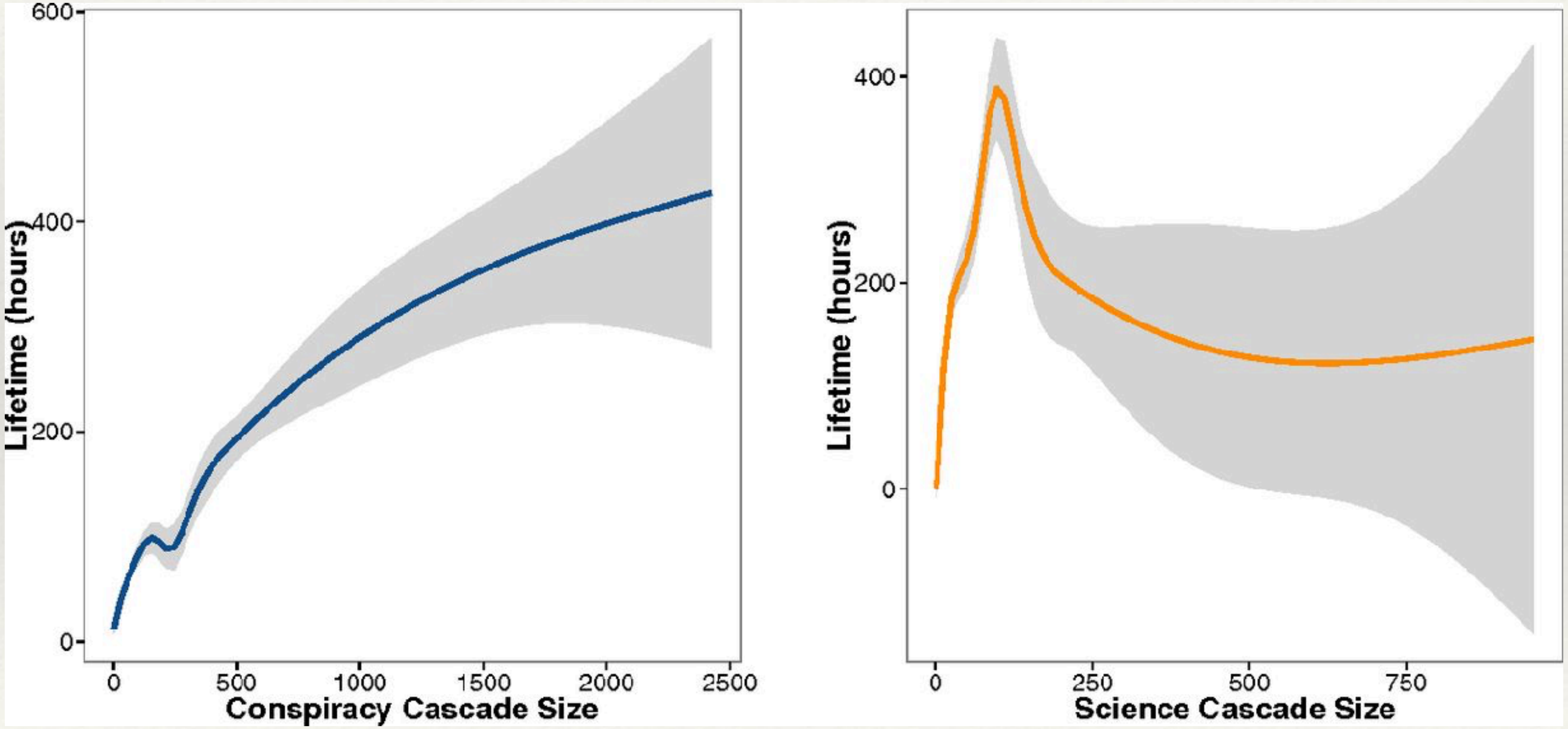
homophily!



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

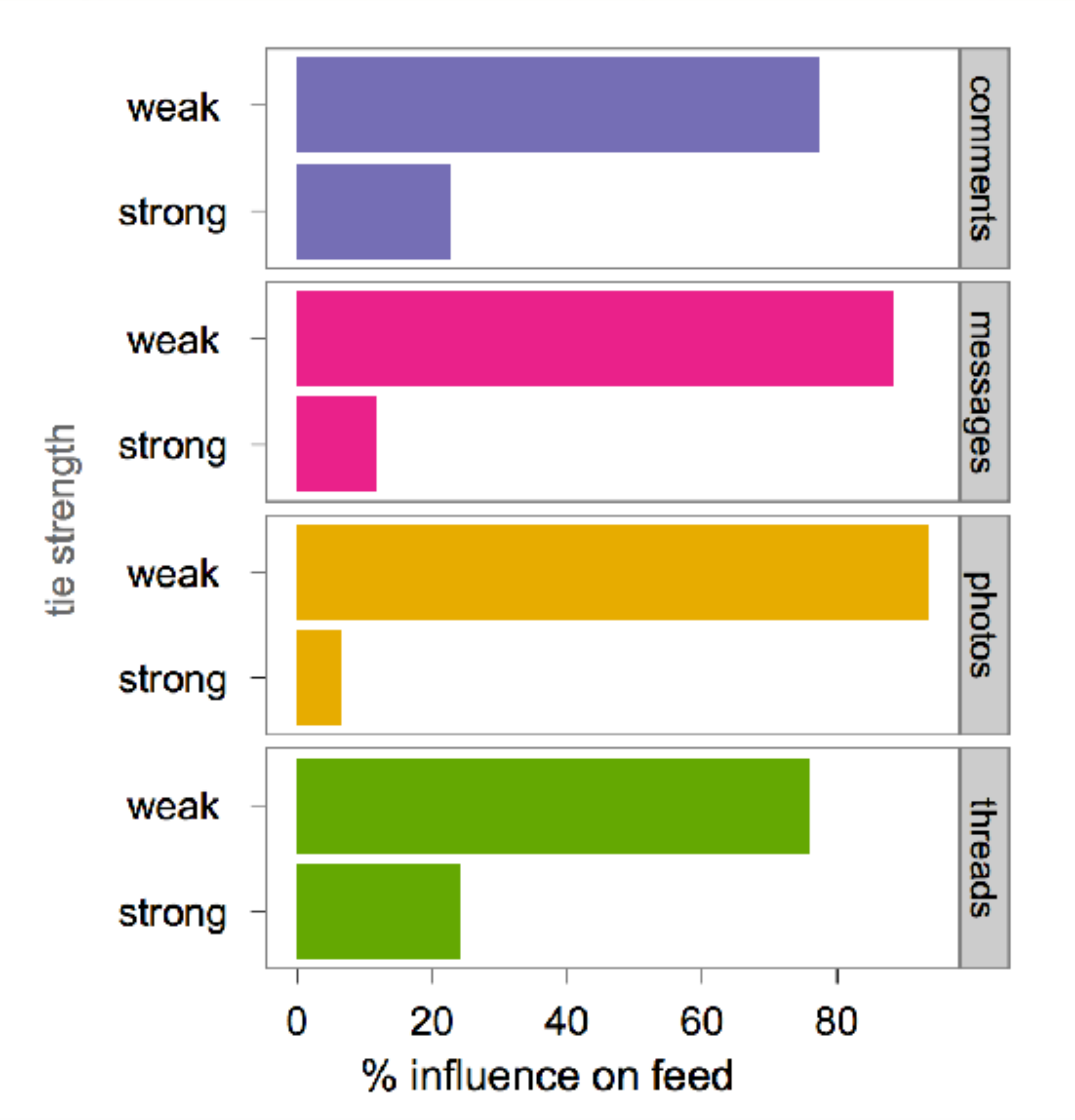
… and polarization fuels misinformation spread

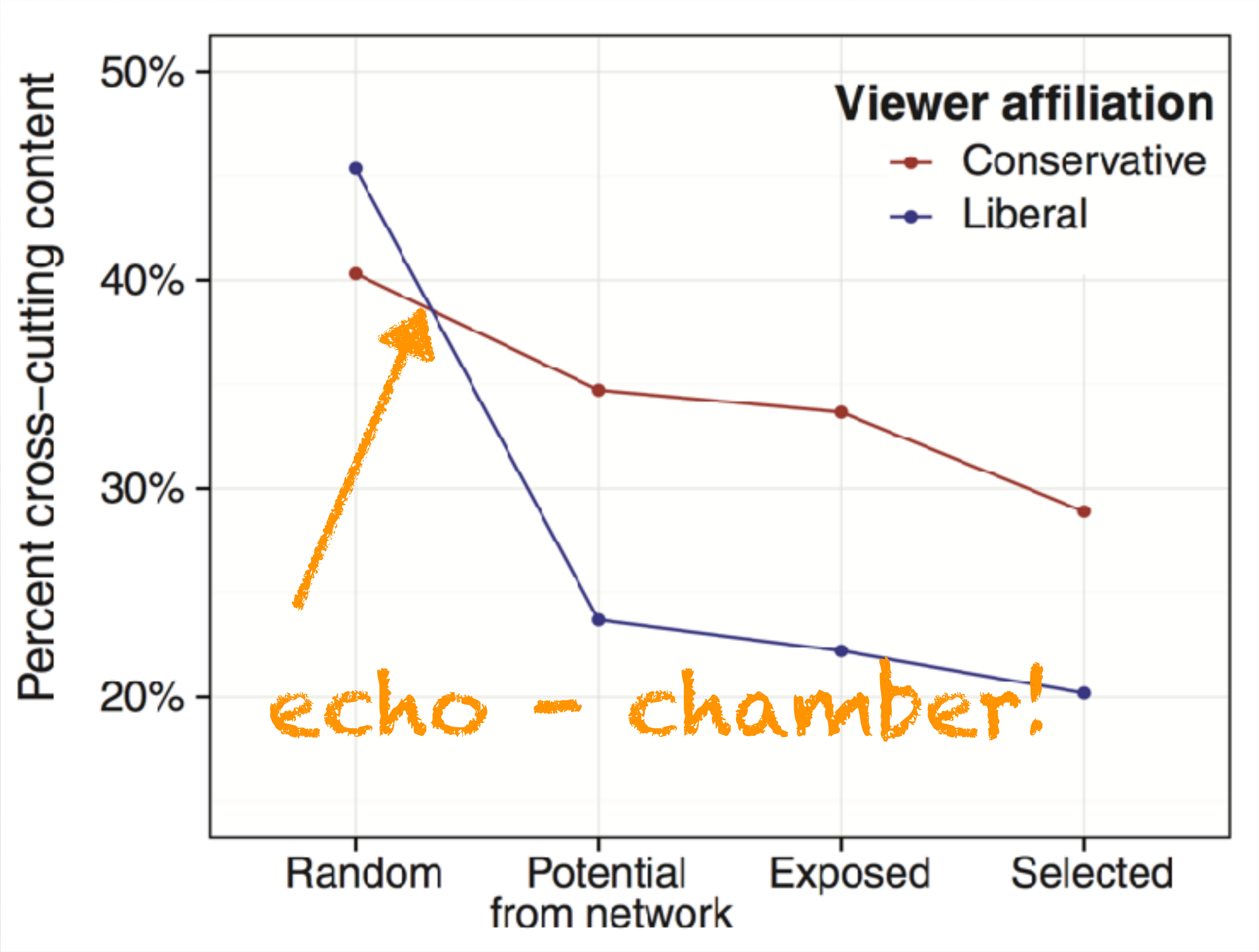
A data-driven percolation model of rumor spreading that demonstrates that homogeneity and polarization are the main determinants for predicting cascades' size



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)

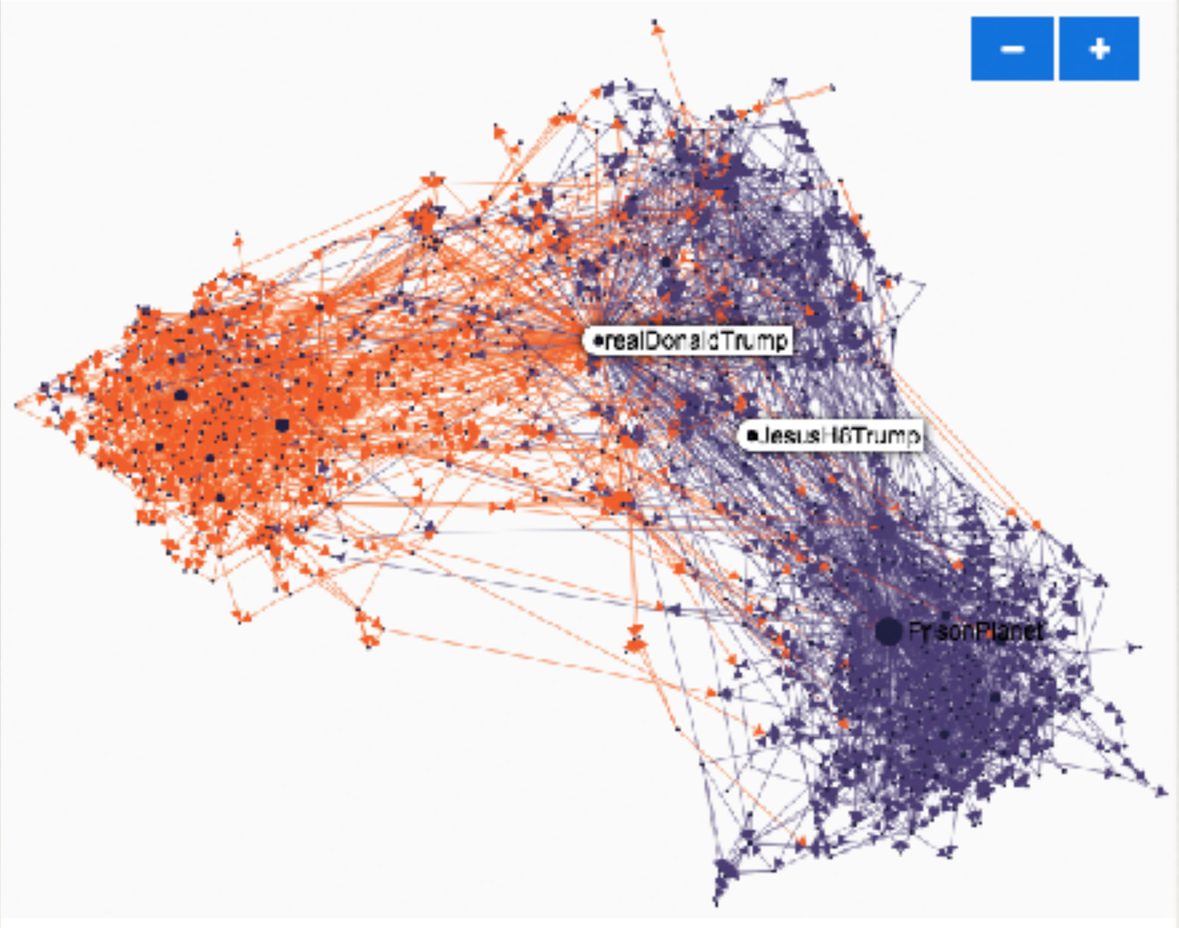
Analyzing the structure of a misinformation network

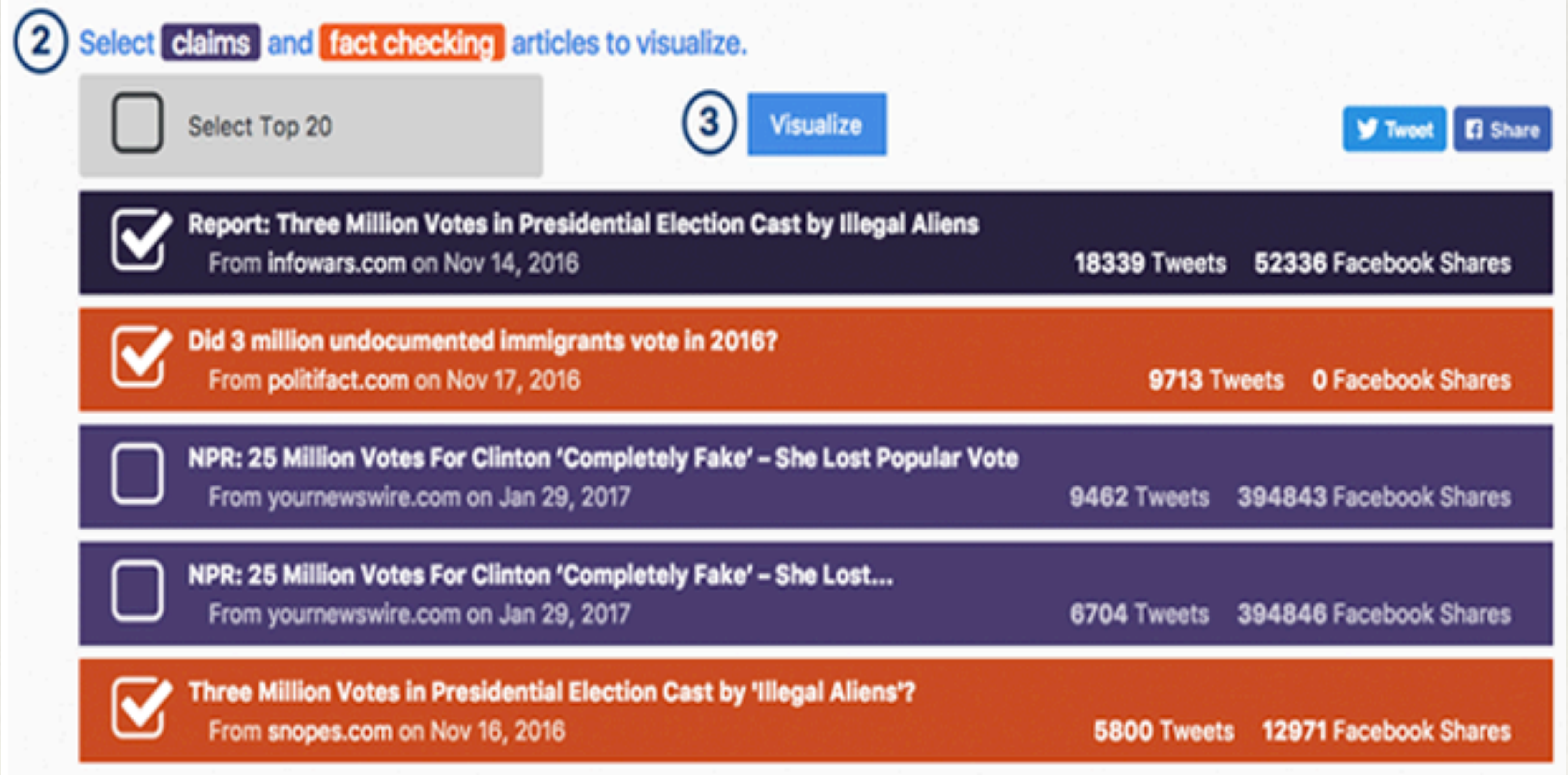
- *What are the structural and dynamic characteristics of the core of the misinformation diffusion network, and who are its main purveyors?*
- "As we move from the periphery to the core of the network, fact-checking nearly disappears, while social bots proliferate."

HOAXY beta

Visualize the spread of claims and fact checking.

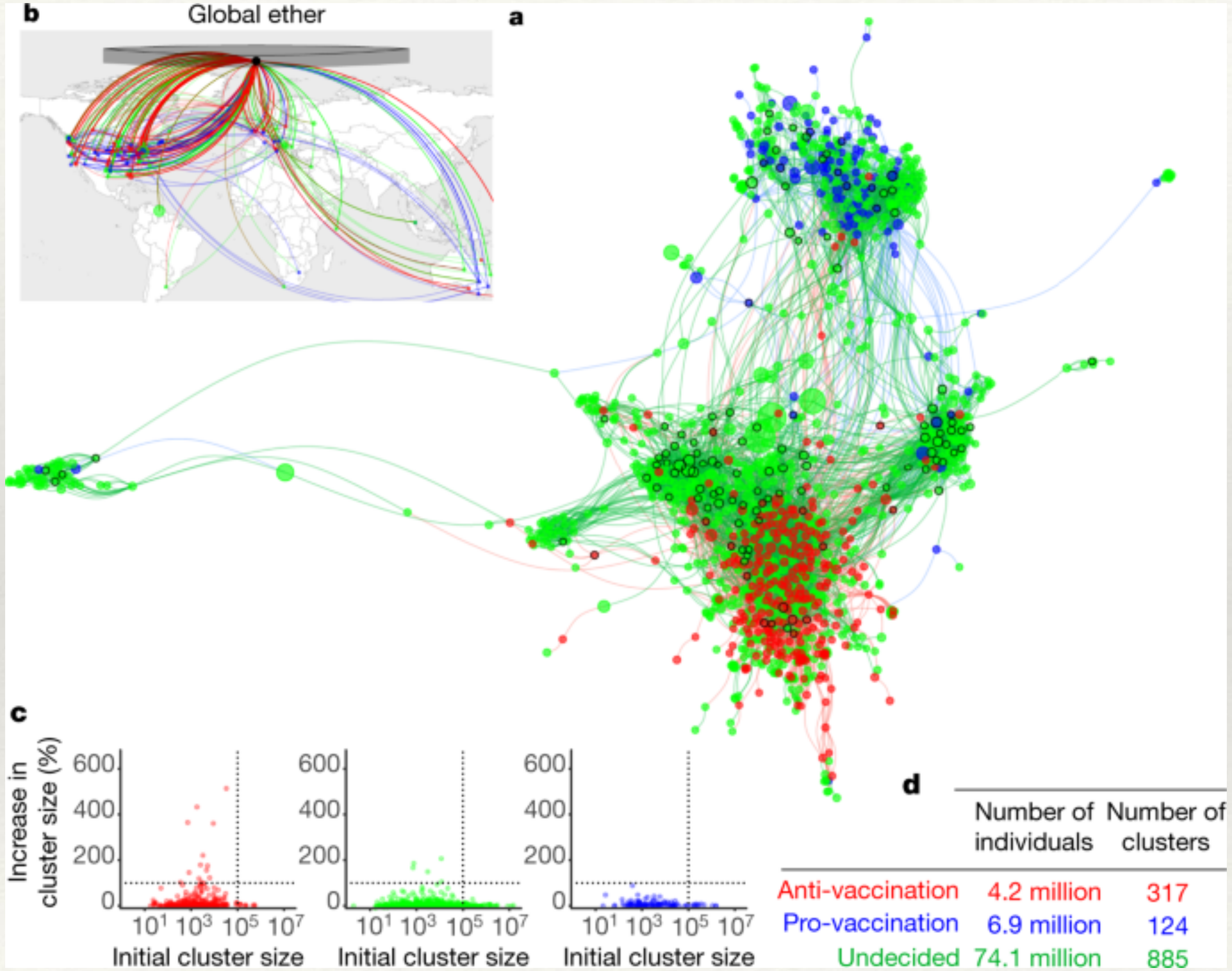
https://hoaxy.iuni.iu.edu



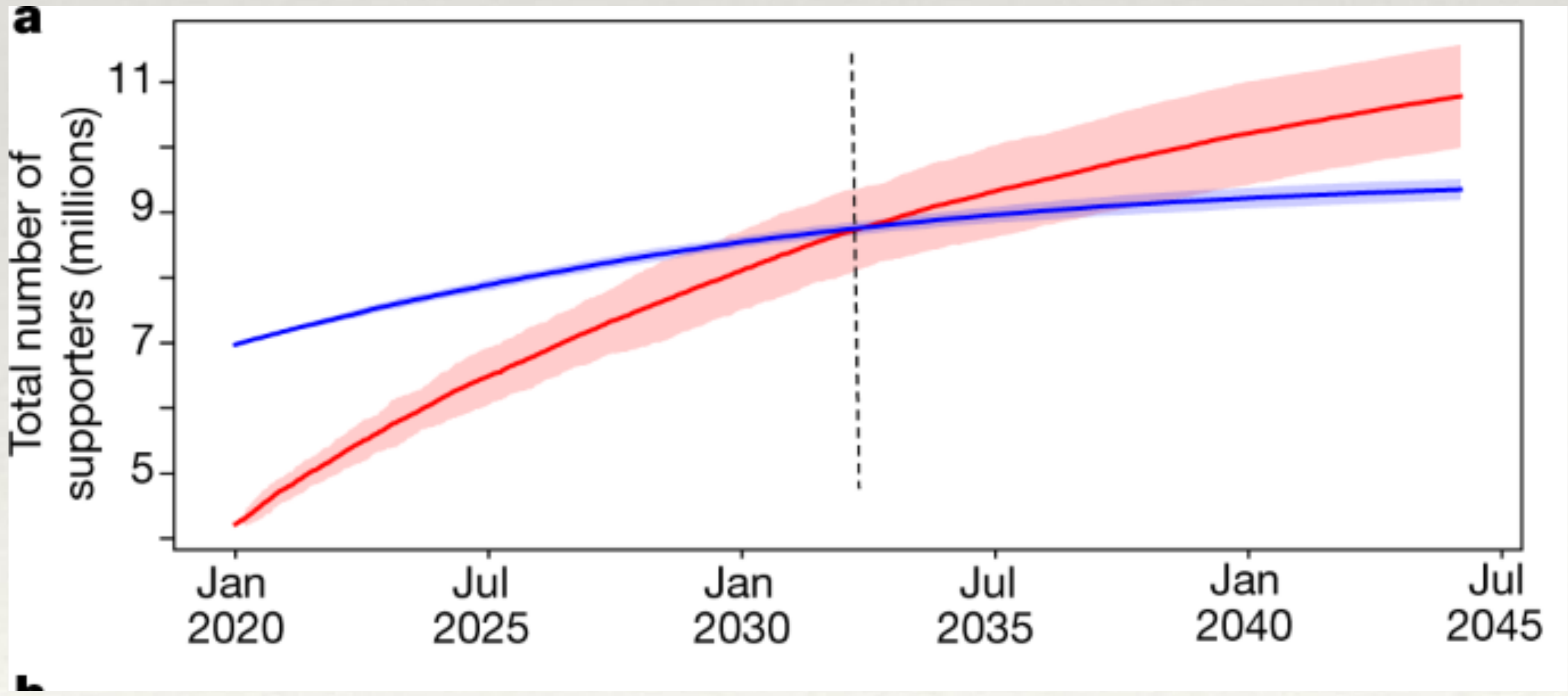


Shao C, Hui P-M, Wang L, Jiang X, Flammini A, Menczer F, et al. (2018) Anatomy of an online misinformation network. PLoS ONE 13(4):e0196087. https://doi.org/10.1371/journal.pone.0196087

