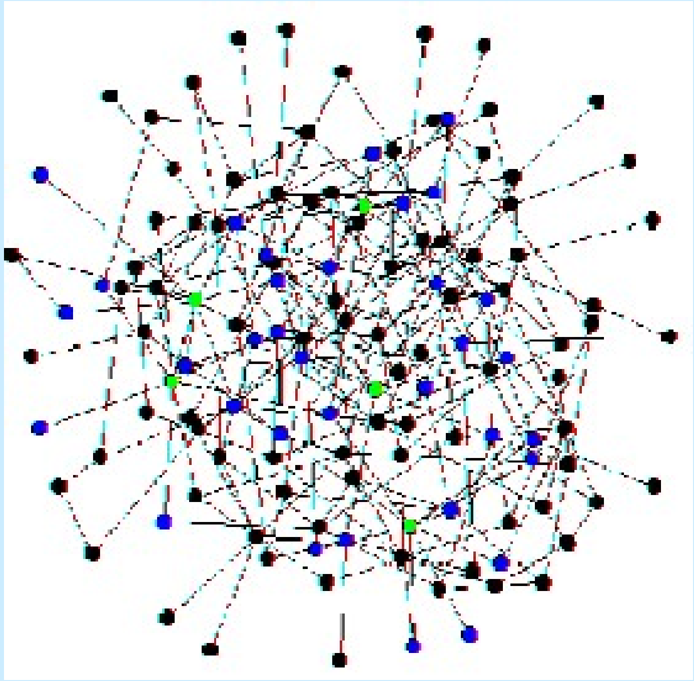
Complex Networks

- Network is a structure of N nodes and 2M links (or M edges)
- Called also graph – in Mathematics
- Many examples of networks



Internet: nodes represent computers

links the connecting cables

Social network: nodes represent people

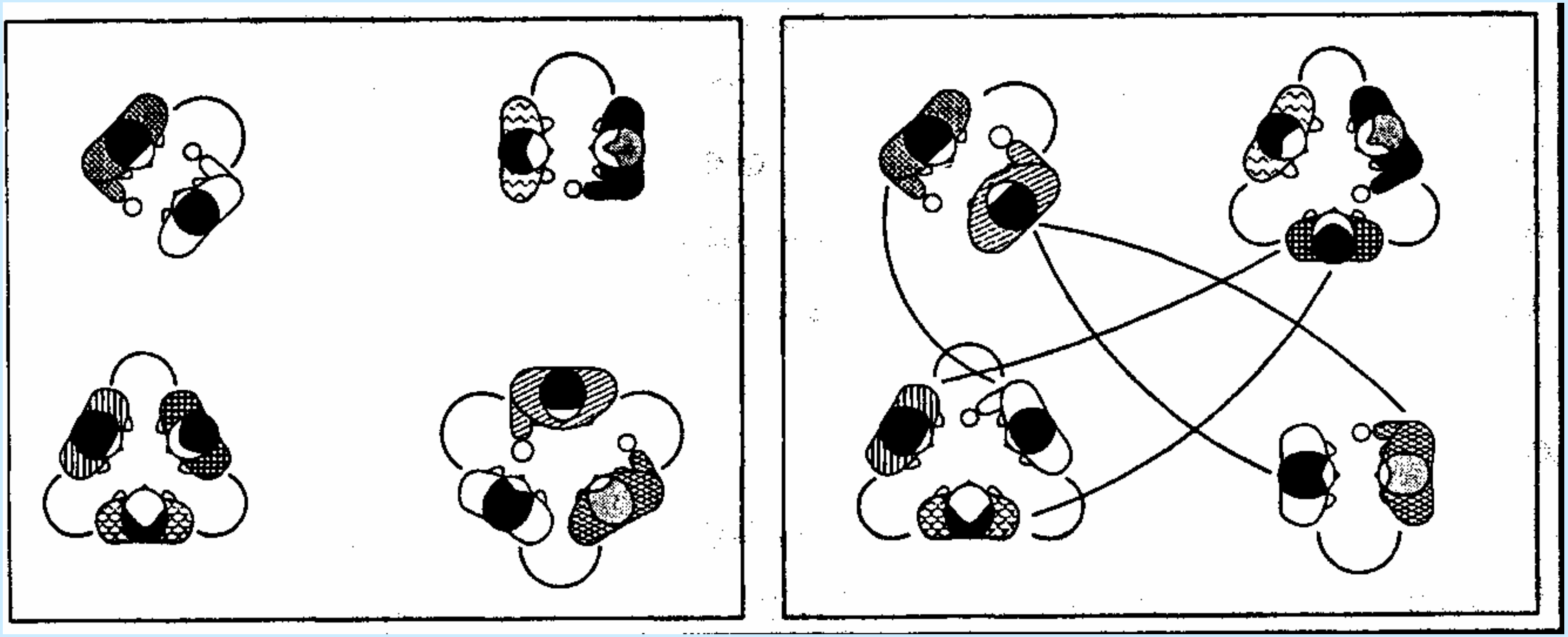
links their relations

Cellular network: nodes represent molecules

links their interactions

- Weighted networks each link has a weight determining the strength or cost of the link

Social Networks- Stanley Milgram (1967)

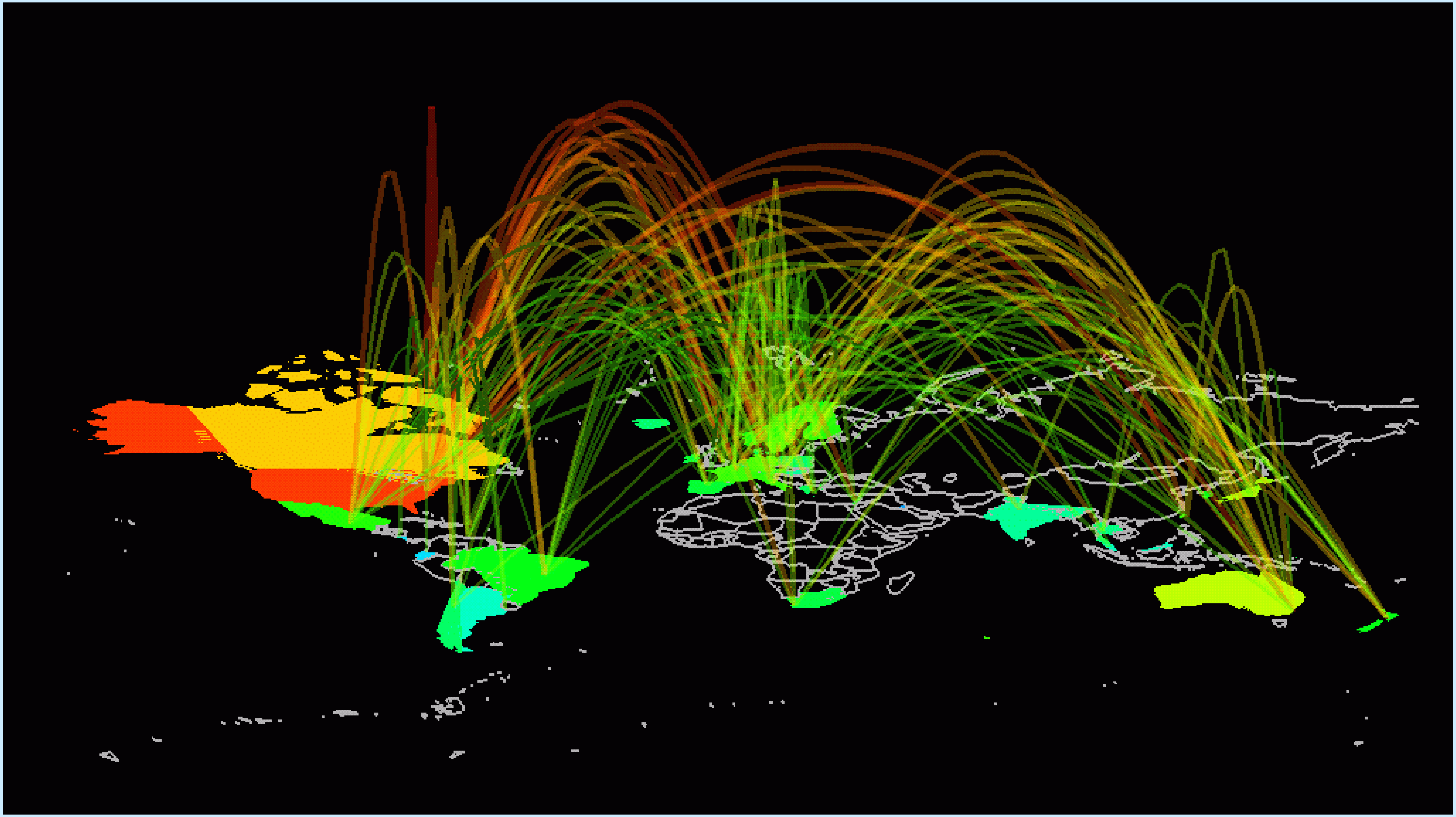


Nodes: individuals

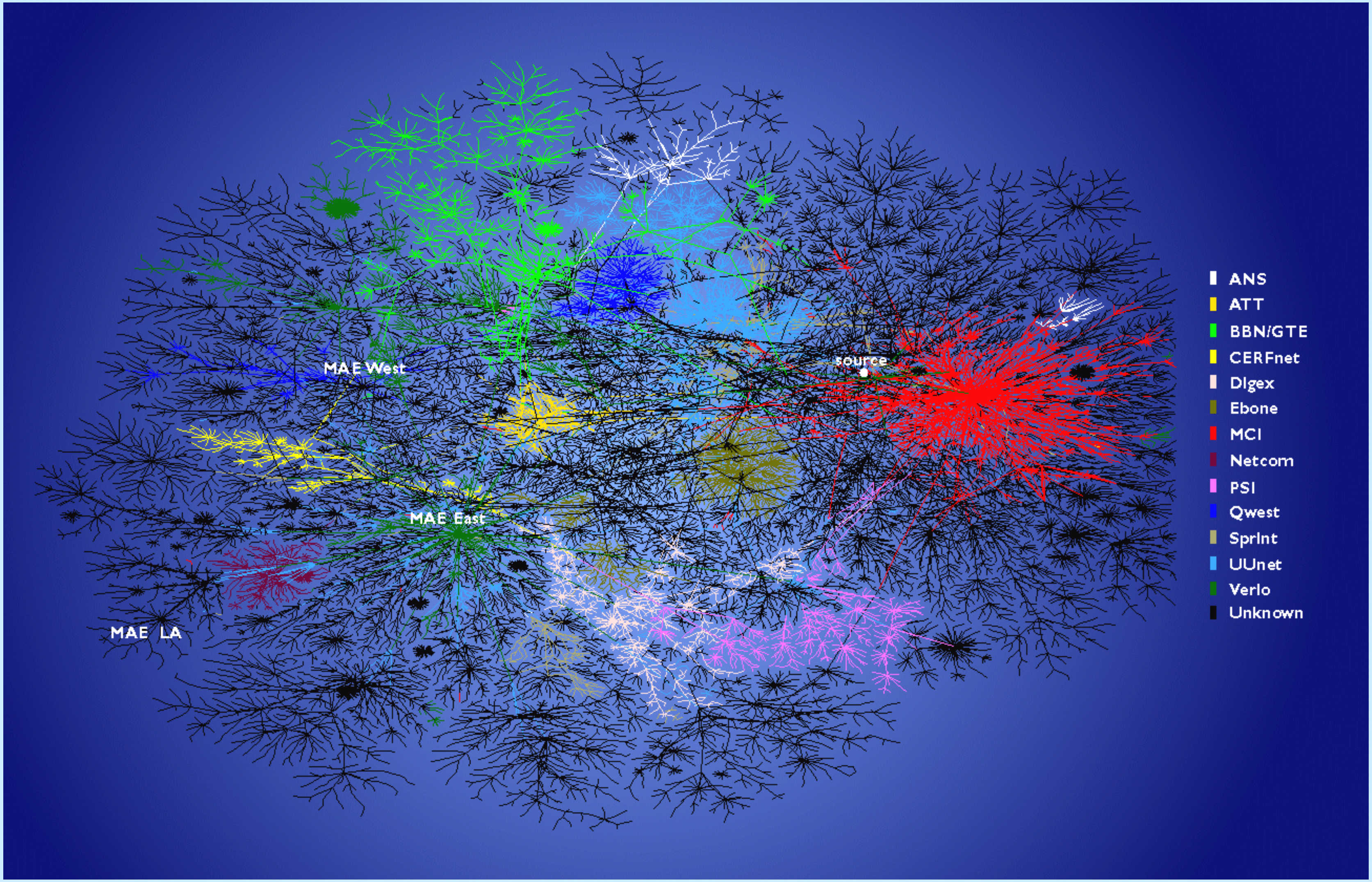
Links: social relationship (family/work/friendship/etc.)

