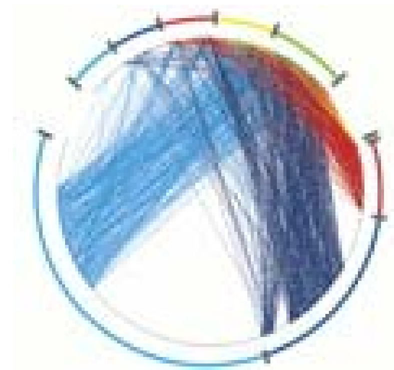
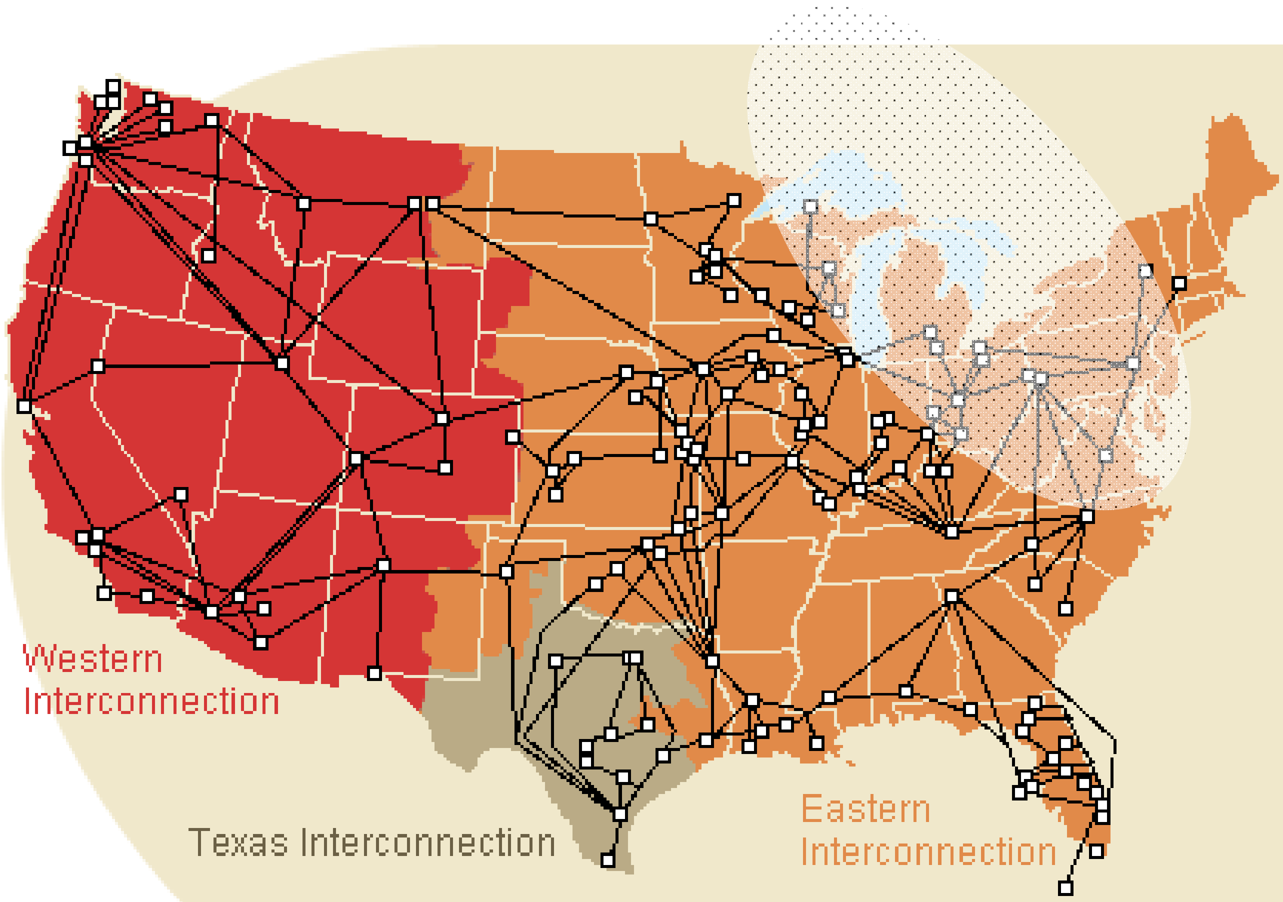
Lecture 19&20&21: Information Cascade, Cascaded Behavior and Epidemics on Complex Networks



Blackout in eastern US in Aug03



6,000 power generating units
500,000 miles of transmission lines

Following the crowd

- When people are connected by a network, they influence each other's behavior and decisions:
  - in the opinions they hold
  - the products they buy
  - the political positions they support
  - The activities they participate in
  - the technologies they use
  - and ….

An example

- Consider you are in unfamiliar town and you want to eat in a restaurant
- You want to go to restaurant A (based on your own information)
- When you arrive, you see there is no one in A
- Instead, restaurant B is crowded
- What would you do?
- It may be rational to join the crowd at B rather than to follow your own information

An example

- To see how this is possible, suppose that each diner has obtained independent but imperfect information about which of the two restaurants is better
- Then, if there are already many diners in B, the information that you can infer from their choices may be more powerful than your own private information.
- In this case, we say that **herding**, or an **information cascade**, has occurred.

So,

- An information cascade has the potential to occur when people make decisions sequentially, with later people watching the actions of earlier people, and from these actions inferring something about what the earlier people know.
- In the restaurant example, when the first diners to arrive chose restaurant B, they conveyed information to later diners about what they knew.
- A cascade then develops when people abandon their own information in favor of inferences based on earlier people's actions.

An experiment by Milgram, Bickman, and Berkowitz in 60s

- The experimenters had groups of people ranging in size from just one person to as many as fifteen people stand on a street corner and stare up into the sky
- They observed how many passersby stopped and also looked up at the sky
- They found that with only one person looking up, very few passersby stopped
- If five people were staring up into the sky, then more passersby stopped, but most still ignored them.
- Finally, with fifteen people looking up, they found that 45% of passersby stopped and also stared up into the sky.

An experiment by Milgram, Bickman, and Berkowitz in 60s

- The experimenters interpreted this result as demonstrating a social force for conformity that grows stronger as the group conforming to the activity becomes larger.
- Another possible explanation could be based on the idea of information cascades.
- It could be that initially the passersby saw no reason to look up (they had no private or public information that suggested it was necessary)
- But with more and more people looking up, future passersby may have rationally decided that there was good reason to also look up (since perhaps those looking up knew something that the passersby did not know)

Direct benefit effects

- Sometimes, you directly benefit from choosing an option that has a large user population.
- Consider a fax machine: if you have one and no one has, it is useless
- This is different from information cascade
- In direct benefit effects, the actions of others are affecting your payoffs directly, rather than indirectly by changing your information that is the case in information cascade

A simple herding experiment

- The experiment created by Anderson and Holt
- The experiment is designed to model:
  - There is a decision to be made, for example, whether to adopt a new technology, wear a new style of clothing, eat in a new restaurant, or support a particular political position
  - People make the decision sequentially, and each person can observe the choices made by those who acted earlier
  - Each person has some private information that helps guide their decision
  - A person can not directly observe the private information that other people **know**, but he or she can make inferences about this private information from what they **do**

A simple herding experiment

- We imagine the experiment taking place in a classroom, with a large group of students as participants.
- The experimenter puts an urn at the front of the room with three marbles hidden in it
- She announces that there is a 50% chance that the urn contains two red marbles and one blue marble, and a 50% chance the urn contains two blue marbles and one red marble.
- In the former case, we will say that it is a **majority-red** urn, and in the latter case, we will say that it is a **majority-blue** urn.
- Now, one by one, each student comes to the front of the room and draws a marble from the urn.

A simple herding experiment

- He looks at the color and then places it back in the urn without showing it to the rest of the class. The student then guesses whether the urn is majority-red or majority-blue and **publicly announces** this guess to the class.
- The public announcement is the key part of the set-up: the students who have not yet had their turn do not get to see which colors the earlier students draw, but they do get to hear the guesses that are being made.

A simple herding experiment

- The First Student:
  - The first student should follow a simple decision rule for making a guess: if she sees a red marble, it is better to guess that the urn is majority-red; and if she sees a blue marble, it is better to guess that the urn is majority-blue
- The Second Student:
  - If the second student sees the same color that the first student announced, then her choice is simple: she should guess this color as well
  - If not, he should rely on his own observation
- The Third Student:
  - Things start to get interesting here. If the first two students have guessed opposite colors, then the third student should just guess the color she sees, since it will effectively **break the tie** between the first two guesses.

A simple herding experiment

- The Fourth Student and Onward:
  - …
- If we want to build a mathematical model for how information cascades occur, it will necessarily involve people asking themselves questions like, “What is the probability this is the better restaurant, given the reviews I've read and the crowds I see in each one?”
- We can use the **Bayes' conditional probability Rule** for modeling “Decision-Making Under Uncertainty”

$$\Pr\left[A \mid B\right] = \frac{\Pr\left[A\right] \cdot \Pr\left[B \mid A\right]}{\Pr\left[B\right]}$$

Bayes's rule in herding experiment

- Consider the previous example of the students
- Notice that each student's decision is intrinsically based on determining a conditional probability:
  - Each student is trying to estimate the conditional probability that the urn is majority-blue or majority-red, given what she has seen and heard
- To maximize her chance of winning the monetary reward for guessing correctly, she should guess majority-blue if

$$\Pr\left[\text{majority-blue} \mid \text{what she has seen and heard}\right] > \frac{1}{2}$$

- and guess majority-red otherwise.
- If the two conditional probabilities are both exactly 0.5, then it doesn't matter what she guesses.

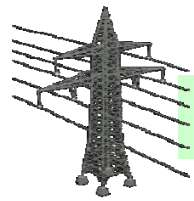
Cascading reactions

- The phenomenon of cascading reactions, in particular cascading failures, is quit similar to virus spreading over various complex networks (we will see the spreading in coming lectures)
- **Similarity:** they typically lead to collapse of a large portion of the network or even the entire network, called avalanche.
- **Differences:** as long as the network falls apart into pieces, the network is considered failed, while in virus spreading infected individuals are the ultimate concerns. Also, in cascading failures, loads (weights) on nodes or edges are taken into consideration, but usually not so in virus spreading.

Real examples of cascading failure

Many real networks show cascading effect:

Internet



Power grids

October 1986:
the first documented Internet congestion collapse

↓

Drop in speed of a factor 100

August 1996:
sag of just one electrical line in Oregon

↓

Blackout for 4 million people in 9 different States

August 2003:
initial disturbance in Ohio

↓

Largest blackout in the US history

Cascading failures

Assumptions and Definitions

1. Each node exchanges information with all the others, using shortest paths.

2. Load of node i at time t: $\mathbf{LOAD_i(t)}$ = $\text{Betweenness}_i(t)$
   (Goh, Kahng, Kim, PRL 87 (2002) 278701)

   total number of shortest paths passing through i

3. Capacity of node i: $\mathbf{C_i} = \alpha \, \text{LOAD}_i(0)$ $\quad \forall i : i \in G$

   where $\alpha>1$ is the tolerance parameter of nodes
   (Motter and Lai, PRE 66 (2002) 065102)

Dynamics of the model

Crucitti, Latora, Marchiori, cond-mat/0309141 (2003)

Just 1 node is removed

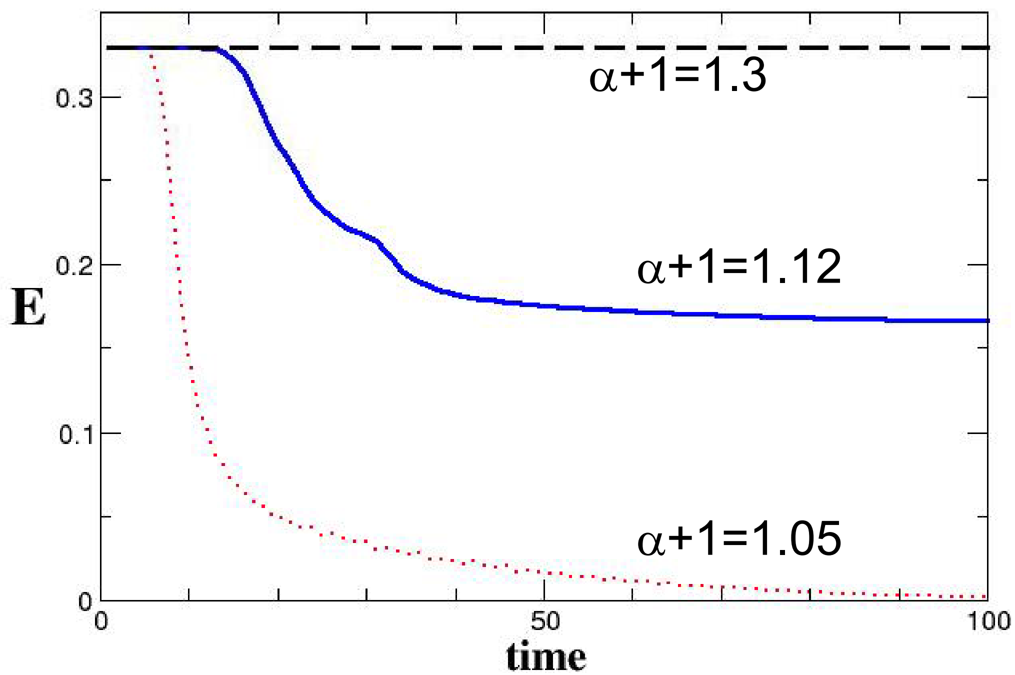
- a random node (<u>error</u>)
- a node with high betweenness (<u>attack</u>).

Load redistribution

Weights Update

$$\begin{cases} \tau_{ij}(t+1) = \tau_{ij}(0) \cdot \dfrac{LOAD_i(t)}{C_i} & if\ \ LOAD_i(t) > C_i \\ \tau_{ij}(t+1) = \tau_{ij}(0) & if\ \ LOAD_i(t) \le C_i \end{cases} \qquad \forall j : j \in G_i$$

Time evolution of efficiency



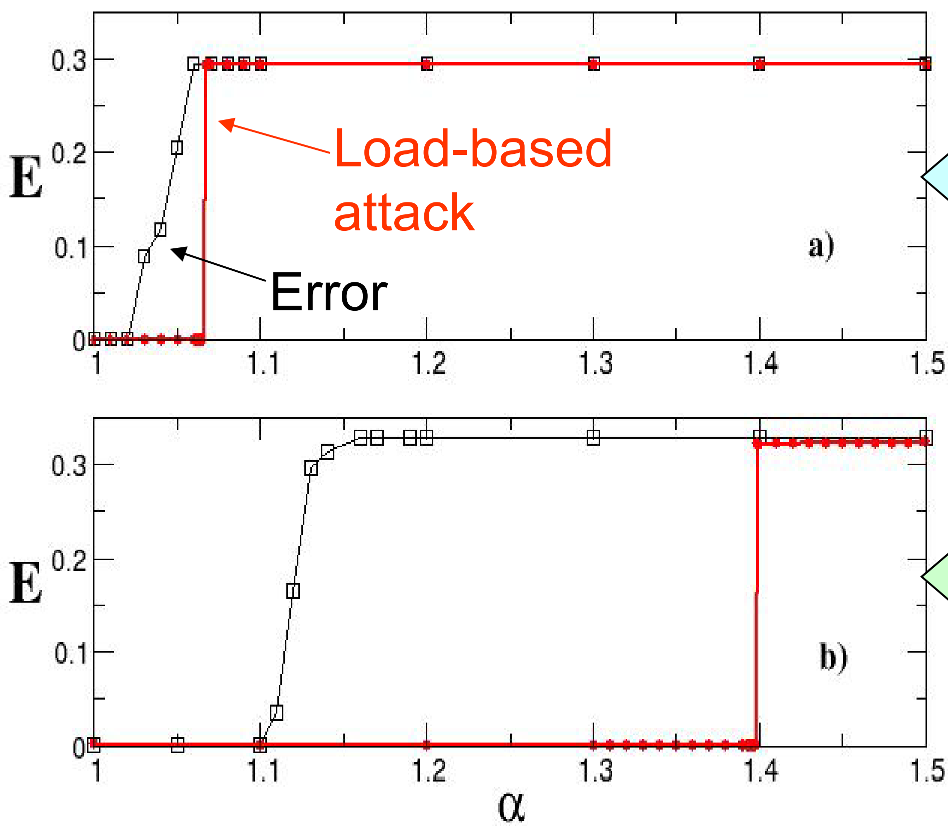
BA scale-free with a random removal for 3 different value of the tolerance parameter

Source: Crucitti et al, 2003

Time evolution of efficiency

Results: ER and BA models

N=2000



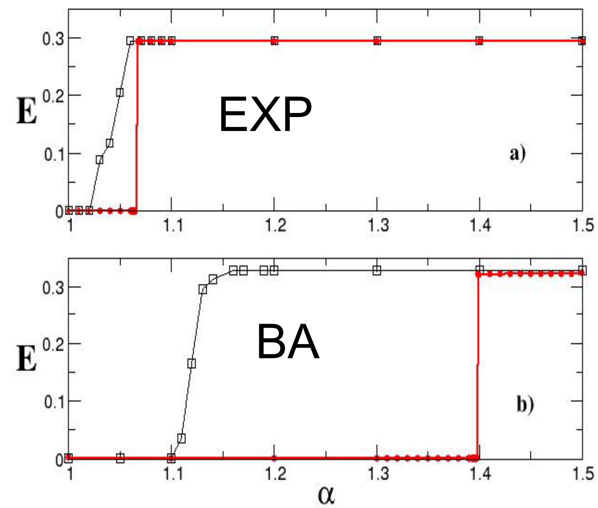
<u>Erdös-Rényi Random graph</u>
(Homogeneous)

<u>Scale-Free (BA model)</u>
(Heterogeneous)

Source: Crucitti et al, 2003

So,

Differences



1. Homogeneous networks (EXP) are more resistant to cascading failures

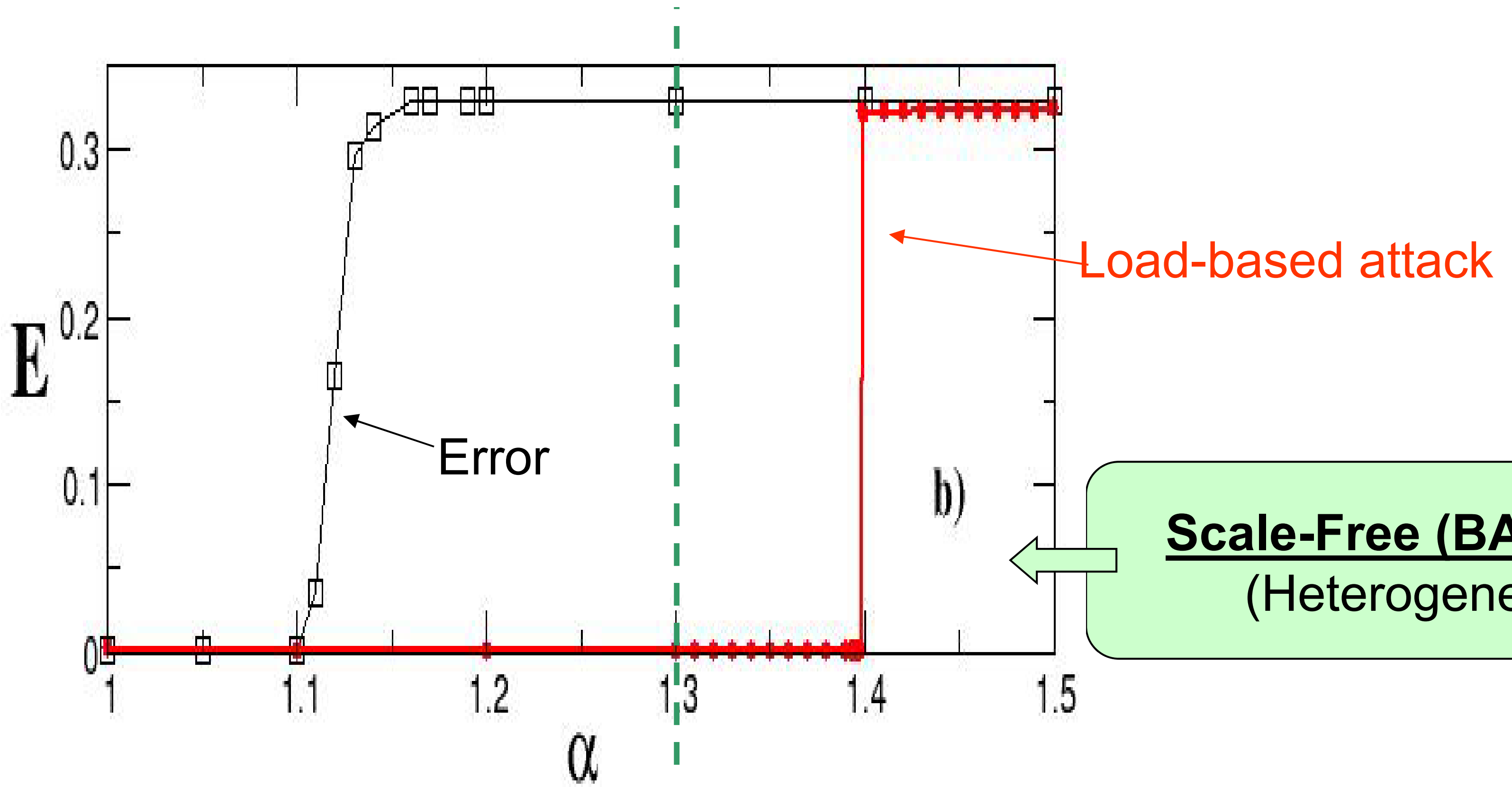
2. Region of α where the system { is stable against errors / collapses under attacks }

is wide for heterogeneous networks (BA).

Source: Crucitti et al, 2003

So,

Results: BA model



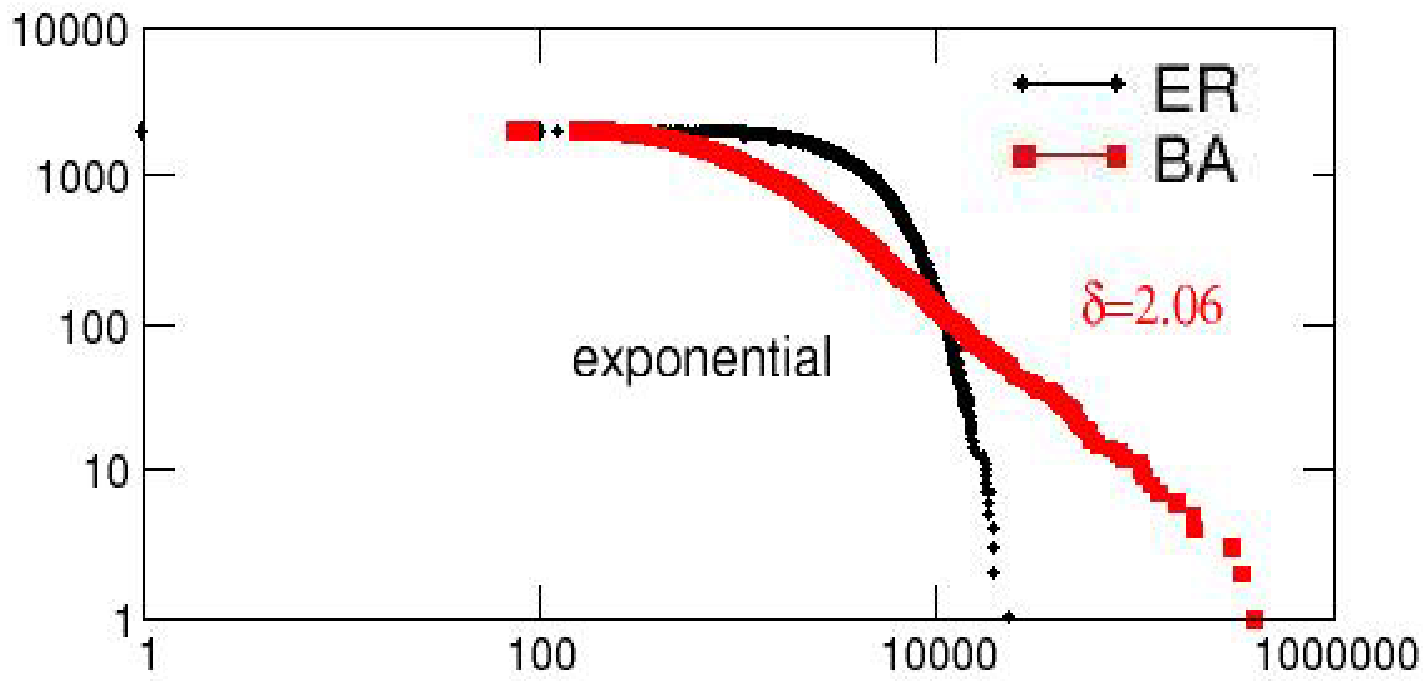
<u>Scale-Free (BA model)</u>
(Heterogeneous)

Source: Crucitti et al, 2003

So,

A glimpse of distributions: ER and BA models

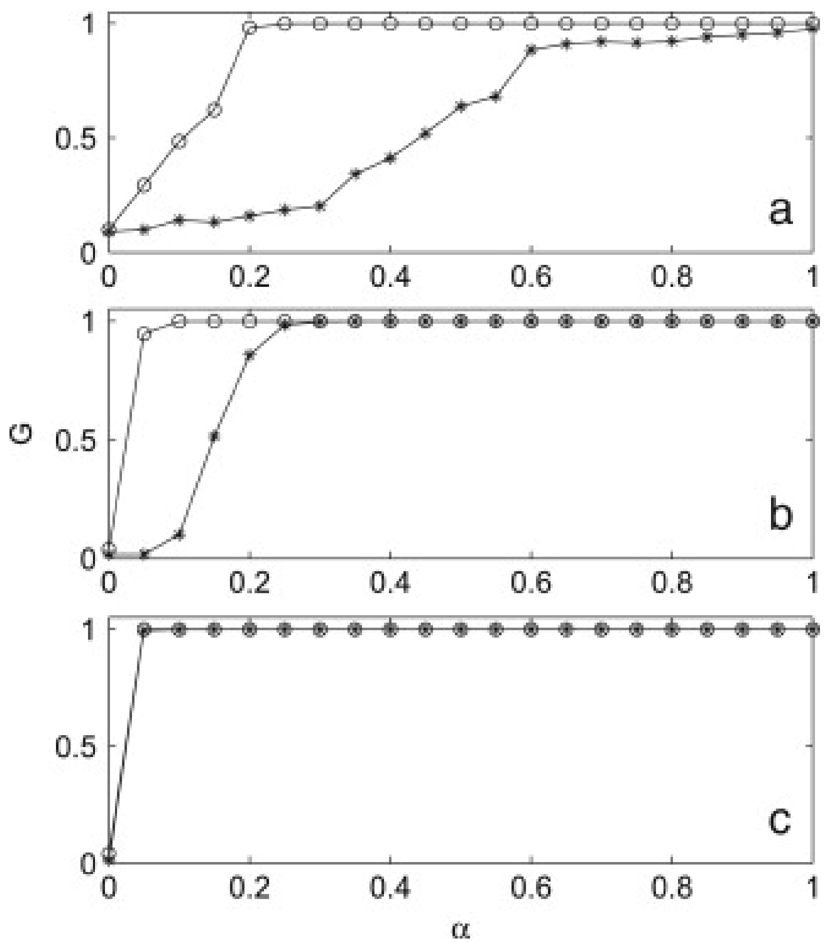
P(LOAD)



LOAD

Source: Crucitti et al, 2003

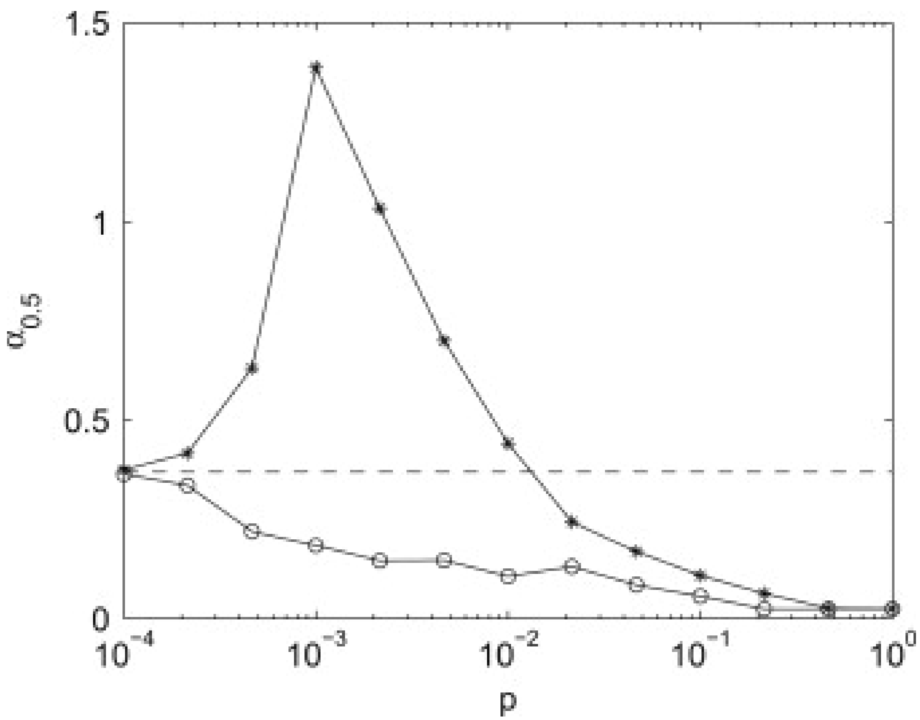
Cascading failure in Watts-Strogatz networks



Size of the giant component as a function of the tolerance parameter in the model for cascading failures in different network models.
(a) Watts-Strogatz network with P = 0.01;
(b) Barabasi-Albert scale-free network;
(c) Erdos-Renyi random graph.
The curve with asterisks is the result under intentional attack, and the curve with circles is the result under random error.

Source: Xia et al, Physica A, 2010

Cascading failure in Watts-Strogatz networks



$\alpha_{0.5}$, defined as the value of α when the size of the giant component in half of the network size N.

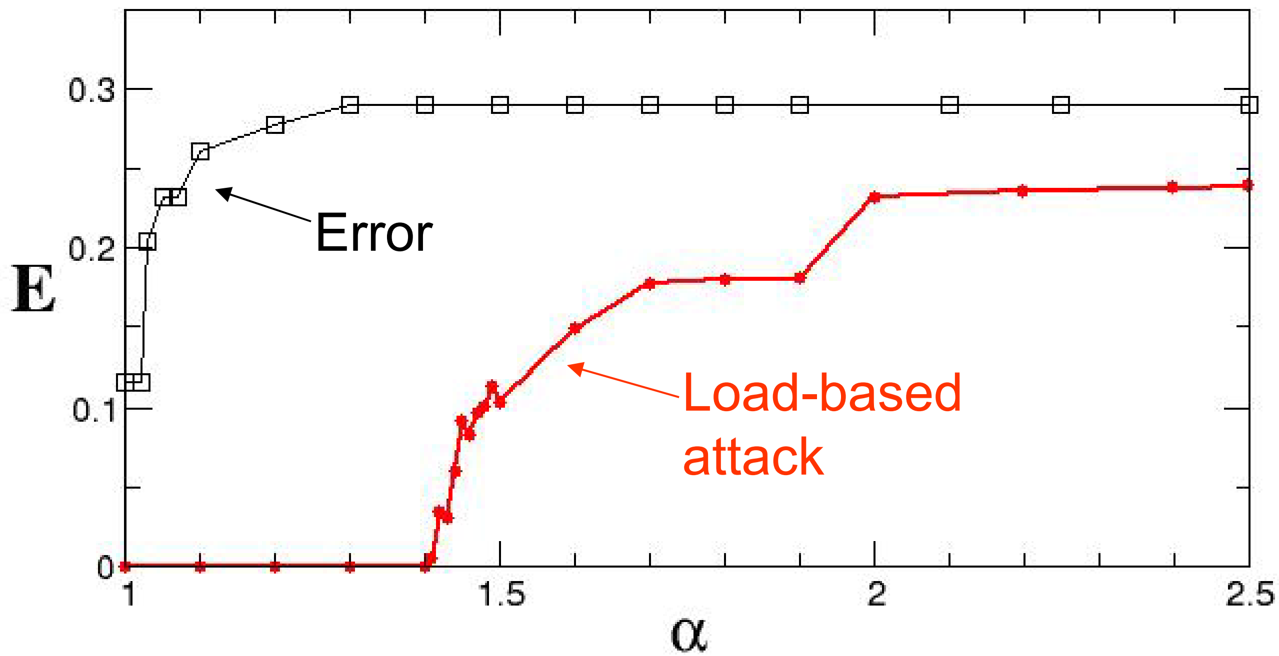
The networks are Wattas-Strogatz networks with different rewiring probability P.

The curve with asterisks is the result under intentional attack, and the curve with circles is the result under random error.

The dashed line is the result for a regular lattice ring (i.e., P = 0).

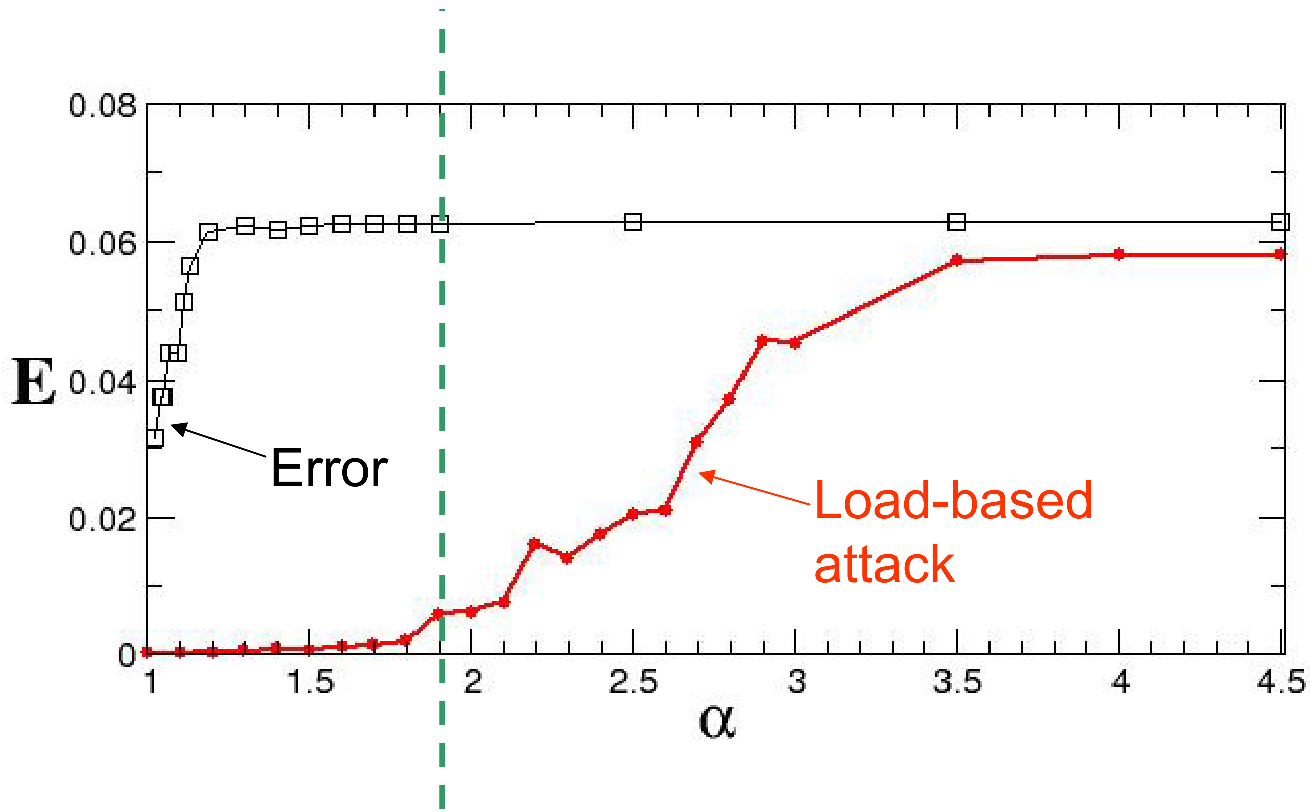
Source: Xia et al, Physica A, 2010

Real networks: the Internet



Source: Crucitti et al, 2003

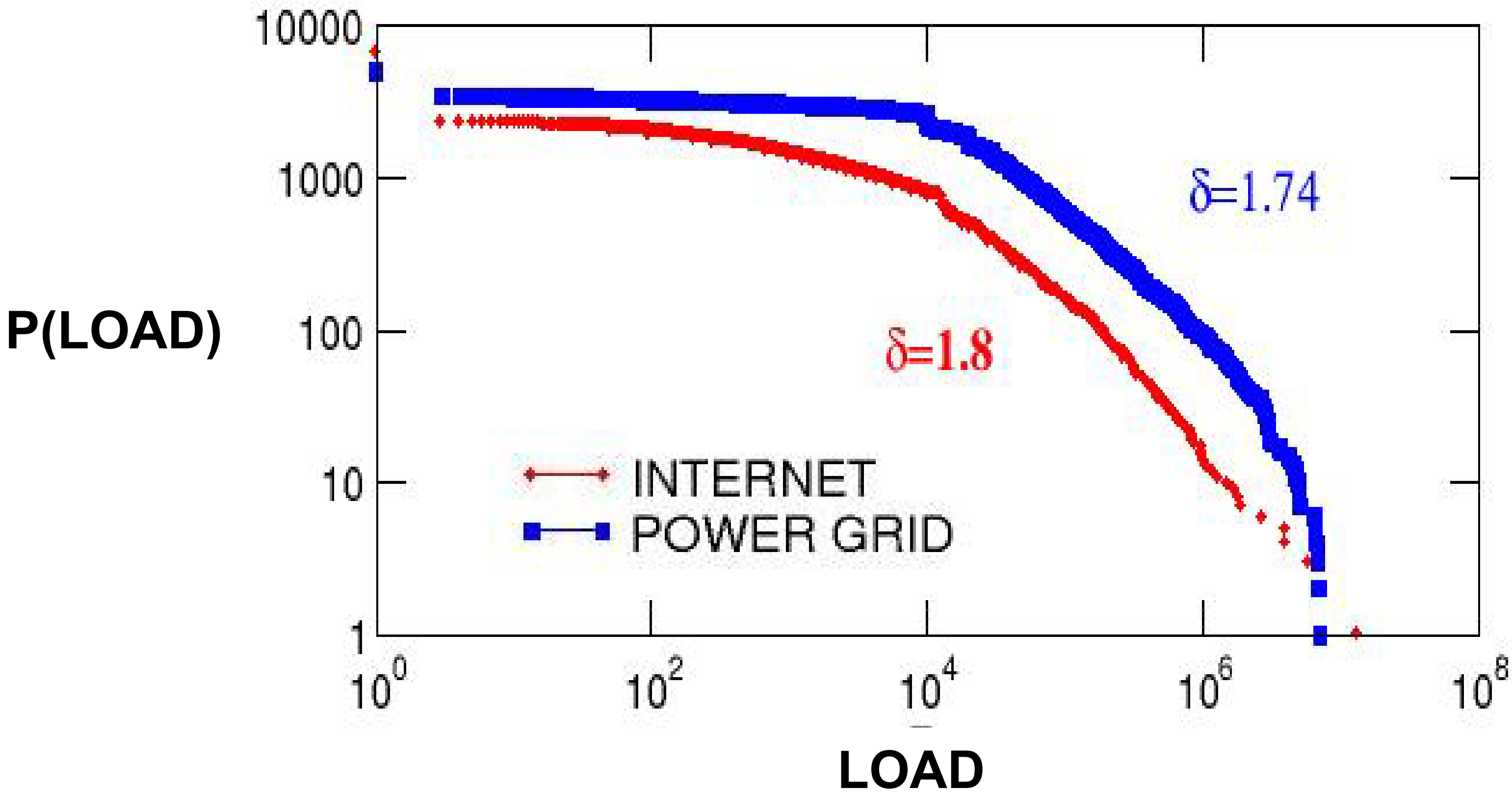
Real networks: US power grid



Source: Crucitti et al, 2003

And,

A glimpse of distributions: Internet and US Power Grid



Source: Crucitti et al, 2003

So,

Simple model → Dynamical redistribution of load

↓

large but rare cascading effects of real networks.

