Network Science

Class 4: Scale-free property

Albert-László Barabási
with Emma K. Towlson, Michael M. Danziger, Sebastian Ruf and Louis Shekhtman

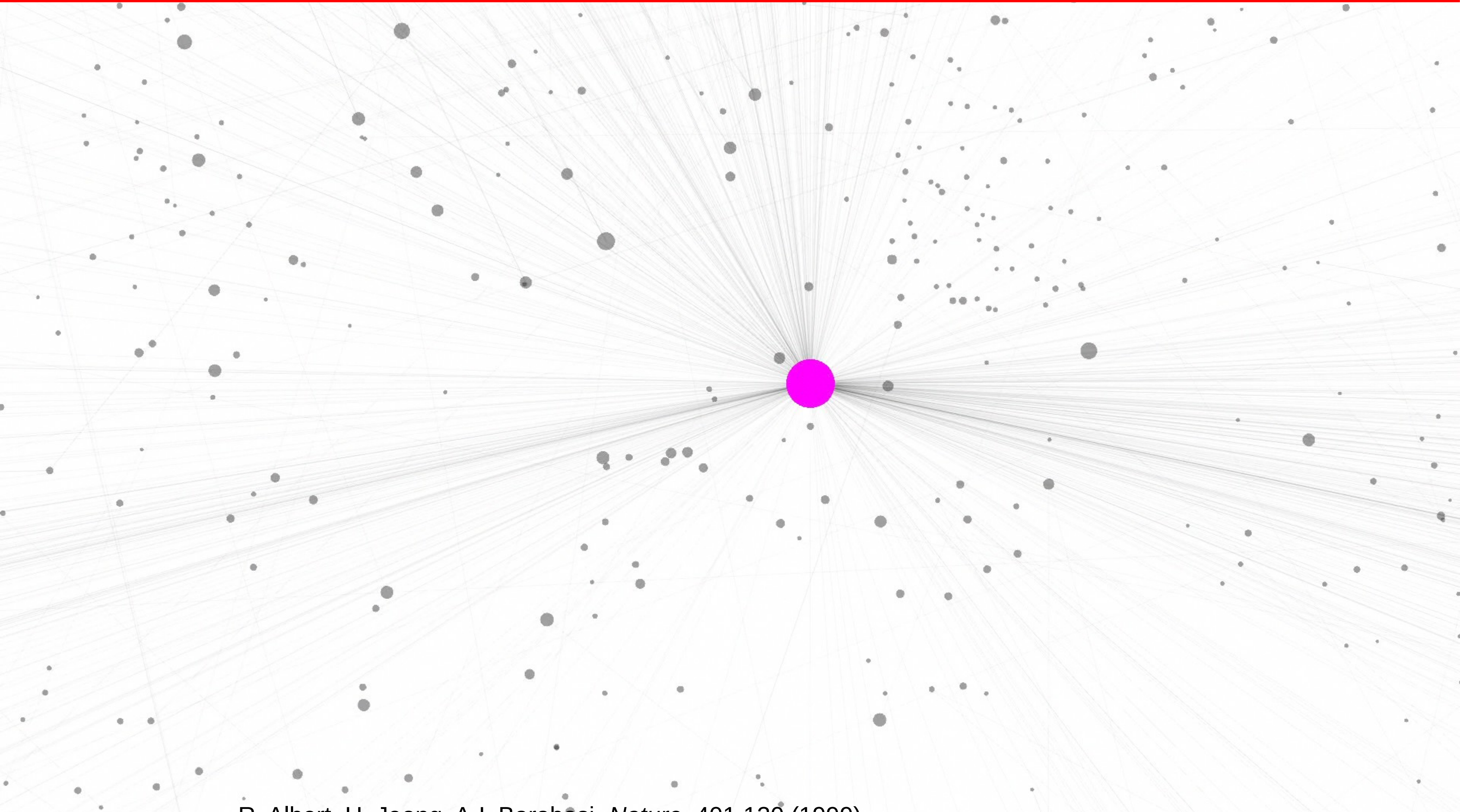
www.BarabasiLab.com

Questions | Scale-free Property

1. From the WWW to Scale-free networks. Definition.
2. Discrete and continuum formalism. Explain its meaning.
3. Hubs and the maximum degree.
4. What does ‘scale-free’ mean?
5. Universality. Are all networks scale-free?
6. From small worlds to ultra small worlds.
7. The role of the degree exponent.

Introduction

WORLD WIDE WEB



R. Albert, H. Jeong, A-L Barabasi, *Nature*, 401 130 (1999).

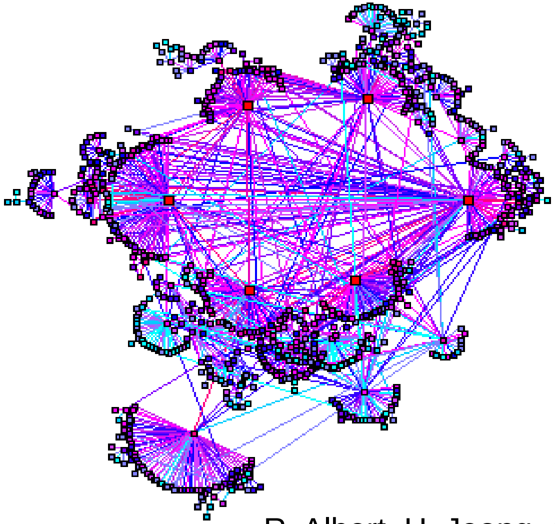
Power laws and scale-free networks

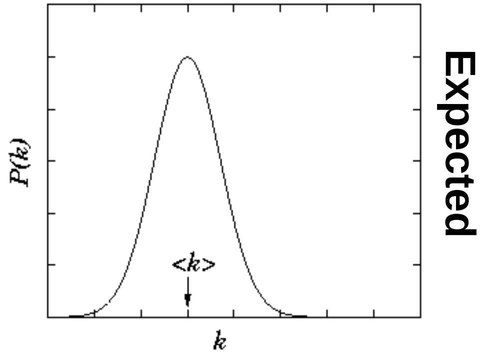
WORLD WIDE WEB

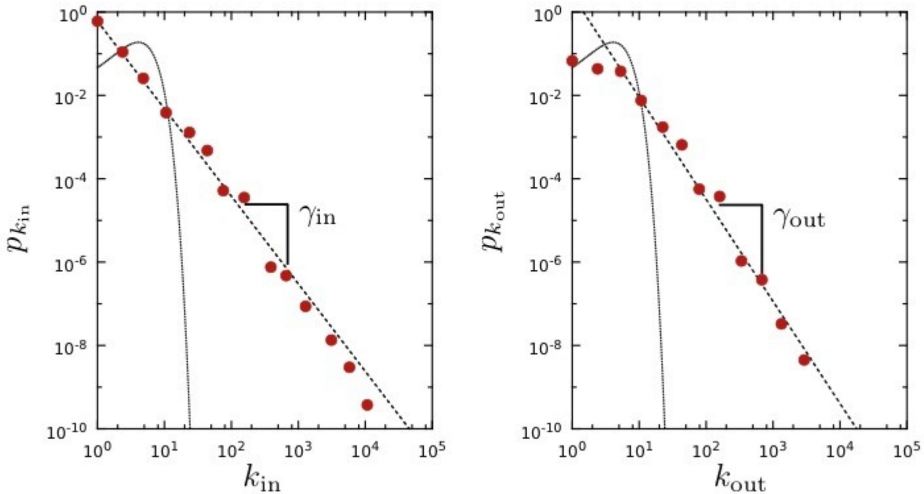
Nodes: **WWW documents**
Links: **URL links**

Over 3 billion documents

ROBOT: collects all URL's found in a document and follows them recursively







R. Albert, H. Jeong, A-L Barabasi, *Nature*, 401 130 (1999).

Discrete vs. Continuum formalism

Discrete Formalism

As node degrees are always positive integers, the discrete formalism captures the probability that a node has exactly k links:

$$p_k = Ck^{-\gamma}.$$

$$\sum_{k=1}^{\infty} p_k = 1.$$

$$C\sum_{k=1}^{\infty} k^{-\gamma} = 1 \qquad C = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)},$$

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$$

Continuum Formalism

In analytical calculations it is often convenient to assume that the degrees can take up any positive real value:

$$p(k) = Ck^{-\gamma}.$$

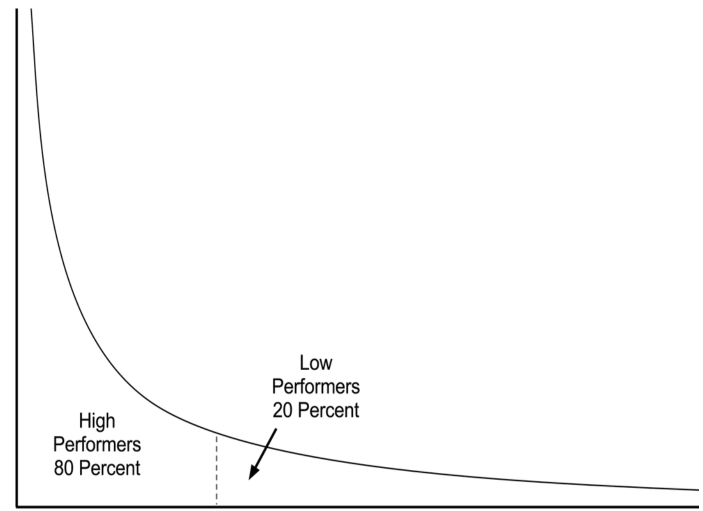
$$\int_{k_{\min}}^{\infty} p(k)dk = 1$$

$$C = \frac{1}{\int_{k_{\min}}^{\infty} k^{-\gamma}dk} = (\gamma - 1)k_{\min}^{\gamma-1}$$

$$p(k) = (\gamma - 1)k_{\min}^{\gamma-1}k^{-\gamma}.$$

INTERPRETATION: p_k $\int_{k_1}^{k_2} p(k)dk$

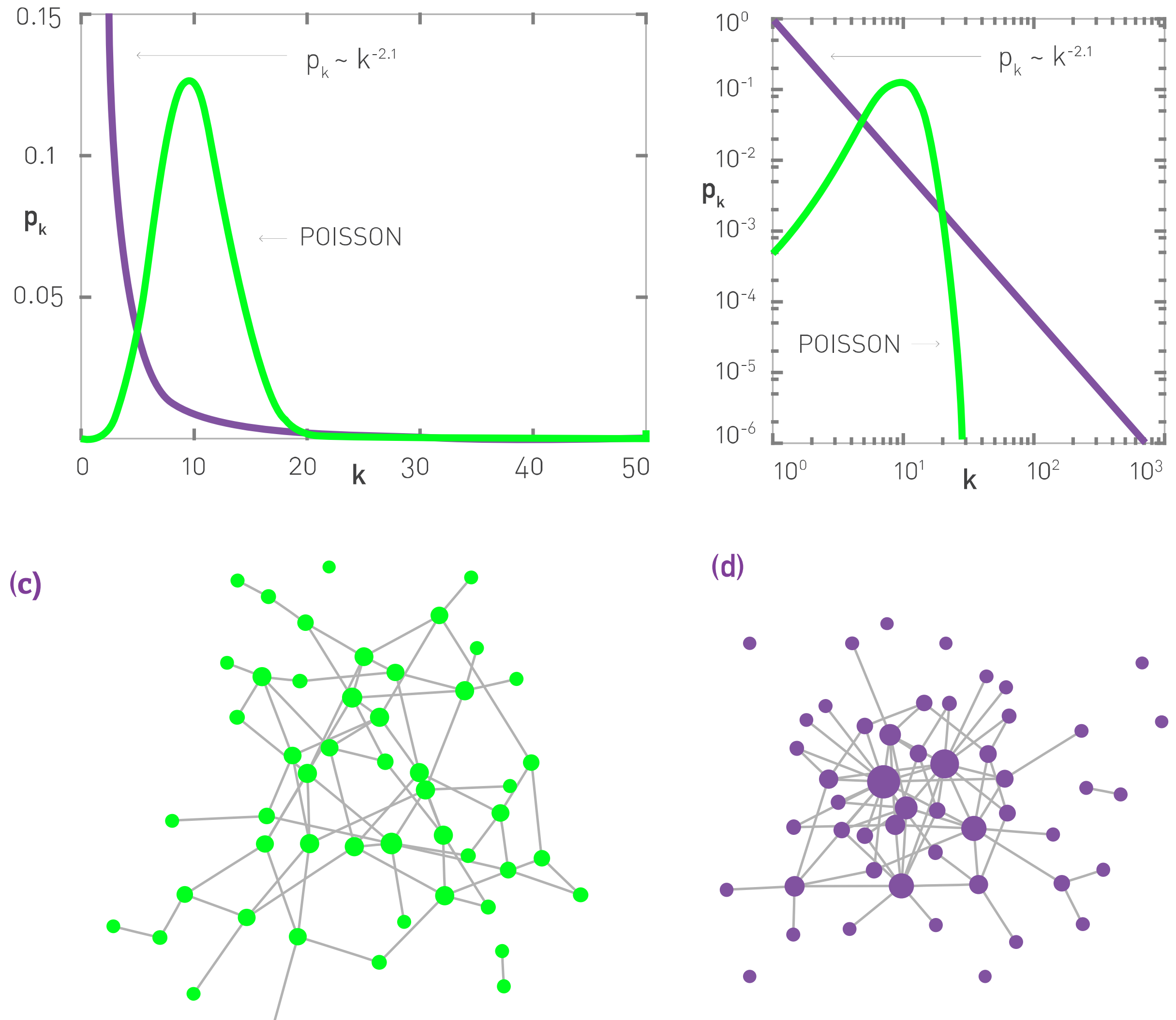
80/20 RULE



Vilfredo Federico Damaso Pareto (1848 – 1923), Italian economist, political scientist and philosopher, who had important contributions to our understanding of income distribution and to the analysis of individuals choices. A number of fundamental principles are named after him, like Pareto efficiency, Pareto distribution (another name for a power-law distribution), the Pareto principle (or 80/20 law).

Hubs

The difference between a power law and an exponential distribution



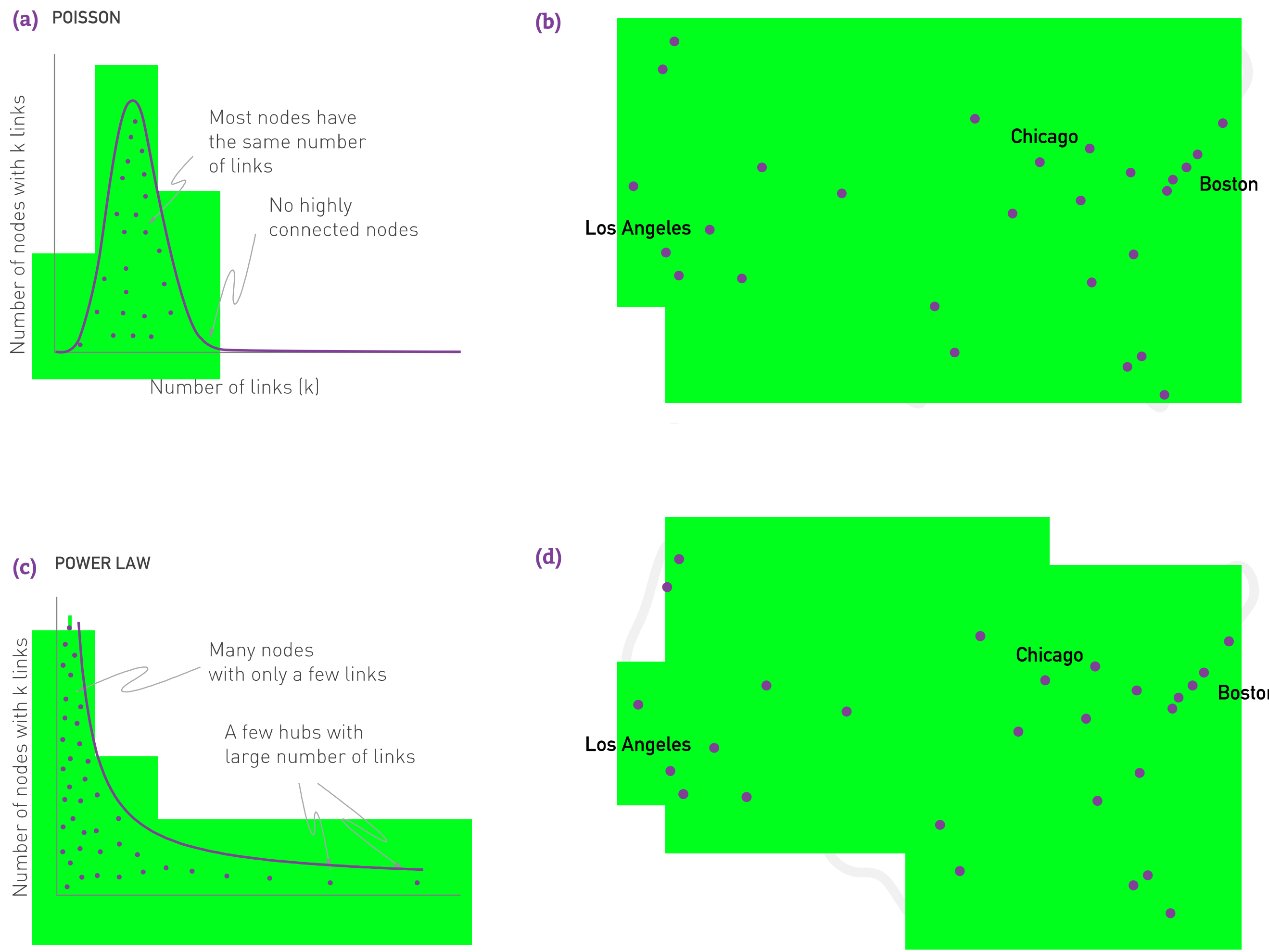
The difference between a power law and an exponential distribution

Let us use the WWW to illustrate the properties of the high-*k regime.*
The probability to have a node with k~100 is

- *About* $p_{100} \simeq 10^{-30}$ *in a Poisson distribution*
- *About* $p_{100} \simeq 10^{-4}$ *if* p_k *follows a power law.*
- *Consequently, if the WWW were to be a random network, according to the Poisson prediction we would expect* 10^{-18} *k>100 degree nodes, or none.*

$$N_{k>100} = 10^9$$

- *For a power law degree distribution, we expect about k>100 degree nodes*


(a) POISSON
Number of nodes with k links
Most nodes have the same number of links
No highly connected nodes
Number of links (k)
(b)
Chicago
Boston
Los Angeles
(c) POWER LAW
Number of nodes with k links
Many nodes with only a few links
A few hubs with large number of links
(d)
Chicago
Boston
Los Angeles

The size of the biggest hub

All real networks are finite → let us explore its consequences.
→ We have an expected maximum degree, k_{max}

Estimating k_{max}

$$\int_{k_{\max}}^{\infty} P(k)\,dk \approx \frac{1}{N}$$

Why: the probability to have a node larger than k_{max} should not exceed the prob. to have one node, i.e. *1/N* fraction of all nodes

$$\int_{k_{\max}}^{\infty} P(k)\,dk = (\gamma-1)k_{\min}^{\gamma-1}\int_{k_{\max}}^{\infty} k^{-\gamma}\,dk = \frac{(\gamma-1)}{(-\gamma+1)}k_{\min}^{\gamma-1}\left[k^{-\gamma+1}\right]_{k_{\max}}^{\infty} = \frac{k_{\min}^{\gamma-1}}{k_{\max}^{\gamma-1}} \approx \frac{1}{N}$$

$$k_{\max} = k_{\min} N^{\frac{1}{\gamma-1}}$$

The size of the biggest hub

$$k_{\max} = k_{\min} N^{\frac{1}{\gamma - 1}}$$

To illustrate the difference in the maximum degree of an exponential and a scale-free network let us return to the WWW sample of Figure 4.1, consisting of $N \simeq 3 \times 10^5$ nodes. As $k_{min} = 1$, if the degree distribution were to follow an exponential, (4.17) predicts that the maximum degree should be $k_{max} \simeq 13$. In a scale-free network of similar size and $\gamma = 2.1$, (4.18) predicts $k_{max} \simeq 85{,}000$, a remarkable difference. Note that the largest in-degree of the WWW map of Figure 4.1 is 10,721, which is comparable to k_{max} predicted by a scale-free network. This reinforces our conclusion that *in a random network hubs are effectivelly forbidden, while in scale-free networks they are naturally present.*

Finite scale-free networks

Expected maximum degree, k_{max}

$$k_{\max} = k_{\min} N^{\frac{1}{\gamma-1}}$$

- k_{max}, increases with the size of the network
  → the larger a system is, the larger its biggest hub

  - For $\gamma>2$ k_{max} increases slower than N

→ the largest hub will contain a decreasing fraction of links as N increases.