John Guare (1992)

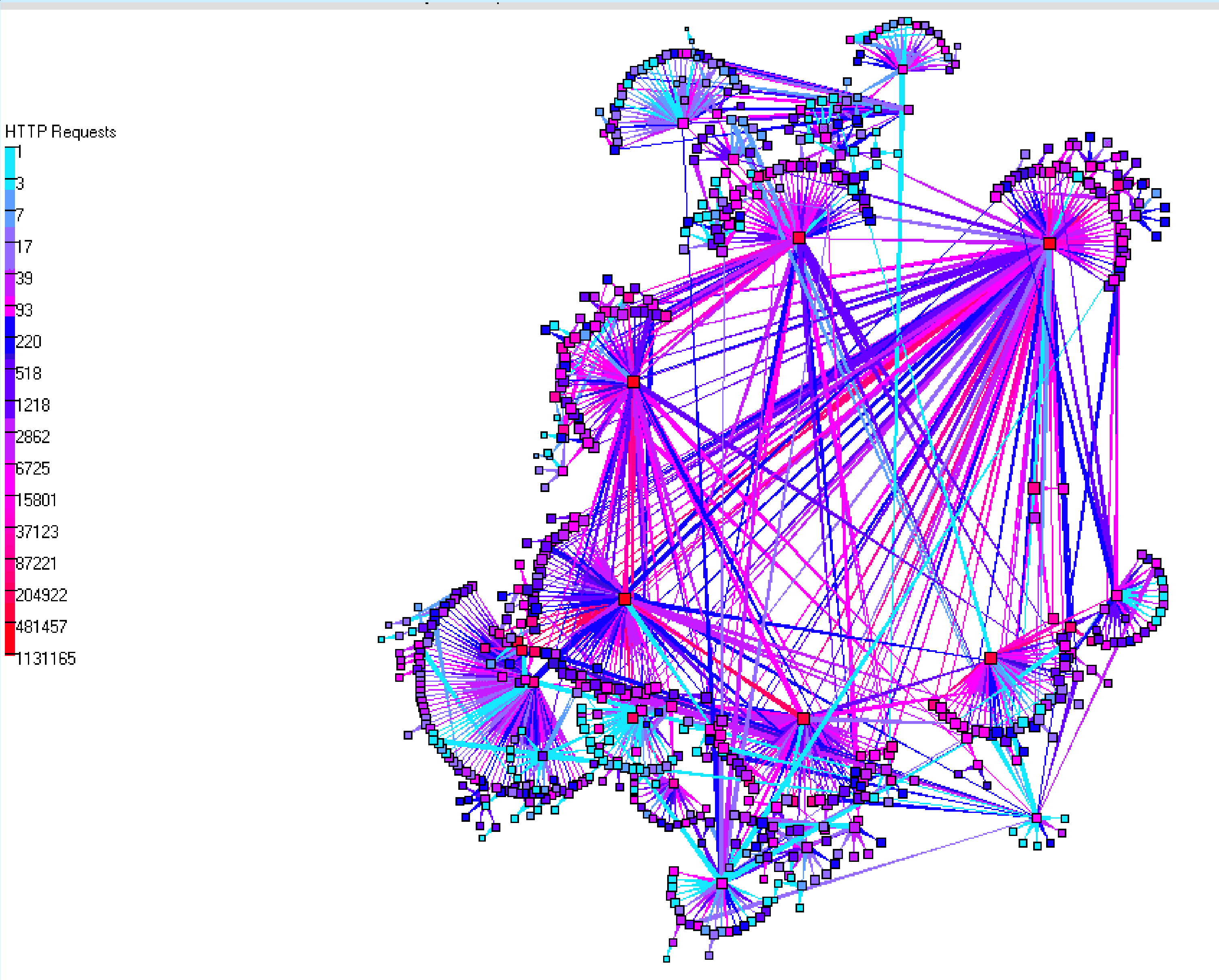
Six Degrees of Separation



Map showing the world-wide internet traffic



Skitter data depicting a macroscopic snapshot of Internet connectivity, with selected backbone **ISP**s (Internet Service Provider) colored separately



Hierarchical topology of the international web cache

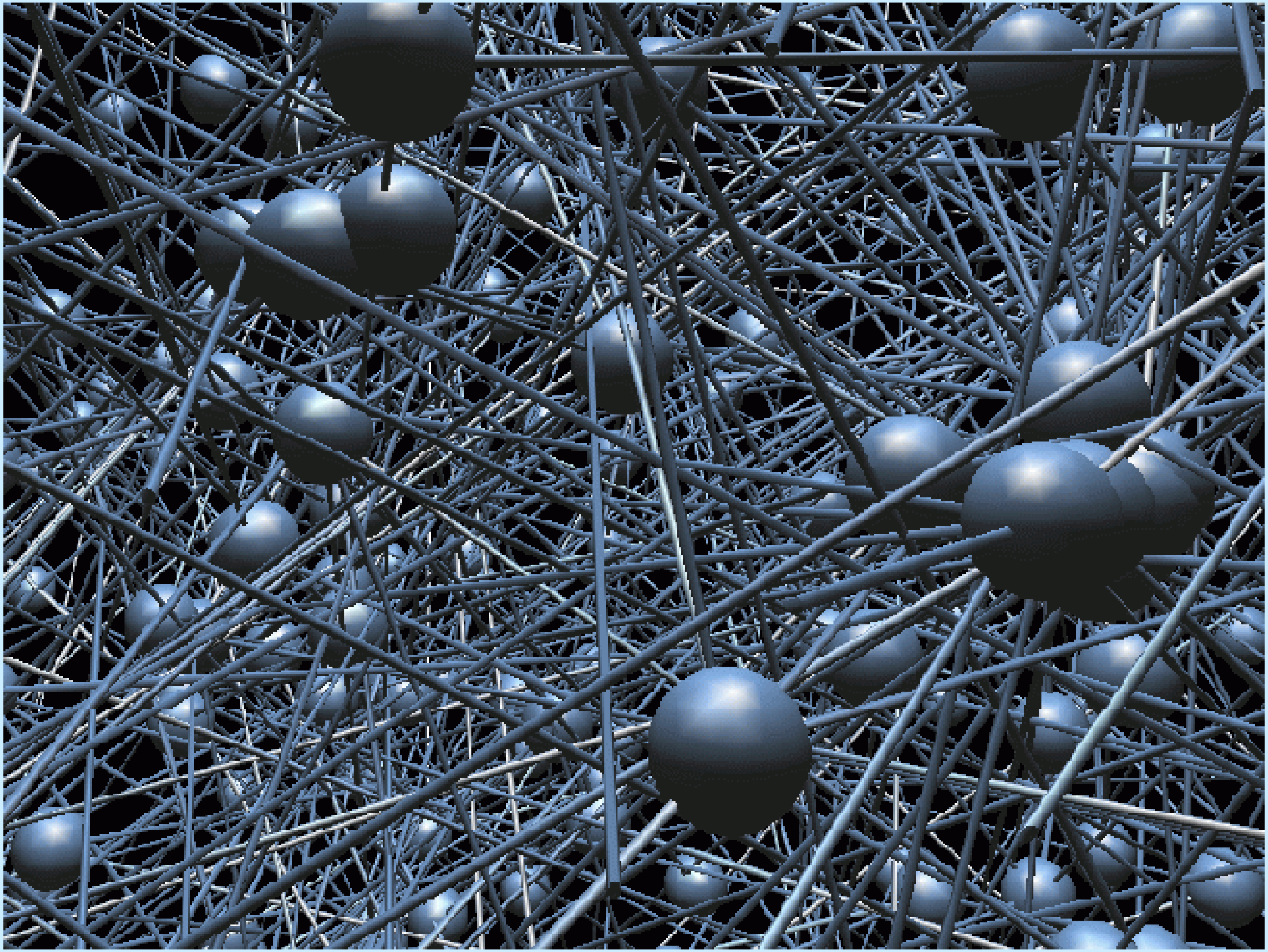
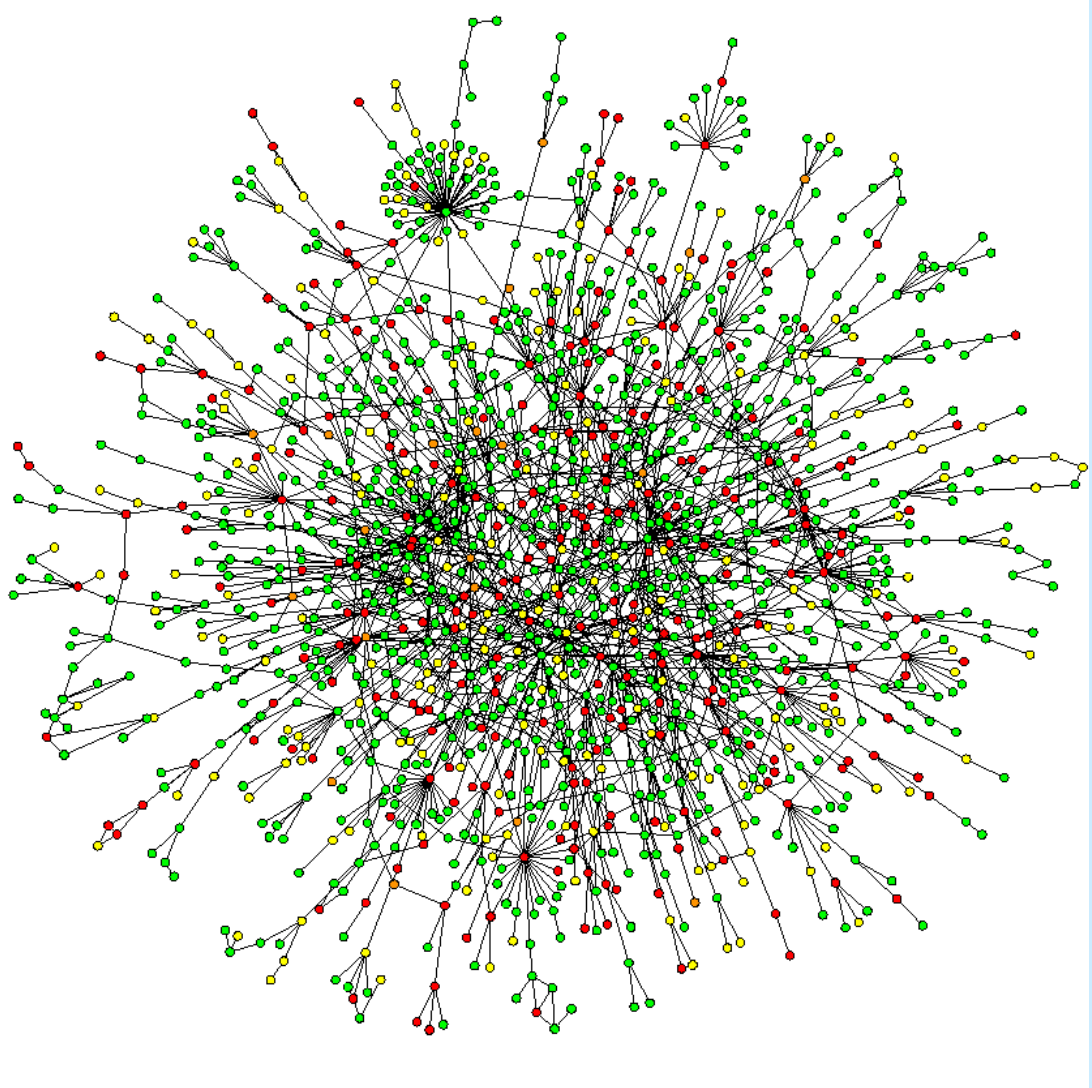
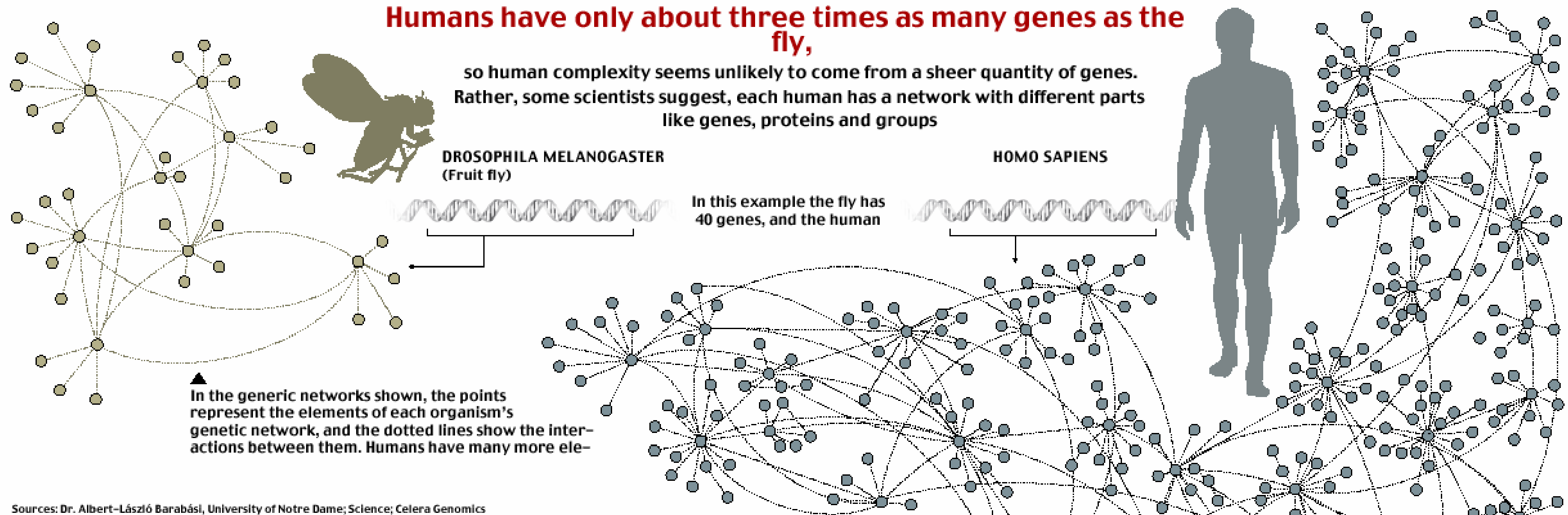
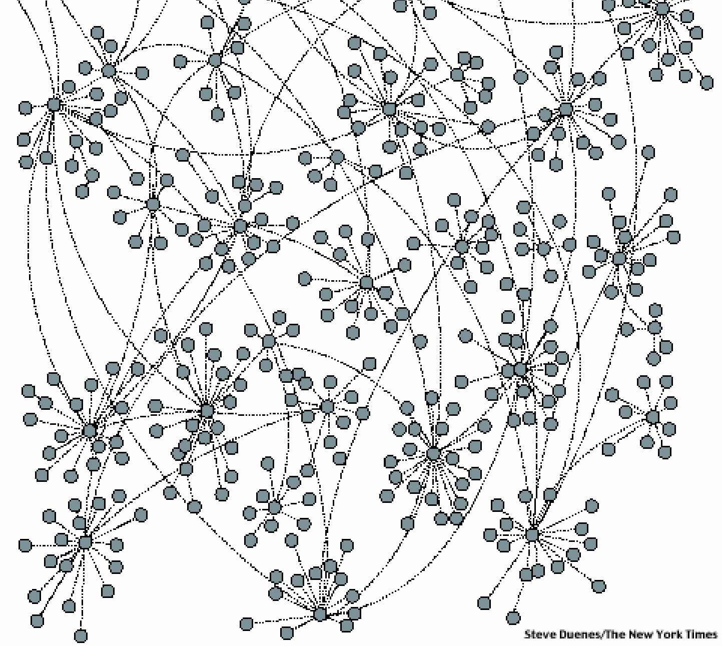


