# Network Science

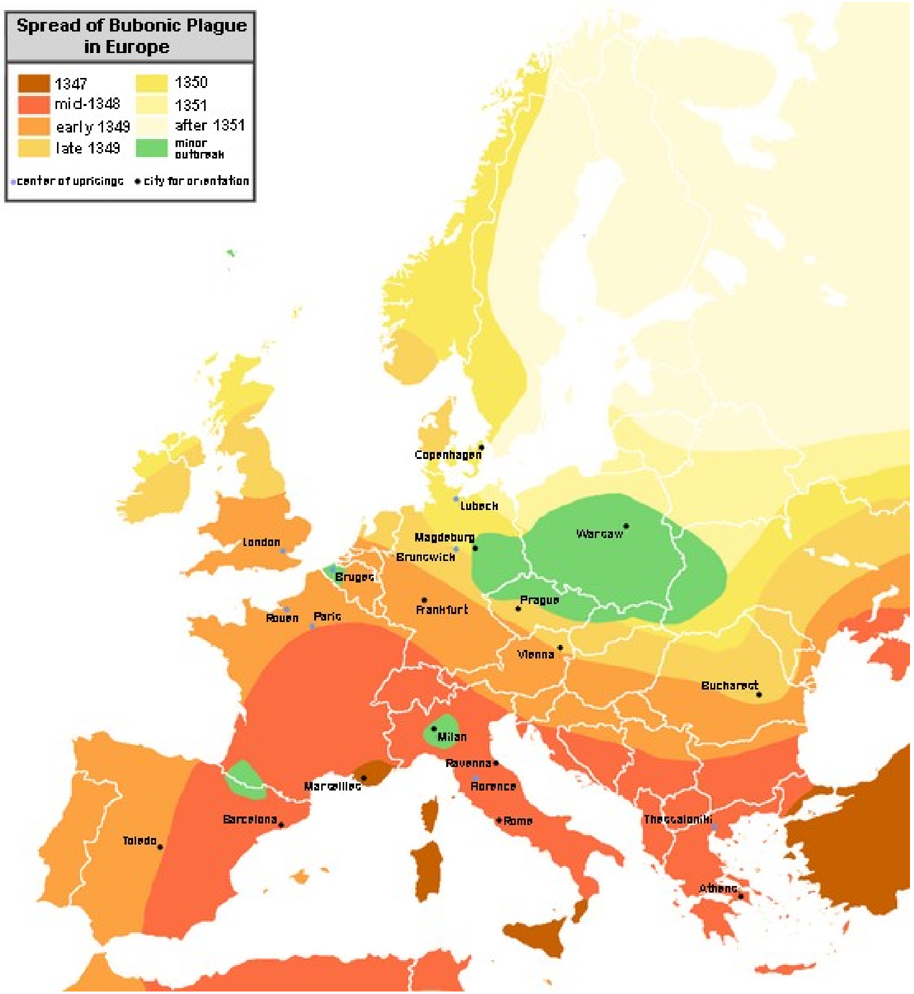
# Spreading Phenomena


Albert-László Barabási
with
Emma K. Towlson and Sean P. Cornelius


www.BarabasiLab.com

# Network Epidemics

# 14th Century – The Great Plague




4 years from France to Sweden

Limited by the speed of human travel

http://en.wikipedia.org/wiki/Black_Death
http://de.wikipedia.org/wiki/Schwarzer_Tod

# Section 10.3 Network Epidemics

The approaches described above have not considered explicitly that the spreading takes place on a network– they 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.

# SI model on a network: Degree Block Approximation


Split nodes by their degrees

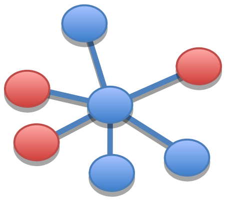
$$i_k = \frac{I_k}{N_k}, \quad i = \sum_k P(k) i_k$$

SI model:

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

Proportional to *k*

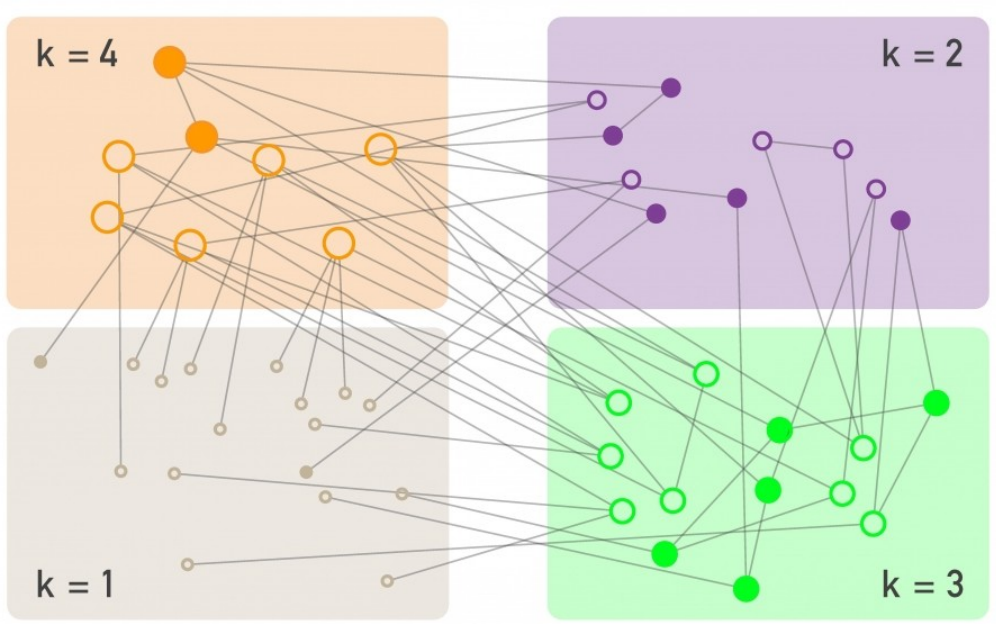
Density of infected neighbors of nodes with degree *k*



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

(Vespignani)

# SIS model on a network: Degree Block Approximation




Split nodes by their degrees

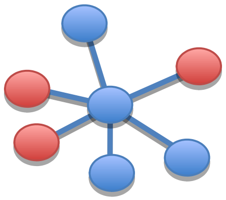
$$i_k = \frac{I_k}{N_k}, \quad i = \sum_k P(k) i_k$$

SIS model:

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

Proportional to *k*

Density of infected neighbors of nodes with degree *k*



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

# Early time behavior

# Early time behavior of an epidemic

*Why do we care about the early behavior of an epidemic?*

- vaccines, cures, and other medical interventions take months to years to develop

- the best way to stop or slow down an epidemic:
  → early quarantine
  → early vaccination

- SI model is the most relevant for early stages

# Network Epidemics: Early Time Behavior

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

$$\frac{di_k}{dt} \approx \beta k \Theta_k(t).$$

# Network Epidemics: Density Function

The density function $\Theta_k(t)$ provides the fraction of infected nodes in the vicinity of a node with degree $k$.

the probability that a link points from a node with degree $k$ to a node with degree $k'$ is independent of $k$. Hence the probability that a randomly chosen link points to a node with degree $k'$ is the excess degree (7.3),

$$\frac{k' p_{k'}}{\sum_k k' p_{k'}} = \frac{k' p_{k'}}{\langle k \rangle}.$$

At least one link of each infected node is connected to another infected node, the one that transmitted the infection. Therefore the number of links available for future transmission is $(k'-1)$, allowing us to write

$$\Theta_k(t) = \frac{\sum_{k'} (k'-1) p_{k'} i_{k'}(t)}{\langle k \rangle} \equiv \Theta(t). \qquad (10.34)$$

In other words, in the absence of degree correlations $\Theta_k(t)$ is independent of $k$. Differentiating (10.34) we obtain

$$\frac{d\Theta(t)}{dt} = \sum_k \frac{(k-1) p_k}{\langle k \rangle} \frac{d i_k(t)}{dt}. \qquad (10.35)$$