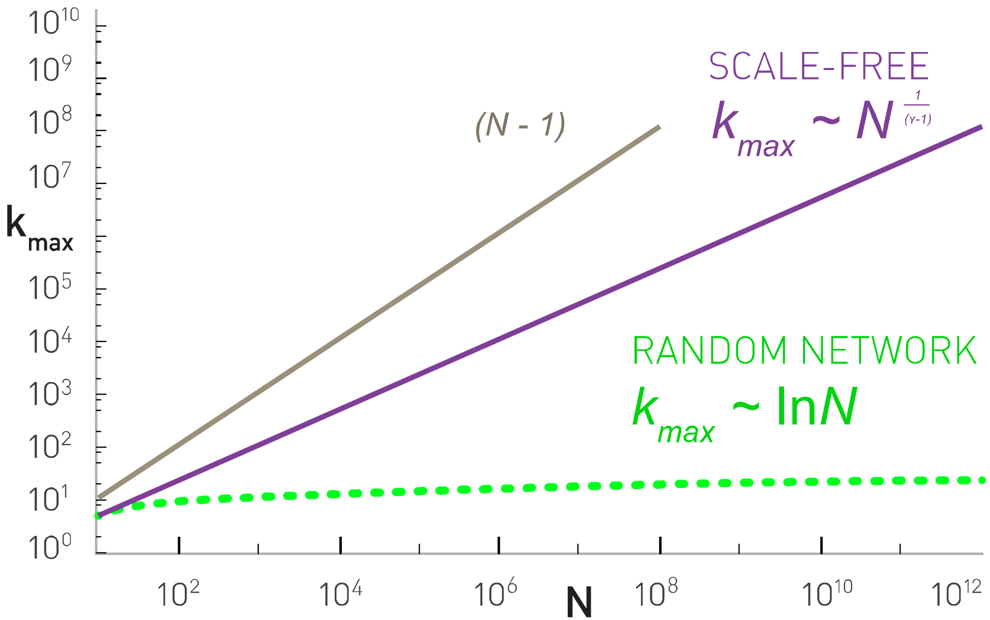
  - For $\gamma=2$ $k_{max} \sim N$.

  → The size of the biggest hub is O(N)

  - For $\gamma<2$ k_{max} increases faster than N: condensation phenomena

  → the largest hub will grab an increasing fraction of links. Anomaly!

The size of the largest hub



$$k_{\max} = k_{\min} N^{\frac{1}{\gamma - 1}}$$

The meaning of scale-free

Scale-free networks: Definition

Definition:

Networks with a power law tail in their degree distribution are called 'scale-free networks'

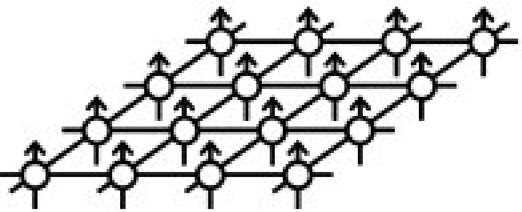
Where does the name come from?

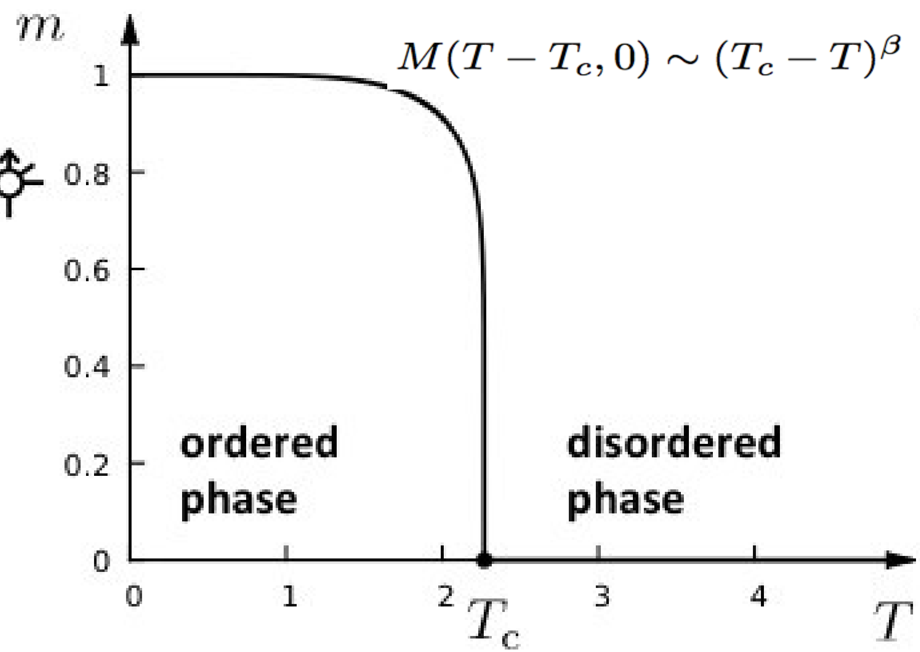
Critical Phenomena and scale-invariance

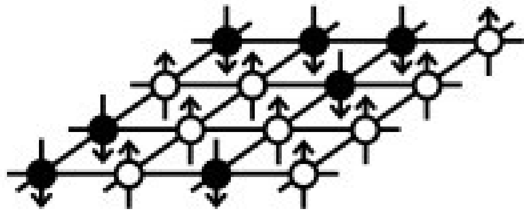
(a detour)

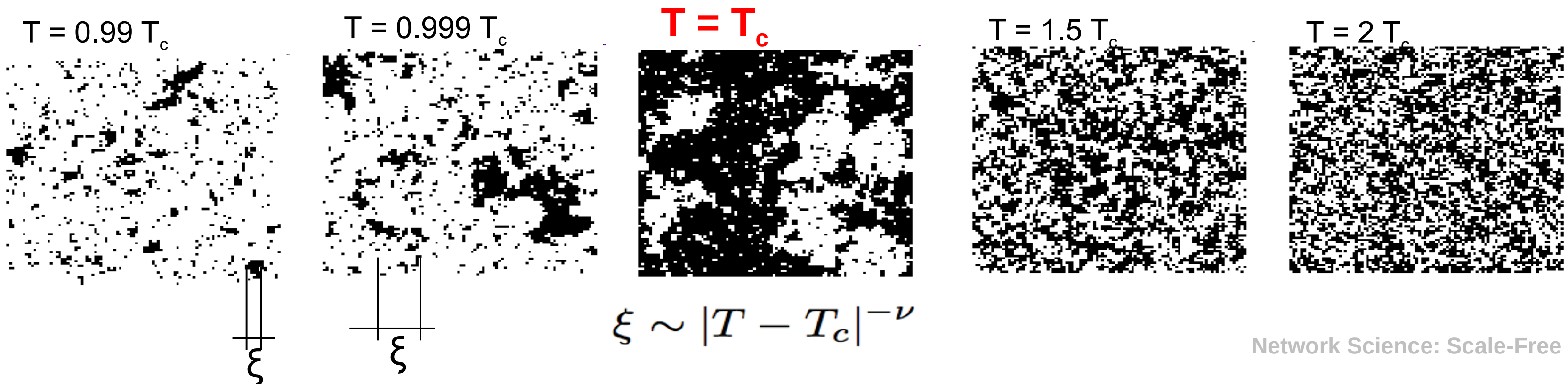
Slides after **Dante R. Chialvo**

Phase transitions in complex systems I: Magnetism









Scale-free behavior in space

$$\xi \sim |T - T_c|^{-\nu}$$

At T = Tc:

correlation length diverges

Fluctuations emerge at all scales:

scale-free behavior

CRITICAL PHENOMENA

- Correlation length diverges at the critical point: the whole system is correlated!
- **Scale invariance**: there is no characteristic scale for the fluctuation **(scale-free behavior).**
- **Universality**: exponents are independent of the system's details.

Divergences in scale-free distributions

$$P(k) = Ck^{-\gamma} \qquad k = [k_{\min}, \infty) \qquad \int_{k_{\min}}^{\infty} P(k) dk = 1 \qquad C = \frac{1}{\int_{k_{\min}}^{\infty} k^{-\gamma}\, dk} = (\gamma - 1) k_{\min}^{\gamma - 1}$$

$$P(k) = (\gamma - 1) k_{\min}^{\gamma - 1} k^{-\gamma}$$

$$< k^m >= \int_{k_{\min}}^{\infty} k^m P(k) dk \qquad < k^m >= (\gamma - 1) k_{\min}^{\gamma - 1} \int_{k_{\min}}^{\infty} k^{m-\gamma}\, dk = \frac{(\gamma - 1)}{(m - \gamma + 1)} k_{\min}^{\gamma - 1} \left[k^{m-\gamma+1} \right]_{k_{\min}}^{\infty}$$

If m-γ+1<0: $$< k^m >= -\frac{(\gamma - 1)}{(m - \gamma + 1)} k_{\min}^{m}$$

If m-γ+1>0, the integral diverges.

For a fixed γ this means that all moments with m>γ-1 diverge.

DIVERGENCE OF THE HIGHER MOMENTS

$$< k^m > = (\gamma - 1) k_{\min}^{\gamma - 1} \int_{k_{\min}}^{\infty} k^{m-\lambda}\, dk = \frac{(\gamma - 1)}{(m - \gamma + 1)} k_{\min}^{\gamma - 1} \left[k^{m-\gamma+1} \right]_{k_{\min}}^{\infty}$$

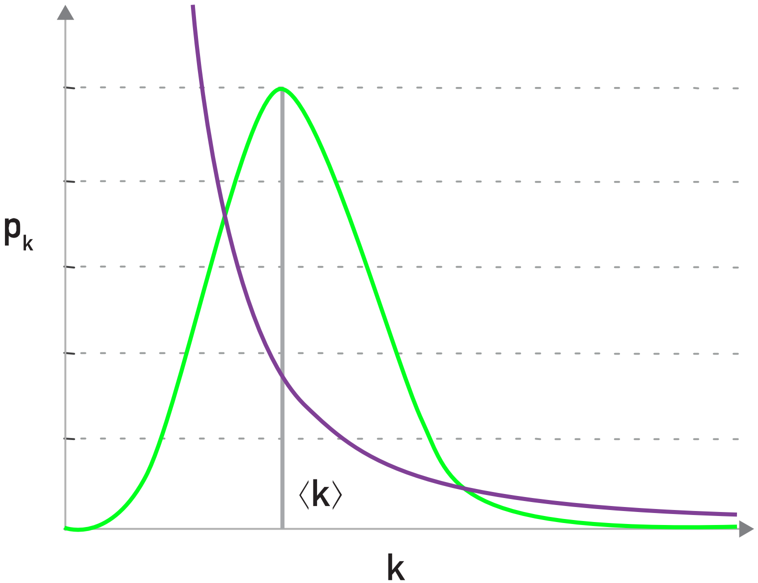
For a fixed λ this means all moments m>γ-1 diverge.

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}
WWW	325 729	4.51	900	2.45	2.1
WWW	4×10^7	7		2.38	2.1
WWW	2×10^8	7.5	4000	2.72	2.1
WWW, site	260 000				1.94
Internet, domain*	3015–4389	3.42–3.76	30–40	2.1–2.2	2.1–2.2
Internet, router*	3888	2.57	30	2.48	2.48
Internet, router*	150 000	2.66	60	2.4	2.4
Movie actors*	212 250	28.78	900	2.3	2.3
Co-authors, SPIRES*	56 627	173	1100	1.2	1.2
Co-authors, neuro.*	209 293	11.54	400	2.1	2.1
Co-authors, math.*	70 975	3.9	120	2.5	2.5
Sexual contacts*	2810			3.4	3.4
Metabolic, *E. coli*	778	7.4	110	2.2	2.2
Protein, *S. cerev.**	1870	2.39		2.4	2.4
Ythan estuary*	134	8.7	35	1.05	1.05
Silwood Park*	154	4.75	27	1.13	1.13
Citation	783 339	8.57			3
Phone call	53×10^6	3.16		2.1	2.1
Words, co-occurrence*	460 902	70.13		2.7	2.7
Words, synonyms*	22 311	13.48		2.8	2.8

Many degree exponents are smaller than 3

→ $<k^2>$ diverges in the N→∞ limit!!!

