Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Network Theory
Lecture 4.08

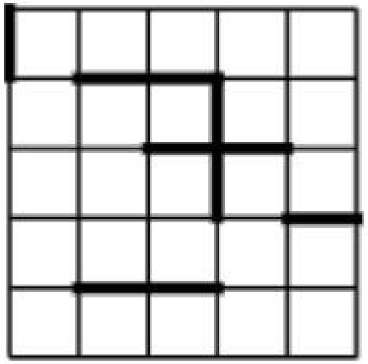
EEU45C09 / EEP55C09
Self Organising Technological Networks

Nicola Marchetti
nicola.marchetti@tcd.ie

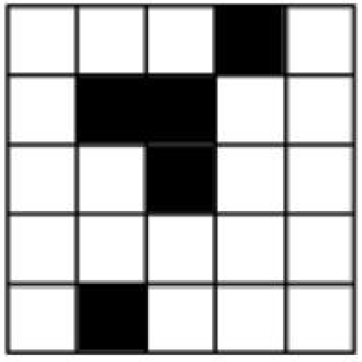
Network robustness and resillience

- Q: If a given fraction of nodes or edges are removed…
  - how large are the connected components?
  - what is the average distance between nodes in the components
- Related to percolation

Movement and filtering of fluids through porous materials. Broader applications have since been developed that cover connectivity of many systems

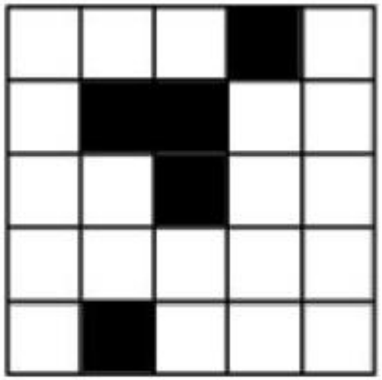


bond percolation



site percolation

Node removal and site percolation

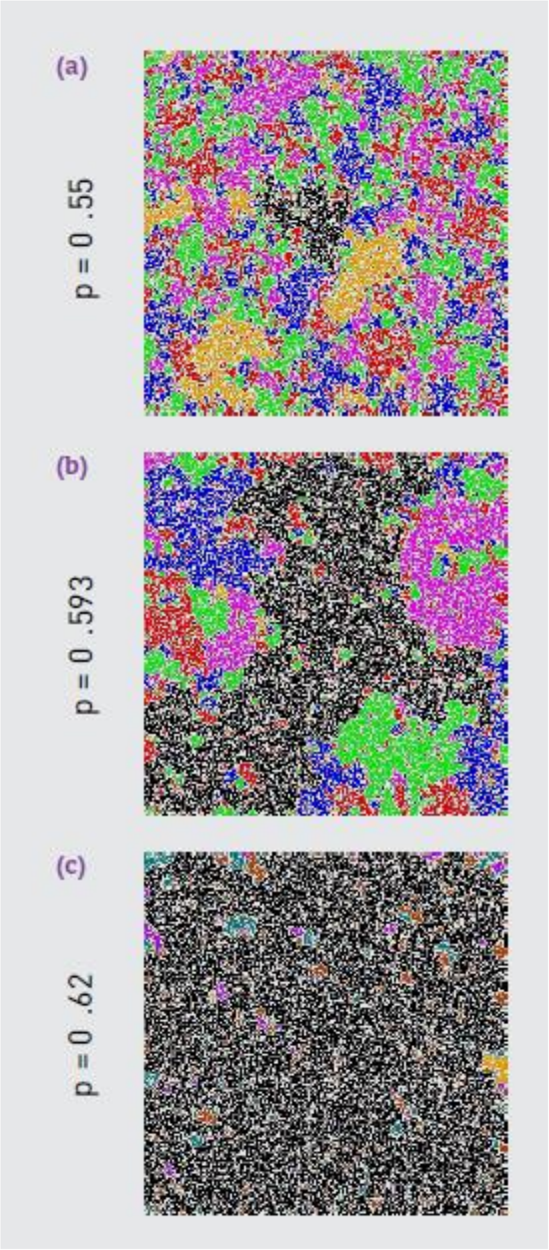


site percolation

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

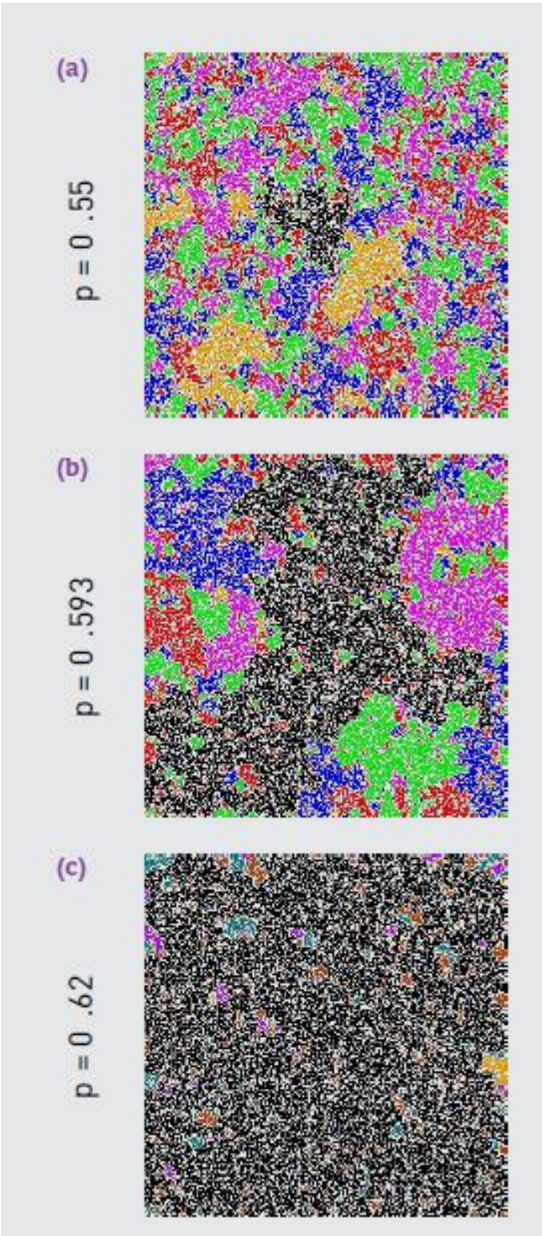
- **low p**: small islands of connected components.
- **p critical**: giant component forms, occupying finite fraction of infinite lattice. Other component sizes are power-law distributed
- **p above critical value**: giant component occupies an increasingly large fraction of the system. Mean size of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

From Forest Fires to Percolation Theory



- We can use the spread of a fire in a forest to illustrate the basic concepts of **percolation theory**
- Let us assume that each pebble in the Figure is a tree and that the lattice describes a forest
- If a tree catches fire, it ignites the neighbouring trees; these, in turn ignite their neighbours
- The fire continues to spread until no burning tree has a non-burning neighbour
- We must therefore ask: If we randomly ignite a tree, what fraction of the forest burns down? And how long it takes the fire to burn out?

