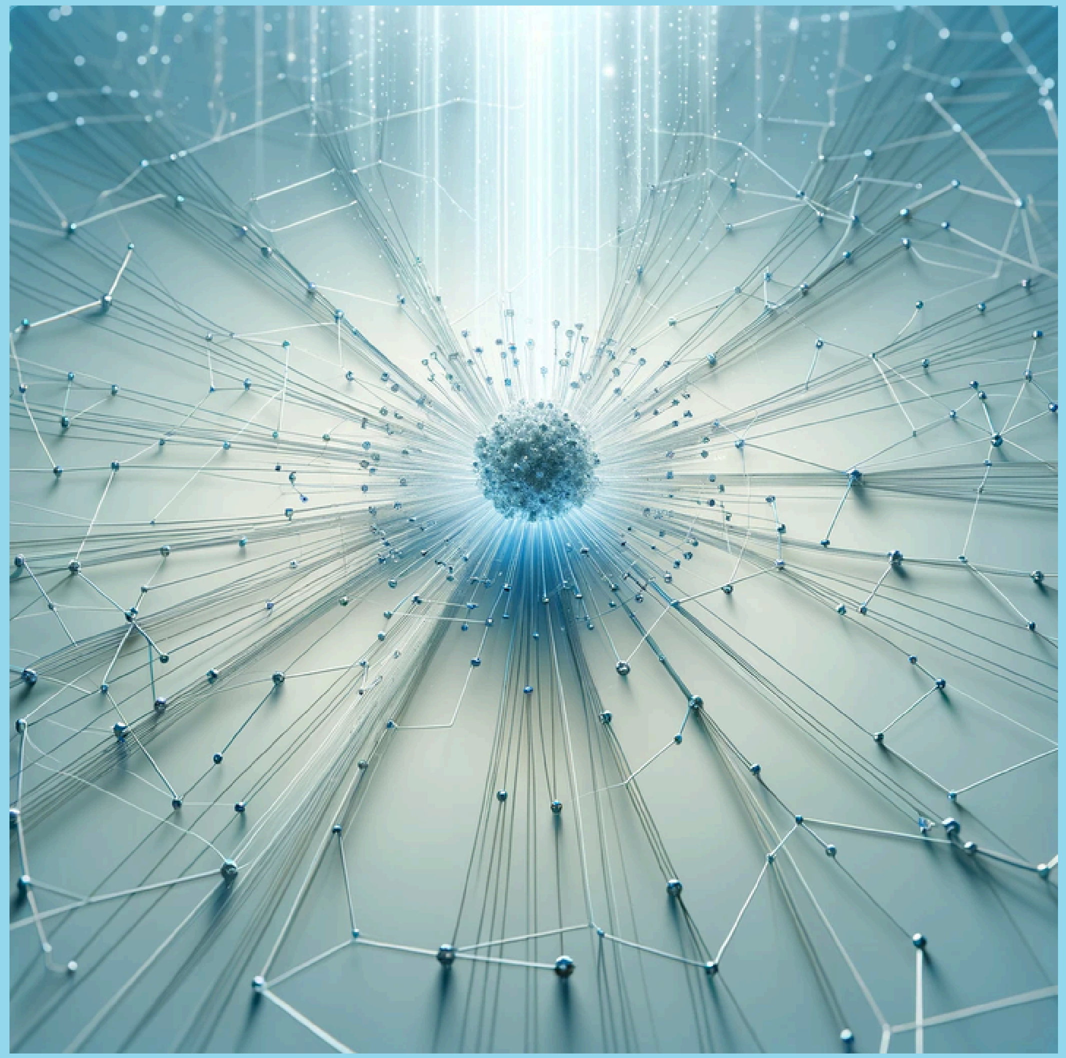


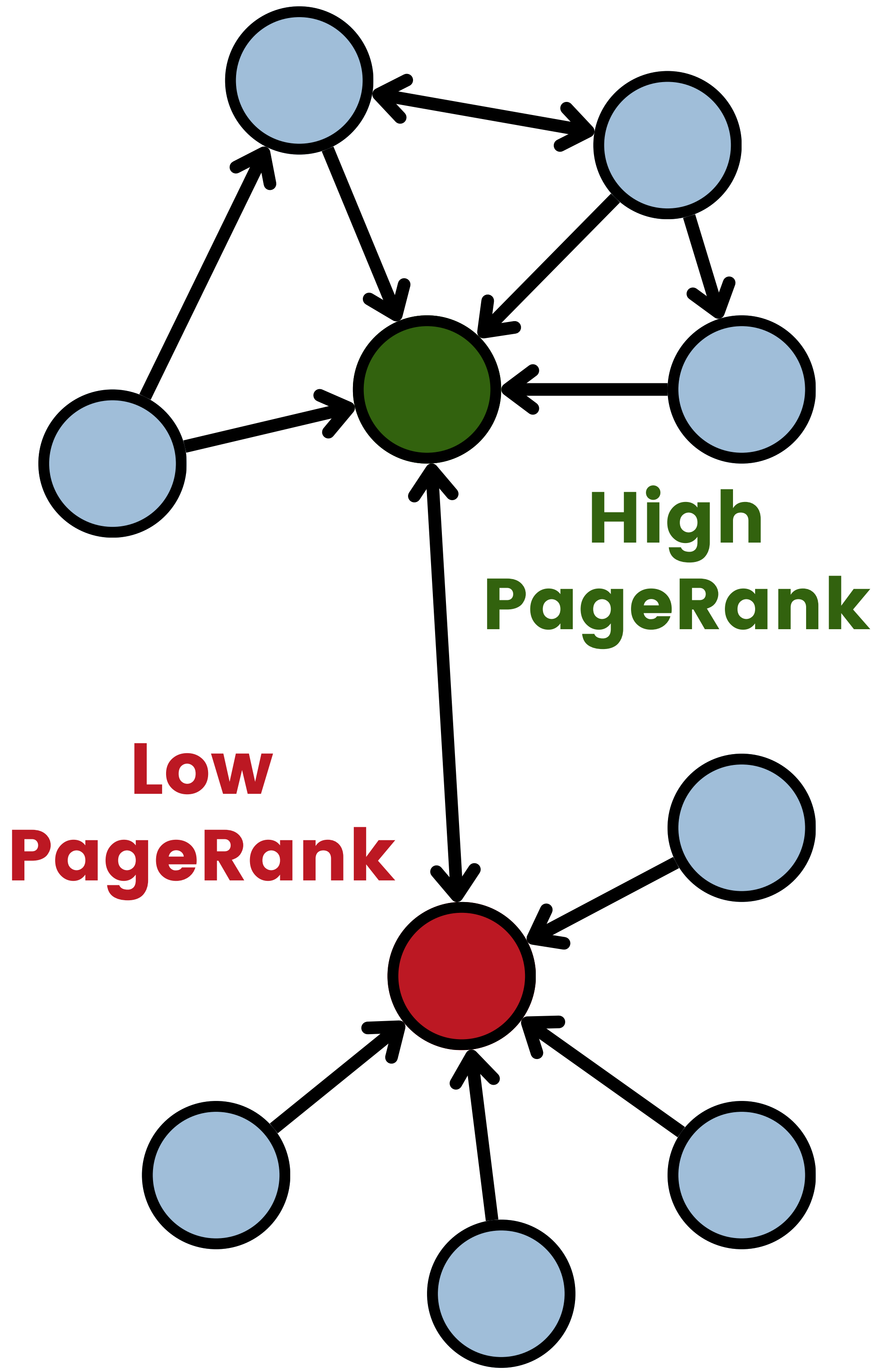
UNIVERSITÄT KONSTANZ

Processes on Networks

Computational Modelling of Social Systems

Giordano De Marzo
Max Pellert

Universität Konstanz



Recap

Networks Basics

We introduced important network measures such as degree, centralities, clustering, diameter.

Real World Networks

Real World Networks are characterized by the small world property, an high clustering and often a scale free degree distribution.

Network Formation Models

We introduced Watts-Strogatz and Barabasi-Albert models, showing how they can reproduce the features of real networks.

Outline

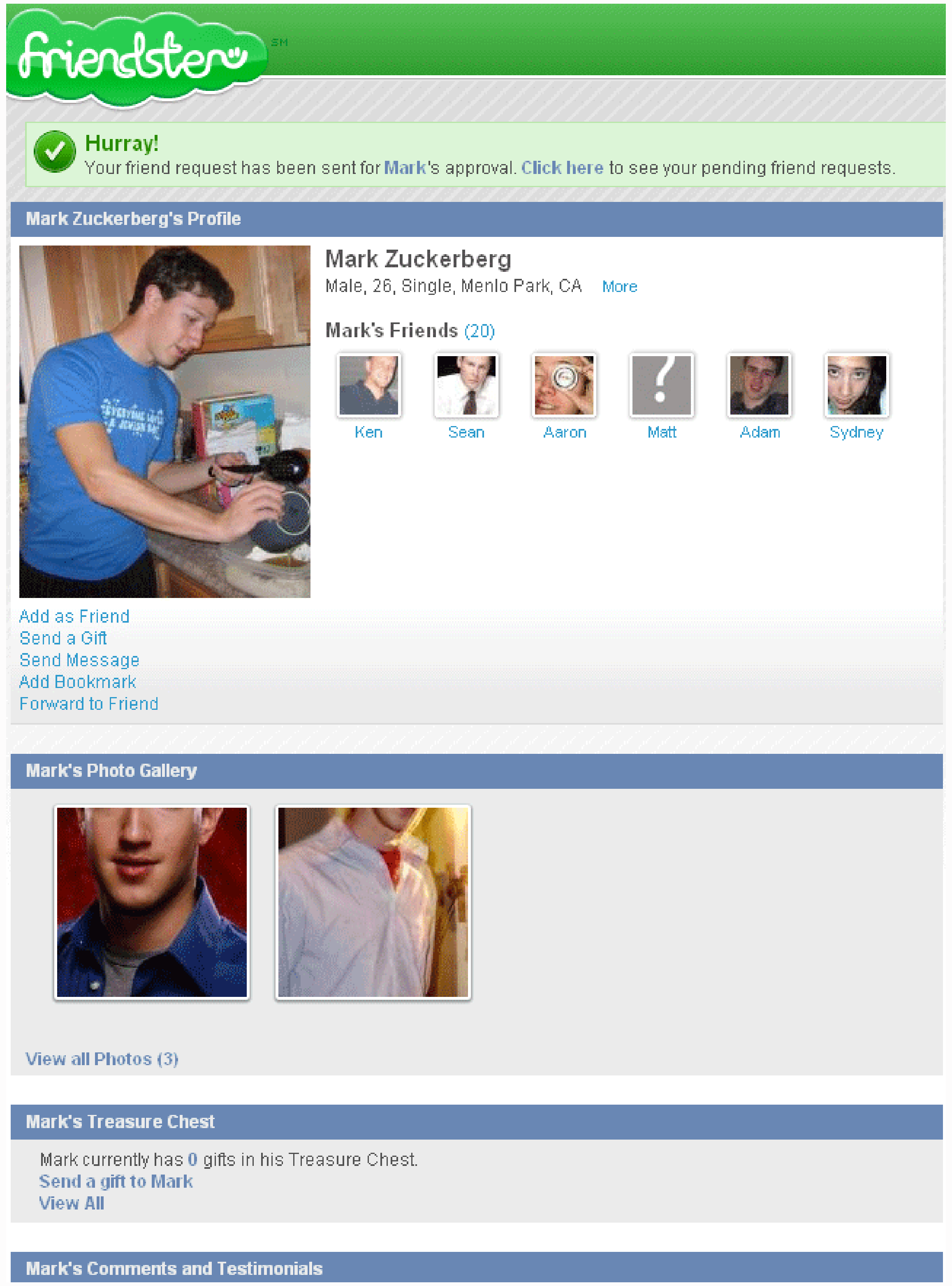
1. Resilience in Social Networks
2. K-Coreness and Social Networks Collapse
3. Epidemic Spreading Processes
4. Spreading on Social Networks

Resilience in Social Networks

The Cemetery of Social Networks

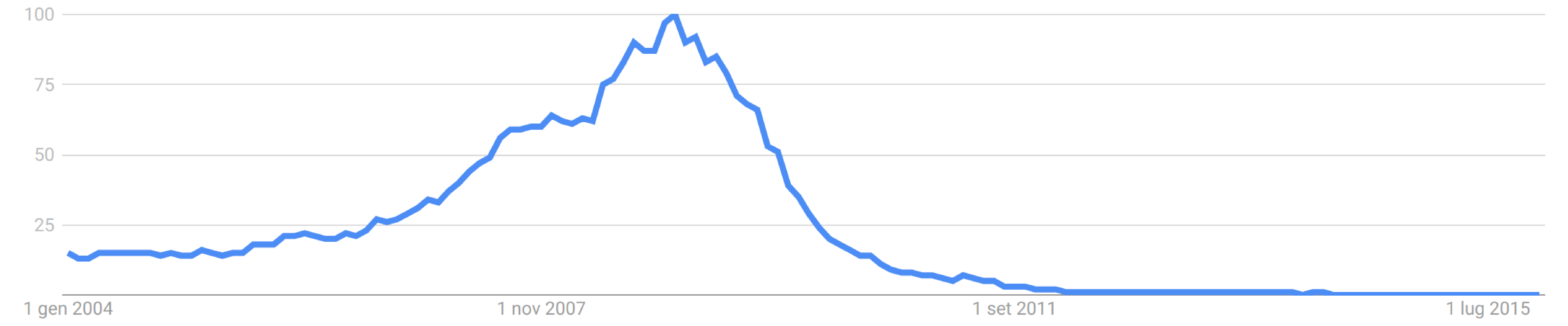
Friendster was a super successful social network, but probably nobody in this room has ever used it. Even Mark Zuckerberg was using it!

- it had a good success in the USA, but its main market was Asia
- it was created in 2003 and got 3 millions of active users just after few months
- this was before Facebook (2004) and other popular social networks



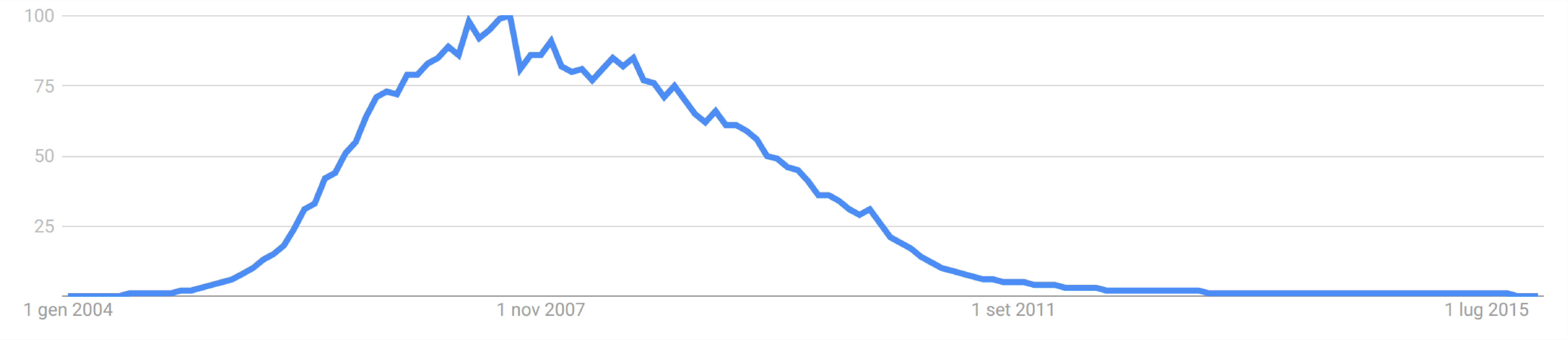
Rise and Fall of Friendster

Friendster went from 80 Million active users to disappear completely. At its peak it had more than 100M users, but it took just a couple of years for the site to be completely forgotten.

