Frontiers of Network Science
Fall 2021

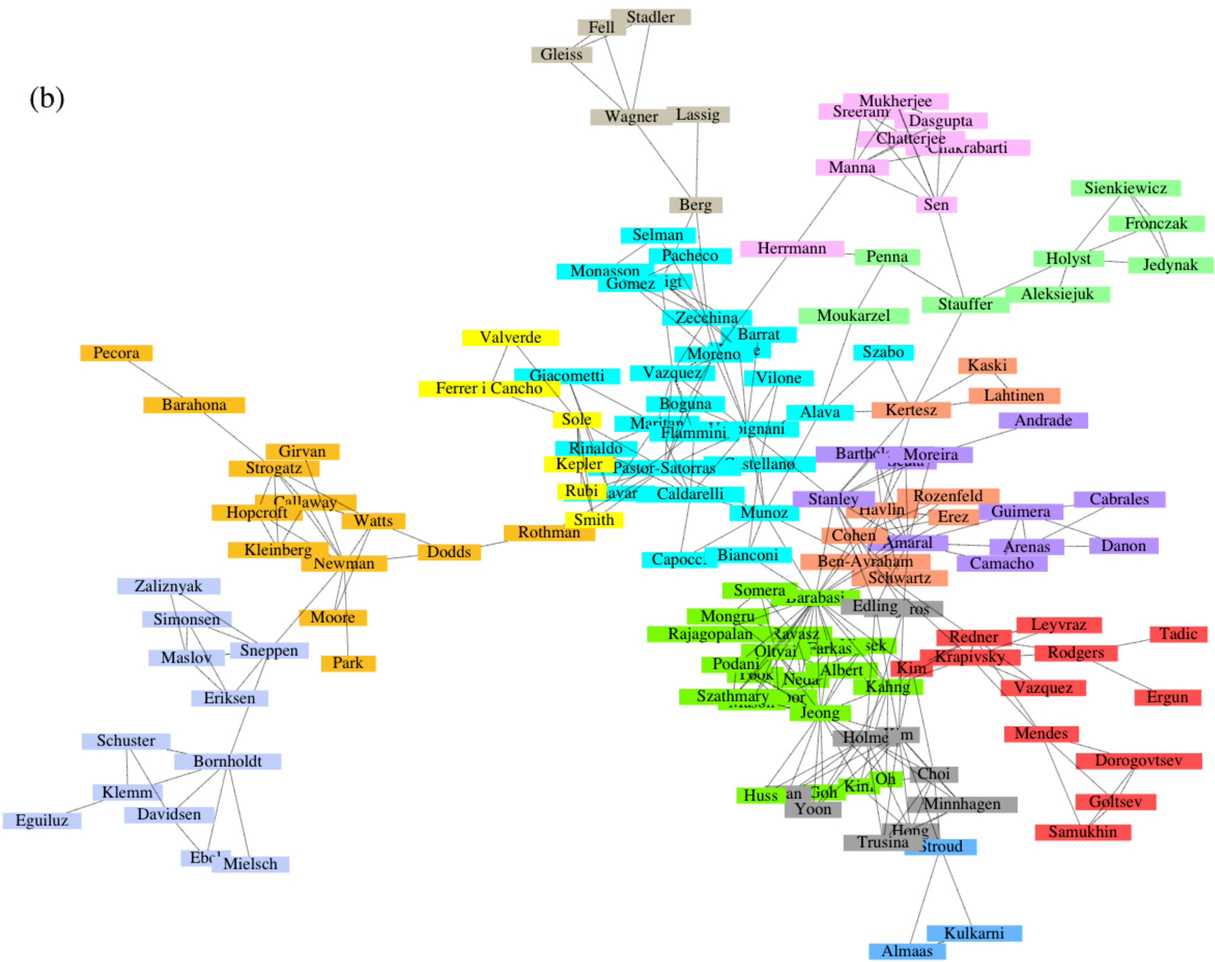
Class 8: Scale Free Networks
(Chapter 4 in Textbook)

Boleslaw Szymanski

based on slides by Albert-László Barabási & Roberta Sinatra

NETWORK DATA: SCIENCE COLLABORATION NETWORKS

(b)

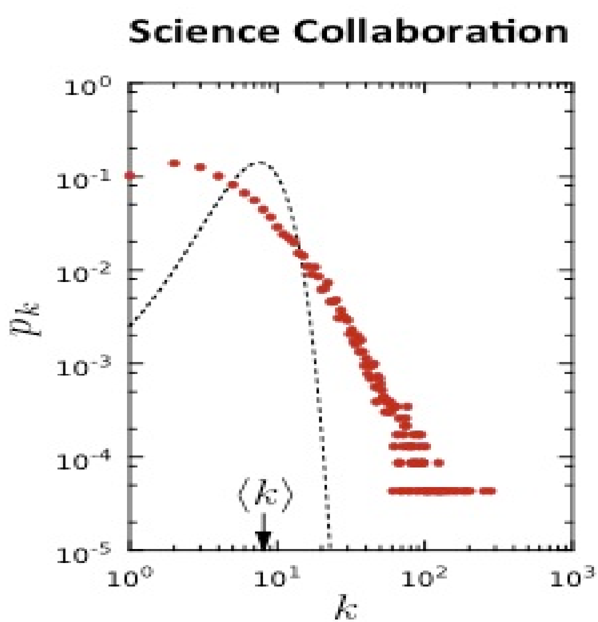


Collaboration Network:
[Nodes](#): Scientists
[Links](#): Joint publications

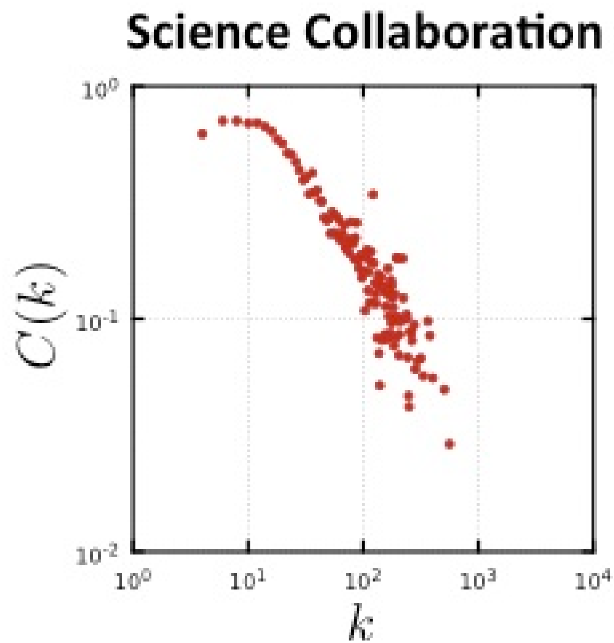
Physical Review:
1893 – 2009.

N=449,673
L=4,707,958

See also Stanford Large Network database
http://snap.stanford.edu/data/#canets.