The meaning of scale-free



Random Network

Randomly chosen node: $k = \langle k \rangle \pm \langle k \rangle^{1/2}$

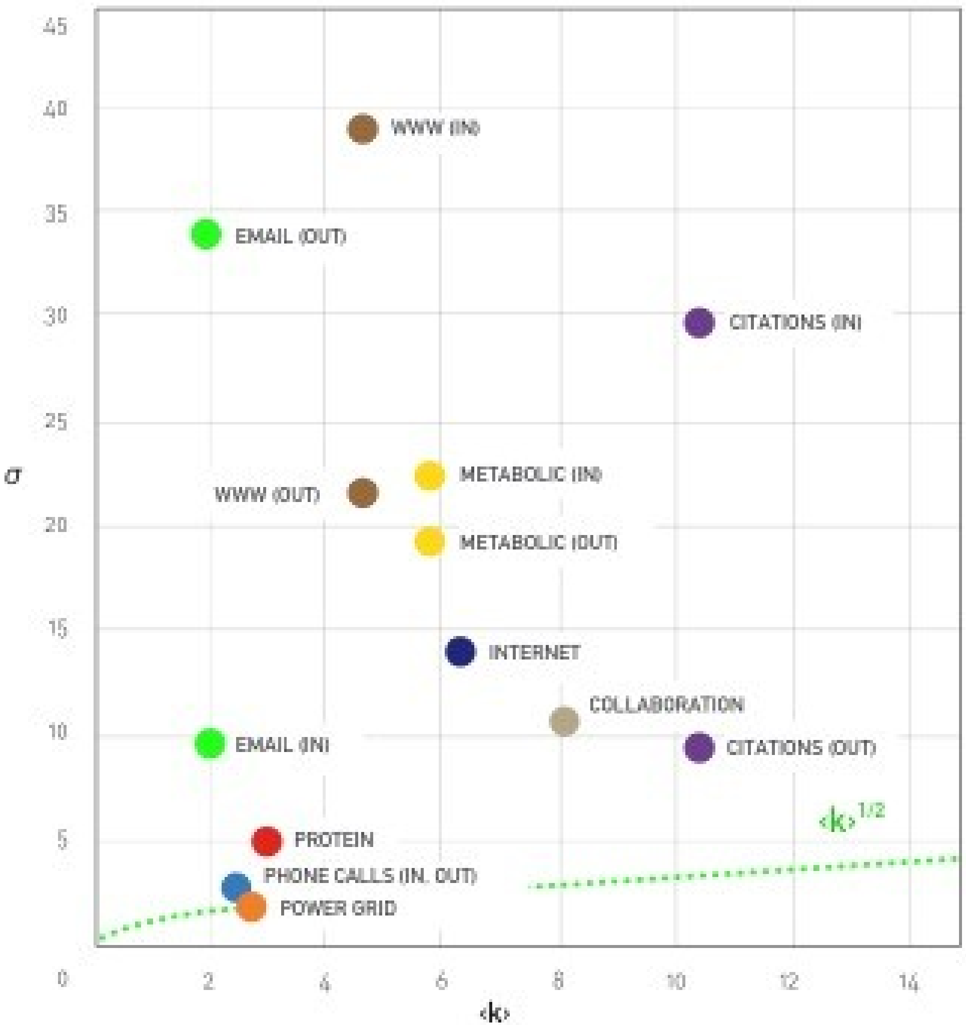
Scale: $\langle k \rangle$

Scale-Free Network

Randomly chosen node: $k = \langle k \rangle \pm \infty$

Scale: none

The meaning of scale-free

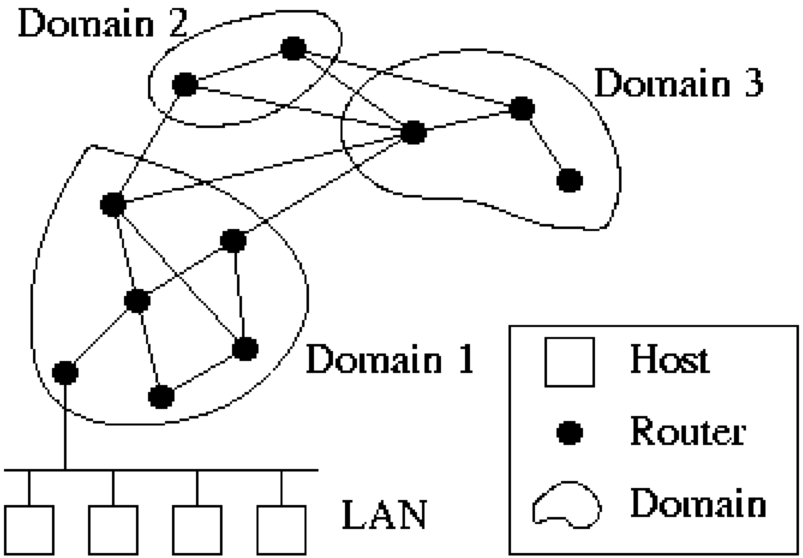


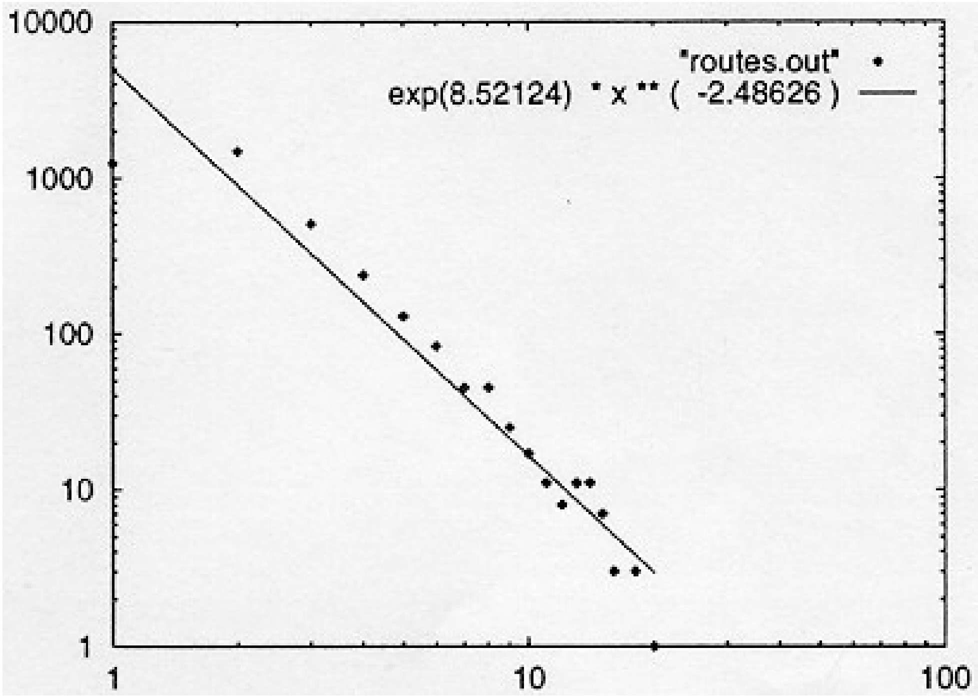
$$k = \langle k \rangle \pm \sigma_k$$

universality

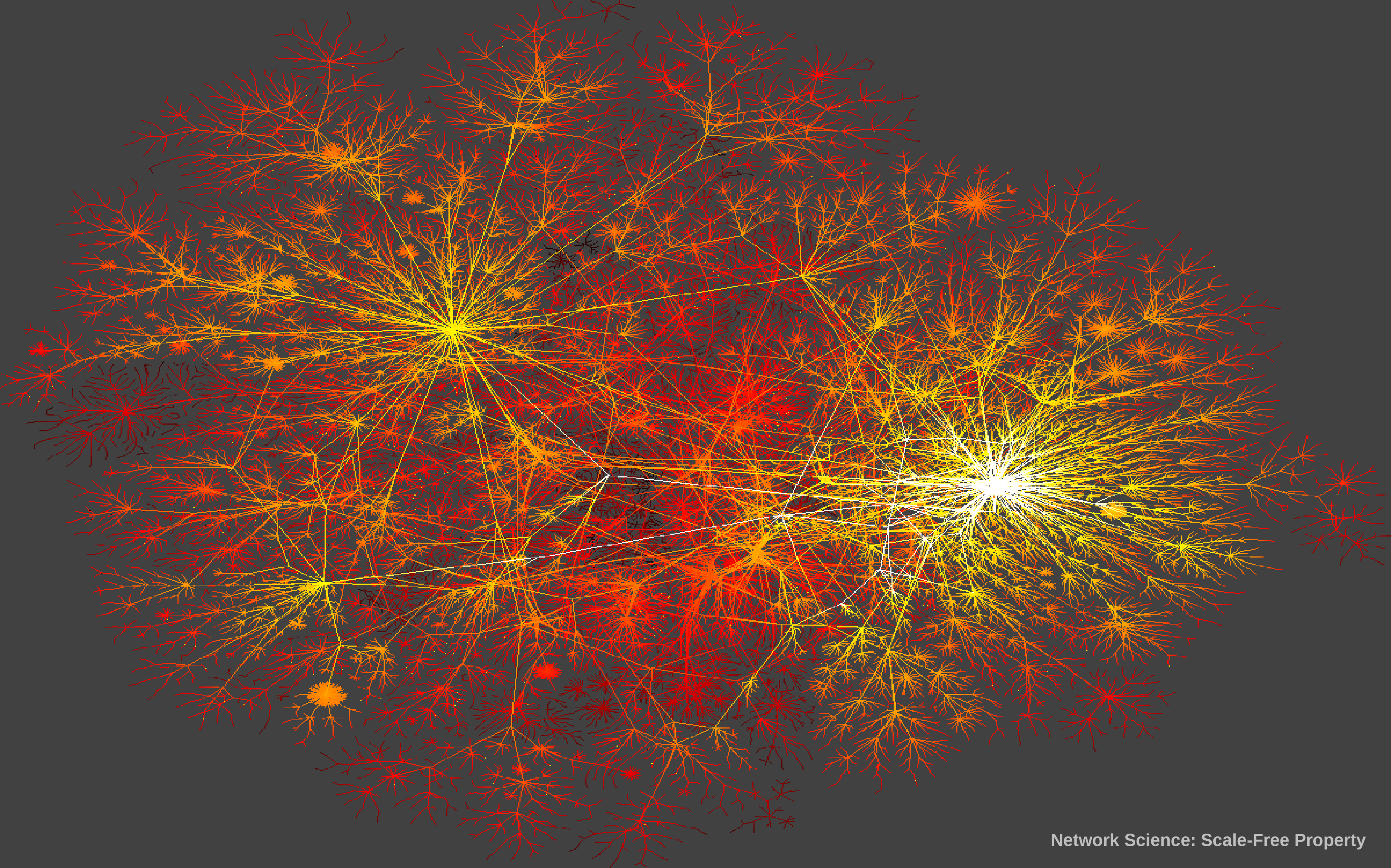
INTERNET BACKBONE

Nodes: computers, routers
Links: physical lines





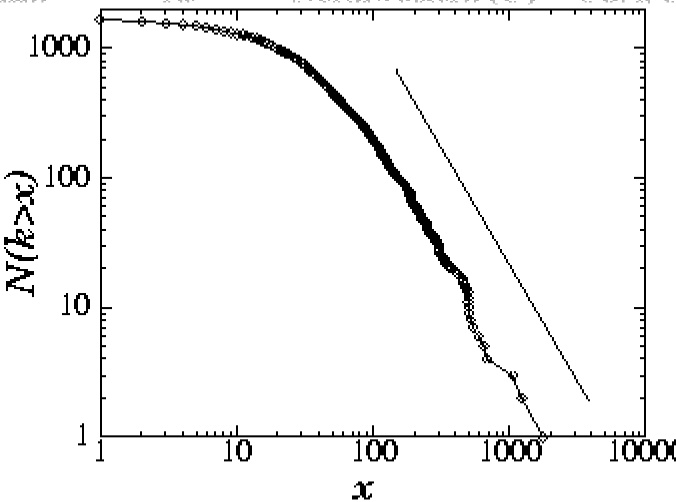
(Faloutsos, Faloutsos and Faloutsos, 1999)

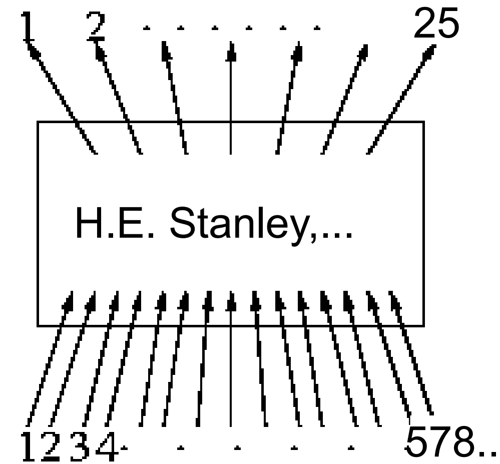


SCIENCE CITATION INDEX

Nodes: papers
Links: citations

1736 PRL papers (1988)





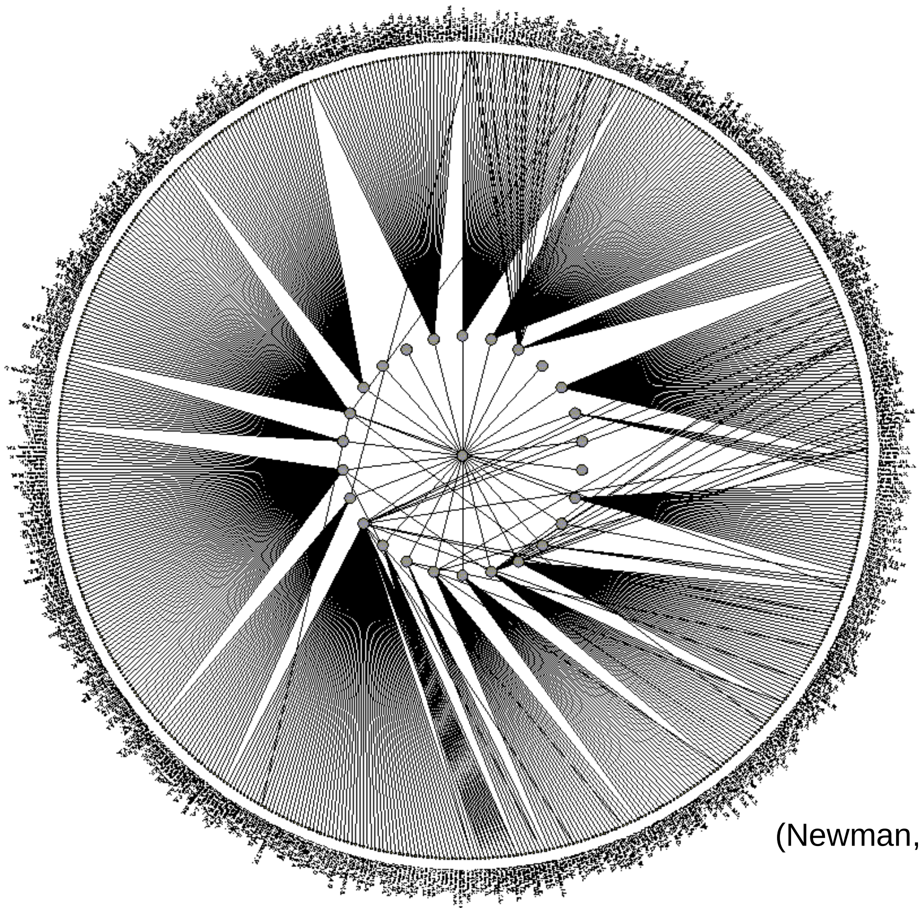
$P(k) \sim k^{-\gamma}$

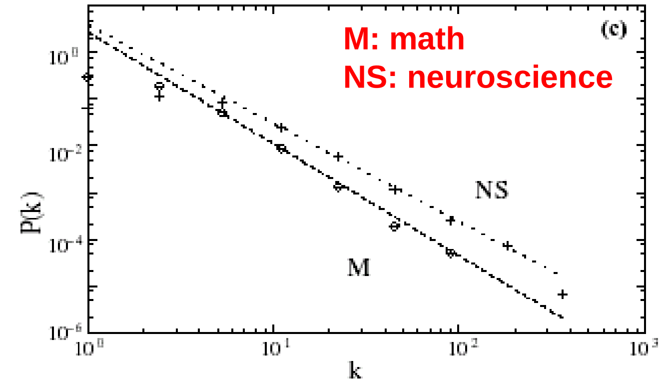
$(\gamma = 3)$

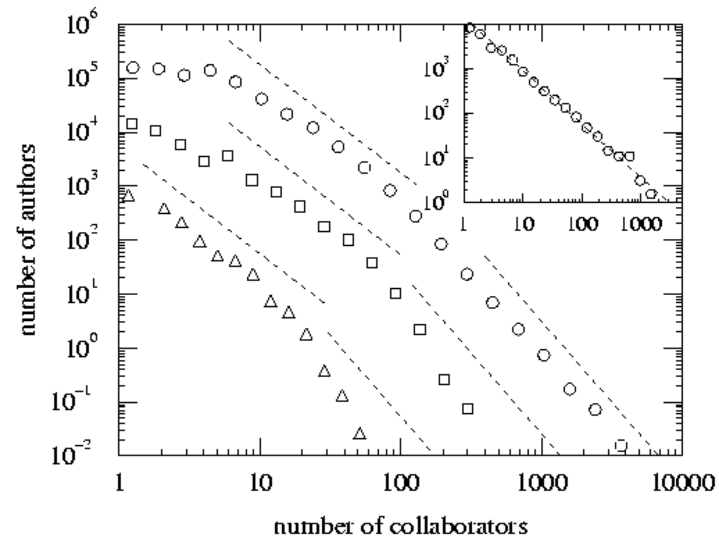
(S. Redner, 1998)

SCIENCE COAUTHORSHIP

Nodes: scientist (authors)
Links: joint publication





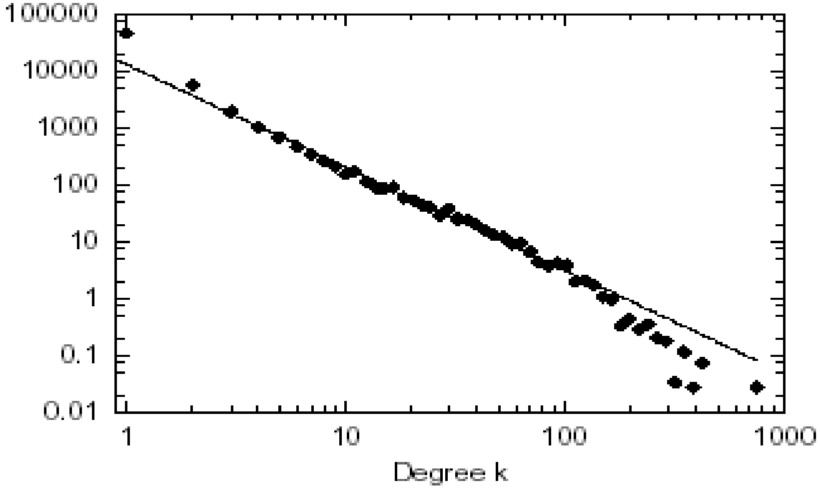


(Newman, 2000, Barabasi et al 2001)

ONLINE COMMUNITIES

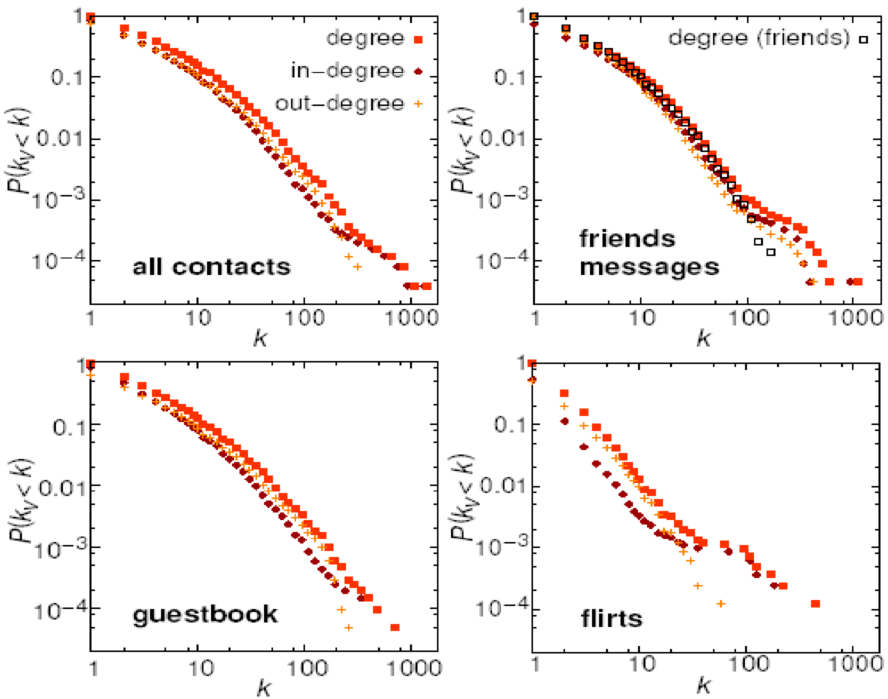
Nodes: online user
Links: email contact

Kiel University log files
112 days, N=59,912 nodes



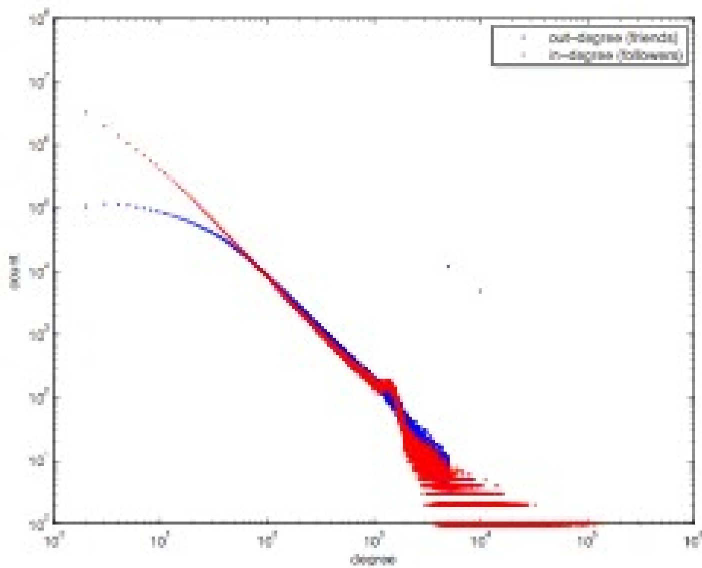
Ebel, Mielsch, Bornholdtz, PRE 2002.

Pussokram.com online community;
512 days, 25,000 users.

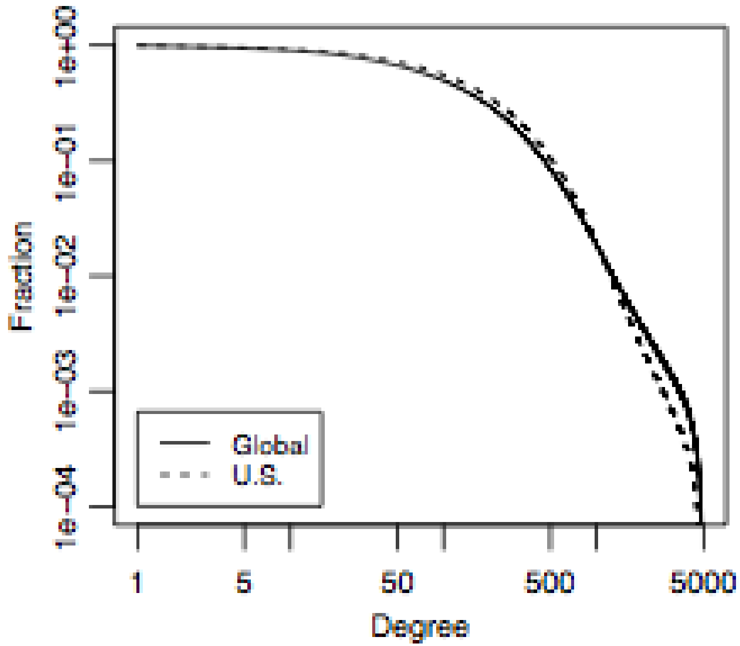


Holme, Edling, Liljeros, 2002.

