

(Social) Network Analysis



Giulio Rossetti

Knowledge Discovery and Data Mining Laboratory (KDD) @ ISTI-CNR

gjulio.rossetti@isti.cnr.it

@GiulioRossetti



Lecture 5

Dynamics on Networks:

Diffusive phenomena

Where next?

Two kind of dynamics:

- **Dynamics of Networks**
(topological perturbations)
- **Dynamics on Networks**
(diffusive phenomena: epidemics,
opinion dynamics...)

Of course they can happen at the same time...

Dynamics of Networks

Assumption:
Topology evolution is faster than diffusive processes unfolding (if any)

Applications:

- Link Prediction
- Dynamic Community Discovery
- ...

Dynamics on Networks

Assumption:
Diffusive processes unfolding is faster than topology evolution (if any)

Applications:

- Epidemic spreading
- Opinion Dynamics
- ...

Mixed Dynamics

Assumption:
Diffusive processes unfolding and topology evolution have comparable rates

Applications:

- Diffusion shape topology
- Topology shape diffusion
- Feedback loops

Chapter 11

Diffusion: Epidemics

Summary

- Probabilistic Epidemic Models
 - SI/SIS/SIR
- Mean Field Formulation
- Epidemics on Networks

Reading

- Chapter 21 of Kleinberg's book
- Chapter 10 of Barabasi's book



Epidemic

Biological:

- Airborne diseases (flu, SARS, ...)
- Venereal diseases (HIV, ...)
- Other infectious diseases (HPV, ...)
- Parasites (bedbugs, malaria, ...)

Digital:

- Computer viruses, worms
- Mobile phone viruses

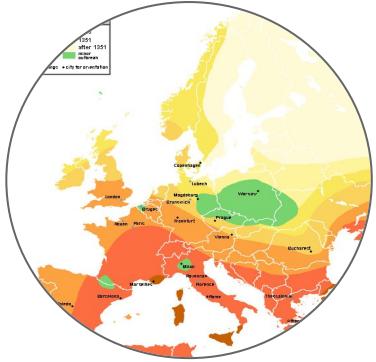
Conceptual/Intellectual:

- Diffusion of innovations
- Rumors
- Memes
- Business practices

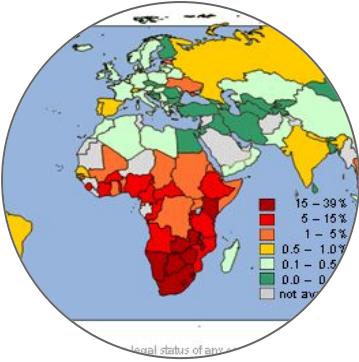
*Epi + demos
upon people*



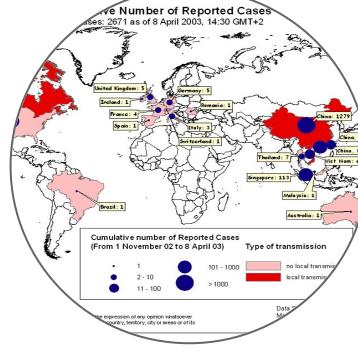
The Great Plague



HIV



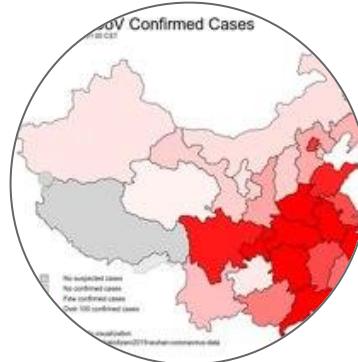
SARS



1918 Spanish flu

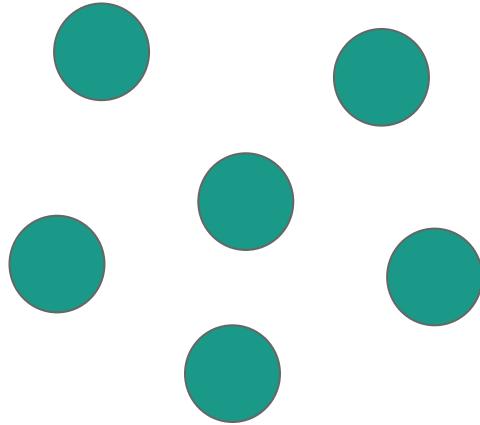


H1N1 flu

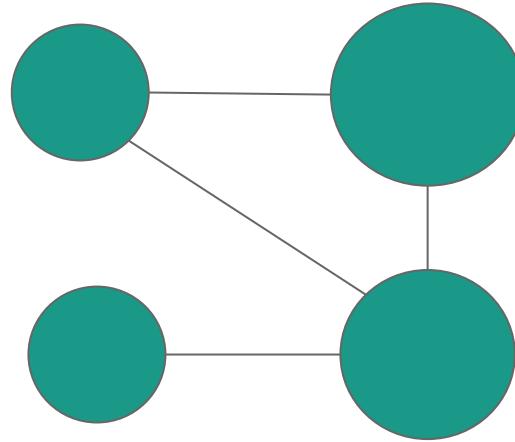


2019-nCoV

Biological: Notable Epidemic Outbreaks



Separate, small population
(hunter-gatherer society, wild animals)



Connected, highly populated areas
(cities)

Human societies have “**crowd diseases**”, which are the consequences of large, interconnected populations
(Measles, tuberculosis, smallpox, influenza, common cold, ...)

Large population can provide the “fuel”

Probabilistic Epidemic Models



Compartmental Models

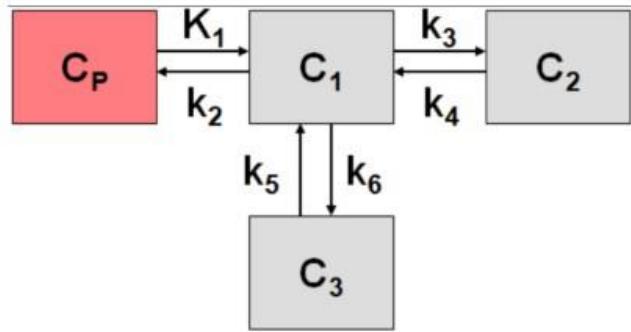
The framework is based on two hypotheses:

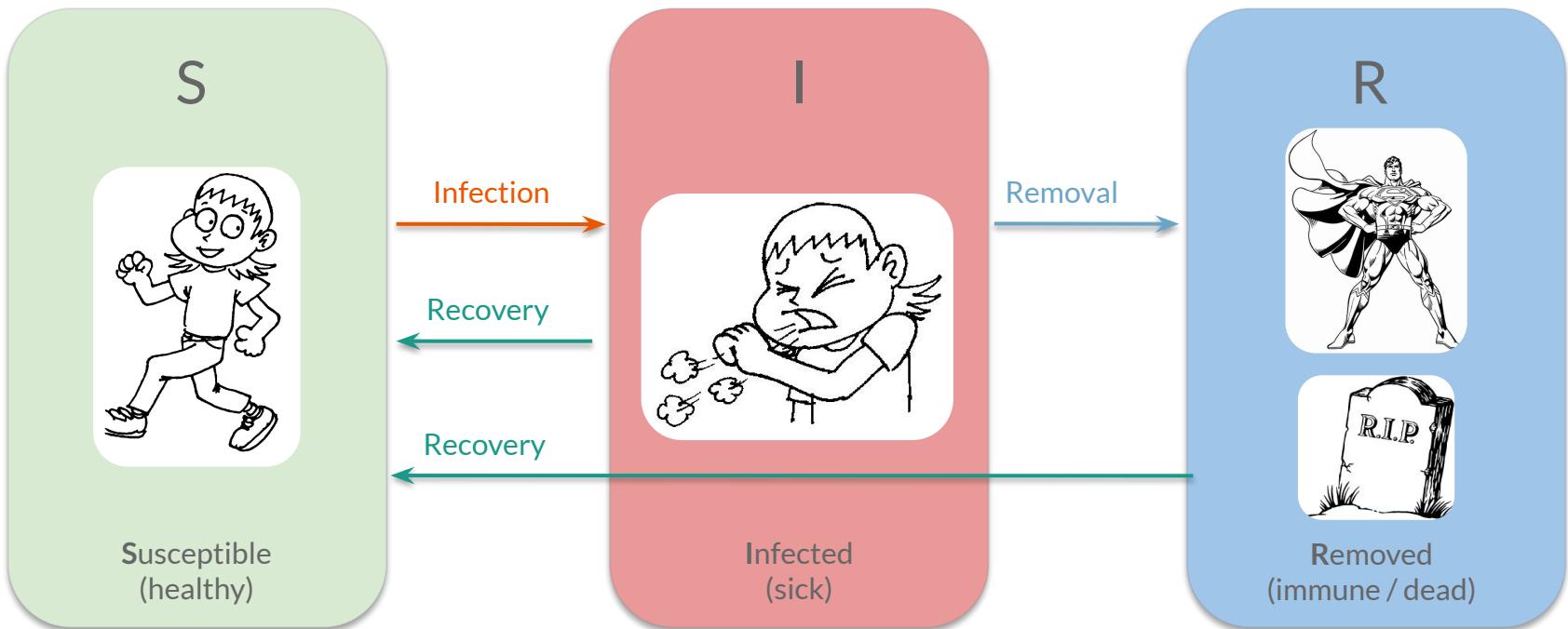
Compartmentalization:

each individual is classified into distinct statuses. The simplest classification assumes that an individual can be in one of the states.

