



Social and Technological Network Analysis

Lecture 6: Network Robustness and Applications

Dr. Cecilia Mascolo

In This Lecture

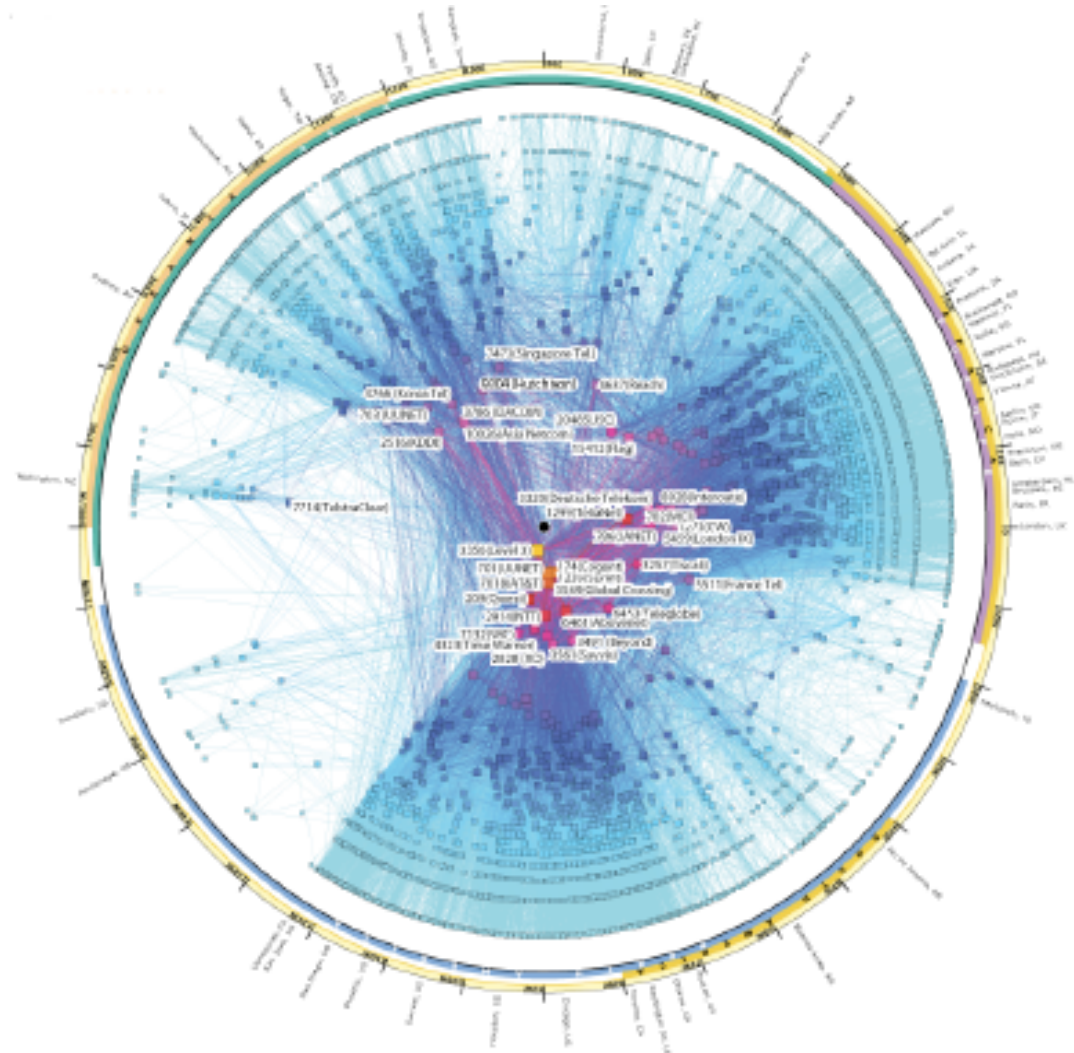


- We revisit power-law networks and define the concept of robustness
- We show the effect of random and targeted attacks on power law networks versus random networks
- We discuss applications to various networks

Internet AS topology



- Autonomous System (AS): a collection of networks under the same administration
- 2009: 25,000 ASs in the Internet

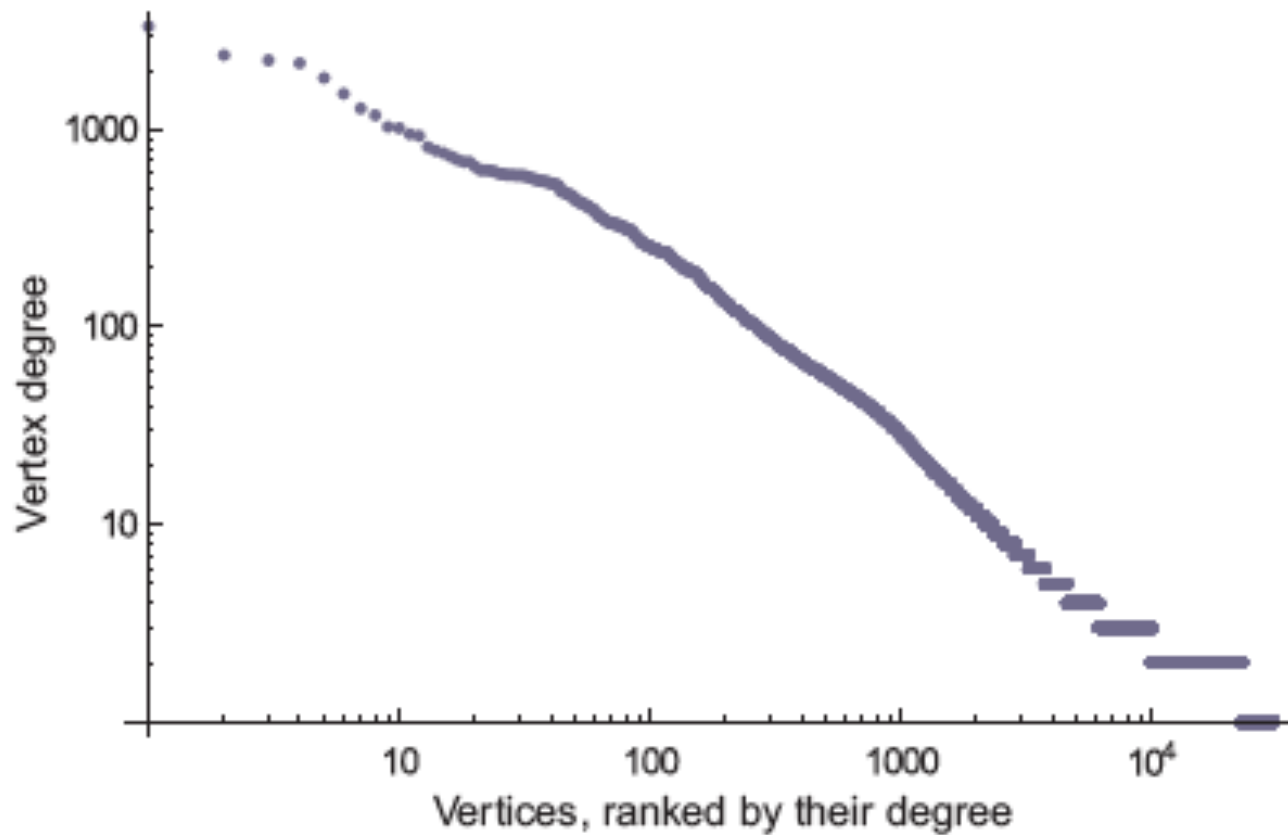


Topology Information



- By reading the routing tables of some gateways connected ASs, Internet topology information could be gathered
- October 08:
 - Over 30,000 ASs (including repeated entries)
 - Over 100,000 edges

Degree distribution of ASs: A scale free network!



