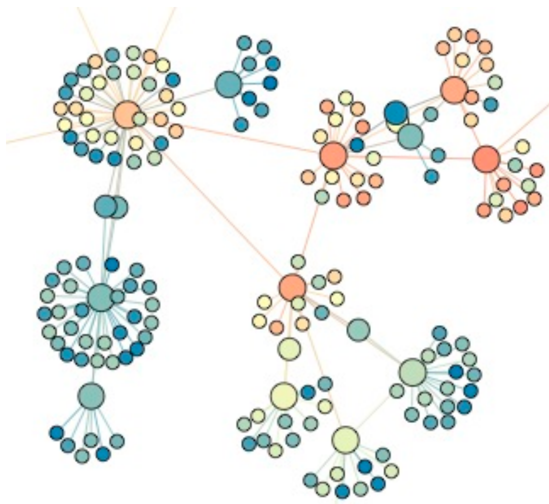


SNA 8: network resilience

Lada Adamic

school of information

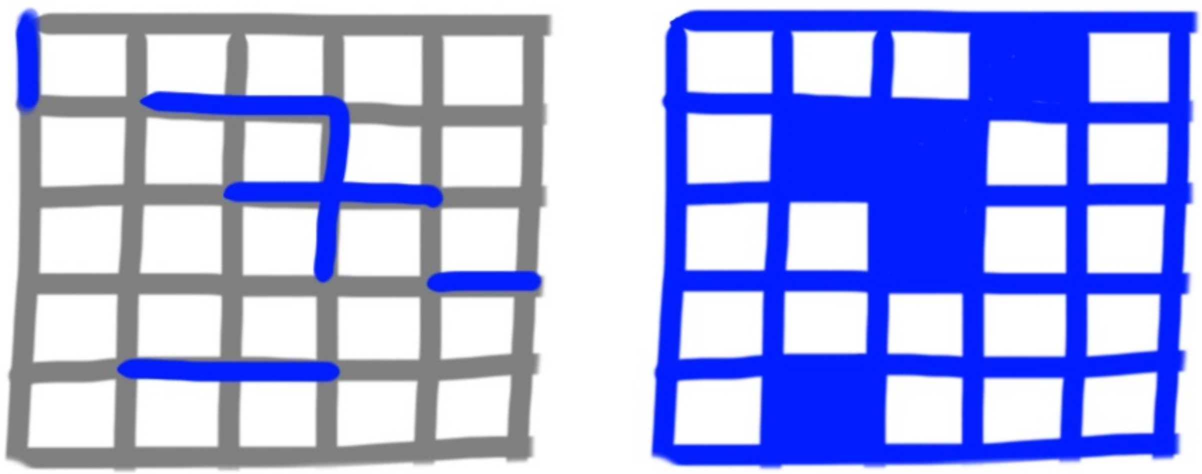
Outline

- Node vs. edge percolation
- Resilience of randomly vs. preferentially grown networks
- Resilience in real-world networks

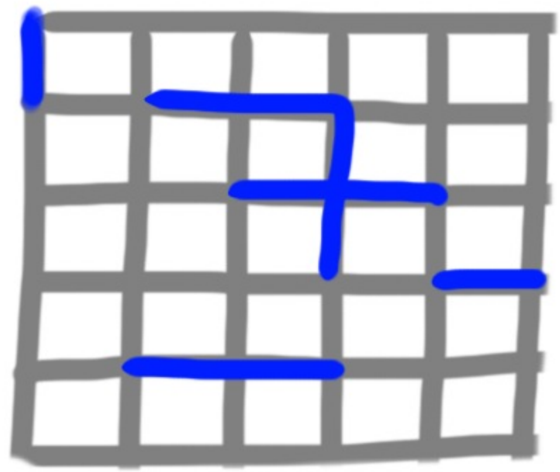
network resilience

- 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 (previously studied on lattices):

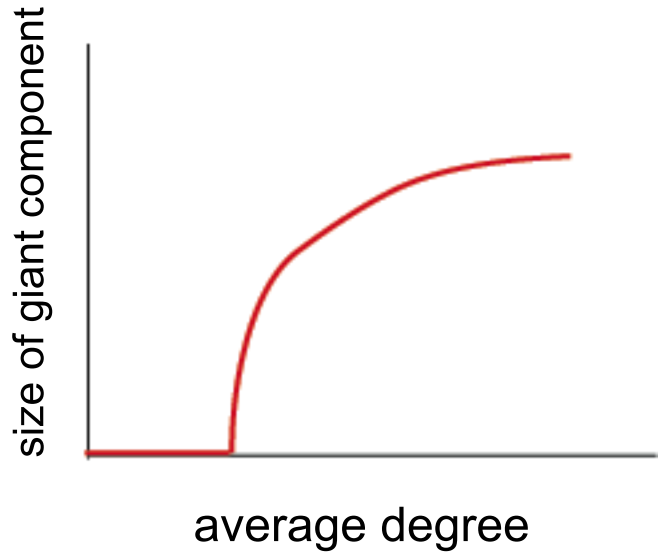


edge percolation

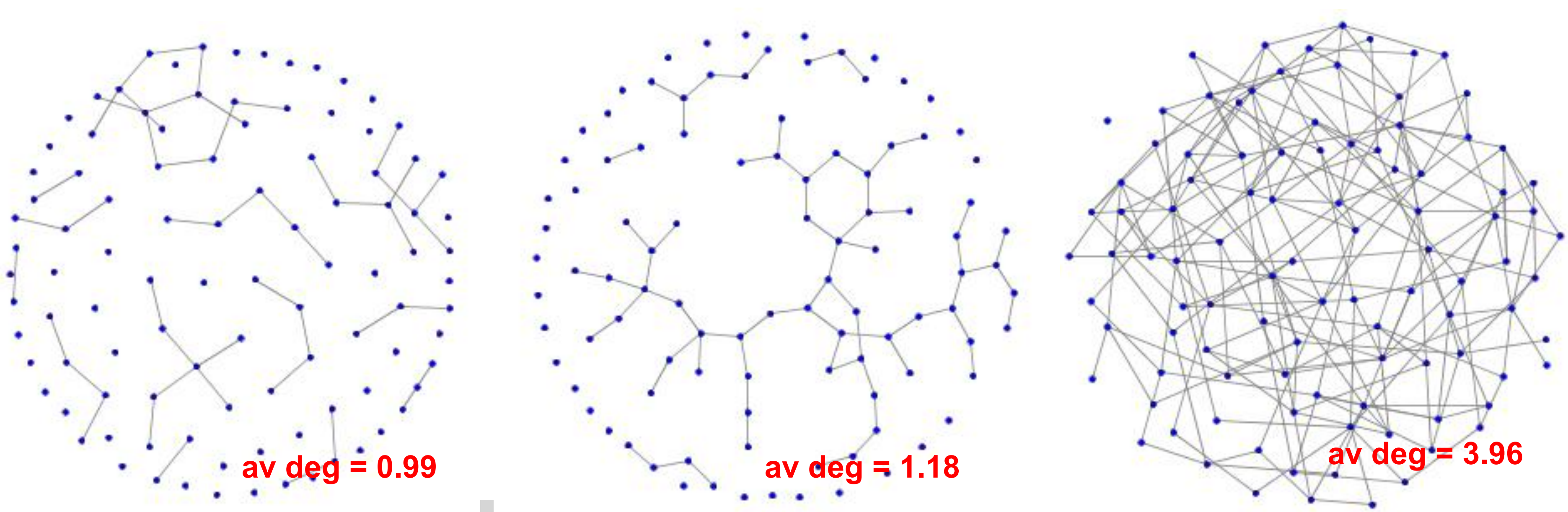


- Edge removal
  - bond percolation: each edge is removed with probability (1-p)
    - corresponds to random failure of links
  - targeted attack: causing the most damage to the network with the removal of the fewest edges
    - strategies: remove edges that are most likely to break apart the network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