The role of the undecided



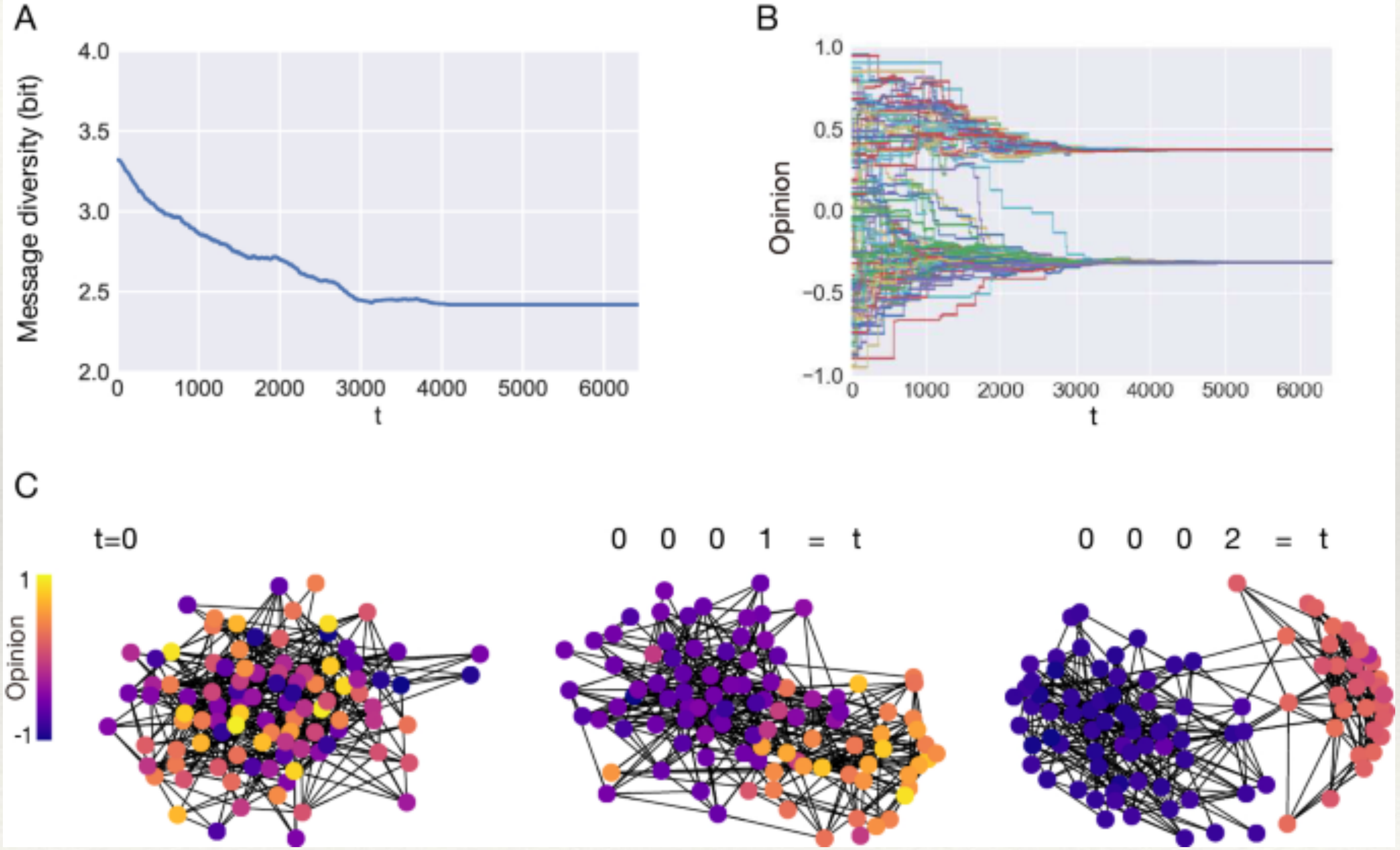
- Theoretical prediction for the future total size of anti-vaccination and pro-vaccination support
- Under the present conditions, it predicts that total anti-vaccination support reaches dominance in around 10 years



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

The Science of Fake-News

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**

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

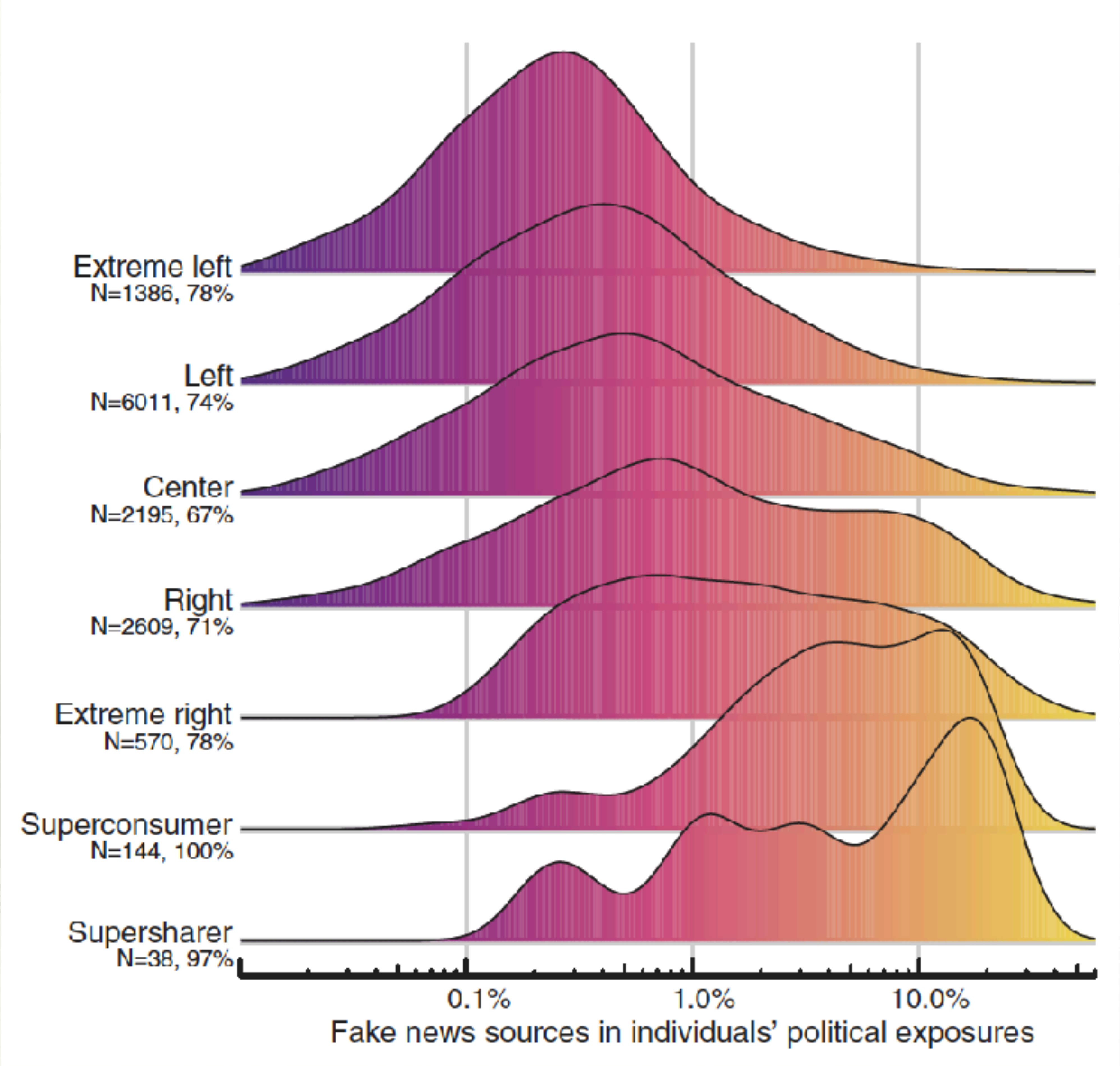
S.Vosoughi, D. Roy, S. Aral, The spread of true and false news online, in Science, 09 mar 2018: 1146-1151

- 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

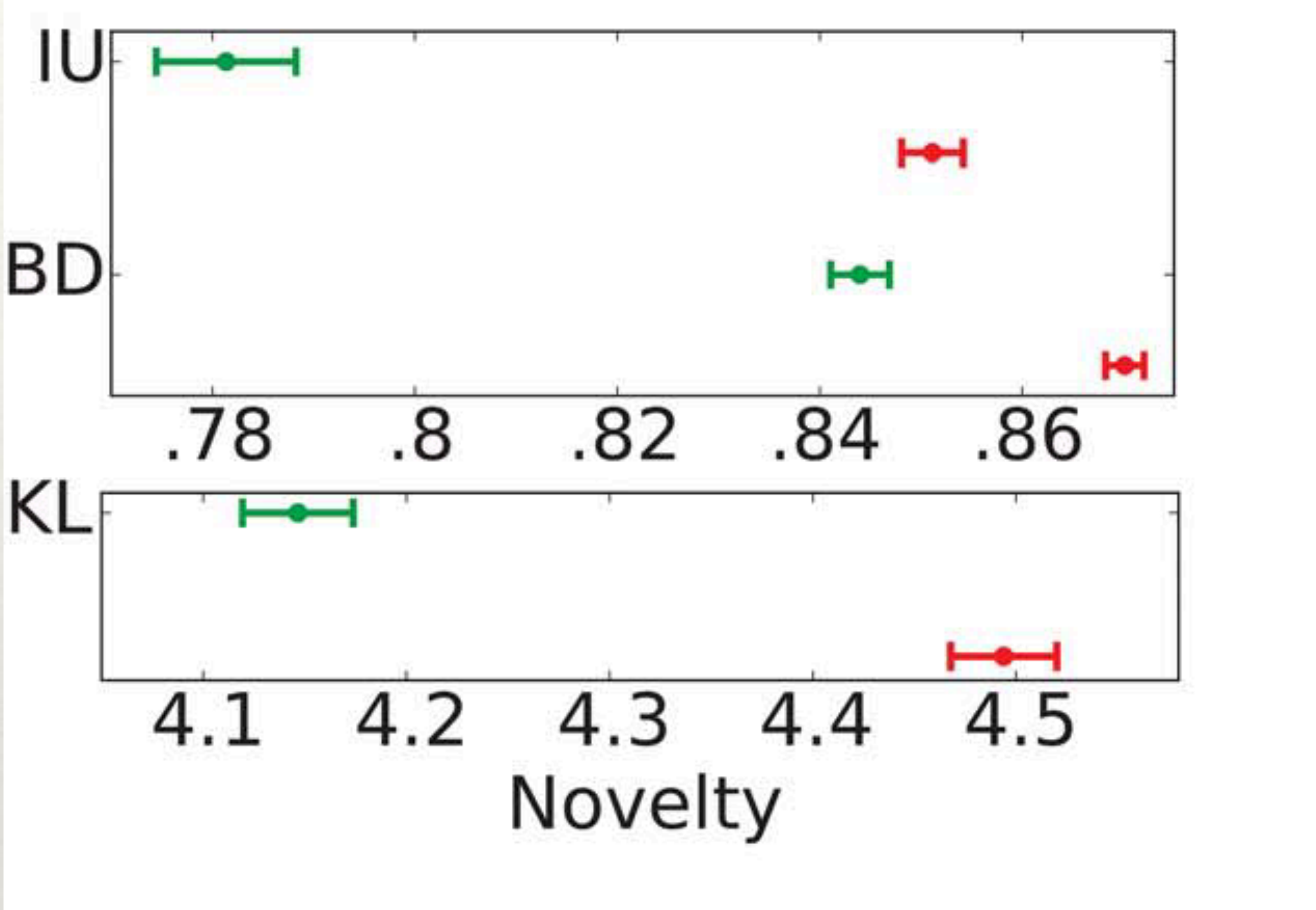
Fake-News and elections

- 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.



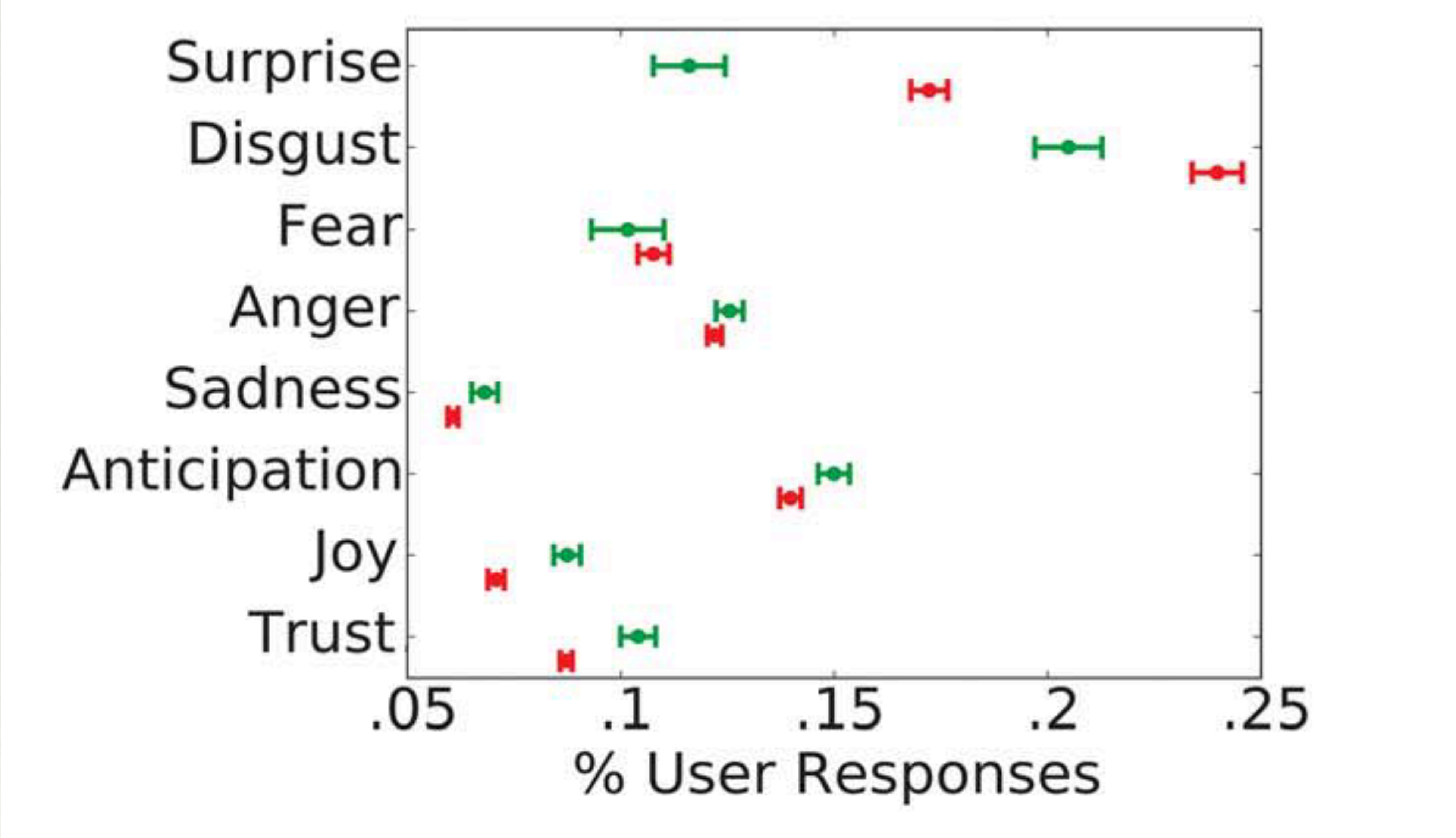
N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Fake news on Twitter during the 2016 U.S. presidential election, Science 25 Jan 2019:Vol. 363, Issue 6425, pp. 374-378 DOI: 10.1126/science.aau2706

Novelty and emotions



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

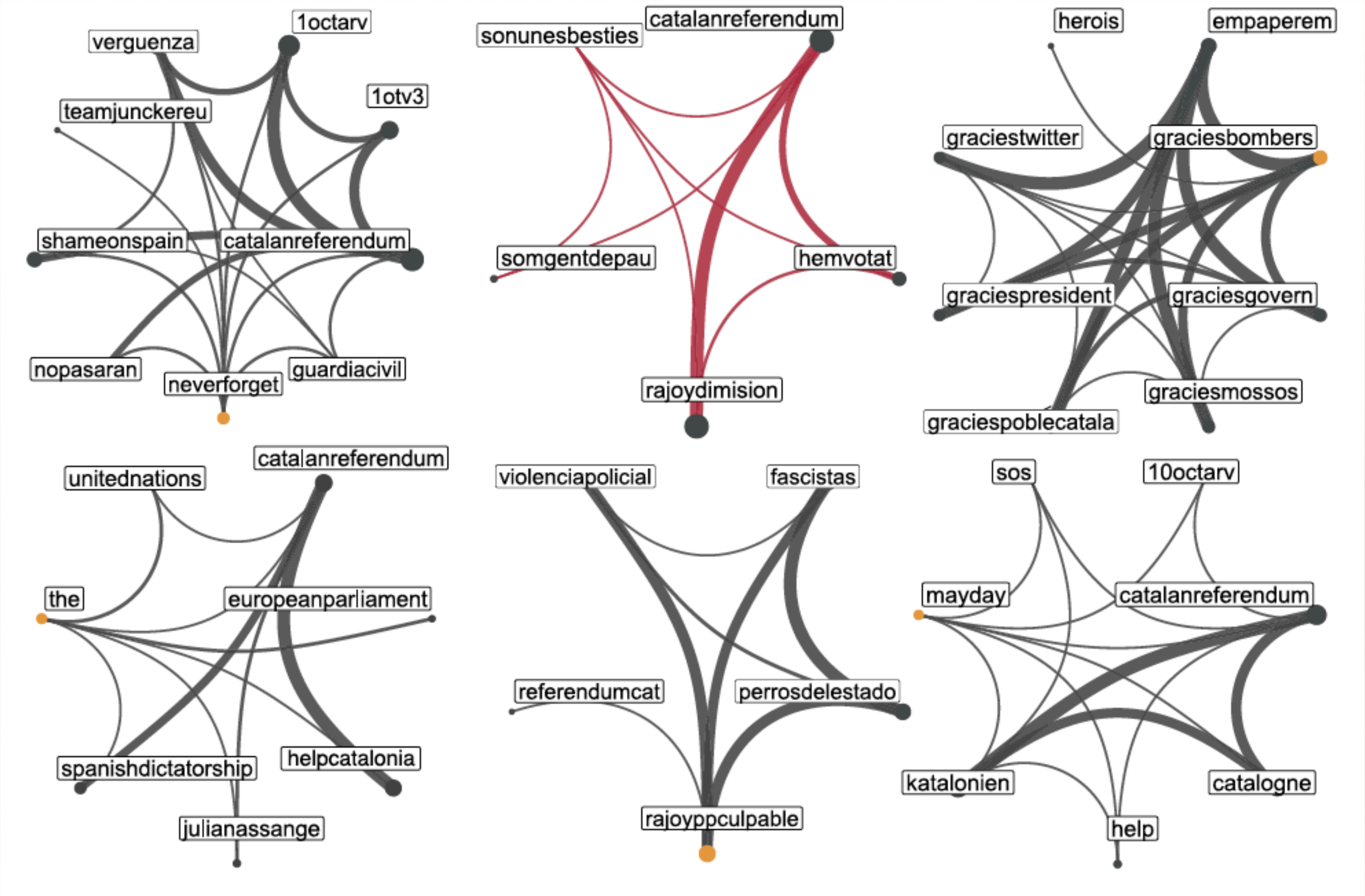
- False stories inspired **fear**, **disgust**, and **surprise** in replies, true stories inspired anticipation, sadness, joy, and trust.



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.