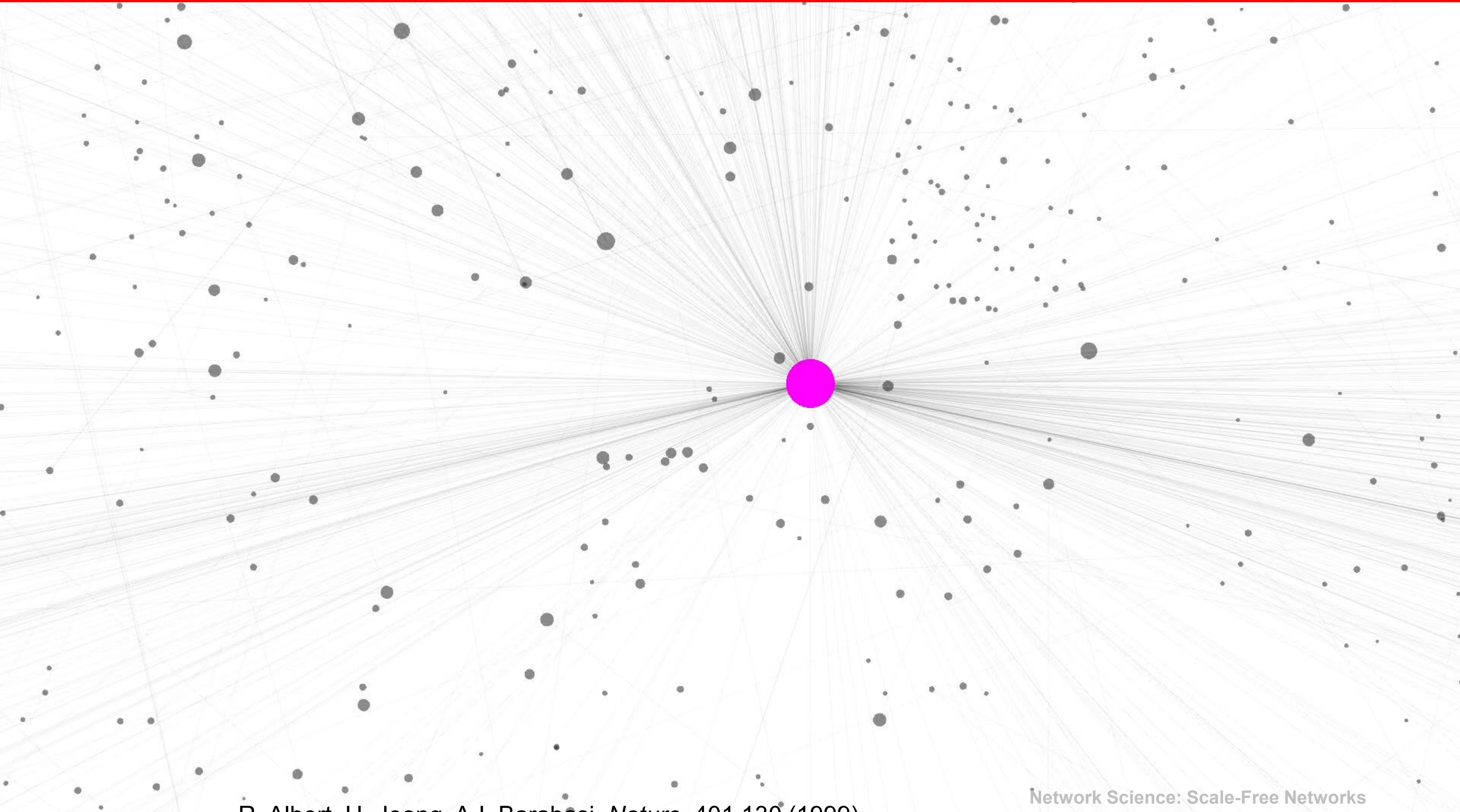


Scale-free



Hierarchical

WORLD WIDE WEB



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

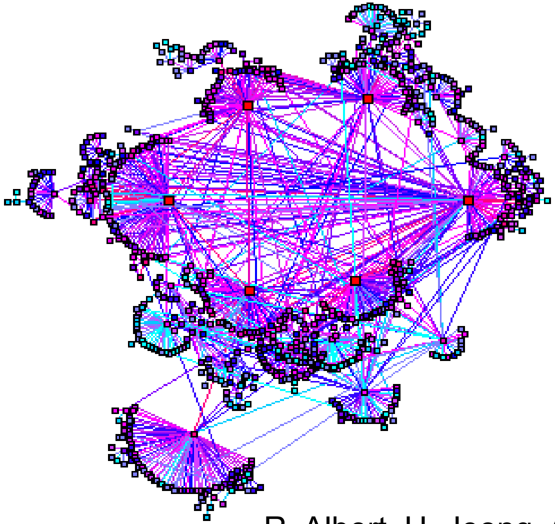
Power laws and scale-free networks

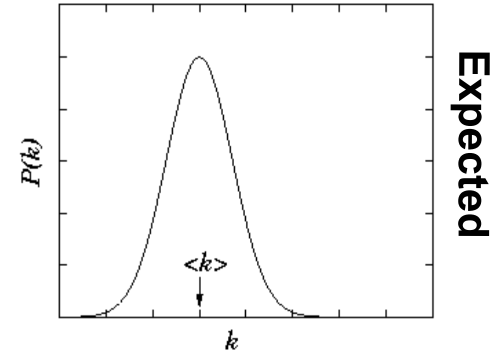
WORLD WIDE WEB

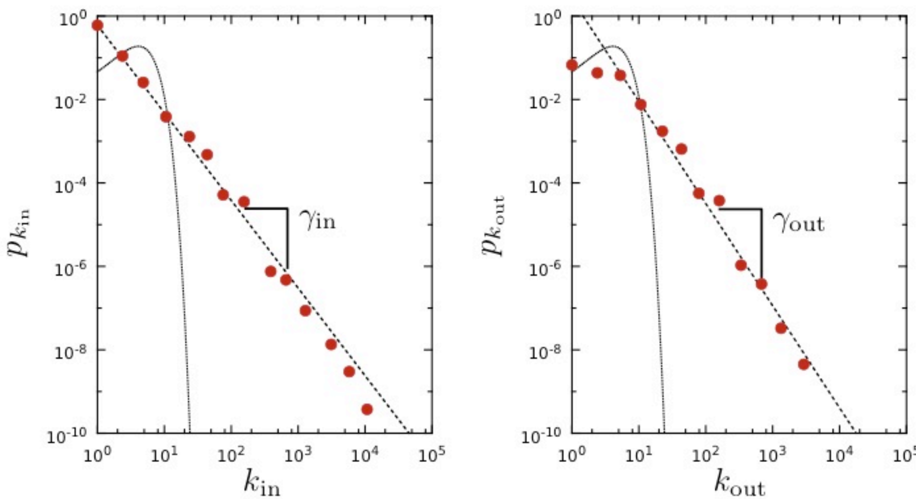
Nodes: **WWW documents**
Links: **URL links**