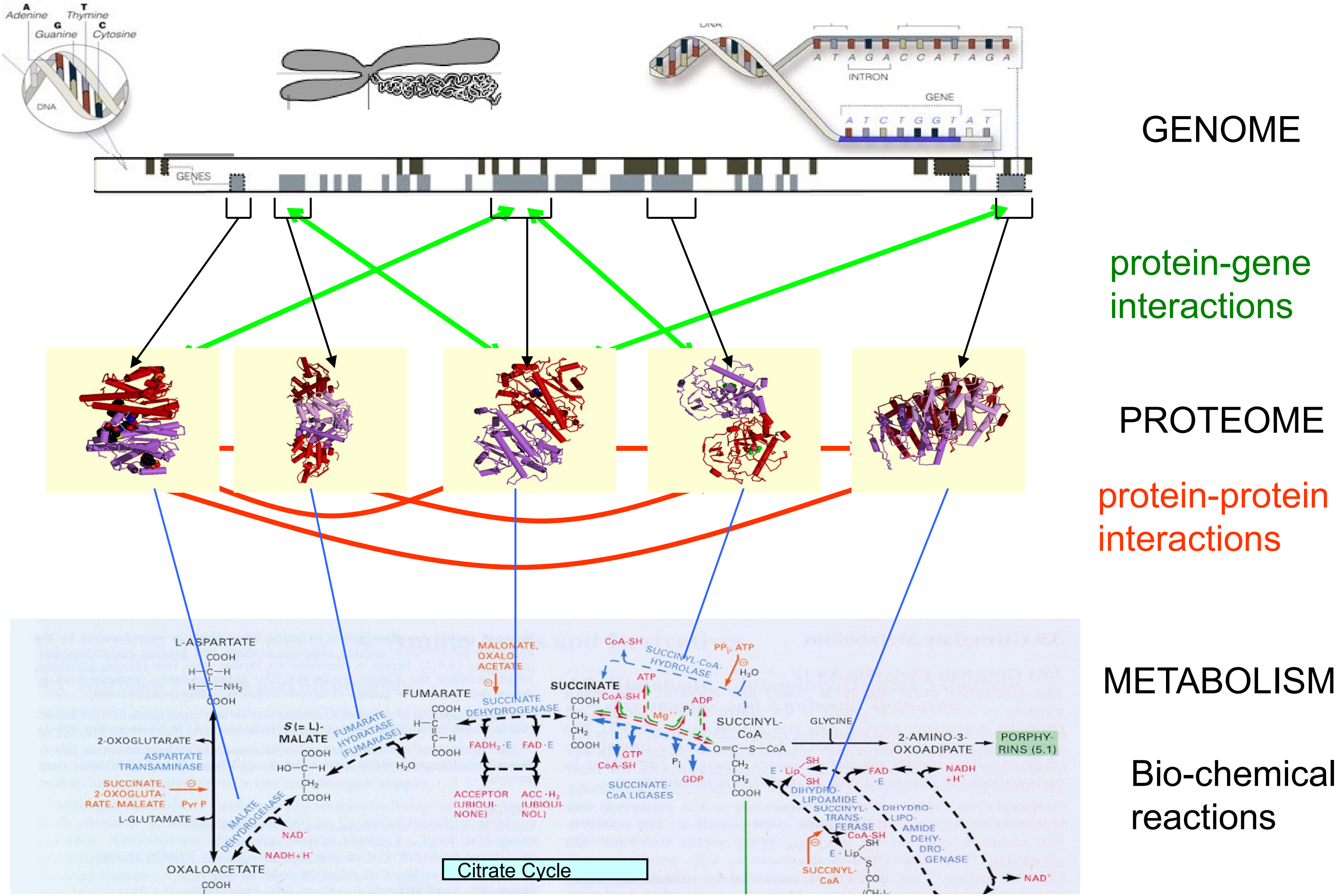
Twitter:



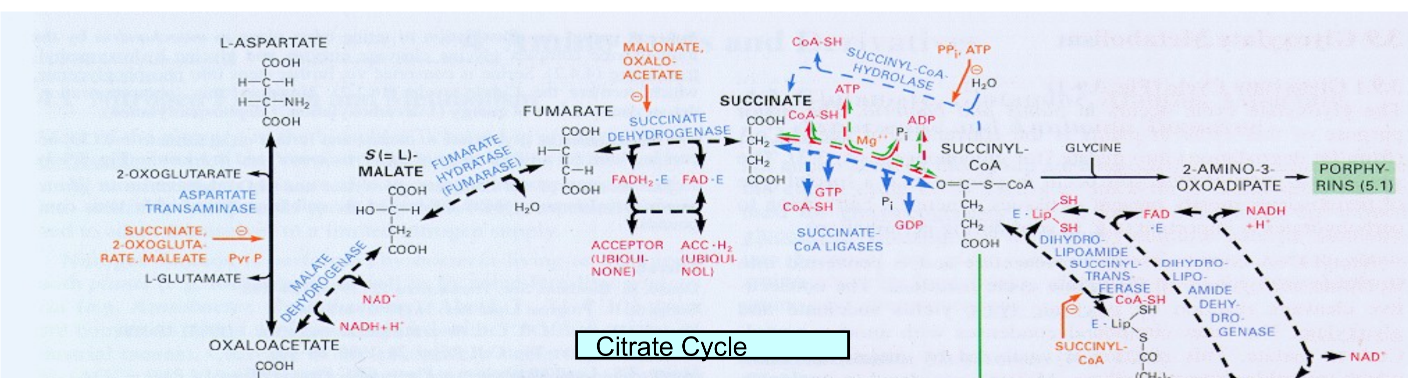
Facebook



Brian Karrer, Lars Backstrom, Cameron Marlowm 2011



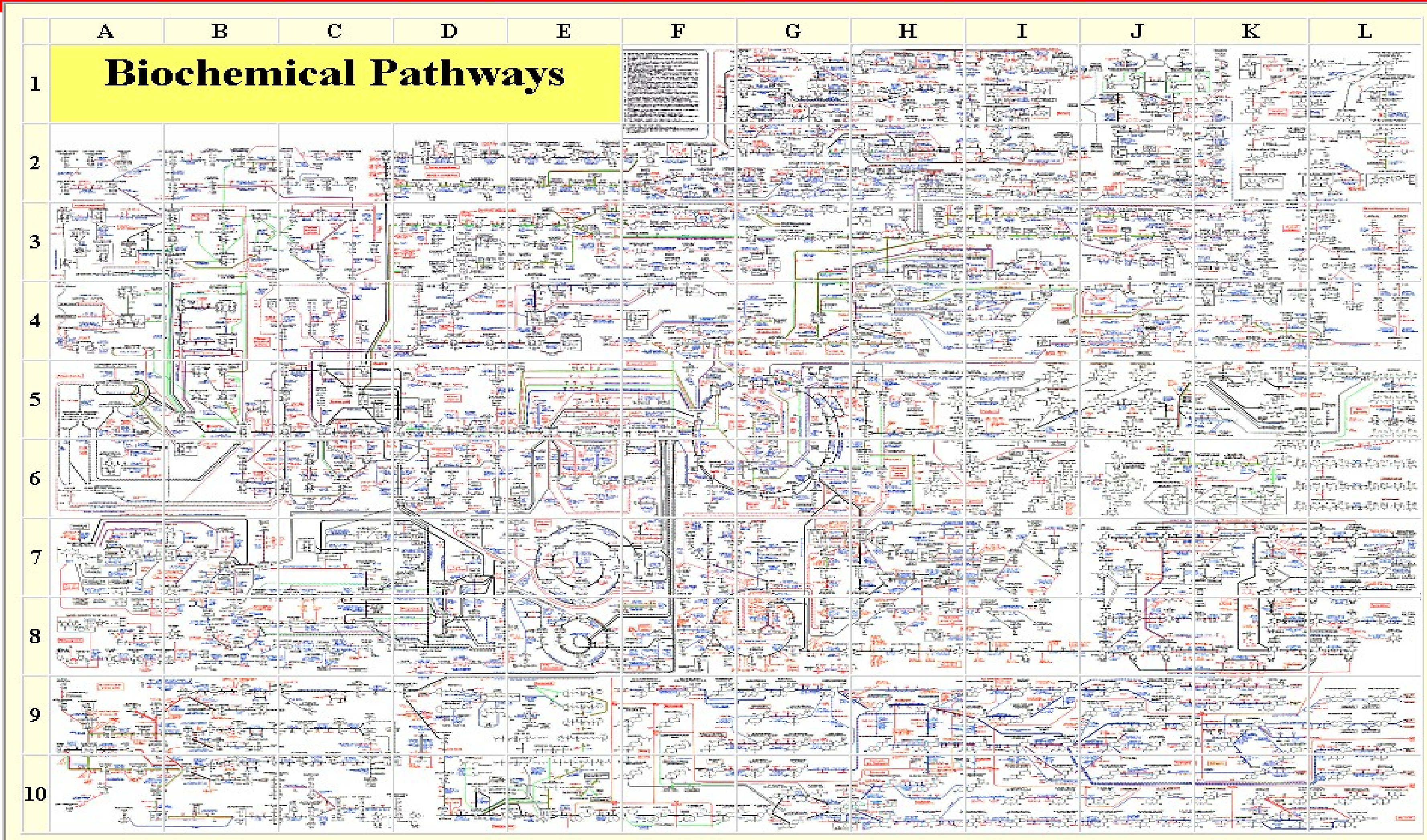
A Adenine
T Thymine
G Guanine
C Cytosine
DNA
A T A G A C C A T A G A
INTRON
GENE
A T C T G G T A T
GENES
GENOME
protein-gene interactions
PROTEOME
protein-protein interactions
METABOLISM
Bio-chemical reactions
L-ASPARTATE
2-OXOGLUTARATE
ASPARTATE TRANSAMINASE
SUCCINATE, 2-OXOGLUTARATE, MALEATE
Pyr P
L-GLUTAMATE
OXALOACETATE
MALATE DEHYDROGENASE
NAD+
NADH+H+
S(= L)-MALATE
FUMARATE HYDRATASE (FUMARASE)
H2O
FUMARATE
MALONATE, OXALOACETATE
SUCCINATE DEHYDROGENASE
FADH2-E
FAD-E
ACCEPTOR (UBIQUINONE)
ACC-H2 (UBIQUINOL)
SUCCINATE
CoA-SH
SUCCINYL-CoA HYDROLASE
PPi, ATP
H2O
ATP
ADP
GTP
GDP
SUCCINATE-CoA LIGASES
SUCCINYL-CoA
GLYCINE
2-AMINO-3-OXOADIPATE
PORPHYRINS (5.1)
DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE
DIHYDROLIPOAMIDE DEHYDROGENASE
NADH +H+
NAD+
SUCCINYL-CoA
Citrate Cycle



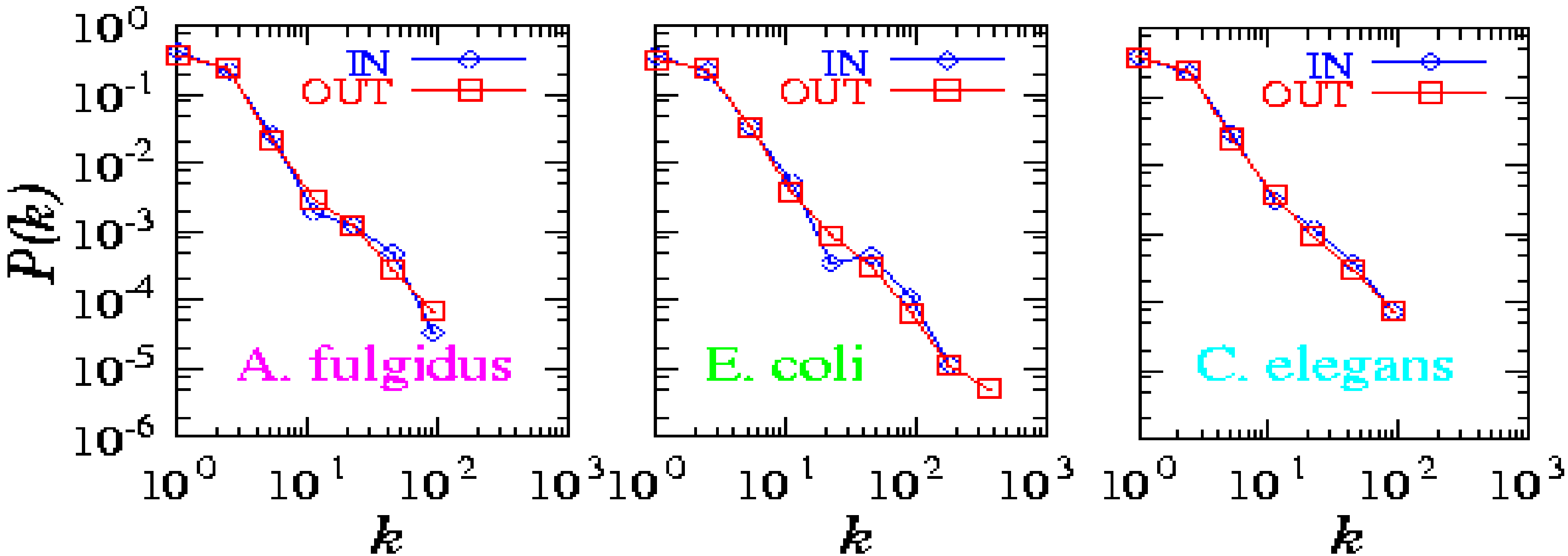
METABOLISM

Bio-chemical reactions

BOEHRING-MENNHEIN



METABOLIC NETWORK



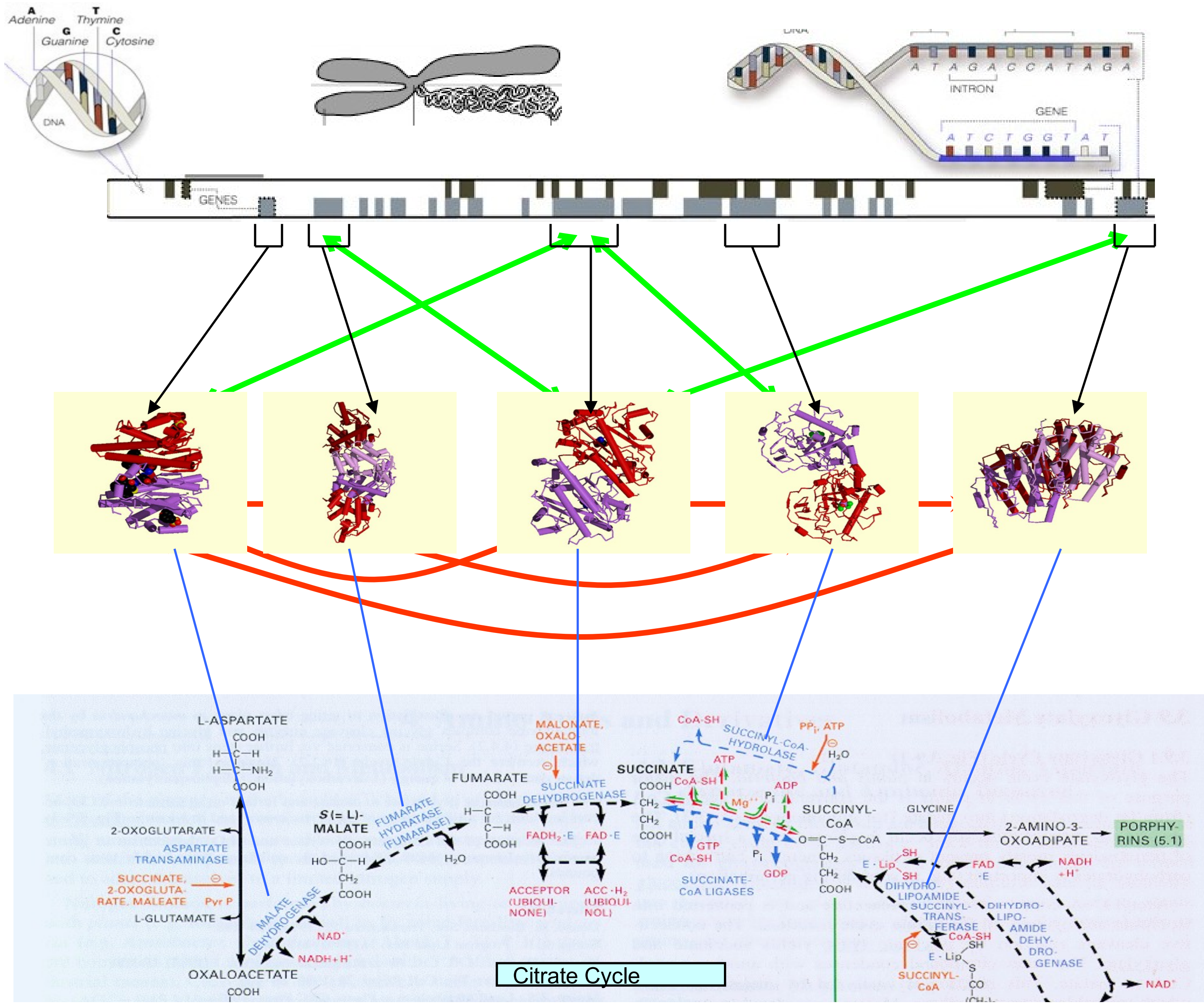
Archaea Bacteria Eukaryotes

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

$$P_{in}(k) \approx k^{-2.2}$$

$$P_{out}(k) \approx k^{-2.2}$$

H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, *Nature*, 407 651 (2000)