Image of Social links in Canberra, Australia



Network of protein-protein interactions. The color of a node signifies the phenotypic effect of removing the corresponding protein (red, lethal; green, non-lethal; orange, slow growth; yellow, unknown).



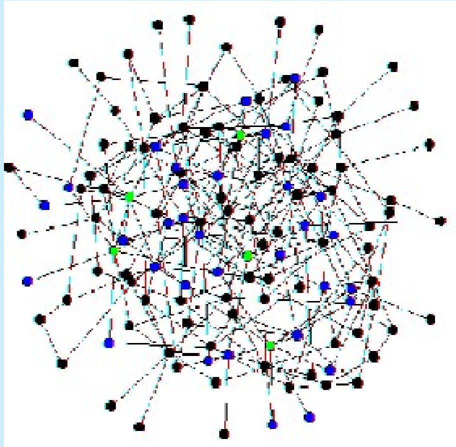


Complex systems

Made of
many non-identical **elements**
connected by diverse **interactions**.

NETWORK

Network Properties



- ❖ Degree distribution P(k) -- k- degree of a node
- ❖ Diameter or distance – Average distance between nodes--d
- ❖ Clustering Coefficient $c(k)=\frac{\text{no. of links between k neighbors}}{k(k-1)/2}$

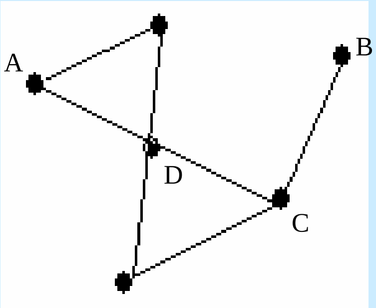
  How many of my friends are also friends?
- ❖ Centrality or Betweeneess -- b

  Number of times a bond or a node is relatively used for the shortest path
- ❖ Critical Threshold: The concentration of nodes that are removed and the network collapses

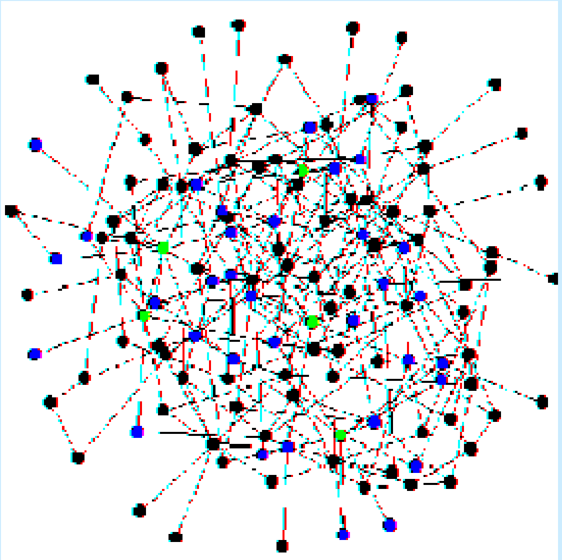
$$d(AB)=3$$

$$c(D)=\frac{2}{6}=\frac{1}{3}$$

$$b(BC)=\frac{N-1}{N-1}=1$$



Random Graph Theory



- **Developed in the 1960's by Erdos and Renyi.** **(Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 1960).**

- **Discusses the ensemble of graphs with N vertices and M edges (2M links).**

- **Distribution of connectivity per vertex is Poissonian (exponential), where k is the number of links :**

$$P(k) = e^{-c}\frac{c^k}{k!}, \qquad c = \langle k \rangle = \frac{2M}{N}$$

- **Distance d=log N -- SMALL WORLD**

More Results

- **Phase transition at average connectivity, $\langle k \rangle = 1$:**

  $\langle k \rangle < 1$ **No spanning cluster (giant component) of order** *logN*

  $\langle k \rangle > 1$ **A spanning cluster exists (unique) of order** N

  $\langle k \rangle = 1$ **The largest cluster is of order** $N^{2/3}$

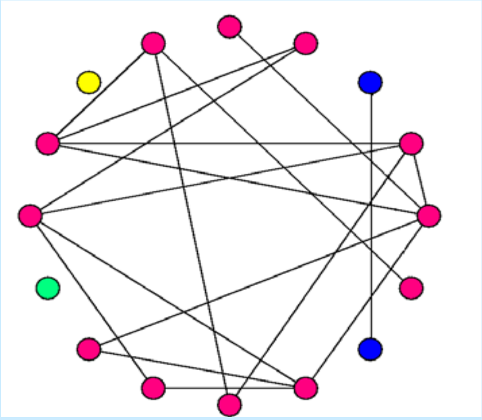
- **Size of the spanning cluster is determined by the self-consistent equation:**

  $$P_\infty = 1 - e^{-\langle k \rangle P_\infty}$$

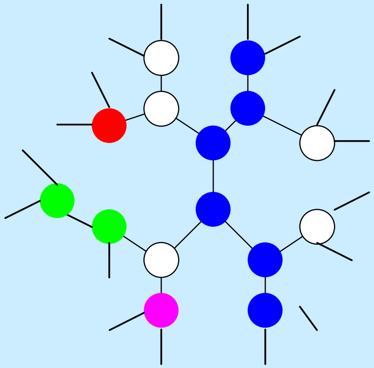
- **Behavior of the spanning cluster size near the transition is linear:**

  $P_\infty \propto (p_c - p)^\beta$, $\beta = 1$, **where** p **is the probability of deleting a site,**