# Network Epidemics: Density Function for SI Model

The density function $\Theta_k(t)$ provides the fraction of infected nodes in the vicinity of a node with degree $k$.

$$\frac{d\Theta(t)}{dt} = \sum_k \frac{(k-1)p_k}{\langle k \rangle} \frac{di_k(t)}{dt}.$$

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

$$\frac{d\Theta(t)}{dt} = \beta \sum_k \frac{(k^2-k)p_k}{\langle k \rangle}[1-i_k(t)]\Theta(t). \qquad (10.36)$$

To predict the early behavior of the epidemics, we consider the fact that for small $t$ the fraction of infected individuals is much smaller than one. Therefore we can neglect the second order terms in (10.36), obtaining

$$\frac{d\Theta(t)}{dt} = \beta\left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1\right)\Theta(t). \qquad (10.37)$$

This has the solution

$$\Theta(t) = Ce^{t/\tau}, \qquad (10.38)$$

where

$$\tau = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}. \qquad (10.39)$$

Using the initial conditions $\Theta(t=0) = C = i_0 \frac{\langle k \rangle - 1}{\langle k \rangle}$

we obtain the time dependent $\Theta(t)$ as

$$\Theta(t) = i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} e^{t/\tau}. \qquad (10.40)$$

# Network Epidemics: SI Model – Early Time Behavior

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

$$\frac{di_k}{dt} \approx \beta k\Theta_k(t).$$

$$\frac{di_k}{dt} \approx \beta k i_0 \frac{\langle k \rangle - 1}{\langle k \rangle} e^{t/\tau},$$

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

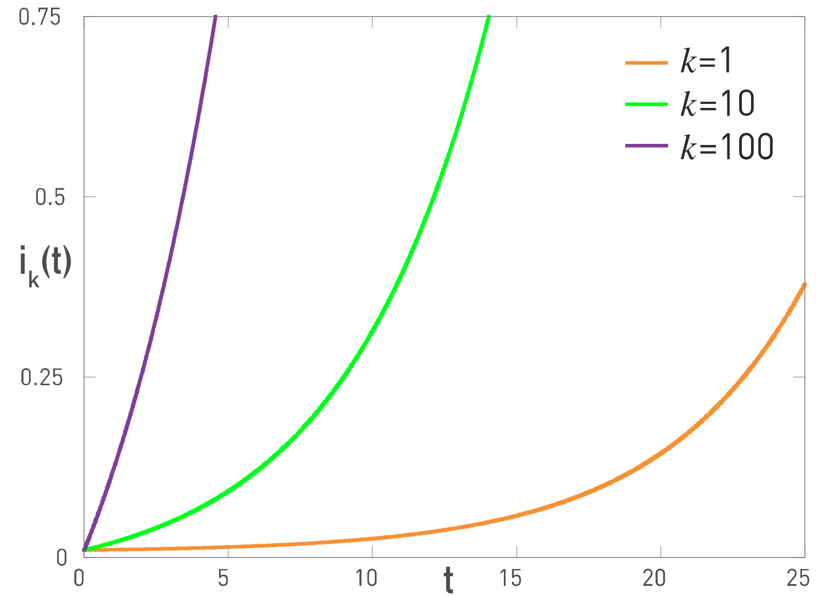
$$i_k = i_0\left(1 + \frac{k\langle k \rangle - 1}{\langle k^2 \rangle - \langle k \rangle}(e^{t/\tau} - 1)\right).$$

# Network Epidemics: SI Model – Early Time Behavior

$$i_k = i_0\left(1 + \frac{k\langle k\rangle - 1}{\langle k^2\rangle - \langle k\rangle}(e^{t/\tau} - 1)\right).$$

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

$$i_k = g(t) + kf(t),$$




Equation (10.17) predicts that the a pathogen spreads with different speed on nodes with different degrees. To be specific, we can write $i_k = g(t) + kf(t)$, indicating that at any time the fraction of high degree nodes that are infected is higher than the fraction of low degree nodes. The figure shows the fraction of infected nodes with degrees $k$=1, 10 and 100 in an Erdős-Rényi network with average degree $\langle k\rangle$=2. We find that at $t$=5 less than 1% of the $k$=1 degree nodes are infected, but close to 10% of the $k$=10 nodes and over 75% of the $k$=100 nodes.

**The higher the degree of a node, the more likely that it becomes infected.**

# Network Epidemics: SI Model – Random vs Scale-Free Network

Random Network: $\tau = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}$. $\tau_{ER} = \frac{1}{\beta\langle k \rangle}$,

Scale-Free Networks

- **Scale-free Network with γ ≥3** — Similar to random, with altered $\tau$
- **Scale-free Networks with γ ≤3** — For $N \to \infty$, we have $\langle k^2 \rangle \to \infty$ and $\tau \to 0$
  Instantaneous spread! The hubs get the disease fast, and distribute it to all.
- **Inhomogenous Networks** — For $N \to \infty$, as long as $\langle k^2 \rangle > \langle k \rangle(\langle k \rangle - 1)$, we have a reduced $\tau$ and higher speed.

# Network Epidemics: SIS model and the Vanishing Epidemic Threshold

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

$$\tau = \frac{\langle k \rangle}{\beta\langle k^2 \rangle - \langle k \rangle \mu}.$$

There is a small but important difference in the density function of the SIS model. For the SI and the SIR models, if a node is infected, then at least one of its neighbors must also be infected or recovered, hence at most ($k$–1) of its neighbors are susceptible, the origin of the (-1) term in the paranthesis of (10.34) . However, in the SIS model the previously infected neighbor can become susceptible again, therefore all $k$ links of a node can be available to spread the disease. Hence we modify the definition (10.34) to obtain

$$\Theta_k^{SIS}(t) = \frac{\sum_{k'} k' p_{k'} i_{k'}}{\langle k \rangle} \equiv \Theta_{SIS}(t). \tag{10.49}$$

Again keeping only the first order terms we obtain

$$\frac{di_k}{dt} = \beta k \Theta_{SIS}(t) - \mu i_k. \tag{10.50}$$

Multiplying the equation with $(k–1)p_k/\langle k \rangle$ and summing over $k$ we have

$$\frac{d\Theta(t)}{dt} = \left(\beta \frac{\langle k^2 \rangle}{\langle k \rangle} - \mu\right)\Theta(t). \tag{10.51}$$

This again has the solution

$$\Theta(t) = Ce^{t/\tau}, \tag{10.52}$$

# Network Epidemics: SIS model and the Vanishing Epidemic Threshold

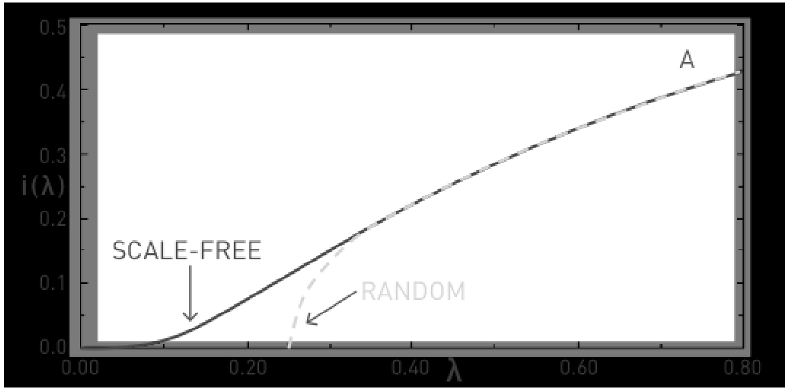
$$\frac{di_k}{dt} = \beta(1 - i_k)k\Theta_k(t) - \mu i_k \, . \quad \tau = \frac{\langle k \rangle}{\beta\langle k^2 \rangle - \langle k \rangle \mu} .$$