GENOME

protein-gene interactions

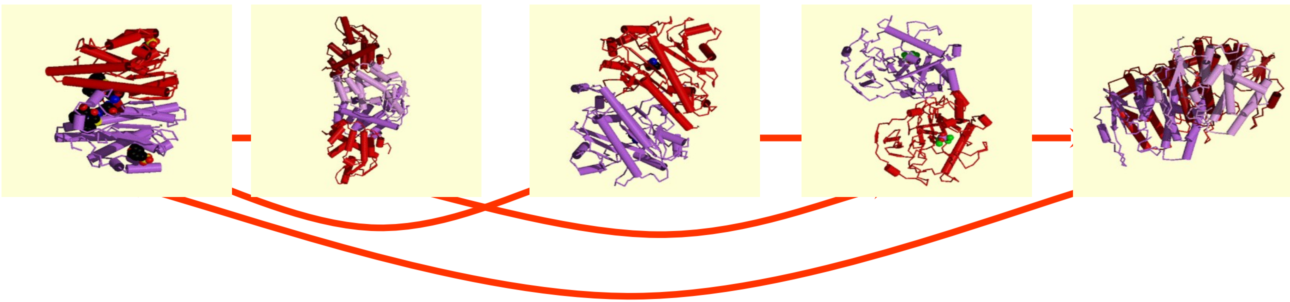
PROTEOME

protein-protein interactions

METABOLISM

Bio-chemical reactions

METABOLIC NETWORK

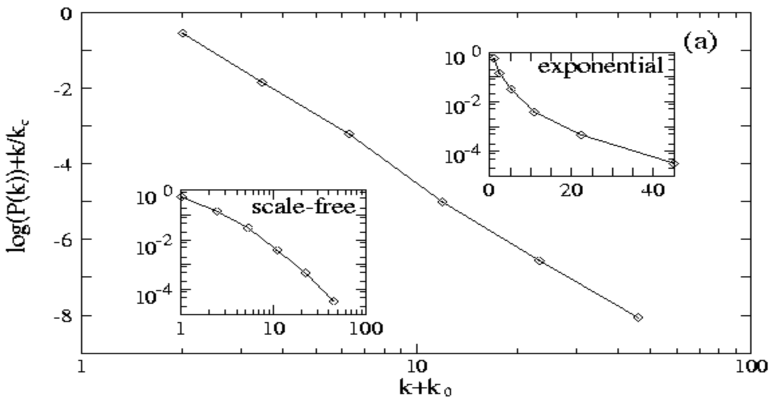


PROTEOME

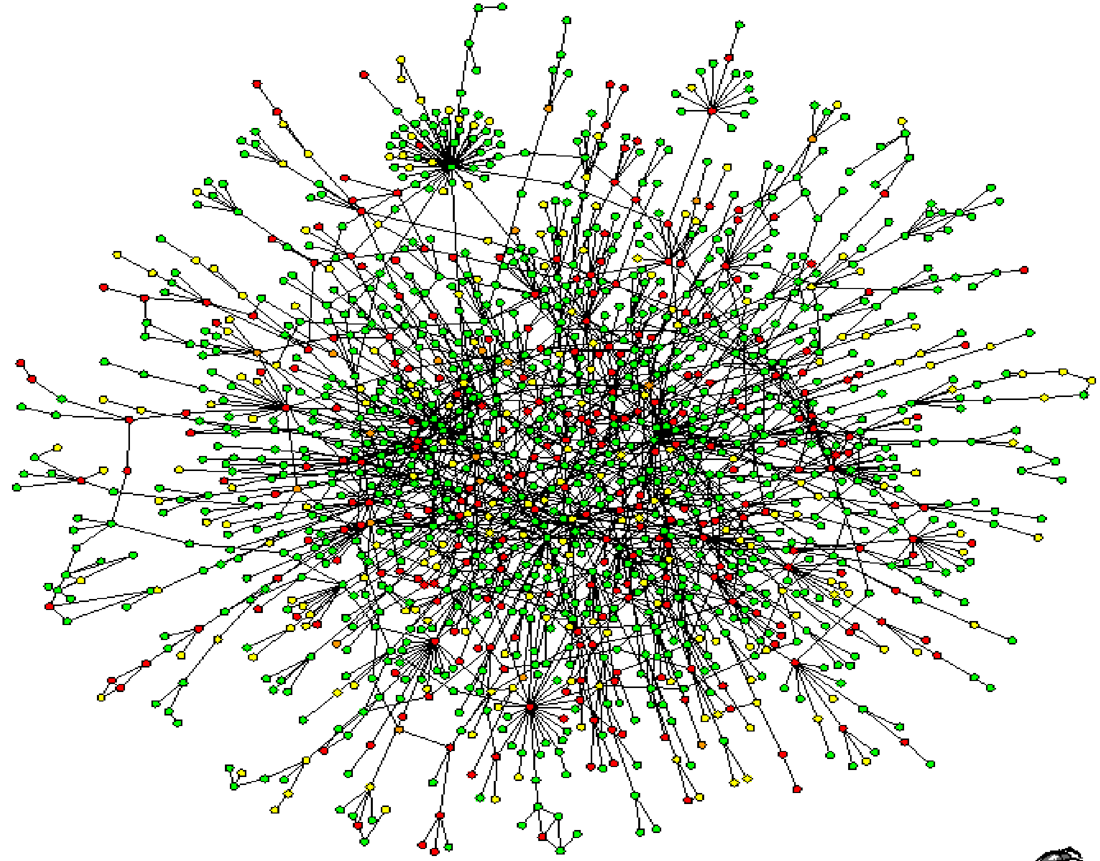
protein-protein interactions

TOPOLOGY OF THE PROTEIN NETWORK

Nodes: proteins
Links: physical interactions-binding

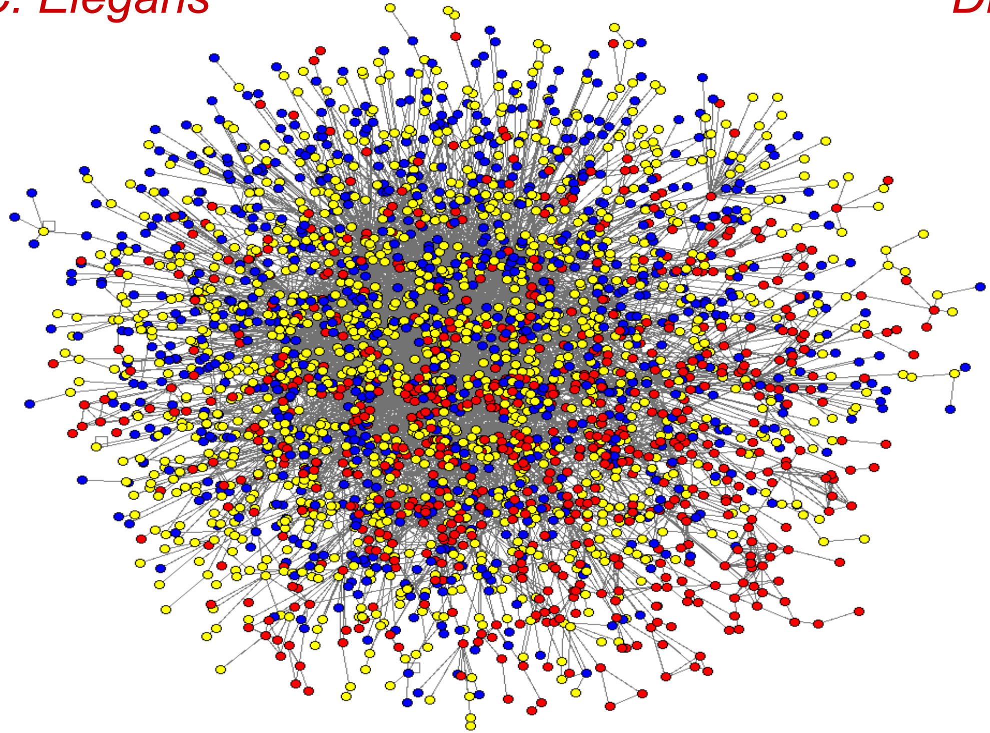


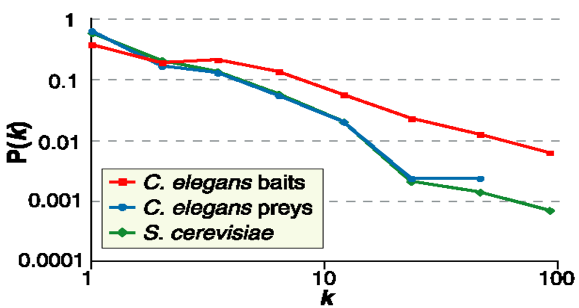
$$P(k) \sim (k + k_0)^{-\gamma} \exp\left(-\frac{k + k_0}{k_\tau}\right)$$



H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Nature 411, 41-42 (2001

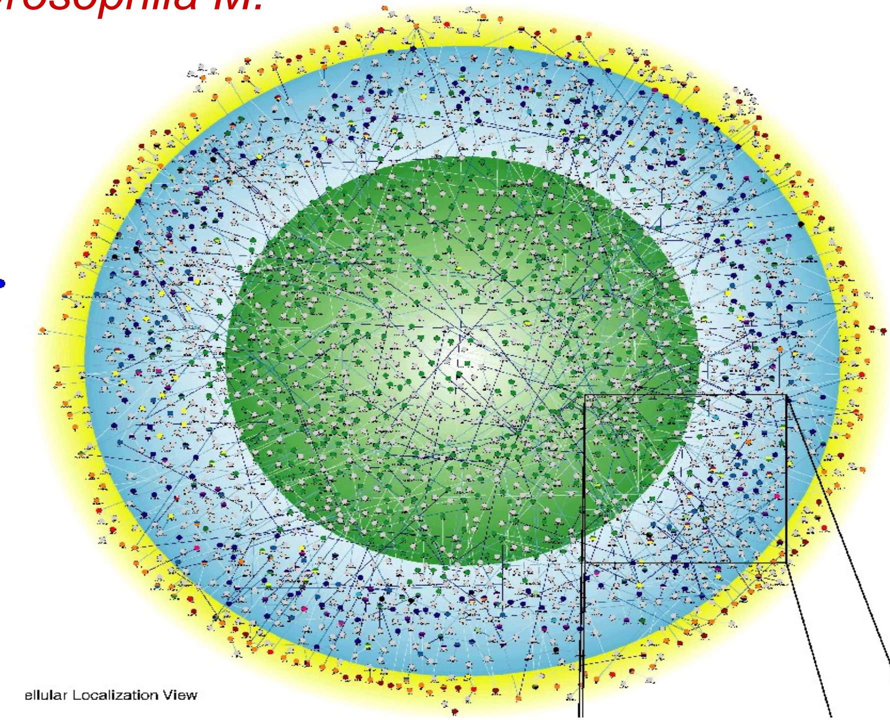
C. Elegans

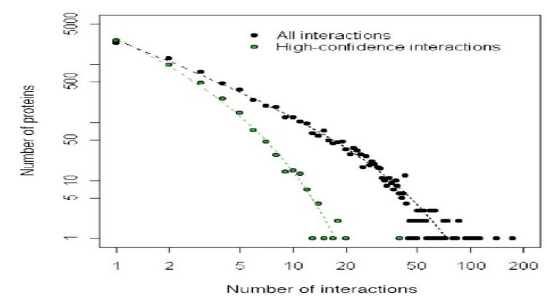




Li *et al.* Science 2004

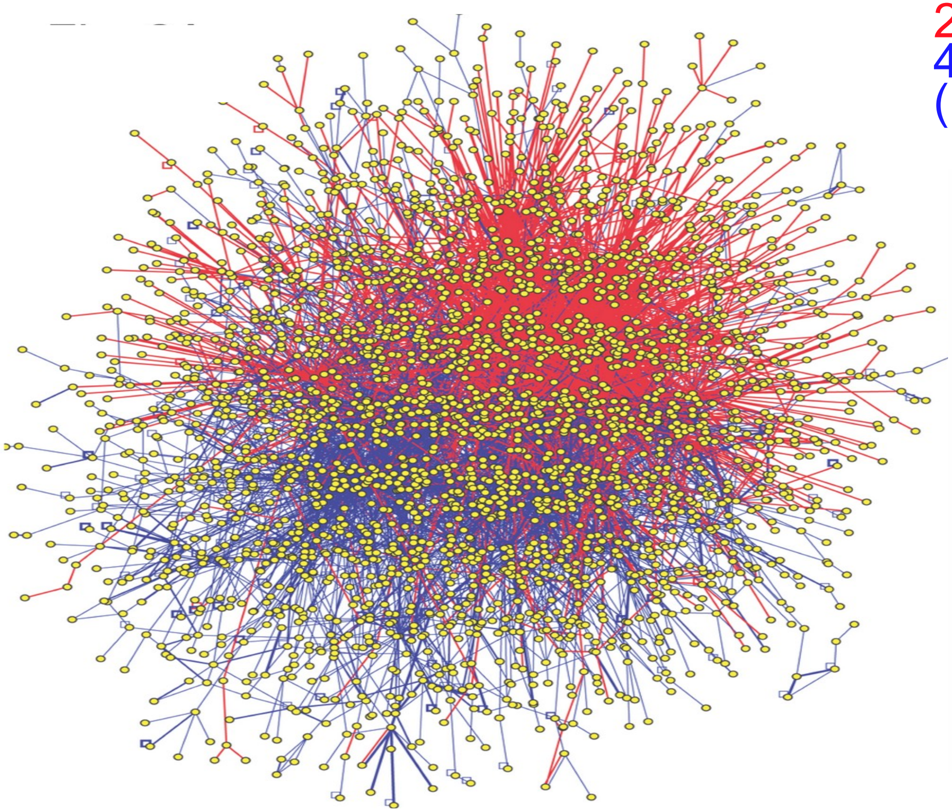
Drosophila M.



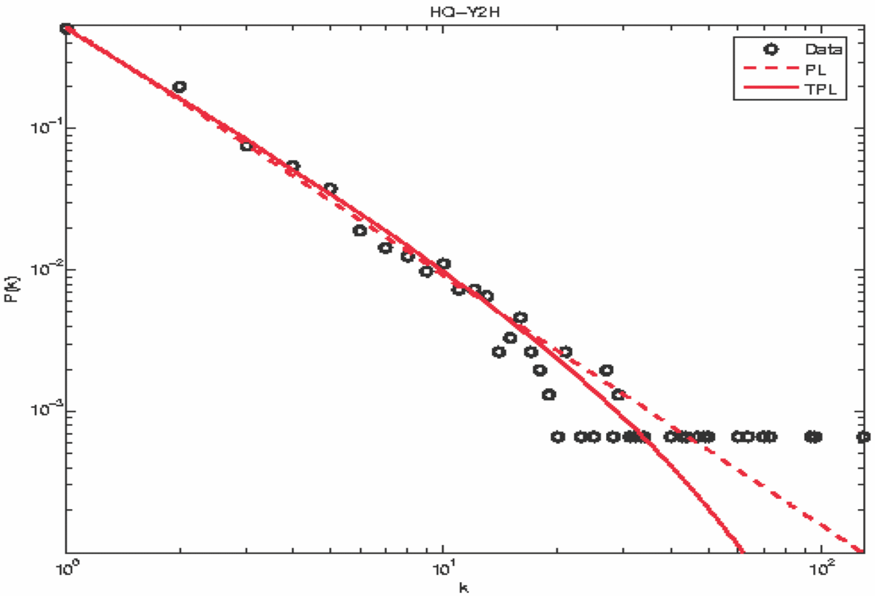


Giot *et al.* Science 2003

HUMAN INTERACTION NETWORK



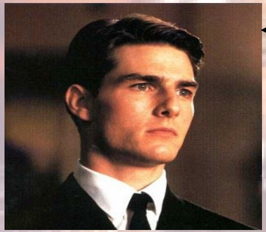
2,800 Y2H interactions
4,100 binary LC interactions
(HPRD, MINT, BIND, DIP, MIPS)



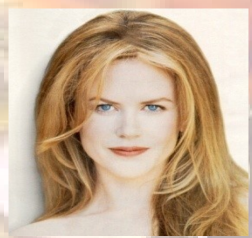
Rual *et al.* Nature 2005; Stelzl *et al.* Cell 2005

ACTOR NETWORK

Nodes: actors
Links: cast jointly



Days of Thunder (1990)
Far and Away (1992)
Eyes Wide Shut (1999)

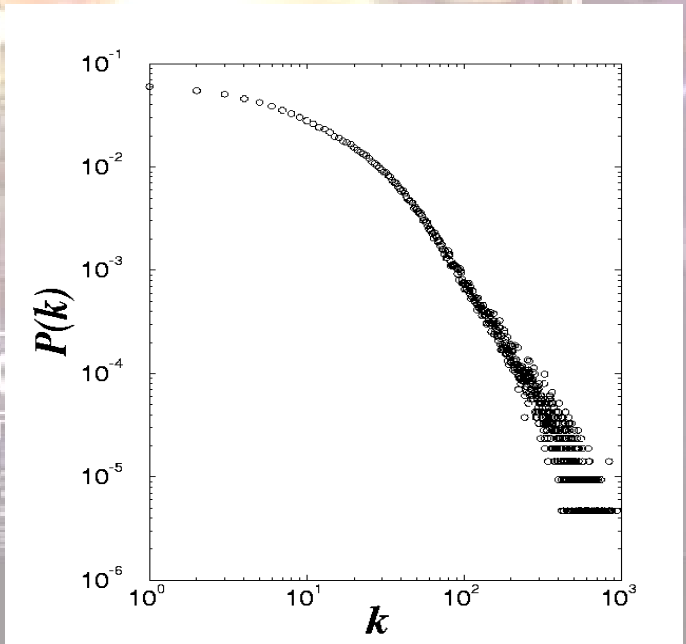


N = 212,250 actors

$\langle k \rangle = 28.78$

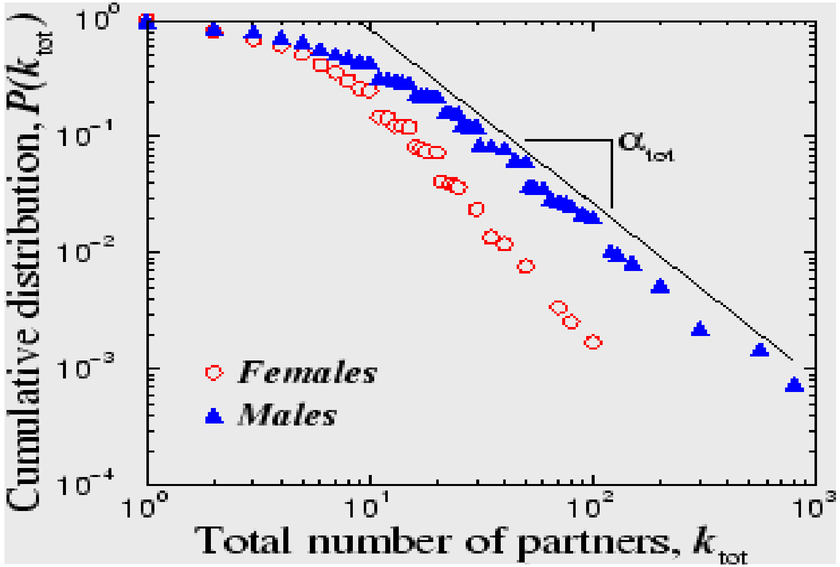
$P(k) \sim k^{-\gamma}$

$\gamma = 2.3$



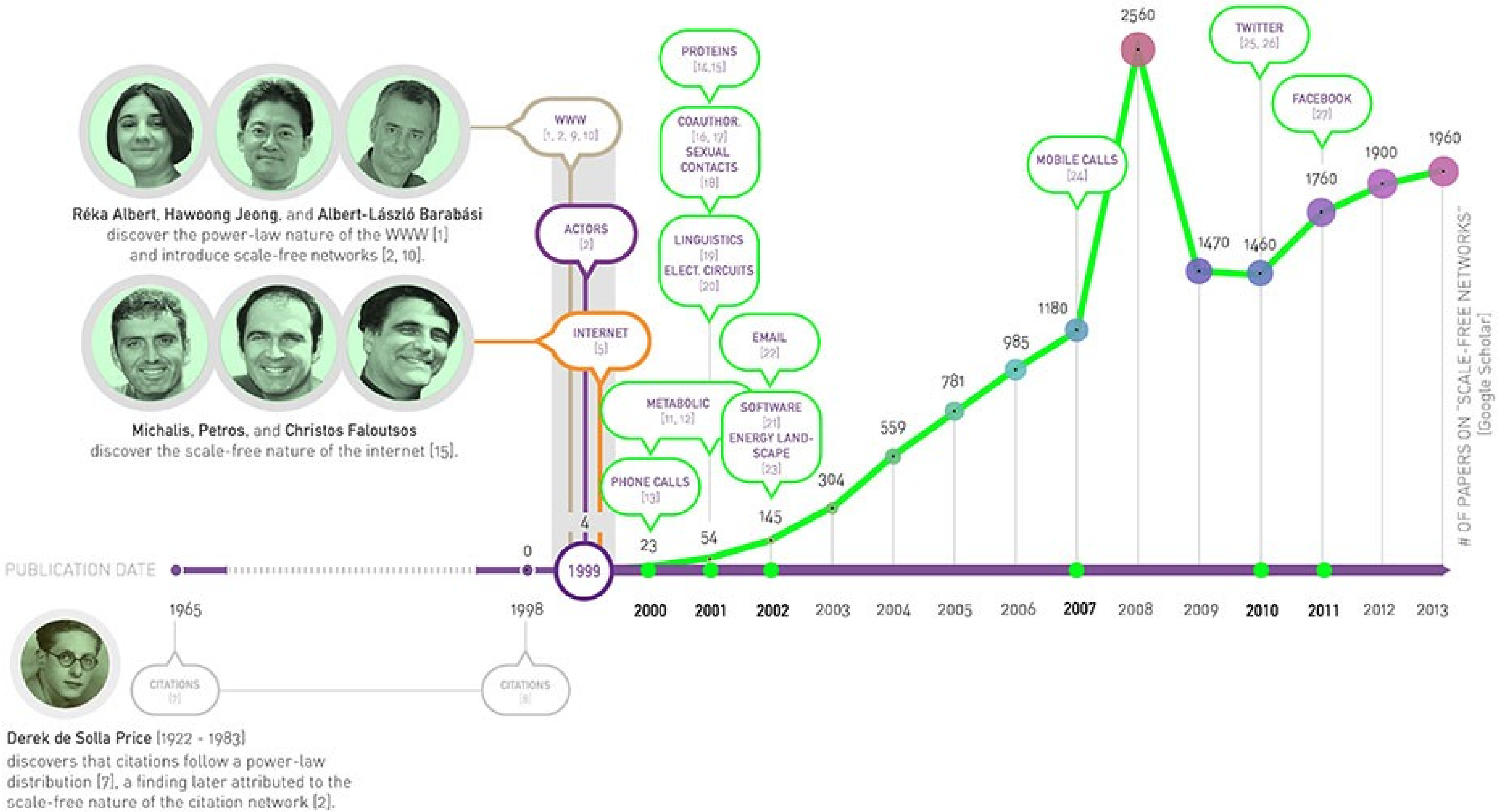
SWEDISH SE-WEB

Nodes: people (Females; Males)
Links: sexual relationships



4781 Swedes; 18-74;
59% response rate.

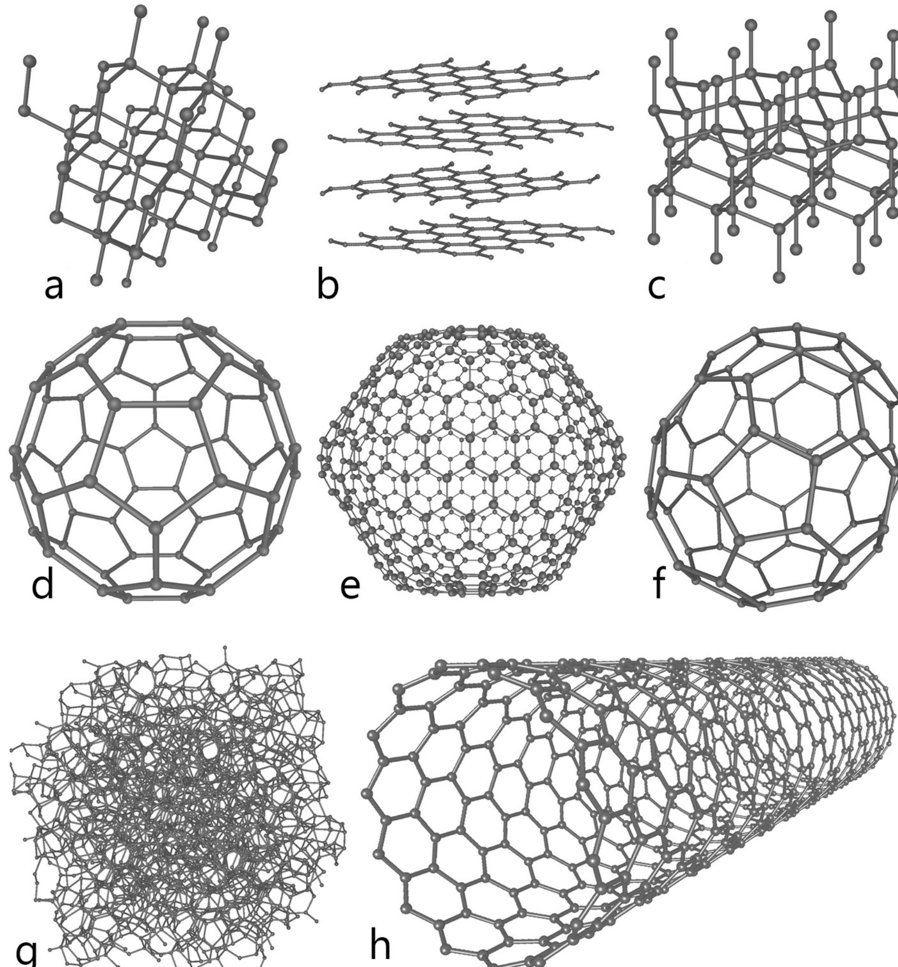
Liljeros et al. Nature 2001



WWW
[1, 2, 9, 10]
ACTORS
[2]
INTERNET
[5]
PROTEINS
[14,15]
COAUTHOR.
[16, 17]
SEXUAL CONTACTS
[18]
LINGUISTICS
[19]
ELECT. CIRCUITS
[20]
METABOLIC
[11, 12]
PHONE CALLS
[13]
EMAIL
[22]
SOFTWARE
[21]
ENERGY LANDSCAPE
[23]
MOBILE CALLS
[24]
TWITTER
[25, 26]
FACEBOOK
[27]
Réka Albert, Hawoong Jeong, and Albert-László Barabási
discover the power-law nature of the WWW [1]
and introduce scale-free networks [2, 10].
Michalis, Petros, and Christos Faloutsos
discover the scale-free nature of the internet [15].
PUBLICATION DATE
0
4
23
54
145
304
559
781
985
1180
2560
1470
1460
1760
1900
1960
1965
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
CITATIONS
[7]
CITATIONS
[8]
OF PAPERS ON "SCALE-FREE NETWORKS"
[Google Scholar]
Derek de Solla Price [1922 - 1983]
discovers that citations follow a power-law
distribution [7], a finding later attributed to the
scale-free nature of the citation network [2].

