Epidemics in Social Networks

Maksim Kitsak
Cooperative Association for Internet Data Analysis (CAIDA)
University of California, San Diego

Q1: How to model epidemics?

Q2: How to immunize a social network?

Q3: Who are the most influential spreaders?`

University of Nevada, Reno *December 2nd, 2010*

Pieter Bruegel's "The Triumph of Death," depicting plague in the 16th century
Image courtesy Museo del Prado, Madrid

430 B.C. Plague of Athens
25% population

1300-1700 Plague
~75-200 million died

1816-1923 Cholera (7 outbreaks)
~38 million died

1918-1920 Spanish Flu
20-100 million died

2003 S.A.R.S.
775 deaths

2009 H1N1 (Swine) Flu
18000 deaths

tim

Pieter Bruegel's "The Triumph of Death," depicting plague in the 16th century
Image courtesy Museo del Prado, Madrid

Other examples of epidemics

Rumor.
Ideas

Email Virus

MMS Virus

How can we model epidemics? Compartmental models!

SI

Susceptible $\xrightarrow{\beta}$ Infected

$$dI/dt = \langle k \rangle \beta S I$$
$$I + S = 1$$

SIS

Susceptible $\xrightarrow{\beta}$ Infected $\xrightarrow{\mu}$ Susceptible

$$dI/dt = \langle k \rangle \beta S I - \mu I$$
$$I + S = 1$$

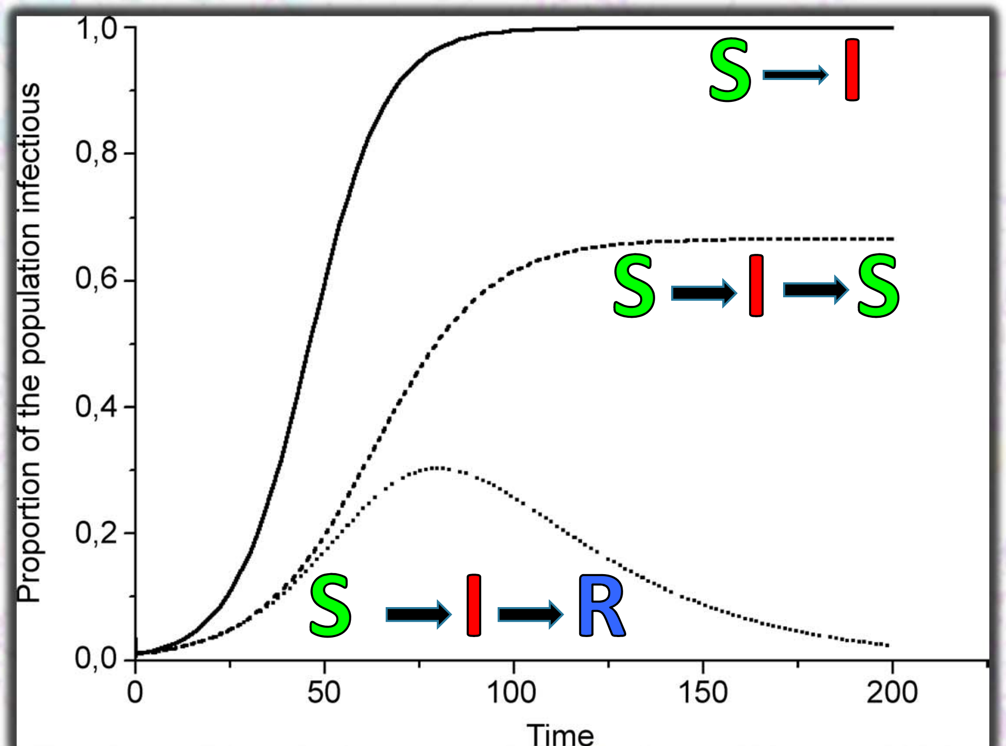
SIR

Susceptible $\xrightarrow{\beta}$ Infected $\xrightarrow{\mu}$ Recovered (immune)

$$dS/dt = -\langle k \rangle \beta S I$$
$$dI/dt = \langle k \rangle \beta S I - \mu I$$
$$I + S + R = 1$$

Assumption: Random Homogeneous Mixing!

How can we model epidemics? Compartmental models!



Everyone Infected

Endemic (equilibrium)
Recovery rate = infectious rate

Everyone Recovers

Critical threshold: **$\beta_c = \mu / \langle k \rangle$**

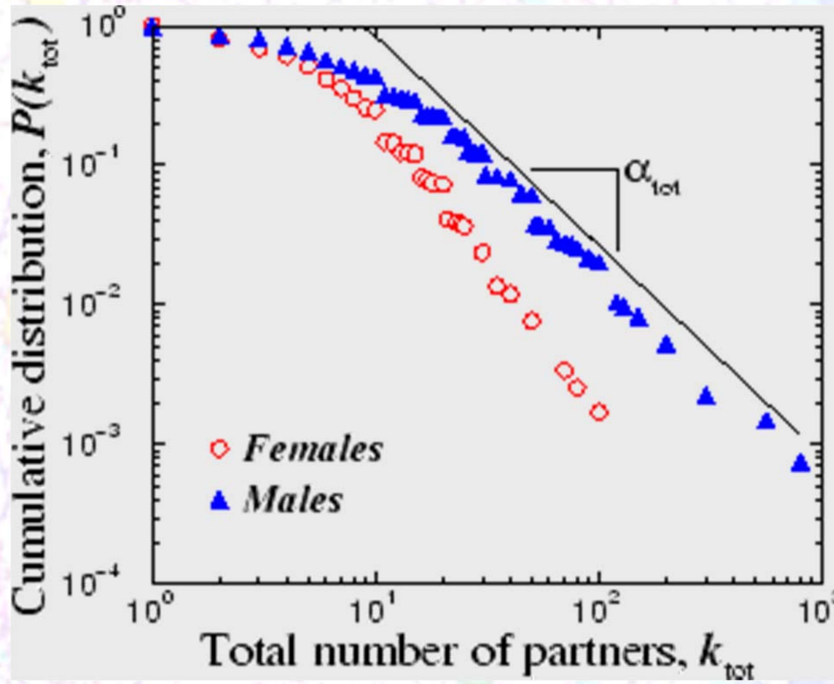
Disease extinct β_c Disease prevails **β**

Compartmental models surprisingly well reproduce highly contagious diseases.

Human sexual contacts

Nodes: people (Females; Males)

