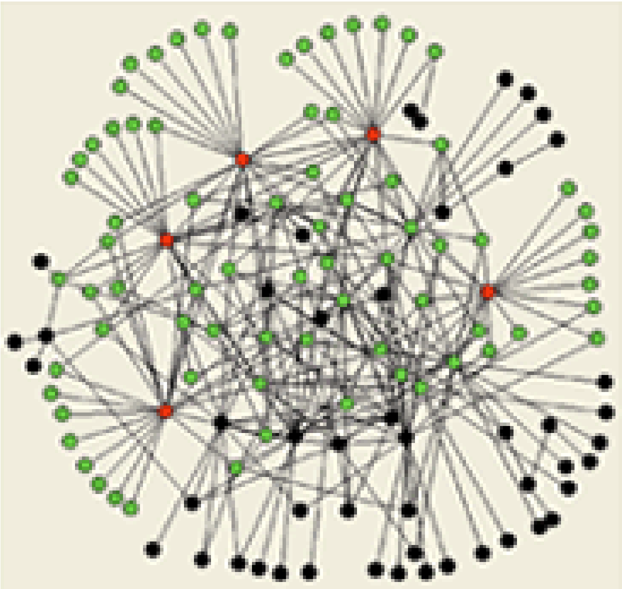
Percolation in Complex Networks: Optimal Paths and Optimal Networks



Shlomo Havlin

Bar-Ilan University

Israel

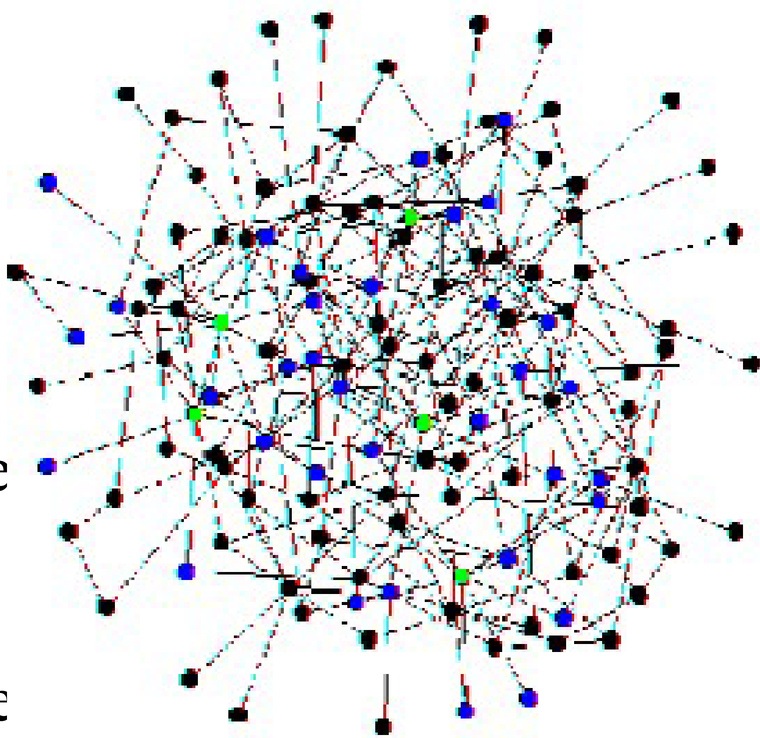
Complex Networks

- 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 mole
  links their interactions



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

Outline

- **Percolation – Graph Theory: Introduction**
- **Degree Distribution, Critical Concentration, Distance, Optimal Distance**
- **Complex Networks: Theory vs Experiment**
- **Generalized Networks: Broad Degree Distribution – Scale Free: Anomalous physics, including percolation**

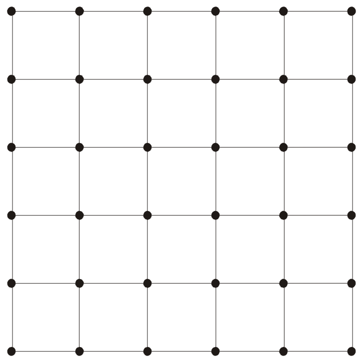
Applications:

- **Efficient Immunization Strategy**
- **Optimal path- Optimal transport**
- **Optimize Network Robustness**

References

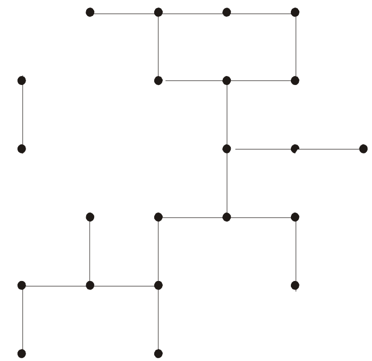
Cohen et al	**Phys. Rev. Lett. 85, 4626 (2000); 86, 3682 (2001)**
Rozenfeld et al	**Phys. Rev. Lett. 89, 218701 (2002)**
Cohen and Havlin	**Phys. Rev. Lett. 90, 58705 (2003)**
Braunstein et al	**Phys. Rev. Lett. 91, 247901 (2003)**
Cohen et al	**Phys. Rev. Lett. 90, 58705 (2003)**

Percolation and Immunization



$p = 0.3$

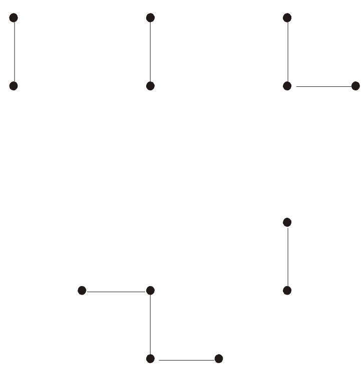
remove
or
immune



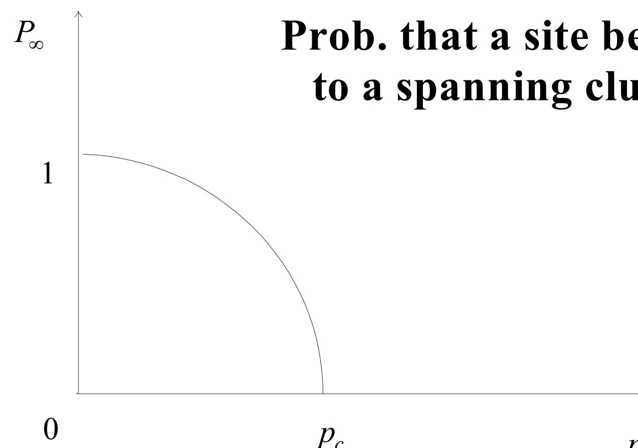
Network exists

$p = 0.5$

remove
or
immune



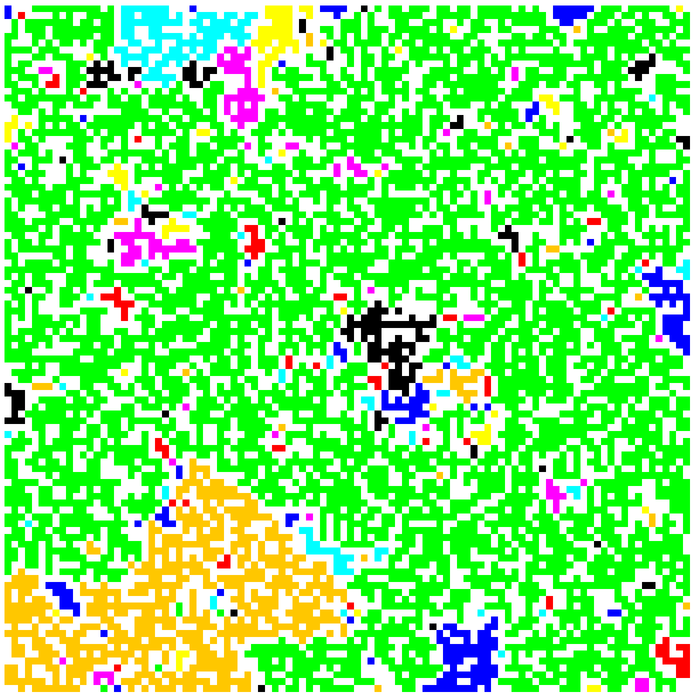
Network collapse



Prob. that a site belongs to a spanning cluster

Percolation Theory

Sites or bonds are randomly removed from a lattice/graph with probability p.



Below threshold

Above threshold

- **A phase transition exists at p_c.**
  - **The correlation length scales like: $\xi \propto |p - p_c|^{-\nu}$.**
    - **The size of the spanning cluster near criticality behaves as: $P_\infty \propto (p_c - p)^\beta$**
      - **The number of clusters of size s behaves at criticality as: $n(s) \propto s^{-\tau}$**

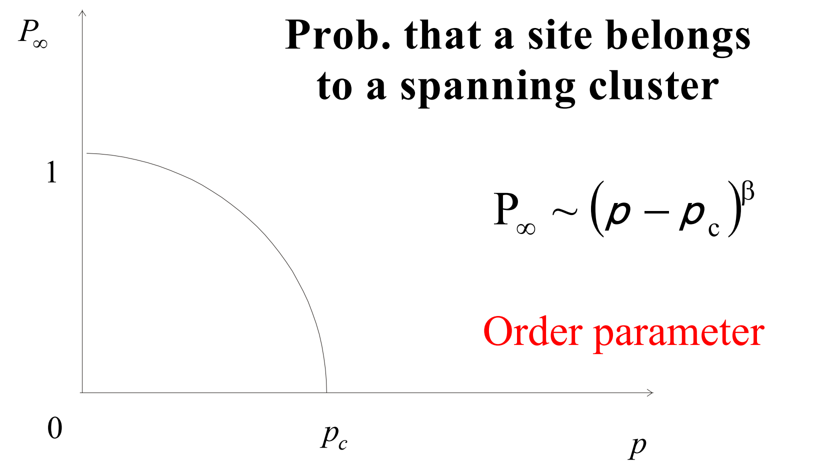
Percolation – critical exponents

- ✓p – same role as T in thermal phase transitions
- ✓P_∞ - probability that a site (or bond) belongs to ∞ cluster order parameter $P_\infty \propto (p-p_c)^\beta$- similar to magnetization

- ✓ ξ - correlation length – mean distance between two sites on the same cluster

$$\xi \propto |p-p_c|^{-\nu}$$

- ✓The average size of finite clusters $S \sim |p-p_c|^{-\gamma}$ (analogous to susceptibility)



- ✓ ν and γ are the same for $p>p_c$ and $p<p_c$
- ✓For ξ and S take into account all finite clusters
- ✓ β, ν and γ called critical exponents $\Rightarrow$ describe critical behavior near the transition
- ✓The exponents are universal
- ✓Universality – property of second order phase transition (order parameter →0 continuously)

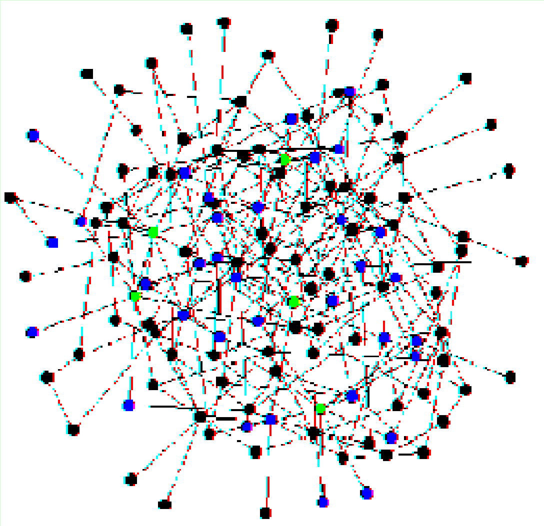
  All magnets in d=3 have same β

  independent on the lattice and type of interactions
- ✓T_c – depends on details (interactions, lattice) – same for p_c

Applications of Percolation

- Oil recovery – Porous media
- Gelation
- Galaxy formation
- Polymerization
- Amorphous materials
- Epidemics and Forest fires

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 :**

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

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

Known 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** $N^{2/3}$

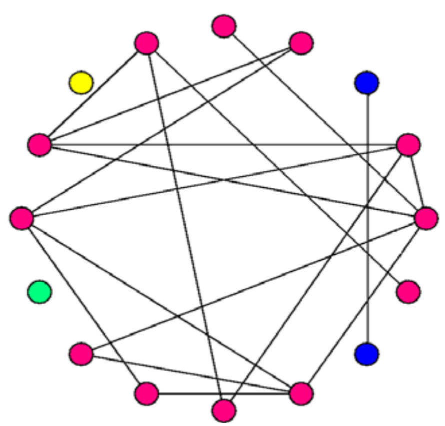
- **Size of the spanning cluster is determined by the self-consistent equation:**

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

- **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 links)**
- **Critical threshold:**

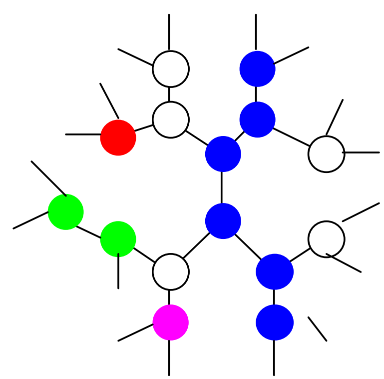
$$p_c = \frac{1}{z-1}$$

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

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

$$\gamma = 1, \ \nu = 1/2, \ \tau = 5/2$$

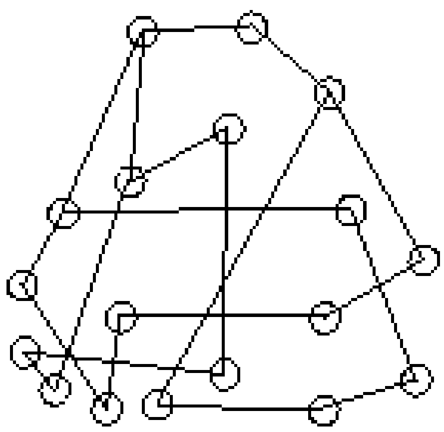
Exactly as for ER networks



Mean Field Exponents -- valid for $d \geq d_c = 6$ Upper critical dimension

In Real World - Many Networks are non-Poissonian

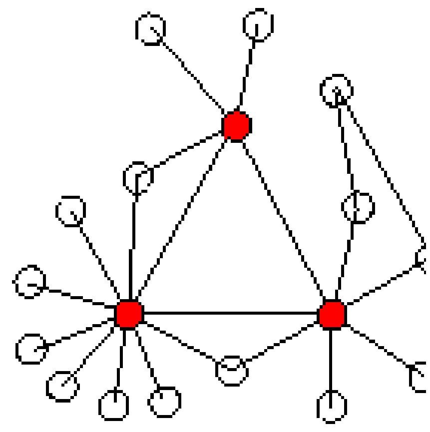
Exponential



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

Homogeneous

Scale-free

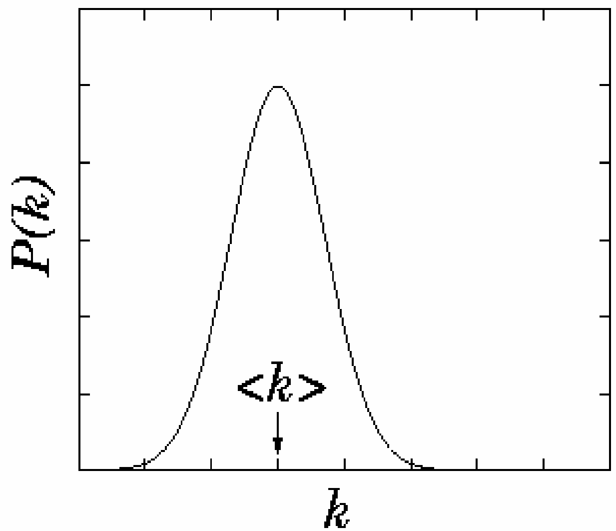


$$P(k) = \begin{cases} ck^{-\lambda} & m \le k \le K \\ 0 & otherwise \end{cases}$$