A global outbreak is possible if $\tau>0$, which yields the critical spreading rate

$$\lambda \equiv \frac{\beta}{\mu} > \frac{\langle k \rangle}{\langle k^2 \rangle}, \qquad (10.54)$$

and the epidemic threshold for the SIS model as (**Table 10.3**)

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} . \qquad (10.55)$$



# SIS Model – Scale-free Exponents

**Degree distribution**

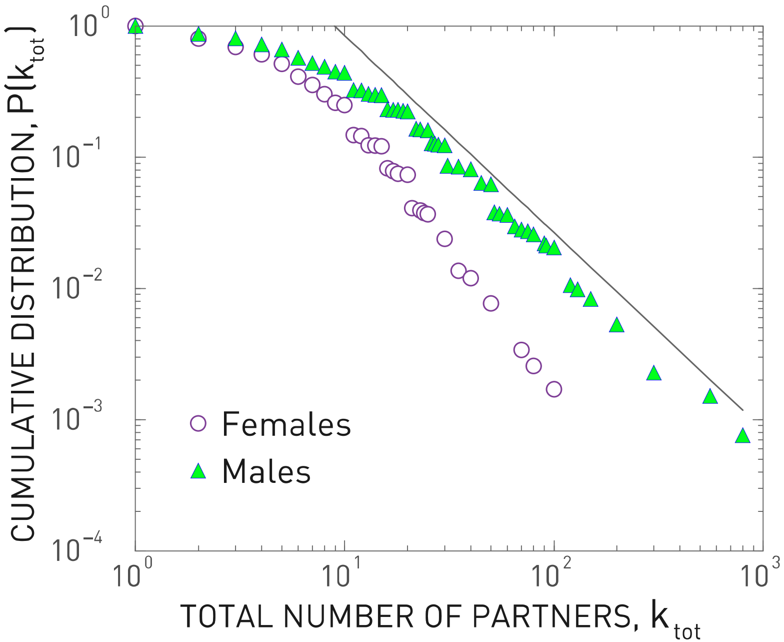
$$P(k) = (1+\gamma)m^{1+\gamma}k^{-2-\gamma}$$

| | | |
|---|---|---|
| $2 < \gamma < 3$ | $\lambda_c=0$<br>$\Theta(\lambda) \approx (m\lambda)^{(\gamma-2)/(3-\gamma)}$<br>$i(\lambda) \approx \lambda^{1/(3-\gamma)}$ | Non- zero prevalence for all β/μ. The exponent 1/(3−γ ), is larger than one → for small β/μ the prevalence is growing very slowly, i.e. there exists a wide region of spreading rates in which *i∞ ≪ 1.* |
| $\gamma = 3$ | $\lambda_c=0$<br>$\Theta(\lambda) \approx \frac{e^{-1/m\lambda}}{\lambda m}(1-e^{-1/m\lambda})^{-1}$<br>$i(\lambda) \approx 2e^{-1/m\lambda}$ | Prevalence approaches zero in a continuous and smooth way, exhibiting an exponentially small value for a wide range of spreading rates (*i∞ ≪ 1).* |
| $3 < \gamma < 4$ | $\lambda_c>0$<br>$i(\lambda) \approx \left(\lambda - \frac{\gamma-3}{m(\gamma-2)}\right)^{1/(\gamma-3)}$ | |
| $4 < \gamma$ | $\lambda_c>0$<br>$i(\lambda) \approx \lambda - \frac{\gamma-3}{m(\gamma-2)}$ | Only in this regime reduces to the homogenous mixing model. |

# Network Epidemics: Summary

| Model | Continuum Equation | $\tau$ | $\lambda_c$ |
| --- | --- | --- | --- |
| SI | $\frac{di_k}{dt} = \beta\left[1 - i_k\right] k\theta_k$ | $\frac{\langle k\rangle}{\beta\left(\left\langle k^2\right\rangle - \langle k\rangle\right)}$ | 0 |
| SIS | $\frac{di_k}{dt} = \beta\left[1 - i_k\right] k\theta_k - \mu i_k$ | $\frac{\langle k\rangle}{\beta\left\langle k^2\right\rangle - \mu\langle k\rangle}$ | $\frac{\langle k\rangle}{\left\langle k^2\right\rangle}$ |
| SIR | $\frac{di_k}{dt} = \beta s_k\theta_k - \mu i_k$ <br> $s_k = 1 - i_l - r_k$ | $\frac{\langle k\rangle}{\beta\left\langle k^2\right\rangle - (\mu+\beta)\langle k\rangle}$ | $\frac{1}{\frac{\left\langle k^2\right\rangle}{\langle k\rangle} - 1}$ |

# Contact Networks




Through interviews and questionnaires, researchers collected information from 4,781 randomly chosen Swedes of ages 18 to 74. The participants were not asked to reveal the identity of their sexual partners, but only to estimate the number of sexual partners they had during their lifetime.

# Are Sexual Hubs Real?

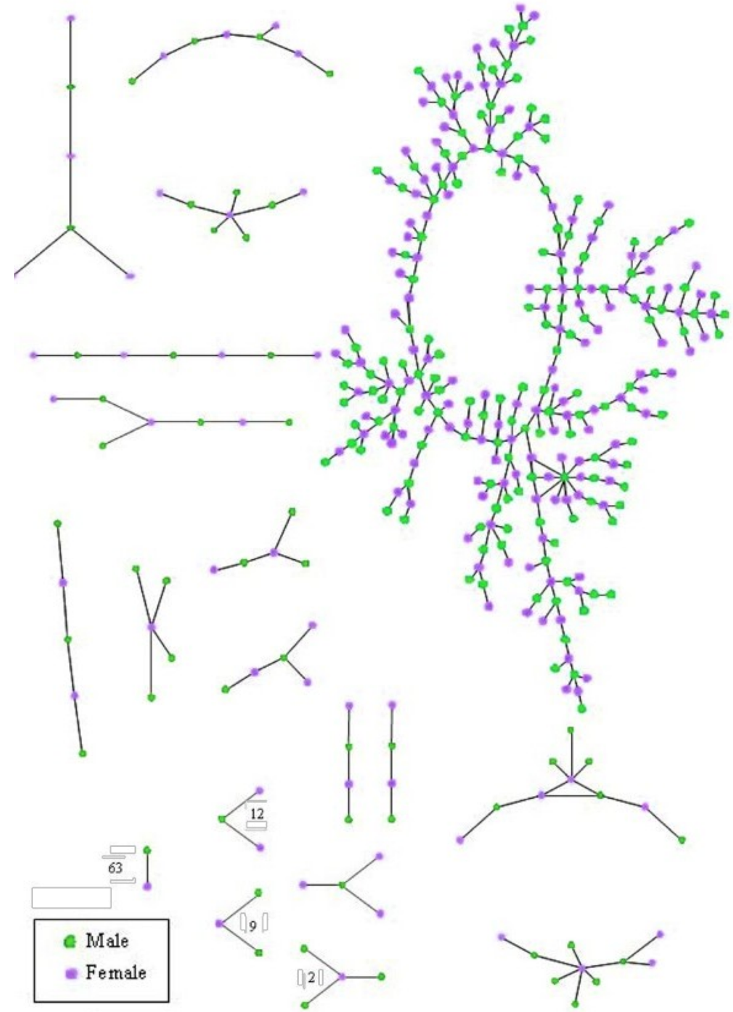
**Wild Chamberlain**, a Hall of Fame basketball player in the 1980s, who claimed havi

sex with a staggering number of 20,000 partners.