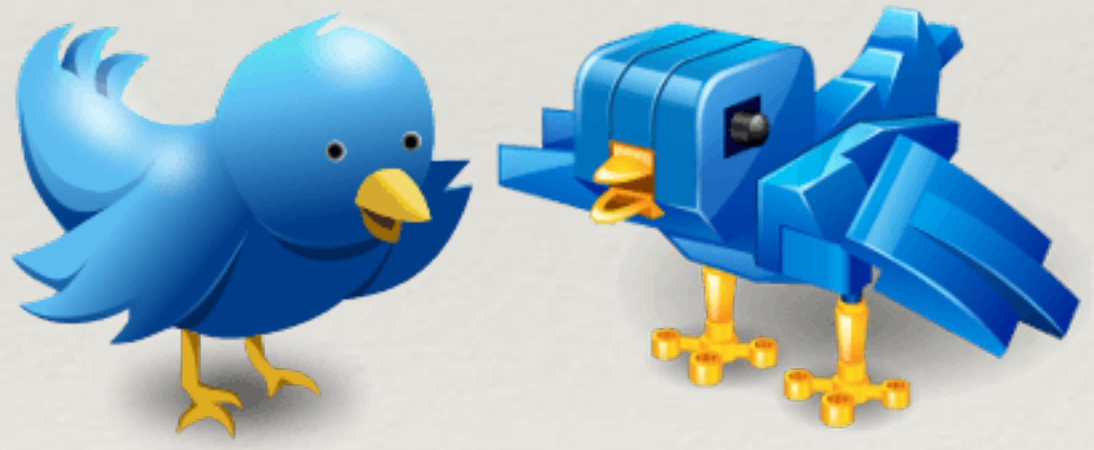
Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F., The spread of low-credibility content by social bots. *Nat Commun* **9**, 4787 (2018). https://doi.org/10.1038/s41467-018-06930-7

BotSlayer and Botometer (IU)

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

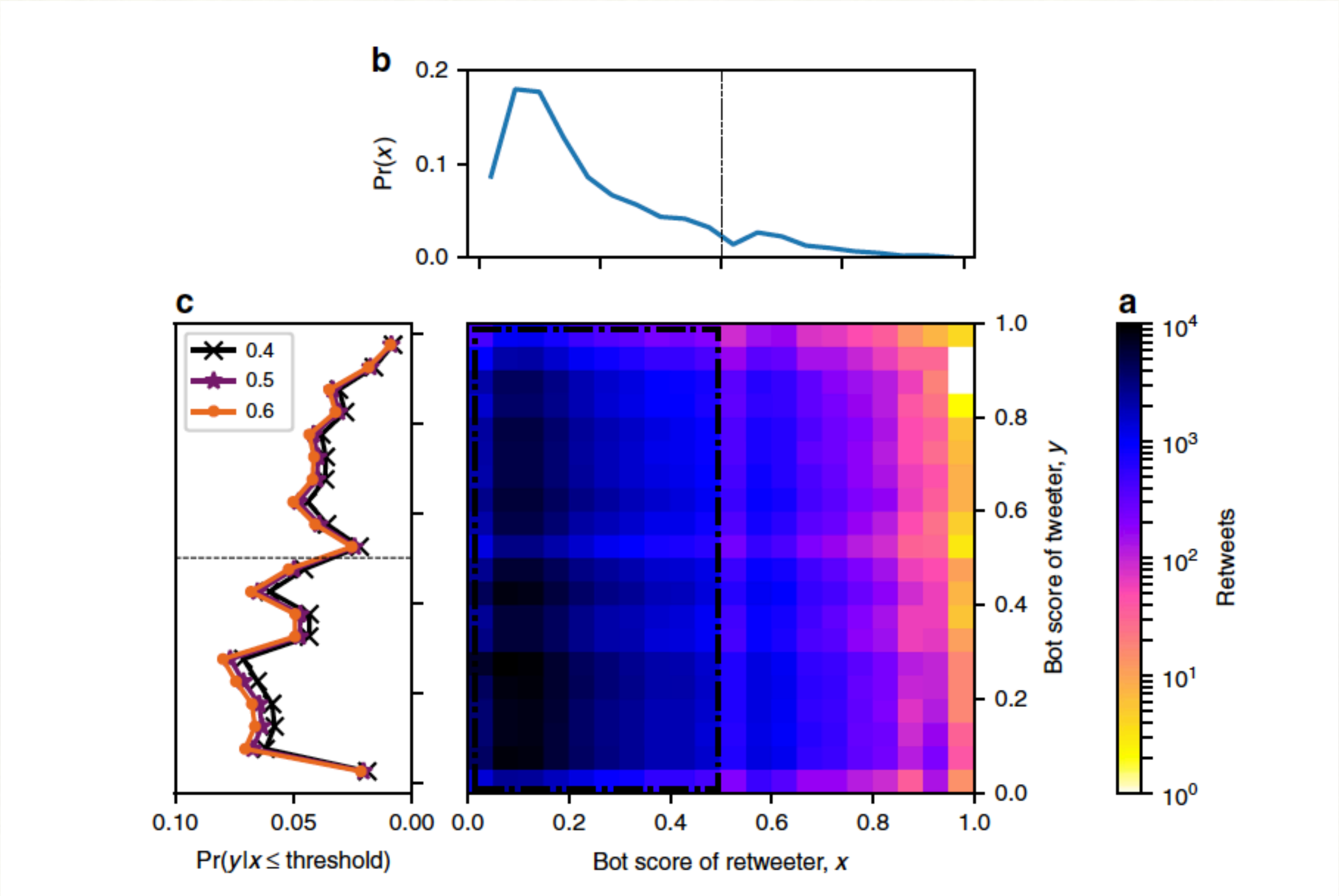
  https://osome.iuni.iu.edu/tools/botslayer/

- **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

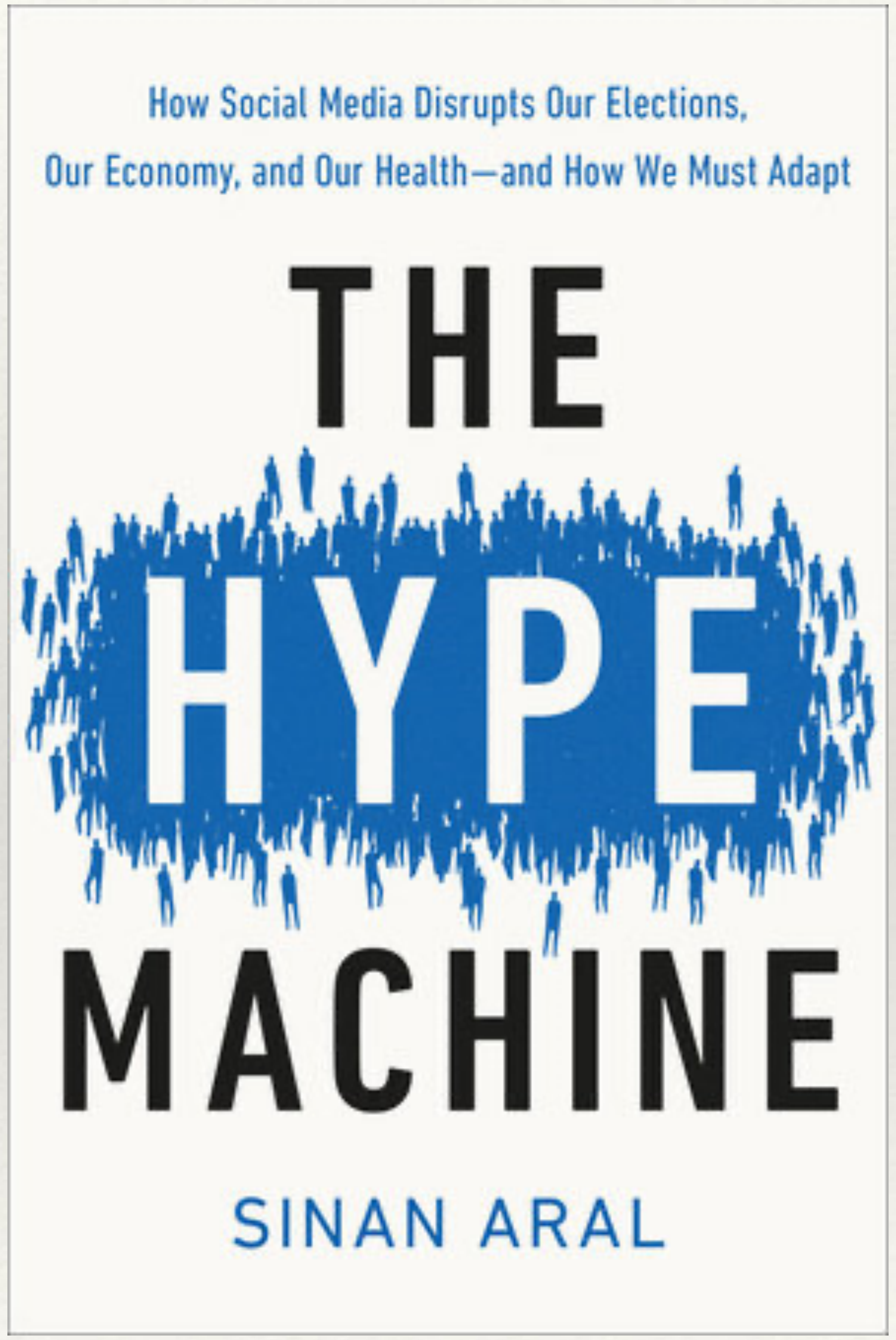
…but humans should be blamed the most



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/

Evaluating Fact-checking

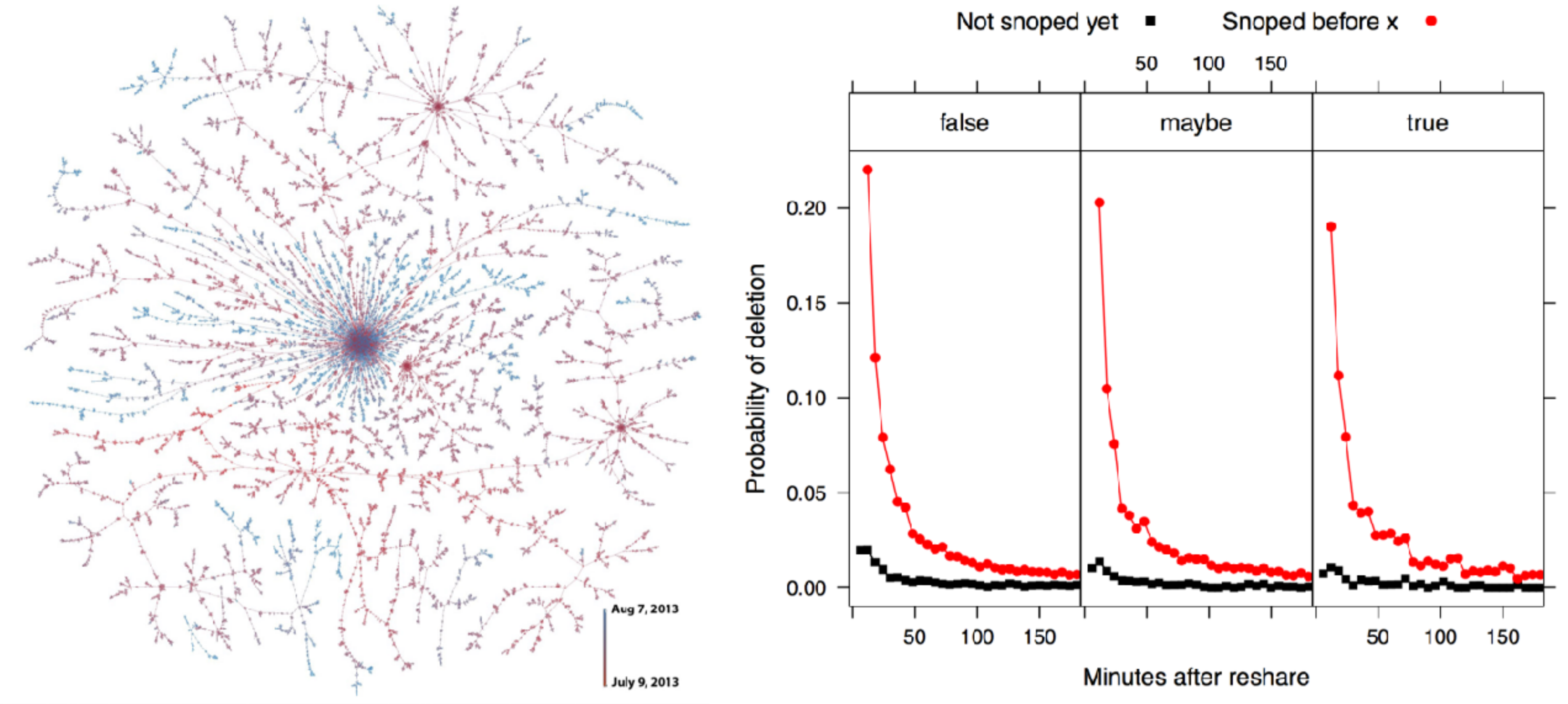
Modeling the spread of misinformation

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?*
- *Which is the best fact-checking strategy?*

Fact checking can be effective

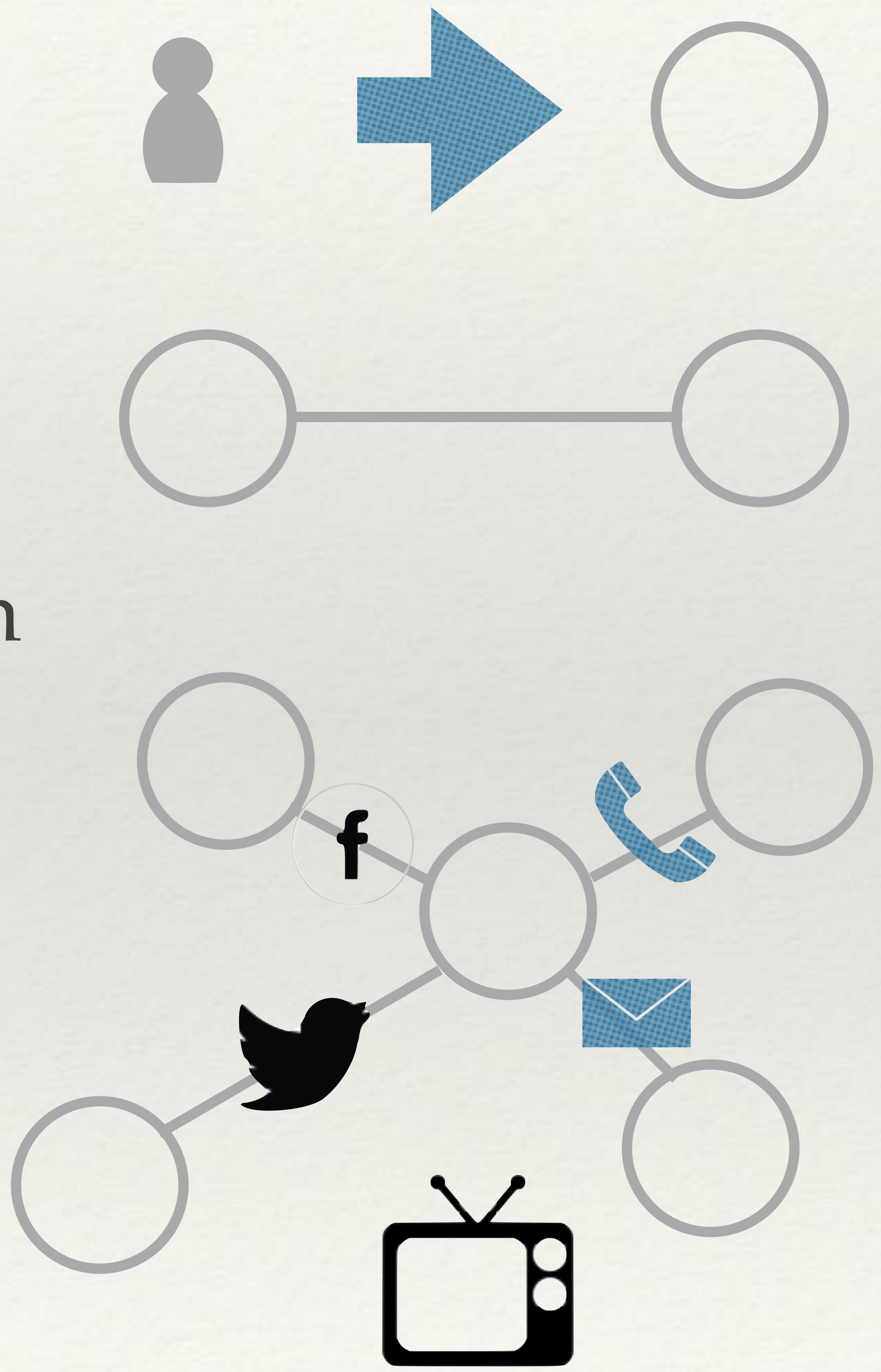
- Cascades of thousands of rumors from FB analyzed
- Reshares that are *snoped* shortly after posting are very likely to be deleted



A Friggeri, L Adamic, D Eckles, J Cheng, Rumor Cascades, in Proc. of 8th int. AAAI conf. on weblogs and and social media (ICWSM), 2014

Networks and their context

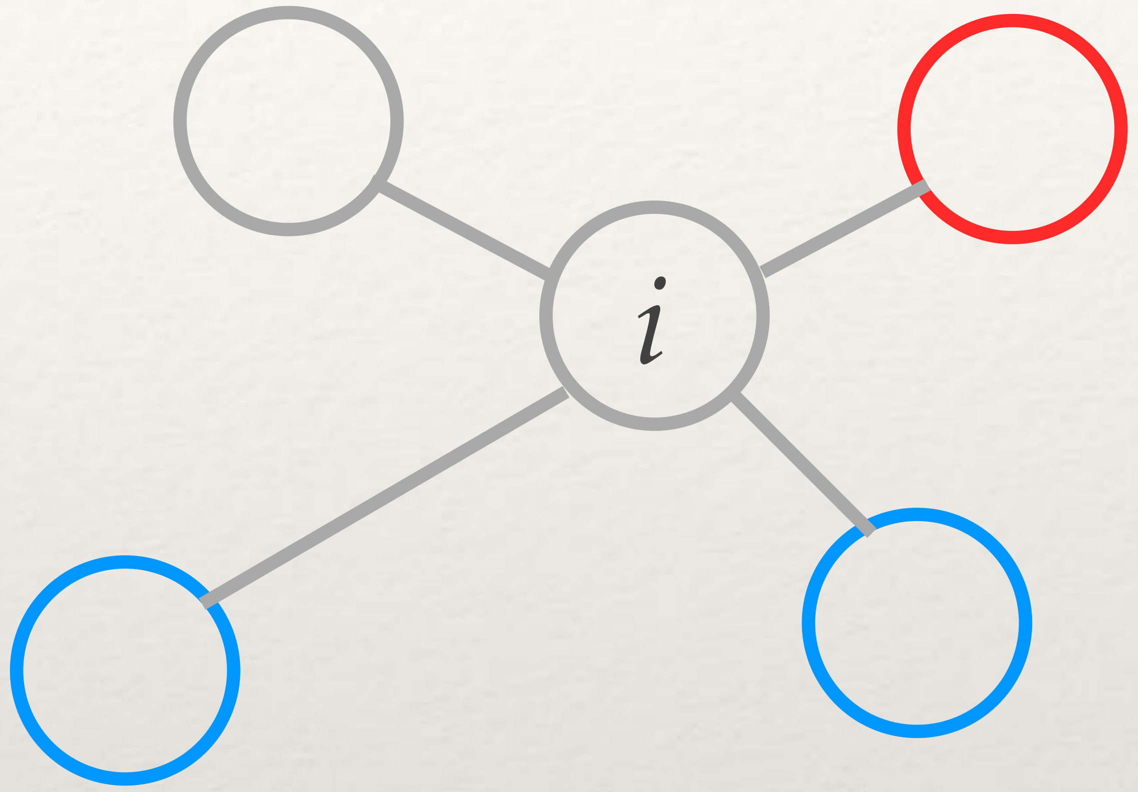
- nodes are **actors** involved in a **generic** social network (no assumption is given)
- links are **social relationships**
- nodes can be exposed to news from both **internal and external sources** and via different communication devices



- **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

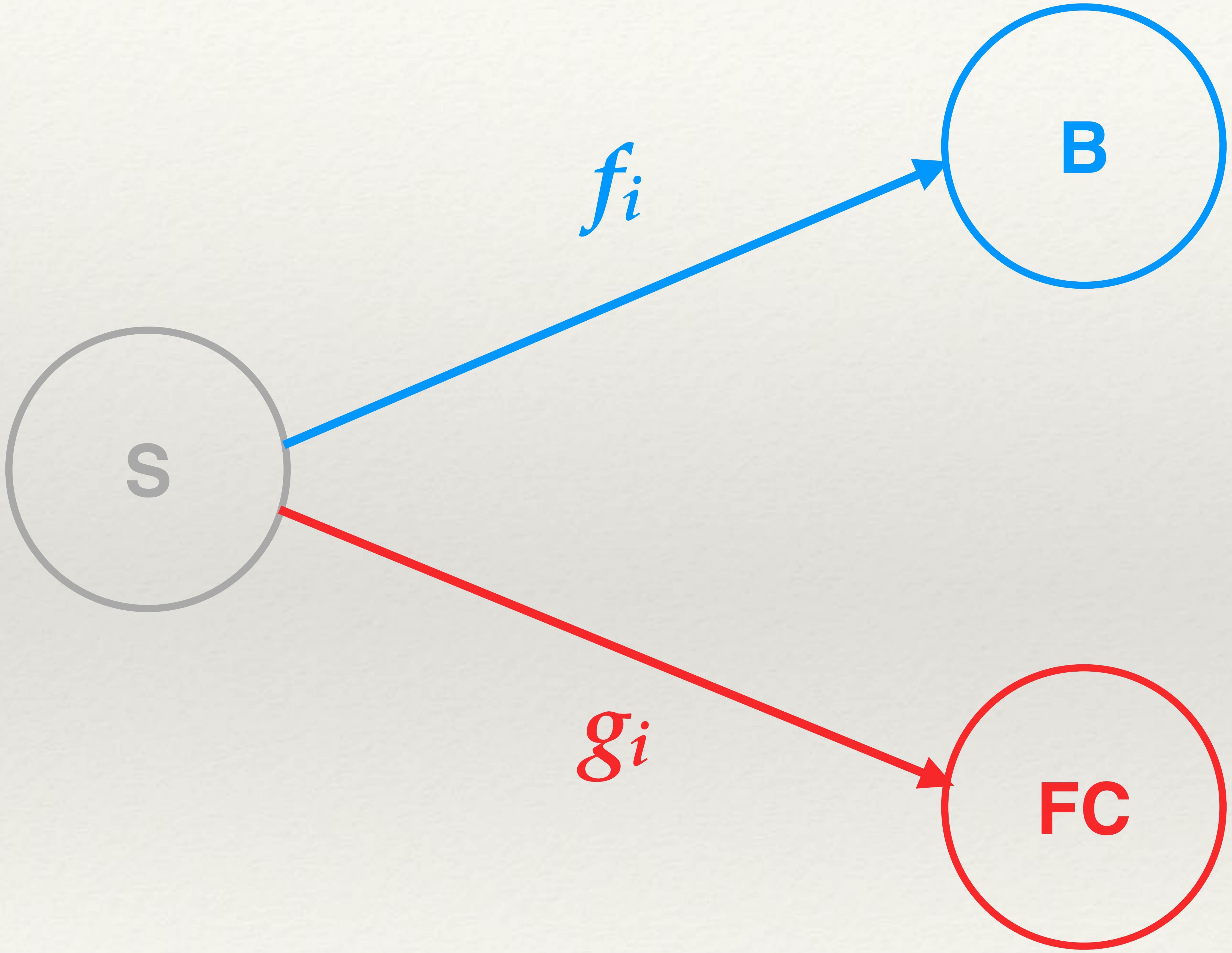


neighbors of i: $\boldsymbol{n_i}$

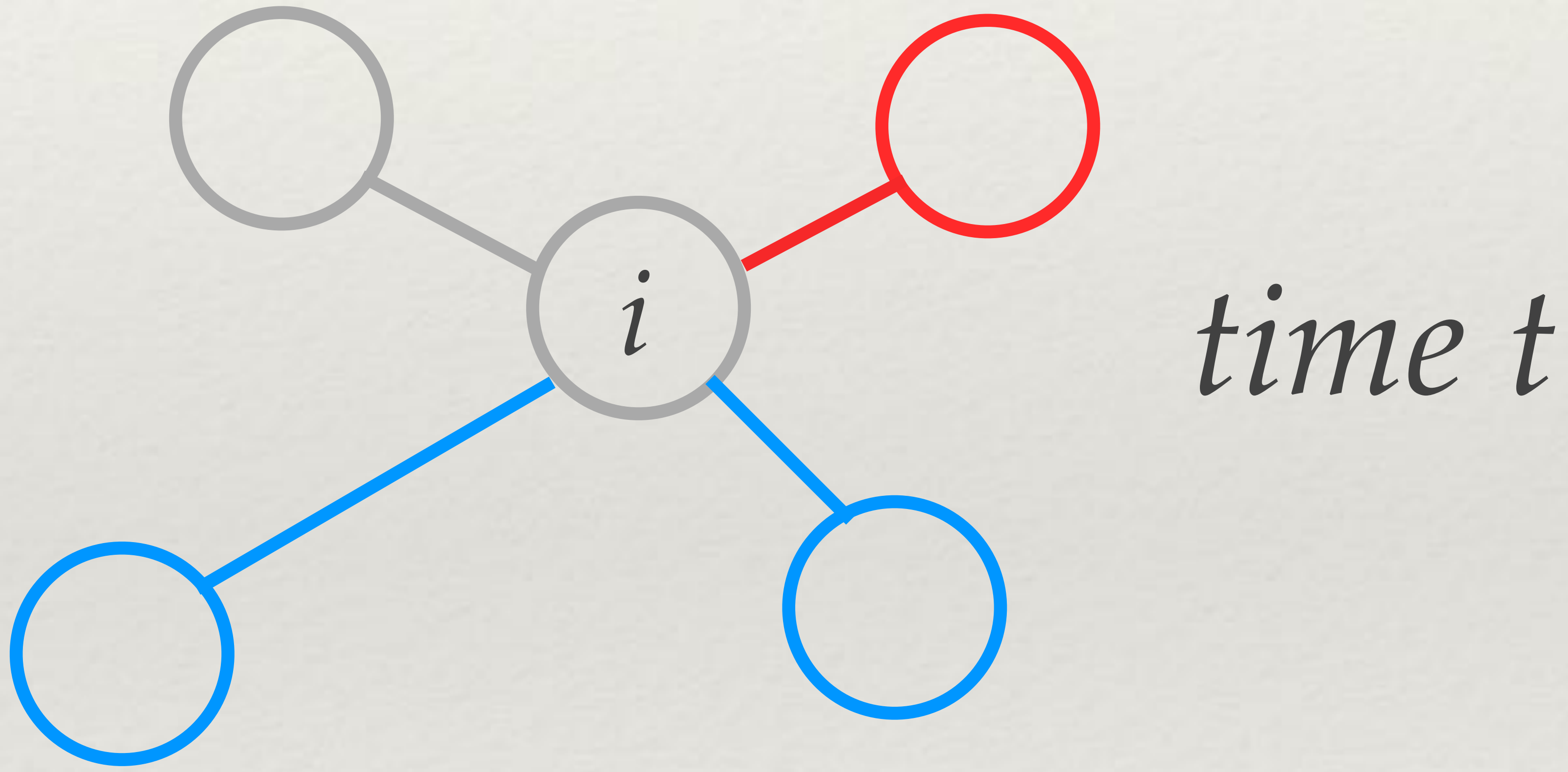
credibility of the hoax: $\boldsymbol{\alpha}$