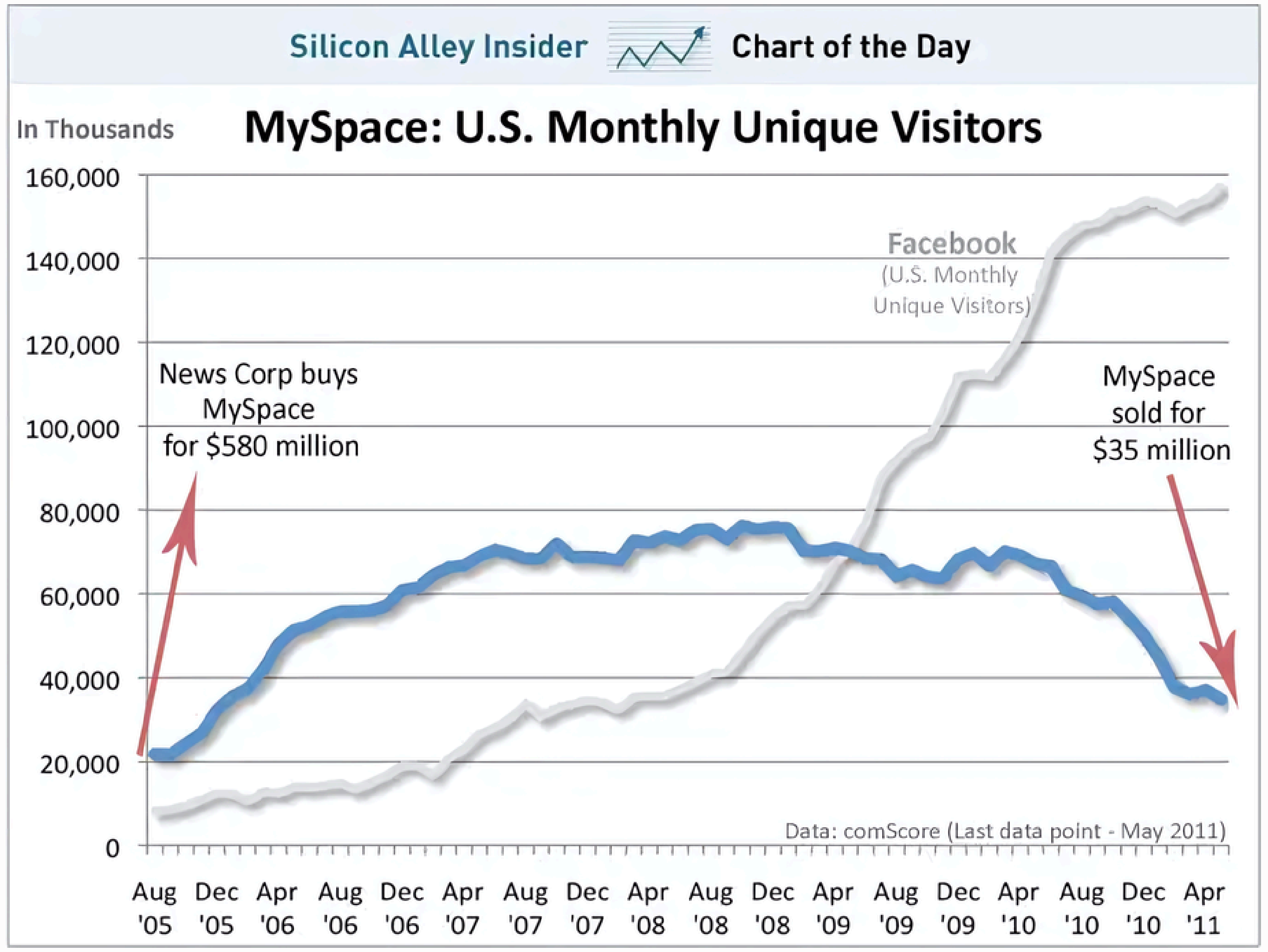


Rise and Fall of MySpace

MySpace followed a very similar trajectory. It went from being valued more than 12 Billion USD in 2008 to be bought by Justin Timberlake for 35 Million USD.



Why do Social Networks Fall?



Friendster had a huge user base and the first mover advantage, why did it failed?

- there are many other similar examples
- why do some social networks have success, while other decline?
- what are the features that make social networks very resilient or fragile?
- can we understand Friendster decline using networks theory?

Rational Users: Benefit vs Cost

We consider social network users as rational, they will balance costs and benefits

- **Benefits:** the content users receive from their friends (shares, comments) and the attention and support given by their friends (likes, votes). A simple way to model monotonic benefits is proportionally to the active friends of a user k_u.

- **Costs:** costs associated with being active, are for instance the time spent on the platform instead of doing something else or potential membership fees. A common assumption about costs is that they are relatively similar for all users, thus modeling them as a constant c.

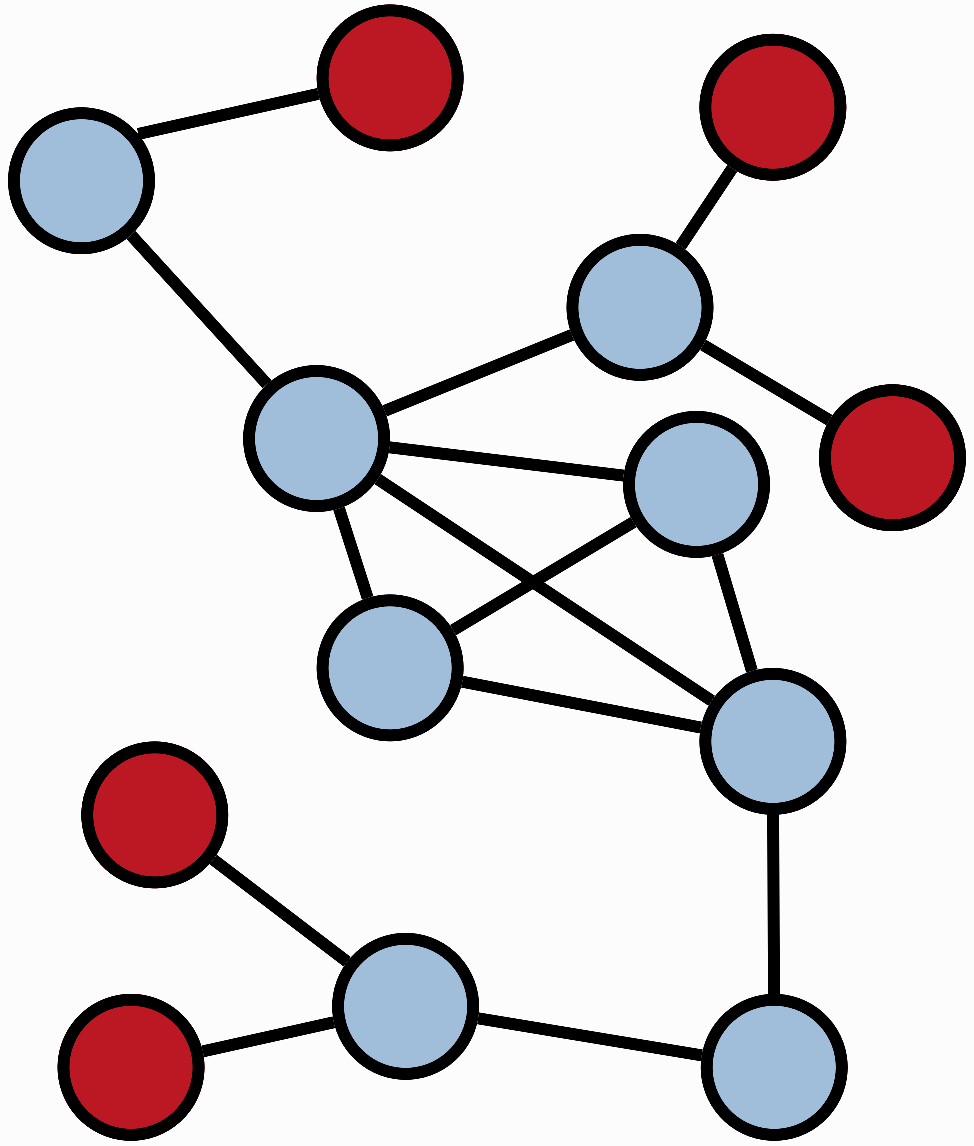
Why do Users Leave?

Users will leave the social network if their benefits are smaller than the cost of using the social network

$b \cdot k_u < c$

As a consequence users with a small degree will tend to leave the social network before users with high degree.

Note that once the low degree users leave, this may lead to other users leave, generating a chain reaction that only ends with the disruption of the social network.



K-Coreness and Social Networks Collapse

The K-Core Decomposition

We can formalize the cascade process we mentioned using the k-core decomposition of a network.

- We start setting $k_s=1$
- We then remove all nodes with degree less than or equal to k_s and their links
- We iterate until we can't remove any additional node
- We then increase k_s by a unit and we repeat the process until no nodes are left

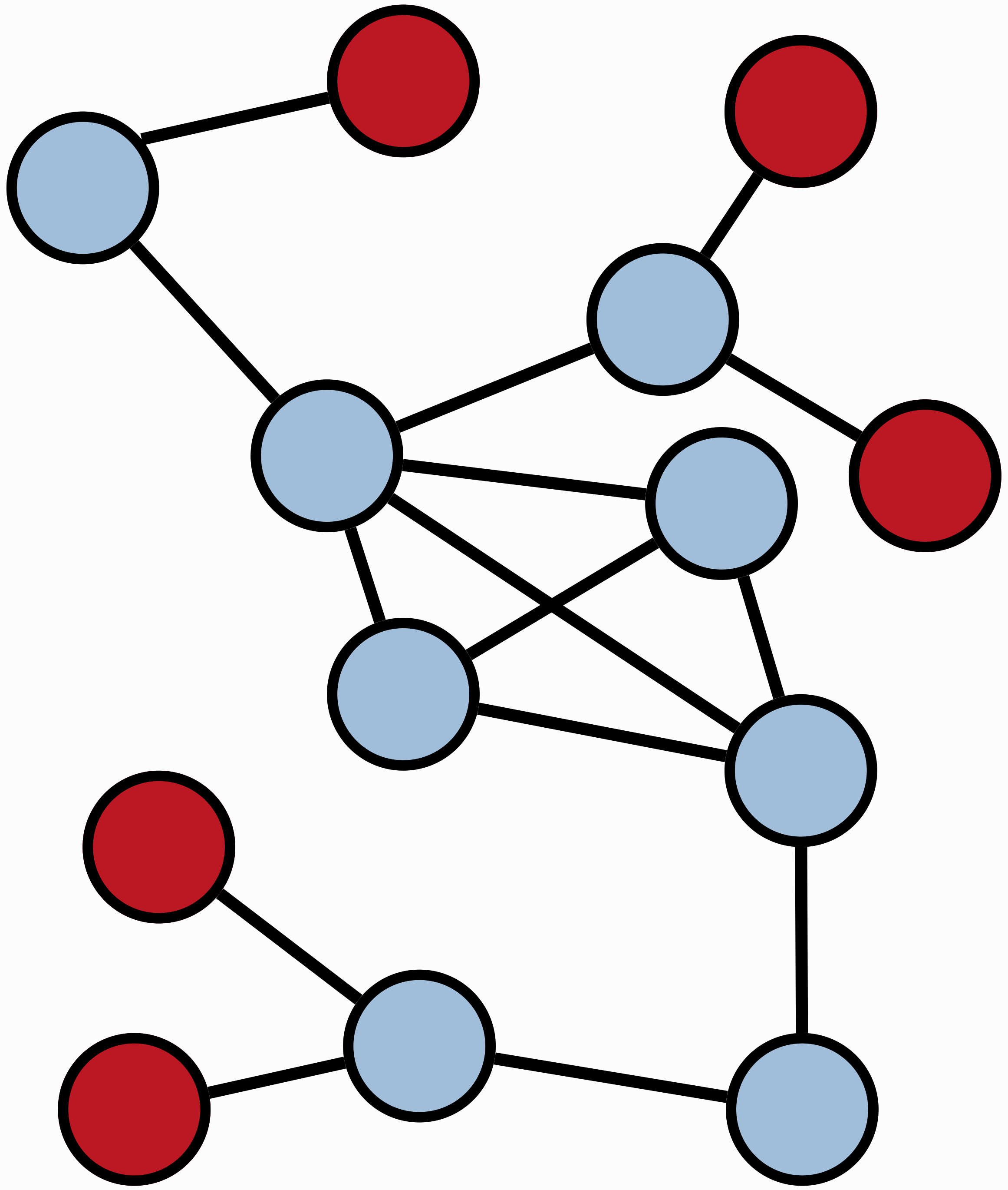
We denote

- k-shell of the network the set of all nodes and edges removed for $k_s=k$
- k-core of the network the set of all k-shells with $k_s \geq k$, so all the nodes that survive after a cascade starting from $k_s=k$

K-Core **Decomposition** Example

Let's consider a practical example with k_s=1

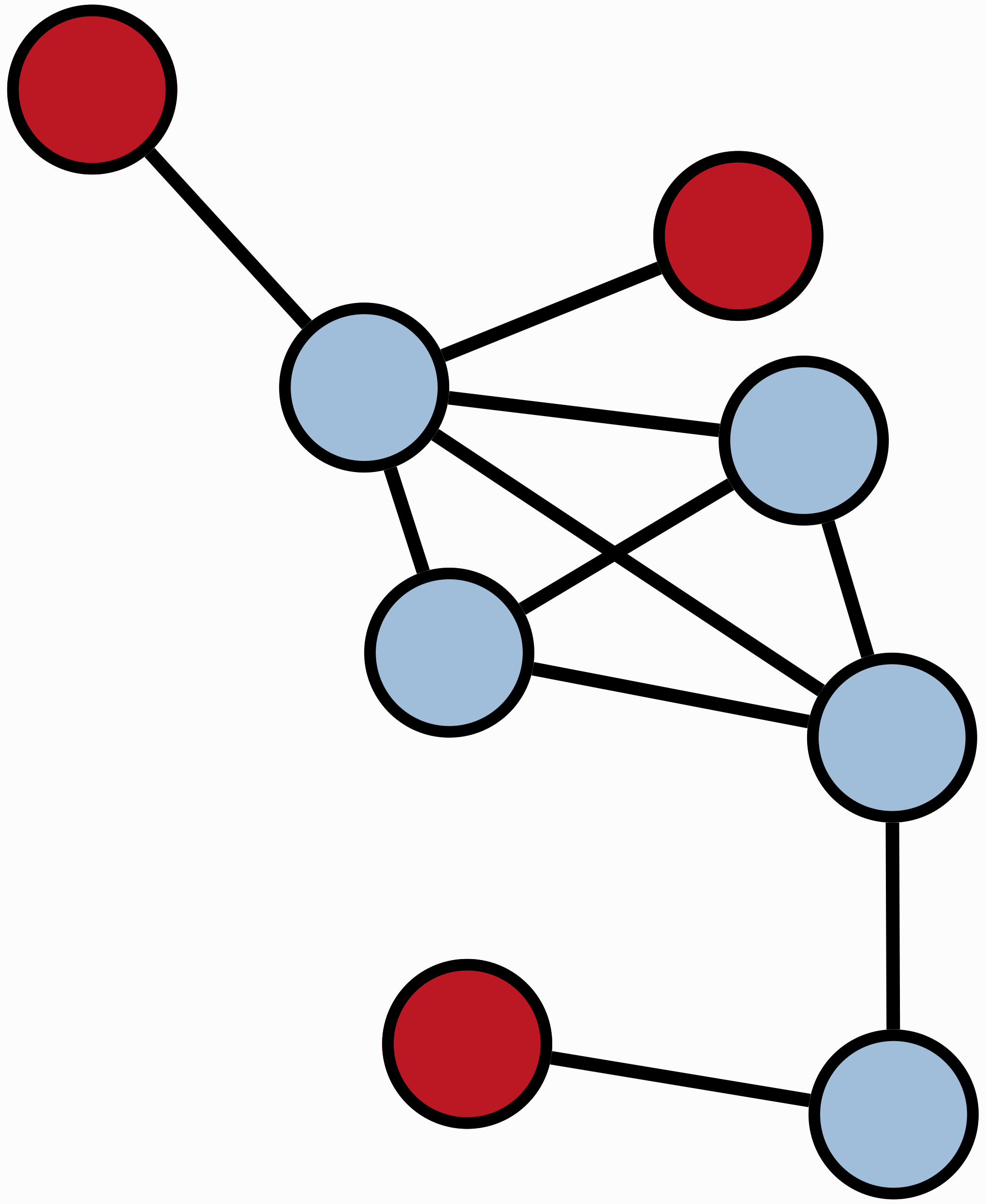
- red nodes have degree equal to one
- we remove all of them