Most failures emerge and dissolve locally

A few failures spread over the whole network through an avalanche mechanism.

Cascading Failure Tolerance of Modular Small-World Networks

- Many real-world networks have modular structure and their components may undergo random errors and/or intentional attacks

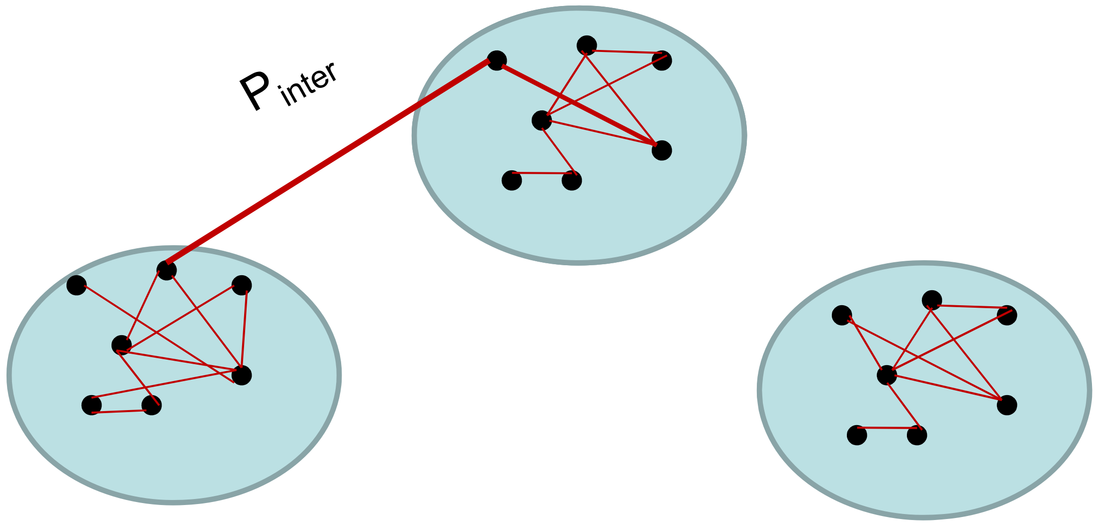
- Many real-world systems show modular structure
  - dense intra-modular connections
  - Sparse inter-modular connections

- We consider modular Watts-Strogatz networks & Barabasi-Albert networks
  - each module has WS network structure
  - the nodes in each module communicate with the nodes of other models with a uniform inter-modular probability

Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

Modular Networks

- Average degree is constant

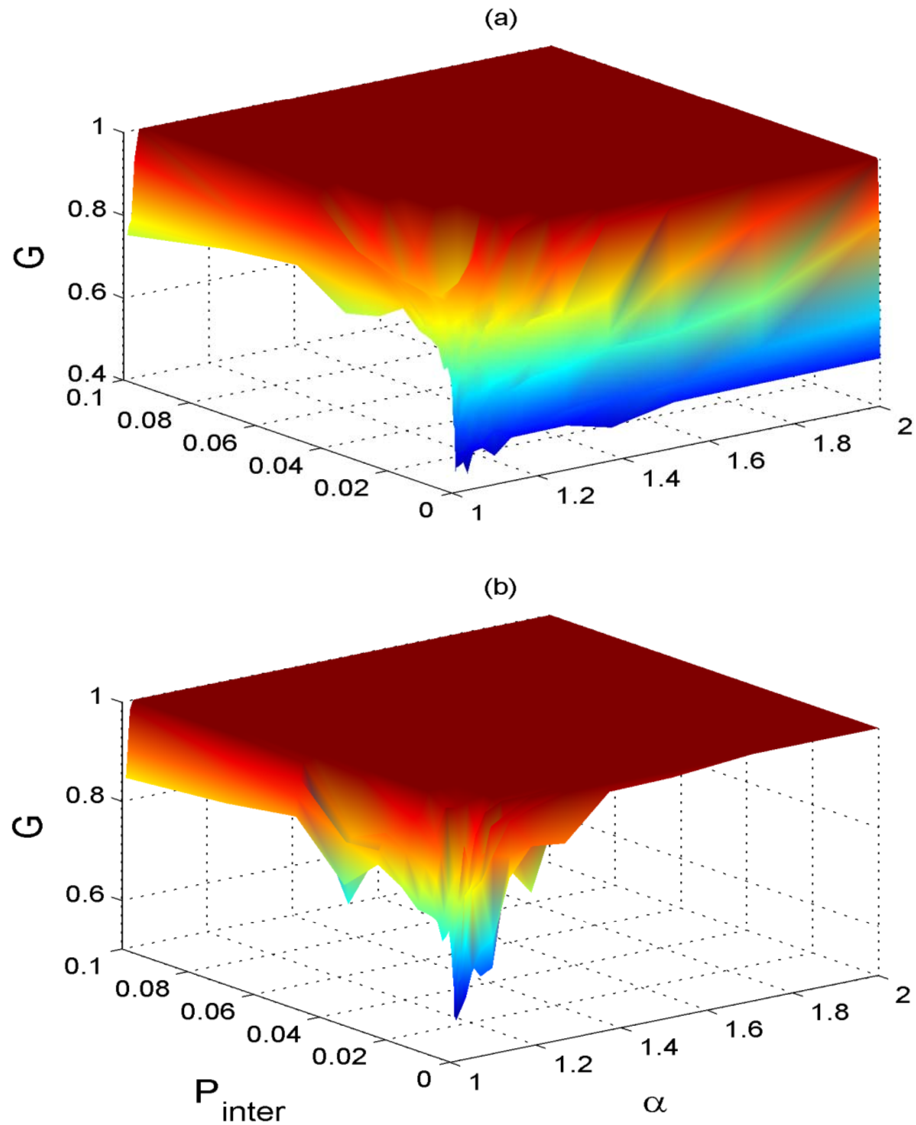


Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

Cascading Failure Tolerance of Modular Small-World Networks

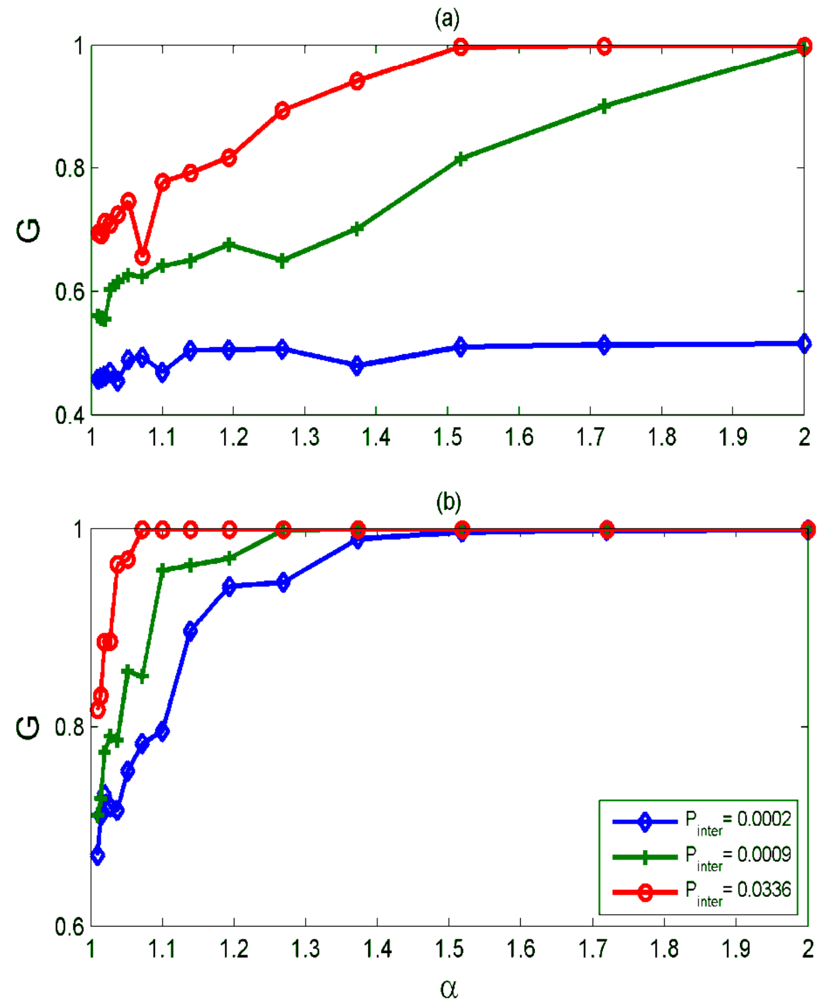
The normalized size of the largest connected component *G* as a function of the tolerance parameter α and the inter-modular connection P_{inter} for modular small-world networks.
The modular networks have 5 modules with 150, 200, 250, 300, and 350 nodes and 300, 400, 500, 600, and 700 edges, respectively. The rewiring probability of Watts-Strogatz model in each module is fixed at $P = 0.2$. The results are
(a) for cascading intentional attack and
(b) cascading random error

Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011



Cascading Failure Tolerance of Modular Small-World Networks

- The normalized size of the largest connected component *G* as a function of the tolerance parameter α and small/large values of P_{inter}. The network parameters are as Figure 1 (a) for cascading intentional attack and (b) cascading random error



Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

Cascading Failure Tolerance of Modular Small-World Networks

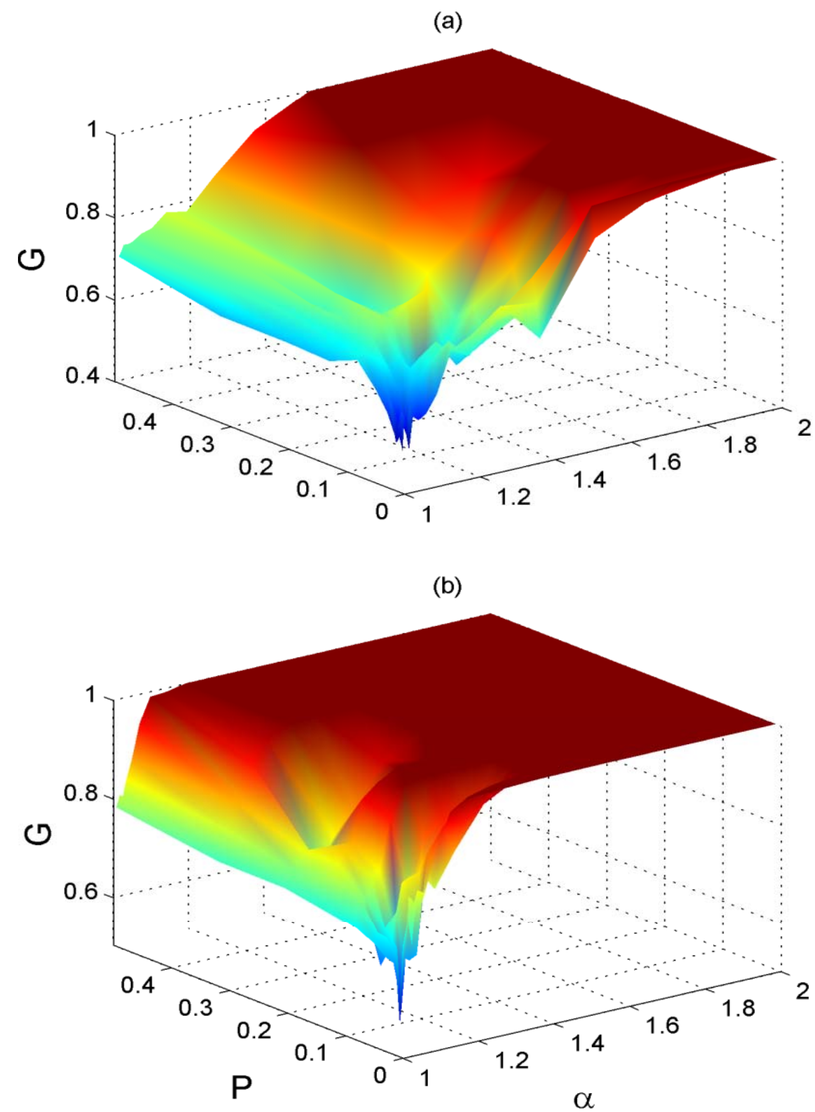
- The normalized size of the largest connected component G as a function of the tolerance parameter α and the rewiring probabilities P of the Watts-Strogatz model.

The network parameters are as Figure 1, and the inter-modular connection Probability is fixed at $P_{inter} = 0.002$.

The results are

(a) for cascading intentional attack and

(b) cascading random error



Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

Cascading Failure Tolerance of Modular Small-World Networks

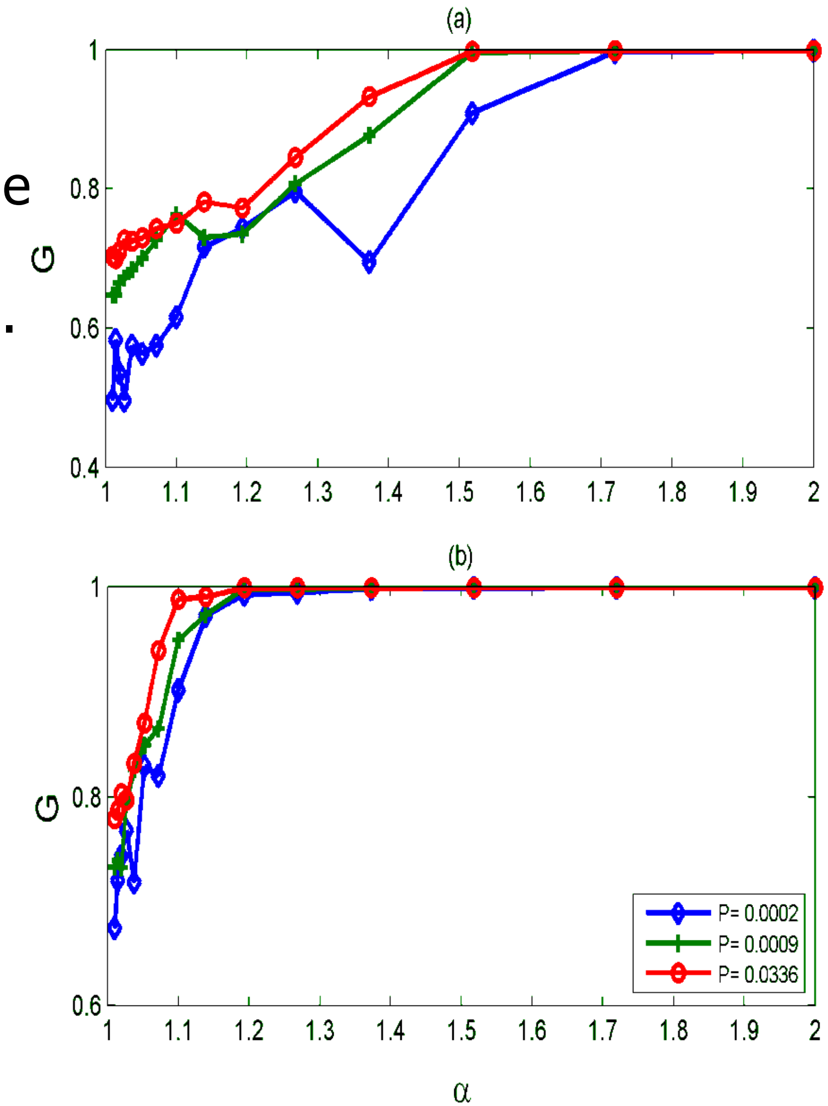
- The normalized size of the largest connected component *G* as a function of the tolerance parameter α and small/large values of *P*.

  The network parameters are as Figure 3.

  The results are

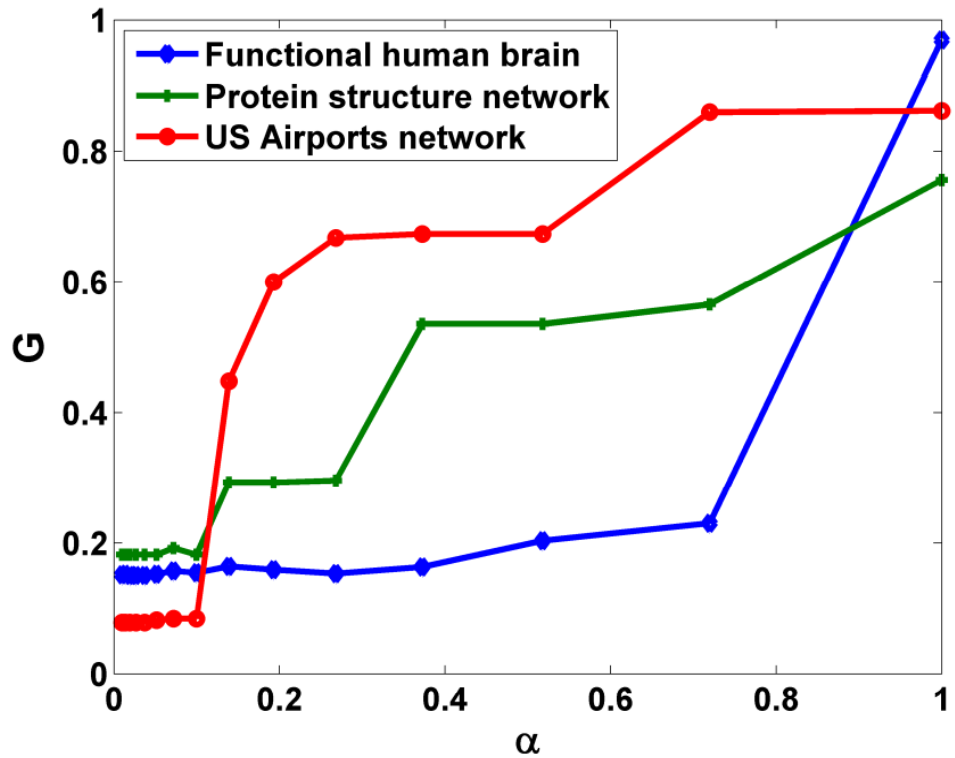
(a) for cascading intentional attack and

(b) cascading random error



Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

Cascading Failure Tolerance of Modular Real Networks



The normalized size of the largest connected component *G* as a function of the tolerance parameter α for modular real networks

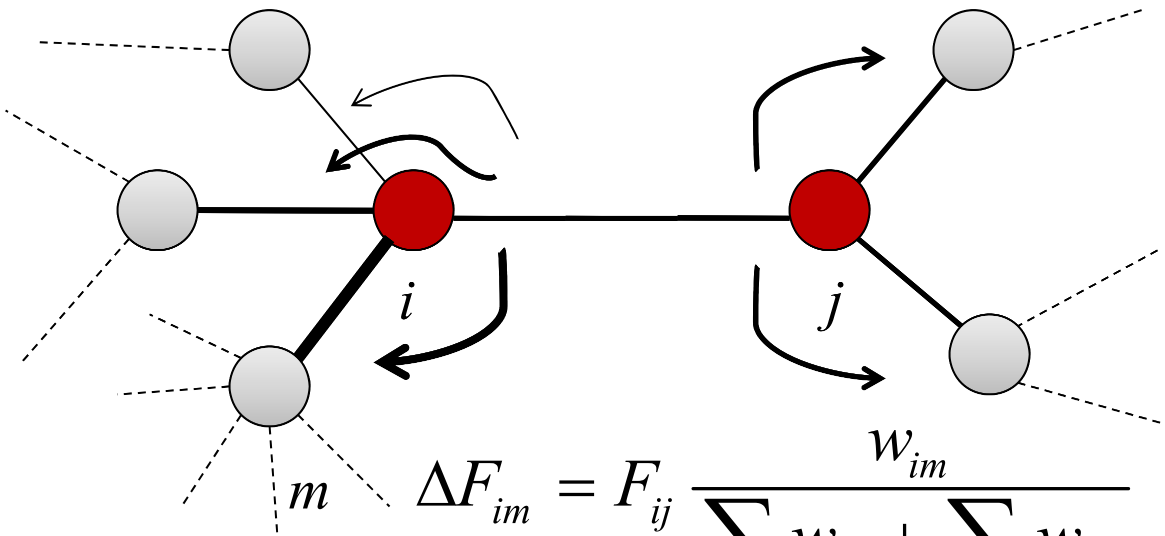
Source: Babaie, Ghassemyieh, and Jalili, IEEE Transactions CASII, 2011

So, …

- P_{inter} and P do not have the same influence on the structural parameters on the network
- same influence on the network robustness against cascading random errors and intentional attacks
- the influence of P_{inter} was more pronounced as compared to P
- As a general observation, the robustness of the modular networks against random errors was better than that of intentional attacks.

Cascaded Failure in Weighted Networks

- In technological networks, it is natural that the load passing through a failed component is redistributed among its neighboring components
- Local Weighted Flow Redistribution Rule (LWFRR)

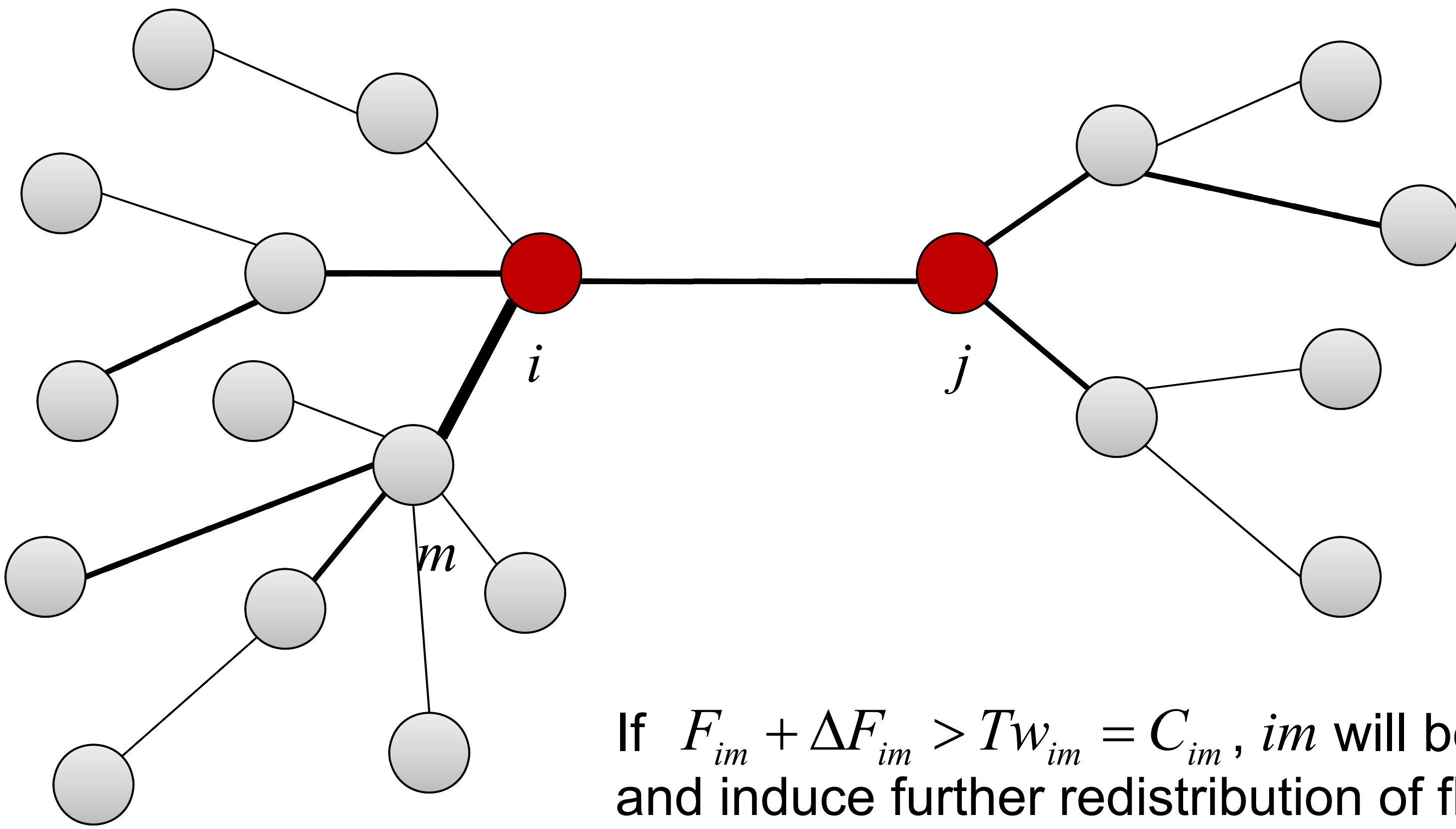


$$\Delta F_{im} = F_{ij} \frac{w_{im}}{\sum_{a \in \Gamma_i} w_{ia} + \sum_{b \in \Gamma_j} w_{jb}}$$

The additional flow ΔF_{im} received by edge im is proportional to its weight

Cascaded Failure in Weighted Networks

- Local Weighted Flow Redistribution Rule (LWFRR)



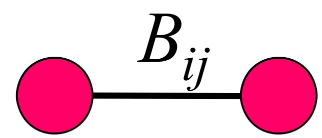
If $F_{im} + \Delta F_{im} > Tw_{im} = C_{im}$, im will be broken and induce further redistribution of flow and potentially further edge breaking

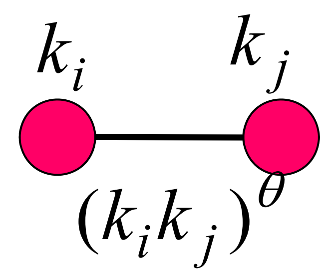
Cascaded Failure in Weighted Networks

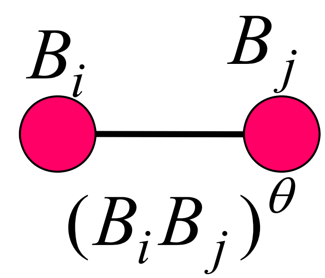
- A phase transition occurs at the critical threshold T_c, which can be used as a measure of the robustness of the network against cascading failure.
- When $T > T_c$, no cascading failure occurs and the system maintains its normal and efficient functioning
- For $T < T_c$, S_N (the number of failed components) suddenly increases from 0 and cascading failure emerges, causing the whole or part of the network to stop working.
- Hence T_c is the least value of protection strength to avoid cascading failure

Cascaded Failure in Weighted Networks

- Weighting Methods
  - Weighting method based on edge betweenness centrality
  - Weighting method based on a function of degrees of the two ends of an edge
  - Weighting method based on a function of betweenness centrality of the two ends of an edge



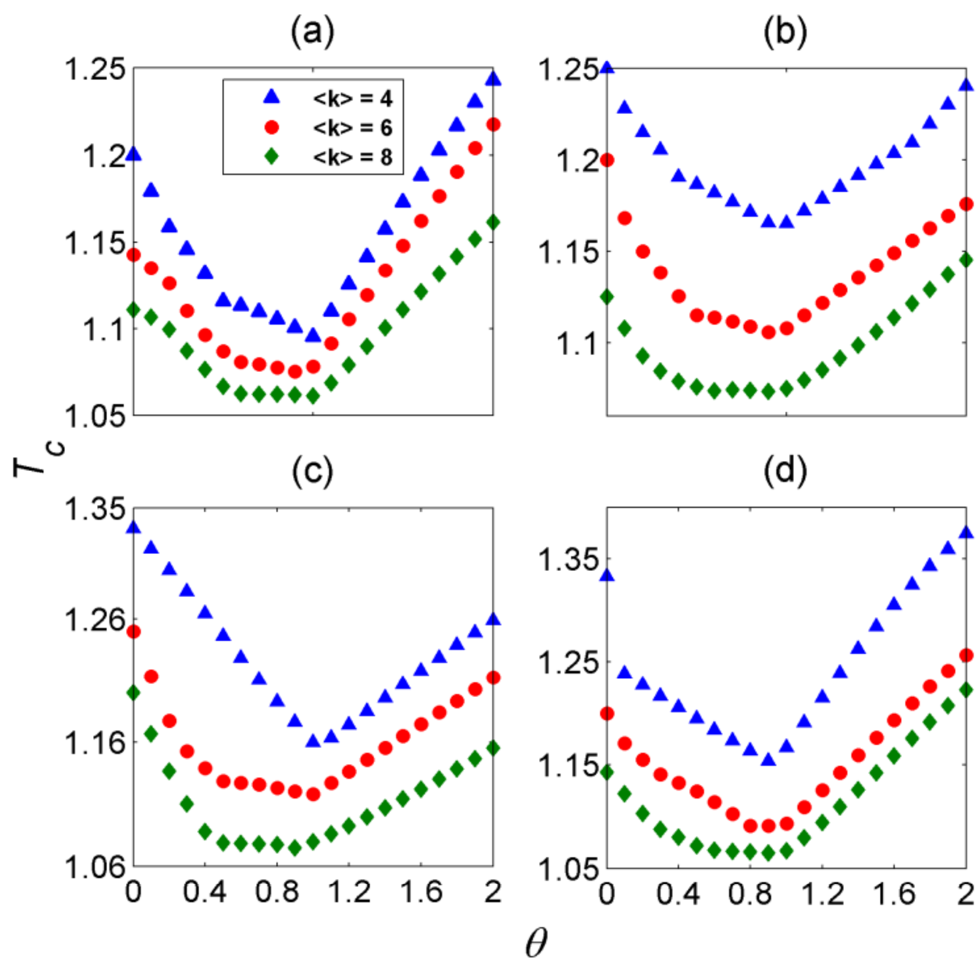




Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

Cascaded Failure in Weighted Networks

The critical threshold T_c as a function of θ for different average degrees <k> on (a), scale-free networks with 1000 nodes (b), Newman-Watts networks with 1000 nodes and p = 0.3 (c) Erdös-Rényi network with 1000 nodes (d) modular networks that has three modules with 200, 300, and 500 nodes. Data shows averages over 10 realizations.



Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

Cascaded Failure in Weighted Networks

- The weighting based on the product of the betweenness centrality of the end nodes resulted in the least critical threshold for these networks.

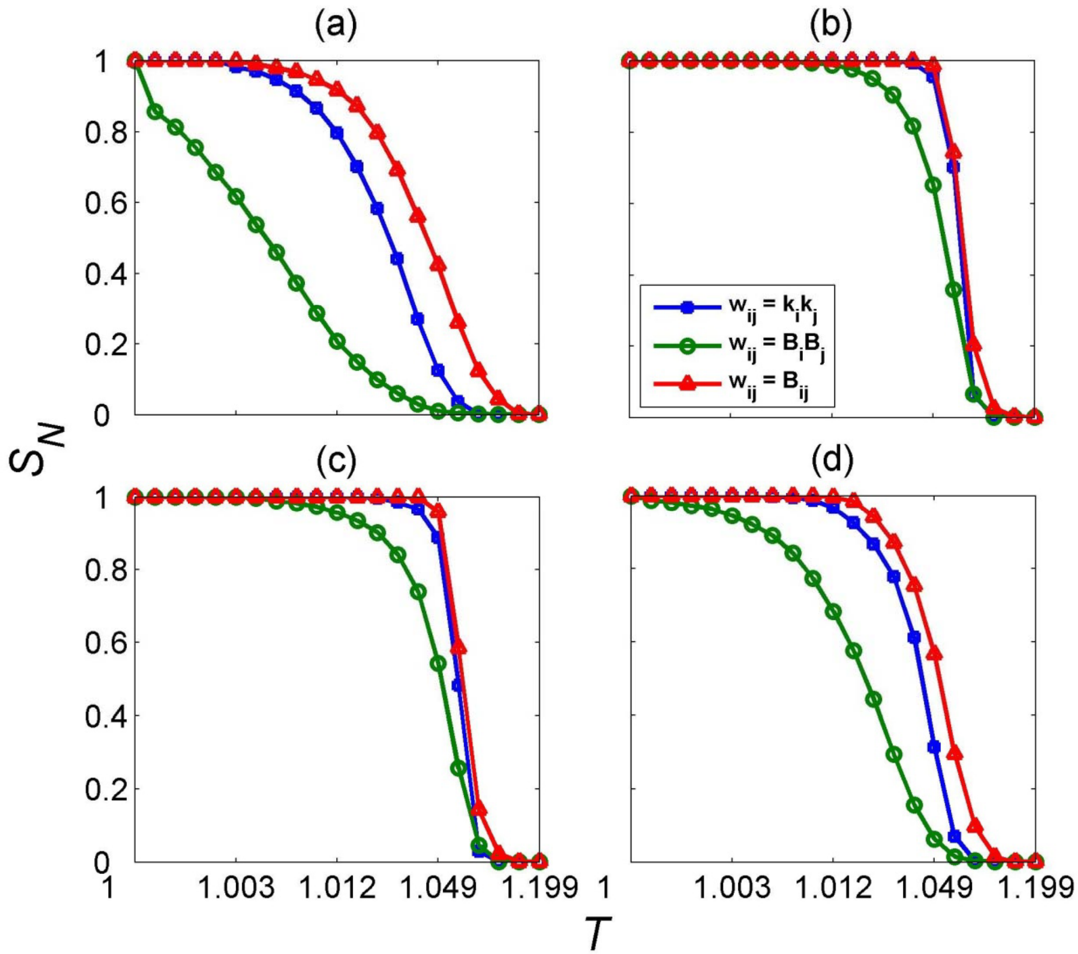
Network	$w_{ij} = B_i B_j$	$w_{ij} = B_{ij}$	$w_{ij} = k_i k_j$
Scale-free	1.091	1.142	1.109
Newman-Watts	1.108	1.174	1.112
Erdős–Rényi	1.117	1.223	1.129
Modular Scale-free	1.093	1.185	1.113

Critical threshold of SF networks with N=1000 nodes and m=3, NW networks with N=1000 nodes and k=3 and p=0.3, ER networks with N=1000 nodes and p=0.006 and modular network that has three modules with 200,300,500 nodes and 3000 edges. The results are shown at $\theta = 1$ for weighting strategies 1-3.

Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

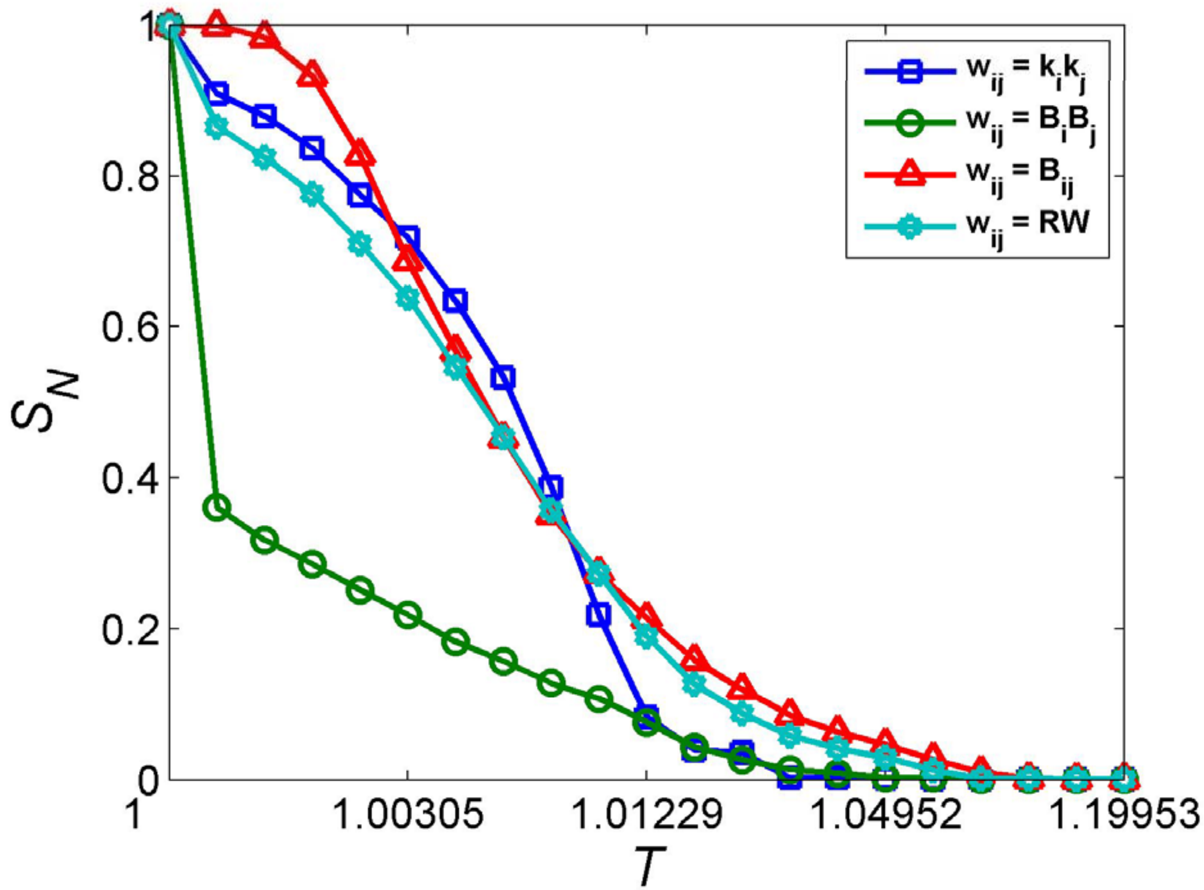
Cascaded Failure in Weighted Networks

Normalized average size of the removed edges (S_N) as a function of the threshold parameter (T) for (a) scale-free network with 1000 nodes and m = 3, (b) Newman-Watts network with 1000 nodes, k = 3 and p = 0.3 (c) Erdös-Rényi network with 1000 nodes and p = 0.006, (d) modular networks that has three modules with 200, 300, and 500 nodes and 3000 edges. The red triangle, blue square and green circle lines show the changes in the for weighing methods based on edge Betweenness centrality, degree multiplication, and multiplication of node betweenness centrality, respectively



Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

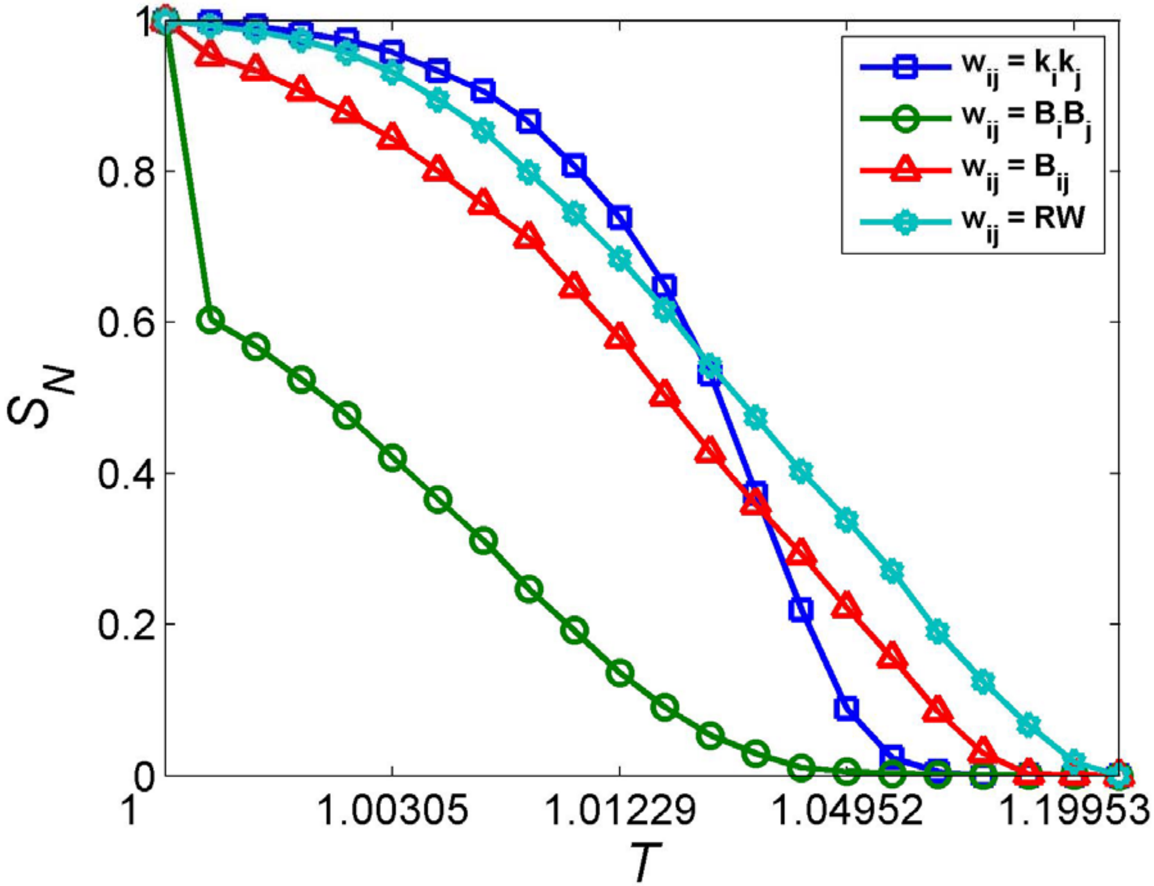
Cascaded Failure in Weighted Networks



S_N as a T for the US airports network with 500 nodes and 2980 edges. The cyan diamond line shows the changes in the for real weights (RW)

Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

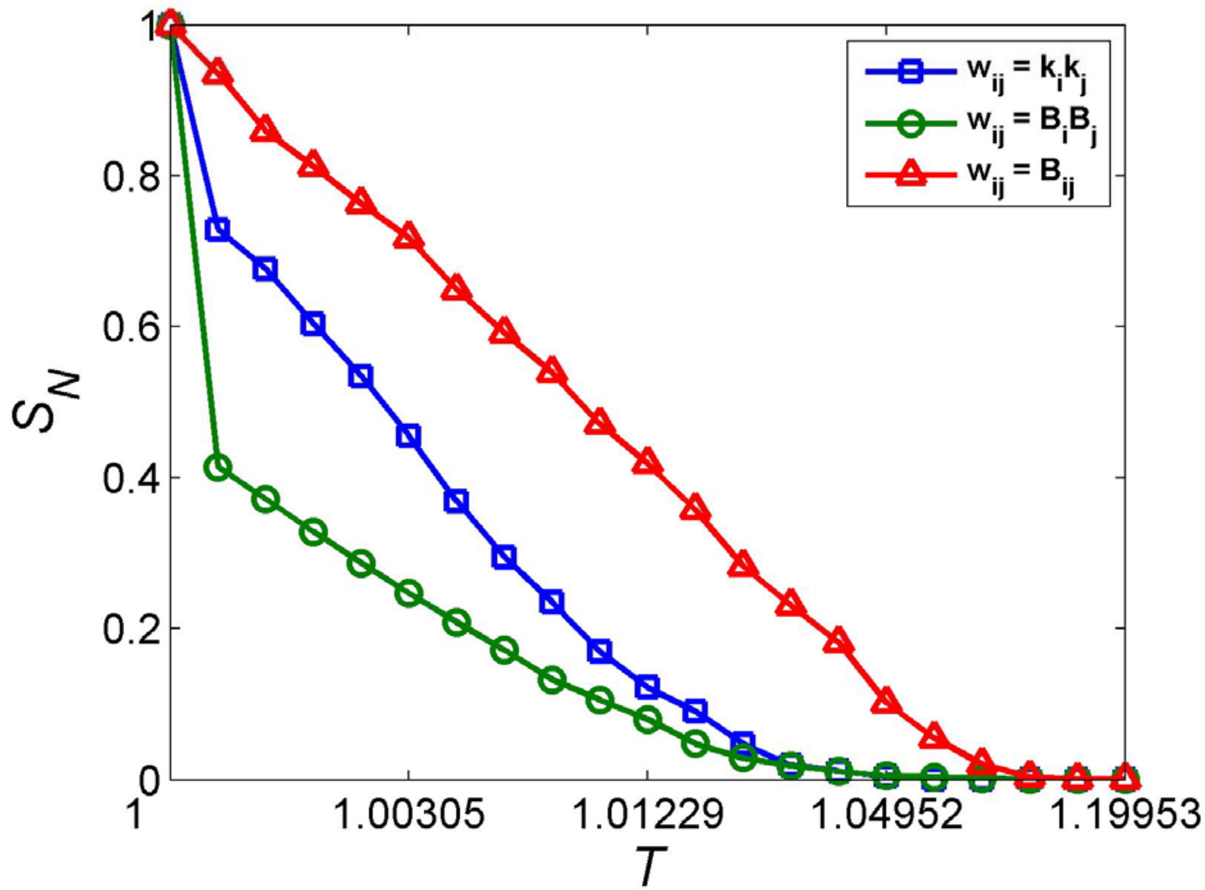
Cascaded Failure in Weighted Networks



S_N as a T for the central European rail network with 2488 nodes and 6691 edges

Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

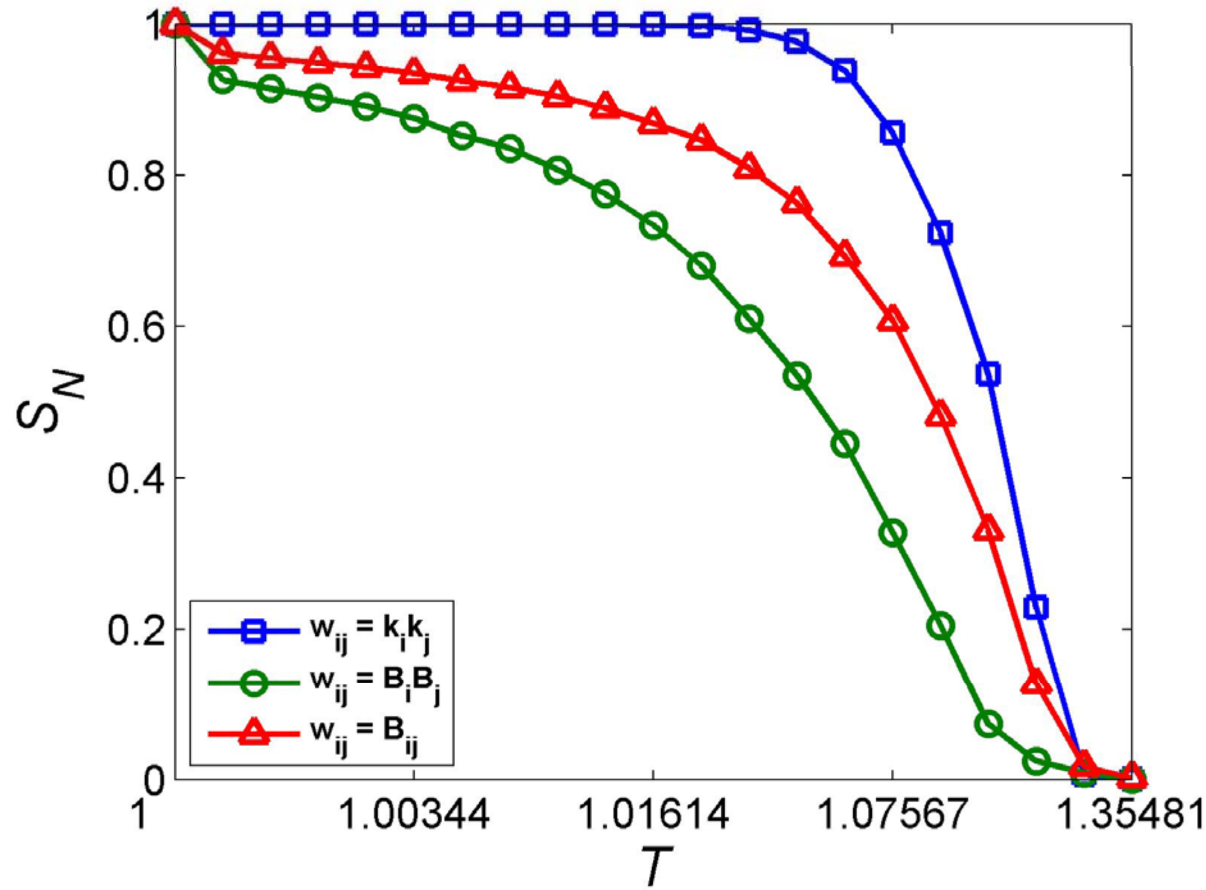
Cascaded Failure in Weighted Networks



S_N as a T for the network of the Internet in the level of autonomous system with 2062nodes and 4233 edges

Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

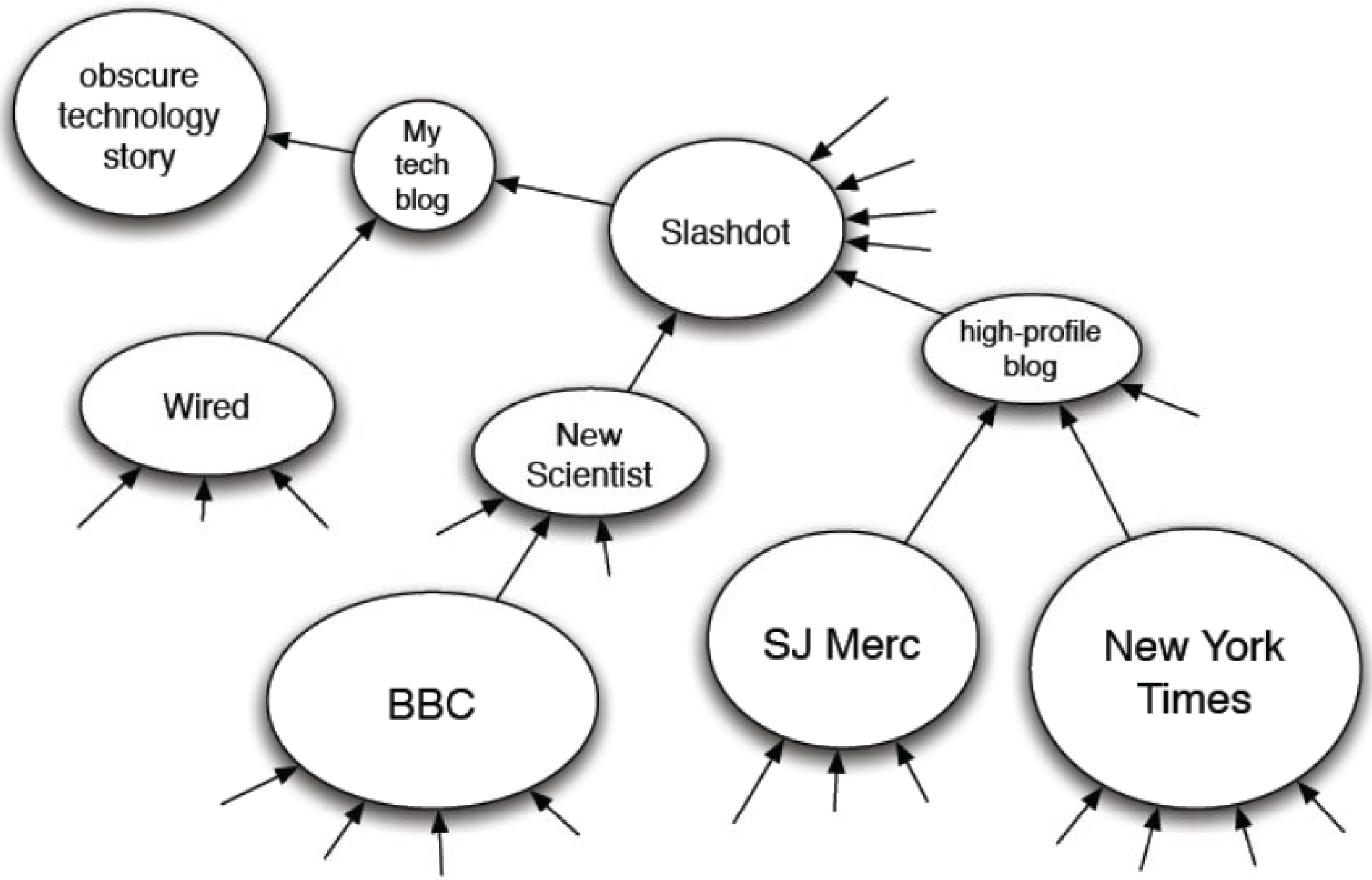
Cascaded Failure in Weighted Networks



S_N as a T for the Electrical power grid with 4941 nodes and 6549 edges

Source: Mirzasoleiman, Babaie, Jalili, and Safari, PRE, 2011

Information diffusion in networks



Diffusion in networks

- A fundamental process in social networks:

  Behaviors that cascade from node to node like an epidemic
- News, opinions, rumors, fads, urban legends, …
- Word-of-mouth effects in marketing: rise of new websites, free web-based services
- Virus, disease propagation
- Change in social priorities: smoking, recycling
- Saturation news coverage: topic diffusion among bloggers
- Internet-energized political campaigns
- Localized effects: riots, people walking out of a lecture, people sleeping in a lecture

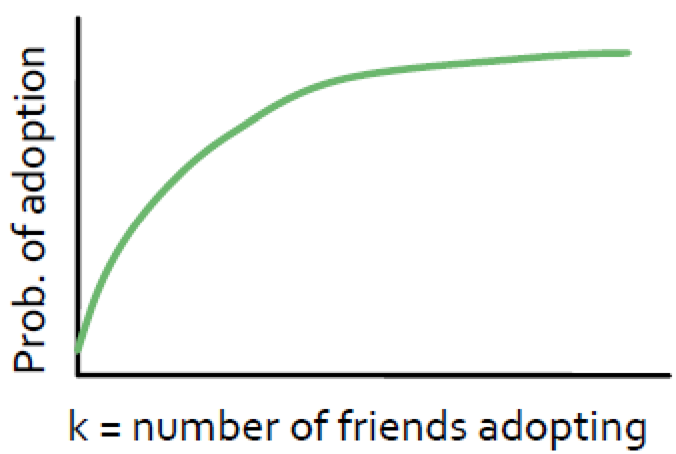
Diffusion in networks

- Many of our interactions with the rest of the world happen at a local, rather than a global, level
- There are two distinct kinds of reasons why imitating the behavior of others can be beneficial:
  - **Informational effects**, based on the fact that the choices made by others can provide indirect information about what they know
  - **Direct-benefit effects**, in which there are direct payoffs from copying the decisions of others, for example, payoffs that arise from using compatible technologies instead of incompatible ones

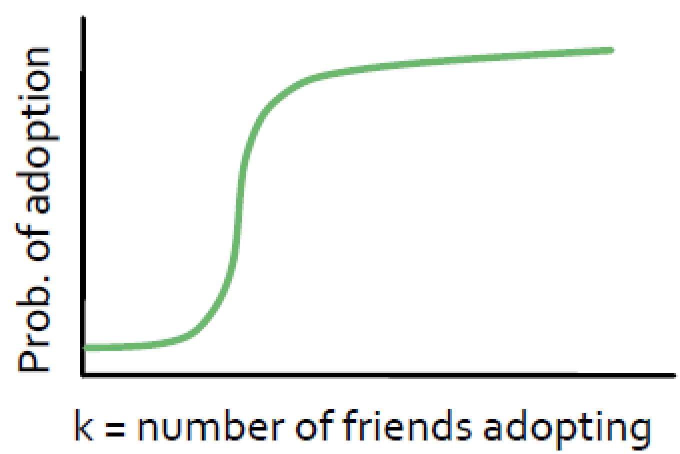
Diffusion curves

Basis for models:

- Probability of adopting new behavior depends on the number of friends who have adopted
- What's the dependence?

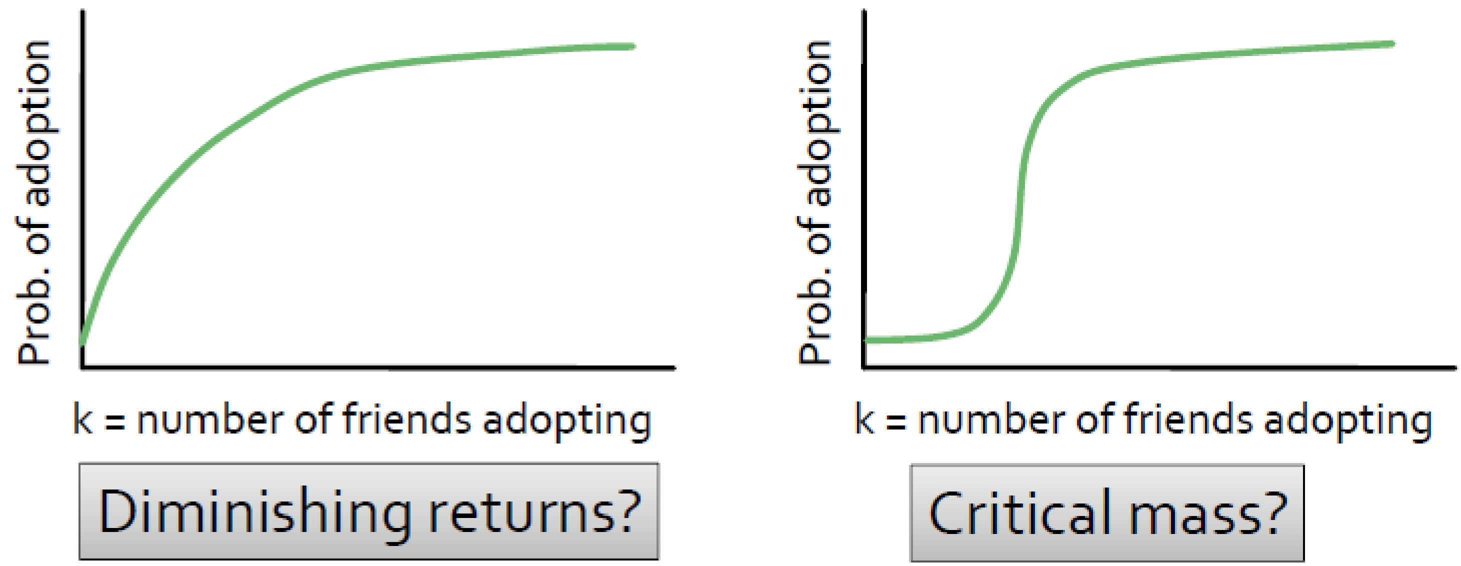


Diminishing returns?



Critical mass?

Diffusion curves



- Key issue: qualitative shape of diffusion curves
- Diminishing returns? Critical mass?
- Distinction has consequences for models of diffusion at population level

Modeling diffusion

- There are, in general, two methodologies to model diffusion
- **Decision-based models:**
  - Adopt new behaviors if k of your friends do
- **Probabilistic models:**
  - “catch” a disease with some probability from neighbors in the network

Diffusion of innovation

- How new behaviors, practices, opinions, conventions, and technologies spread from person to person through a social network, as people influence their friends to adopt new ideas?
- The works in this field are known as **diffusion of innovations** in sociology
- Some of the works in person-to-person influence was due to informational effects:
  - People observed that the decisions of their neighbors provided indirect information that led them to try the innovation as well
- Other important studies are based on direct-benefit effects
  - The spread of technologies such as the telephone and e-mail is based on this fact

Modeling diffusion through a net

There are four methods for modeling diffusion through a network, in general:

- Linear Threshold Model
- Networked Coordination Games
- Independent Cascade Model
- Voter Model

Linear threshold model

- Let us consider N people
- Everyone can observe all actions
- Each person i has threshold t_i:
  - Node i will take part in the behavior iff at least t_i other people are already doing it:
  - Small t_i: early adopter
  - Large t_i: late adopter
- Easy to simulate:
  - Given $t_1, \ldots, t_n$
  - $F(x)$ = fraction of people whose threshold $\leq x$
  - $s(t)$ = fraction of participants at time t
  - Then:
  - $S(0) = 0$, $s(1) = F(s(o)) = F(0)$, $s(2) = F(s(1))$, …, $s(t+1) = F(s(t))$

Discontinuous transition

- Each threshold t_i is drawn independently from some distribution $F(x) = \Pr[\text{threshold} \leq x]$
- Suppose: Normal with mean = N/2, variance σ
- As we continuously vary the parameter σ, something discontinuous happens:
  - Bigger variance lets you build a bridge from early adopters to mainstream
  - But, if we increase the variance even further we move one higher fixed point lower

Personal threshold

- Personal threshold k:
- Node i takes part in a behavior if at least k other nodes (including i) also take part
- "I will show up to the protest if I am sure at least k people in total (including myself) will show up"
- Each node in the network knows the thresholds of all their friends

Weakness of the model

It does not take into account:

- No notion of social network – more influential users
- It matters who the early adopters are, not just how many
- Models people’s awareness of size of participation not just actual number of people participating
- Modeling thresholds
  - Richer distributions
  - Deriving thresholds from more basic assumptions
- Modeling perceptions of who is adopting the behavior/who you believe is adopting
- Non monotone behavior: dropping out if too many people adopt
- Similarity: thresholds not based only on numbers
- People get “locked in” to certain choice over a period of time

Networked coordination game

- Consider a social network
- Each node has a choice between two possible behaviors, labeled A and B
- If nodes v and w are linked by an edge, then there is an incentive for them to have their behaviors match
- Let us represent this using a game in which v and w are the players and A and B are the possible strategies
- The payoffs are defined as follows:
  - If v and w both adopt behavior A, they each get a payoff of a > 0
  - If they both adopt B, they each get a payoff of b > 0
  - If they adopt opposite behaviors, they each get a payoff of 0

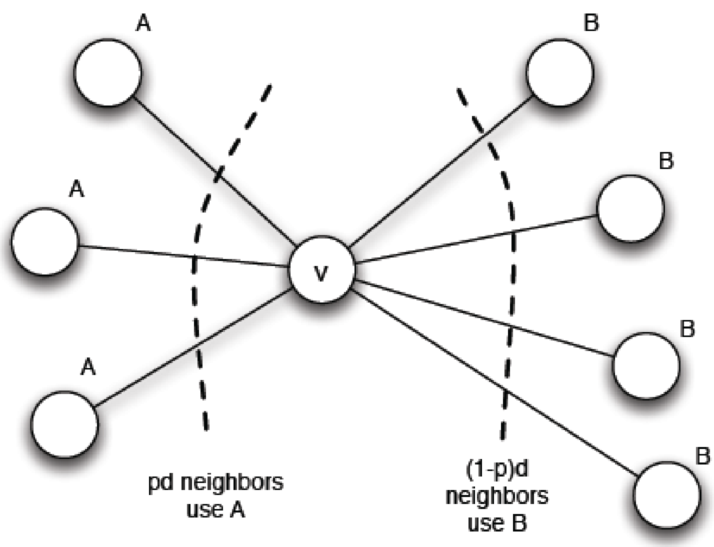
	A	B
A	a, a	$0, 0$
B	$0, 0$	b, b

Networked coordination game

- The payoff matrix describes what happens on a single edge of the network
- But, the point is that each node v is playing a copy of this game with each of its neighbors
- Its payoff is the sum of its payoffs in the games played on each edge
- Hence, v's choice of strategy will be based on the choices made by all of its neighbors, taken together.
- The basic question faced by v:
  - suppose that some of its neighbors adopt A, and some adopt B; what should v do in order to maximize its payoff?
- This depends on the relative number of neighbors doing each, and on the relation between values a and b.

Networked coordination game

- So, For node v,
- Assume v has d neighbors
- Fraction p of its neighbors adopt A, then Payoff of v = a.p.d, if v chooses A
- Fraction 1-p of the neighbors adopt B, then, if v chooses B, its payoff would be b.(1-p).d
- Behavior A is better choice if:



p.d.a ≥ (1-p).b.d

p ≥ b/(b+a) = q

Networked coordination game

- So, there is a simple threshold rule:
- If at least $q = b/(a+b)$ fraction of the neighbors follow behavior A, then you should too
- When q is small, then A is the much more attractive behavior, and it only takes a small fraction of your neighbors engaging in A for you to do so as well
- If q is large, then the opposite holds: B is the attractive behavior, and you need a lot of your friends to engage in A before you switch to A.
- There is a tie-breaking question when exactly q fraction of a node's neighbors follow A; in this case, we will adopt the convention that the node chooses A rather than B.

Cascading behavior in the game

- In any network, there are two obvious equilibria to this network:
  - one in which everyone adopts A
  - another in which everyone adopts B
- We want to understand how easy it is, in a given situation, to **tip** the network from one of these equilibria to the other.
- We also want to understand what other **intermediate** equilibria look like, states of coexistence where A is adopted in some parts of the network and B is adopted in others.

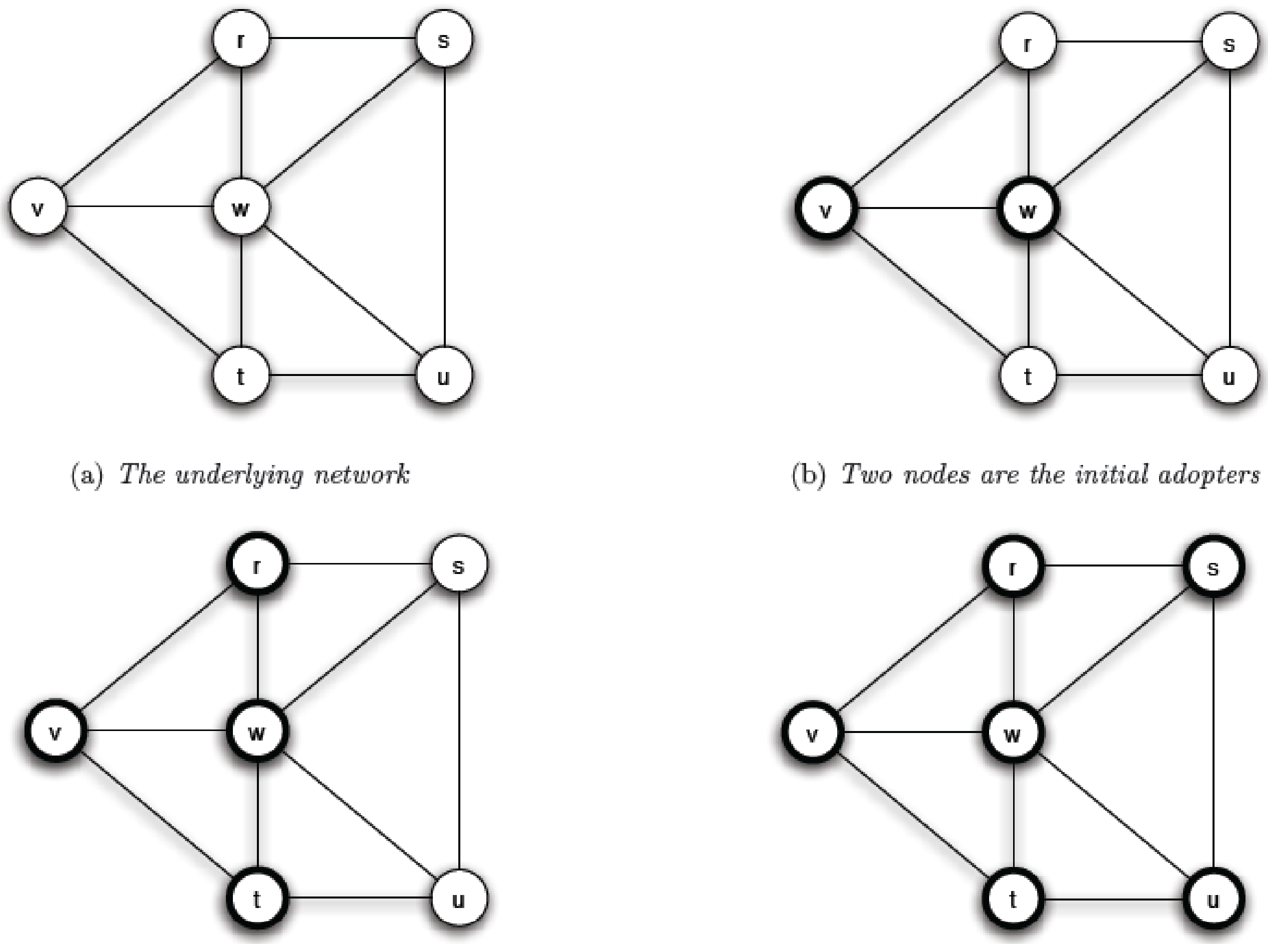
Cascading behavior in the game

- Suppose everyone is initially using B as default behavior
- Then, a small set of **initial adopters** decide to use A
  - Assume that the initial adopters have switched to A for some reason outside the definition of the game
- Given the fact that the initial adopters are now using A, some of their neighbors may decide to switch to A as well, and then some of their neighbors might, and so forth, in a potentially cascading fashion
- When everyone switches to A?
- The answer will depend on
  - the network structure
  - the choice of initial adopters
  - the value of the threshold q

Cascading behavior in the game

- So, the models is described as:
  - An initial set of nodes adopts A while everyone else adopts B
  - Time then runs forward in unit steps
  - In each step, each node uses the threshold rule to decide whether to switch from B to A
  - The process stops either when every node has switched to A, or when we reach a step where no node wants to switch
  - At this point things have stabilized on coexistence between A and B

Example



(a) *The underlying network*

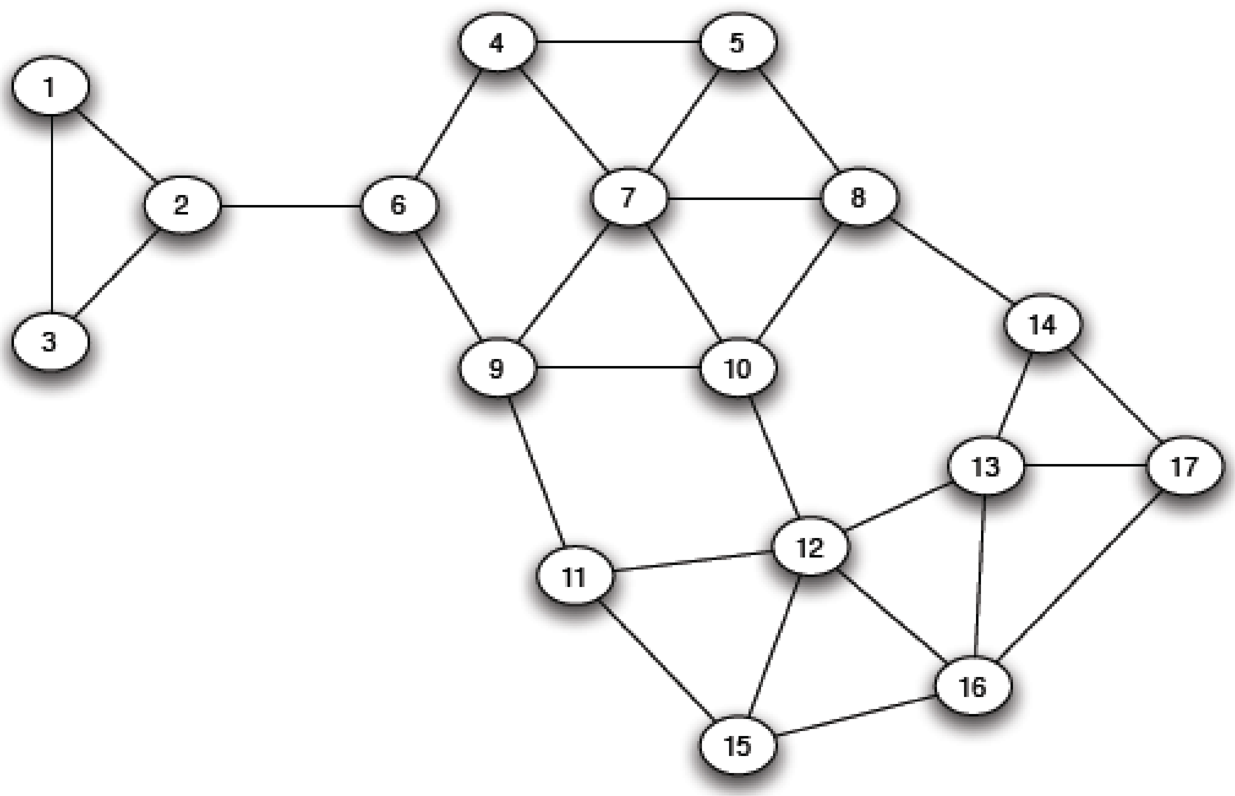
(b) *Two nodes are the initial adopters*

(c) *After one step, two more nodes have adopted*

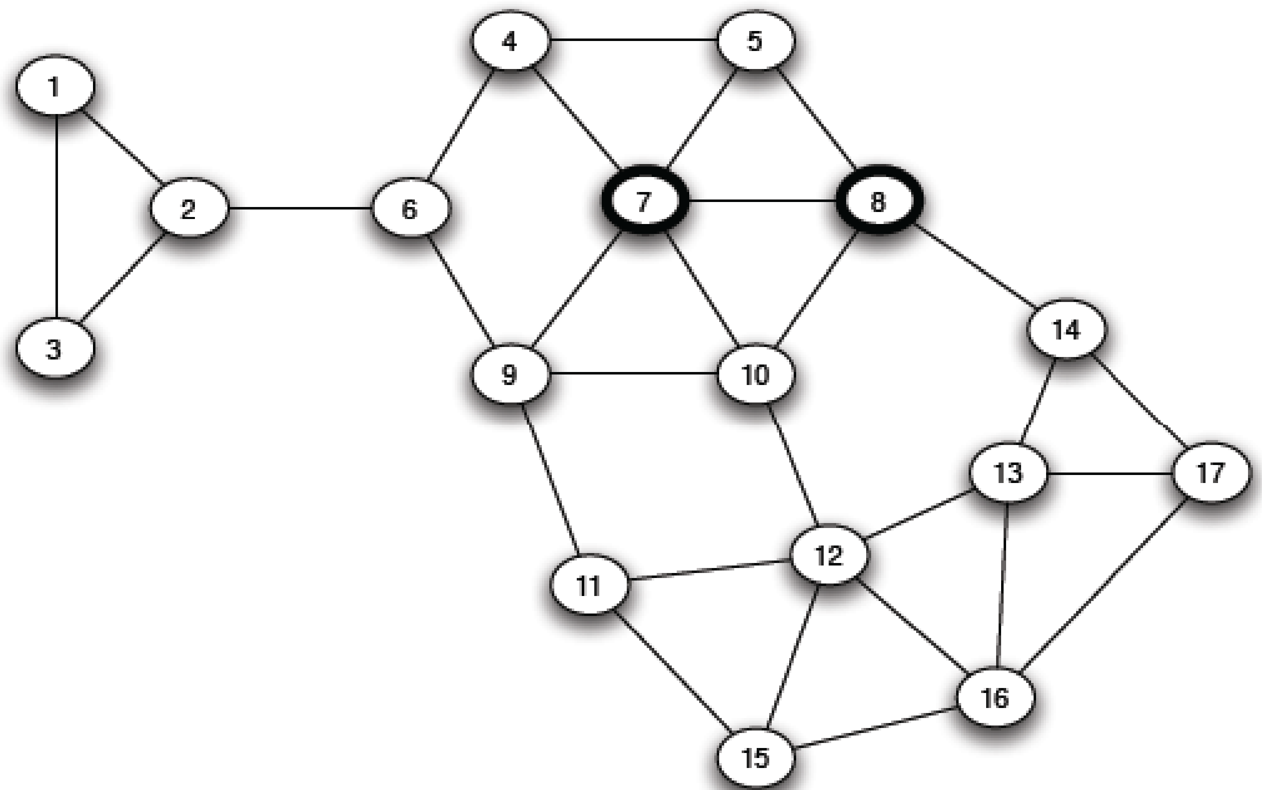
(d) *After a second step, everyone has adopted*

- v and w as the initial adopters
- payoffs a=3 and b=2 the new behavior, so the threshold is q=2/5
- A spreads to all nodes in two steps
- Nodes adopting A in a given step are drawn with dark borders
- nodes adopting B are drawn with light borders

Example

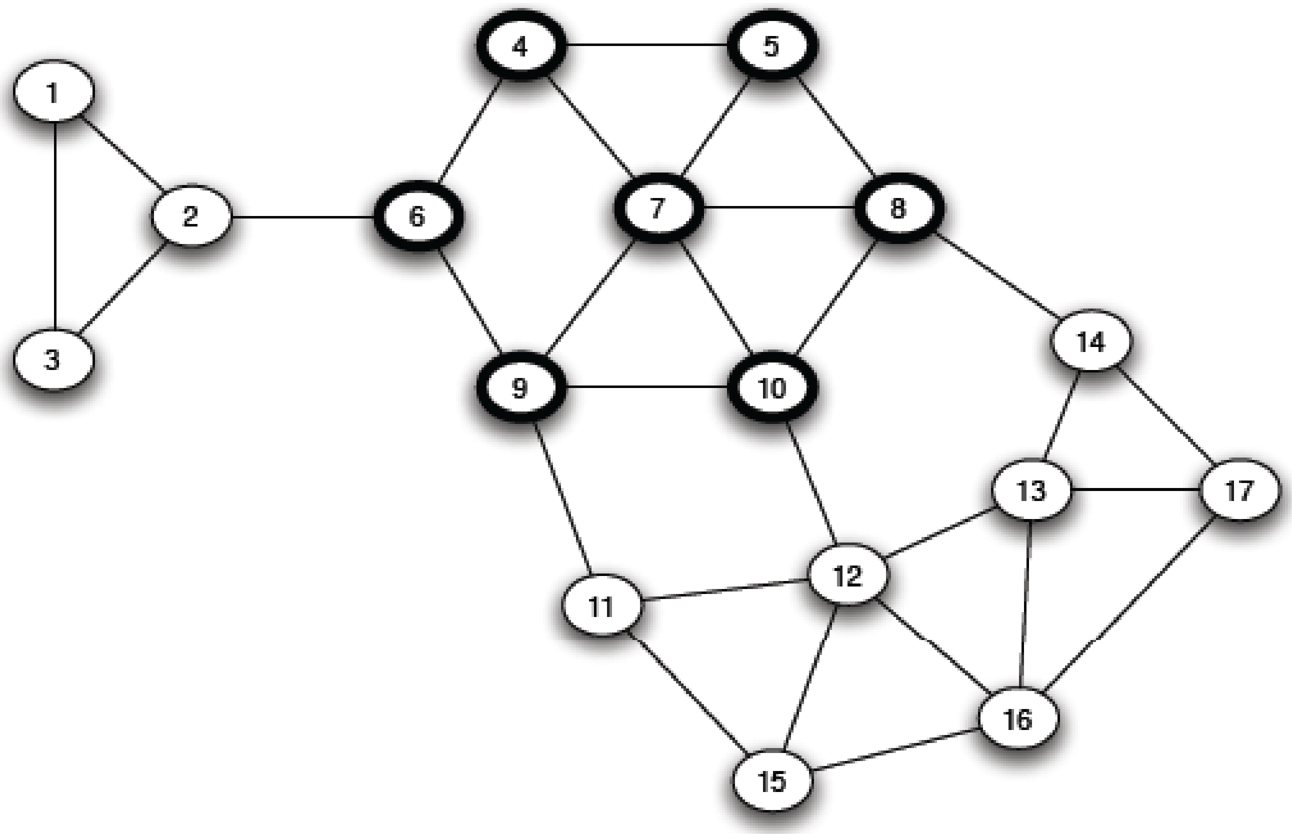


Example



- nodes 7 and 8 as the initial adopters
- payoffs a=3 and b=2 the new behavior

Example



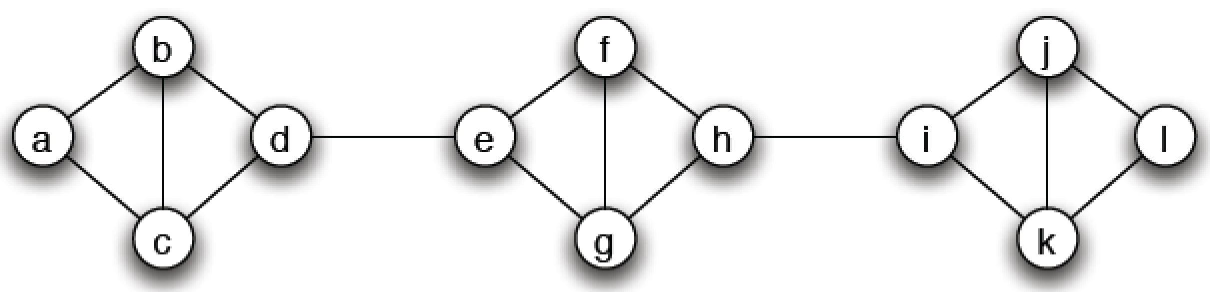
- In the next three steps, first nodes 5 and 10 switch to A
- Then, nodes 4 and 9
- Then, node 6
- At this point, no further nodes will be willing to switch

Cascading behavior in the game

- The chain reaction of switches to A is called a cascade of adoptions of A
- There are two possibilities:
  - (i) the cascade runs for a while but stops while there are still nodes using B
  - (ii) there is a complete cascade, in which every node in the network switches to A.
- We introduce the following terminology for referring to the second possibility.
  - Consider a set of initial adopters who start with A, while every other node starts with B. Nodes then repeatedly evaluate the decision to switch from B to A using a threshold of q. If the resulting cascade of adoptions of A eventually causes every node to switch from B to A, then we say that the set of initial adopters causes a **complete cascade** at threshold q.