From Forest Fires to Percolation Theory

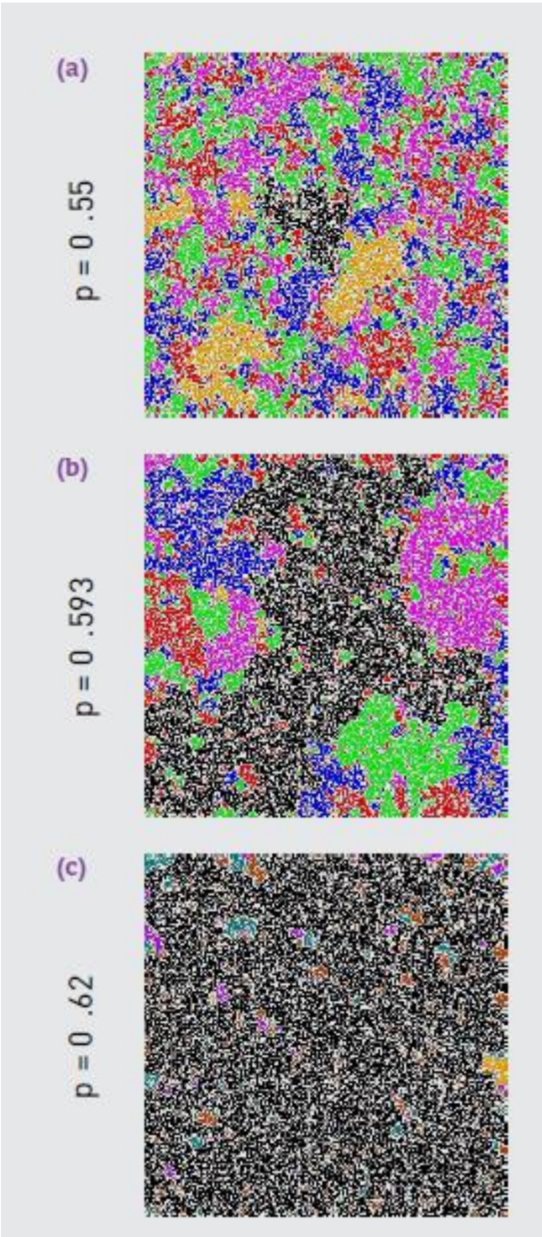


- The answer depends on the *tree density*, controlled by the parameter p

- For small p the forest consists of many small islands of trees (p =0.55, Figure (a)), hence igniting any tree will at most burn down one of these small islands. Consequently, the fire will die out quickly

- For large p most trees belong to a single large cluster, hence the fire rapidly sweeps through the dense forest (p = 0.62, Figure (c))

From Forest Fires to Percolation Theory

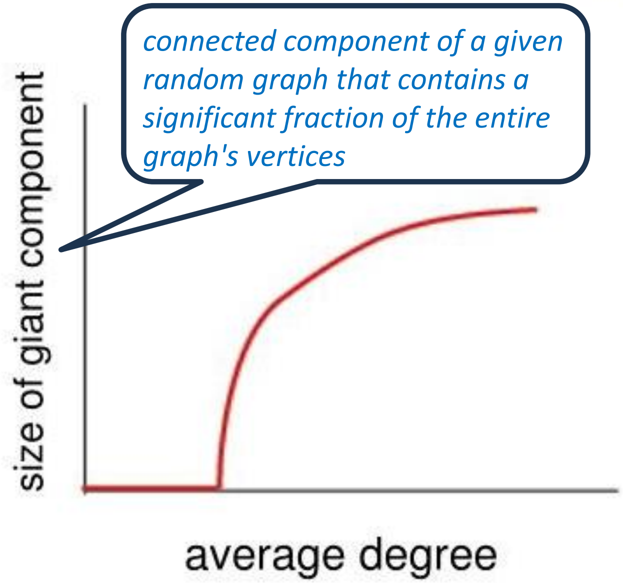


- The simulations indicate that there is a critical p_c at which it takes an extremely long time for the fire to end

- This p_c is the **critical threshold** of the percolation problem

- Indeed, at $p = p_c$ the giant component just emerges through the union of many small clusters (Figure (b))

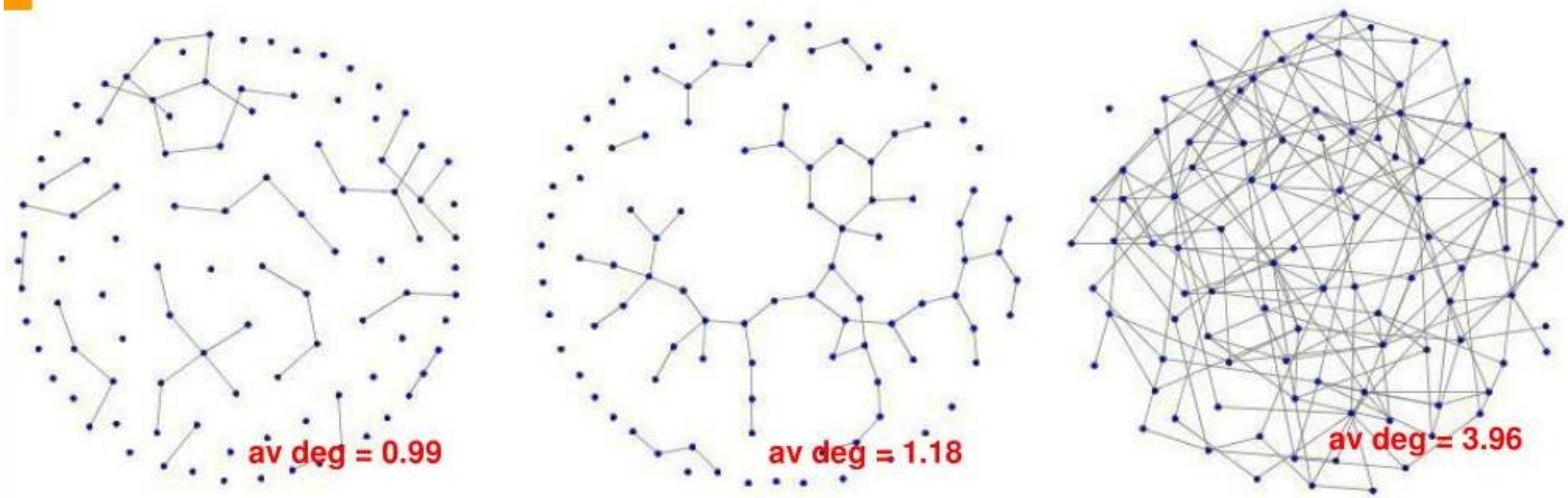
Percolation threshold in Erdos-Renyi Graphs



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

As the average degree increases to z = 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



Random Graphs: Diameter (d)

Reminder from Lecture 4.01

The diameter of a graph is the maximal distance between any pair of its nodes.

The number of nodes at a distance l is not much smaller than $\langle k \rangle^l$. When all nodes are within this distance, we can say that

$$\langle k \rangle^d \sim N,$$

where d is the diameter of the graph.

Random Graphs: Diameter (d)

Reminder from Lecture 4.01

Therefore, we can find an expression for the diameter,

$$d \log \langle k \rangle \sim \log N,$$

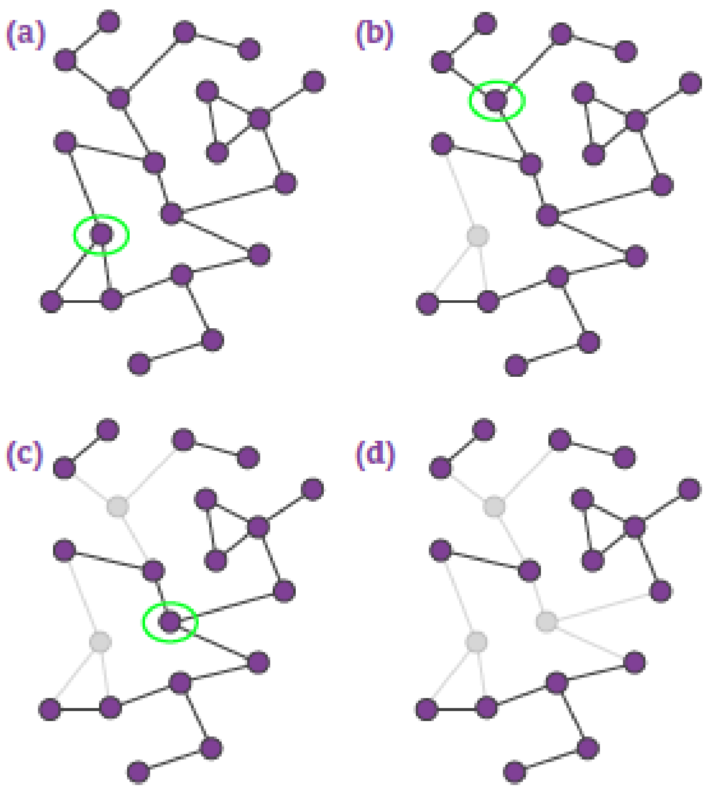
$$d \sim \frac{\log N}{\log \langle k \rangle}.$$