Not all networks are scale-free

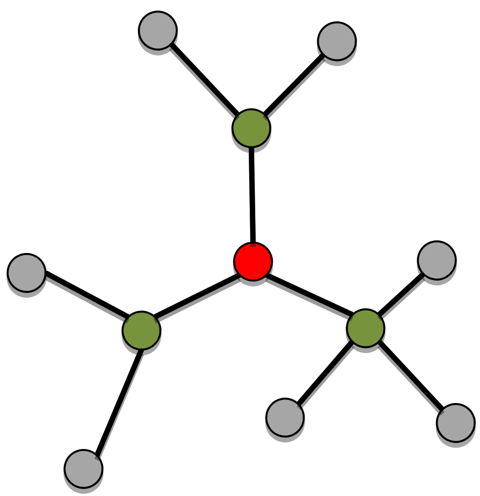
- *Networks appearing in material science, like the network describing the bonds between the atoms in crystalline or amorphous materials, where each node has exactly the same degree.*
- *The neural network of the C.elegans worm.*
- *The power grid, consisting of generators and switches connected by transmission lines*



Ultra-small property

DISTANCES IN RANDOM GRAPHS

Random graphs tend to have a tree-like topology with almost constant node degrees.



- nr. of first neighbors: $N_1 \cong \langle k \rangle$
- nr. of second neighbors: $N_2 \cong \langle k \rangle^2$
- nr. of neighbours at distance d: $N_d \cong \langle k \rangle^d$
- estimate maximum distance:

$$1 + \sum_{l=1}^{l_{max}} \langle k \rangle^i = N \Rightarrow l_{max} = \frac{\log N}{\log \langle k \rangle}$$

SMALL WORLD BEHAVIOR IN SCALE-FREE NETWORKS

$$k_{max} = k_{min} N^{\frac{1}{\gamma - 1}}$$

Ultra Small World

$$< l > \sim \begin{cases} const. & \gamma = 2 \\ \dfrac{\ln \ln N}{\ln(\gamma - 1)} & 2 < \gamma < 3 \\ \dfrac{\ln N}{\ln \ln N} & \gamma = 3 \\ \ln N & \gamma > 3 \end{cases}$$

Size of the biggest hub is of order O(N). Most nodes can be connected within two layers of it, thus the average path length will be independent of the system size.

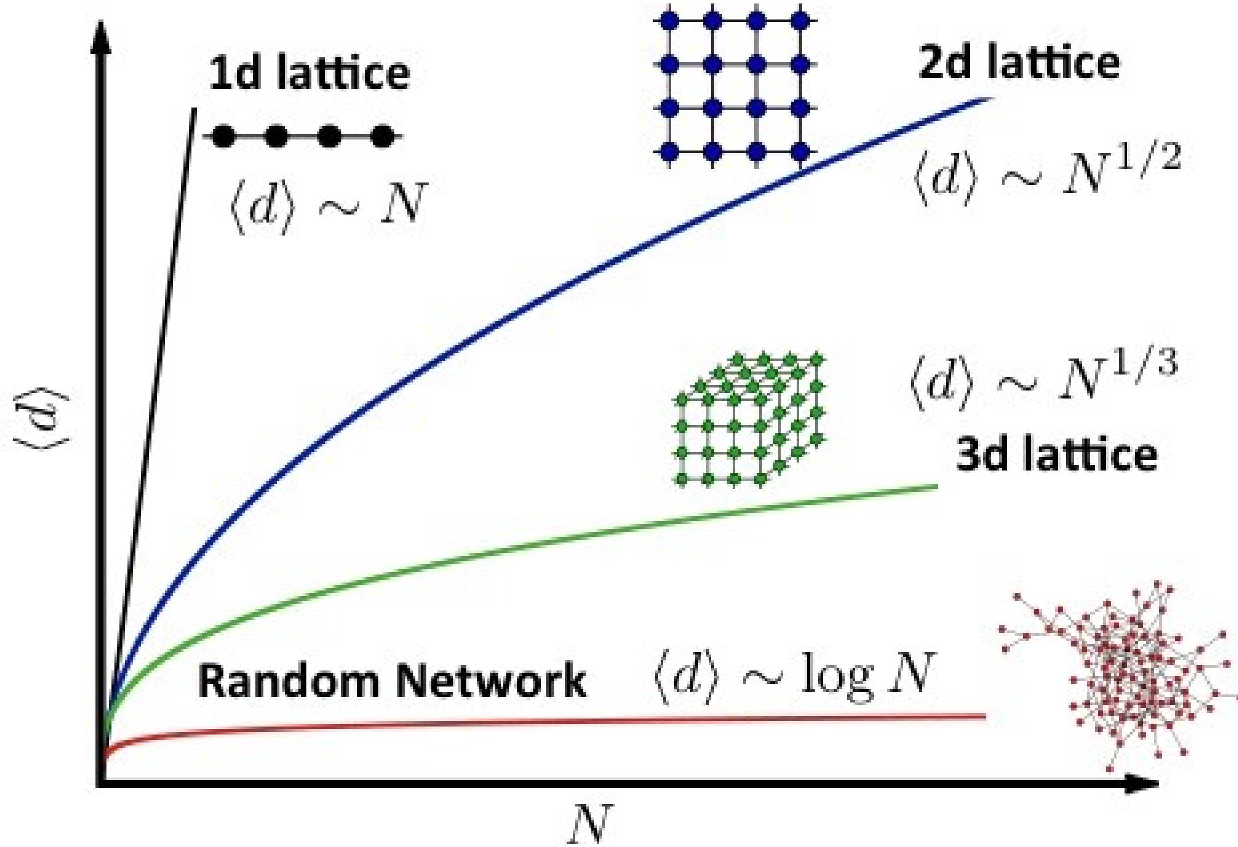
The average path length increases slower than logarithmically. In a random network all nodes have comparable degree, thus most paths will have comparable length. In a scale-free network the vast majority of the path go through the few high degree hubs, reducing the distances between nodes.

Some key models produce γ=3, so the result is of particular importance for them. This was first derived by Bollobas and collaborators for the network diameter in the context of a dynamical model, but it holds for the average path length as well.

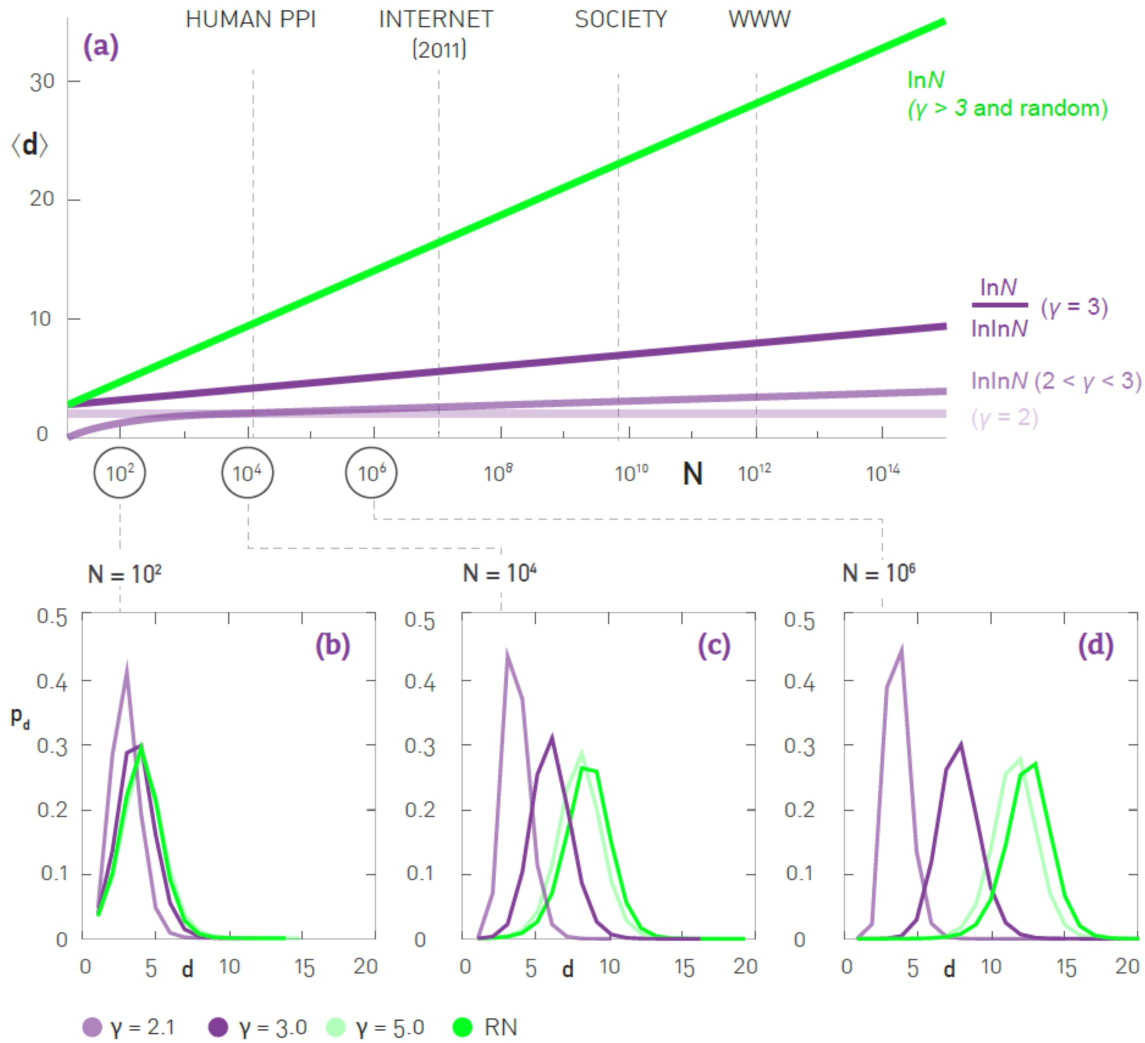
The second moment of the distribution is finite, thus in many ways the network behaves as a random network. Hence the average path length follows the result that we derived for the random network model earlier.

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); (Bollobas, Riordan, 2002; Bollobas, 1985; Newman, 2001

Why are small worlds surprising? Suprising compared to what?



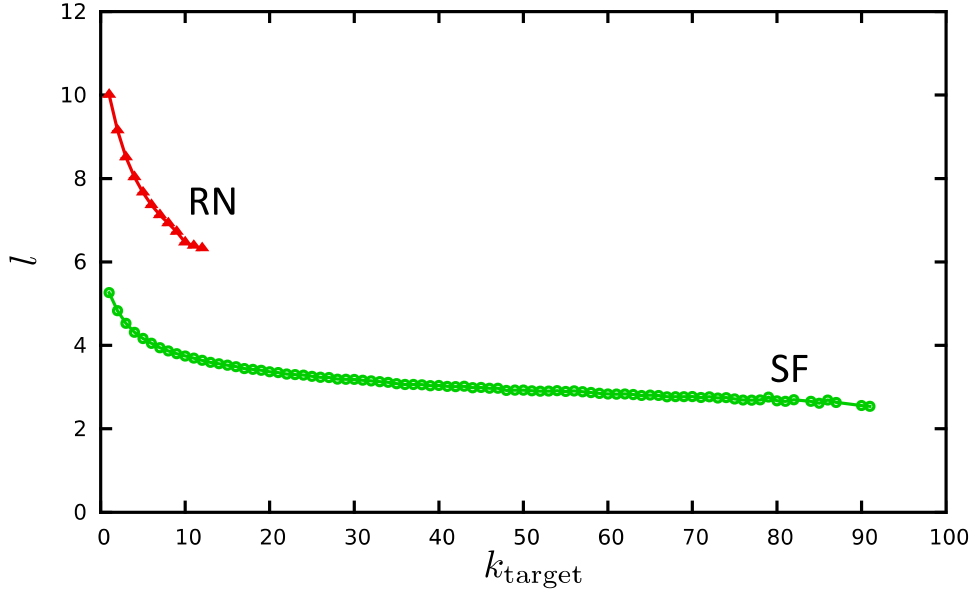
SMALL WORLD BEHAVIOR IN SCALE-FREE NETWORKS



$$
\langle d \rangle \sim \begin{cases} \text{const.} & \gamma = 2, \\ \dfrac{\ln \ln N}{\ln(\gamma - 1)} & 2 < \gamma < 3, \\ \dfrac{\ln N}{\ln \ln N} & \gamma = 3, \\ \ln N & \gamma > 3. \end{cases}
$$

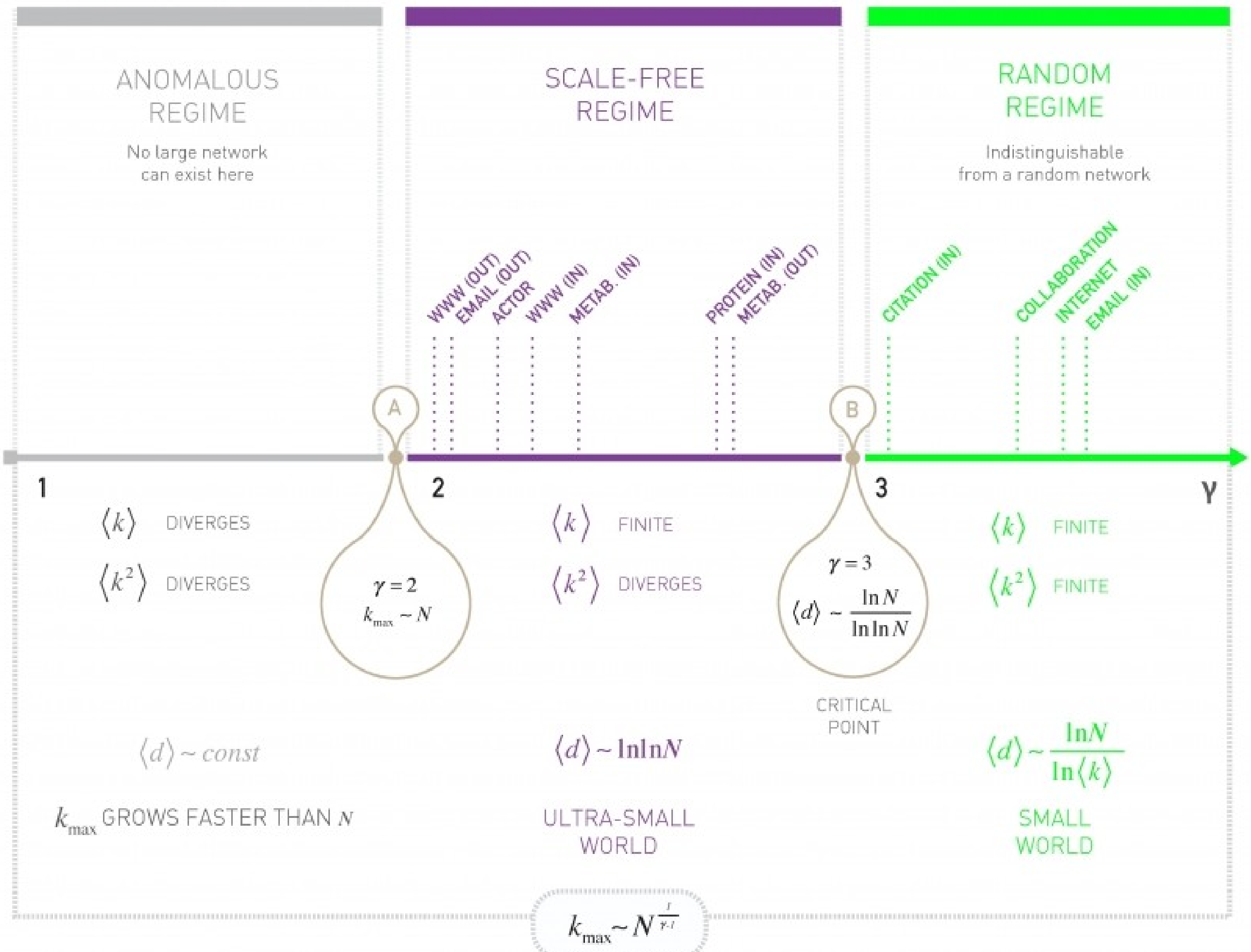
We are always close to the hubs

" it's always easier to find someone who knows a famous or popular figure than some run-the-mill, insignificant person.”
(Frigyes Karinthy, 1929)

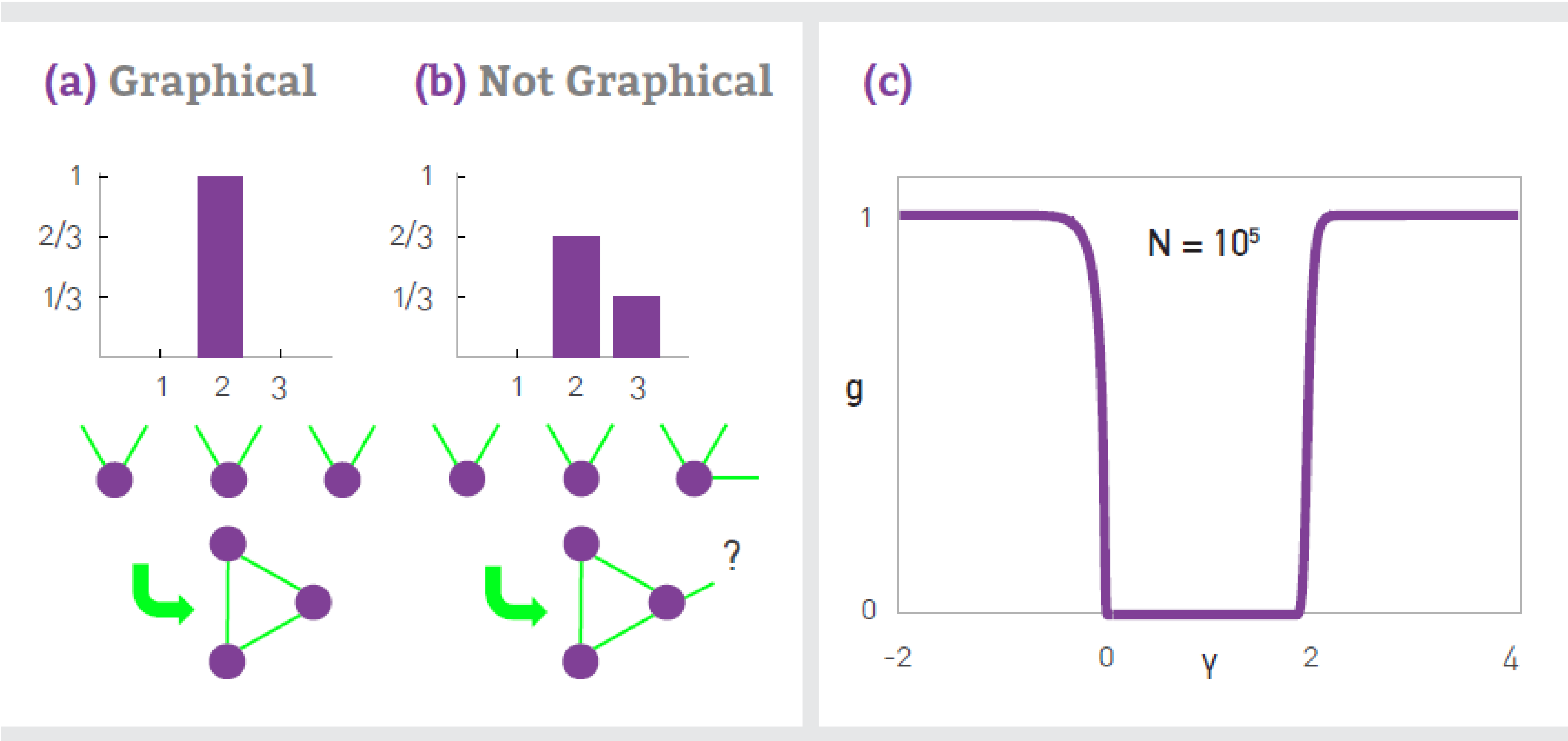


The role of the degree exponent

SUMMARY OF THE BEHAVIOR OF SCALE-FREE NETWORKS



Graphicality: No large networks for γ<2



In scale-free networks: $k_{\max} = k_{\min} N^{\frac{1}{\gamma - 1}}$ For γ<2: 1/(γ-1)>1

Why don't we see networks with exponents in the range of γ=4,5,6, etc?