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:

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

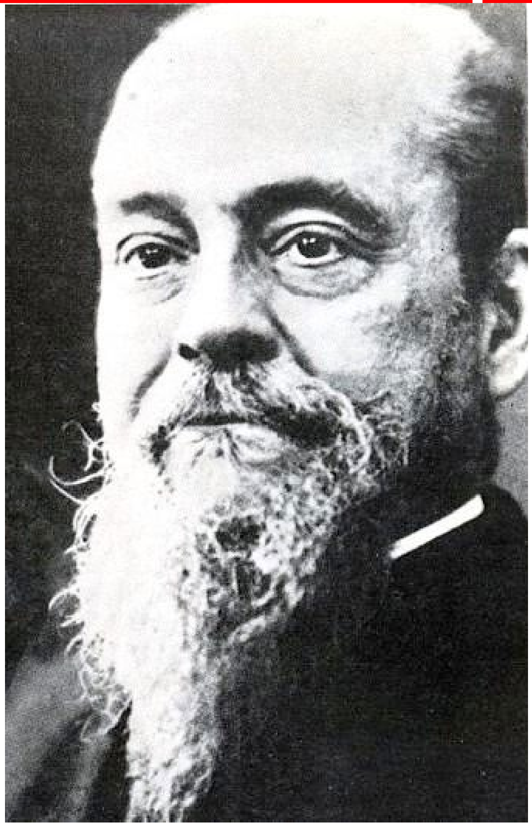
$$\int_{k_{\min}}^{\infty} p(k)dk = 1$$

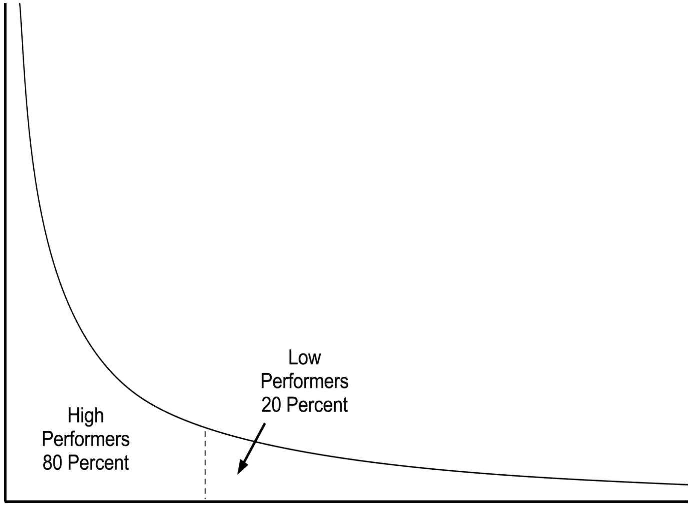
$$C = \frac{1}{\int_{k_{\min}}^{\infty} k^{-\gamma} dk} = (\gamma - 1)k_{\min}^{\gamma-1}$$

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

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

80/20 RULE

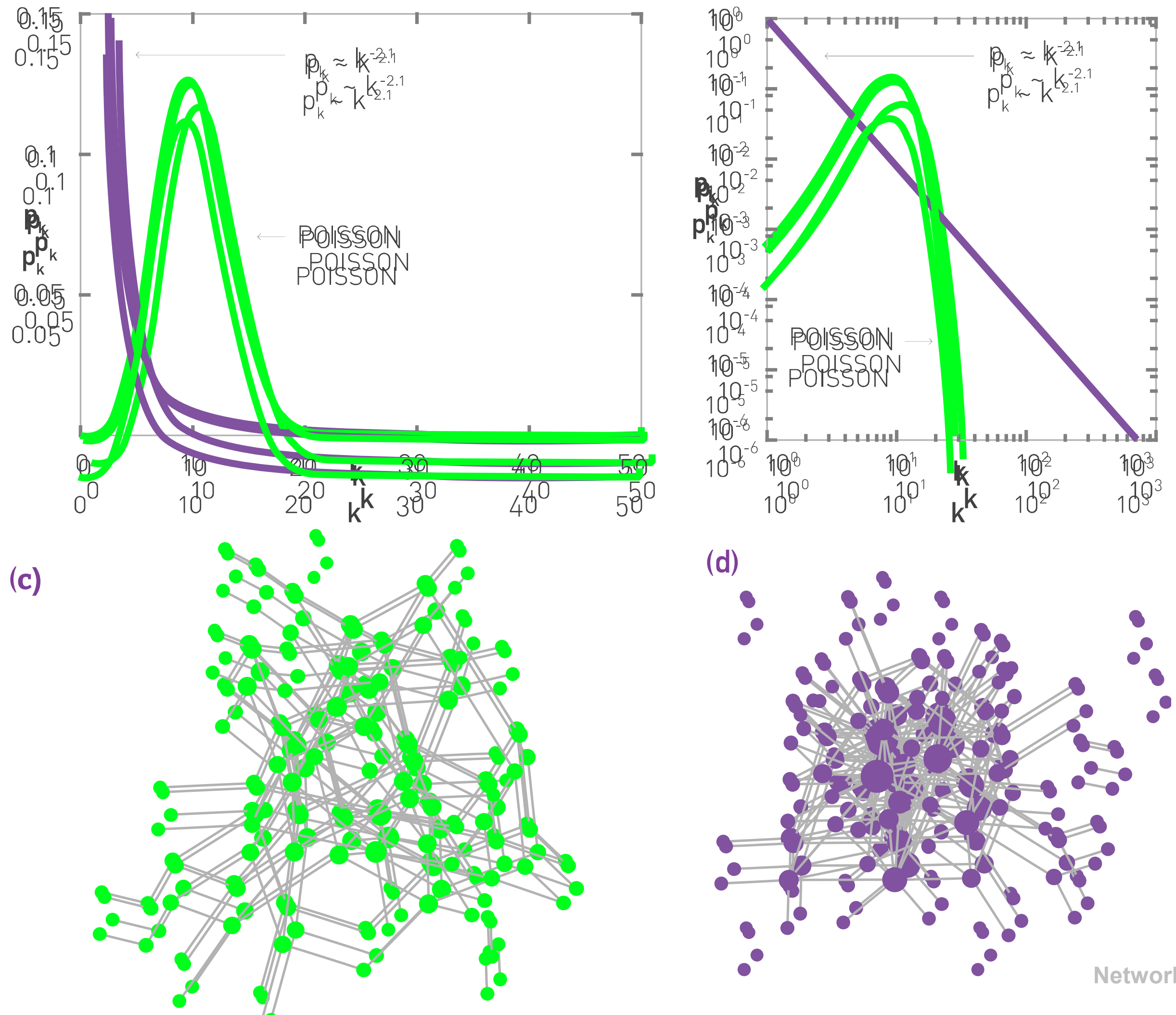




Vilfredo Federico Damaso Pareto (1848 – 1923), Italian economist, political scientist and philosopher, who had important contributions to our understanding of income distribution and to the analysis of individuals choices. A number of fundamental principles are named after him, like Pareto efficiency, 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.*