Random Graphs: Diameter (d)

Since this is an estimate of the diameter, we should look at a few important cases:

- If $\langle k \rangle = pN < 1$, a typical random graph is composed of isolated trees and its diameter equals that of a tree.
- If $\langle k \rangle > 1$, a giant cluster appears. The diameter of the graph equals the diameter of the giant cluster if $\langle k \rangle \geq 3.5$, and is proportional to $\log N / \log \langle k \rangle$.
- If $\langle k \rangle \geq \log N$, almost every graph is totally connected. The diameters of the graphs having the same N and $\langle k \rangle$ are concentrated around a few values near $\log N / \log \langle k \rangle$.

Impact of node removal



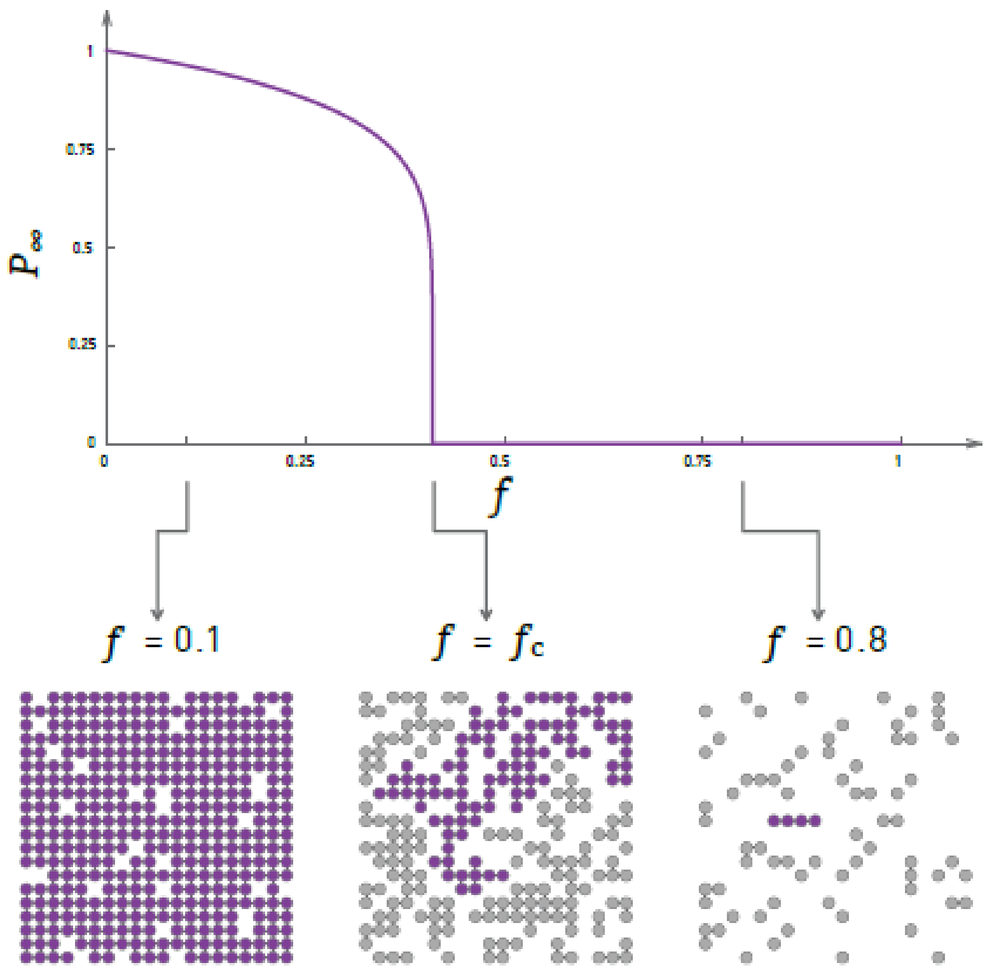
- Gradual fragmentation of a small network following the breakdown of its nodes

- In each panel we remove a different node (highlighted with a green circle), together with its links

- While the removal of the first node has only limited impact on the **network's integrity**, the removal of the second node isolates two small clusters from the rest of the network

- Finally, the removal of the third node fragments the network, breaking it into five non-communicating clusters of sizes s = 2, 2, 2, 5, 6

Network breakdown as inverse percolation



$0 < f < f_c$:
There is a giant component.

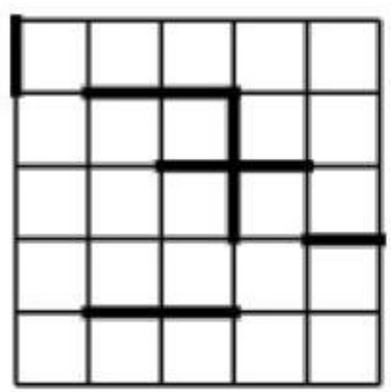
$f = f_c$:
The giant component vanishes.

$f > f_c$:
The lattice breaks into many tiny components.

- We start from a square lattice
- We randomly select and remove a fraction *f* of nodes and measure the size of the largest component formed by the remaining nodes
- This size is accurately captured by P_∞, which is the probability that a randomly selected node belongs to the largest component