Links: sexual relationships

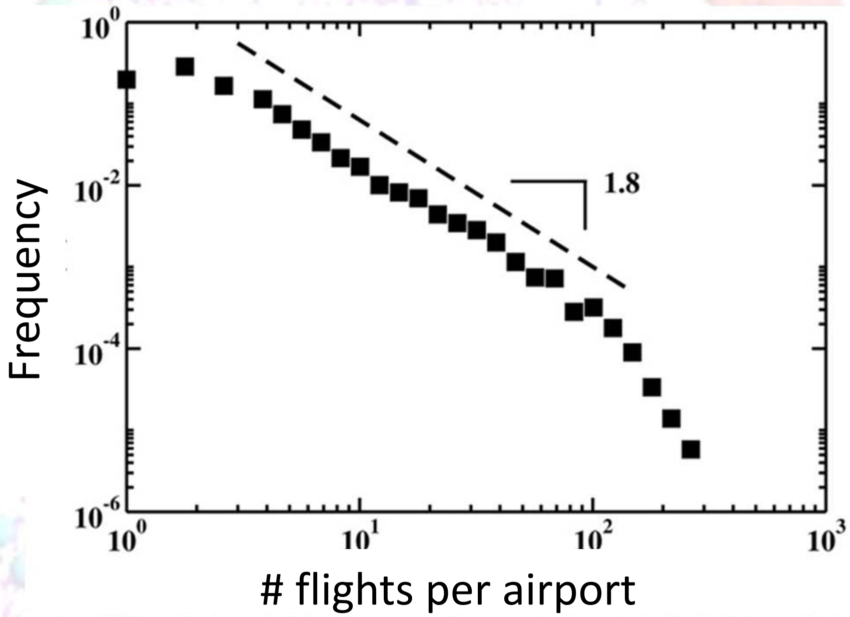


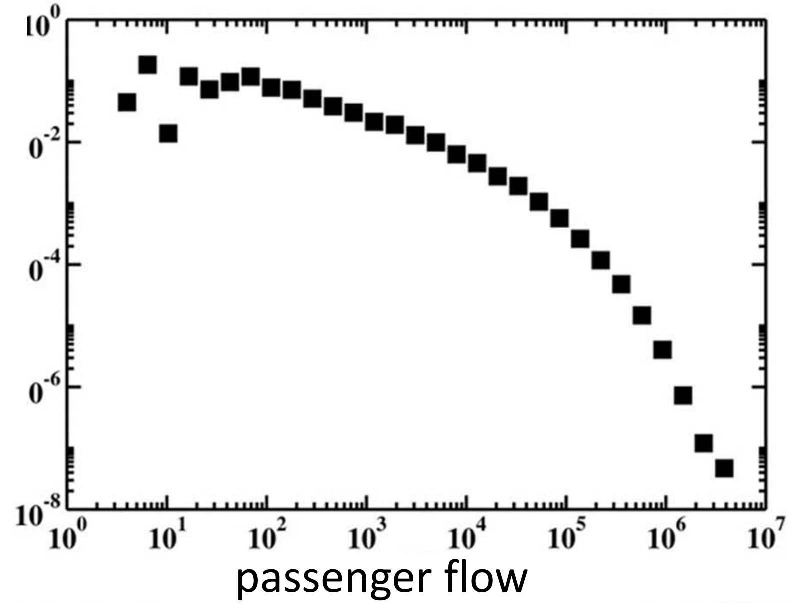
4781 Swedes; 18-74;
59% response rate.

Liljeros et al. Nature 2001

Worldwide Airport Network

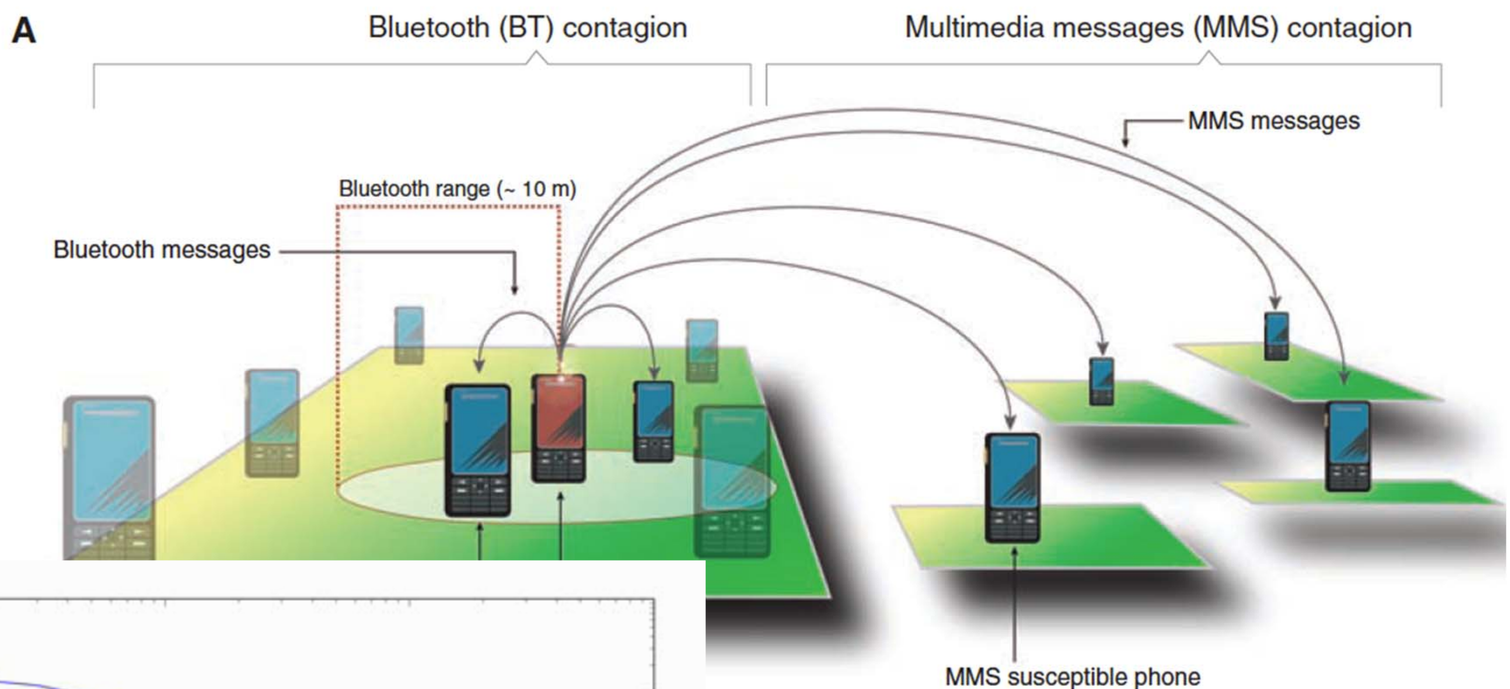
3100 airports
17182 flights
99% worldwide traffic