  $p_c = 1 - 1/\langle k \rangle$



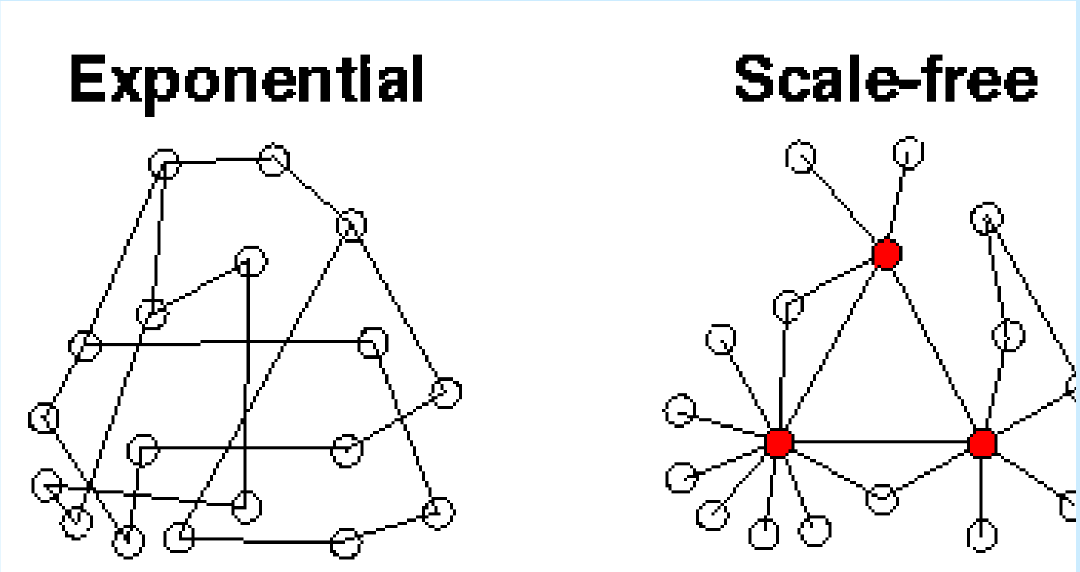
Percolation on a Cayley Tree



- **Contains no loops**
- **Connectivity of each node is fixed (*z* connections)**
- **Critical threshold:** $$p_c = \frac{1}{z-1}$$
- **Behavior of the spanning cluster size near the transition is linear:**

$$P_\infty \propto (p_c - p)^\beta, \quad \beta = 1$$

In Real World - Many Networks are non-Poissonian

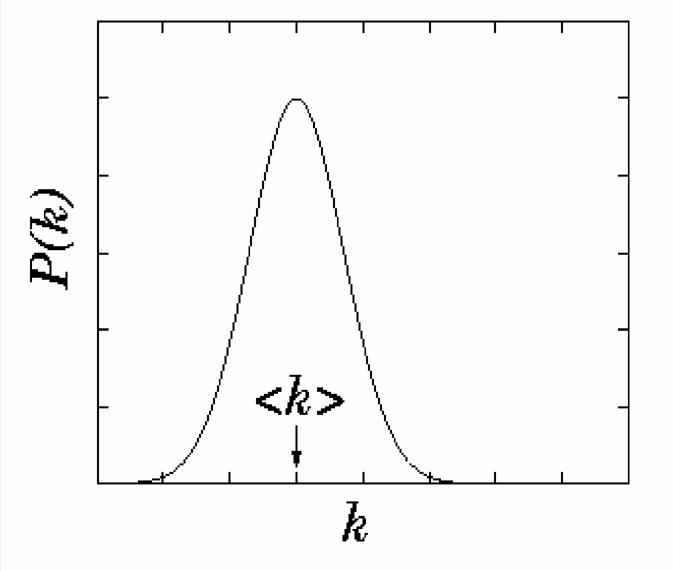


$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

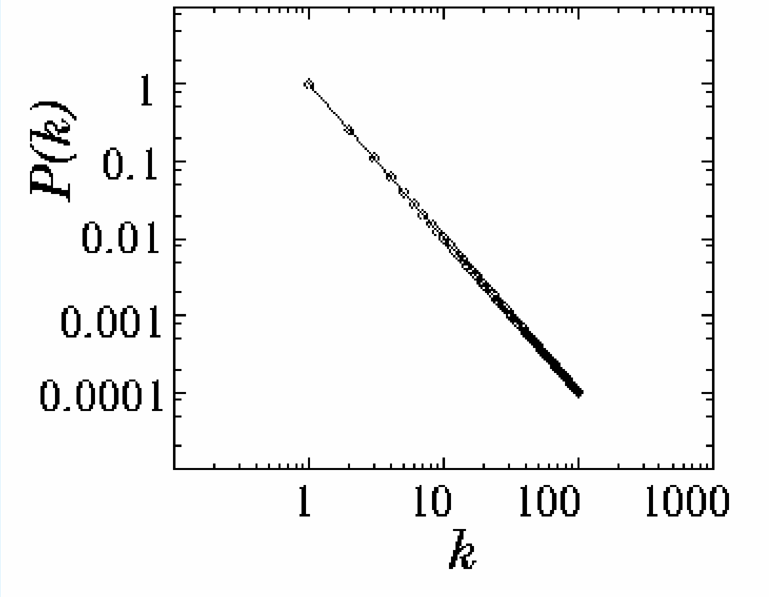
$$P(k) = \begin{cases} ck^{-\lambda} & m \le k \le K \\ 0 & otherwise \end{cases}$$

