

Redes y Sistemas Complejos
Cuarto Curso del Grado en Ingeniería Informática

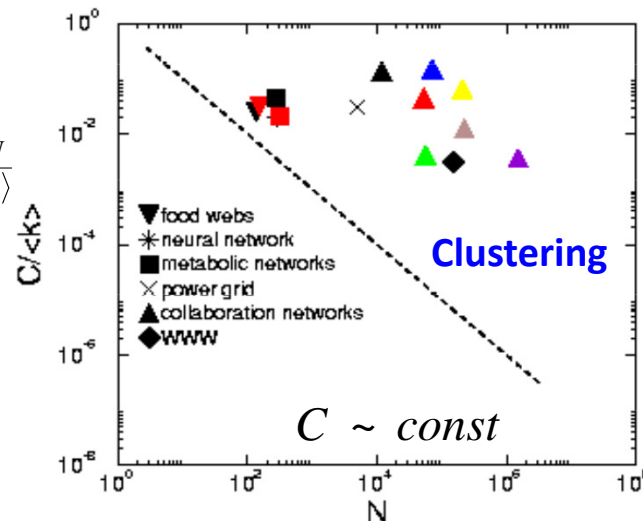
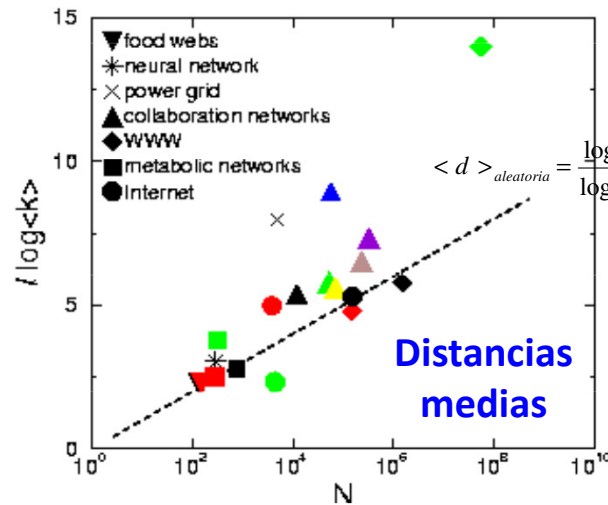
Tema 5: Modelos de Redes
5.3. Redes Libres de Escala

Oscar Cordon García

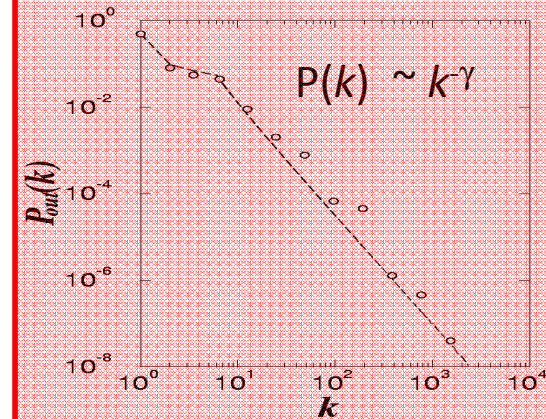
Dpto. Ciencias de la Computación e Inteligencia Artificial
ocordon@decsai.ugr.es

¿ES LA WWW UNA
RED ALEATORIA?

VALIDEZ DE LOS MODELOS EXISTENTES PARA REDES REALES



Distribución de los grados



Redes
regulares

$$\langle d \rangle \approx N^{1/D}$$



$$C \sim \text{const}$$



$$P(k) = \delta(k - k_d)$$



Redes ER
(aleatorias)

$$\langle d_{\text{aleatoria}} \rangle \approx \frac{\log N}{\log \langle k \rangle}$$



$$C_{\text{aleatoria}} = p = \frac{\langle k \rangle}{N}$$



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



Mundos
pequeños

$$\langle d_{\text{aleatoria}} \rangle \approx \frac{\log N}{\log \langle k \rangle}$$



$$C \sim \text{const}$$



Exponencial



LA RED COMPLEJA DE LA WORLD WIDE WEB (1)

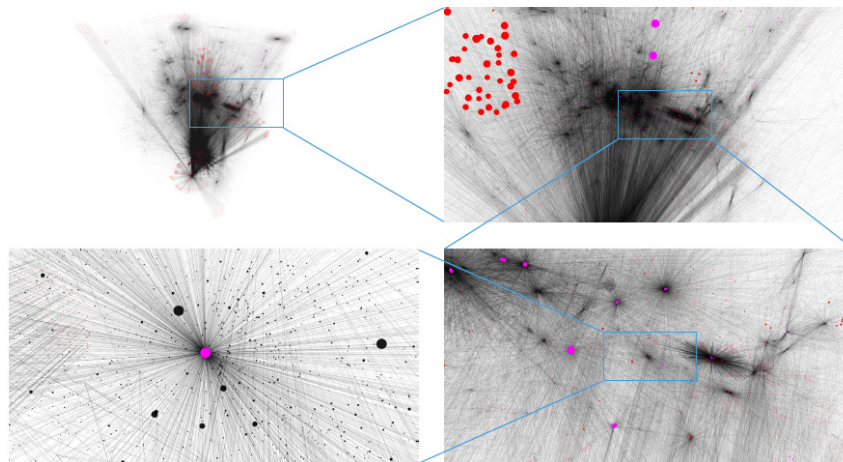
Nodos: **documentos WWW**

Enlaces: **hiperenlaces URL**

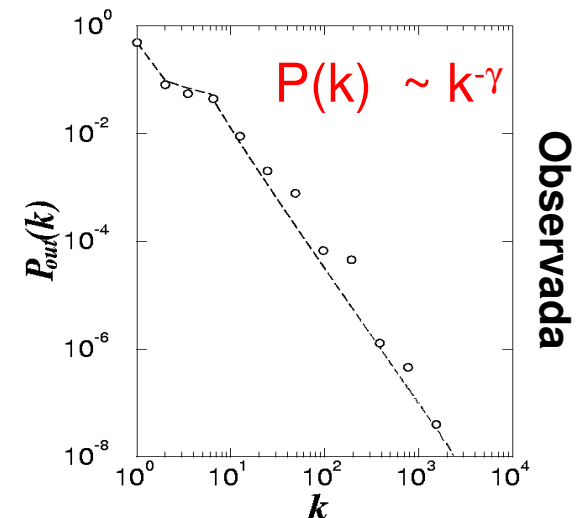
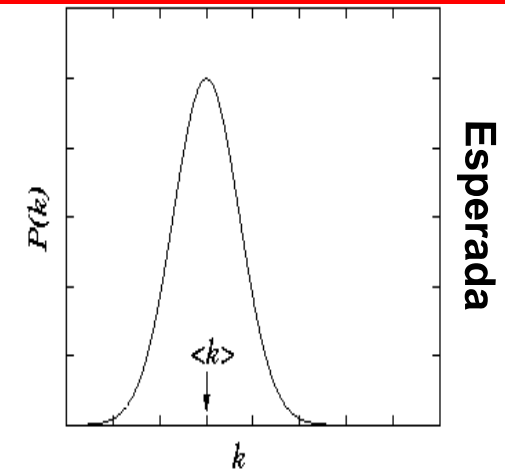
En torno a 1 billón ($N \approx 10^{12}$) de documentos

ROBOT: recopila todas las URLs encontradas en una página web y las sigue recursivamente

En 1998 se pensaba que era una red aleatoria:



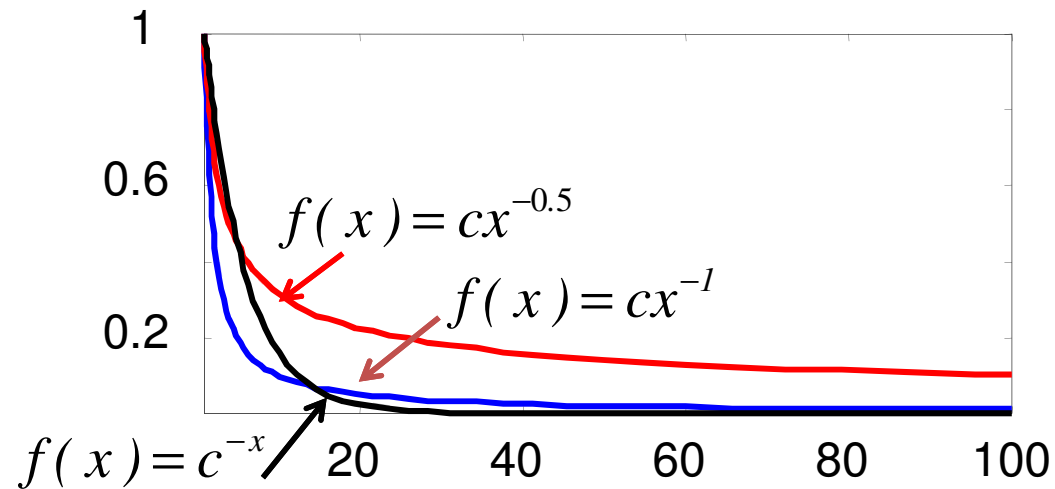
Mapa de la web (dominio nd.edu) con 325,725 páginas (1998)
Nodos rojos $\rightarrow$ grado ≥ 50 ; nodos morados $\rightarrow$ grado ≥ 500



H. Jeong, R. Albert, A-L Barabasi, *Nature*, 401:130-131 (1999)

DIFERENCIA ENTRE UNA LEY DE LA POTENCIA Y UNA EXPONENCIAL (1)

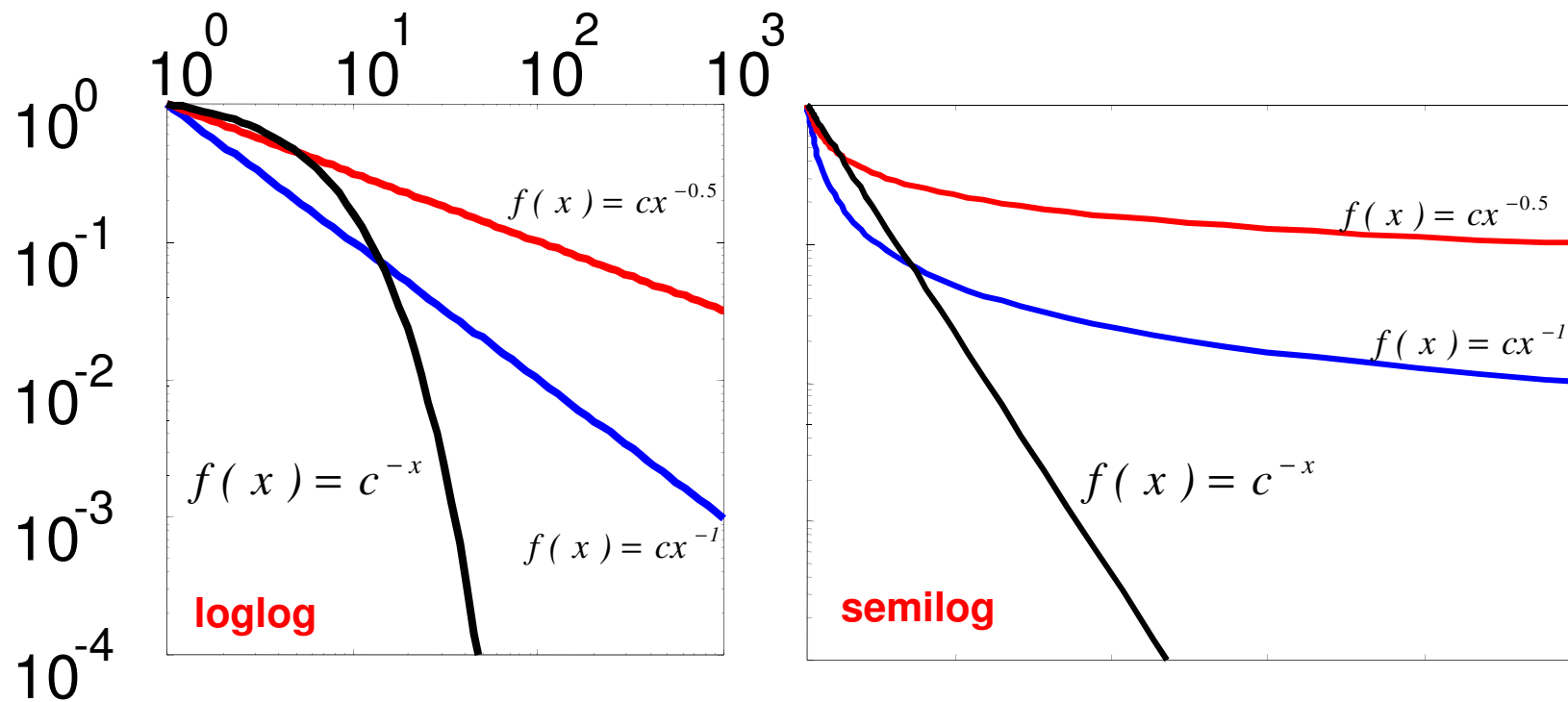
La distribución de los grados observada en la WWW sigue la llamada **Ley de la Potencia**



Por encima de un cierto umbral x , la ley de la potencia es siempre mayor que la exponencial

DIFERENCIA ENTRE UNA LEY DE LA POTENCIA Y UNA EXPONENCIAL (2)

La diferencia es muy clara si se representa en una escala logarítmica para el eje Y: para valores grandes de x existen diferencias de varios órdenes de magnitud entre las dos funciones



$\log p_k \sim -\gamma \log k \rightarrow \text{Log } p_k \text{ depende linealmente de } \log k$
La pendiente de la recta viene dada por el exponente de grado γ

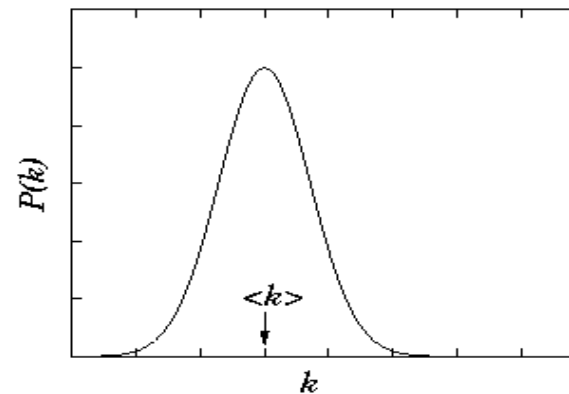
¿QUÉ SIGNIFICA REALMENTE ESTA DIFERENCIA?