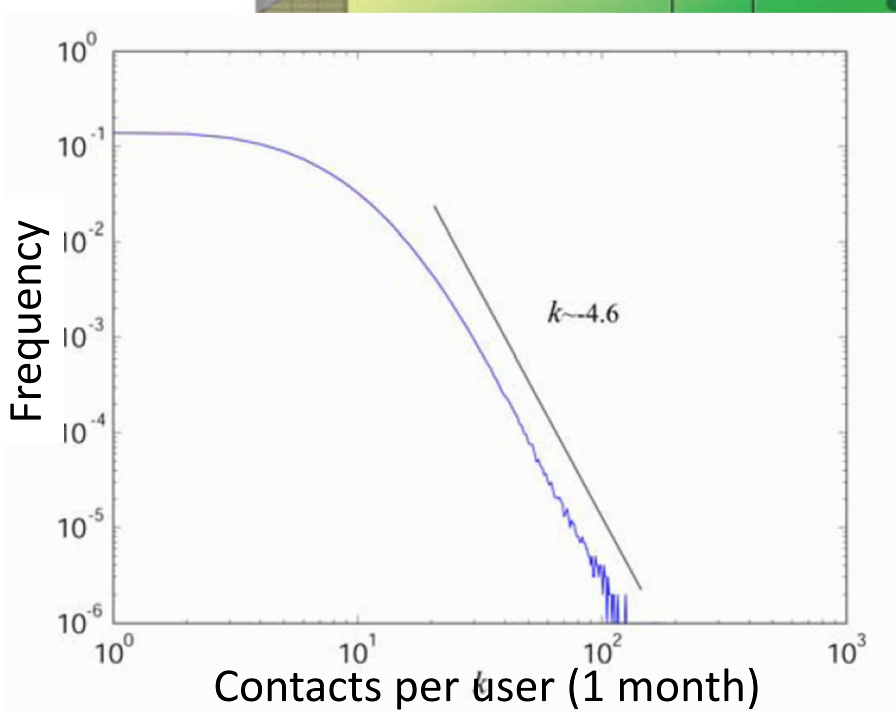




Colizza et al. PNAS 2005

Mobile Phone Contact Network





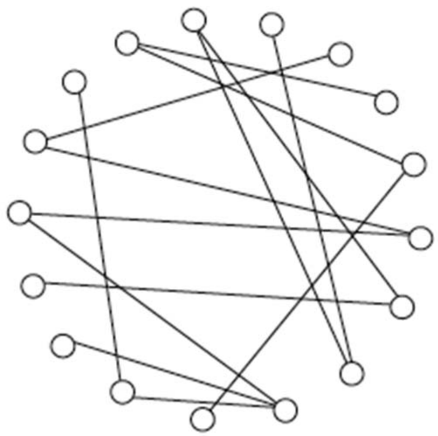
6.8 million users

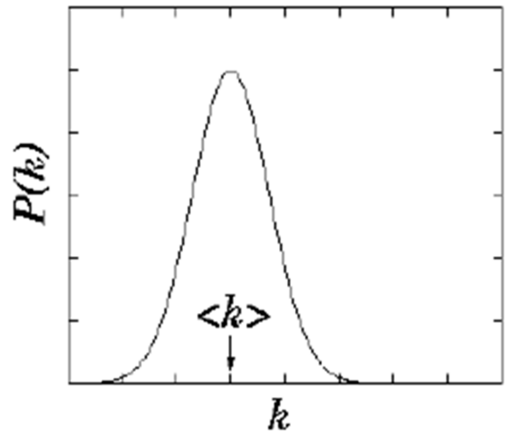
1 month observation

Wang et al. Science 2009

Random vs. scale-free networks

(a) Erdös Rényi

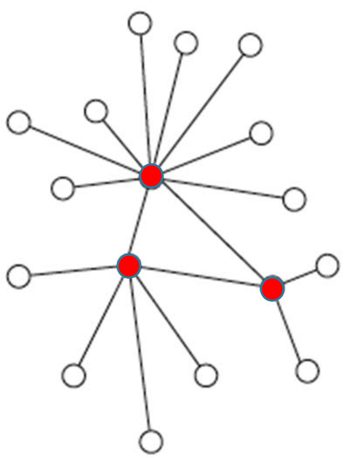


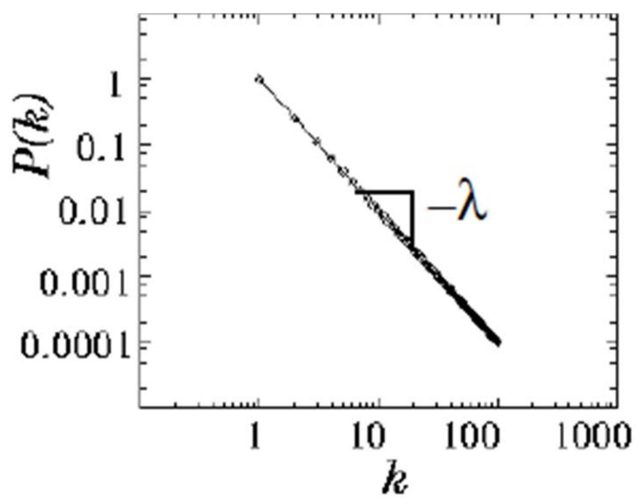


Poisson distribution
(Exponential tail)

$$P(k) = e^{-k} < k >^{k} / k!$$

(b) Scale-Free





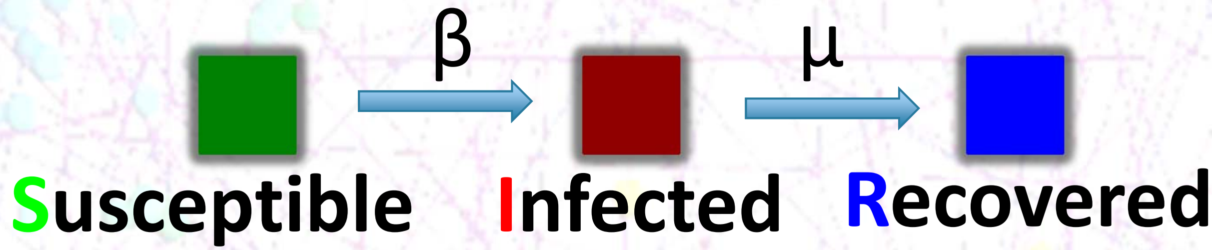
Power-law distribution

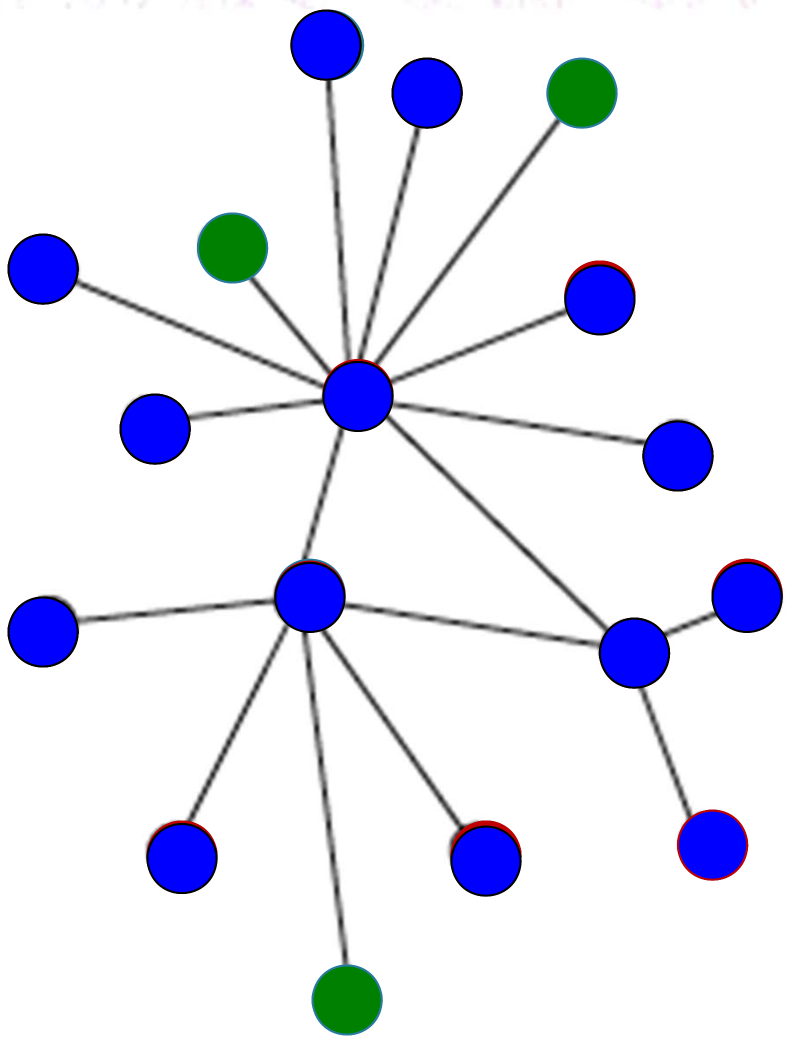
$$P(k) \sim k^{-\lambda}$$

$$\lambda \in (2,3)$$

Social networks are scale-free! Need *stochastic* epidemic models.