spreading rate: $\boldsymbol{\beta}$

From Susceptible to Believer/Fact-Checker

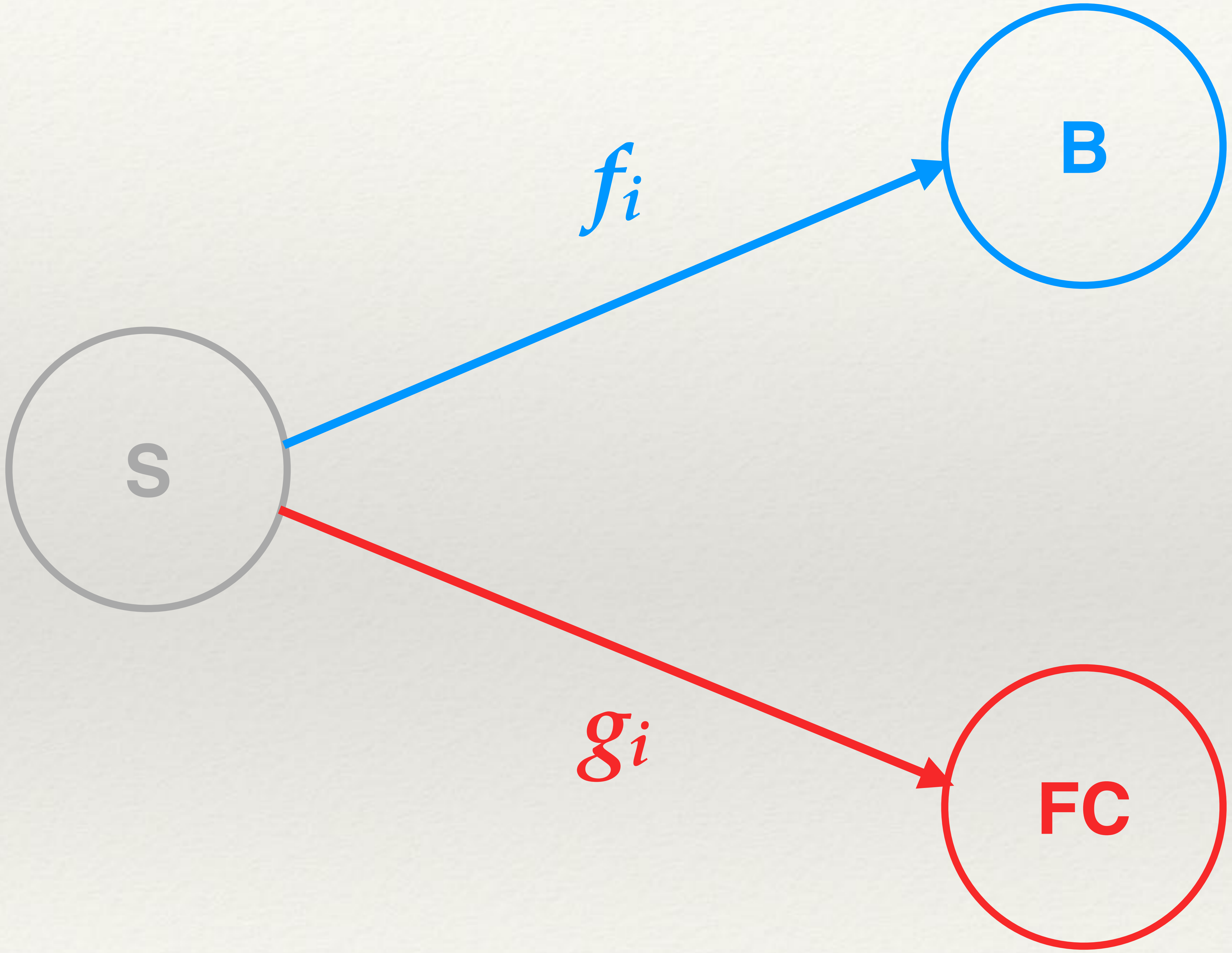


$$f_i(t) = \beta \frac{n_i^B(t)(1+\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$

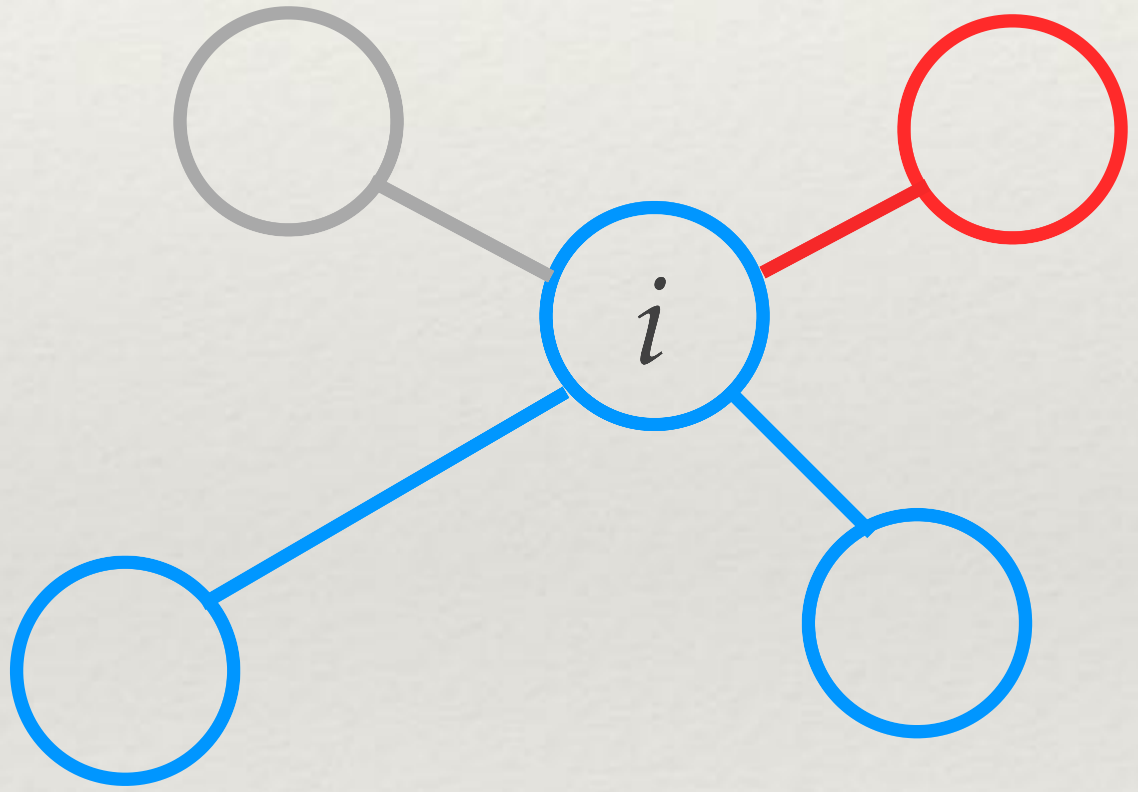


$$g_i(t) = \beta \frac{n_i^F(t)(1-\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$

From Susceptible to Believer/Fact-Checker



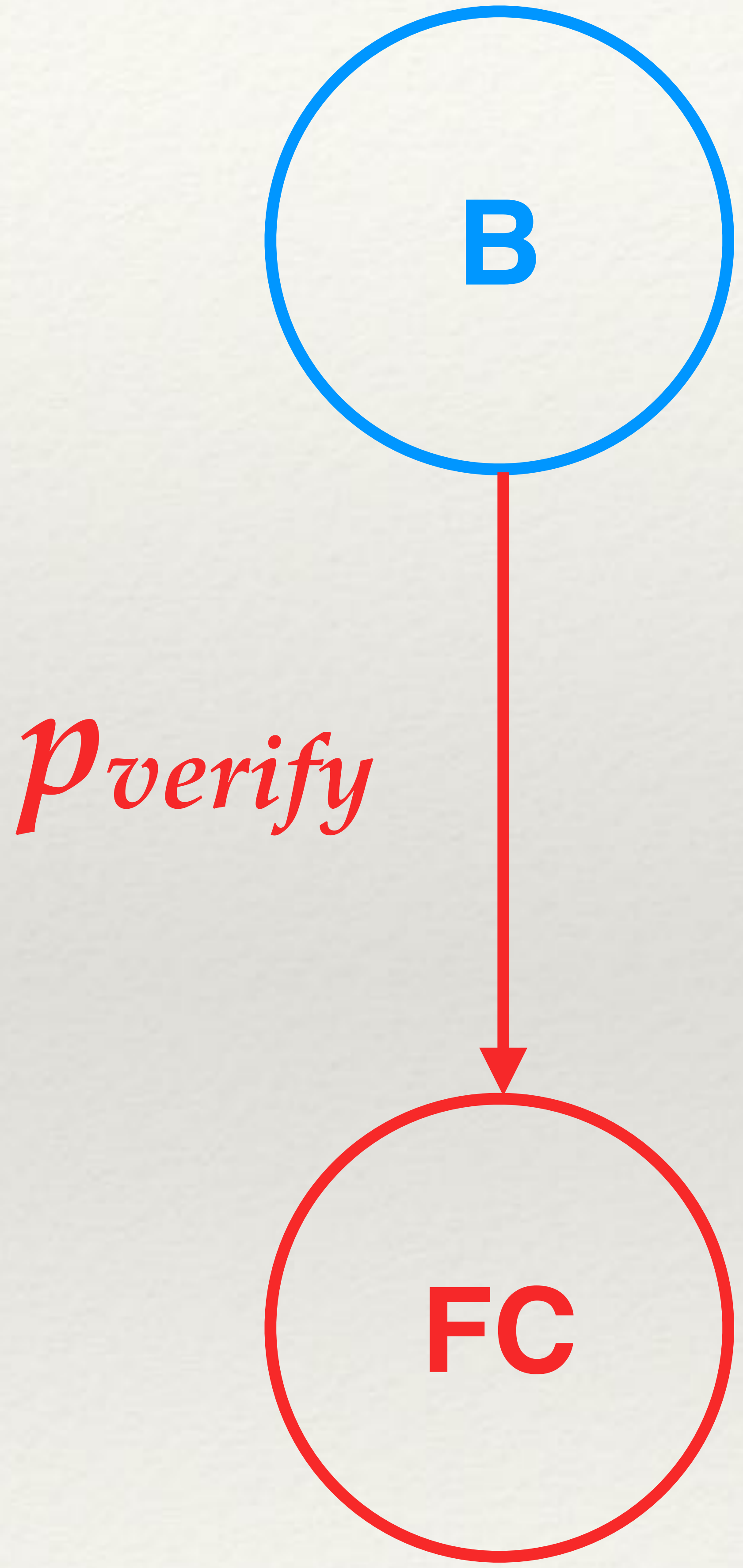
$$f_i(t) = \beta \frac{n_i^B(t)(1+\alpha)}{n_i^B(t)(1+\alpha) + n_i^F(t)(1-\alpha)}$$



time t+1

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

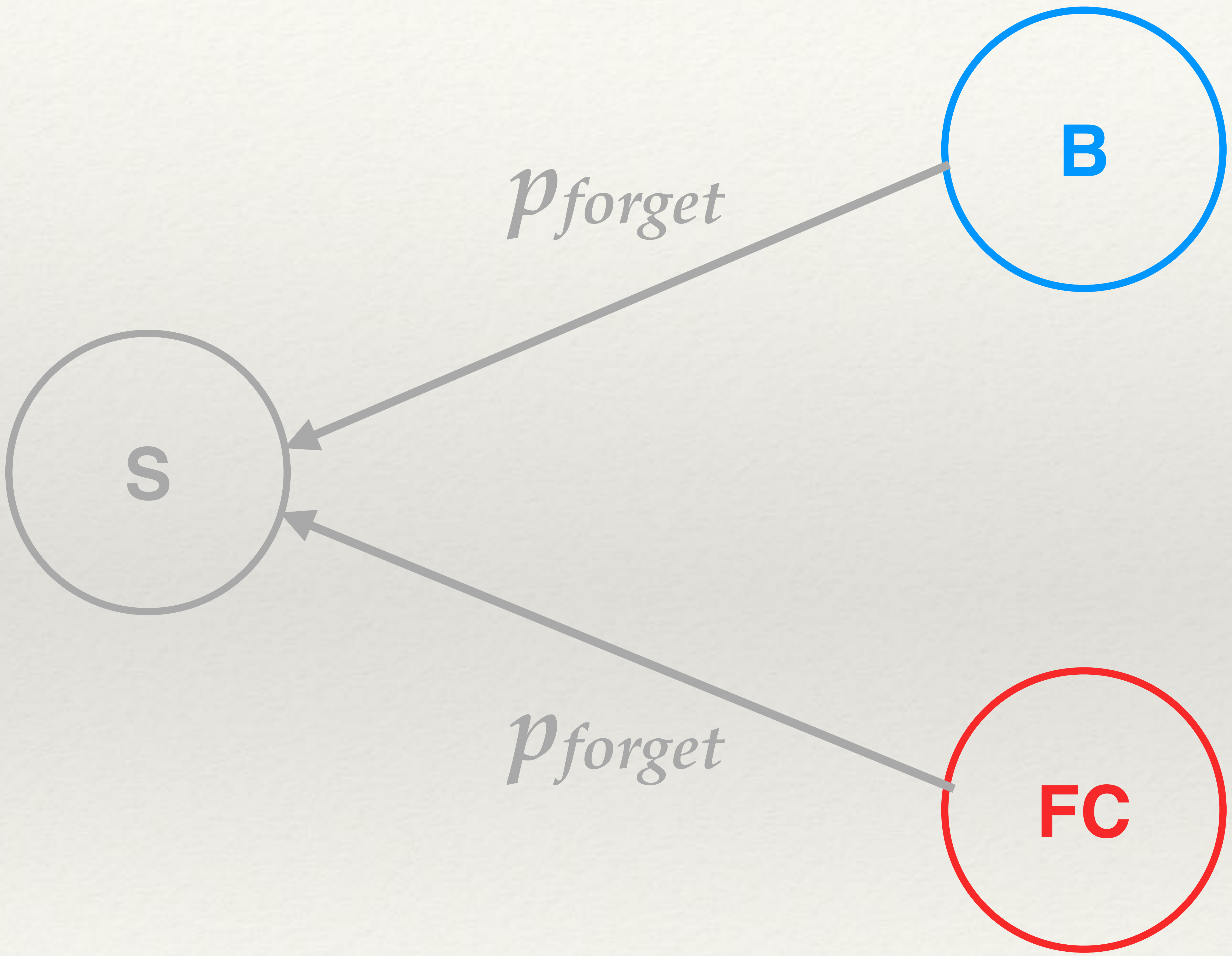
From Believer to Fact-Checker



VERIFYING

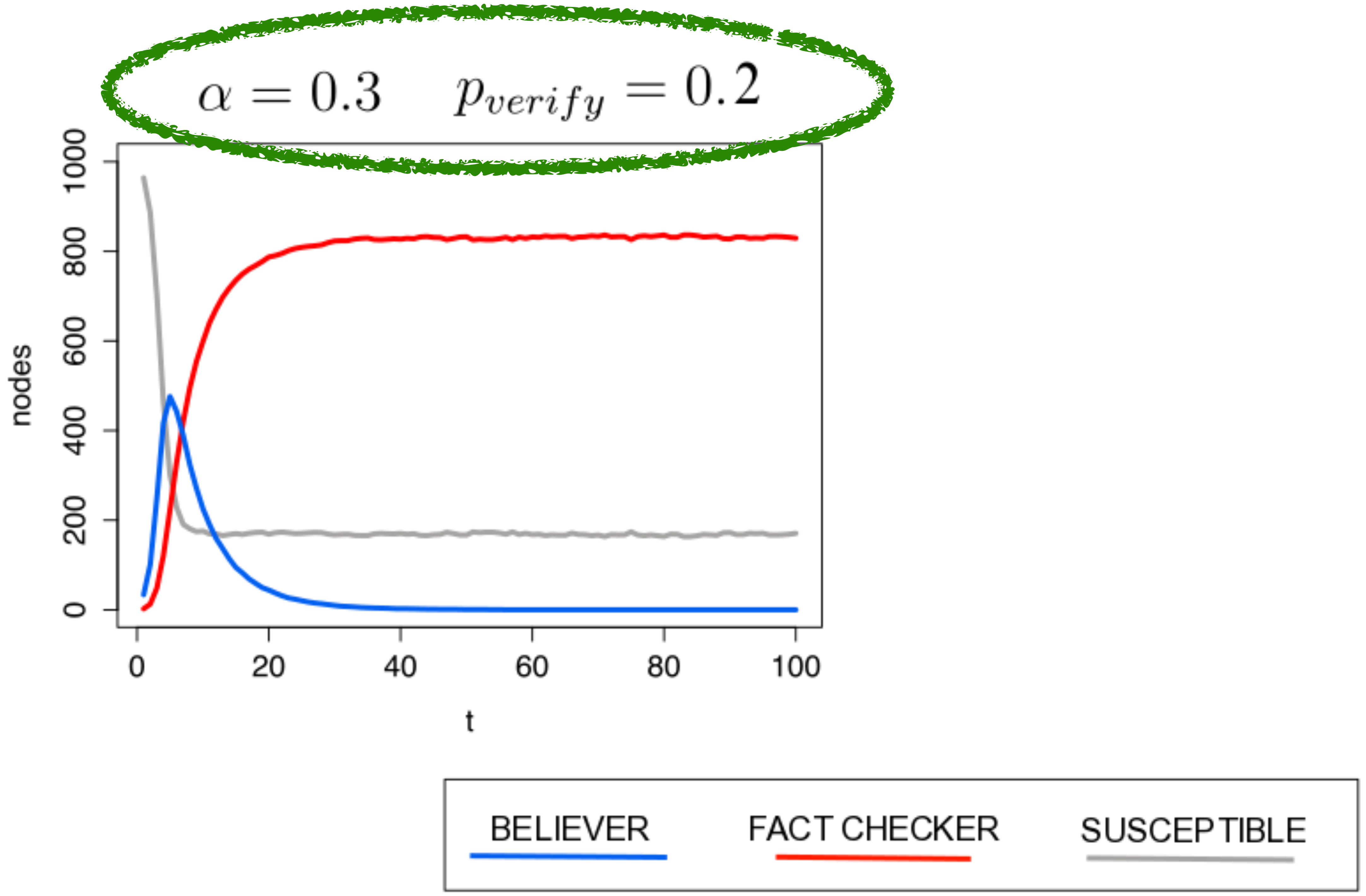
probability of fact-checking (or just deciding not to believe)

From Believer/Fact-Checker to Susceptible

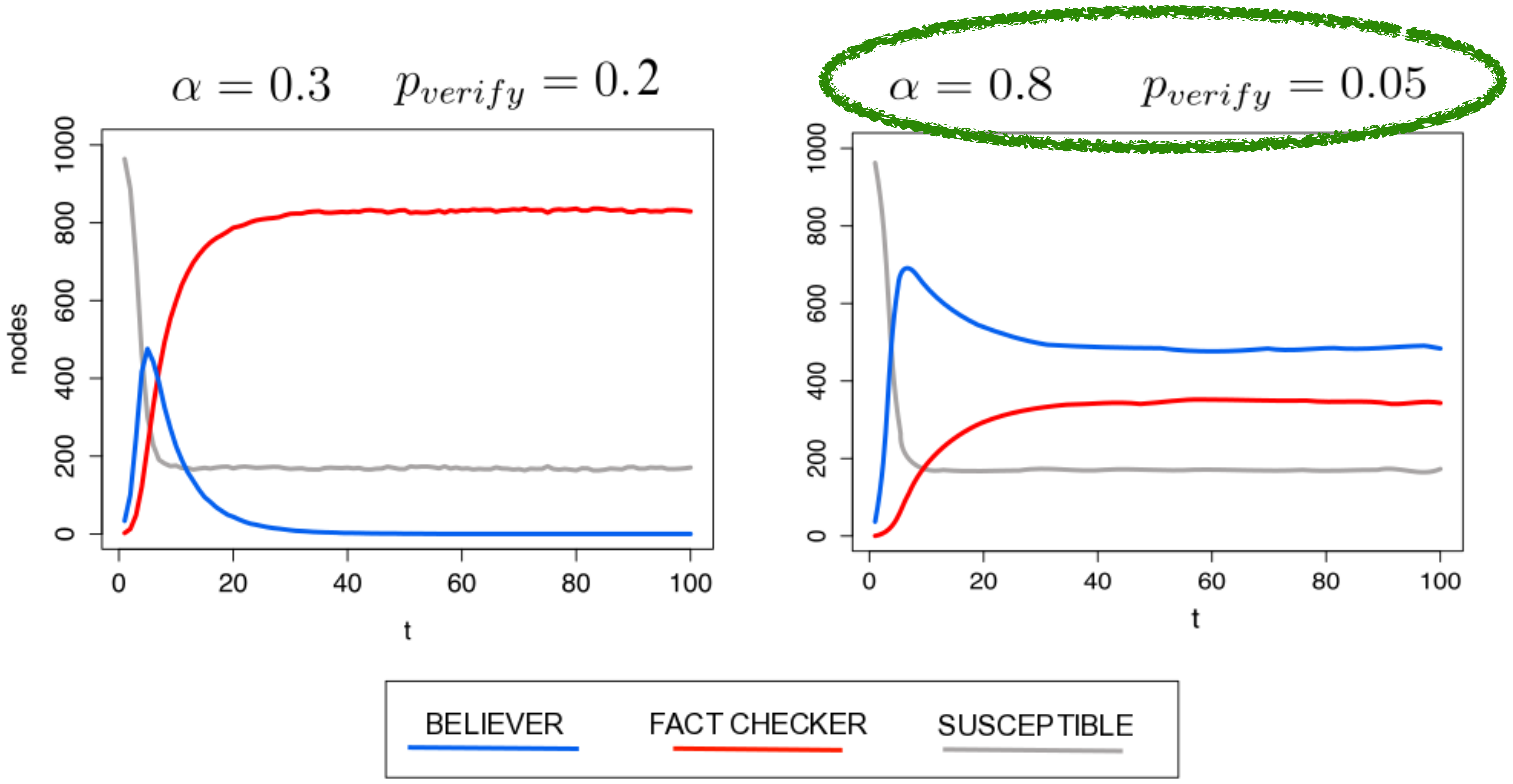


FORGETTING

Dynamics (agent-based simulations)