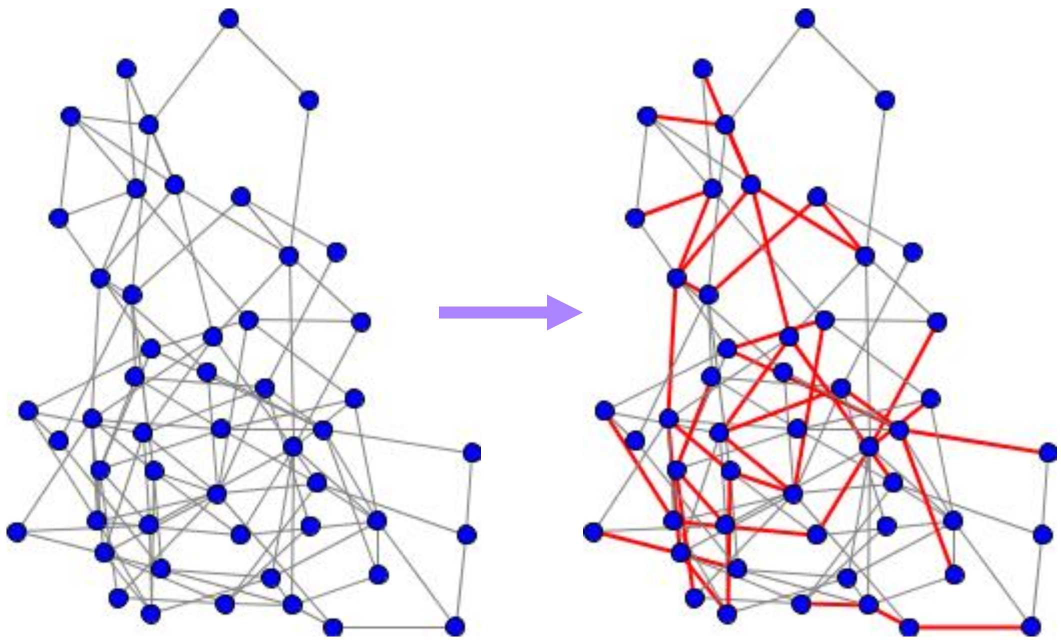
reminder: percolation in ER graphs



- 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

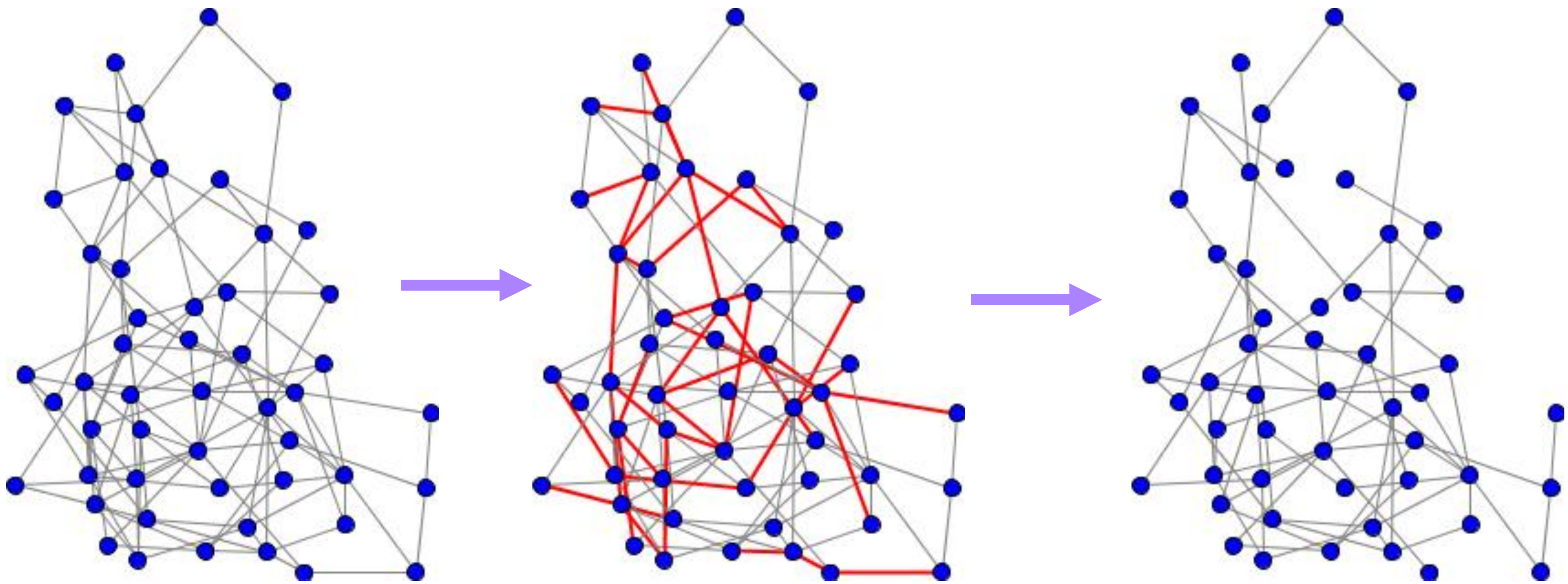


Quiz Q:



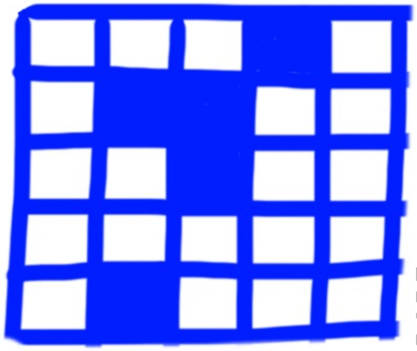
In this network each node has average degree 4.64, if you removed 25% of the edges, by how much would you reduce the giant component?

edge percolation

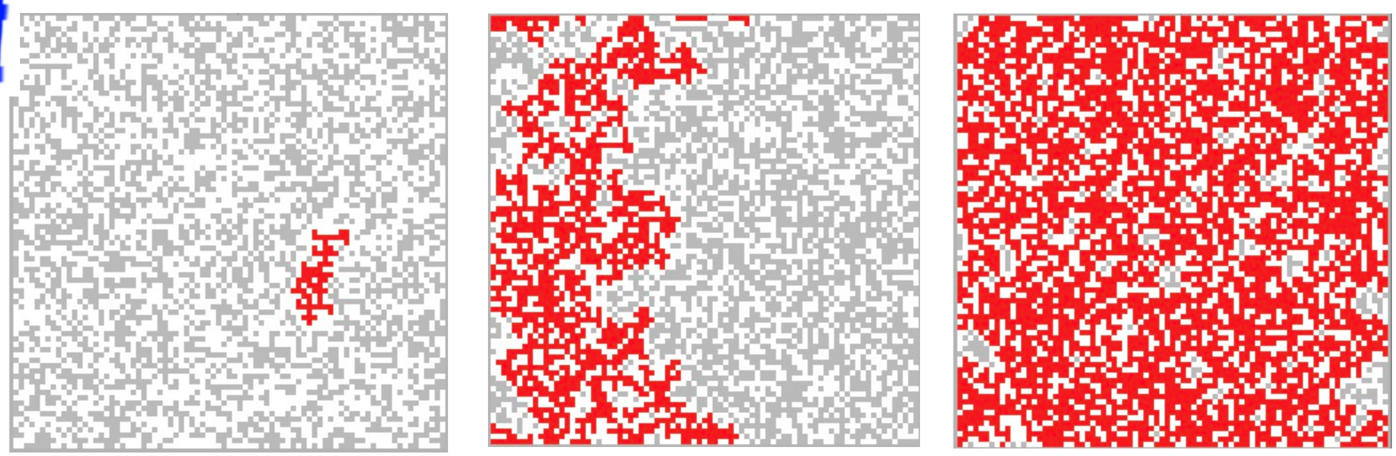


50 nodes, 116 edges, average degree 4.64
after 25 % edge removal
76 edges, average degree 3.04 – still well above percolation threshold