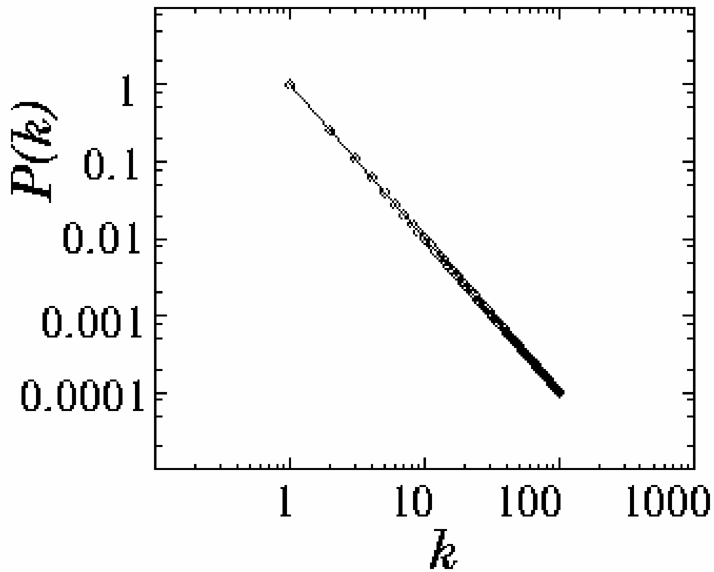
Heterogeneous

New Type of Networks

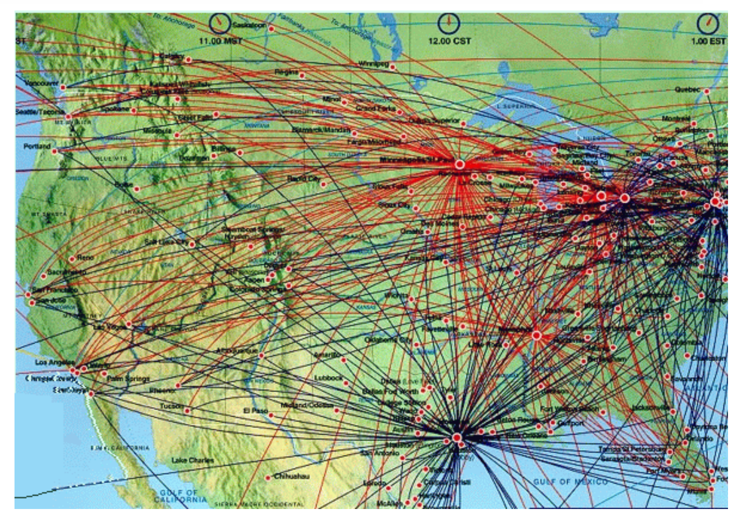
Poisson distribution



Power-law distribution

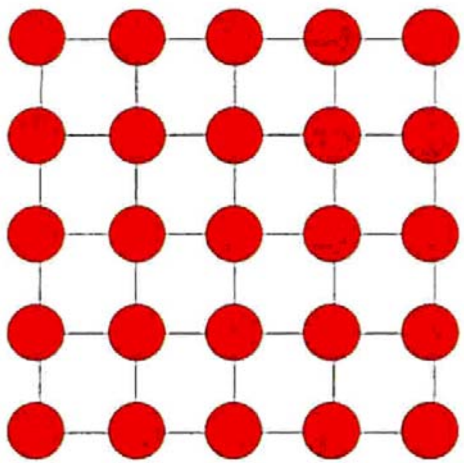


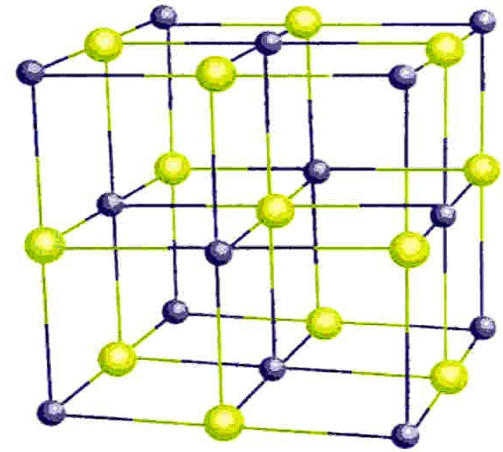
Exponential Network

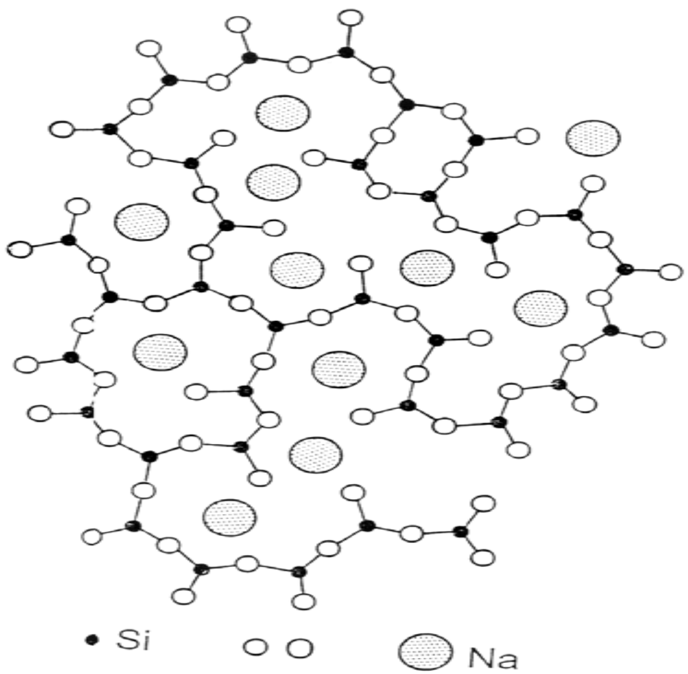


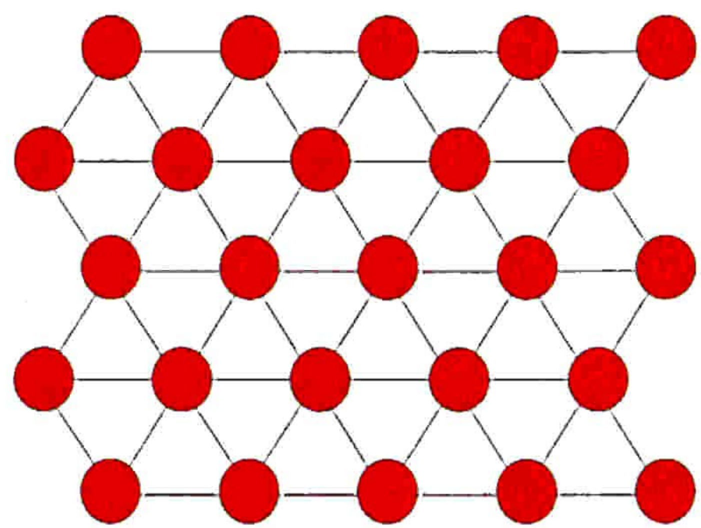
Scale-free Network

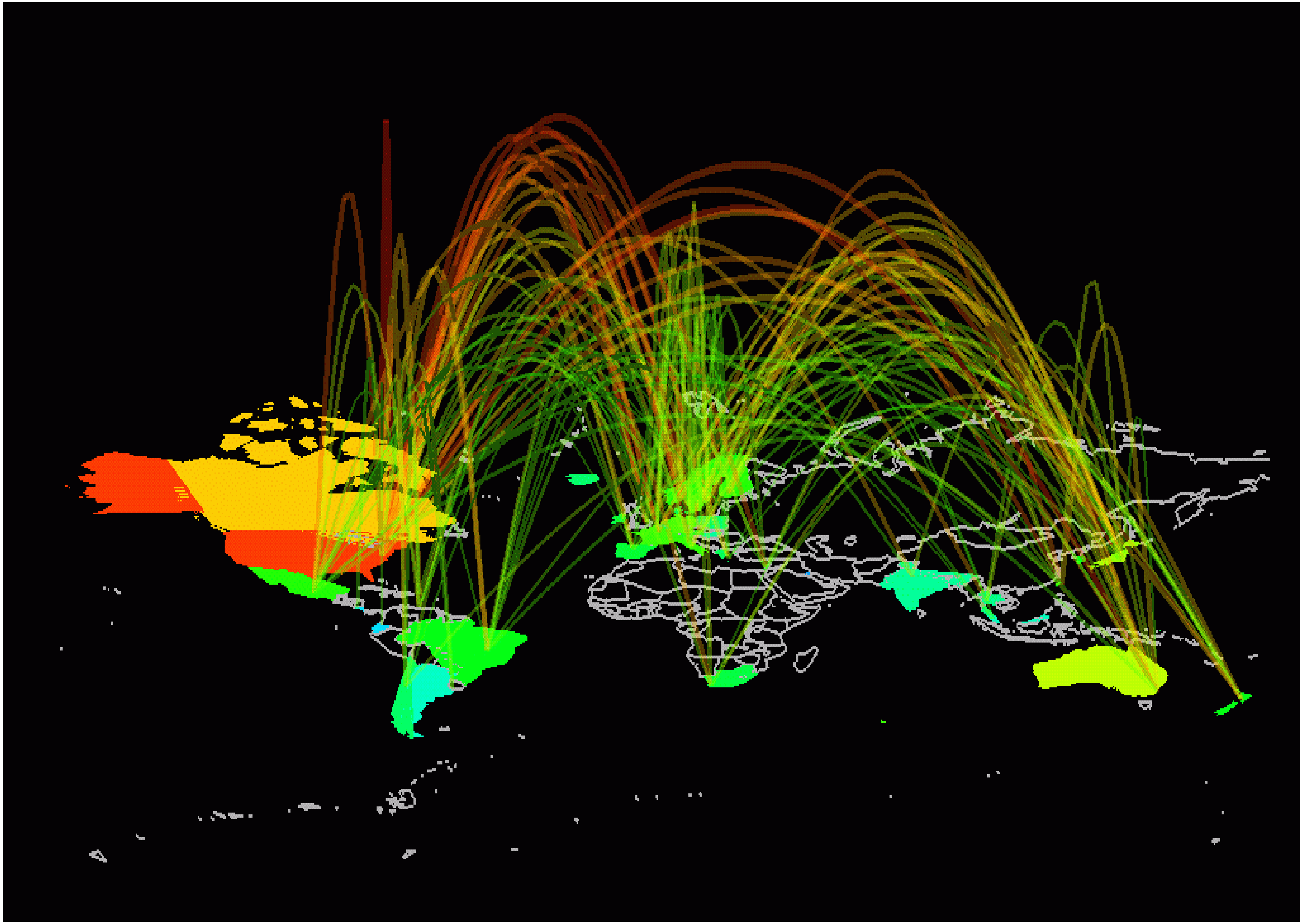
Networks in Physics







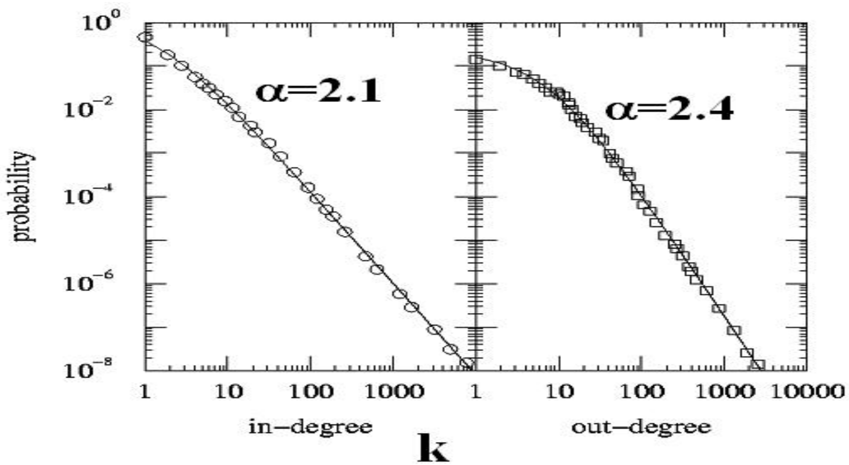




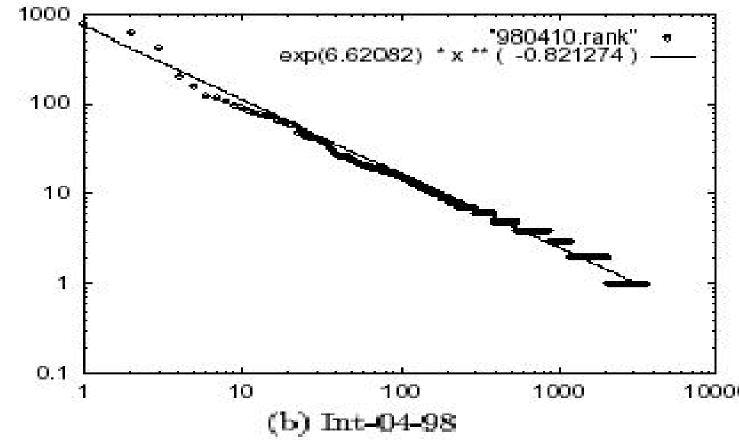
Examples

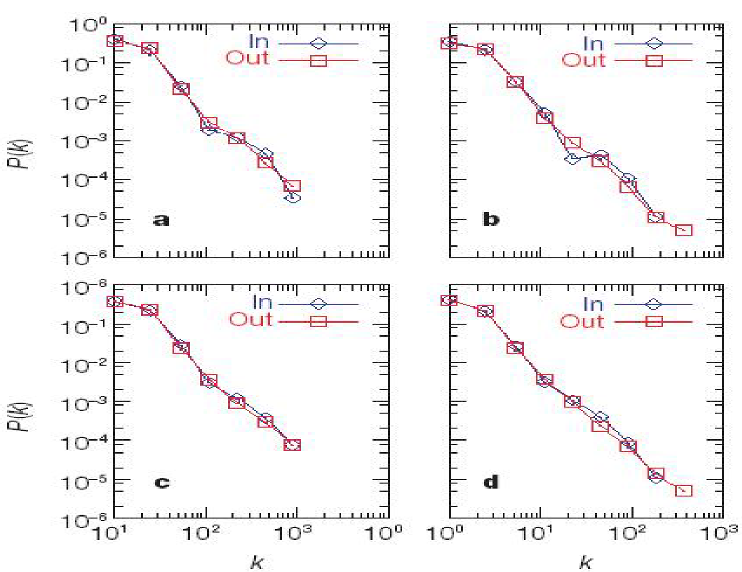
- Internet, WWW (Faloutsos et. al., SIGCOMM ’99, Border et. al. IBM-Altavista research,Albert, Jeong and Barabasi, Nature 2000)
- Protein Families (Delisi et al. PRL (2000), Unger et al, preprint)
- Metabolic cellular networks (Barabasi et. al., Nature 2000)
- Telephone & Power Networks
- Science collaboration networks (Newman, 2001, Barabasi et. al., 2001)
- Ecological Networks (Sole & Montoya, 2001)

P(k)

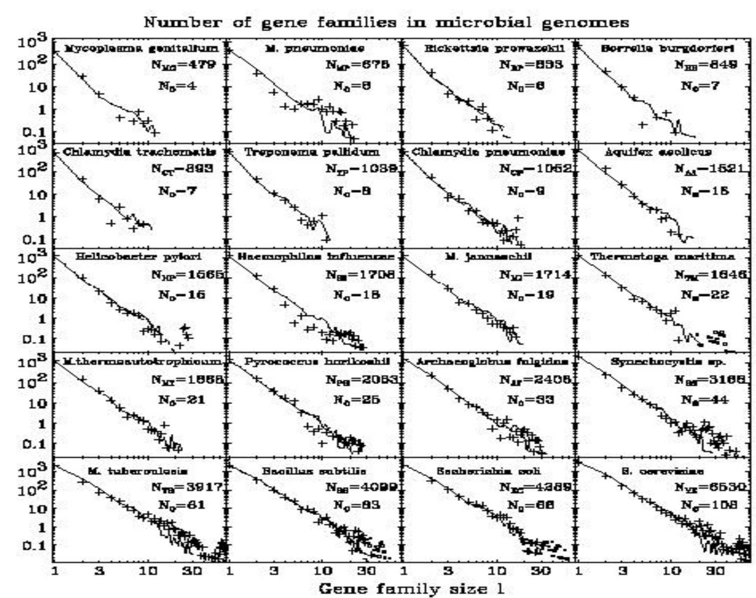


P(k)

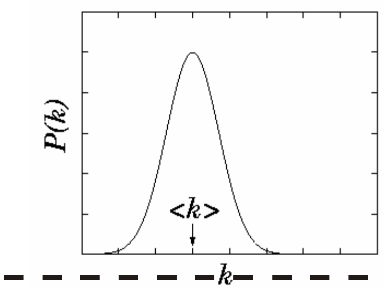
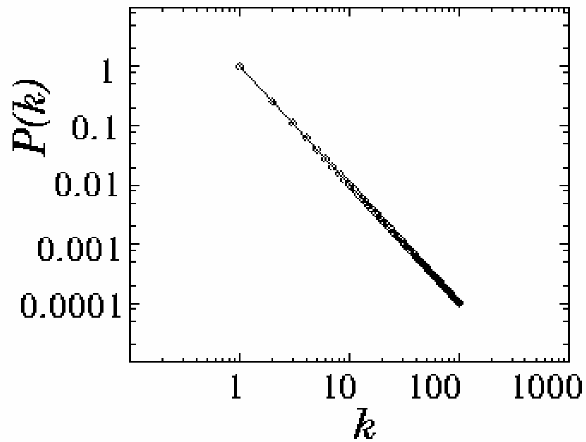




P(k)



Does Percolation Theory Valid?

Stability and Immunization		Distribution	
$$p_c = 1 - \frac{1}{\langle k \rangle}$$ Critical concentration 30-70%	HOMOGENEOUS		Theory
(see table below)	HETEROGENEOUS NEW THEORY IS NEEDED!		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%

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

Modelling: Albert, Jeong, Barabasi (Nature 2000)

Experimental Data: Virus survival

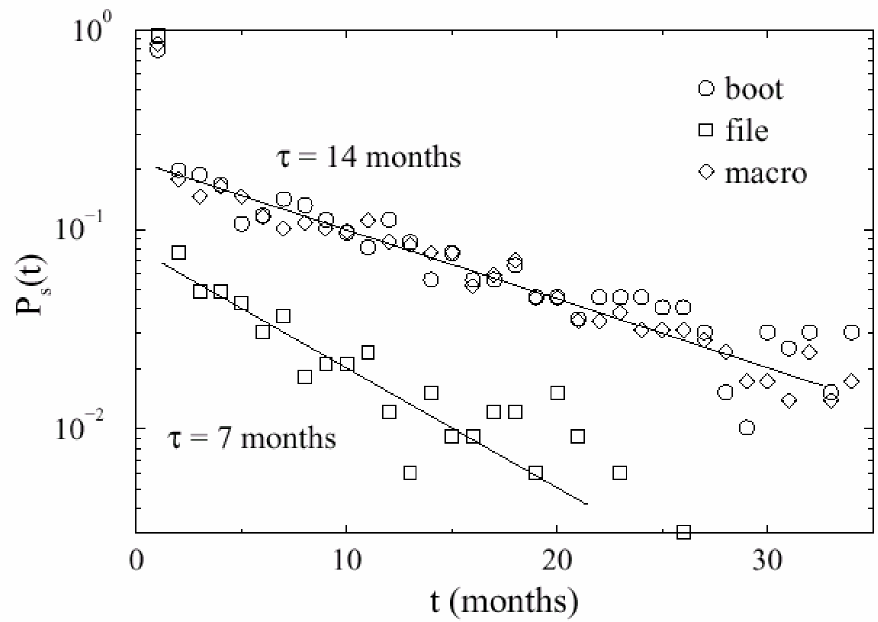


FIG. 1. Surviving probability for viruses in the wild. The 814 different viruses analyzed have been grouped in three main strains [9]: file viruses infect a computer when running an infected application; boot viruses also spread via infected applications, but copy themselves into the boot sector of the 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.