Stochastic SIR model





Transmission rate: $\beta = 0.5$

Recovery rate: $\mu = 0.5$

Quantities of interest:

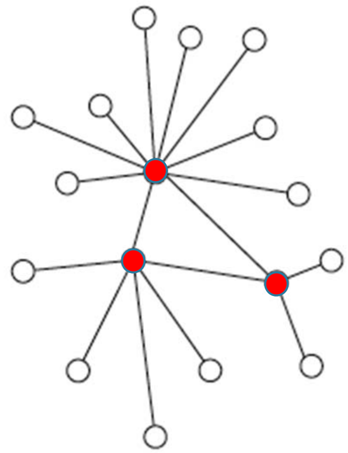
Total Recovered:

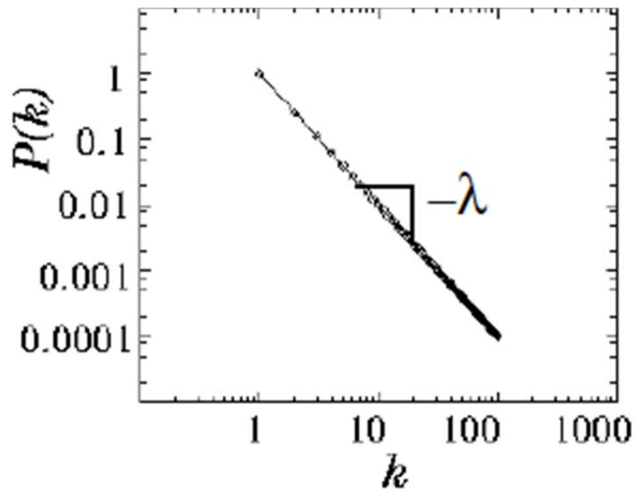
***M*=14**

Survivors: ***S*=3**

Total time: ***T*=5**

Epidemics in scale-free networks





Power-law distribution

$$P(k) \sim k^{-\lambda}$$

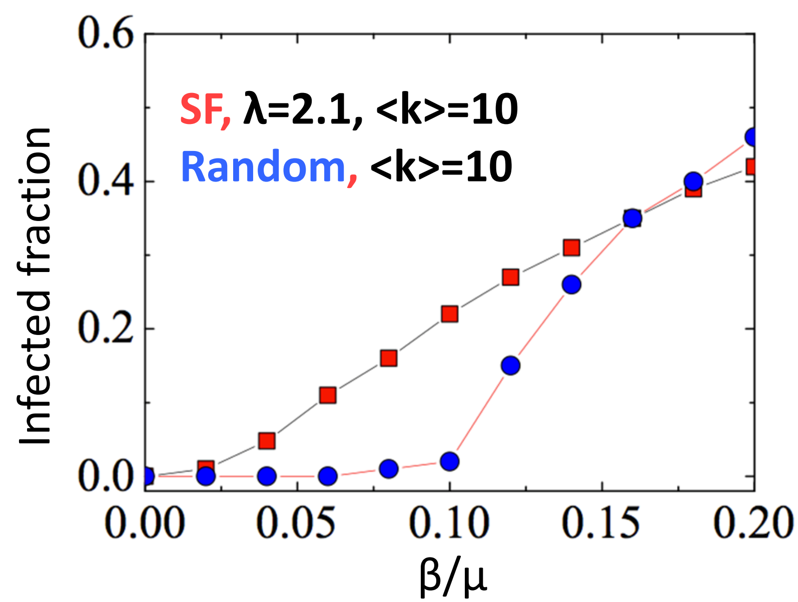
$$\lambda \in (2,3)$$

Anderson, May (1991)

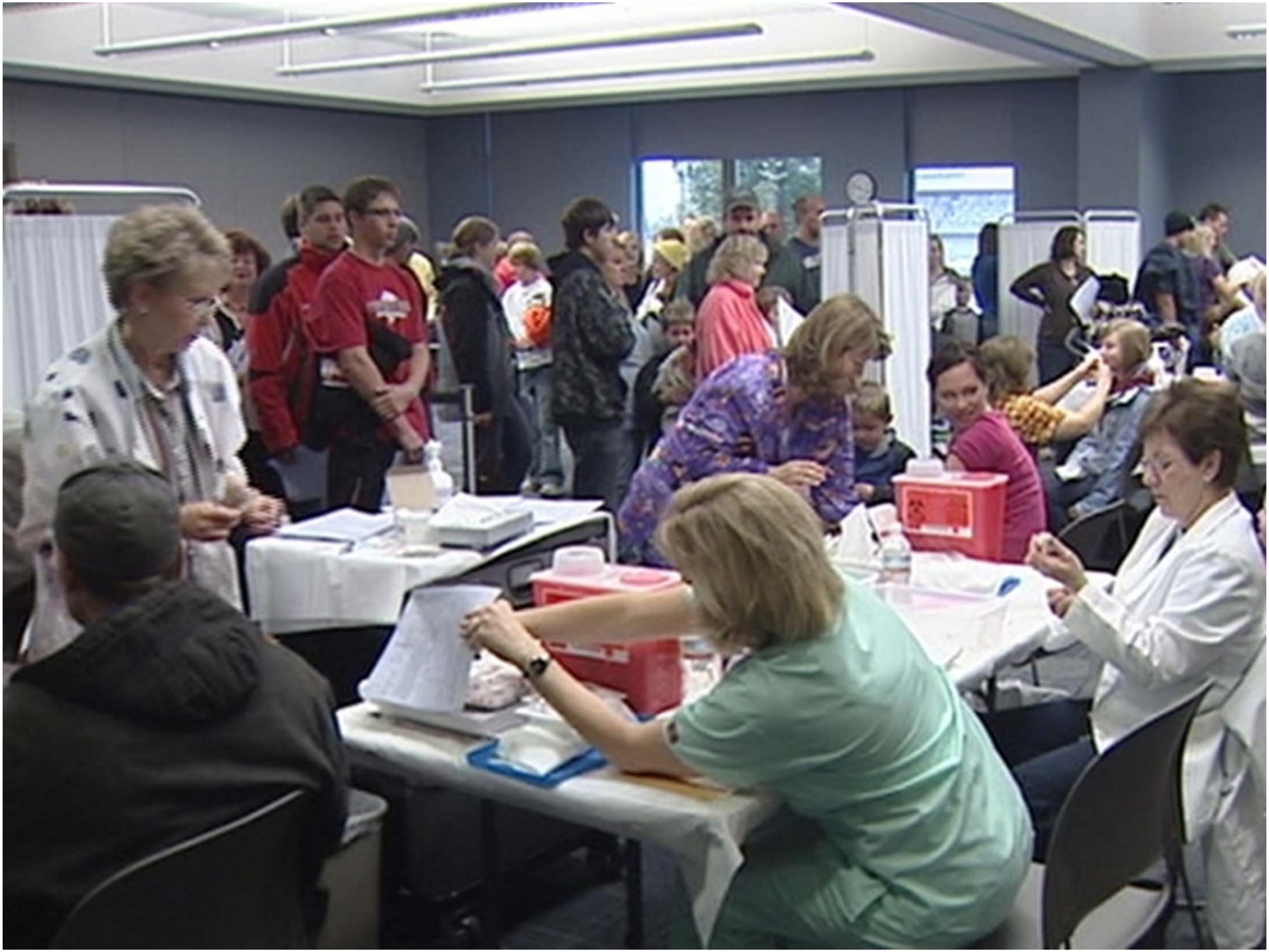
Epidemic threshold: $\beta_c = \mu \dfrac{\langle k \rangle}{\langle k^2 \rangle}$

$$\langle k^2 \rangle = \sum k^2 P(k) = \infty \quad (\lambda < 3)$$

$$\boxed{\beta_c = 0}$$



No epidemic threshold in Scale-free networks!