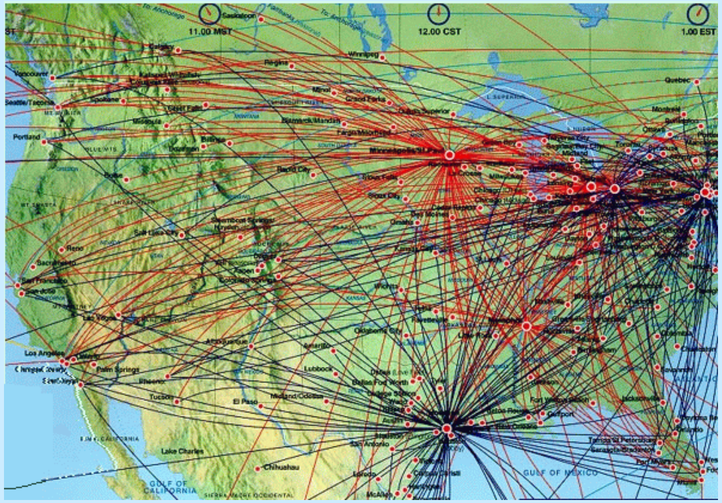
New Type of Networks

Poisson distribution



Power-law distribution

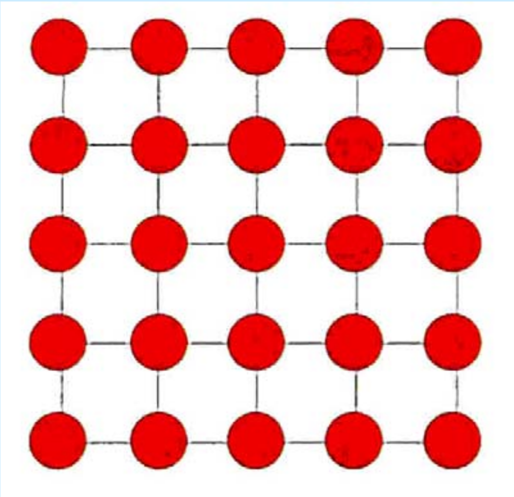


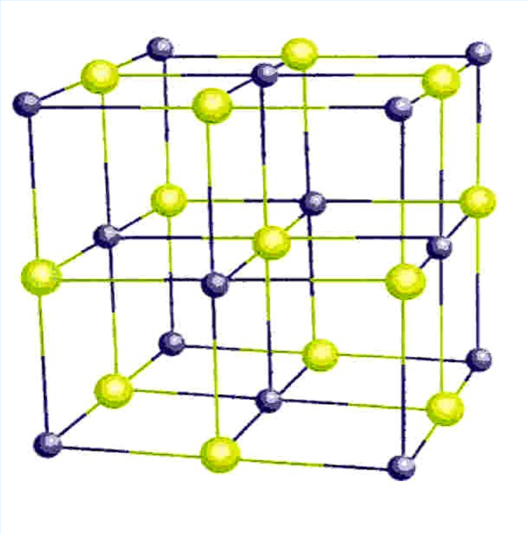


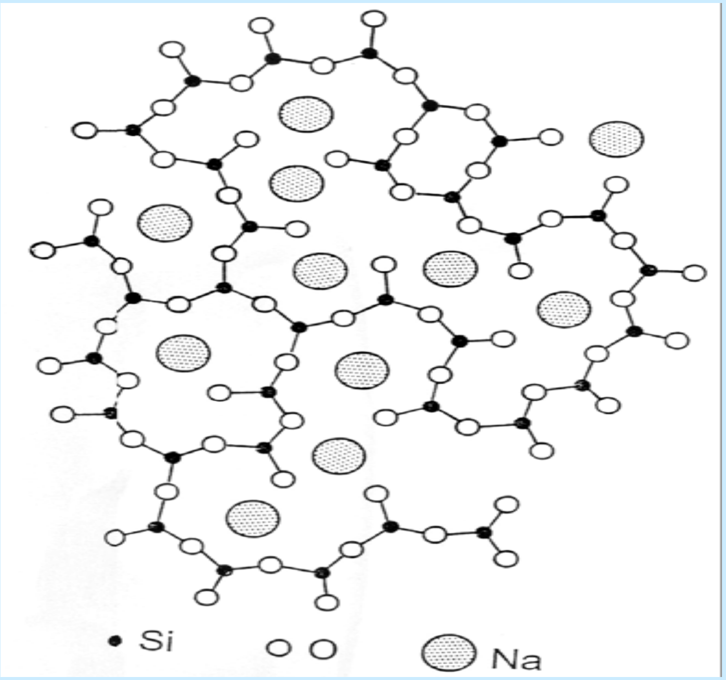
Exponential Network

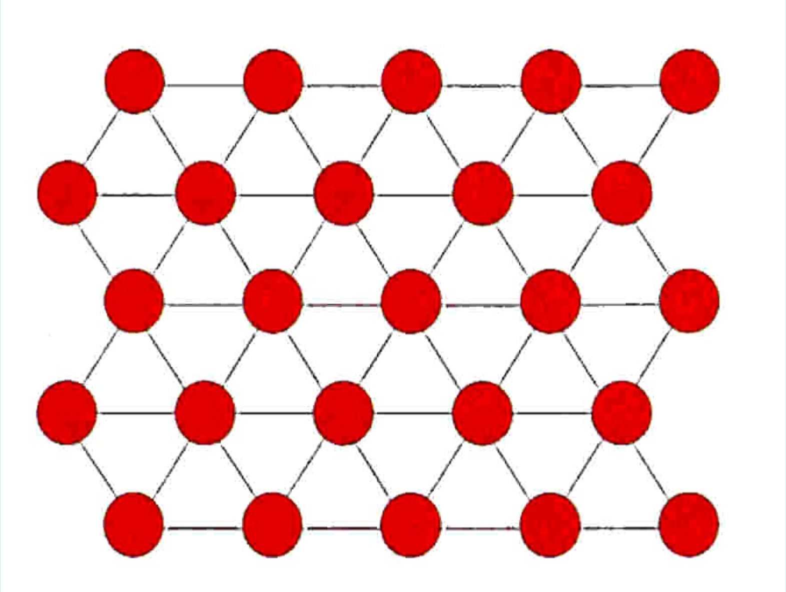
Scale-free Network

Networks in Physics



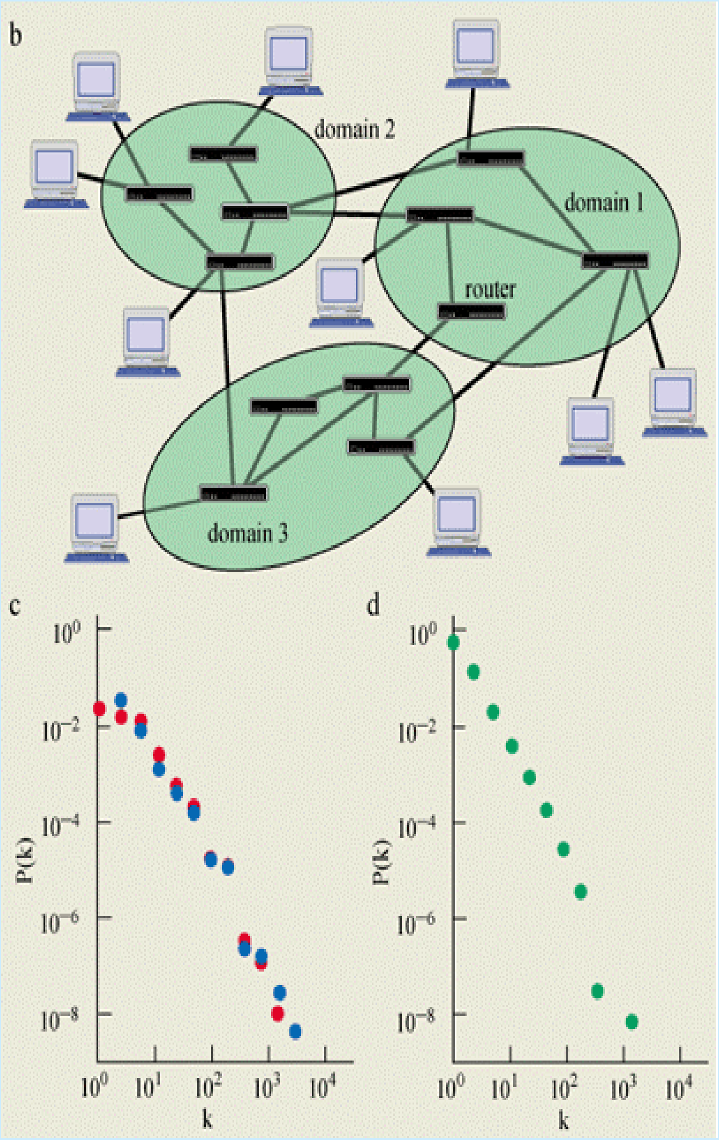


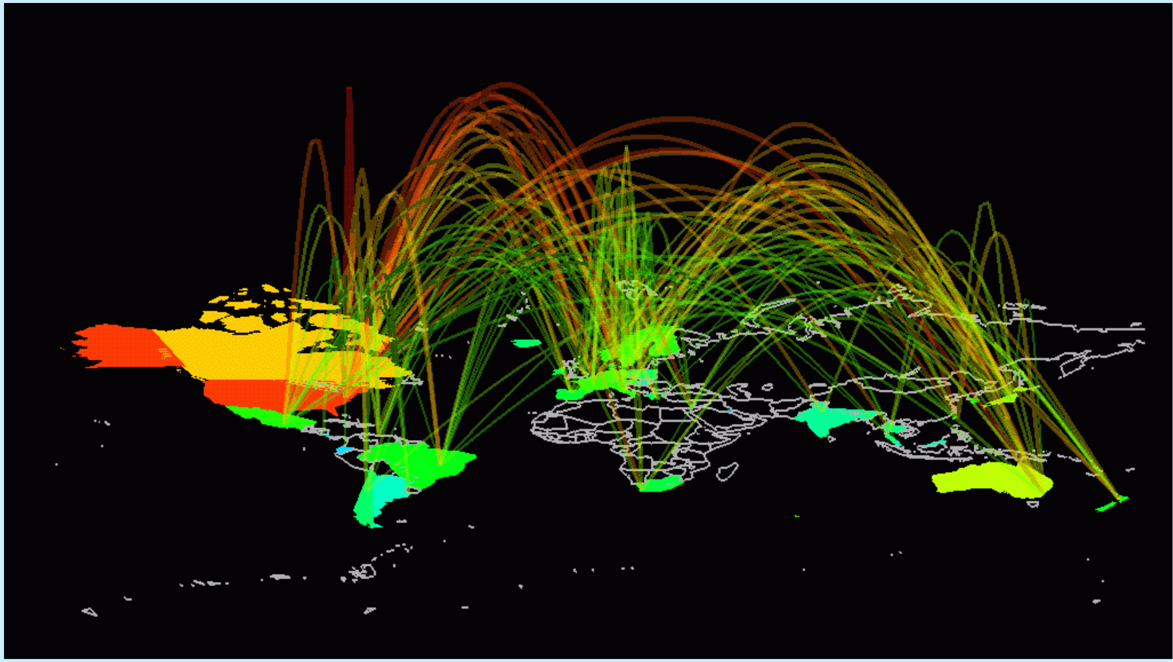


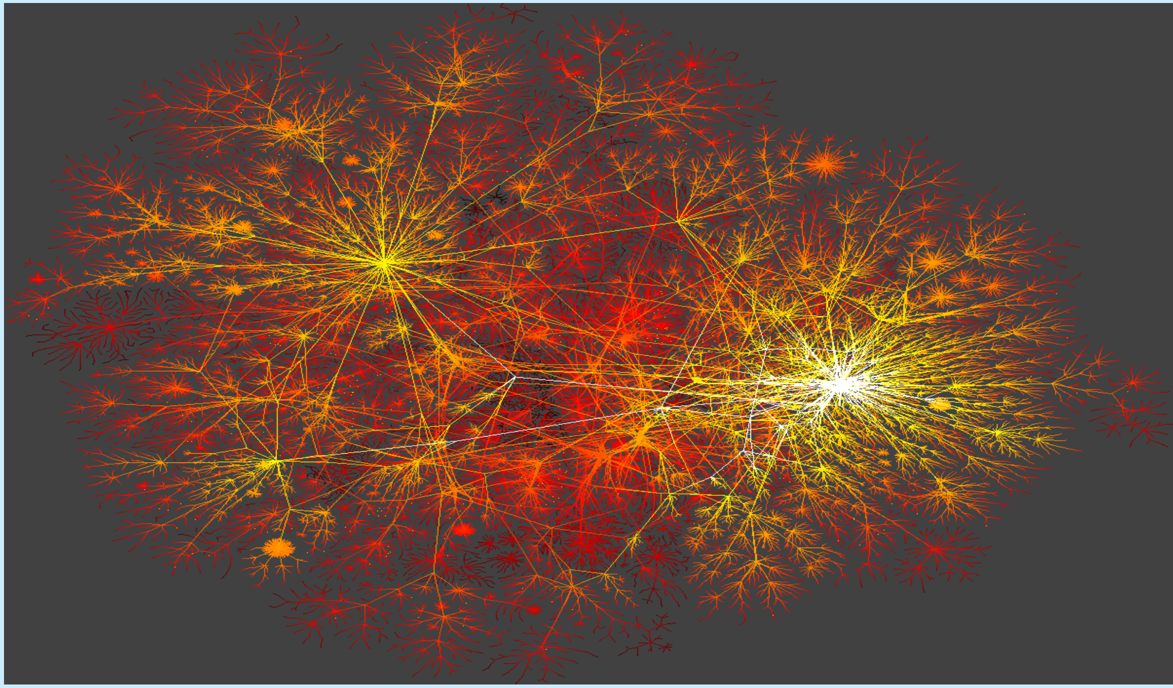


Internet Network

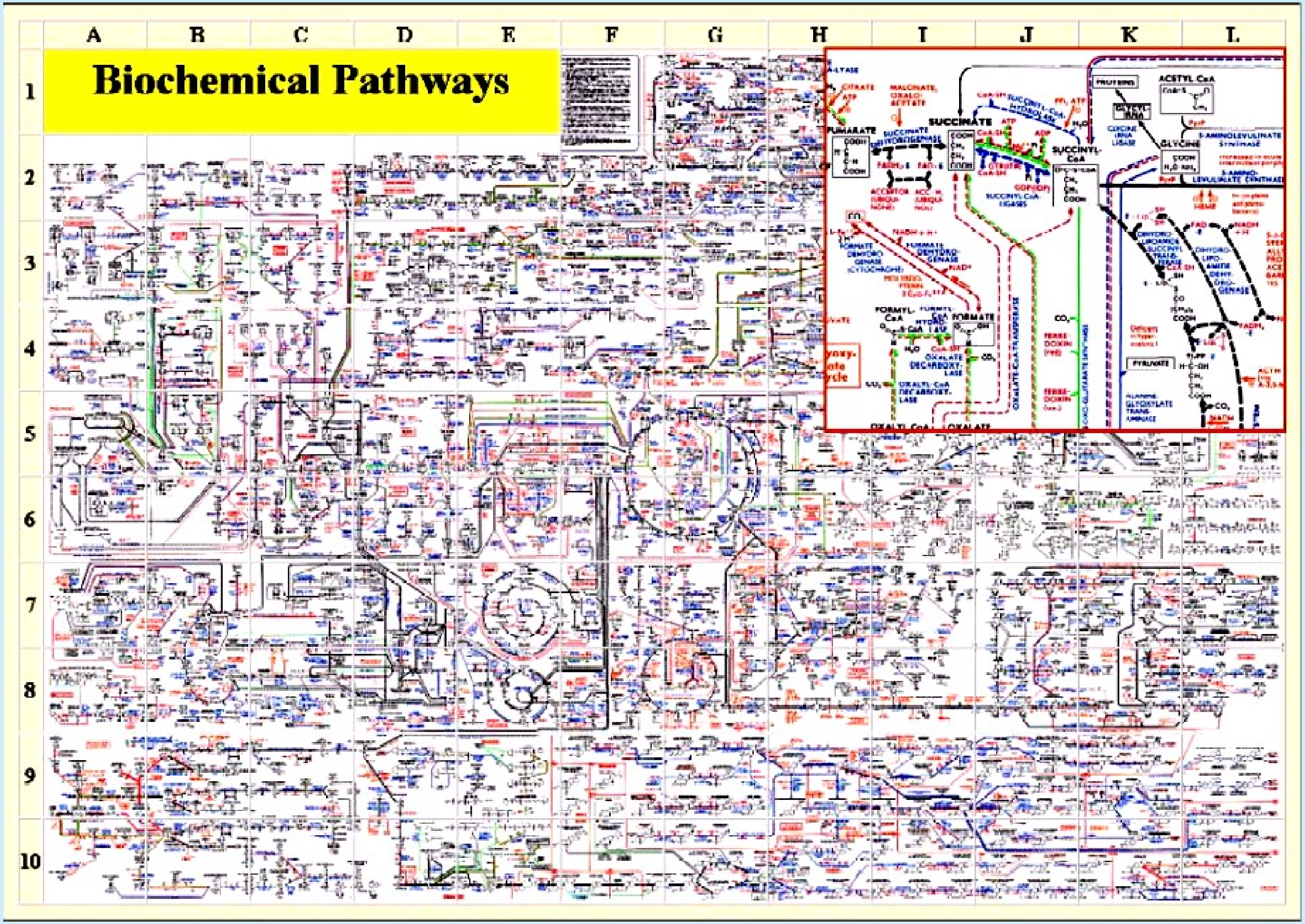
Faloutsos et. al., SIGCOMM '99







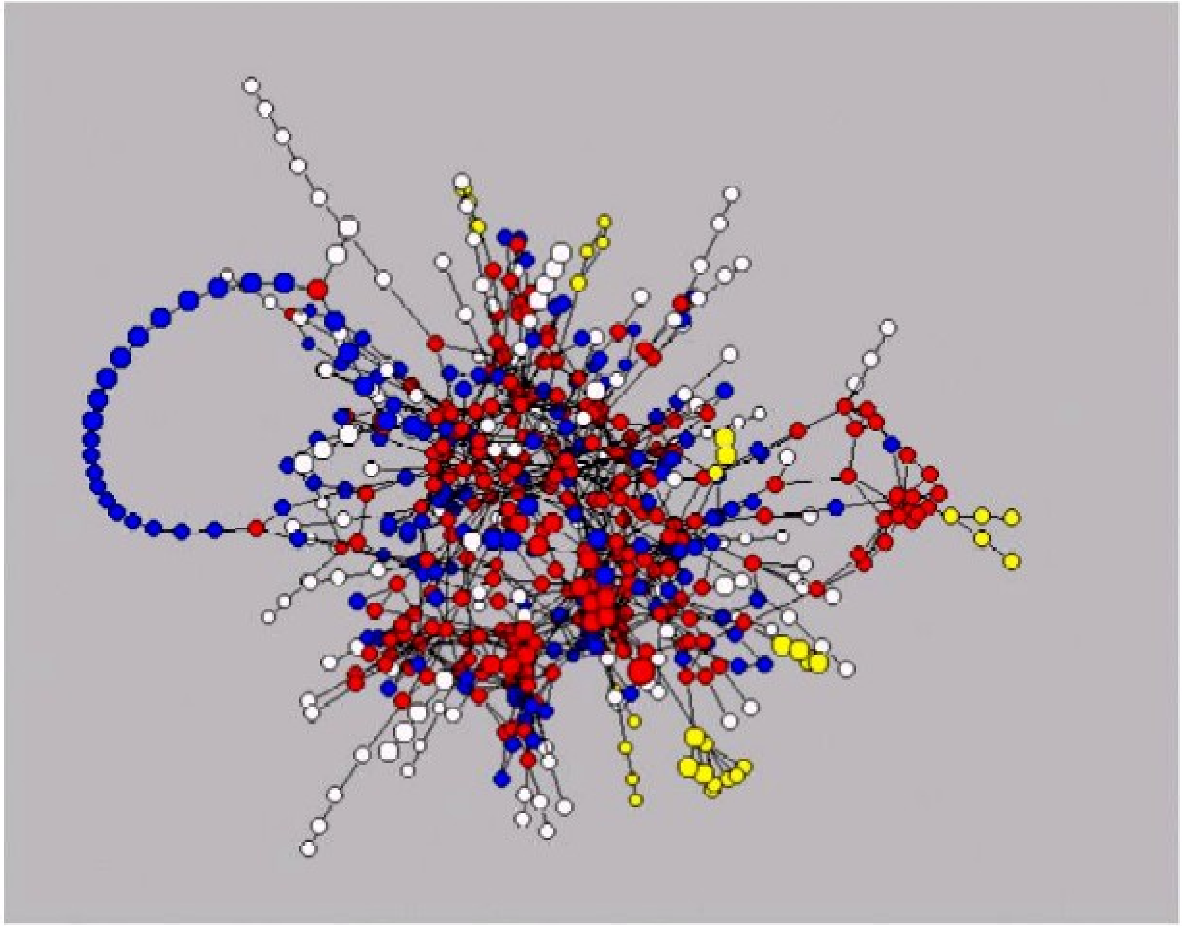
Metabolic network



Metabolic Network

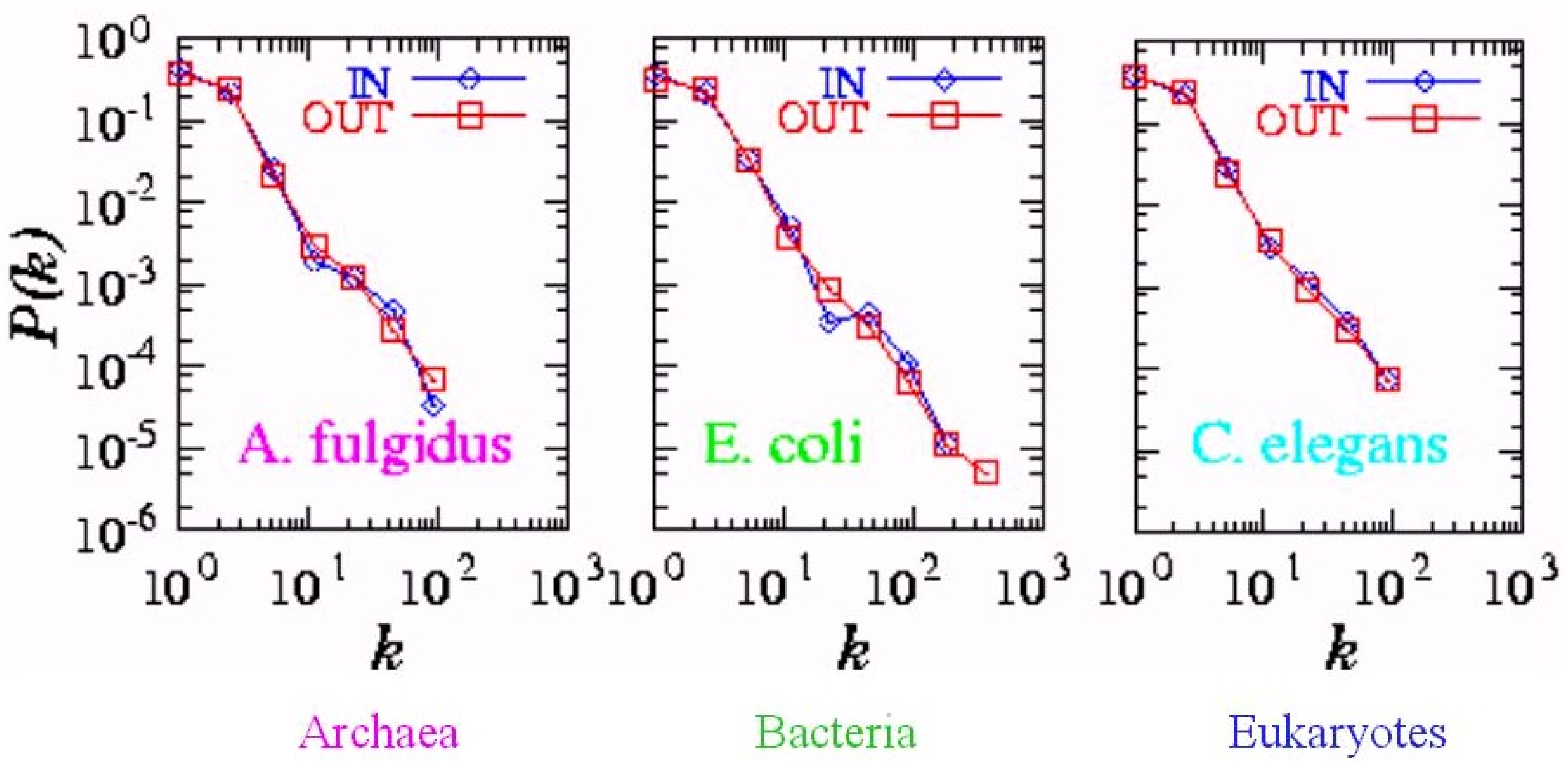
Nodes: chemicals (substrates)

Links: bio-chemical reactions



Jeong et all Nature 2000

Metabolic network



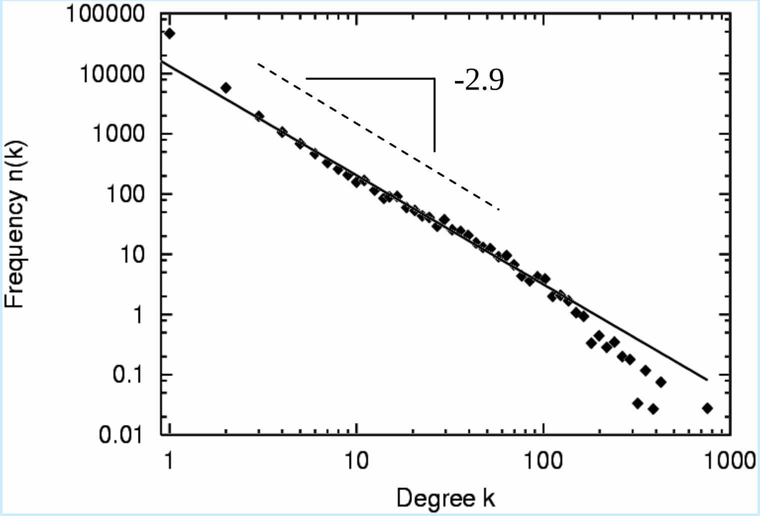