node removal and site percolation



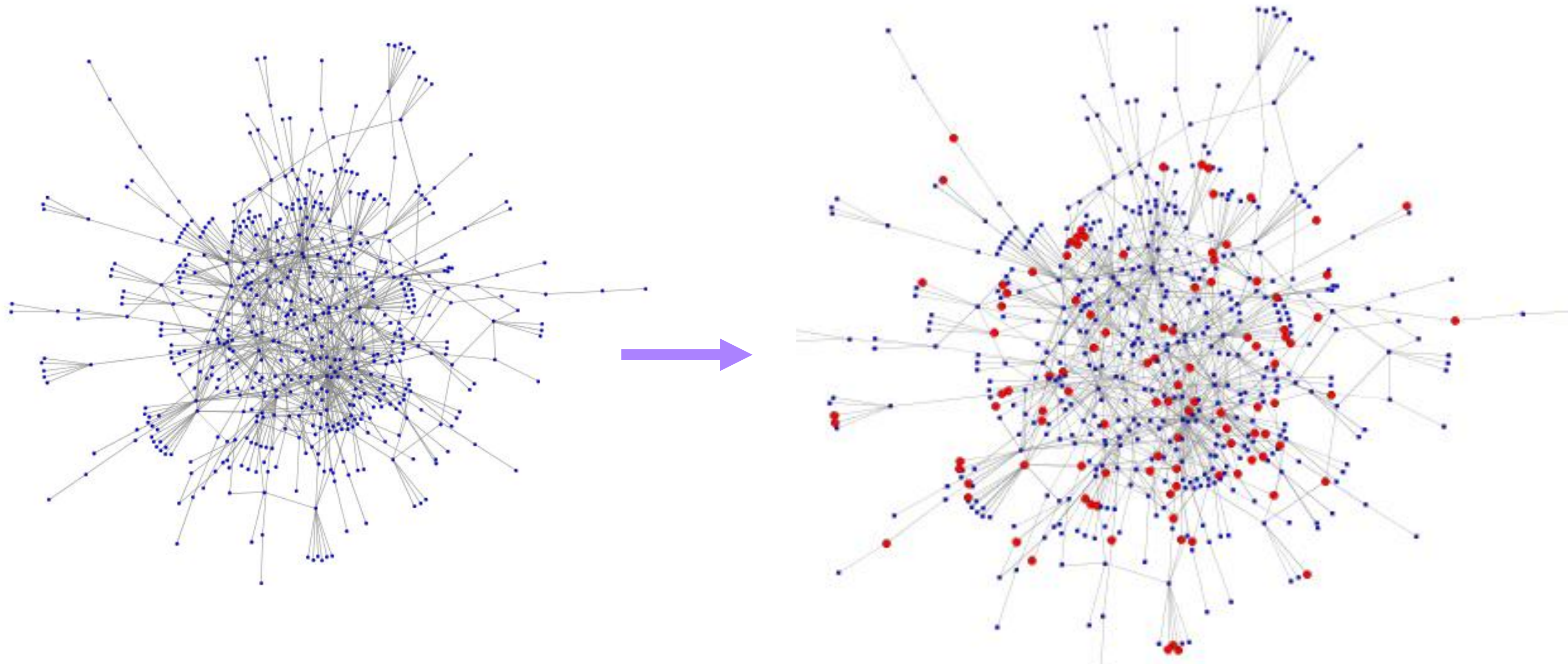
Ordinary Site Percolation on Lattices:
Fill in each site (site percolation) with probability p



- **low p**: small islands
- **p critical**: giant component forms, occupying finite fraction of infinite lattice.
  p above critical value: giant component occupies an increasingly larger portion of the graph

http://www.ladamic.com/netlearn/NetLogo501/LatticePercolation.html

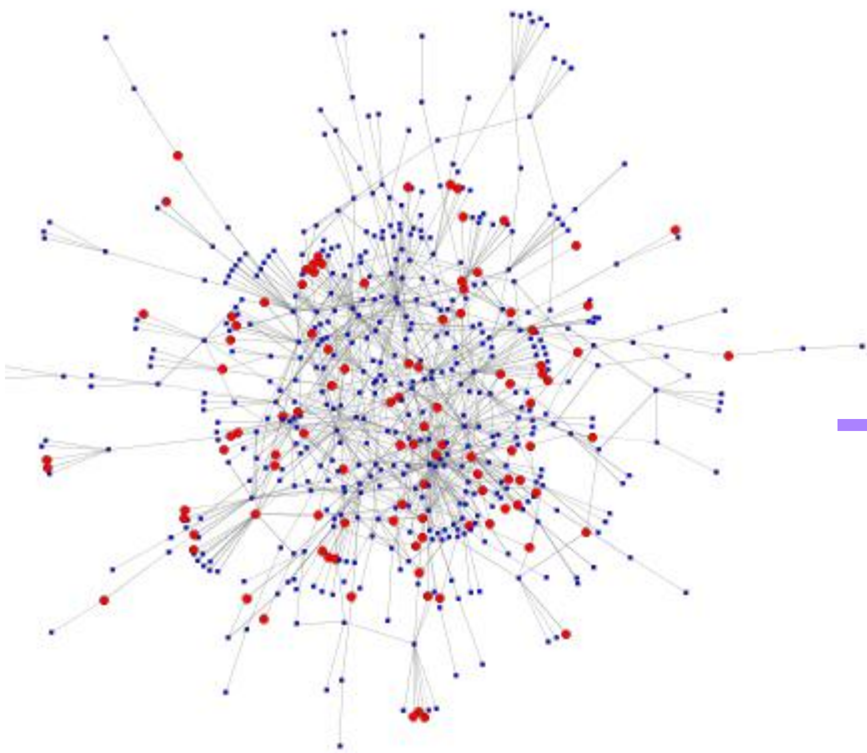
Percolation on networks



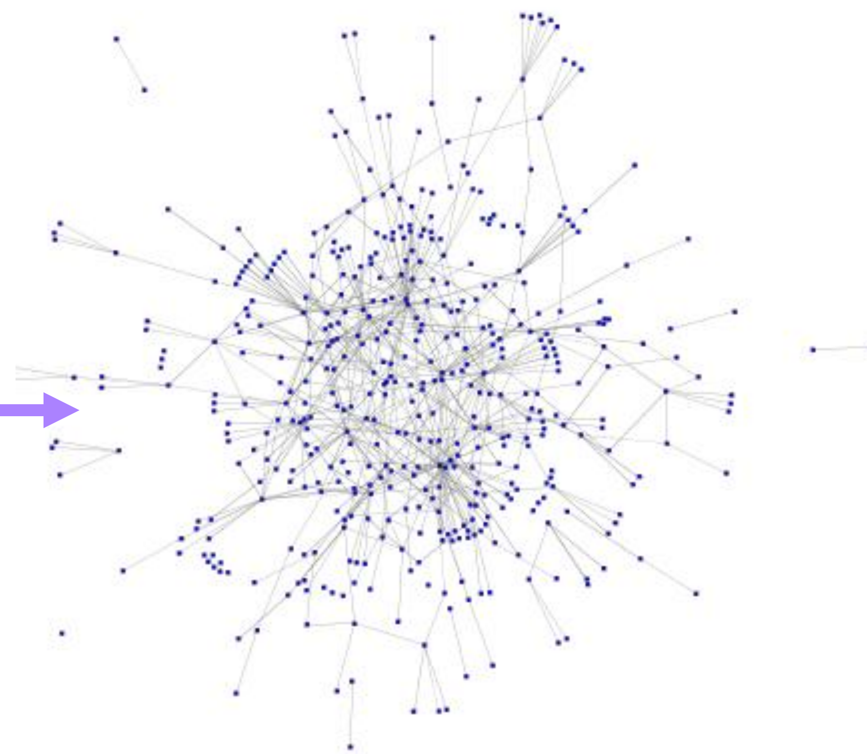
- Percolation can be extended to networks of arbitrary topology.
- We say the network percolates when a giant component forms.