- *For a power law degree distribution, we expect about* $N_{k>100} = 10^9$ *k>100 degree nodes*

geles

The size of the biggest hub

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

<u>Estimating k_{max}</u>

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

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}

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

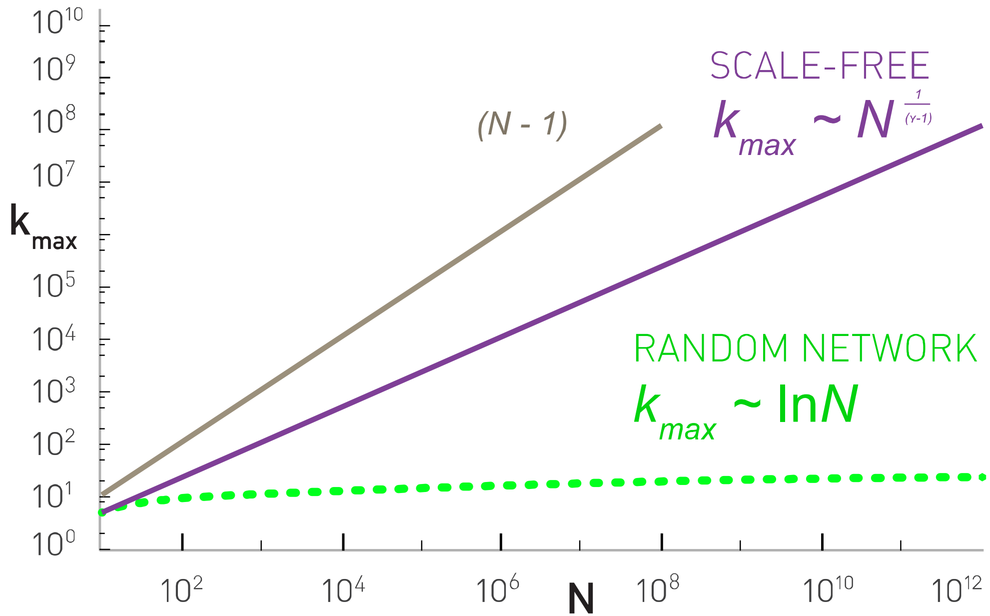
- k_{max}, increases with the size of the network
  →the larger a system is, the larger its biggest hub

- 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}~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'

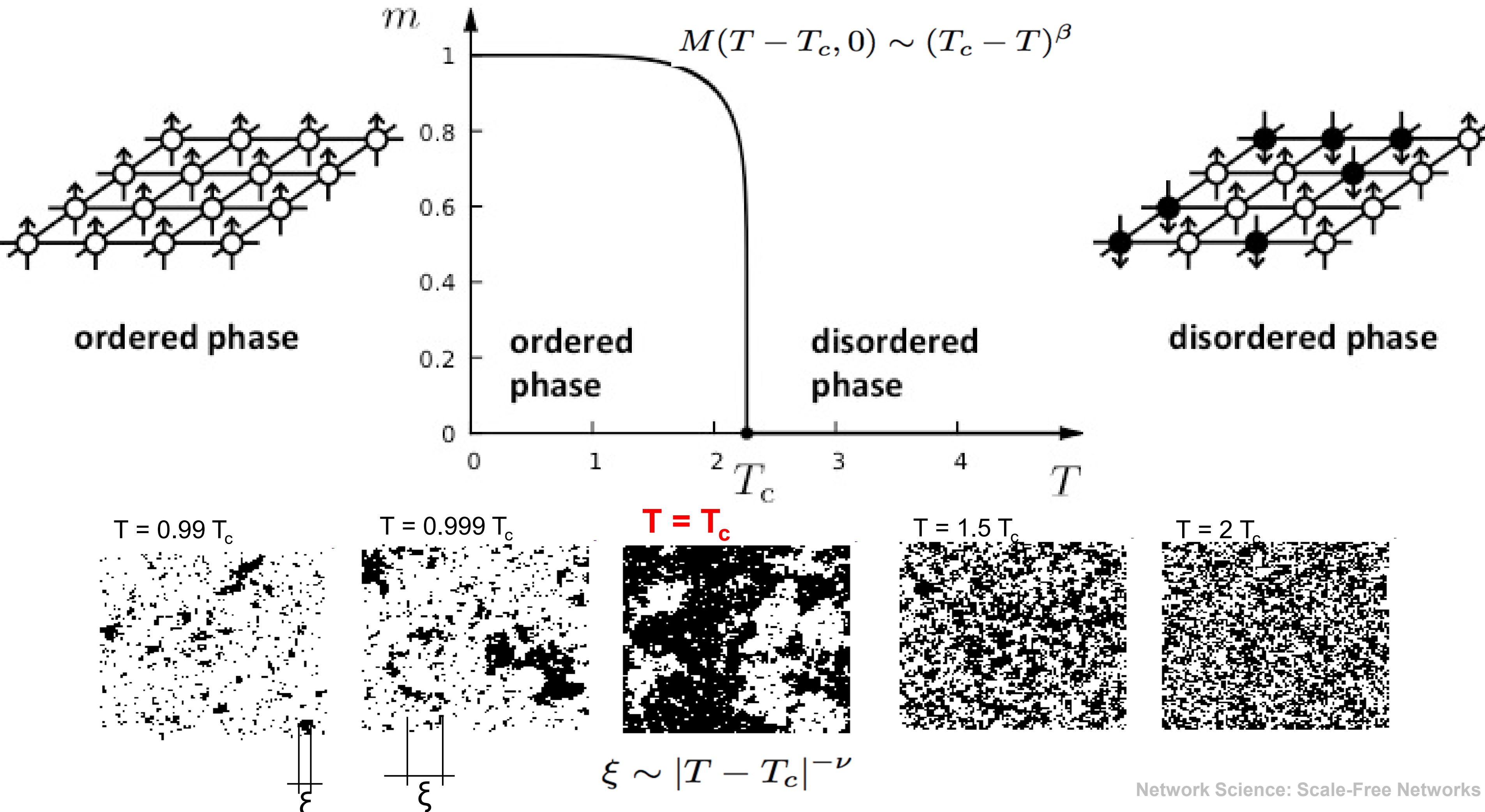
Where does the name come from?

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

Scale invariance at the critical point

by Douglas Ashton

www.kineticallyconstrained.com

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 γ this means all moments m>γ-1 diverge.