Properties



- The top AS is connected to almost 10% of all ASs
- This connectedness drops rapidly
- Very high clustering coefficient for top 1000 hubs: an almost complete graph
- Most paths no longer than 3-4 hops
- Most ASs separated by shortest paths of maximum length of 6

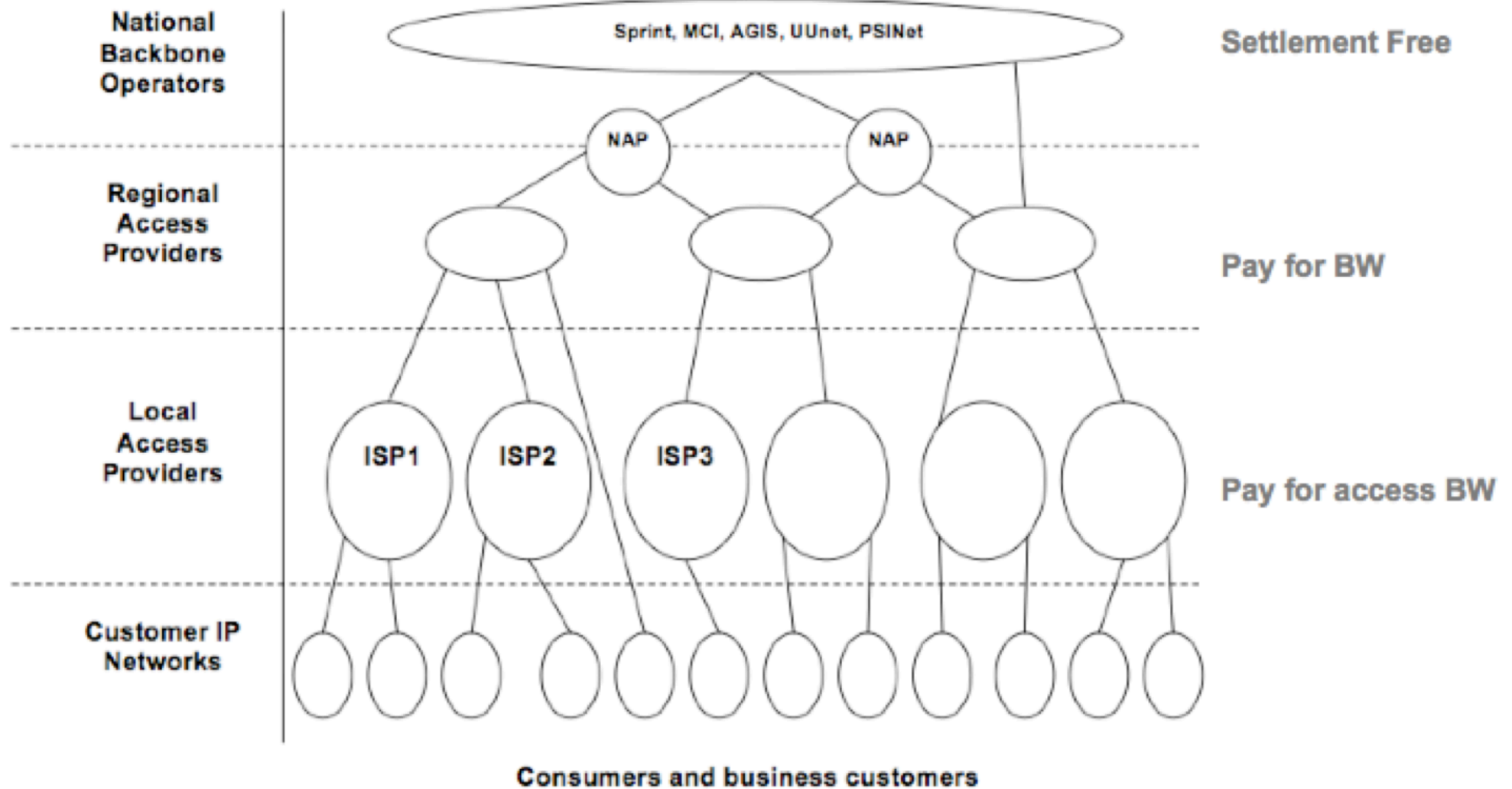
Rank:	1	2	3	4	5	6	7	8	9	10
Degree:	3309	2371	2232	2162	1816	1512	1273	1180	1029	1012

The Internet Now [Sigcomm10]

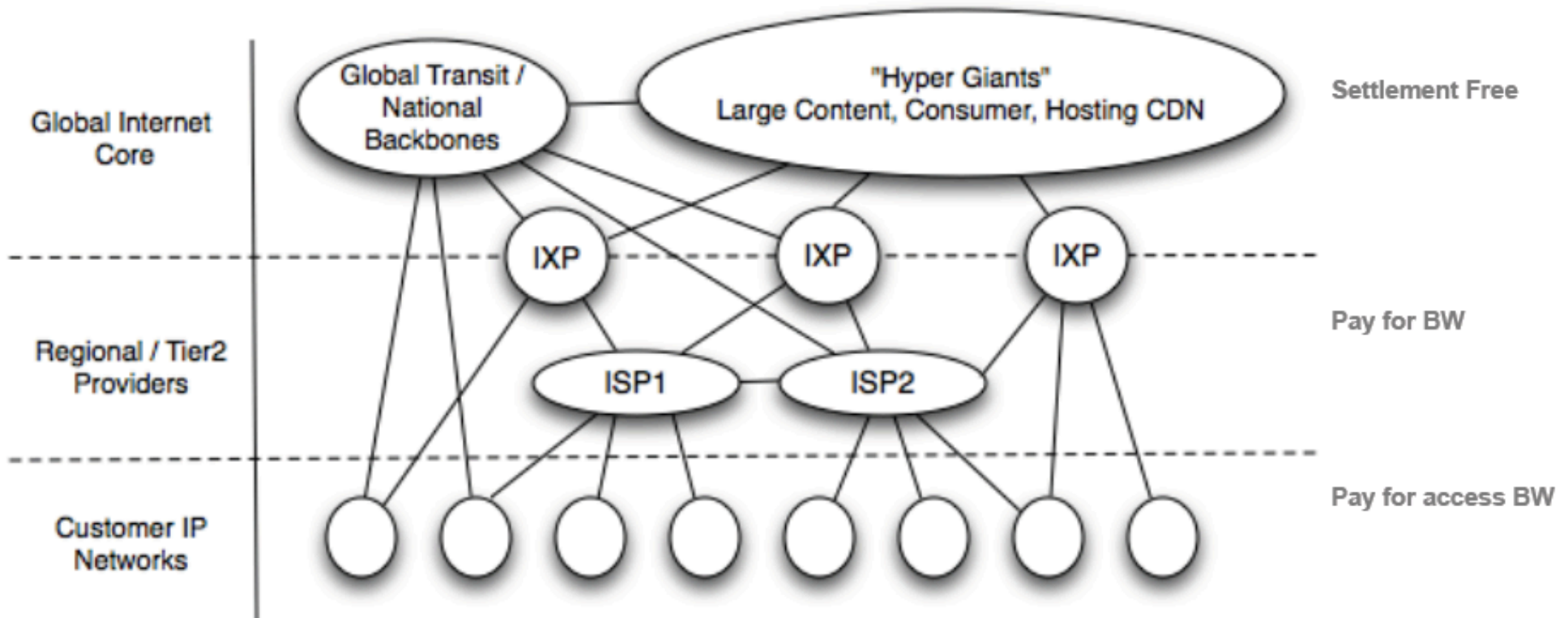
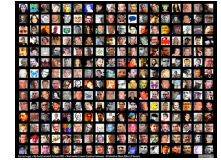


- They monitored inter-domain traffic for **2 years**
 - 3095 Routers
 - 110 ISPs
 - 18 Global
 - 38 Regional
 - 42 Consumer
 - 12 Terabits per second
 - 200 Exabytes total (200,000,000,000,000,000,000)
 - ~25% all inter-domain traffic.
- Inspect packets and classify them.