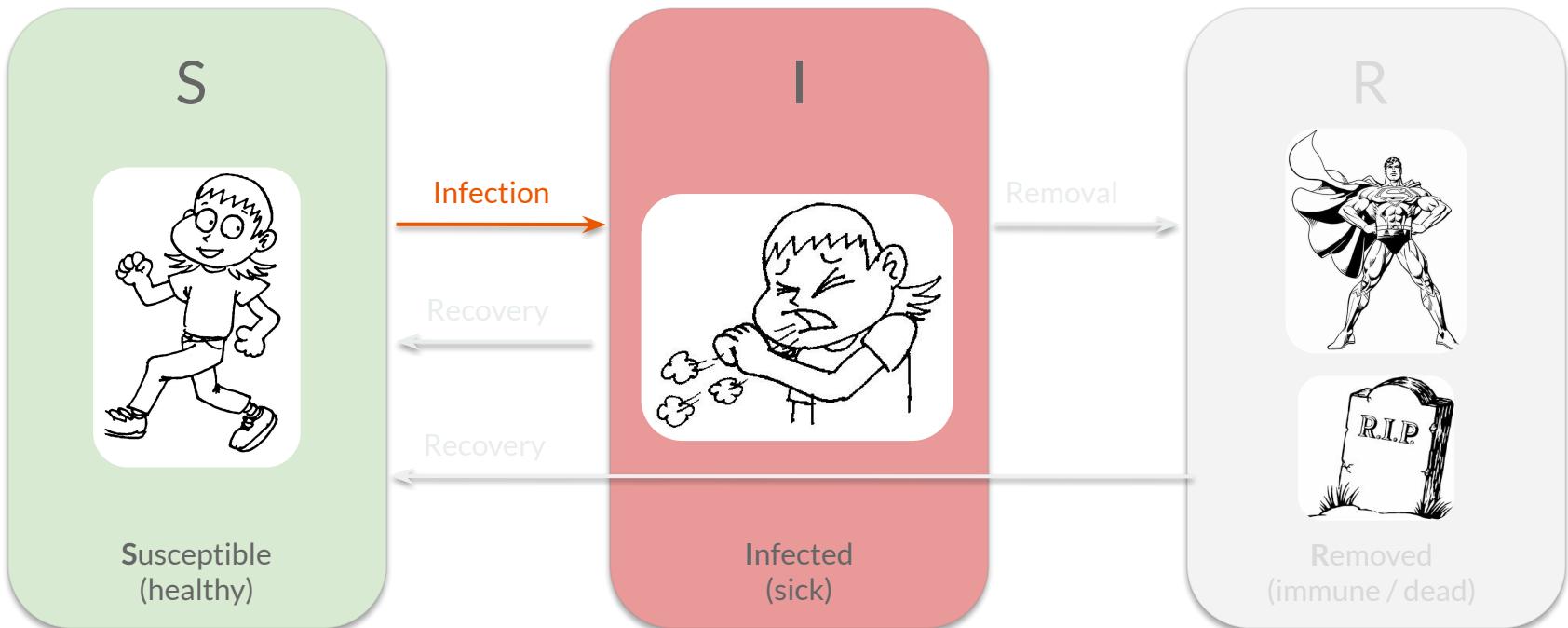
Homogeneous Mixing:

each individual has the same chance of coming into contact with an infected individual.

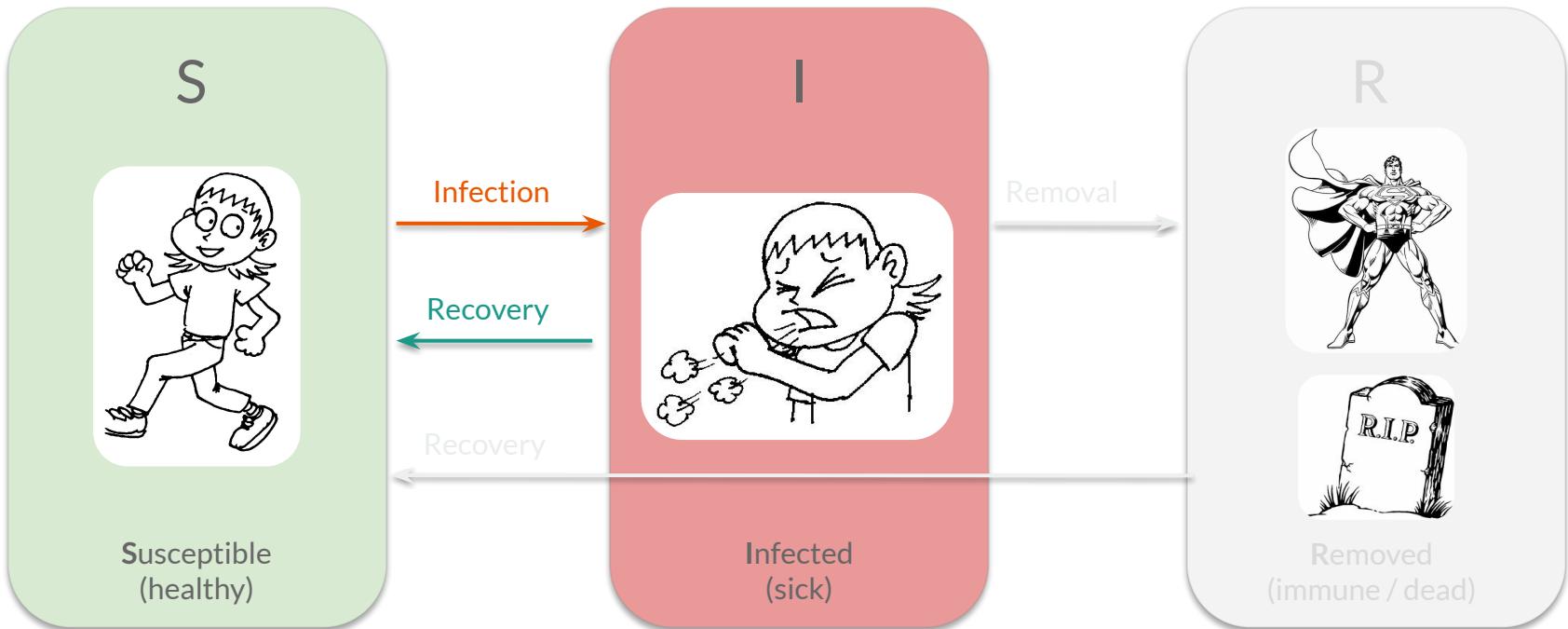




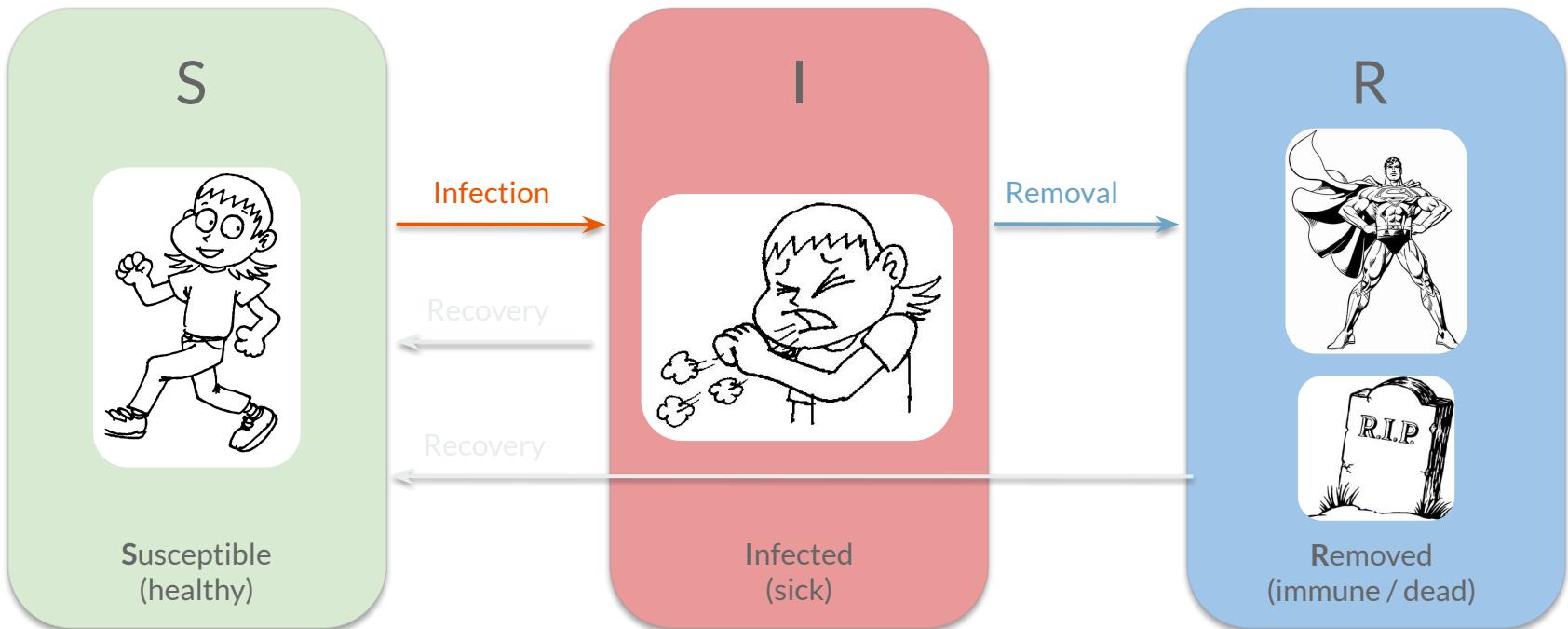
Classic Models Compartments



SI: The simplest model



SIS: Common Cold



SIR: Flu, SARS, Plague

Mean Field formulation

(Homogeneous mixing)



