Identifying Influential Spreaders in Complex Networks

Maksim Kitsak, Lazaros K. Gallos, Shlomo Havlin,
Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley, and Hernan A. Makse

http://arxiv.org/abs/1001.5285

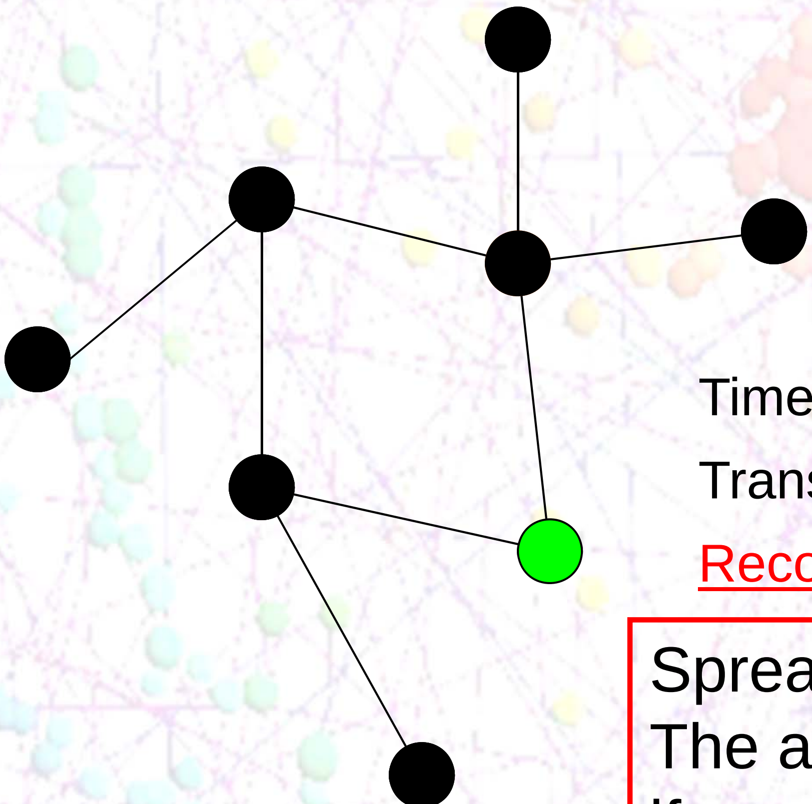
Q1: Definition of spreading efficiency?

Q2: What determines spreading efficiency?

Q3: Who are the most efficient spreaders?

Spreading Processes: Examples and Models

- **Examples:**
- **Infectious Diseases*(smallpox, influenza…)***
- **Innovations, Rumor, Ideas**
- ***Computer* Viruses *(spreading via email)***



The SIR Model

“**S**usceptible” (unaffected) individual.

“**I**nfected” (affected) individual.

“**R**ecovered” individual.

