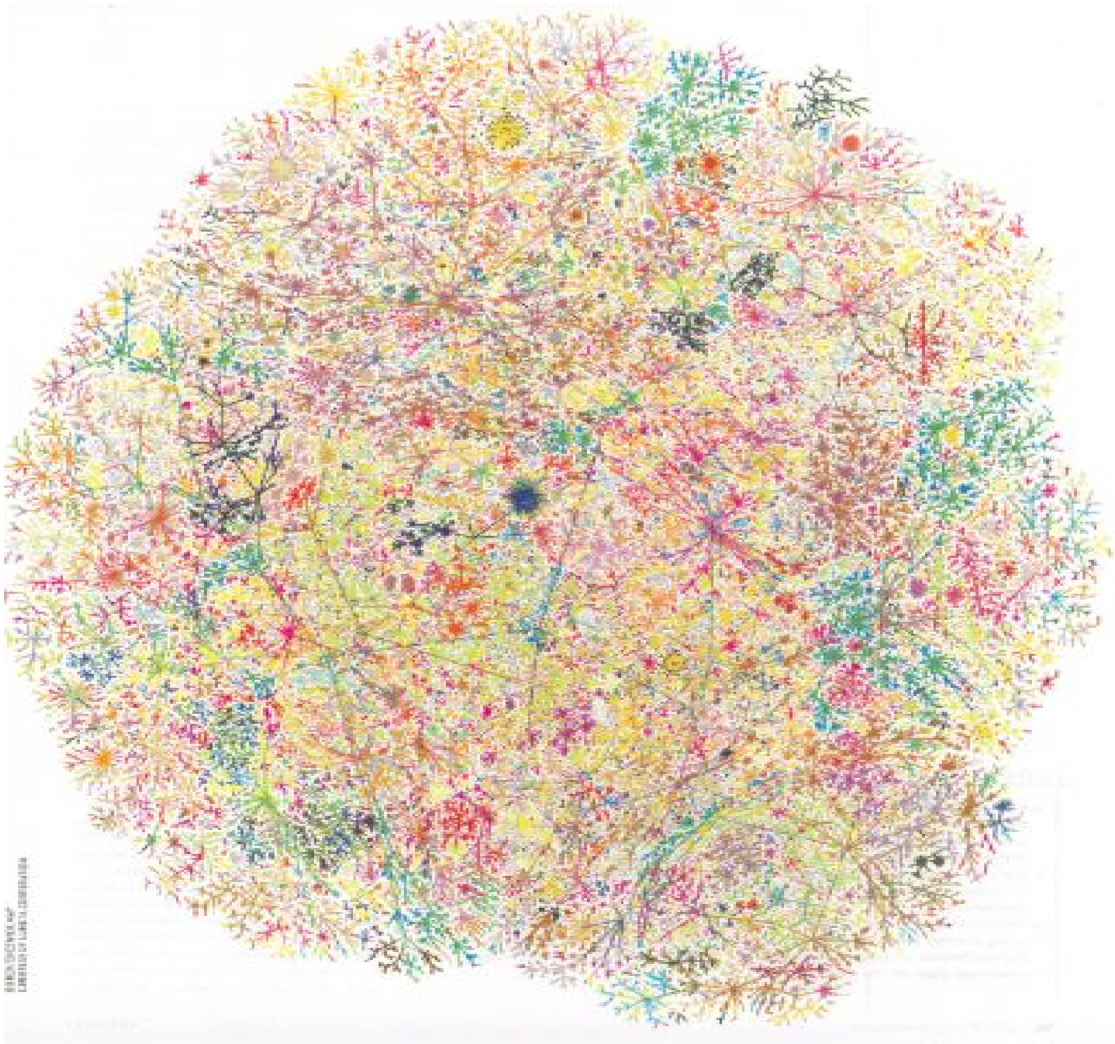
Scale free



An important question regarding the average path length is whether the onset of small-world behavior is dependent on the system size. It was Watts (1999) who first noticed that ℓ does not begin to decrease until $p \geqslant 2/NK$, guaranteeing the existence of at least one shortcut. This implies that the transition p depends on the system size, or conversely, there exists a p-dependent crossover length (size) N^* such that if $N < N^*$, $\ell \sim N$, but if $N > N^*$, $\ell \sim \ln(N)$. The concept of the crossover size was introduced by Barthélémy and Amaral (1999), who conjectured that the characteristic path length scales as (see Fig. 17)

$$\ell(N,p) \sim N^* F\left(\frac{N}{N^*}\right), \tag{67}$$

where

$$F(u) = \begin{cases} u & \text{if} \quad u \ll 1 \\ \ln(u) & \text{if} \quad u \gg 1. \end{cases} \tag{68}$$

Numerical simulations and analytical arguments (Barrat 1999; Barthélémy and Amaral, 1999; Newman and Watts, 1999a; Argollo de Menezes *et al.*, 2000; Barrat and Weigt, 2000) concluded that the crossover size N^* scales with p as $N^* \sim p^{-\tau}$, where $\tau = 1/d$ and d is the dimension of the original lattice to which the random edges are added (Fig. 18). Thus for the original WS model, defined on a circle ($d=1$), we have $\tau = 1$, the onset of small-world behavior taking place at the rewiring probability $p^* \sim 1/N$.