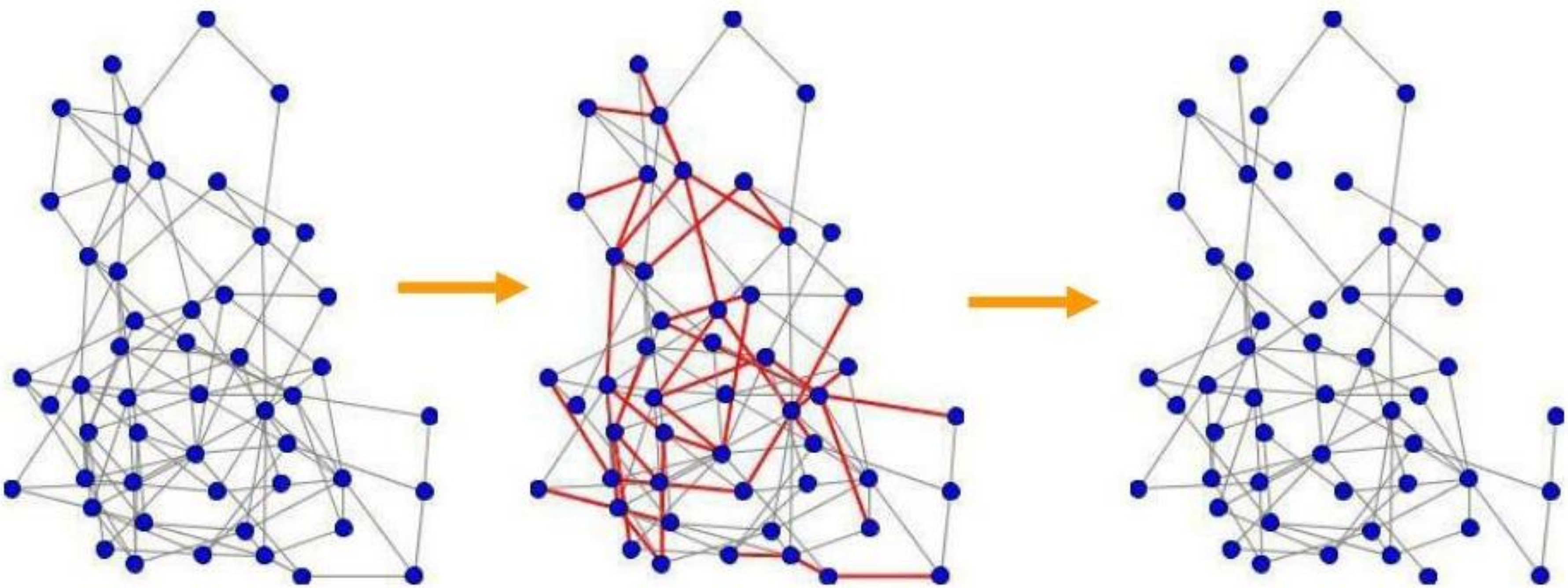
Bond percolation in Networks

- Edge removal
  - bond percolation: each edge is removed with probability (1-p)
    - corresponds to random failure of links
  - targeted attack: causing the most damage to the network with the removal of the fewest edges
    - strategies: remove edges that are most likely to break apart the network or lengthen the average shortest path
    - e.g. usually edges with high betweenness



bond percolation

Edge percolation



50 nodes, 116 edges, average degree 4.64

after 25 % edge removal

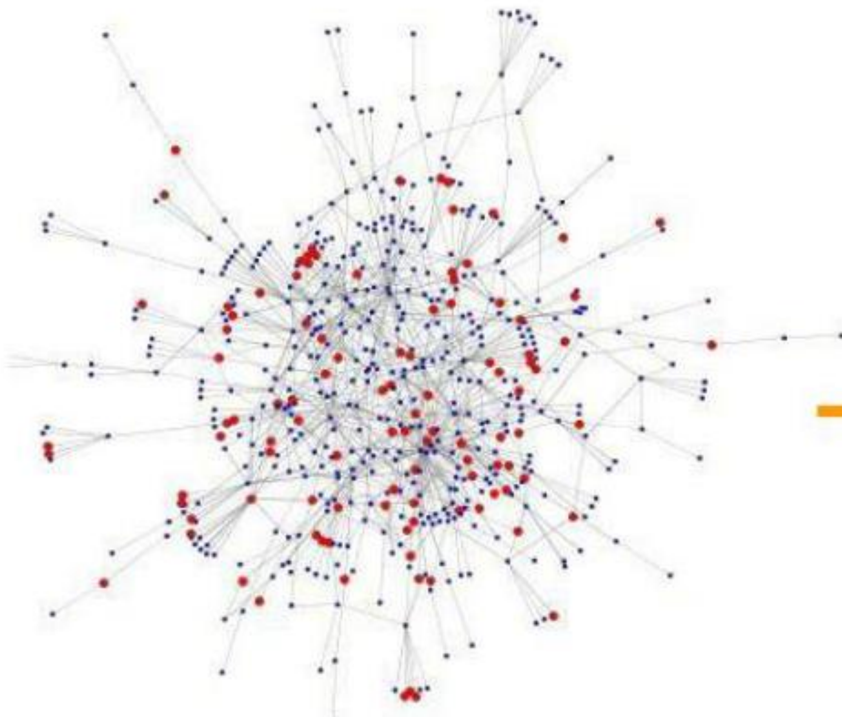
76 edges, average degree 3.04 – still well above percolation threshold

Describes the formation of long-range connectivity in random systems

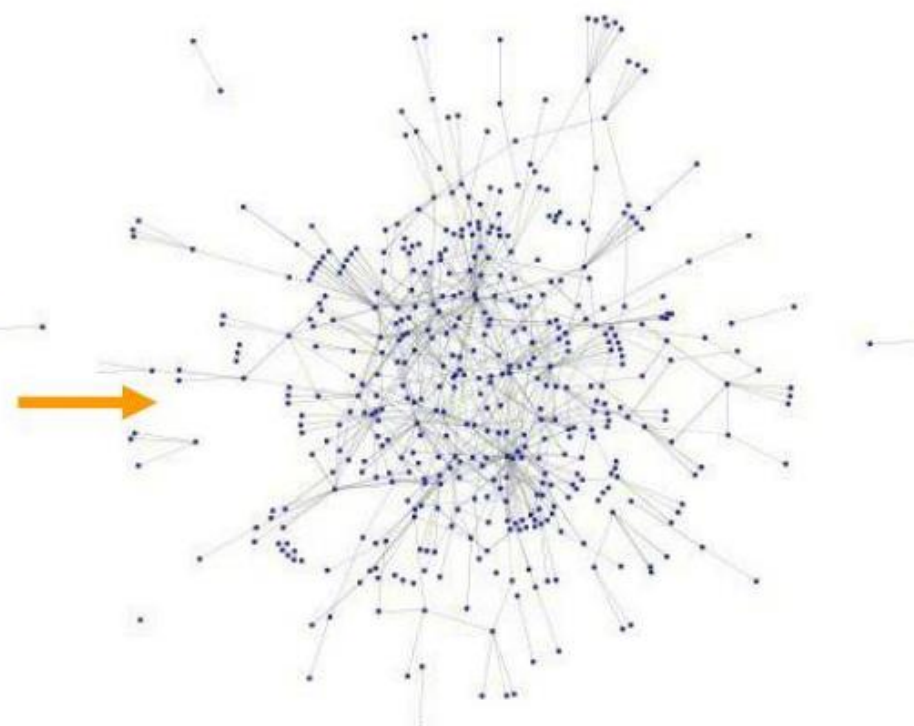
Below the threshold a giant connected component does not exist; while above it, there exists a giant component of the order of *log N*