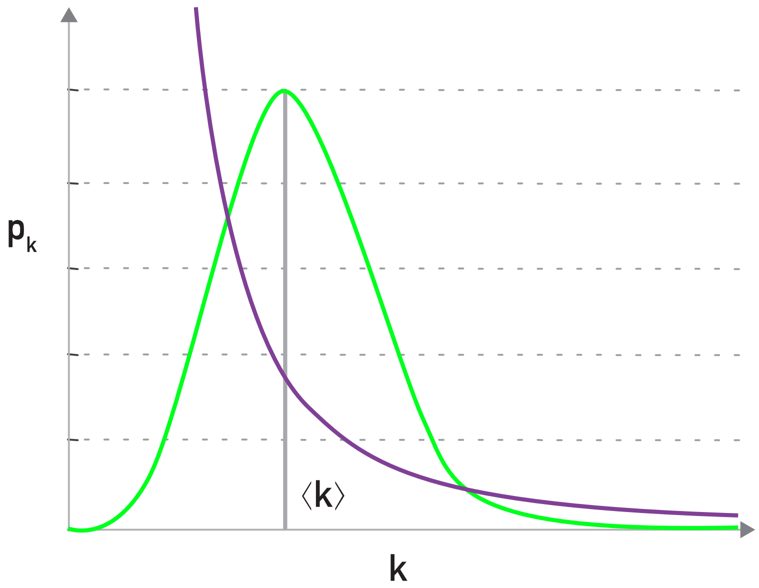
Network	N	L	$\langle k \rangle$	$\langle k_{in}^2 \rangle$	$\langle k_{out}^2 \rangle$	$\langle k^2 \rangle$	γ_{in}	γ_{out}	γ
Internet	192,244	609,066	6.34	-	-	240.1	-	-	3.42*
WWW	325,729	1,497,134	4.60	1546.0	482.4	-	2.00	2.31	-
Power Grid	4,941	6,594	2.67	-	-	10.3	-	-	Exp.
Mobile-Phone Calls	36,595	91,826	2.51	12.0	11.7	-	4.69*	5.01*	-
Email	57,194	103,731	1.81	94.7	1163.9	-	3.43*	2.03*	-
Science Collaboration	23,133	93,437	8.08	-	-	178.2	-	-	3.35*
Actor Network	702,388	29,397,908	83.71	-	-	47,353.7	-	-	2.12*
Citation Network	449,673	4,689,479	10.43	971.5	198.8	-	3.03*	4.00*	-
E. Coli Metabolism	1,039	5,802	5.58	535.7	396.7	-	2.43*	2.90*	-
Protein Interactions	2,018	2,930	2.90	-	-	32.3	-	-	2.89*-

Many degree exponents are smaller than 3

→ $\langle k^2 \rangle$ diverges in the N→∞ limit!!!

→ $\langle k \rangle$ diverges in the N→∞ limit!!!