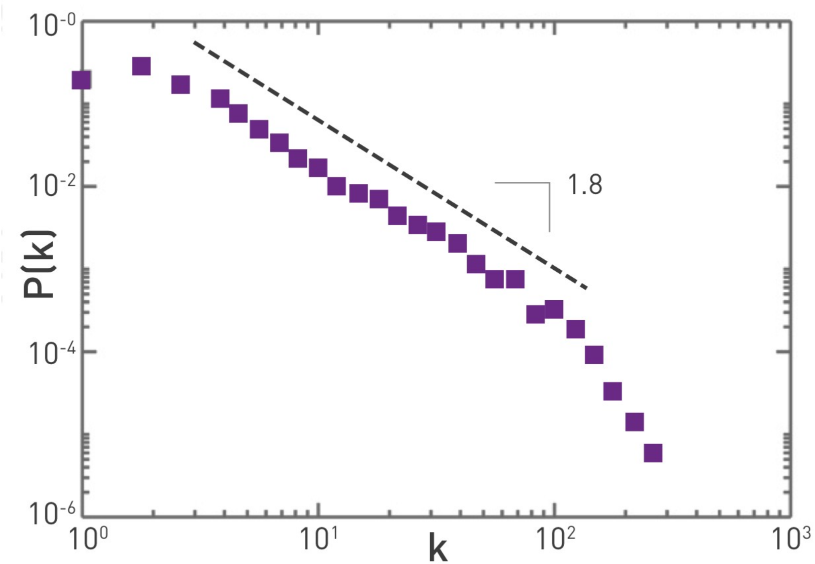
*"Yes, that's correct, twenty thousand different ladies,*

*"At my age, that equals to having sex with 1.2 woman a day, every day, since I was fifteen years old."*

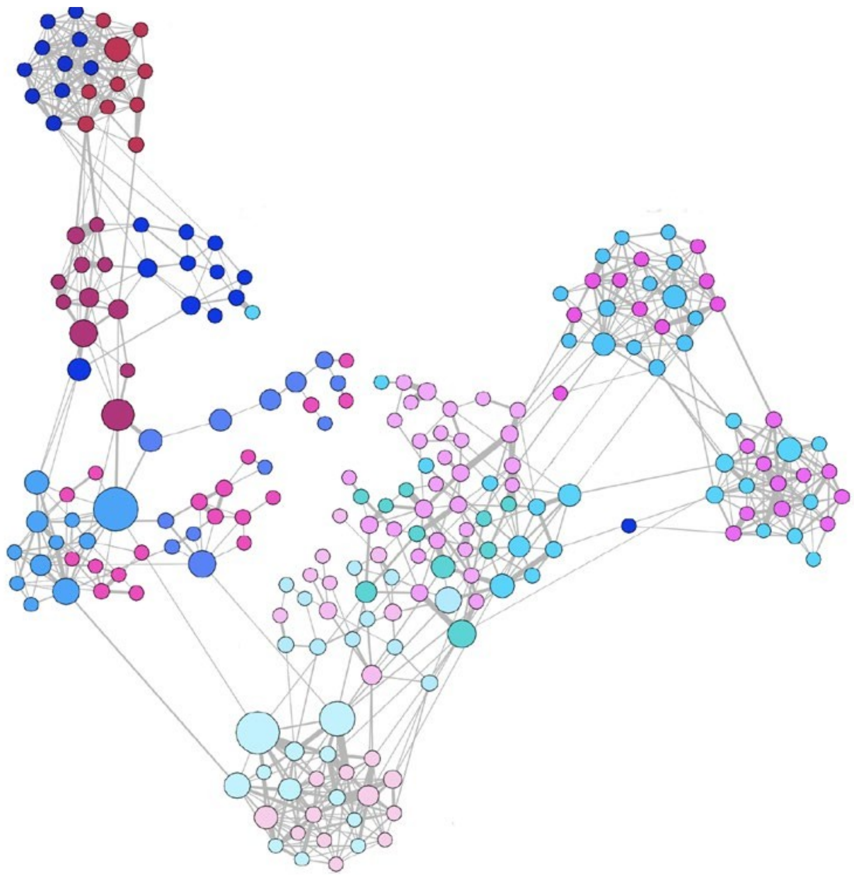
Within the AIDS literature the story of **Geetan Dugas**, a flight attendant with approximately 250 homosexual partners, is well documented.




Romantic and sexual links between high school students in a study focusing on over 800 adolescents living in a midwestern United States. Each circle represents a student and the links represent romantic relationships during the 6 months preceding the interview. The numbers indicate the frequency of each subgraph: for example 63 pairs of individuals (couples) are isolated from the rest of the network. After [20].



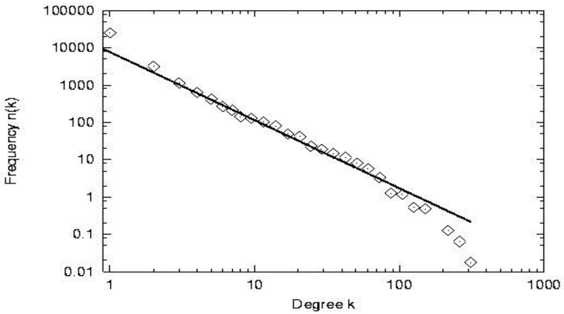

The degree distribution of the air transportation network is well approximated by a power-law for almost two decades with $\gamma$=1.8±0.2. The map was built using the International Air Transport Association database that contains the world list of airport and the direct flights between them during 2002. The resulting worldwide air-transportation network is a weighted graph containing the N=3,100 largest airports as nodes that are connected by L=17,182 direct flights as links, together accounting for 99% of the worldwide traffic. After [25].



A face-to-face contact network mapped out using RFA tags, capturing interactions between 232 students and 10 teachers across 10 classes in a school [31]. The structure of the obtained map depends on the context in which it is collected: the school network shown here reveals the presence of clear communities. In contrast, a study capturing the interactions between individuals that visited a museum reveal an almost linear network [29]. Finally, a network of attendees of a small conference is rather dense, as most participants interact with most others [29]. After [31].

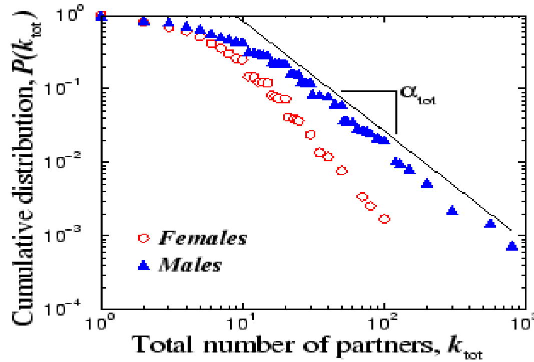
# SIS Model – Absence of Epidemic Threshold

Email network



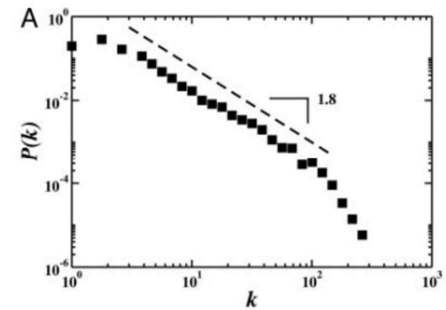

**Ebel et al. (2002)**

Human sexual network




**Liljeros et al., Nature (2001), Schneeberger et al. STD (2004)**

Air transportation network

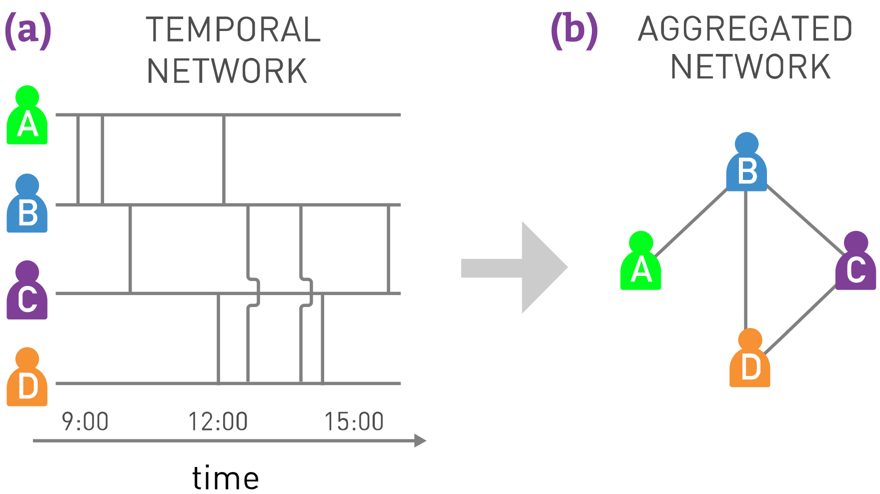



Colizza et al., PNAS 2006

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

Many networks will have vanishing epidemic threshold!