Internet 2007

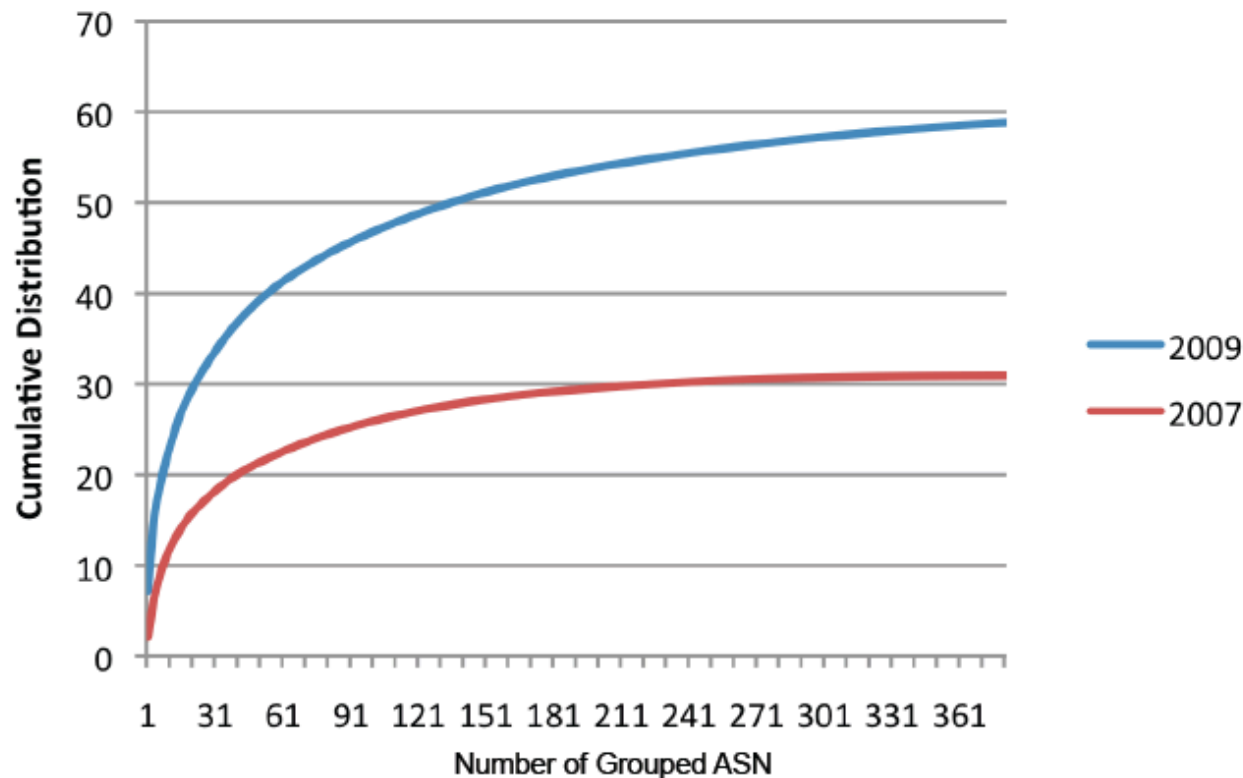


Internet 2009



- Flatter and much more densely interconnected Internet
- Disintermediation between content and "eyeball" networks
- New commercial models between content, consumer and transit

Internet traffic: responsibility to few



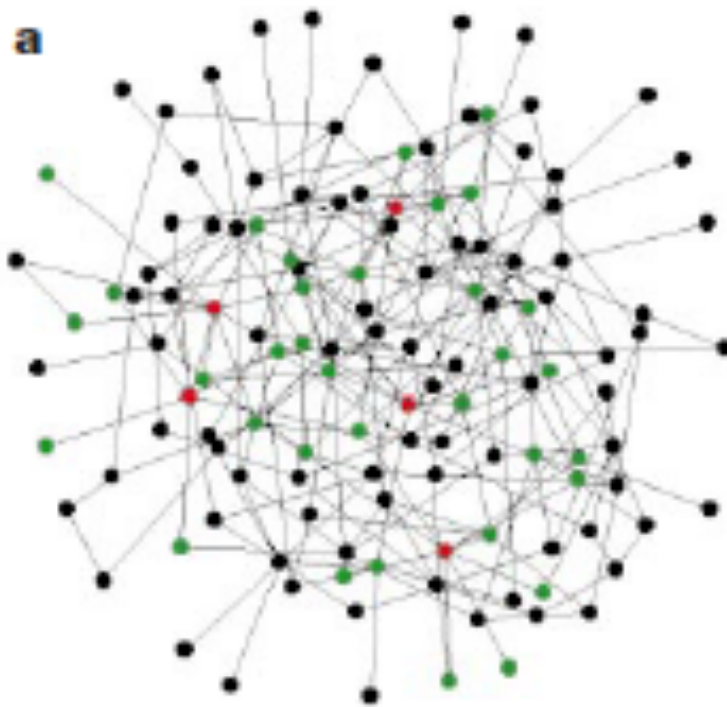
- In 2007, thousands of ASNs contributed 50% of content
- In 2009, 150 ASNs contribute 50% of all Internet traffic

Robustness

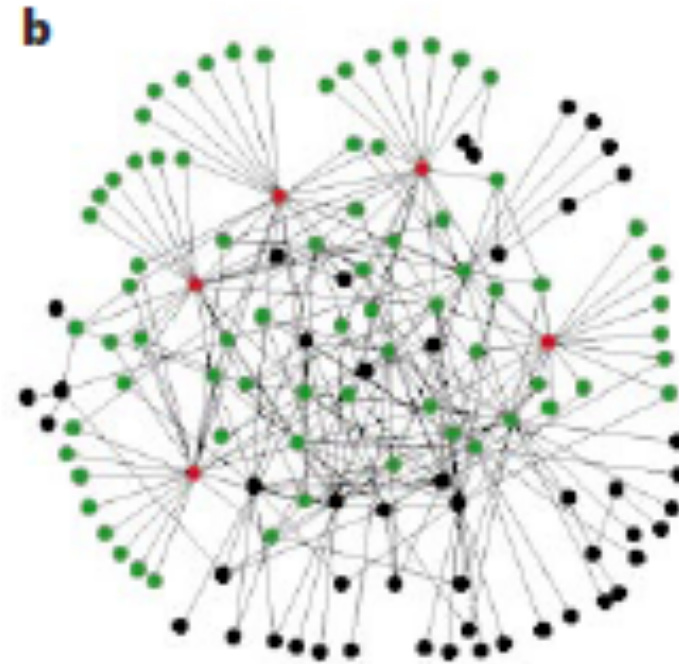


- If a fraction of nodes or edges are removed:
 - How large are connected components?
 - What is the average distance between nodes in the components?
- This is related to *Percolation*
 - each edge/node is removed with probability $(1-p)$
 - Corresponds to random failure
 - Targeted attacks: remove nodes with high degree, or edges with high betweenness.
- The formation or dissolution of a giant component defines the percolation threshold

How Robust are These?

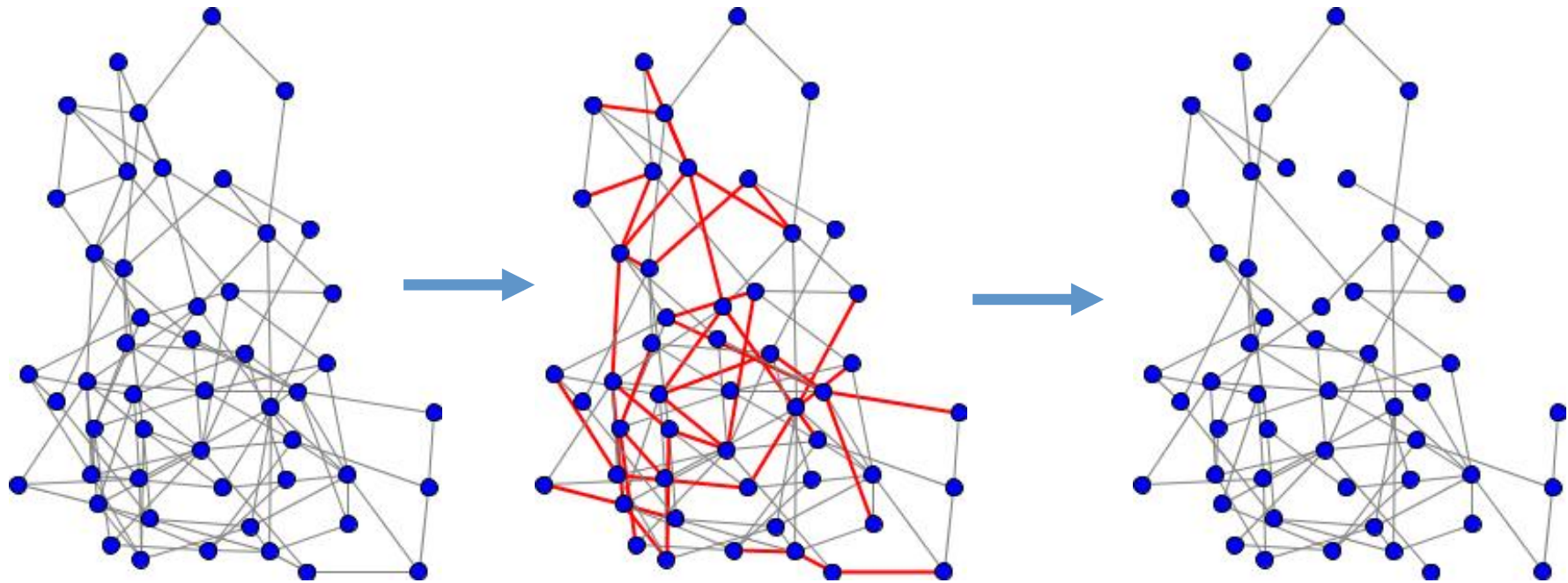


Exponential



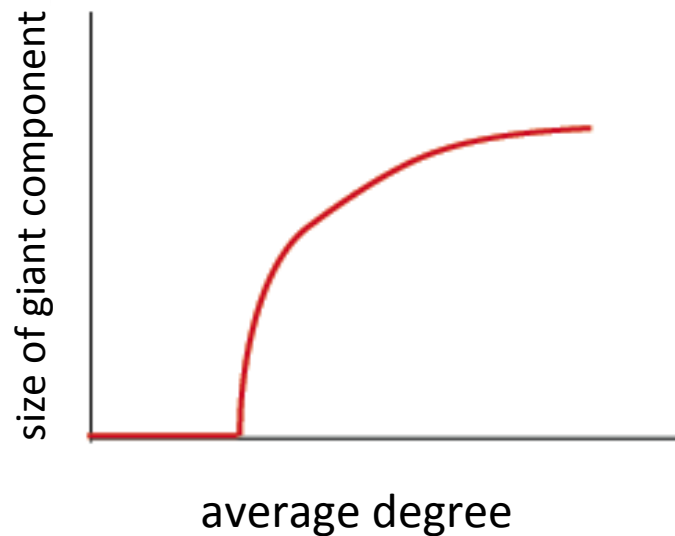
Scale-free

Edge (or Bond) Percolation



- 50 nodes, 116 edges, average degree 4.64
- after 25% edge removal
- 76 edges, average degree 3.04 – still well above percolation threshold