Scale-free networks are resilient with respect to random attack

- Example: gnutella network, 20% of nodes removed



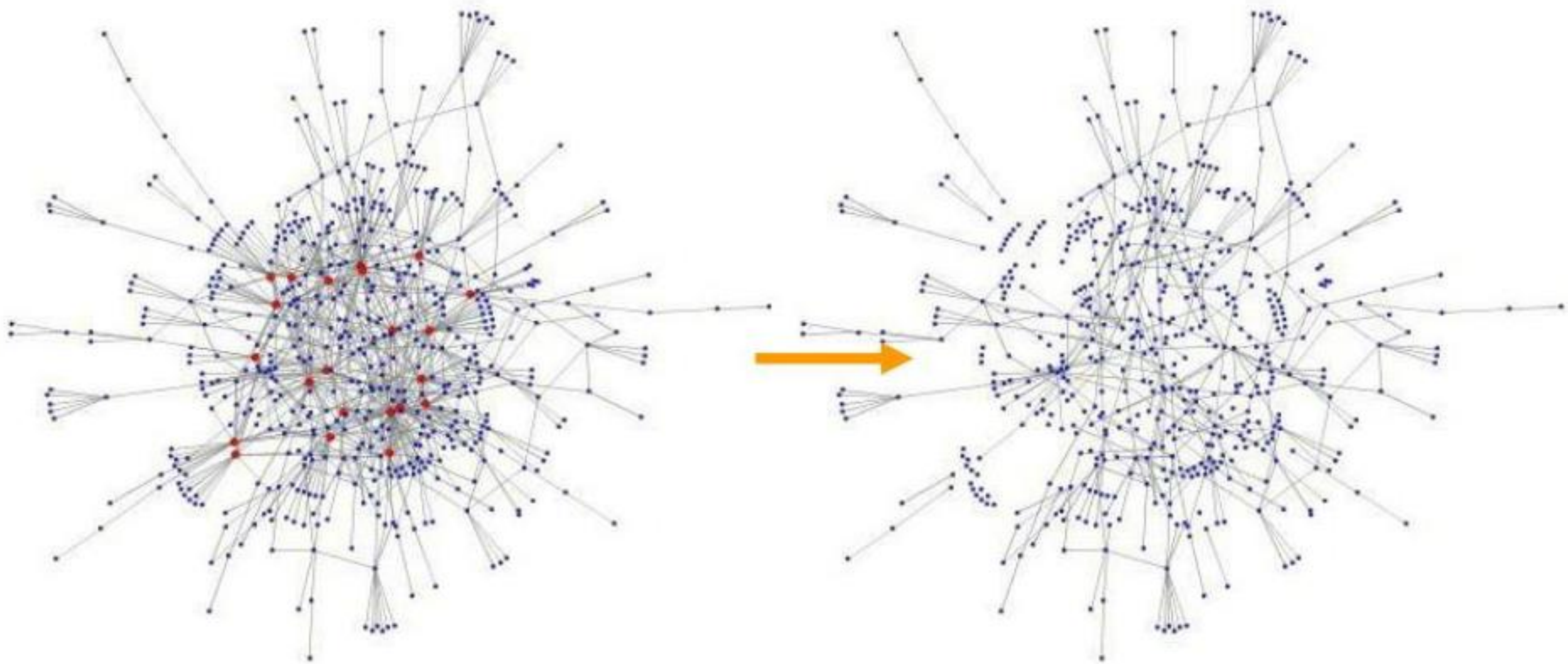
574 nodes in giant component



427 nodes in giant component

Targeted attacks are affective against scale-free networks

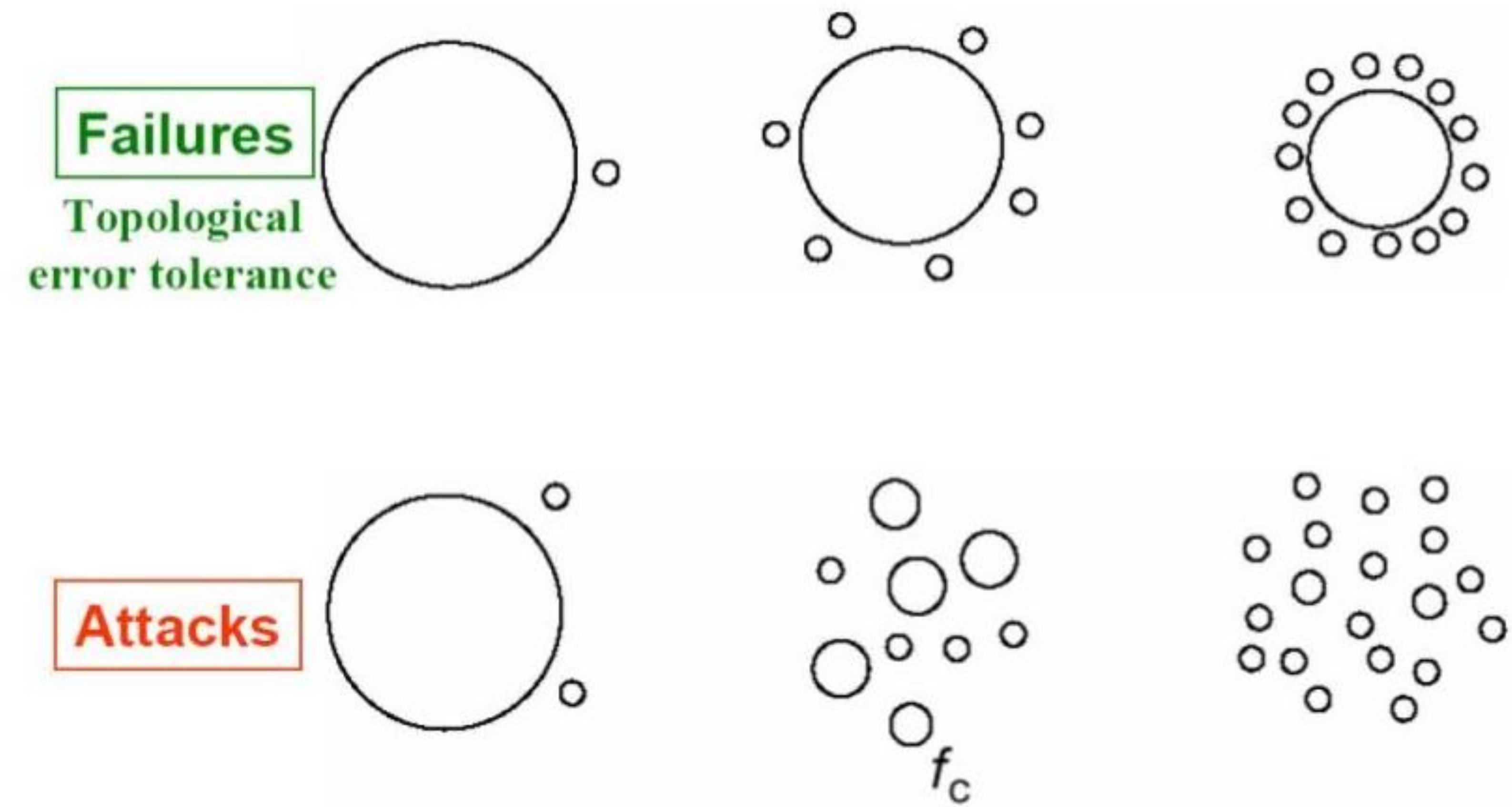
- Example: same gnutella network, 22 most connected nodes removed (2.8% of the nodes)



574 nodes in giant component

301 nodes in giant component

random failures vs. attacks



Critical fraction f_c of the removed nodes

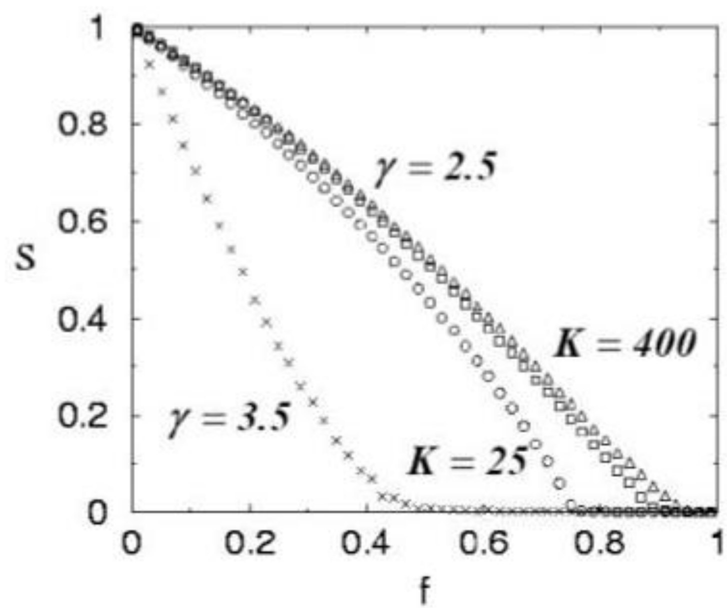
Error tolerance : nodes are removed randomly
Attacks : the most connected nodes are removed

Percolation Threshold scale-free networks

- For scale free networks there is always a giant component, unless nearly all nodes are down

Percolation transition for networks with power-law connectivity distribution

Fraction *S* of nodes that remain in the giant cluster after breakdown of a fraction *f* of all nodes