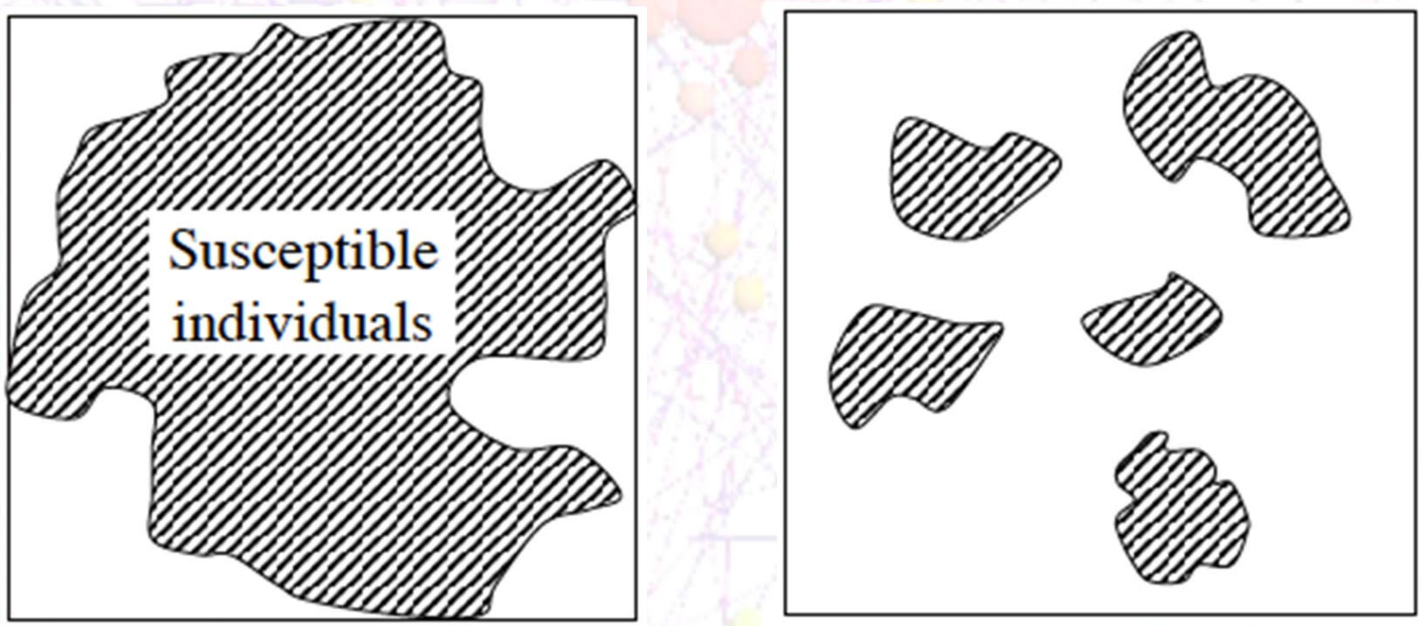
Network Immunization Strategies

Goal of efficient immunization strategy:

Immunize at least critical fraction f_c of nodes
so that only isolated clusters of susceptible individuals remain .
If possible, without detailed knowledge of the network.

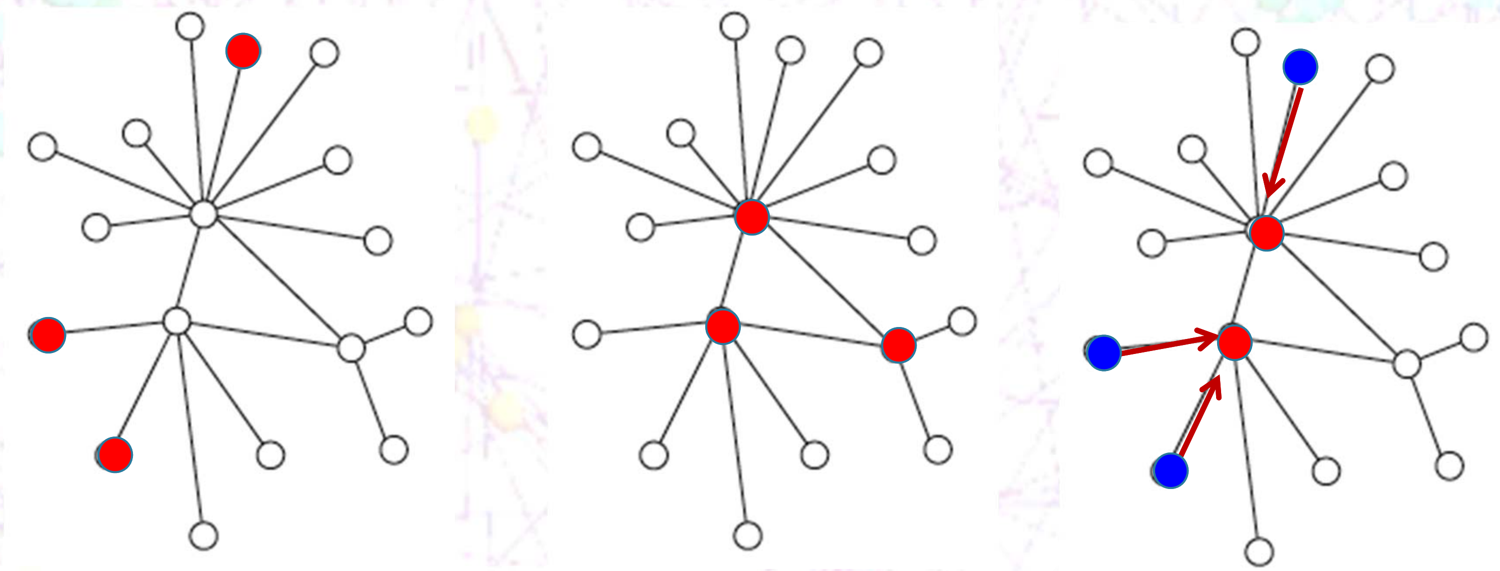
Large global cluster of susceptible individuals | Small (local) clusters of susceptible individuals

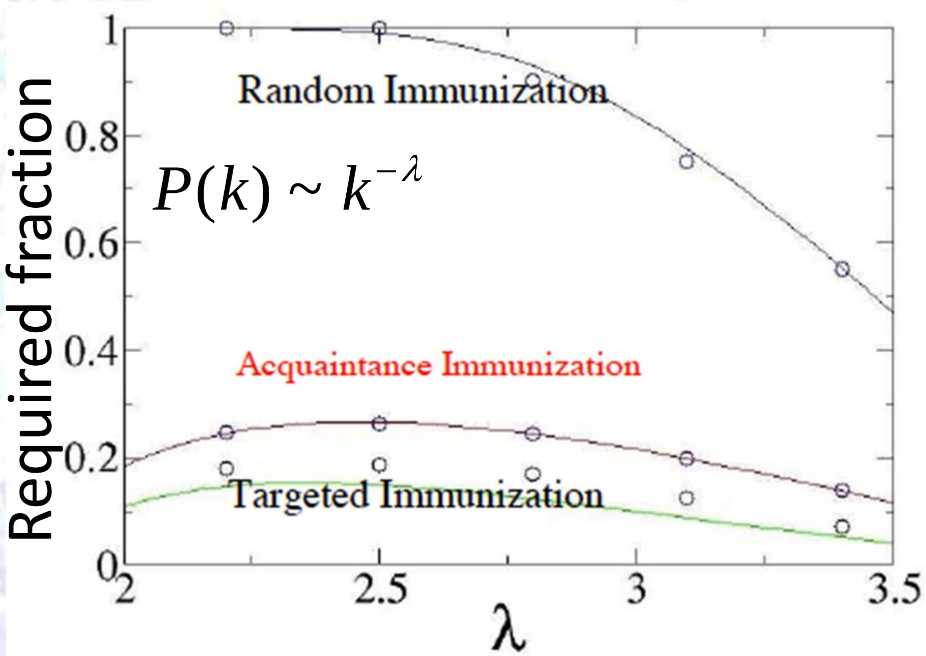
f=0 — f_c — **f=1**



Network Immunization Strategies

Random: **Targeted:** **Acquaintance:**





Random:
High threshold, no topology knowledge required.
Targeted:
Low threshold, knowledge of Connected nodes required.
Acquaintance:
Low threshold, no topology knowledge required.