# Beyond the degree distribution

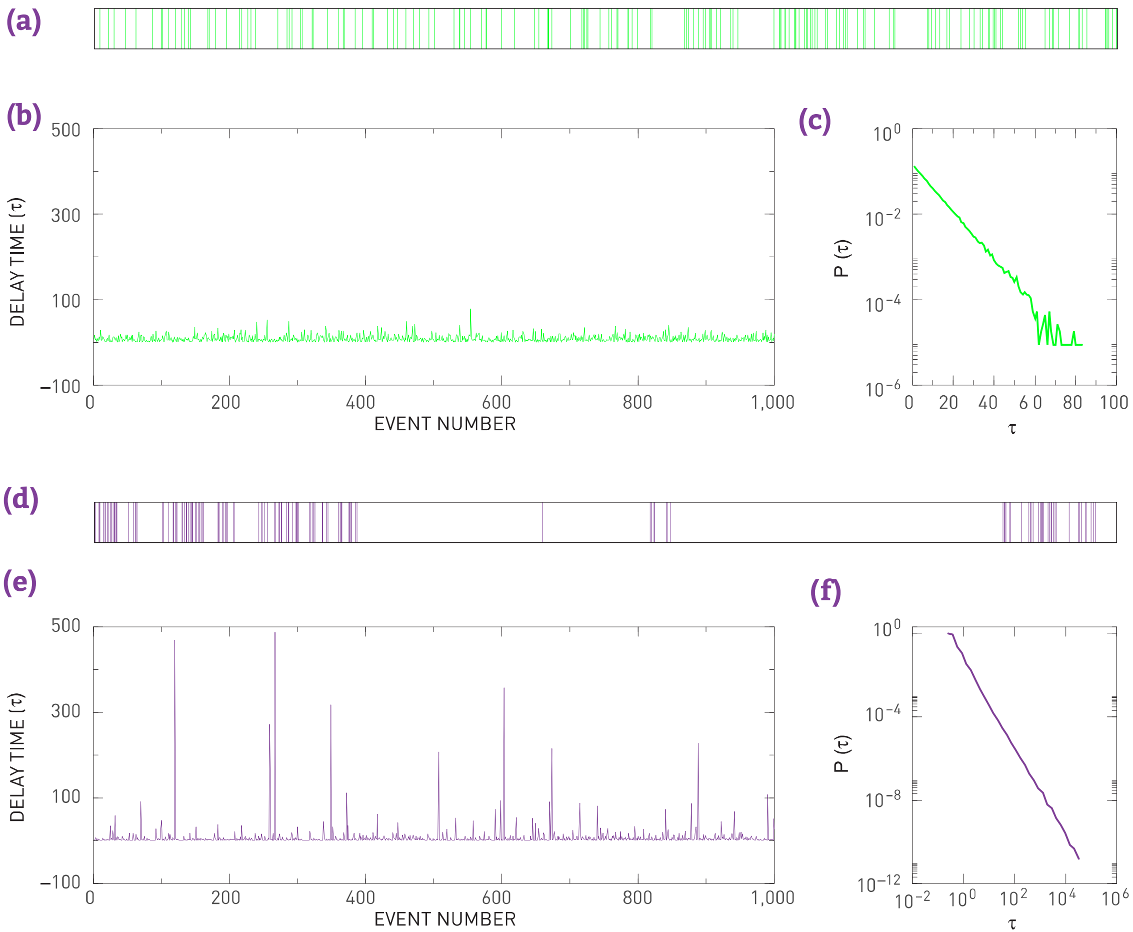



**(a) Temporal Network**

The timeline of the interactions between four individuals. Each vertical line marks the moment when two individuals come into contact with each other. If A is the first to be infected, the pathogen can spread from A to B and then to C, eventually reaching D. If, however, D is the first to be infected, the disease can reach C and B, but not A. This is because there is a temporal path from A to D.

**(b) Aggregated Network**

The network obtained by merging the temporal interactions shown in (a). If we only have access to this aggregated representation, the pathogen can reach all individuals, independent of its starting point. After [40].




For example, if the time between consecutive emails would follow a Poisson distribution, an email virus would follow

$i(t) \sim \exp(-t/\tau)$

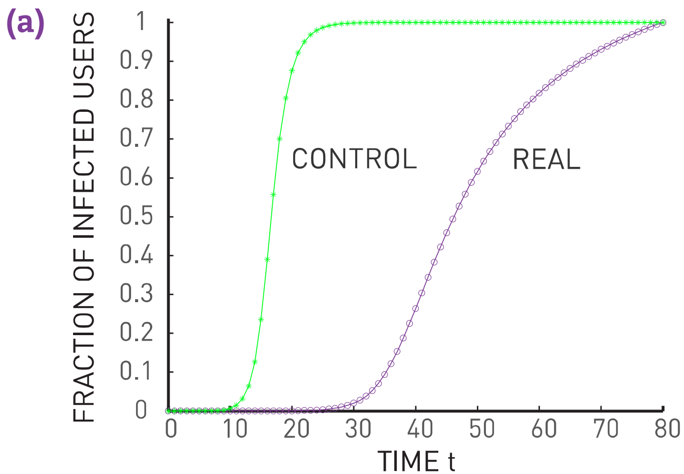
with a decay time of

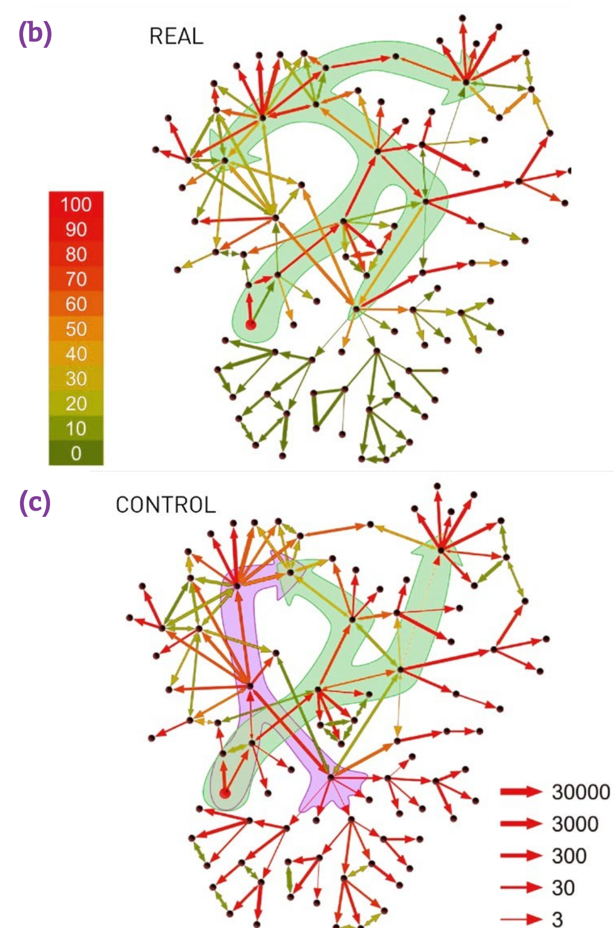
$\tau \simeq 1$ day.

In the real data, however, the decay time is

$\tau \simeq 21$ days,

a much slower process, correctly predicted by the theory if we use power law inter- event times.