It is now widely accepted that the characteristic path length obeys the general scaling form

$$\ell(N,p) \sim \frac{N^{1/d}}{K} f(pKN), \tag{69}$$

where $f(u)$ is a universal scaling function that obeys

$$f(u) = \begin{cases} \text{const} & \text{if} \quad u \ll 1 \\ \ln(u)/u & \text{if} \quad u \gg 1. \end{cases} \tag{70}$$

Distribucion de Grados en un Network

Grado de un nodo= numero de links que “salen de un nodo” (para networks dirigidos hay que ver “in” - “out”

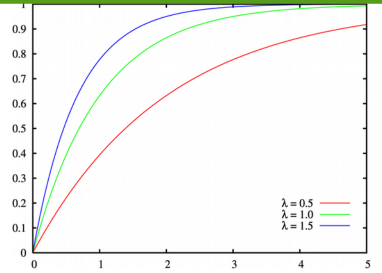
Para el random network:
siendo k el grado
La densidad de probabilidad del grado k es:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$ (binomial)

Tambien se puede trabajar con la Cumulativa

$$P_k = \sum_{k'=k}^{\infty} P(k)$$

exp



distribution function

$$\widetilde{P}_k = \sum_{k'=1}^{k} P(k)$$

Supongamos que la cola de una densidad es tipo power law

$$P(k) = C\,k^{-\alpha} \qquad \text{Para } k \geq k_{min}$$

Entonces

$$P_k = C\sum_{k'=k}^{\infty} k'^{-\alpha} \simeq \int_{k'}^{\infty} k'^{-\alpha}\, dk' = \frac{C}{\alpha-1} k^{-(\alpha-1)}$$

De nuevo power

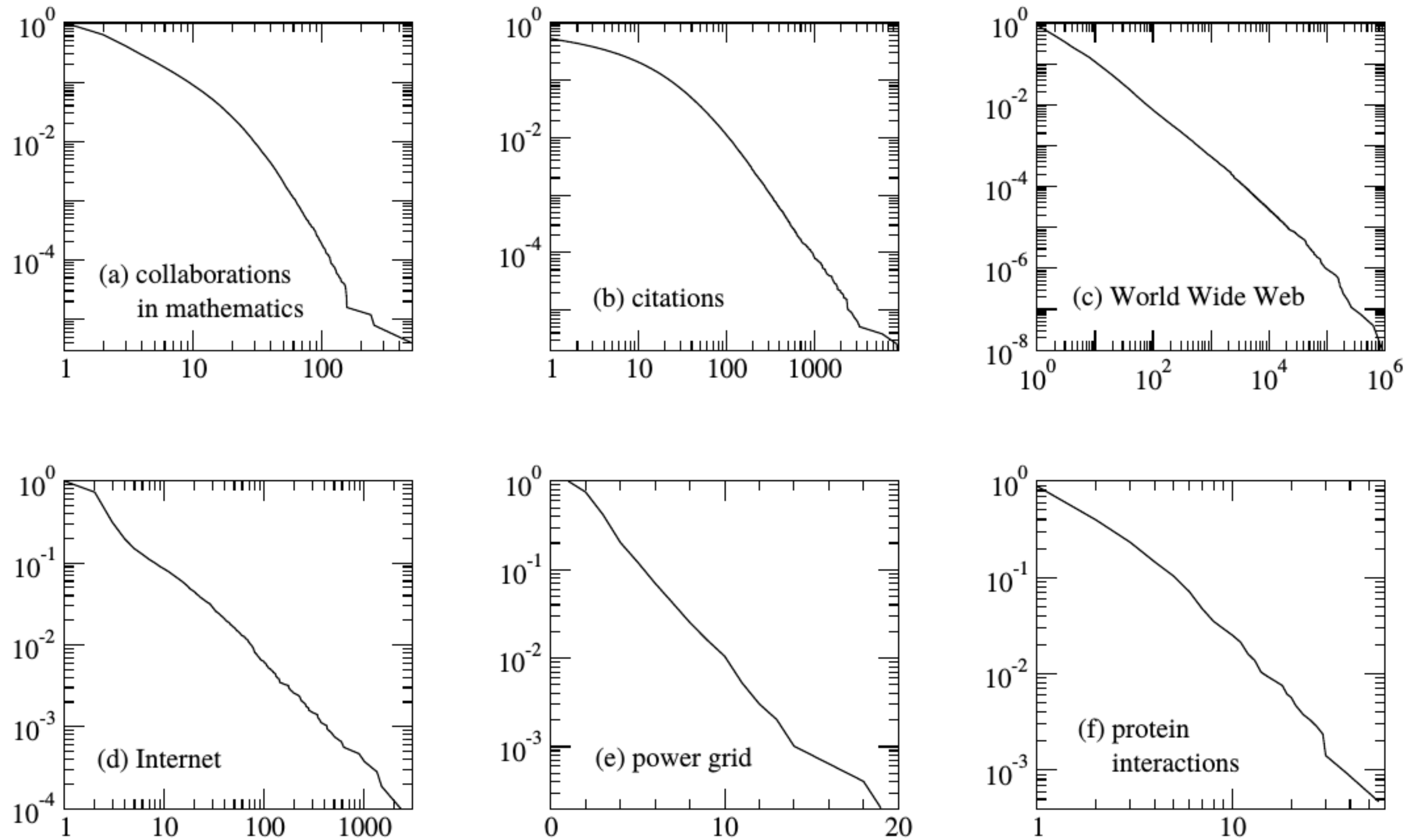


FIG. 6 Cumulative degree distributions for six different networks. The horizontal axis for each panel is vertex degree k (or in-degree for the citation and Web networks, which are directed) and the vertical axis is the cumulative probability distribution of degrees, i.e., the fraction of vertices that have degree greater than or equal to k. The networks shown are: (a) the collaboration network of mathematicians [182]; (b) citations between 1981 and 1997 to all papers cataloged by the Institute for Scientific Information [351]; (c) a 300 million vertex subset of the World Wide Web, *circa* 1999 [74]; (d) the Internet at the level of autonomous systems, April 1999 [86]; (e) the power grid of the western United States [416]; (f) the interaction network of proteins in the metabolism of the yeast *S. Cerevisiae* [212]. Of these networks, three of them, (c), (d) and (f), appear to have power-law degree distributions, as indicated by their approximately straight-line forms on the doubly logarithmic scales, and one (b) has a power-law tail but deviates markedly from power-law behavior for small degree. Network (e) has an exponential degree distribution (note the log-linear scales used in this panel) and network (a) appears to have a truncated power-law degree distribution of some type, or possibly two separate power-law regimes with different exponents.