Organisms from all three domains of life are **scale-free** networks!

Jeong et al, Nature (2000)

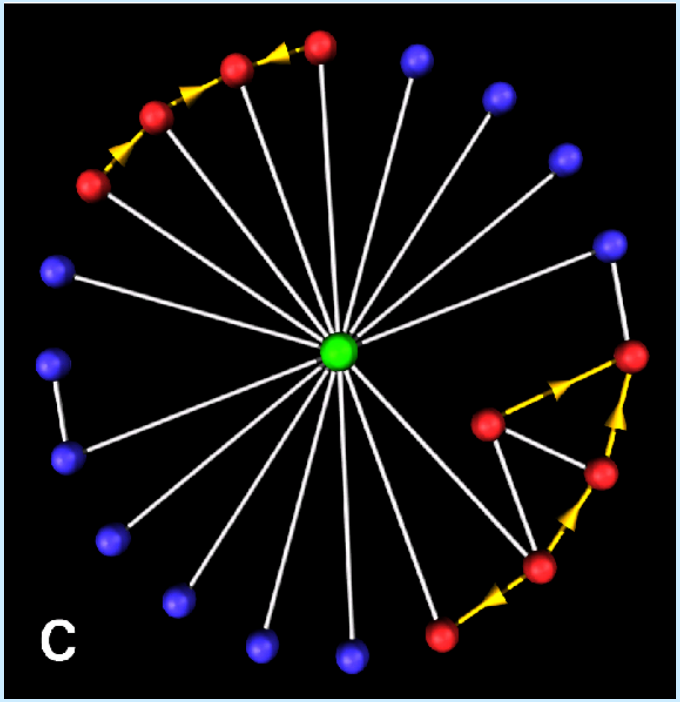
More Examples

- Trust networks: Guardiola et al (2002)
- Email networks: Ebel etal PRE (2002)

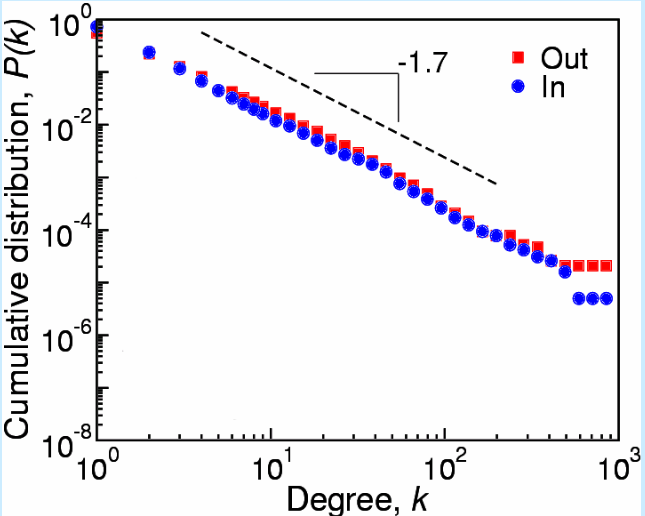
Email



Trust



Trust



Erdös Theory is Not Valid

Stability and Immunization

Theory

$$p_c = 1 - \frac{1}{\langle k \rangle}$$

Critical concentration 30-50%

Experiment

Infectious disease	Critical concentration
Malaria	99%
Measles	90-95%
Whooping cough	90-95%
Fifths disease	90-95%
Chicken pox	85-90%
Mumps	85-90%
Rubella	82-87%
Poliomyelitis	82-87%
Diphtheria	82-87%
Scarlet fever	82-87%
Smallpox	70-80%
INTERNET	99%

Distance

Theory

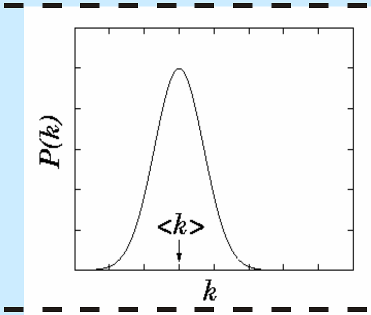
$$\boldsymbol{d} \sim \log \boldsymbol{N}$$

