

Spreading processes on networks

In some cases is not enough to specify the nodes (components) and edges (interactions).

What are the **dynamical aspects** of the interaction?

What is the characteristic quantity changed by the interaction?

- Social relations - spreading of information - knows or not
- Internet - packet routing - travel times
- Molecular networks - chemical reactions – concentrations

The spread of epidemic disease

- Many diseases spread through populations by contact between infective and susceptible individuals
- The pattern of the disease-causing contacts forms a network
- Questions asked:
 - what is the typical size of an outbreak?
 - will an epidemic occur? – outbreak affects a finite fraction of the population
 - what determines the probability of an epidemic?

The fully mixed model

- Assume that an infective individual is equally likely to spread the disease to any other member of the population or subpopulation to which they belong
- Individuals can be in one of three states:
 - Susceptible (S)
 - Infective (I)
 - Removed (R) (either recovered or dead)
- Constant rate of infection (infective contacts per unit time) β
- Constant rate of removal (recoveries per unit time) γ

Assumptions: rapidly spreading disease
recovered individuals become immune
individuals have the same number of contacts

Lowell Reed and Wade Hampton Frost (1920)

Solving the fully mixed model

- Fraction of individuals in one of three states:

1. Susceptible (s)
2. Infective (i)
3. Removed (r)

$$s = \frac{S}{N}, i = \frac{I}{N}, r = \frac{R}{N} \quad s + i + r = 1$$

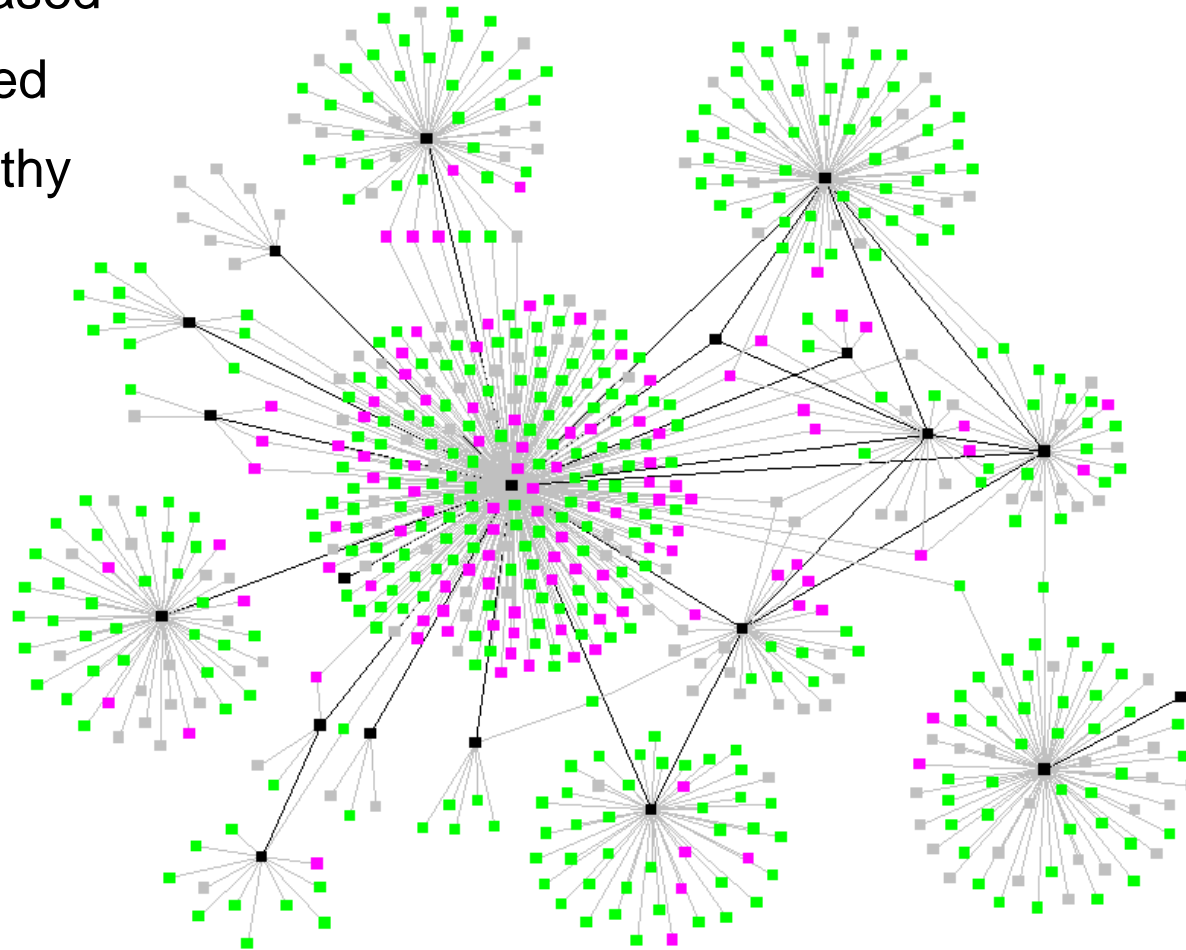
- rate of contact β , rate of removal γ

$$\frac{ds}{dt} = -\beta is, \quad \frac{di}{dt} = \beta is - \gamma i, \quad \frac{dr}{dt} = \gamma i$$

- the number of infected individuals decreases if $\beta/\gamma < 1$ – no epidemic
- $R_0 = \beta / \gamma$ is the basic reproductive ratio – the number of secondary cases produced by an infectious individual in a totally susceptible population

