Network Science

Class 4: Scale-free property

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

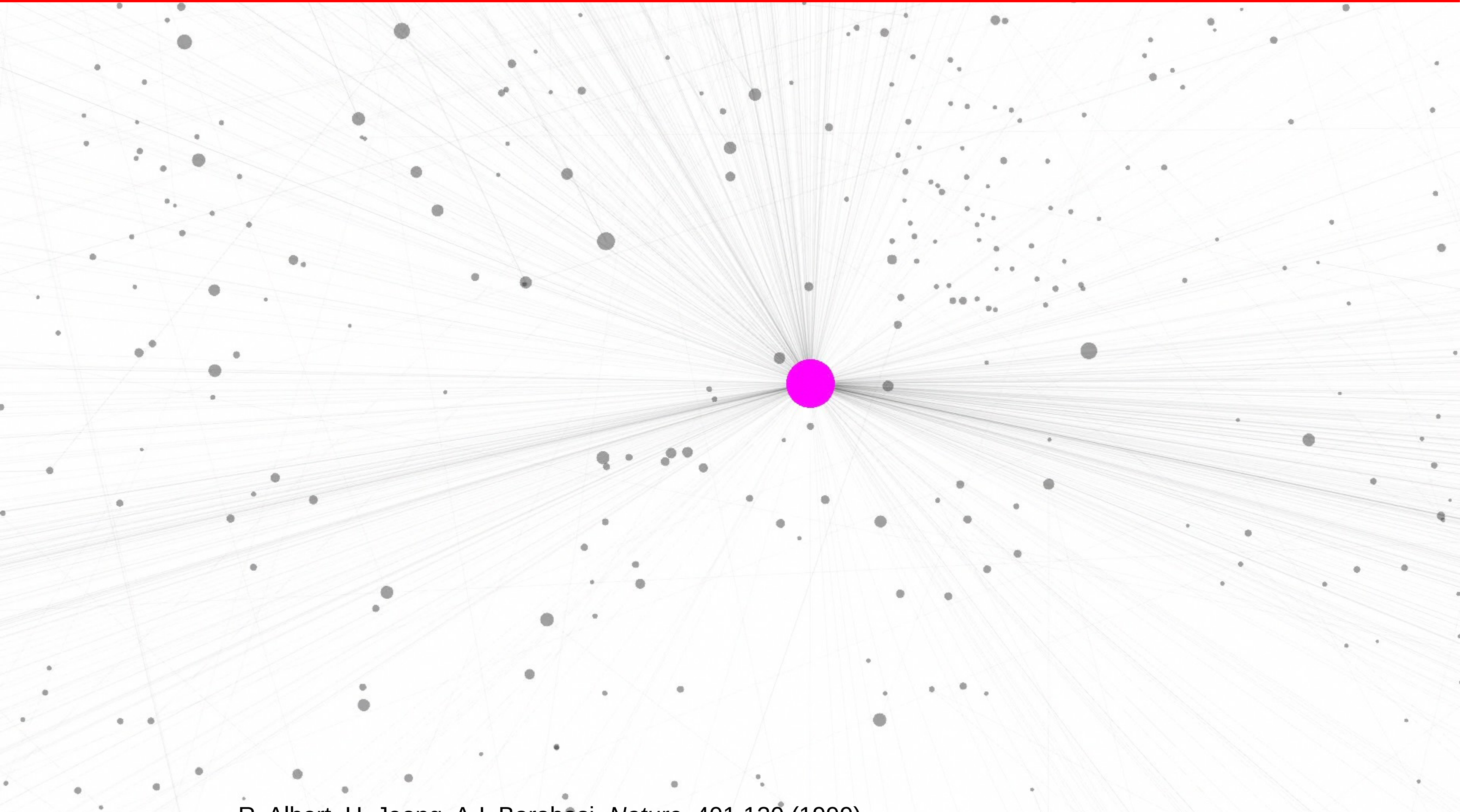
www.BarabasiLab.com

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).