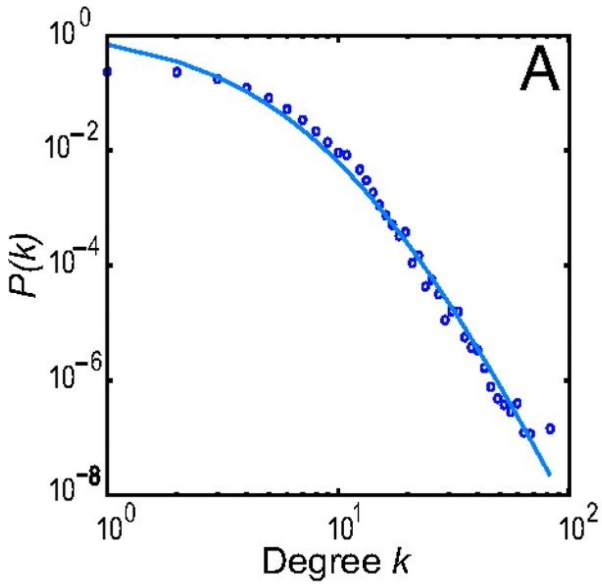
In order to document scale-free networks, we need 2-3 orders of magnitude scaling.
That is, $K_{max} \sim 10^2 K_{min}$ to $10^3 K_{min}$

However, that constrains on the system size we require to document it.
For example, to measure an exponent $\gamma=5$,we need to maximum degree a system size of the order of

$$K_{\max} = K_{\min} N^{\frac{1}{\gamma-1}}$$

$$N = \left(\frac{K_{\max}}{K_{\min}}\right)^{\gamma-1} \approx 10^8$$

Mobile Call Network



$N=4.6\text{x}10^6$
$\gamma=8.4$

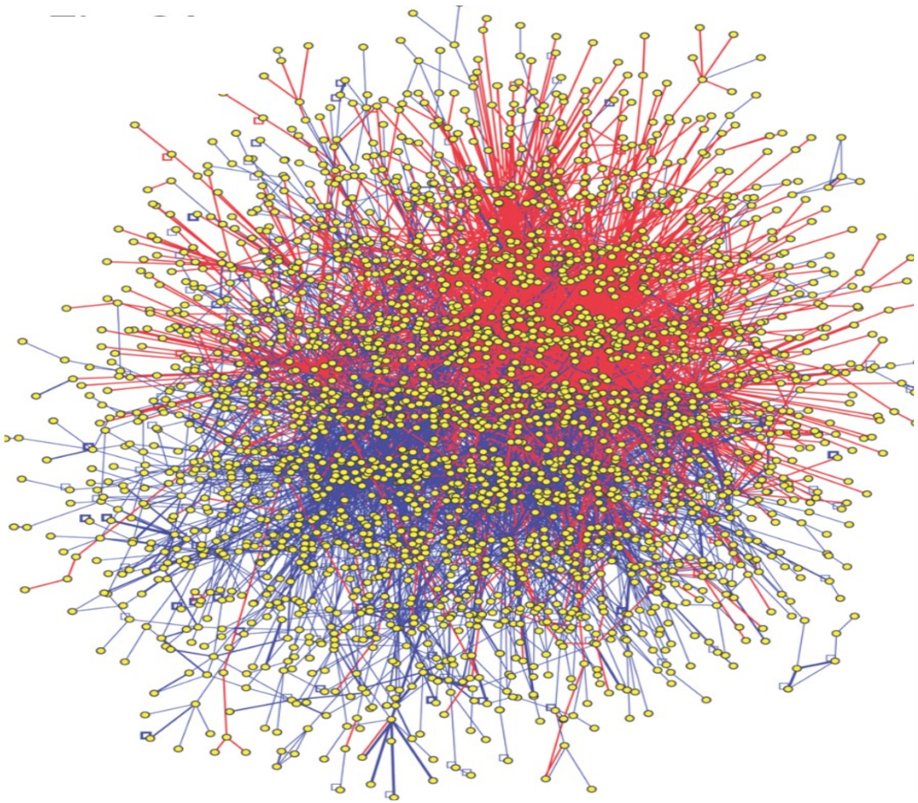
Fig. 1.

Characterizing the large-scale structure and the tie strengths of the mobile call graph. (A and B) Vertex degree (A) and tie strength distribution (B). Each distribution was fitted with $P(k) = a(k + k_0)^{-\gamma} \exp(-k/k_c)$.

Onella et al. PNAS 2007

PLOTTING POWER LAWS

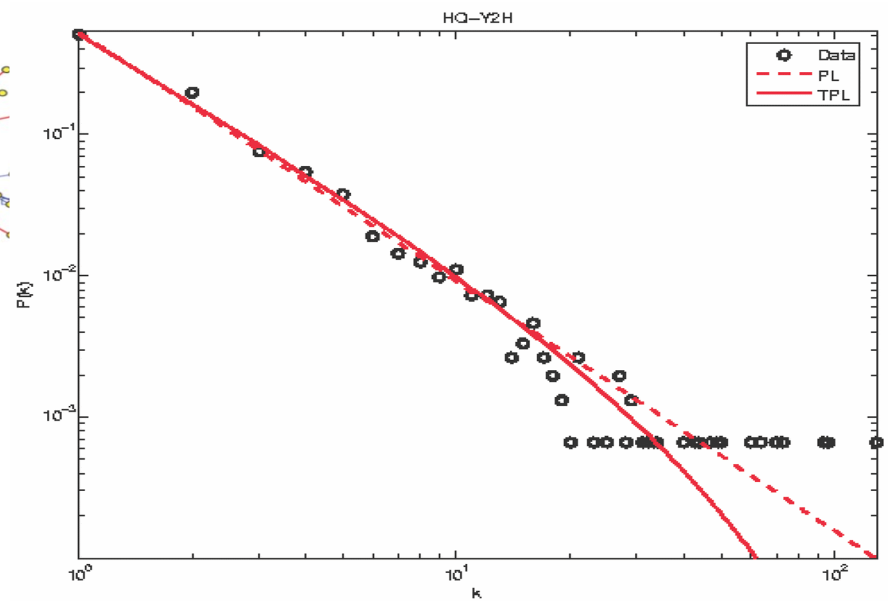
HUMAN INTERACTION NETWORK



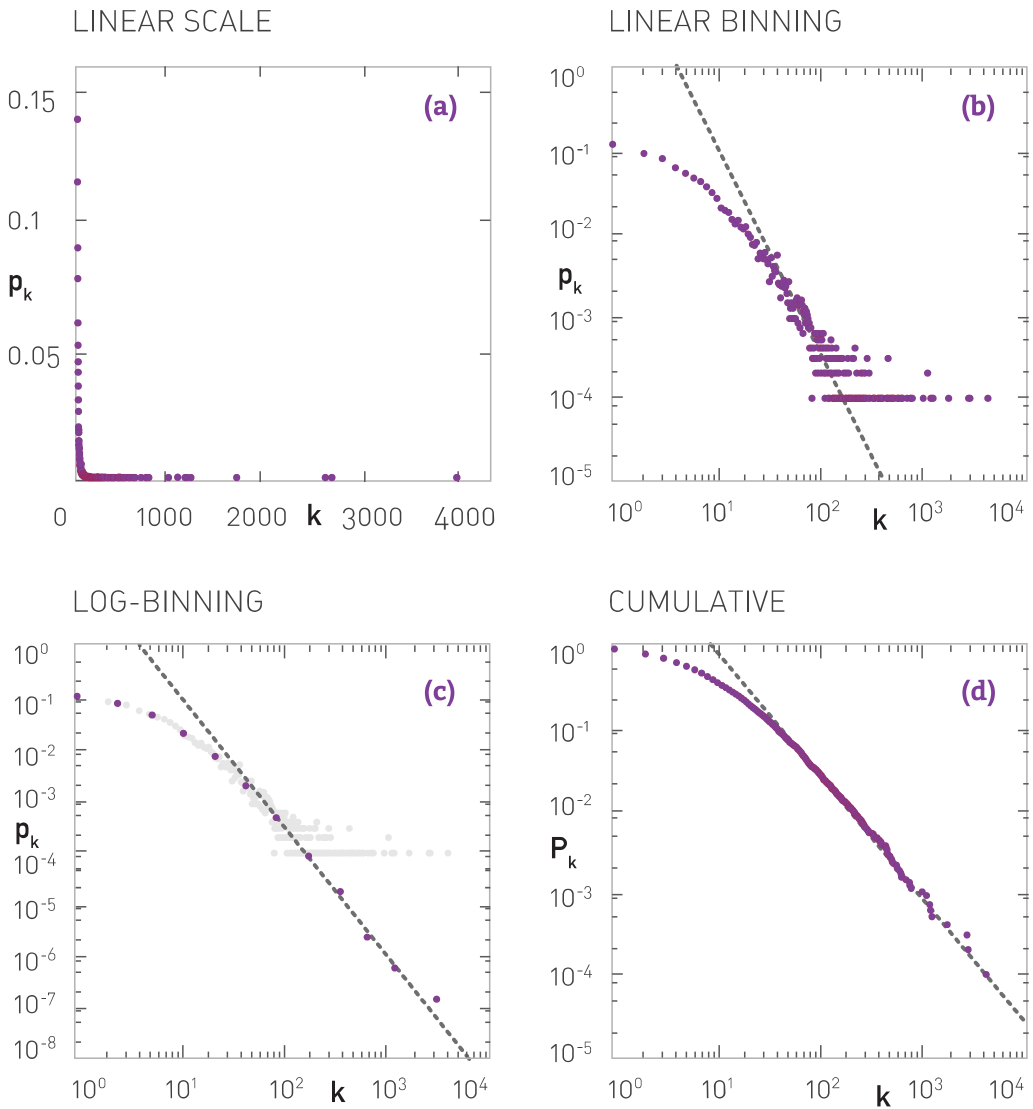
2,800 Y2H interactions

4,100 binary LC interactions

(HPRD, MINT, BIND, DIP, MIPS)



Rual et al. Nature 2005; Stelzl et al. Cell 2005



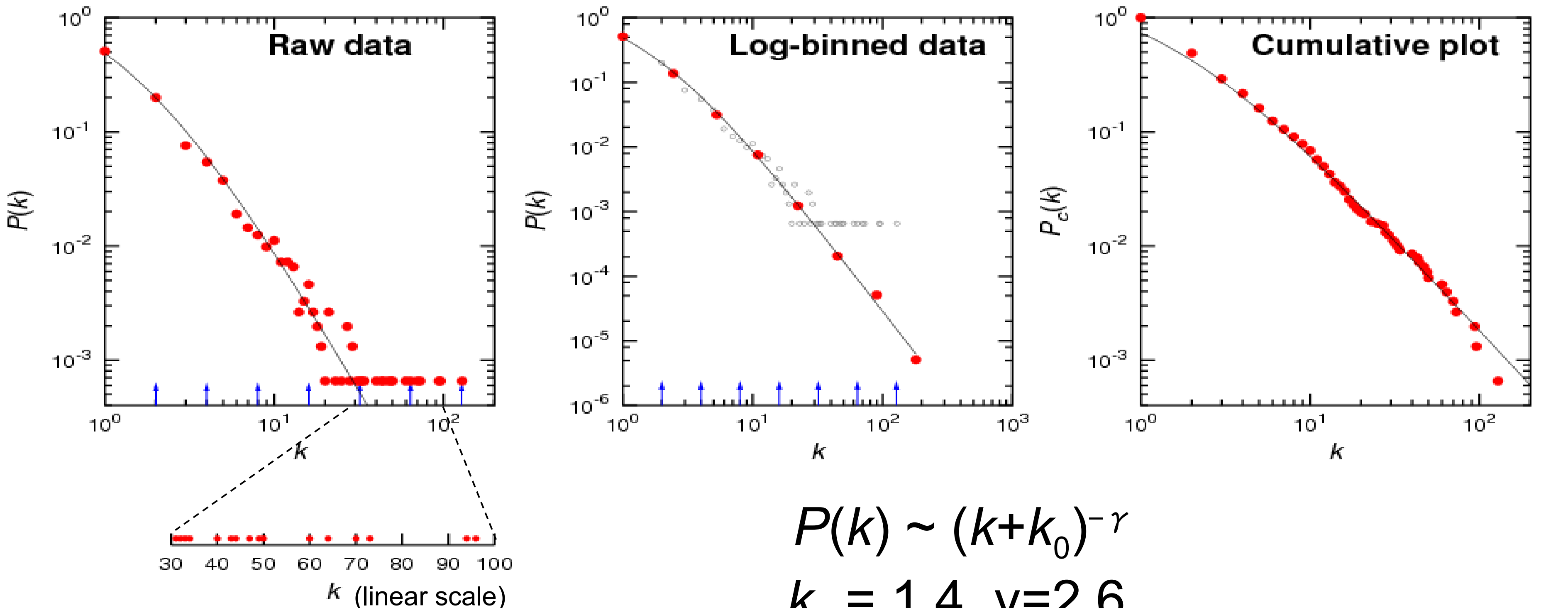
Use a Log-Log Plot

Avoid Linear Binning

Use Logarithmic Binning

Use Cumulative Distribution

HUMAN INTERACTION DATA BY RUAL ET AL.

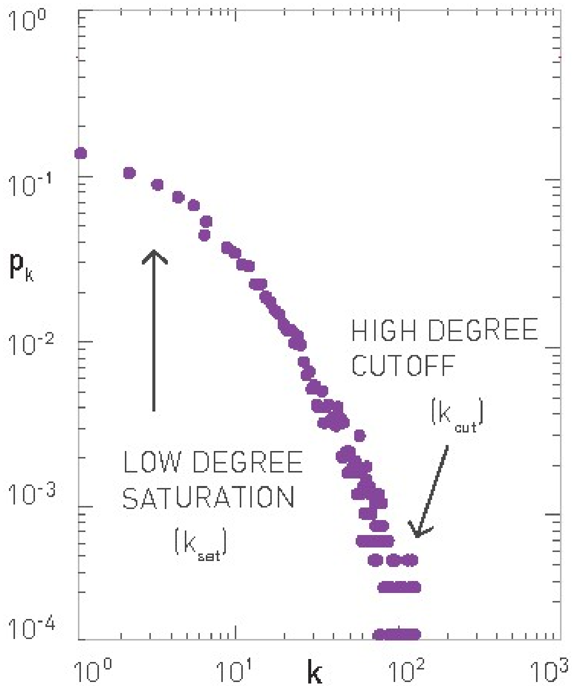


$$P(k) \sim (k+k_0)^{-\gamma}$$

$$k_0 = 1.4,\ \gamma = 2.6.$$

COMMON MISCONCEPTIONS

Http://www.nd.edu/~networks

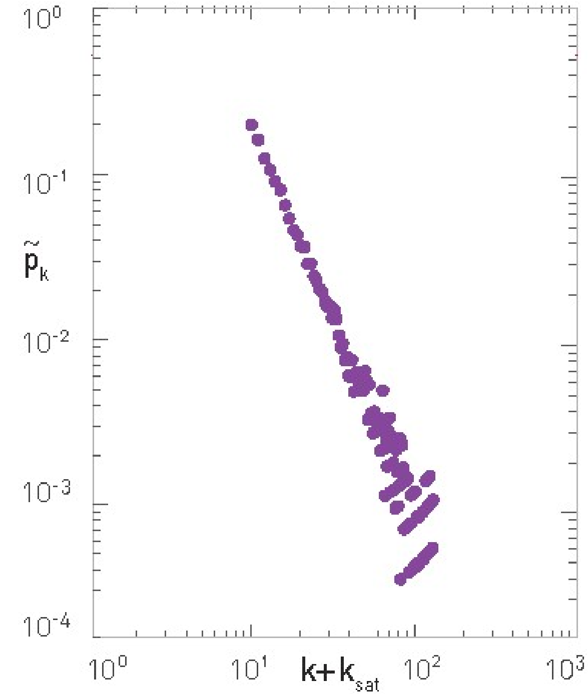


$$p_k = a\left(k + k_{sat}\right)^{-\gamma} \exp\left(-\frac{k}{k_{cut}}\right)$$