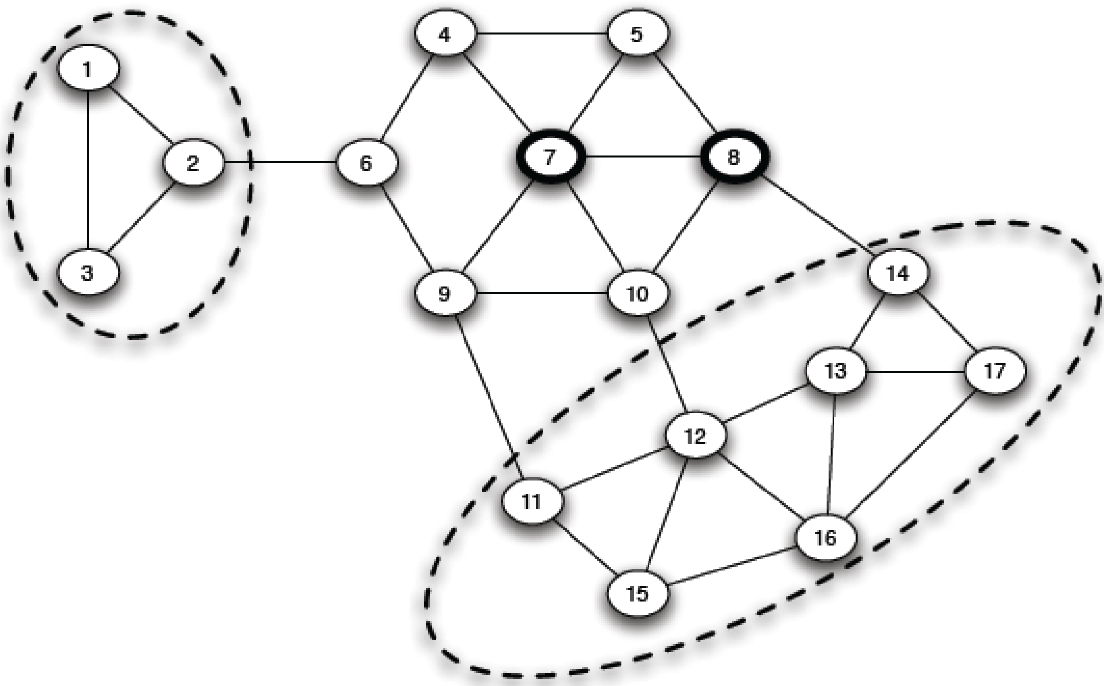
Cascades and clusters

- Communities have obvious effect on the spreading of information
- Spread of a new behavior can stop when it tries to break in to a tightly-knit community within the network
- Let's think about **densely connected community**
- We say that a cluster of density p is a set of nodes such that each node in the set has at least p fraction of its network neighbors in the set.
- Example: the set of nodes a, b, c, and d forms a cluster of density 2/3 in the following network



Cascades and clusters

- Consider the previous example
- There are two communities each of density 2/3
- These correspond precisely to the parts of the network that the cascading behavior A was unable to break into, starting from nodes 7 and 8 as initial adopters

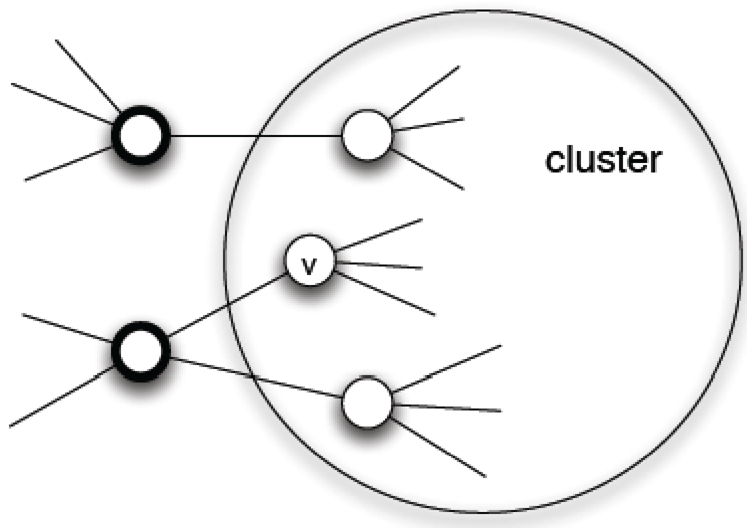


Cascades and clusters

- **Claim:** Consider a set of initial adopters of behavior A, with a threshold of q for nodes in the remaining network to adopt behavior A.
- (i) If the remaining network contains a cluster of density greater than 1-q, then the set of initial adopters will not cause a complete cascade
- (ii) Moreover, whenever a set of initial adopters does not cause a complete cascade with threshold q, the remaining network must contain a cluster of density greater than 1-q

Proof of part (i)

- Assume the opposite that some node v inside the cluster does adopt A at the earliest time step t

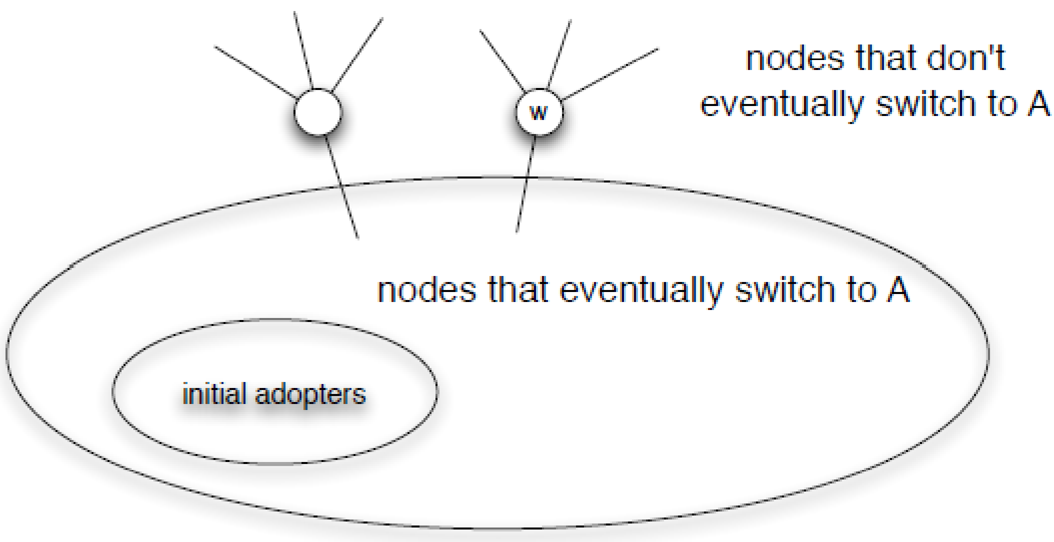


Proof of part (i)

- At the time that v adopted A, its decision was based on the set of nodes who had adopted A by the end of the previous time step, t-1.
- Since no node in the cluster adopted before v did, the only neighbors of v that were using A were outside the cluster
- But, the cluster has density greater than 1-q, more than a 1-q fraction of v's neighbors are inside the cluster, and hence less than q fraction of v's neighbors are outside the cluster who could have been using A
- The threshold rule requires at least q fraction of neighbors using v, this is a contradiction.
- Hence our original assumption, must be false.

Proof of part (ii)

- Consider running the process by which A spreads, starting from the initial adopters, until it stops
- It stops because there are still nodes using B, but none of the nodes in this set want to switch
- The following figure illustrated this



Proof of part (ii)

- Let S denote the set of nodes using B at the end of the process
- We claim that S is a cluster of density greater than 1-q
- Consider any node w in this set S
- Since w doesn't want to switch to A, it must be that the fraction of its neighbors using A is less than q
- Hence, the fraction of its neighbors using B is greater than 1-q
- But, the only nodes using B in the whole network belong to the set S, so the fraction of w's neighbors belonging to S is greater than 1-q
- This holds for all nodes S, it follows that S is a cluster of density greater than 1- q

To wrap up

- The punch-line is that in this model, a set of initial adopters can cause a complete cascade at threshold q if and only if the remaining network contains no cluster of density greater than (1-q).
- So in this sense, cascades and clusters truly are natural opposites:
  - **Clusters block the spread of cascades, and whenever a cascade comes to a stop, there's a cluster that can be used to explain why**

Heterogeneous threshold

- For each node v, we define a payoff a_v and b_v
- When two nodes v and w interact across an edge in the network, they are thus playing the following coordination game

	A	B
A	a_v, a_w	$0, 0$
B	$0, 0$	b_v, b_w

- Almost all of the previous analysis carries over with only small modifications

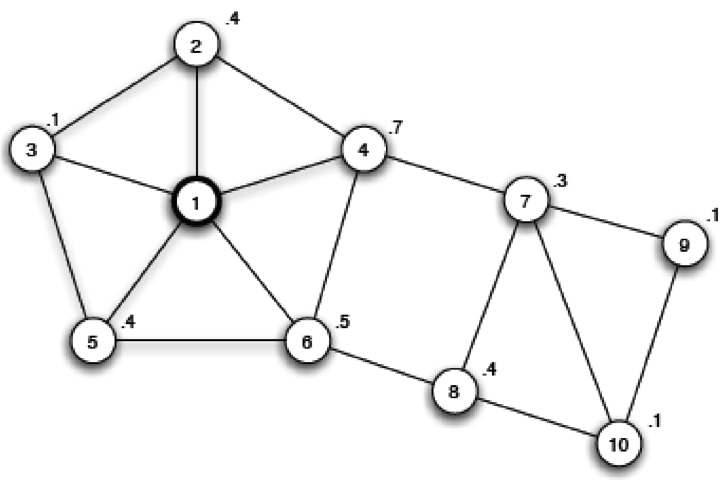
Heterogeneous threshold

- If v has d neighbors, of whom p fraction have behavior A and 1-p fraction have behavior B, then
  - **the payoff from choosing A is p.d.a_v**
  - **the payoff from choosing B is (1-p).d.b_v**
- Thus, A is better choice if:

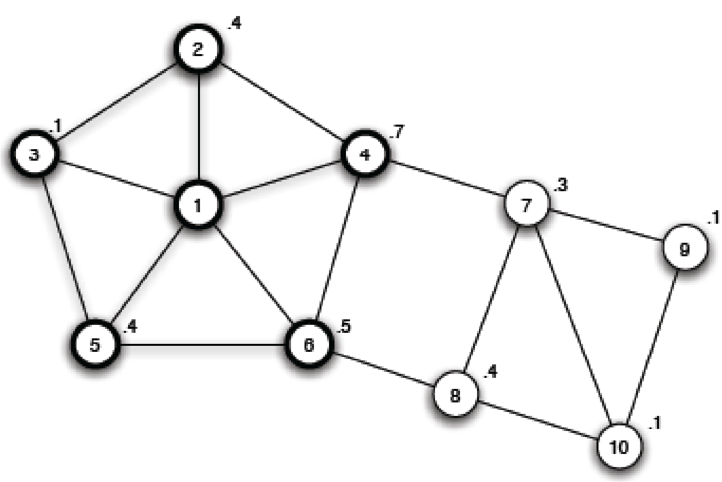
$$\mathbf{p \geq b_v/(b_v+a_v)=q_v}$$

- Each node v has its own personal threshold q_v
- v chooses A if at least q_v fraction of its neighbors have done so
- Moreover, if a node values A more highly relative to B, its threshold q_v is correspondingly lower

Heterogeneous threshold



(a) *One node is the initial adopter*



(b) *The process ends after four steps*

- Each node's threshold is drawn to the upper-right of the node itself
- The diversity in node thresholds clearly plays an important role that interacts in complex ways with the structure of the network
- e.g., despite node 1's **central position**, it would not have succeeded in converting anyone at all to A were it not for the extremely low threshold on node 3
- So, for understanding the spread of behaviors in social networks, we need to take into account not just the power of influential nodes, but also the extent to which these influential nodes have access to easily influenceable people.

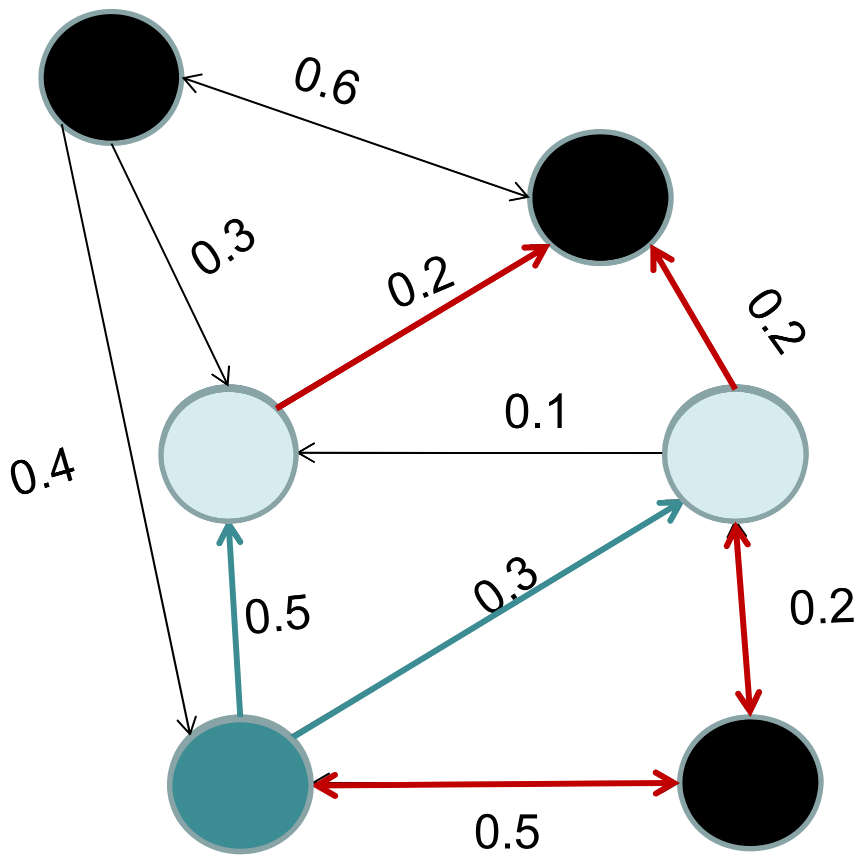
The cascade capacity

- What is the largest threshold at which any small set of initial adopters can cause a complete cascade?
- This maximum threshold is an inherent property of the network, indicating the outer limit on its ability to support cascades
- We refer to it as the **cascade capacity** of the network

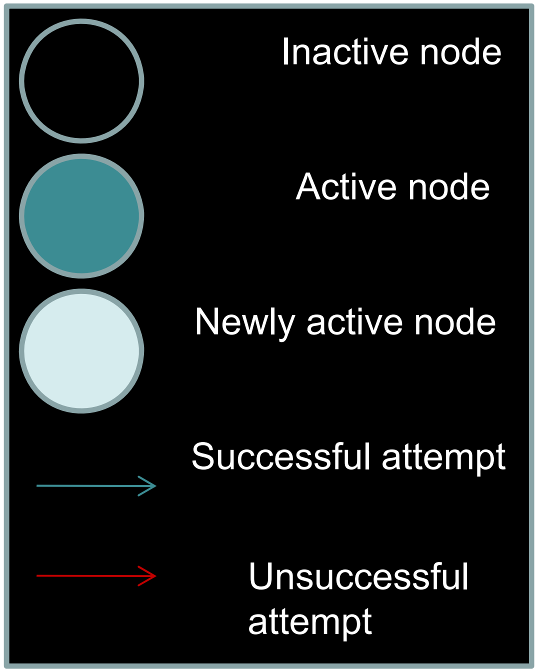
Independent Cascade Model

- Start with an initial set of active nodes A_0
- The diffusion process unfold in discrete steps
  - When node v first becomes active in step t, it is given a signal chance to active each currently inactive neighbor
    - It succeeds at a certain probability
  - If v succeed, then for example some of its neighbors become active at t+1

Example



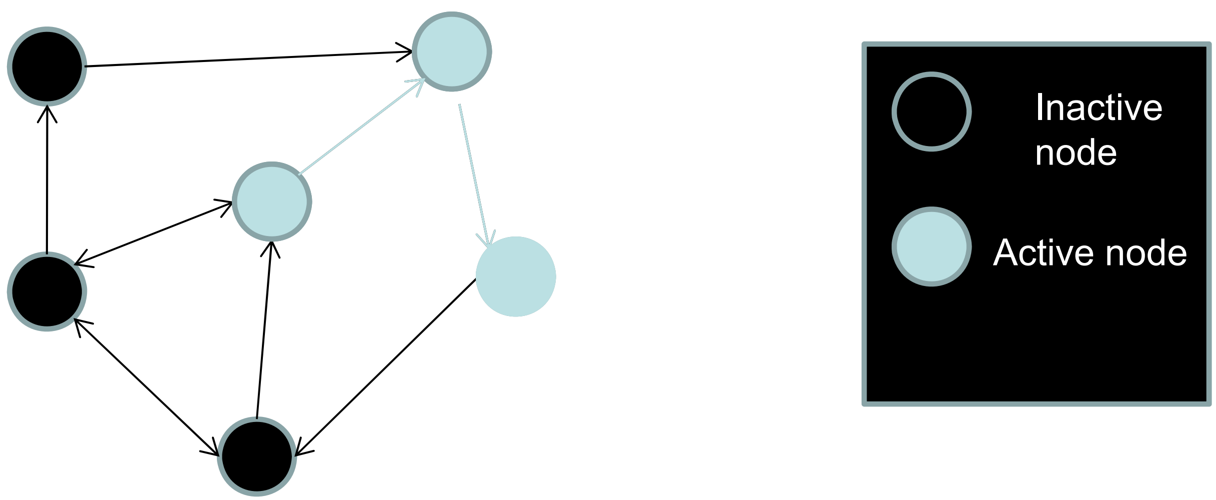
0.6
0.3
0.2
0.2
0.1
0.4
0.5
0.3
0.2
0.5



Inactive node
Active node
Newly active node
Successful attempt
Unsuccessful attempt

Voter Model

- Start with an initial set of active nodes A_0
- At each time step
  - For each node i
  - One of its neighbors, j say, chosen at random
  - i assume the opinion of j



Opinion formation

- Sznajd (S)
- Deffuant (D)
- Karause and Hegselman (KH)

Sznajd

- Discrete opinion : +1, -1
- Neighbors : e.g. D-dimensional lattice
- Meeting : one neighbor per cycle
- Ising Model:
  - At each time step, two randomly selected neighboring agents transfer their opinion to their neighbors if and only if they share the same opinion

Deffuant

- Continuous opinions, e.g. in the range [0,1] or [-1,1]
- Neighbors : any network topology
- Meeting : one neighbor per cycle
- Bounded confidence model
  - Two interacting agents have opinion X_i and X_j
  - If $|X_i - X_j| > \varepsilon$, nothing happens ($\varepsilon$ is threshold)
  - If $|X_i - X_j| < \varepsilon$, the agents can negotiate → they update their opinions as

$$\begin{cases} x_i(n+1) = x_i(n) + \mu\left[x_j(n) - x_i(n)\right] \\ x_j(n+1) = x_j(n) + \mu\left[x_i(n) - x_j(n)\right] \end{cases}$$

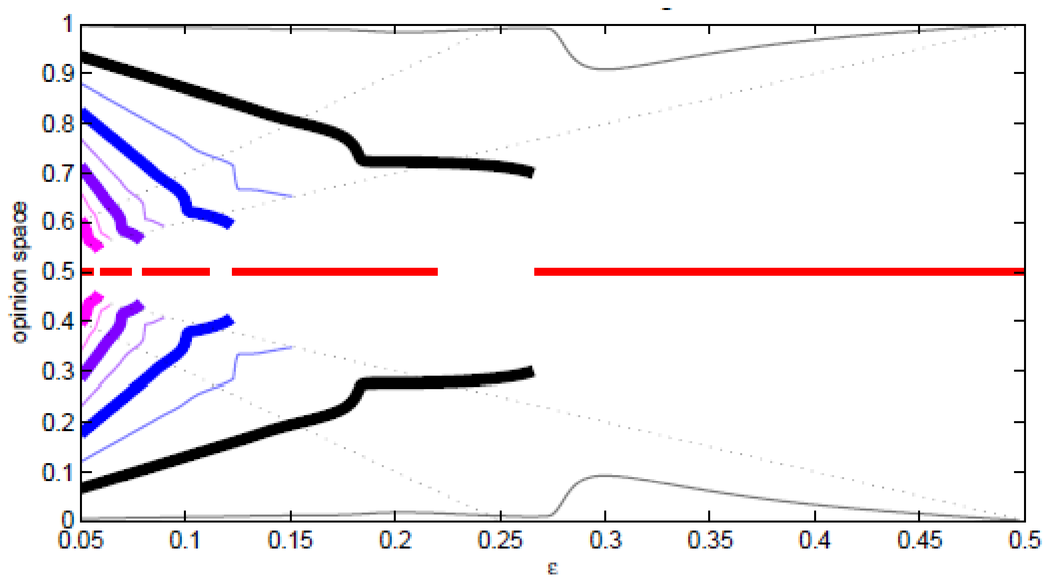
  - μ is convergence parameter

Karause and Hegselman (KH)

- Continuous opinions
- Neighbors : any network topology
- Meeting : one neighbor with all neighbors in its confidence bound
- Model
  - Node i is chosen at random
  - Changes its opinion into the arithmetic average of the opinions of all the neighboring agents that are within a confidence bound

Bounded-confidence model

- An interesting behavior is when the agents reach consensus in their opinions
- This clearly depends on both the network structure and the threshold parameter (the consensus cannot be attained for small value of thresholds)
- Opinions, e.g. in the range [0,1]
- Figure shows bifurcation diagram of opinion as a function of the threshold



Source: Lorenz, IJMPC, 2007

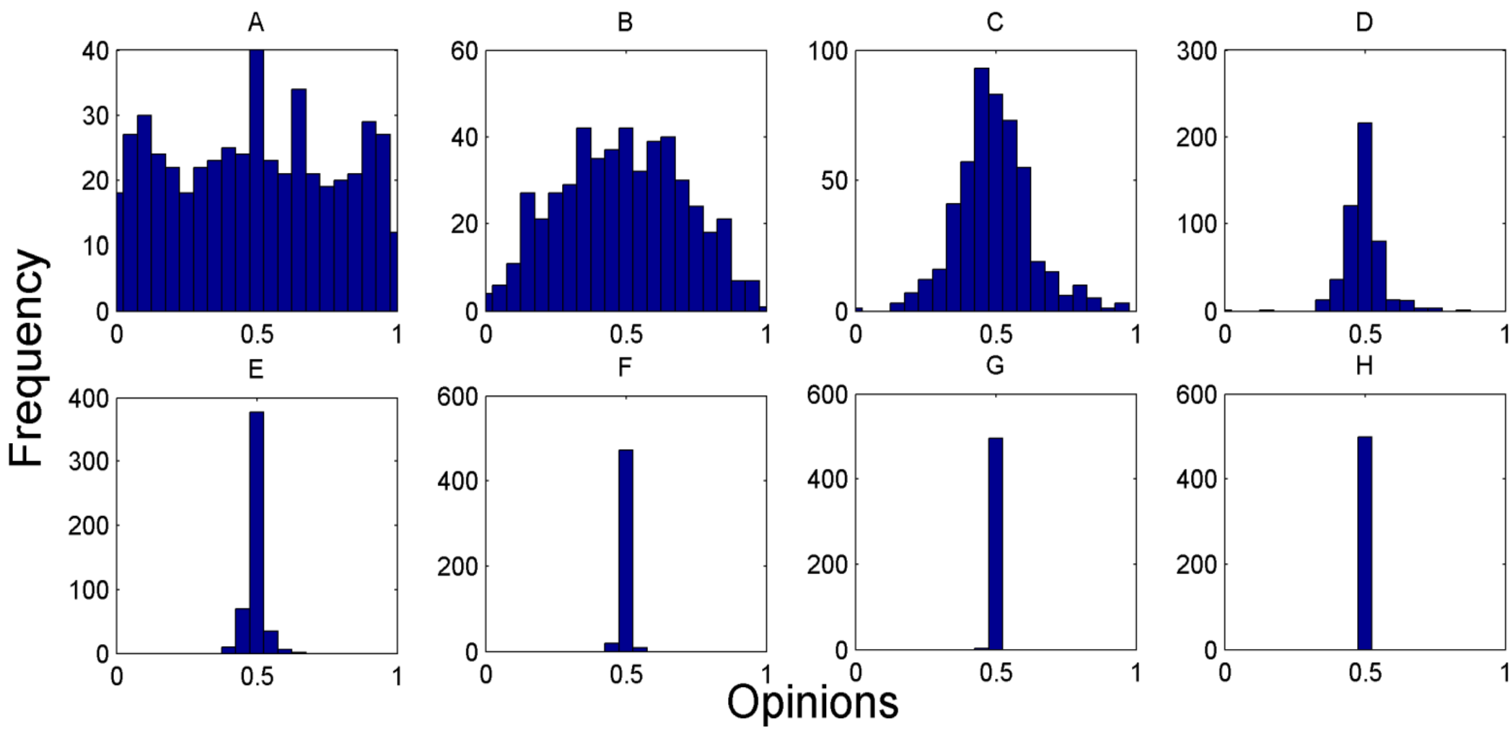
Bounded-confidence model

- In Practice, the consensus can be checked numerically
- At each time t, average error is calculated as

$$E(t) = \frac{2}{N(N-1)} \sum_{i<j} \left\| X_i(t) - X_j(t) \right\|^2$$

- The time the network needs until the average error reaches a threshold, e.g. $\delta = 10^{-6}$, and stays below thereafter is interpreted as consensus time
- Indeed, the time T, where $E(T) = \delta$ and $E(t) < \delta$ for $t > T$, is consensus time
- Less T means faster consensus
- In some applications it is desired to have consensus as fast as possible

An example



Histogram of the opinions in increasing simulation steps from A-H. The network is scale-free with size $N = 500$ and average degree $<k> = 4$. The threshold is set at 0.4

Source: Jalili, 2011

BC model with social power

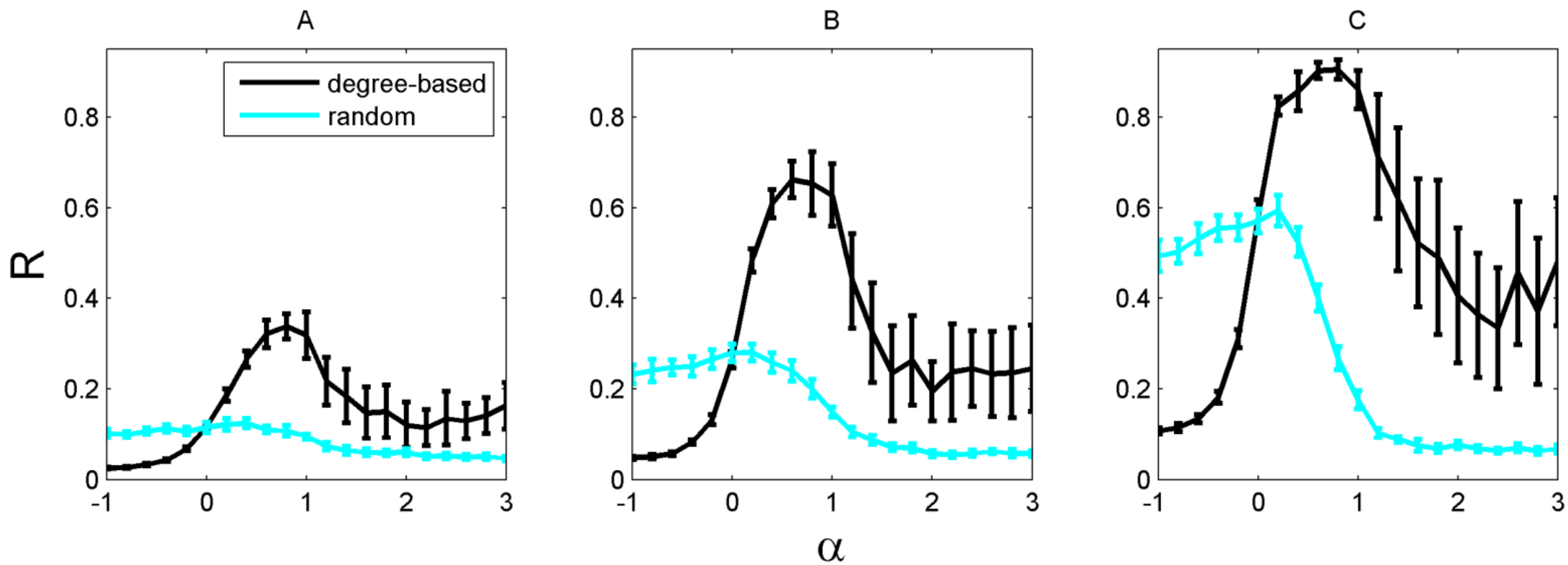
- In reality, not all the nodes have the same influence on each other
- Some have social power → have greater influence
- Let us assume social power of node i as

$$SP_i = (C_i)^\alpha$$

- where C_i is a quantity proportional to its centrality
- The simplest centrality notion is degree
- So, the update equations get a form as

$$\begin{cases} x_i(n+1) = x_i(n) + \mu \dfrac{SP_j}{SP_i + SP_j}\left[x_j(n) - x_i(n)\right] \\ x_j(n+1) = x_j(n) + \mu \dfrac{SP_i}{SP_j + SP_i}\left[x_i(n) - x_j(n)\right] \end{cases}$$

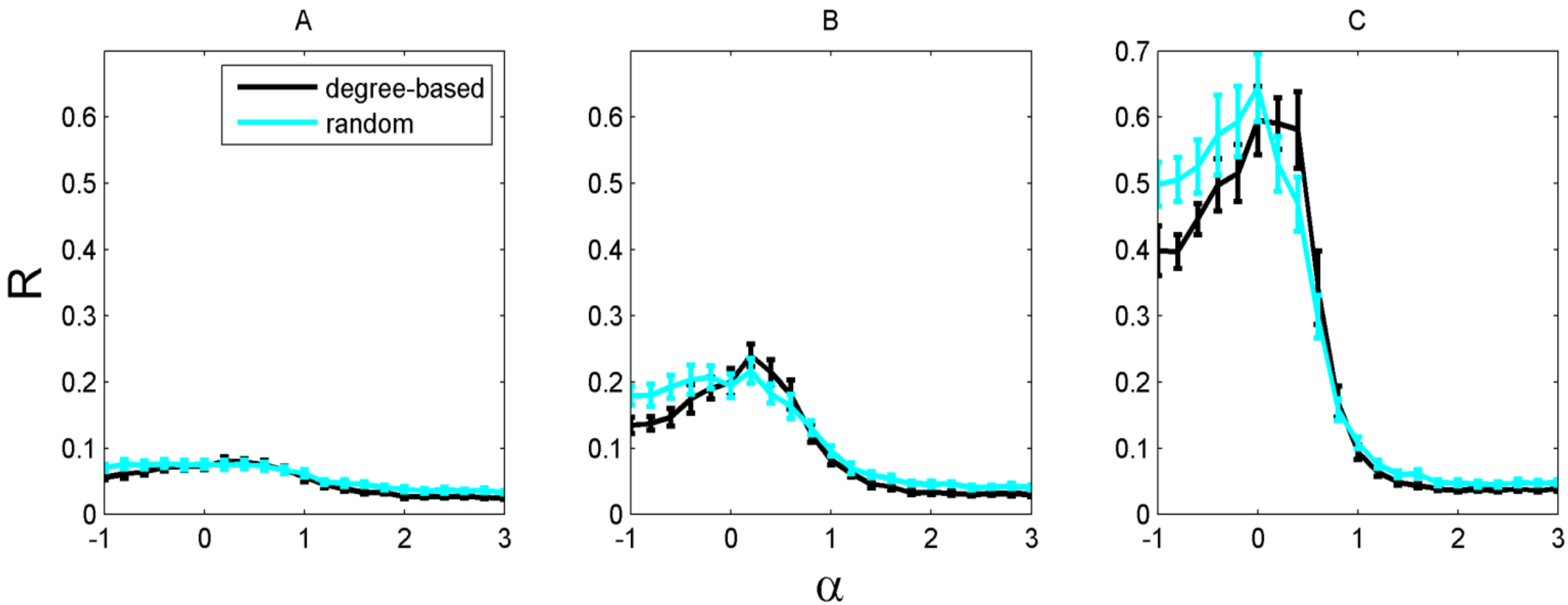
Simulation results



The normalized number of the agents reached a consensus in their opinions (R) as a function of α. The networks are BA scale-free with N = 512 nodes, and average degree as A) $<k>$ = 4, B) $<k>$ = 6, and C) $<k>$ = 10. 5% of the hub nodes or random selection of nodes are chosen to have social power and the simulations are repeated for 20000 time steps

Source: Jalili, 2011

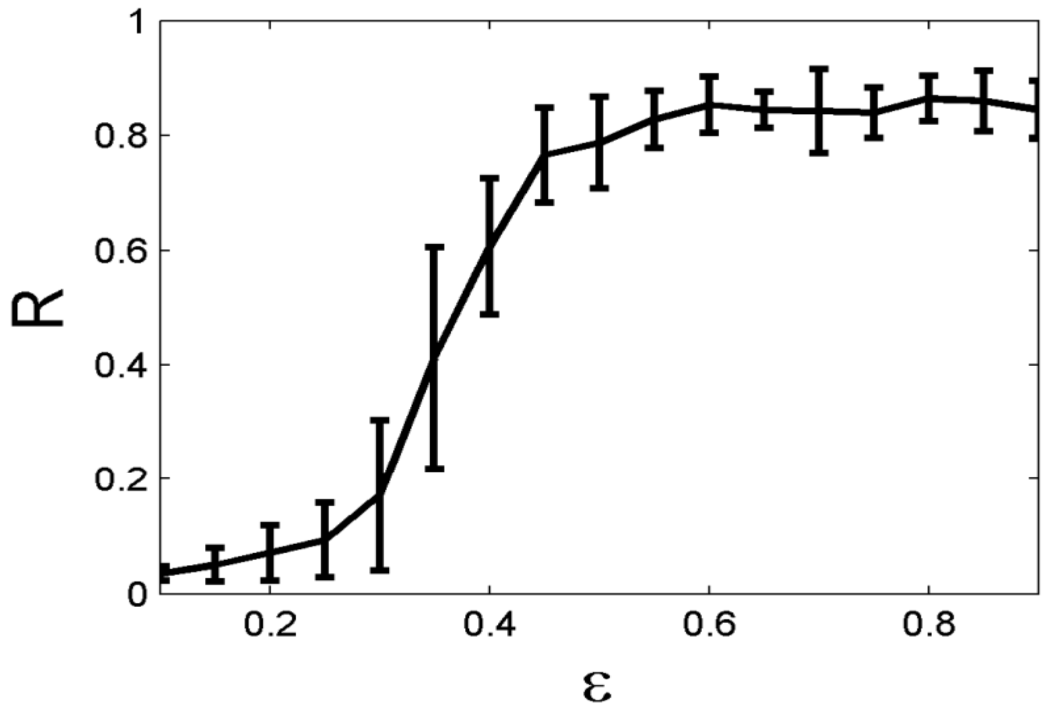
Simulation results



R as a function of α in Watts-Strogatz networks with $N = 512$ nodes, rewiring probability $P = 0.2$, and average degree as A) $<k> = 4$, B) $<k> = 6$, and C) $<k> = 10$. 5% of the hub nodes or random selection of nodes are chosen to have social power and the simulations are repeated for 20000 time steps

Source: Jalili, 2011

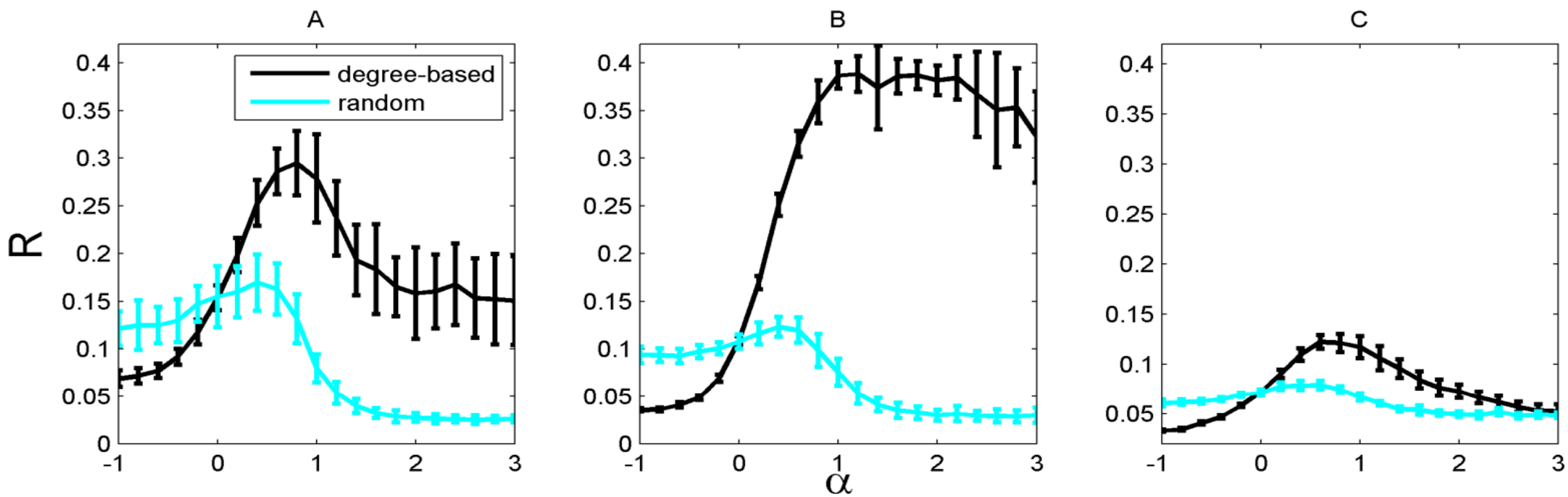
Simulation results



R as a function of threshold ε in BA scale-free networks with $N = 512$ and $\langle k \rangle = 6$. The social power parameter is set to $\alpha = 1$ and the simulations are repeated for 20000 time steps