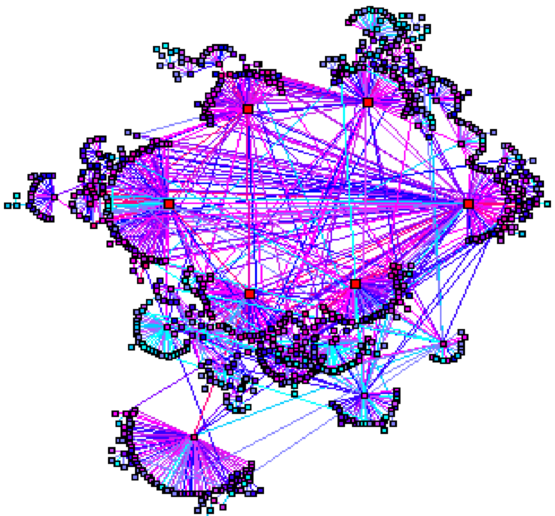
Power laws and scale-free networks

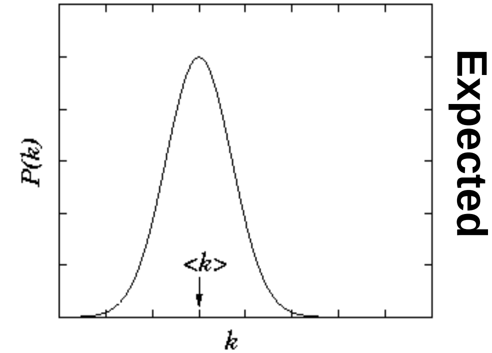
WORLD WIDE WEB

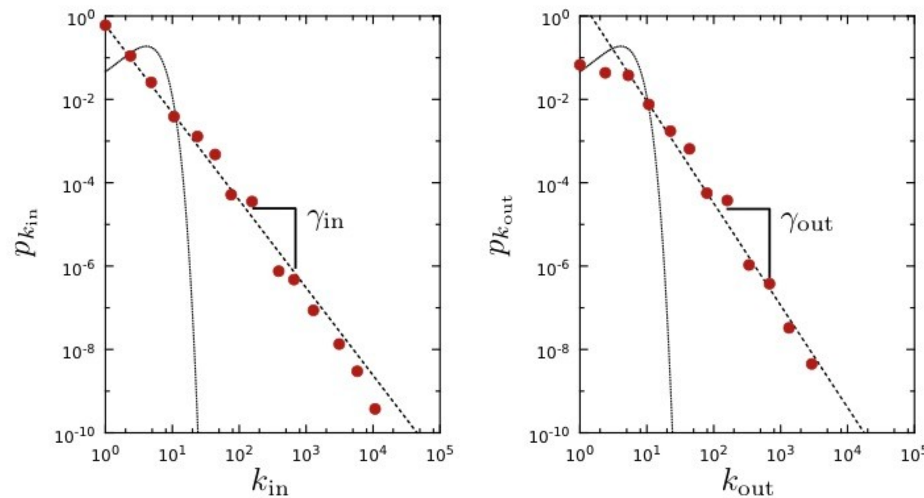
Nodes: **WWW documents**
Links: **URL links**

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:

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

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

$$C\sum_{k=1}^{\infty} k^{-\gamma} = 1$$

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

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

Continuum Formalism

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

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

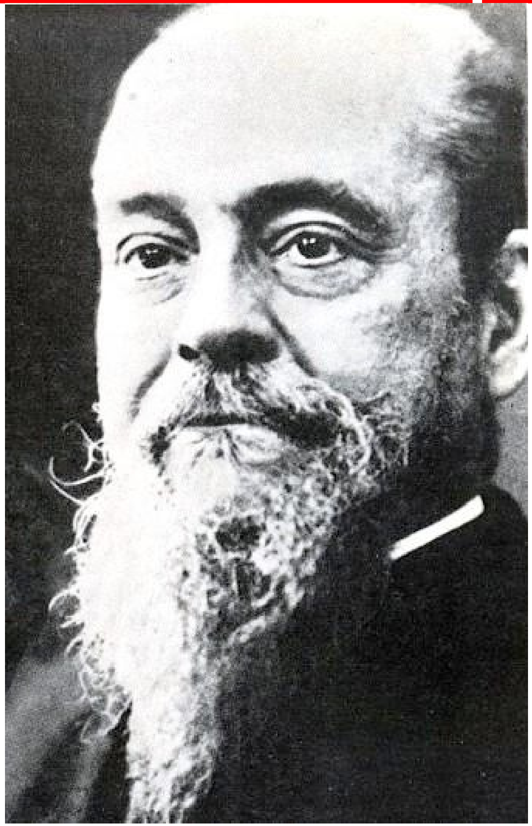
$$\int_{k_{\min}}^{\infty} p(k)dk = 1$$

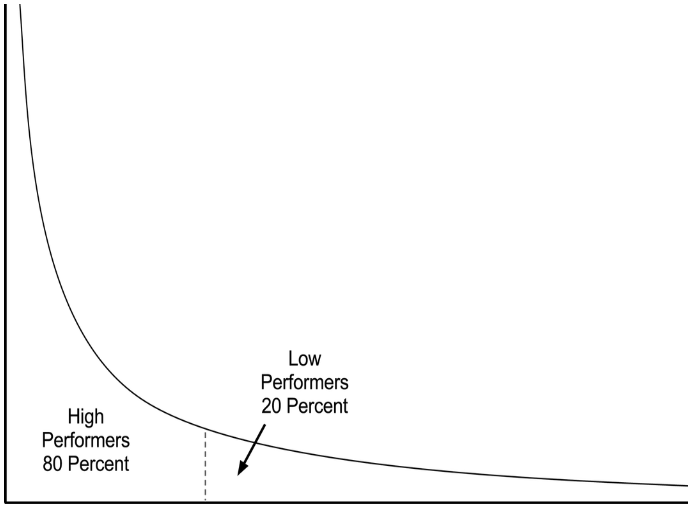
$$C = \frac{1}{\int_{k_{\min}}^{\infty} k^{-\gamma} dk} = (\gamma - 1)k_{\min}^{\gamma-1}$$

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

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

80/20 RULE

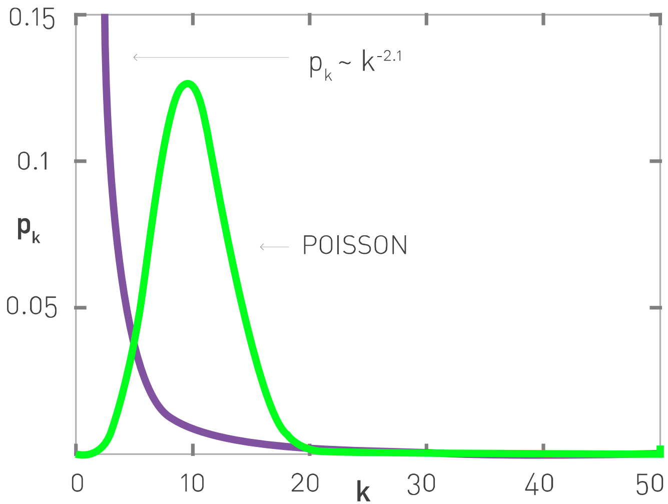


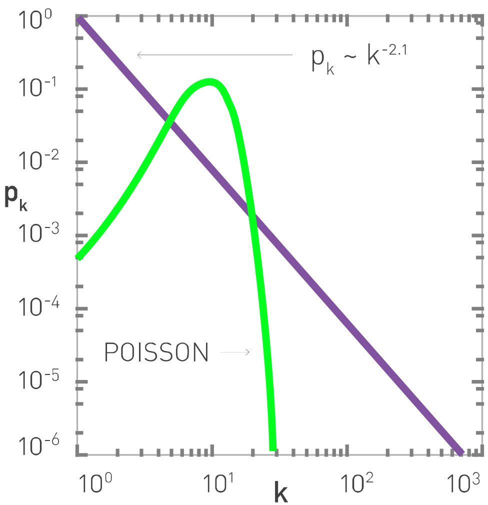


Vilfredo Federico Damaso Pareto (1848 – 1923), Italian economist, political scientist and philosopher, who had important contributions to our understanding of income distribution and to the analysis of individuals choices. A number of fundamental principles are named after him, like Pareto efficiency,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