Experiment

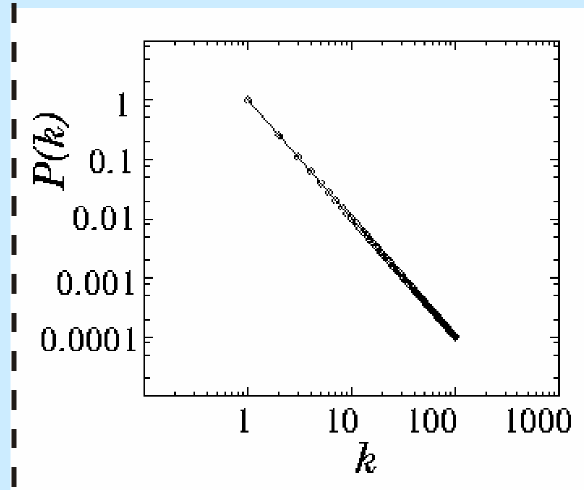
Almost constant
(Metabolic Networks,
Jeong et. al.
(Nature, 2000))

Distribution

Theory



Experiment



Generalization of Erdös Theory:
Cohen, Erez, ben-Avraham, Havlin, PRL **85**, 4626 (2000)
Epidemiology Theory: Vespignani, Pastor-Satoral, PRL (2001), PRE (2001)

Cohen, Havlin,
Phys. Rev. Lett. 90, 58701(2003)

Modelling: Albert, Jeong, Barabasi (Nature 2000)

Experimental Data: Virus survival

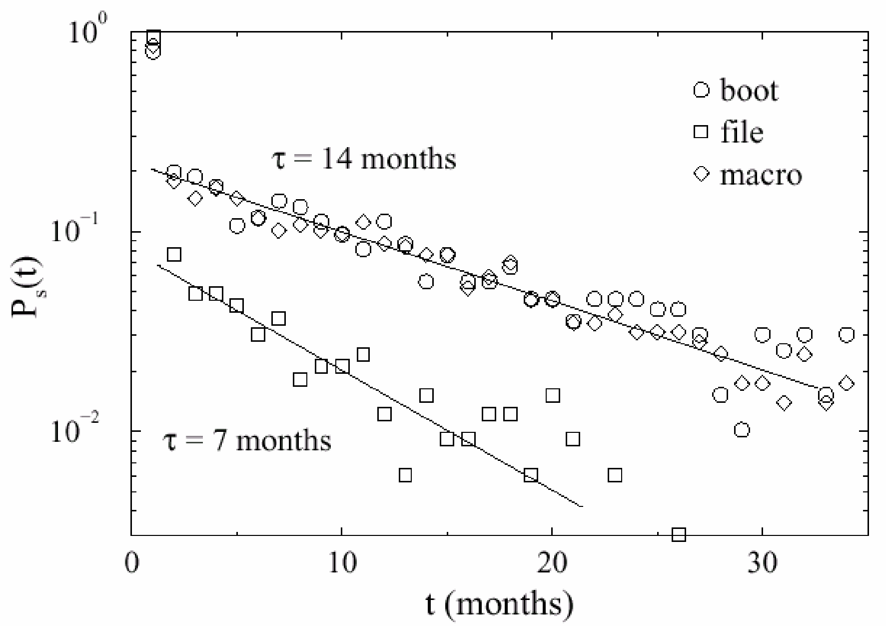
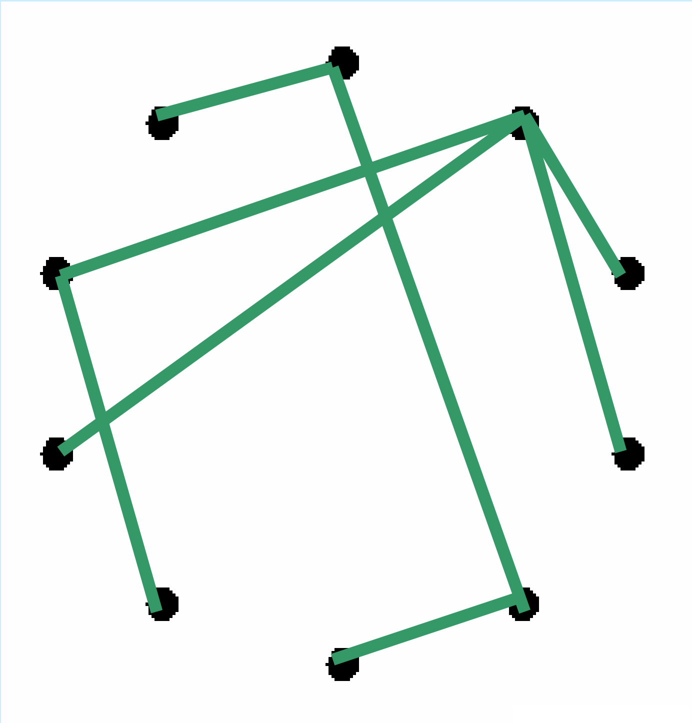
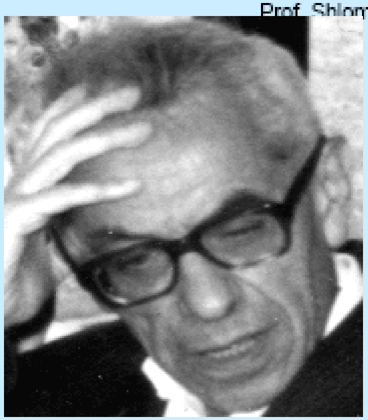


FIG. 1. Surviving probability for viruses in the wild. The 814 different viruses analyzed have been grouped in three main strains [9]: file viruses infect a computer when running an infected application; boot viruses also spread via infected applications, but copy themselves into the boot sector of the hard-drive and are thus immune to a computer reboot; macro viruses infect data files and are thus platform-independent. It is evident in the plot the presence of an exponential decay, with characteristic time $\tau \simeq 14$ months for macro and boot viruses and $\tau \simeq 7$ months for file viruses.

(Pastor-Satorras and Vespignani, Phys. Rev Lett. 86, 3200 (2001))

Erdös-Rényi model
(1960)

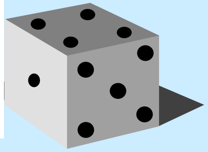




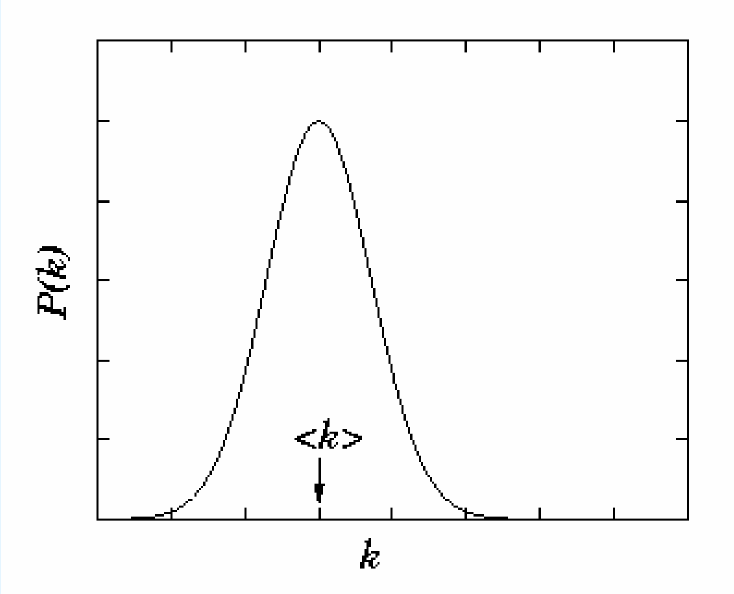
Pál Erdös
(1913-1996)

Connect with probability p

p=1/6
N=10
$\langle k \rangle \sim 1.5$



Poisson distribution



- **Democratic**
- **Random**

Scale-free model

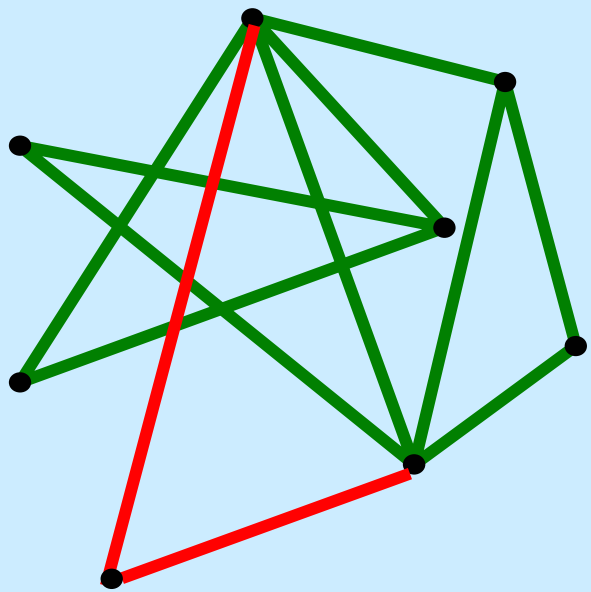
(1) GROWTH :

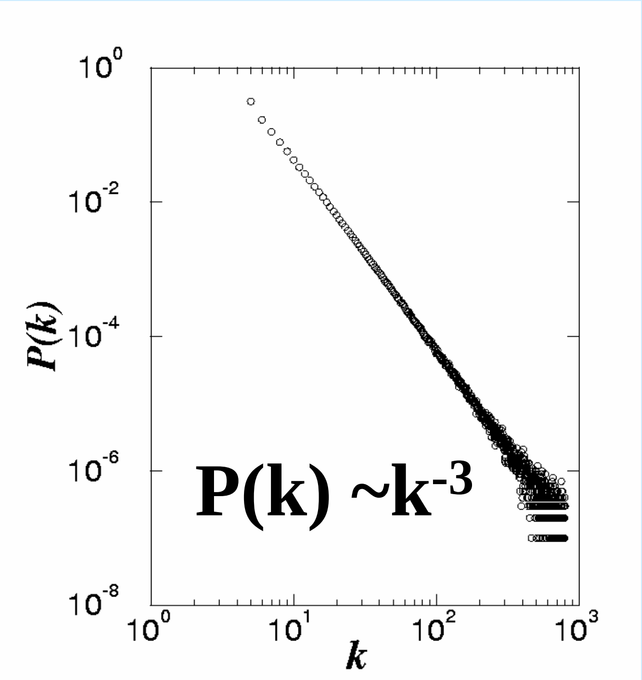
At every time step we add a new node with *m* edges (connected to the nodes already present in the system).