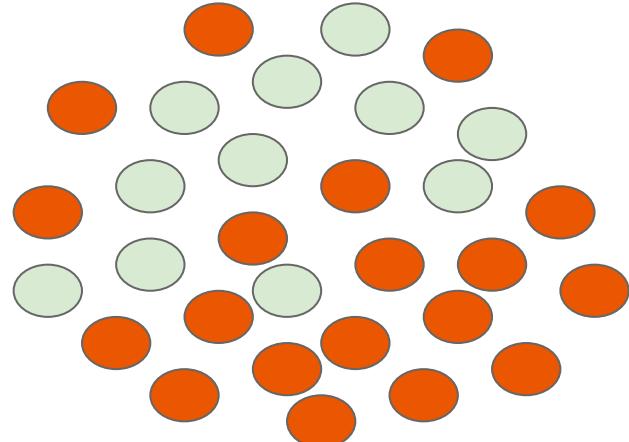
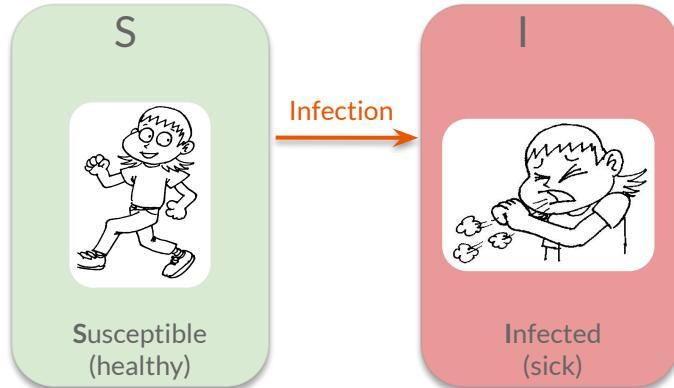
SI model

Each individual has β contacts with randomly chosen others individuals per unit time.

If there are I infected individual and S susceptible individuals, the average rate of new infection is $\beta si/N$

with $s = S/N$, $i = I/N$

$$\frac{di}{dt} = \beta si = \beta i(1 - i)$$



SI model

Dynamics

Logistic equation:
a basic model of population growth.

$$\frac{di}{dt} = \beta i(1 - i)$$

| S

$$\frac{di}{i} + \frac{di}{(1-i)} = \beta dt$$
$$\ln i - \ln(1-i) + c = \beta t$$

$$\frac{i}{1-i} = C \exp(\beta t)$$
$$C = \frac{i_0}{1-i_0}$$

$$\ln \frac{i}{1-i} = c + \beta t$$

$$\therefore i(t) = \frac{i_0 \exp(\beta t)}{1 - i_0 + i_0 \exp(\beta t)}$$

http://en.wikipedia.org/wiki/Logistic_function
<http://mathworld.wolfram.com/LogisticEquation.html>

SI model

Behaviour

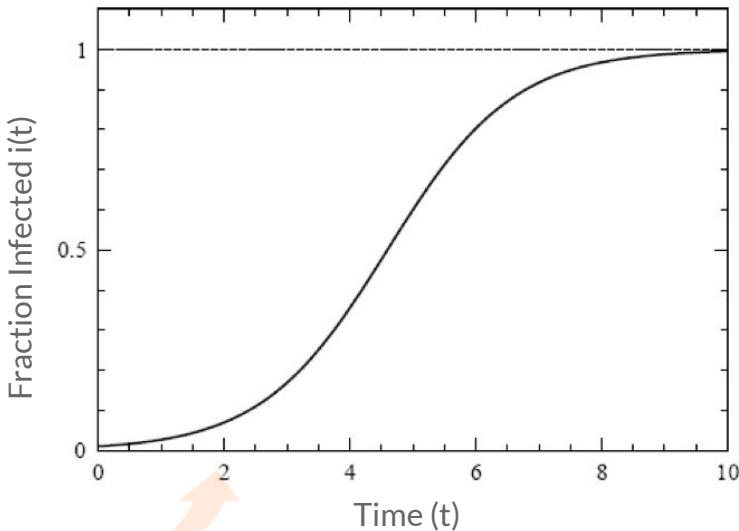
If $i(t)$ is small,

$$\frac{di}{dt} \approx \beta i$$

$$i \approx i_0 \exp(\beta t)$$

exponential outbreak

$$i(t) = \frac{i_0 \exp(\beta t)}{1 - i_0 + i_0 \exp(\beta t)}$$



As $i(t) \sim 1$.

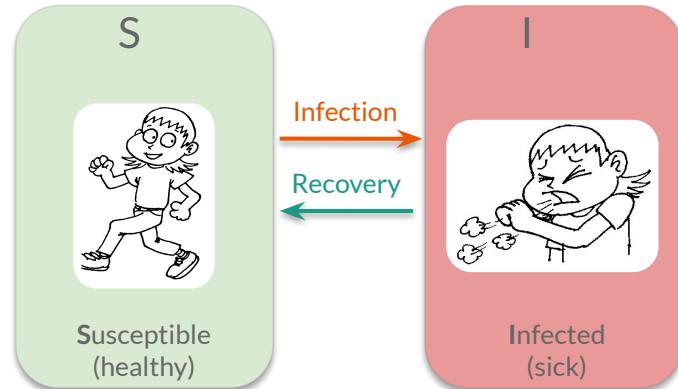
$$\frac{di}{dt} \rightarrow 0$$

saturation

SI model:

the fraction infected increases until everyone is infected.

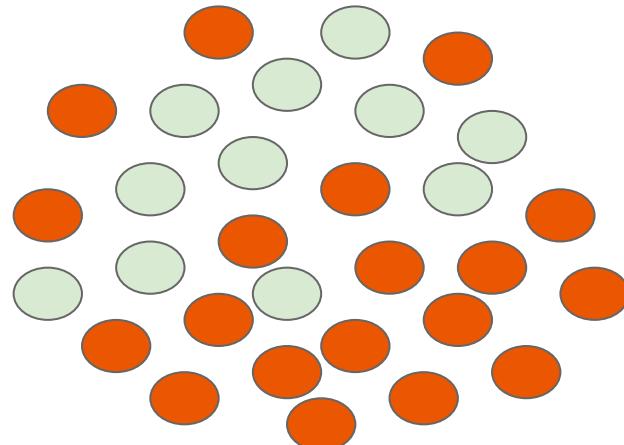
SIS model



Modeling Common Cold

Each individual has β contacts with randomly chosen others individuals per unit time.

Each infected individual has μ probability of revert its status to susceptible



SIS model

Behaviour

$$\frac{di}{dt} = \frac{\beta i(1-i)}{I} - \frac{\mu i}{S} = i(\beta - \mu - \beta i)$$

$$\frac{di}{i} + \frac{di}{1 - \mu/\beta - i} = (\beta - \mu)dt$$

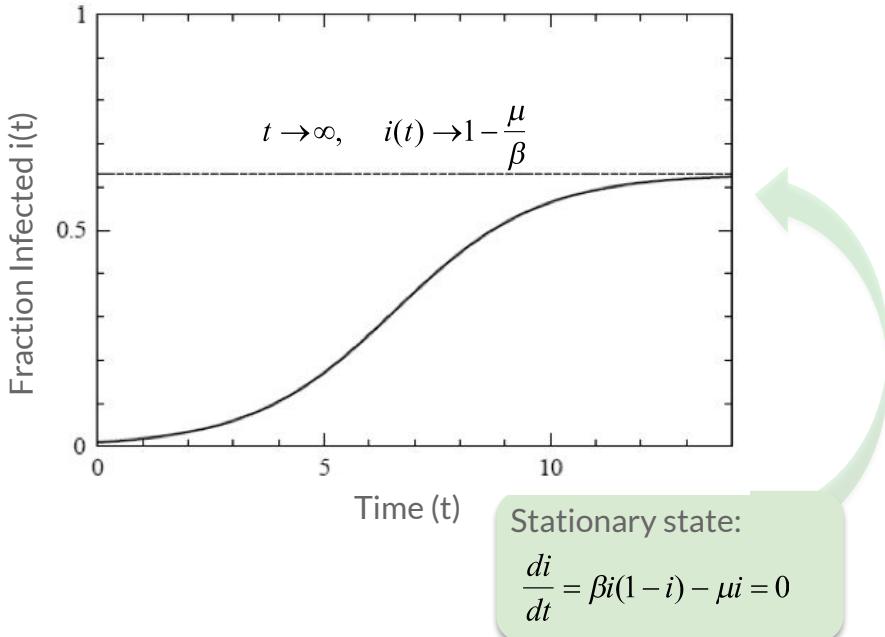
$$\ln(i) - \ln(1 - \mu/\beta - i) = (\beta - \mu)t + c$$

$$\frac{i}{1 - \mu/\beta - i} = Ce^{(\beta - \mu)t}$$

$$\therefore i(t) = \left(1 - \frac{\mu}{\beta}\right) \frac{Ce^{(\beta - \mu)t}}{1 + Ce^{(\beta - \mu)t}}$$