Source: Jalili, 2011

Simulation results



R as a function of α for a number of real networks including A) email communication, B) facebook-like, and C) collaboration networks. The simulations are repeated for 50000 time steps

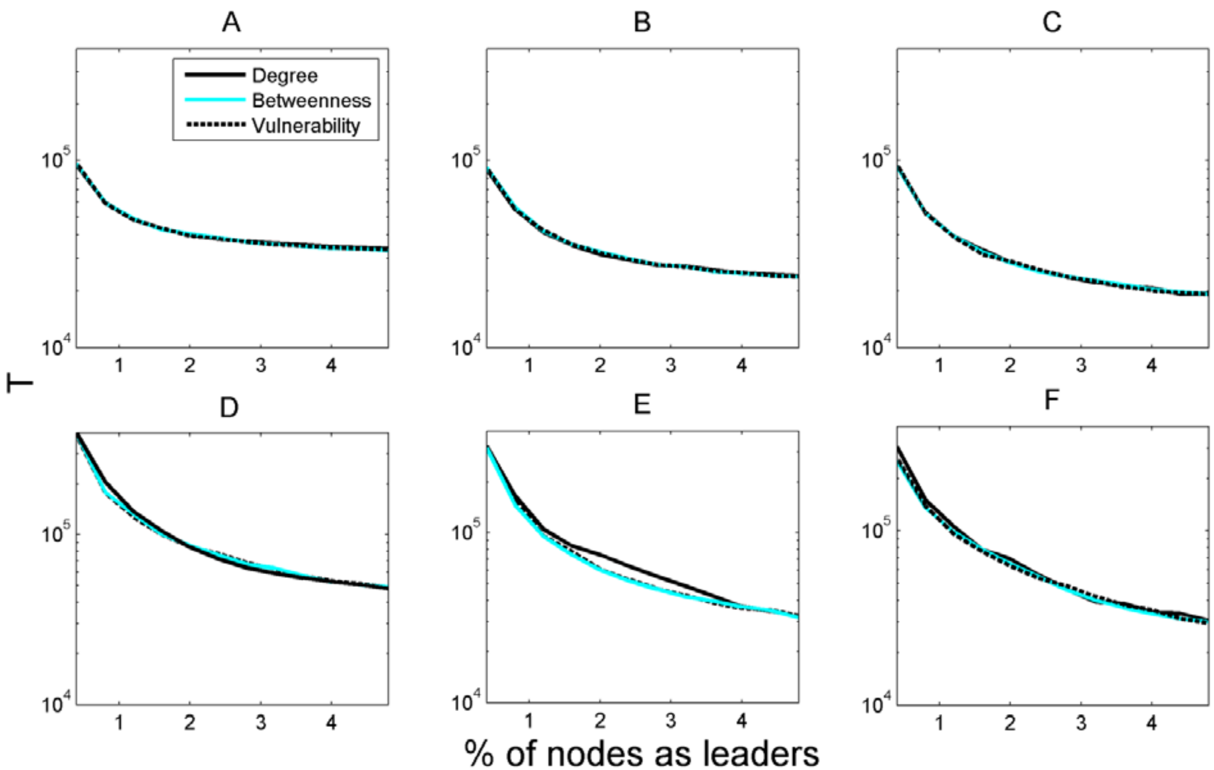
Source: Jalili, 2011

BC model with leaders

- Again, a more realistic way is to have social leaders
- Leaders have great influence on other individuals in the society
- But, their opinions are hardly influenced by others
- Let us suppose fist K agents are leaders whose opinions are unchanged
- We also consider social power
- So,

$$
\begin{cases}
\begin{cases}
x_i(n+1) = x_i(n) & \text{if} \quad 1 \le i \le K \\
x_i(n+1) = x_i(n) + \mu \dfrac{SP_j}{SP_i + SP_j}\left[x_j(n) - x_i(n)\right] & \text{if} \quad \mathrm{K} < i \le N
\end{cases} \\
\begin{cases}
x_j(n+1) = x_j(n) & \text{if} \quad 1 \le j \le K \\
x_j(n+1) = x_j(n) + \mu \dfrac{SP_i}{SP_j + SP_i}\left[x_i(n) - x_j(n)\right] & \text{if} \quad \mathrm{K} < j \le N
\end{cases}
\end{cases}
$$

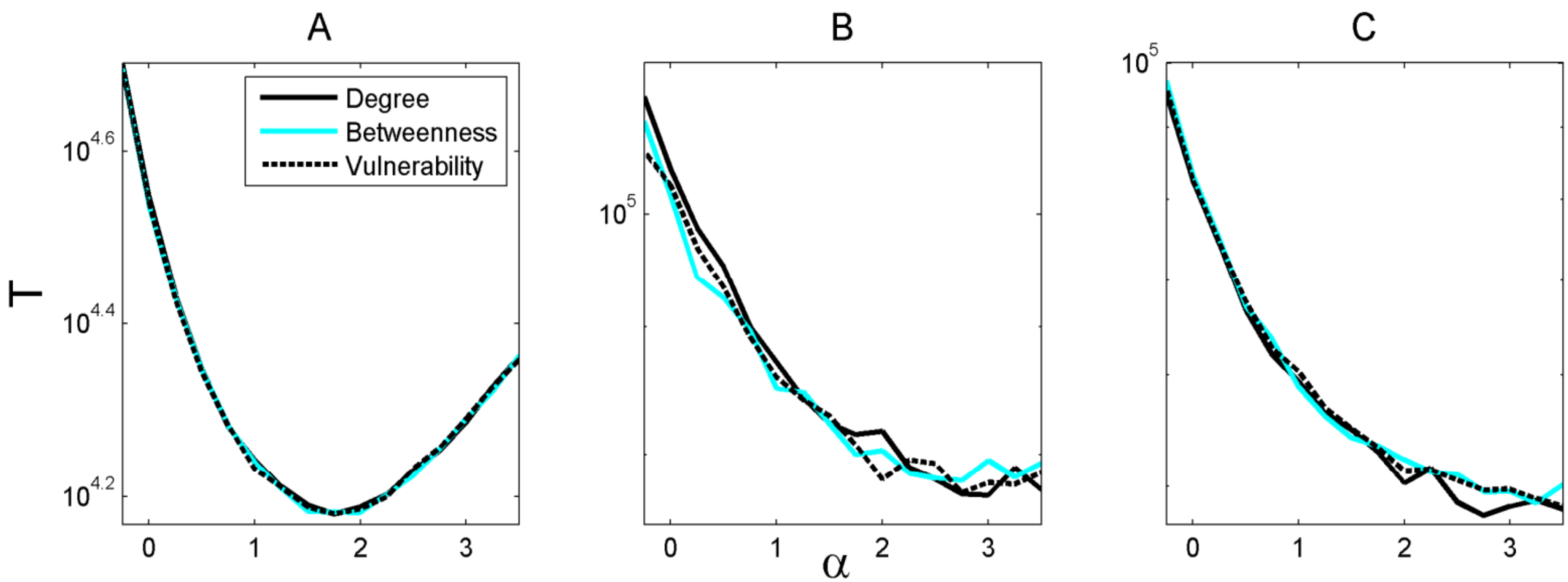
Simulation results



Consensus time (T). Different criteria are considered to chose the leaders. The networks are BA with size $N = 500$ and average degree A) $<k> = 4$, B) $<k> = 6$, and C) $<k> = 10$; or Watts-Strogatz with size $N = 500$, rewiring probability $P = 0.2$ and average degree D) $<k> = 4$, E) $<k> = 6$, and F) $<k> = 10$

Source: Jalili, 2011

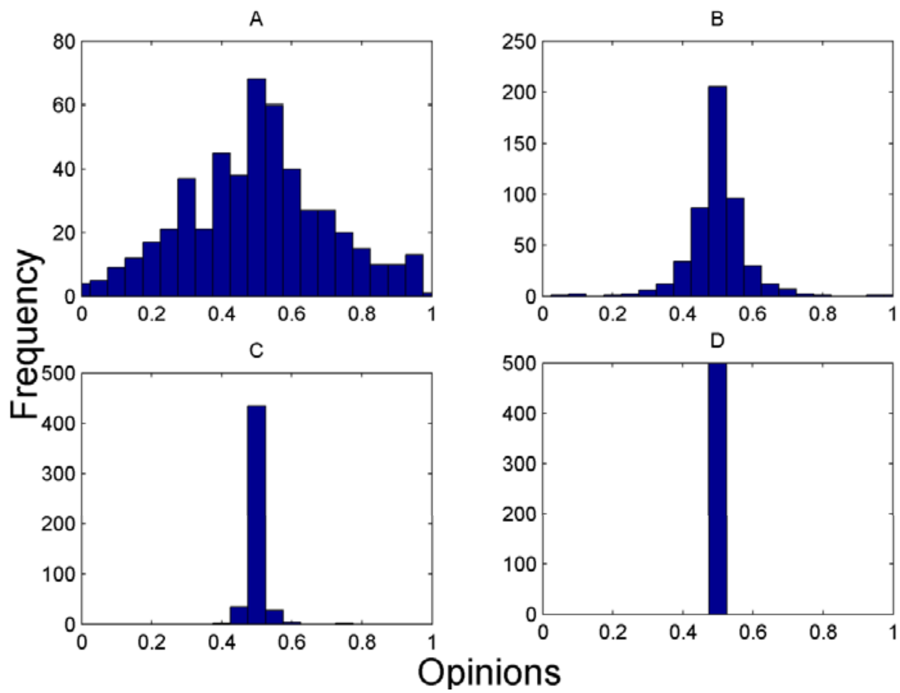
Simulation results



- T as a function of α controlling the social power. 1% of the nodes were considered to be leaders while others were normal agents. The networks are A) BA scale-free with $N = 500$ and $<k> = 6$, B) Watts-Strogatz with $N = 500$, $<k> = 6$, and $P = 0.2$, and C) Erdős-Rényi with $N = 500$ and $P = 0.02$

Source: Jalili, 2011

Simulation results

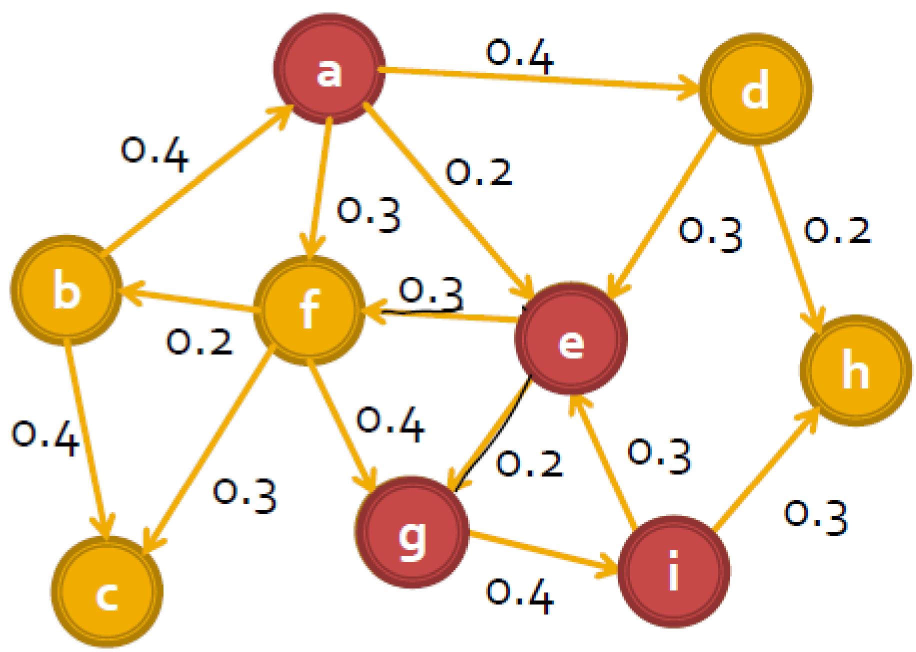


- Histogram of the opinions in different simulation steps as A) 1000, B) 3000, C) 6000, and D) 10000. The network is BA scale-free with size $N = 500$ and average degree $\langle k \rangle = 4$, and 2.4% of the high-degree nodes are considered to be the leaders with opinion value of 0.5.

Source: Jalili, 2011

K-most influential person problem

- Initially some nodes S are active
- Each edge (a,b) has probability (weight) p_{ab}
- Node a becomes active:
  - activates node b with probability p_{ab}
- Activations spread through the network



K-most influential person problem

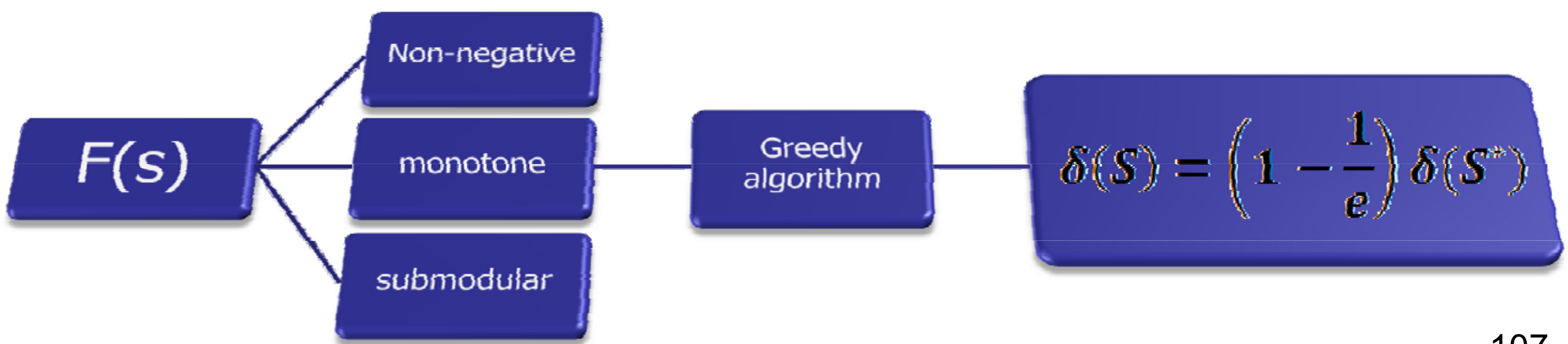
- Influence of node set S (S is the initial active set):
  - f(S): expected number of final active nodes (cascade size)
- S is more influential if f(S) is larger
- Problem:
  - Given a parameter k (budget), find a k-node set S to maximize f(S)
  - Constrained optimization problem with f(S) as the objective function

$$S^* = argmax_{s \subseteq V} f(s)$$

  - Influence maximization is NP-hard
  - There exists an approximation algorithm!
  - Greedy hill-climbing to find a good set S

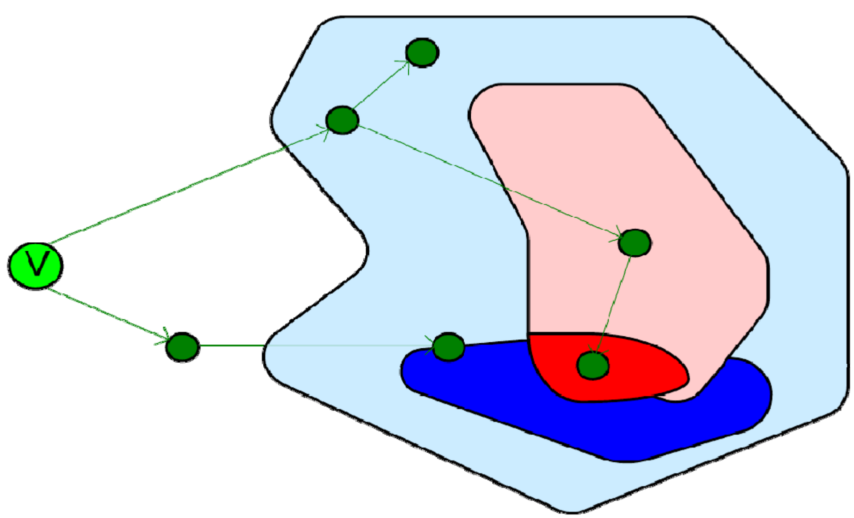
Hill climbing algorithm

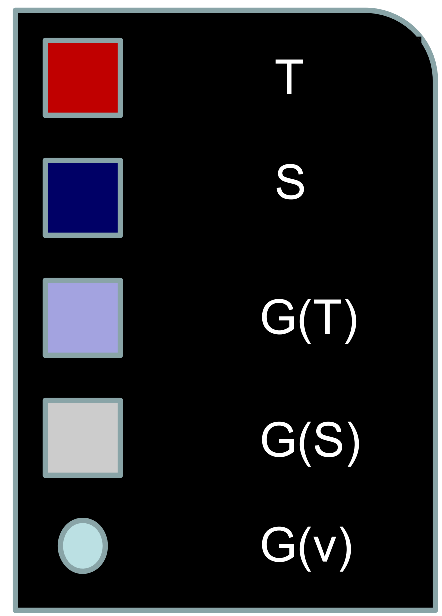
- Hill climbing
  - Start with $S_0 = \{\}$
  - For $i = 1:k$
  - Choose node v that max $f\{S_{i-1} \cup v\}$
  - Let $S_i = S_{i-1} \cup v$
- What is the runtime?
  - Each step just runs N time
  - steps for each node v



Approximation guarantee

- Hill climbing produces a solution S where f(s) ≥ (1-1/e) of optimal value (~63%)
- Claim holds for functions f with 2 properties:
- f is monotone: if S ≤ T, then f(S) ≤ f(T) and f({})=0
- f is submodular: adding element to a set gives less improvement than adding to one of subsets





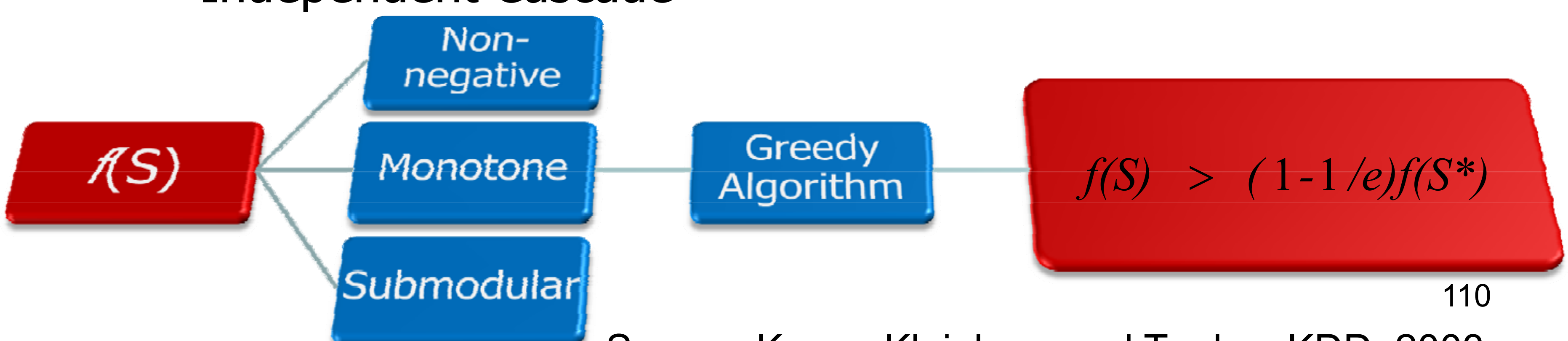
F(S) is submodularity

f is submodular: $\forall S \subseteq T$

$$\underbrace{F(S \cup \{u\}) - f(S)}_{\text{Gain of adding a node to a small set}} \geq \underbrace{f(T \cup \{u\}) - f(T)}_{\text{Gain of adding a node to a large set}}$$

Maximizing the Spread of Influence through a Social Network

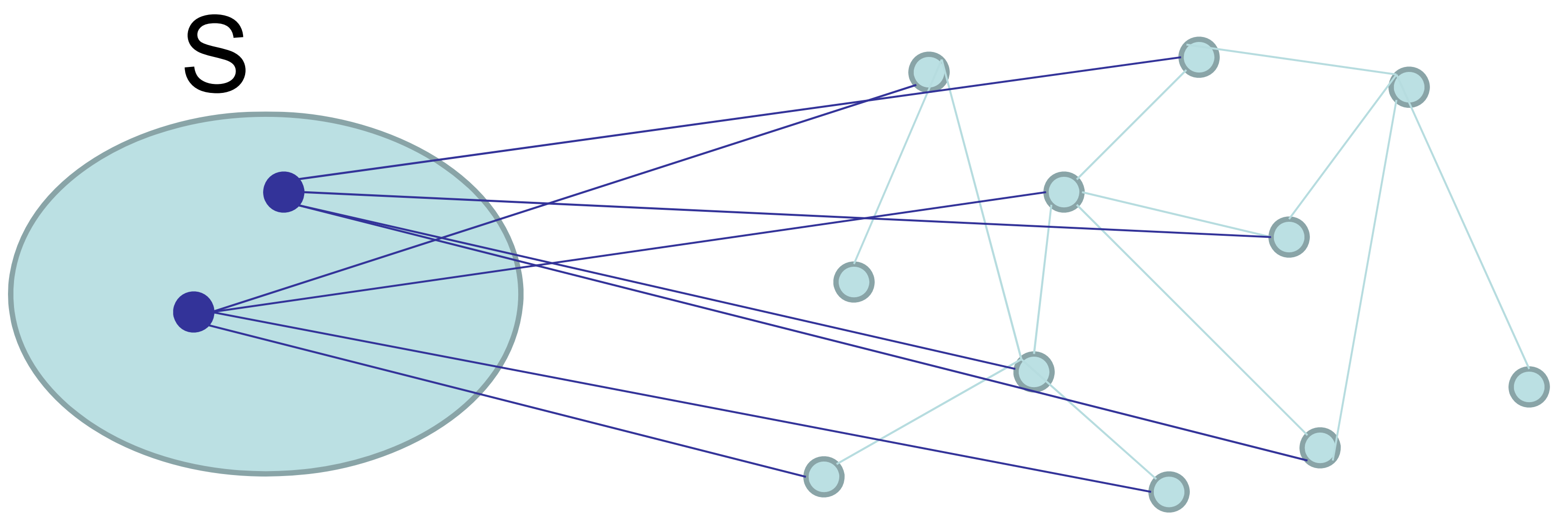
- Given
  - a limited budget B for initial advertising (e.g. give away free samples of product)
- Goal
  - trigger a large cascade of influence (e.g. further adoptions of a product)
- Models of influence
  - Linear Threshold
  - Independent Cascade



Source: Kemp, Kleinberg,and Tardos, KDD, 2003

Optimal Marketing Strategies over Social Networks

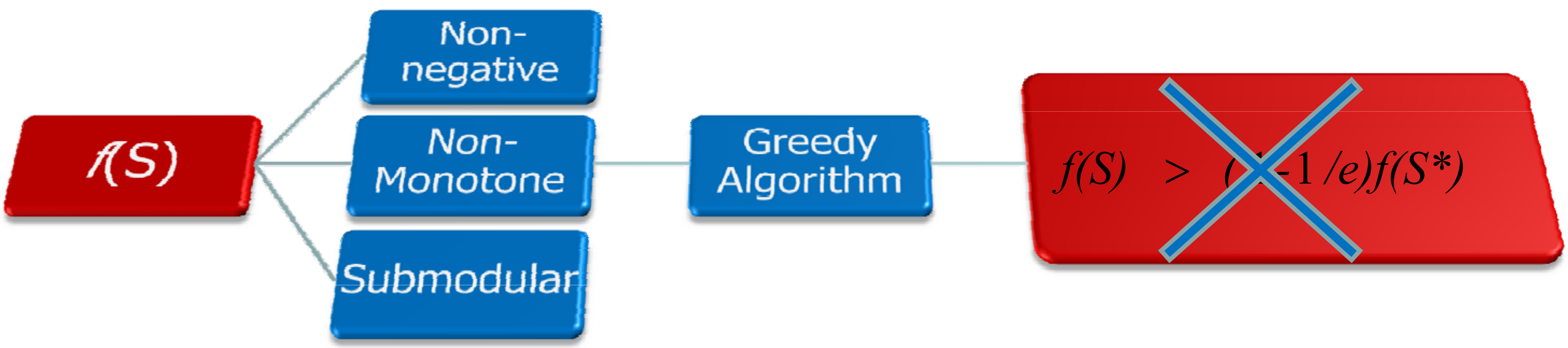
- Given
  - Adoption is based on price and network effects
- Goal
  - maximize expected revenue
- Identify a family of strategies called influence and exploit strategies that are easy to implement



Source: J. Hartline, V. S. Mirrokni et al. 2008

Optimal Marketing Strategies over Social Networks

- IE (information extraction) Strategy

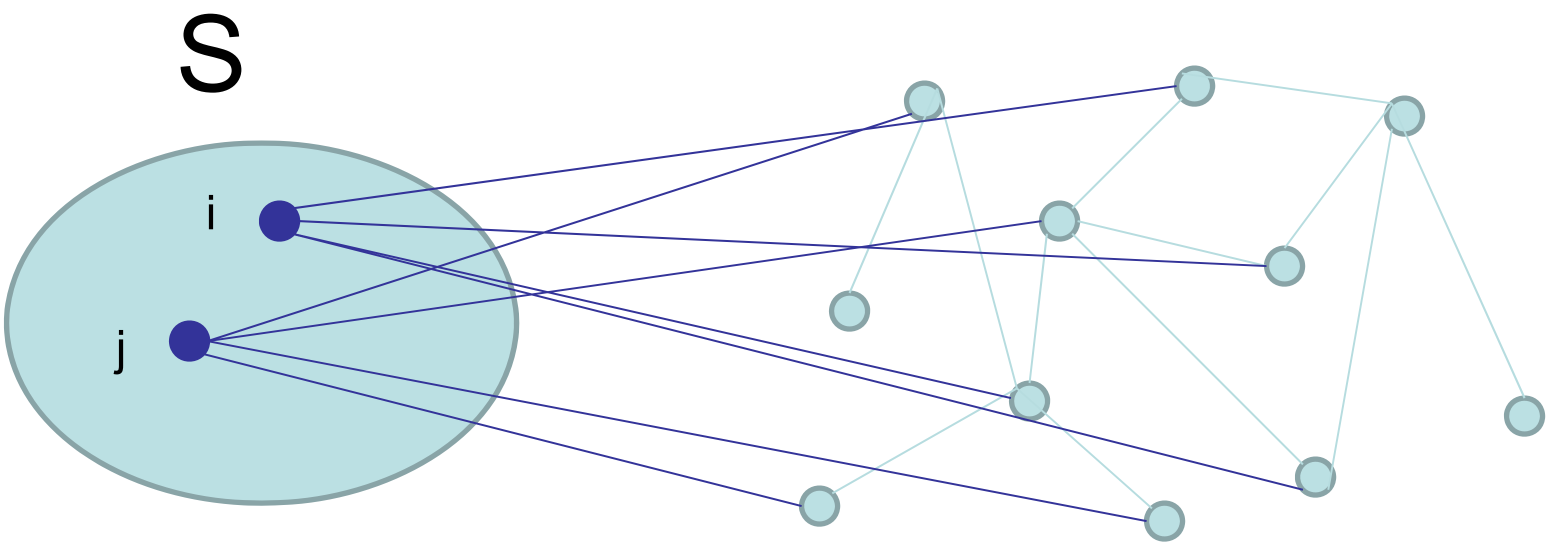


- Deterministic Local Search: 1/3-approximation
- Randomized Local Search : 2/5-approximation

Source: U. Feige, V. S. Mirrokni et al. 2007

Revenue Maximization

- Offering discount instead of offering the item for free to an initial set of buyers



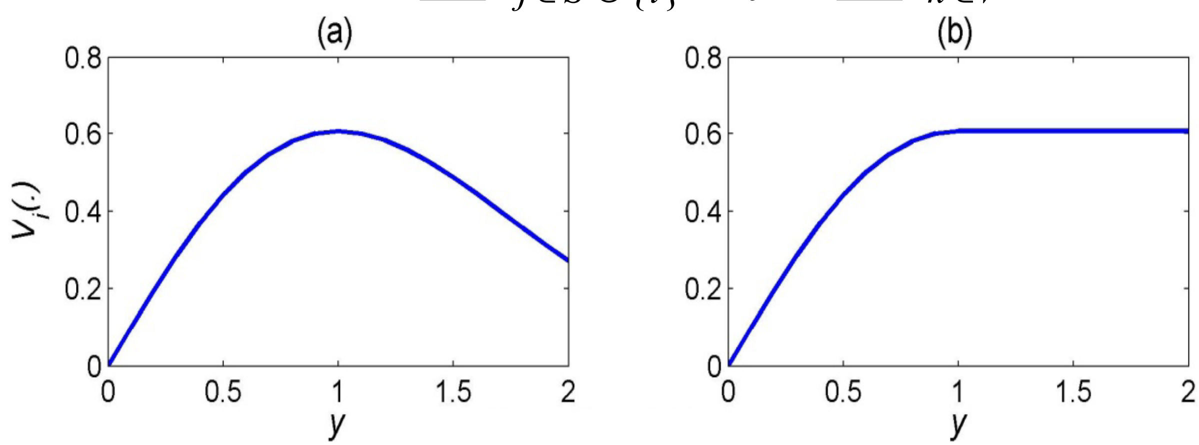
$$g(S) = f(S) + v_i + v_j$$

Source: Babaie , Mirzasoleiman, Jalili, and Safari, 2011

Revenue Maximization

- Influence Models
  - Monotone Concave Model
  - Non-monotone Concave Model

$$v_i(S) = f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik})$$



$v_i(.)$ as a function of y=2×normalized influence for a) non-monotone concave and b) monotone concave influence models.

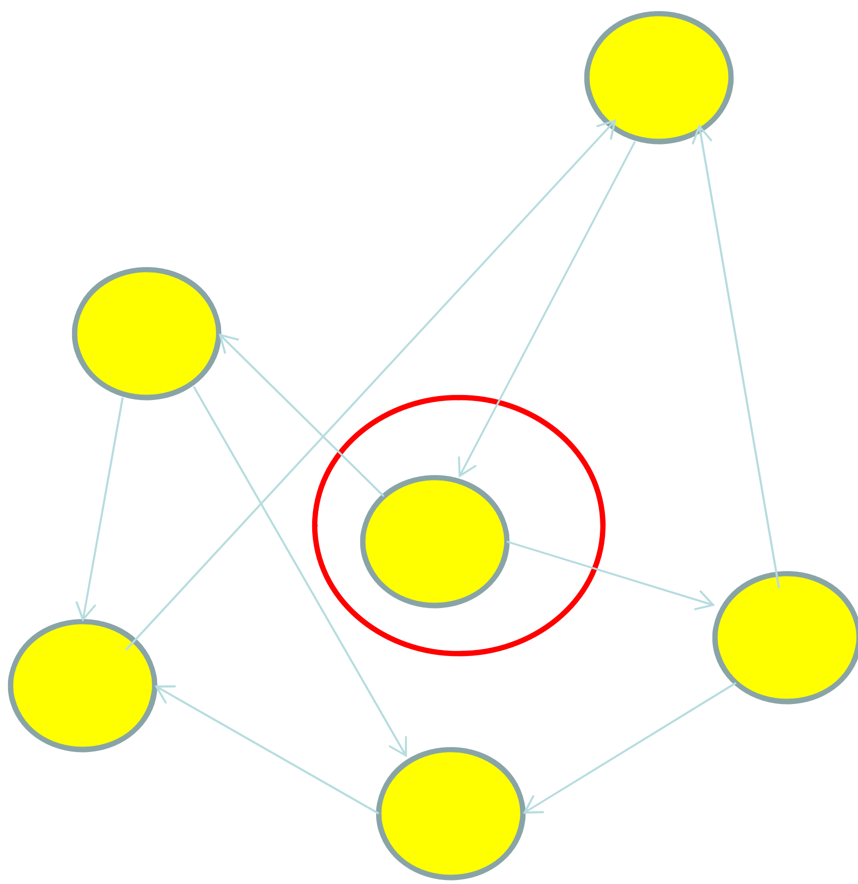
Source: Babaie , Mirzasoleiman, Jalili, and Safari, 2011

Revenue Maximization

- Local Search
  - Initialize set S = {v}, where v = argmax$_i$ g\{i\}
  - If neither of the following two steps apply (there is no local improvement), output S.
  - For any buyer that accepts the offered price, $g(S \cup \{i\}) > (1+\frac{\varepsilon}{n^2})g(S)$,then let $S = S \cup \{i\}$.
  - For any buyer $i \in S$, if $g(S \setminus \{i\}) > (1+\frac{\varepsilon}{n^2})g(S)$, then let $S = S \setminus \{i\}$ and go to step 2.

Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Local Search



Add to S/Delete from S, if F(S) improves

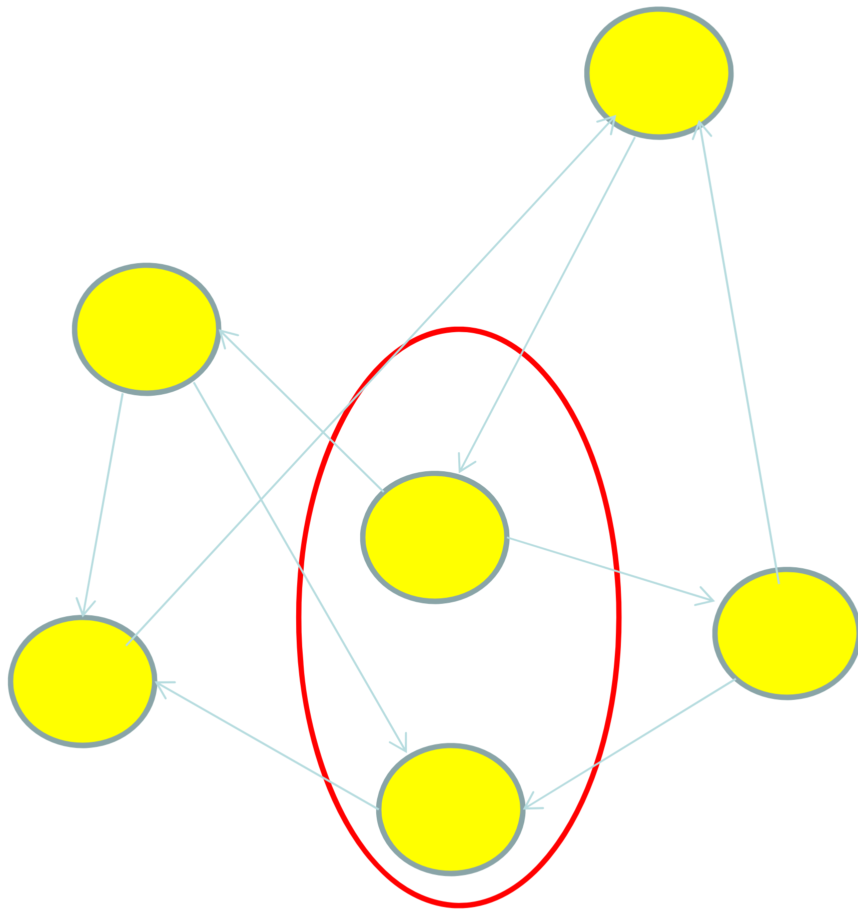
S = {5}

F(S) = 5

Maximizing non-monotone sub-modular functions

Source: Feige et. al., 2008

Local Search



Add to S/Delete from S, if F(S) improves

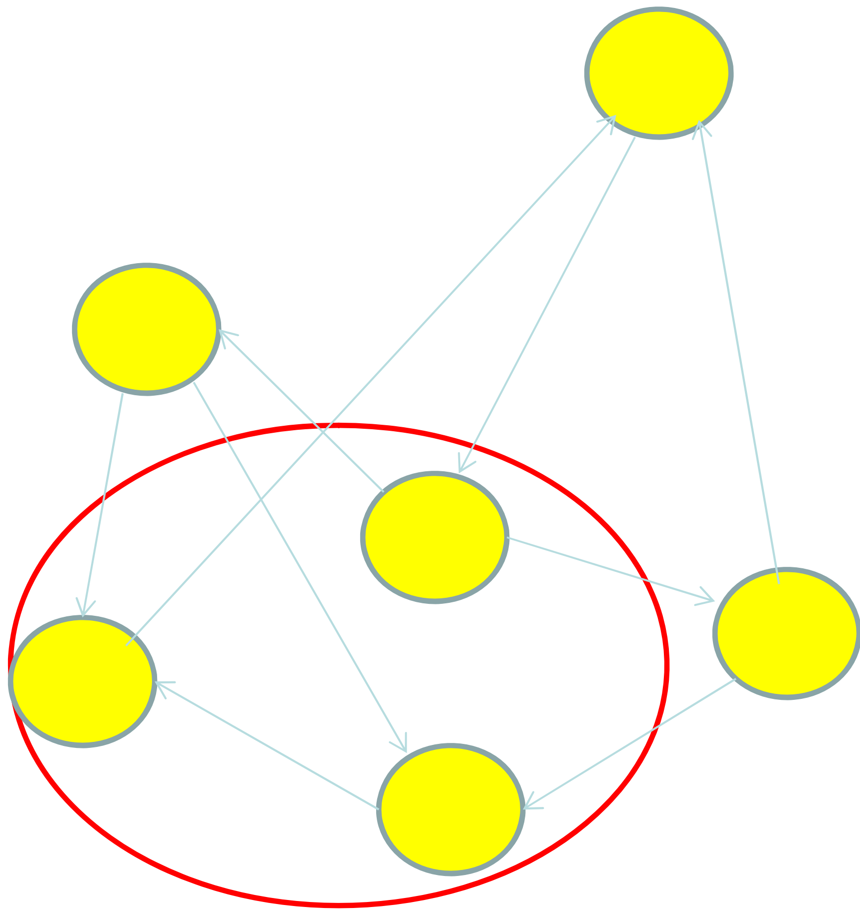
S = {3,5}

F(S) = 10

Maximizing non-monotone sub-modular functions

Source: Feige et. al., 2008

Local Search



Add to S/Delete from S, if F(S) improves

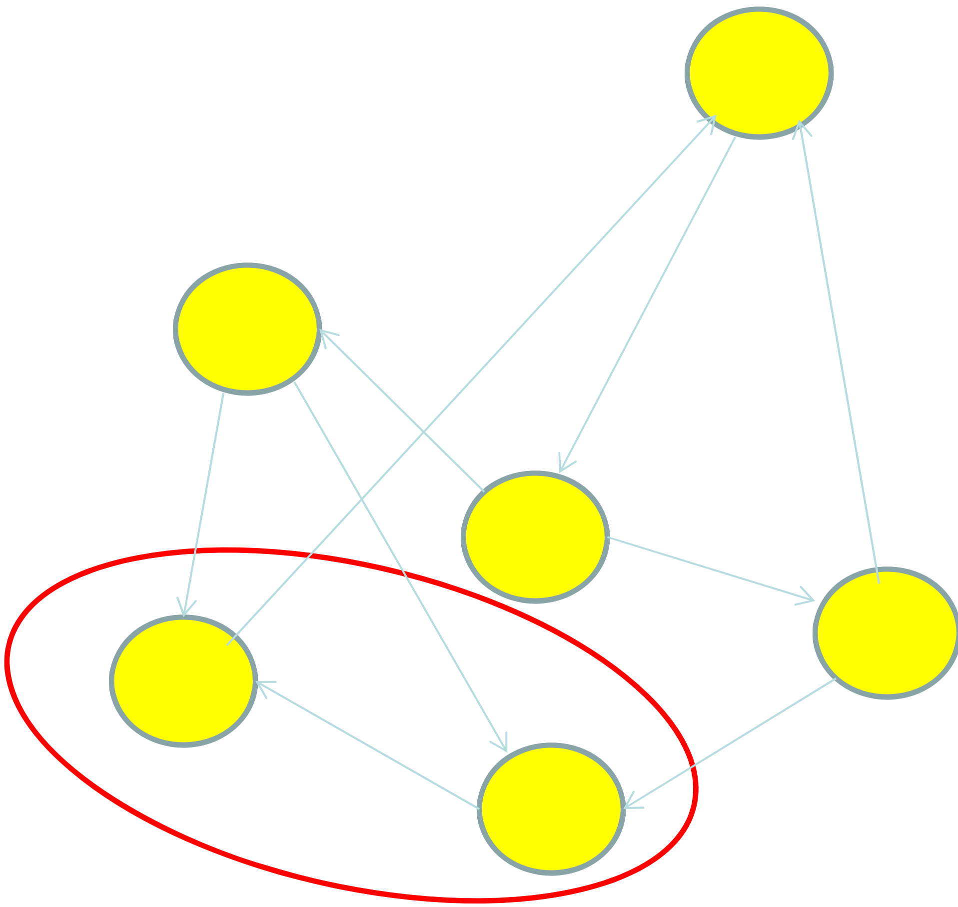
S = {2, 3, 5}

F(S) = 11

Maximizing non-monotone sub-modular functions

Source: Feige et. al., 2008

Local Search



Add to S/Delete from S, if F(S) improves

S = {2, 5}

F(S) = 12

Maximizing non-monotone sub-modular functions

Source: Feige et. al., 2008

Marketing Strategies

- Simple Greedy Strategy
  - Initialize set $S = \phi.$
  - Among buyers who accept the offered price choose buyer *i* so that $i = \arg\max_i g(S \cup \{i\}) - g(S).$
  - If $g(S \cup \{i\}) \le g(S),$ output *S*.
  - $S = S \cup \{i\}$ and go to step 2.

Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Finding Appropriate Offer Sequence

- Discount based on average degree
- Greedy Discount Approach
- Discount based on standard deviation of the degree distribution

Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Discount based on average degree

- In step *j*, where $j = 1 \ldots n,\ n = \lfloor \mu \rfloor + 1$, we offer the item with the price of $f(j-1/\mu)$ to all potential buyers.
- We prove that offering the item to $\lfloor N/\mu \rfloor$ buyers in each step, the revenue will be increased in networks with any structure.
- Consider $X(j) = \begin{cases} 1 & if\ w_{ij} > 0 \\ 0 & if\ w_{ij} = 0 \end{cases}$

$$E(\sum_{j \in S \cup \{i\}} [X(j)]) = \sum_{j \in S \cup \{i\}} E([X(j)]) \approx \sum_{j \in S \cup \{i\}} \frac{d_i.\overline{F}_{ij}}{N.\overline{F}_{ij}} = k \frac{d_i}{\mu} .$$

- Thus, the value of buyer *j* is

$$f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{iik}) \approx f_i(\frac{k.d_i.\overline{F}_{ij}}{\mu.d_i.\overline{F}_{ij}}) = f_i(\frac{k}{\mu})$$

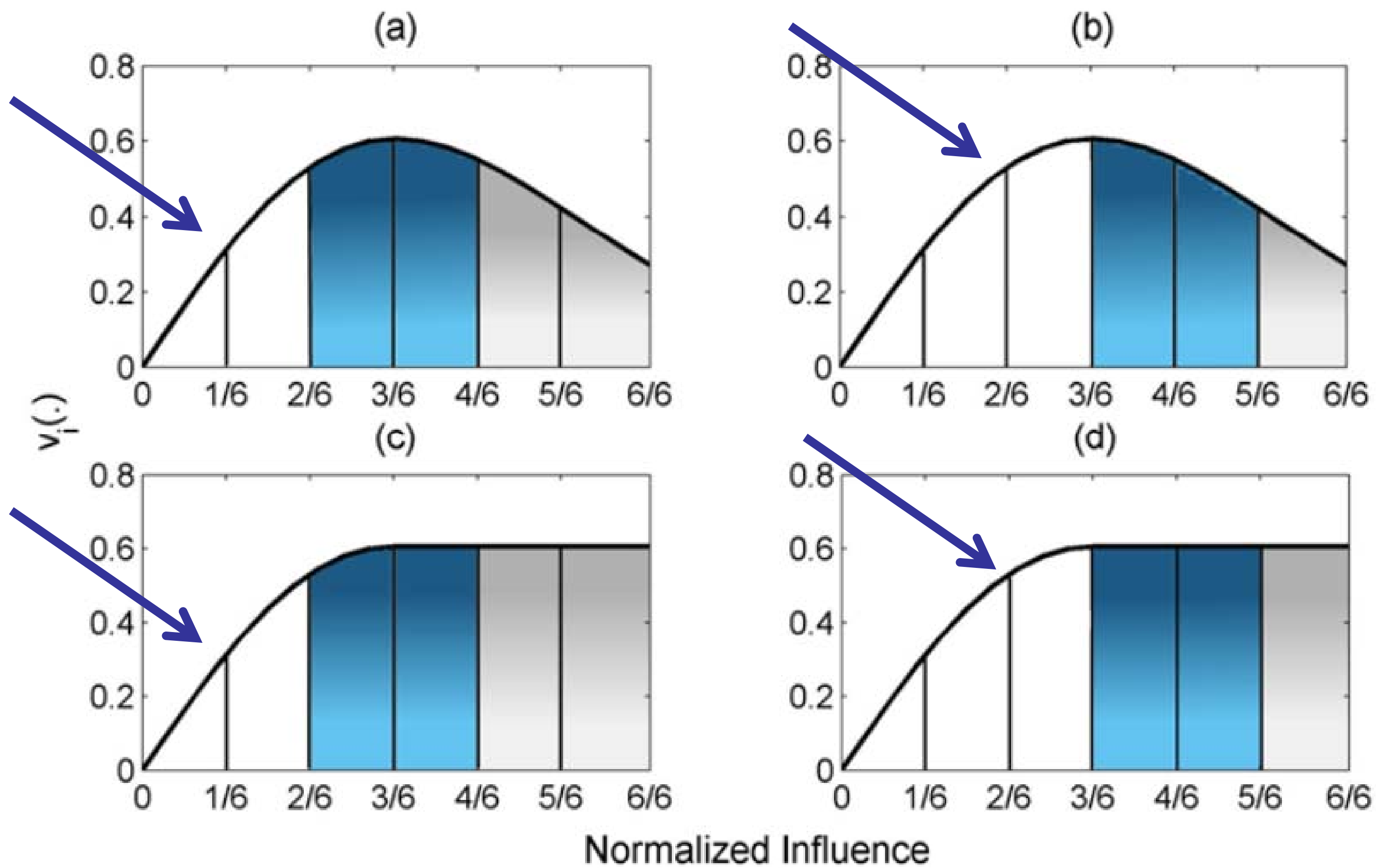
Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Discount based on Greedy Discount Approach

- Sort the nodes of a network in decreasing order of their degrees in array *D*.
- Find $k_1 = \arg\max_k (\sum_{S=D(1:k),\ i \in V/S \mid d_i > \mu/2} X(i))$ where $X(i) = \begin{cases} 1 & \frac{2}{6} \le v_i(S) < \frac{4}{6} \\ 0 & otherwise \end{cases}$ for all $i \in V \setminus S$.
- Find $k_2 = \arg\max_k (\sum_{S=D(1:k),\ i \in V/S \mid d_i > \mu/2} X(i))$ where $X(i) = \begin{cases} 1 & \frac{3}{6} \le v_i(S) < \frac{5}{6} \\ 0 & otherwise \end{cases}$ for all $i \in V \setminus S$
- Give away the item for free to the first k_1 buyers. Next, offer the item with the price of *f*(1/6) to the following buyers until k_2 - k_1 buyers in set accept the offer. Then, offer the item with the price of *f*(2/6) to the remaining potential buyers. The buyers are chosen by revenue maximization algorithm.

Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Greedy Discount Approach



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

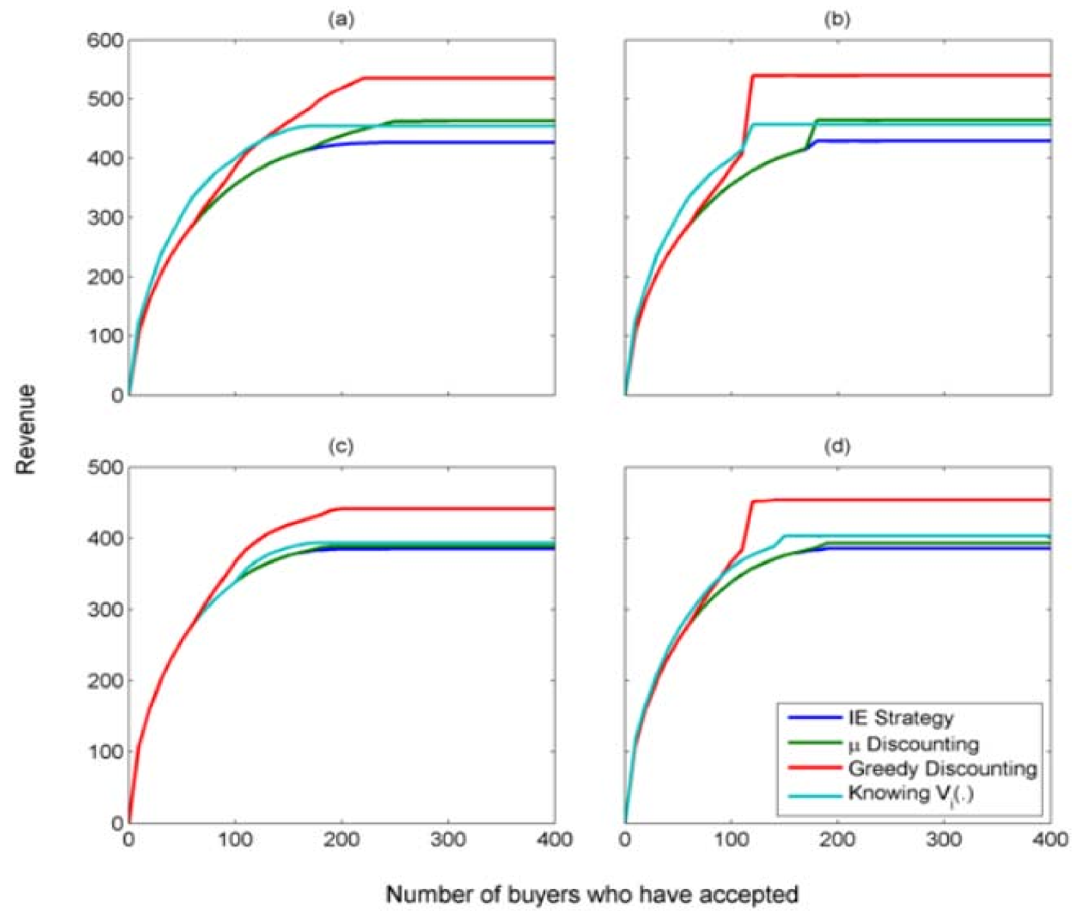
Discount based on standard deviation of the degree distribution

- Giving the item for free to all nodes i for which $d_i > \mu+$ as the nodes with very high degree or the most influential nodes in a network.
- Then, we can offer the item with the price of $f(1/6)$ to all nodes i for which $\mu+ > d_i > \mu$.
- From this point on, we offer the item with the price of $f(2/6)$ to all remaining buyers in set

Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

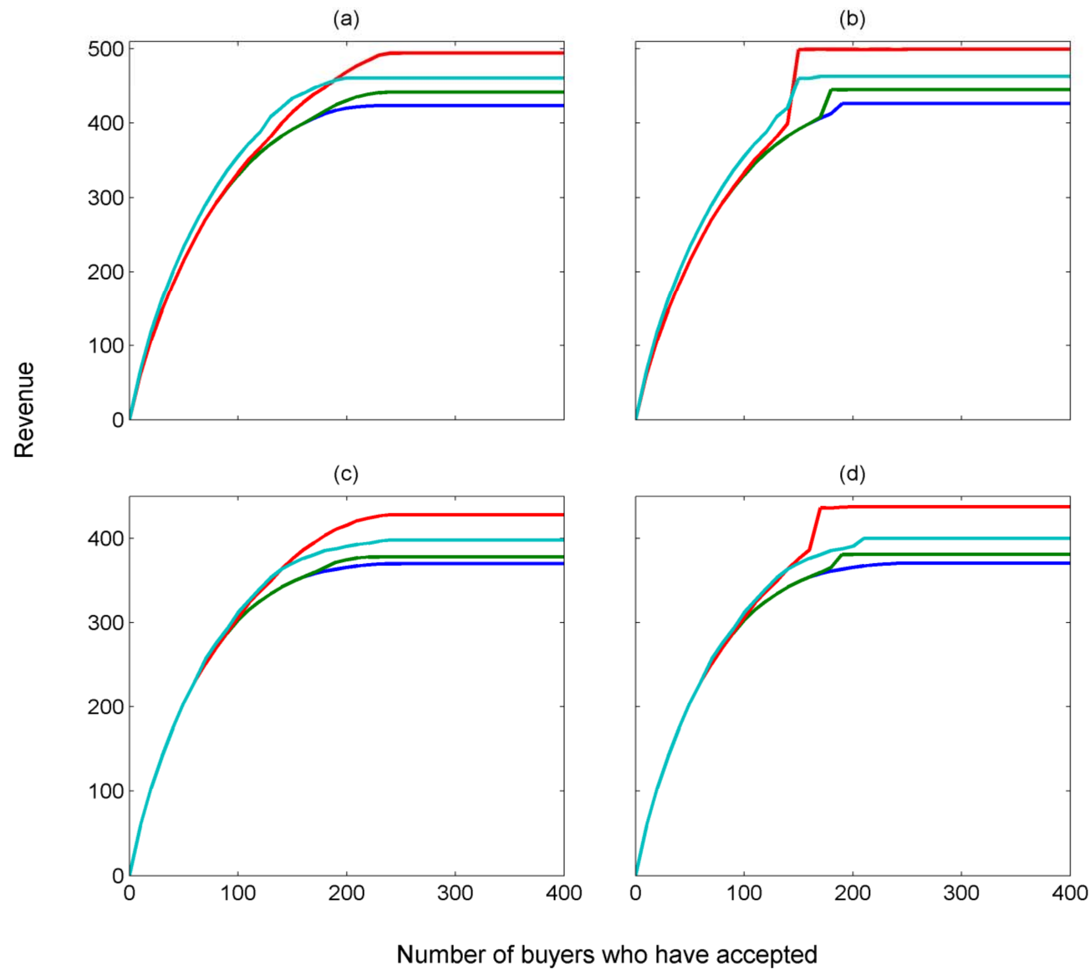
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the forest-fire network with 1000 nodes, p=.37 and p_b=.32.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

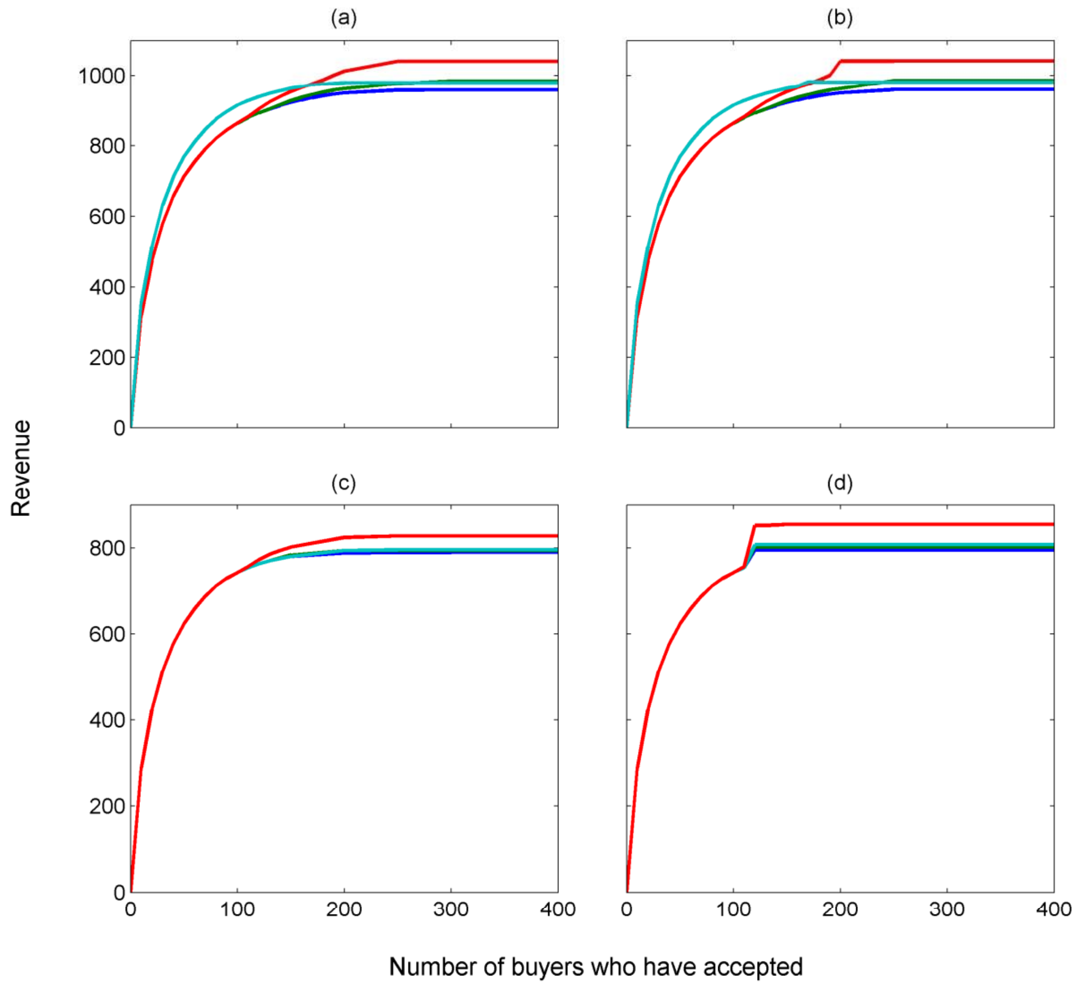
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the modular forest-fire network that has three modules with 200,300 and 500 node. Each module constructed with p=.37 and p_b=.32. The inter-modular rewiring probability is P=.01.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

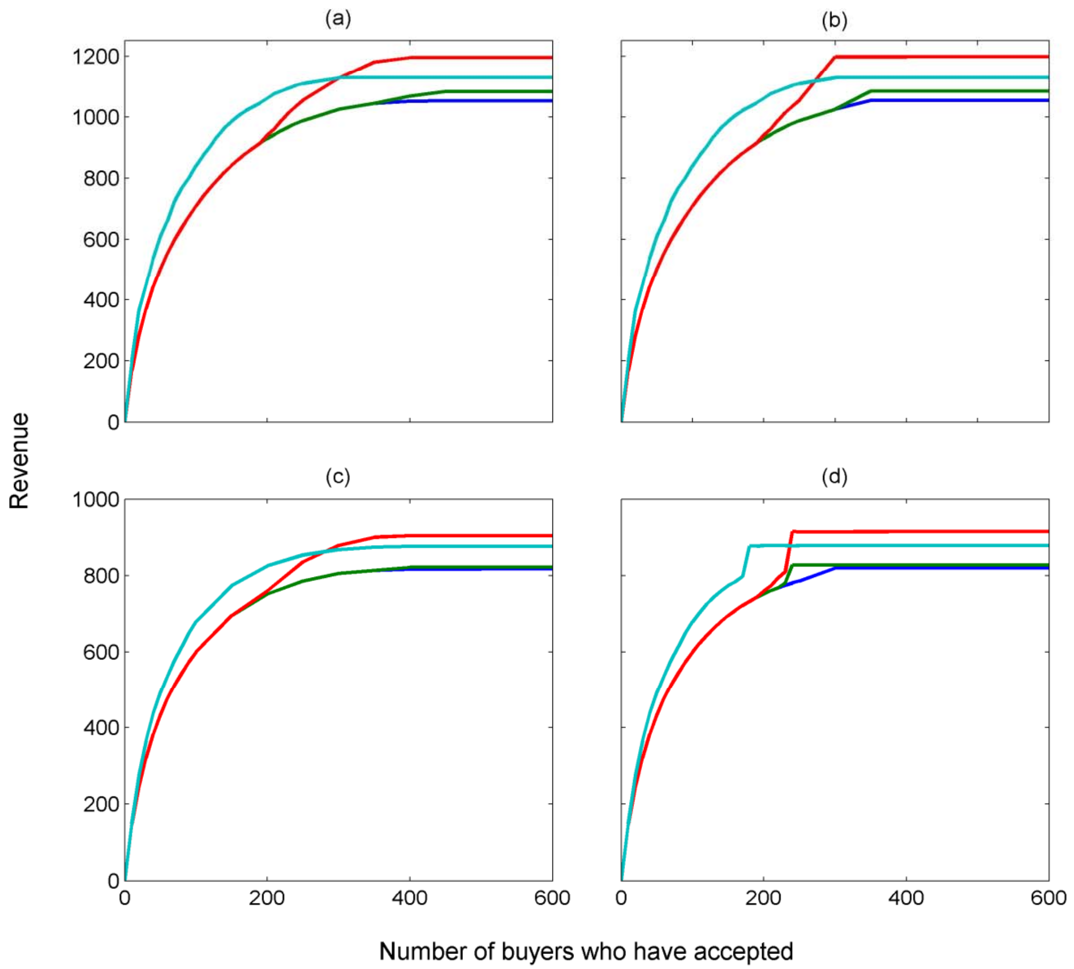
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the Facebook-like social network with 1,899 nodes and 20,296 edges.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

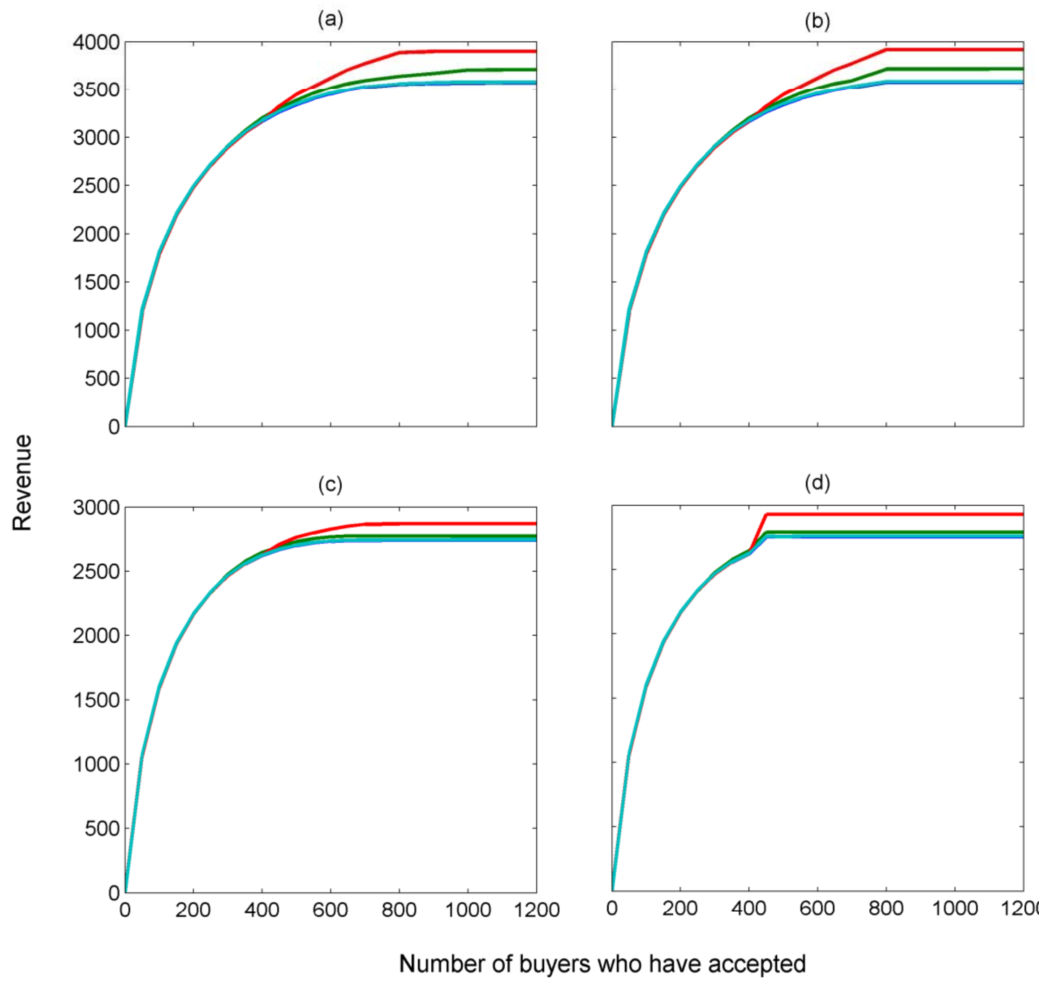
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the yeast protein interaction network with 2,224 nodes and 6,829 edges.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

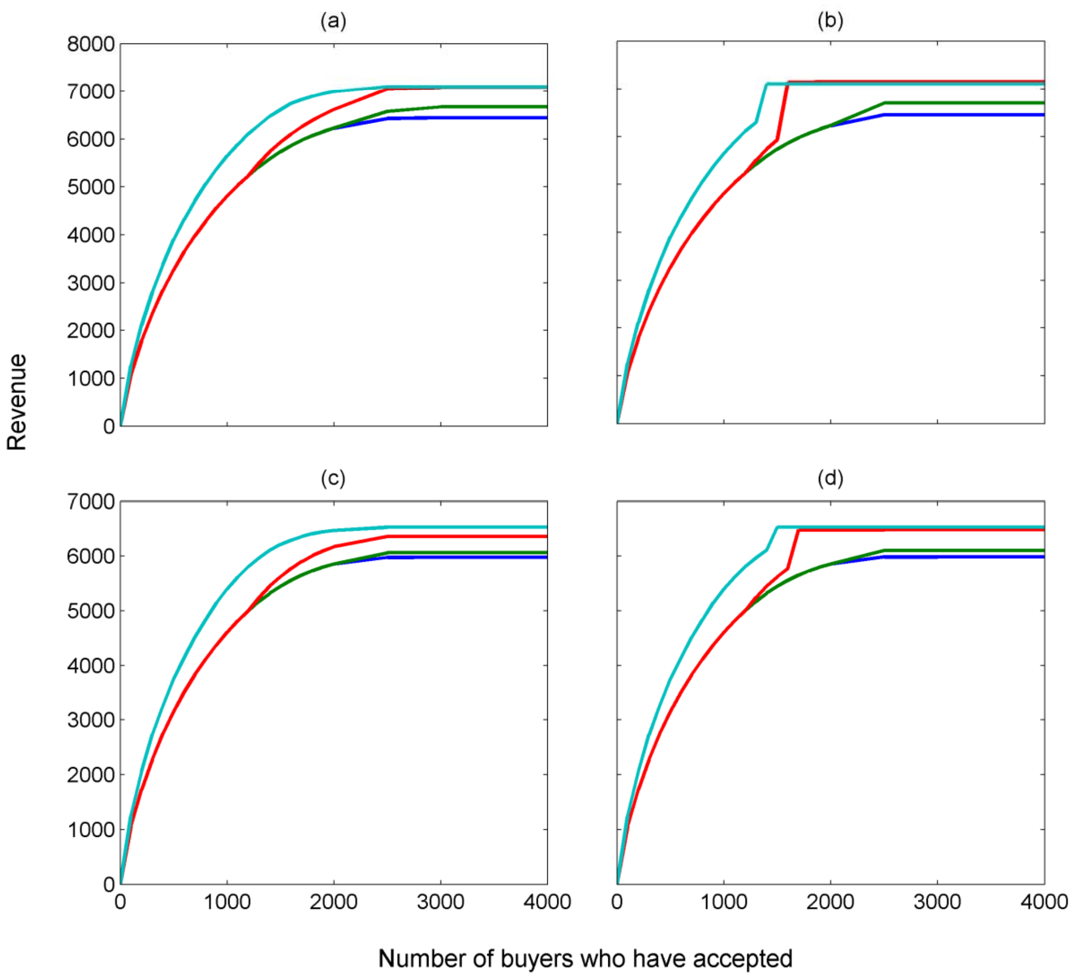
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the wiki-vote network with 7,115 nodes and 103,689 edges.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

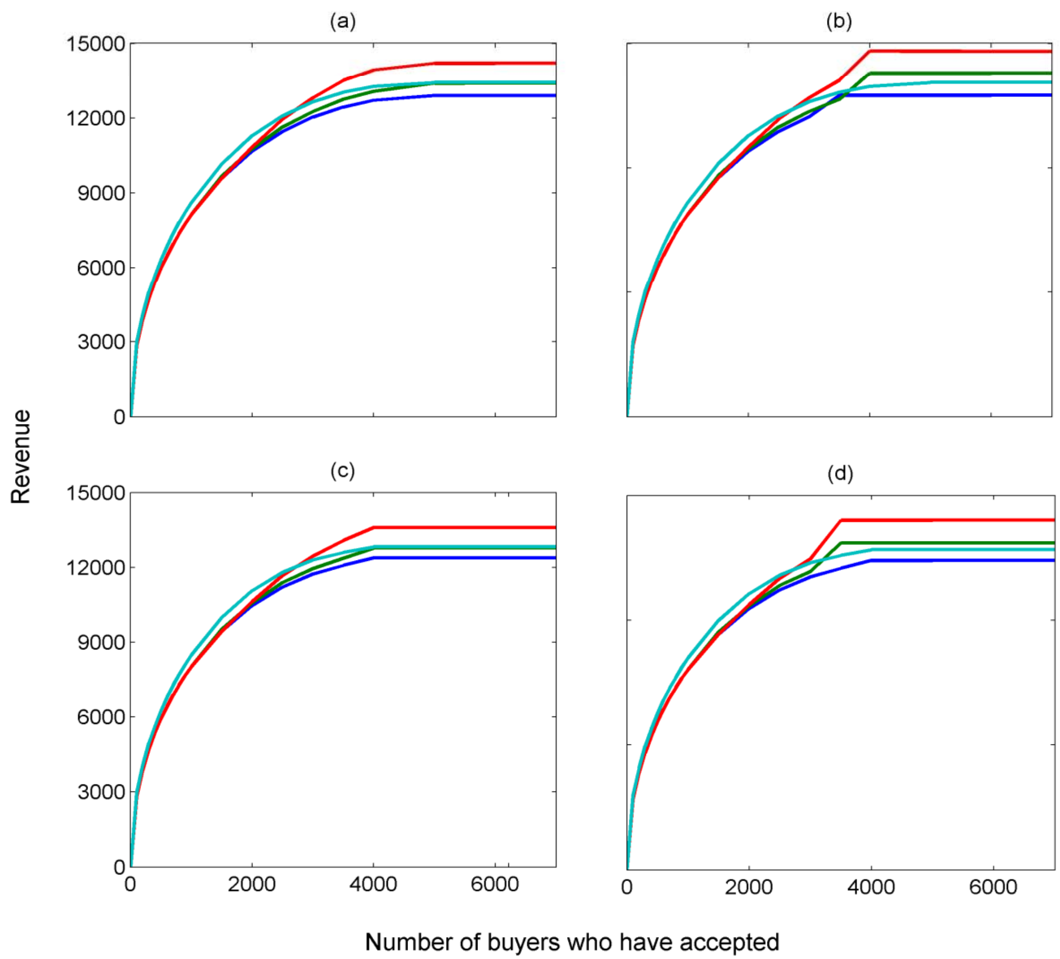
Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the Newman scientific collaboration network with 16,726 nodes and 47,594 edges.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Experiments

Revenue from the monotone concave influence model (top row) and non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set *S*, for the High-energy physics theory citation network with 27,770 nodes and 352,807 edges.



Source: Babaei , Mirzasoleiman, Jalili, and Safari, 2011

Epidemics spreading

- The patterns of epidemics spreading is determined by
  - the properties of the pathogen carrying it
  - the length of its infectious period
  - its severity
  - network structures within the population it is affecting: **contact network**
- there is a node for each person, and an edge if two people come into contact with each other in a way that makes it possible for the disease to spread from one to the other.
- how travel patterns within a city or via the worldwide airline network could affect the spread of a fast-moving disease?
- how diseases spread through animal populations?
- ...

Connections to the diffusion

- There are clear connections between epidemic disease and the diffusion of ideas through social networks
- Both diseases and ideas can spread from person to person, across similar kinds of networks that connect people
- They exhibit very similar structural mechanisms
- The difference between the two:
  - People are making decisions to adopt a new idea
  - With diseases, not only is there a lack of decision-making in the transmission of the disease from one person to another, but it is sufficiently complex and unobservable at the person-to-person level
  - We will generally assume that when two people are directly linked in the contact network, and one of them has the disease, there is a given probability that he or she will pass it to the other.

Two levels

•**Microscopic level**
Researchers who disassemble and try to kill off new viruses.

<u>Corresponds to the quest for new vaccines and medicines</u>

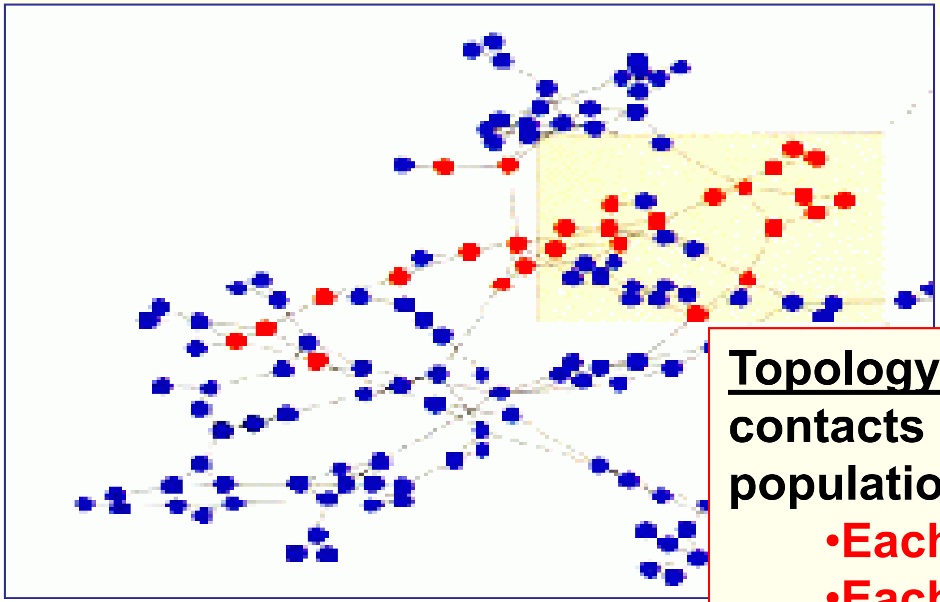
•**Macroscopic level**
Statistical analysis and modeling of epidemiological data in order to find informations and policies aimed at lowering epidemic outbreaks

<u>Macroscopic prophylaxis , Vaccination campaigns</u>

Mathematical models of epidemics

Coarse grained description of individuals and their state

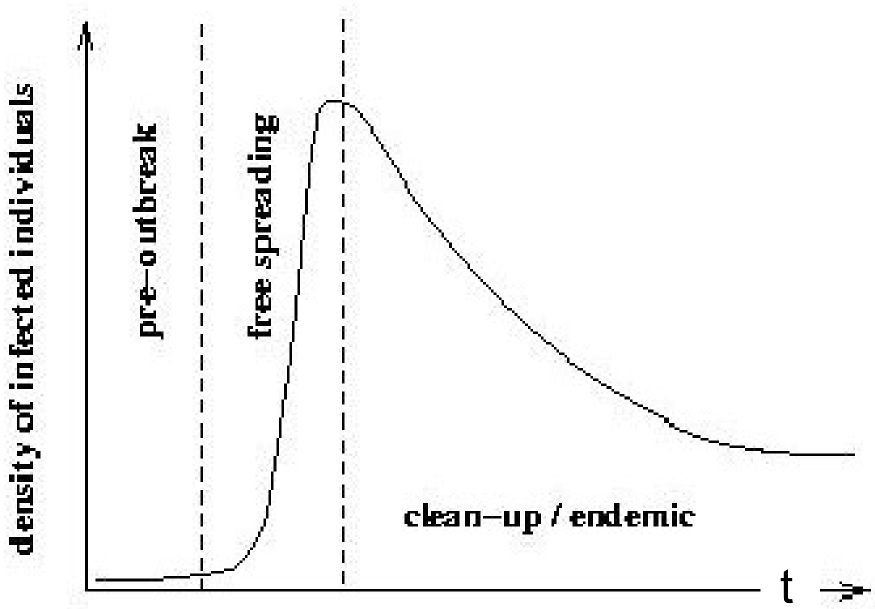
- **Individuals exist only in few states:**
- **Healthy or Susceptible * Infected * Immune * Dead**
- **Particulars on the infection mechanism on each individual are neglected.**



Topology of the system: the pattern of contacts along which infections spread in population is identified by a network

- **Each node represents an individual**
- **Each link is a connection along which the virus can spread**

Stages of an epidemic outbreak



Infected individuals => prevalence/incidence

Branching process

First wave:

- Suppose that a person carrying a new disease enters a population
- He transmits it to each person he meets independently with a probability of p.
- Further, suppose that he meets k people while he is contagious; let's call these k people the first wave of the epidemic.
- Some of the people in the first wave may be infected with the disease, while the rest are not.