K-Core **Decomposition** Example

Let's consider a practical example with k_s=1

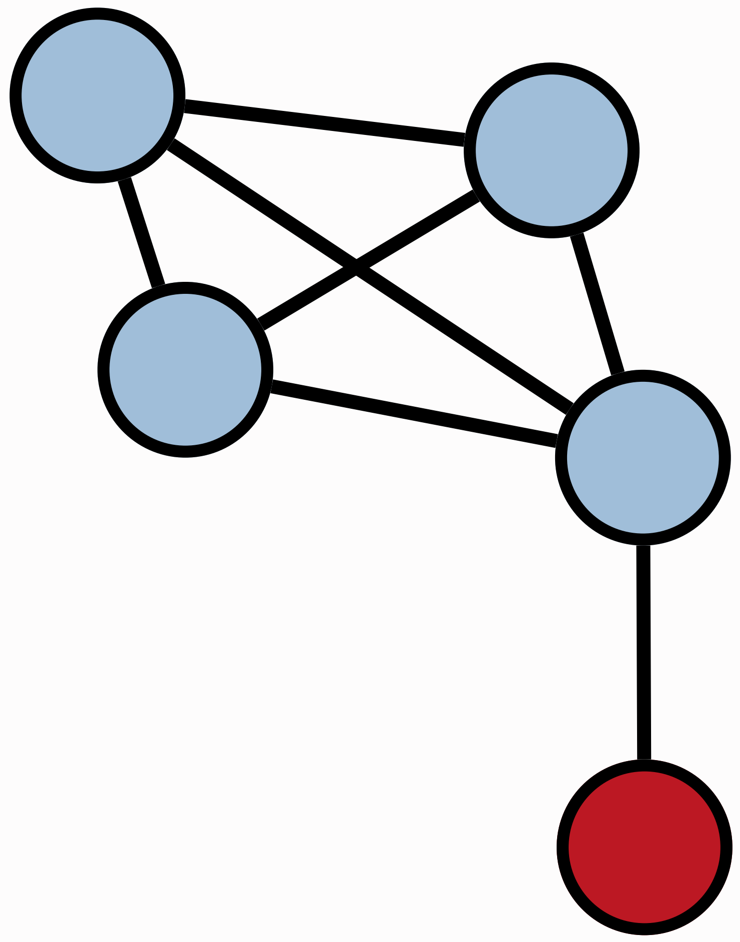
- red nodes have degree equal to one
- we remove all of them
- after removing the first set of nose, some of the surviving nodes now have k=1
- we remove also them



K-Core Decomposition Example

Let's consider a practical example with k_s=1

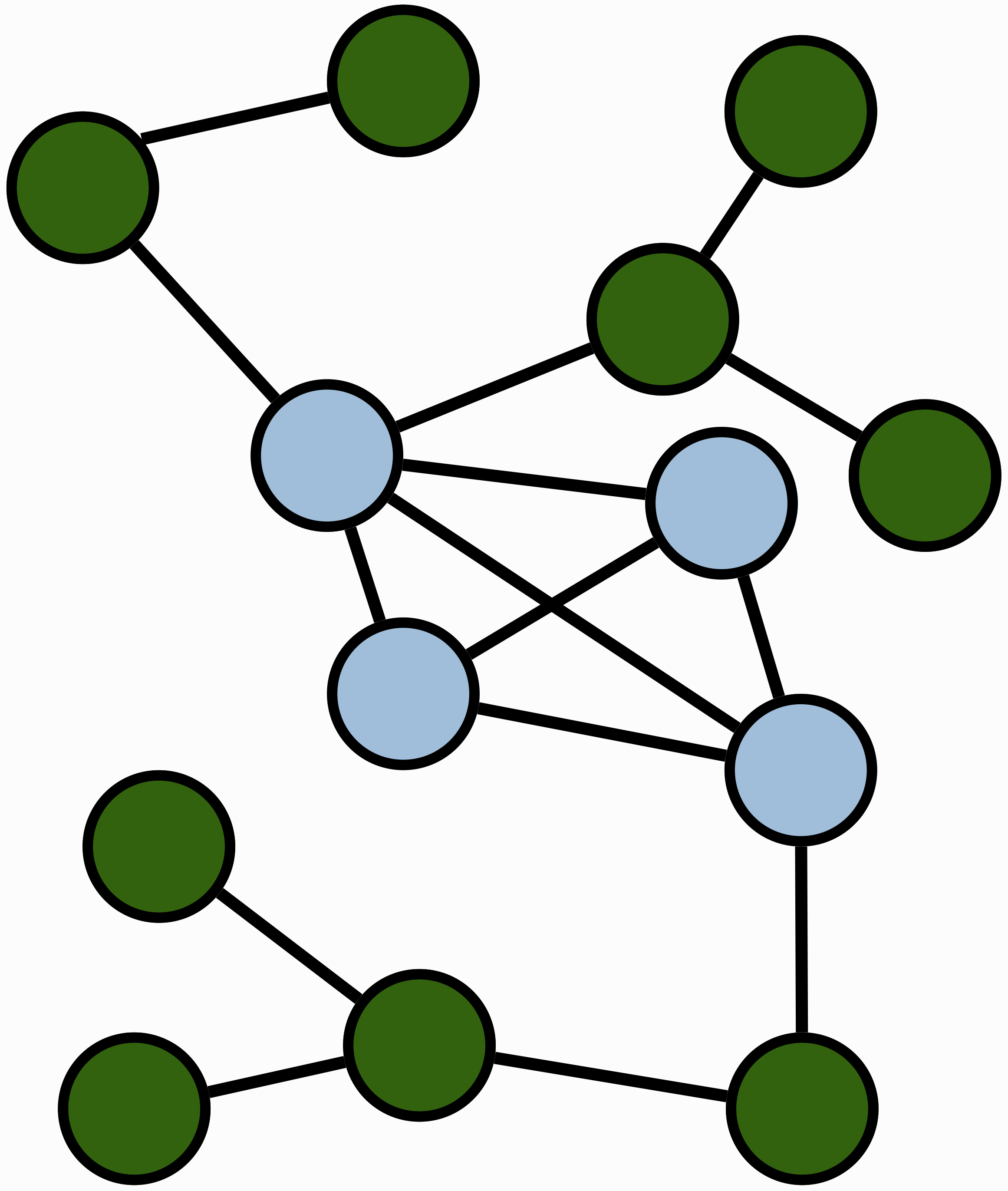
- red nodes have degree equal to one
- we remove all of them
- after removing the first set of nose, some of the surviving nodes now have k=1
- we remove also them
- we iterate the process



K-Core **Decomposition** Example

In the end we obtain:

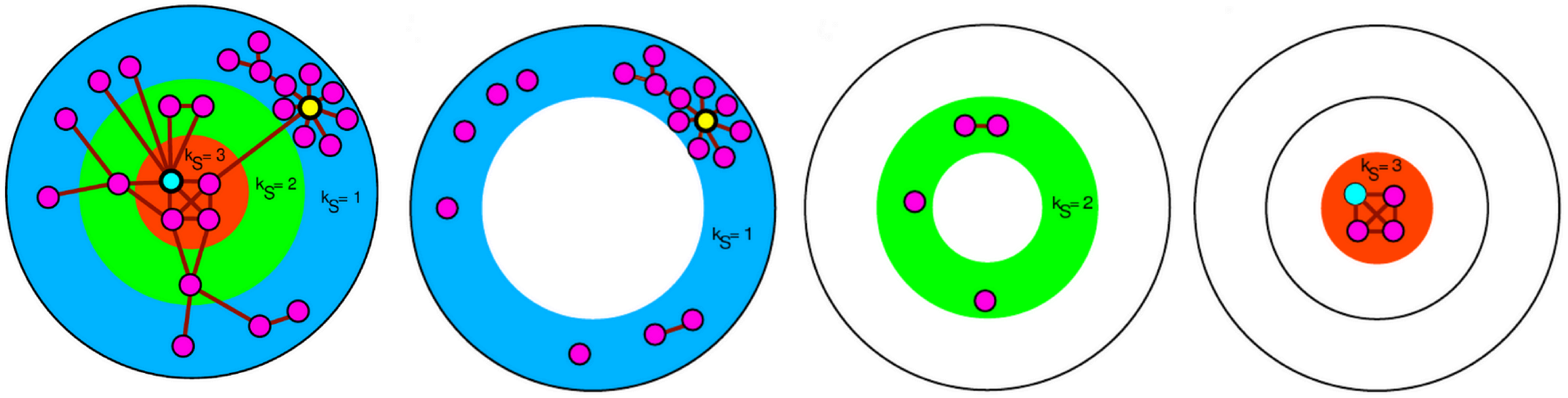
- the 2-core of the network (blue nodes). All the nodes in the 2-core have at least degree 2 so they will never be removed
- the 1-shell of the network (green nodes). The are all the nodes that get removed in the cascade chain process starting from k_s=1



Coreness Centrality

Nodes with the same degree can have very different properties. Some of them are captured by the coreness-centrality, that is defined as the k-shell number of a node

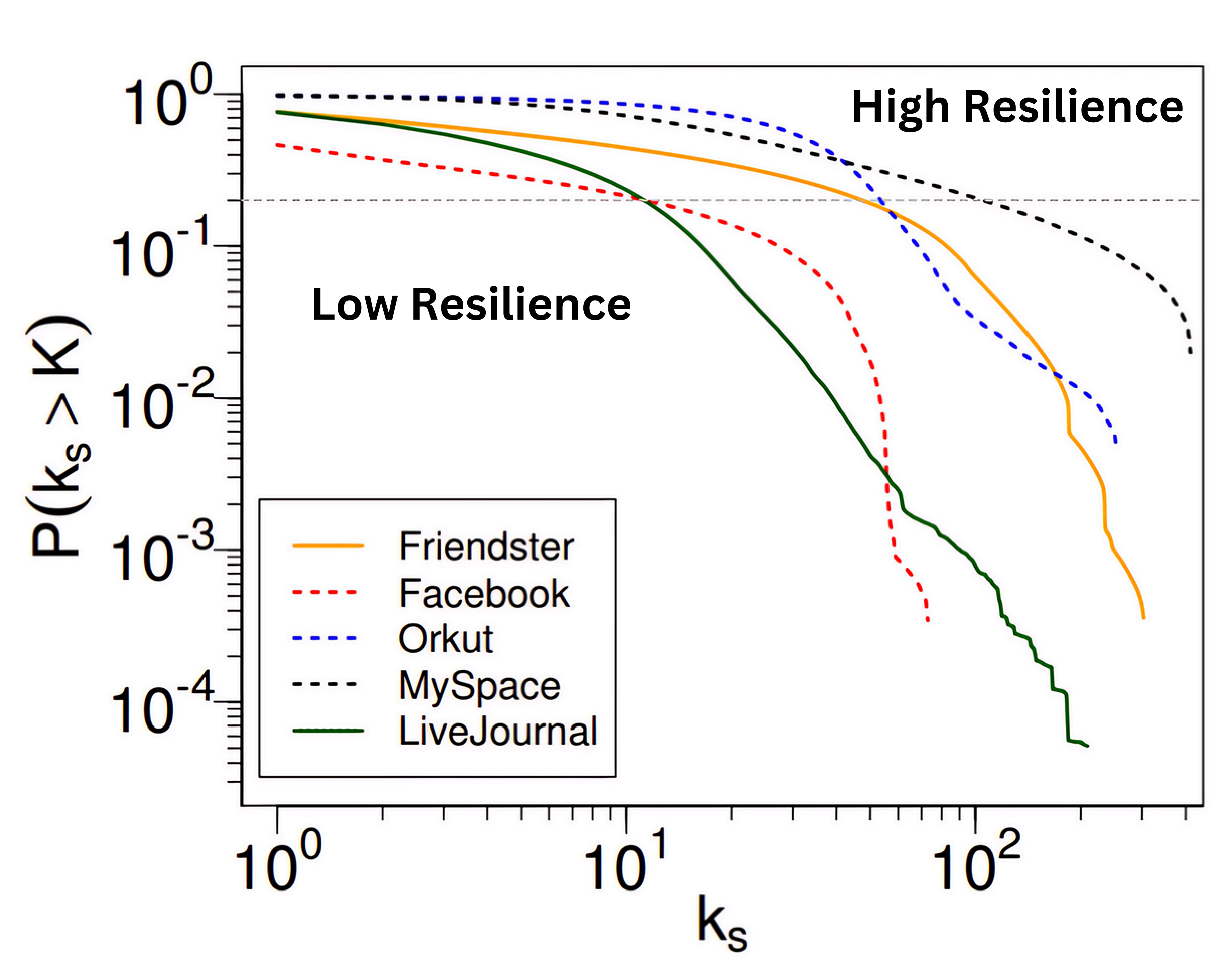
- the light node and the yellow node have the same degree, but the yellow one has a lower coreness-centrality