Random attack on scale-free networks

- Example: gnutella filesharing network, 20% of nodes removed at random



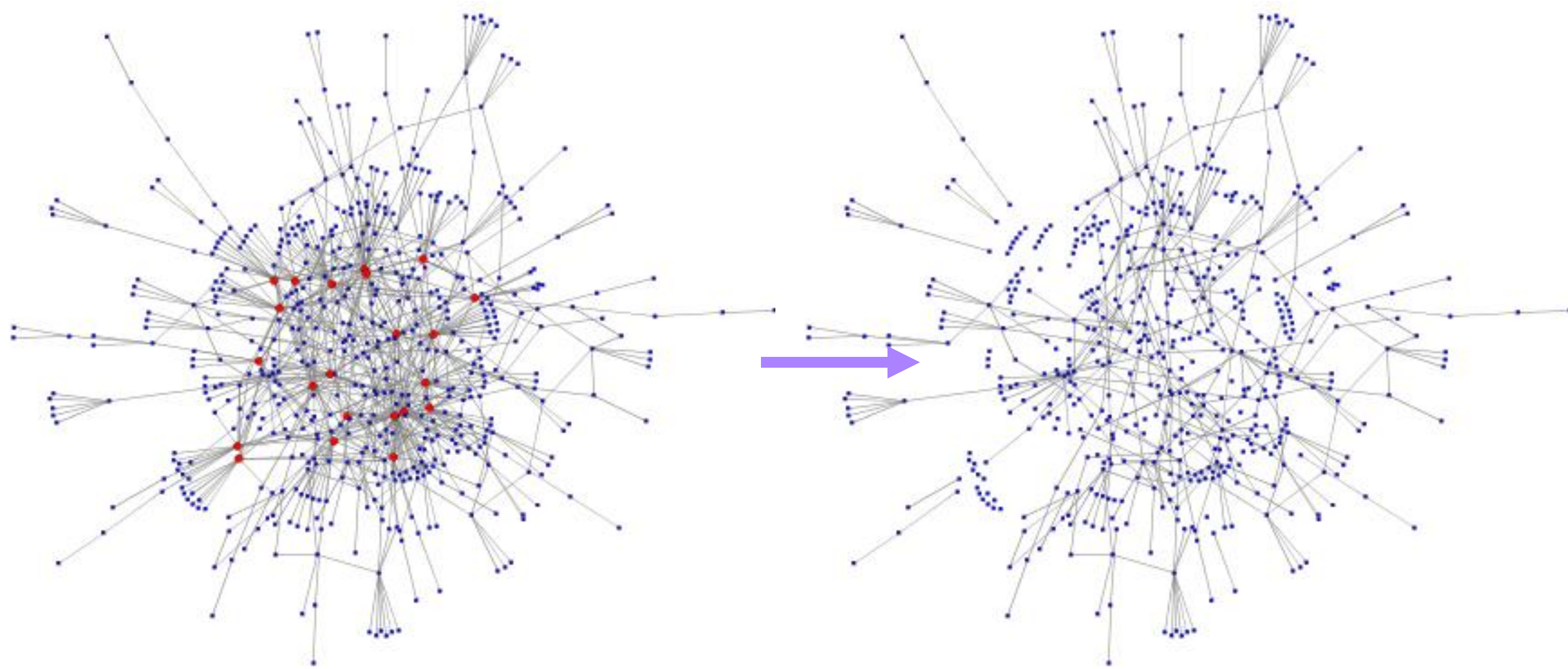
574 nodes in giant component



427 nodes in giant component

Targeted attacks on power-law networks

- Power-law networks are vulnerable to targeted attack
- Example: same gnutella network, 22 most connected nodes removed (2.8% of the nodes)



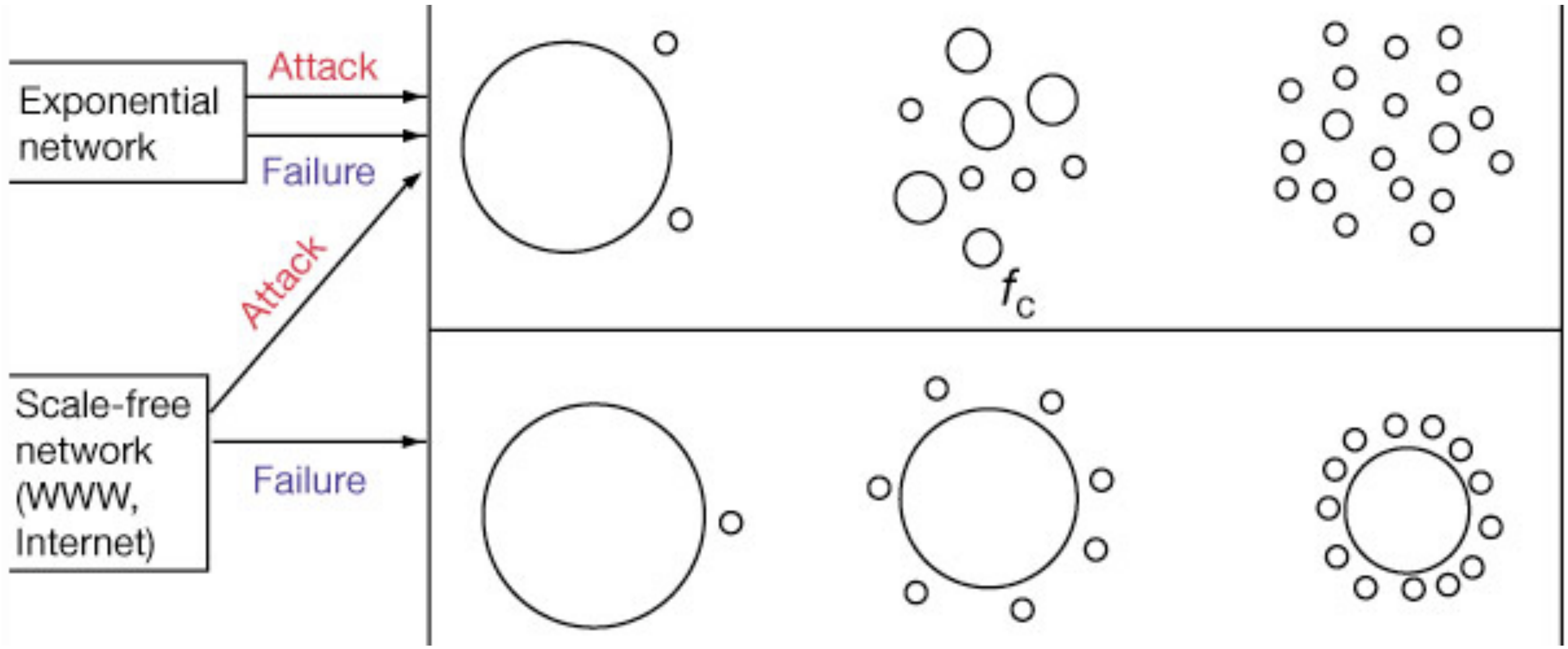
574 nodes in giant component

301 nodes in giant component

Quiz Q:

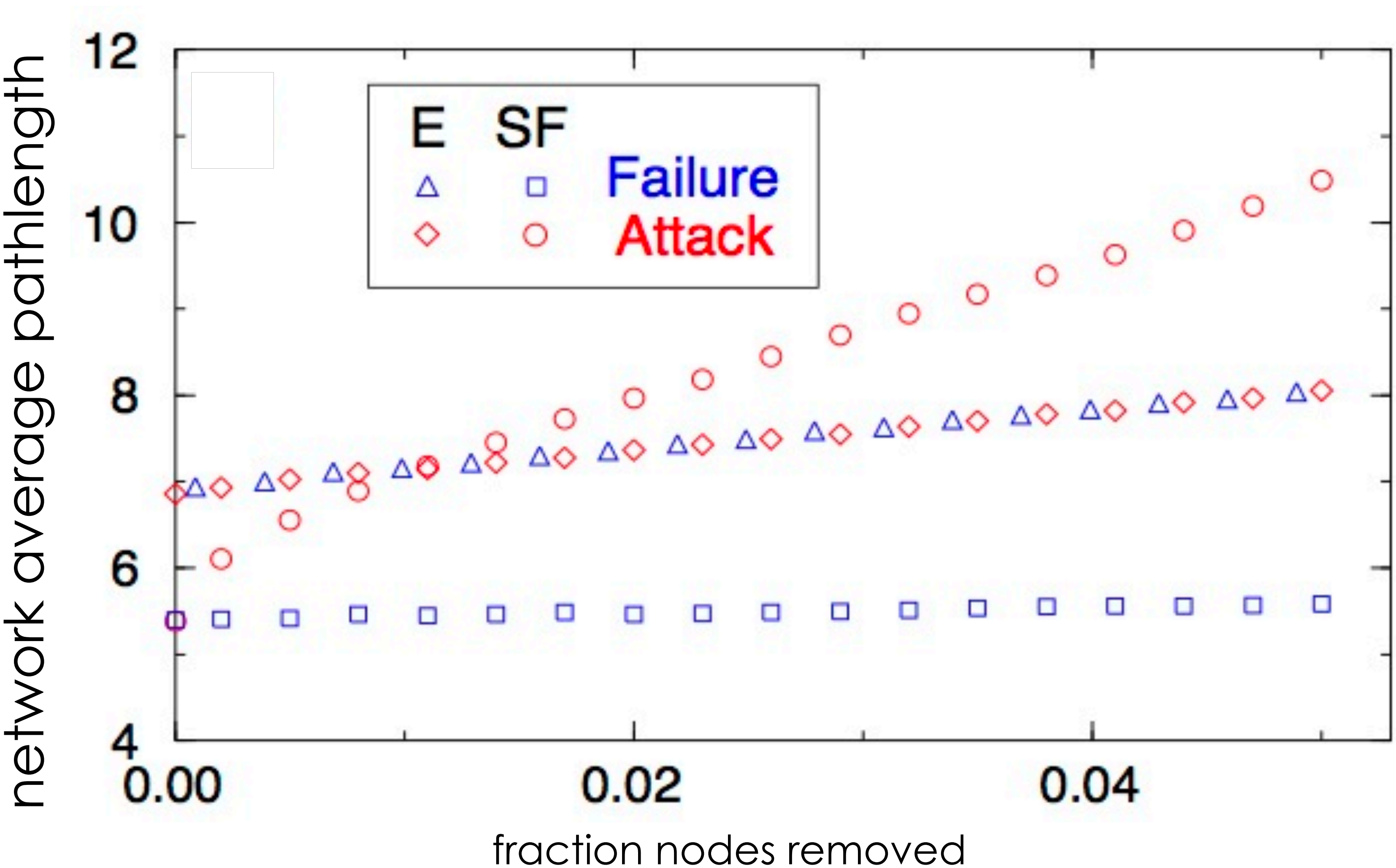
- Why is removing high-degree nodes more effective?
  - it removes more nodes
  - it removes more edges
  - it targets the periphery of the network