SIS model

Dynamics



$$\therefore i(t) = \left(1 - \frac{\mu}{\beta}\right) \frac{Ce^{(\beta-\mu)t}}{1 + Ce^{(\beta-\mu)t}}$$

SIS model:
the fraction of infected
individual saturates below 1

SIS model

Basic Reproductive Number

λ (also identified with R_0):

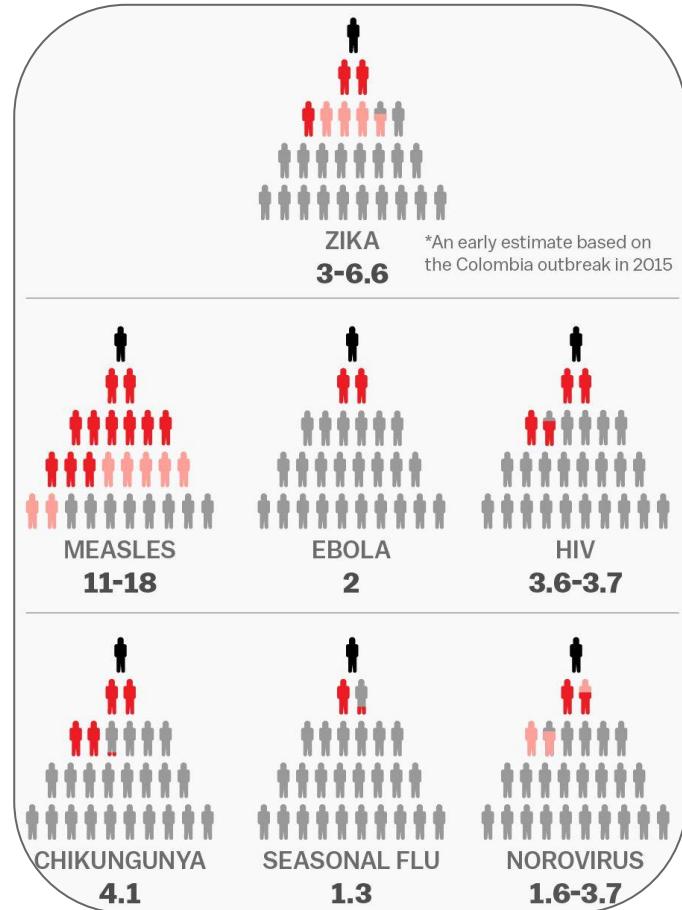
average # of infectious individuals generated by one infected in a fully susceptible population.

$$\lambda \equiv \frac{\beta}{\mu}$$

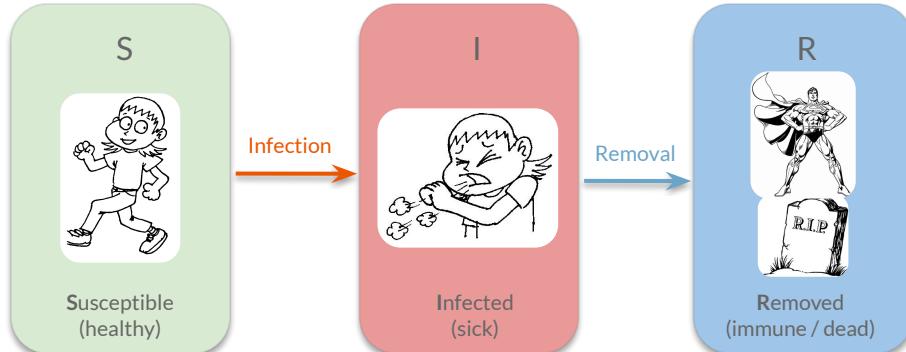
$\lambda > 1$: Outbreak

$\lambda < 1$: Die Out

Epidemic Threshold
if $\mu = \square$ then $i \rightarrow 0$



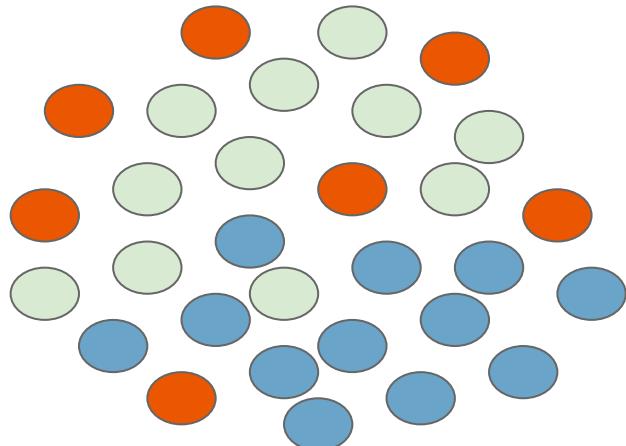
SIR model



Modeling Flu-like disease

Each individual has β contacts with randomly chosen others individuals per unit time.

Each infected individual has μ probability of becoming immune after being infected



SIS model

Behaviour

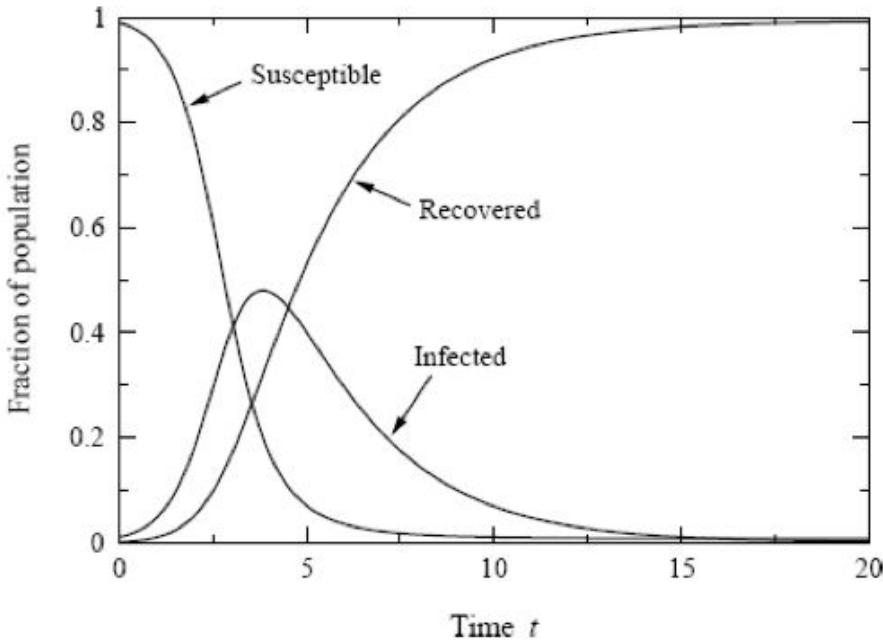
$$\frac{ds(t)}{dt} = \beta\langle k \rangle i(t) [1 - r(t) - i(t)]$$

$$\frac{di(t)}{dt} = -\mu i(t) + \beta\langle k \rangle i(t) [1 - r(t) - i(t)]$$

$$\frac{dr(t)}{dt} = \mu i(t).$$

SIR model:

the fraction infected peaks and the fraction recovered saturates.



	SI	SIS
1 Early Behaviour Exponential growth of infected individuals	$i(t) = \frac{i_0 \exp(\beta t)}{1 - i_0 + i_0 \exp(\beta t)}$	$i(t) = \left(1 - \frac{\mu}{\beta}\right) \frac{C e^{(\beta - \mu)t}}{1 + C e^{(\beta - \mu)t}}$
2 Late Behaviour Saturation at $t \rightarrow \infty$	$i(t) \rightarrow 1$	$i(t) \rightarrow 1 - \frac{\mu}{\beta}$
3 Epidemic Threshold Disease not always spread	No Threshold	$\lambda_c = 1$

Recap: Basic Features of Epidemic Models

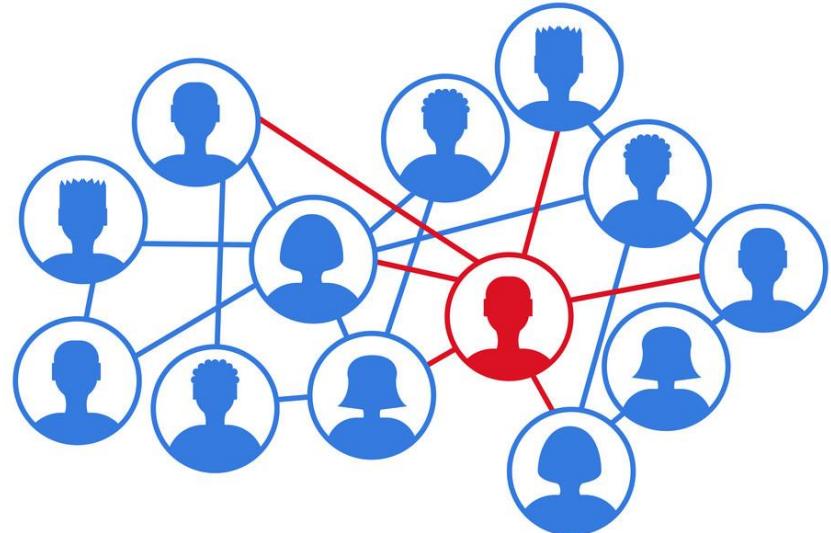
Epidemics on Networks



Topology matters

The described approaches assumed *homogenous mixing*, which means that each individual can infect *any* other individual.