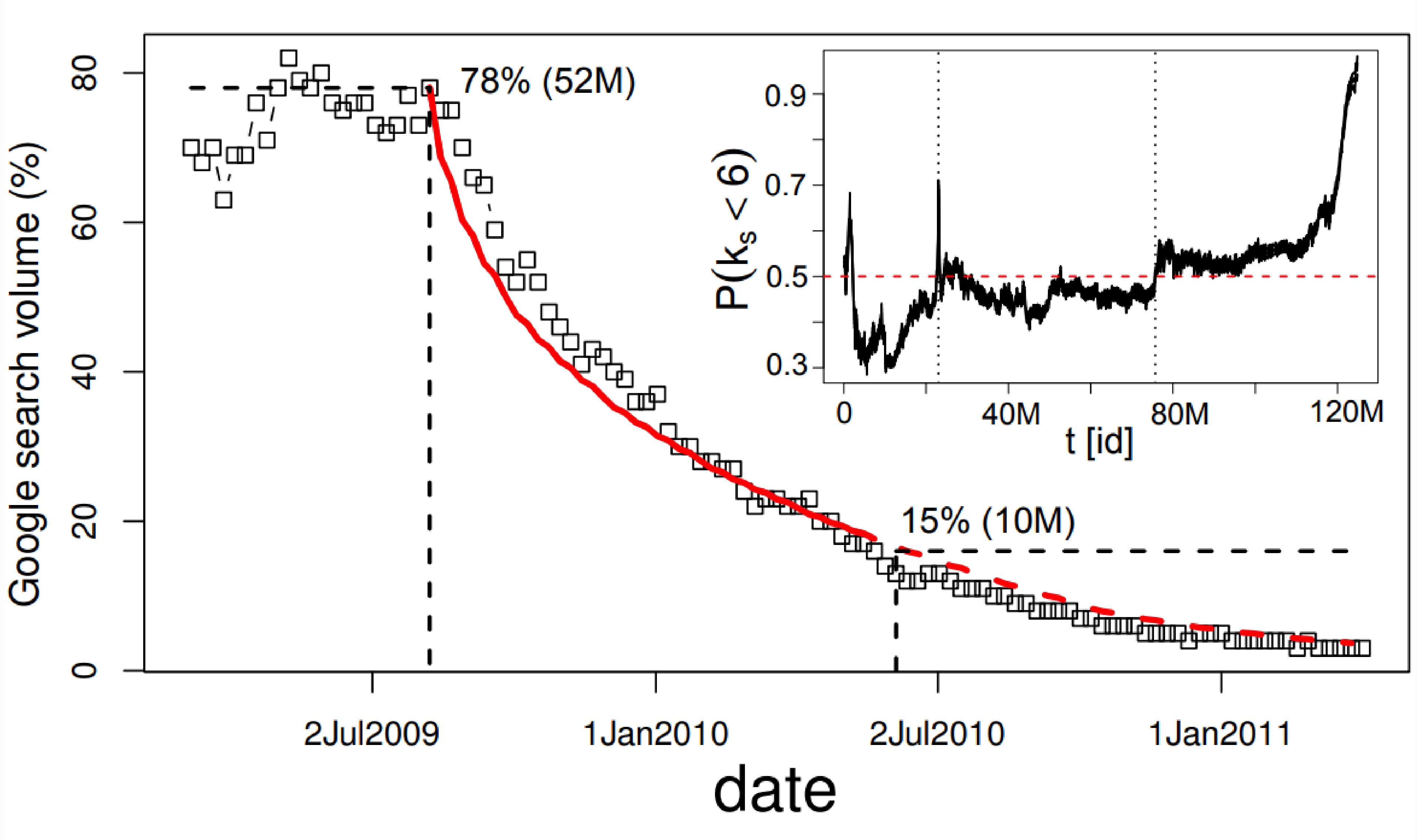
Coreness and Resilience

The cumulative distribution of coreness-centrality measures the resilience of a social network

- The cost to benefit ratio defines a critical value of the degree $K=c/b$ below which users with degree $k_u<K$ will leave
- The remaining active social network is the k-core corresponding to K



Modeling Friendster **Collapse**



Garcia, David, Pavlin Mavrodiev, and Frank Schweitzer. "Social resilience in online communities: The autopsy of friendster." Proceedings of the first ACM conference on Online social networks. 2013.

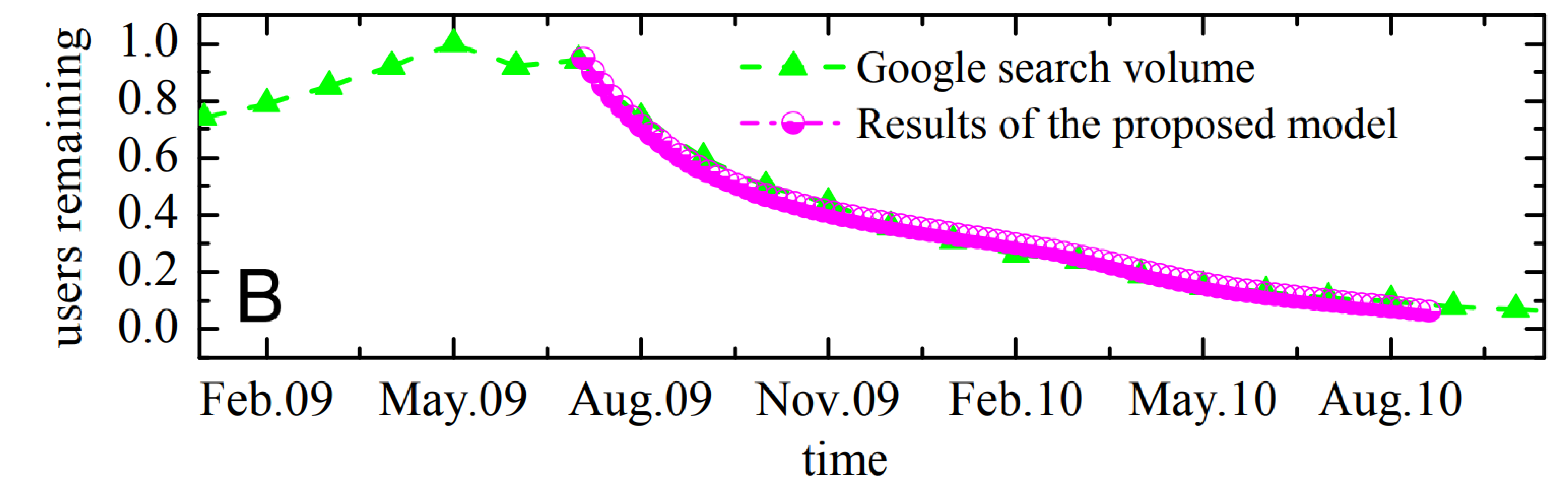
We can model Friendster collapse as an iterative k-core cascade process

- we use google data to get the number of active users over time
- we assume the cost to increase linearly over time
- the red line shows the fit to the Friendster collapse
- the inset shows the fraction of nodes with coreness below the median

A More Realistic Model

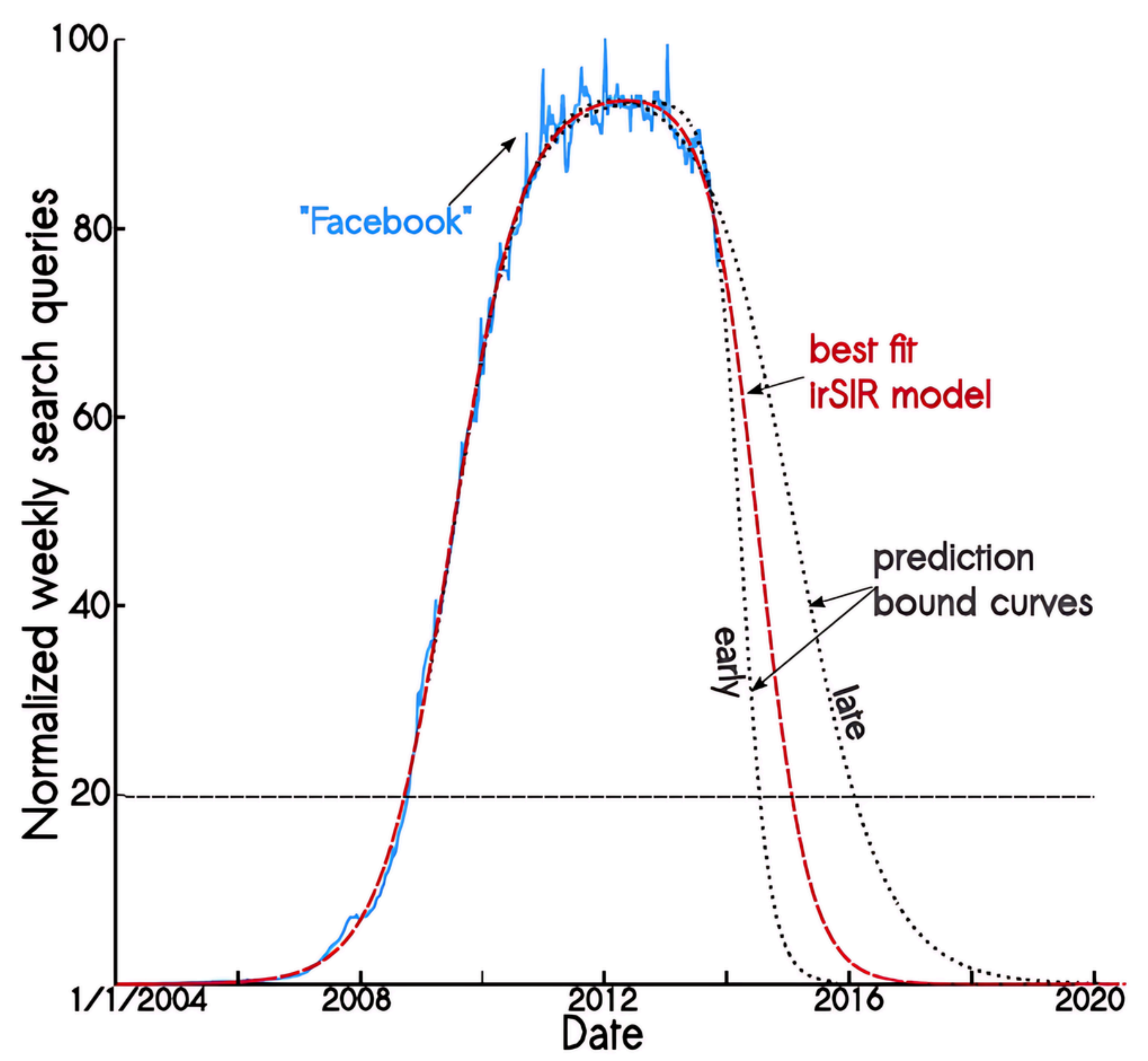
The previous model is not realistic, it requires the cost to increase linearly over time. We can get a similar fit with fixed cost, provided that

- users leave also when a fraction of their friends becomes inactive, even if their degree is still high (if many of your friends leave, you also leave)
- users don’t always leave, but they do so with a given probability



Yu, Yi, et al. "System crash as dynamics of complex networks." Proceedings of the National Academy of Sciences 113.42 (2016): 11726-11731.

The Death of Facebook



Motivated by the success of using Google trends data to model the collapse of Friendster, researchers applied this to Facebook

- looking at data in 2014 researchers observed a decline in the search volume
- they applied an epidemic spreading process to model this
- they came to the conclusion that Facebook was gonna fall very shortly

Facebook is not in great shape, but it is still there!

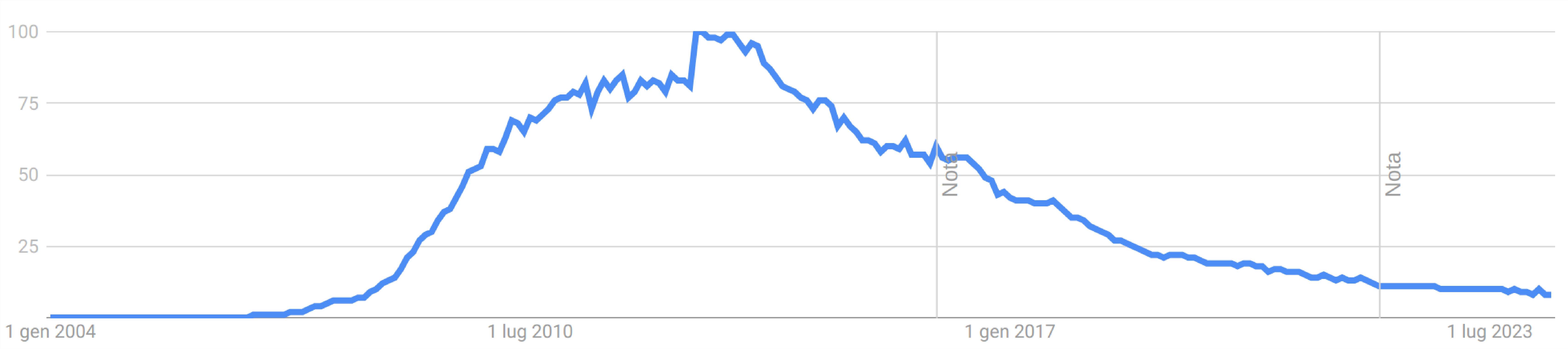
Epidemiological modeling of online social network dynamics. John Cannarella, Joshua A. Spechler. Arxiv preprint (2014)

Limits of Google Trends

Researchers at Facebook where not very happy and replied showing that Google Trends can give very dubious results. For instance below you see the trend for the keyword “Princeton“.

Limits of Google Trends

The Facebook examples show that decrease in search volume is a decrease in information searching about the social network, not a decrease in access and use. Nowadays there are not much searches including the keyword “Facebook”, but the social network is still very popular!



Epidemic Spreading Processes

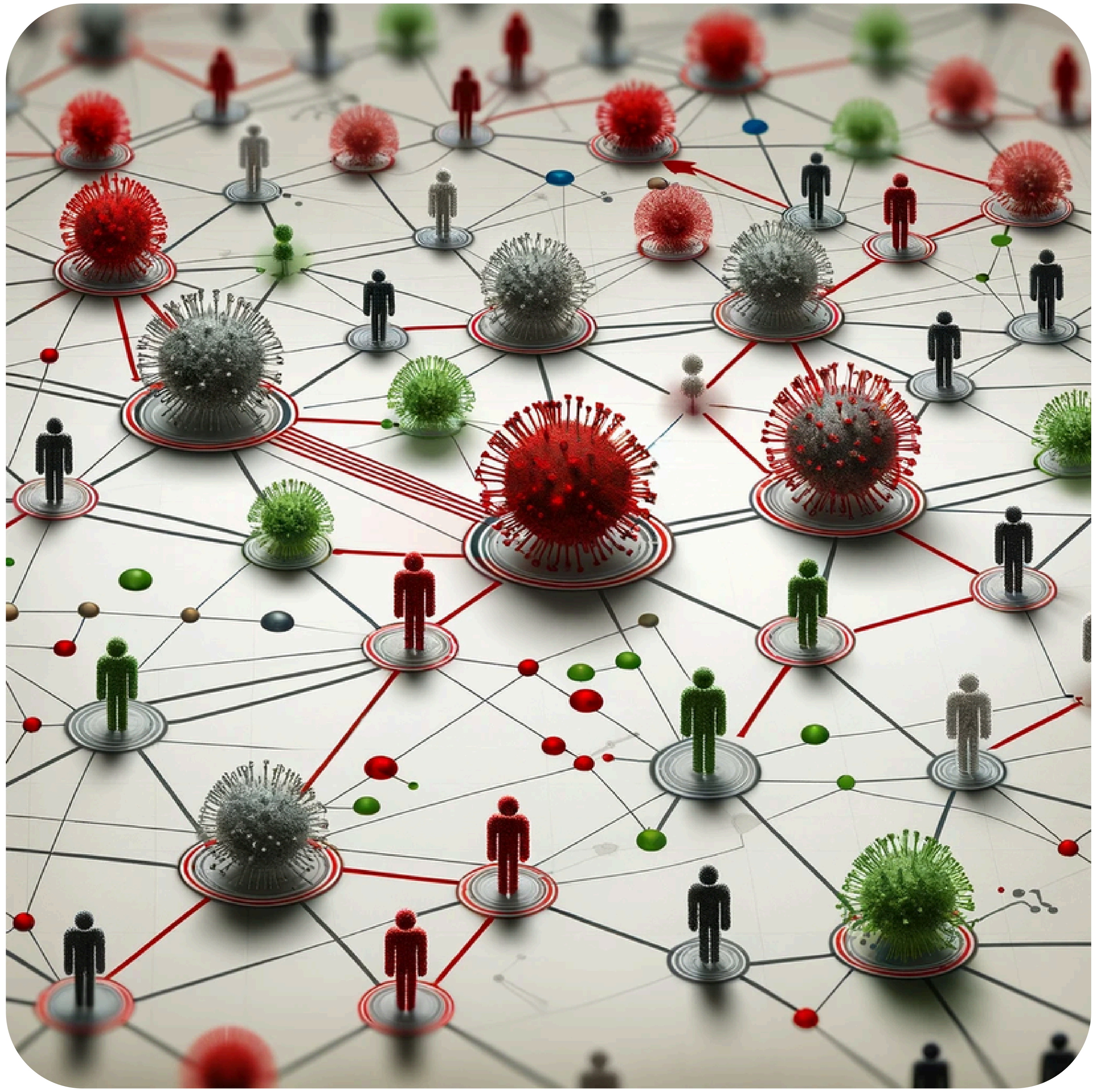
Epidemic Spreading

A very relevant process taking place on networks is epidemic spreading

- epidemic spreading models describe how an illness spread in a group of individuals connected on a network
- they have been crucial in mitigating the effects of Covid and in guiding policies