The meaning of scale-free



Random Network

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

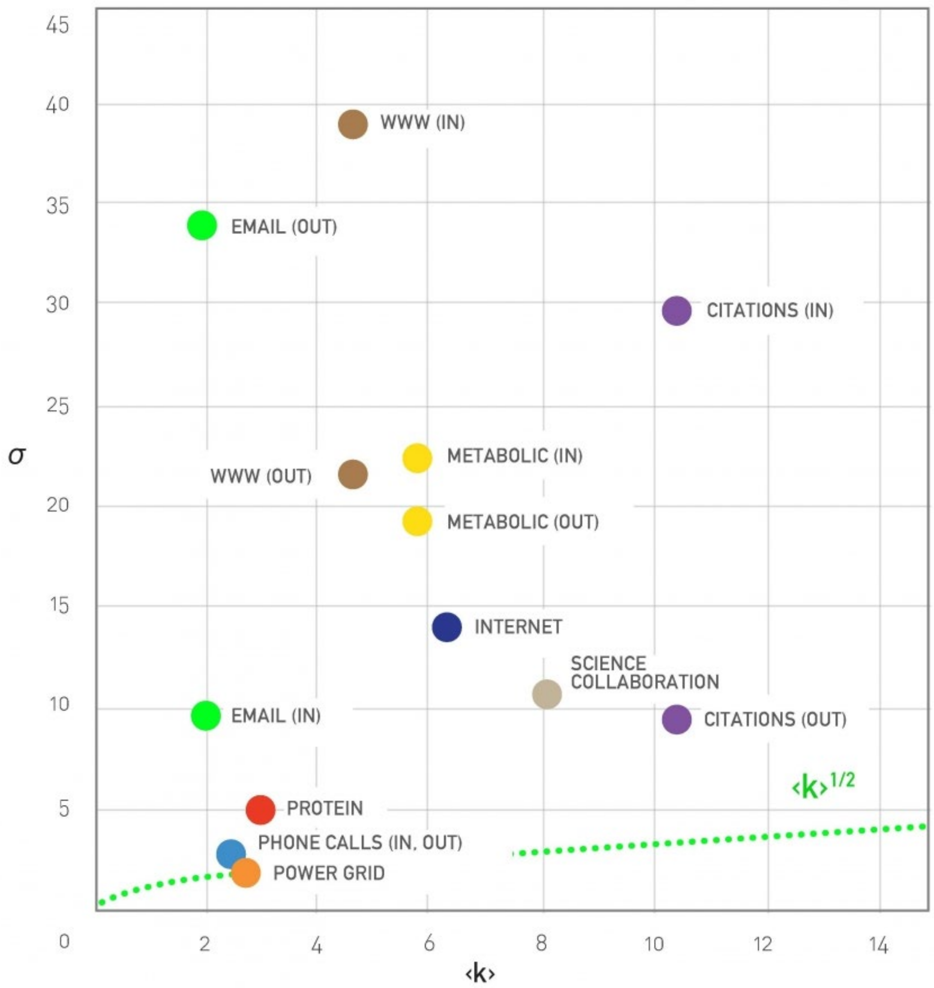
Scale: $\langle k \rangle$

Scale-Free Network

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

Scale: none

The meaning of scale-free

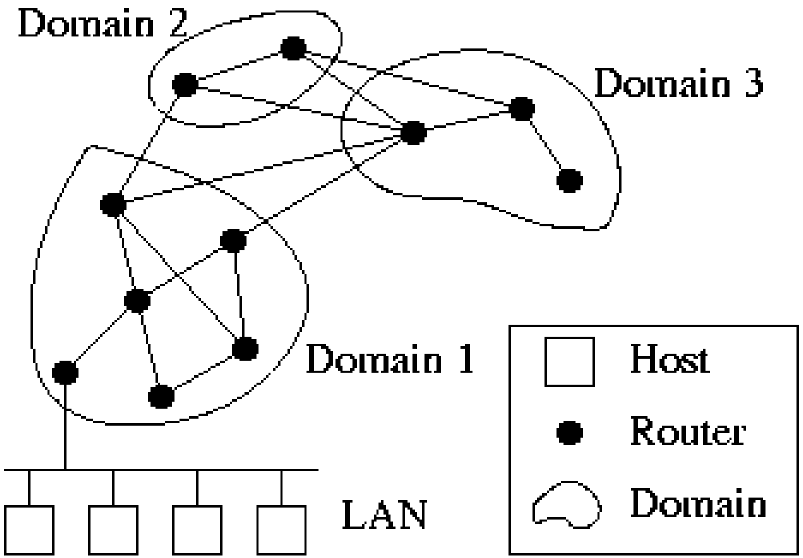


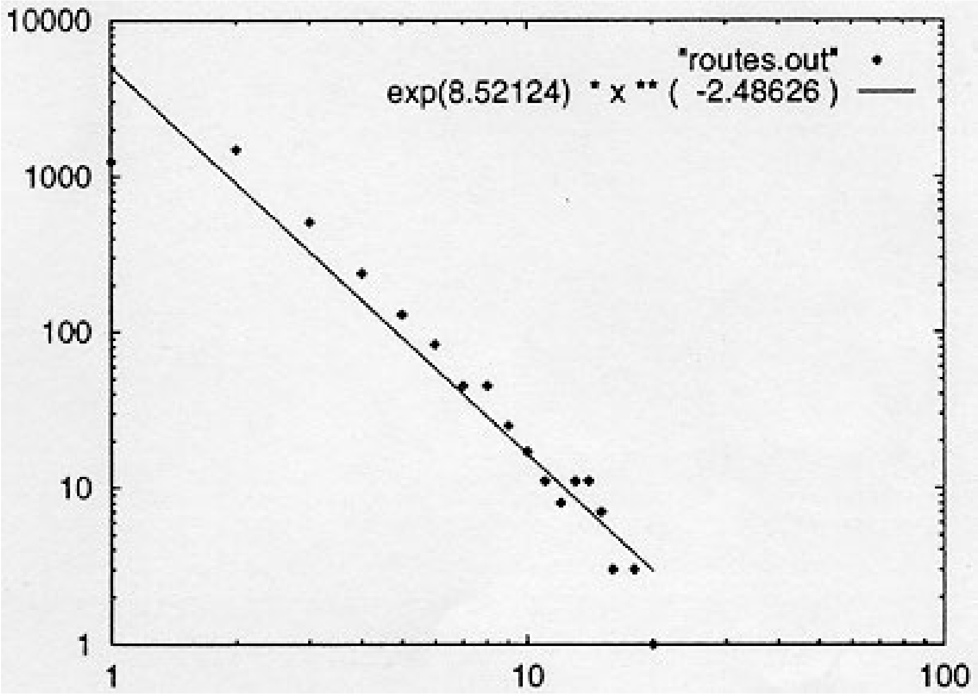
$$k = \langle k \rangle \pm \sigma_k$$

universality

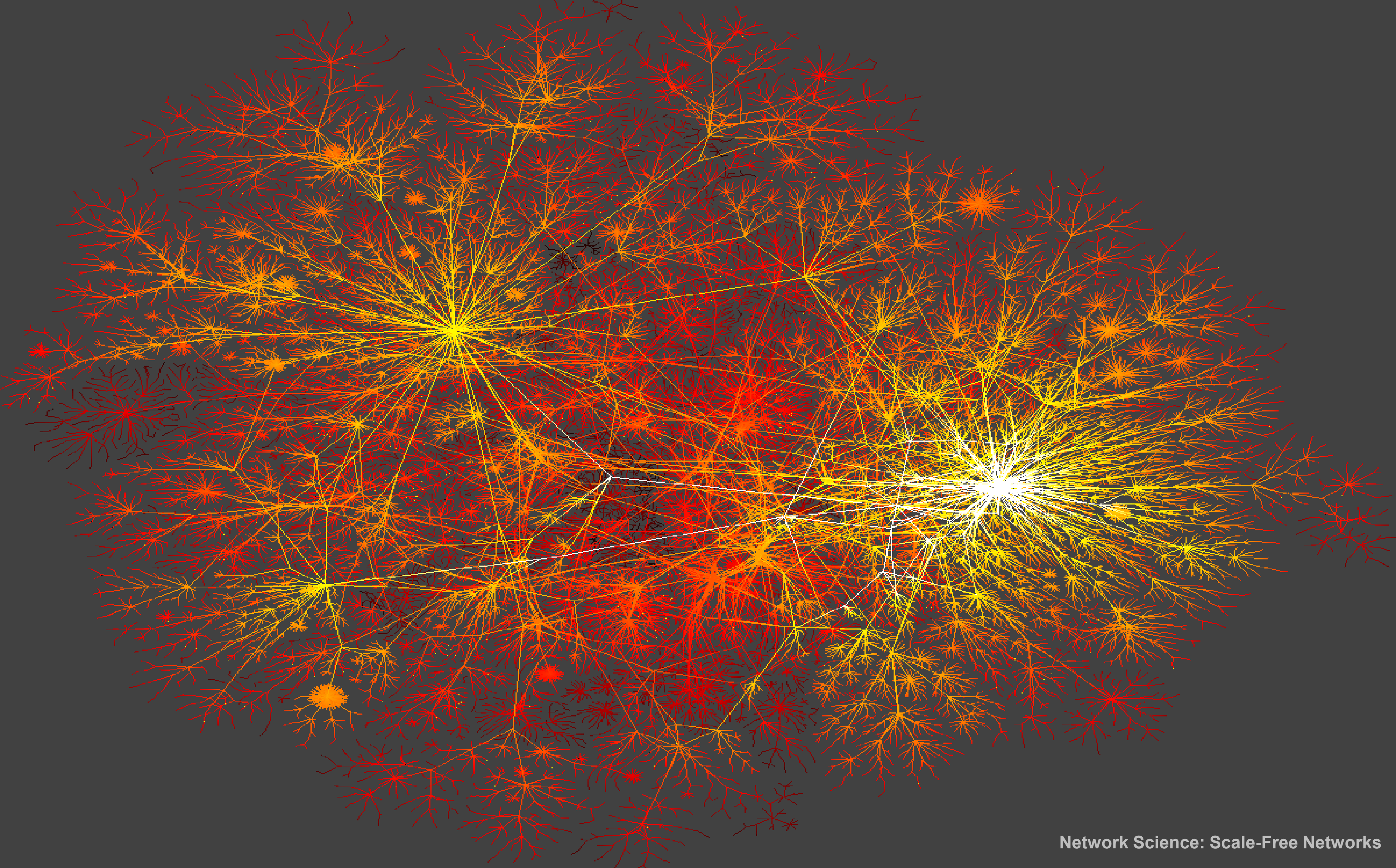