In reality, epidemics spread along *links in a network*: we need to explicitly account for the role of the network in the epidemic process.



Modeling choices

Degree based representation:

split nodes by degree

$$i_k = \frac{I_k}{N_k}, \quad i = \sum_k P(k)i_k$$

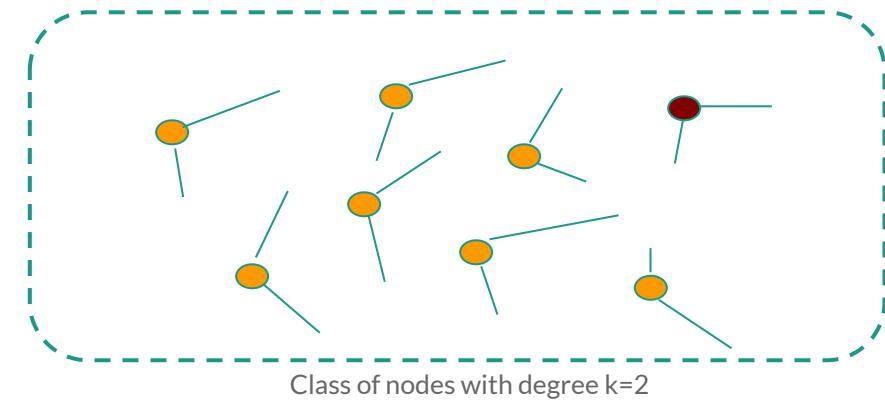
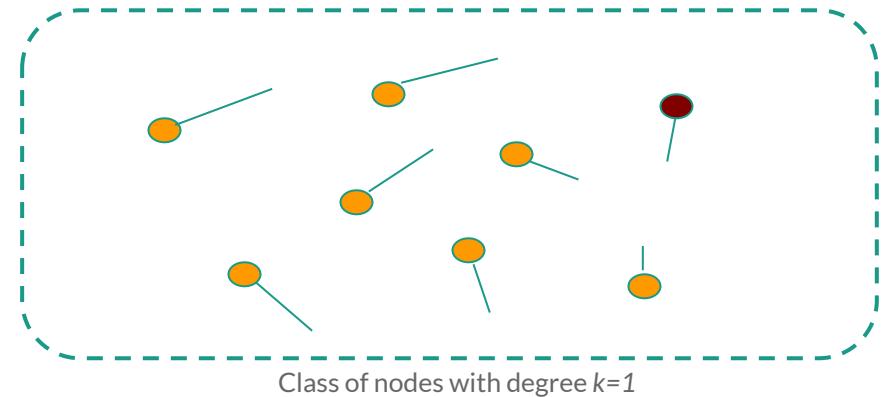
Example SIS:

I am susceptible with k neighbors, and $\Theta_k(t)$ of my neighbors are infected.

$$\frac{di_k(t)}{dt} = \beta(1 - i_k(t))k\Theta_k(t) - \mu i_k(t)$$

Proportional to k

Density of infected neighbors of nodes with degree k



Modeling choices

Agent based representation:

Each node is an agent having a current status (S/I/R...) and subject to probabilistic transition rules

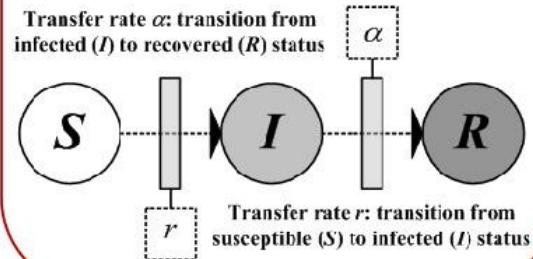
Example SIR:

- Current node status S:
Applicable rules: $S \rightarrow I$
If at least one of my neighbors is infected, with probability β change my status to infected.
- Current node status I:
Applicable rules: $I \rightarrow R$
With probability μ turn my status to removed.

Example SIR model

Three node statuses:
(S)usceptible,
(I)nfected,
(R)ecovered

Two transition rules:
 $S \rightarrow I$;
 $I \rightarrow R$

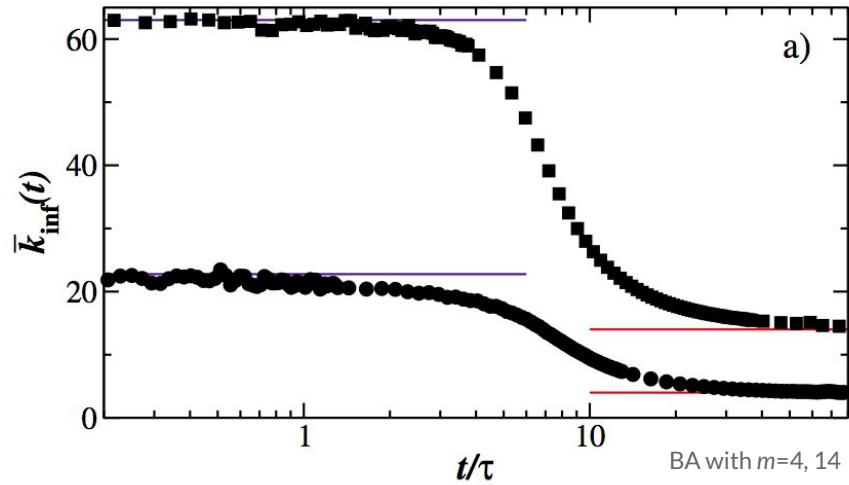


Early time behaviour

$$\tau = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}$$

The **timescale** it takes for an epidemics to grow. The smaller is τ , the faster it grows.

ER model	BA model
$\langle k^2 \rangle = \langle k \rangle (\langle k \rangle - 1)$ The more connected the network is, the faster does the epidemic spread.	$\langle k^2 \rangle \propto \infty$ for $N \propto \infty$ $\Rightarrow \tau \approx 0$ The characteristic time vanishes : the epidemic becomes instantaneous. Reason: the hubs get infected first, which then rapidly reach most nodes.



Numerical Test:
The average degree of newly infected nodes at time t :

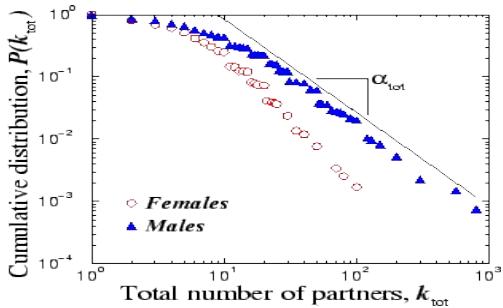
$$\bar{k}_{\text{inf}}(t) = \frac{\sum_k k(I_k(t) - I_k(t-1))}{I(t) - I(t-1)}$$

SIS Model

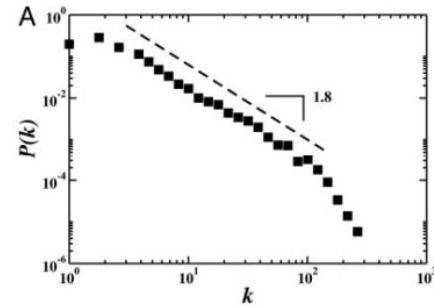
No Epidemic Threshold

Many networks will have small or vanishing epidemic threshold.

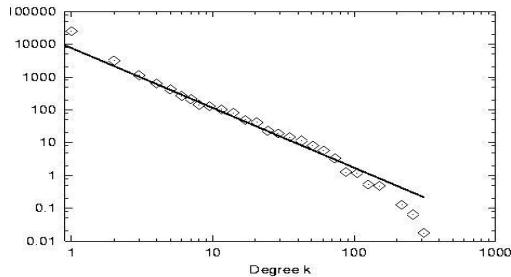
Diffusion will not die out.



Human Sexual network



Air Transportation network



Email network

Chapter 11

Conclusion

Take Away Messages

1. Viruses spread over a population
2. Populations can be modeled with social networks topologies
3. Stochastic epidemic models can be leveraged to simulate and reason upon real world viruses diffusion

Suggested Readings

- Chapter 21 of Kleinberg's book
- Chapter 10 of Barabasi's book

What's Next

Chapter 12:
Diffusion: Opinion Dynamics



Chapter 12

Diffusion: Opinion Dynamics

Summary

- Modeling Opinions
- Polarization, Fragmentation, Bias
- Modeling Algorithmic Bias
- Opinions and Topology

Reading

- “Opinion dynamics: models, extensions and external effects.”
Sirbu et al.