Representación Visual (1)

Red Exponencial

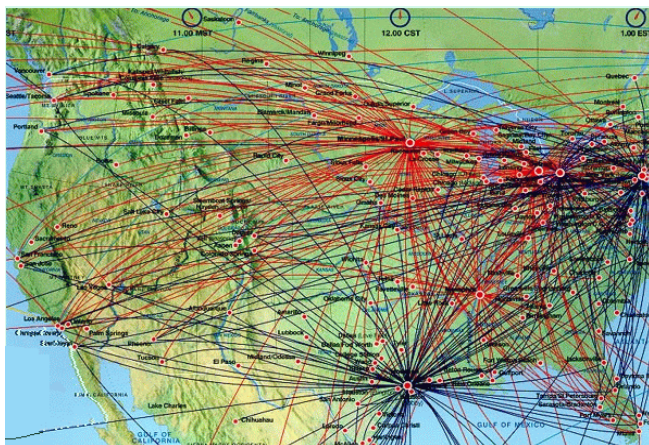


Red de nacional de autovías

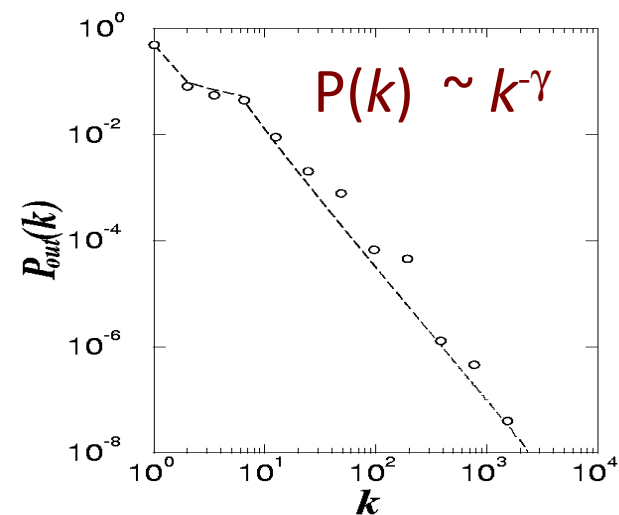


Esperada

Red con Ley de la Potencia



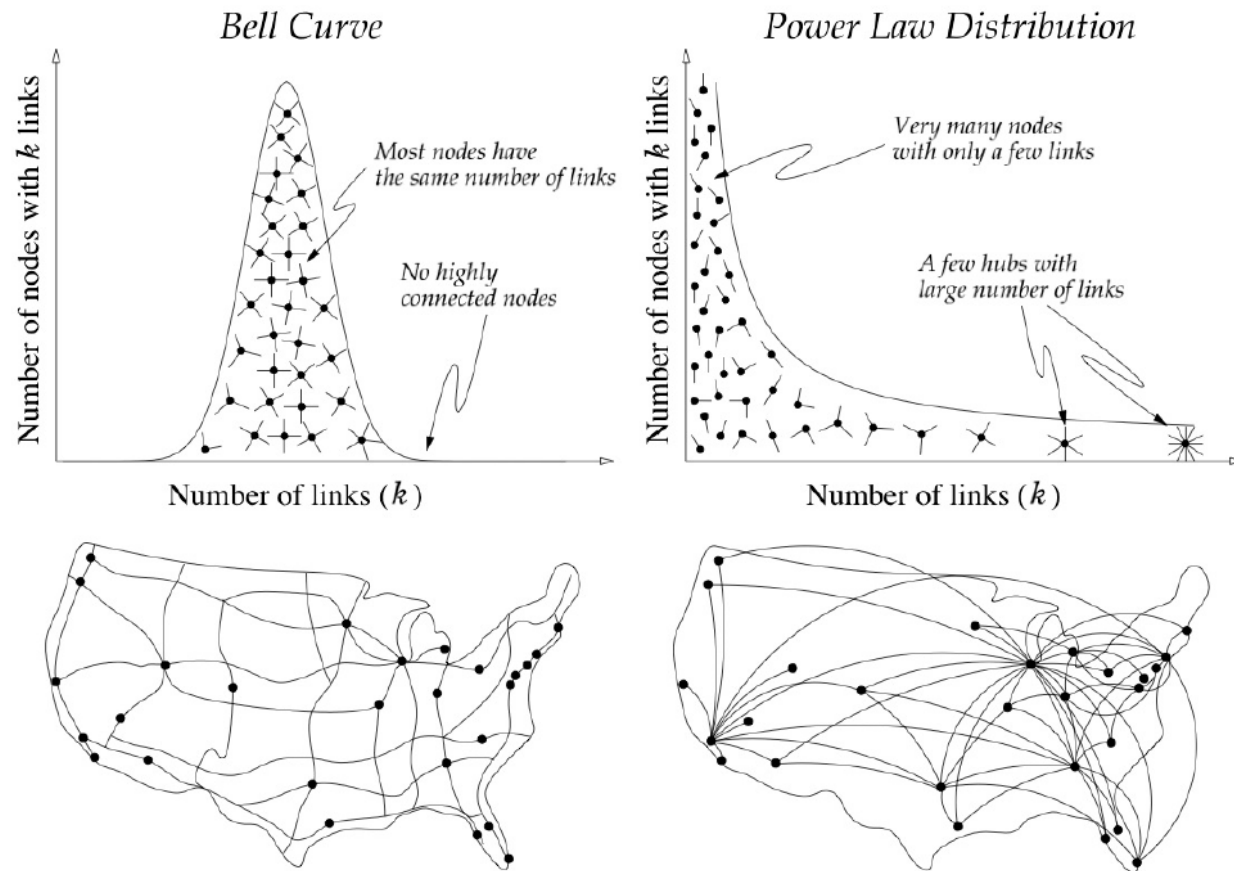
Red de tráfico aéreo



Observada

¿QUÉ SIGNIFICA REALMENTE ESTA DIFERENCIA?

Representación Visual (2)



Once hubs are present, they change the way we navigate the network. For example, if we travel from Boston to Los Angeles by car, we must drive through many cities (nodes). On the airplane network, however, we can reach most destinations via a single hub, like Chicago. After [4].

[4] A.-L. Barabási. *Linked: The New Science of Networks*. (Plume, New York). ISBN 0-452-28439-2, 2002.

LA RED COMPLEJA DE LA WORLD WIDE WEB NO ES ALEATORIA

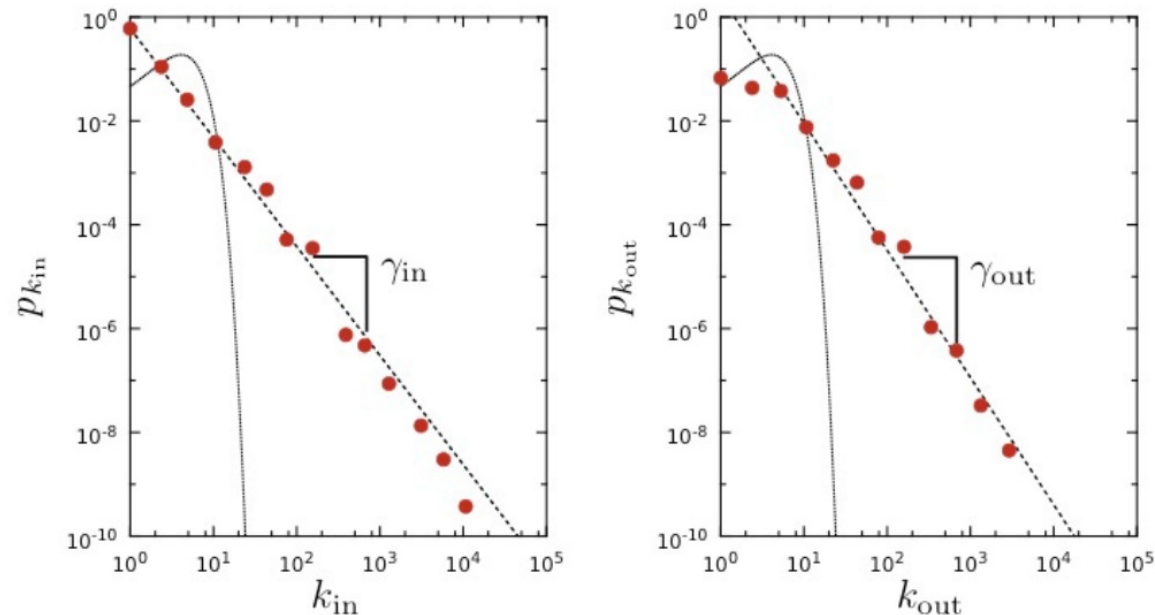


Figure 4.2: The degree distribution of the WWW.

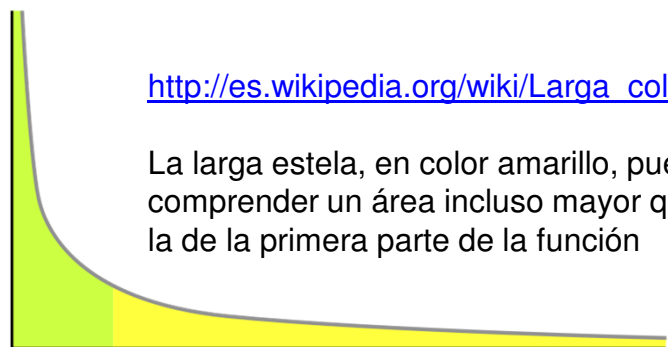
The incoming (left panel) and outgoing (right panel) degree distribution of the WWW sample mapped in the 1999 study of Albert *et al.* [1]. The degree distribution is shown on double logarithmic axis (log-log plot), in which a power law is expected to follow a straight line. The symbols correspond to the empirical data and the dotted line corresponds to the power law fit, with degree exponents $\gamma_{in} = 2.1$ and $\gamma_{out} = 2.45$. The degree distribution predicted by a Poisson function with average degree $\langle k_{in} \rangle = \langle k_{out} \rangle = 4.60$, representing the observed values for the WWW sample, is shown as a dotted line.

REDES LIBRES DE ESCALA

DEFINICIÓN DE REDES LIBRES DE ESCALA

Las redes cuya distribución de grados sigue la ley de la potencia se denominan “redes libres de escala” (*scale-free*)

Estas redes presentan una distribución de grados de larga estela o larga cola (*heavy-tailed, long-tailed or power-law-tailed distributions*)



http://es.wikipedia.org/wiki/Larga_cola

La larga estela, en color amarillo, puede comprender un área incluso mayor que la de la primera parte de la función

- Distribuciones de Pareto: regla 80/20
- Modelos económicos: *Amazon.com*

Si la red es dirigida, se puede analizar por separado la “**propiedad de libertad de escala**” sobre las distribuciones de grados de entrada y salida

Se pueden emplear dos formulaciones matemáticas distintas para estudiar las propiedades de estas redes aunque **la propiedad en sí es independiente del formalismo considerado**

FORMULACIÓN

Formalismo discreto (1)

Considera que los grados son valores enteros positivos. Captura la probabilidad $p_k > 0$ de que un nodo tenga exactamente k enlaces:

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

C es una constante que se calcula a partir de la condición de normalidad:

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

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \dots \quad \sigma = \Re(s) > 1.$$

función Zeta de Riemann

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)} \quad \text{para } k > 0 \quad (\text{se asume que no hay nodos desconectados en la red, } p_0 \text{ se calcula aparte})$$

FORMULACIÓN

Formalismo discreto (2)

En algunas aplicaciones, sólo se tiene en cuenta la cola de la distribución de grados:

$$p_k = Ck^{-\gamma} \quad k = [k_{\min}, \infty)$$

$$\sum_{k=k_{\min}}^{\infty} p_k = 1$$

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

$$\zeta(s, a) \equiv \sum_{k=0}^{\infty} \frac{1}{(k+a)^s}$$

función Zeta generalizada o incompleta

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

FORMULACIÓN

Formalismo continuo

En cálculos analíticos conviene asumir que los grados pueden tomar cualquier valor real positivo. En ese caso, la distribución de grados de la ley de la potencia toma la forma:

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

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

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

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

$k_{\min}$ es el menor valor de grado para el que se cumple que: $p_k = \frac{k^{-\gamma}}{\zeta(\gamma)}$

LOS HUBS EN LAS REDES LIBRES DE ESCALA (1)

Las diferencias principales entre una red aleatoria y otra libre de escala se encuentran en la cola de la distribución:

- Para valores pequeños de k , la ley de la potencia está por encima de la de Poisson: **la red libre de escala tiene un número alto de nodos con grado pequeño**
- Para k 's cercanos a $\langle k \rangle$, la distribución de Poisson está por encima de la ley de la potencia: **en una red aleatoria hay muchos nodos con grado $k \approx \langle k \rangle$**
- Para valores altos de k , la ley de la potencia vuelve a estar por encima de la de Poisson: **la posibilidad de encontrar un nodo de grado alto, un **hub**, es varios órdenes de magnitud mayor en una red libre de escala**

Ejemplo (WWW): $p_{\text{aleatoria}}(100) = 10^{-30}$; $p_{\text{libre-escala}}(100) = 10^{-4}$

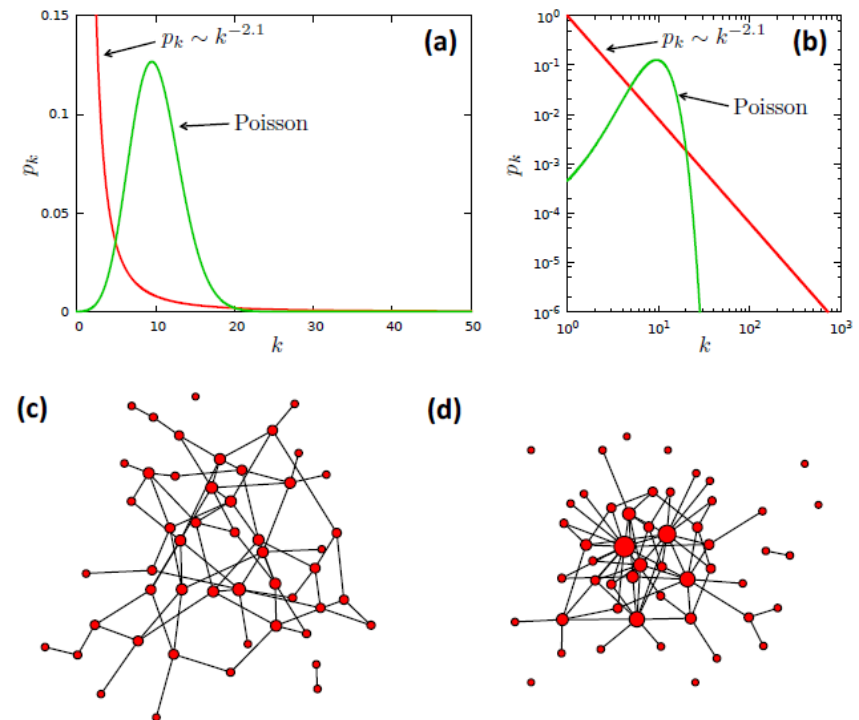


Figure 4.3: Poisson vs. power-law distributions.

(a) A Poisson function and a power law function with $\gamma = 2.1$. Both distributions have $\langle k \rangle = 10$.

(b) The curves in (a) shown on a log-log plot, offering a better view of the difference between the two functions in the high- k regime.

(c) A random network with $\langle k \rangle = 3$ and $N = 50$, illustrating that most nodes have comparable degree $k \approx \langle k \rangle$.

(d) A scale-free network with $\langle k \rangle = 3$, illustrating that numerous small degree nodes coexist with a few highly connected hubs.

LOS HUBS EN LAS REDES LIBRES DE ESCALA (2)

¿Cómo afecta el tamaño de la red N al tamaño de sus hubs?

Para cualquier distribución de grados p_k es posible calcular el grado máximo esperado $k_{\max}$:

- Red aleatoria, distribución exponencial $p_k = Ce^{-\lambda k}$: $k_{\max} = k_{\min} + \frac{\ln N}{\lambda}$

Depende de $\ln N$, crece suavemente y hay poca diferencia entre $k_{\min}$ y $k_{\max}$. **No hay hubs**

- Red aleatoria, distribución de Poisson: crece aún más suavemente.

- Red libre de escala, distribución de larga cola: $k_{\max} \sim k_{\min} N^{\frac{1}{\gamma-1}}$

Depende de N . **Cuanto mayor es el tamaño de la red, mayor es el grado del hub más grande. Hay diferencias significativas entre $k_{\min}$ y $k_{\max}$**

Ejemplo (muestra WWW $N=3 \cdot 10^5$): $k_{\max}(\text{red aleatoria}) \approx 13$; $k_{\max}(\text{red libre de escala}) \approx 85,000$

LOS HUBS EN LAS REDES LIBRES DE ESCALA (3)

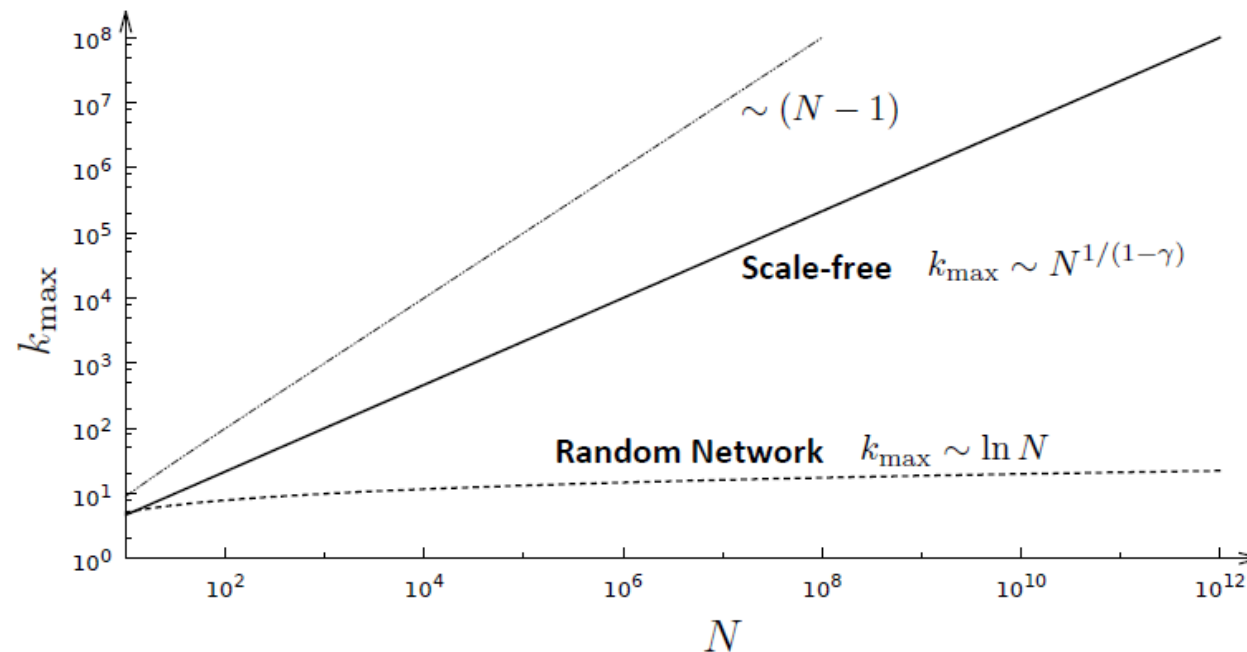


Figure 4.5: Hubs are large in scale-free networks.