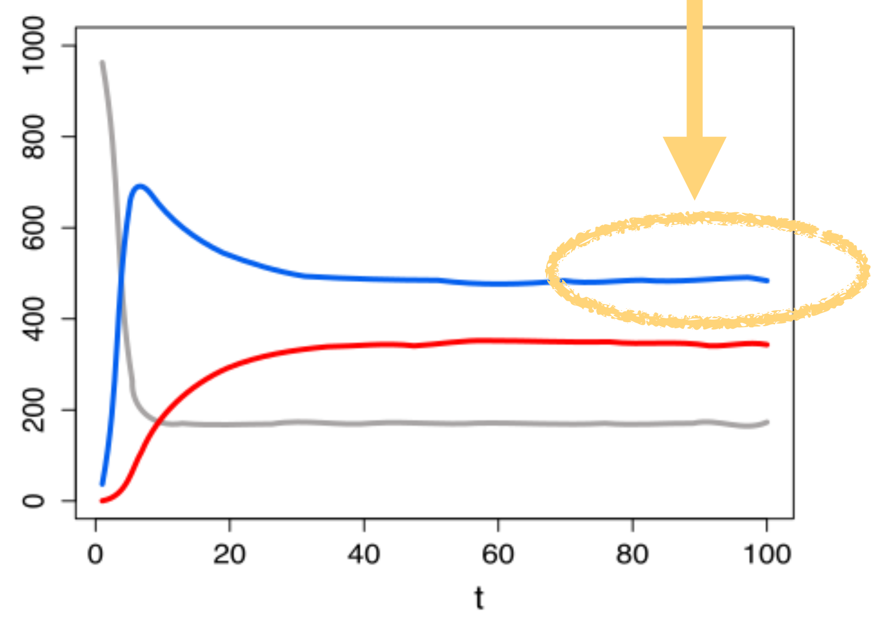
Dynamics (agent-based simulations)

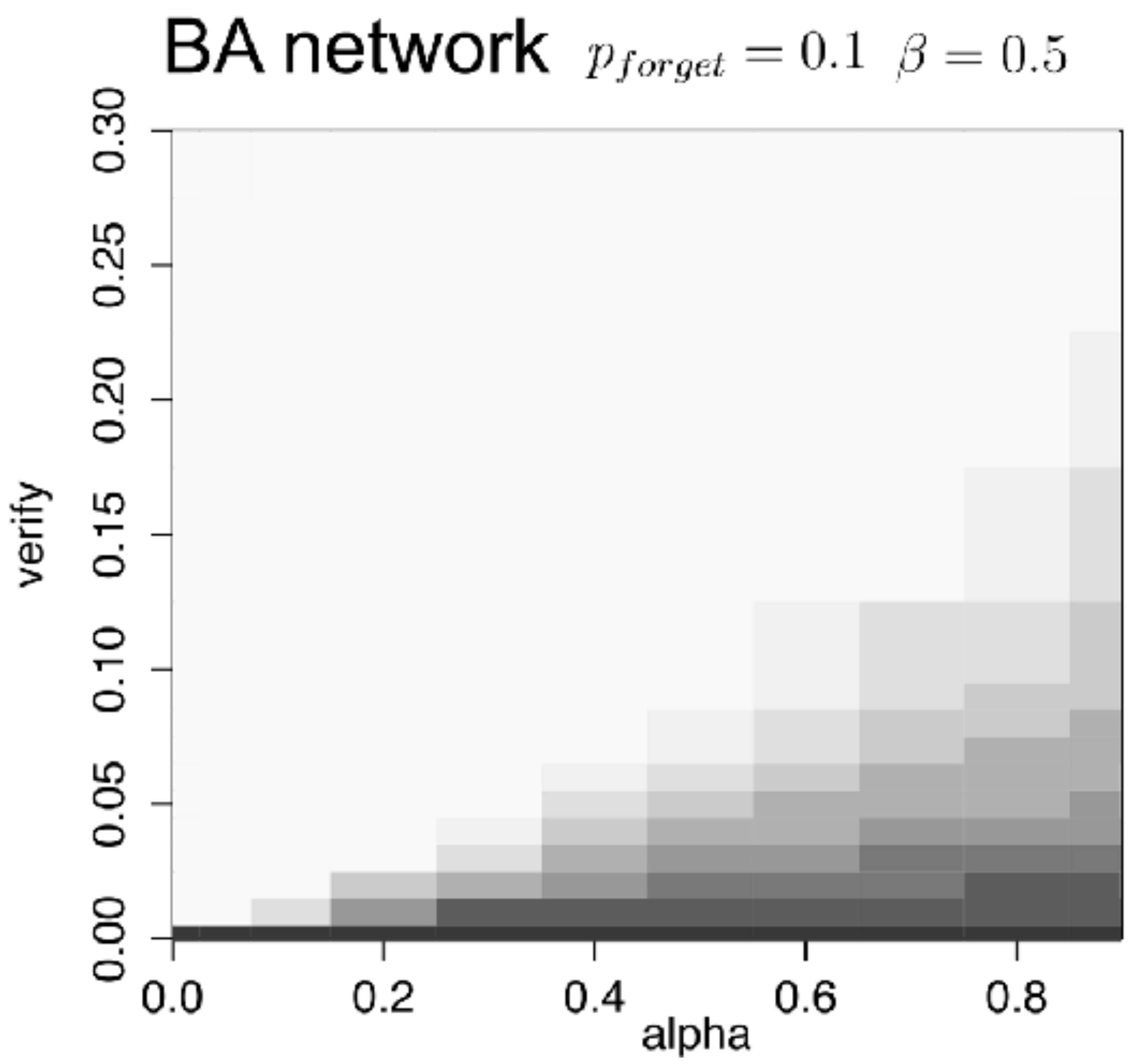


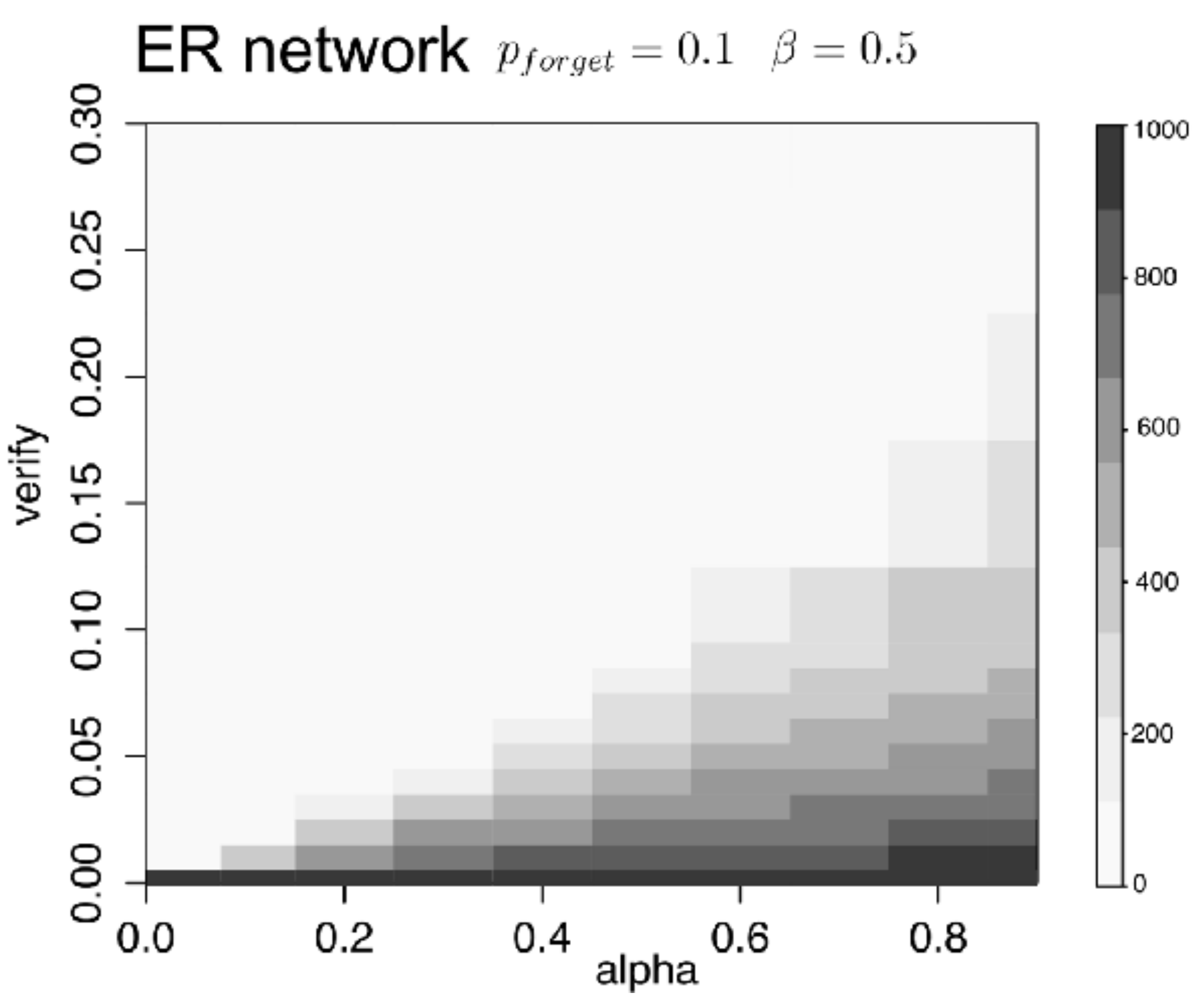
hoax **credibility** and **fact-checking probability** rule hoax persistence in the network

Dynamics (agent-based simulations)

number of 'believers' at the equilibrium





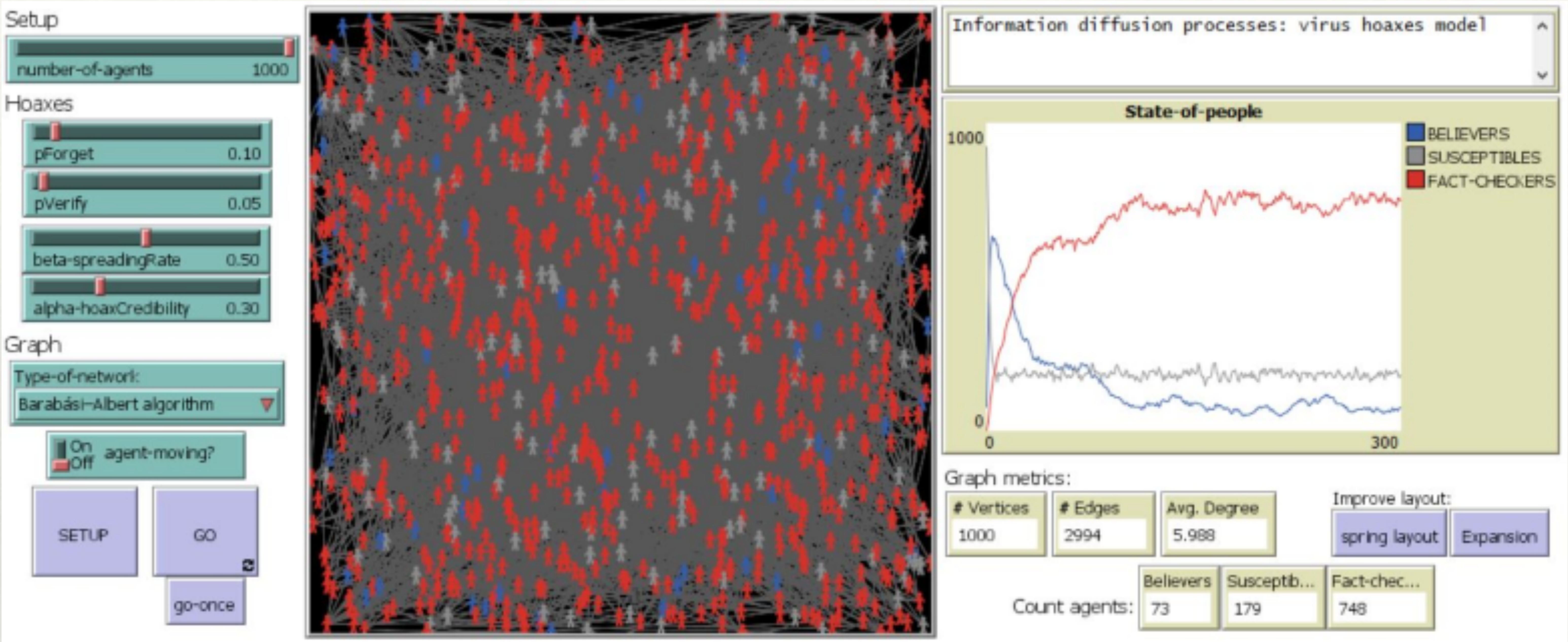


First step toward "good practices" understanding

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)

NetLogo simulator

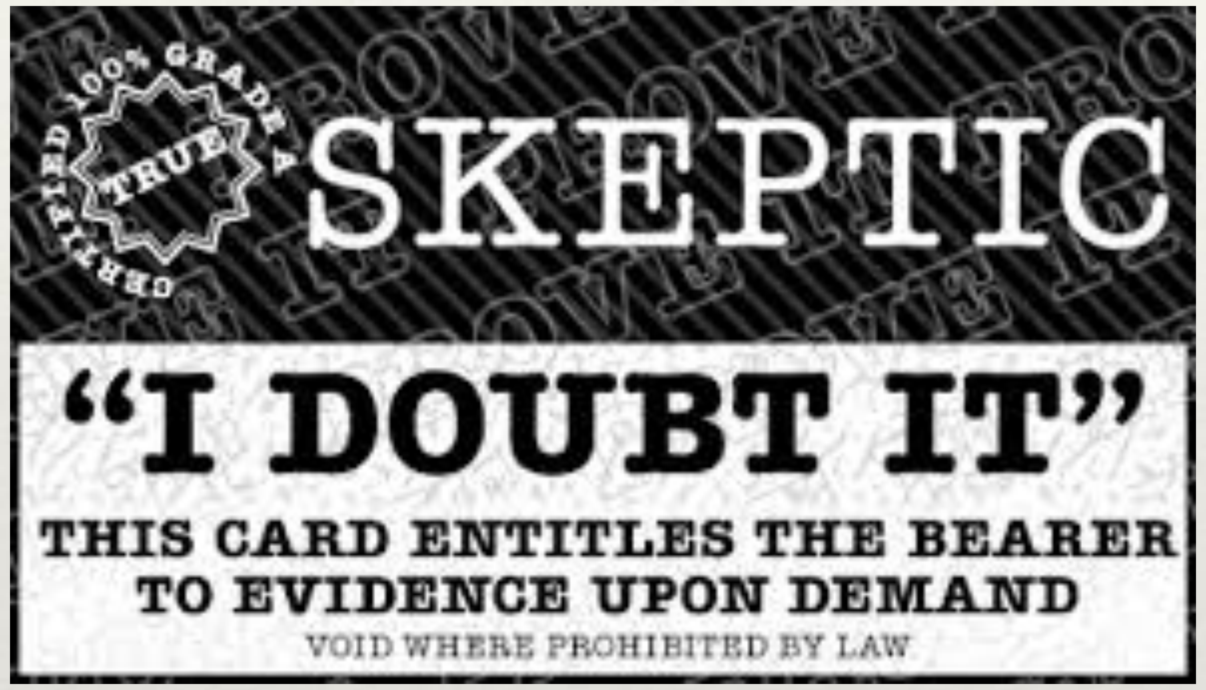


http://modelingcommons.org/browse/one_model/6403#model_tabs_browse_info

Emilio Sulis and Marcella Tambuscio, Simulation of misinformation spreading processes in social networks: an application with NetLogo, in Proc. of IEEE DSAA2020 - Special Session "Fake News, bots, and trolls"

The role of segregation

Skeptical and gullible agents



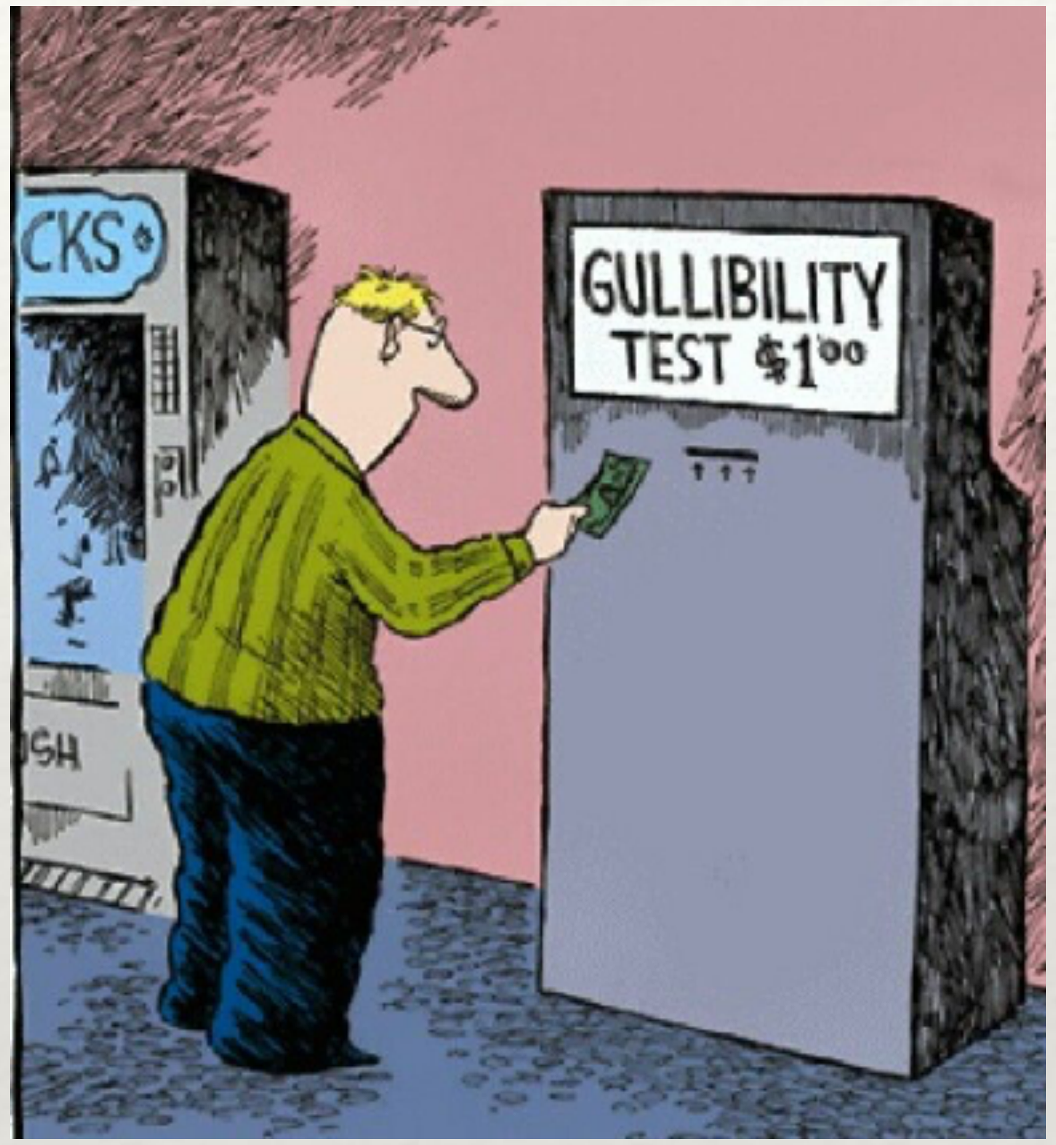
let's tune **credibility** accordingly

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

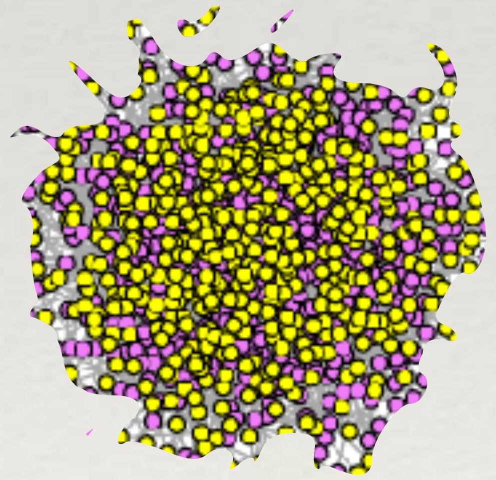
segregation ($0.5 <$ **s** < 1)

fraction of edges within same community

[Gu-Gu, Sk-Sk]