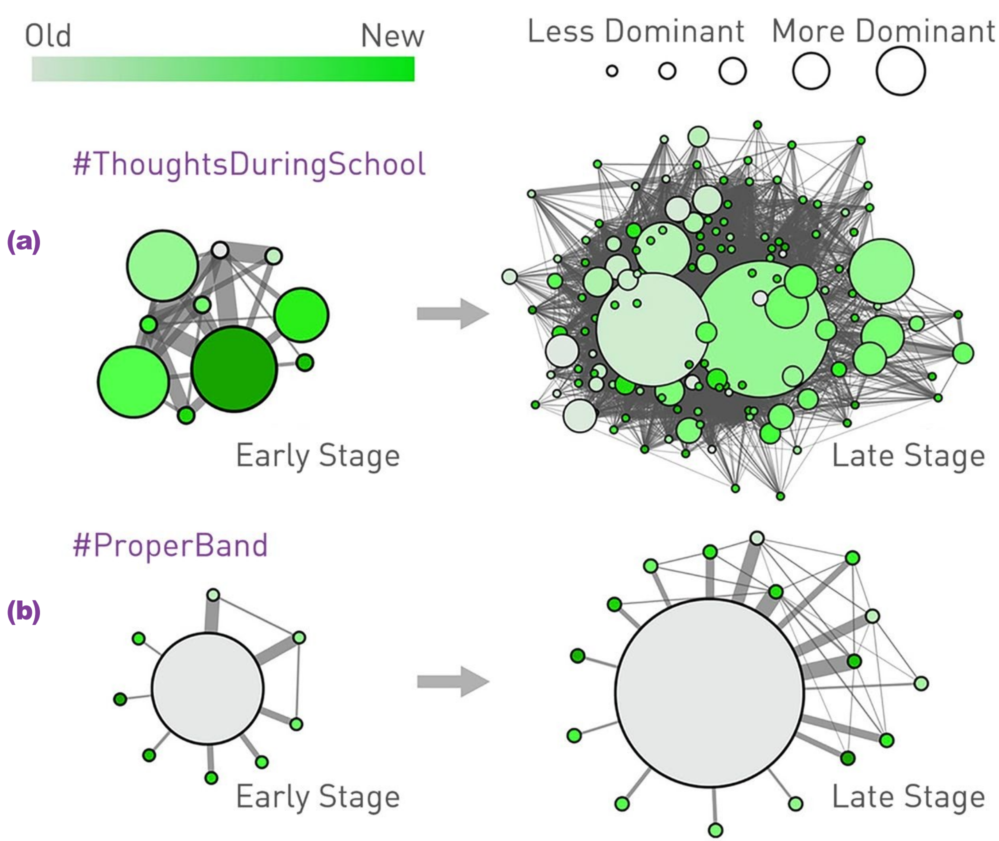



$$p_{ii}\sim\beta w_{ii},$$

# Section 10.5 Complex contagion

*Simple contagion* is the process we explored so far: It is sufficient to come into contact with an infected individual to be infected. The spread of behavior is often described by *complex contagion*, capturing the fact that most individuals do not adopt a new behavioral pattern at the first contact. Rather, adoption requires reinforcement [64], i.e. contact with *several* individuals who have already adopted. For example, the higher is the fraction of a person's friends that have a mobile phone, the more likely that she also buys one.




The community structure of the Twitter follower network. Each circle corresponds to a community and its size is proportional to the number of tweets produced by the respective community. The color of a community represents the time when the studied hashtag (meme) is first used in the community. Lighter colors denote the first communities to use a hashtag, darker colors denote the last community to adapt it.

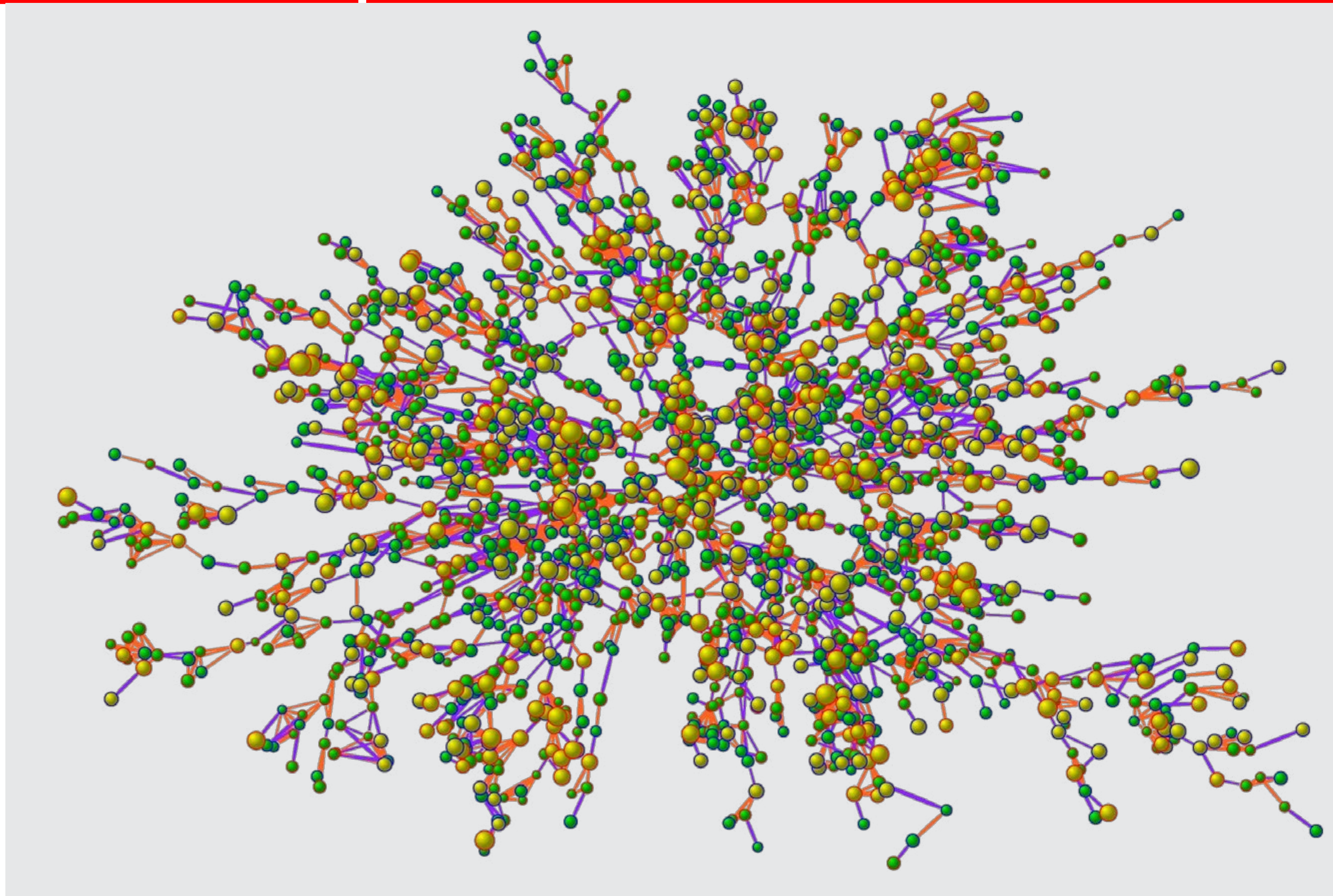
**(a) Simple Contagion**

The evolution of the viral meme captured by the #ThoughtsDuringSchool hashtag from its early stage (30 tweets, left) to the late stage (200 tweets, right). The meme jumps easily between communities, infecting many of them, following a contagion pattern characterizing pathogens as well.

**(b) Complex Contagion**

The evolution of a non-viral meme caputed by the #ProperBand hashtag from the early stage (left) to the final stage (65 tweets, right). The tweet is trapped in a few of communities, having difficulty to escape them. This is a signature of reinforcement, an indication that the meme follows complex contagion. After [54].