Spread of disease in a social network

- black: diseased
- pink: infected
- green: healthy



Network models

- Individuals are part of social networks
- Disease-causing contacts are only possible along the edges of this network
- There is a constant disease transmission probability T along an edge
- Nodes can only be susceptible or infected
- An outbreak starts from an infected node, and spreads with probability T to the first neighbors of the node, then to the second,
- Study whether the infection stops spreading or spreads to the whole network

Network representation

- At first, each node (individual) is susceptible
- mark (or occupy) each edge in the social network with probability T .
- The ultimate size of an outbreak would be the cluster of nodes that can be reached from an infective node by traversing marked edges.
- Thus, we only need to determine the sizes of the clusters formed by marked edges. We know that infection of any of the nodes in that cluster will cause an outbreak equal to the size of the cluster.
- If a marked cluster contains a large fraction of the nodes (it is a giant connected component), an infection of any of the nodes in that cluster will cause an **epidemic**.

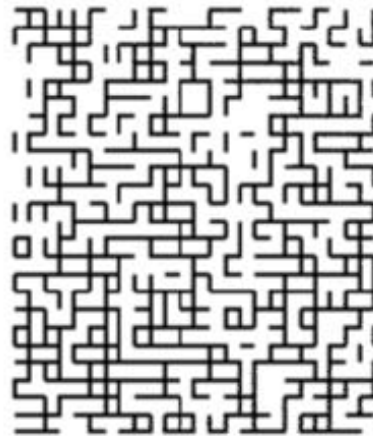
This is an example of bond percolation on a network!

Bond percolation

- Start with a lattice (or network)
- Draw (or mark) the edges with a certain probability p
- The remaining edges are open (unmarked)
- At a critical probability p_c a spanning cluster appears



$p=0.315$



$p=0.525$

H. L Frisch and J. M. Hammersley, J. SIAM 11, 894 (1963)

Percolation on a general random network

Start with a random network with a given degree distribution $P(k)$.

The network has a large connected component if

$$\sum_k k(k-2)P(k) > 0 \quad \text{or} \quad \frac{\langle k^2 \rangle}{\langle k \rangle} > 2$$

Mark edges with probability T . Disregard the unmarked edges.

Expectation: if $T > T_c$ there will exist a large connected component of marked edges.