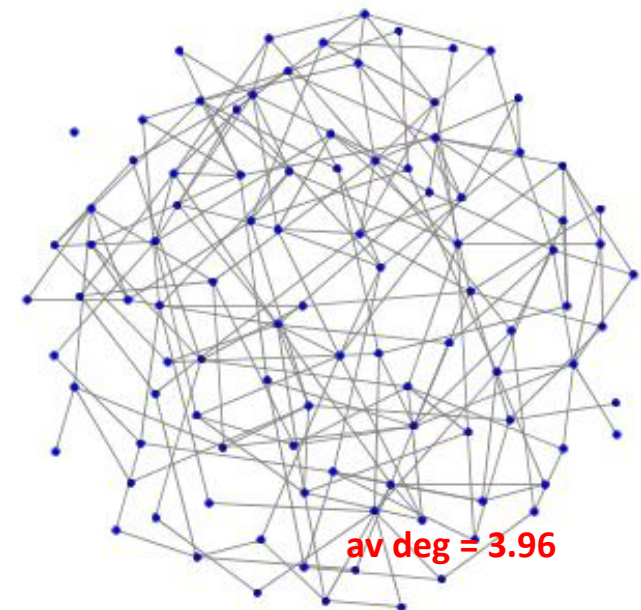
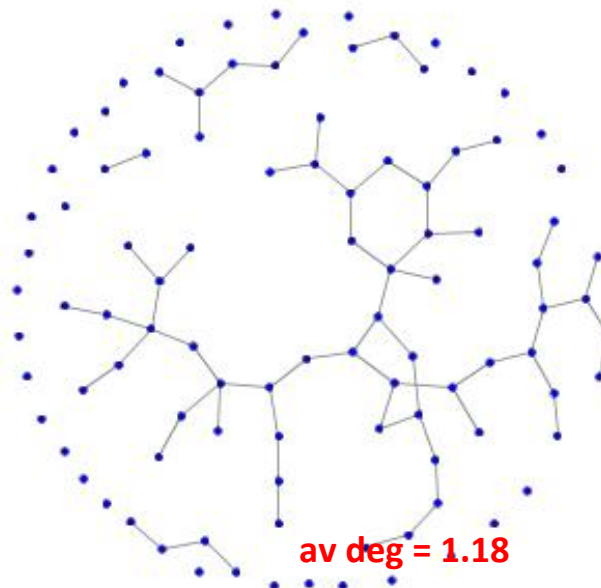
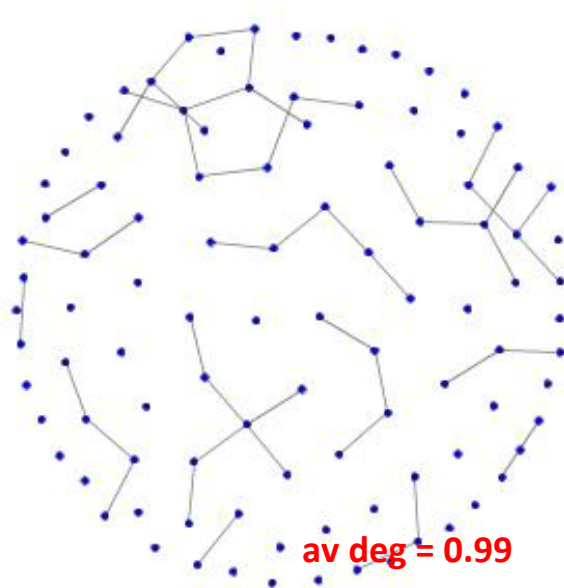
Percolation threshold in Random Graphs



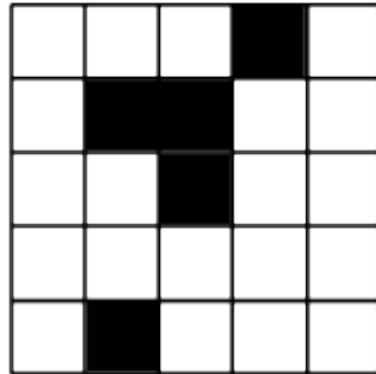
Percolation threshold: how many edges have to be removed before the giant component disappears?

As the average degree increases to 1, a giant component suddenly appears

Edge removal is the opposite process – at some point the average degree drops below 1 and the network becomes disconnected

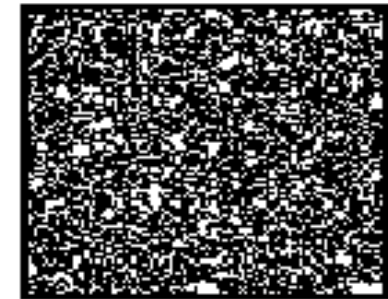


Site Percolation



site percolation

Ordinary Site Percolation on Lattices:
Fill in each site (site percolation) with probability p



- **low p :** small islands of connected components.
- **p critical:** giant component forms, occupying finite fraction of infinite lattice. Other component sizes are power-law distributed
- **p above critical value:** giant component occupies an increasingly large fraction of the system.

Barabasi-Yeong-Albert's study (2000)

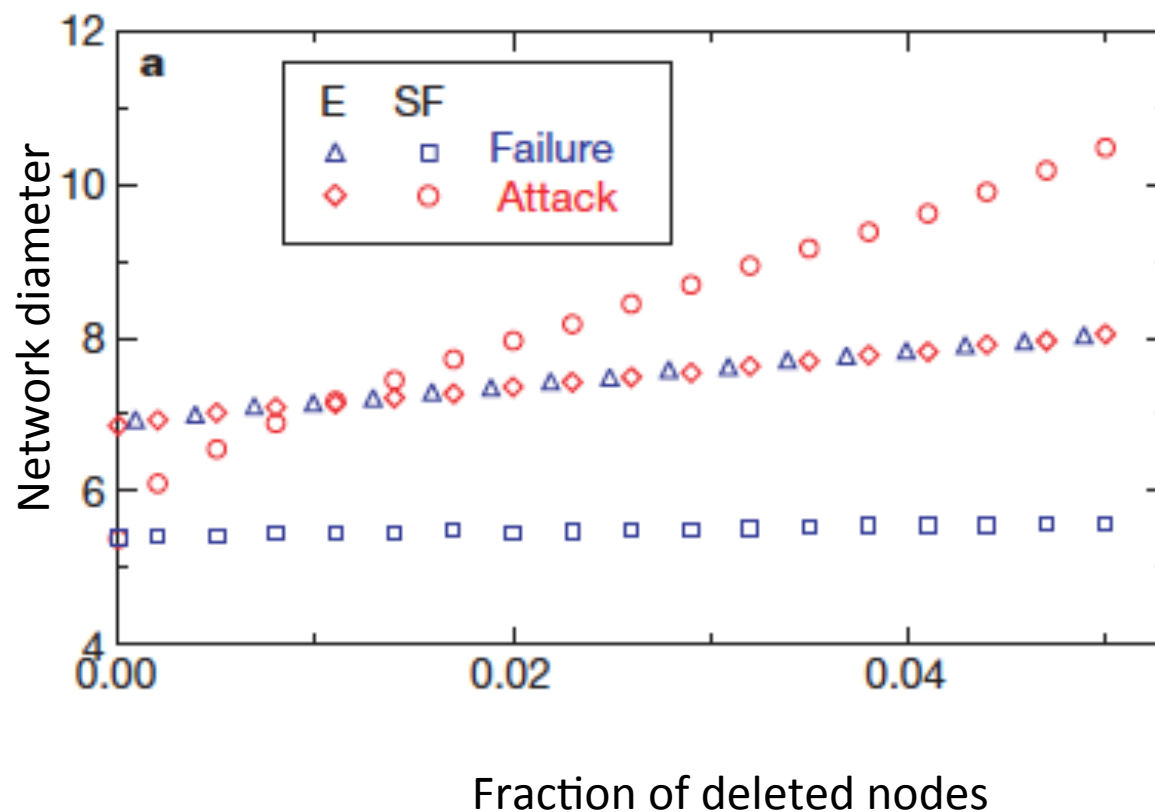


- Given 2 networks (one exponential one scale free) with same number of nodes and links
- Remove a small number of nodes and study changes in average shortest path to see if information communication has been disrupted and how much.