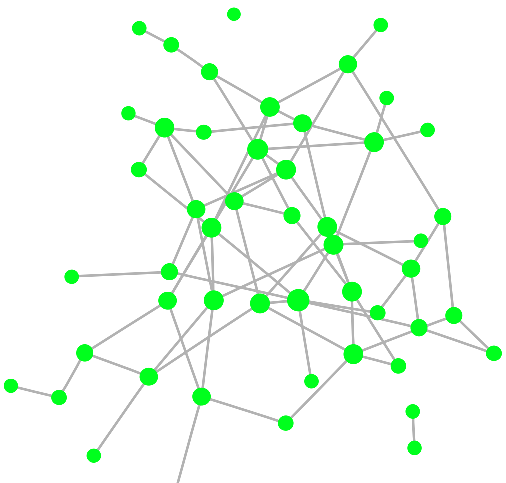
Hubs

The difference between a power law and an exponential distribution

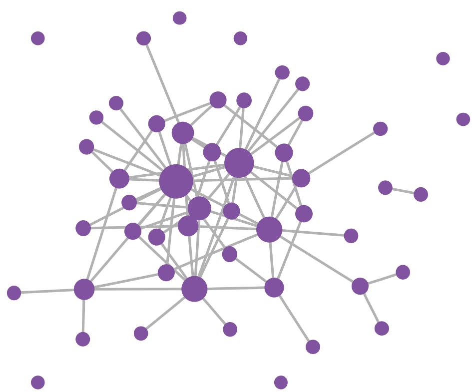




(c)



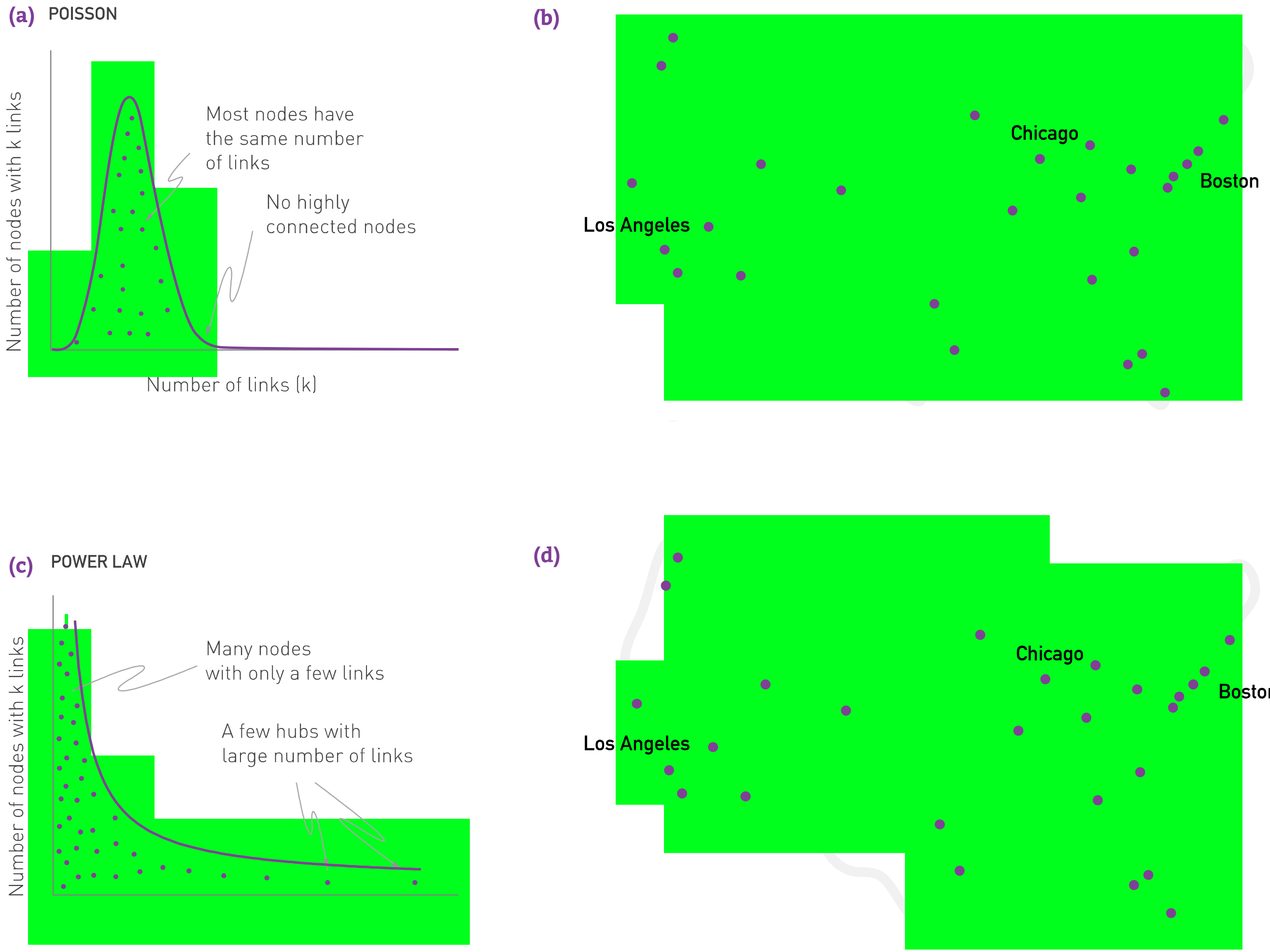
(d)