Modelling Opinions



Opinion Dynamics

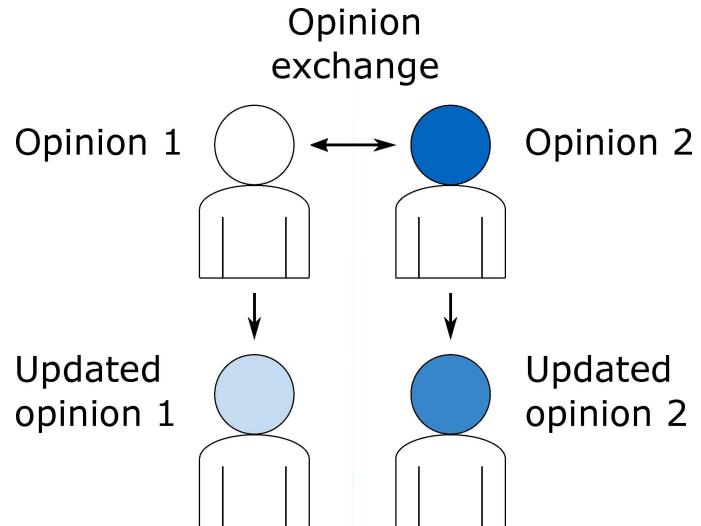
Model evolution of opinions in a population

Opinions are at the base of human behaviour

- understand behaviour - which mechanisms are important?
- trigger changes in behaviour ~ intervention methods in spreading, less explored

Broadly part of complex contagion modelling:
peer effects through social network.

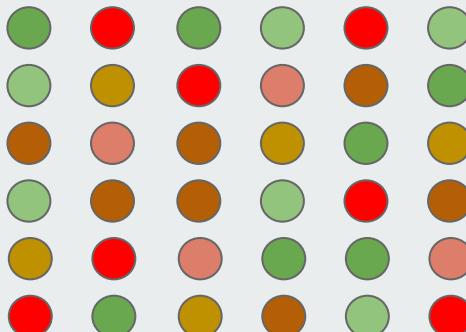
Simple representations of opinions - one variable.



Opinions

Continuous

Individual status is identified by a (bounded) real value:
- e.g., opinions, beliefs,...



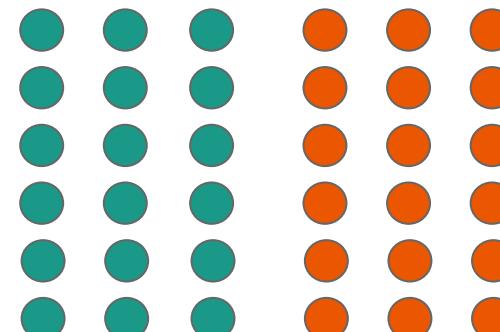
0

1

Opinions

Discrete

Individual status is identified by a discrete value:
- e.g., political party affiliation...



0

1

Models

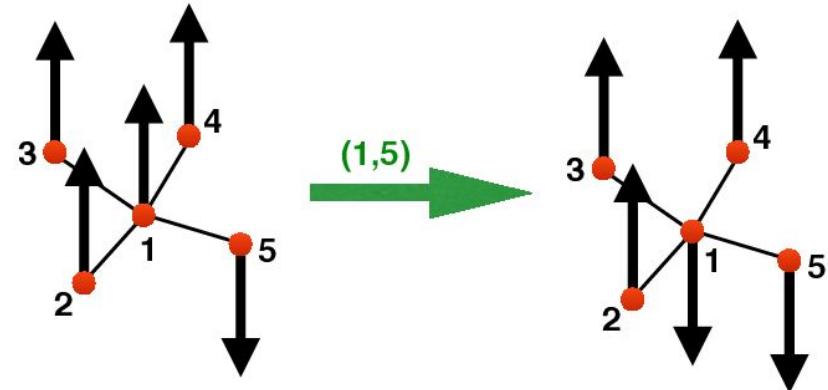
Voter

Originally introduced to analyse competition of species, then applied to electoral competitions.

Discrete opinions: $\{-1, 1\}$

Iteration:

- A random agent i is selected with one of its neighbors j
- i takes j 's opinion



R. Holley and T. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model," Ann. Probab., (1975).

Models

Majority Rule

Originally introduced to describe public debates (e.g., global warming, H1N1 pandemic).

Discrete opinions: $\{-1, 1\}$

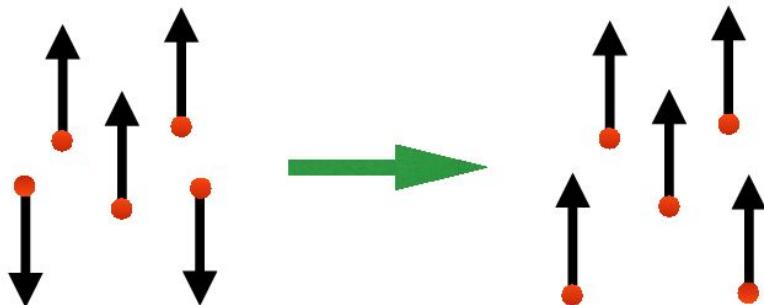
Iteration:

- A random group of r agents is selected
- The agents take the majority opinion within the group

r odd: majority always exists

r even: possibility of tied configurations.

To address them, bias toward an opinion is introduced (social inertia)



S.Galam, "Minority opinion spreading in random geometry." *Eur.Phys. J. B.*, (2002).
R.Friedman and M.Friedman, "The Tyranny of the Status Quo." Harcourt Brace Company (1984).

Models

Deffuant Model

Simple model of opinion formation, with
bounded confidence

Opinions $x_i \in [0,1]$ (Continuous values)

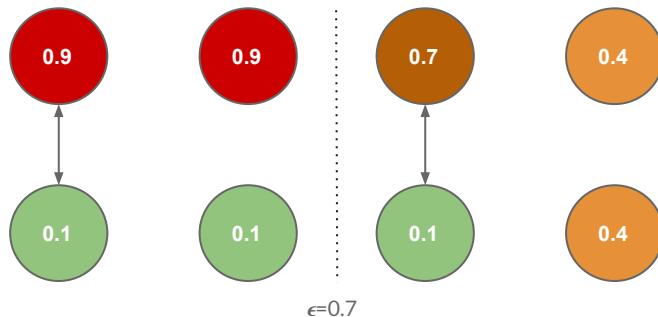


Discrete time steps

Iteration:

Two random individuals interact with bounded confidence ϵ (open-mindedness)

- $x_i(t+1) = x_j(t+1) = (x_i(t) + x_j(t))/2$
- only if $|x_i(t+1) - x_j(t+1)| < \epsilon$



Deffuant G, Neau D, Amblard F, Weisbuch G. Mixing beliefs among interacting agents. *Advances in Complex Systems*. (2000).

Behaviour

Continuous

One or more clusters
(depending on the bounded confidence parameter.)

- Extreme information → segregation
- Mild information → consensus

Extensions:

- Noise, heterogeneous bounds of confidence → consensus
- Contrarians → fragmentation, extremism, agreement with external information



Behaviour

Discrete

Consensus on one of the two opinions

Questions:

- Exit probability: prob. to obtain consensus on +1/-1
 - Consensus time for a population of size N

Extensions:

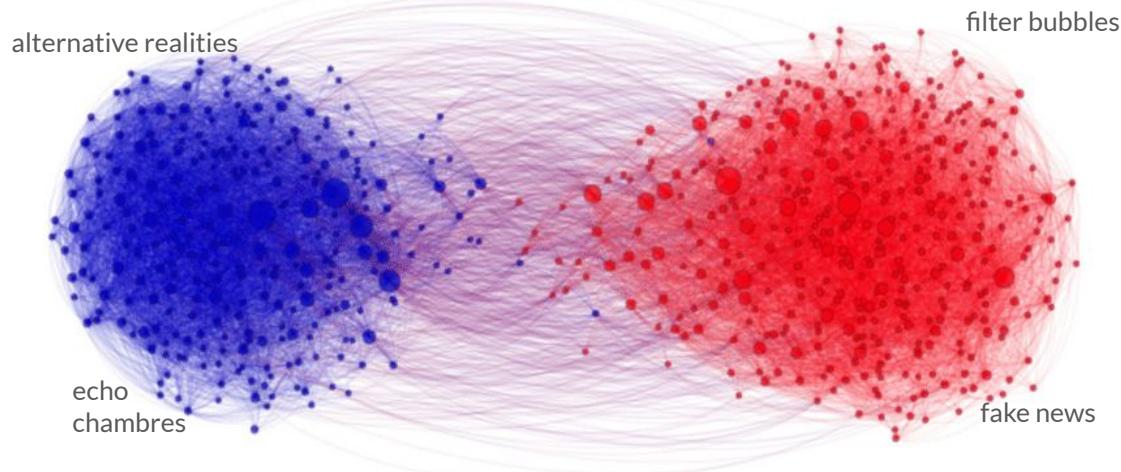
- contrarians, inflexibles (zealots), independents (noise)
- Consensus breaks → clusters of opinion



Polarization and Fragmentation in Social Media



Polarization of the public debate



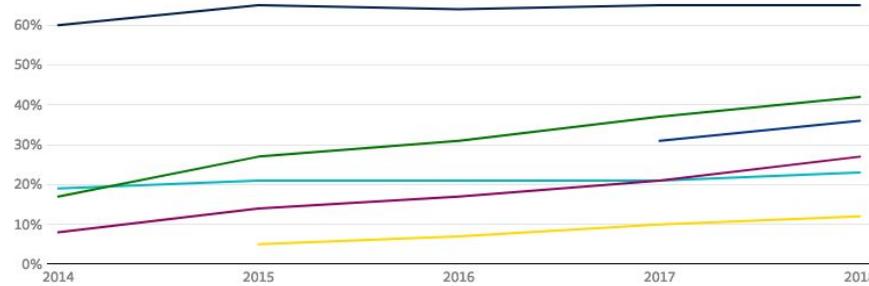
Adamic, Lada A., and Natalie Glance. "The political blogosphere and the 2004 US election: divided they blog." ACM (2005).