random failures vs. attacks



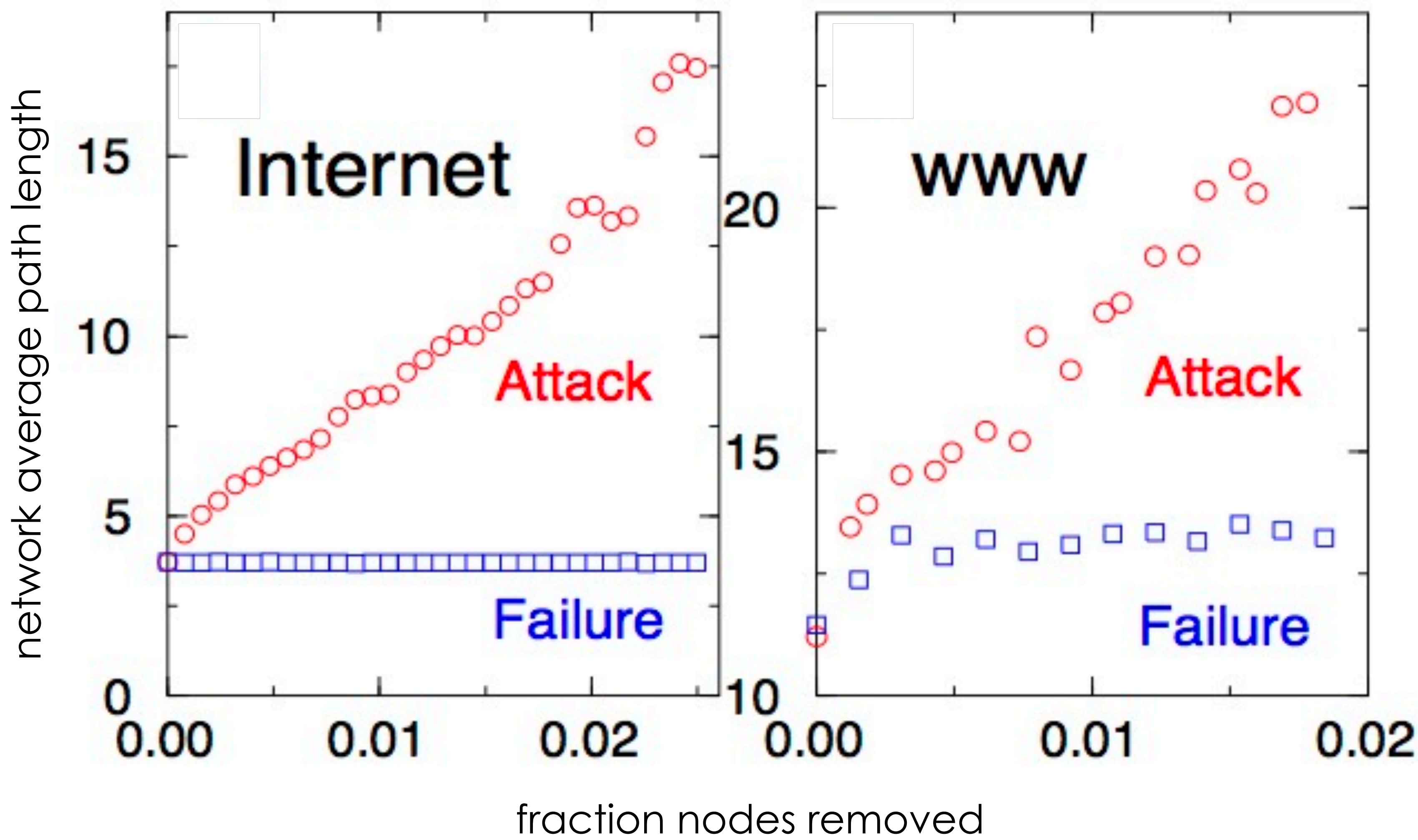
Source: Error and attack tolerance of complex networks. Réka Albert, Hawoong Jeong and Albert-László Barabási. Nature 406, 378-382(27 July 2000); http://www.nature.com/nature/journal/v406/n6794/abs/406378A0.html

effect on path length



Source: Error and attack tolerance of complex networks. Réka Albert, Hawoong Jeong and Albert-László Barabási. Nature 406, 378-382(27 July 2000); http://www.nature.com/nature/journal/v406/n6794/abs/406378A0.html

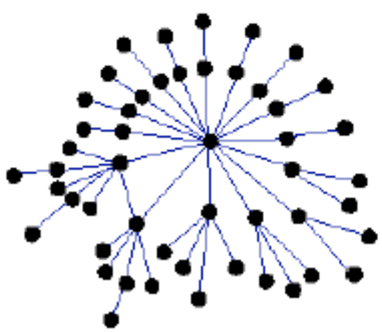
applied to empirical networks



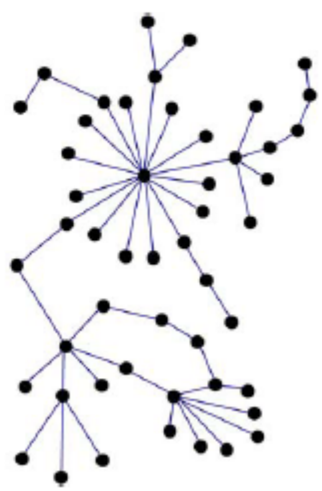
Source: Error and attack tolerance of complex networks. Réka Albert, Hawoong Jeong and Albert-László Barabási. Nature 406, 378-382(27 July 2000); http://www.nature.com/nature/journal/v406/n6794/abs/406378A0.html

Assortativity

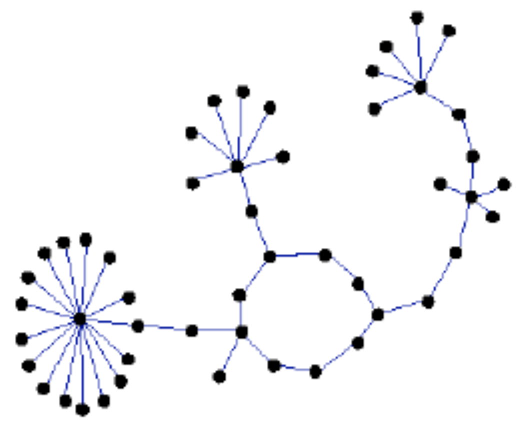
- Social networks are assortative:
  - the gregarious people associate with other gregarious people
  - the loners associate with other loners
- The Internet is disassortative:



Assortative:
hubs connect to hubs



Random

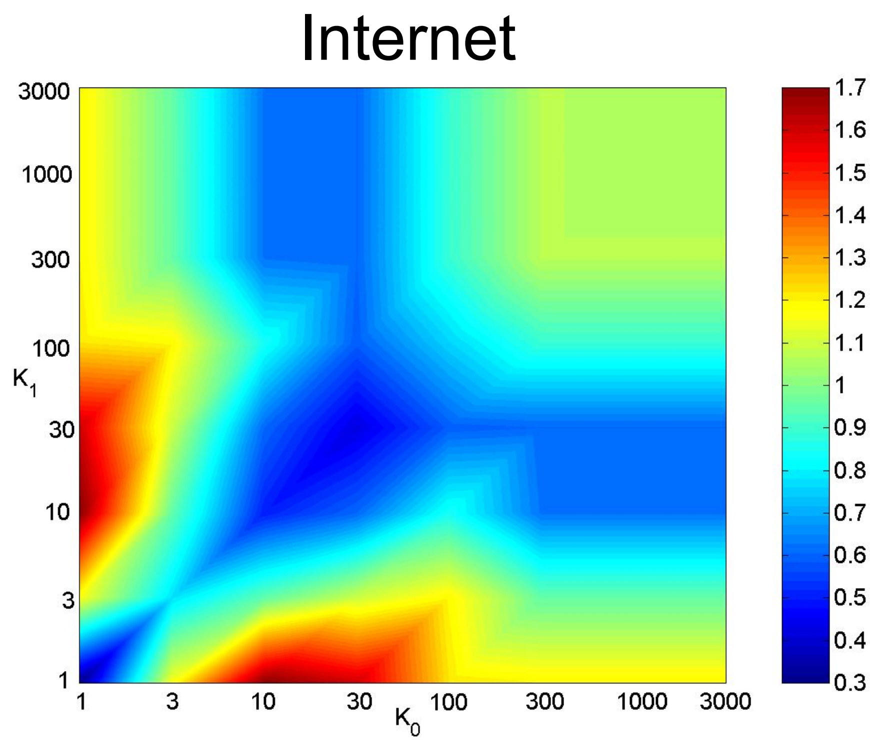


Disassortative:
hubs are in the periphery

Correlation profile of a network

- Detects preferences in linking of nodes to each other based on their connectivity
- Measure $N(k_0,k_1)$ – the number of edges between nodes with connectivities k_0 and k_1
- Compare it to $N_r(k_0,k_1)$ – the same property in a properly randomized network
- Very noise-tolerant with respect to both false positives and negatives