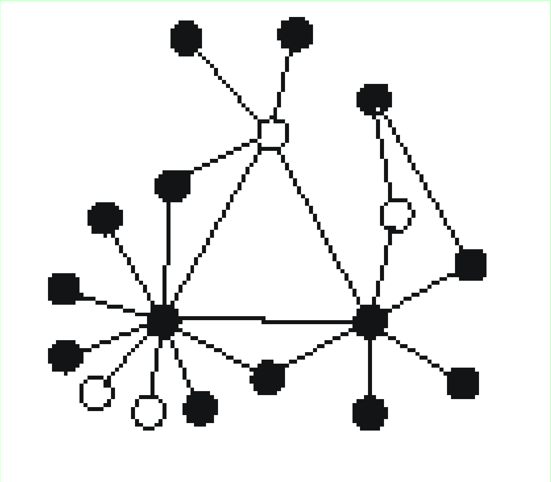
(Pastor-Satorras and Vespignani, Phys. Rev Lett. 86, 3200 (2001))

Percolation Model I

Random Breakdown (Immune)

The Internet and many other real networks are scale-free network, where

$$P(k) \propto k^{-\lambda},\ 2 \le \lambda \le 3$$



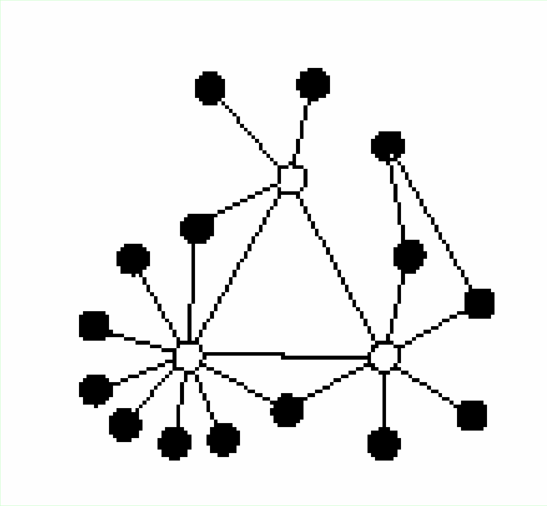
Nodes are randomly removed (or immune) with probability p

Where does the phase transition occur ?

Percolation Model II

Intentional Attack (Immune)

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



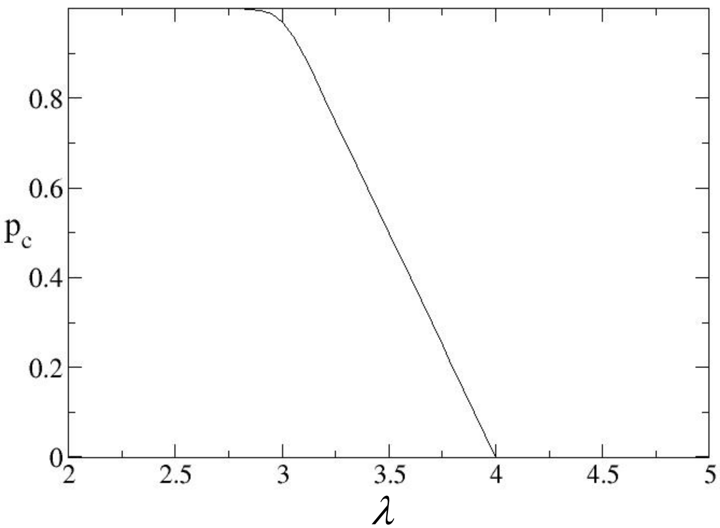
Is this fundamentally different from random breakdown?

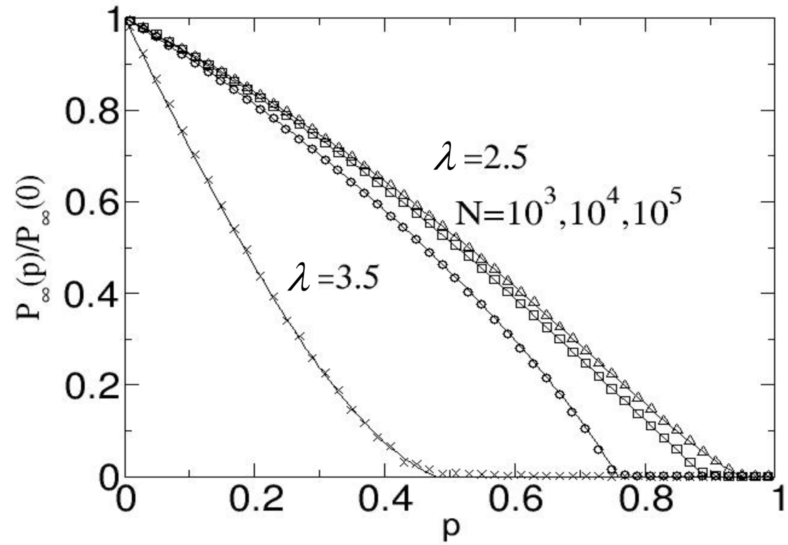
We find that not only critical thresholds but also critical exponents are different !

THE UNIVERSALITY CLASS DEPENDS ON THE WAY CRITICALITY REACHED

Results of Simulations and Theory

Random Breakdown (Immune)

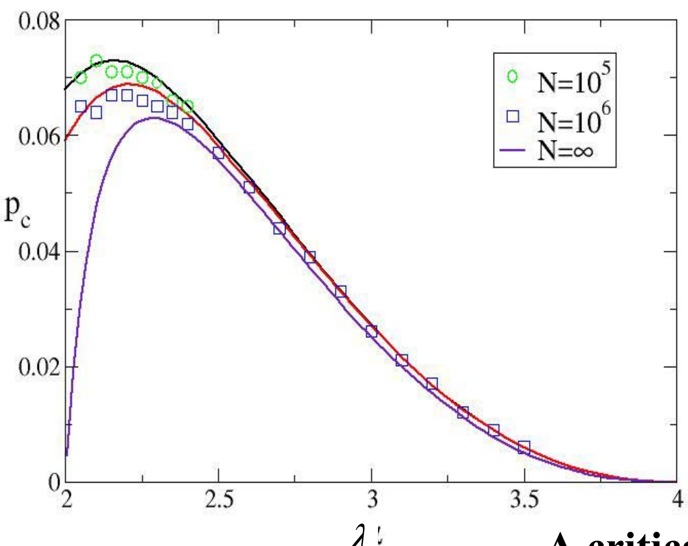


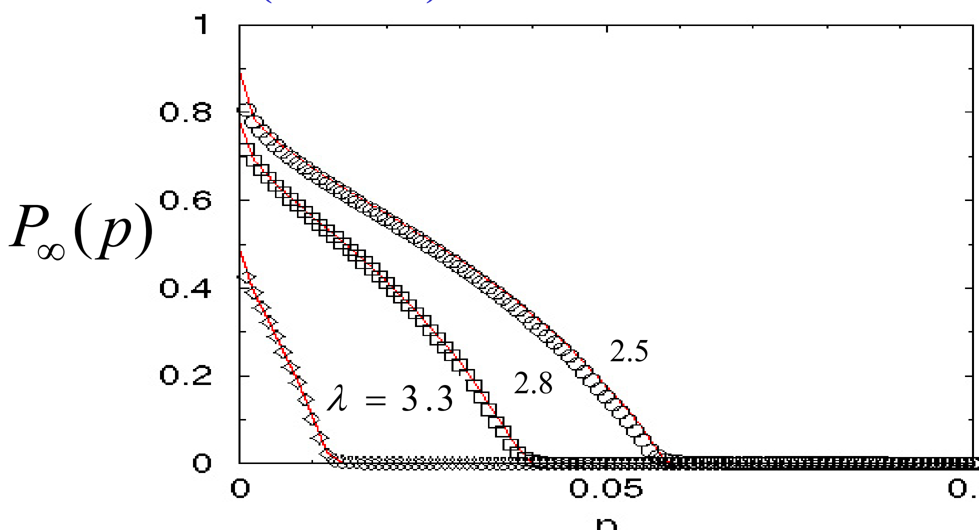


$\lambda \leq 3$ - no critical threshold – a spanning cluster always exists

$\lambda > 3$ - a critical threshold exists

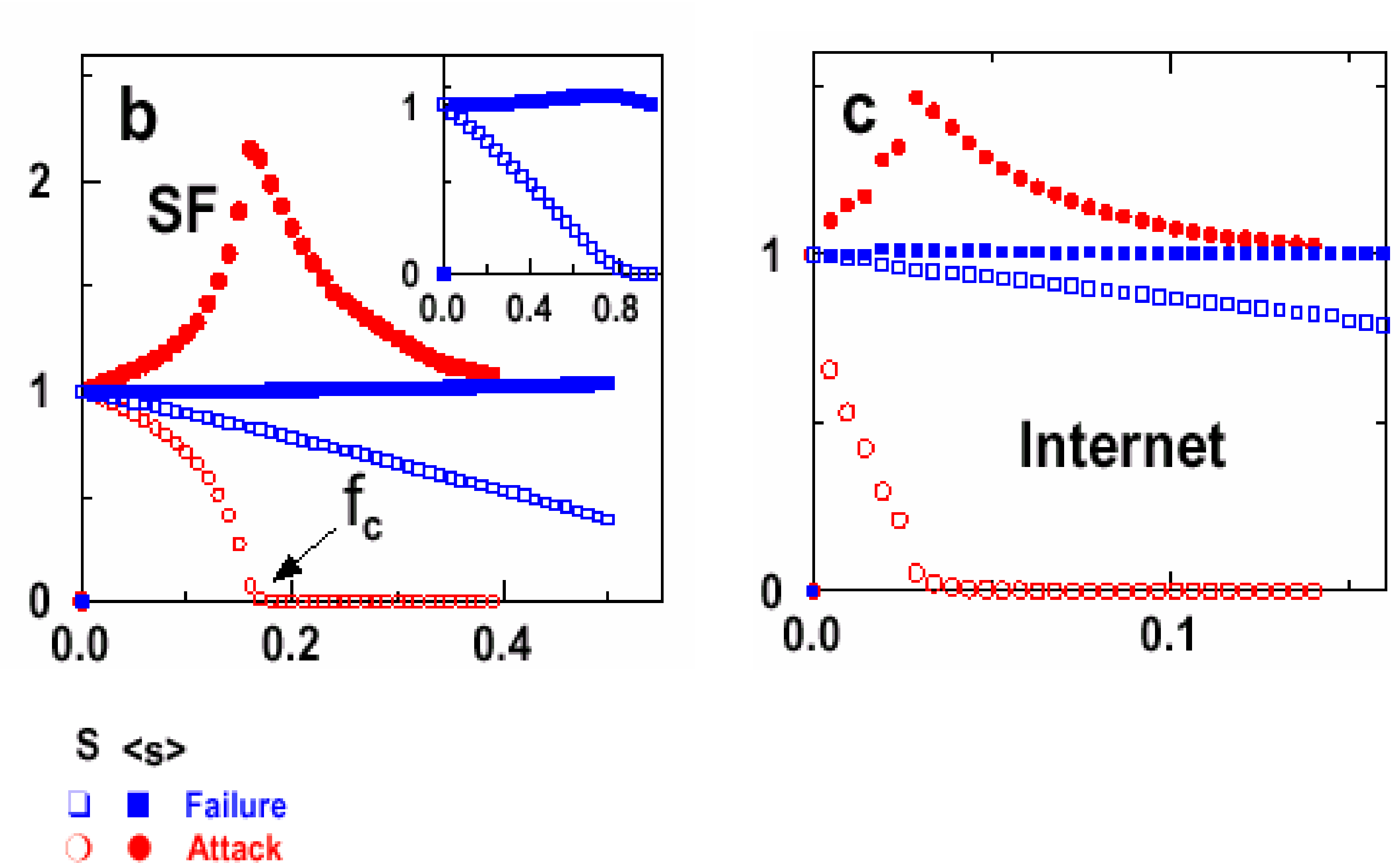
Intentional Attack (Immune)





A critical threshold exist for every λ

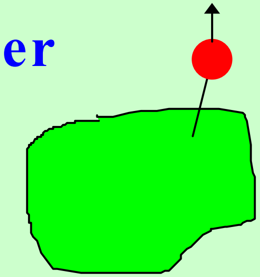
Experimental Data: Internet Stability



(Albert, Jeong and barabasi, Nature 406, 378 (2000))

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: $$P(k_i | i \leftrightarrow j) = \frac{P(i \leftrightarrow j | k_i) P(k_i)}{P(i \leftrightarrow j)}$$

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: $$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

(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 change in κ **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** $p_c = 1 - 1/ < k_0 >$ **for Erdos Renyi**

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: $1 - \bar{p}_c = \dfrac{1}{\bar{\kappa} - 1}$**,**

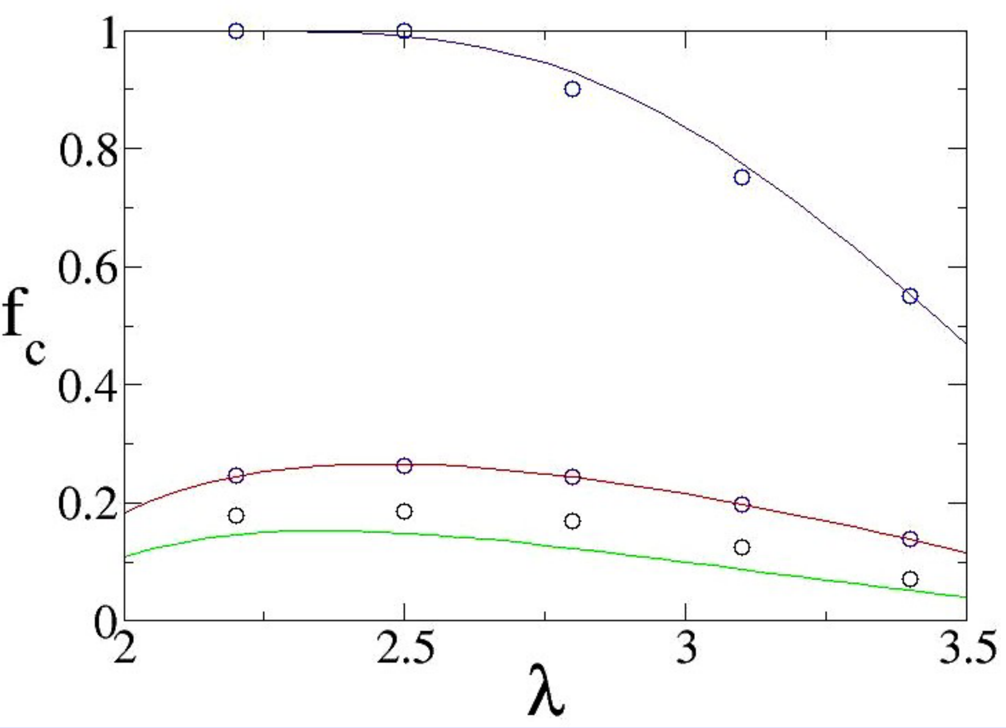
where $\bar{\kappa} = \left(\dfrac{2-\alpha}{3-\alpha}\right)\dfrac{\bar{K}^{3-\lambda} - m^{3-\lambda}}{\bar{K}^{2-\lambda} - m^{2-\lambda}}$ **, gives the critical threshold.**

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

Efficient Immunization Strategies:

Acquaintance Immunization

Critical Threshold
Scale Free



•Random immunization is inefficient in scale free graphs, while targeted immunization requires knowledge of the degrees.

• In *Acquaintance Immunization* one immunizes random neighbors of random individuals.

• One can also do the same based on *n* neighbors.

•The threshold is finite and no global knowledge is necessary.

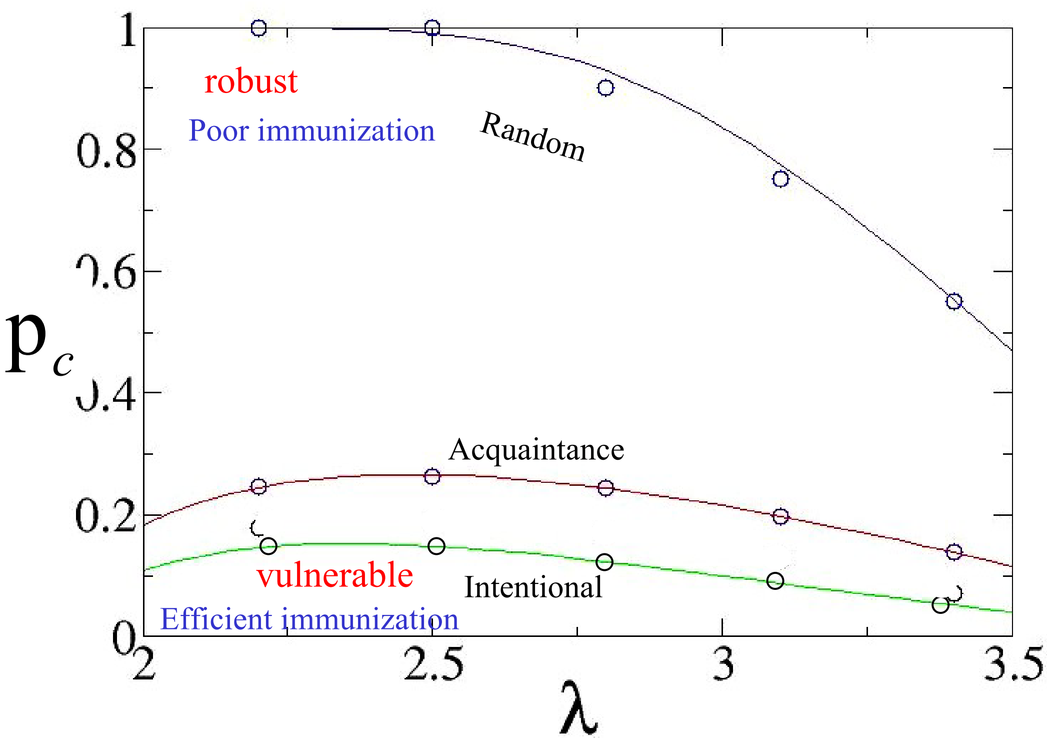
$$\sum_k P(k)(k-1)\left(k/\langle k\rangle\right) v_{p_c}^{k-2} e^{-2p_c/k} = 1$$

$$f_c = 1 - \sum_k P(k) v_{p_c}^k$$

$$v_p = \sum_k kP(k)\exp(-p/k)/\langle k\rangle$$

Cohen et al PRL (2003); cond-mat/0207387

Critical Threshold
Scale Free



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

General result:

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = 2 \leftrightarrow criticality$$

$$p_c = 1 - \frac{1}{\kappa_0 - 1}$$

$$\kappa_0 \equiv \frac{\langle k^2 \rangle}{\langle k \rangle}$$

For Poisson :

$$\kappa_0 = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\langle k \rangle^2 + \langle k \rangle}{\langle k \rangle}$$

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

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 classical mean field percolation. For intentional attack- same as classical mean-field.