INTERNET BACKBONE

Nodes: computers, routers
Links: physical lines





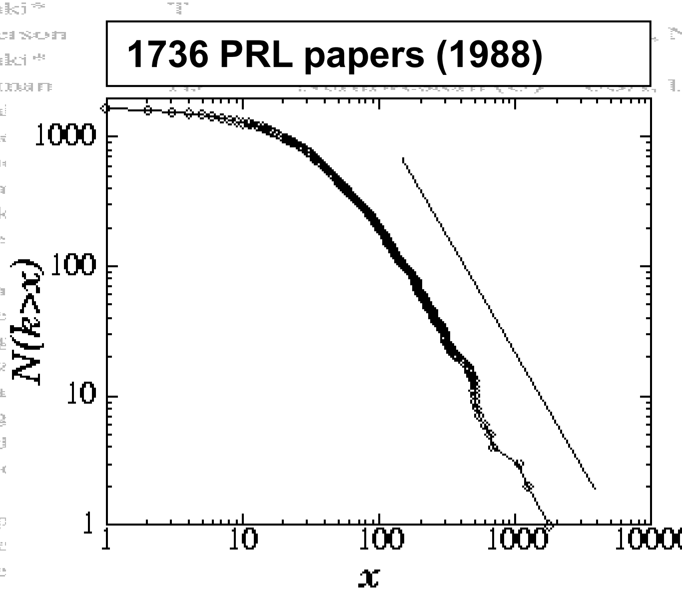
(Faloutsos, Faloutsos and Faloutsos, 1999)

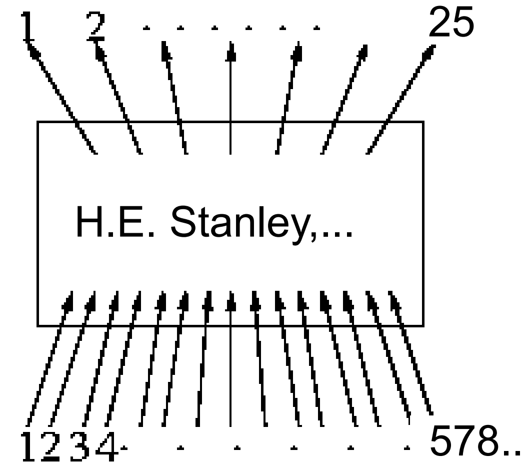


Network Science: Scale-Free Networks

SCIENCE CITATION INDEX

Nodes: papers
Links: citations



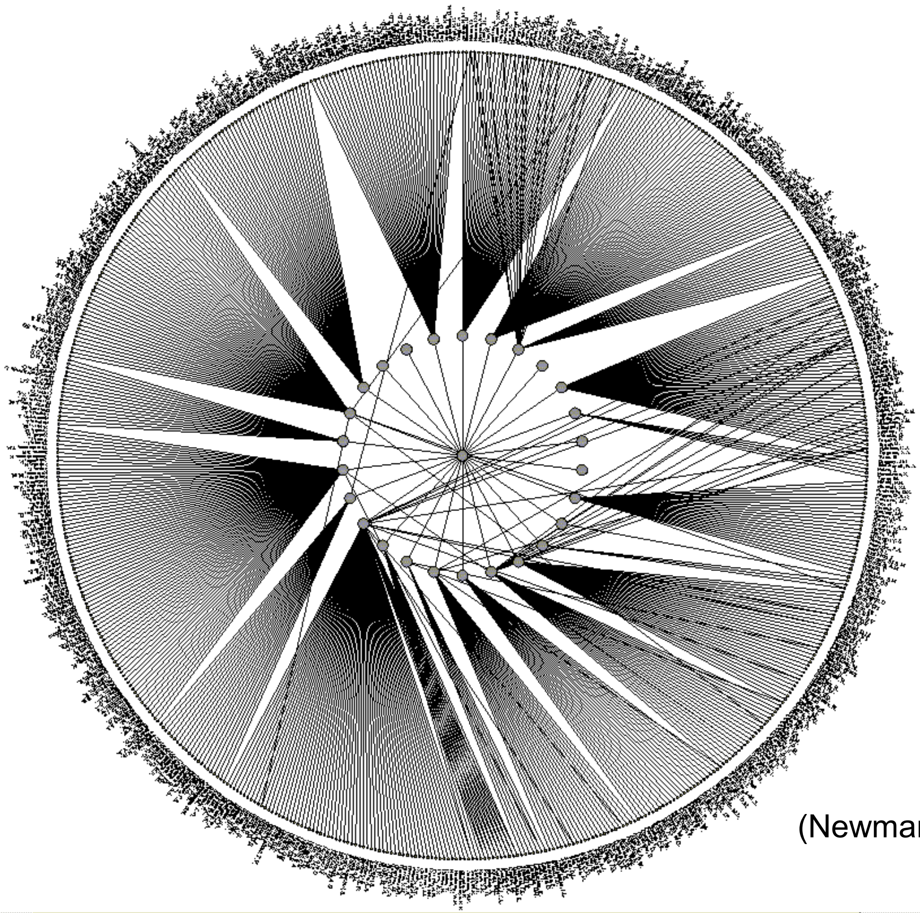


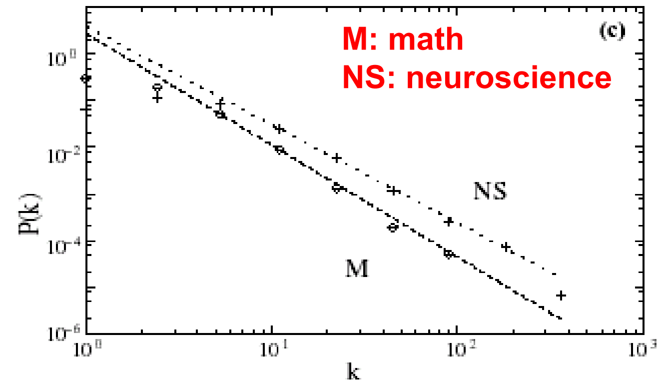
$P(k) \sim k^{-\gamma}$
($\gamma = 3$)

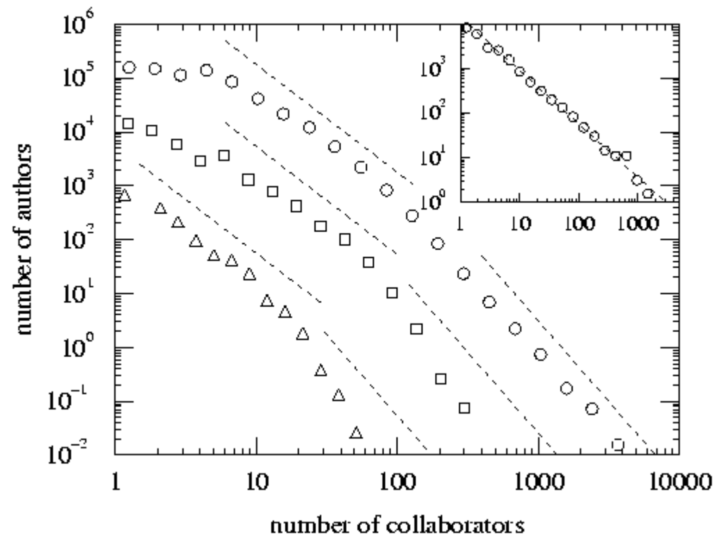
(S. Redner, 1998)