distribucion de grados nodos salientes

distribucion de grados nodos entrantes

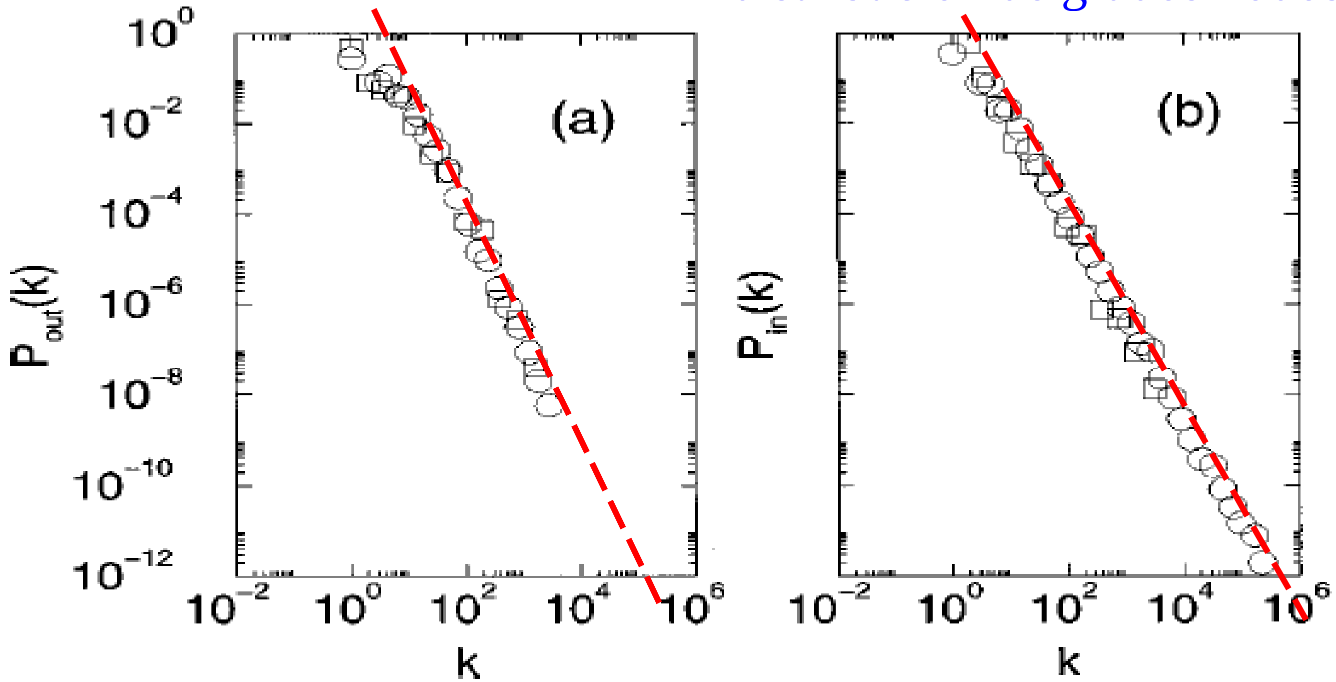


FIG. 2. Degree distribution of the World Wide Web from two different measurements: □, the 325 729-node sample of Albert *et al.* (1999); ○, the measurements of over 200 million pages by Broder *et al.* (2000); (a) degree distribution of the outgoing edges; (b) degree distribution of the incoming edges. The data have been binned logarithmically to reduce noise. Courtesy of Altavista and Andrew Tomkins. The authors wish to thank Luis Amaral for correcting a mistake in a previous version of this figure (see Mossa *et al.*, 2001).