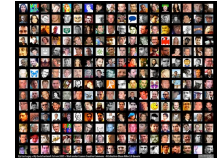
Let's look at the blue lines



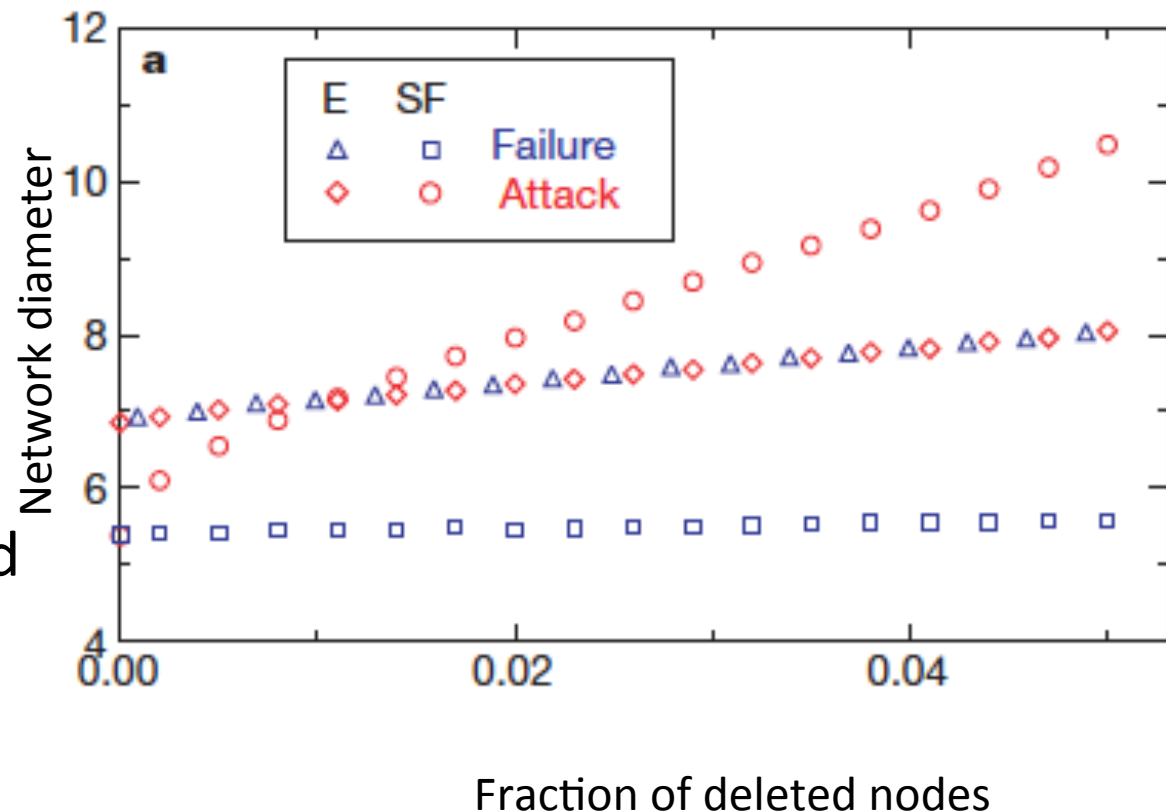
- Random graph: increasing monotonically
- SF: remains unchanged until at least 5%