T_c depends on $P(k)$

To find the exact relation, we need to use **generating functions**

H. Wilf, *Generatingfunctionology* (1994)

Generating functions in a graph

Node degree generating function

$$G_0(x) = \sum_{k=0}^{\infty} P(k)x^k, \quad |x| \leq 1$$

Finding $P(k)$ and degree moments from the generating function

$$P(k) = \frac{1}{k!} \left. \frac{d^k G_0}{dx^k} \right|_{x=0} \quad \langle k^n \rangle = \sum_k k^n P(k) = \left[\left(x \frac{d}{dx} \right)^n G_0(x) \right]_{x=1}$$

If a certain property is described by a gen. function, then its sum over m independent realizations is generated by the m th power of the gen. function

Ex. The generating function for the sum of the degrees of two nodes is $[G_0(x)]^2$

Generating functions in a graph

Generating function for the degree of nodes at the end of a random edge

probability to find that node

discount the edge we arrived along

$$G_1(x) = \frac{\sum_k kP(k)x^{k-1}}{\sum_k kP(k)} = \frac{G'_0(x)}{G'_0(1)}$$

normalization

M. E. J. Newman, S. H. Strogatz, D. J. Watts, Phys. Rev. E 64, 026118 (2001)

Generating functions for marked edges

Probability that m of the k edges of a node are marked

$$C_k^m T^m (1-T)^{k-m}$$

Generating function for marked edges

$$G_0(x; T) = \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} P(k) C_k^m T^m (1-T)^{k-m} x^m = G_0(1 - (1-x)T)$$

Generating function for nodes we arrive at following a random edge

$$G_1(x; T) = G_1(1 - (1-x)T)$$

M. E. J. Newman, Phys. Rev. E 66, 016128 (2002)

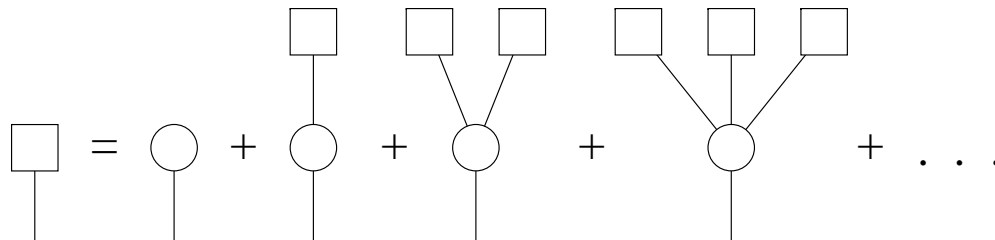
Generating functions for clusters

Generating function for clusters connected by marked edges

$P_s(T)$ - distribution of the marked cluster sizes

$$H_0(\mathbf{x}) = \sum_{s=0}^{\infty} P_s(T) \mathbf{x}^s, \quad |\mathbf{x}| \leq 1$$

Generating function for the marked cluster we reach by following a randomly chosen edge



$$H_1(\mathbf{x}; T) = \mathbf{x} G_1(H_1(\mathbf{x}; T); T)$$

$$H_0(\mathbf{x}; T) = \mathbf{x} G_0(H_1(\mathbf{x}; T); T)$$

Existence of a giant connected cluster

Average size of marked clusters

$$\langle s \rangle = H'_0(\mathbf{1}; T) = 1 + \frac{TG'_0(\mathbf{1})}{1 - TG'_1(\mathbf{1})}$$

Diverges when

$$T = T_c = \frac{1}{G'_1(\mathbf{1})} = \frac{\sum_k kP(k)}{\sum_k k(k-1)P(k)}$$

Giant connected cluster – epidemic when $T > T_c$

$$T_c = \frac{1}{\langle k^2 \rangle / \langle k \rangle - 1}$$

The more heterogeneous the network, the smaller T_c is, i.e. the easier it is for an epidemic to occur.

Breakdown transition in general random graphs

Consider a random graph with arbitrary $P(k_0)$

A giant cluster exists if each node is connected to at least two other nodes.