Distribucion De grados Networks reales

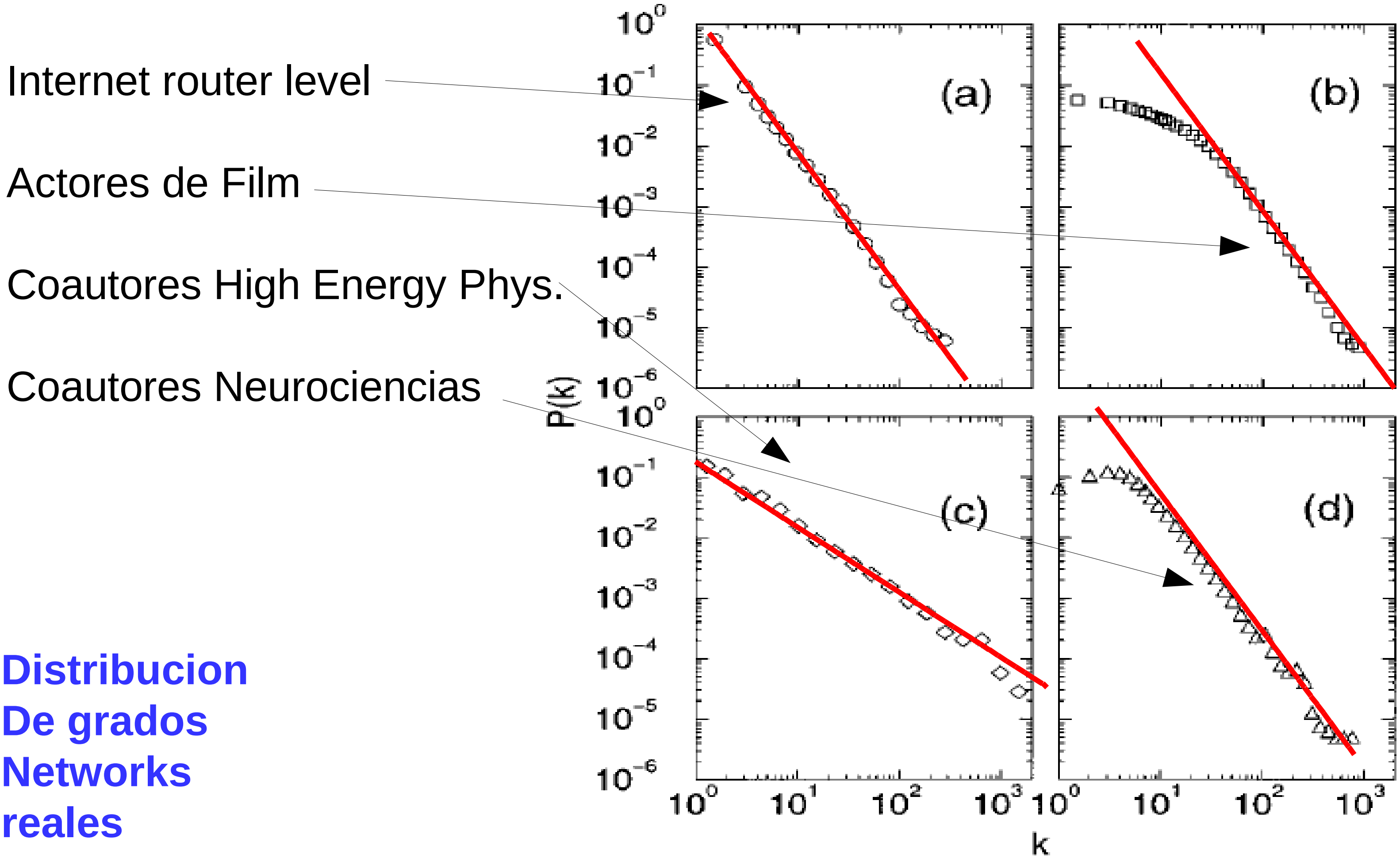


FIG. 3. The degree distribution of several real networks: (a) Internet at the router level. Data courtesy of Ramesh Govindan; (b) movie actor collaboration network. After Barabási and Albert 1999. Note that if TV series are included as well, which aggregate a large number of actors, an exponential cutoff emerges for large k (Amaral *et al.*, 2000); (c) co-authorship network of high-energy physicists. After Newman (2001a, 2001b); (d) co-authorship network of neuroscientists. After Barabási *et al.* (2001).

Growing Scale free networks

REVIEWS OF MODERN PHYSICS, VOLUME 74, JANUARY 2002

Statistical mechanics of complex networks

Réka Albert* and Albert-László Barabási

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

Los metodos usuales hast digamos 1999 tomaban un numero dado de nodos y luego imponian condiciones varias sobre los links

Modelo de Albert_Barabasi

(para ellos SW effect $\Rightarrow l$ pequeño y C grande)

Sea un Sistema de links no dirigidos.

Caracteristicas del sistema (propiedades genericas propuestas por Barabasi y Albert)

a) crece, con el agregado de nuevos nodos

b) preferential attachment, los nodos agregados no se conectan al azar sino que responde a propiedades del netowork existente como ser la distribucion de grados.

En el caso de Albert-Barabasi

Cuales son los ingredientes fundamentales?

a) El crecimiento → se comienza con un numero pequeño de nodos m_0 y a cada paso se agrega un nuevo nodo con m nuevos links $(m \leq m_0)$

b) Preferential attachment → los nodos a los que se liga el nuevo nodo se eligen con probabilidad proporcional a

$$\Pi(k) = \frac{k_i}{\sum k_j}$$

luego de t pasos el network tiene:

a) nodos : $N = t + m_0$

b) lados : mt

Vemos un crecimiento

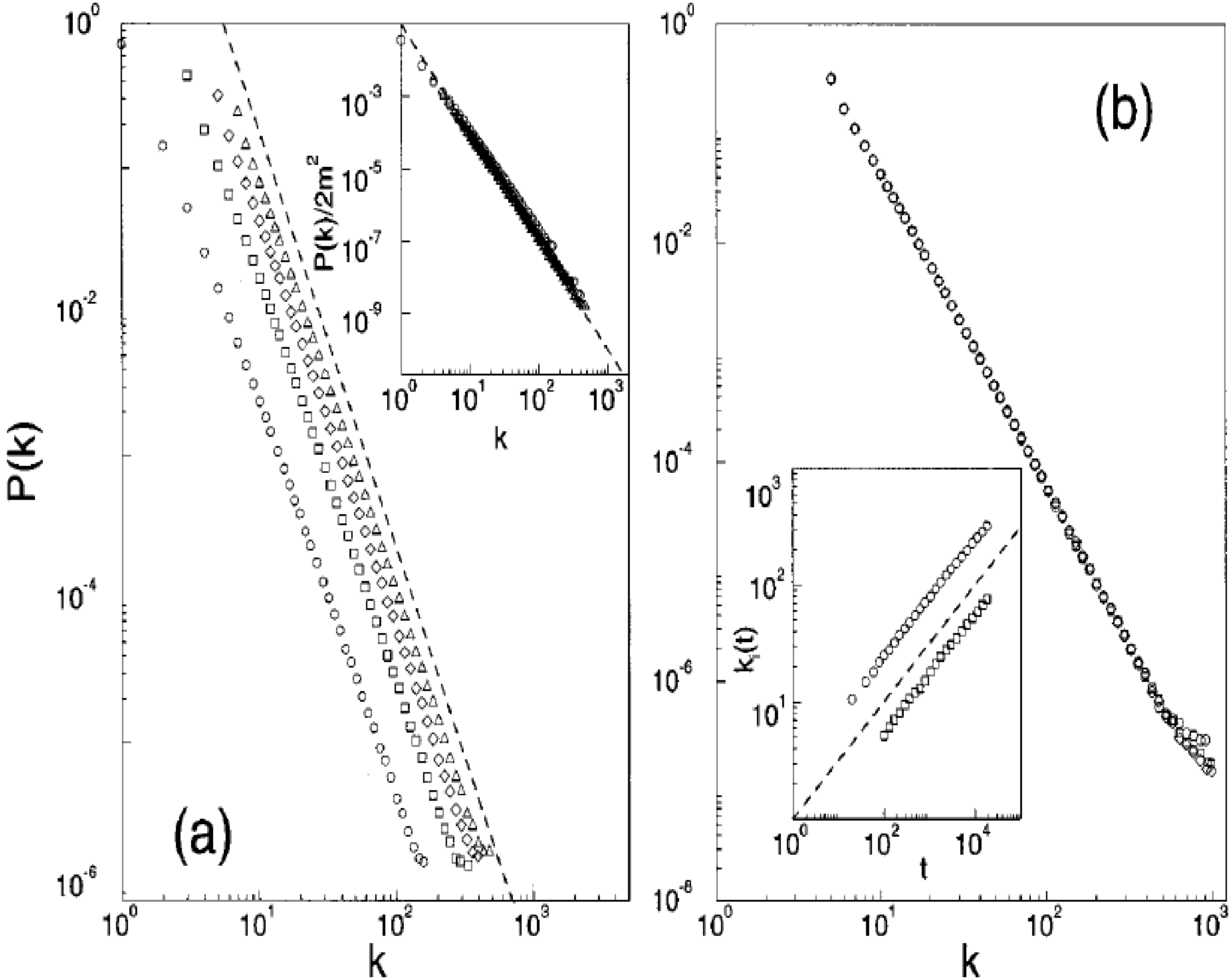
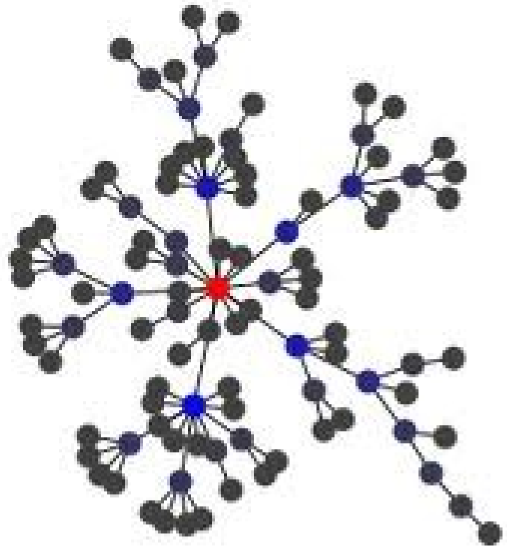
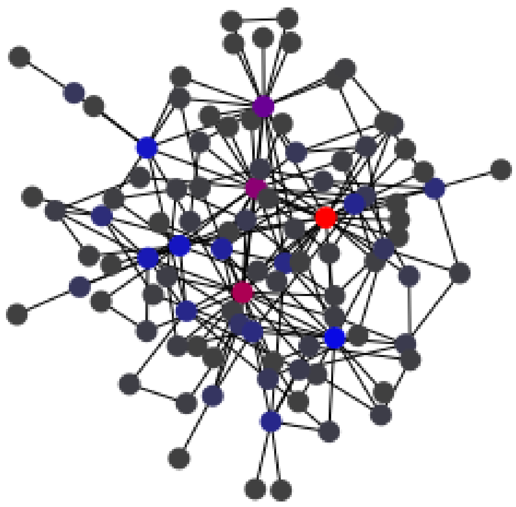


FIG. 21. Numerical simulations of network evolution: (a) Degree distribution of the Barabási-Albert model, with $N=m_0$ $=300\,000$ and ○, $m_0=m=1$; □, $m_0=m=3$; ◇, $m_0=m=5$; and △, $m_0=m=7$. The slope of the dashed line is $\gamma=2.9$, provid the best fit to the data. The inset shows the rescaled distribution (see text) $P(k)/2m^2$ for the same values of m, the slope of t dashed line being $\gamma=3$; (b) $P(k)$ for $m_0=m=5$ and various system sizes, ○, $N=100\,000$; □, $N=150\,000$; ◇, $N=200\,000$. T inset shows the time evolution for the degree of two vertices, added to the system at $t_1=5$ and $t_2=95$. Here $m_0=m=5$, and t dashed line has slope 0.5, as predicted by Eq. (81). After Barabási, Albert, and Jeong (1999).