Even for networks with $3 < \lambda < 4$, where $\langle k \rangle$ and $\langle k^2 \rangle$ are finite, the critical exponents differ 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:

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

(classical mean field, for d>d, and ER)

Distribution of finite clusters at criticality:

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

(classical mean field, d>6 or ER)

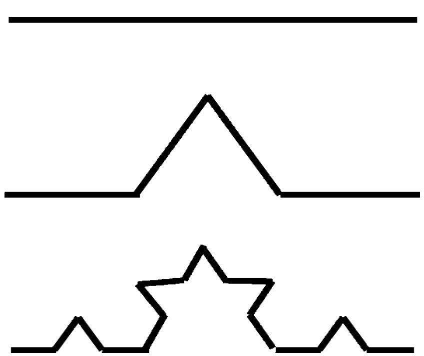
(a) Intentional attack same exponents as for ER, (b) Percolation clusters at criticality are fractals

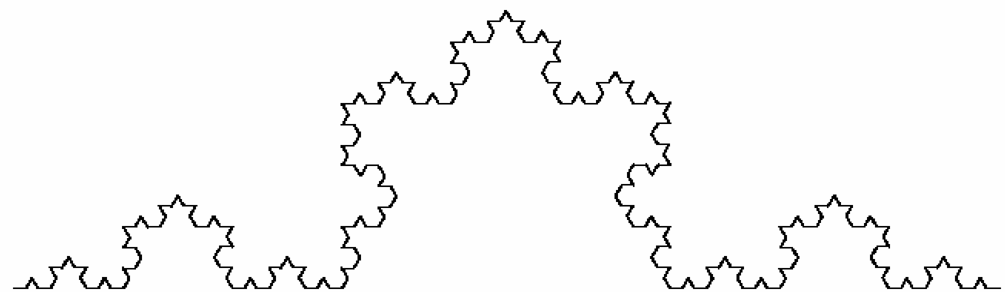
Fractals

Fractal geometry describes Nature better than classical geometry
Two types of fractals: deterministic and random.

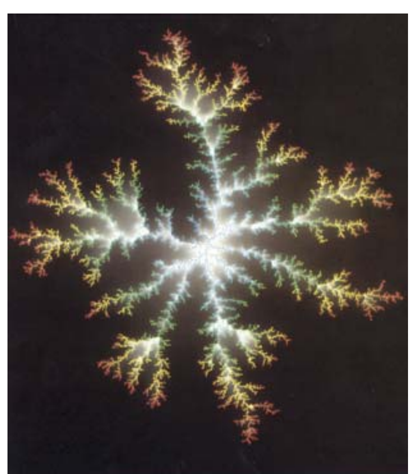
Deterministic fractals and Random fractals

Ideal fractals are self-similar.
Every small part of the picture when magnified properly, is the same as the whole picture.





Koch curve

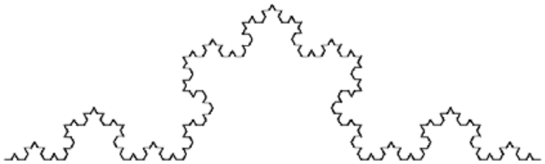


DLA

Definition of fractal dimension $M(bL) = b^{d_f} M(L)$

generalization to non-integer dimension d_f

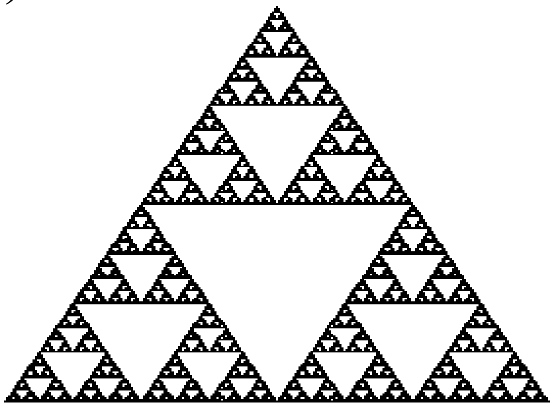
Solution: $M(L) = AL^{d_f}$



Example: Koch curve

$$M\left(\frac{1}{3}L\right) = \frac{1}{4}M(L) = \left(\frac{1}{3}\right)^{d_f} M(L) \Rightarrow \left(\frac{1}{3}\right)^{d_f} = \frac{1}{4} \quad or \quad d_f = \frac{\log 4}{\log 3} \approx 1.262$$

d_f - non integer – between 1 and 2 dimensions. Koch curve is not a line (d=1) but doesn't fill a plane (d=2).

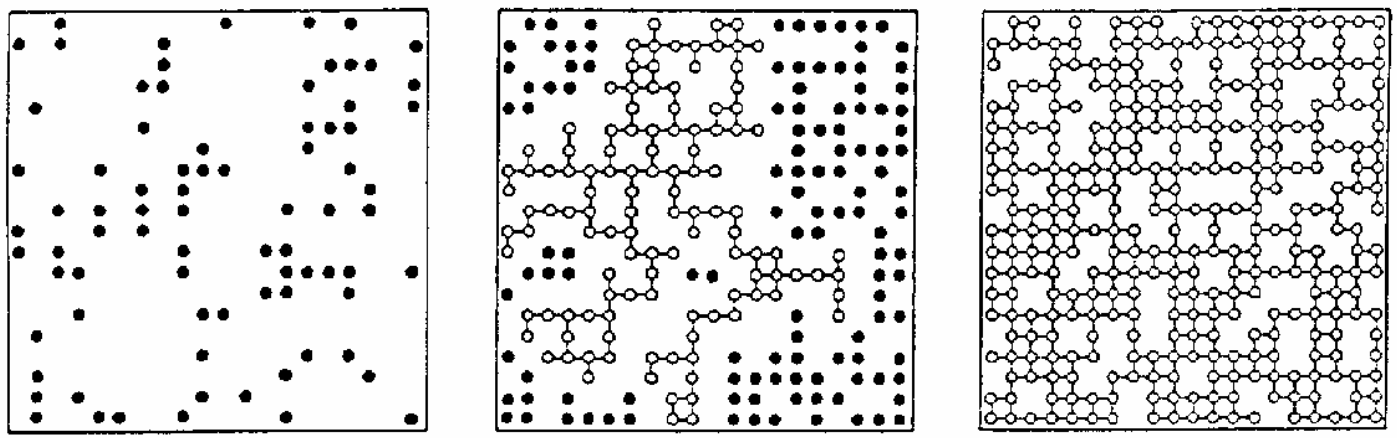


Example: Sierpinski gasket

$$M\left(\frac{1}{2}L\right) = \frac{1}{3}M(L) = \left(\frac{1}{2}\right)^{d_f} M(L) \Rightarrow \left(\frac{1}{2}\right)^{d_f} = \frac{1}{3} \quad or \quad d_f = \frac{\log 3}{\log 2} \approx 1.585$$

Non integer dimension between 1 and 2 dimensions.

Fractal Dimension at Critical Percolation



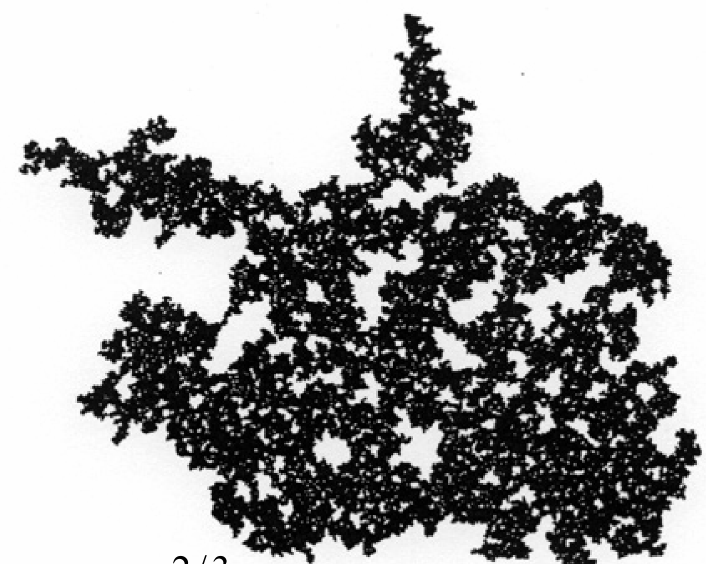
p=0·2 p=0·59 p=0·8

for $d = 2$

$$S \sim R^{d_f},\ d_f = \frac{91}{48} = 1.896$$

$$S \sim \ell^{d_\ell},\ d_\ell \approx 1.68$$

$$N = R^d = R^2,\ S \sim N^{d_f/d} \sim N^{0.95}$$



For $d = d_c = 6,\ N = R^6,\ d_f = 4,\ d_\ell = 2 \rightarrow S = N^{2/3}$

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 $\lambda > 3$.

Short path dimension:

$$d_\ell = \begin{cases} \dfrac{\lambda-2}{\lambda-3} & \lambda < 4 \\ 2 & \lambda \geq 4 \end{cases}$$

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

Fractal dimension:

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

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

Embedding dimension:

(upper critical dimension)

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

Random Graphs – Erdos Renyi(1960)

Largest cluster at criticality

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

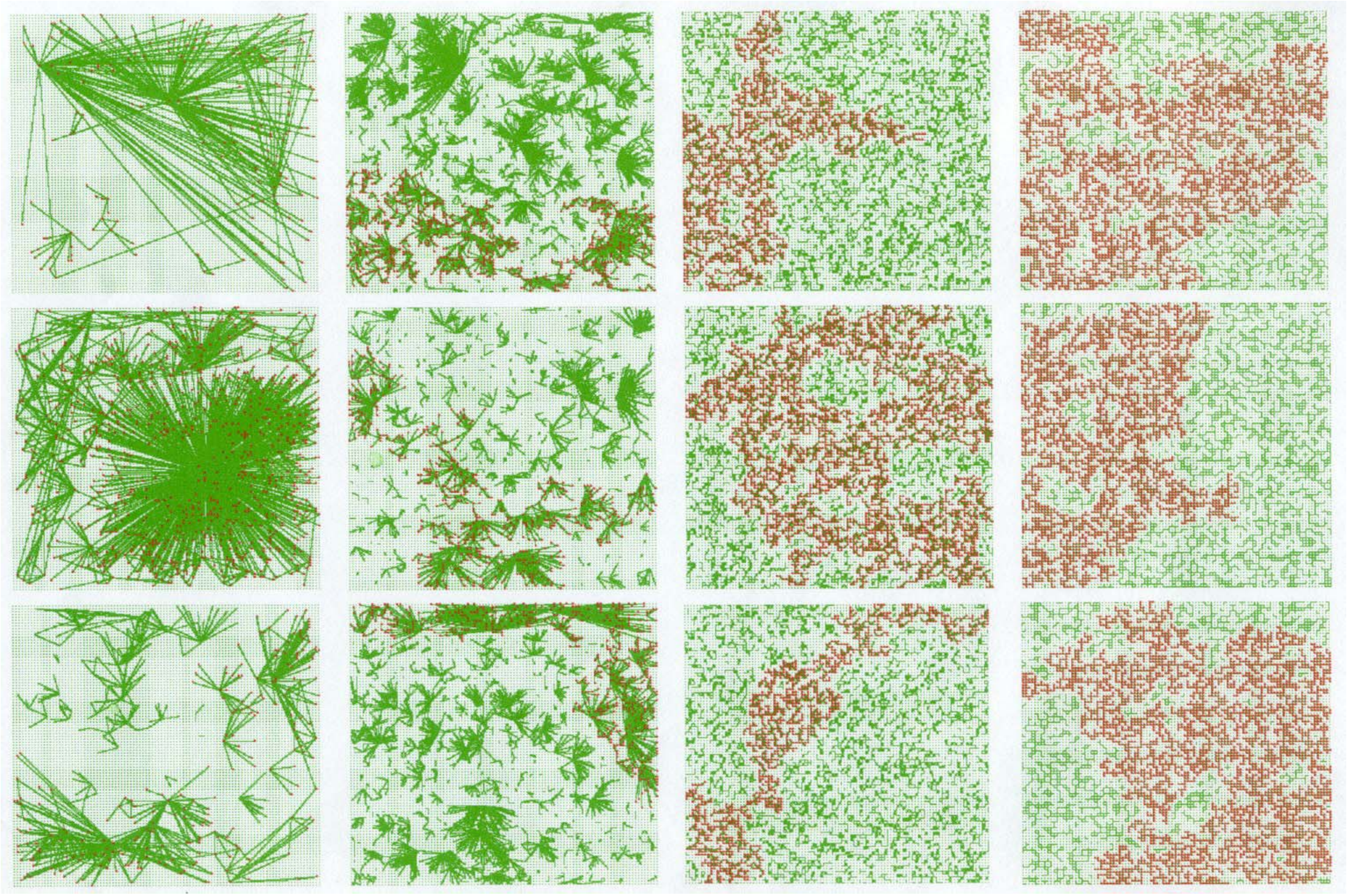
Scale Free networks

$$N = R^d$$

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

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

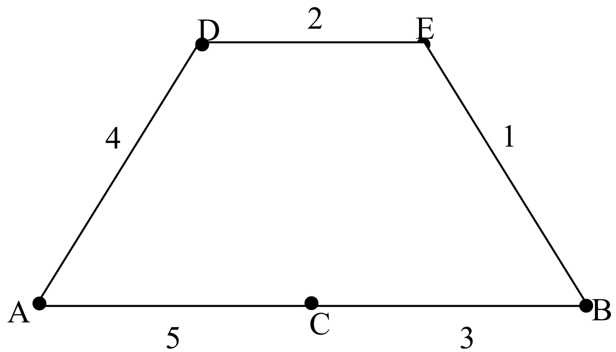
The dimensionality of the graphs at criticality depends on the distribution!



λ=2 λ=2.5 λ=5 λ=50

Rozenfeld et al PRL (2002), Manna et al (2002)

Problem: Optimal Distance



Path from **A** to **B**

l_{min}= 2(ACB)

l_{opt}= 3(ADEB)

$\boldsymbol{w_i}$ = weight = price, quality, time…..
$$\sum_i w_i = \text{minimal} \Rightarrow \text{optimal path}$$

Weak disorder (WD) – all $\boldsymbol{w_i}$ contribute to the sum (narrow distribution)

Strong disorder (SD)– a single term dominates the sum (broad distribution)

THE PATH IN STRONG DISORDER=THE PATH ON A MINIMAL SPANNING TREE (MST)

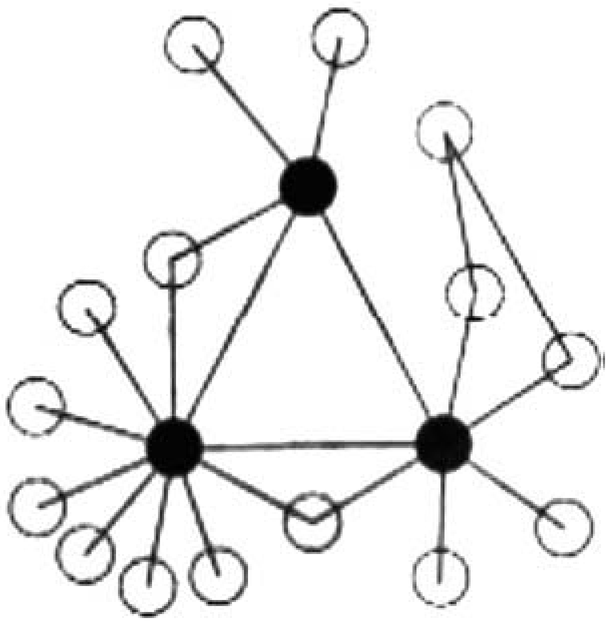
SD – example: Broadcasting video over the Internet,

a transmission at constant high rate is needed.

The narrowest band width link in the path

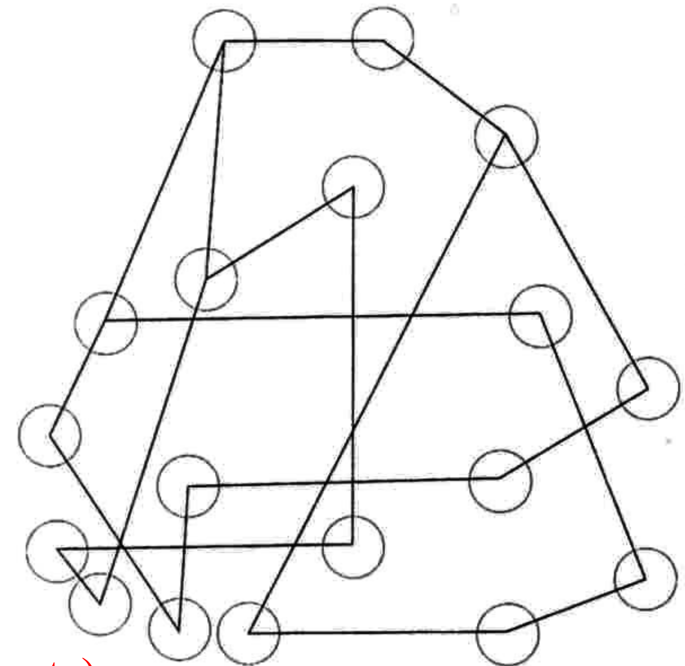
between transmitter and receiver controls the rate.