Time to “recover” $T_R = 2$

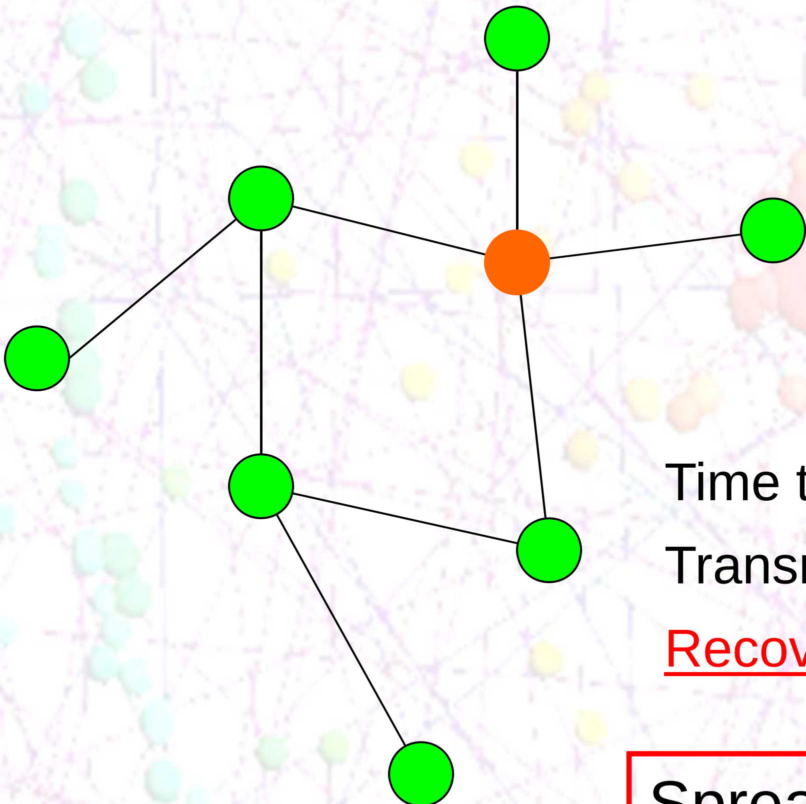
Transmission probability $\beta = 0.5$

Recovered individuals can not be infected!!!

Spreading efficiency: ***<M_i>***
The average number of infected nodes if spreading starts at node ***i***

Spreading Processes: Examples and Models

- **Infectious Diseases*(smallpox, influenza…)***
- **Innovations, Rumor, Ideas**
- **Computer Viruses *(spreading via email)***



The SIS Model

"Susceptible" (unaffected) individual.

"Infected" (affected) individual.