Un link por paso



tres links por pas



Deduccion teorica (la de Albert y Barabasi) modelo continuo

Proba de ligarse
Al k_i

para un nodo i

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

Numero de links a agregar

Donde el $N-1$ es porque no se incluye el nodo recien agregado.

$$\sum_i k_i = 2N$$

El denominador resulta entonces:

$$\sum_{j=1}^{N-1} k_j = \sum_{j=1}^{N} k_j - 2m = 2mt - 2m$$

que tiende a $2mt$ cuando t es grande , entonces

luego de t pasos el network tiene:

a) nodos : $N = t + m_0$

b) lados : mt

$$\frac{\partial k_i}{\partial t} = m \frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

$$\frac{\partial k_i}{\partial t} = m \frac{k_i}{2mt} = \frac{k_i}{2t}$$

tomando en cuenta que al tiempo de introduccion del nodo al sistema, en el tiempo t_i,

para un nodo i

$$k_i(t_i) = m$$

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t} \Rightarrow$$

$$\ln(k_i(t)) = \frac{1}{2}\ln t + c$$

en t_i $$\ln(m) = \frac{1}{2}\ln t_i + c \Rightarrow$$

t_i es el tiempo de introducción

$$c = \ln(m) - \frac{1}{2}\ln t_i \Rightarrow$$

$$\ln(k_i(t)) = \frac{1}{2}\ln t + \ln(m) - \frac{1}{2}\ln t_i$$

$$= \ln\left[m\left(\frac{t}{t_i}\right)^{1/2}\right]$$

Resulta que

depende del tiempo de inicio t_i

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}$$

Con $\beta = \frac{1}{2}$

A partir de esta ecuacion se puede escribir la probabilidad de que un nodo tenga $k_i(t)$ menor que un cierto valor. k

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}$$

$$k_i(t) < k \Rightarrow m\left(\frac{t}{t_i}\right)^{\beta} < k \Rightarrow$$

$$\left(\frac{k}{m}\right)^{\frac{1}{\beta}} > \frac{t}{t_i} \Rightarrow$$

$$P(k_i(t) < k) = P\left(t_i > \frac{m^{1/\beta}}{k^{1/\beta}}t\right)$$

Si los nodos se agregan a intervalos iguales (pura formalidad) es uniforme

$$P(t_i) = \frac{1}{m_0 + t}$$

(pues los t_i son 1 para cada nodo agregado, los iniciales no tienen t_i)

entonces

$$P\left(t_i \leq \frac{m^{1/\beta}}{k^{1/\beta}}t\right) = \frac{m^{1/\beta}}{k^{1/\beta}}t\frac{1}{m_0 + t} = \int_0^k \cdots$$

como $P\left(t_i > \frac{m^{1/\beta}}{k^{1/\beta}}t\right) + P\left(t_i \leq \frac{m^{1/\beta}}{k^{1/\beta}}t\right) = 1$

entonces

$$P\left(t_i > \frac{m^{1/\beta}}{k^{1/\beta}}t\right) = 1 - \frac{m^{1/\beta}}{k^{1/\beta}}t\frac{1}{m_0 + t}$$

Pero ahora $P(k) = \frac{\partial}{\partial k}P(k_i(t) < k)$ (la densidad de proba es la derivada de la distribucion)

$$\frac{\partial}{\partial k}P\left(t_i > \frac{m^{1/\beta}}{k^{1/\beta}}t\right) = \frac{\partial}{\partial k}P\left(t_i > \frac{m^{1/\beta}}{k^{1/\beta}}t\right)$$

$$= \frac{\partial}{\partial k}\left(1 - \frac{m^{1/\beta}}{m_0 + t}t\frac{1}{k^{1/\beta}}\right)$$

$$= \frac{m^{1/\beta}}{m_0 + t}t\frac{1}{k^{(1/\beta)+1}}$$

Cuando $t \to \infty$ resulta

$$P(k) = \frac{m^{1/\beta}}{1}\frac{1}{k^{(1/\beta)+1}} = m^{1/\beta}k^{-\gamma}$$

Con

$$\gamma = \frac{1}{\beta} + 1 = 3$$

Power law con exponente fijo

Este es el estado estacionario que no depende de N ni de m_0.

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}	ℓ_{real}	ℓ_{rand}	ℓ_{pow}
WWW	325 729	4.51	900	2.45	2.1	11.2	8.32	4.77
WWW	4×10^7	7		2.38	2.1			
WWW	2×10^8	7.5	4000	2.72	2.1	16	8.85	7.61
WWW, site	260 000				1.94			
Internet, domain*	3015–4389	3.42–3.76	30–40	2.1–2.2	2.1–2.2	4	6.3	5.2
Internet, router*	3888	2.57	30	2.48	2.48	12.15	8.75	7.67
Internet, router*	150 000	2.66	60	2.4	2.4	11	12.8	7.47
Movie actors*	212 250	28.78	900	2.3	2.3	4.54	3.65	4.01
Co-authors, SPIRES*	56 627	173	1100	1.2	1.2	4	2.12	1.95
Co-authors, neuro.*	209 293	11.54	400	2.1	2.1	6	5.01	3.86
Co-authors, math.*	70 975	3.9	120	2.5	2.5	9.5	8.2	6.53
Sexual contacts*	2810			3.4	3.4			
Metabolic, *E. coli*	778	7.4	110	2.2	2.2	3.2	3.32	2.89
Protein, *S. cerev.**	1870	2.39		2.4	2.4			
Ythan estuary*	134	8.7	35	1.05	1.05	2.43	2.26	1.71
Silwood Park*	154	4.75	27	1.13	1.13	3.4	3.23	2
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			

Casos limite del modelo de Albert-Barabasi

i) El sistema crece pero sin preferential attachment

entonces

$$\Pi(k_i) \cong \frac{1}{m_0 + t - 1}$$

1 / (el numero de nodos)

Entonces

$$\frac{\partial k_i}{\partial t} = m\frac{1}{m_0 + t - 1} = \frac{m}{(m_0 - 1) + t}$$