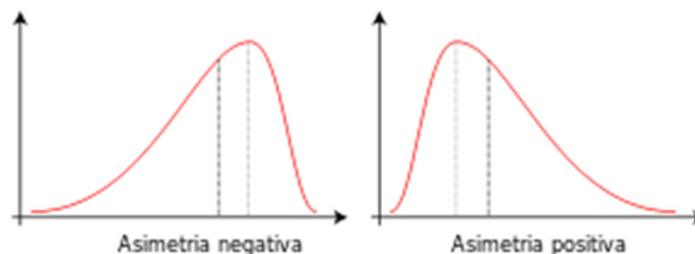
The expected degree of the largest node (natural cutoff) in scale-free and random networks with the same average degree $\langle k \rangle = 3$. For the scale-free network we chose $\gamma = 2.5$. For comparison, we also show the linear behavior, $k_{\max} \sim N - 1$, expected for a complete network. Overall, hubs in a scale-free network are several orders of magnitude larger than the biggest node in a random network with the same N and $\langle k \rangle$.

ORIGEN DEL TÉRMINO “LIBRE DE ESCALA” (1)

El término “libre de escala” está relacionado con los **momentos de la distribución de probabilidad de grados**:

$$\langle k^n \rangle = \sum_{k_{min}}^{\infty} k^n p_k = \int_{k_{min}}^{\infty} k^n p(k) dk$$

- $n=1$: el primer momento es el **grado medio** $\langle k \rangle$
- $n=2$: el segundo momento $\langle k^2 \rangle$ es la **varianza** $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ que mide la dispersión de los grados. Su raíz cuadrada σ es la **desviación típica**
- $n=3$: el tercer momento $\langle k^3 \rangle$ determina la **asimetría**, indicando cómo de simétrica es p_k alrededor de la media (en distribuciones simétricas, vale 0)



ORIGEN DEL TÉRMINO “LIBRE DE ESCALA” (2)

En una red libre de escala, el n -ésimo momento de la distribución de grados es:

$$\langle k^n \rangle = \int_{k_{\min}}^{k_{\max}} k^n p(k) dk = C \frac{k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1}}{n - \gamma + 1}$$

Como $k_{\max}$ crece con el tamaño de la red, hay que analizar su comportamiento en el límite $k_{\max} \rightarrow \infty$ para determinar las propiedades de las redes muy grandes:

$$\langle k^n \rangle = (\gamma - 1) k_{\min}^{\gamma-1} \int_{k_{\min}}^{\infty} k^{n-\gamma} dk = \frac{(\gamma - 1)}{(n - \gamma + 1)} k_{\min}^{\gamma-1} \left[k^{n-\gamma+1} \right]_{k_{\min}}^{\infty}$$

- Si $n - \gamma + 1 \leq 0$, el n -ésimo momento está acotado. **Todos los momentos en los que $n \leq \gamma + 1$ son finitos**
- Si $n - \gamma + 1 \geq 0$, el n -ésimo momento tiende a infinito según la red crece. **Todos los momentos en los que $n \geq \gamma + 1$ divergen**

ORIGEN DEL TÉRMINO “LIBRE DE ESCALA” (3)

La mayoría de los exponentes de grado γ están entre 2 y 3: $\langle k \rangle$ es finito pero $\langle k^2 \rangle$ y el resto de momentos tienden a infinito. Eso implica que:

$$\sigma_k = (\langle k^2 \rangle - \langle k \rangle^2)^{1/2} \rightarrow \infty$$

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

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}		
WWW	325 729	4.51	900	2.45	2.1	WWW:	$\langle k \rangle = 7 \pm \infty$
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:	$\langle k \rangle = 3.5 \pm \infty$
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	Redes metabólicas:	$\langle k \rangle = 7.4 \pm \infty$
Metabolic, <i>E. coli</i>	778	7.4	110	2.2	2.2		
Protein, <i>S. cerev.</i> *	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	Llamadas telefónicas:	$\langle k \rangle = 3.16 \pm \infty$
Phone call	53×10^5	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		

¡Los valores medios no tienen sentido al ser las fluctuaciones demasiado grandes!
 La escala interna no es coherente. La red es “libre de escala”

ORIGEN DEL TÉRMINO “LIBRE DE ESCALA” (4)

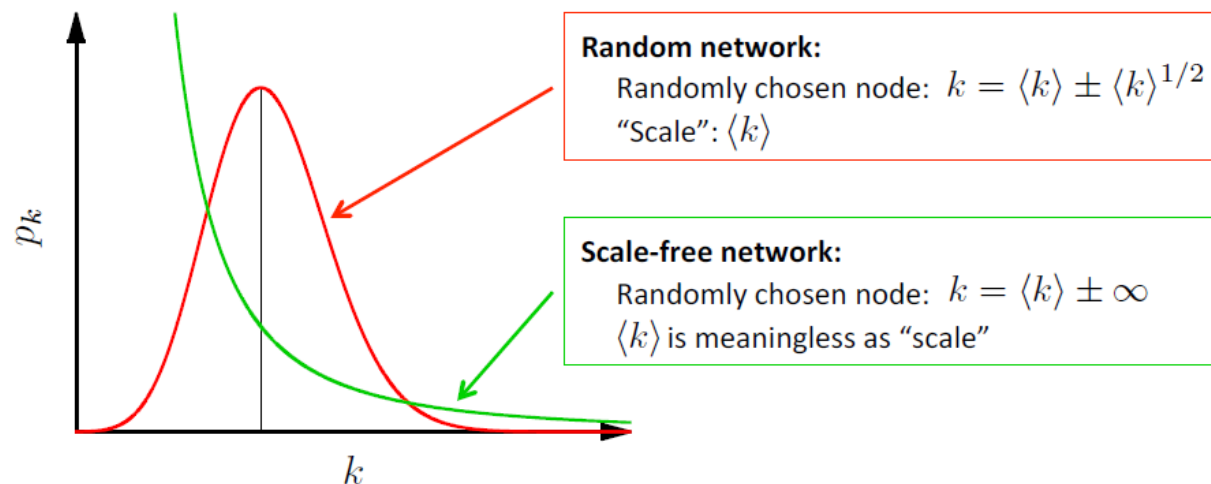


Figure 4.6: Scale-free networks lack an internal scale.

For any bounded distribution (e.g. a Poisson or a Gaussian distribution) the degree of a randomly chosen node will be in the vicinity of $\langle k \rangle$. Hence $\langle k \rangle$ serves as the network’s “scale”. In a scale-free network the second moment diverges, hence the degree of a randomly chosen node can be arbitrarily different from $\langle k \rangle$. As a scale-free network lacks an intrinsic scale, is it “scale-free.”

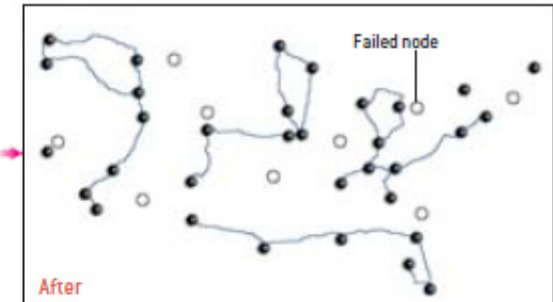
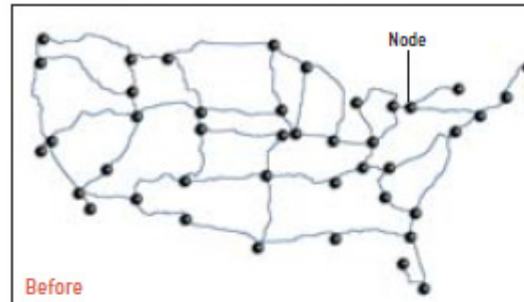
Esta divergencia de las redes libres de escala es el origen de muchas de las propiedades más interesantes de estas redes, como su robustez a fallos/ataques aleatorios y su comportamiento anómalo en la propagación de virus

PROPIEDADES DE LAS REDES LIBRES DE ESCALA

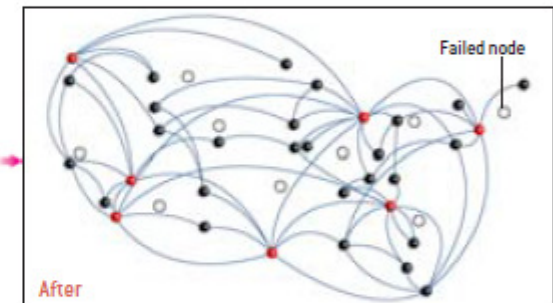
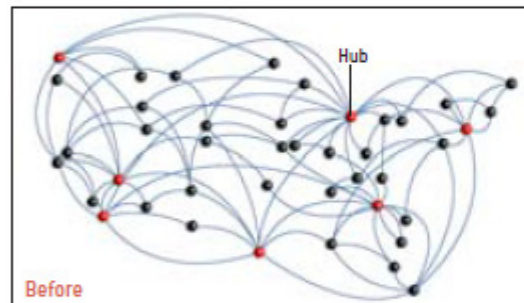
Robustez

THE ACCIDENTAL FAILURE of a number of nodes in a random network (*top panels*) can fracture the system into non-communicating islands. In contrast, scale-free networks are more robust in the face of such failures (*middle panels*). But they are highly vulnerable to a coordinated attack against their hubs (*bottom panels*).

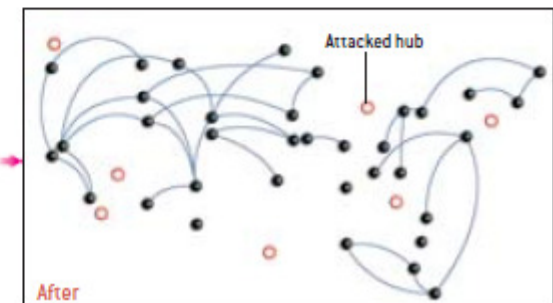
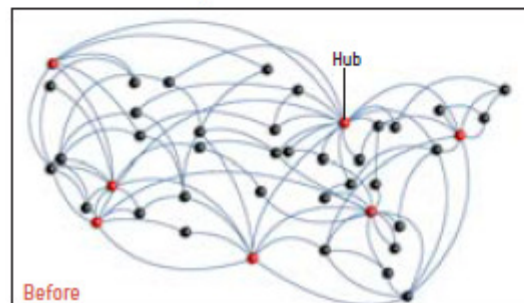
Random Network, Accidental Node Failure



Scale-Free Network, Accidental Node Failure



Scale-Free Network, Attack on Hubs



The Potential Implications of Scale-Free Networks for...

Computing

- Computer networks with scale-free architectures, such as the World Wide Web, are highly resistant to accidental failures. But they are very vulnerable to deliberate attacks and sabotage.
- Eradicating viruses, even known ones, from the Internet will be effectively impossible

Medicine

- Vaccination campaigns against serious viruses, such as smallpox, might be most effective if they concentrate on treating hubs—people who have many connections to others. But identifying such individuals can be difficult.
- Mapping out the networks within the human cell could aid researchers in uncovering and controlling the side effects of drugs. Furthermore, identifying the hub molecules involved in certain diseases could lead to new drugs that would target those hubs.

Business

- Understanding how companies, industries and economies are interlinked could help researchers monitor and avoid cascading financial failures.
- Studying the spread of a contagion on a scale-free network could offer new ways for marketers to propagate consumer buzz about their products.

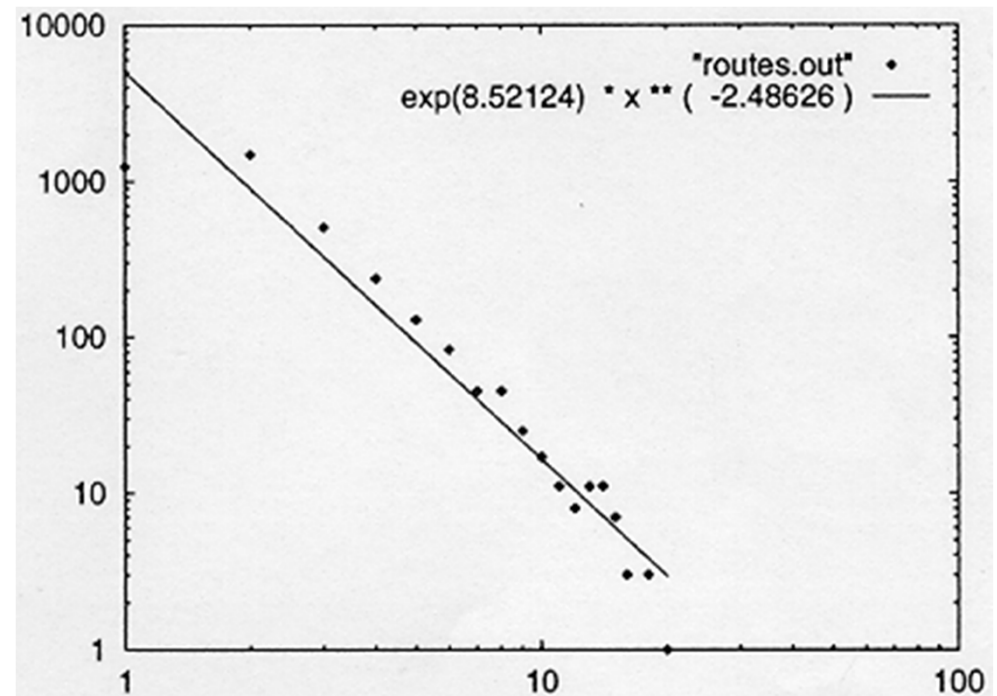
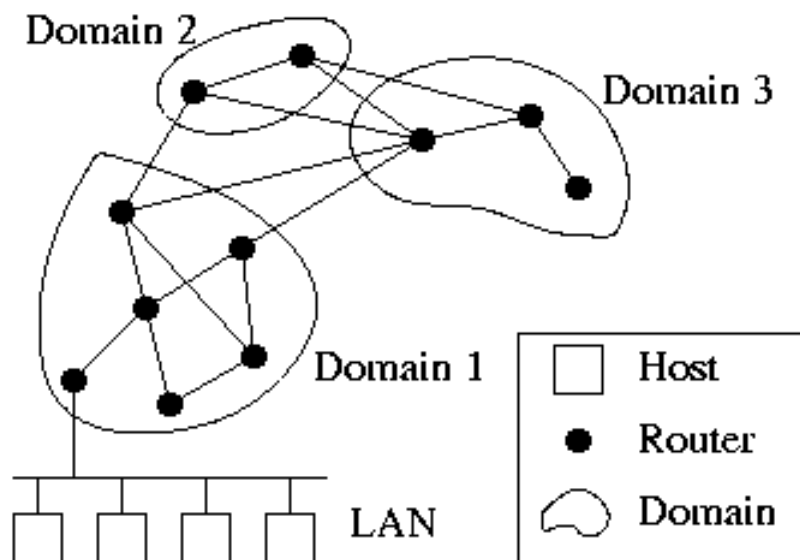
UNIVERSALIDAD

¿Cómo de general es el descubrimiento de una distribución de grados que sigue la ley de la potencia?