~~"Recovered" individual.~~

Time to "recover" $T_R = 2$

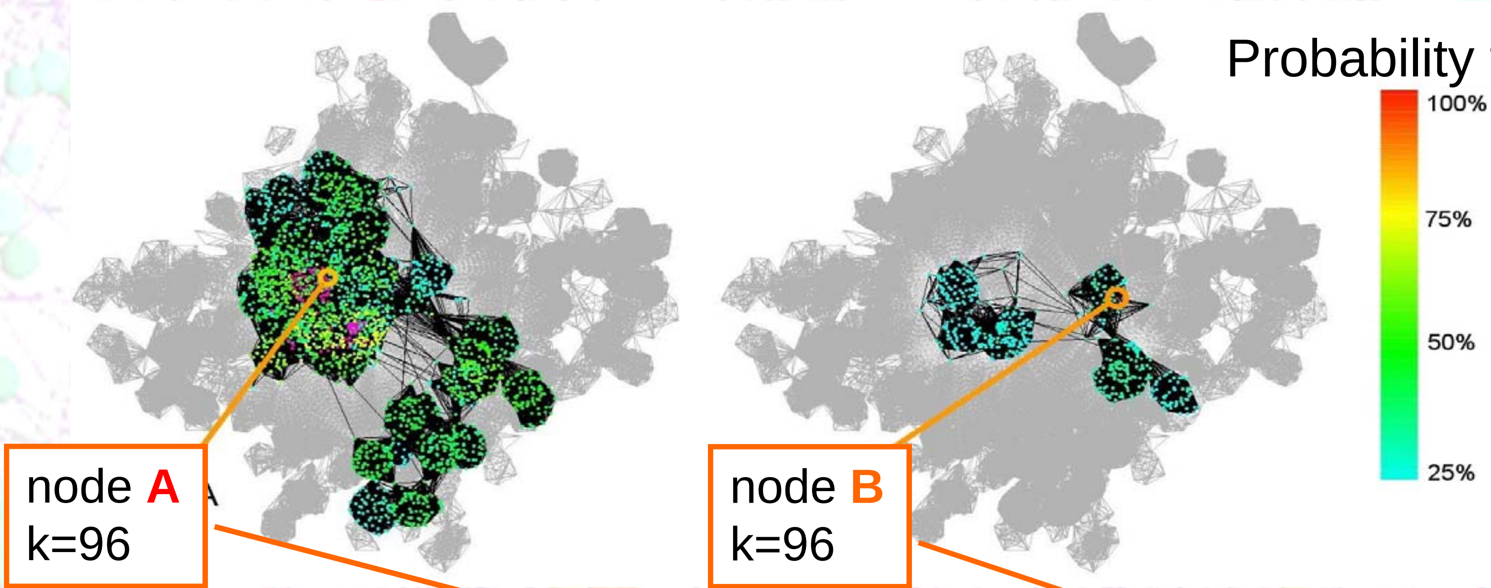
Transmission probability $\beta = 0.5$

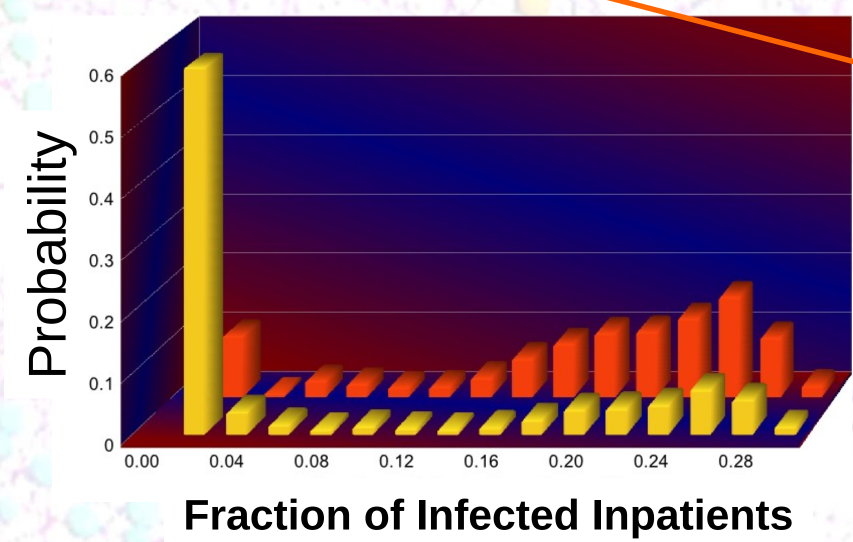
Recovered individuals can be infected again!!!

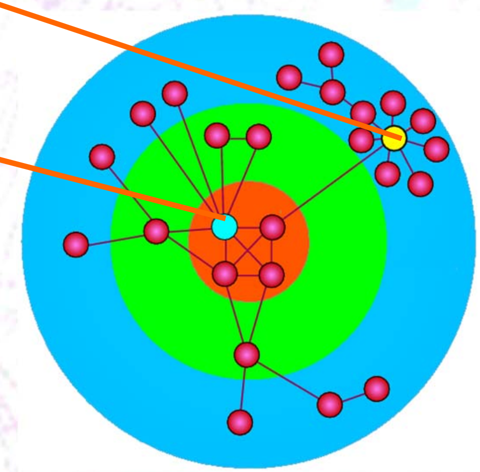
Spreading efficiency: ***$\rho_i(t)$***
Probability node ***i*** is infected at time ***t***

Spreading efficiently 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:

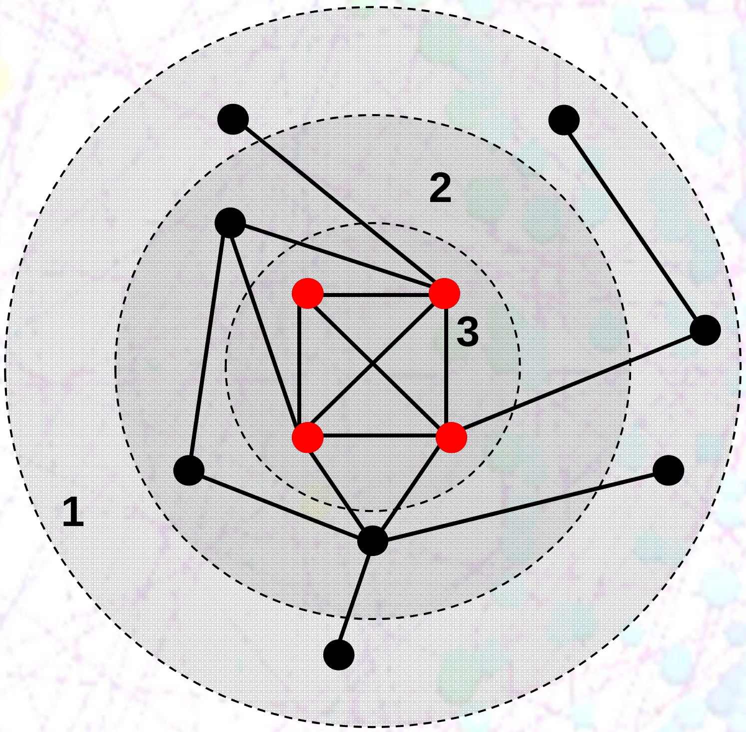
1) Remove all nodes with k=1.