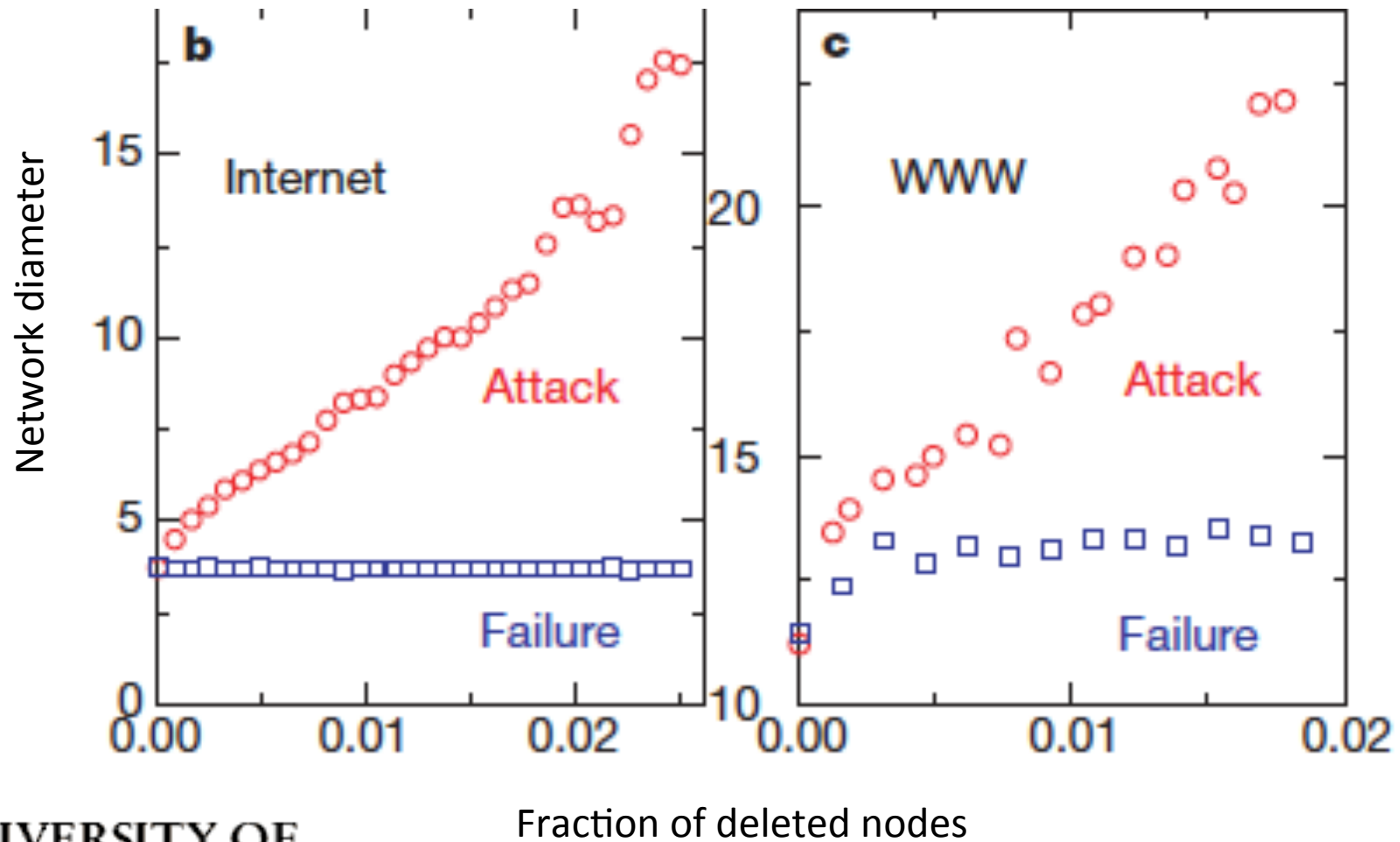
Let's look at the red lines



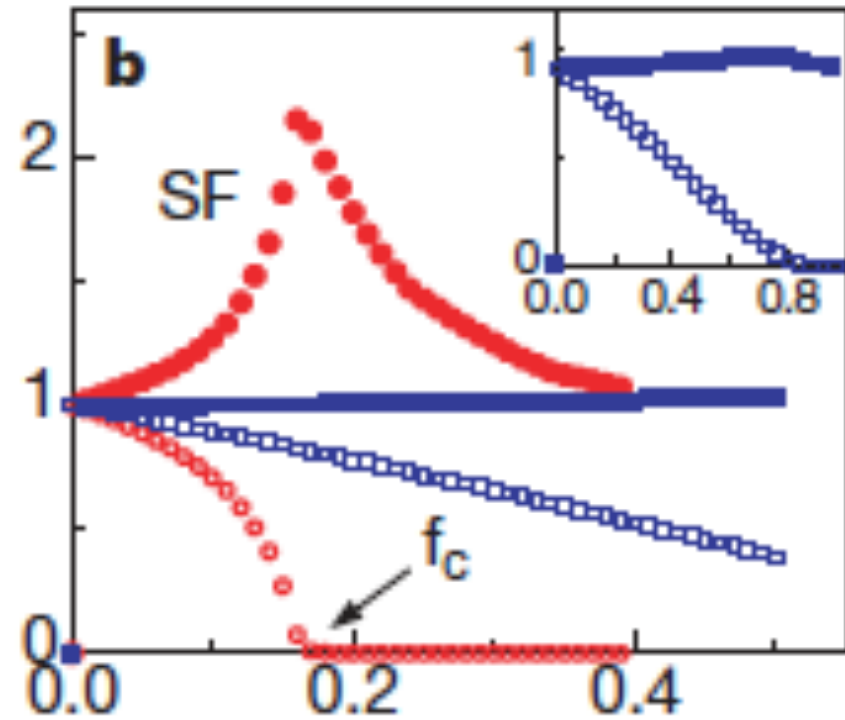
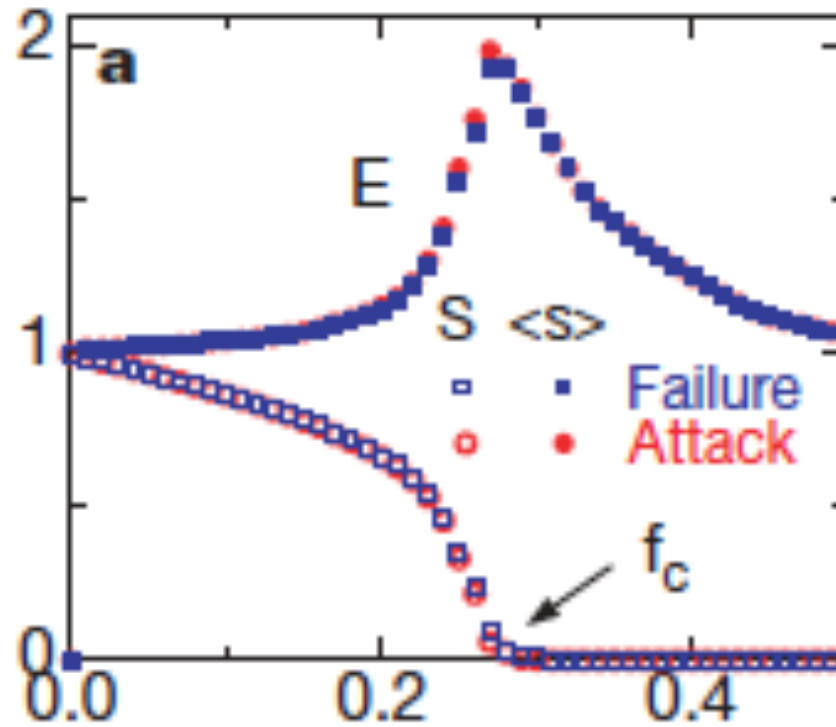
- Random graph: same behaviour if nodes with most links are chosen first
- SF: with 5% nodes removed the diameter is doubled