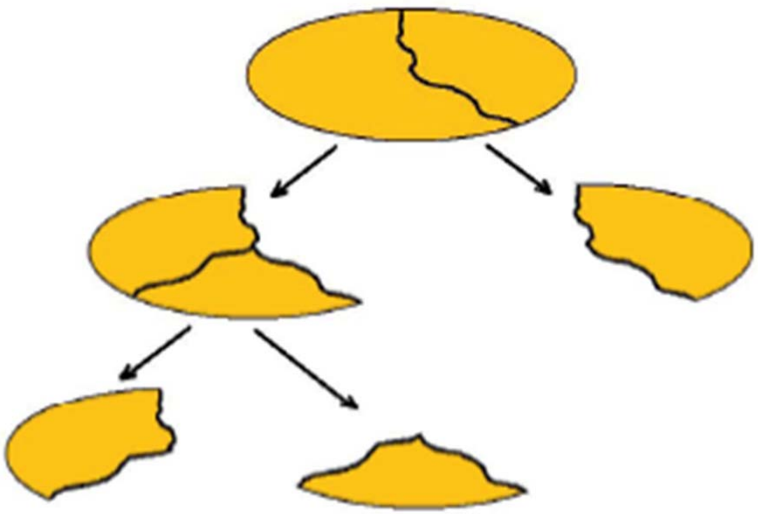
R. Cohen et al, Phys. Rev. Lett. (2003)

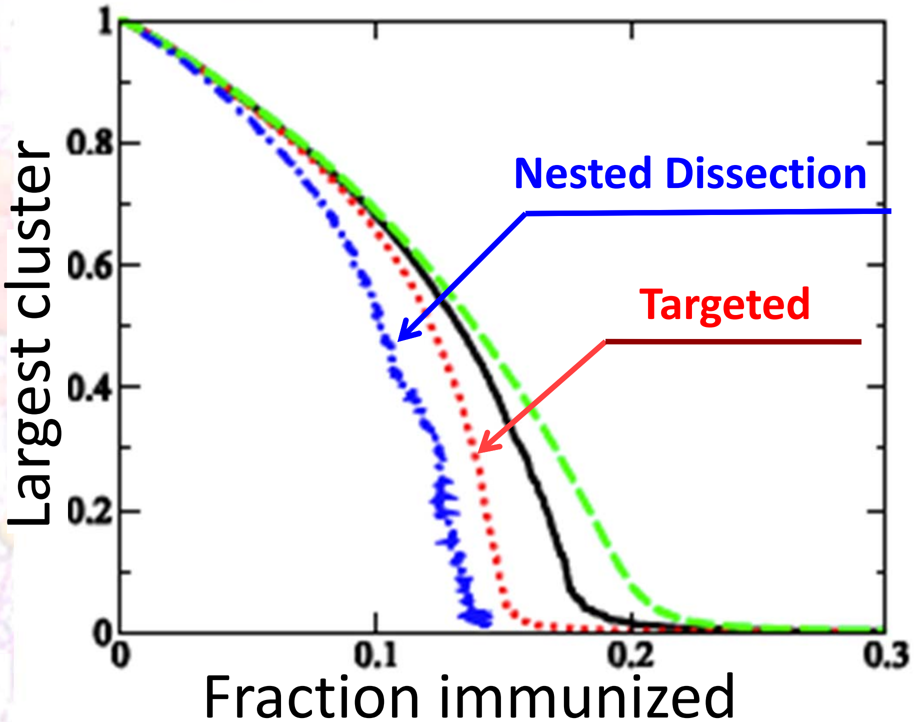
Graph Partitioning Immunization Strategy

Partition network into arbitrary number of same size clusters

Based on the Nested Dissection Algorithm

R.J. Lipton, SIAM J. Numer. Anal.(1979)





5% to 50% fewer immunization doses required

Y. Chen et al, Phys. Rev. Lett. (2008)

Who are the most influential spreaders?

SIR:

Who infects/influences the largest fraction of population?

SIS:

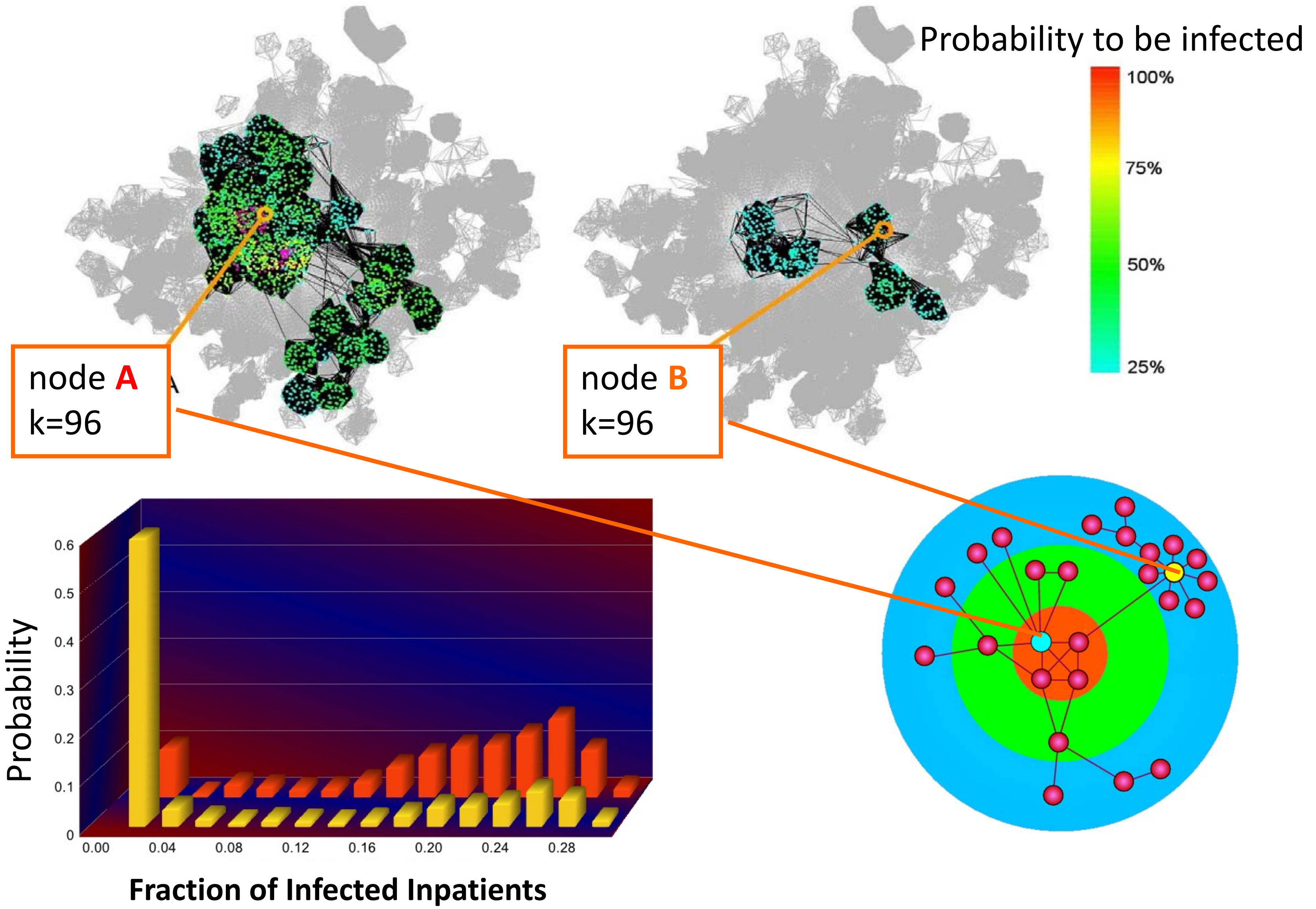
Who is the most persistent spreader? Who stays the most in the Infected state?

Not necessarily the most connected people!
Not the most central people!

M. Kitsak et al. Nature Physics 2010

Spreading efficiency determined by node placement!