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*


(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}

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

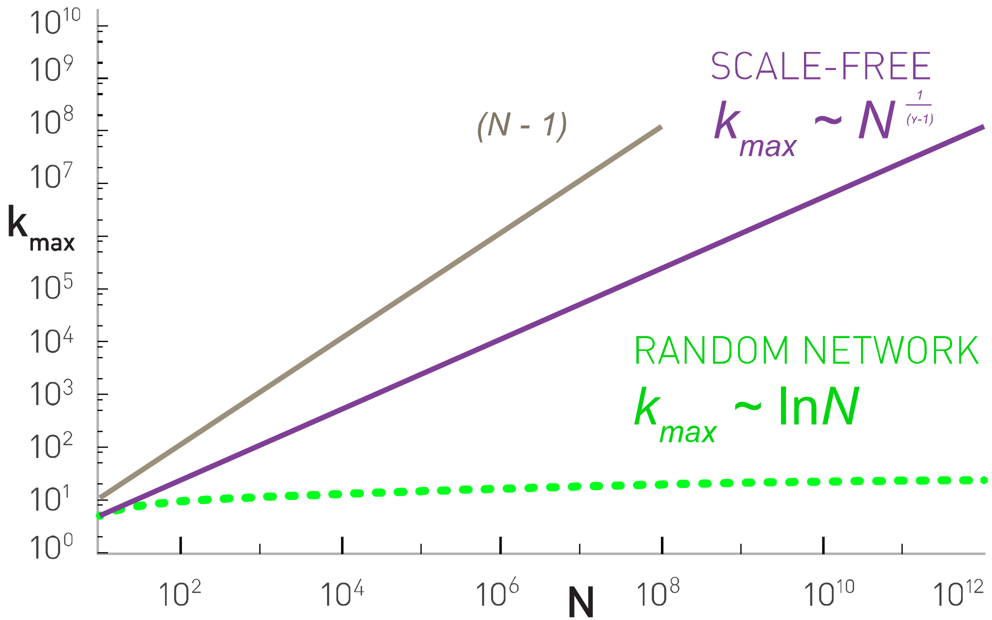
- k_{max}, increases with the size of the network
  → the larger a system is, the larger its biggest hub

- 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’