Effect of attacks and failure on WWW and Internet

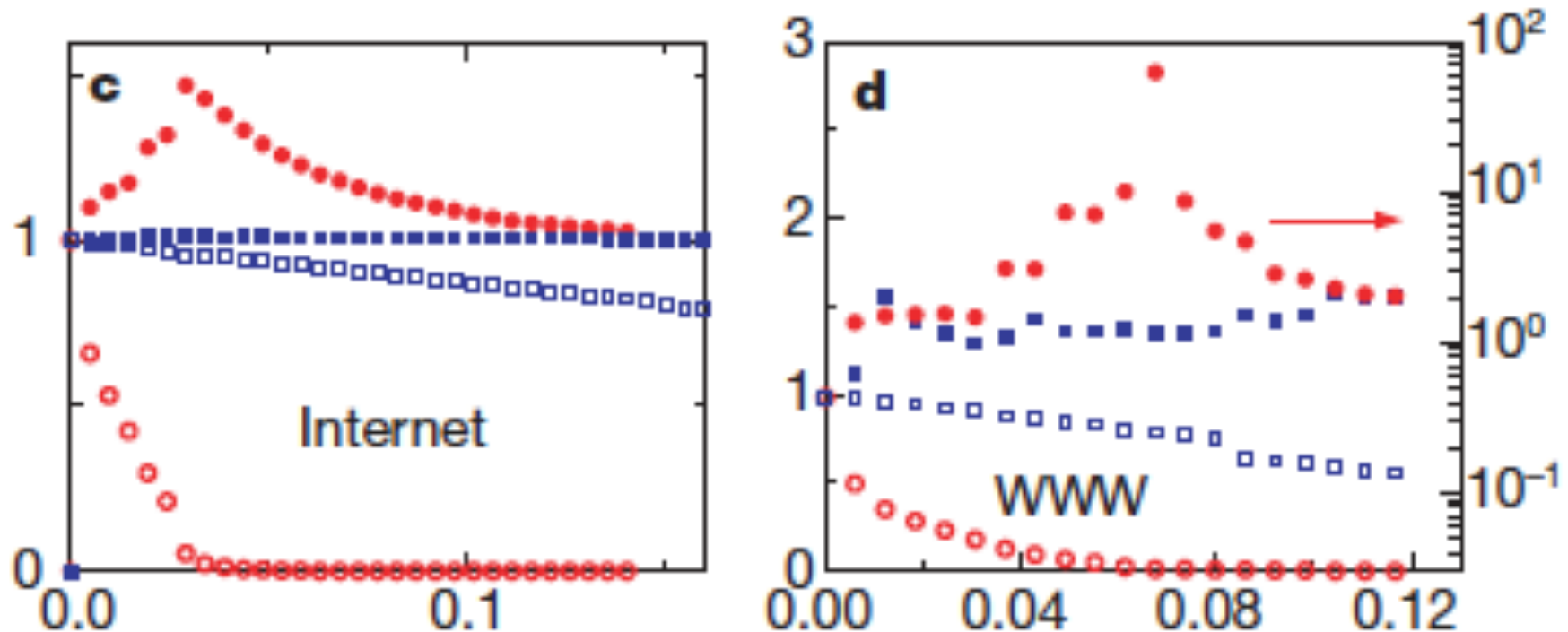


Effect on Giant Component



Fraction of deleted nodes

Internet and WWW: Effect on Giant Component

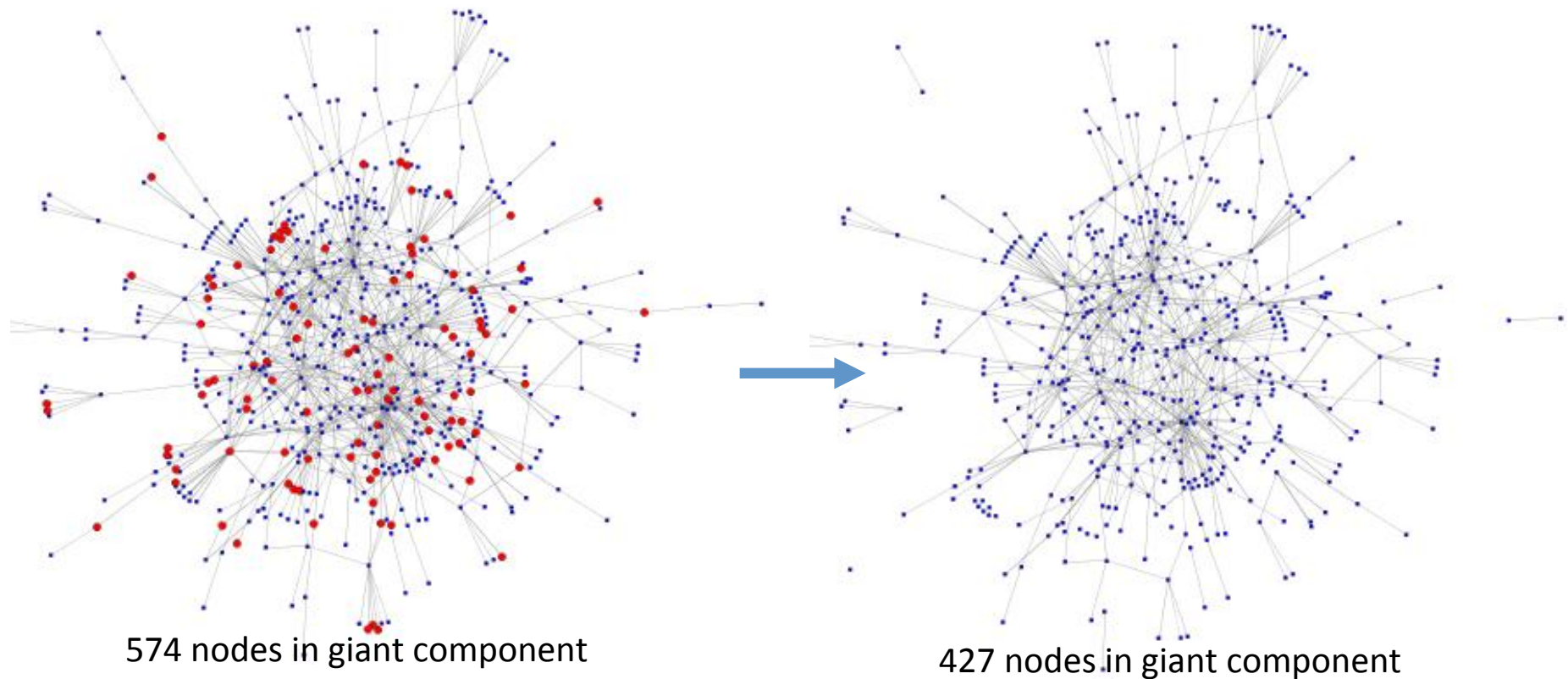


Fraction of deleted nodes

Scale-free networks are resilient with respect to random attack



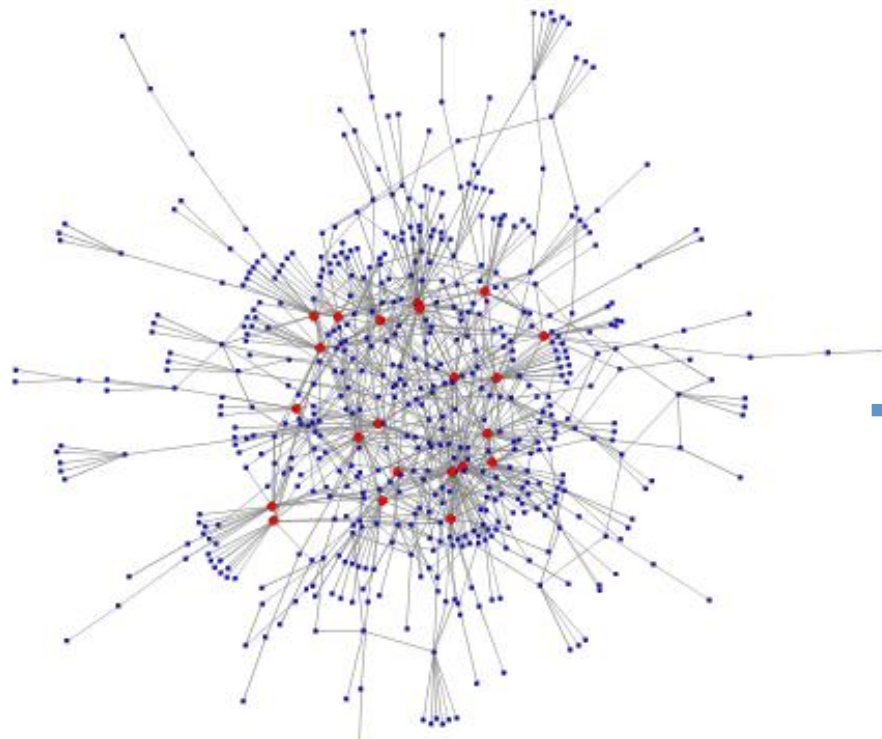
- Example: Gnutella network, 20% of nodes



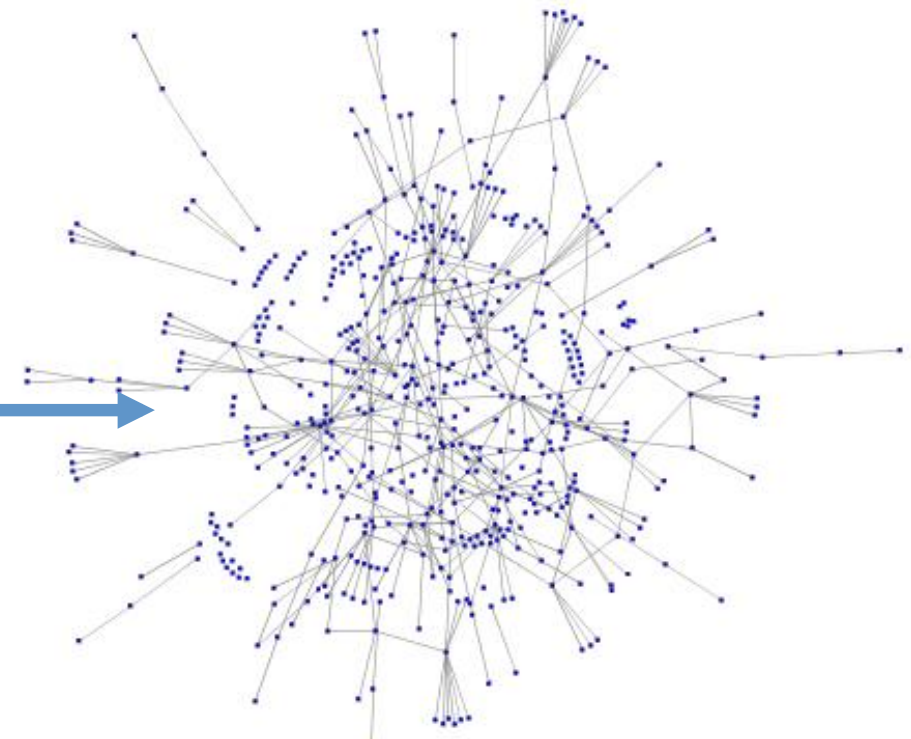
Targeted attacks are effective against scale-free networks



- Example: same Gnutella network, 22 most connected nodes removed (2.8% of the nodes)



574 nodes in giant component

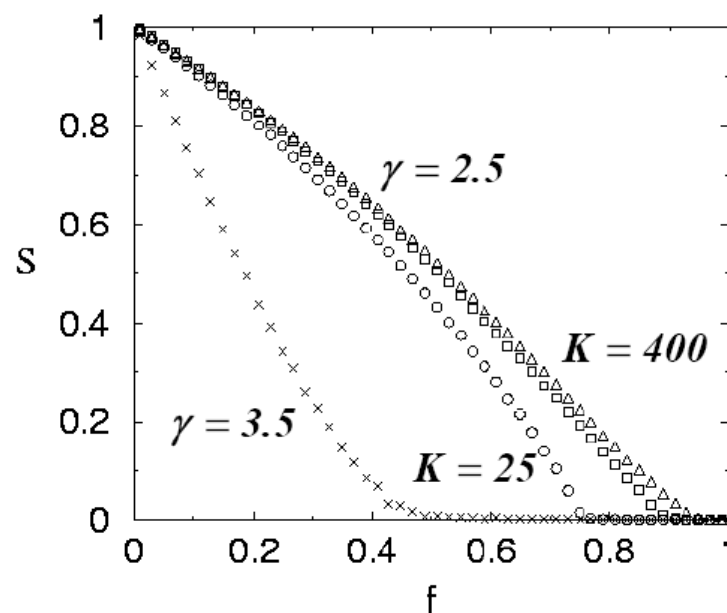


301 nodes in giant component

Another study of power-laws



- Graph shows fraction of GC size over fraction of nodes randomly removed
- Robustness of the Internet (γ is the exponent of PL).
 - $\gamma = 2.5$ Virtually no threshold exists which means a GC is always present
 - For $\gamma = 3.5$ there is a threshold around .0.4
- K indicates the connectivity network considered