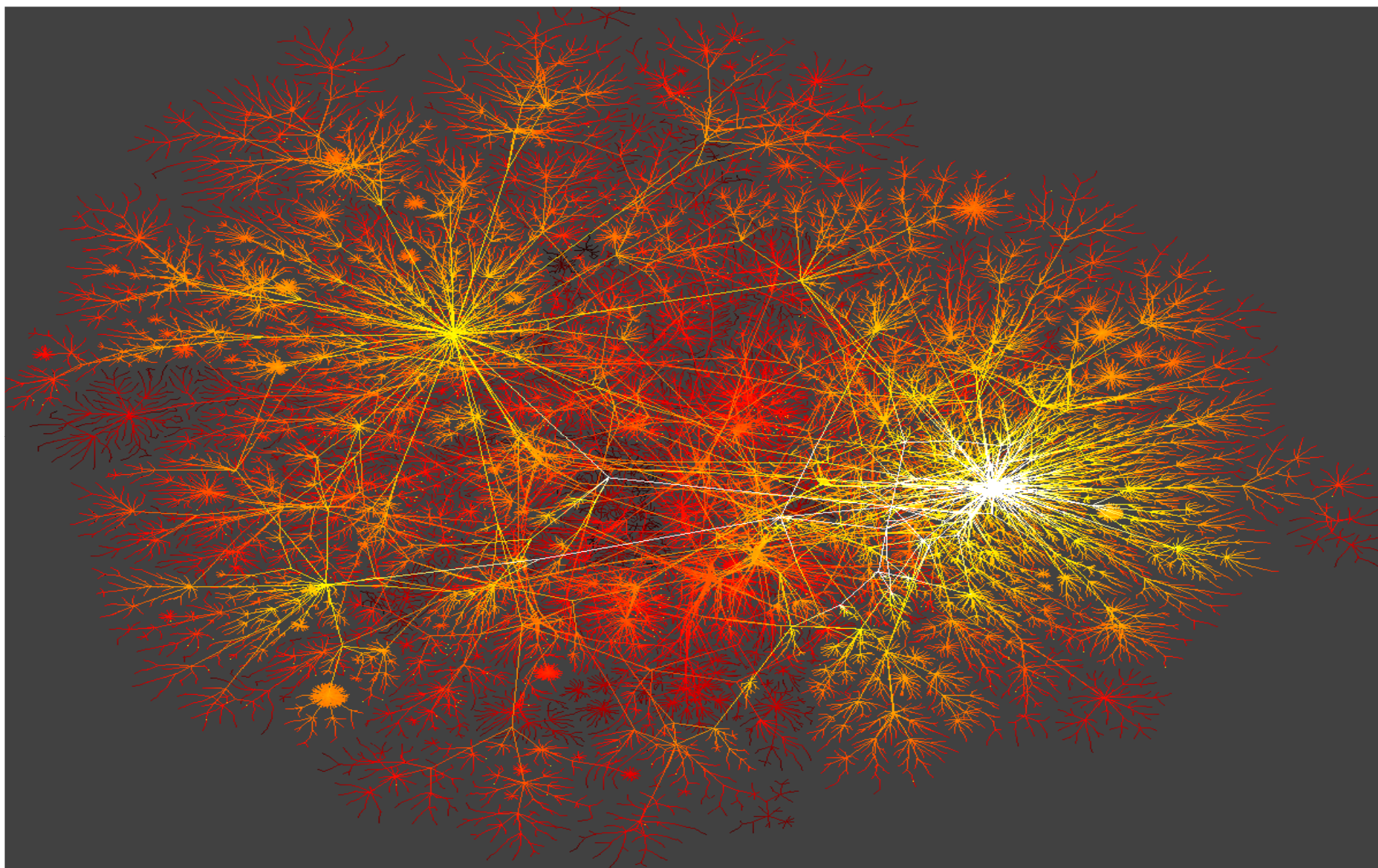
RED COMPLEJA DEL BACKBONE DE INTERNET

Nodos: ordenadores, routers

Enlaces: conexiones físicas



Faloutsos, Faloutsos y Faloutsos, Comput. Commun. Rev. 29:251-262 (1999)



REDES CIENTÍFICAS

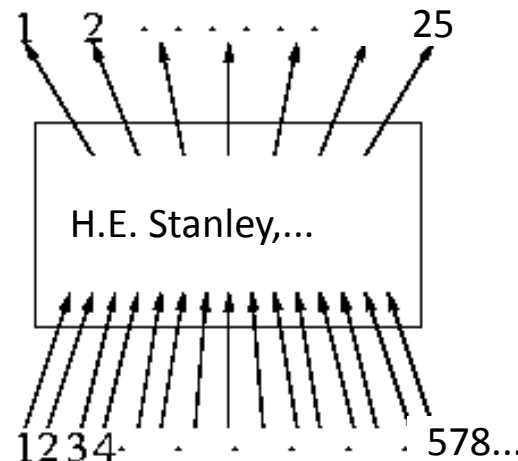
Science Citation Index

Out of over 500,000 Examined
(see <http://www.sst.nrel.gov>)

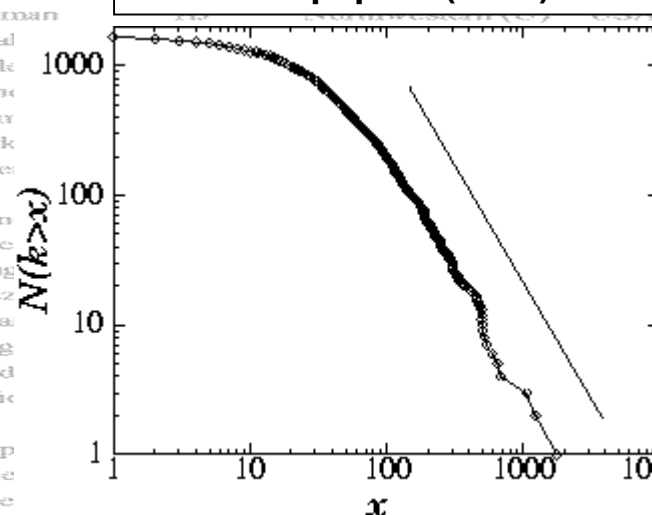
Nodos: artículos científicos

Enlaces: citas

Author	Country	Field	avg. cites	total art.	total cites	rank by total cit.
Witten	USA, NJ	High-energy (T)	168	138	23235	1
David	USA, CA	Semie				2
Cava	USA, NJ	Supern				3
Battlogg	USA, NJ	Supern				4
Ploog	Germany	Semie				5
Ellis	Switzerland	Astrop				6
Fisk	USA, FL	Solid				7
Cardona	Germany	Semie				8
Nanopoulos	USA, TX	High-e				9
Heeger	USA, CA	Polym				10
Lee*	PA					11
Suzuki*	T					12
Anderson	USA, NJ	Solid				13
Suzuki*	IL	Solid				14
Freeman						15
Tanal						16
Mulk	land	Superc				17
Schn	land	Superc				18
Chen	land	Optics				19
Mork	land	Semie				19
Mille	land	Semie				21
Chu	land	Supetconductivity (E)	44	213	9453	22
Bedn	land	Supercond	85	83	9311	23
Coh	land	Solid State (T)	284	284	9311	23
Metg	land	Supetconductivity (E)	86	108	9300	25
Wasz	land	Superconductivity (E)	57	162	9170	26
Shira	land	Superconductivity (E)	33	269	8841	27
Wieg	land	Semiconductors (E)	85	104	8822	28
Vand	land	Magnetism (E)	67	129	8686	29
Uchik	land		28	301	8520	30
Hor	land	Supetconductivity (E)	72	119	8512	31
Murp	land	Astronomy (E)	111	76	8439	32
Birge	land	Superconductivity (E)	41	286	8375	33
Jorge	land	Superconductivity (E)	50	167	8298	34
Hinks	land	Supetconductivity (E)	37	223	8263	35



1736 PRL papers (1988)



S. Redner, Eur. Phys. J. B 4:131 (1998)

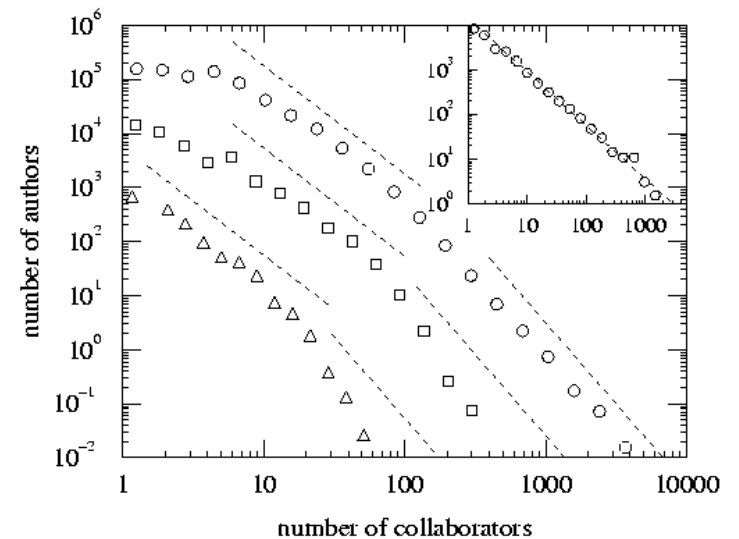
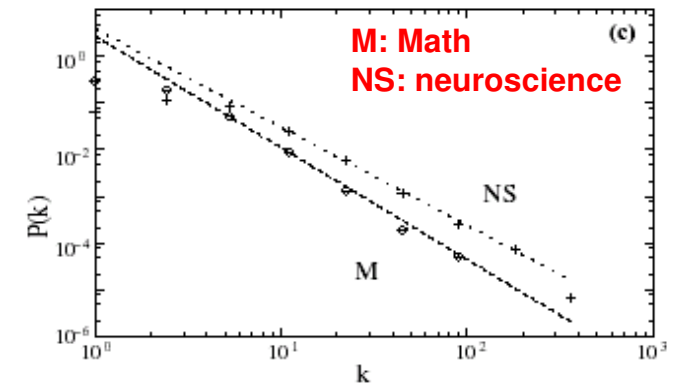
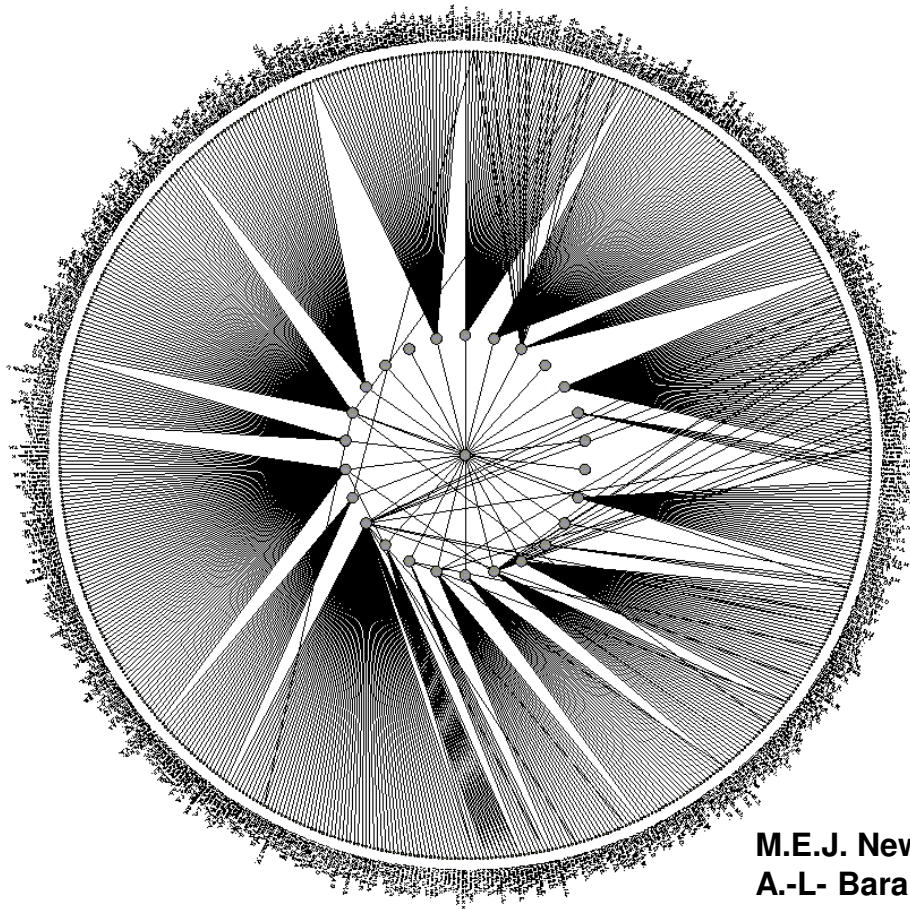
Citation total may be skewed because of multiple authors with the same name

REDES CIENTÍFICAS

Coautoría científica

Nodos: científicos (autores)

Enlaces: publicaciones conjuntas

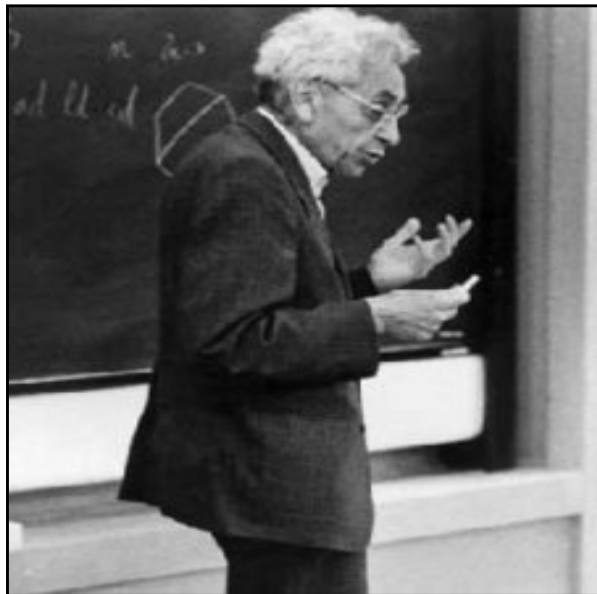


M.E.J. Newman, PNAS 98(2):404-409 (2001)

A.-L. Barabasi et al. Physica A 311:590-614 (2002)

EL NÚMERO ERDOS

(Juego de Bacon para cerebritos ☺)



Paul Erdős (1913-1996)

Erdos tiene un mayor valor de centralidad en su red que Kevin Bacon en la suya

Número de enlaces requeridos para conectar los investigadores a Erdős via la co-autoría de artículos científicos

Erdős escribió más de 1500 artículos con 507 co-autores

La web de Jerry Grossman (Oakland Univ.) permite a los matemáticos calcular su número Erdos:

<http://www.oakland.edu/enp/>

Conectando las longitudes de los caminos entre los matemáticos únicamente :