SCIENCE COAUTHORSHIP

Nodes: scientist (authors)
Links: joint publication







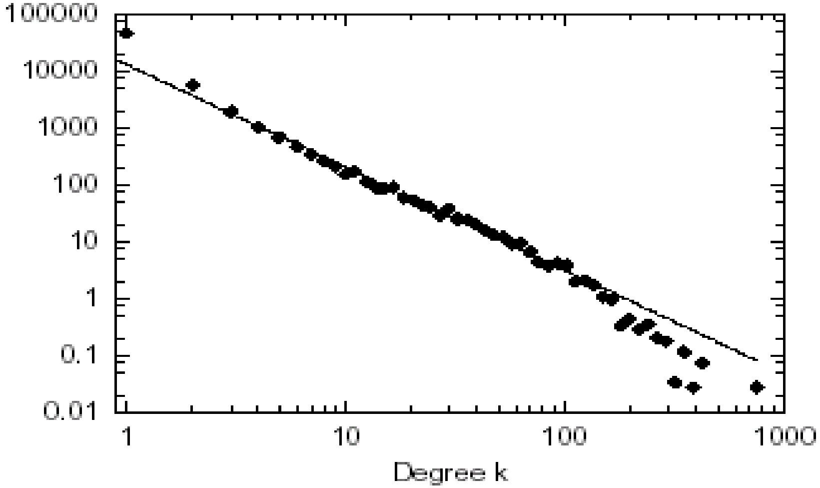
(Newman, 2000, Barabasi et al 2001)

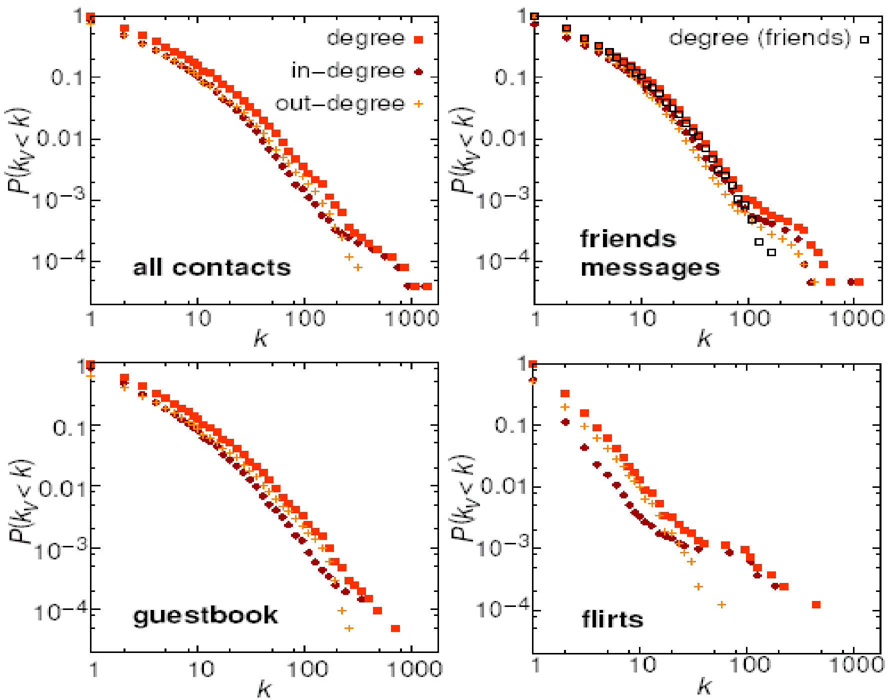
ONLINE COMMUNITIES

Nodes: online user
Links: email contact

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

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



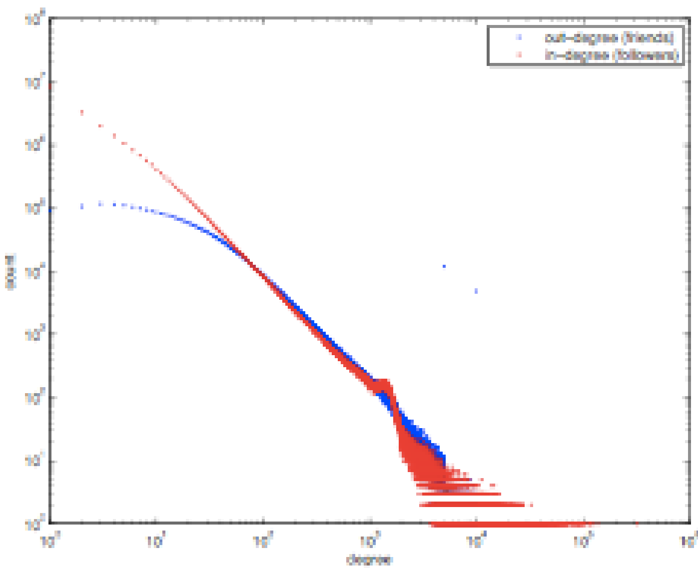


Ebel, Mielsch, Bornholdtz, PRE 2002.

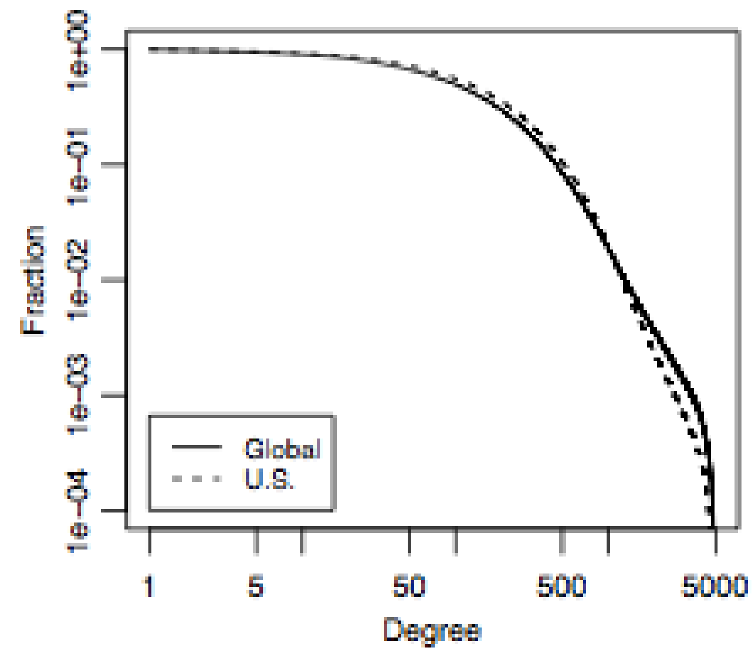
Holme, Edling, Liljeros, 2002.

ONLINE COMMUNITIES

Twitter:



Facebook



Brian Karrer, Lars Backstrom, Cameron Marlowm 2011