There are two macro epidemic model

- SIR model (virus with immunity eg. measles)
- SIS model (virus without immunity eg. seasonal influenza or flu)



The SIR Model

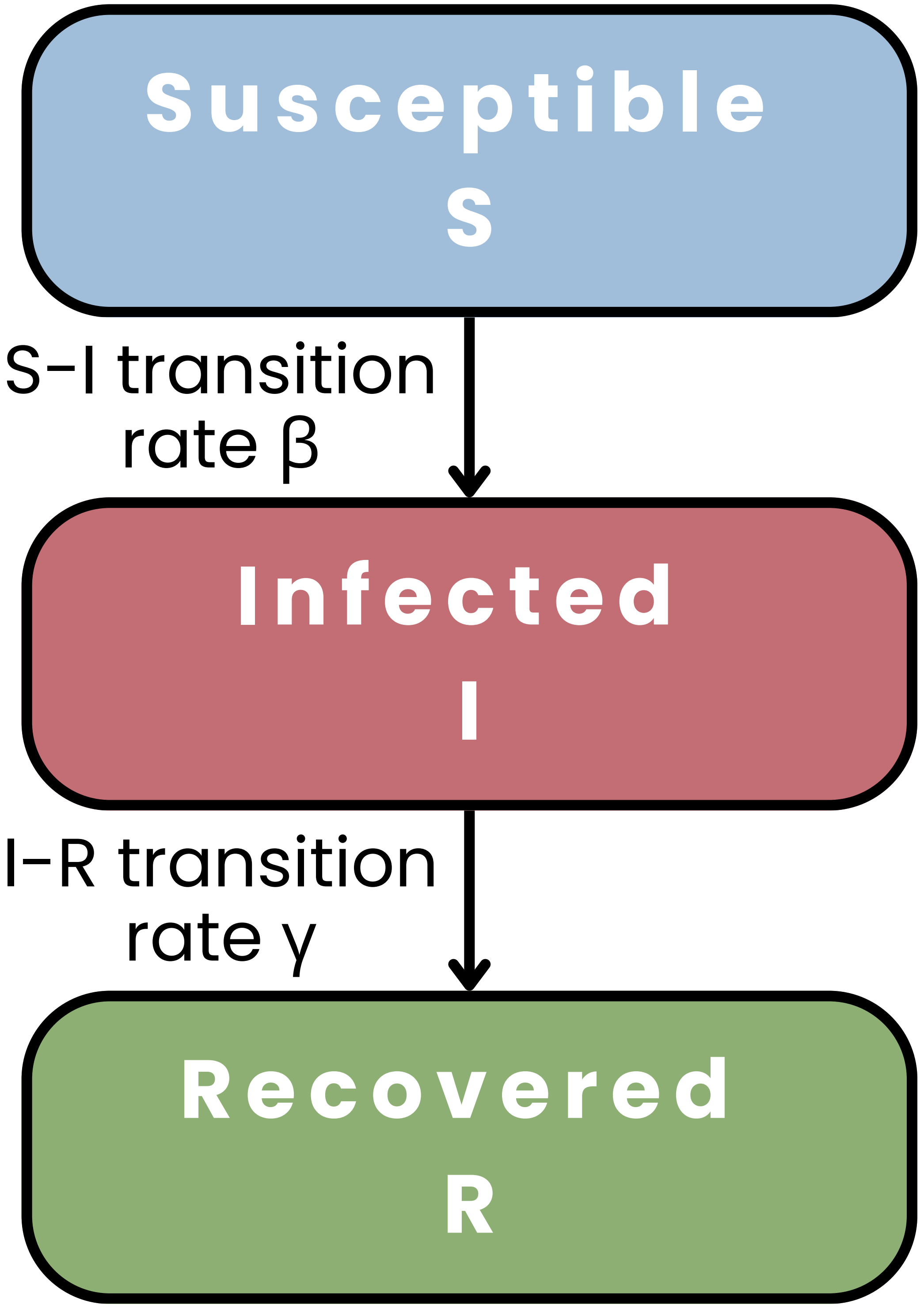
In the SIR model individuals can be in 3 possible different states

- **Susceptible S:** not infected, could be infected
- **Infected I:** has disease and is contagious
- **Recovered R:** not contagious and immune

There are (probabilistic) transitions between states:

- From S to I: infection from another infected individual.
- From I to R: recovery from disease, death, or permanent isolation.

Transitions happen based on parameters β and γ respectively.



SIR Transition Equations

On a fully connected network the SIR model is described by a set of 3 (coupled) differential equations

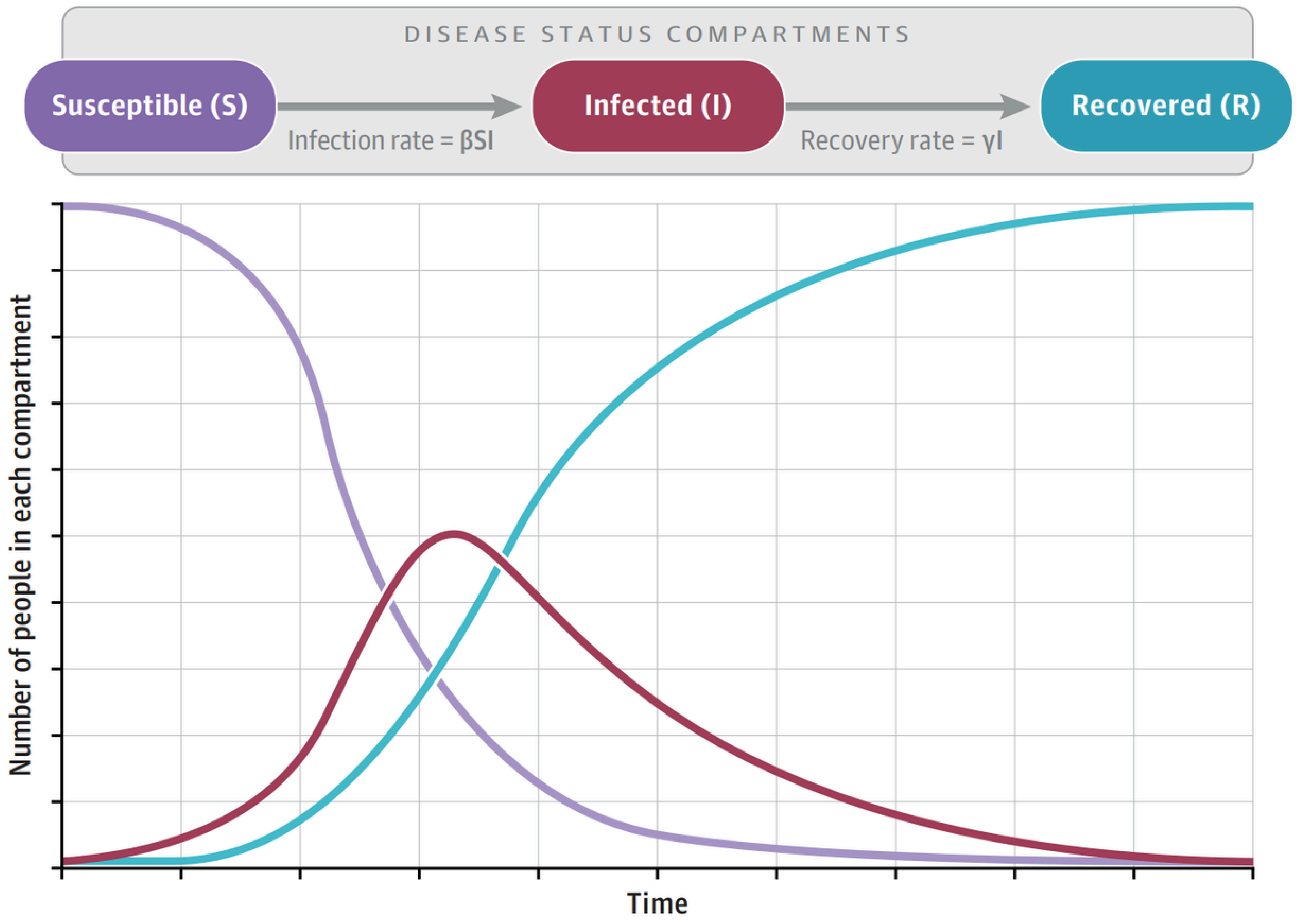
- Susceptibles diminish as they get Infected with a rate $-\beta IS/N$
- Infected increase the same way and diminish at a rate γI
- Recovered only increase from Infected with the same rate
- Parameters are not just a biological property of the disease (vaccines, lockdown etc)
- The basic reproduction number $R_0=\beta/\gamma$ is the mean number of new infections caused by a single infected individual.

$$\frac{dS}{dt} = -\frac{\beta IS}{N}$$

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

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

SIR Epidemic Curves



The SIS Model

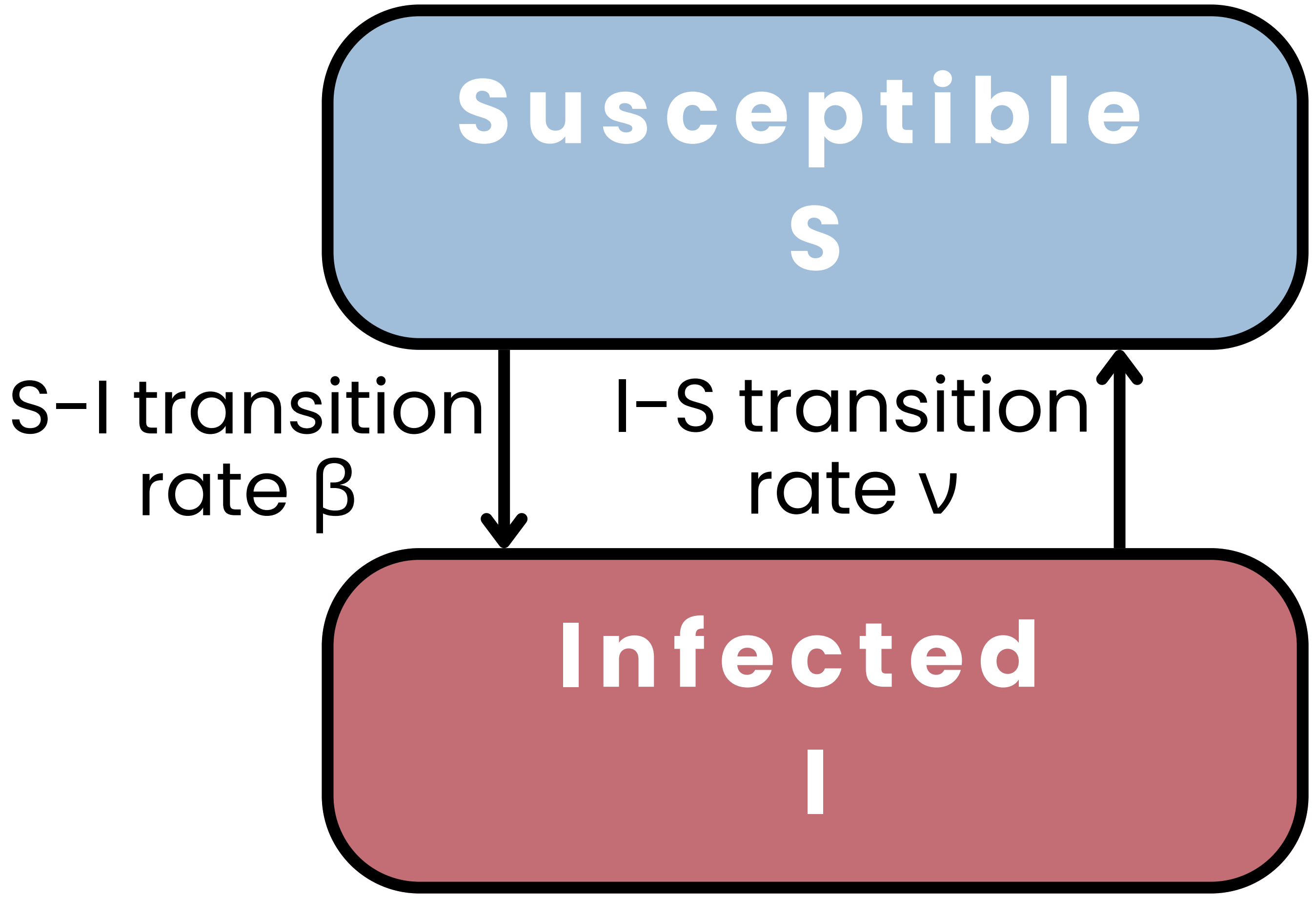
In the SIS model individuals can only be in 2 possible different states

- **Susceptible S:** not infected, could be infected
- **Infected I:** has disease and is contagious

Also in this case there are (probabilistic) transitions between states:

- From S to I: infection from another infected individual.
- From I to S: recovery from disease, but a new infection is possible

Transitions happen based on parameters β and ν respectively.