- La longitud media es 4.65
- La longitud máxima es 13

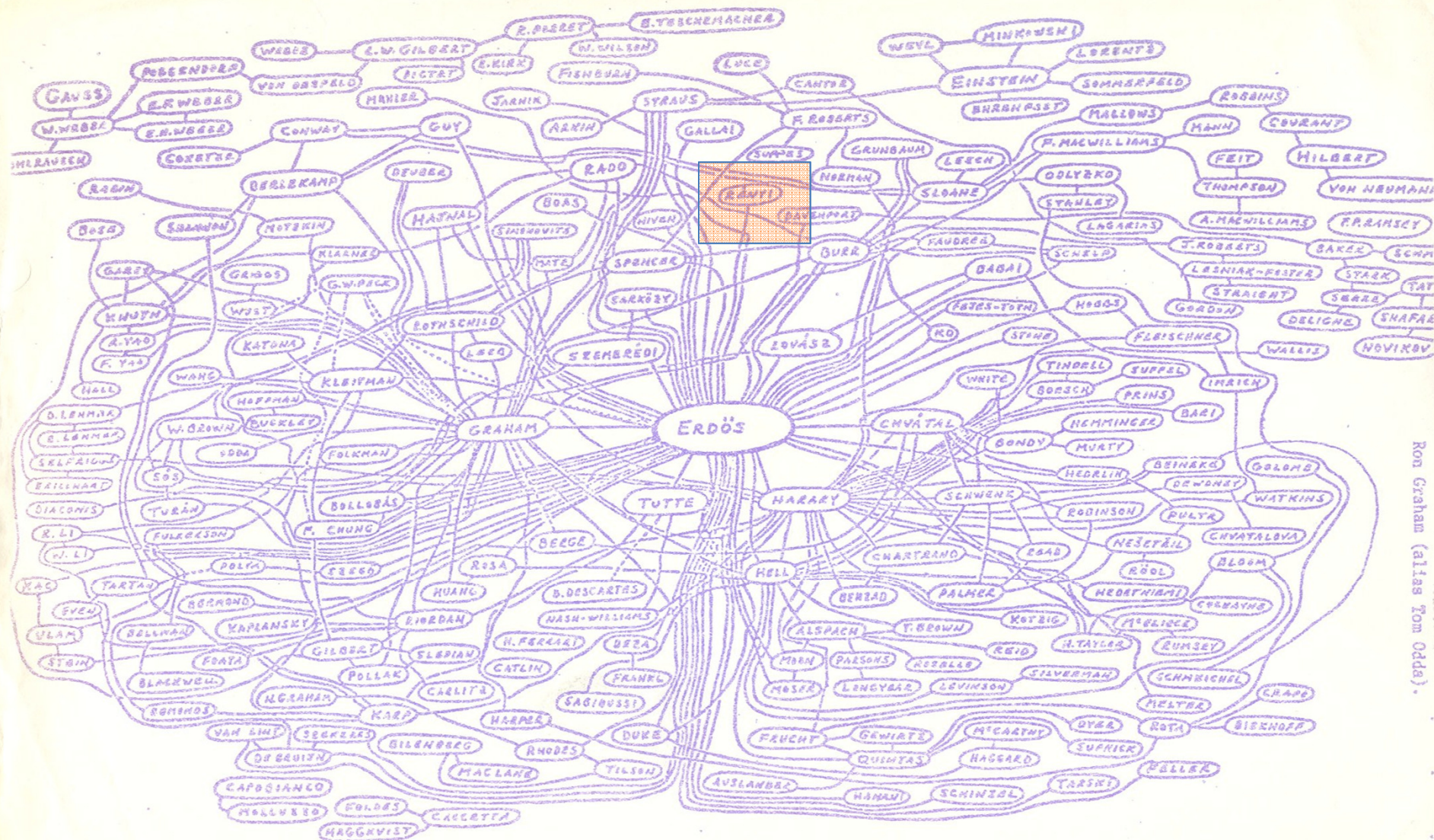


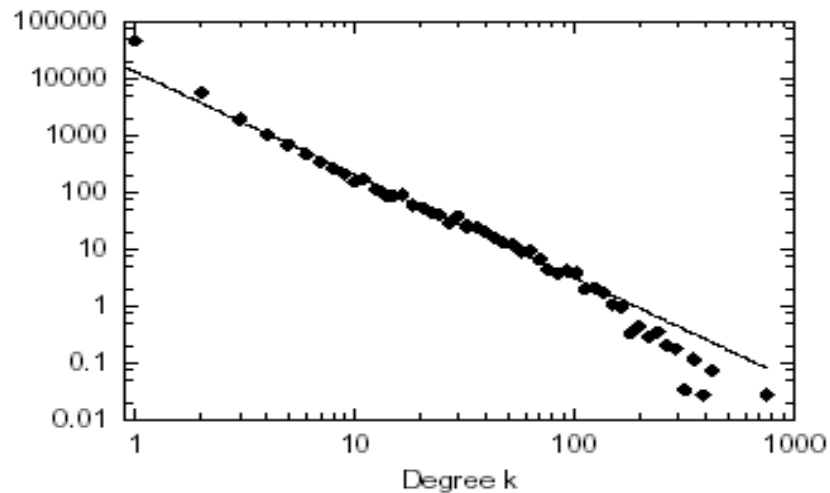
Figure 1
To appear in Topics in Graph Theory (P. Harary, ed.), New York Academy of Sciences (1979).

COMUNIDADES VIRTUALES (1)

Nodos: usuarios

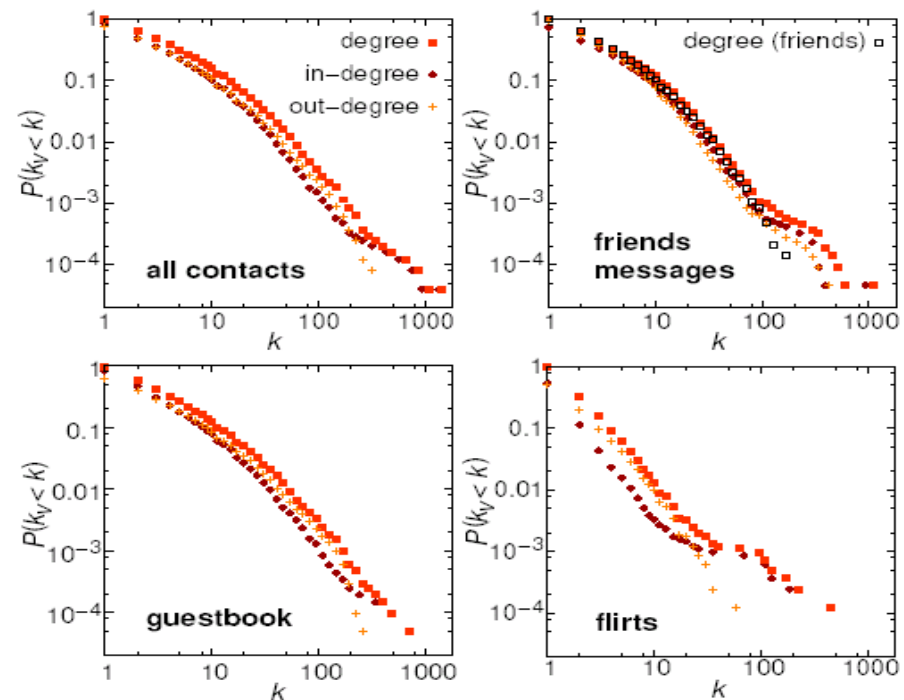
Enlaces: contactos por e-mail

Ficheros de log de la Universidad de Kiel
112 días, $N=59,912$ nodos



Ebel, Mielsch y Bornholdtz, *Physical Reviews E* 66:035103(R), (2002)

Comunidad on-line Pussokram.com
512 días, $N=25,000$ usuarios



Holme, Edling y Liljeros, *Social Networks* 26(2):155-174 (2004)

COMUNIDADES VIRTUALES (2)

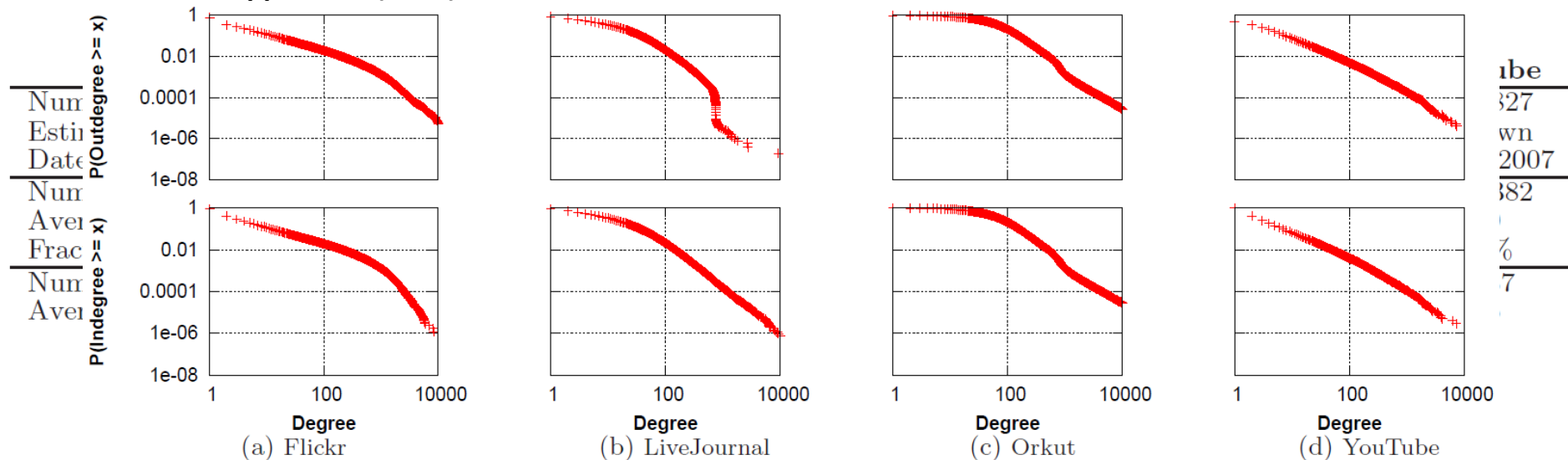
Nodos: usuarios

Enlaces: relaciones en redes sociales

Todas las distribuciones muestra un comportamiento de **cola amplia**: hay desviaciones de varios órdenes de magnitud en los grados

Network	Outdegree		Indegree	
	α	D	α	D
Web [12]	2.67	-	2.09	-
Flickr	1.74	0.0575	1.78	0.0278
LiveJournal	1.59	0.0783	1.65	0.1037
Orkut	1.50	0.6319	1.50	0.6203
YouTube	1.63	0.1314	1.99	0.0094

Mislove et al., Measurement and Analysis of Online Social Networks,
Proc. IMC'07, pp. 29-42 (2007)

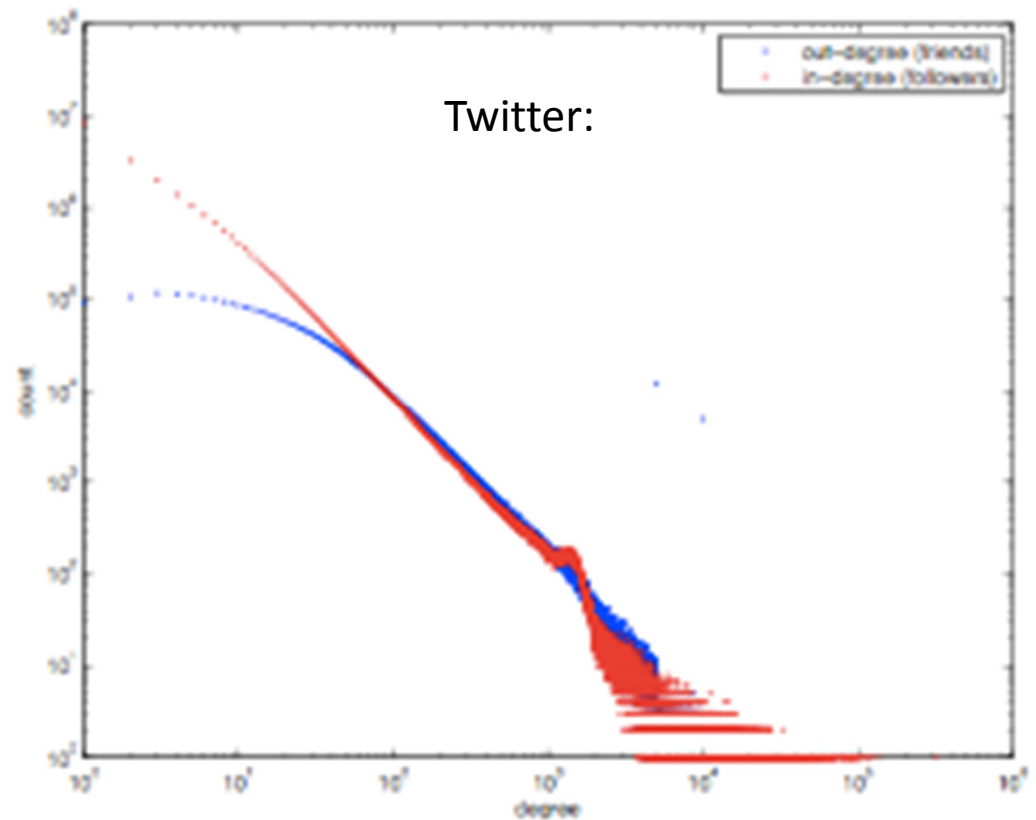


COMUNIDADES VIRTUALES (3)

Nodos: usuarios

Enlaces: relaciones en redes sociales

Todas las distribuciones muestran un comportamiento de **cola amplia**: hay desviaciones de varios órdenes de magnitud en los grados



Jake Hoffman, Yahoo

REDES GENÉTICAS (1)

GENOMA

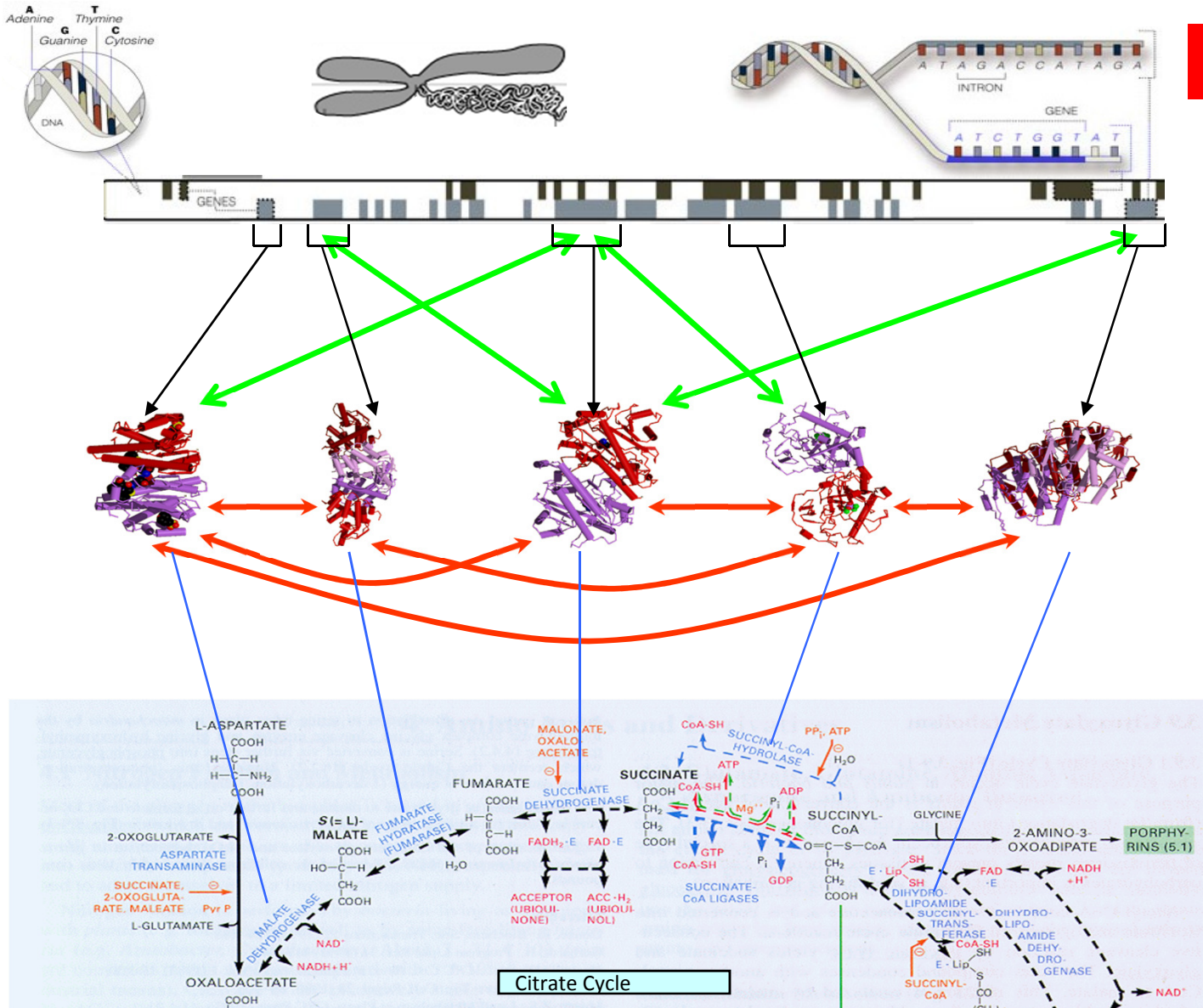
Interacciones
proteína-gen

PROTEOMA

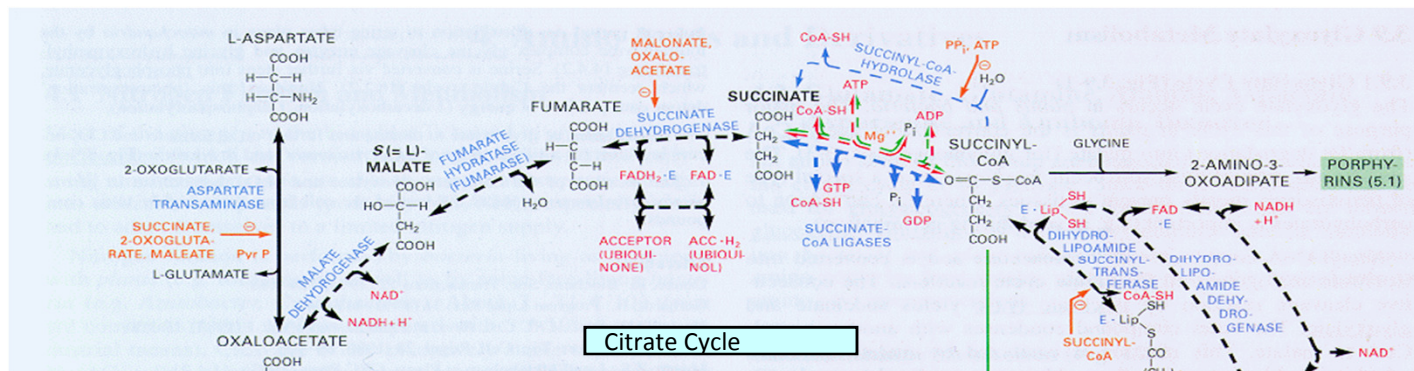
Interacciones
proteína-proteína

METABOLISMO

Reacciones
bio-químicas



REDES GENÉTICAS (2)