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

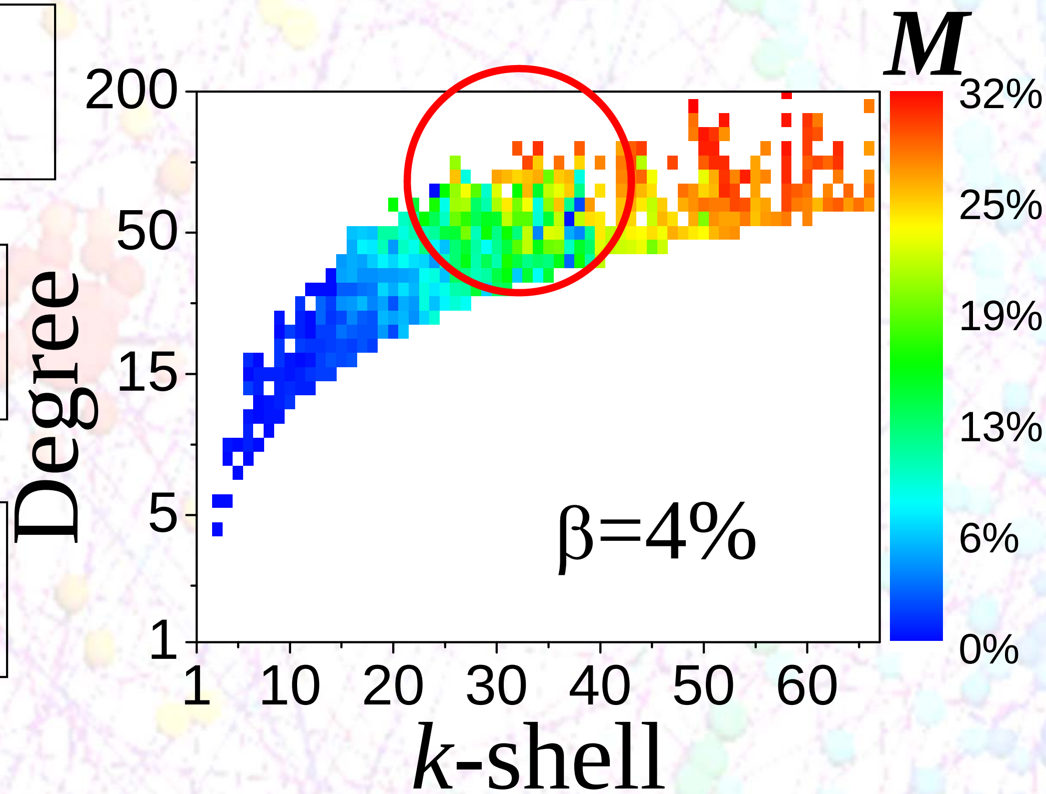
Identifying efficient spreaders in the hospital network (SIR)

(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

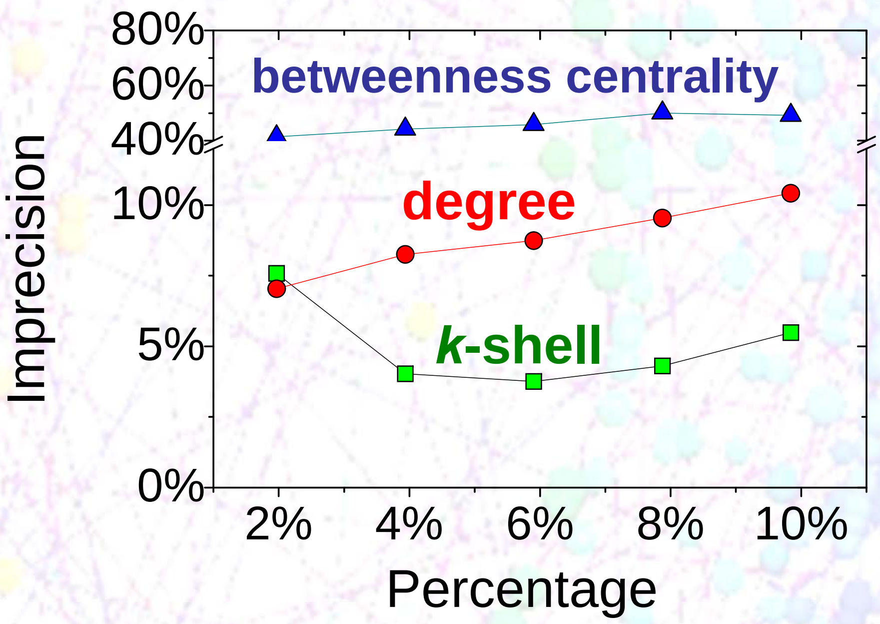
For given percentage ***p***

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

Imprecision function:

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

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

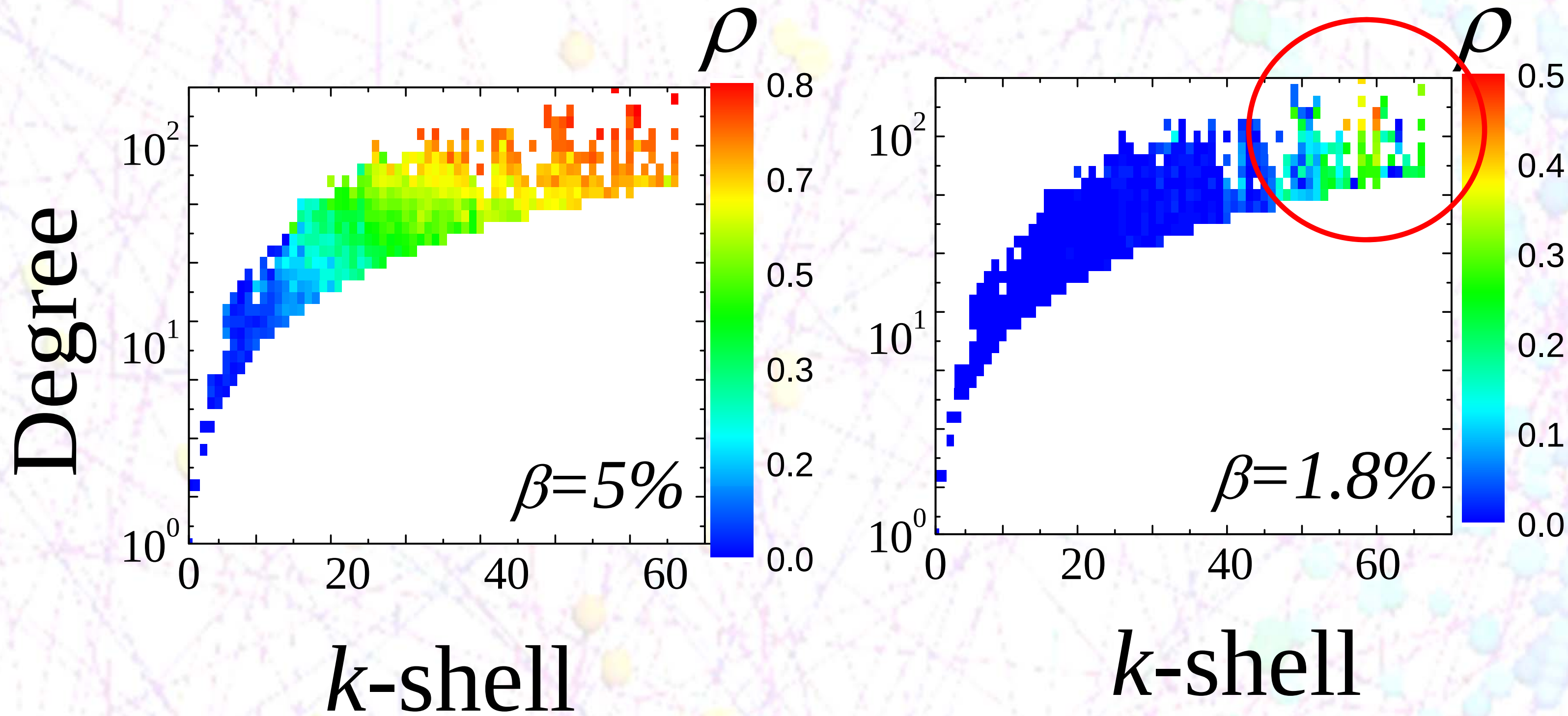


k-shell is the most robust spreading efficiency indicatior.
(followed by degree and betweenness centrality)

Identifying efficient spreaders in the hospital network (SIS)

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

Persistence $\rho_i(t)$ (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)

Summary

SIR

1) k-shell value is a reliable indicator of spreading efficiency. The most efficient spreaders occupy the innermost k-shells.

2) Multiple source spreading is enhanced when one “repels” sources. (*Discussed in the paper*)

SIS

3) High k-shells form a reservoir where virus can survive locally and infect neighbor nodes.

4) High k-shells may decrease epidemic threshold.

5) Immunization/Removal of high k-shells helps to suppress virus persistence.

http://arxiv.org/abs/1001.5285