Skewness of power-law networks and effects and targeted attack

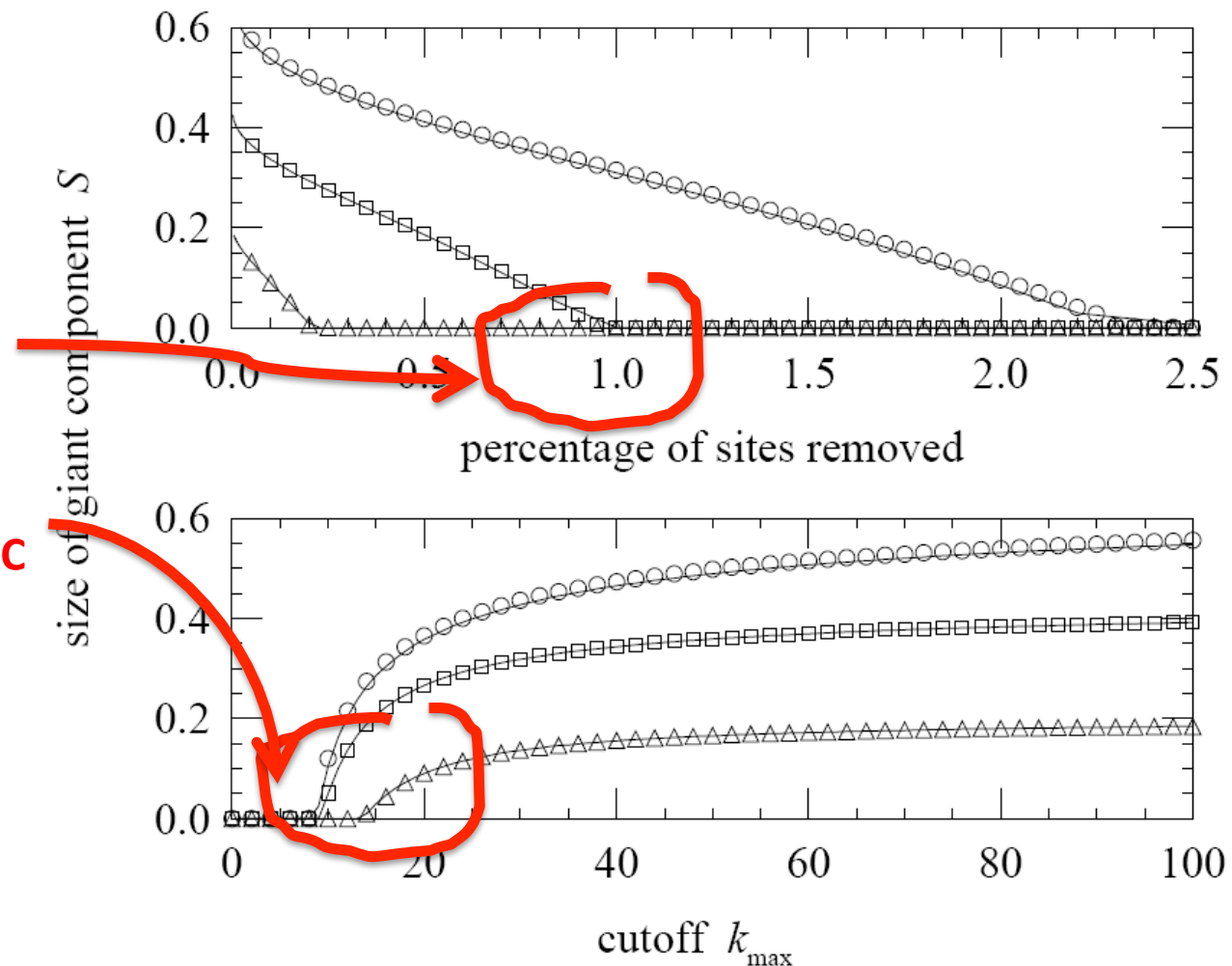


% of nodes removed, from highest to lowest degree

$\gamma = 2.7$ only 1% nodes removed leads to no GC

k_{\max} needs to be very low (10) to destroy the GC

k_{\max} is the highest degree among the remaining nodes



Percolation: let's get formal



- Percolation process:
 - Occupation probability ϕ = number of nodes in the network [ie not removed]
- It can be proven that the critical threshold depends on the degree:

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

- This tells us the minimum fraction of nodes which must exist for a GC to exist.

Threshold for Random Graphs



- For Random networks $\varphi_{\text{critical}} = 1/c$ where c is the mean degree
 - If c is large the network can withstand the loss of many vertices
 - $c=4$ then $\frac{1}{4}$ of vertices are enough to have a GC [3/4 of the vertices need to be destroyed to destroy the GC]

Threshold for Scale Free Networks



- For the Internet and Scale Free networks with $2 < \alpha < 3$
 - Finite mean $\langle k \rangle$ however $\langle k^2 \rangle$ diverges (in theory)
 - Then $\varphi_{\text{critical}}$ **is zero**: no matter how many vertices we remove there will always be a GC
 - In practice $\langle k^2 \rangle$ is never infinite for a finite network, although it can be very large, resulting in very small $\varphi_{\text{critical}}$, so still highly robust networks

Non random removal



- The threshold models we have presented hold for random node removal but not for targeted attacks [ie removal of high degree nodes first]
- The equation for non random removal cannot be solved analytically

Robustness Study and Improvements



- A method to improve network resilience
- Percolation threshold q ignores situation when the network is very damaged but not collapsing.
- Robustness:

$$R = \frac{1}{N} \sum_{Q=1}^N s(Q)$$

$S(Q)$ = nodes in the connected component after removing $Q=qN$ nodes

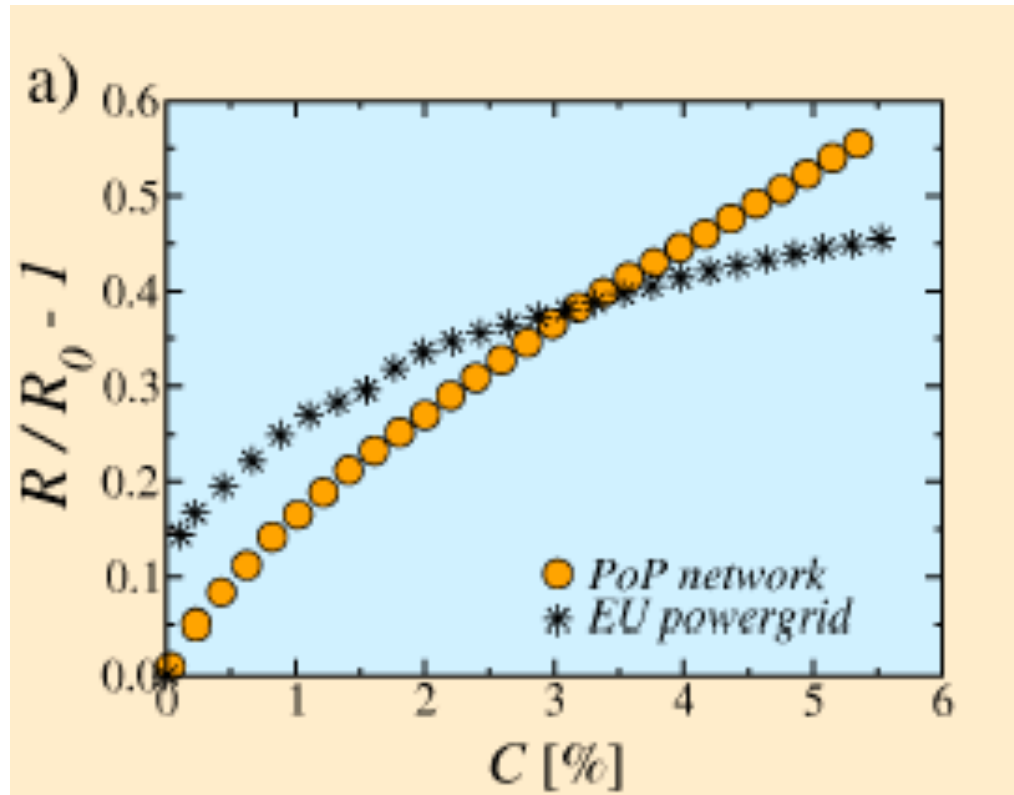
- R ranges from $1/N$ to 0.5 (star and fully connected graph).

Improve Robustness



- Add links until network is fully connected: not practical
- Swap edges of 2 random nodes so that $R' > R$
 - Repeat until no substantial improvement (a value delta)
- Some additional constraints could be (limit the geographical length of new edges for economic reasons)

Robustness Improvement over edge changes

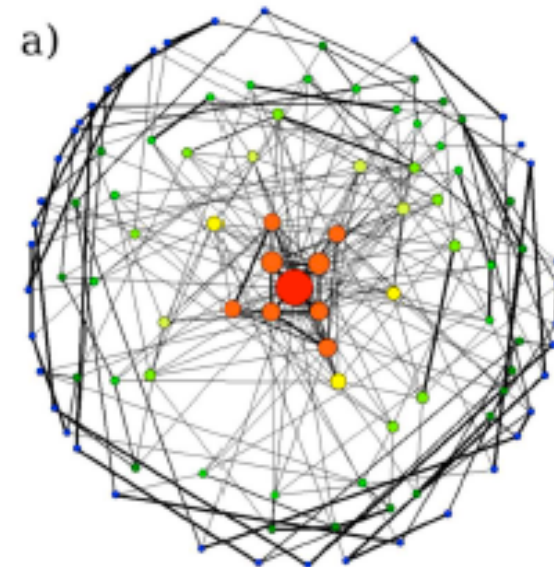


Robustness improved
by 55-45% with ~5%
link change.
Percolation threshold
remains unchanged.

Best Network for Robustness



- How do we design a robust network with a fixed degree distribution?
- Scale free $N=100 = 300$, exponent=2.5
- Onion-like structure!



Robustness of Technological and Social Network



- Targeted attacks on high degree nodes are lethal to a technological and a biological as well as transport network.
- However as seen in Lecture 2, for social systems it is the bridges and weak ties which make a difference...

References



- R. Albert, H. Jeong, A.-L. Barabási. *Error and attack tolerance of complex networks*. Nature 406, 378-482 (2000).
- Cohen et al., *Resilience of the Internet to Random Breakdowns*. Phys. Rev. Lett. 85, 4626 (2000).
- D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Network robustness and fragility: Percolation on random graphs*, Phys. Rev. Lett., 85 (2000), pp. 5468–5471.
- C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, F. Jahanian. *Internet inter-domain traffic*. Proceedings of ACM SIGCOMM 2010 conference. Pages 75-86. ACM.
- C. Schneider, A. Moreira, J. S. Andrade, Jr., S. Havlin, and H. J. Herrmann. Mitigation of malicious attacks on networks. PNAS 2011 108 (10) 3838-3841.