(2) PREFERENTIAL ATTACHMENT :

The probability **Π** that a new node will be connected to node *i* depends on the connectivity k_i of that node

$$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j}$$





A.-L.Barabási, R. Albert, Science **286,** 509 (1999)

Shortest Paths in Scale Free Networks

$$P(k) \sim k^{-\lambda}$$

$$d = const. \qquad \lambda = 2$$

Ultra Small World

$$d = \log\log N \qquad 2 < \lambda < 3$$

$$d = \frac{\log N}{\log\log N} \qquad \lambda = 3$$

(Bollobas, Riordan, 2002)

Small World

$$d = \log N \qquad \lambda > 3$$

(Bollobas, 1985)
(Newman, 2001)

Cohen, Havlin Phys. Rev. Lett. 90, 58701(2003)

Cohen, Havlin and ben-Avraham, in Handbook of Graphs and Networks eds. Bornholdt and Shuster (Willy-VCH, NY, 2002) chap.4

Confirmed also by: Dorogovtsev et al (2002), Chung and Lu (2002)

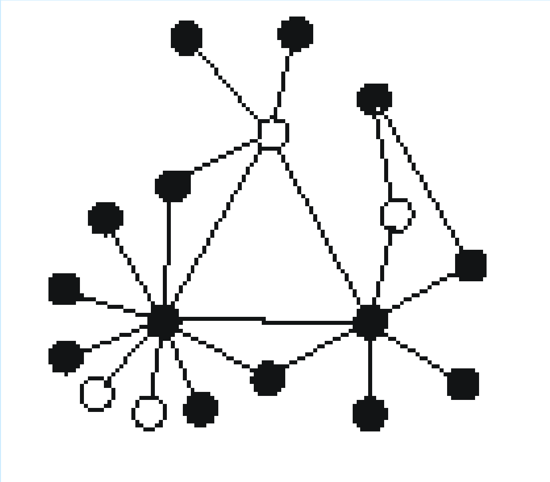
Model of Stability

Random Breakdown (Immune)

The Internet is believed to be a randomly connected scale-free network, where

$$P(k) \propto k^{-\lambda}, \lambda \approx 2.5$$

Nodes are randomly removed (or immune) with probability p

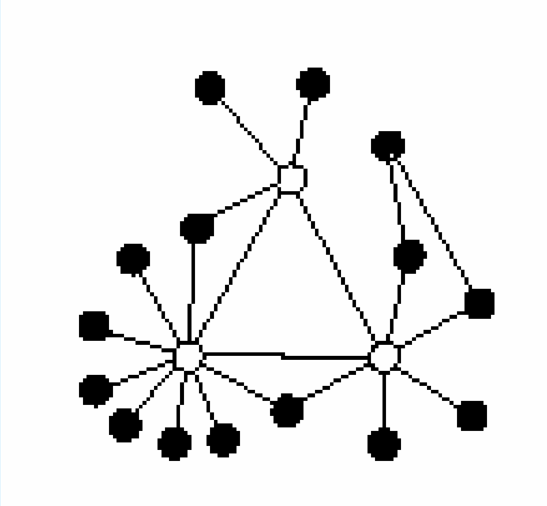


Where does the phase transition occur ?

Model for Stability

Intentional Attack (Immune)

The fraction, p, of nodes with the highest connectivity are removed (or immune).



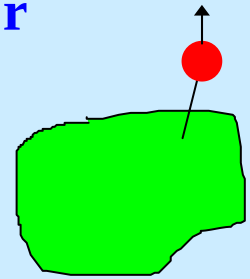
Is this fundamentally different from random breakdown?

We find that not only critical thresholds but also critical exponents are different !

THE UNIVERSALITY CLASS DEPENDS ON THE WAY CRITICALITY REACHED

THEORY FOR ANY DEGREE DISTRIBUTION

Condition for the Existence of a Spanning Cluster



If we start moving on the cluster from a single site, in order that the cluster does not die out, we need that each site reached will have, on average, at least 2 links (one "in" and one "out").

This means: $\langle k_i | i \leftrightarrow j \rangle = \sum_{k_i} k_i P(k_i | i \leftrightarrow j) \geq 2$, **where** $i \leftrightarrow j$ **means that site** i **is connected to site** j.

But, by Bayes rule: $P(k_i | i \leftrightarrow j) = \dfrac{P(i \leftrightarrow j | k_i) P(k_i)}{P(i \leftrightarrow j)}$

We know that $P(i \leftrightarrow j | k_i) = \dfrac{k_i}{N-1}$ **and** $P(i \leftrightarrow j) = \dfrac{\langle k \rangle}{N-1}$

Combining all this together: $\kappa \equiv \dfrac{\langle k^2 \rangle}{\langle k \rangle} = 2$

(for every distribution) at the critical point.

Exponential graph:

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

$$\Rightarrow \langle k \rangle = 1$$

Cayley Tree:

$$p_c = \frac{1}{z-1}$$

Cohen et al, PRL 85, 4626 (2000): PRL 86, 3862 (2001)

Percolation for Random Breakdown

If percolation is considered the connectivity distribution changes according to the law: $\bar{P}(k) = \sum_{k'>k} P(k') \binom{k'}{k} p^{k'-k} (1-p)^k$

Calculating the condition $\kappa = 2$ gives the percolation threshold:

$1 - p_c = \frac{1}{\kappa_0 - 1}$, **where** $\kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle}$. **compared to** $p_c = 1 - 1/\langle k_0 \rangle$ **for Erdos Renyi**

and

$$p_c = 1 - \frac{1}{z-1} \text{ for Cayley tree}$$

For scale-free distribution with lower cutoff m, and upper cutoff K, gives

$$\kappa_0 = \left(\frac{2-\lambda}{3-\lambda}\right)\frac{K^{3-\lambda} - m^{3-\lambda}}{K^{2-\lambda} - m^{2-\lambda}}, \quad K \sim N^{\frac{1}{\lambda-1}}.$$

For scale-free graphs with $\lambda \le 3$ the second moment diverges.

No critical threshold!

Network is stable (or not immuned) even for $p \to 1$.

Percolation for Intentional Attack (Immune)

Attack has two kinds of influence on the connectivity distribution:

- **Change in the upper cutoff**

  Can be calculated by $\sum_{k=\bar{K}}^{K} P(k) = p$ **,**

  or approximately: $\bar{K} = mp^{1/(1-\lambda)}$ **.**

- **Change in the connectivity of all other sites due to possibility of a broken link (which is different than in random breakdown). The probability of a link to be removed can be calculated by:**

$$\bar{p} = \frac{1}{\langle k_0 \rangle} \sum_{k=\bar{K}}^{K} kP(k),$$

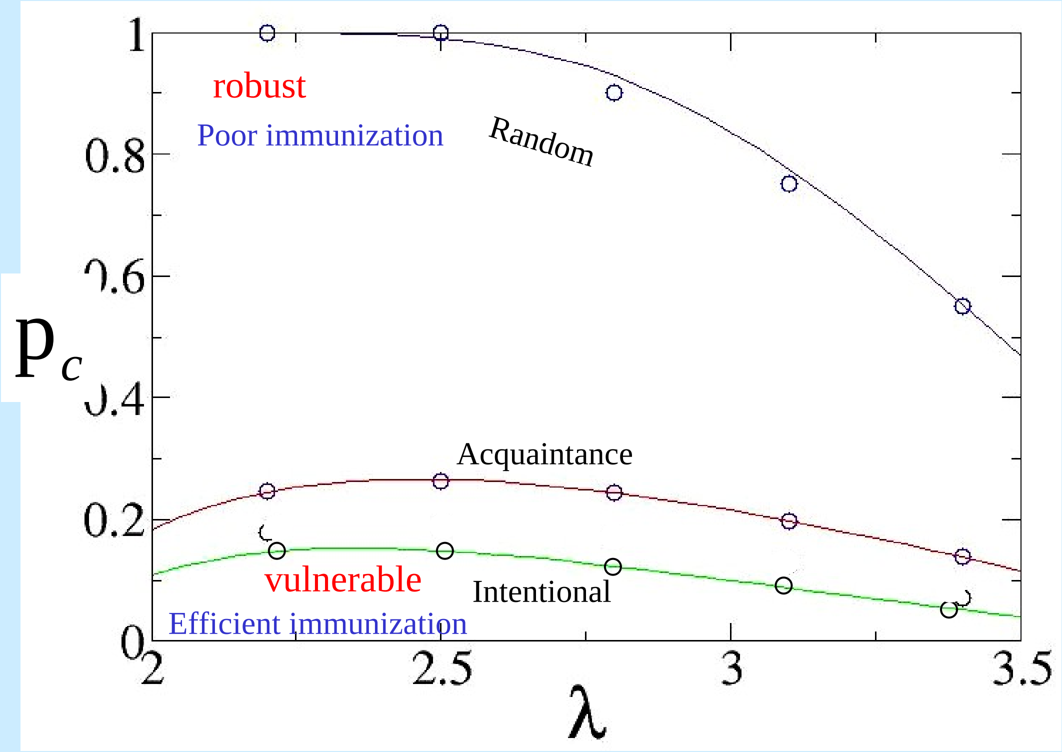
or approximately: $\bar{p} = p^{(2-\lambda)/(1-\lambda)}$ **.**

Substituting this into: $1 - \bar{p}_c = \frac{1}{\bar{k} - 1}$ **,**

where $\bar{\kappa} = \left(\frac{2-\alpha}{3-\alpha}\right) \frac{\bar{K}^{3-\lambda} - m^{3-\lambda}}{\bar{K}^{2-\lambda} - m^{2-\lambda}}$ **, gives the critical threshold.**

There exists a finite percolation threshold even for networks resilient to random error!

Critical Threshold
Scale Free



Cohen et al. Phys. Rev. Lett. 91 , 168701 (2003)

General result:

$$p_c = 1 - \frac{1}{K_0 - 1}$$

$$K_0 \equiv \frac{\langle k^2 \rangle}{\langle k \rangle}$$

For Poisson:

$$K_0 = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\langle k \rangle^2 + \langle k \rangle}{\langle k \rangle}$$

$$p_c = 1 - \frac{1}{\langle k \rangle}$$

Efficient Immunization Strategies:

Acquaintance Immunization

Critical Exponents

Using the properties of power series (generating functions) near a singular point (Abelian methods), the behavior near the critical point can be studied.

(Diff. Eq. Melloy & Reed (1998) Gen. Func. Newman Callaway PRL(2000), PRE(2001))

For random breakdown the behavior near criticality in scale-free networks is different than for random graphs or from mean field percolation. For intentional attack-same as mean-field.

Even for networks with $3<\lambda<4$, where $\langle k \rangle$ and $\langle k^2 \rangle$ are finite, the critical exponents change from the known mean-field result $\beta=1$. The order of the phase transition and the exponents are determined by $\langle k^3 \rangle$.

Size of the infinite cluster:

$$P_\infty \sim (p-p_c)^\beta \qquad \beta = \begin{cases} \frac{1}{3-\lambda} & 2<\lambda<3 \\ \frac{1}{\lambda-3} & 3<\lambda<4 \\ 1 & \lambda>4 \end{cases}$$

(known mean field)

Distribution of finite clusters at criticality:

$$n_s \sim s^{-\tau} \qquad \tau = \begin{cases} \frac{2\lambda-3}{\lambda-2} & \lambda<4 \\ 2.5 & \lambda \geq 4 \end{cases}$$

(known mean field)

Fractal Dimensions

From the behavior of the critical exponents the fractal dimension of scale-free graphs can be deduced.

Far from the critical point - the dimension is infinite - the mass grows exponentially with the distance.

At criticality - the dimension is finite for $\lambda>3$.

Chemical dimension:

$$S \sim \ell^{d_\ell}$$

$$d_l = \begin{cases} \frac{\lambda-2}{\lambda-3} & \lambda < 4 \\ 2 & \lambda \geq 4 \end{cases}$$

Fractal dimension:

$$S \sim R^{d_f}$$

$$d_f = \begin{cases} 2\frac{\lambda-2}{\lambda-3} & \lambda < 4 \\ 4 & \lambda \geq 4 \end{cases}$$

Embedding dimension:

(upper critical dimension)

$$d_c = \begin{cases} 2\frac{\lambda-1}{\lambda-3} & \lambda < 4 \\ 6 & \lambda \geq 4 \end{cases}$$

Random Graphs – Erdos Renyi(1960)

Largest cluster at criticality

$$S \sim N^{\frac{2}{3}}$$

Scale Free networks

$$S \sim R^{d_f} \sim N^{\frac{d_f}{d_c}} \sim N^{\frac{\lambda-2}{\lambda-1}} \qquad \lambda \leq 4$$

$$S \sim N^{\frac{2}{3}} \qquad \lambda \geq 4$$

The dimensionality of the graphs depends on the distribution!