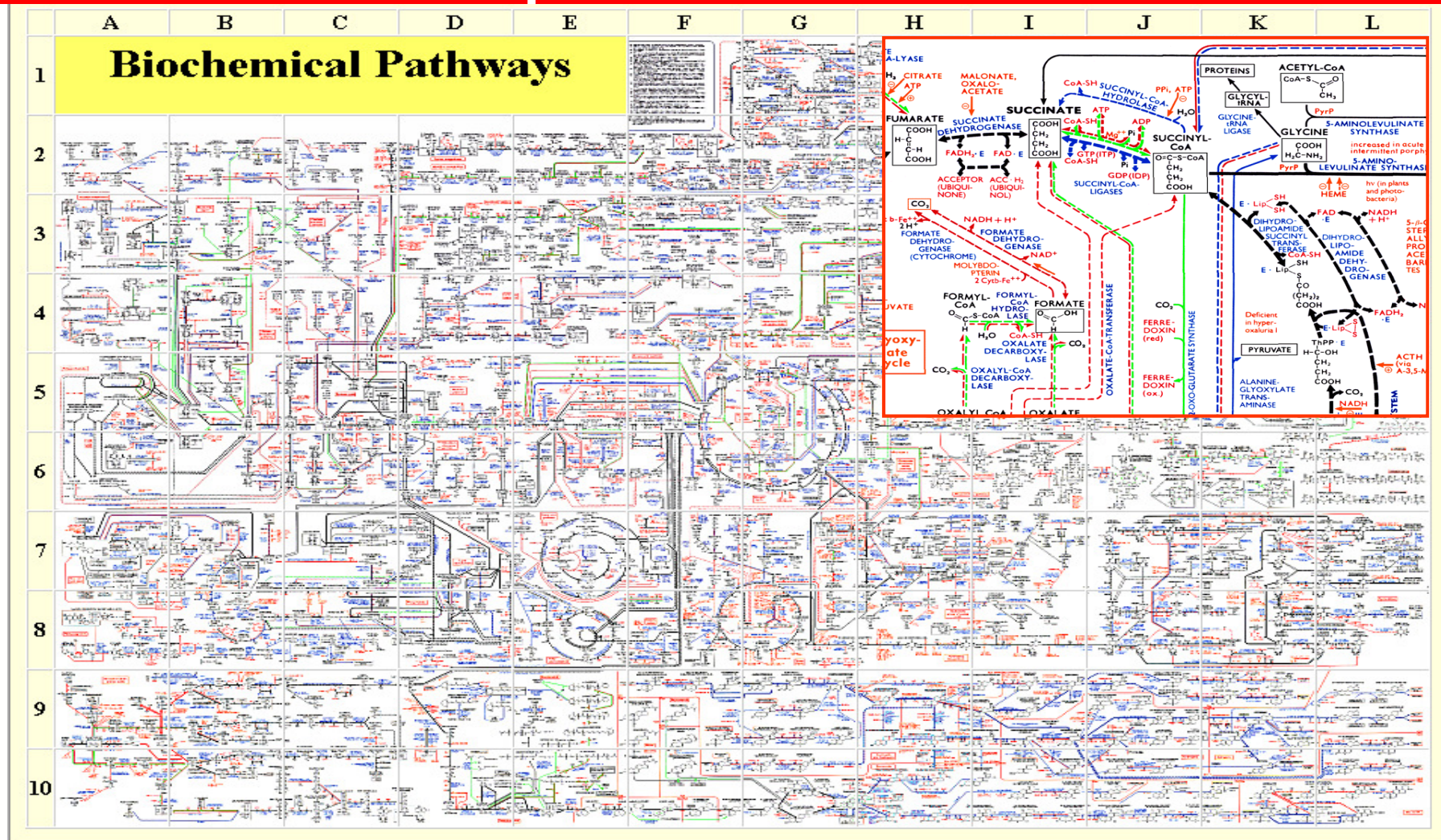


METABOLISMO

Reacciones
bio-químicas

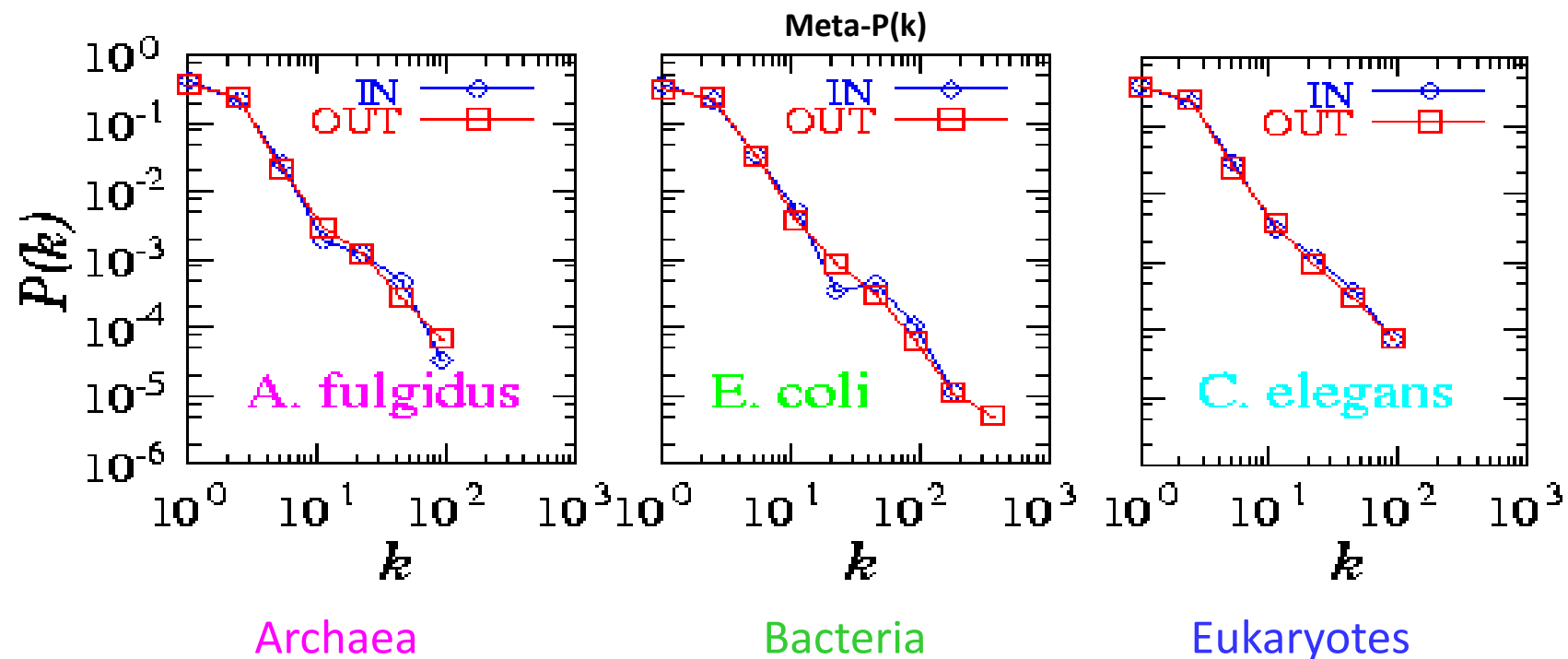
REDES GENÉTICAS (3)

BOEHRING-MENNHEIN



REDES GENÉTICAS (4)

Redes metabólicas



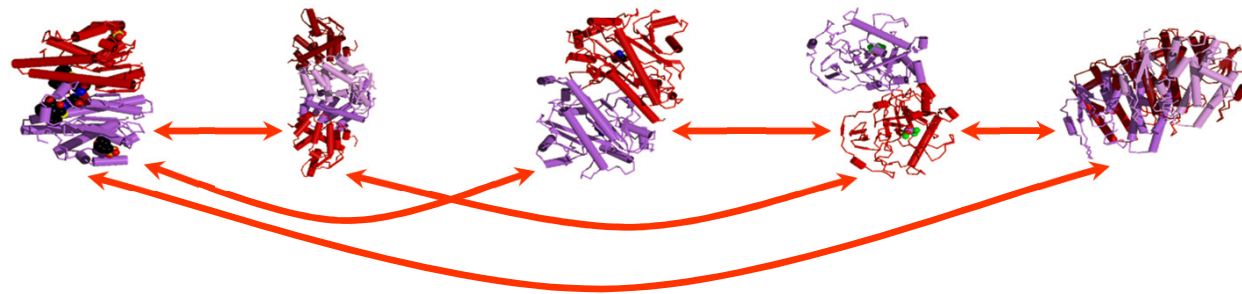
¡Las redes de los organismos de los tres dominios de la vida son **libres de escala**!

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

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

H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, *Nature* 407:651-654 (2000)

REDES GENÉTICAS (5)



PROTEOMA

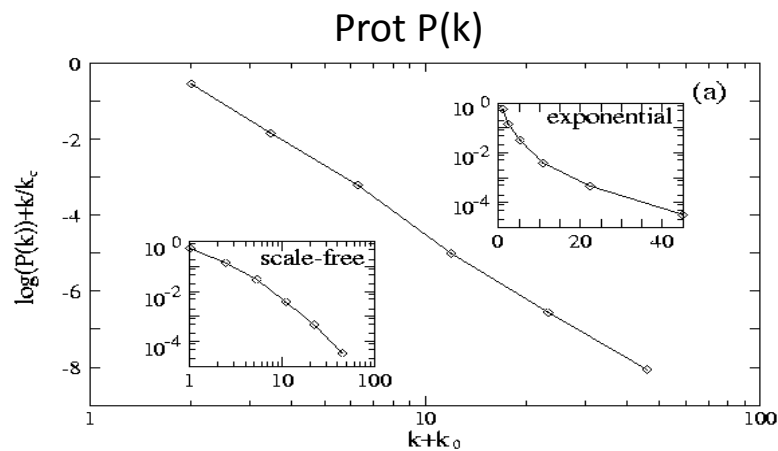
Interacciones
proteína-proteína

REDES GENÉTICAS (6)

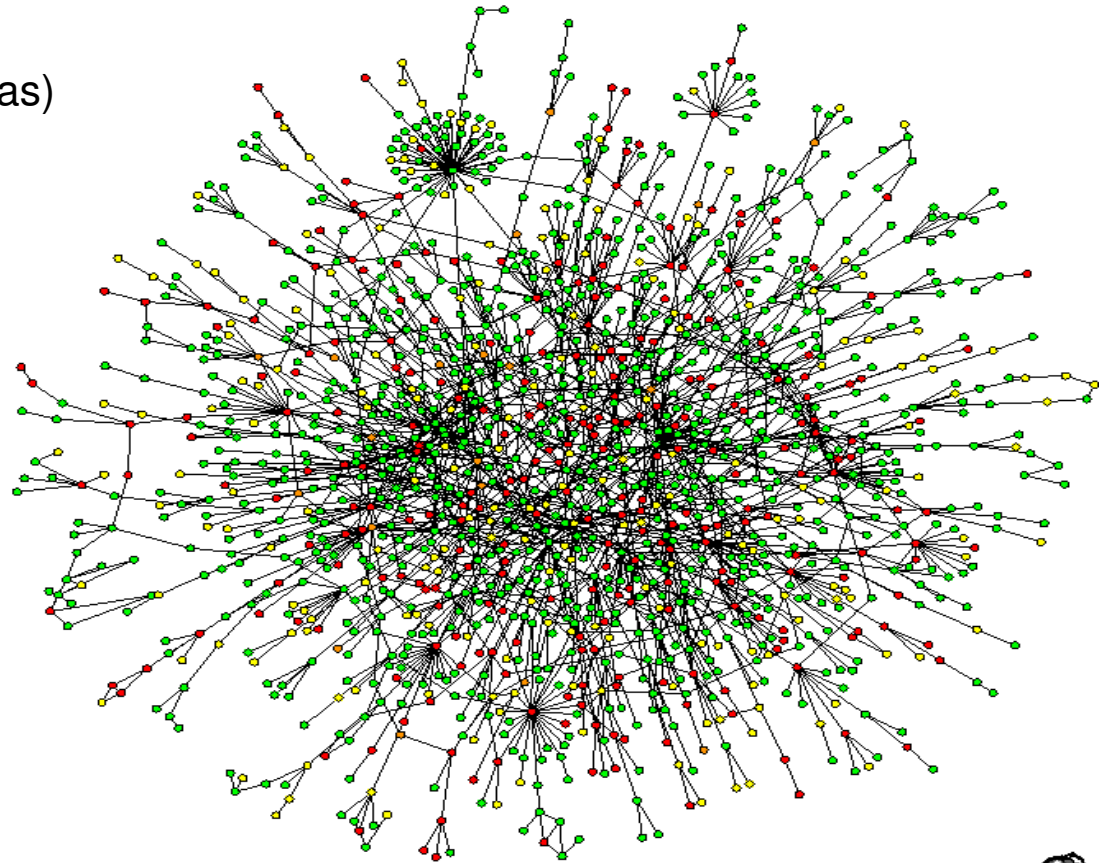
Topología de las redes de proteínas (1)

Nodos: proteínas

Enlaces: interacciones físicas (ligaduras)



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

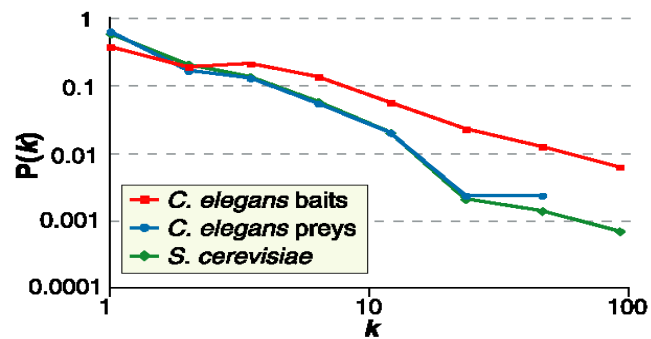
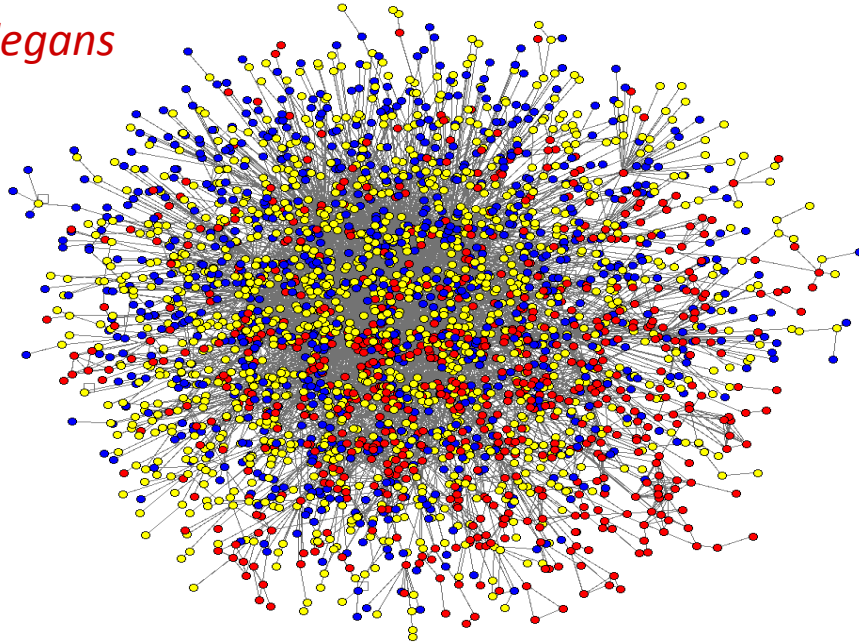


H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, *Nature* 407:651-654 (2000)

REDES GENÉTICAS (7)