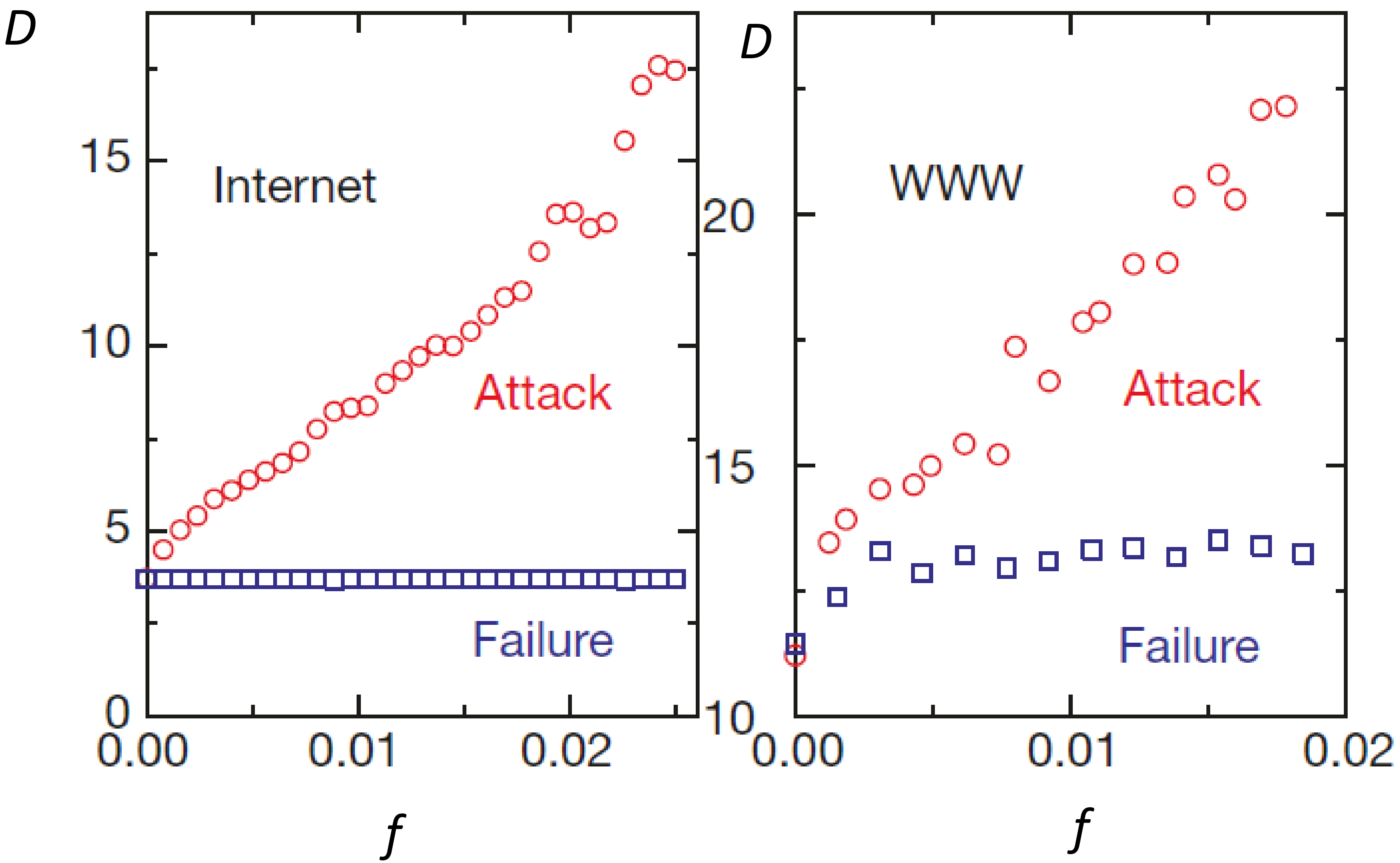


γ : degree distribution power law exponent

K : max degree

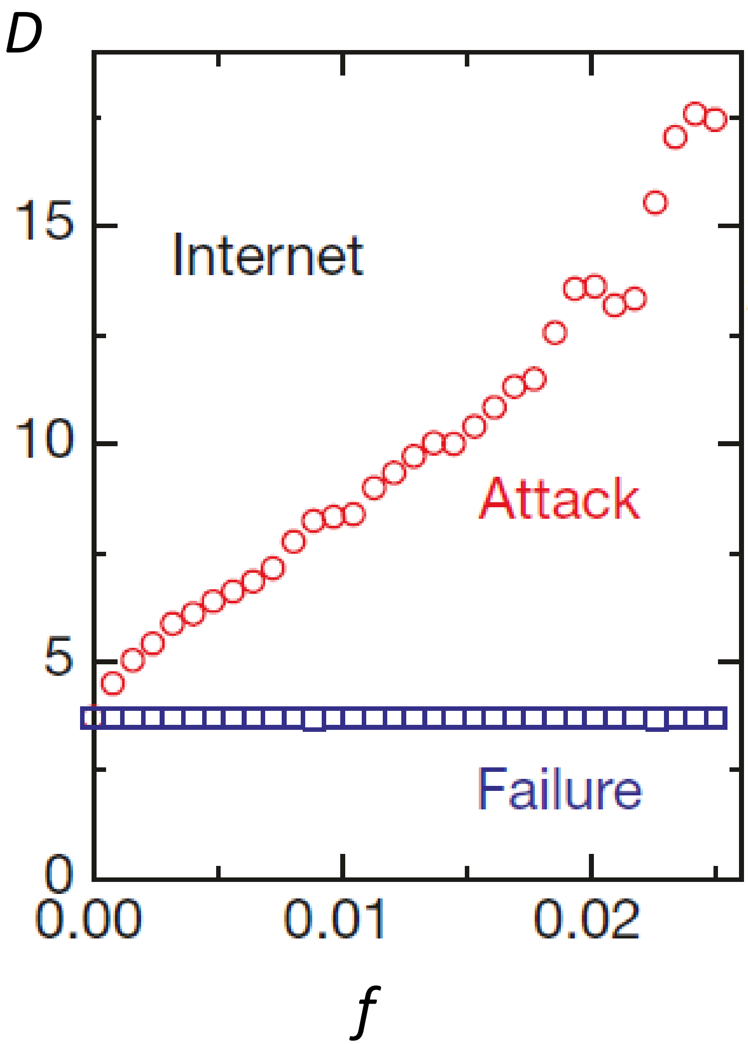
Cohen et al., Phys. Rev. Lett. 85, 4626 (2000)

Changes in the diameter *D* of the network as a function of the fraction *f* of the removed nodes