Que es

$$\frac{1}{m}k_i = \ln((m_0 - 1) + t)$$

de donde

$$P(k) = \frac{e}{m}\exp(-k/m)$$

ii) El sistema no crece y tiene preferential attachment

Tiene un limite para T $\approx N^2$

$$k_i(t) \simeq \frac{2}{N} t$$

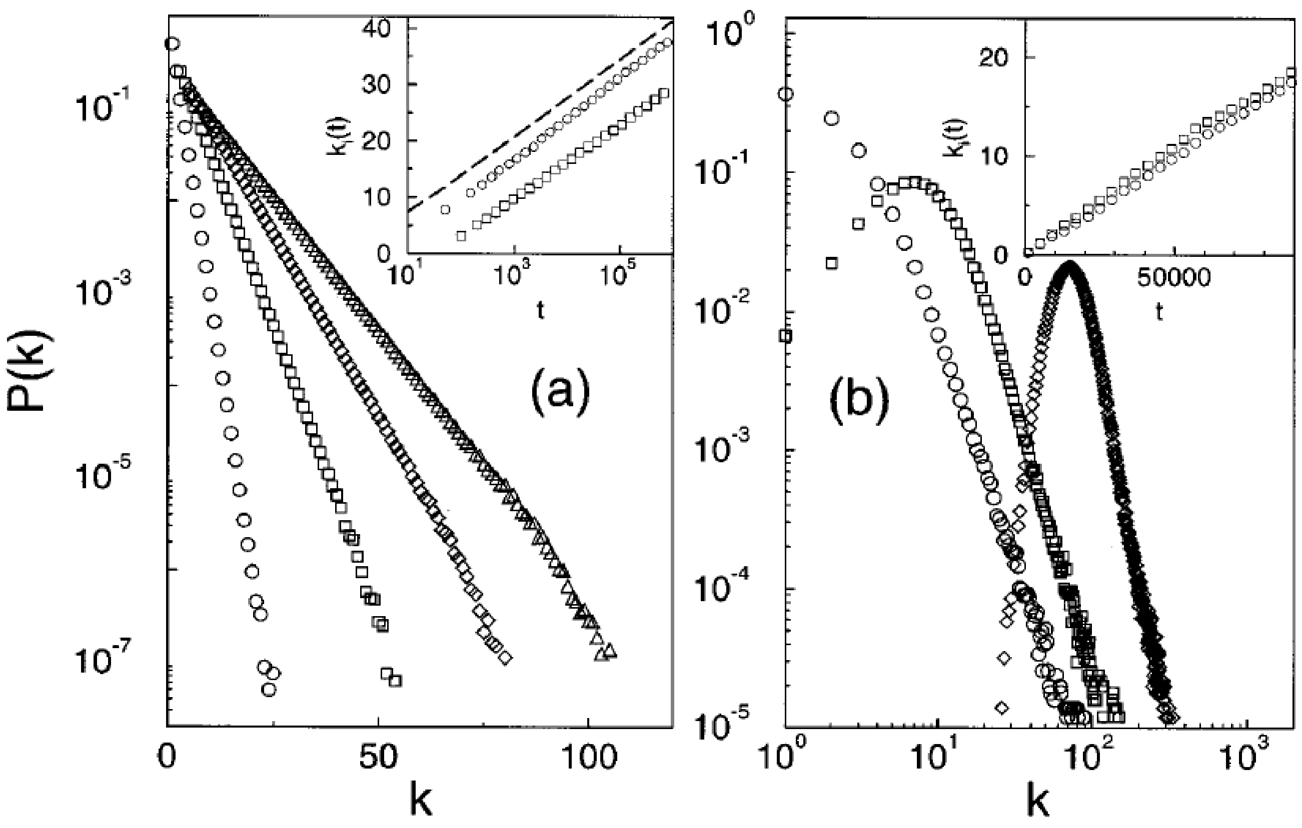


FIG. 22. Degree distribution for two models: (a) Degree distribution for model A: ○, $m_0 = m = 1$; □, $m_0 = m = 3$; ◇, $m_0 = m = 5$; △, $m_0 = m = 7$. The size of the network is $N = 800\,000$. Inset: time evolution for the degree of two vertices added to the system at $t_1 = 7$ and $t_2 = 97$. Here $m_0 = m = 3$. The dashed line follows $k_i(t) = m \ln(m_0 + t - 1)$; (b) the degree distribution for model B for $N = 10\,000$: ○, $t = N$; □, $t = 5N$; and ◇, $t = 40N$. Inset: time dependence of the degrees of two vertices. The system size is $N = 10\,000$. After Barabási, Albert, and Jeong (1999).