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

After the random removal of a fraction f of the nodes,

Breakdown threshold:

$$f_c = 1 - \frac{1}{\frac{\langle k_0^2 \rangle}{\langle k_0 \rangle} - 1}$$

A giant connected cluster exists if $f < f_c$

Parallels between epidemics on graphs and breakdown of graphs

Consider a random graph with arbitrary $P(k)$

If a fraction $f < f_c$ of nodes is lost, a giant component still exists

$$f_c = 1 - \frac{1}{\langle k^2 \rangle / \langle k \rangle - 1}$$

To model an epidemic process on this graph, assume that the edges transmit a disease with probability T

A giant connected component of disease-carrying edges exists if $T > T_c$

$$T_c = \frac{1}{\langle k^2 \rangle / \langle k \rangle - 1}$$

Complete equivalence between T and $1-f$

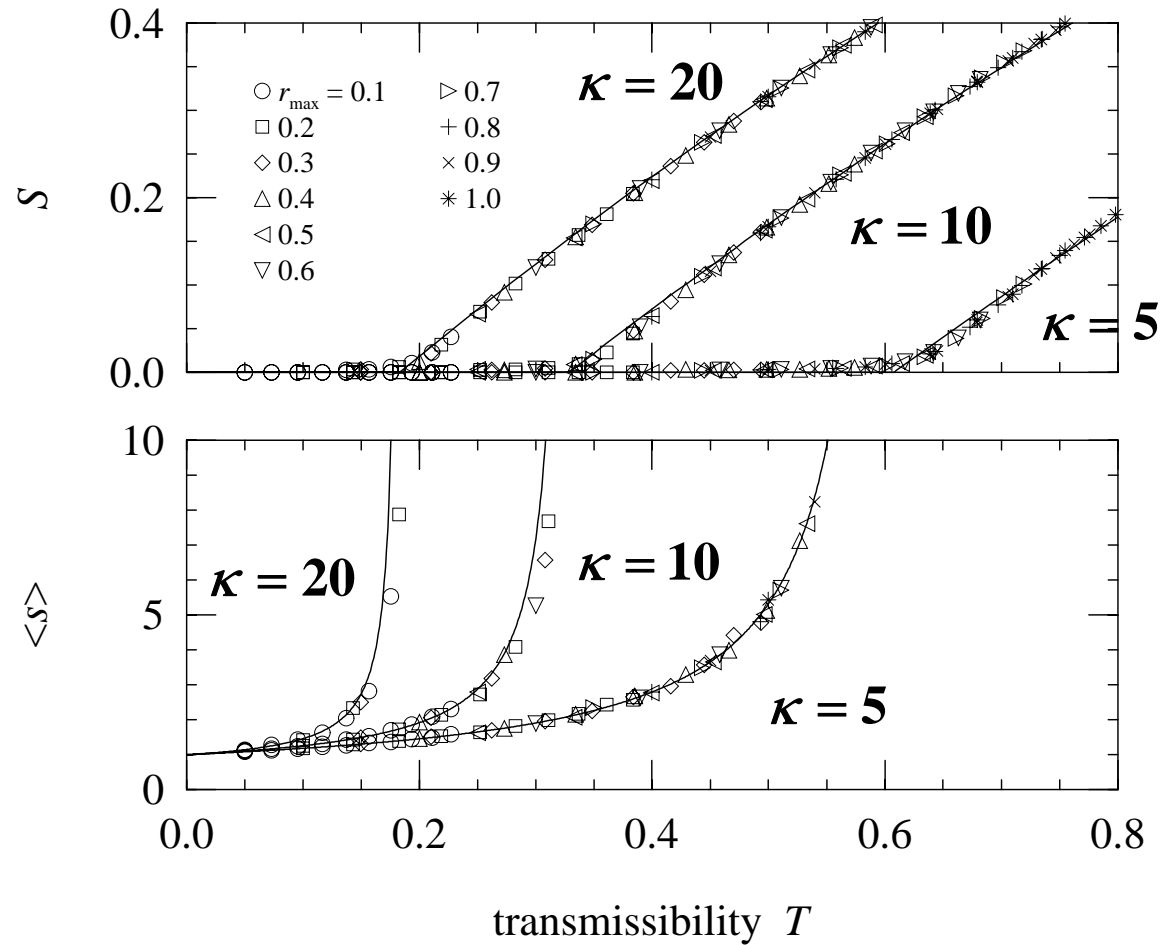
(Is this an obvious parallel?)