Degree correlation profiles: 2D

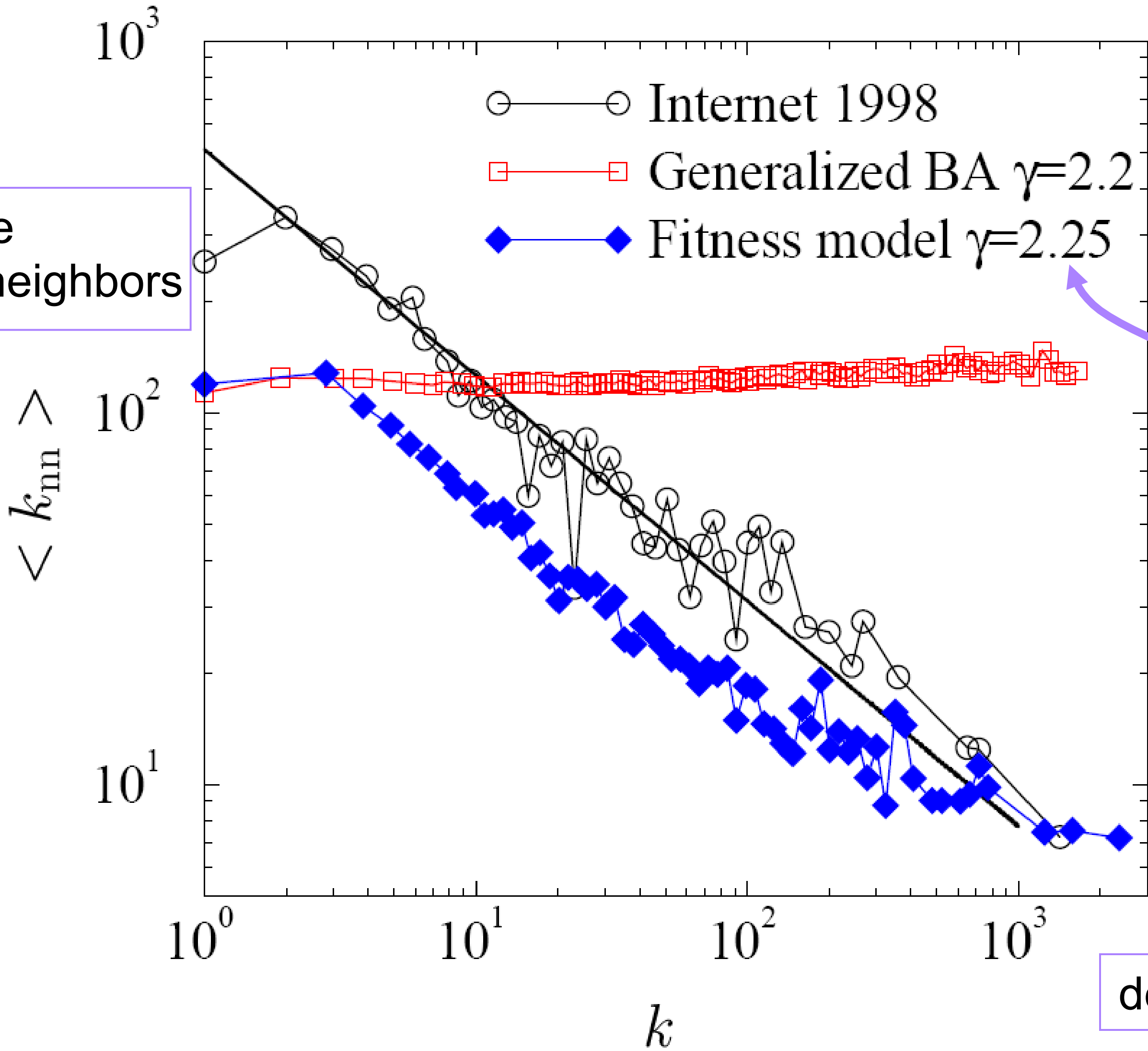


source: Sergei Maslov

Average degree of neighbors

- Pastor-Satorras and Vespignani: 2D plot

average degree
of the node's neighbors



probability
of aquiring
edges is
dependent
on 'fitness'
+ degree
Bianconi &
Barabasi

degree of node

Single number

- cor(deg(i),deg(j)) over all edges {ij}

$$\rho_{internet} = -0.189$$

The Pearson correlation coefficient of nodes on each side on an edge

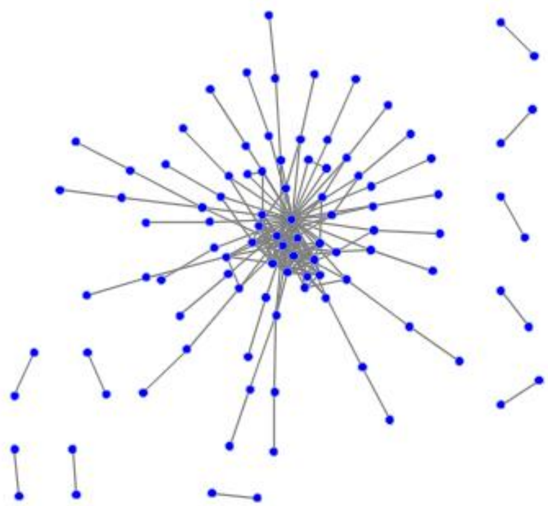
assortative mixing more generally

- Assortativity is not limited to degree-degree correlations other attributes
  - social networks: race, income, gender, age
  - food webs: herbivores, carnivores
  - internet: high level connectivity providers, ISPs, consumers

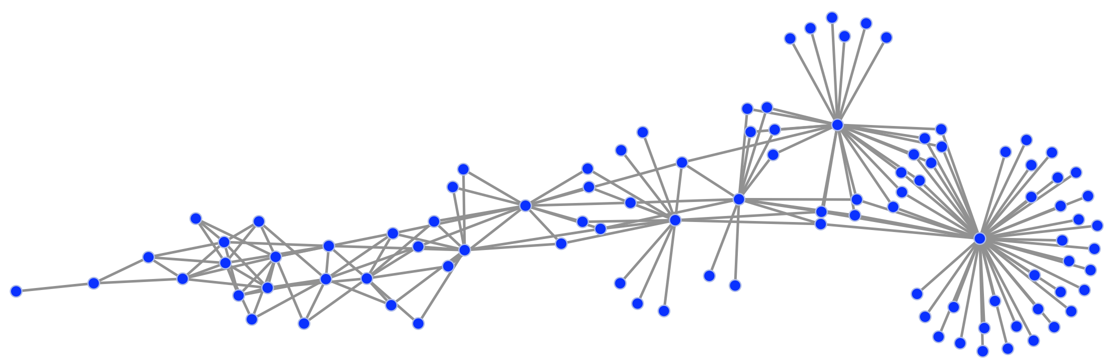
- Tendency of like individuals to associate = ‘homophily’

Quiz Q:

will a network with positive or negative degree assortativity be more resilient to attack?