Gullible

α **large**

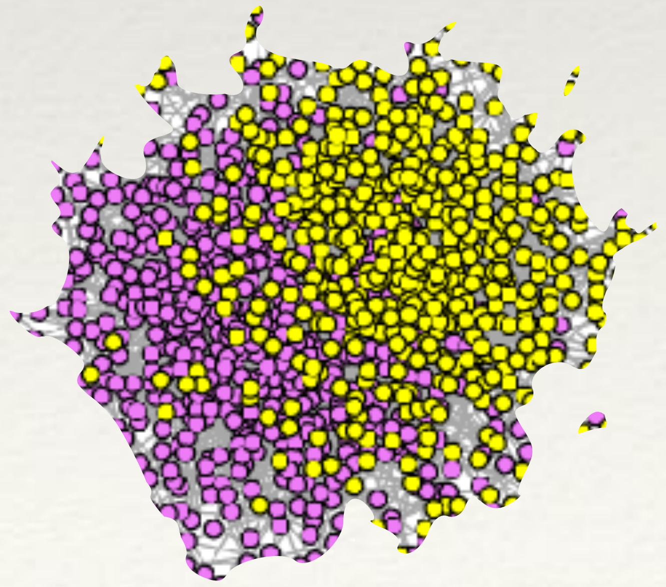


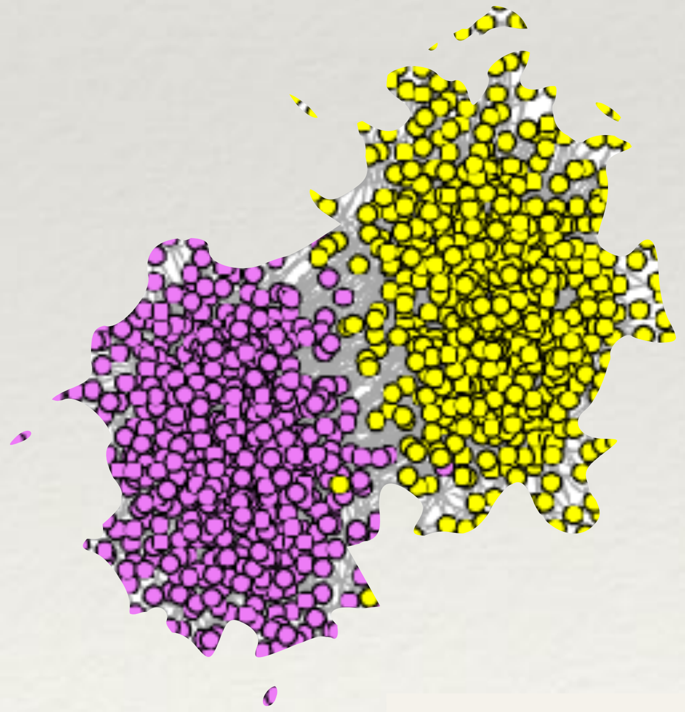
s=0.55

γ=500

s=0.8

γ=500

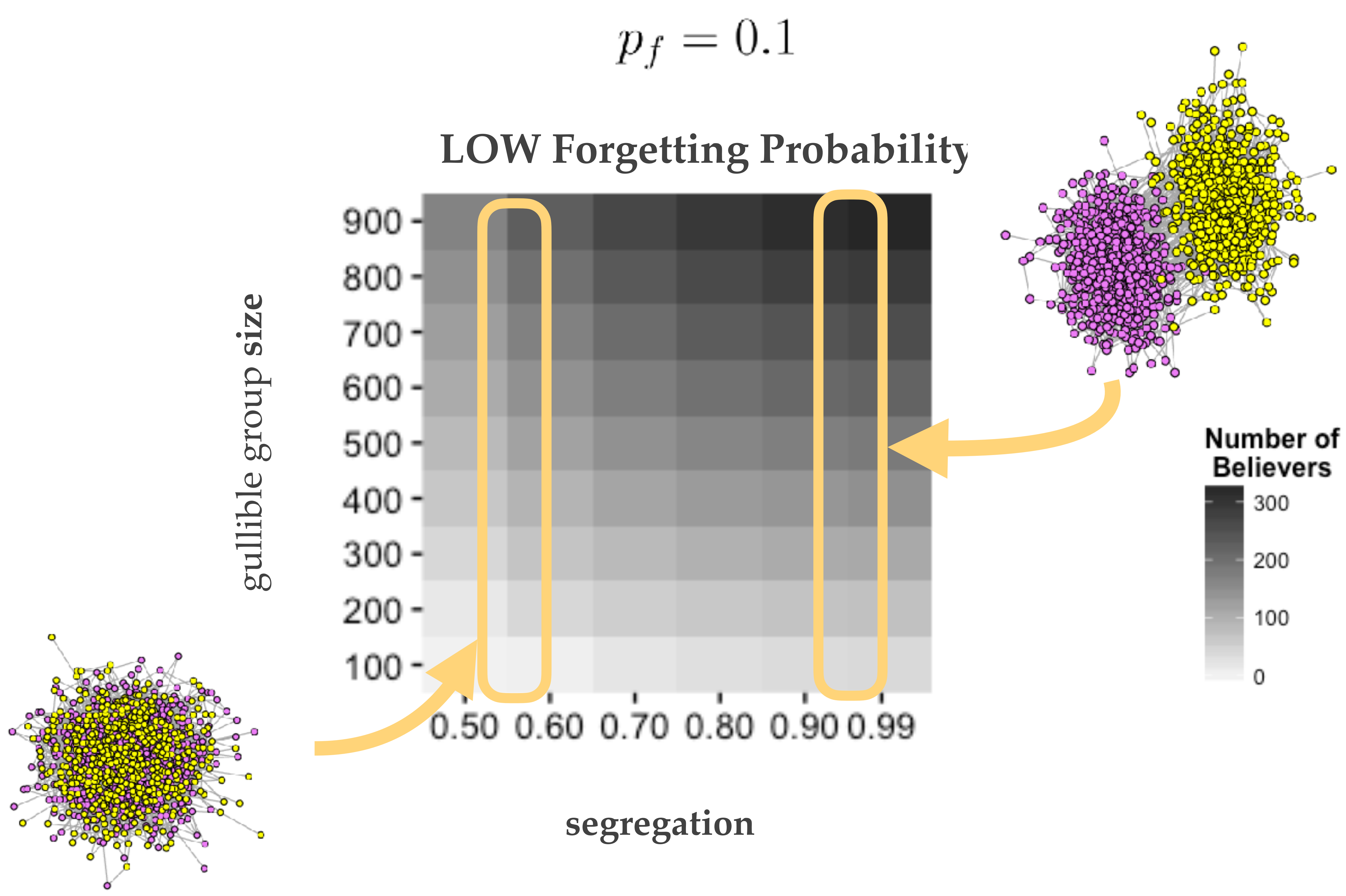




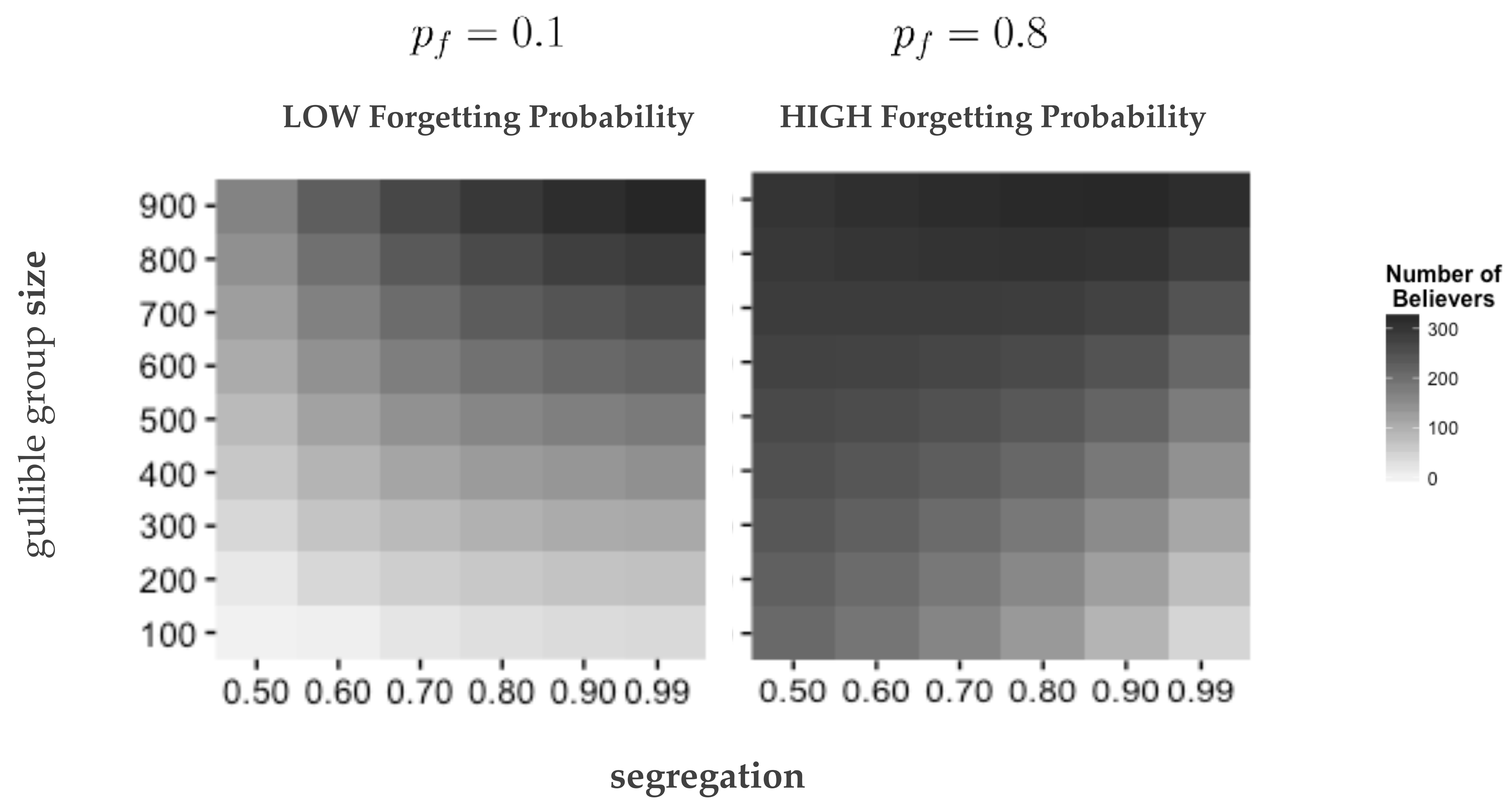
s=0.95

γ=500

Size vs segregation



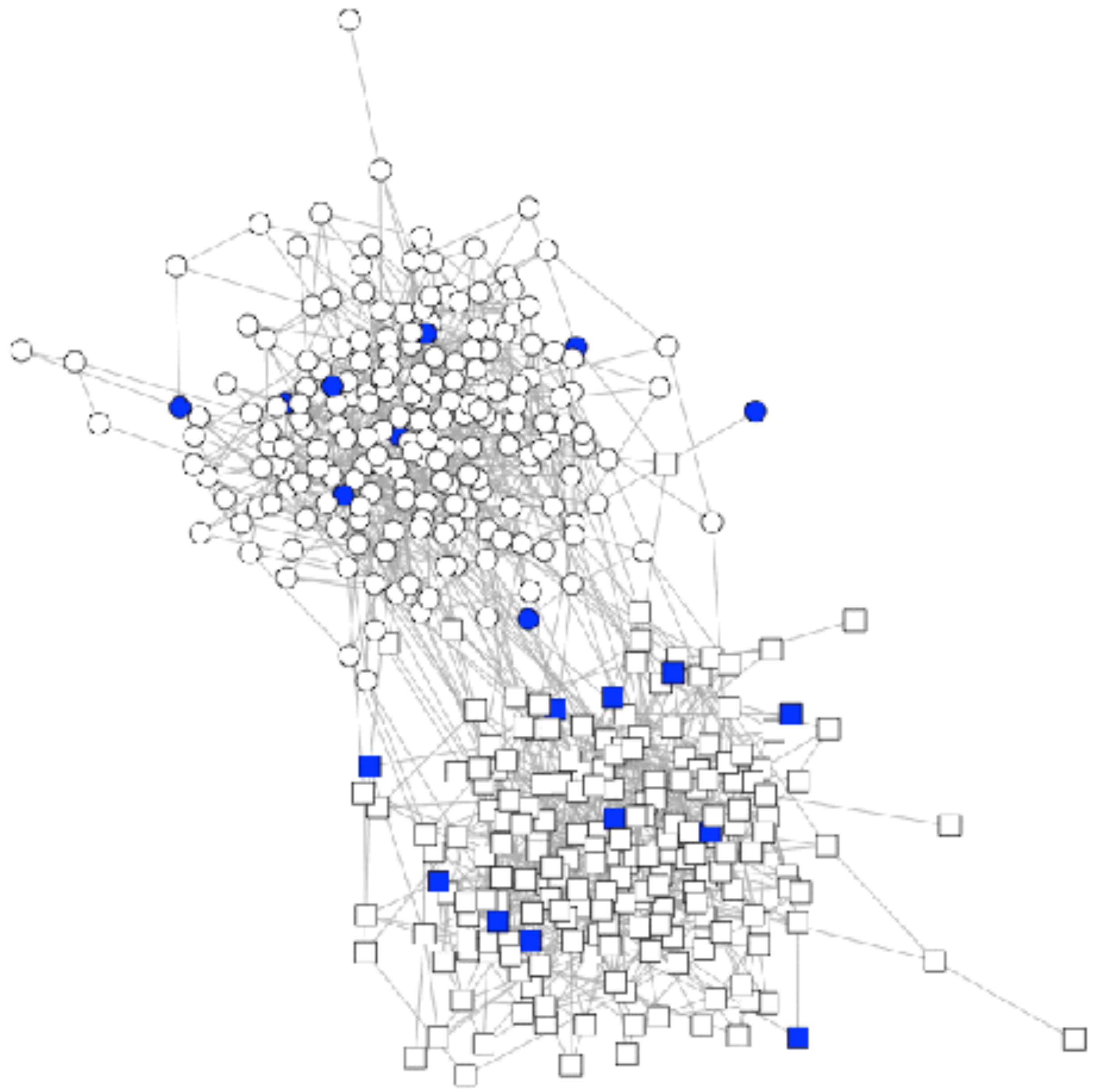
Size vs segregation



Role of forgetting

LOW Forgetting Rate

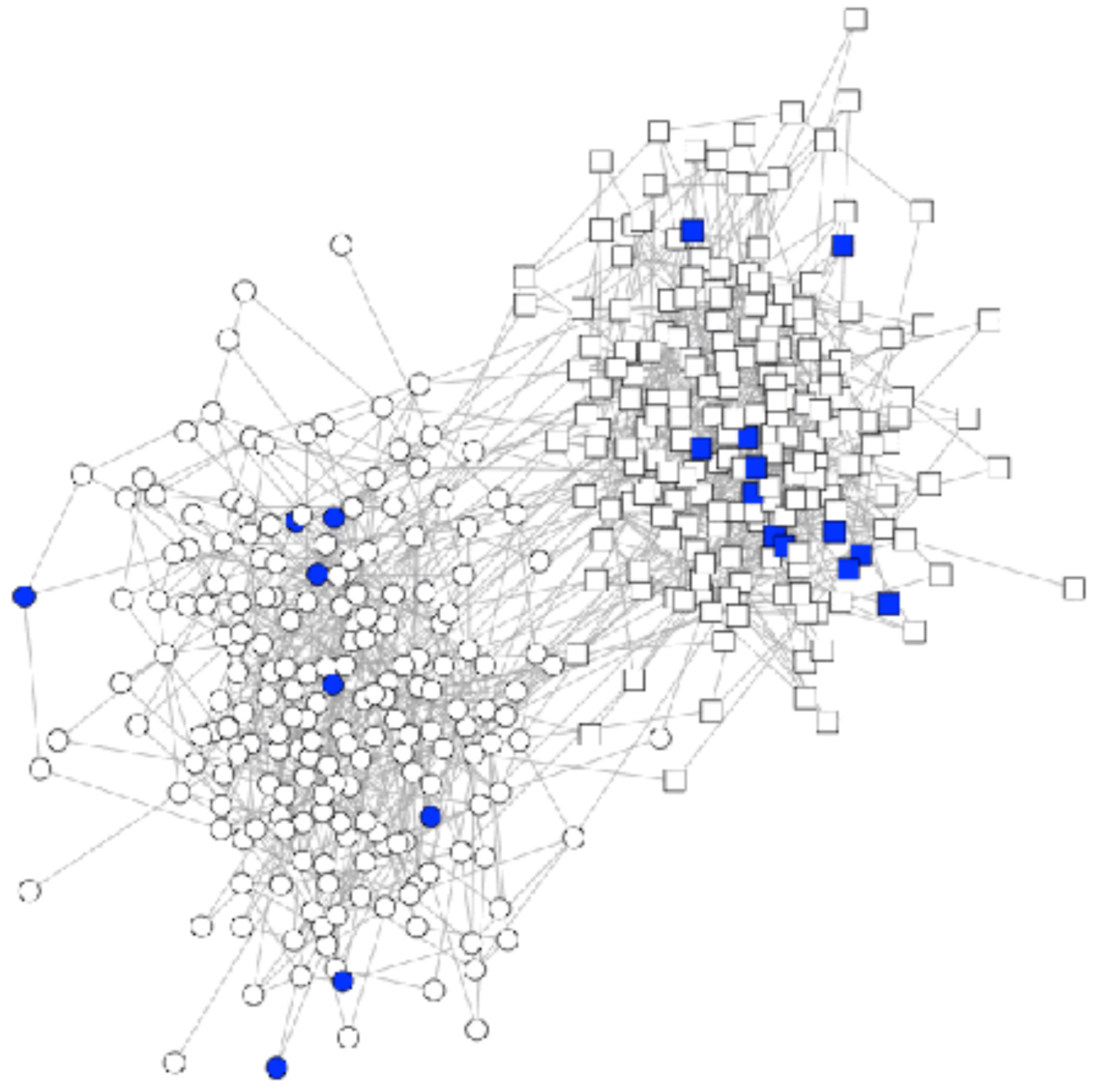
$$p_f = 0.1$$



Time = 1

HIGH Forgetting Rate

$$p_f = 0.8$$



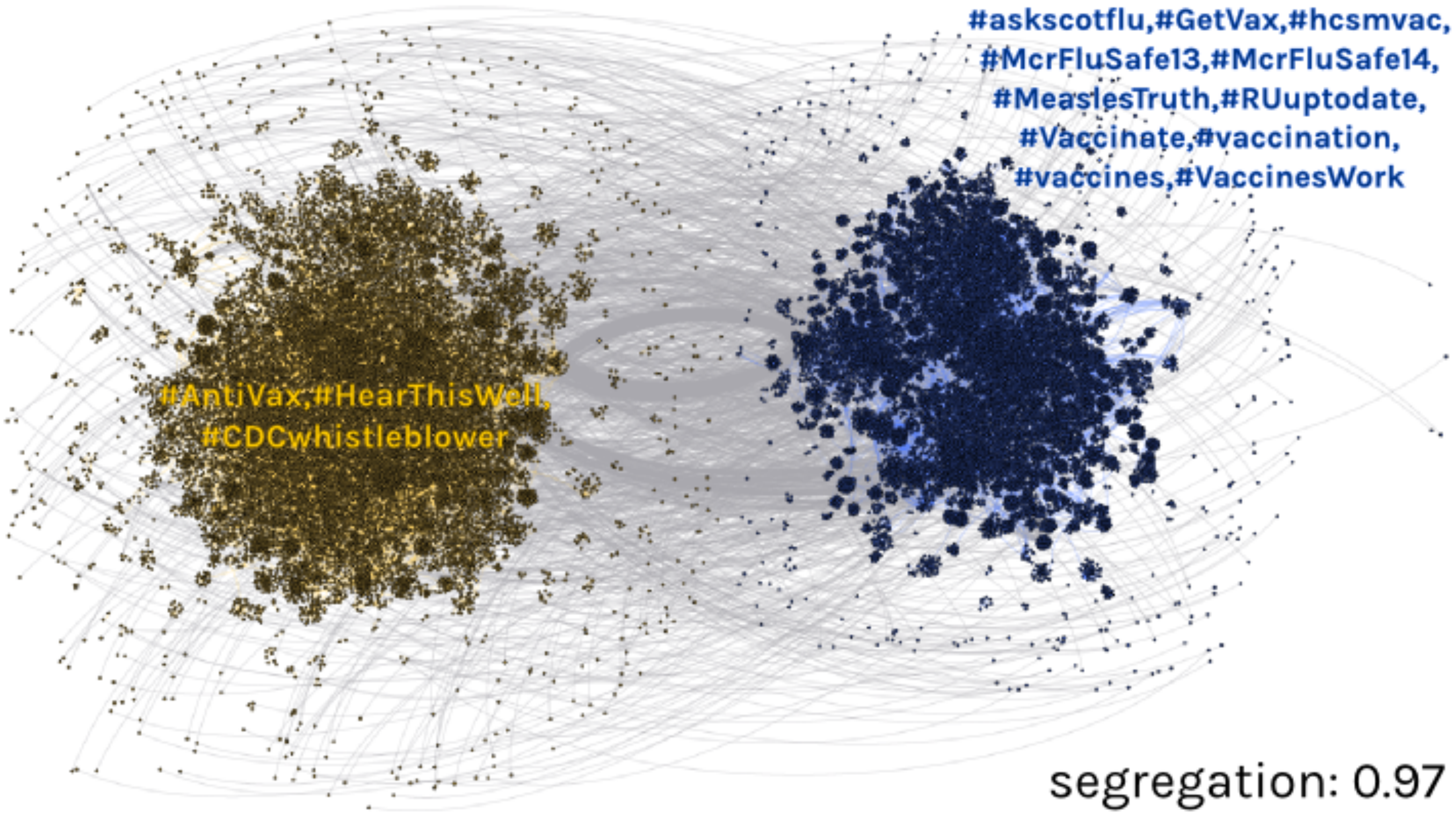
Time = 1

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**

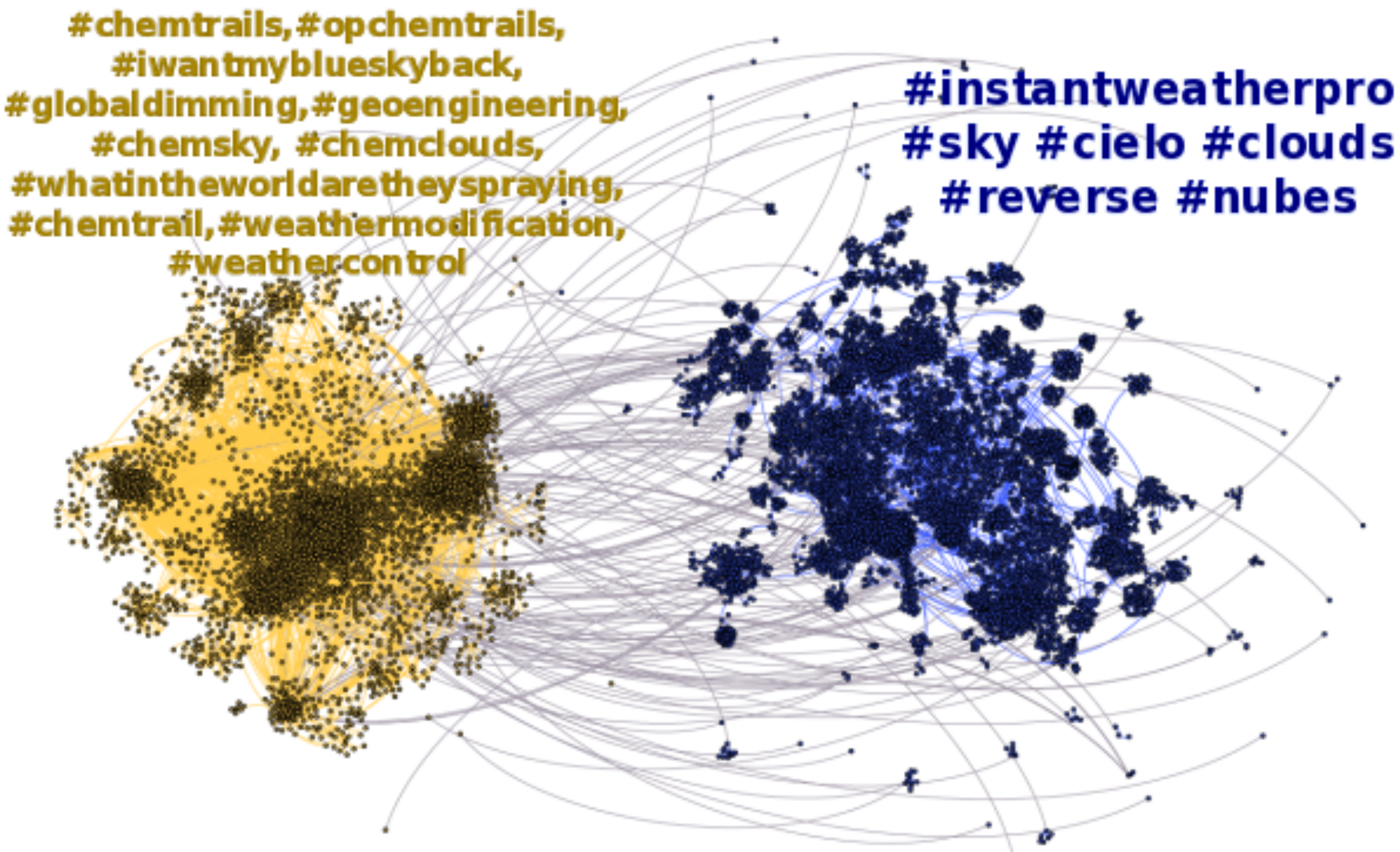
M Tambuscio, D F M Oliveira, G L Ciampaglia, G Ruffo, Network segregation in a model of misinformation and fact-checking, Journal of Computational Social Science (2018) 1: 261.