SIS Transition Equations

On a fully connected network the SIS model is described by a set of 2 (coupled) differential equations

- Susceptibles diminish as they get Infected with a rate $-\beta IS/N$ and increase due to recovery with rate νI
- Infected increase and diminish with the opposite rates

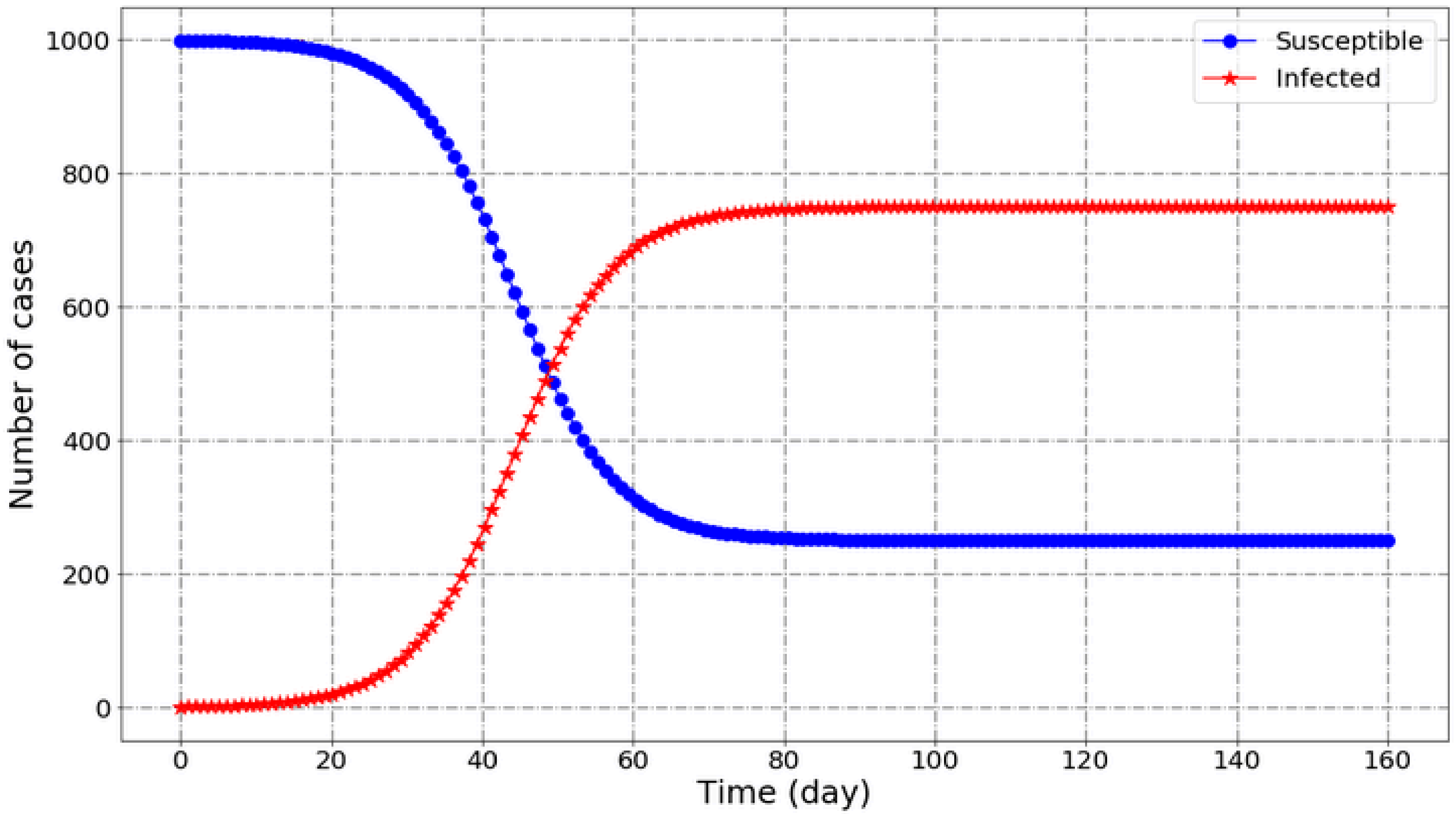
The fate of the infection depends on the values of the parameters

- if $\beta<\nu$ the infection will eventually day
- if $\beta>\nu$ instead the infection will never die and will keep surviving in the population, infecting, on average $(1-\nu/\beta)N$

$$\frac{dS}{dt} = -\frac{\beta IS}{N} + \nu I$$

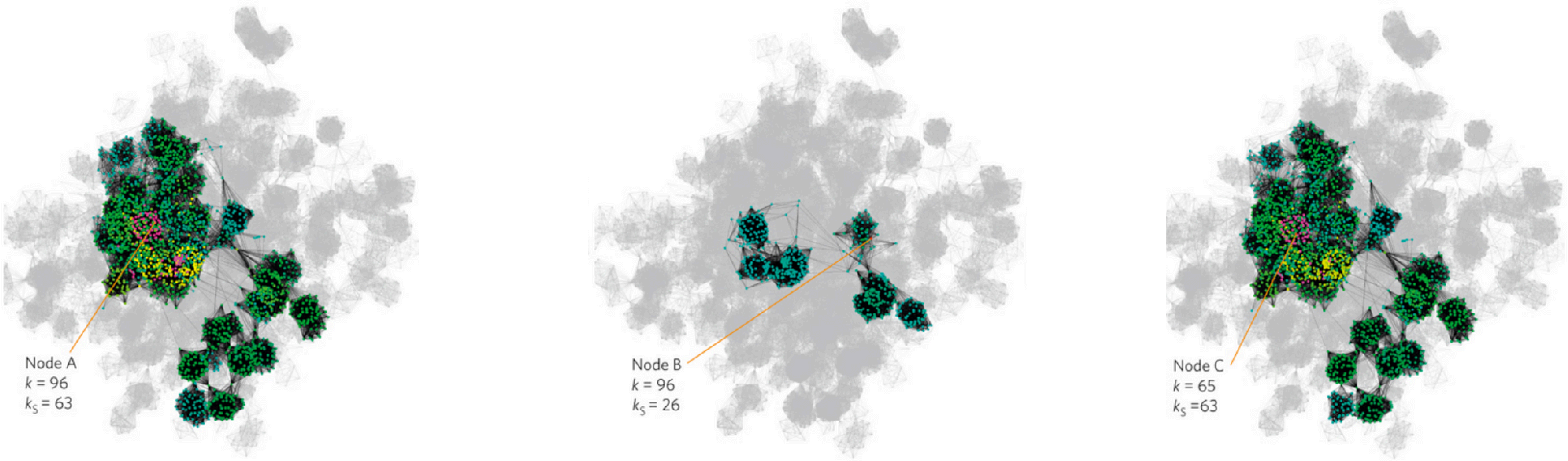
$$\frac{dI}{dt} = \frac{\beta IS}{N} - \nu I$$

SIS Epidemic Curves



Identifying Super-Spreaders

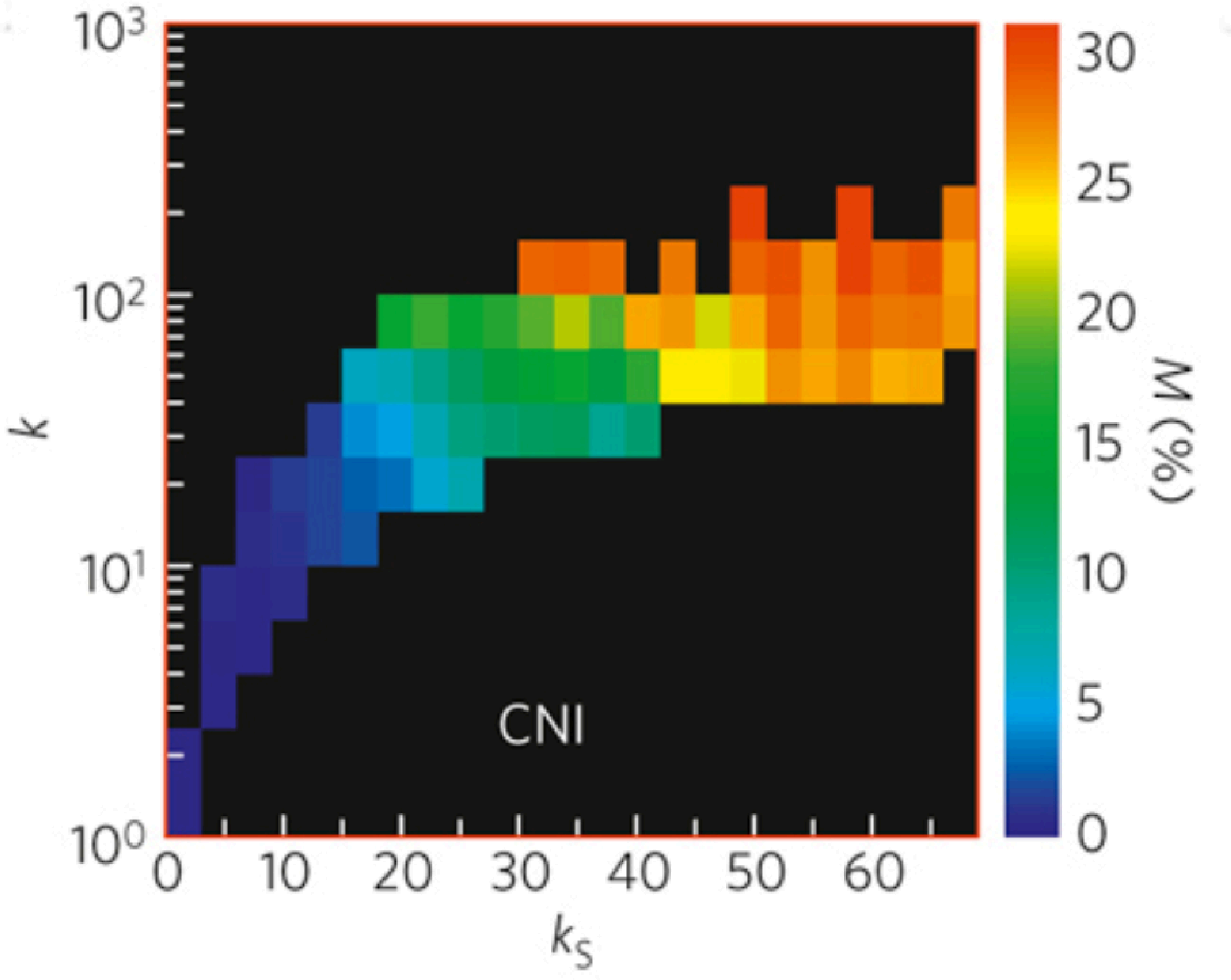
How can we identify super-spreaders in networks? Below you see the epidemic cascade generated by different initial nodes. The coreness centrality seems to work much better than the degree.



Degree vs Coreness

In order to make this more precise we can simulate epidemic spreading on real networks

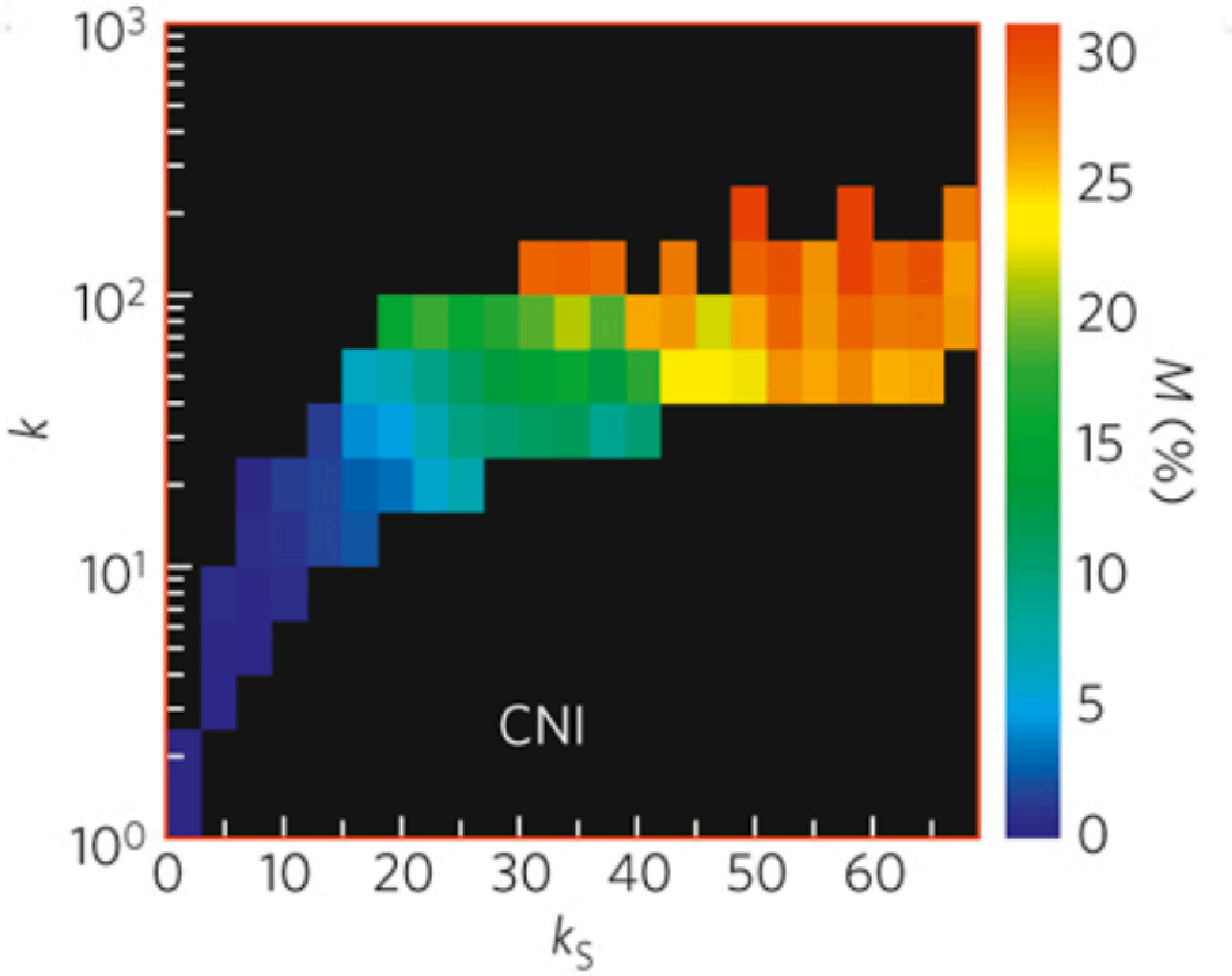
- we start the epidemic with just an infected node
- we compute the relative size M of the average epidemic it produces
- we repeat the process for all nodes in the network
- we study which nodes properties are responsible for larger epidemics



Degree vs Coreness

Example 1: Contacts in Swedish hospital

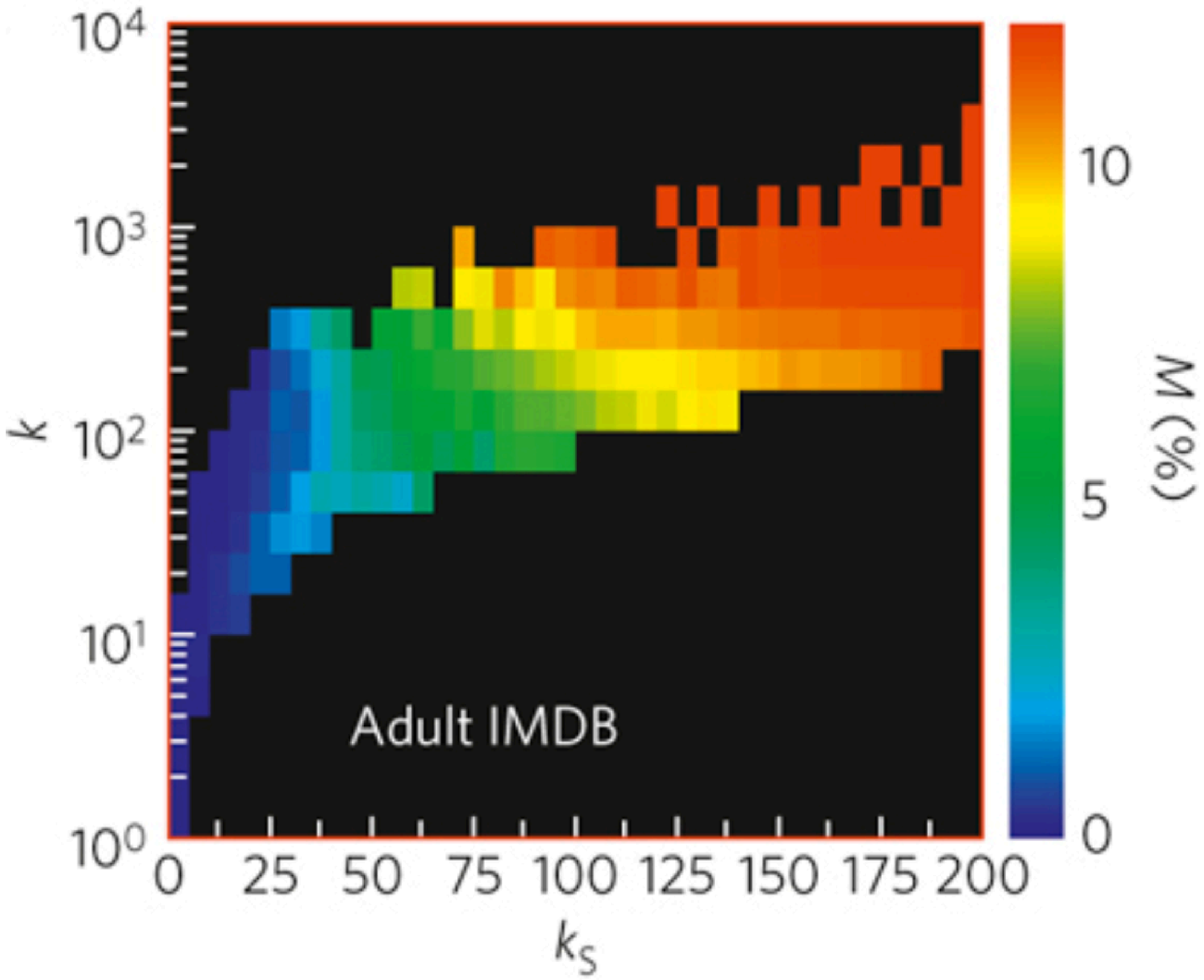
- Nodes are patients
- Edges connect patients that have been in the same room at the same time