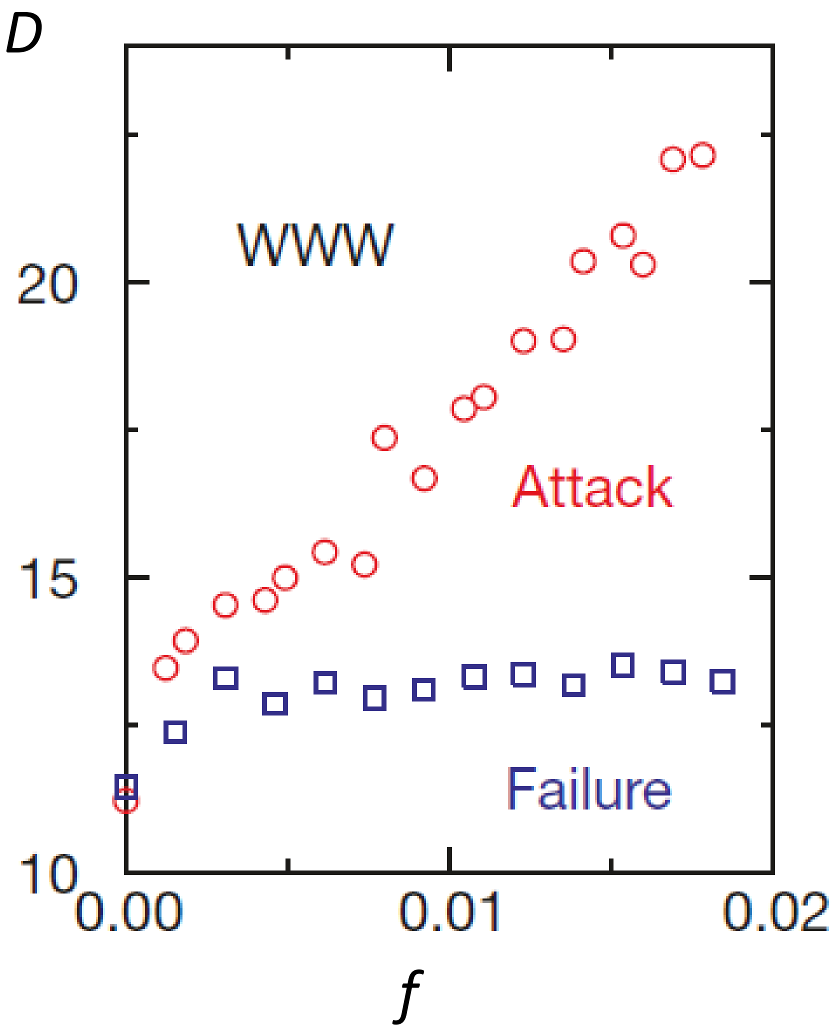
Albert, R., Jeong, H. & Barabási, AL. Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000)

Changes in the diameter *D* of the network as a function of the fraction *f* of the removed nodes



- Diameter *D* : longest shortest path between any two nodes
- Changes in *D* of the Internet under random failures (squares) or attacks (circles)
- Used a topological map of the Internet, containing 6,209 nodes and 12,200 links (<*k*> = 3.4), collected by the National Laboratory for Applied Network Research (http://moat.nlanr.net/Routing/rawdata/)

Changes in the diameter *D* of the network as a function of the fraction *f* of the removed nodes



- Diameter *D* : longest shortest path between any two nodes
- Changes in D of the World-Wide Web under random failures (squares) or attacks (circles)
- **Survivability** of the WWW, measured on a sample containing 325,729 nodes and 1,498,353 links, such that $\langle k \rangle = 4.59$

Breakdown Thresholds Under Random Failures and Attacks

NETWORK	RANDOM FAILURES (REAL NETWORK)	RANDOM FAILURES (RANDOMIZED NETWORK)	ATTACK (REAL NETWORK)
Internet	0.92	0.84	0.16
WWW	0.88	0.85	0.12
Power Grid	0.61	0.63	0.20
Mobile-Phone Call	0.78	0.68	0.20
Email	0.92	0.69	0.04
Science Collaboration	0.92	0.88	0.27
Actor Network	0.98	0.99	0.55
Citation Network	0.96	0.95	0.76
E. Coli Metabolism	0.96	0.90	0.49
Yeast Protein Interactions	0.88	0.66	0.06

- The table shows the estimated f_c for random node failures (2nd column) and attacks (4th column) for 10 reference networks
- The 3rd column (randomized network) shows f_c for an ER network whose size and average path length coincide with the original network

Critical fraction f_c of the removed nodes

Breakdown Thresholds Under Random Failures and Attacks

NETWORK	RANDOM FAILURES (REAL NETWORK)	RANDOM FAILURES (RANDOMIZED NETWORK)	ATTACK (REAL NETWORK)
Internet	0.92	0.84	0.16
WWW	0.88	0.85	0.12
Power Grid	0.61	0.63	0.20
Mobile-Phone Call	0.78	0.68	0.20
Email	0.92	0.69	0.04
Science Collaboration	0.92	0.88	0.27
Actor Network	0.98	0.99	0.55
Citation Network	0.96	0.95	0.76
E. Coli Metabolism	0.96	0.90	0.49
Yeast Protein Interactions	0.88	0.66	0.06

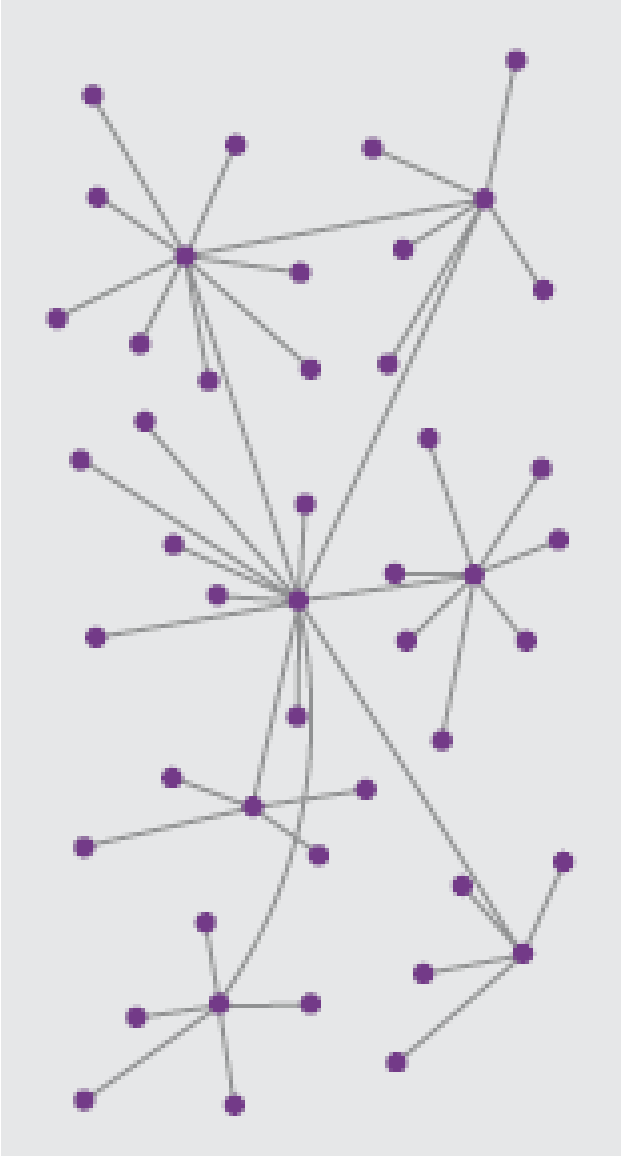
Critical fraction f_c of the removed nodes

- For most networks, f_c for random failures exceeds f_c for the corresponding randomized network, indicating that these networks display enhanced robustness
- **Attacks are much more detrimental than failures**