Branching process

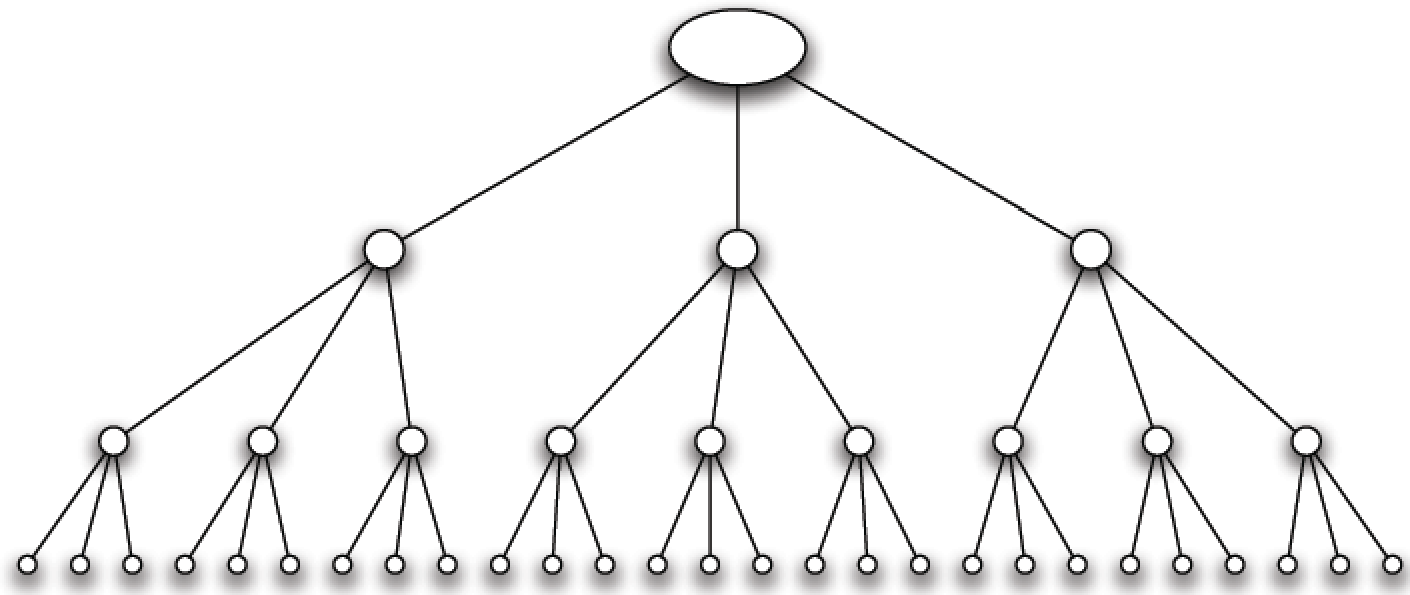
Second wave:

- Now, each person in the first wave goes out into the population and meets k different people, resulting in a second wave of $k \times k = k^2$ people
- Each infected person in the first wave passes the disease independently to each of the k second-wave people they meet, again independently with probability p

Subsequent waves:

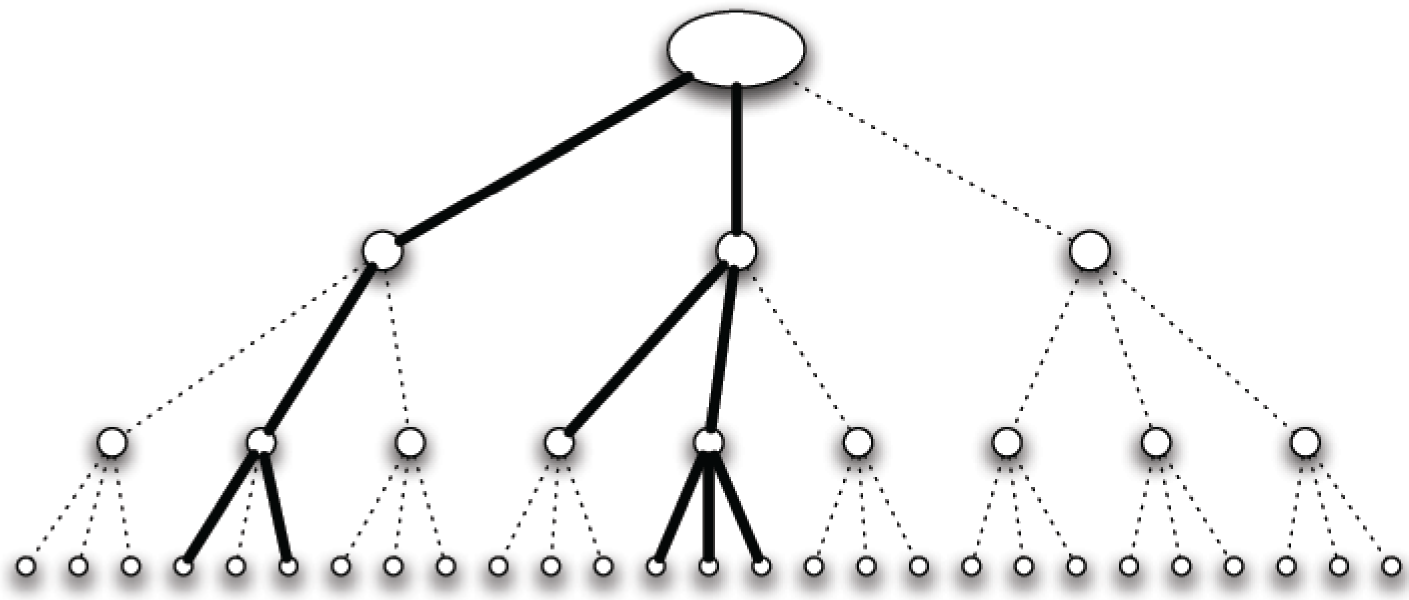
- Further waves are formed in the same way, by having each person in the current wave meet k new people, passing the disease to each independently with probability p

Branching process



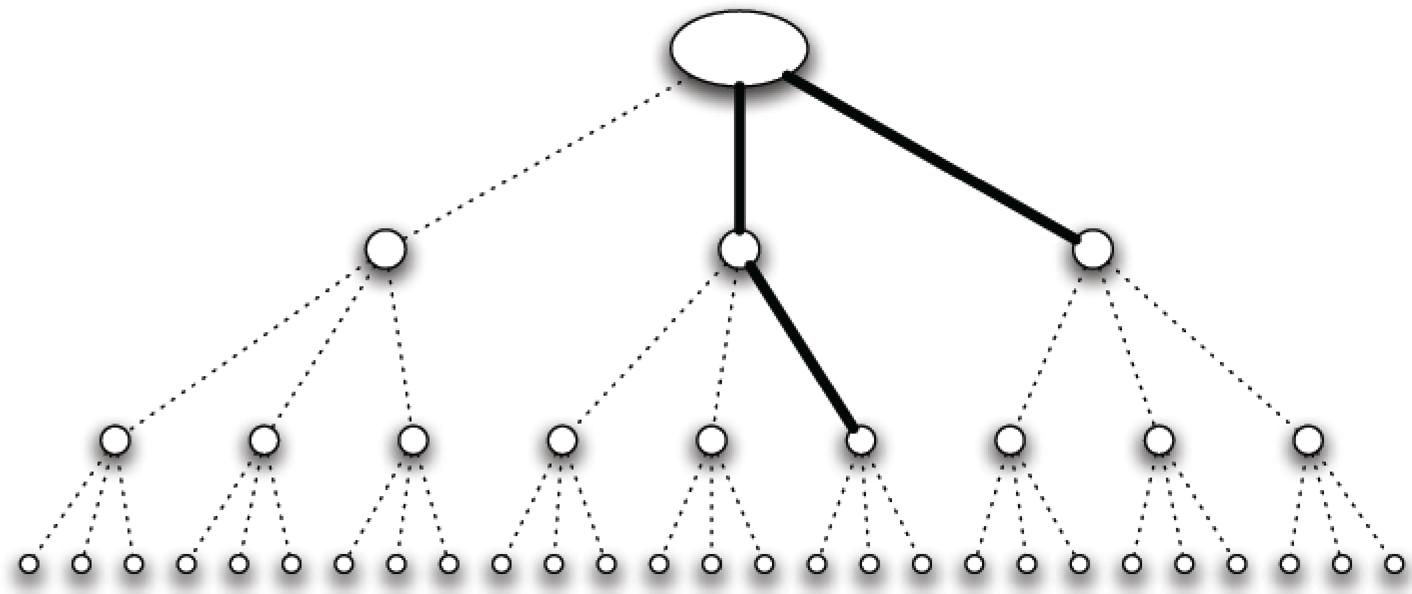
(a) *The contact network for a branching process*

Branching process



(b) *With high contagion probability, the infection spreads widely*

Branching process



(c) *With low contagion probability, the infection is likely to die out quickly*

The fundamental process

- If the disease in the process ever reaches a wave where it fails to infect anyone, then it has died out:
  - Since people in future waves can only catch the disease from others higher up in the tree, no one in any future wave will be infected either.
- So there are really only two possibilities for a disease in the branching process model:
  - It reaches a wave where it infects no one, thus dying out after a finite number of steps
  - Or it continues to infect people in every wave, proceeding infinitely through the contact network
- There is a simple condition to tell these two possibilities apart, based on a quantity called the basic reproductive number of the disease.

The fundamental process

- The basic reproductive number, denoted R_0, is the expected number of new cases of the disease caused by a single individual.
- Since in our model everyone meets k new people and infects each with probability p, the basic reproductive number here is given by $R_0 = pk$.
- The outcome of the disease in a branching process model is determined by whether the basic reproductive number is smaller or larger than 1.

The fundamental process

Claim:

- If $R_0 < 1$, then with probability 1, the disease dies out after a finite number of waves.
- If $R_0 > 1$, then with probability greater than 0 the disease persists by infecting at least one person in each wave.

Rumor Spreading

- Spread as fast as possible

- DK model
  - I : Ignorant
  - S : Spreader
  - R : Stifler

$$I(i) + S(j) \xrightarrow{\lambda} S(i) + S(j)$$
$$S(i) + S(j) \xrightarrow{\alpha} R(i) + S(j)$$
$$S(i) + R(j) \xrightarrow{\alpha} R(i) + R(j)$$

$$\frac{di(t)}{dt} = -\lambda \langle k \rangle i(t) s(t)$$

$$\frac{ds(t)}{dt} = \lambda \langle k \rangle i(t) s(t) - \alpha \langle k \rangle s(t)[s(t) + r(t)]$$

$$\frac{dr(t)}{dt} = \alpha \langle k \rangle s(t)[s(t) + r(t)]$$

SIR model

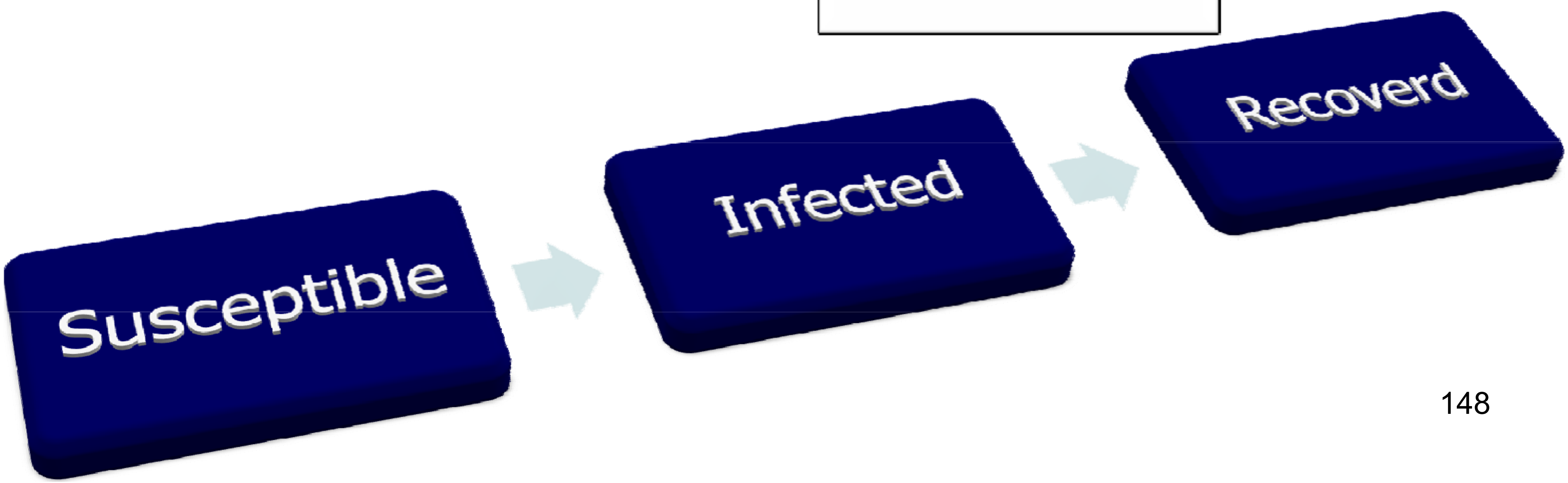
- Each node may be in the following states
  - Susceptible: healthy but not immune
  - Infected: has the virus and can actively propagate it
  - Recovered: (or Removed/Immune/Dead) had the virus but it is no longer active
- Infection rate p: probability of getting infected by a neighbor per unit time
- Immunization rate q: probability of a node getting recovered per unit time
- It can be shown that virus propagation can be reduced to the bond-percolation problem for appropriately chosen probabilities
  - again, there is no percolation threshold for scale-free graphs

SIR Model

- S: Susceptible
- I: Infected
- R:Recovered

$$S(i) + I(j) \xrightarrow{\lambda} I(i) + I(j)$$

$$I(i) \xrightarrow{\mu} R(i)$$



SIR Model

- $s(t)$, $\rho(t)$, $r(t)$ are the fraction of susceptible, infected and removed individuals at time t
- $s(t) + \rho(t) + r(t) = 1$
- $\bar{k}$ is the number of contacts per unit time that is supposed to be constant for the whole population

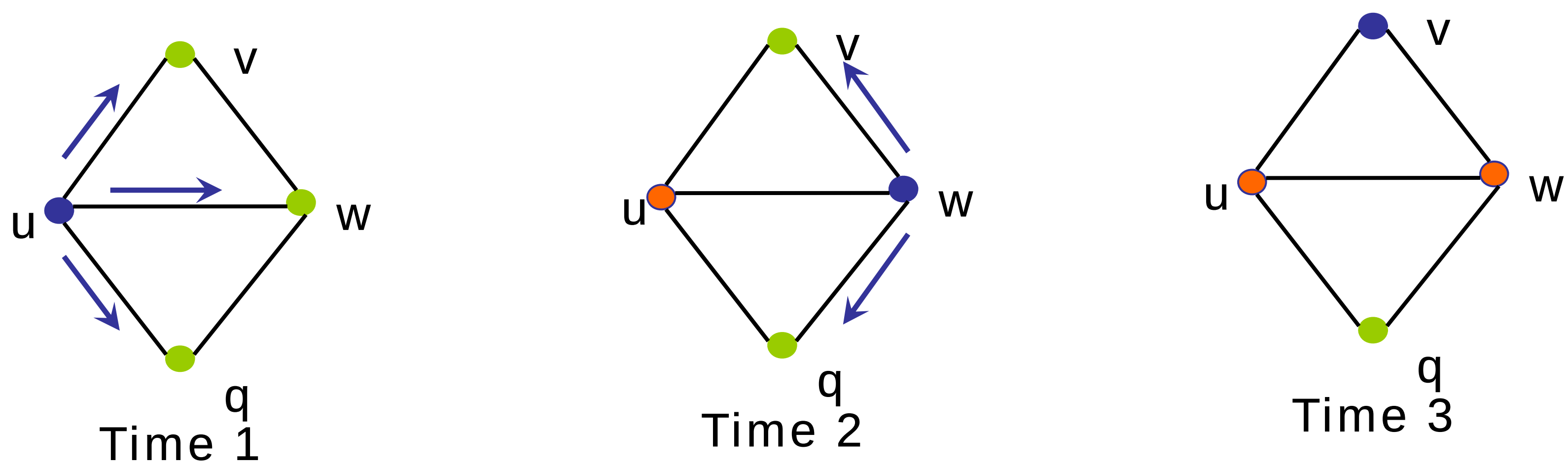
$$\frac{ds(t)}{\mathrm{dt}} = -\lambda \bar{k} \rho(t) s(t)$$

$$\frac{d\rho(t)}{dt} = \lambda \bar{k} \rho(t) s(t) - \mu \rho(t)$$

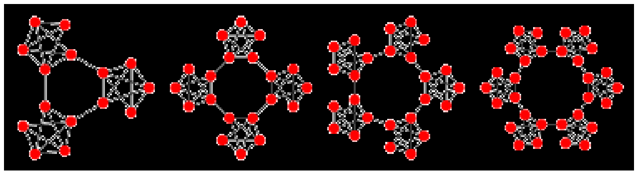
$$\frac{dr(t)}{dt} = \mu \rho(t)$$

A simple SIR model

- Time proceeds in discrete time-steps
- If a node is infected at time t it infects all its neighbors with probability p
- Then the node becomes recovered (q = 1)



The caveman small-world graphs



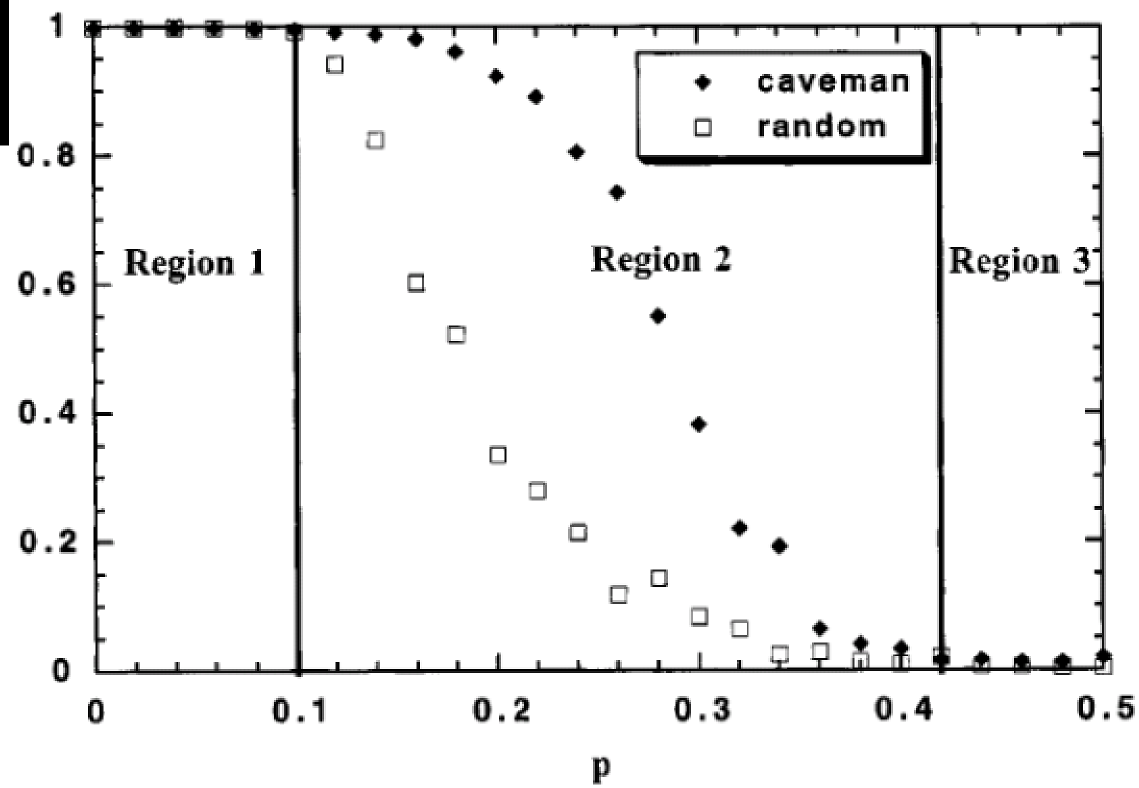


FIG. 11.—Fraction of uninfected survivors (F_s) versus infectiousness (p) for disease spreading dynamics on a network generated by the α-model at clustered and random extremes.

SIS model

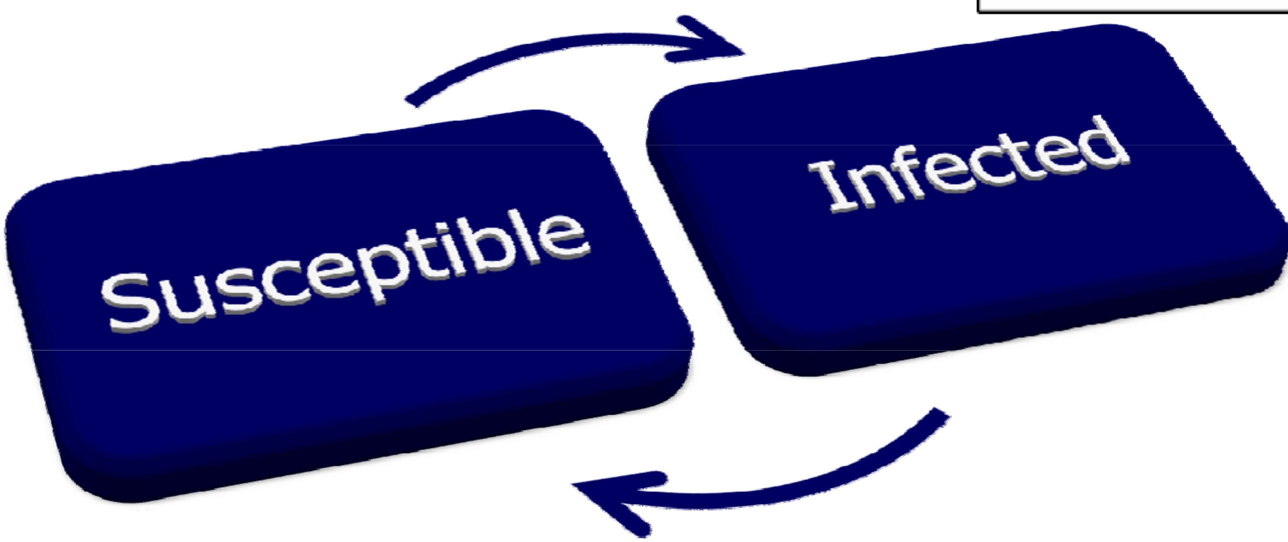
- Susceptible-Infected-Susceptible:
  - each node may be healthy (susceptible) or infected
  - a healthy node that has an infected neighbor becomes infected with probability p
  - an infected node becomes healthy with probability q
  - spreading rate r=p/q

SIS Model

- S: Susceptible
- I: Infected

$$S(i) + I(j) \xrightarrow{\lambda} I(i) + I(j)$$

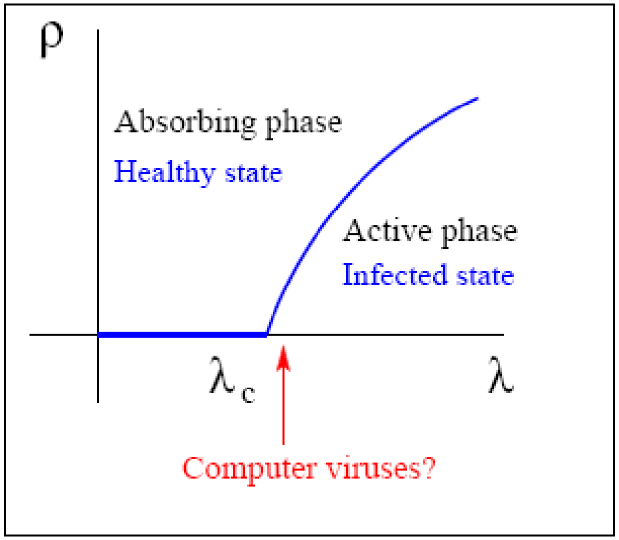
$$I(i) \xrightarrow{\mu} S(i)$$



Epidemic Threshold

- The epidemic threshold for the SIS model is a value r_c such that for $r < r_c$ the virus dies out, while for $r > r_c$ the virus spreads.
- For homogeneous graphs,

$$r_c = \frac{1}{\langle k \rangle}$$

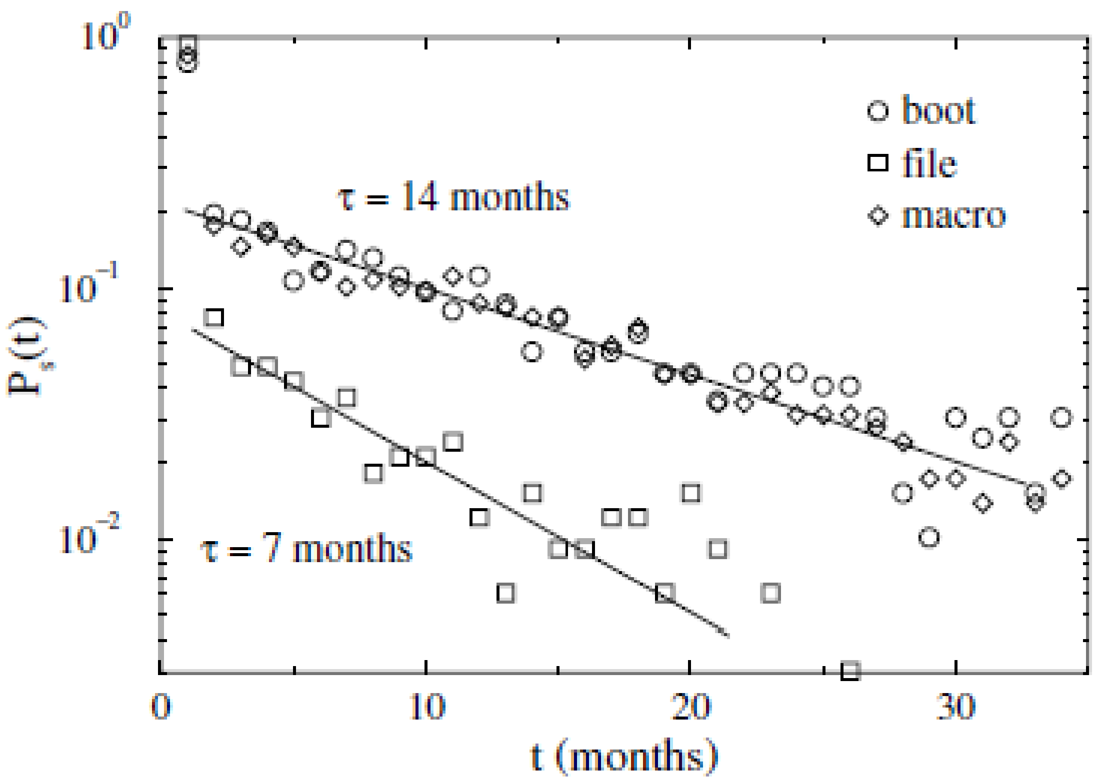


- For scale free graphs

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

- For exponent less than 3, the variance is infinite, and the epidemic threshold is zero

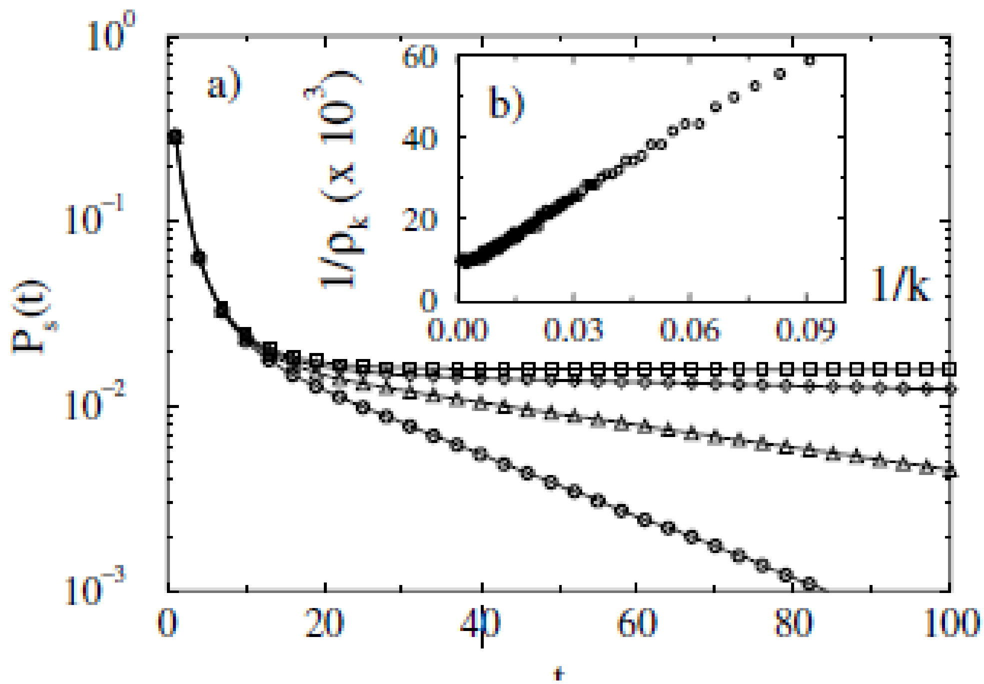
Real systems



Surviving probability for viruses in the wild. The 814 different viruses analyzed . The presence of an exponential decay is evident in the plot, with characteristic time t.

Source: Pastor-Satorras and Vespignani, PRL 2001

Real systems



(a) Surviving probability P_{st} for a spreading rate of **0.065** in scale-free networks of different sizes. The exponential behavior, following a sharp initial drop, is compatible with the data analysis of real systems.

Source: Pastor-Satorras and Vespignani, PRL 2001

An eigenvalue point of view

- Time proceeds in discrete time-steps. At time t,
  - an infected node u infects a healthy neighbor v with probability p.
  - node u becomes healthy with probability q
- If A is the adjacency matrix of the network, then the virus dies out if

$$\lambda_1(A) \leq \frac{q}{p}$$

- That is, the epidemic threshold is $r_c = 1/\lambda_1(A)$

Multiple copies model

- Each node may have multiple copies of the same virus
  - **v**: state vector
    - v_i : number of virus copies at node i

- At time $t = 0$, the state vector is initialized to $\mathbf{v}^0$
- At time t,

  For each node i

  For each of the v_i^t virus copies at node i

  the copy is propagated to a neighbor j with prob p

  the copy dies with probability q

Analysis

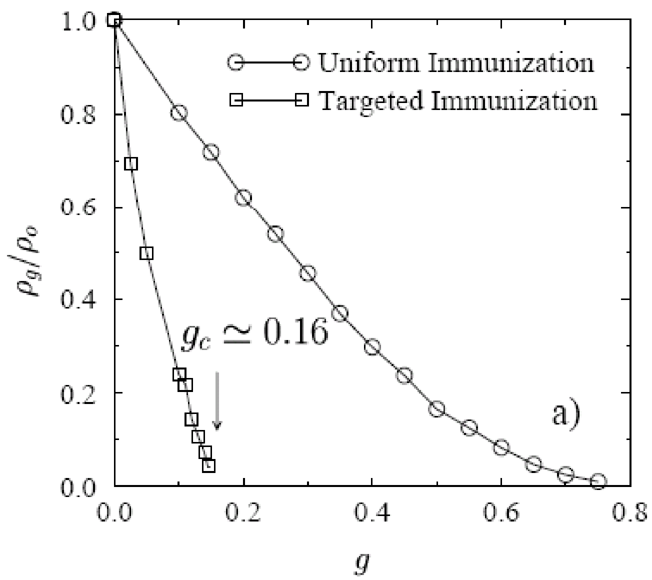
- The expected state of the system at time t is given by

$$\overline{\mathbf{v}^{t}} = \left(p\mathbf{A} + (1-q)\mathbf{I}\right)\overline{\mathbf{v}^{t-1}}$$

- As t → ∞
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) < 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) < q/p$ then $\overline{\mathbf{v}^{t}} \rightarrow 0$
    - the probability that all copies die converges to 1
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) = 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) = q/p$ then $\overline{\mathbf{v}^{t}} \rightarrow \mathbf{c}$
    - the probability that all copies die converges to 1
  - if $\lambda_1\left(p\mathbf{A} + (1-q)\mathbf{I}\right) > 1 \Leftrightarrow \lambda_1\left(\mathbf{A}\right) = q/p$ then $\overline{\mathbf{v}^{t}} \rightarrow \infty$
    - the probability that all copies die converges to a constant < 1

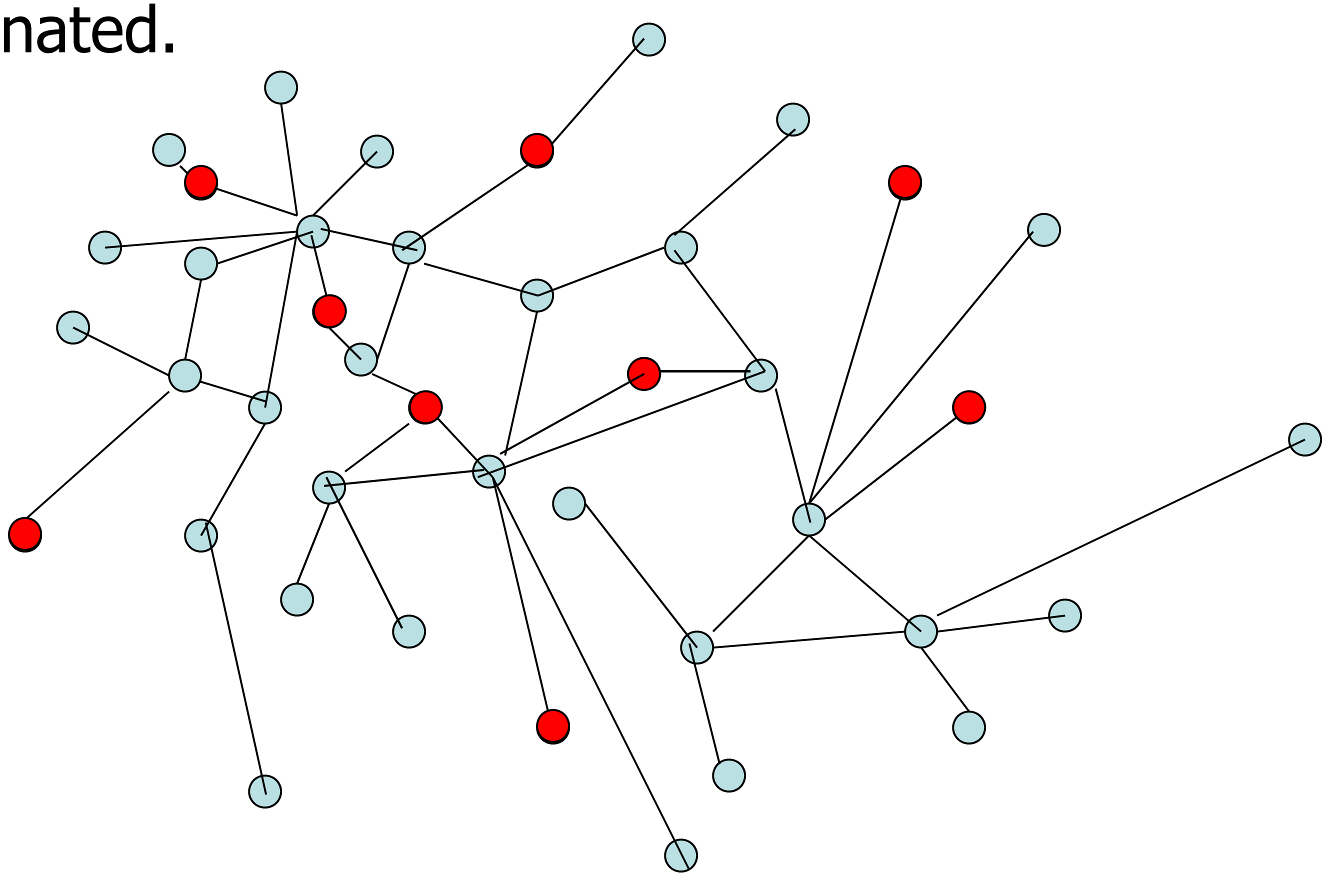
Immunization

- Given a network that contains viruses, which nodes should we immunize in order to contain the spread of the virus?
- The flip side of the percolation theory
- Uniform immunization vs Targeted immunization in scale-free networks:



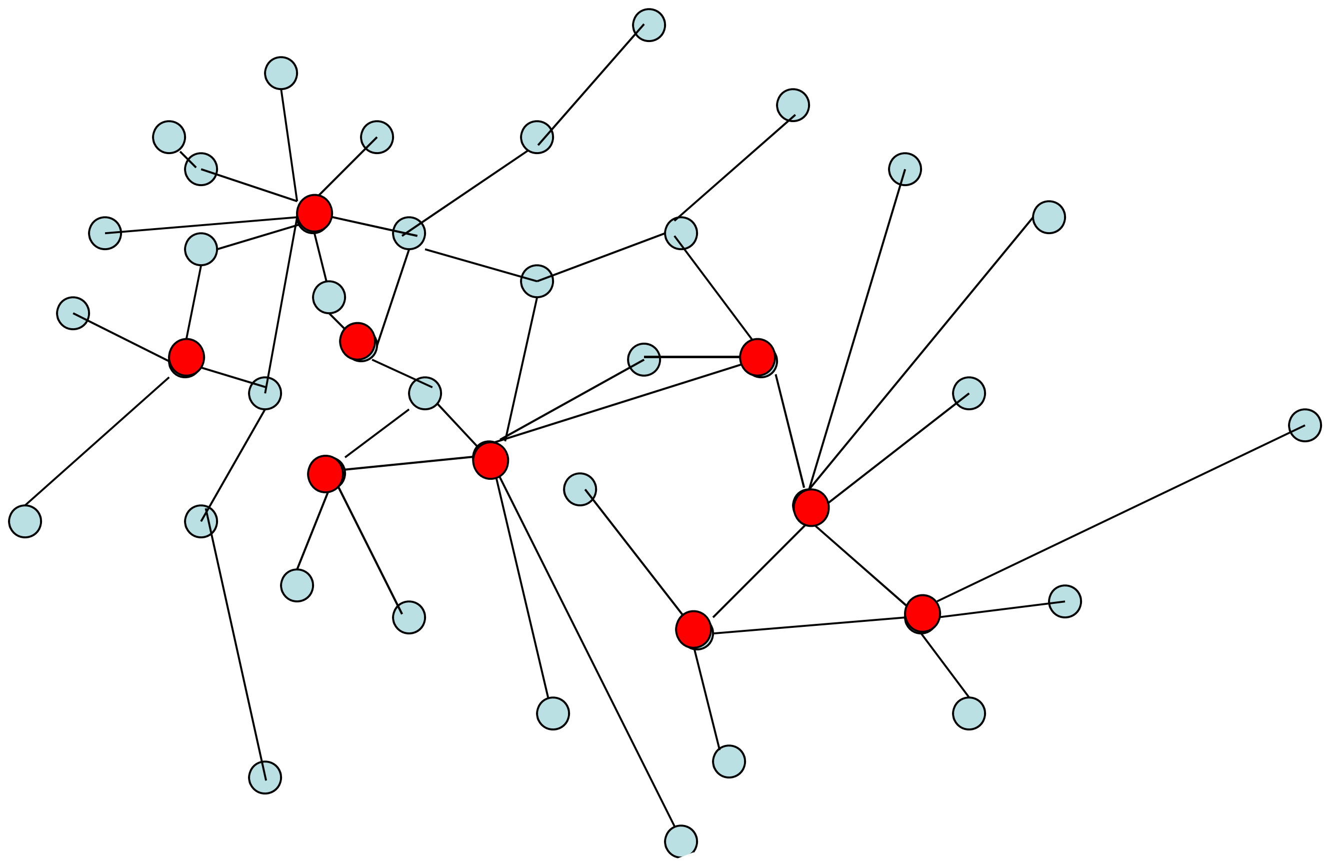
Random Immunization

- Random or uniform strategy selects all the individuals within the population with the same probability to be vaccinated.



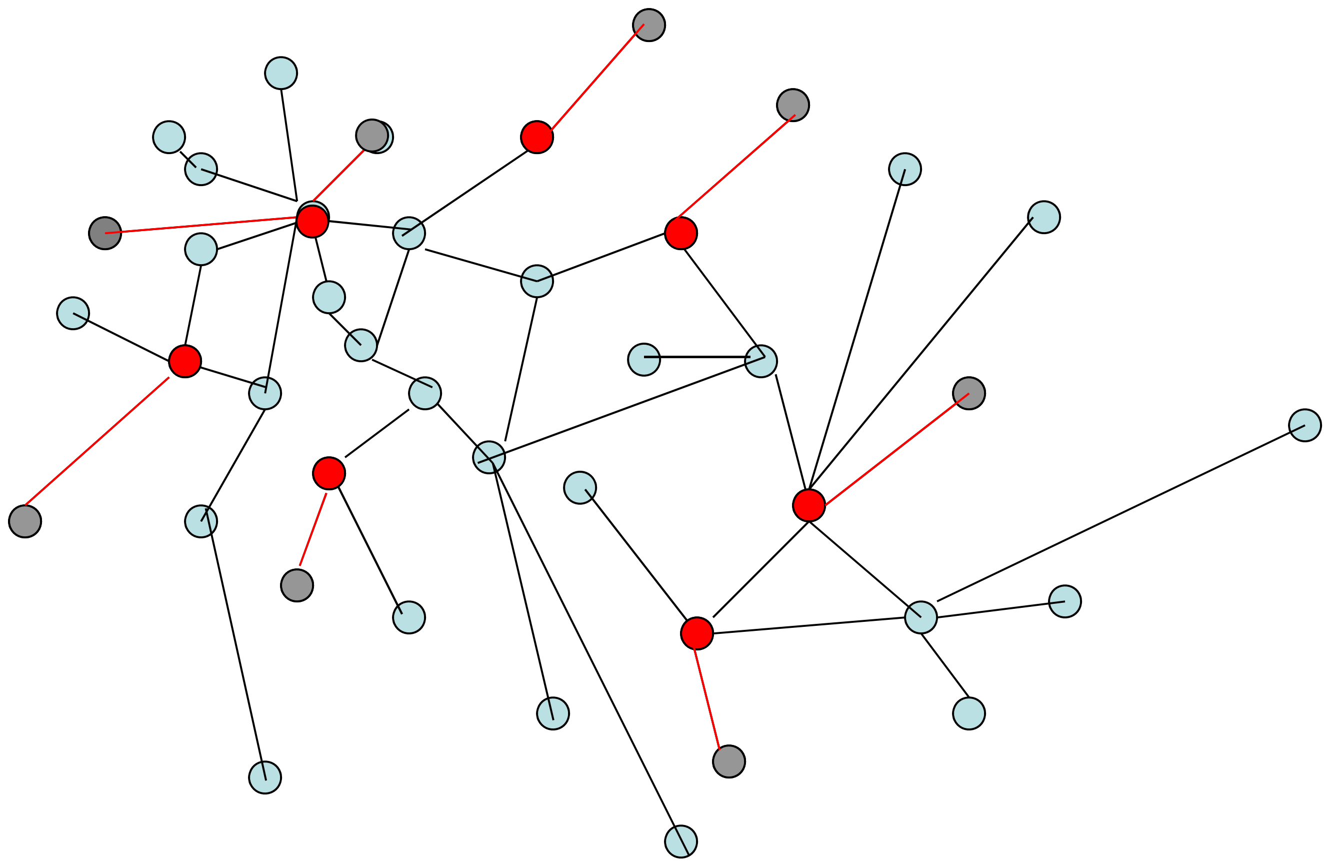
Targeted Immunization

- Targeted Immunization progressively removing the most connected nodes within the population.