Degree vs Coreness

Example 2: Actors of adult movies in IMDB

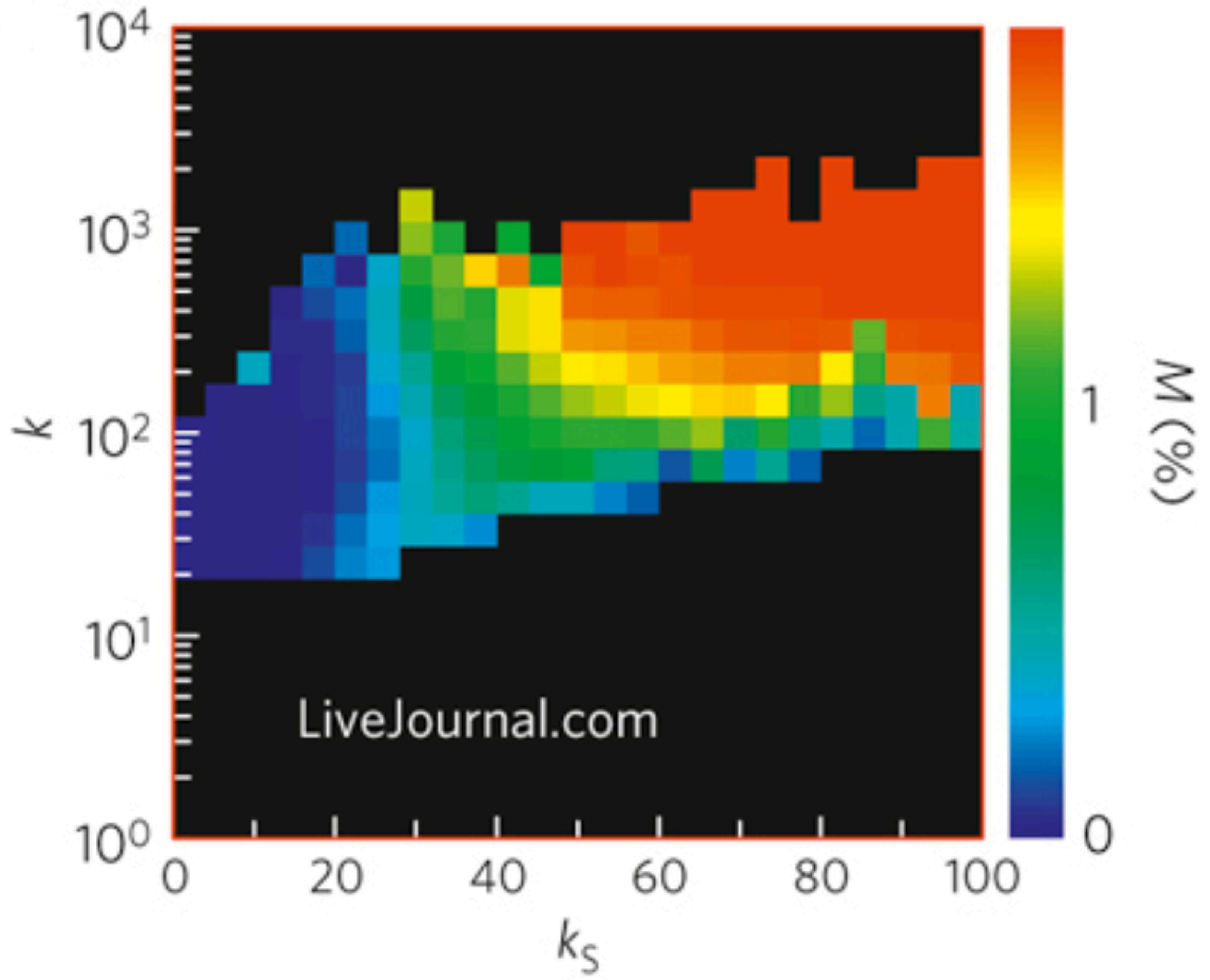
- Nodes are actors of adult movies in IMDB
- Edges connect actors who appear in the same movie (disease spread risk)



Degree vs Coreness

Example 3: Livejournal network for blogs

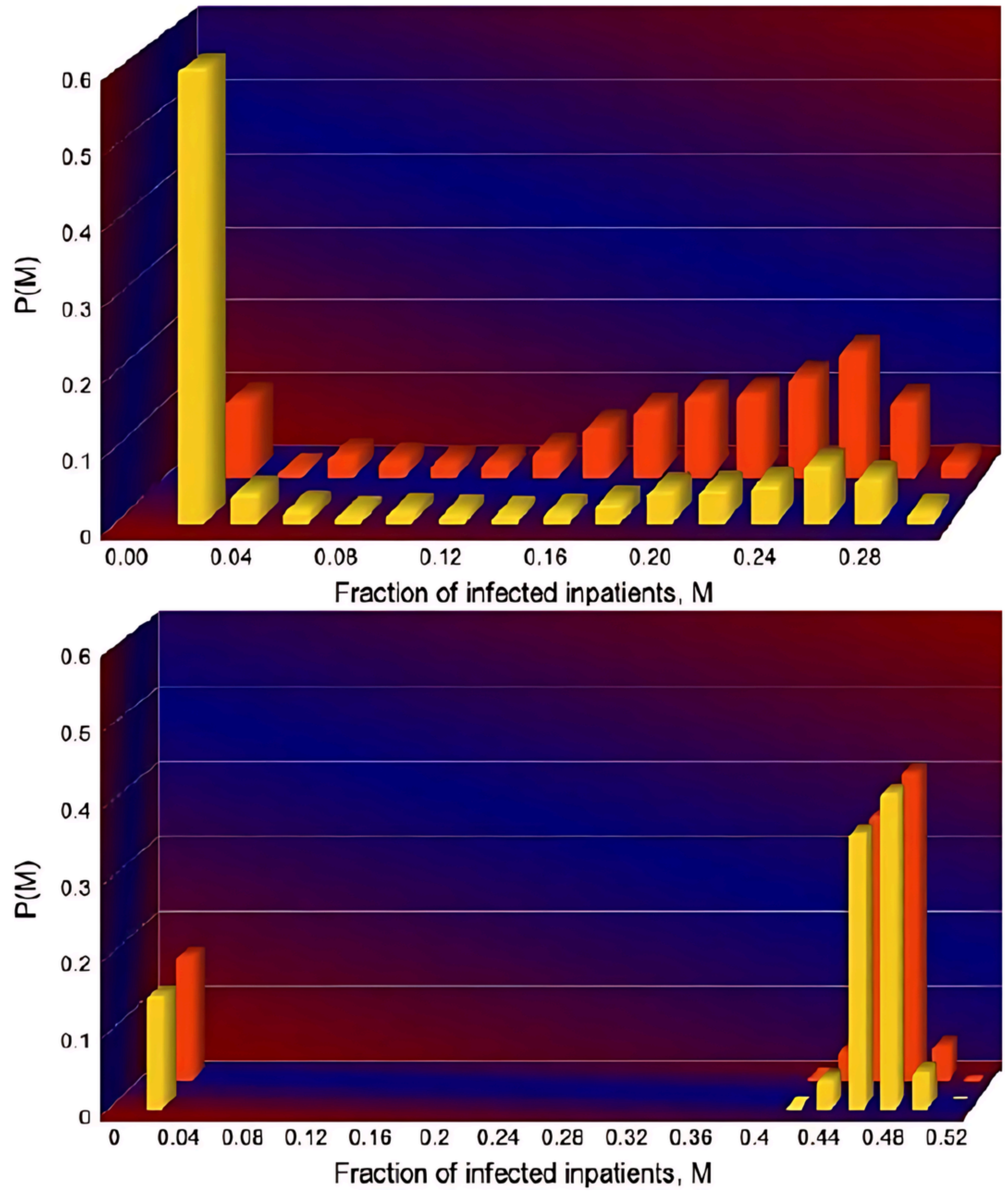
- Nodes are users of Livejournal
- Edges exist between blogs that have links to each other
- Cascades are information or behavior cascades



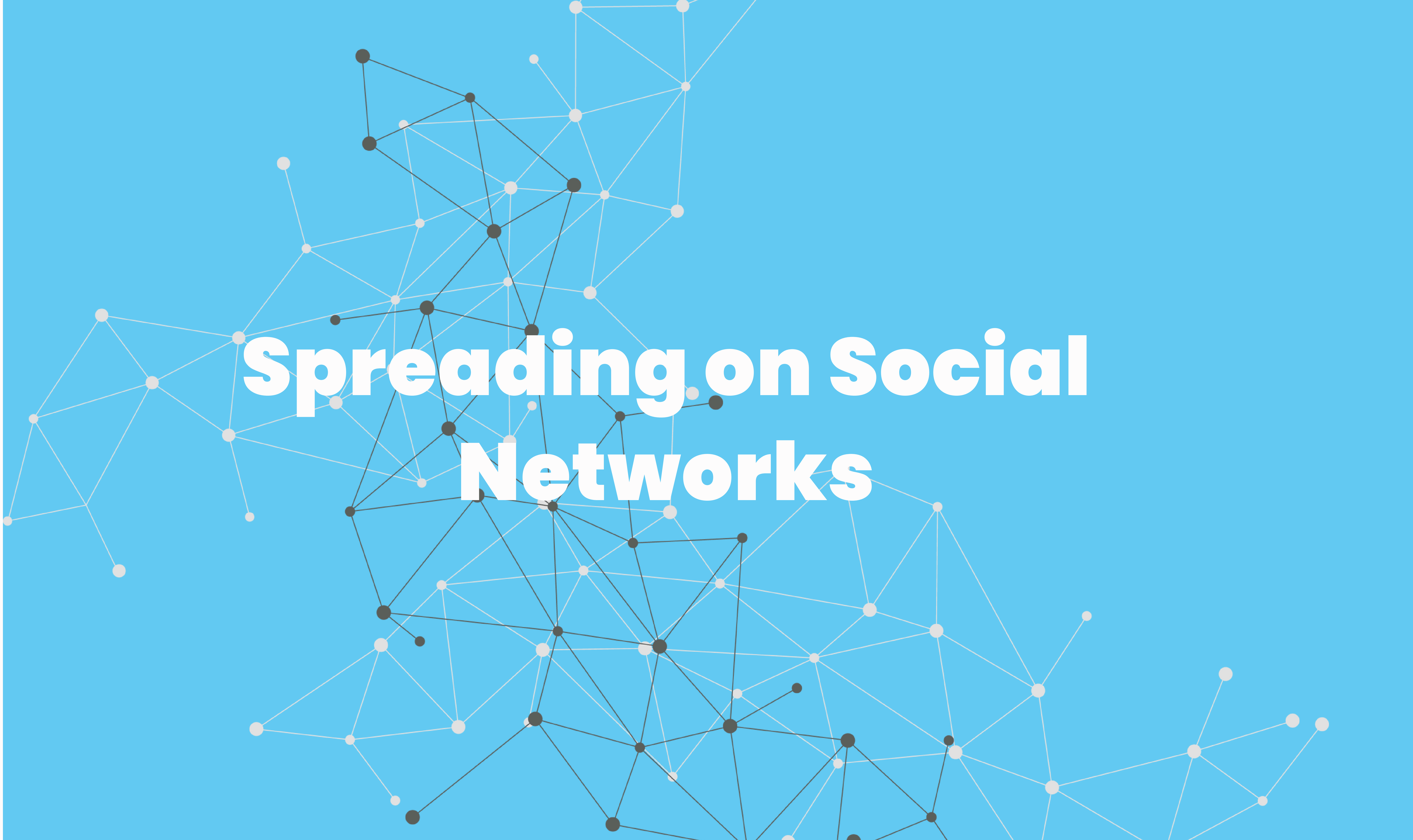
Role of Network Topology

We can look at the epidemic size distribution of two nodes with the same degree (k=96) but different coreness-centrality (k_s=63 orange, k_s=26 yellow)

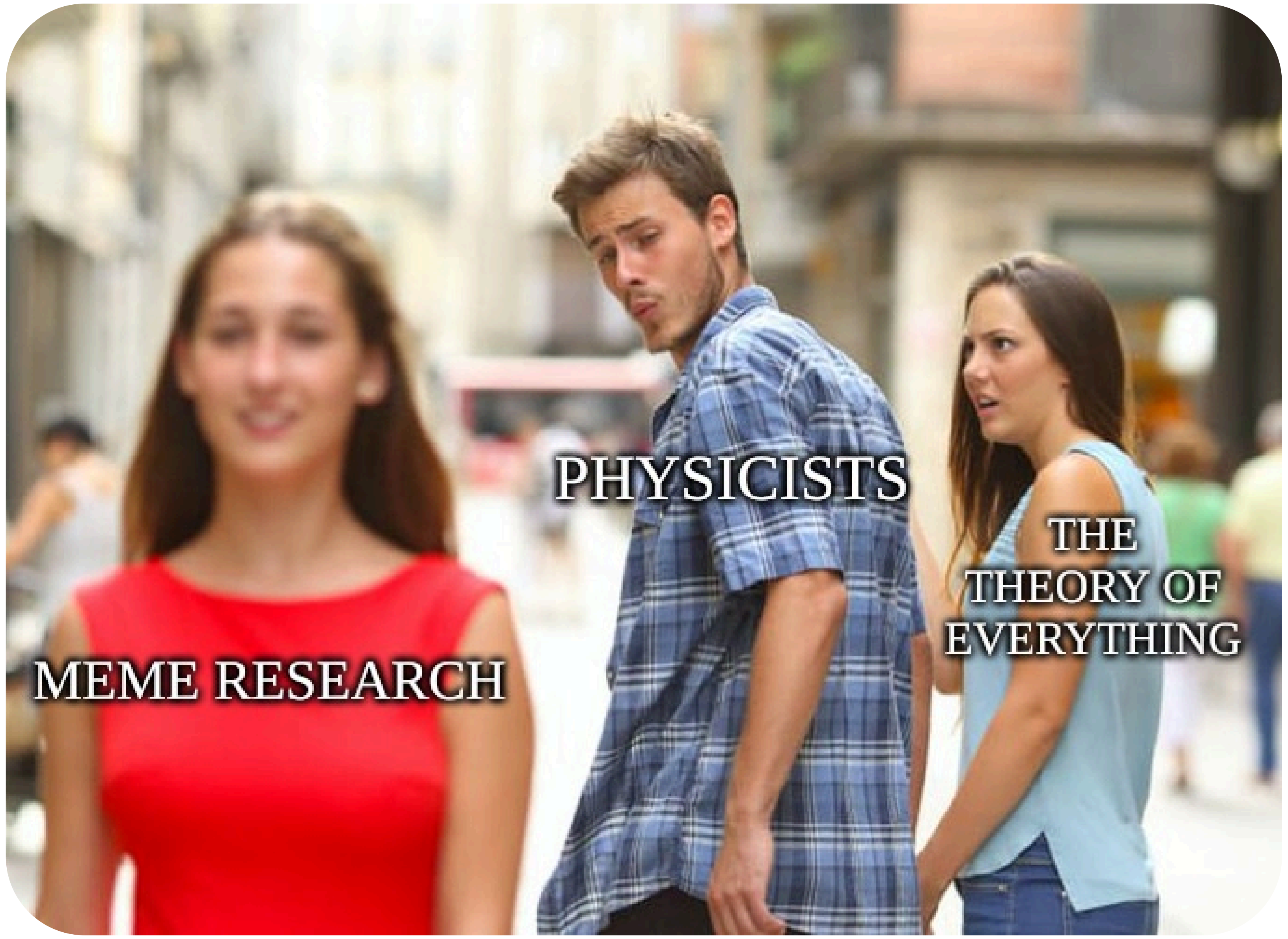
- on a real network the node with high coreness-centrality create much larger epidemics
- in a random network (lower plot) with the same degree sequence, the two nodes show the same behavior
- the distribution of epidemics is bimodal



Spreading on Social Networks



Memes or Microbes?