# Do our friends make us fat?



# Immunization Strategies

# How to control the epidemic?

- Transmission-reducing interventions: face masks, gloves, washing hands – may reduce the transmission rate below the epidemic-causing critical rate

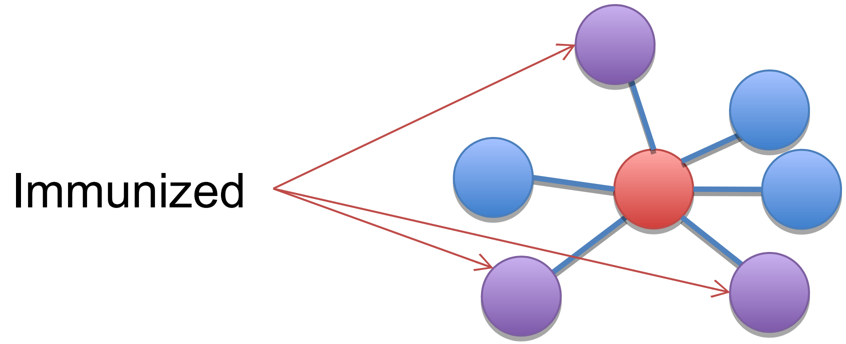
- Contact-reducing interventions: quarantining a patient, closing schools – make the network sparser, may increase the critical transmission rate

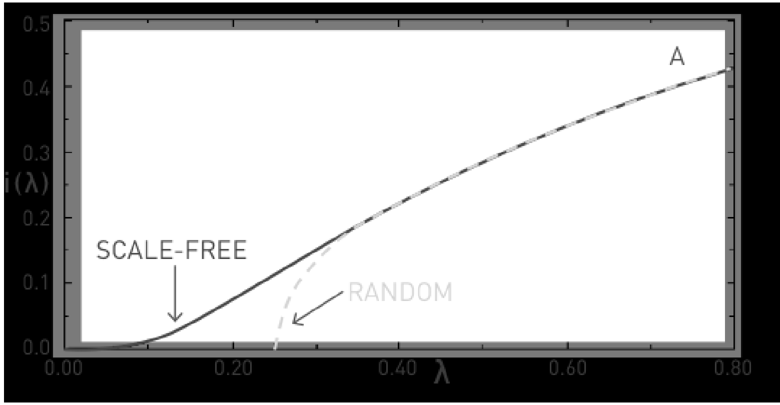
- Vaccinations: remove nodes from the network

- Q: Who should be vaccinated for most effective control?

# Immunization strategies– Random Networks

A fraction g of individuals are randomly chosen to be immunized.



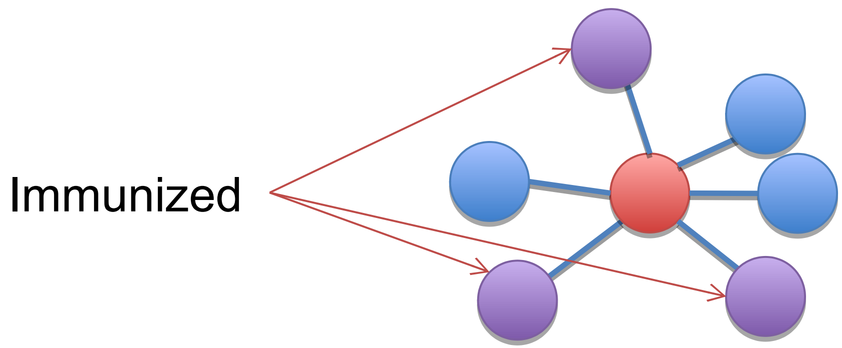

$$\beta \rightarrow \beta(1-g)$$

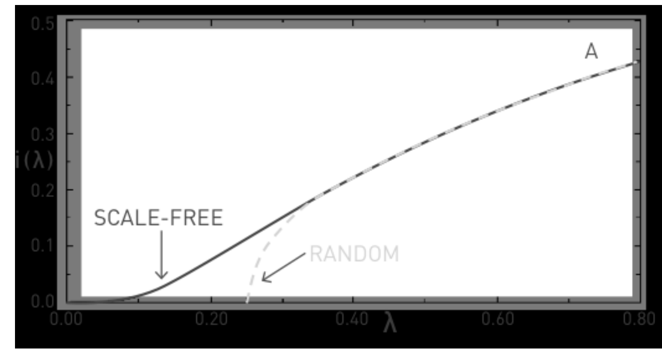
$$\frac{(1-g)\beta}{\mu} = \frac{1}{\langle k \rangle + 1},$$

Consequently, if vaccination increases the fraction of immunized individuals above $g_c$, it pushes the spreading rate under the epidemic threshold $\lambda_c$. In this case $\tau$ becomes negative and the pathogen dies out naturally. This explains why health official

# Immunization strategies

A fraction g of individuals are randomly chosen to be immunized.




$$\beta \rightarrow \beta(1-g)$$

$$\frac{\beta}{\mu}(1-g_c) = \frac{\langle k \rangle}{\langle k^2 \rangle}$$

$$g_c = 1 - \frac{\mu}{\beta}\frac{\langle k \rangle}{\langle k^2 \rangle}.$$

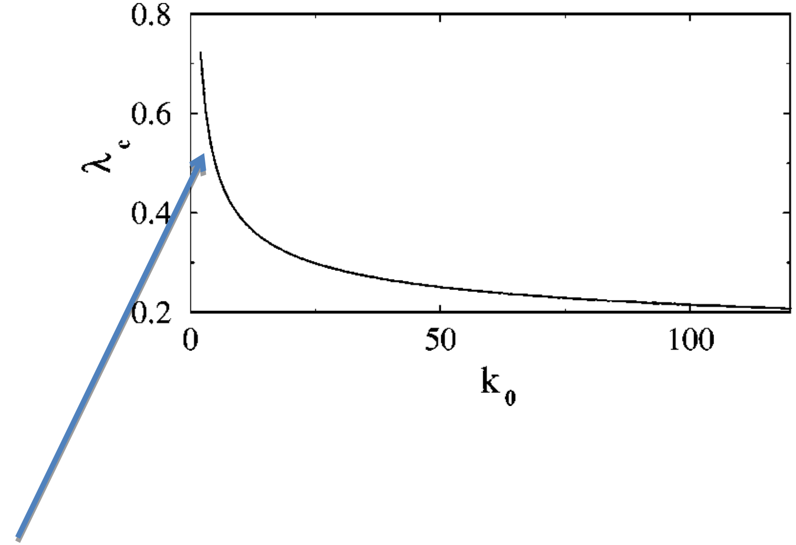
If $\langle k^2 \rangle \rightarrow \infty$, random immunization cannot prevent the outbreak.

# Vaccination strategies in scale-free networks

As hubs are responsible for the spread of the disease → cure the hubs.

Targeted immunization – immunize all nodes with degree $k>k_0$.

$$\lambda'_c \approx \frac{\gamma-2}{3-\gamma}\frac{k_{min}^{2-\gamma}}{(k'_{max})^{\gamma-3}}.$$



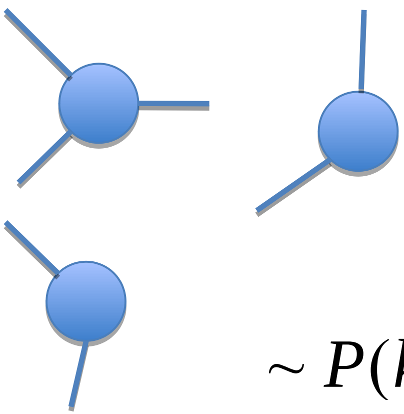

With increasing critical point increases, it is harder for a virus to spread.