Scale Free (Barabasi-Albert)



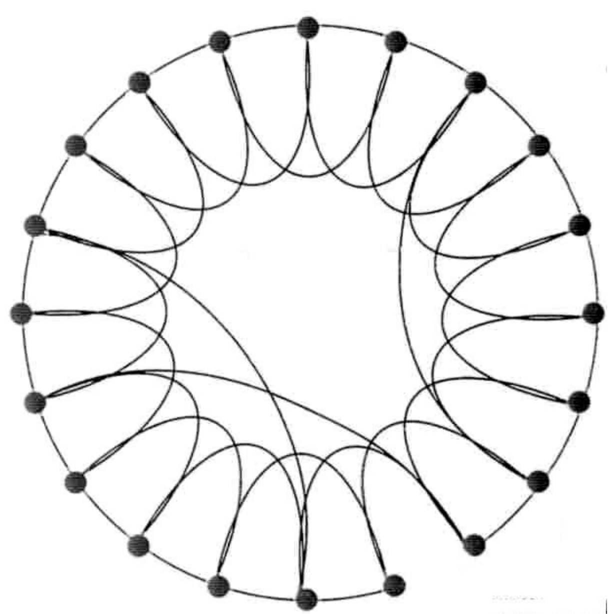
$$P(k) = Ak^{-\lambda}$$

Random Graph (Erdos-Renyi)



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

Small World (Watts-Strogatz)



Z = 4

Shortest Paths in Scale Free Networks

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

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

Ultra Small World

$$\boxed{d = \log\log N \qquad 2 < \lambda < 3}$$

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

(Bollobas, Riordan, 2002)

Small World

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

(Bollobas, 1985)
(Newman, 2001)

Same as for ER and WS

Cohen, Havlin and ben-Avraham, in Handbook of Graphs and Networks

eds. Bornholdt and Shuster (Willy-VCH, NY, 2002) Chap.4

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

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

Optimal path – weak disorder

Random Graphs and Watts Strogatz Networks

$$l_{\min} \sim \log N$$

$$l_{opt} \sim \log N$$

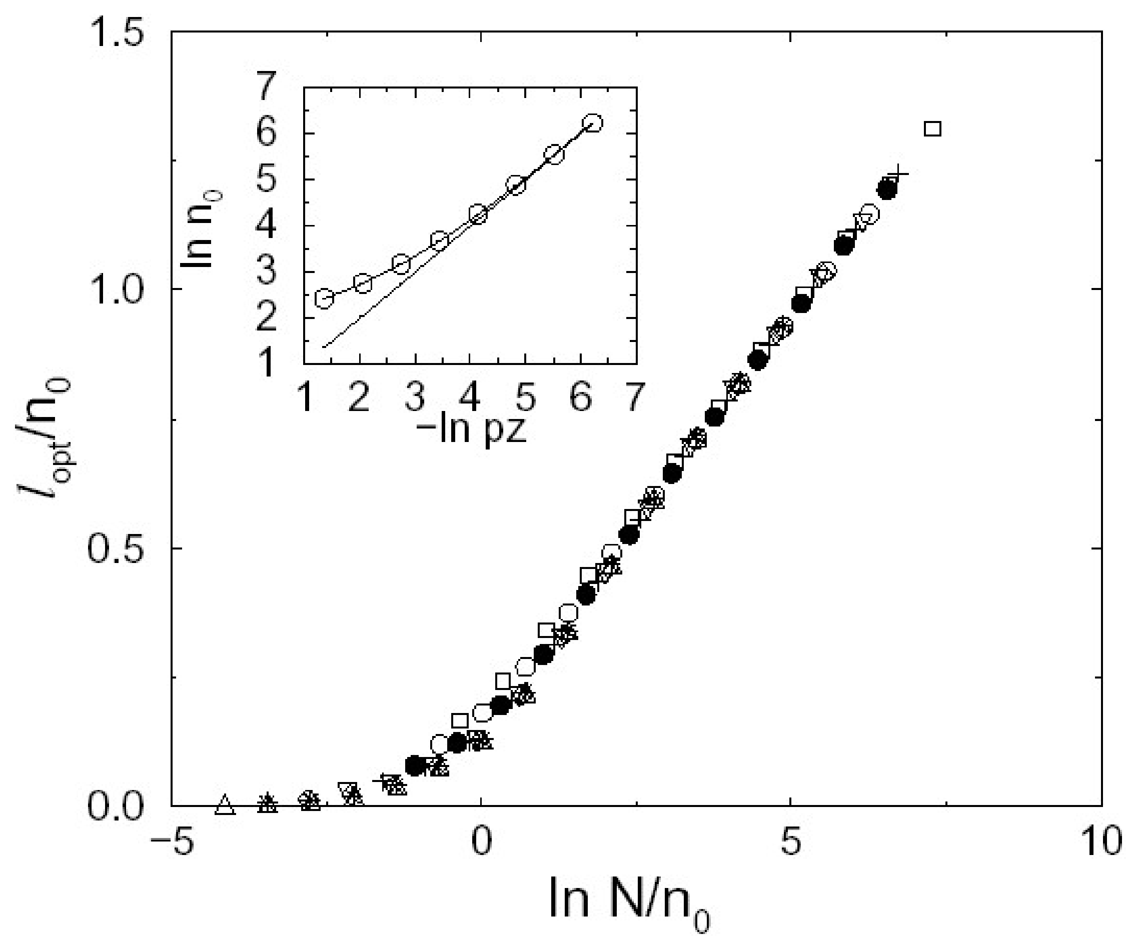
$$n_0 \propto \frac{1}{pz}$$

Typical short range neighborhood

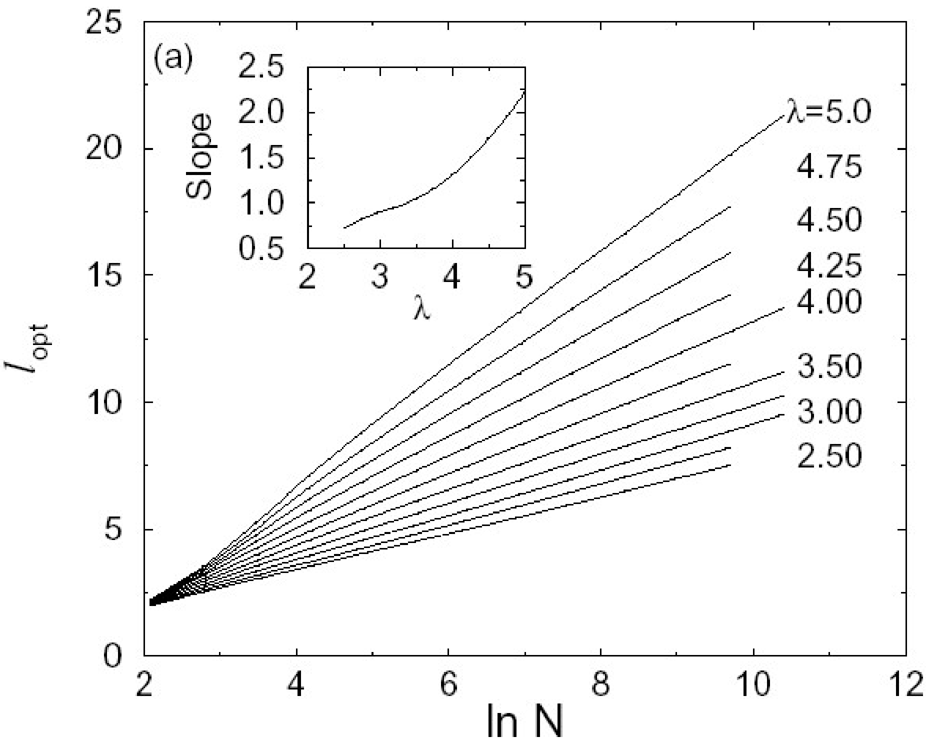
Crossover from large to small world

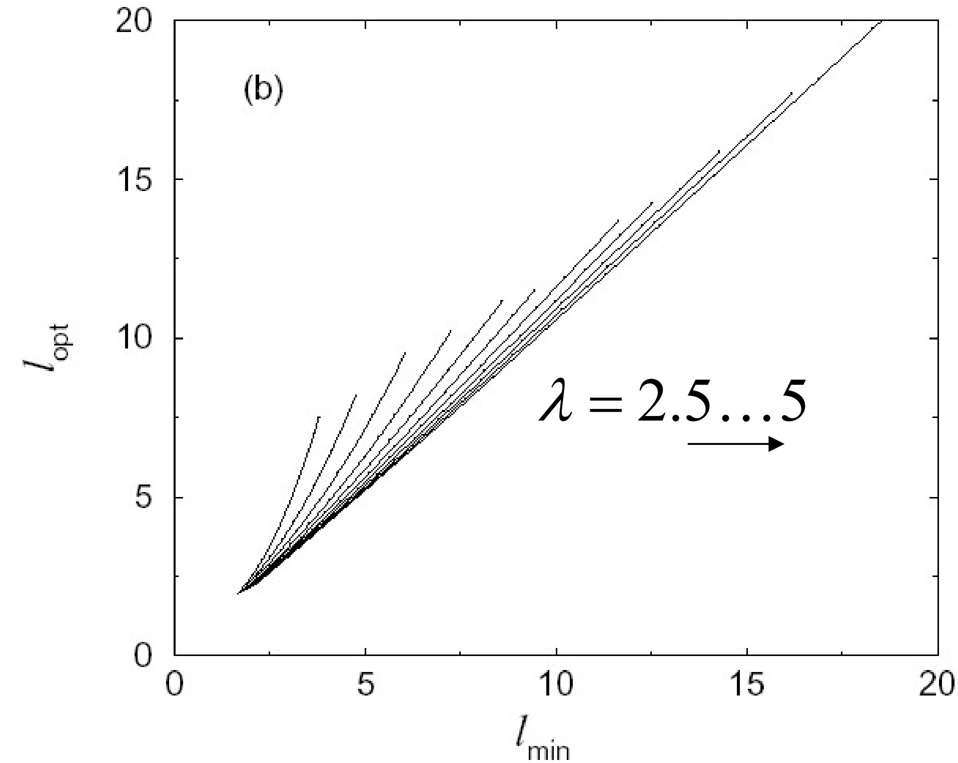
For $N / n_0 < 1$

$$l_{opt} \sim N \sim e^{\log N / n_0}$$



Scale Free – Optimal Path – Weak disorder





$$For \quad \lambda > 3$$

$$l_{opt} \sim A(\lambda) \log N$$

$$l_{\min} \sim B(\lambda) \log N$$

$$Thus \quad \boxed{l_{opt} \sim l_{\min}}$$

$$For \ 2 < \lambda < 3$$

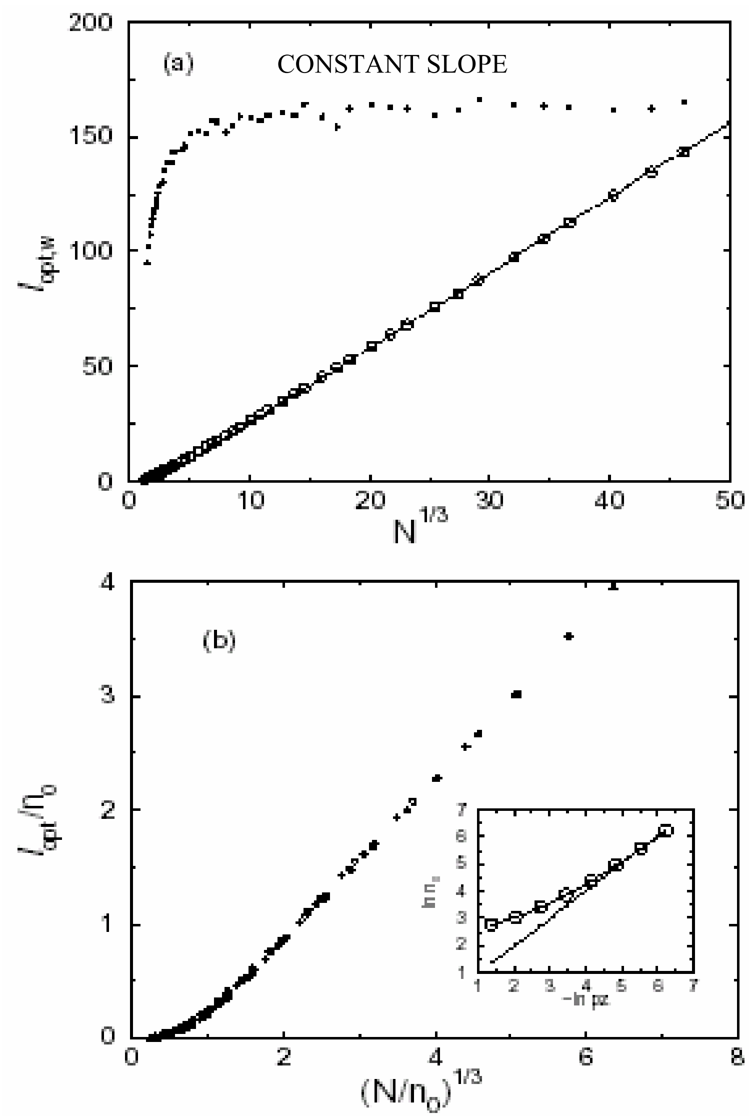
$$l_{opt} \sim \log N$$

$$l_{\min} \sim \log \log N$$

$$Thus \ \boxed{l_{opt} \sim \exp(l_{\min})}$$

Optimal Path – Strong Disorder-Minimal Spanning Trees (MST)

Random Graphs and Watts Strogatz Networks



N – total number of nodes

$$l_{opt} \sim N^{1/3}$$ Analytically and Numerically

LARGE WORLD!!

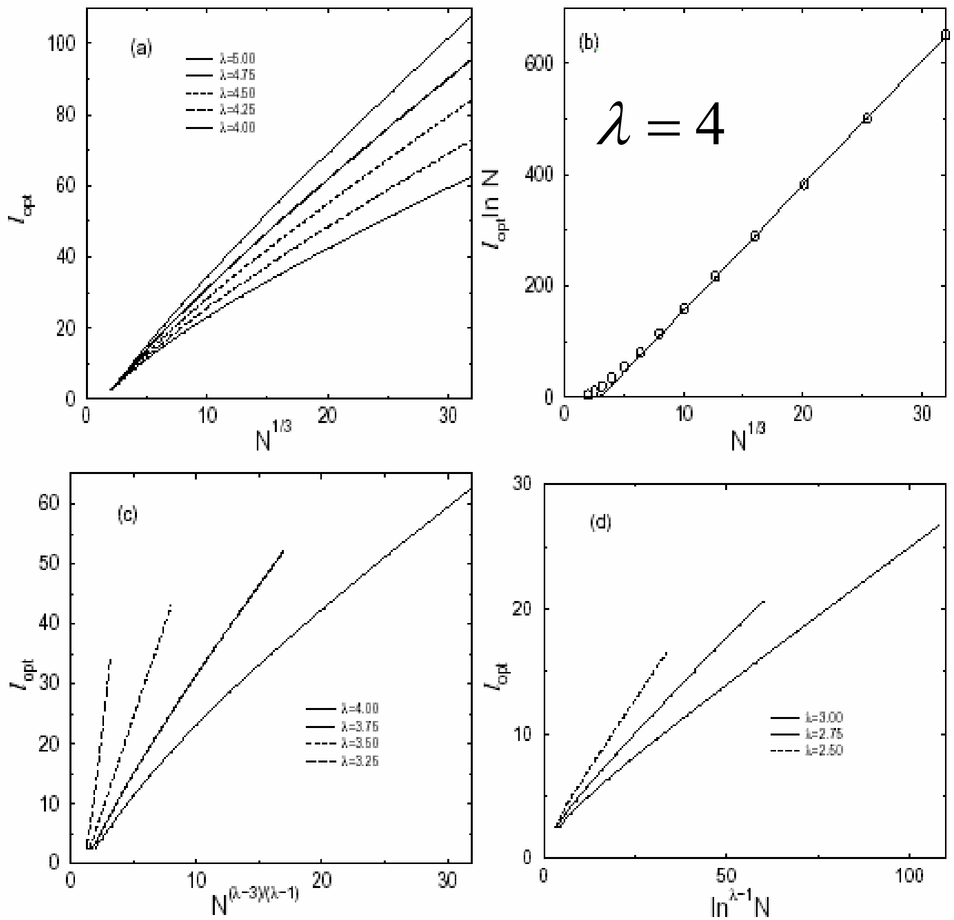
Compared to the diameter or average shortest path or weak disorder

$$l_{\min} \sim \log N$$ (small world)

n_0 - typical range of neighborhood without long range links

$\frac{N}{n_0}$ - typical number of nodes with long range links

Scale Free – Optimal Path



Strong Disorder-Minimal Spanning Trees

Theoretically + Numerically

$$l_{opt} \sim \begin{cases} N^{(\lambda-3)/(\lambda-1)} & 3<\lambda<4 \\ N^{1/3}\log N & \lambda = 4 \\ N^{1/3} & \lambda > 4 \end{cases}$$

LARGE WORLD!!

Numerically $\quad l_{opt} \sim \log^{\lambda-1} N \quad 2<\lambda<3$

SMALL WORLD!!

Weak Disorder

$$l_{opt} \sim \log N \quad \textit{for all } \lambda$$

Diameter – shortest path

$$l_{\min} \sim \begin{cases} \log N & \lambda > 3 \\ \log N / \log\log N & \lambda = 3 \\ \log\log N & 2<\lambda<3 \end{cases}$$

Braunstein et al,
Phys. Rev. Lett. 91, 247901 (2003);
Cond-mat/0305051

Theoretical Approach – Strong Disorder

(i) Distribute random numbers 0<u<1 on the links of the network.

(ii) Strong disorder represented by $\varepsilon_i = \exp(au_i)$ with $a \to \infty$.

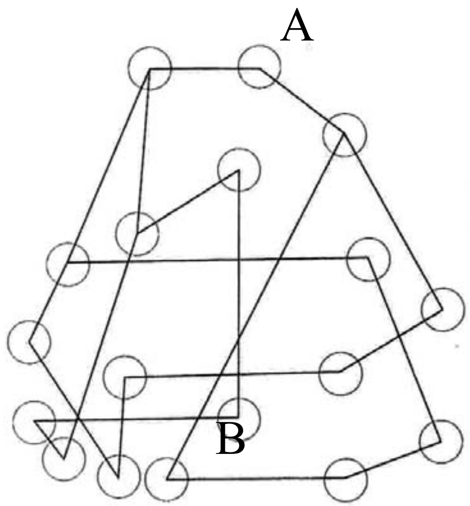
(iii) The largest u_i in each path between two nodes dominates the sum.

(iv) The optimal path is the path with the min-max

(v) Percolation exists if we remove all links with $u_i > 1 - p_c$

(vi) The optimal path must therefore be on the percolation cluster at criticality.

What do we know about percolation clusters at criticality in networks?



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 $\lambda>3$.

Short path dimension:

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

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

Fractal dimension:

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

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

Embedding dimension:

(upper critical dimension)

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

Random Graphs – Erdos Renyi(1960)

Largest cluster at criticality

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

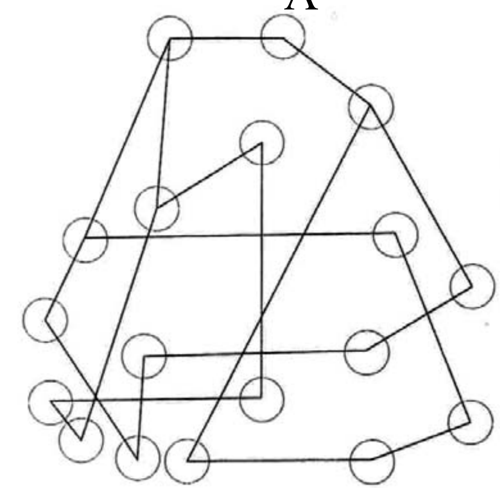
Scale Free networks

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

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

The dimensionality of the graphs depends on the distribution!