$$\widetilde{p}_k = p_k \exp\left(\frac{k}{k_{cut}}\right)$$

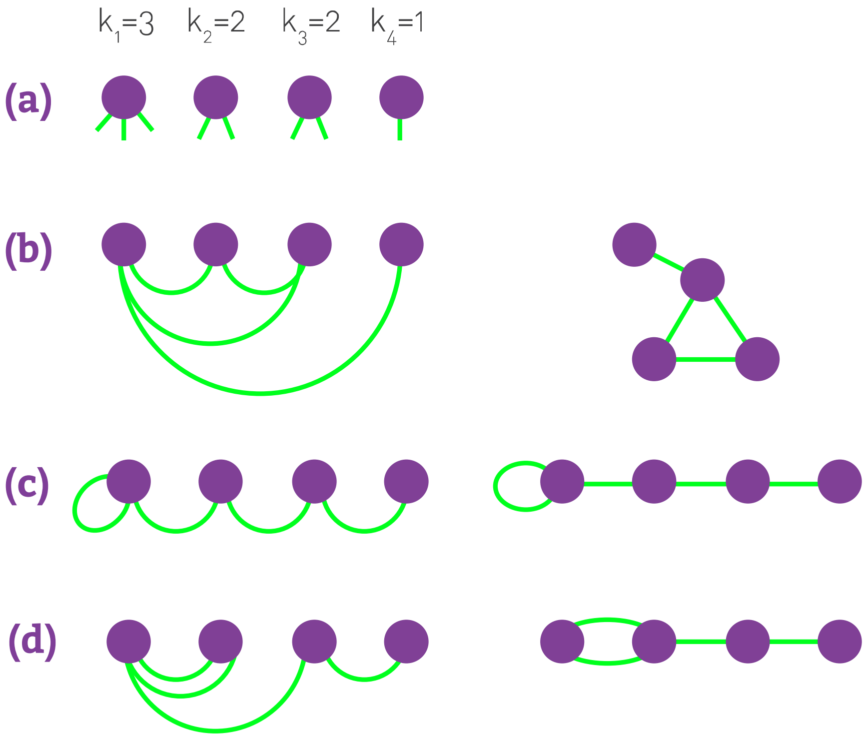
$$\widetilde{k} = k + k_{sat}$$

$$\widetilde{p}_k \sim \widetilde{k}^{-\gamma}$$



Generating networks with a pre-defined p_k

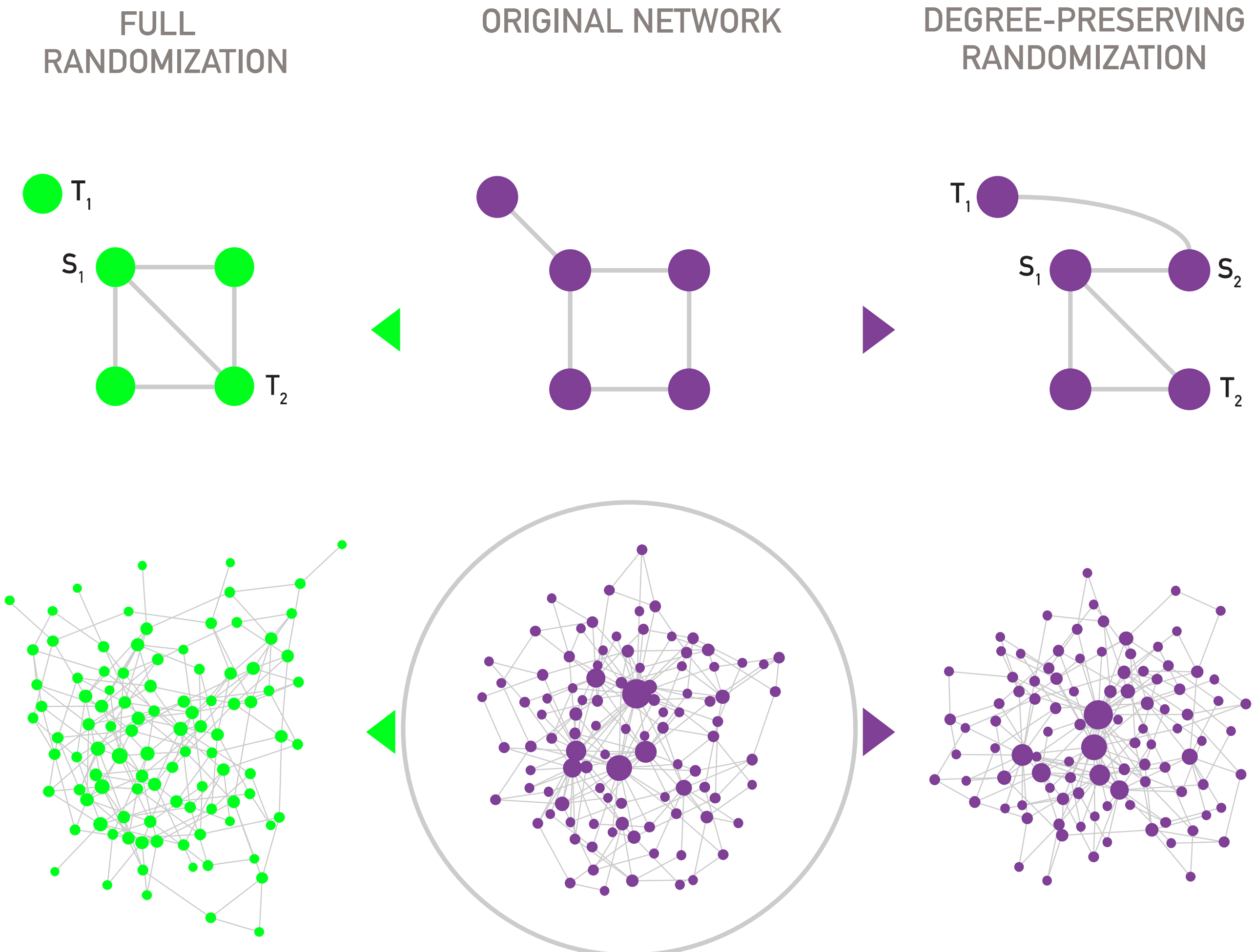
Configuration model



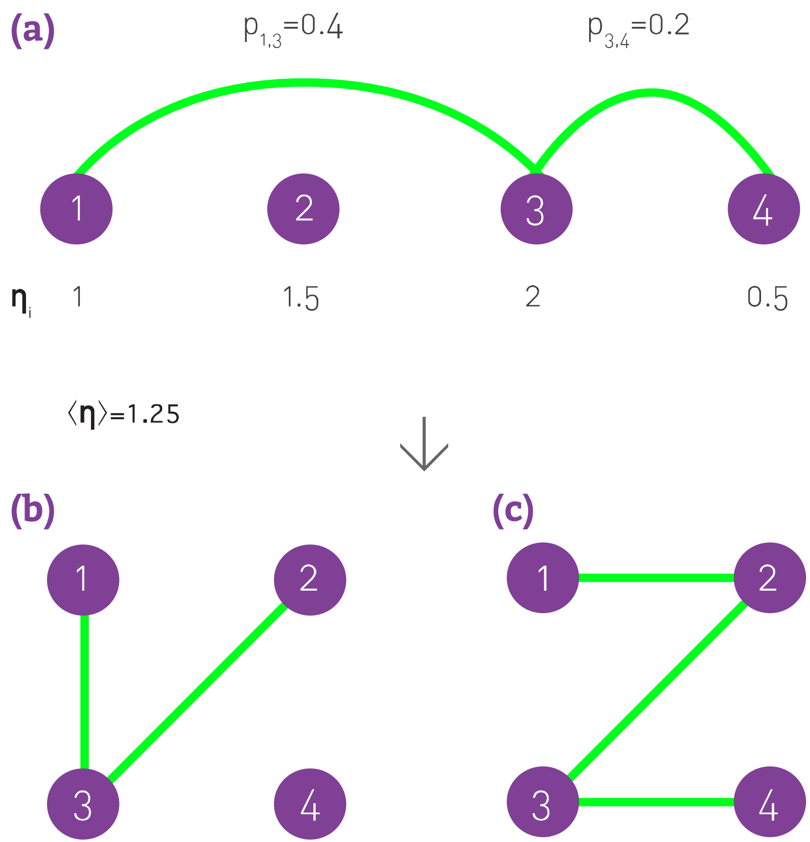
$$p_{ij} = \frac{k_i k_j}{2L - 1}$$

(1) ***Degree sequence****: Assign a degree to each node, represented as stubs or half-links. The degree sequence is either generated analytically from a pre-selected distribution (Box 4.5), or it is extracted from the adjacency matrix of a real network. We must start from an even number of stubs, otherwise we will be left with unpaired stubs.(2)* ***Network assembly****: Randomly select a stub pair and connect them. Then randomly choose another pair from the remaining stubs and connect them. This procedure is repeated until all stubs are paired up. Depending on the order in which the stubs were chosen, we obtain different networks. Some networks include cycles (2a), others self-edges (2b) or multi-edges (2c). Yet, the expected number of self- and multi-edges goes to zero in the limit.*

Degree Preserving randomization



Hidden parameter model



$$p(\eta_i, \eta_j) = \frac{\eta_i \eta_j}{\langle \eta \rangle N}$$

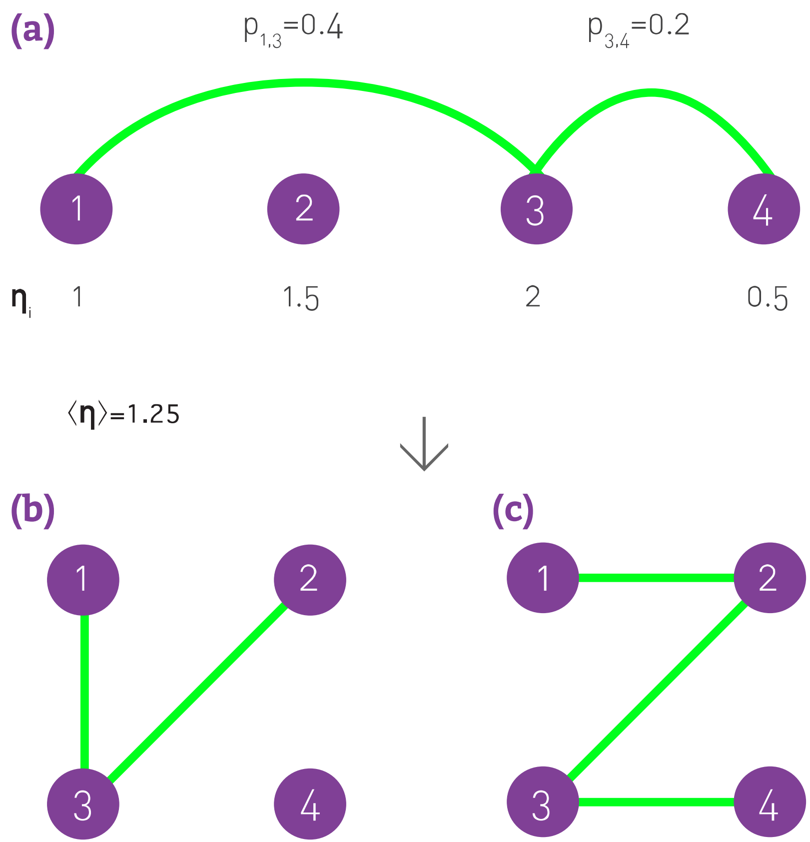
$$p_k = \int \frac{e^{-\eta} \eta^k}{k!} p(\eta) d\eta.$$

$$\{\eta_1, \eta_2, \ldots, \eta_N\}$$

$$p_k = \frac{1}{N} \sum_j \frac{e^{-\eta_j} \eta_j^{\,k}}{k!}.$$

$$\eta_j = \frac{c}{i^{\alpha}}, i = 1, \ldots, N \qquad p_k \sim k^{-(1+\frac{1}{\alpha})}$$

Hidden parameter model



Start with N isolated nodes and assign to each node a "hidden parameter" η , which can be randomly selected from a ρ(η) distribution. We next connect each node pair with probability

$$p(\eta_i,\eta_j)=\frac{\eta_i\eta_j}{\langle\eta\rangle N}$$

For example, the figure shows the probability to connect nodes (1,3) and (3,4). After connecting the nodes, we end up with

the networks shown in (b) or (c), representing two independent realizations generated by the same hidden parameter sequence (a). The expected number of links in the obtained network is

$$L=\frac{1}{2}\sum_{N}^{i,j}\frac{\eta_i\eta_j}{\langle\eta\rangle N}=\frac{1}{2}\langle\eta\rangle N$$

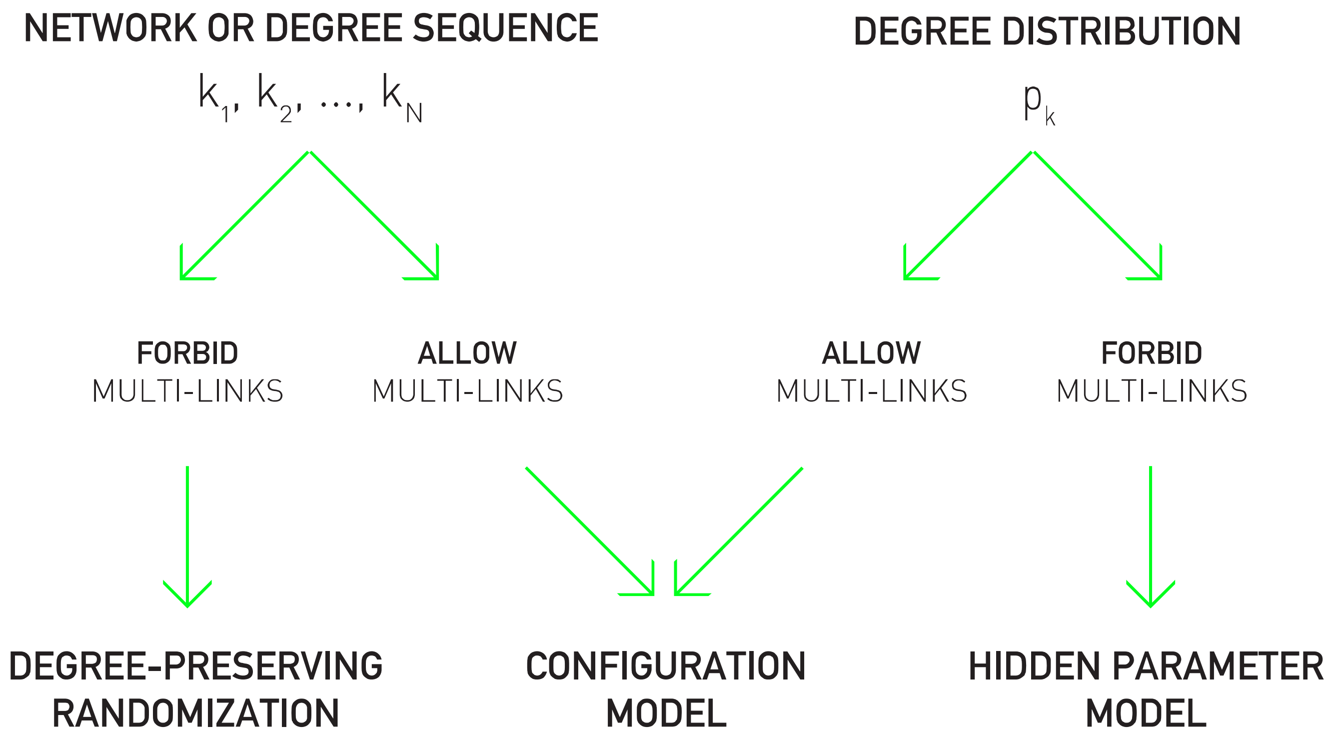
$$\eta_j=\frac{c}{i^{\alpha}},i=1,\ldots,N$$

$$p_k=\int\frac{e^{-\eta}\eta^k}{k!}p(\eta)d\eta.$$

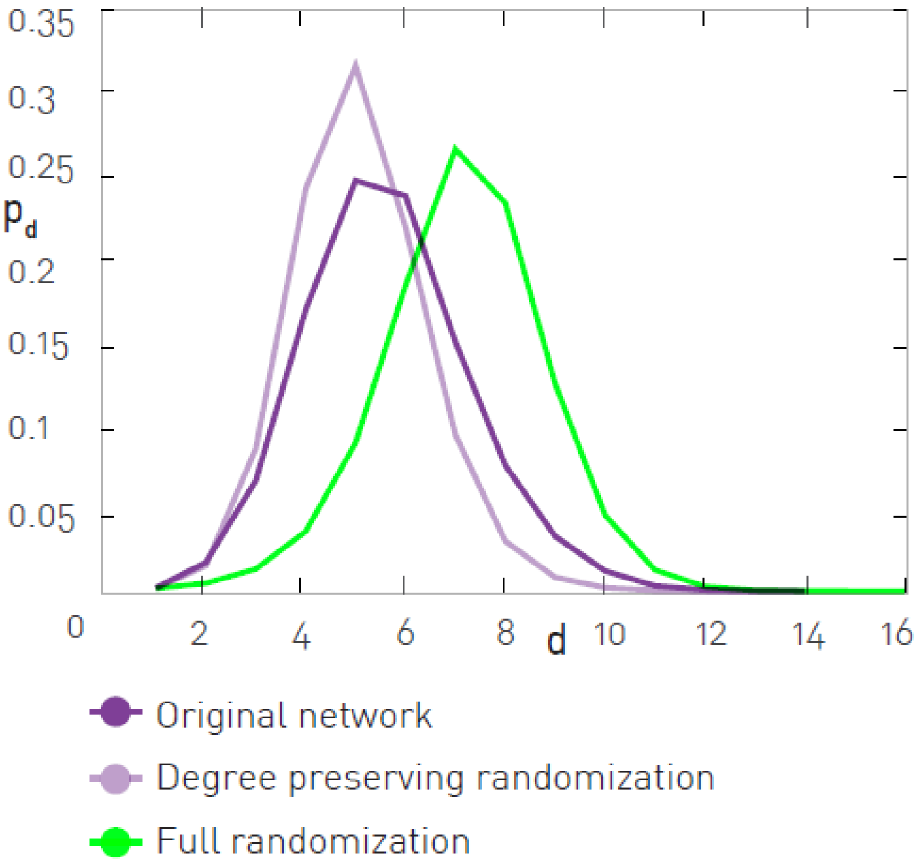
$\{\eta_1, \eta_2, \ldots, \eta_N\}$ $$p_k=\frac{1}{N}\sum_j\frac{e^{-\eta_j}\eta_j^{\,k}}{k!}.$$

$$p_k\sim k^{-(1+\frac{1}{\alpha})}$$

Decision tree



Case Study: PPI Network Distance Distribution



We have: $\langle d \rangle$=5.61±1.64 (original), $\langle d \rangle$=7.13 ± 1.62 (full randomization), $\langle d \rangle$=5.08 ± 1.34 (degree-preserving randomization).

Something to keep in mind

summary

DEGREE DISTRIBUTION

Discrete form:

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}.$$

Continuous form:

$$p(k) = (\gamma - 1)k_{min}^{\gamma-1}\, k^{-\gamma}$$

SIZE OF THE LARGEST HUB

$$k_{max} \sim k_{min} N^{\frac{1}{\gamma-1}}.$$

MOMENTS OF p_k *for* $N \to \infty$

$2 < \gamma < 3$: $\langle k \rangle$ finite, $\langle k^2 \rangle$ diverges.

$\gamma > 3$: $\langle k \rangle$ and $\langle k^2 \rangle$ finite.

DISTANCES

$$\langle d \rangle \sim \begin{cases} \text{const.} & \gamma = 2, \\ \dfrac{\ln\ln N}{\ln(\gamma - 1)} & 2 < \gamma < 3, \\ \dfrac{\ln N}{\ln\ln N} & \gamma = 3, \\ \ln N & \gamma > 3. \end{cases}$$

Bounded Networks

We call a network *bounded* if its degree distribution decrease exponentially or faster for high *k*. As a consequence $<k^2>$ is smaller than $<k>$, implying that we lack significant degree variations. Examples of p_k in this class include the Poisson, Gaussian, or the simple exponential distribution (Table 4.2). The Erdős-Rényi and the Watts-Strogatz networks are the best known network models belonging to this class. Bounded networks lack outliers, consequently most nodes have comparable degrees. Real networks in this class include highway networks and the power grid.

Unbounded Networks

We call a network *unbounded* if its degree distribution has a fat tail in the high-*k* region. As a consequence $<k^2>$ is much larger than $<k>$, resulting in considerable degree variations. Scale-free networks with a power-law degree distribution (4.1) offer the best known example of networks belonging to this class. Outliers, or exceptionally high-degree nodes, are not only allowed but are expected in these networks. Networks in this class include the WWW, the Internet, the protein interaction networks, and most social and online networks.

CLASS INFORMATION