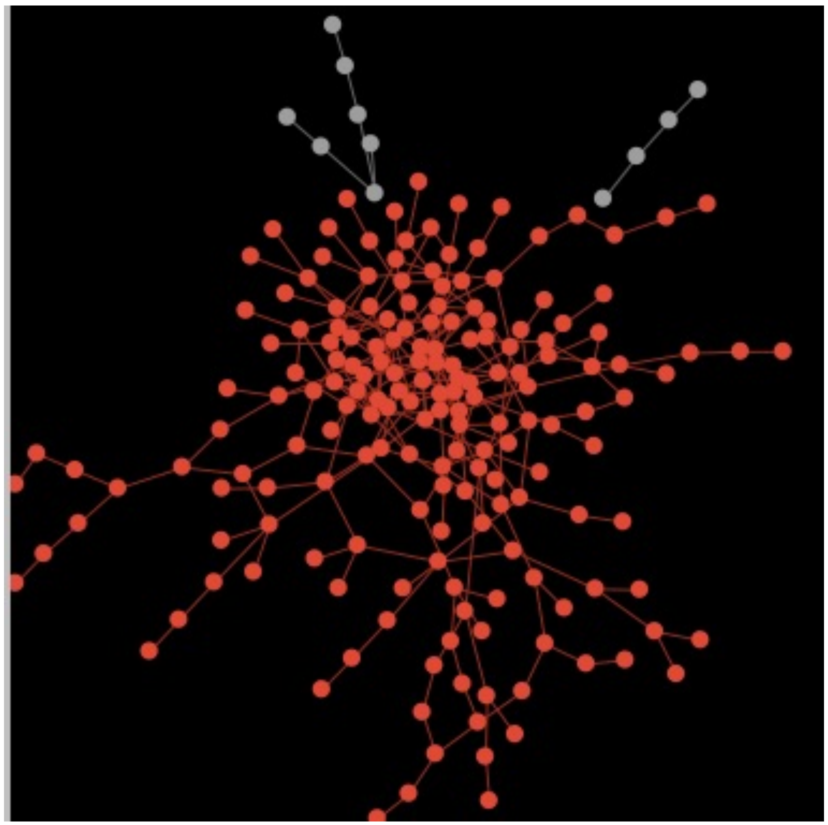
assortative



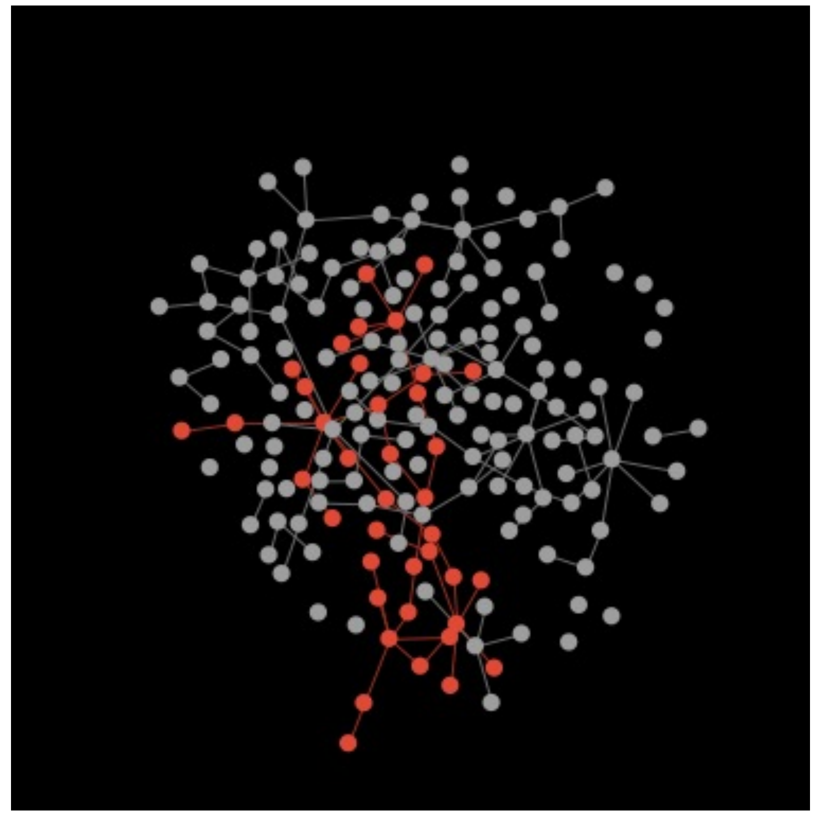
disassortative

Assortativity and resilience

assortative



disassortative

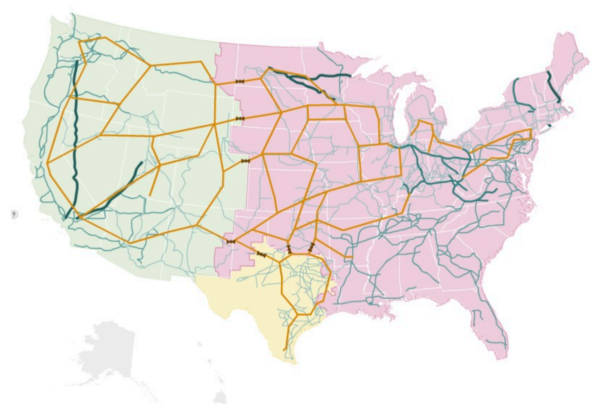


Is it really that simple?

- Internet?
- terrorist/criminal networks?

Power grid

- Electric power flows simultaneously through multiple paths in the network.
- For visualization of the power grid, check out NPR's interactive visualization: http://www.npr.org/templates/story/story.php?storyId=110997398



Cascading failures

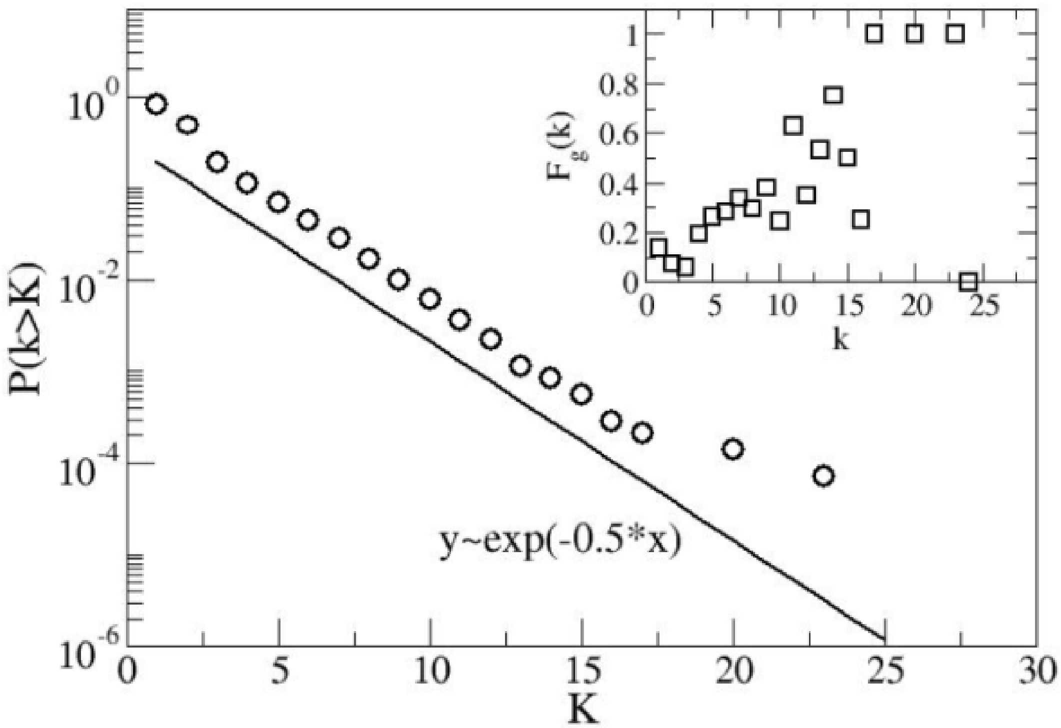
- Each node has a load and a capacity that says how much load it can tolerate.
- When a node is removed from the network its load is redistributed to the remaining nodes.
- If the load of a node exceeds its capacity, then the node fails

Case study: US power grid

Modeling cascading failures in the North American power grid
R. Kinney, P. Crucitti, R. Albert, and V. Latora, Eur. Phys. B, 2005

- Nodes: generators, transmission substations, distribution substations
- Edges: high-voltage transmission lines
- 14099 substations:
  - N_G 1633 generators,
  - N_D 2179 distribution substations
  - N_T the rest transmission substations
- 19,657 edges

Degree distribution is exponential



$$P(k > K) \approx \exp(-0.5K)$$

Efficiency of a path

- efficiency *e* [0,1], 0 if no electricity flows between two endpoints, 1 if the transmission lines are working perfectly

- harmonic composition for a path

$$e_{path} = \left[\sum_{edges} \frac{1}{e_{edge}} \right]^{-1}$$

  - path A, 2 edges, each with e=0.5, e_{path} = 1/4
  - path B, 3 edges, each with e=0.5 e_{path} = 1/6
  - path C, 2 edges, one with e=0 the other with e=1, e_{path} = 0

- simplifying assumption: electricity flows along most efficient path

Efficiency of the network

- Efficiency of the network:
  - average over the most efficient paths from each generator to each distribution station

$$E = \frac{1}{N_G N_D} \sum_{i \in G_G} \sum_{j \in G_D} \epsilon_{ij}$$

ε_{ij} is the efficiency of the most efficient path between *i* and *j*

capacity and node failure

- Assume capacity of each node is proportional to initial load

$$C_i = \alpha L_i(0) \quad i = 1, 2..N$$

  - L represents the weighted betweenness of a node
  - Each neighbor of a node is impacted as follows