Z. Dezso and A-L. Barabasi, Phys. Rev. E 65, 055103 (2002);
R. Pastor-Satorras and A. Vespignani, Phys. Rev. E 65, 036104 (2002)

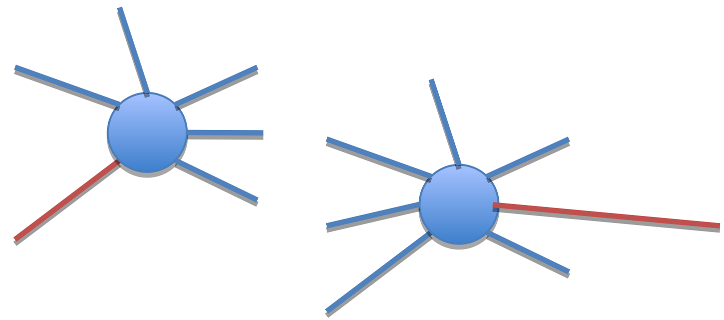
# Immunization strategies – without global knowledge

In many cases, you cannot figure out who are the hubs.

Can we effectively immunize the population when we don't know the node degrees?



$$\sim P(k)$$



$$\sim kP(k)$$

If you follow an edge, you are likely to meet high-degree nodes!

Select a random individual, then immunize **one of its RANDOMLY CHOSEN FRIENDS**.

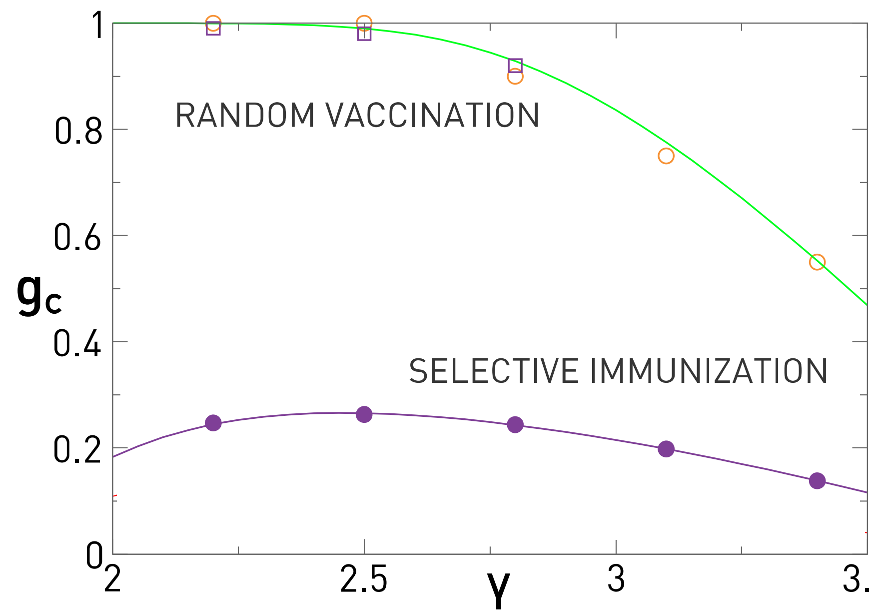
R. Cohen et al., PRL 91, 247901 (2003)

# A local method for vaccination effective in scale-free networks

Contact network described by scale-free random graph

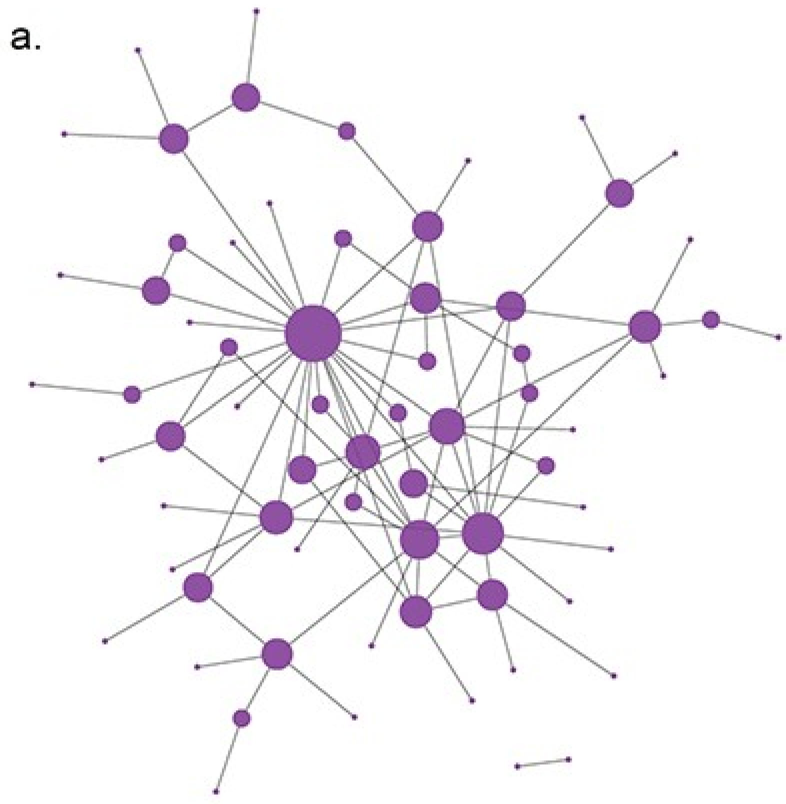
Immunization strategy: select a node randomly, immunize a randomly selected neighbor of it.

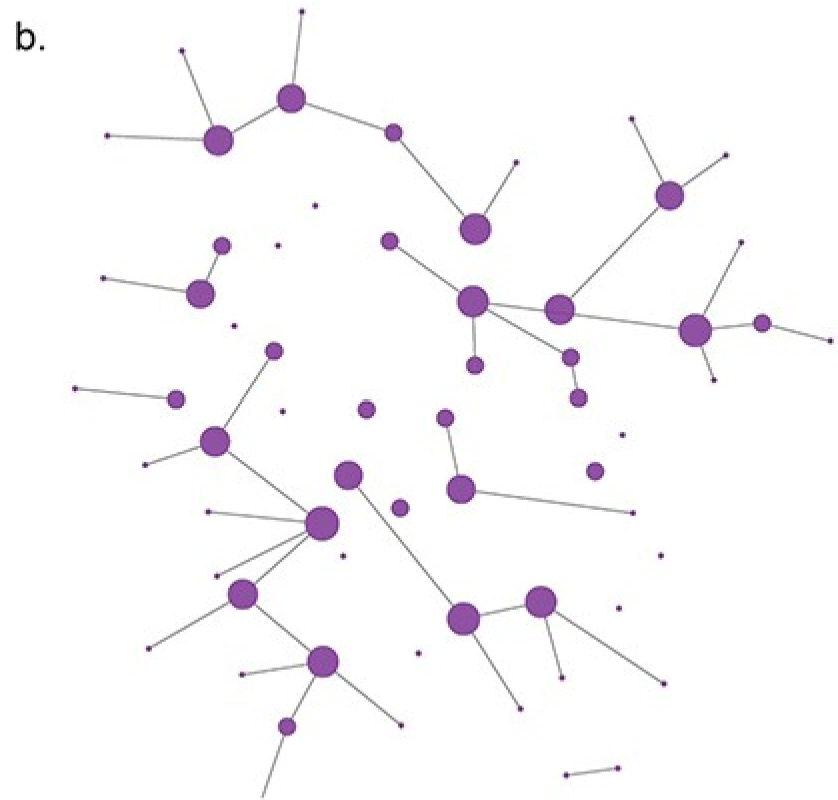
Use theory and simulations to determine the critical immunization fraction for each transmissibility value




R. Cohen et al, Phys. Rev. Lett 91, 247901 (2003)

## *Robustness and Immunisation*



# Robustness and Immunisation

*Eradication* is the complete elimination of a pathogen from the population.

Eradication campaigns had mixed success: smallpox and rinderpest were successfully eradicated, but programs targeting hookworm, malaria, and yellow fever have failed.

**Figure 10.23**
**Eradicating Smallpox**

Rahima Banu, the last smallpox infected patient in Bangladesh in 1976. After [68].