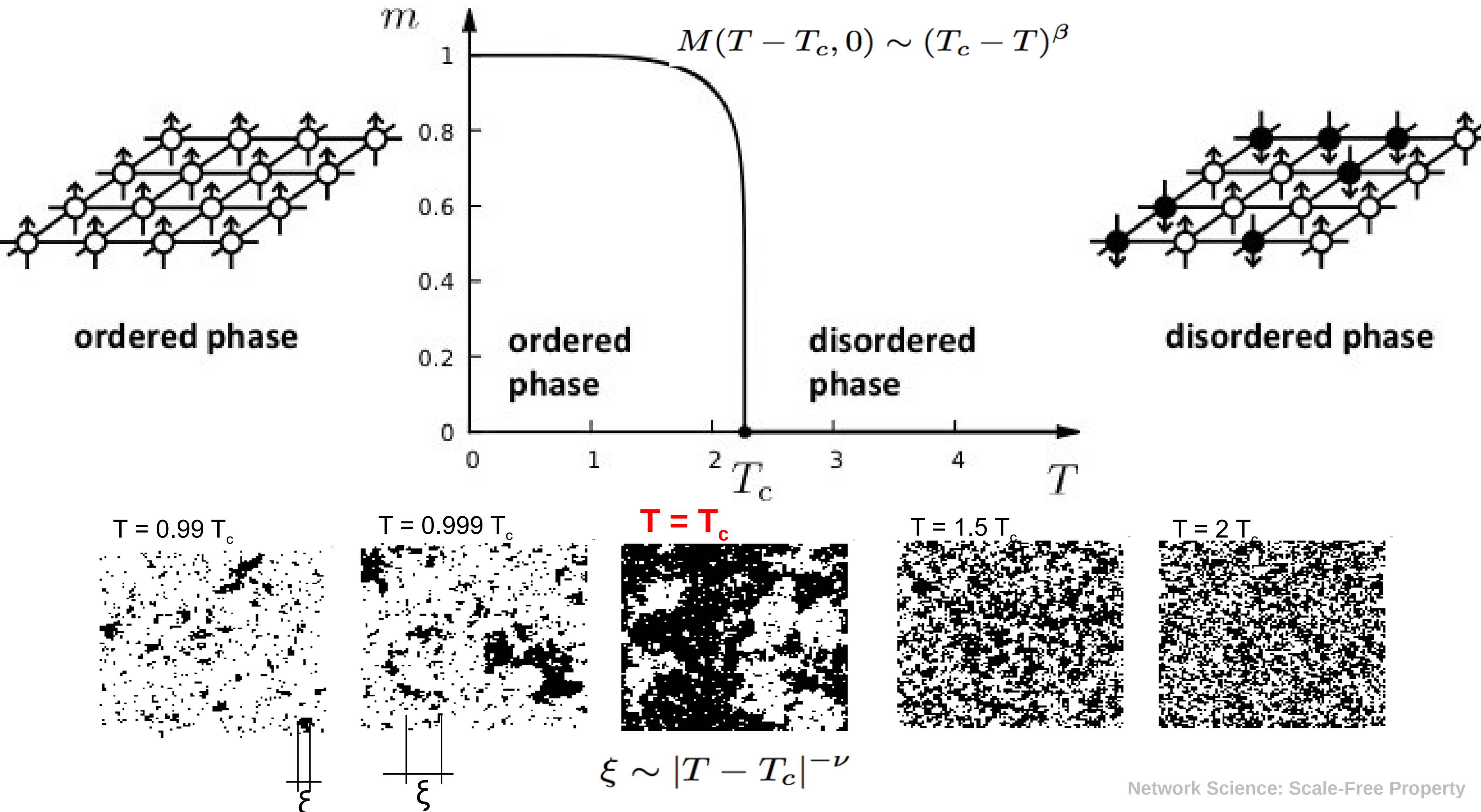
Where does the name come from?

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

$$P(k) = Ck^{-\gamma} \qquad k = [k_{\min}, \infty)$$

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

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

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

$$< k^m >= \int_{k_{\min}}^{\infty} k^m P(k)dk$$

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

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

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

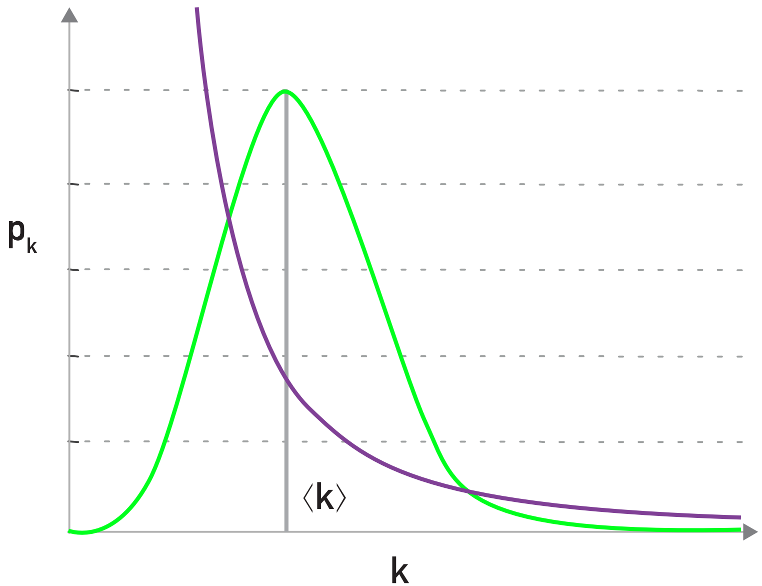
For a fixed λ this means all moments m>γ-1 diverge.