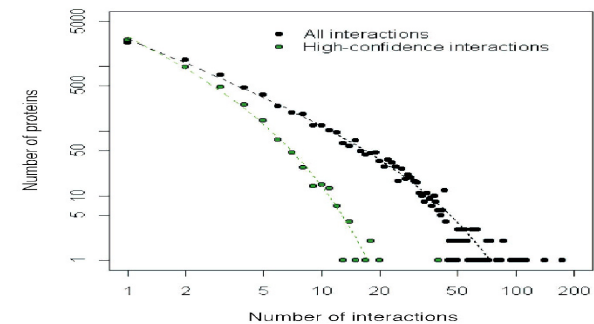
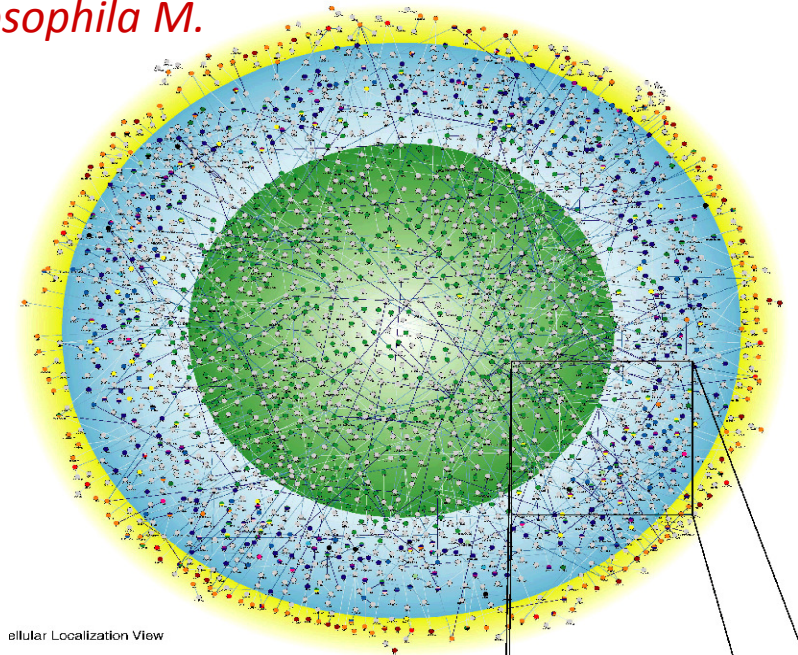
Topología de las redes de proteínas (2)

C. Elegans



Li et al., Science 303(5657):540-5434 (2004)

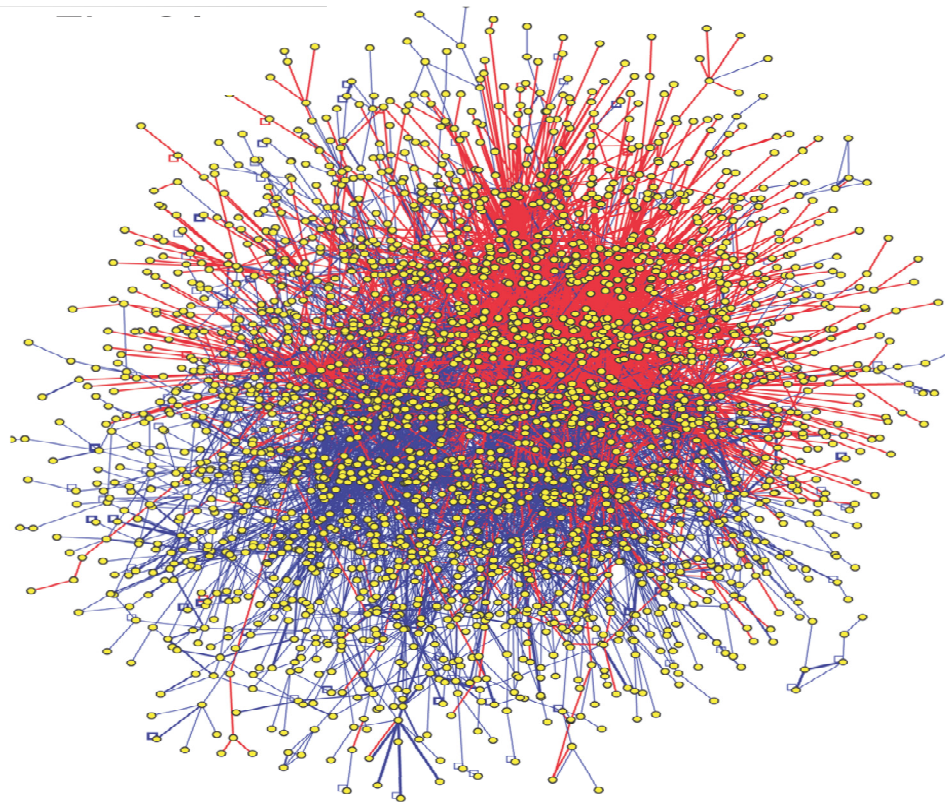
Drosophila M.



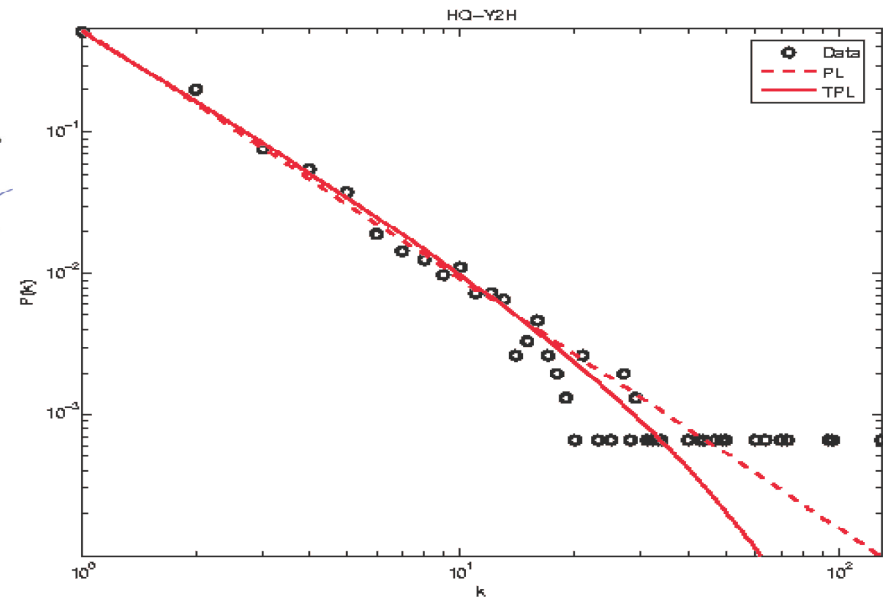
Giot et al., Science 302(5651):1727-1736 (2003)

REDES GENÉTICAS (8)

Red de interacción humana



2,800 Y2H interacciones
4,100 interacciones binarias LC
(HPRD, MINT, BIND, DIP, MIPS)



Rual et al., Nature 43:1173-1178 (2005)

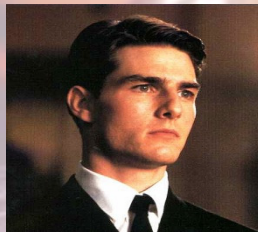
;

Stelzl et al., Cell 122(6):957-968 (2005)

RED DE ACTORES DE HOLLYWOOD

Nodos: actores

Enlaces: actuaciones conjuntas



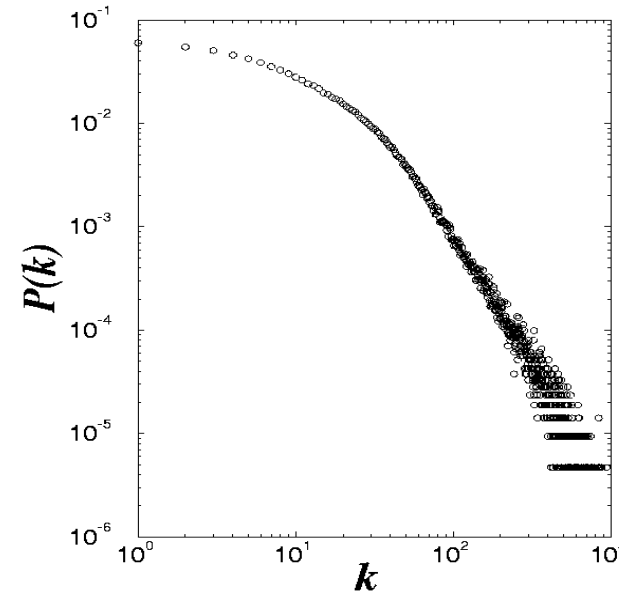
Days of Thunder (1990)
Far and Away (1992)
Eyes Wide Shut (1999)



$N = 212,250$ actores
 $\langle k \rangle = 28.78$

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

$\gamma = 2.3$



EN RESUMEN...

Una gran cantidad de redes reales comparten una arquitectura similar:

Redes libres de escala

WWW, Internet (routers y dominios), circuitos electrónicos, software de ordenador, actores de películas, redes de coautoría, redes sexuales, mensajería instantánea, redes de e-mail, citas, llamadas telefónicas, redes metabólicas, redes de interacciones de proteínas, dominios de proteínas, redes de funciones cerebrales, redes lingüísticas, personajes de comic, comercio internacional, sistemas bancarios, redes de confianza para encriptación, plataformas energéticas, terremotos, redes astrofísicas...

RESUMEN HISTÓRICO

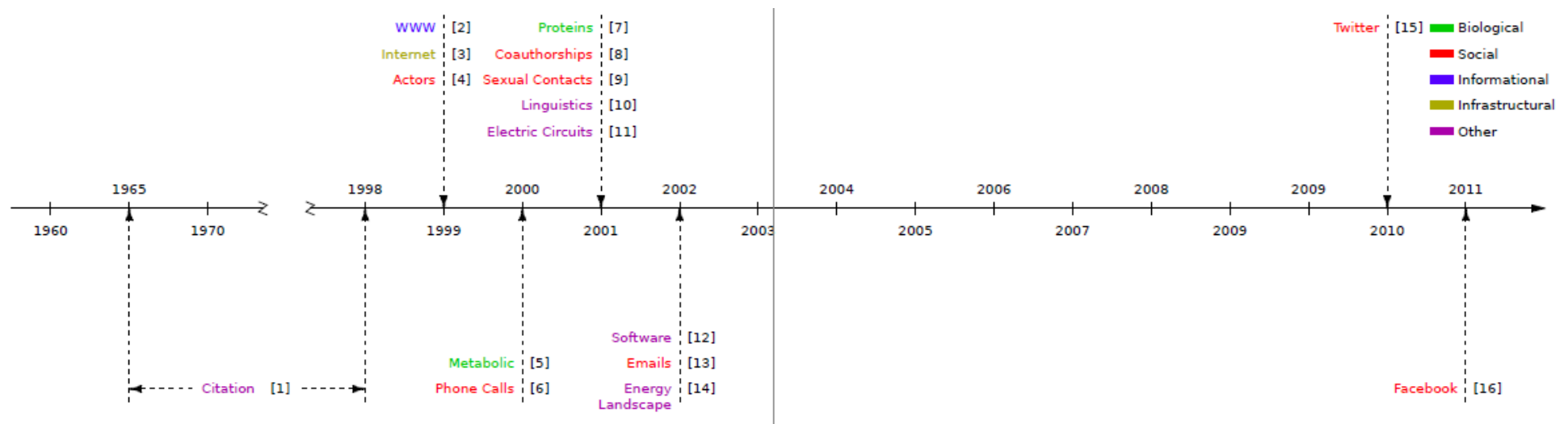


Figure 4.10: Scale-free history.

Many biological, social, and technological networks display the scale-free property. The figure shows the timeline of the discoveries reporting the scale-free nature of various real networks. While there is a clear burst of reports following the 1999 discovery of scale-free networks, in hindsight it is clear that several early papers have reported characteristics that are consistent with what we call today a scale-free topology. For example, Etel de Solla Price reported in 1965 that citations to scientific papers follow a power-law distribution [7], a property independently discovered by Redner in 1998 [8]. This is a consequence of the scale-free nature of citation networks. A common feature of these early works is that they viewed the observed quantities as scalar events, not as a manifestation of some network phenomena. It wasn't until the 1999 that it was understood that power laws are also a fundamental *network* property. Indeed, Barabási and Albert, in their 1999 *Science* paper argued that "we expect that the scale-invariant state observed in all systems for which detailed data has been available to us is a generic property of many complex networks, with applicability reaching far beyond the quoted examples." The 'scale-free network' term was also first used in 1999 [2,9].

The figure shows the timeline of these discoveries, marking the year the evidence about the scale-free property of a particular network has emerged. The number in the parenthesis offers pointers to the respective publications, as follows: **[1]** de Solla Price (1965)[7], Redner(1998)[8]; **[2]** Albert et al. (1999) [1], Kumar et al. (1999) [10]; **[3]** Faloutsos et al. (1999) [5]; **[4]** Barabási et al. (1999) [2]; **[5]** Jeong et al. (2000) [11]; **[6]** Aiello et al. (2000) [12]; **[7]** Jeong et al. (2001) [13]; **[8]** Barabási et al. (2001) [14], Newman (2001) [15]; **[9]** Liljeros et al. (2001) [16]; **[10]** Ferrer i Cancho et al. (2001) [17], Yook et al. (2001) [18]; **[11]** Ferrer i Cancho et al. (2001) [19]; **[12]** Valverde et al. (2002) [20]; **[13]**; Ebel et al. (2002) [21]; **[14]** Doye (2002) [22]; **[15]** Kwak et al. (2010) [23], Cha et al. (2010) [24]; **[16]** Ugander et al. (2011) [25].

UNIVERSALIDAD EN LOS EXPONENTES DE GRADO (1)

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, <i>E. coli</i>	778	7.4	110	2.2	2.2
Protein, <i>S. cerev.</i> *	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

Fenómeno crítico:

La universalidad implica que los exponentes son los mismos para distintos sistemas complejos... son independientes de los detalles concretos

Redes:

Los exponentes varían dependiendo del sistema

La mayoría están entre 2 y 3

Universalidad:

Se da una emergencia de características comunes entre las distintas redes

Ocurre igual que en el caso de la propiedad de libertad de escala

UNIVERSALIDAD EN LOS EXPONENTES DE GRADO (2)

NETWORK	N	L	$\langle k \rangle$ $\langle k_{in} \rangle = \langle k_{out} \rangle$	σ_{in}	σ_{out}	σ	γ_{in}	γ_{out}	γ
Internet	192,244	609,066	6.34	-	-	14.14	-	-	3.42*
WWW	325,729	1,497,134	4.60	39.05	21.48	-	2.31	2.00	-
Power Grid	4,941	6,594	2.67	-	-	1.79	-	-	Exp.
Mobile Phone Calls	36,595	91,826	2.51	2.39	2.32	-	4.69*	5.01*	-
Email	57,194	103,731	1.81	9.56	34.07	-	3.43*	2.03	-
Science Collaboration	23,133	93,439	8.08	-	-	10.63	-	-	3.35
Actor Network	702,388	29,397,908	83.71	-	-	200.86	-	-	2.12
Citation Network	449,673	4,689,479	10.43	29.37	9.49	-	3.03**	4.00	-
E. Coli Metabolism	1,039	5,802	5.58	22.46	19.12	-	2.43	2.90	-
Yeast Protein Interactions	2,018	2,930	2.90	-	-	4.88	-	-	2.89*

Table 4.1: The characteristics of several real network.

The table shows the standard deviation of the degree distribution, $\sigma_k = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}$ (σ_{in} and σ_{out} for directed networks) for our ten reference networks. It indicates that for most networks σ is much larger than $\langle k \rangle$, consequence of their scale-free nature. It also lists the estimated degree exponent, γ , for each network, determined using the procedure discussed in Advanced Topics 4.A. The stars next to the reported values indicate the statistical confidence for a particular fit to the degree distribution. That is, * means that the fit shows statistical confidence for a *power-law* $k^{-\gamma}$ fit; while ** marks datasets that display statistical confidence for a $(k + k_{sat})^{-\gamma} e^{-k/k_{cut}}$ fit. Those with no stars do not show statistical confidence for any of the two forms; the reasons for this are discussed later in the next chapter and in Advanced Topics 4.C. Note that the power grid is not considered scale-free. For this network a degree distribution of the form $e^{-\lambda k}$ offers a statically significant fit.

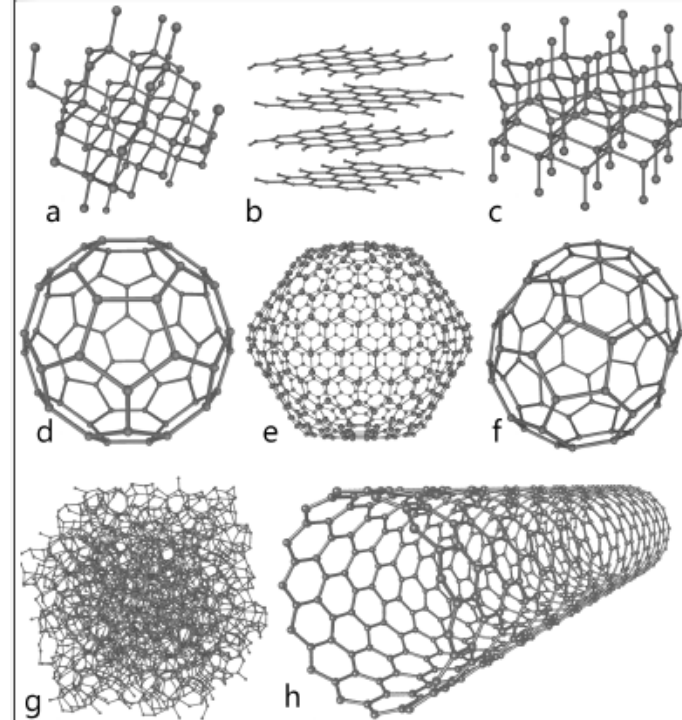
...PERO NO TODAS LAS REDES REALES SON LIBRES DE ESCALA

Box 4.2 Not all networks are scale-free.