Theoretical Approach – Strong Disorder Conclusions



(i) Percolation on random networks is like percolation in $d \to \infty$ or $d = d_c$.

(ii) Since loops can be neglected the optimal path can be identified with the shortest path on percolation-only a single path exist between any nodes.

Calculate the length of shortest path:

Mass of infinite cluster $S \sim R^{d_f}$ where $N \sim R^d$

For ER $d_c = 6$, $S \sim N^{d_f/d} = N^{4/6} = N^{2/3}$ (also Erdos-Renyi, 1960)

From percolation $S \sim l^2, (d_l = 2)$

Thus, for ER, WS and SF with $\lambda > 4$: $l_{opt} \sim l \sim S^{1/2} \sim N^{1/3}$

For SF with $3 < \lambda < 4$ $d_c, d_f,$ and d_l change due to novel topology: $l_{opt} \sim N^{(\lambda-3)/(\lambda-1)}$

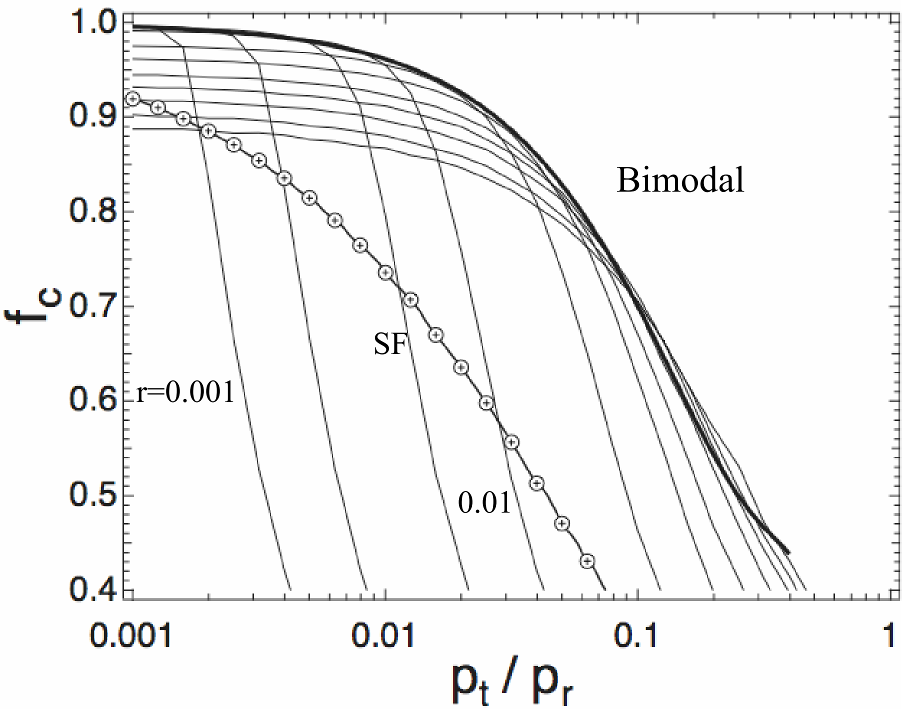
Optimal Networks

Simultaneous waves of targeted and random attacks

Bimodal: fraction of (1-r) having k_1 links and r having $k_2 = (<k> - 1 + r)/r$ links

r = 0.001—0.15 from left to right

Optimal Bimodal : $r \cong 2(p_t / p_r)$



p_t - Fraction of targeted

p_r - Fraction of random

Condition for connectivity: $\kappa = \dfrac{<k^2>}{<k>} \geq 2$

P(k) changes:

$$P(k) = \sum_{k_0}^{K} P_0(k) \binom{k_0}{k} (1 - p_r)^k p_r^{k_0 - k}$$

P(k) changes also due to targeted attack

For $p_t, p_r << 1$, p_t / p_r is the only parameter

f_c - critical threshold

Optimal Bimodal

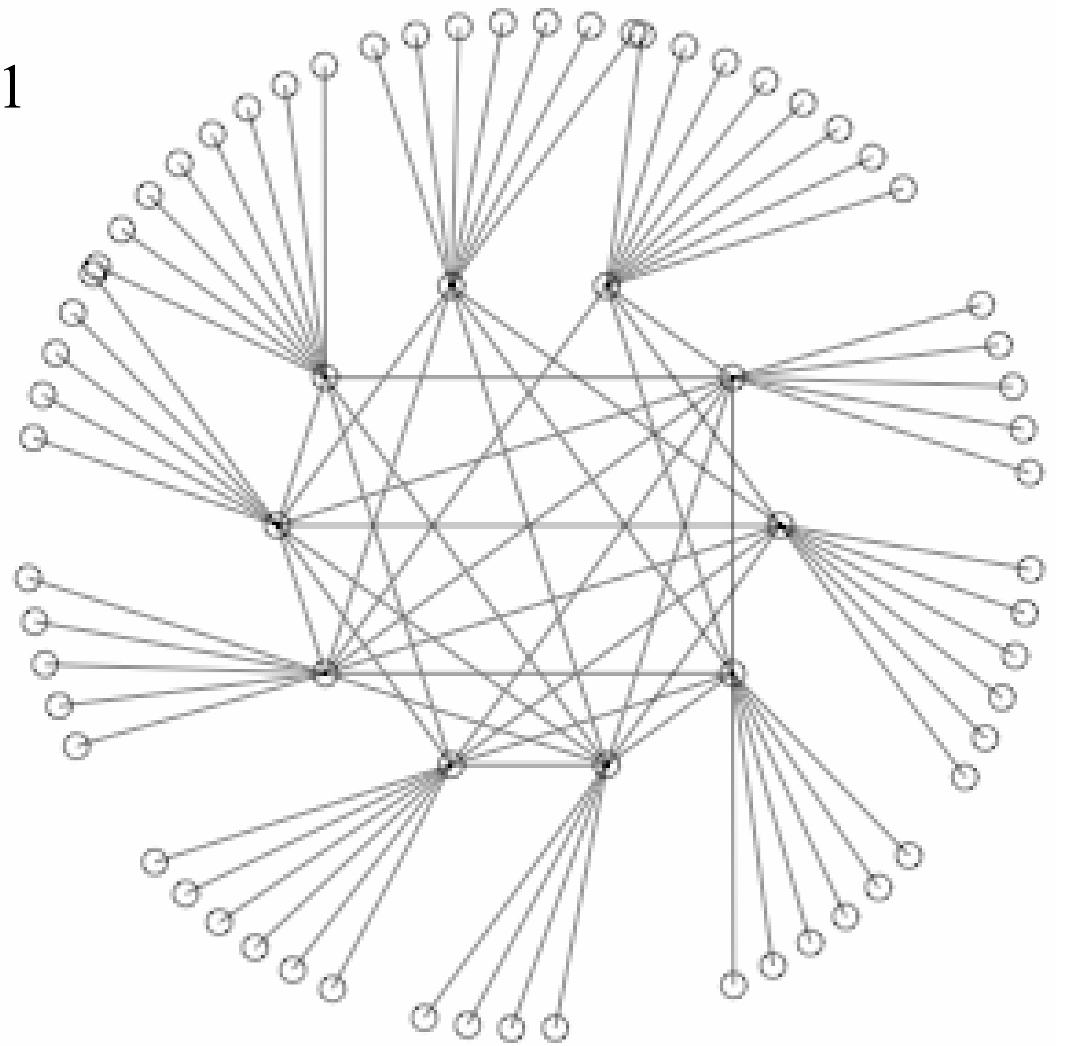
Specific example:

Given: N=100, $<k>=2.1$, $k_1=1$

and $p_t / p_r = 0.05 \rightarrow r = 0.1$

i.e, 10 "hubs" of degree $k_2 = 12$

using $k_2 = (<k> - 1 + r)/r$



Paul et al. Europhys. J. B 38, 187 (2004), (cond-mat/0404331)

Tanizawa et al. Optimization of Network Robustness to Waves of Targeted and Random Attacks (Cond-mat/0406567, Phys. Rev. E (2005)

Conclusions and Applications

- **Novel Condition for criticality** $\kappa_0 = 2,\ p_c = 1 - 1/(\kappa_0 - 1),\ \kappa_0 \equiv \langle k^2 \rangle / \langle k \rangle$
- **Distance in scale free networks λ<3 : d~loglogN - ultra small world, λ>3 : d~logN.**
- **Optimal distance – strong disorder – Random Graphs and WS** $l_{opt} \sim N^{1/3}$ **Large World**

  (Minimal spanning trees)

  scale free $\begin{cases} l_{opt} \sim N^{\frac{\lambda-3}{\lambda-1}} & for \quad \lambda > 3 \quad \Rightarrow \textbf{Large World} \\ l_{opt} \sim \log^{\lambda-1} N & for \quad 2 < \lambda < 3 \quad \Rightarrow \textbf{Small World} \end{cases}$

- **Transition between weak and strong disorder - in both percolation theory is important**
- **Scale Free networks (2<λ<3) are robust to random breakdown.**
- **Scale Free networks are vulnerable to attack on the highly connected nodes.**
- **Efficient immunization is possible without knowledge of topology, using Acquaintance Immunization.**
- **The critical exponents for scale-free directed and non-directed networks are different than those in exponential networks – different universality class!**
- **Large networks can have their topology optimized for maximum robustness to random breakdown and/or intentional attack.**

Conclusions and Applications

- **Generalized random graphs** $P(k) \sim k^{-\lambda},\ \lambda \geq 4$ **– Erdos-Renyi, λ<4 – novel topology.**
- **Scale free networks (2<λ<3) are robust to random breakdown.**
- **Scale free networks are vulnerable to intentional attack on the most highly connected nodes.**
- **Efficient immunization is possible without knowledge of topology, using Acquaintance Immunization.**
- **The critical exponents for scale-free (λ<4) directed and non-directed networks are different than those in exponential networks – different universality class!**
- **Moreover, the critical exponents depend on the strategy of node removal!** ***THE UNIVERSALITY CLASS DEPENDS ON THE WAY CRITICALITY IS REACHED*!**
- **Optimal path and minimal spanning trees on networks can be studied using percolation theory.**
- **Networks topology can be optimized for maximum robustness to various scenarios of failures such as random breakdown and/or intentional attack-using percolation**

Size of the Spanning Cluster

The spanning cluster size can be determined using either differential equations (Molloy & Reed, Combinatorics, Probability & Statistics, (1998)), or by the generating function method (Callaway et. al., Phys. Rev. Lett. 85, 5624 (2000)) .

Using those methods the spanning cluster size can be shown to be:

$$P_\infty = (1-p)(1-\sum_{k=0}^{\infty} P(k)u^k),$$

where u **is the smallest positive root of:**

$$u = p + (1-p)\sum_{k=0}^{\infty} \frac{kP(k)u^{k-1}}{\langle k \rangle}.$$

(Cohen et al., Phys. Rev. Lett. 86, 3682 (2001))

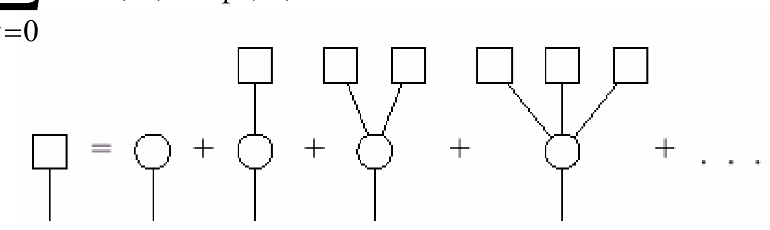
Generating Functions

(Newman et al, PRE 64, 026118 (2001))

Connectivity distribution: $G_0(x) = \sum_{k=0}^{\infty} P(k) x^k$

Probability of reaching: $G_1(x) = \frac{G_0'(x)}{\langle k \rangle} = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} k P(k) x^{k-1}$

Branch size: $H_1(x) = x G_1(H_1(x)) = \frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} k P(k) H_1(x)^{k-1}$



$p_s = s n_s$

Cluster size: $H_0(x) = \sum_{s=0}^{\infty} p_s x^s = G_0(H_1(x)) = \sum_{k=0}^{\infty} P(k) H_1(x)^k$

Infinite cluster size: $P_\infty = 1 - H_0(1)$

With Percolation: (Callaway et al, PRL 85, 5468 (2000))

$$P_\infty(q) = q\left(1 - \sum_{k=0}^{\infty} P(k) u^k\right) \qquad u = 1 - q + \frac{q}{\langle k \rangle} \sum_{k=0}^{\infty} k P(k) u^{k-1}$$

Critical exponents – Derivation of β - *Regular Graphs*

$$P_\infty \propto (q-q_c)^\beta$$

$$P_\infty(q) = q\left(1-\sum_{k=0}^{\infty} P(k)u^k\right)$$

$$q \equiv 1-p$$

$$P_\infty \sim q_c\langle k\rangle\varepsilon \sim (q-q_c)^\beta$$

$$u = 1-\varepsilon \qquad q = q_c + \delta$$

$$1-\varepsilon = 1-q_c-\delta+\frac{q_c+\delta}{\langle k\rangle}\sum_{k=0}^{\infty} k\,P(k)(1-\varepsilon)^{k-1}$$

$$\sum_{k=0}^{\infty} kP(k)u^{k-1} = \langle k\rangle - \langle k(k-1)\rangle\varepsilon + \frac{1}{2}\langle k(k-1)(k-2)\rangle\varepsilon^2+\ldots$$

$$\varepsilon \sim \frac{2\langle k(k-1)\rangle}{\langle k(k-1)(k-2)\rangle}\delta \qquad \Rightarrow \beta = 1$$

Cayley Tree:

$$P_\infty(q) = q\left(1-u^z\right)$$

$$u = 1-q+\frac{q}{z}u^{z-1}$$

$$1-\varepsilon = 1-q_c-\delta+\frac{q_c+\delta}{z}(1-\varepsilon)^{z-1}$$

$$\varepsilon \sim \frac{2}{z-2}\delta$$

$$\Rightarrow \beta = 1$$

Critical exponents – Derivation of β

$$P_\infty \propto (q-q_c)^\beta$$

$$P_\infty(q) = q\left(1-\sum_{k=0}^{\infty} P(k)u^k\right)$$

$$u = 1-q+\frac{q}{\langle k\rangle}\sum_{k=0}^{\infty} kP(k)u^{k-1}$$

$$u = 1-\varepsilon \qquad q = q_c+\delta$$

$$1-\varepsilon = 1-q_c-\delta+\frac{q_c+\delta}{\langle k\rangle}\sum_{k=0}^{\infty} k\,P(k)(1-\varepsilon)^{k-1}$$

$$P(k)\sim k^{-\alpha} \qquad\qquad q_c = \langle k\rangle / \langle k(k-1)\rangle$$

Using Abelian methods:

$$\sum_{k=0}^{\infty} kP(k)u^{k-1} \;\square\; \langle k\rangle - \langle k(k-1)\rangle\varepsilon + \frac{1}{2}\langle k(k-1)(k-2)\rangle\varepsilon^2 + \ldots + c\Gamma(2-\alpha)\varepsilon^{\alpha-2}$$

$$\varepsilon \sim \begin{cases} \left(\dfrac{\langle k(k-1)\rangle}{c\Gamma(2-\alpha)}\right)^{\frac{1}{\alpha-3}} \delta^{\frac{1}{\alpha-3}} & 3<\alpha<4, \\[2ex] \dfrac{2\langle k(k-1)\rangle}{\langle k(k-1)(k-2)\rangle}\delta & \alpha>4. \end{cases}$$

$$\beta = \begin{cases} \dfrac{1}{\alpha-3} & 3<\alpha<4 \\ 1 & \alpha>4 \end{cases}$$

For $2<\alpha<3$:

$$q_c = 0$$

$$\sum_{k=0}^{\infty} kP(k)\; u^{k-1} \sim \langle k\rangle + c\Gamma(2-\alpha)\varepsilon^{\alpha-2}$$

$$\varepsilon = \left(-\frac{c\Gamma(2-\alpha)}{\langle k\rangle}\right)^{\frac{1}{3-\alpha}} \delta^{\frac{1}{3-\alpha}}$$

$$\beta = \frac{1}{3-\alpha},\quad 2<\alpha<3$$

Critical exponents – Derivation of τ

$$n_s \sim s^{-\tau} e^{-s/s^*}$$

$$p_s = s n_s = s^{1-\tau}$$

$$H_0(x) \equiv \sum p_s x^s$$

$$H_0(x) = 1 - q + qx \sum_{k=0}^{\infty} P(k) H_1(x)^k$$

$$H_1(x) = 1 - q + qx \sum_{k=0}^{\infty} kP(k) H_1(x)^{k-1}$$

$$x = 1 - \varepsilon \qquad q = q_c$$

$$\phi(\varepsilon) \equiv 1 - H_1(1-\varepsilon)$$

Tauberian theorems:

$$\phi \sim \varepsilon^y \quad \Rightarrow \quad p_s \sim s^{-1-y}$$

$$-\phi = -q_c + (1-\varepsilon) q_c \left[1 - \frac{\phi}{q_c} + \frac{\langle k(k-1)(k-2) \rangle}{2\langle k \rangle} \phi^2 + \ldots + c \frac{\Gamma(2-\alpha)}{\langle k \rangle} \phi^{\alpha-2} \right]$$

$$\tau = 2 + y$$

$$\tau = \frac{5}{2}, \ \alpha > 4$$

$$\tau = 2 + \frac{1}{\alpha - 2} = \frac{2\alpha - 3}{\alpha - 2}, \ 2 < \alpha < 4$$

Critical exponents – Derivation of σ

$$n(s) \sim s^{-\tau} e^{-s/s^*}, s^* \sim (q - q_c)^{-\sigma}$$

Scaling relation in percolation theory:

$$\tau - 1 = d\sigma\nu$$

$$q_c(\infty) - q_c(N) \sim N^{-\frac{1}{d\nu}} = N^{-\frac{\sigma}{\tau-1}}$$

$$q_c = \frac{1}{\kappa - 1}$$

$$\kappa \approx \left(\frac{2-\alpha}{3-\alpha}\right) \frac{K^{3-\alpha} - m^{3-\alpha}}{K^{2-\alpha} - m^{2-\alpha}} \qquad K \approx N^{1/(\alpha-1)}$$

For $3 < \alpha < 4$:

$$\Delta q_c = q_c(\infty) - q_c(N) \sim \Delta\kappa \sim K^{3-\alpha} \sim N^{\frac{3-\alpha}{\alpha-1}}$$

$$\sigma = \frac{\alpha - 3}{\alpha - 2}, \; 3 < \alpha < 4$$

For $2 < \alpha < 3$:

$$q_c(\infty) = 0$$

$$q_c(N) \sim 1/\kappa(N) \sim K^{\alpha-3}$$

$$\sigma = \frac{3-\alpha}{\alpha - 2}, \; 2 < \alpha < 3$$

Percolation for Intentional Attack

Attack has two kinds of influence on the connectivity distribution:

- **Change in the upper cutoff**

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

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

- **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:

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

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

Substituting this into: $1 - \overline{p_c} = \dfrac{1}{\overline{\kappa} - 1}$**,**

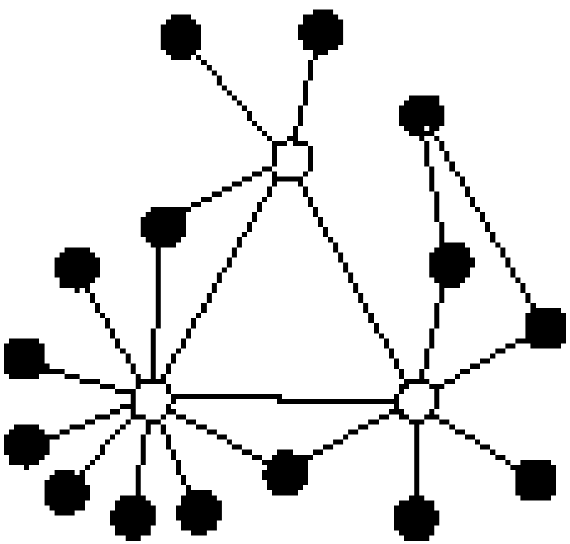
Where $\overline{\kappa} = \left(\dfrac{2-\alpha}{3-\alpha}\right) \dfrac{\overline{K}^{3-\alpha} - m^{3-\alpha}}{\overline{K}^{2-\alpha} - m^{2-\alpha}}$**, gives the critical threshold.**

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

The Model

Intentional Attack

The fraction, p, of nodes with the highest connectivity are removed.