Example: scale-free network

Scale-free

with cutoff at $k=\kappa$

$$P(k) \approx k^{-\alpha} e^{-k/\kappa}$$



Epidemics in scale free networks

Random graph with $P(k) \approx k^{-\alpha} e^{-k/\kappa}$

$$T_c = \frac{Li_{\alpha-1}(e^{-1/\kappa})}{Li_{\alpha-2}(e^{-1/\kappa}) - Li_{\alpha-1}(e^{-1/\kappa})}, \quad Li_n(x) = \sum_k k^{-n} x^k$$

n th polylogarithm of x

T_c decreases with κ

$$\lim_{\kappa \rightarrow \infty} T_c = 0$$

Any infection leads to epidemics in infinite scale-free networks
with no cutoff.

Ex. 1

The network representation discussed here assumes that each node is susceptible. How should it be modified to include immune individuals?

How will the results change?

What is your expectation for the value of T_c ?

Ex. 2

Consider the emergence of a second disease after an epidemic swept through the population. Assume that infected and recovered individuals cannot contract the second disease.

How would you estimate the chances of the second disease to not die out?

Dynamics of epidemic spreading on a network

Susceptible-infected model on a network.

$$s = \frac{S}{N}, i = \frac{I}{N}, s + i = 1$$

The transmission rate β depends on the number of first neighbors.

Define the transmission rate per edge, λ

First approximation:

statistically homogeneous (random) network topology

the fraction of infected neighbors of a susceptible node can be

approximated by $\langle k \rangle i$ - mean field approximation

$$\frac{ds}{dt} = -\lambda \langle k \rangle i s, \quad \frac{di}{dt} = \lambda \langle k \rangle i s$$

Mean field SI, time-dependent behavior

Mean field approximation:

$$\frac{ds}{dt} = -\lambda \langle k \rangle is, \quad \frac{di}{dt} = \lambda \langle k \rangle is$$

Initial spread

$$i(t) \approx i_0 e^{t/\tau_H}, \quad \tau_H = 1/\lambda \langle k \rangle$$

τ_H - time scale governing the growth of the infection in a homogeneous network

Ex. 1

Write the rate of change of the fraction of susceptible, infected and recovered nodes in a susceptible-infected-recovered model in a network using the mean-field approximation.

Ex. 2

Write the same equations for a susceptible-infected-susceptible model.

Ex. 3

What is the condition for the existence of an epidemic in either of the cases above?

Heterogeneous network

Focus on nodes with given degrees

$$i_k = \frac{I_k}{N_k}, \quad N_k = NP(k), \quad s_k = 1 - i_k$$

Assumptions: nodes with degree k are equivalent

the fraction of infected nodes that are neighbors of a susceptible node, $k\Theta_k$, is the same for each node of degree k

$$\frac{di_k(t)}{dt} = \lambda k s_k(t) \Theta_k(t)$$

Heterogeneous network

$$\frac{di_k(t)}{dt} = \lambda k s_k(t) \Theta_k(t)$$

Θ_k - the density of infected neighbors of a vertex of degree k

Uncorrelated network:

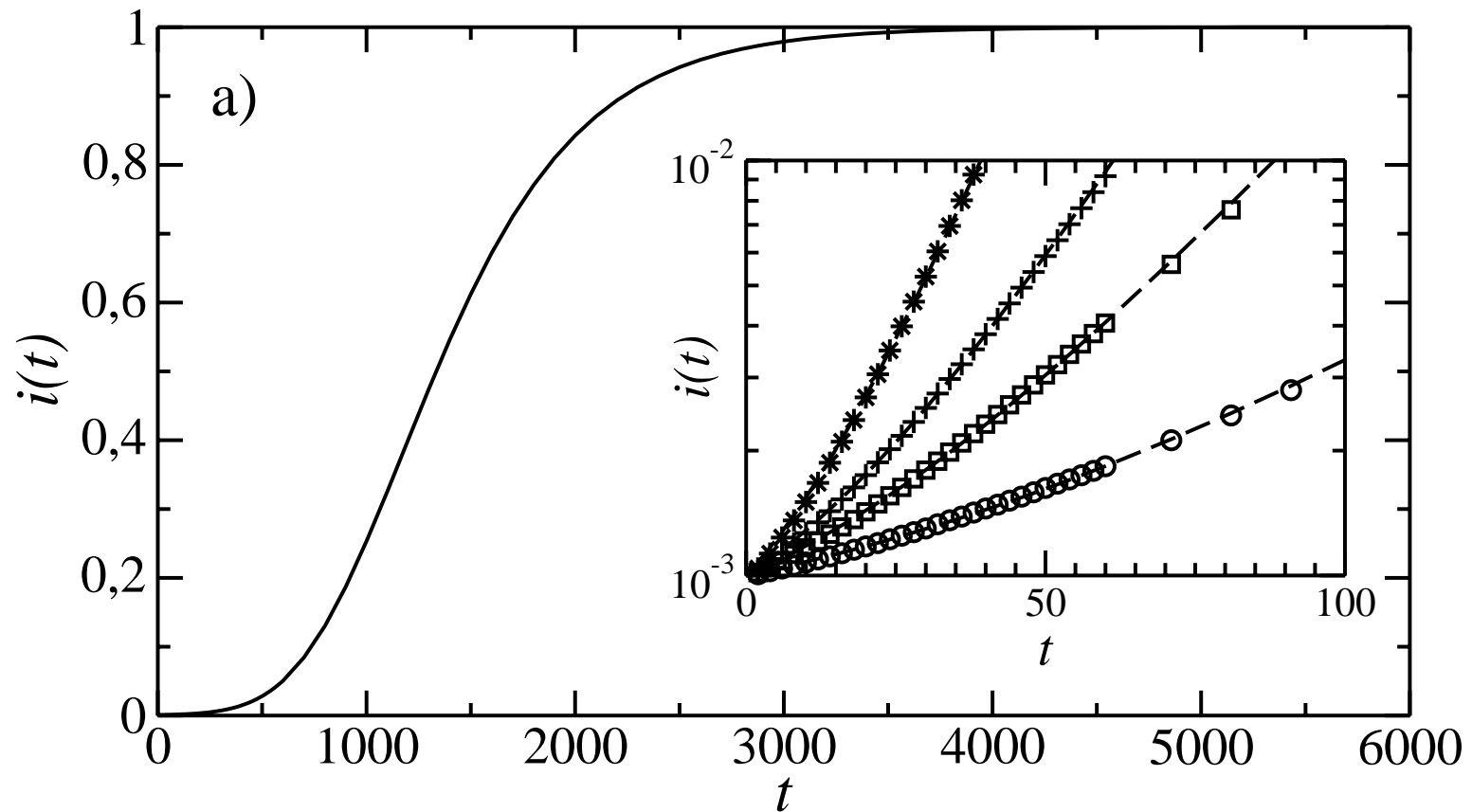
$$\Theta_k \equiv \Theta = \frac{\sum_l l P(l) i_l(t)}{\langle k \rangle}$$

First order:

$$\frac{di_k(t)}{dt} = \lambda k \Theta(t), \quad \frac{d\Theta(t)}{dt} = \lambda \frac{\langle k^2 \rangle}{\langle k \rangle} \Theta(t)$$