The ubiquity of the scale-free property does not mean that *all* real networks are scale-free. Indeed, several important networks do not share this property:

- 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.

For the scale-free property to emerge the nodes need to have the capacity to link to an arbitrary number of other nodes. These links do not need to be simultaneous: we do not constantly chat with each of our acquaintances and a protein in the cell does not simultaneously bind to each of its potential interaction partners. In general the scale-free property is absent in systems that have a limitation in the number of links a node can have, as such limitations limit the size of the hubs. As illustrated in the image, such limitations are common in materials, explaining why they cannot develop a scale-free topology.



The material network.

A carbon atom can share only four electrons with other atoms, hence no matter how we arrange these atoms relative to each other, in the resulting network a node can never have more than four links. Hence, hubs are forbidden and the scale-free property cannot emerge. The figure shows several carbon allotropes, each characterized by a different "network", resulting in materials with different physical characteristics, like (a) diamond; (b) graphite; (c) lonsdaleite; (d) C60 (buckminsterfullerene); (e) C540 (a fullerene) (f) C70 (another fullerene); (g) amorphous carbon; (h) single-walled carbon nanotube. After http://www.thenanoage.com/Figures/Eight_Allotropes_of_Carbon.png

EFECTO DE LOS EXPONENTES DE GRADO EN LAS PROPIEDADES DE LAS REDES LIBRES DE ESCALA

PROPIEDAD DE MUNDOS ULTRA-PEQUEÑOS (1)

¿Cómo afecta la presencia de hubs a la propiedad de mundos pequeños?

Las distancias en una red libre de escala son o bien menores o bien iguales a las existentes en una red aleatoria equivalente

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

Mundo ultra-pequeño	$\langle l \rangle \sim$	$const.$	$\gamma = 2$	El tamaño del mayor hub es de orden $O(N)$. La mayoría de los nodos están conectados a él, haciendo que la distancia media sea independiente del tamaño del sistema
		$\frac{\ln \ln N}{\ln(\gamma - 1)}$	$2 < \gamma < 3$	La distancia media se incrementa en un orden menor que el logarítmico. En una red aleatoria todos los nodos tienen un grado similar, por lo que muchos caminos tienen una longitud similar. En una red libre de escala, la gran mayoría de los caminos pasan a través de unos pocos hubs de grado muy alto, reduciendo las distancias entre los nodos
		$\frac{\ln N}{\ln \ln N}$	$\gamma = 3$	Algunos modelos clave producen $\gamma=3$. Es un resultado particularmente importante, un punto crítico . El segundo momento de la distribución de grados ya no diverge y las distancias vuelven a tener la dependencia logarítmica de los mundos pequeños en las redes aleatorias, aunque suavizado
Mundo pequeño		$\ln N$	$\gamma > 3$	El segundo momento es finito y la red pasa a comportarse como una red aleatoria. Aunque sigue habiendo hubs, no son suficientemente grandes para tener un impacto significativo en las distancias

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

PROPIEDAD DE MUNDOS ULTRA-PEQUEÑOS (2)

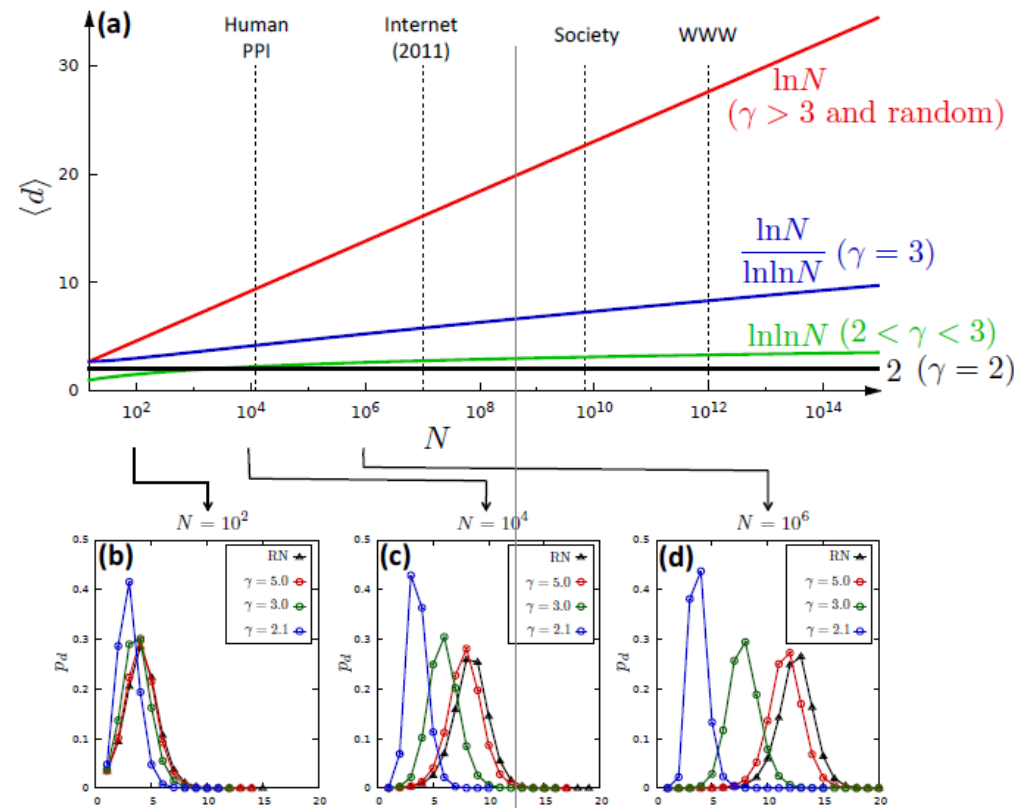
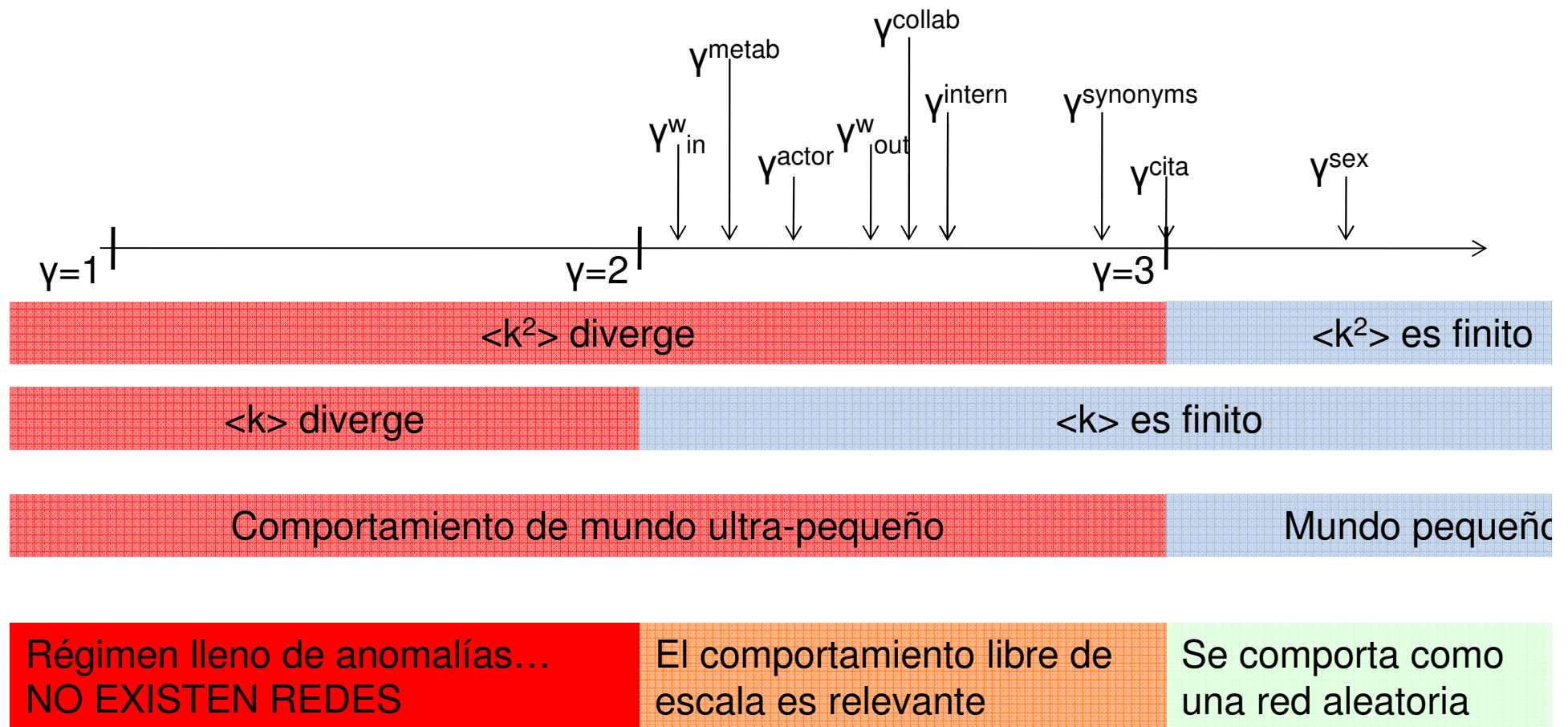


Figure 4.1 I: Distances in scale-free networks.

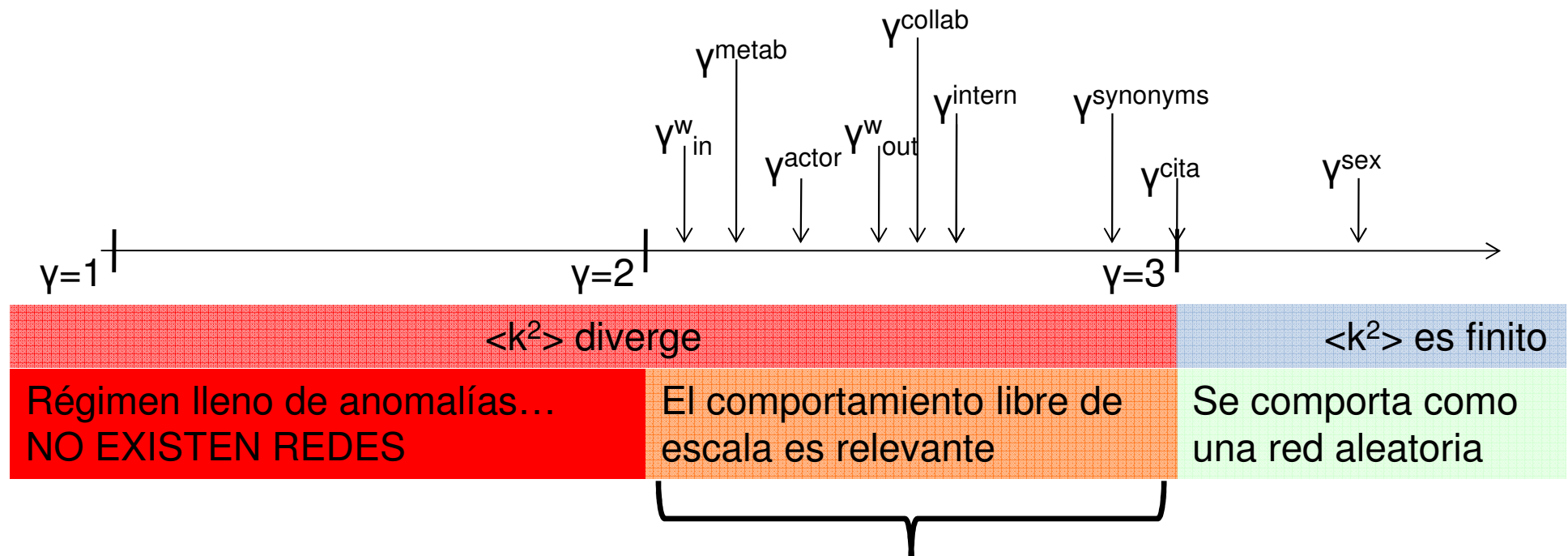
(a) The scaling of the average path length in the four scaling regimes characterizing a scale-free network: $\ln N$ (scale-free networks with $\gamma > 3$ and random networks), $\ln N / \ln \ln N$ ($\gamma = 3$) and $\ln \ln N$ ($2 < \gamma < 3$). The dotted lines mark the approximate size of several real networks of practical interest. For example, given their modest size, in biological networks the differences in the node to node distances are relatively small in the four regimes. The differences become quite relevant for networks of the size of the social network or the WWW. For these the small-world formula considerably underestimates the real value of $\langle d \rangle$.

(b),(c),(d) **Distance distribution** for networks of size $N = 10^2, 10^4, 10^6$, illustrating that while for small N ($= 10^2$) the distance distributions is not too sensitive to γ , for large N ($= 10^6$) p_d and $\langle d \rangle$ changes visibly with γ . As (d) shows, the smaller γ , the shorter are the distances between the nodes. The networks were generated using the static model [29] with $\langle k \rangle = 3$.

REGÍMENES DE LAS REDES LIBRES DE ESCALA (1)



REGÍMENES DE LAS REDES LIBRES DE ESCALA (2)



¿POR QUÉ NO OBSERVAMOS REDES CON EXPONENTE $\gamma=4,5,6$, ETC?

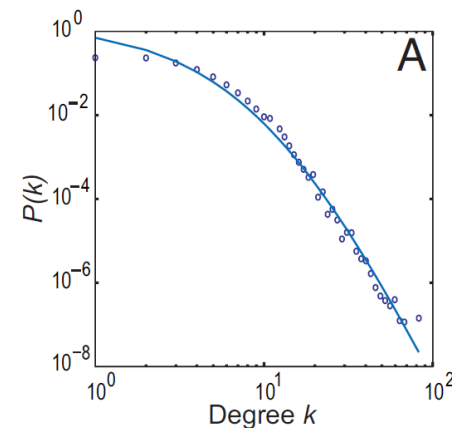
Pueden existir redes libres de escala de este tipo pero es difícil distinguirlas de las redes aleatorias

Para ello, se necesitan 2-3 órdenes de magnitud de escalado, es decir, k_{max} tiene que ser $10^2 - 10^3$ veces más grande que k_{min}

Sin embargo, esto restringe el tamaño del sistema requerido para poder documentar esas redes. P.ej. Para medir un exponente $\gamma=5$, necesitamos que el sistema presente un grado máximo del orden de:

$$k_{max} = k_{min} N^{\frac{1}{\gamma-1}}$$
$$N = \left(\frac{k_{max}}{k_{min}} \right)^{\gamma-1} \approx 10^{12}$$

Onnela et al., PNAS 104(18):7332–7336 (2007)



Red de llamadas
de móviles

$N=4.6 \cdot 10^6$
 $\gamma=8.4$

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 $P(x) = a(x + x_0)^{-\gamma} \exp(-x/x_c)$, shown as a blue curve, where x corresponds to either k or w . The parameter values for the fits are $k_0 = 10.9$, $\gamma_k = 8.4$, $k_c = \infty$ (A, degree), and $w_0 = 280$, $\gamma_w = 1.9$, $w_c = 3.45 \times$

RESUMEN DE PROPIEDADES DE LAS REDES LIBRES DE ESCALA

Box 4.7: At a Glance: Scale-free Networks.

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 :

$2 < \gamma < 3$: $\langle k \rangle$ finite, $\langle k^2 \rangle$ diverges when $N \rightarrow \infty$.

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

Distances in a scale-free network:

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

The Second Law: Scale-free property.

Many real networks are characterized by a fat-tailed degree distribution. This means that many small degree nodes are held together by a few hubs.

Referencias y Agradecimientos

Para diseñar los materiales de este tema, he hecho uso de material desarrollado por expertos en el área disponible en Internet:

- “Network Science Interactive Book Project” del Laszlo Barabasi Lab:
<http://barabasilab.com/networksciencebook>