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

- 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)

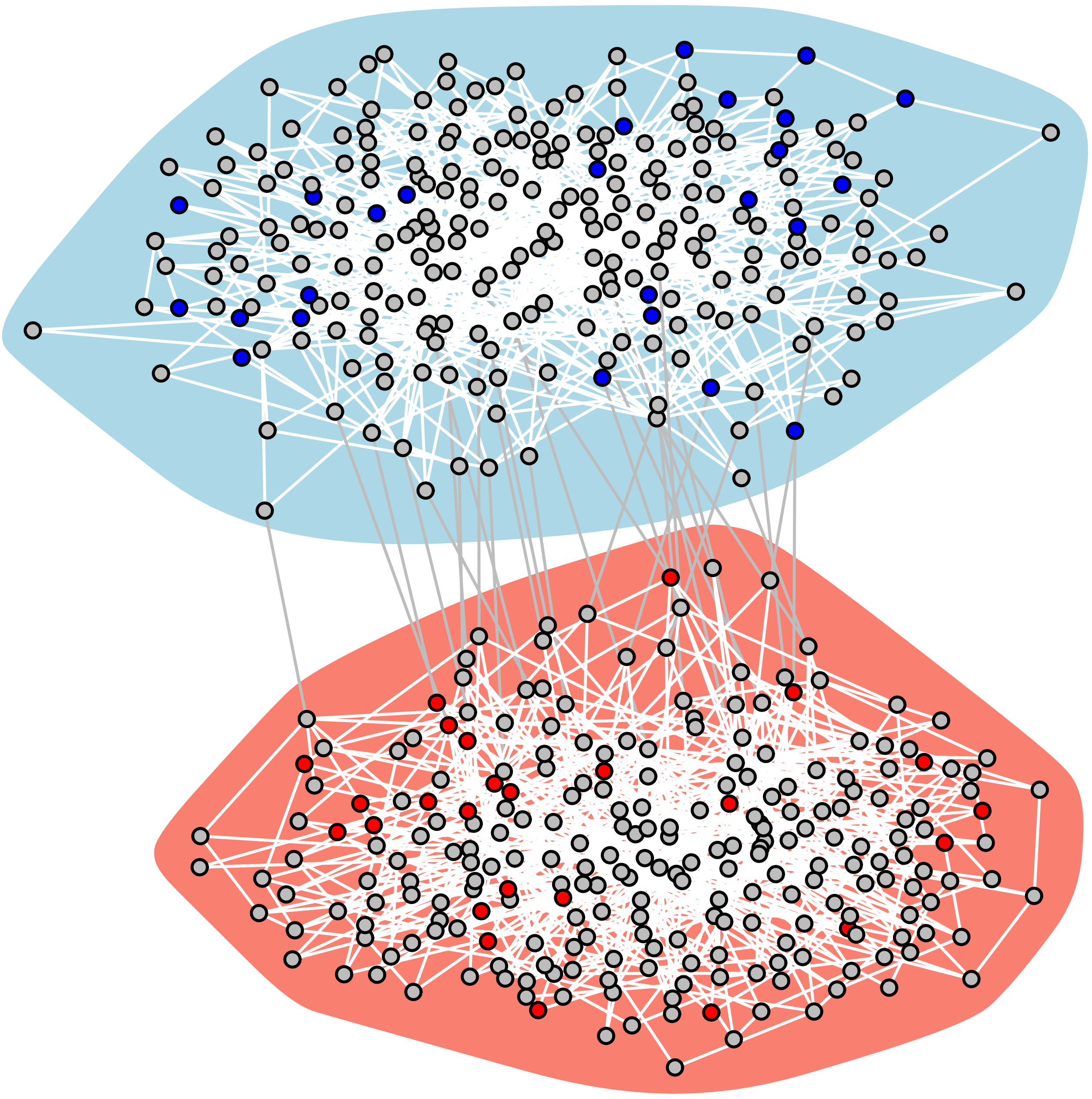
gullible group:

- α: 0.8
- seeders B: 10%

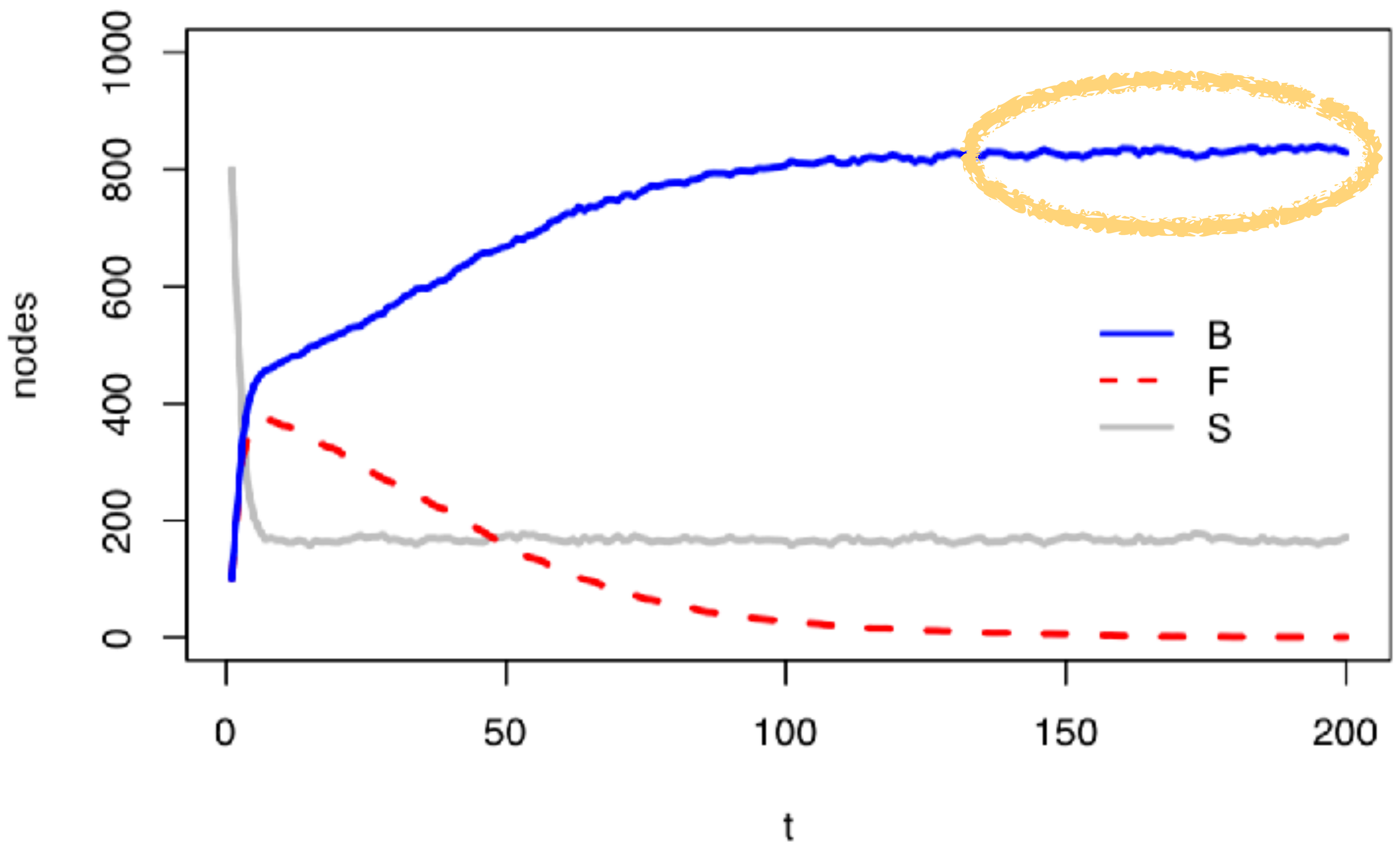
skeptical group:

- 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)

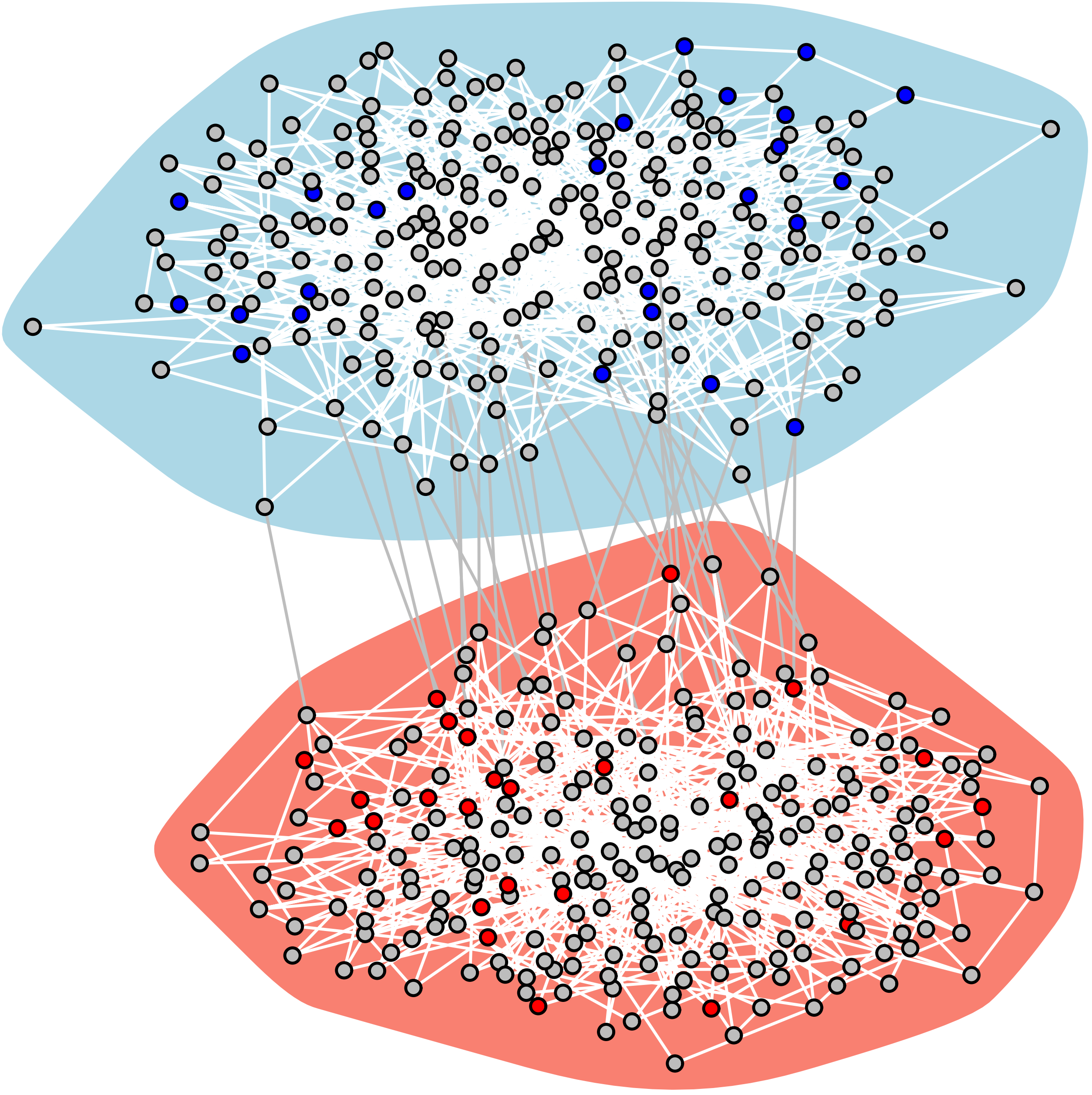
gullible group:

- α: 0.8
- seeders B: 10%

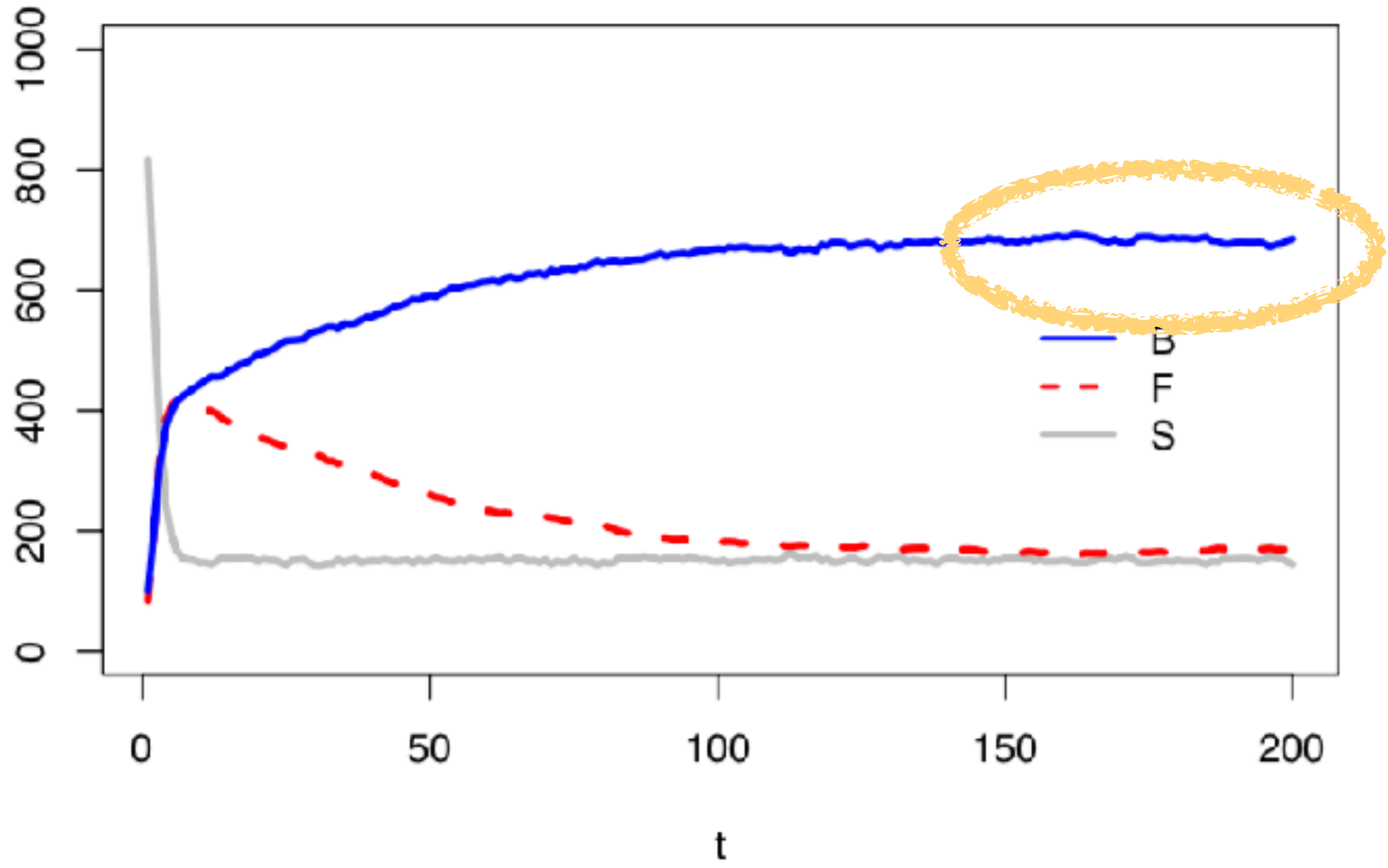
skeptical group:

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

Simulation start



Simulation results



better, but still…

Bridges as eternal fact-checker

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

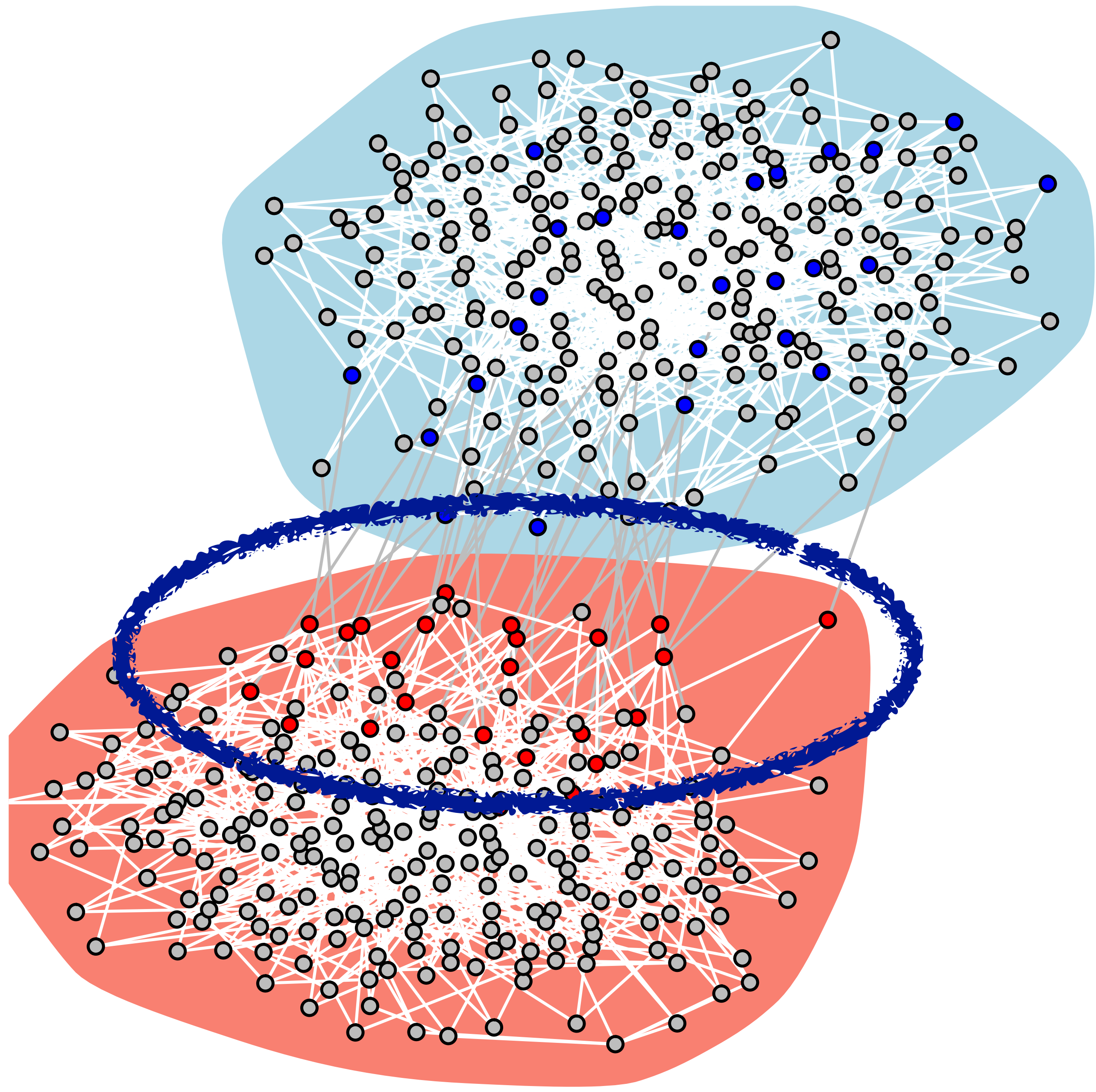
gullible group:

- α: 0.8
- seeders B: 10%

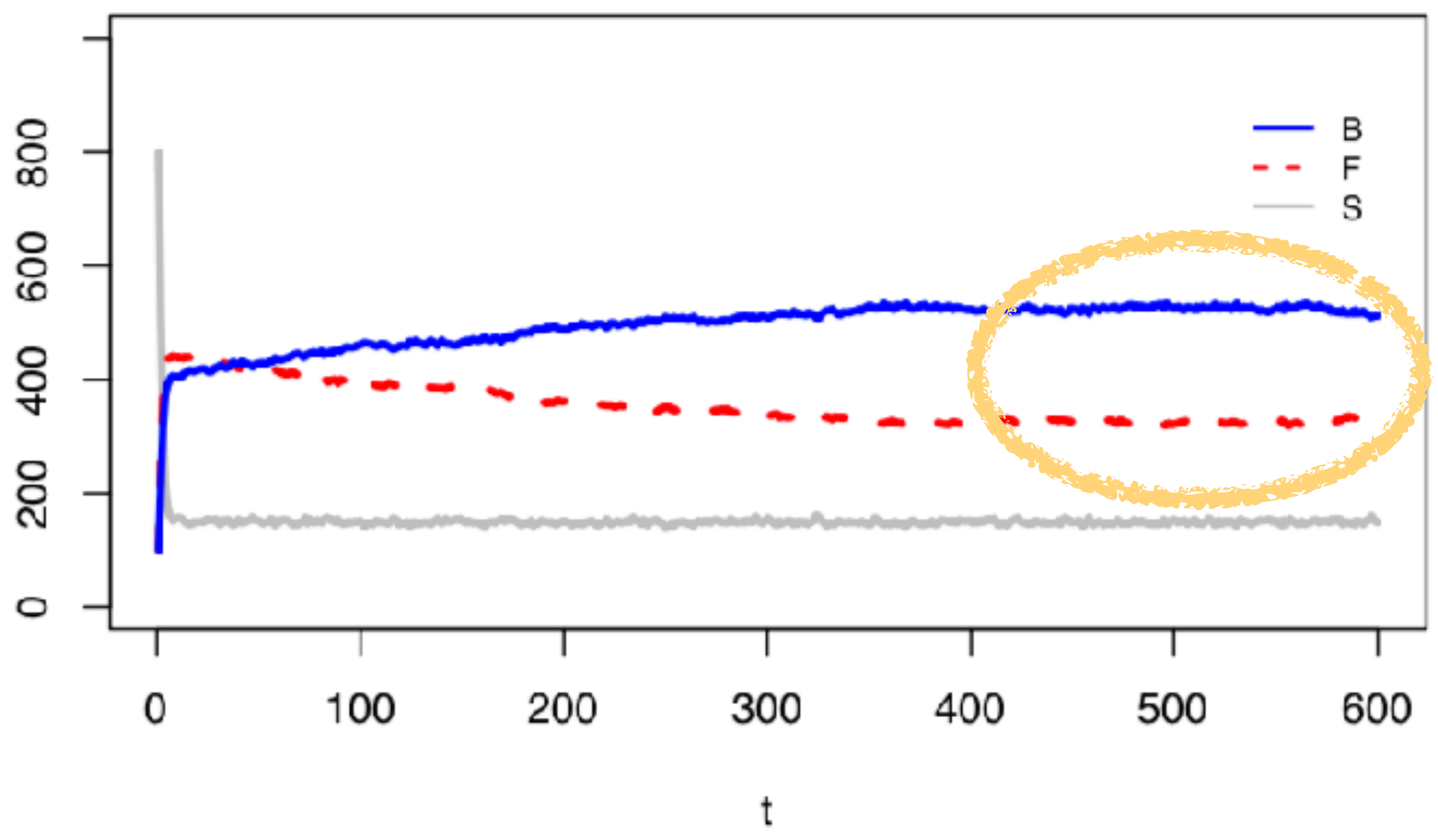
skeptical group:

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

Simulation start



Simulation results



better, but complex contagion

Hubs as eternal fact-checkers

Setting

segregation: 0.92 (high)

forgetting: 0.1 (low)

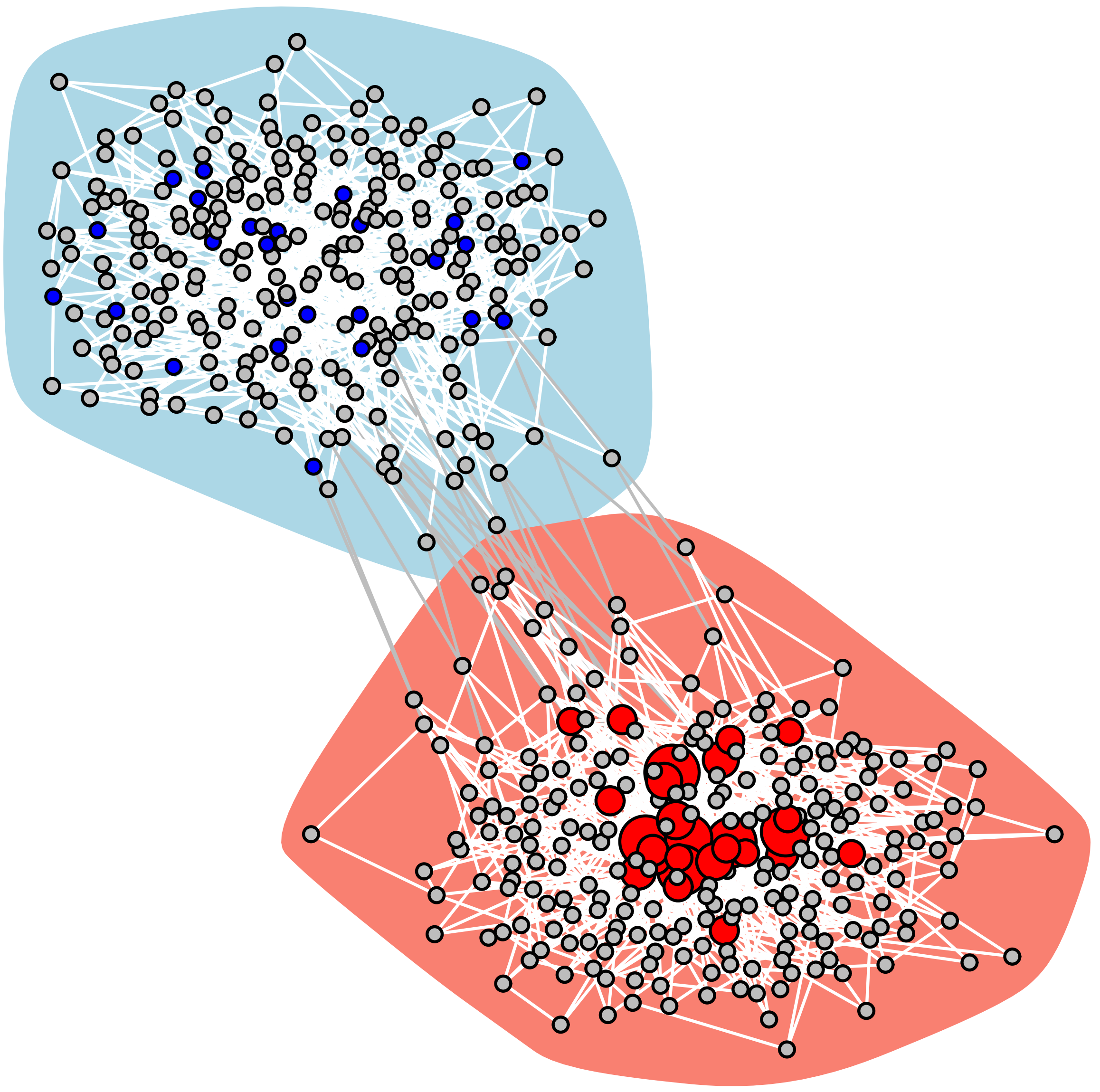
gullible group:

- α: 0.8
- seeders B: 10%

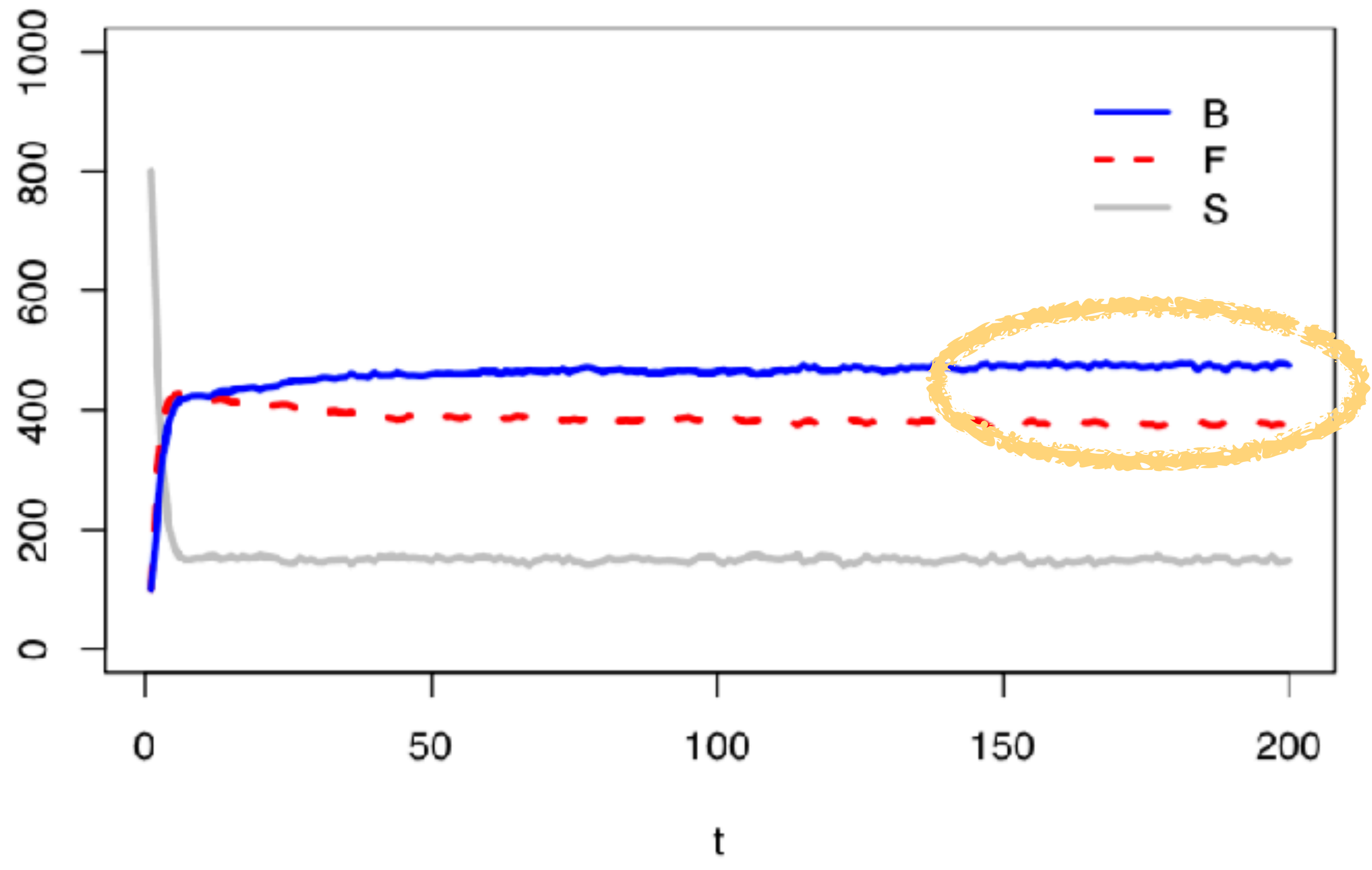
skeptical group:

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

Simulation start



Simulation results



better

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 **gatekeepers** and **local leaders**
- Note: no socio-psychological assumption so far. Real world is much more complicated

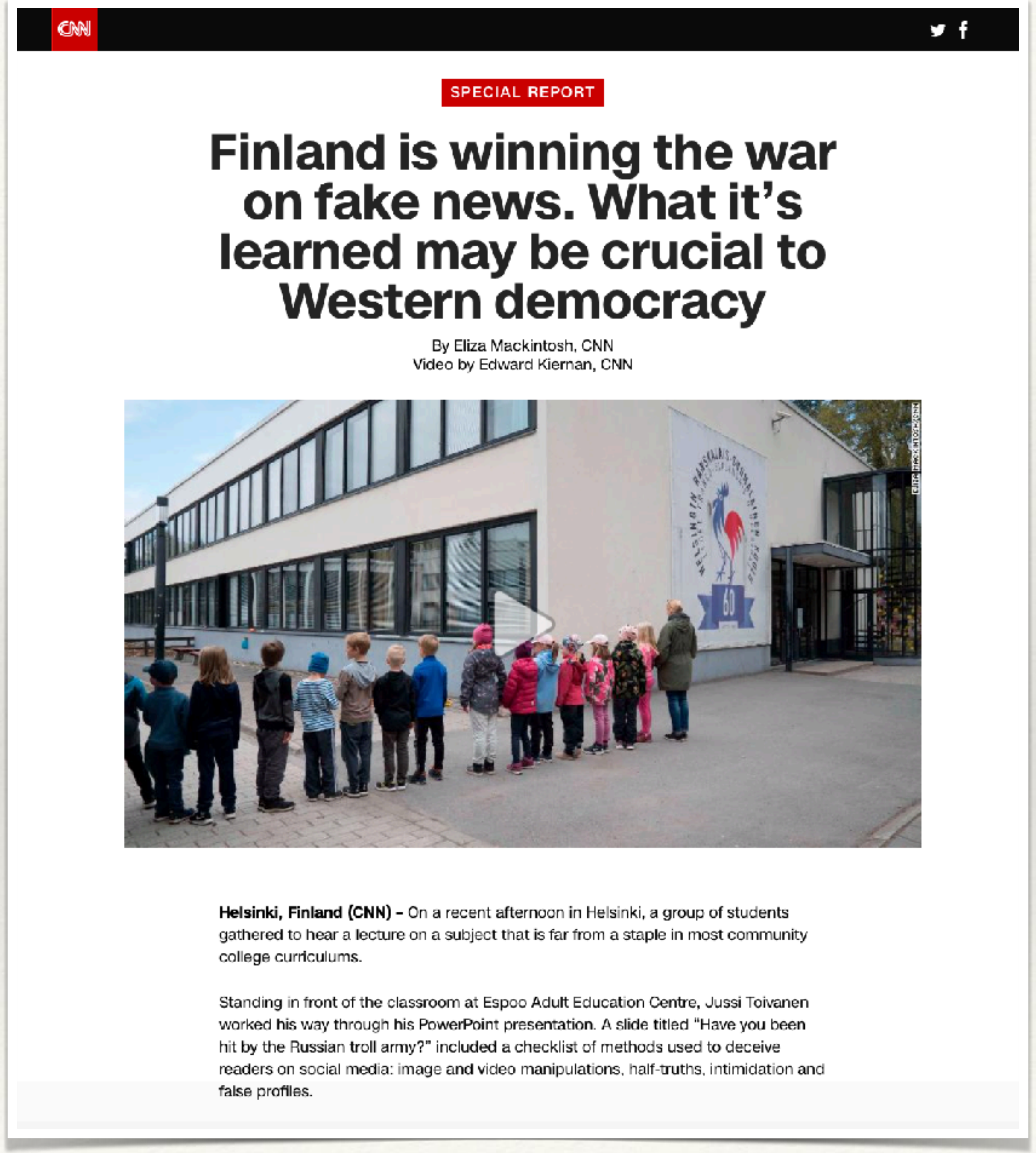
M Tambuscio, G. Ruffo, Fact-checking strategies to limit urban legends spreading in a segregated society, in Applied Network Science 4, 116 (2019), Springer, https://appliednetsci.springeropen.com/articles/10.1007/s41109-019-0233-1

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**



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.

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

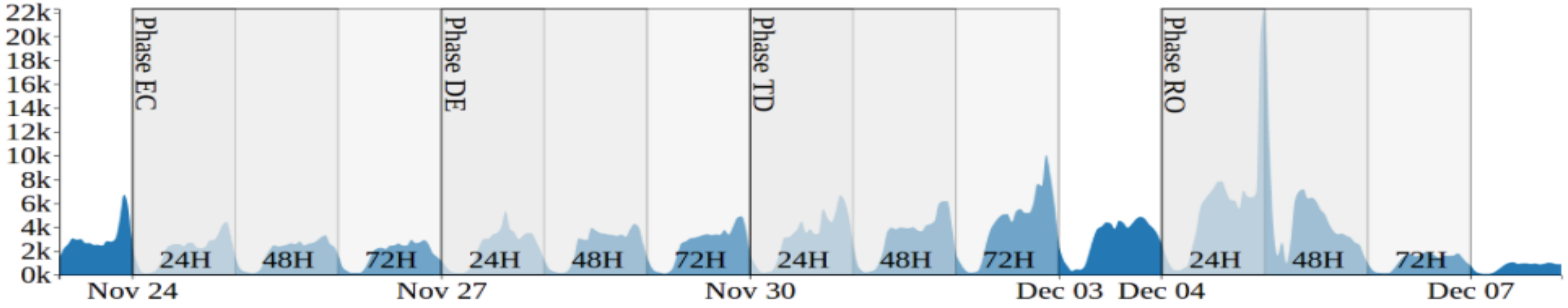
NLP and identification of bots and trolls

- Profiling bots, trolls, fact-checker: moving targets
- "Fake News, Bots, Trolls" Special Session at DSAA2020
  - brilliant keynotes from Preslav Nakov, and Damiano Spina
- Challenge at PAN 2020: "Profiling Fake News Spreaders on Twitter" https://pan.webis.de/clef20/pan20-web/author-profiling.html

Giachanou A, Ríssola EA, Ghanem B, Crestani F, Rosso P. The Role of Personality and Linguistic Patterns in Discriminating Between Fake News Spreaders and Fact Checkers. Natural Language Processing and Information Systems. 2020;12089:181-192. Published 2020 May 26. doi:10.1007/978-3-030-51310-8_17

Italian 2016 Constitutional Referendum

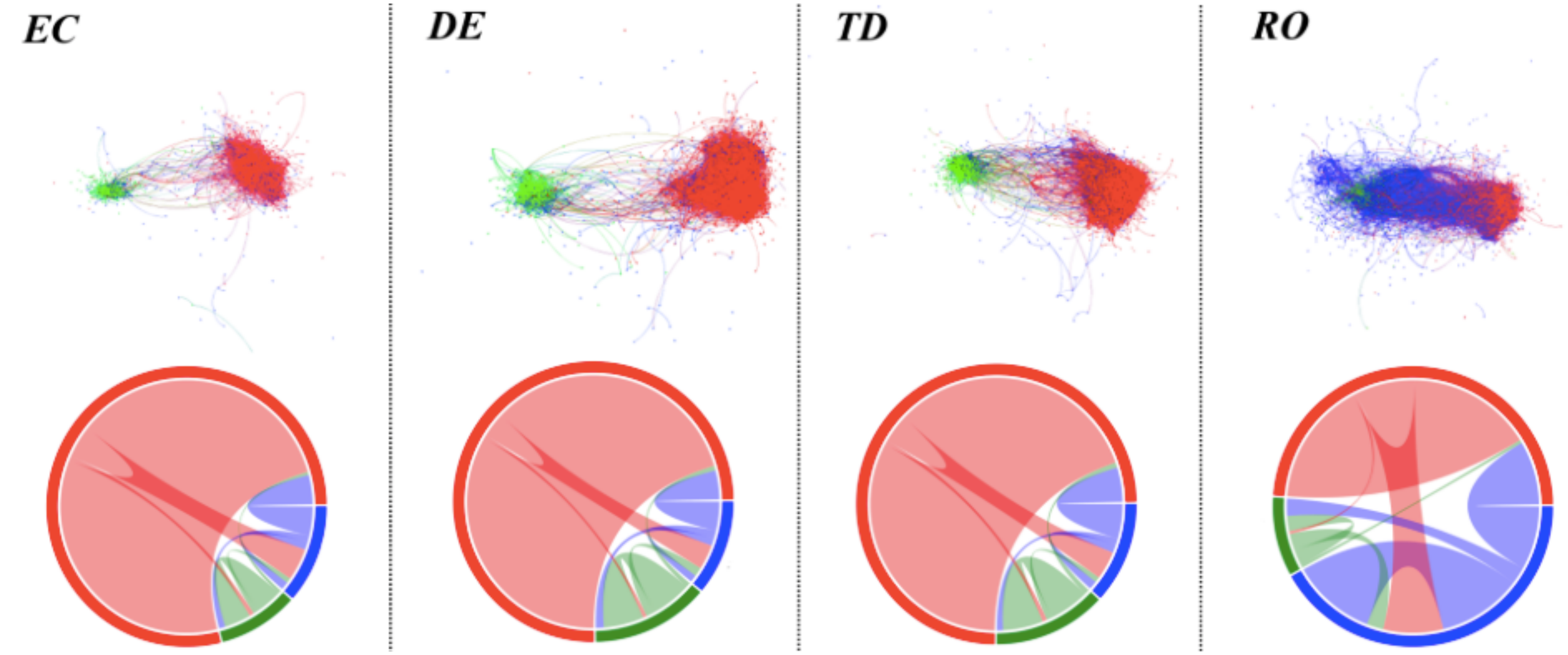
Collected Tweets



stance detected as **AGAINST**

stance detected as **IN FAVOR**

stance detected as **NONE**

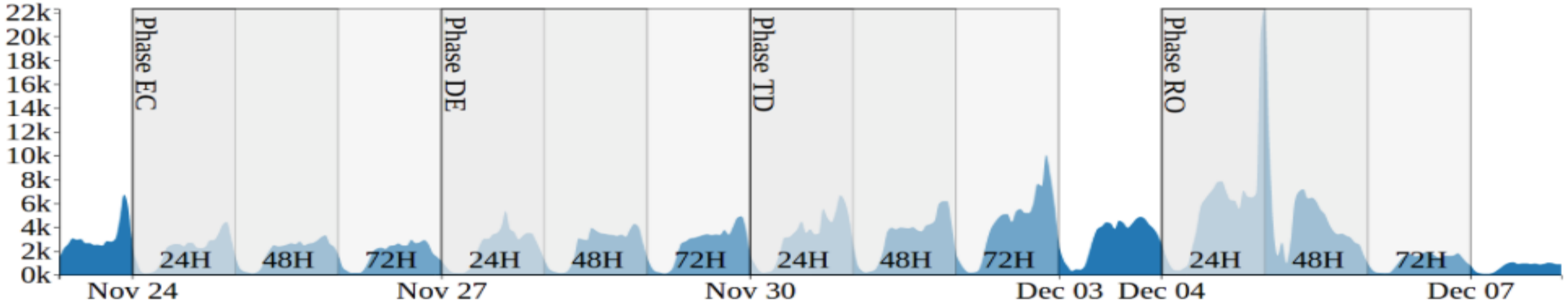


Retweet Network

strong signal of **homophily**

Italian 2016 Constitutional Referendum

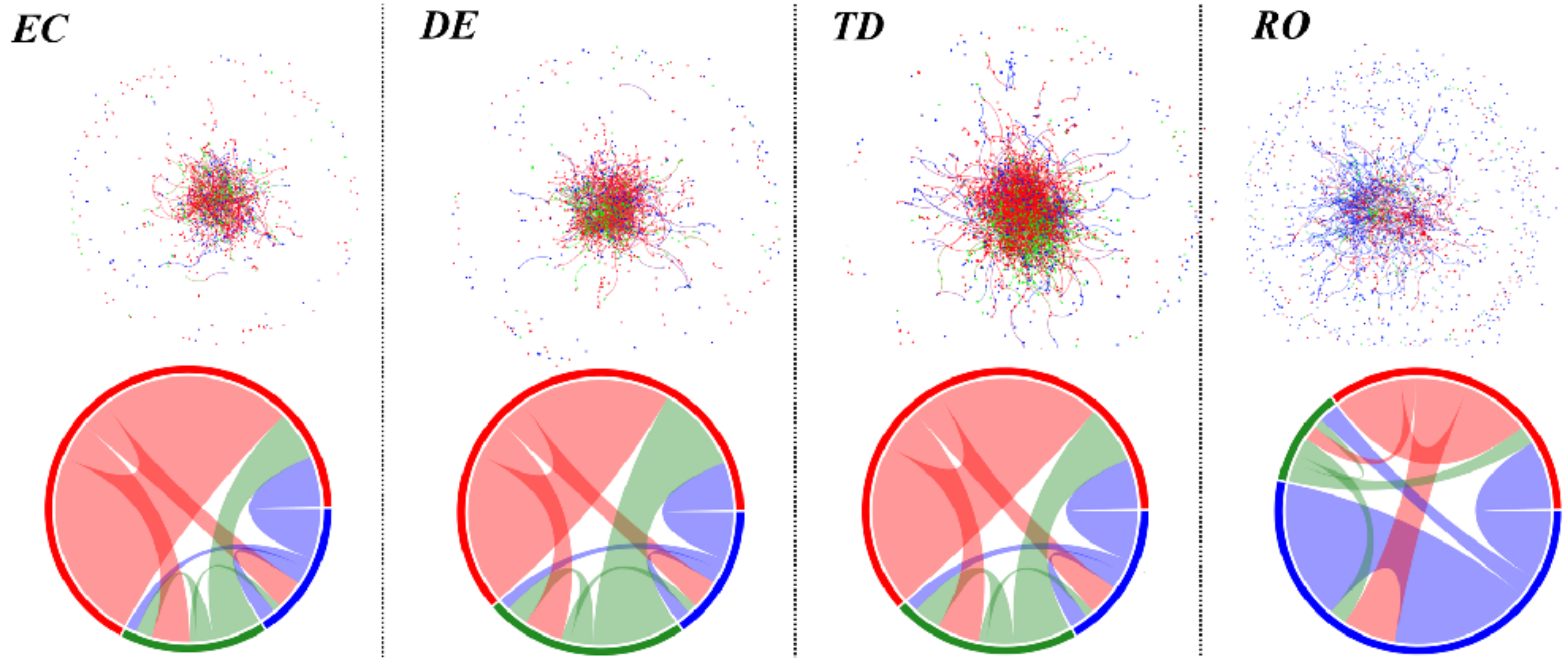
Collected Tweets



stance detected as **AGAINST**

stance detected as **IN FAVOR**

stance detected as **NONE**



Reply-to 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 the reply-to messages**

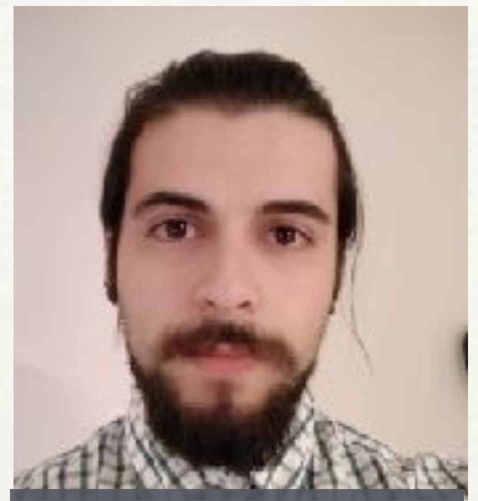
M Lai, M Tambuscio, V Patti, P Rosso, G. Ruffo, Stance Polarity in Political Debates: a Diachronic Perspective of Network Homophily and Conversations on Twitter, Data & Knowledge Engineering Journal, online: September 2019 https://www.sciencedirect.com/science/article/pii/S0169023X19300187

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
  - Profiling bots, trolls, and fact-checker is a big challenge!
- **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

Practical Lab (A. Semeraro)

di.unito.it


Arthur Capozzi


Alfonso Semeraro


Alessandra Urbinati


Salvatore Vilella


Mirko Lai

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


Marcella Tambuscio


André Panisson


Luca Aiello


Rossano Schifanella


Viviana Patti


Edoardo Galimberti


Emilio Sulis


Martina Deplano


Ciro Cattuto


Cristina Bosco

Thanks!

Bookmark this:

(slides and annotated bibliography available soon)

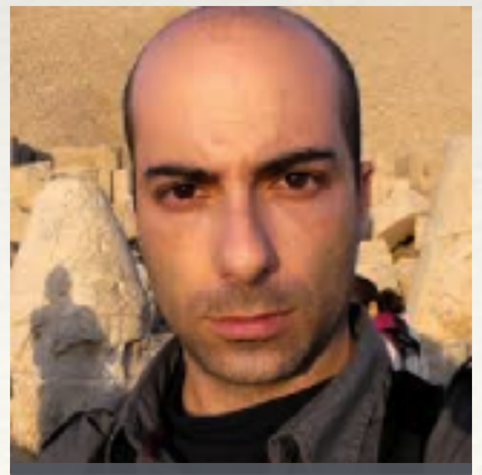
http://arcs.di.unito.it/2020/10/08/dsaatutorial/

Universitat Politècnica de València


Paolo Rosso

Universidad del Desarrollo


Leo Ferres

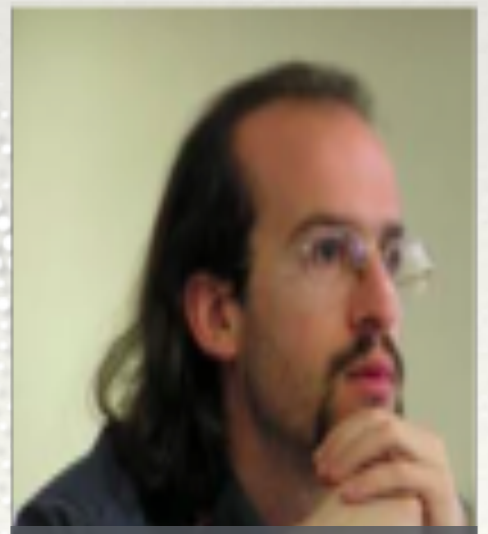

Giovanni Luigi Ciampaglia


Chengcheng Shao

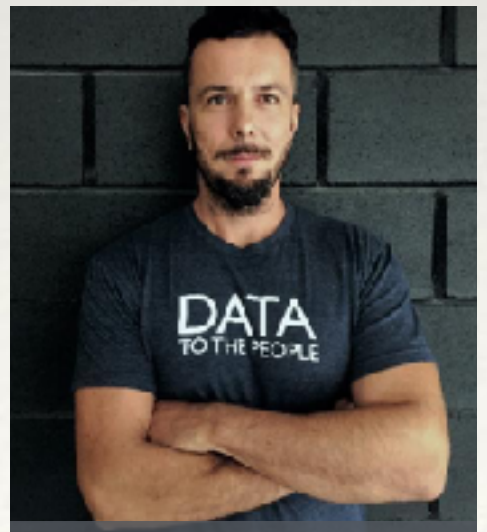

Alessandro Flammini


Fil Menczer

CPT


Alain Barrat

Institute for Scientific Interchange Foundation



Francesco Bonchi


Daniela Paolotti