Hospital Network: Inpatients in the same quarters connected with links



k-cores and *k*-shells determine node placement

***K*-core: sub-graph with nodes of degree at least k inside the sub-graph.**

Pruning Rule:

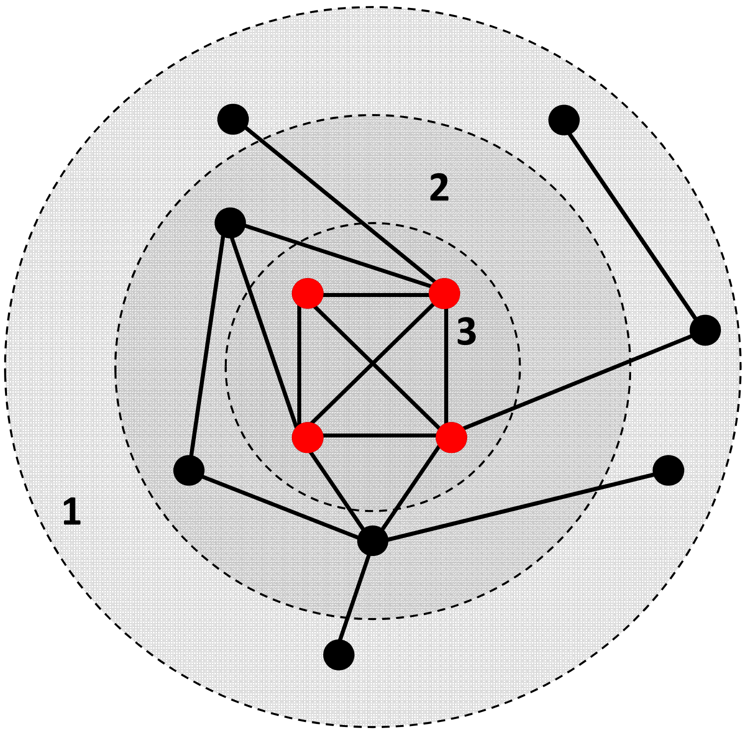
1) Remove all nodes with k=1.

Some remaining nodes may now have k = 1.

2) Repeat until there is no nodes with k = 1.

3) The remaining network forms the 2-core.

4) Repeat the process for higher k to extract other cores



S. B. Seidman, Social Networks, **5**, 269 (1983).

K-shell is a set of nodes that belongs to the *K*-core but NOT to the *K+1*-core

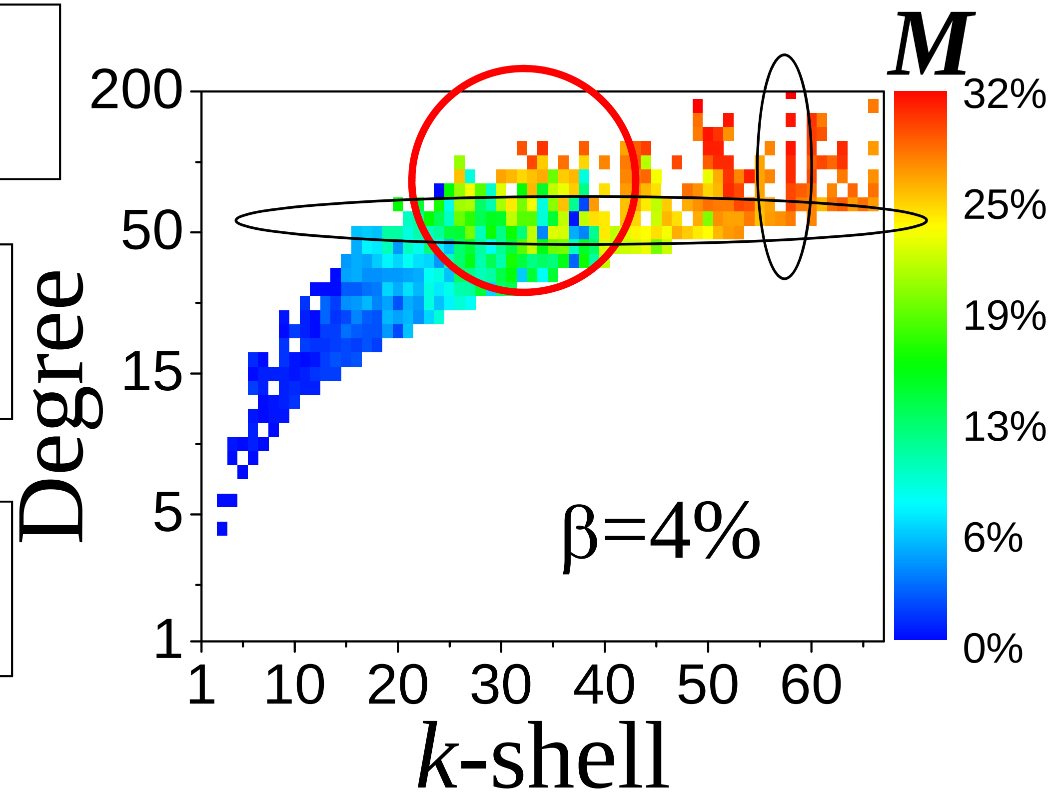
Identifying efficient spreaders in the hospital network (SIR)

(1) For every individual ***i*** measure the average fraction of individuals M_i he or she would infect (spreading efficiency).
(2) Group individuals based on the number of connections and the k-shell value.

A. Most efficient spreaders occupy high *k*-shells.

B. For fixed *k*-shell ***<M>*** is independent of ***k*.**

C. A lot of hubs are inefficient spreaders.



Three candidates:
Degree, *k*-shell, betweenness centrality

Imprecision functions test the merits of degree, k-shell and centrality

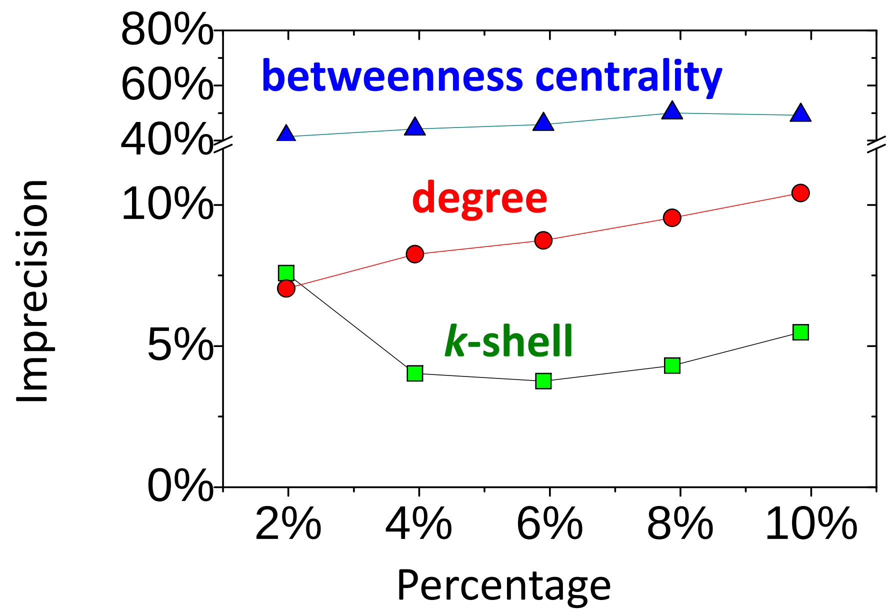
For given percentage ***p***

- Find ***Np*** the most efficient spreaders (as measured by ***M***)
- Calculate the average infected mass $\boldsymbol{M_{EFF}}$.
- Find ***Np*** the nodes with highest ***k-shell*** indices.
- Calculate the average infected mass $\boldsymbol{M_{kshell}}$.