Online News Consumption

PROPORTION THAT USED EACH SOCIAL NETWORK FOR ANY PURPOSE IN THE LAST WEEK (2014–18)

Selected markets

Snapchat Twitter WhatsApp Instagram FB Messenger Facebook



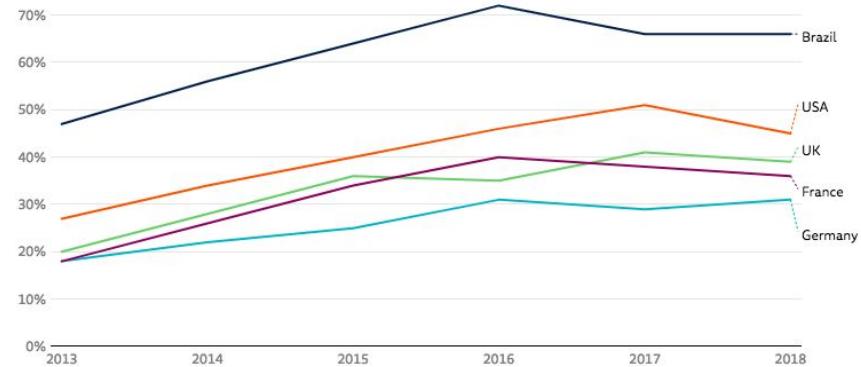
Q12A. Which, if any, of the following have you used for any purpose in the last week?

Base: Total sample across selected markets: 2014 = 18859, 2015 = 23557, 2016 = 24814, 2017 = 24487, 2018 = 24735.

Note: From 2015–18, the 12 markets included are UK, US, Germany, France, Spain, Italy, Ireland, Denmark, Finland, Japan, Australia, Brazil. In 2014, we did not poll in Australia or Ireland.

PROPORTION THAT USED SOCIAL MEDIA AS A SOURCE OF NEWS IN THE LAST WEEK (2013–18)

Selected countries



Q3. Which, if any, of the following have you used in the last week as a source of news?

Base: Total 2013–2018 sample in each market.

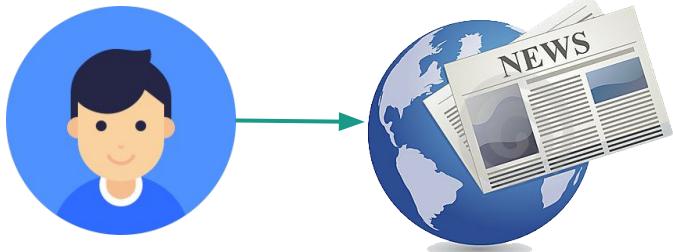
Online consumption of information

Interaction of

- *users*,
- with *media* content

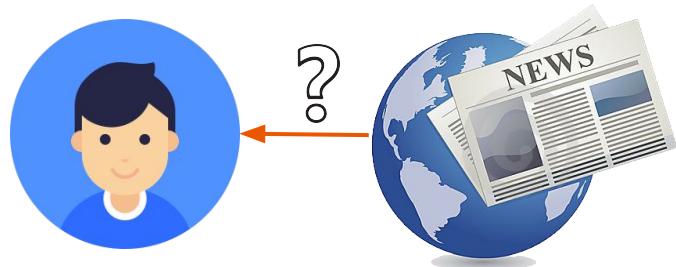
mediated by *computer programs*

1st scenario



Users *actively* search for news

2nd scenario



Users are *passively* fed of news

Online consumption of information

The aim of the computer programs is to **maximise** the **usage of the platform**

To fulfill such goal they carefully **taylor** the information shown to their users



Confirmation Bias

"[is the] tendency to search for, interpret, favor, and recall information in a way that **confirms one's preexisting beliefs or hypotheses.**"

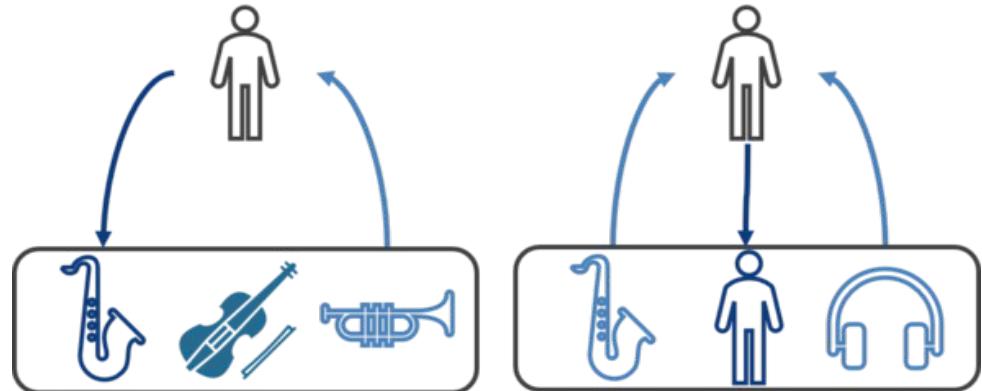
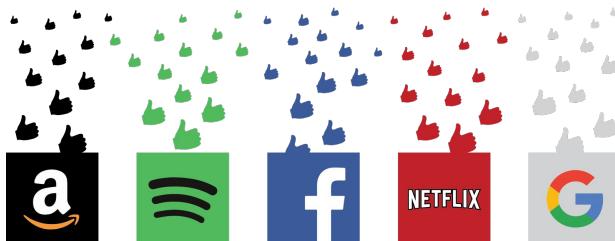
Recommender Systems

Leveraging user's history

Recommendations are built on top of user's past choices...

- type of news searched, product bought...

As well as on top of "similar" users' ones



A product is recommended that is **similar to products** the customer has already looked at.

The customer is shown products that customers with **similar data profiles** have found interesting.

Online consumption of information

Users are mostly shown **opinions** that are **close** to their own (algorithmic bias)

- News about topics we like,
- Posts from close friends,
- ...

Users **do not** even get confronted with narratives **different** from their favorite ones

- or they get in contact with **extreme opposite** narratives



Modeling Algorithmic Bias



Models

Algorithmic Bias

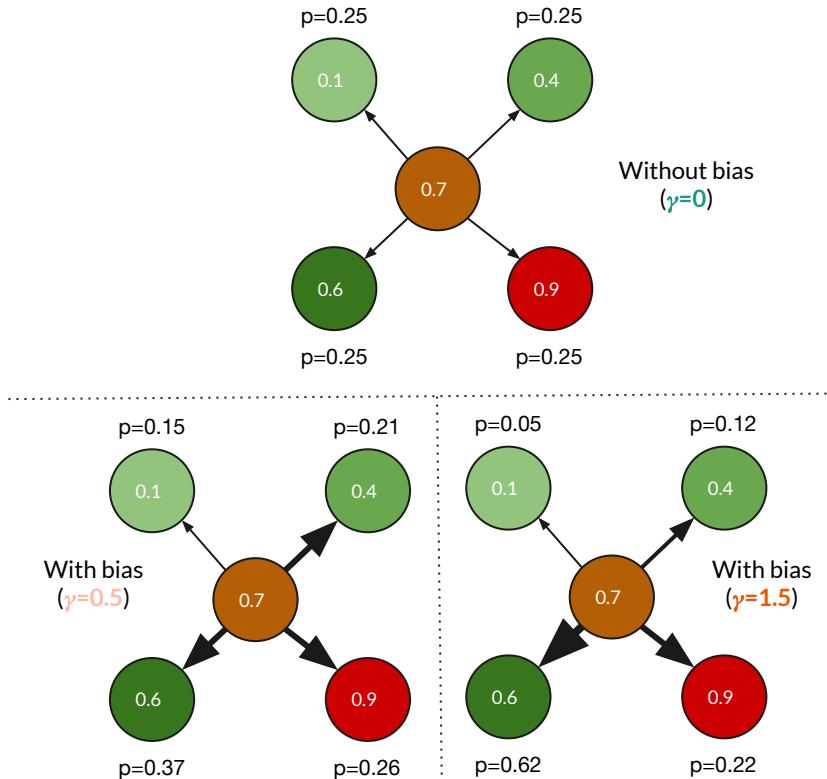
Modified Deffuant model

Probability to select interaction partner depends by

- the **opinion distance**, d_{ij}
- the **bias strength**, γ

$$p_i(j) = \frac{d_{ij}^{-\gamma}}{\sum_{k \neq i} d_{ik}^{-\gamma}}$$

The more similar the opinions, the more likely that the interaction will take place.