The spreading of content online resembles an epidemic process

- we often talk of viral content (shorts, memes, songs)
- memes spread very rapidly, sometimes they die, other times they keep circulating like a disease
- can we model the spreading of content online using epidemic spreading?

Fractional SIR Model

In a standard epidemic process the larger the number of infected contacts, the large the probability of getting infected

- on online platform the situation is different
- we have limited attention, thus we ignore most of the content
- the more friends we have, the harder it is for any of them to “infect” us with a meme

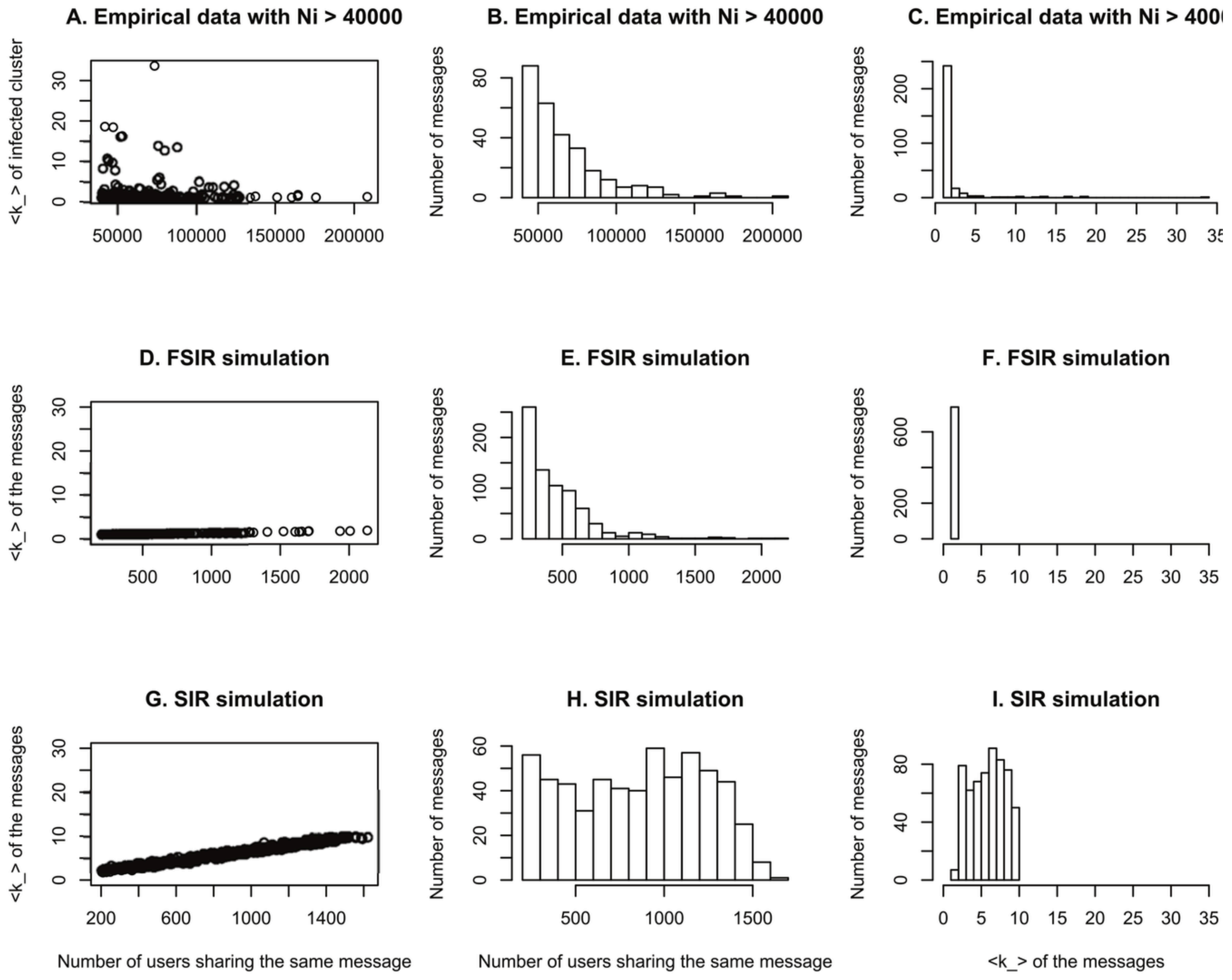
This property is described by the Fractional SIR (FSIR) model

- it is conceptually similar to the SIR model
- individuals recover after a time τ
- instead of the infection rate γ we use γ/k_u
- when k_u is large, an individual is infected only if many of its contacts are infected

Testing the **Model**

Authors compared the results of epidemic cascades in real data (Weibo social network) with those obtained using the SIR model and the FSIR model.

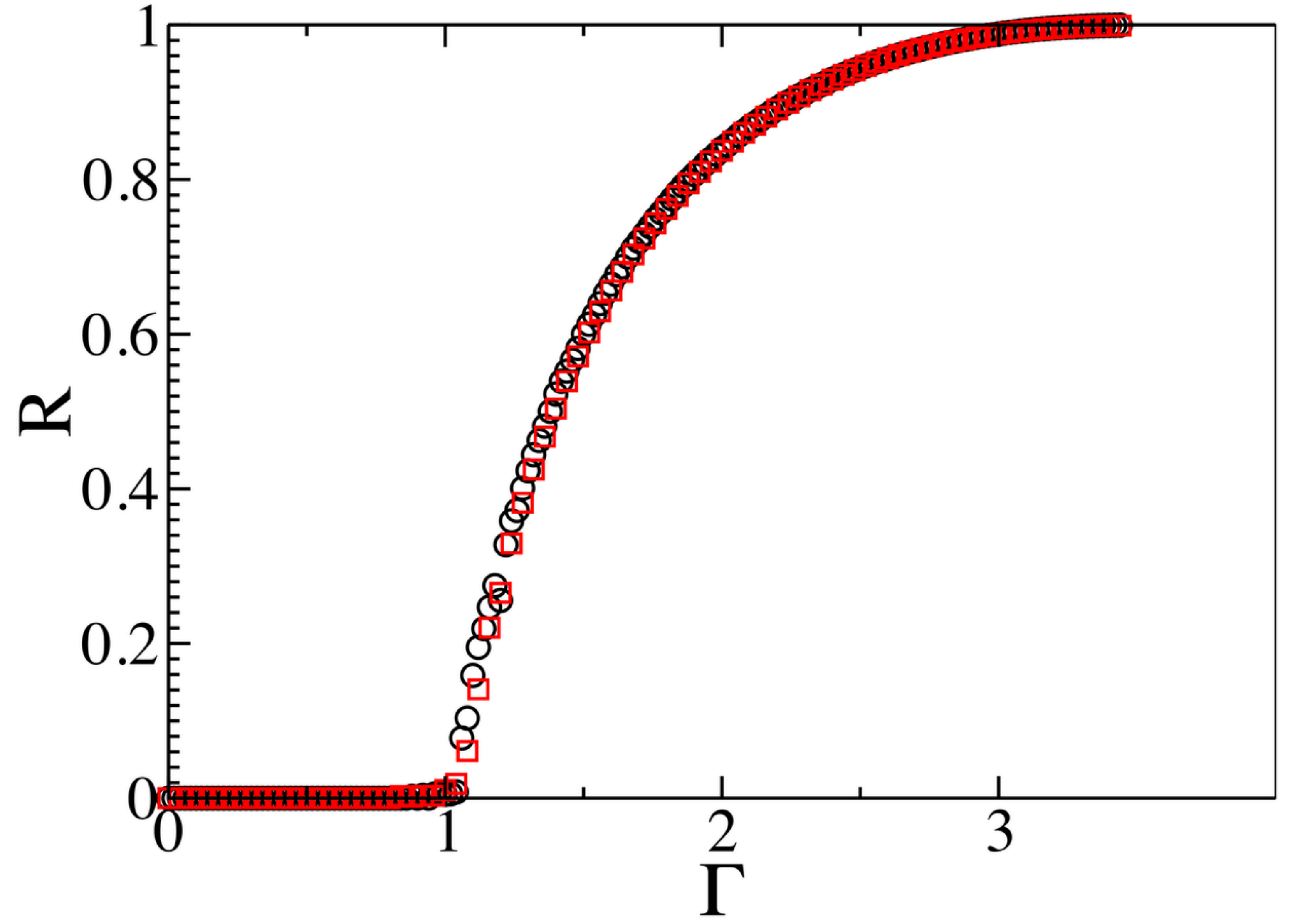
- The FSIR model much better describes the data
- The SIR model create cascades that are larger on average, but with much less viral content



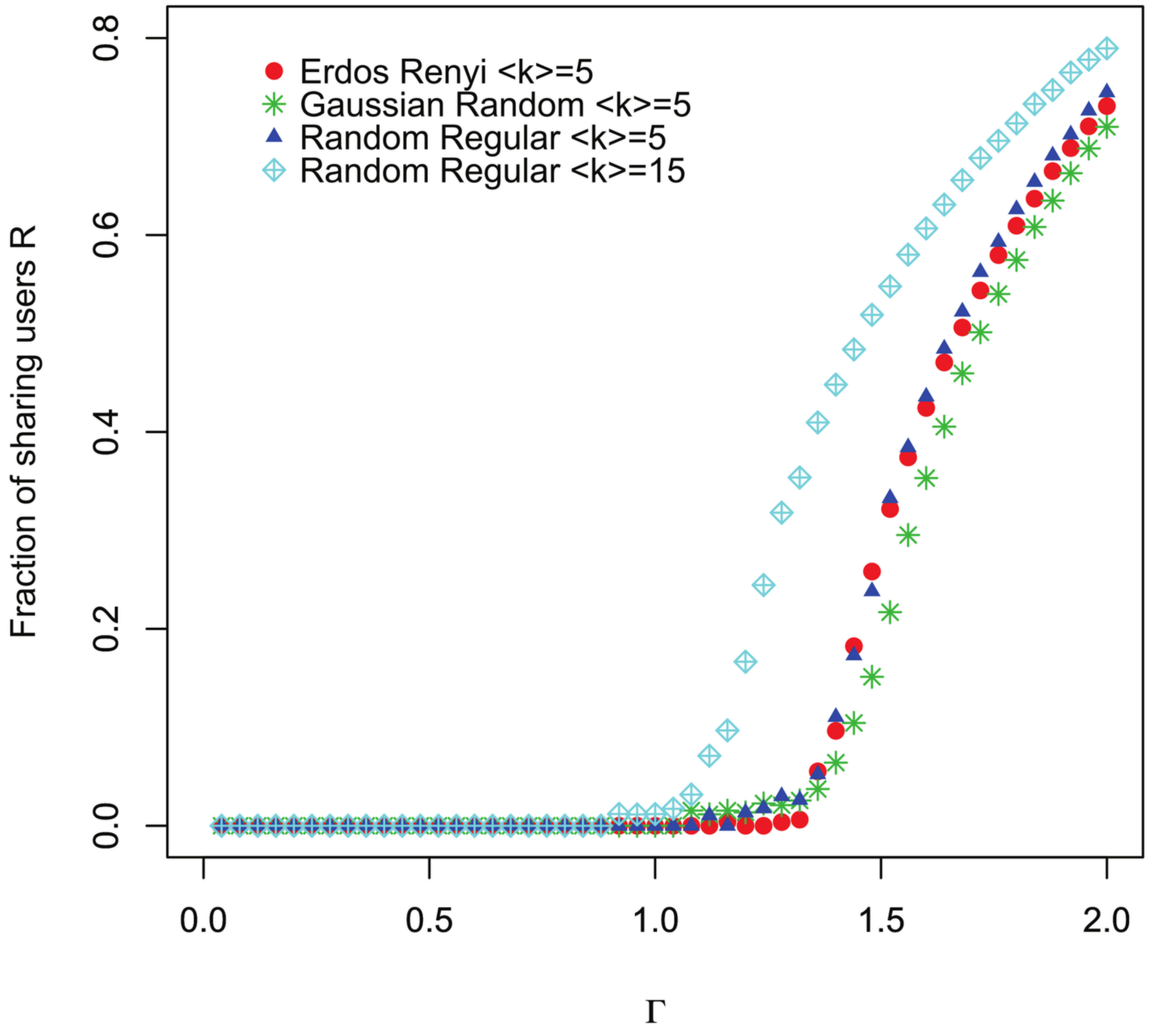
Phase Transition in the FSIR Model

The FSIR model shows a continuous phase transition in the variable $\Gamma = \gamma\tau$

- for small values of Γ there are no viral messages
- the ratio R of infected individuals is null
- for $\Gamma = 1$ there is a phase transition and viral content appears
- when Γ is large some messages spread in the whole network



The Role of Topology



It is possible to compute the expression of the critical point analytically

$$\Gamma_c = \frac{\langle k \rangle}{\langle k \rangle - 1}$$

The critical point only depends on the average degree, not on the network topology. For large values of the degree, it tends to one as we already saw.

Conclusions

Resilience in Social Networks

Some social networks collapse despite being very popular, network science can help understand the reasons behind this.

K-Coreness and Social Networks Collapse

The K-Coreness decomposition of a network allows to quantify its resilience and its ability to recover from mass leave from the platform

Epidemic Spreading

We can model spreading of diseases on networks using the SIR and SIS model. The coreness centrality works better in identifying super spreaders

Spreading on Social Networks

The spreading of content online resembles epidemic spreading. However we have to consider the presence of limited attention.