Is this fundamentally different from random breakdown?

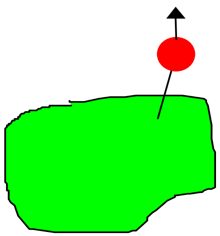
Condition for the Existence of a Spanning Cluster

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:

$$P(k_i | i \leftrightarrow j) = \frac{P(i \leftrightarrow j | k_i) P(k_i)}{P(i \leftrightarrow j)}$$

We know that

$$P(i \leftrightarrow j | k_i) = \frac{k_i}{N-1}$$

and

$$P(i \leftrightarrow j) = \frac{\langle k \rangle}{N-1}.$$

Combining all this together:

$$\kappa \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} = 2$$

(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:

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

Percolation for Random Breakdown

If percolation is considered the connectivity distribution changes according to the law:

$$\overline{P}(k) = \sum_{k'>k} P(k') \binom{k'}{k} p^{k'-k} (1-p)^k$$

Calculating the change in κ gives the percolation threshold:

$$1 - p_c = \frac{1}{\kappa_0 - 1}, \text{ where } \kappa_0 = \frac{\langle k_0^2 \rangle}{\langle k_0 \rangle}.$$

Plugging in the scale-free distribution with lower cutoff m, and upper cutoff K, gives:

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

For scale-free graphs with $\alpha \leq 3$ the second moment diverges –
No critical threshold!
Network is stable even for $p \to 1$.

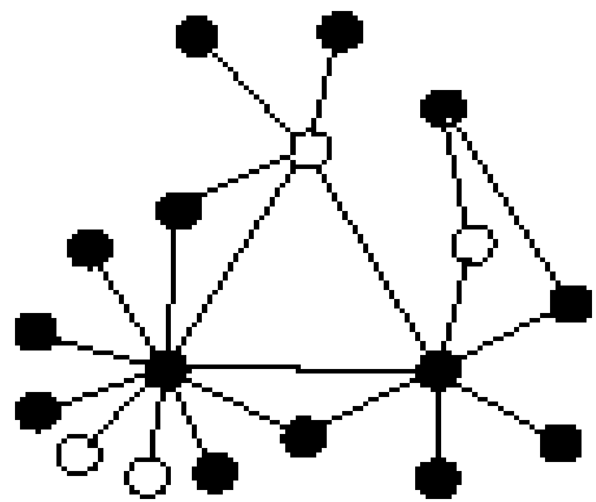
The Model

Random Breakdown

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

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

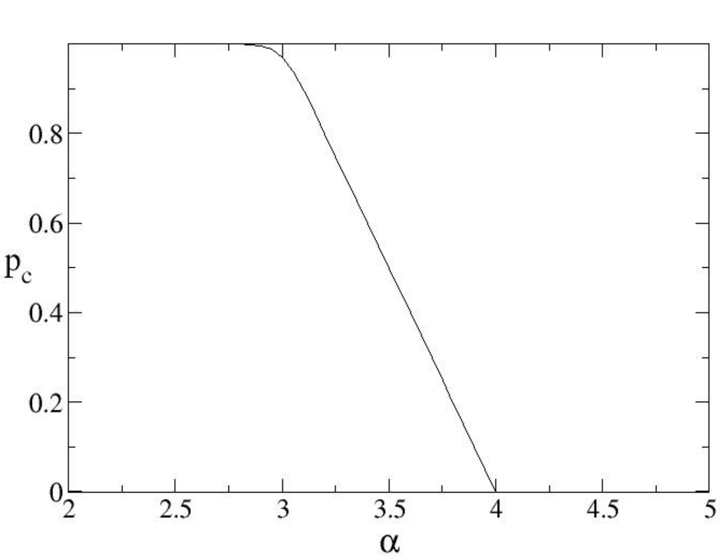
Nodes are randomly removed with probability p**.**

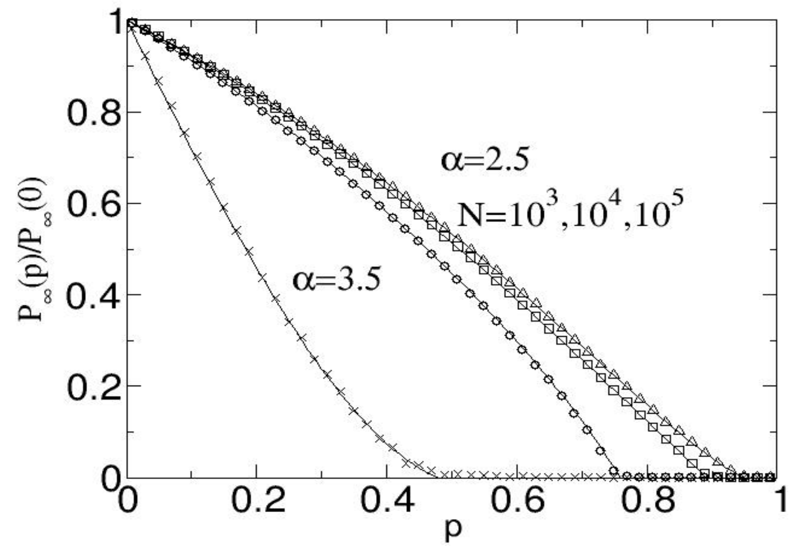


Where does a phase transition occur?

Results of Simulations and Theory

Random Breakdown

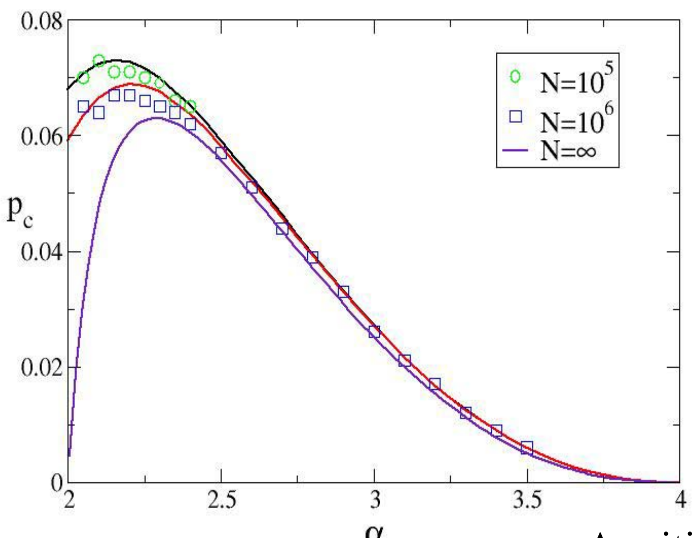


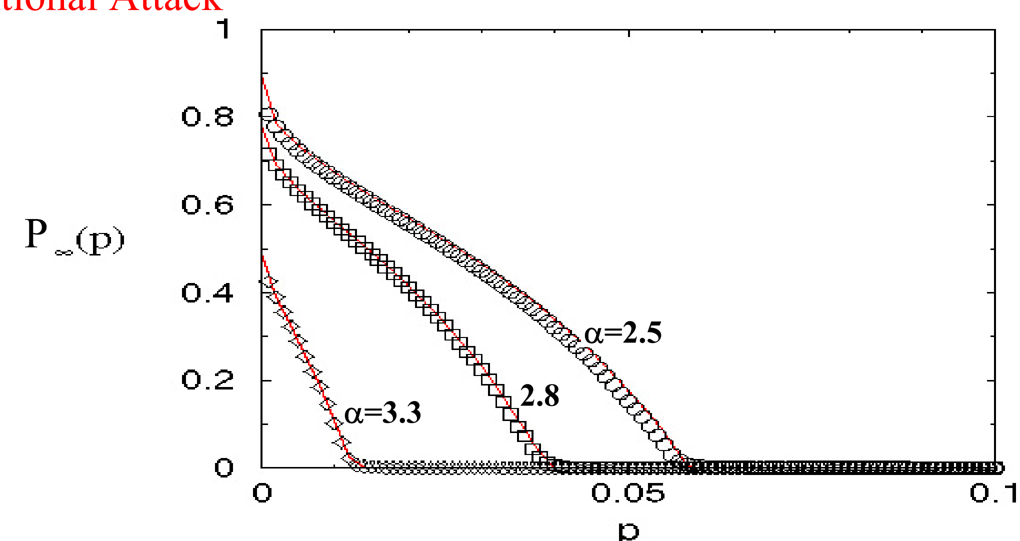


$\alpha \leq 3$ - no critical threshold – a spanning cluster always exists

$\alpha > 3$ - a critical threshold exists

Intentional Attack





A critical threshold exist for every α

Critical Exponents

Using the properties of power series near a singular point (Abelian methods), the behavior near the critical point can be studied.
The behavior near criticality in scale-free networks is different than for exponential ones!

Even for networks with $3 < \alpha < 4$, in which $\langle k \rangle$ 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:

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

Distribution of finite clusters at criticality:

$$n_s \sim s^{-\tau} \qquad \tau = \begin{cases} \dfrac{2\alpha - 3}{\alpha - 2} & \alpha < 4 \\ 2.5 & \alpha \ge 4 \end{cases}$$

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.

Chemical dimension:

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

Fractal dimension:

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

Embedding dimension:

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

The dimensionality of the graphs depends on the distribution!

Behavior of the Site to Site Distance near Criticality

In standard random graph models the average distance between sites scales as: $r \sim \log_k N$,

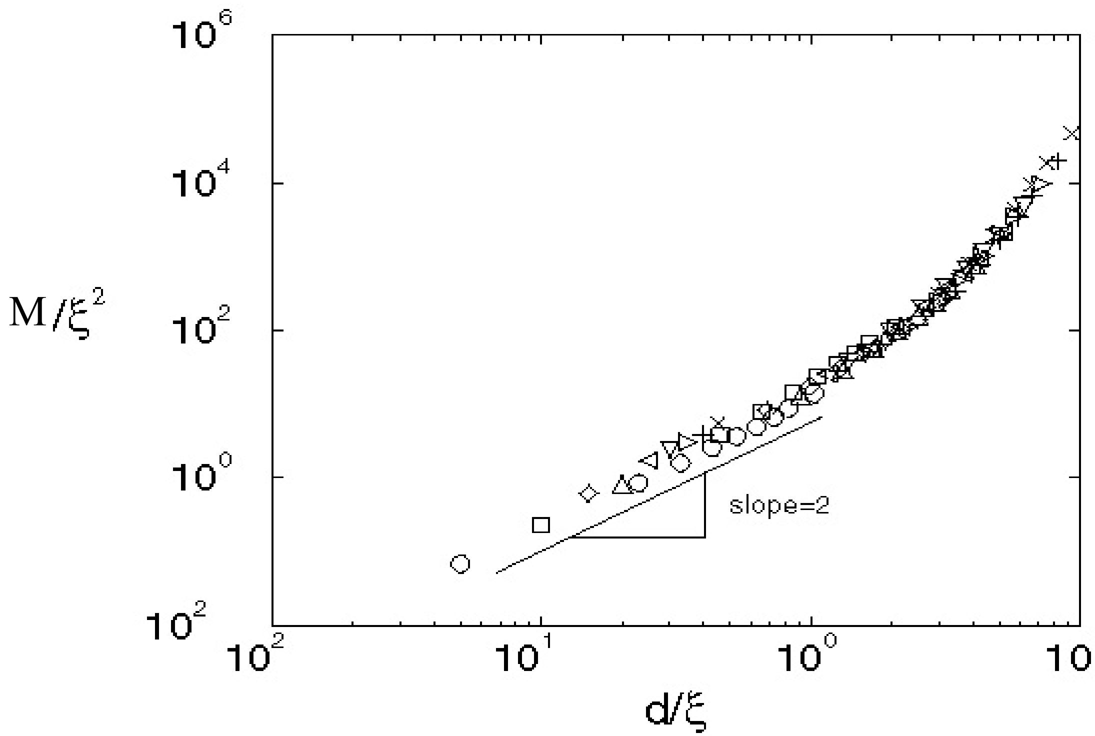
where N **is the system size.**

That is, the mass behaves as: $M \sim k^r$.

From infinite dimensional percolation theory it is known: $M \sim r^2$, or $r \sim \sqrt{M}$ at $p = p_c$.

This is also true for scale-free networks with $\alpha \geq 4$.

A crossover exists between those behaviors where the correlation length: $\xi \sim (p - p_c)^{-1}$.



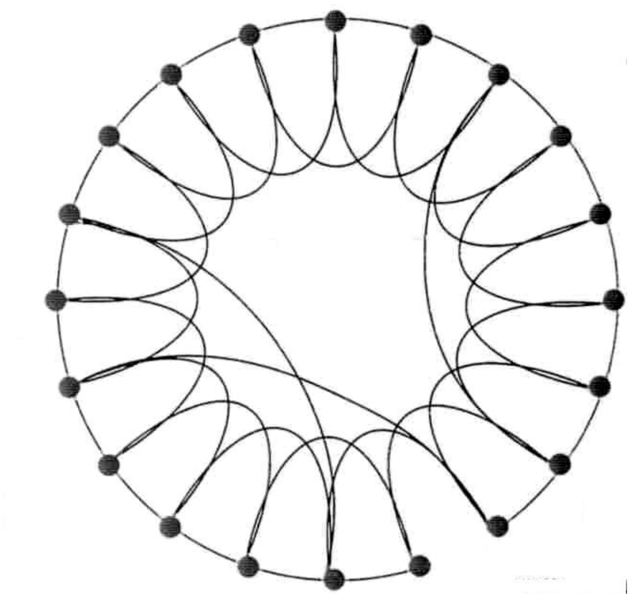
Communication becomes inefficient even before the breakdown

Conclusions and Applications

- **The Internet is resilient to random breakdown.**
- **The Internet is sensitive to intentional attack on the most highly connected nodes.**
- **Large networks can have their connectivity distribution optimized for maximum resilience to random breakdown and/or intentional attack.**
- **A virus has a finite probability of infecting a large portion of the Internet, regardless of how low is the probability of infection.**
- **However, if a finite fraction of the most highly connected routers of the Internet block the virus, it cannot infect a finite portion of the Internet.**
- **Even before breakdown the diameter of the spanning cluster becomes large – making communication inefficient.**
- **The critical exponents for scale-free networks are different than those in exponential networks – different universality class!**

Small World

A small world network is a regular lattice with added random links.

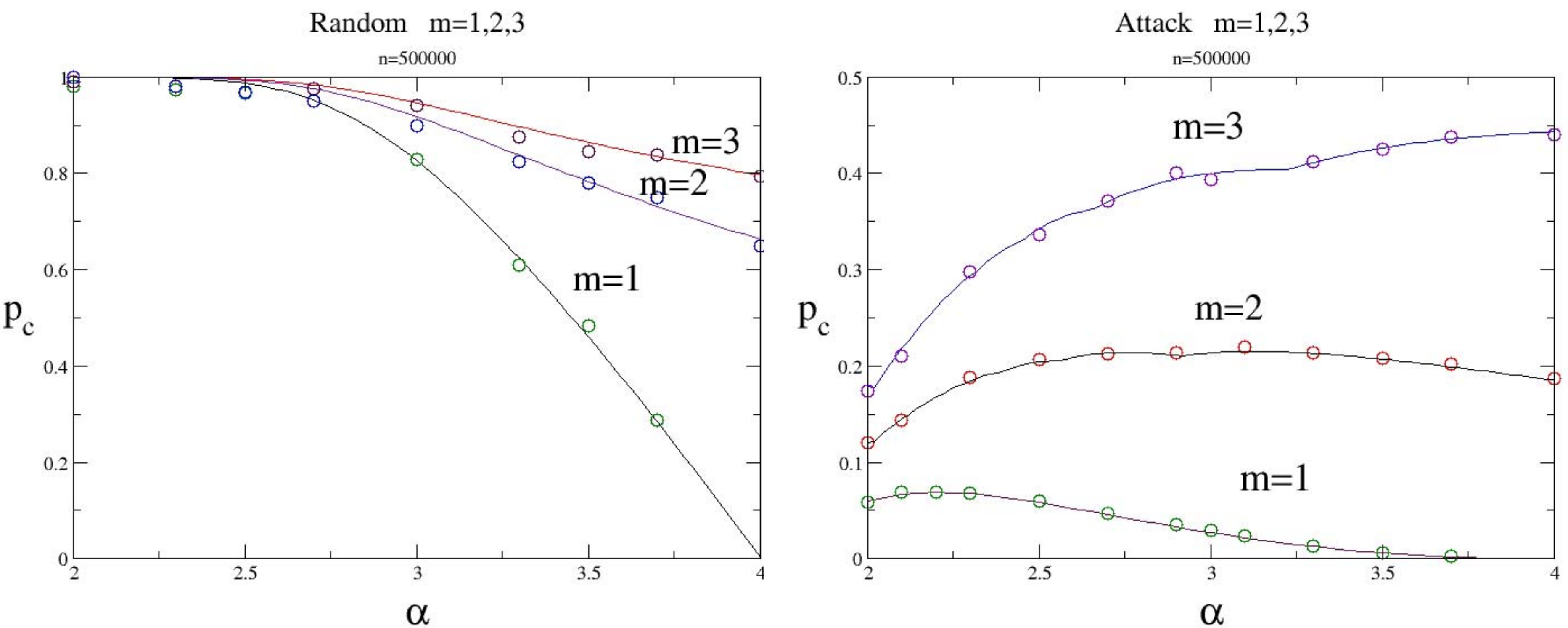


Examples:

- Movie actors
- Polymer chains configuration space
- Acquaintance networks
- Neural networks

Exponents are the same as in mean-field percolation.