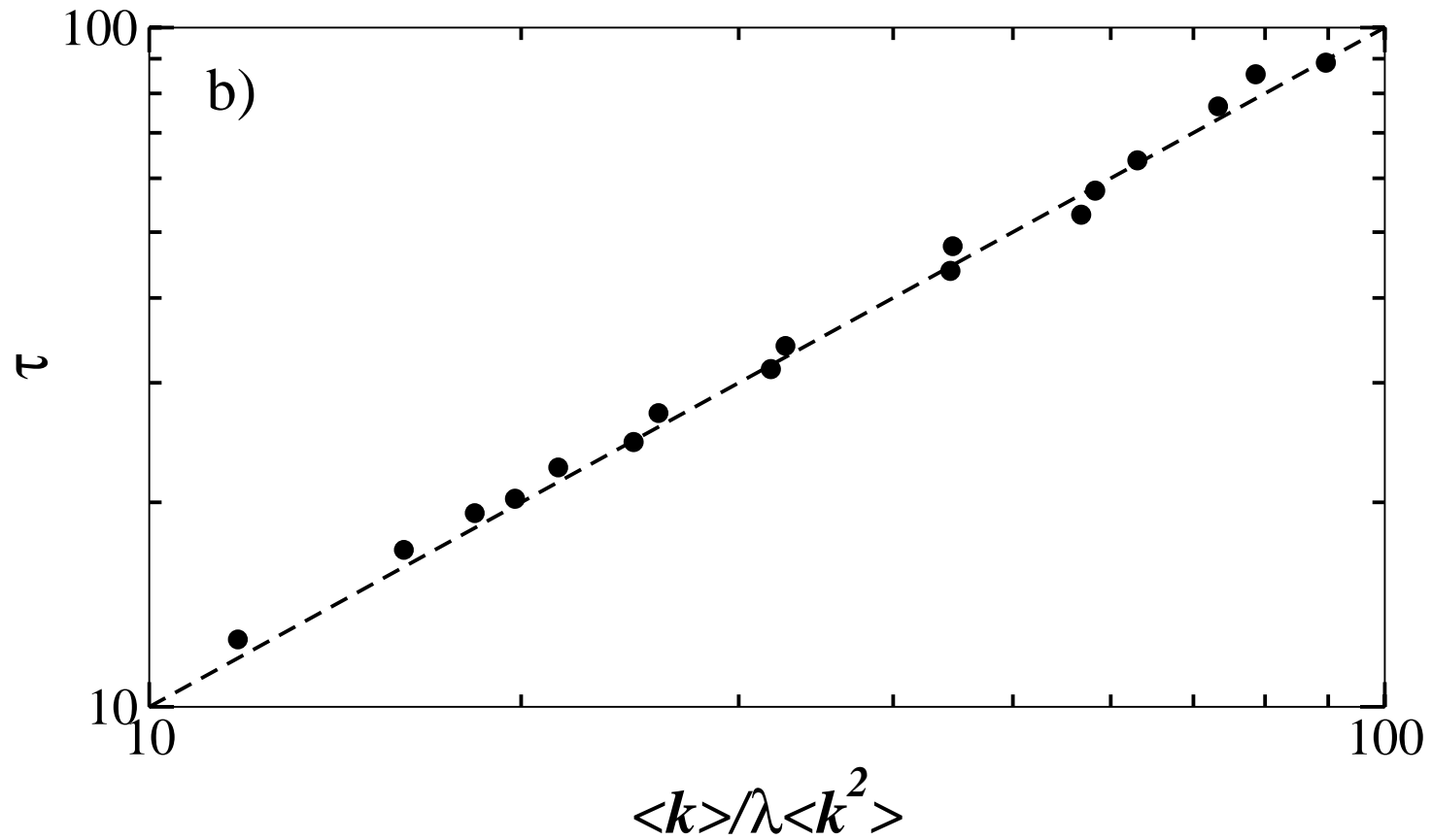
$$i(t) = i_0 \left[1 + \frac{\langle k \rangle^2 - \langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} (e^{t/\tau} - 1) \right] \quad \tau = \frac{\langle k \rangle}{\lambda \langle k^2 \rangle}$$