Paul Baran and the Internet

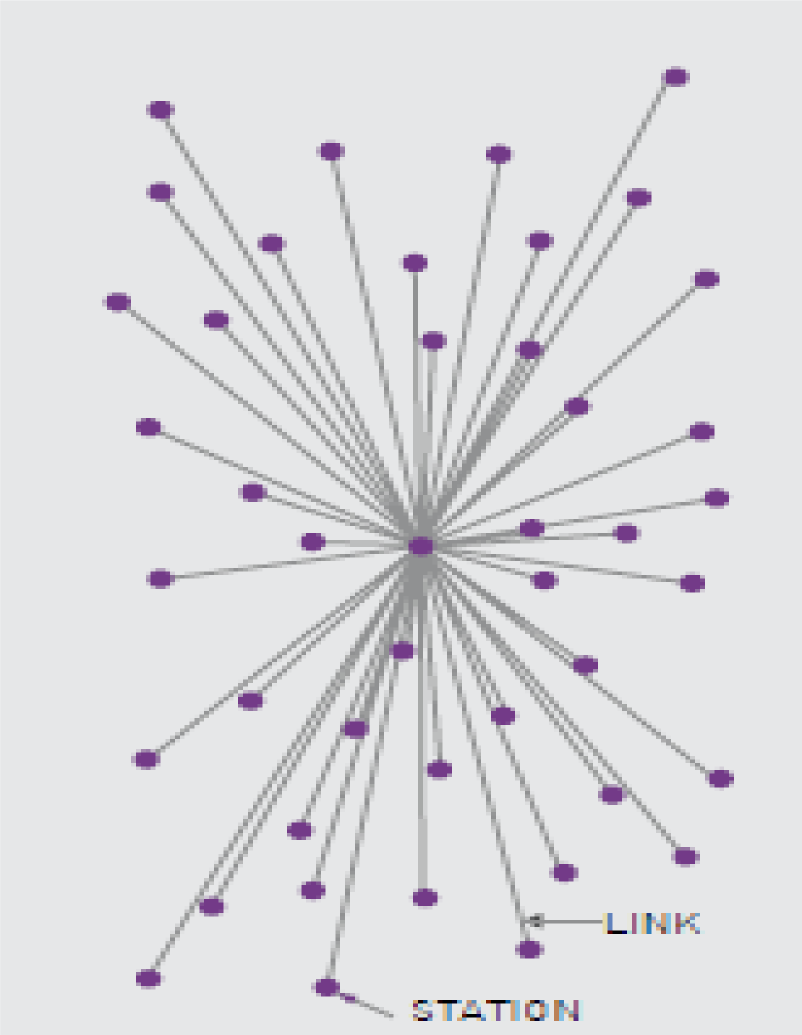
https://en.wikipedia.org/wiki/RAND_Corporation

- In 1959 RAND, a Californian think-tank, assigned **Paul Baran**, a young engineer at that time, to develop a communication system that can survive a Soviet nuclear attack
- As a nuclear strike handicaps all equipment within the range of the detonation, Baran had to design a system whose users outside this range do not loose contact with one another
- He described the communication network of his time as a "hierarchical structure of a set of stars connected in the form of a larger star," offering an early description of what we call today a **scale-free network**
- He concluded that this topology is too centralized to be viable under attack



Paul Baran and the Internet

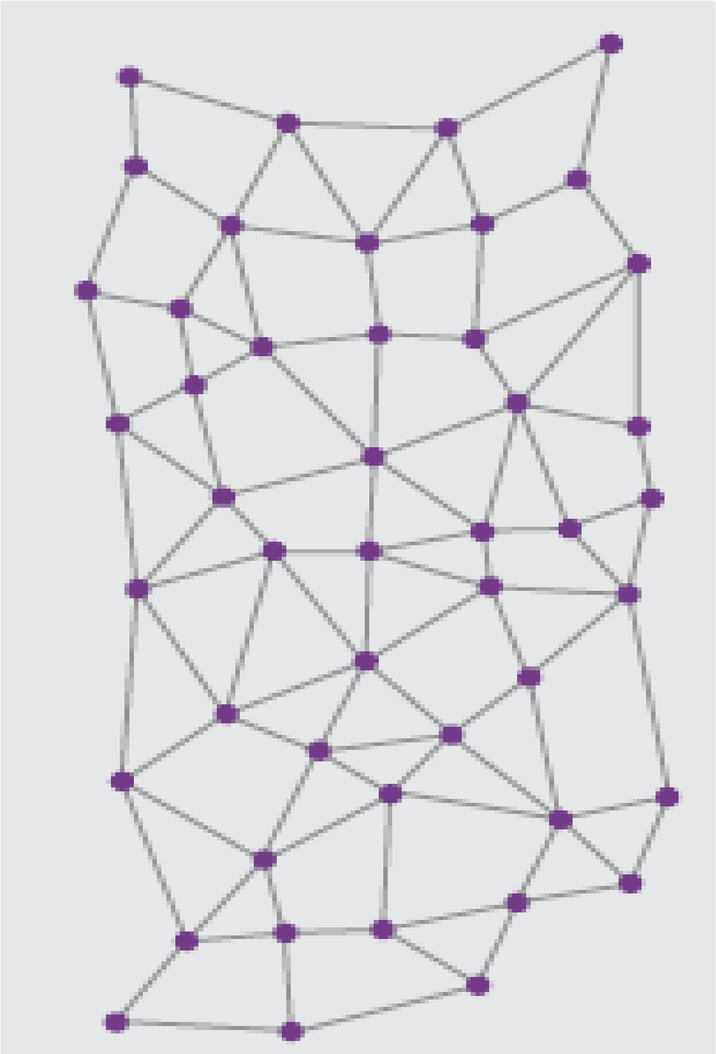
He also discarded the hub-and-spoke topology shown here, noting that the "**centralized network** is obviously vulnerable as destruction of a single central node destroys communication between the end stations."



Paul Baran and the Internet

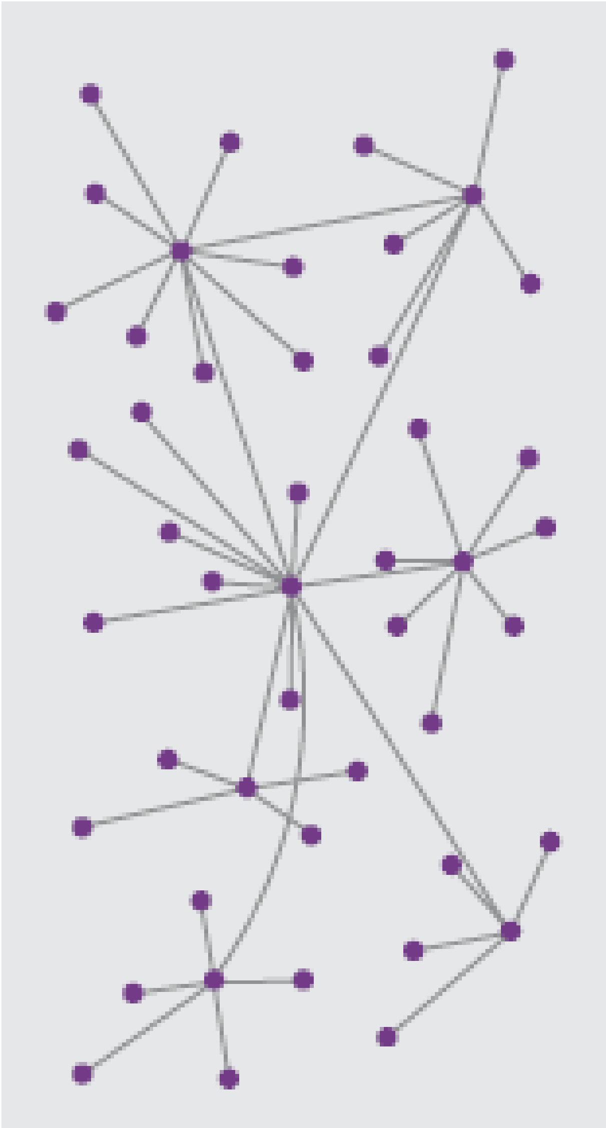
- Baran decided that the ideal survivable architecture was a distributed **mesh-like network**

- This network is sufficiently redundant, so that even if some of its nodes fail, alternative paths can connect the remaining nodes



Paul Baran and the Internet

- **Baran's ideas were ignored by the military**, so when the Internet was born a decade later, it instead relied on distributed protocols that allowed each node to decide where to link
- This decentralized philosophy paved the way to the emergence of a scale-free Internet, rather than the uniform mesh-like topology envisioned by Baran



Ackowledgement

- Albert-László Barabási, "Network science," *Cambridge University Press*, 2016

- "Network resilience," School of Information, *University of Michigan*, 2016 (available online)