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

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

The meaning of scale-free



Random Network

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

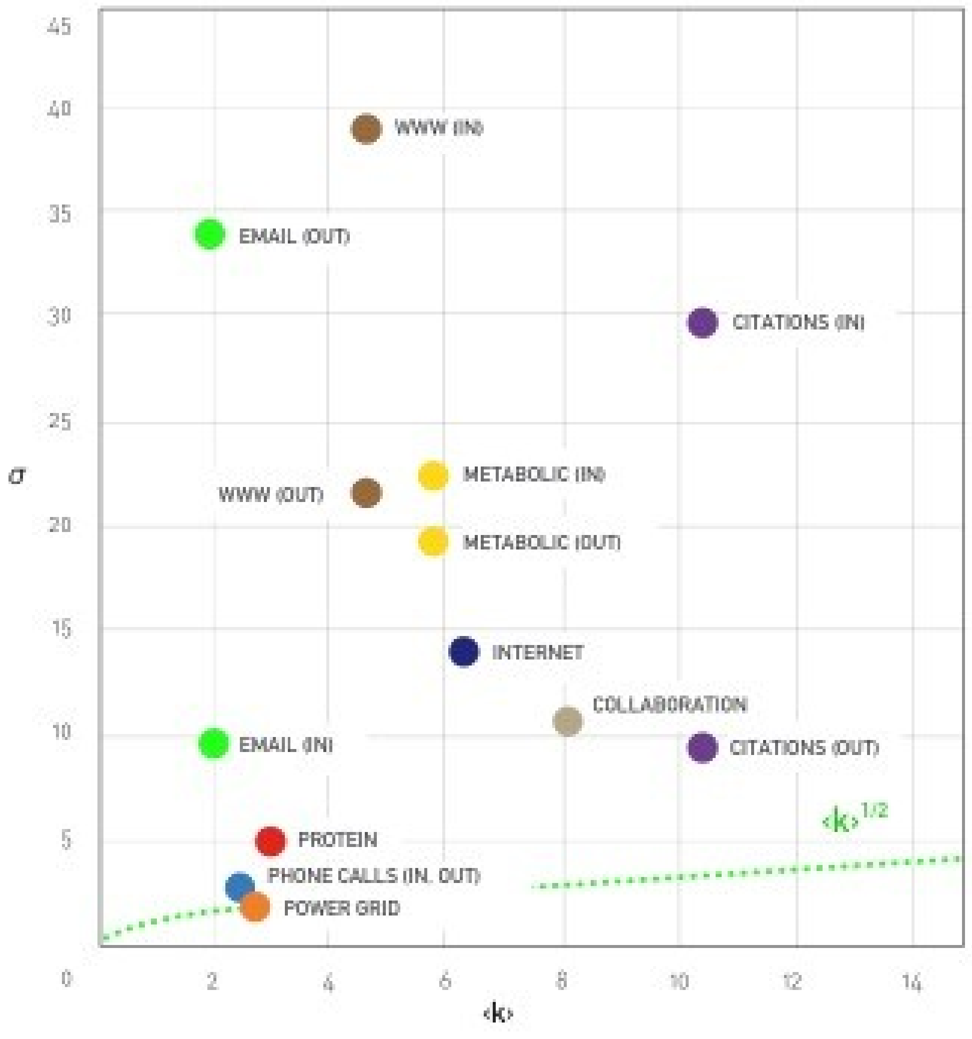
Scale: $\langle k \rangle$

Scale-Free Network

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

Scale: none

The meaning of scale-free



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