Increase in the fraction of infected nodes



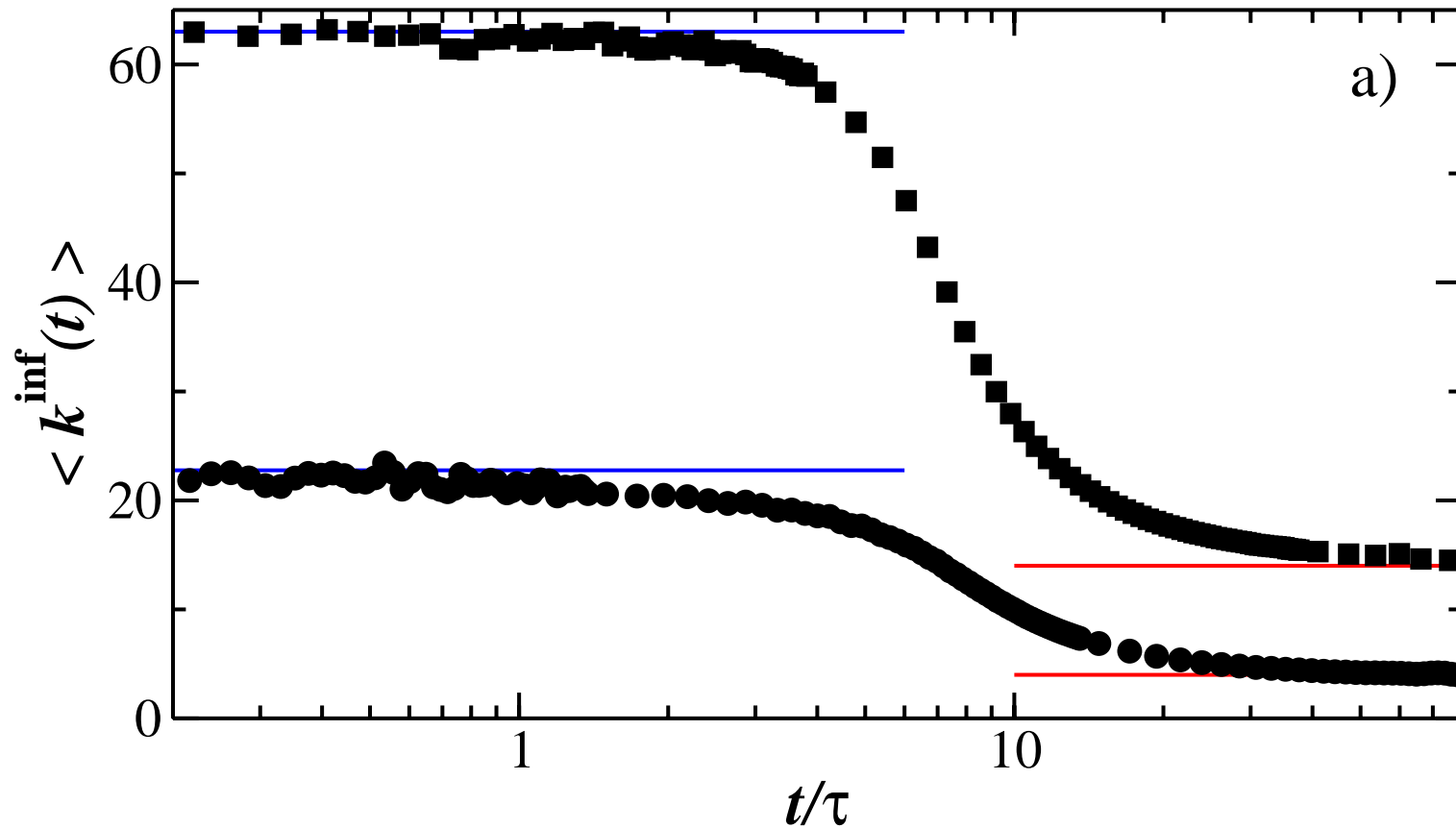
Barthelemy et al, cond-mat/0311501 (2004)

Time scale of the infection



The timescale of the initial spreading process depends inversely on the heterogeneity of the network.

High degree nodes catch the infection sooner



The average degree of newly infected nodes $\langle k^{\text{inf}} \rangle$ is high at the beginning of the spreading process.

Conclusions

- Infinite scale-free networks do not have an epidemic threshold – any disease can become an epidemic
- Diseases are able to spread efficiently in highly heterogeneous networks, the high degree nodes are rapidly infected
- This analysis focuses on the initial spread and does not describe the recovery process