Sîrbu, Alina, et al. "Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model." PloS One (2019)

Deffuant

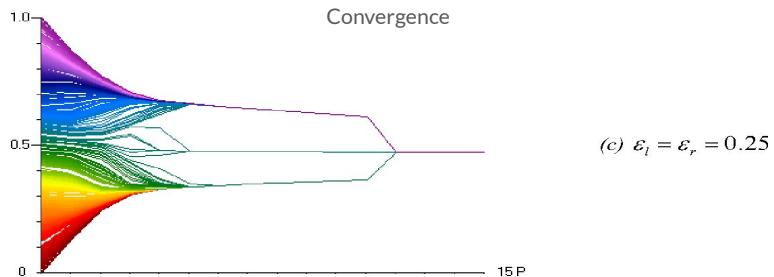
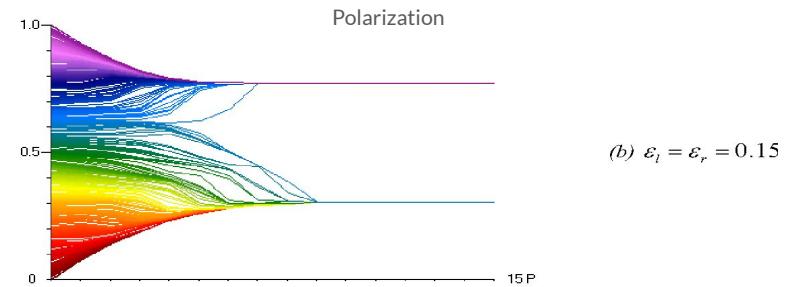
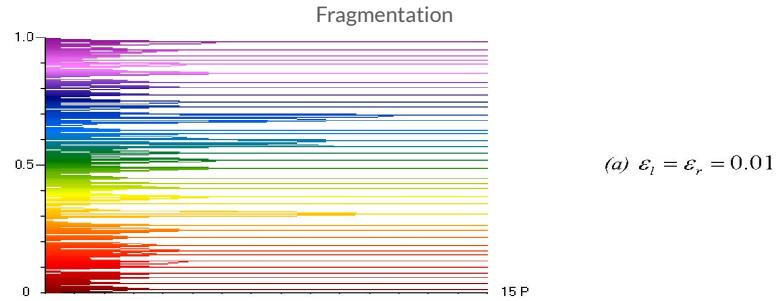
Simulations

Recap:

Reducing the bounded confidence threshold value opinion fragmentation (polarization) intensifies

Interpretation:

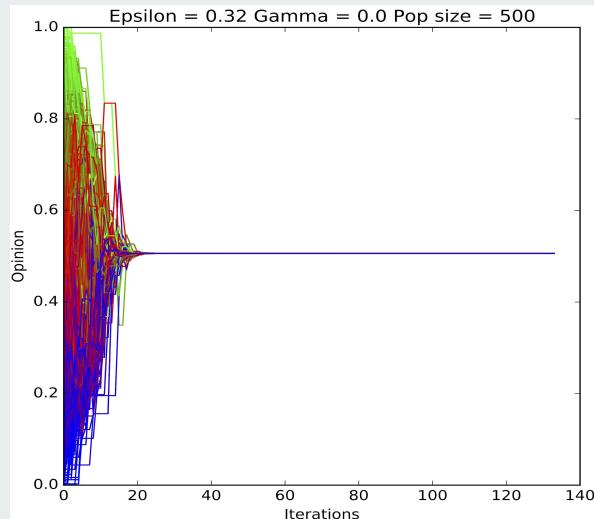
The **larger** the open-mindedness value, the **more likely** that **consensus** will be reached



Deffuant

Without Bias

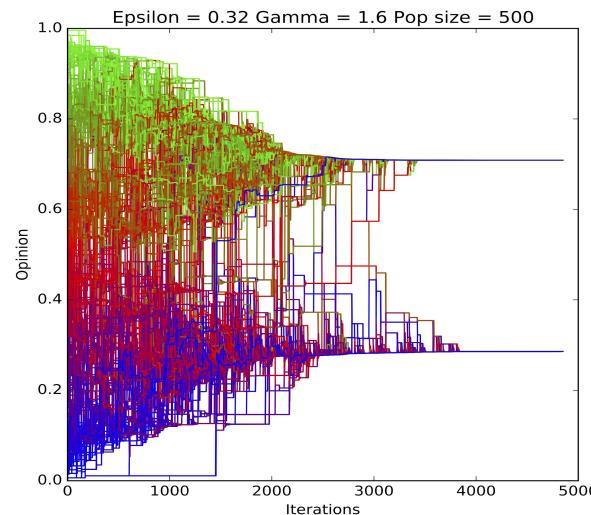
Convergence to common opinion



Deffuant

With Bias

Opinion Polarization, Fragmentation,
Convergence slow-down (instability)



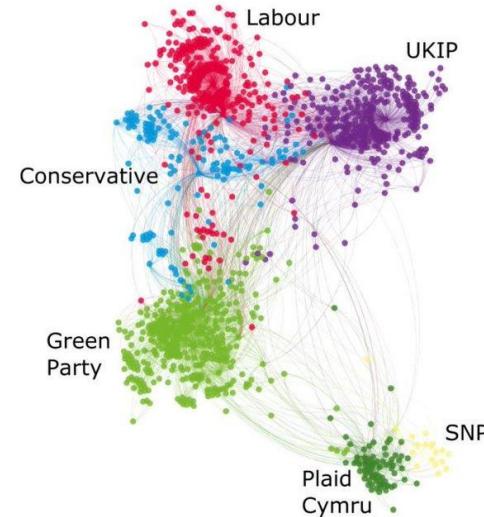
Algorithmic Bias

Is this the whole story?

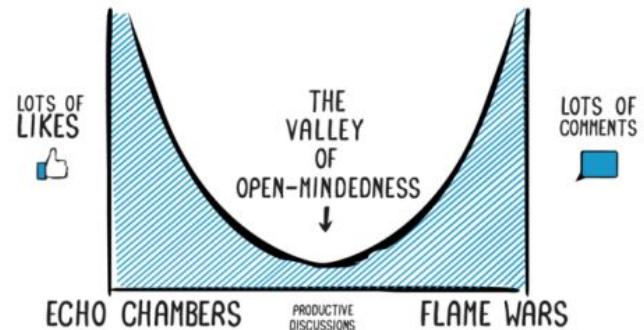
Unfortunately, it is not.

The situation in reality is even **worse**

- Simulations performed in mean field
- The observed effects can be exacerbated by the **topology** of the social network



POLITICAL DISCUSSIONS ON THE FACEBOOK



Social Tissue as Shaped By Opinions

A feedback loop model



Models

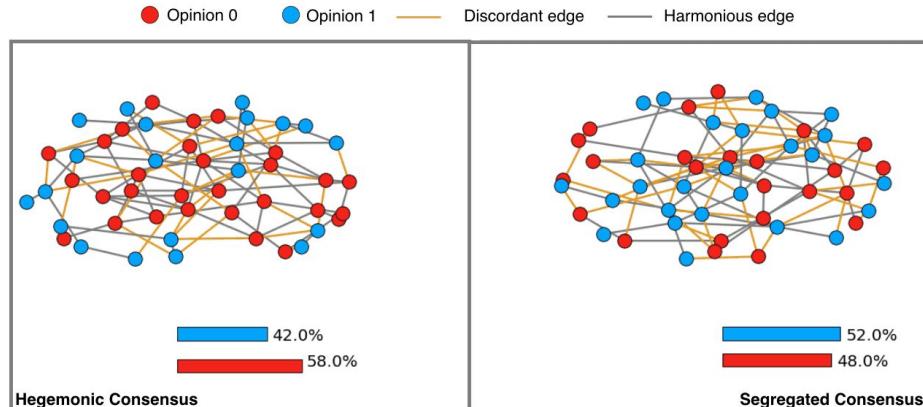
Co-Evolving Voter Model

Opinion dynamics may affect network topology

Discrete opinions: $\{-1, 1\}$

Iteration:

- A random agent i is selected with one of its neighbors j
 - If they share the same opinion nothing happens.
- Otherwise,
- with probability p :
 i detaches from j and attaches randomly to a node z that shares i 's opinion;
 - with probability $1-p$:
 i adopts j 's opinion



F. Vazquez, V.M. Eguíluz and M. San Miguel. Generic absorbing transition in coevolution dynamics. Phys. Rev. Lett., 2008).

Summarizing

Conclusion

Opinions, as well as viruses, are “objects” that spread over a social tissue.

Different assumptions on how they diffuse allow the design of (simplified and controllable) “what if” scenarios so to study specific social phenomena.

01	Discrete Opinions	<ul style="list-style-type: none">• Voter, Q-Voter• Majority• Sznajd
02	Continuous Opinions	<ul style="list-style-type: none">• Deffuant• Algorithmic Bias
03	Dynamic On & Off	<ul style="list-style-type: none">• Co-Evolving Voter

Chapter 12

Conclusion

Take Away Messages

1. Opinions diffuse through social interactions
2. They can be modeled as continuous as well as discrete variables
3. Different models design adhoc "what if" scenarios that allow analyzing real phenomena

Suggested Readings

- "Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model." Sirbu et al.