Results of Simulations and Theory



Experimental Data: Virus survival

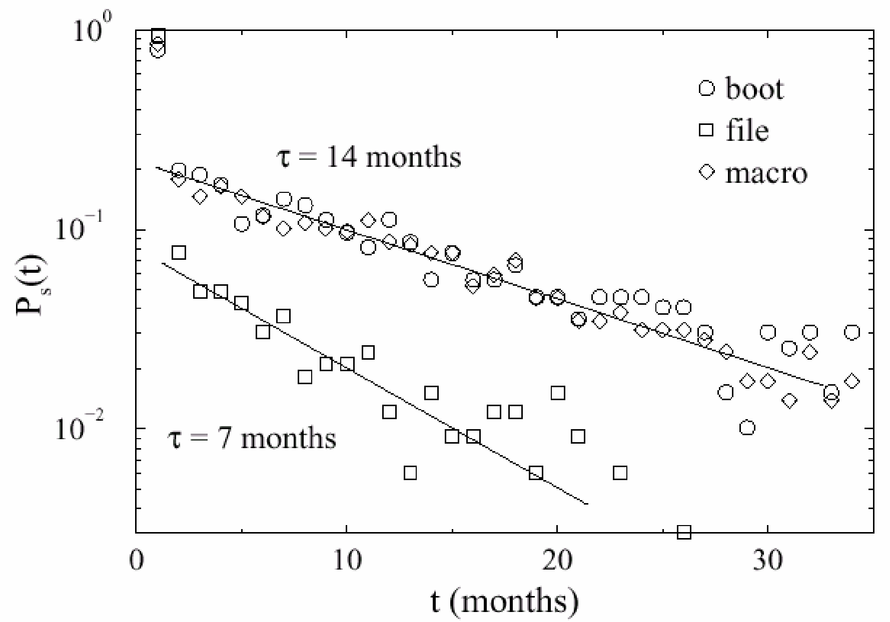
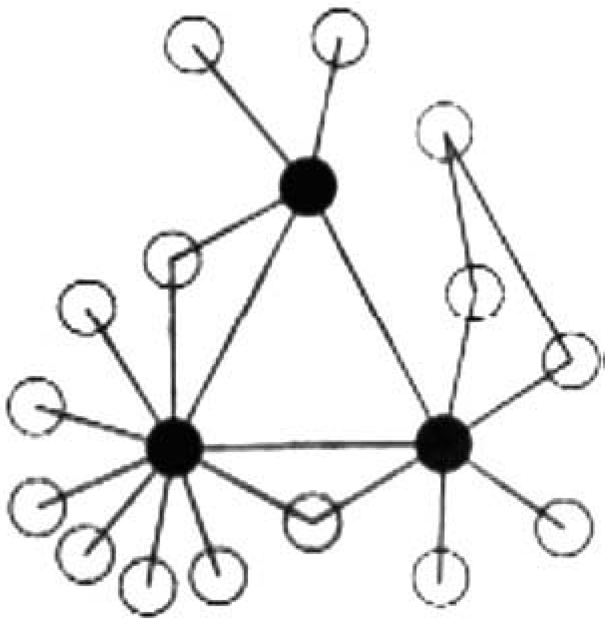


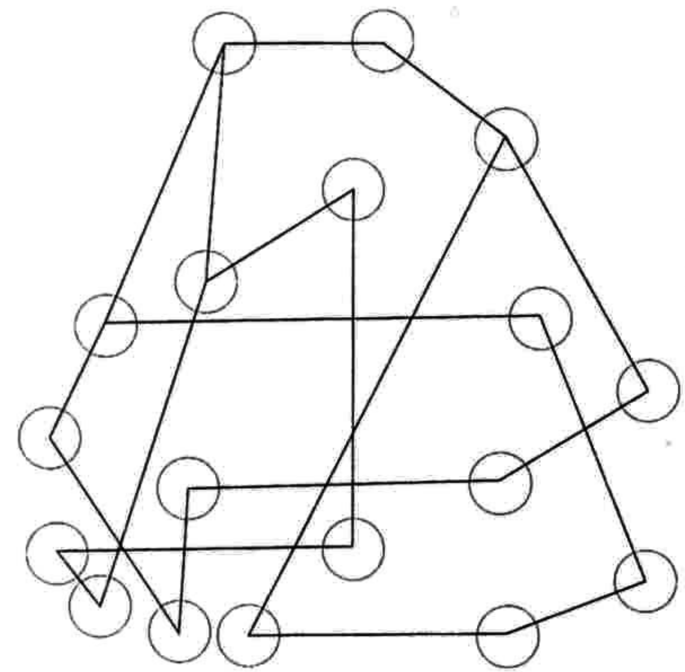
FIG. 1. Surviving probability for viruses in the wild. The 814 different viruses analyzed have been grouped in three main strains [9]: file viruses infect a computer when running an infected application; boot viruses also spread via infected applications, but copy themselves into the boot sector of the 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))

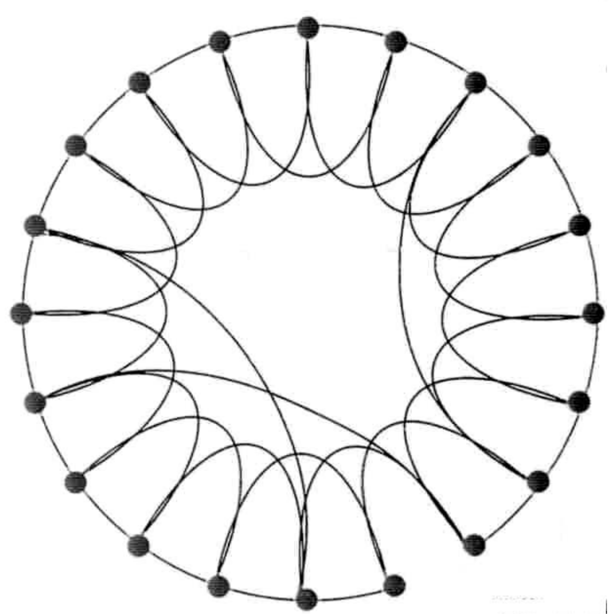
Scale Free



Random Graph

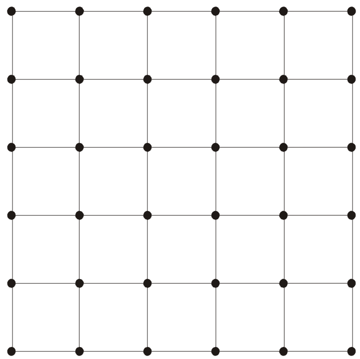


Small World

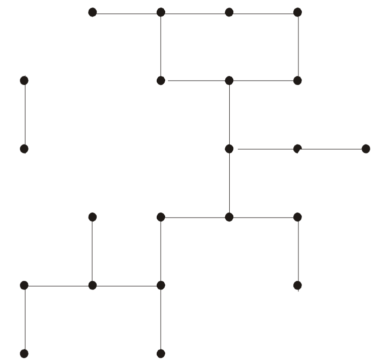


Percolation and Immunization

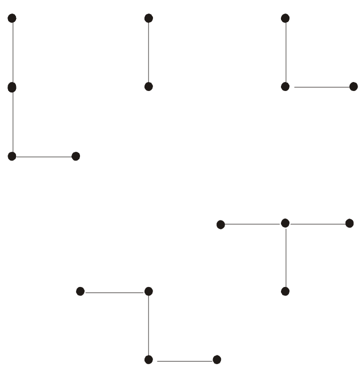
Percolation theory



$p = 0.3$
remove or immune

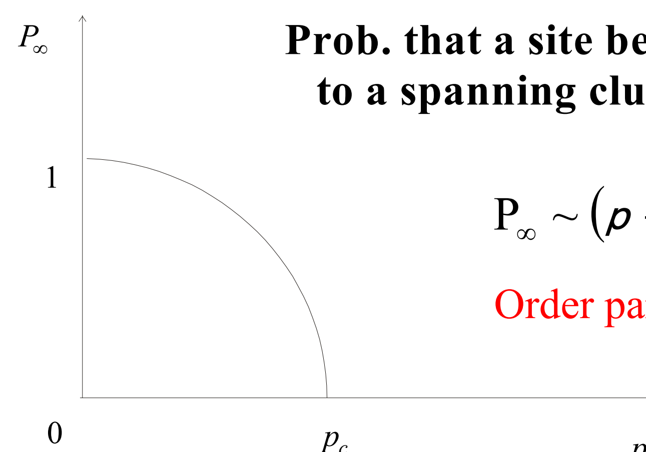


Network exists



$p = 0.5$
remove or immune

Network collapse



Prob. that a site belongs to a spanning cluster

$$P_\infty \sim (p - p_c)^\beta$$

Order parameter

Size of the Spanning Cluster

The spanning cluster size can be determined using either differential equations (Molloy & Reed, Combinatorics, Probability & Statistics, (1998)), or by the generating function method (Callaway et. al., Phys. Rev. Lett. 85, 5624 (2000)) .

If After time t, tN links have been followed, the change in the number of unexposed nodes of connectivity k is: $\frac{dP(k,t)}{dt} = -\frac{kP(k,t)}{\langle k \rangle - 2t - 1}$.

The number of open links in the cluster is: $X(t) = \langle k \rangle - 2t - \sum kP(k,t)$**. When the entire cluster has been exposed.**

Solving this gives the cluster size:

$$P_\infty = (1-p)(1-\sum_{k=0}^{\infty} P(k)u^k),$$

where u is the smallest positive root of:

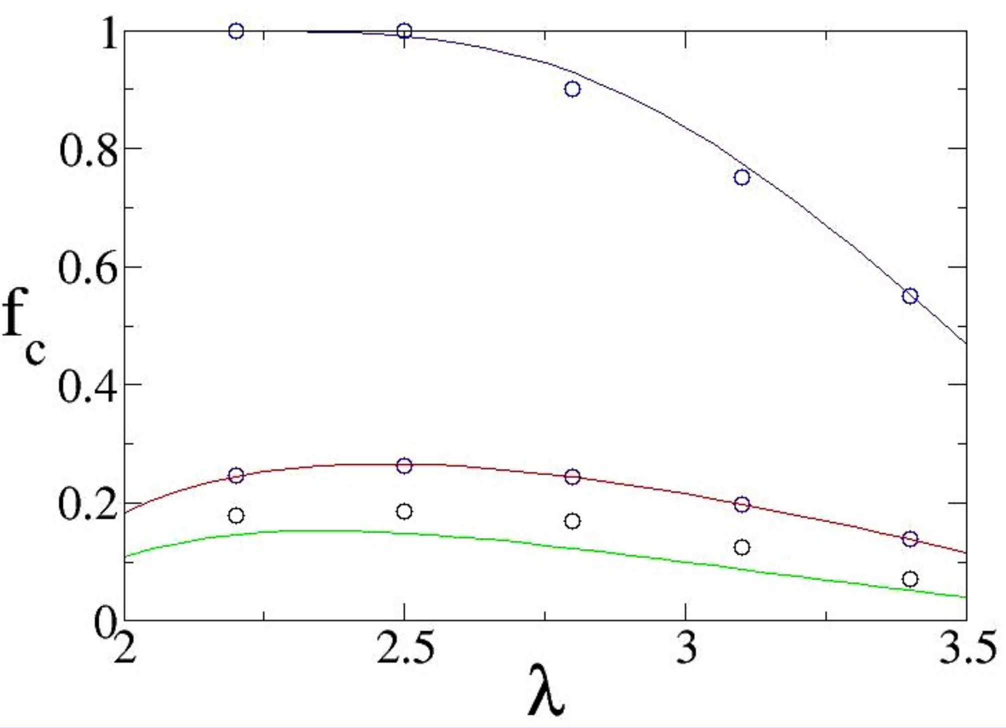
$$u = p + (1-p)\sum_{k=0}^{\infty} \frac{kP(k)u^{k-1}}{\langle k \rangle}$$

(Cohen et al., Phys. Rev. Lett. 86, 3682 (2001))

Efficient Immunization Strategies:

Acquaintance Immunization

Critical Threshold
Scale Free



- Random immunization is inefficient in scale free graphs, while targeted immunization requires knowledge of the degrees.
- In *Acquaintance Immunization* one immunizes random neighbors of random individuals.
- One can also do the same based on *n* neighbors.
- The threshold is finite and no global knowledge is necessary.

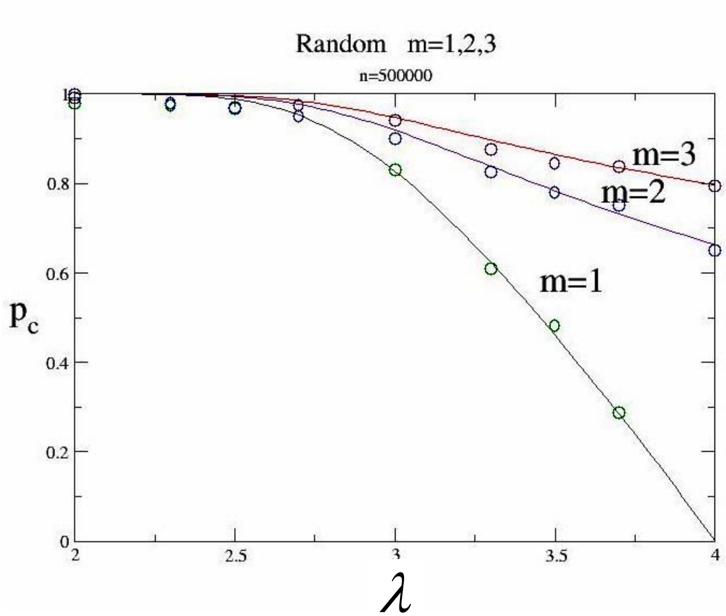
$$\sum_k P(k)(k-1)\left(k/\langle k\rangle\right) v_{p_c}^{k-2} e^{-2p_c/k} = 1$$

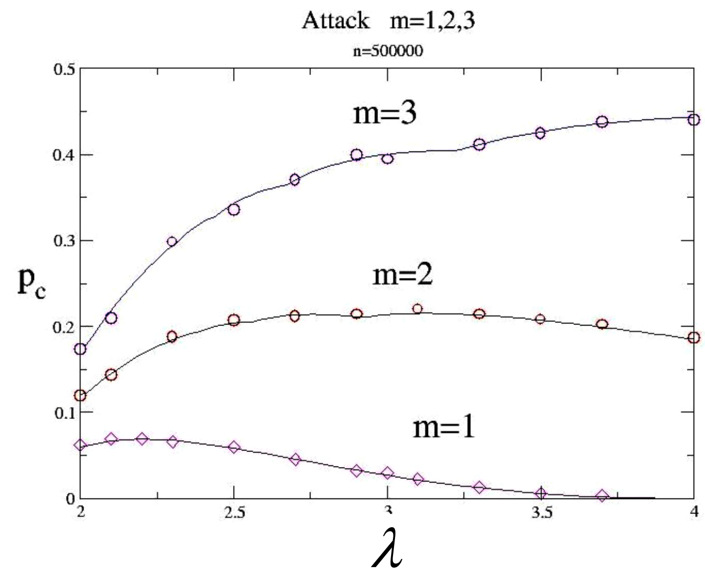
$$f_c = 1 - \sum_k P(k) v_{p_c}^k$$

$$v_p = \sum_k kP(k)\exp(-p/k)/\langle k\rangle$$

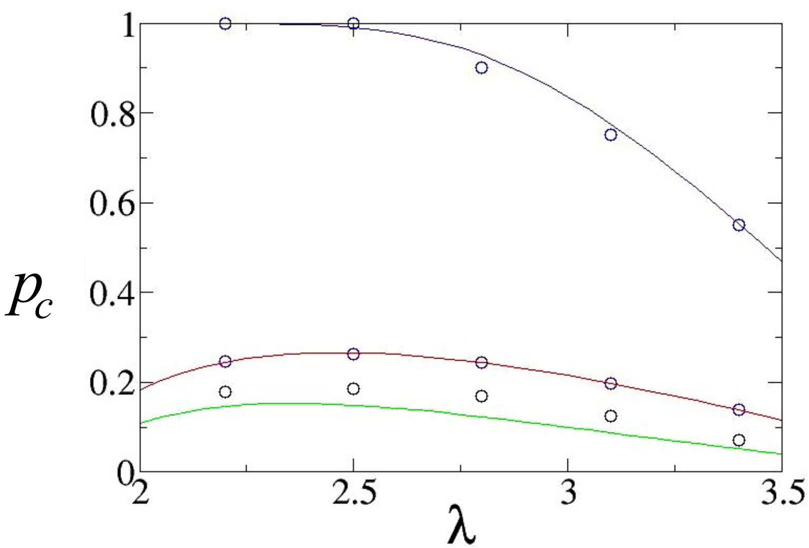
Cohen et al cond-mat/0207387

Results of Simulations and Theory





Efficient Immunization Strategy



- Random immunization is inefficient in scale free graphs, while targeted immunization requires knowledge of the degrees.

- In *Acquaintance Immunization* one immunizes random neighbors of random individuals.