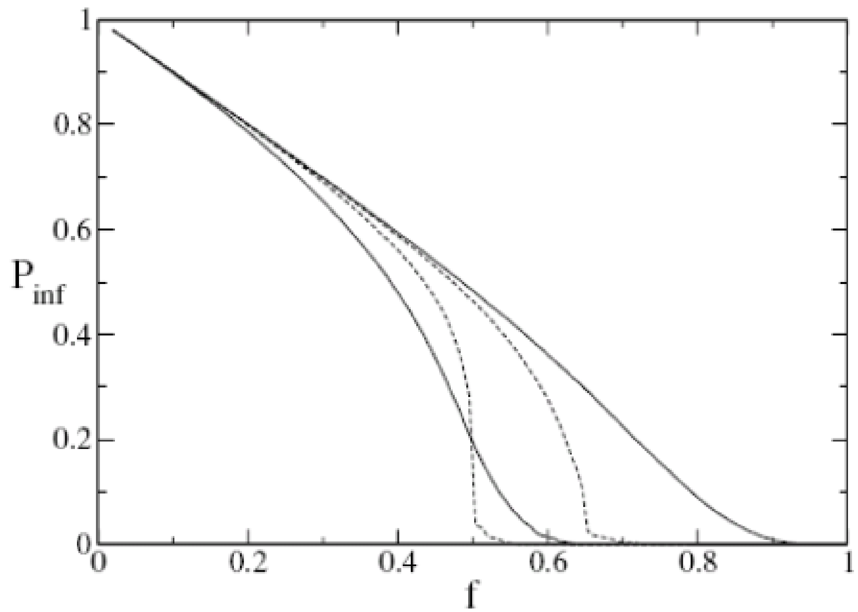
Aquaintance Immunization

- Pick a fraction f of nodes in the graph, and immunize one of their acquaintances



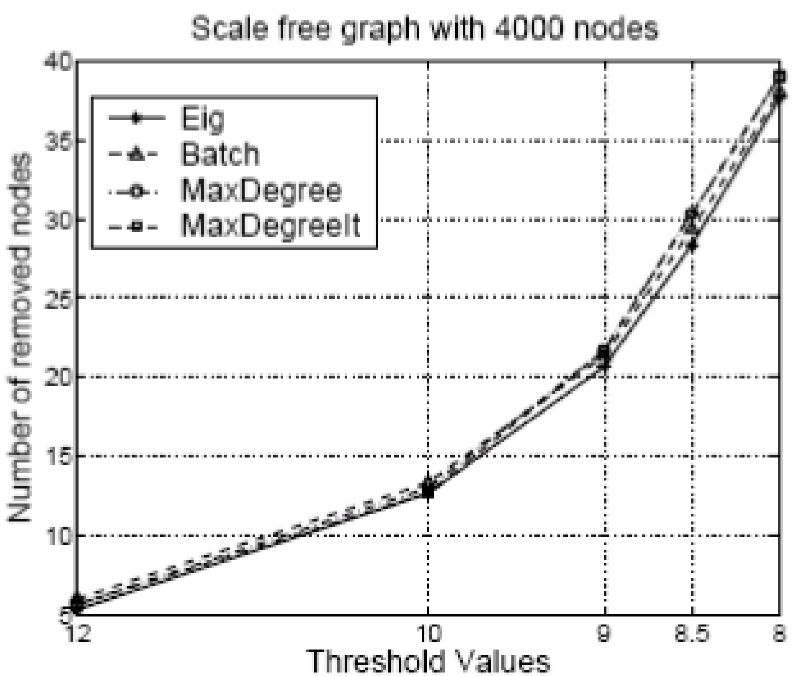
Aquaintance Immunization

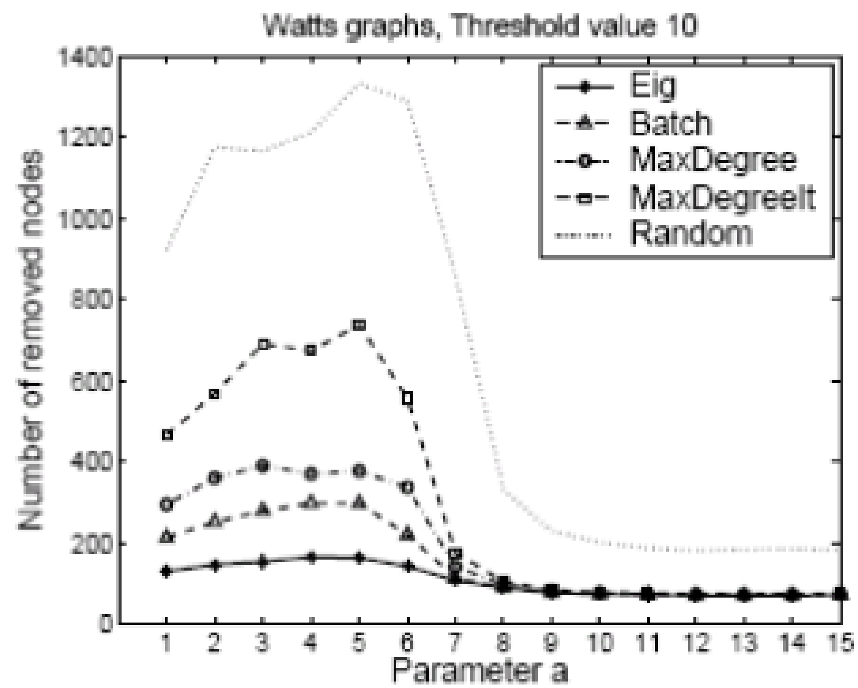
- you should gravitate towards nodes with high degree



Reducing the eigenvalue

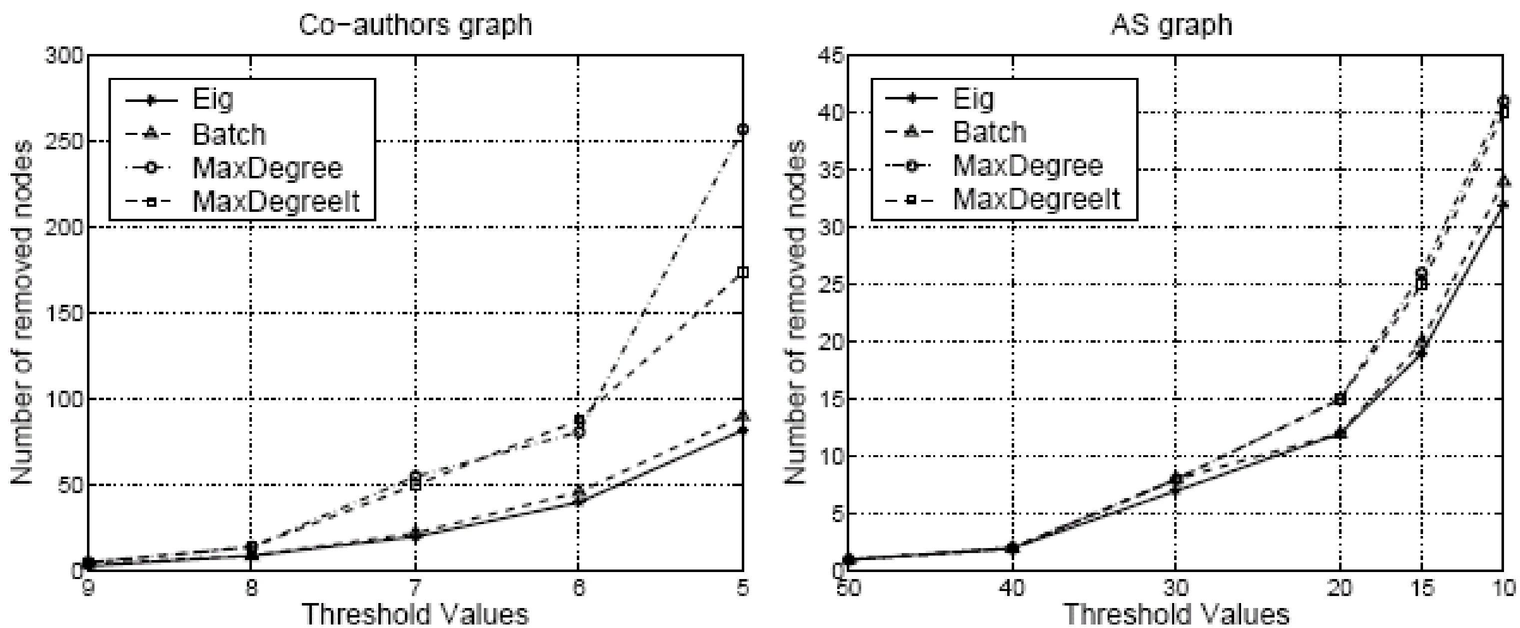
- Repeatedly remove the node with the highest value in the principal eigenvector





Reducing the eigenvalue

- Real graphs

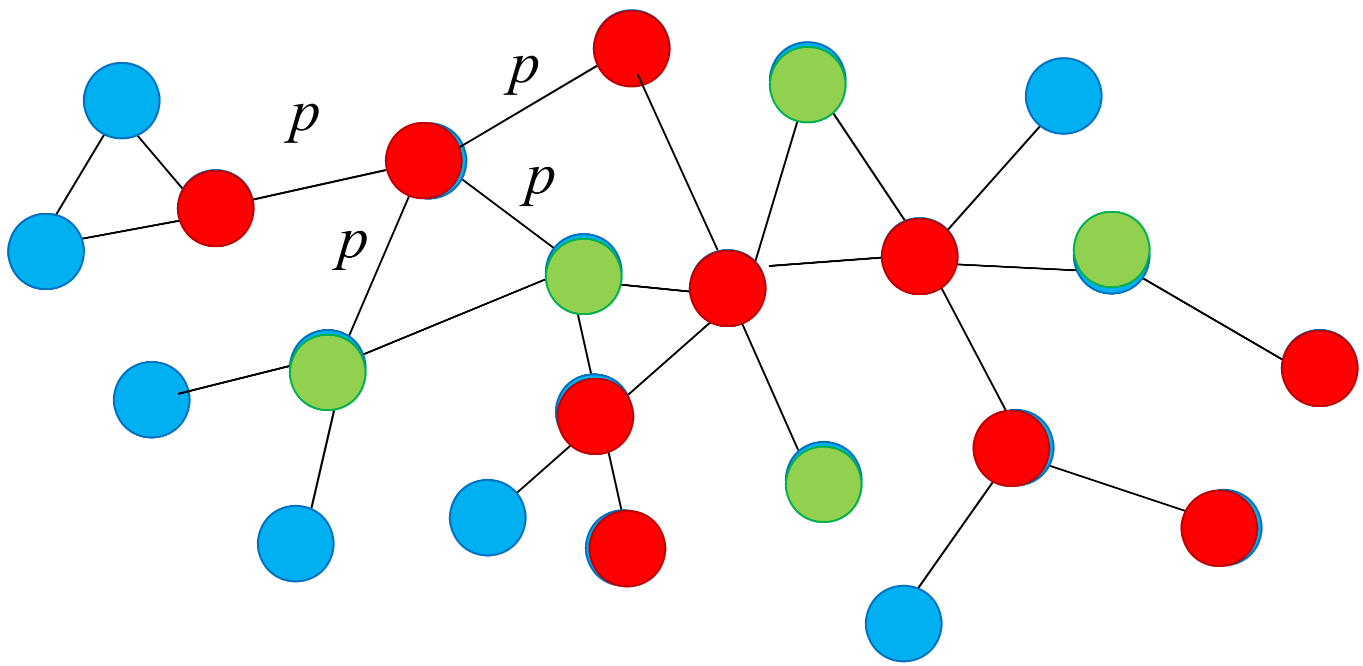


Immunizing Complex Networks with Limited Budget

- The effect of immunization strategies are often studied by ignoring the vaccination costs or assuming an infinite budget; however, this might not be the case in real situations.
- Here, we aim at a more natural objective by finding an effective immunization strategy considering a limited budget.
- Indeed, many people might be eager to pay a price to become immunized against an epidemic disease.

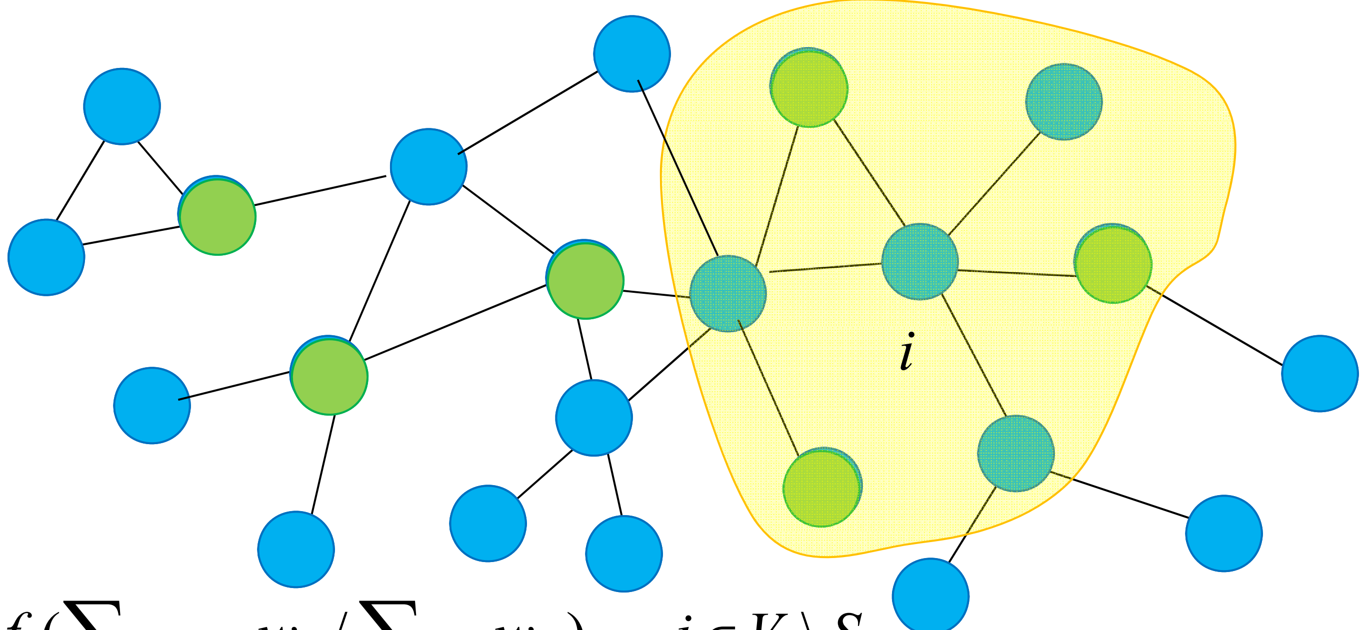
Models

- The standard SIR and SIS models assume equal transmission probability over all links of the contact network. However, this is not a realistic assumption in real scale-free networks.



Disease transmission and Immunity model

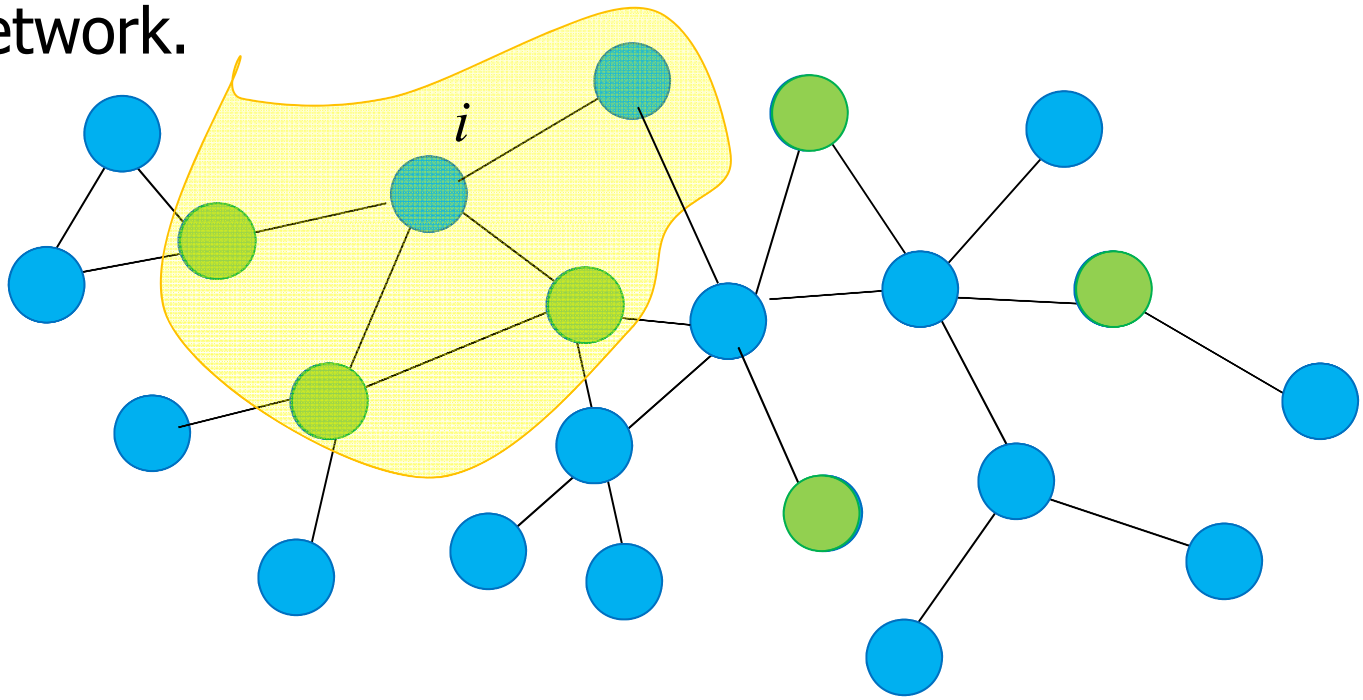
- The immunized individuals do not become infected, and thus, do not transmit the disease to healthy individuals. Therefore, we proposed a general model in which the immunity of an individual against the disease depends on the set of its immunized friends in the network.



$$q_i(S) = f_i\left(\sum_{j\in s\cup\{i\}} w_{ij} / \sum_{k\in V} w_{ik}\right), \quad i \in V \setminus S$$

Influence model

- Motivated by a recent study on viral marketing, we modeled the willingness of each individual to buy the vaccine as a function of the immunized individual in the network.



$$q_i(S) = f_i\left(\sum_{j \in s \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik}\right), \quad i \in V \setminus S$$

Immunization strategies

- Each immunized individual can be regarded as a node which is removed from the network. The goal is to break apart the underlying structure of the network in order to prevent the transmission of the disease from one part to the others.

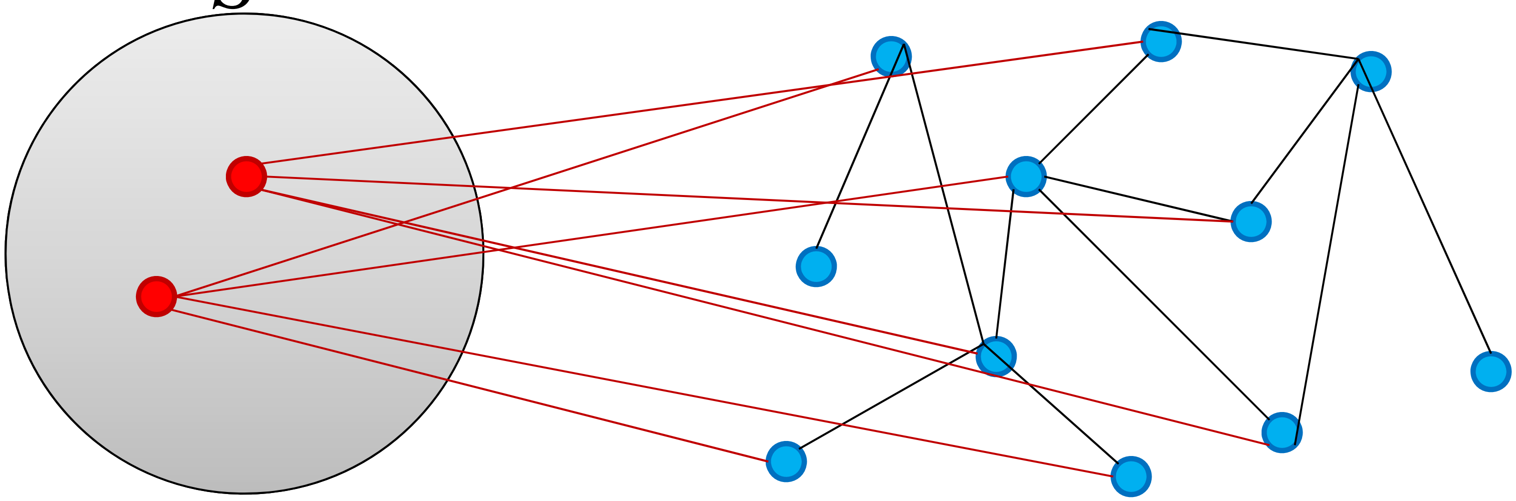
- Immunization Strategies
  - Random Immunization
  - Targeted Immunization
  - Greedy Hill-Climbing

Greedy hill-climbing

- Finding an optimal immunization strategy can be considered as a problem of finding the initial set of nodes *S* to be immunized in order to maximize the global immunity *g*(*S*) in the network.

S

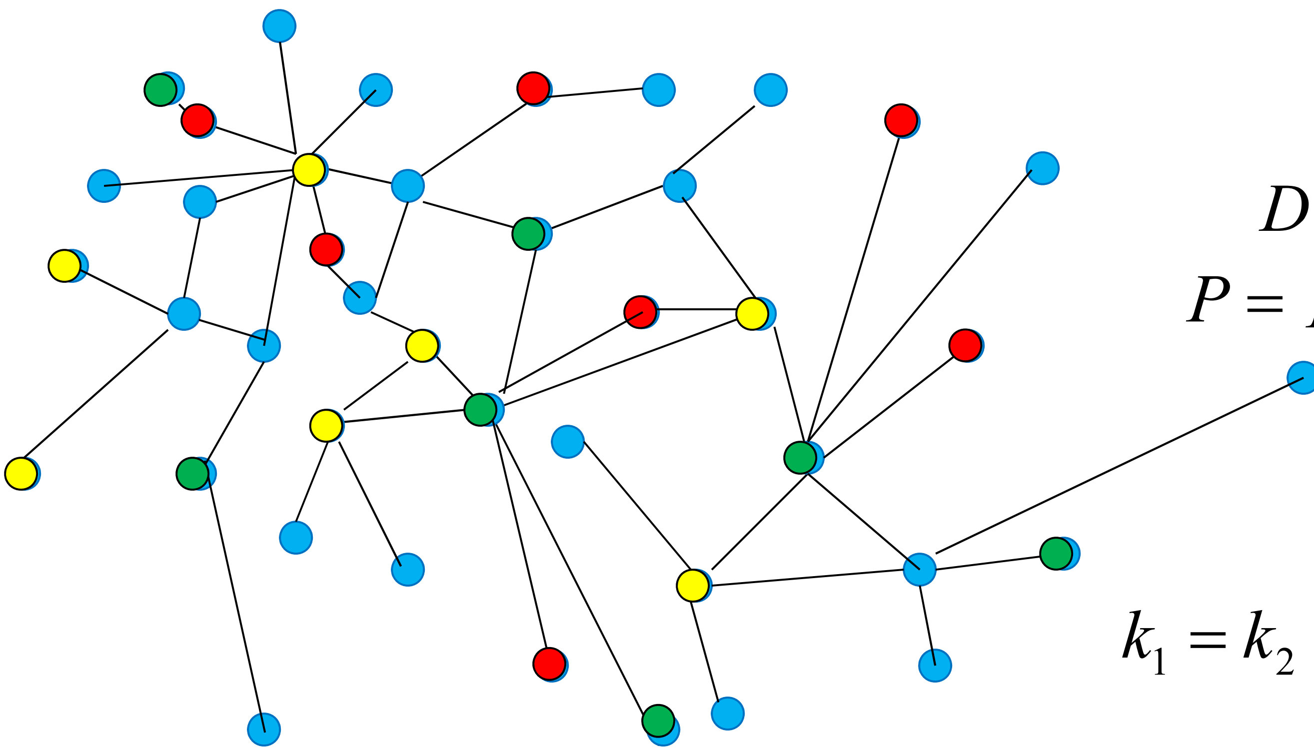
$$i = \arg\max_i g(S \cup \{i\}) - g(S)$$



$$g(S) = f(\{i\})$$

Finding the optimal marketing strategy

- Pricing based on average degree



$$D = k_1 \;\; k_2 \;\; k_3 \;\; \dots \; k_{n-1}$$

$$P = P_1 \;\; P_2 \;\; P_3 \;\; \dots P_{n-1} \;\; P_n$$

$$k_1 = k_2 = \dots \; k_{n-1} = \lfloor N / \mu \rfloor$$

$$P_1 = 0, P_2 = f(1/\mu), \dots \; P_{n-1} = f(\frac{\mu - 1}{\mu}), P_n = f(\frac{\mu}{\mu})$$

Source: Mirzasoleiman, Babaei, and Jalili, 2011

Finding the optimal marketing strategy

- Greedy pricing approach
  - Sort the nodes of a network in decreasing order of their degrees in array *D*.
  - Initialize $k_0 = 0$, $i = 1$.
  - For $j = 1 \ldots n$ repeats the following three steps.
  - If $j > \sum_{r=1}^{i} \left\lfloor \frac{\mu}{2^r} \right\rfloor$ then $i=i+1$.
  - $Z = \left\lfloor \frac{\mu}{2^i} \right\rfloor$.
  - Find $k_m = \arg\max_k \left(\sum_{S=D(1:k),\, i\in V/S \,\mid\, d_i \ge Z} X(i)\right)$ where for all $i \in V \setminus S$

$$X(i) = \begin{cases} 1 & \frac{j}{r} \le \sum_{j\in S\cup\{i\}} w_{ij} \Big/ \sum_{k\in V} w_{ik} \\ 0 & otherwise \end{cases}.$$

  - $k_j = k_m - k_{m-1}$.

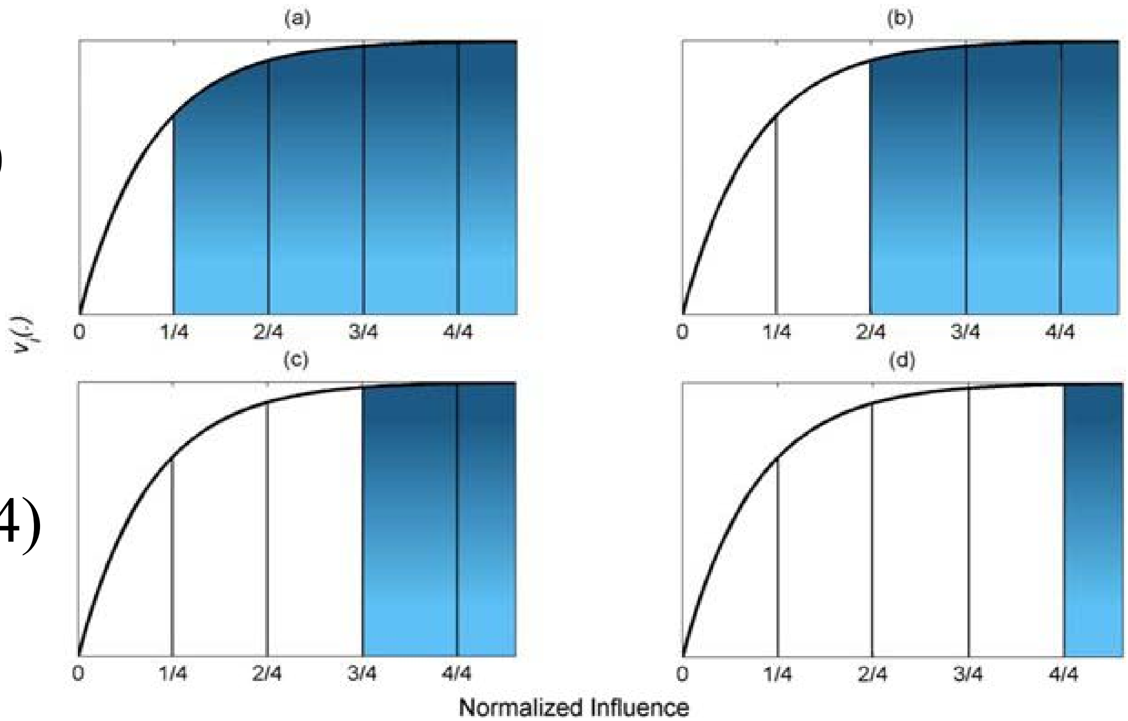
Source: Mirzasoleiman, Babaei, and Jalili, 2011

Finding the optimal marketing strategy

- Offer the vaccine with the price of to the nodes chosen by the targeted strategy until k_j nodes accept the offer and buy the vaccine.

$P = 0$

$P = f(1/4)$



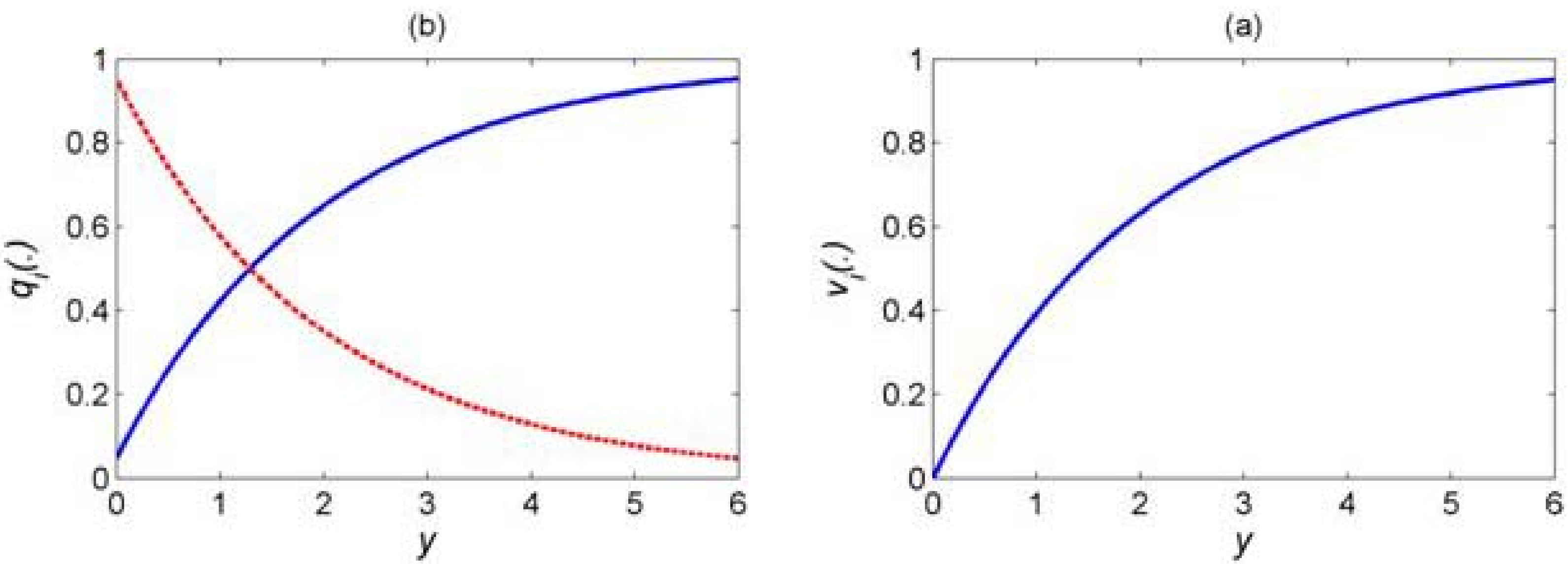
$P = f(2/4)$

$P = f(3/4)$

$v(i)$ as a function of normalized influence for a network with $4 \leq \mu \leq 5$

Experiments

- Immunity and Influence models

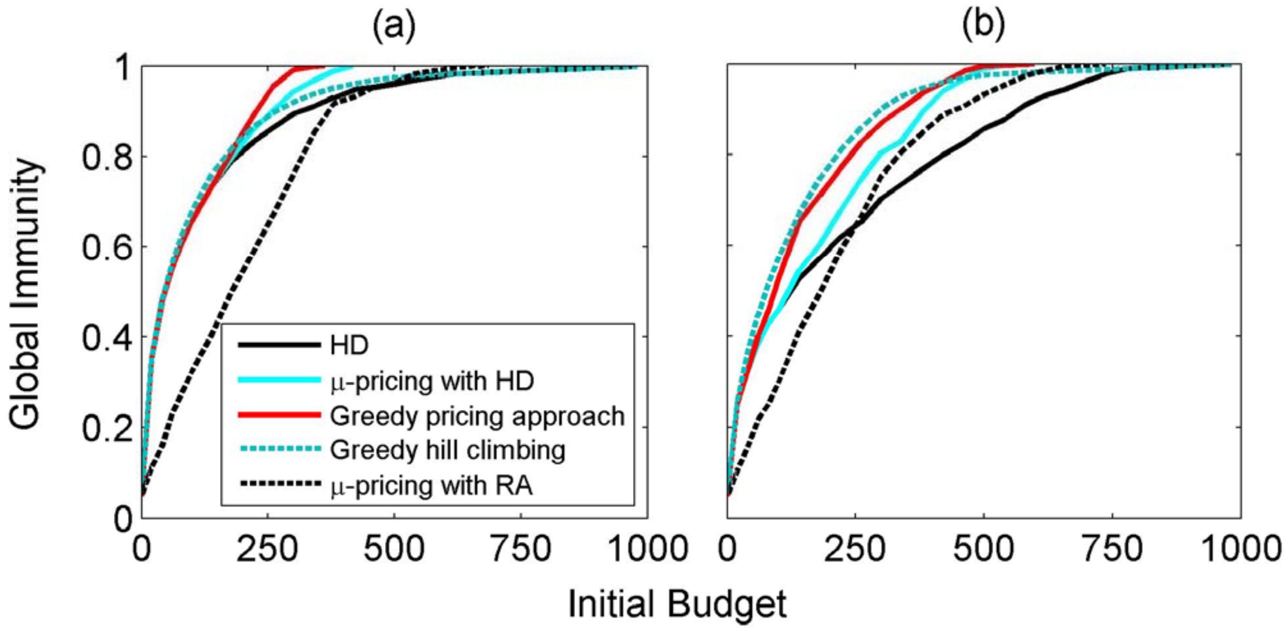


$v(i)$ and $q(i)$ as a function of y=6*normalized influence on each individual. The red dashed line in b) shows the probability of being exposed to disease.

Source: Mirzasoleiman, Babaei, and Jalili, 2011

Results

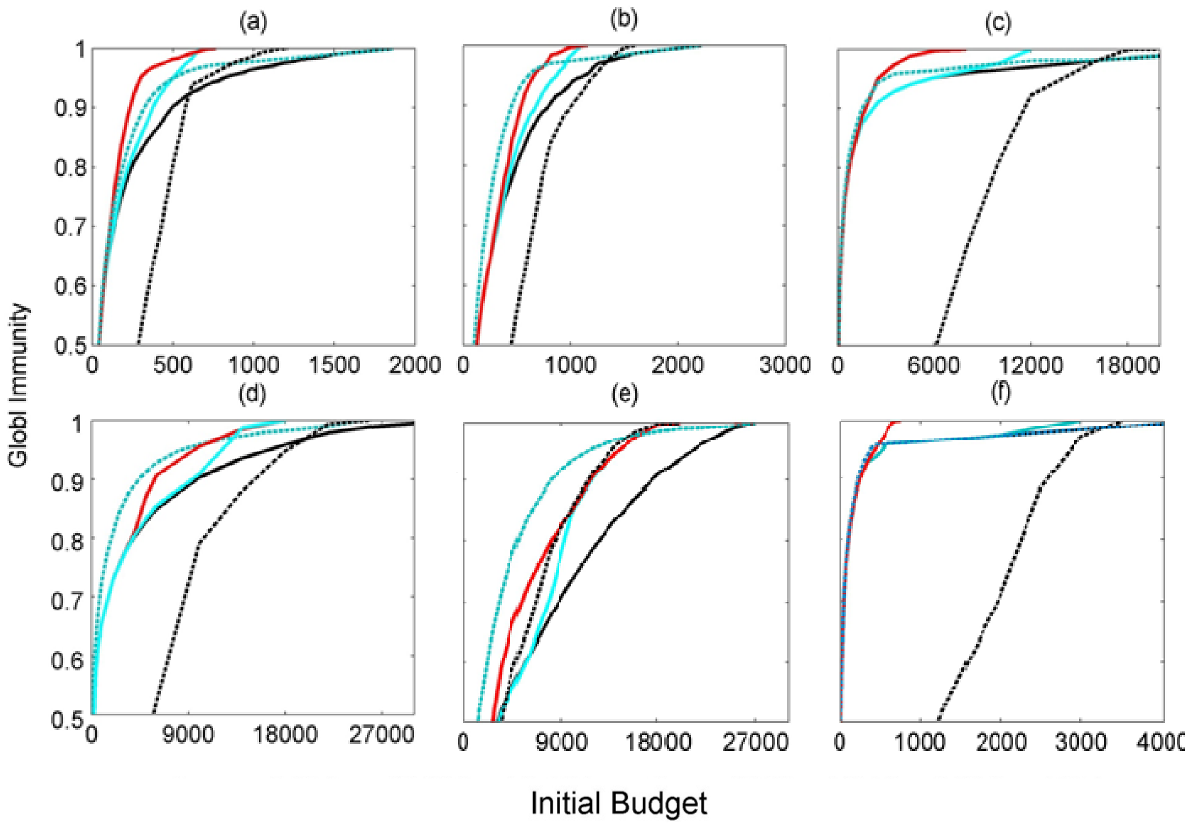
- The greedy hill-climbing algorithm provides a (1–1/*e*)-approximation guarantee for the non-negative, monotone and submodular functions.



Source: Mirzasoleiman, Babaei, and Jalili, 2011

Global immunity of the network (the fraction of immunized nodes) as a function of initial budget for immunization, for a) the Barabasi-Albert and b) the forest-fire network with 1000 nodes

Results



Global immunity of the network as a function of initial budget for immunization, for a) Facebook-like social network, b) yeast protein interaction network, c) High-energy physics theory citation networks, d) CAIDA AS relationship network, and e) Enron email network.

Source: Mirzasoleiman, Babaei, and Jalili, 2011

Readings

- “Networks, Crowds, and Markets” by Easley and Kleinberg (Chapters 16, 19, and 21)
- and many nice papers searchable in the Internet