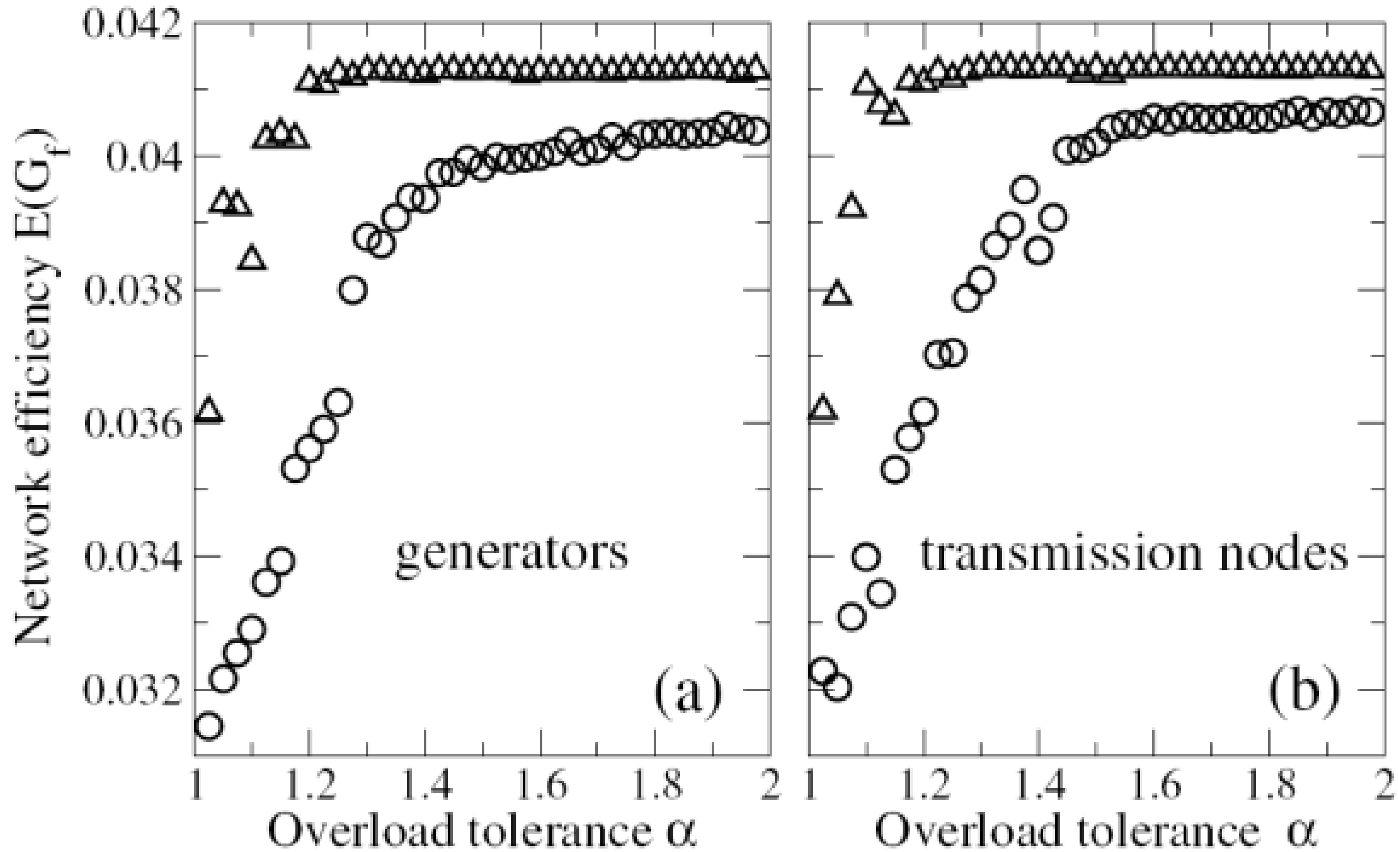
$$e_{ij}(t+1) = \begin{cases} e_{ij}(0) / \frac{L_i(t)}{C_i} \text{ if } L_i(t) > C_i \\ e_{ij}(0) \text{ if } L_i(t) \leq C_i \end{cases}$$

load exceeds capacity

- Load is distributed to other nodes/edges
- The greater a (reserve capacity), the less susceptible the network to cascading failures due to node failure

power grid structural resilience

- efficiency is impacted the most if the node removed is the one with the highest load



○ highest load generator/transmission station removed

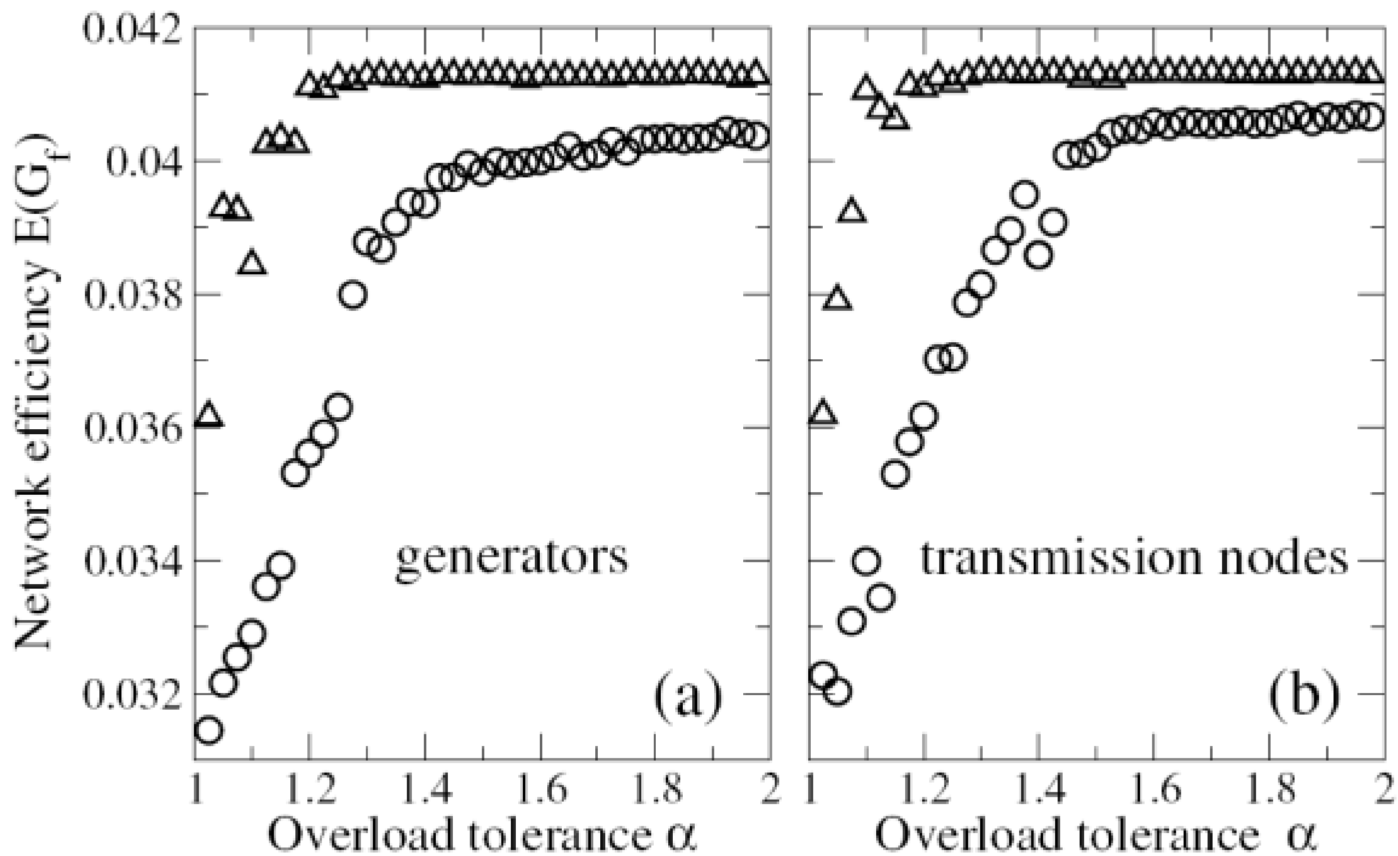
Source: Modeling cascading failures in the North American power grid; R. Kinney, P. Crucitti, R. Albert, V. Latora, Eur. Phys. B, 2005

Quiz Q:

- Approx. how much higher would the capacity of a node need to be relative to the initial load in order for the network to be efficient? (remember capacity $C = \alpha * L(0)$, the initial load).

power grid structural resilience

- efficiency is impacted the most if the node removed is the one with the highest load



○ highest load generator/transmission station removed

Source: Modeling cascading failures in the North American power grid; R. Kinney, P. Crucitti, R. Albert, V. Latora, Eur. Phys. B, 2005

recap: network resilience

- resilience depends on topology
- also depends on what happens when a node fails
  - e.g. in power grid load is redistributed