Imprecision function:

$$\varepsilon(p) = 1 - \frac{M_{kshell}(p)}{M_{EFF}(p)}$$

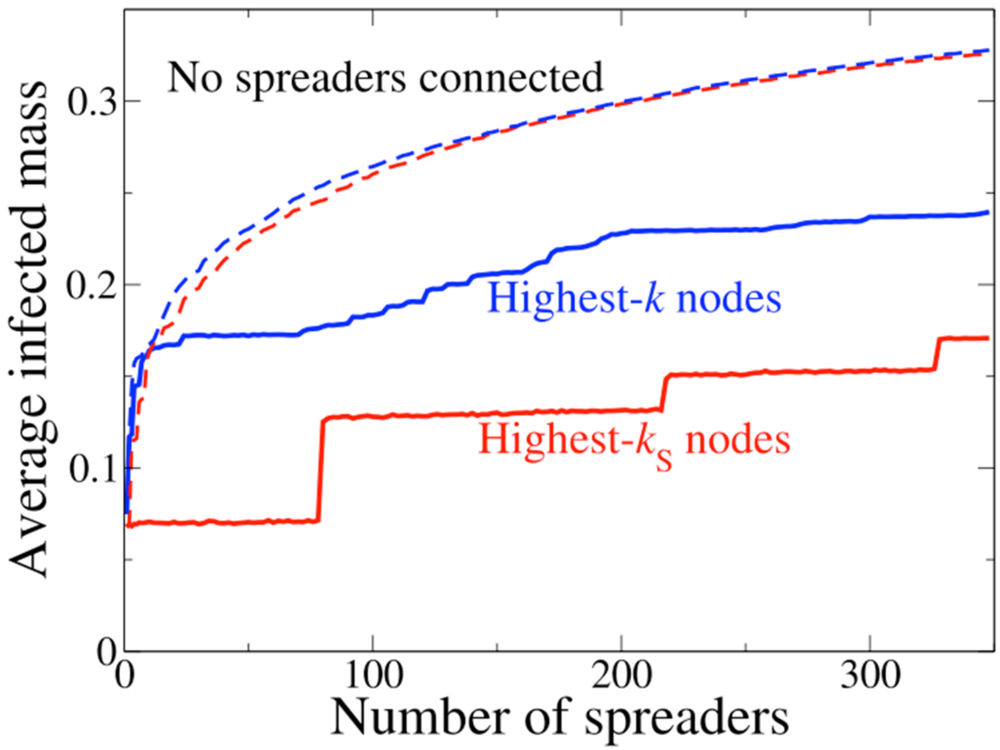
Measure the imprecision for K-shell, degree and centrality.

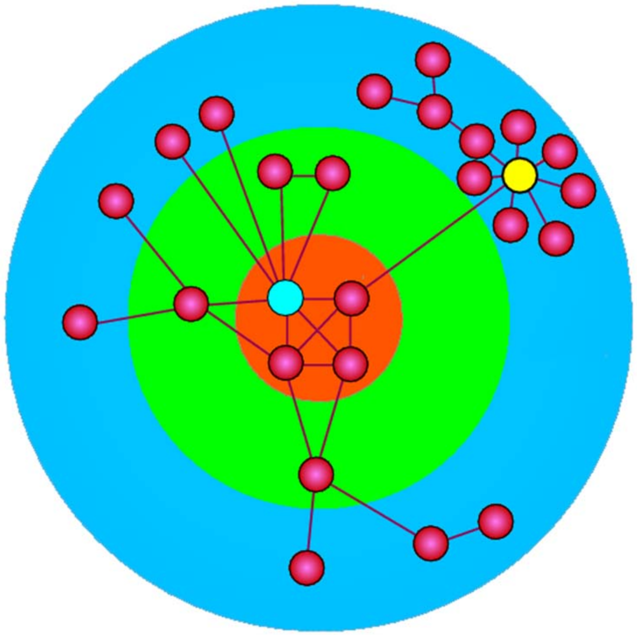


k-shell is the most robust spreading efficiency indicatior.
(followed by degree and betweenness centrality)

Multiple Source Spreading

What happens if virus starts from several origins simultaneously?



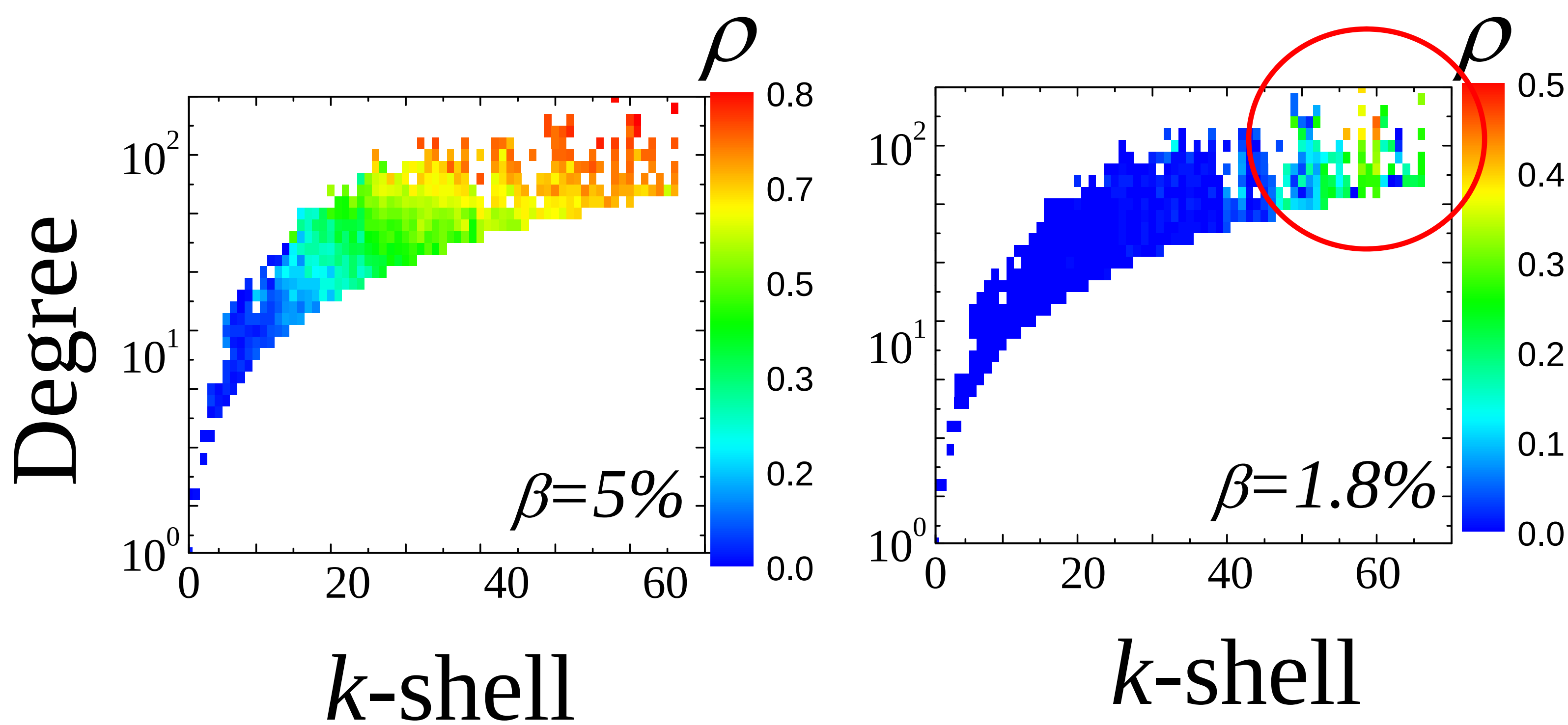


Multiple source spreading is enhanced when one “repels” sources.

Identifying efficient spreaders in the hospital network (SIS)

SIS: Number of infected nodes reaches endemic state (equilibrium)

Persistence $\boldsymbol{\rho_i(t)}$ (probability node ***i*** is infected at time ***t***)



High k-shells form a reservoir where virus can exist locally.

Consistent with core groups (H. Hethcote et al 1984)

Take home messages

1) *(Almost)* No epidemic threshold in Scale-free networks!

2) **Efficient immunization strategy:**

Immunize at least critical fraction f_c of nodes so that only isolated clusters of susceptible individuals remain

3) **Immunization strategy is not reciprocal to spreading strategy!**

4) **Influential spreaders (not necessarily hubs) occupy the innermost k-cores.**

Collaborators

Lazaros K. Gallos
CCNY, New York, NY

Lev Muchnik
NYU, New York, NY

Shlomo Havlin
Bar-Ilan University
Israel

Fredrik Liljeros
Stockholm University
Sweden

H. Eugene Stanley
Boston University,
Boston, MA

Hernán A. Makse
CCNY,
New York, NY