Propiedades del Modelo

Para Small World $\Rightarrow L \propto \ln(N)$

1) Camino minimo medic

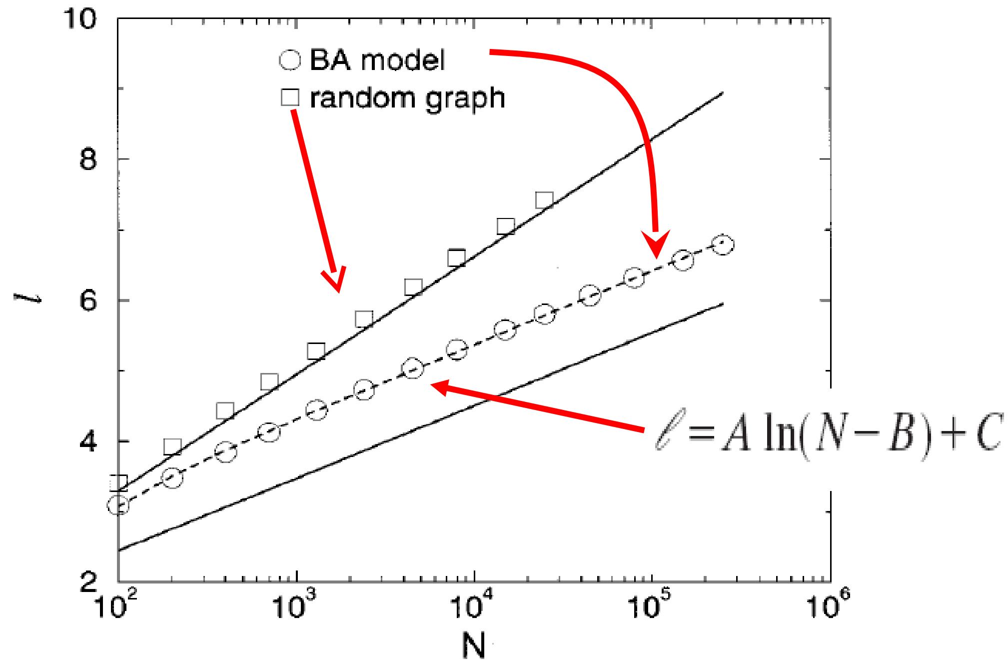


FIG. 23. Characteristic path length ℓ versus network size N in a Barabási-Albert (BA) network with $\langle k \rangle = 4$ (○), compared with a random graph of the same size and average degree generated with the algorithm described in Sec. III.A (□). The dashed line follows Eq. (94), and the solid lines represent Eq. (60) with $z_1 = \langle k \rangle$ and z_2 the numerically obtained number of next-nearest neighbors in the respective networks.

Recent analytical results indicate that there is a double logarithmic correction to the logarithmic N dependence, i.e., $\ell \sim \ln(N)/\ln\ln(N)$ (Bollobás and Riordan, 2001).

2) Clustering

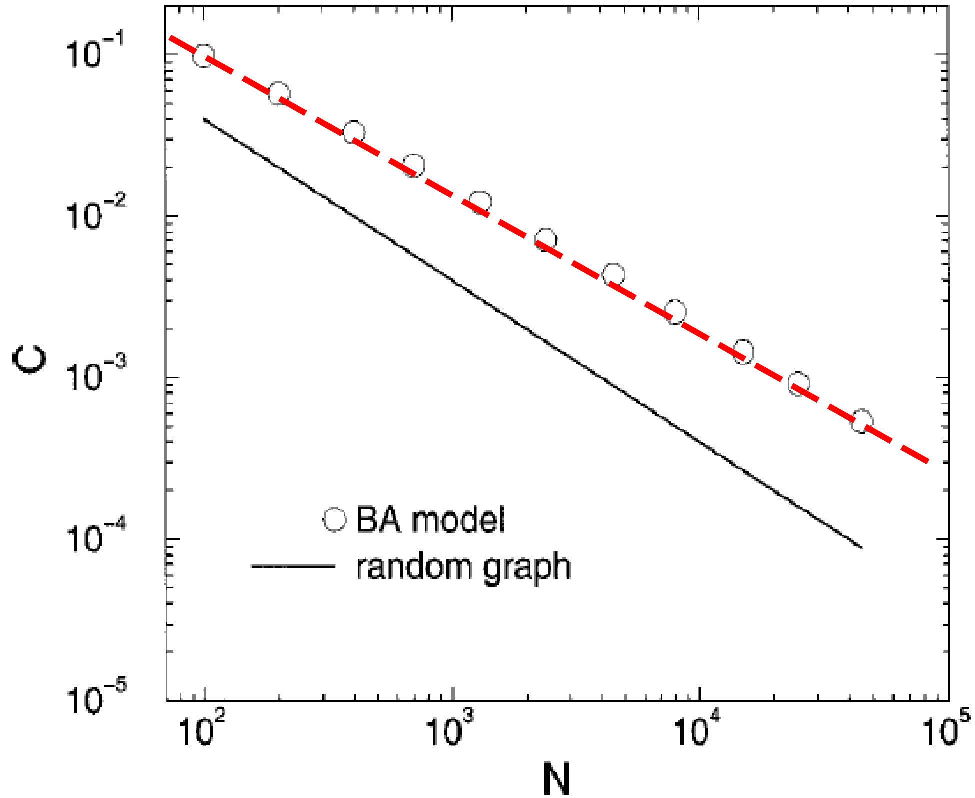


FIG. 24. Clustering coefficient versus size of the Barabási-Albert (BA) model with $\langle k \rangle = 4$, compared with the clustering coefficient of a random graph, $C_{rand} \simeq \langle k \rangle / N$.

Se encuentra que $C \simeq N^{-0.75}$, mientras que el random va como $C = \langle k \rangle N^{-1}$

O sea que es mejor que el random pero no lo suficiente!

Volvamos al Preferential Attachment

Estudiar esto en networks reales necesita de conocer el tiempo al cual el nodo fue agregado y con cuantos links

Esto se conoce para una variedad de networks, (citas, el de actores, internet, etc..)

La cuestion es entonces

a) contar los nodos presentes y sus links a tiempo t ("nodos viejos")

b) calcular el crecimiento del grado de los "nodos viejos" durante un intervalo de tiempo Δt

como

$$\Pi(k_i) = \frac{k_i}{\sum k_j}$$

Para calcular la distribucion se calcula $\Delta k_i / \Delta k$ en funcion de los "viejos" k_i (con Δk la variacion total en Δt con $\Delta t \ll t_{total}$ del grafo) y

se grafica para cada k_i y de alli la distribucion.

Como ya vimos puede ser mas apropiado calcular la cumulativa para reducir fluctuaciones.

Se obtiene:

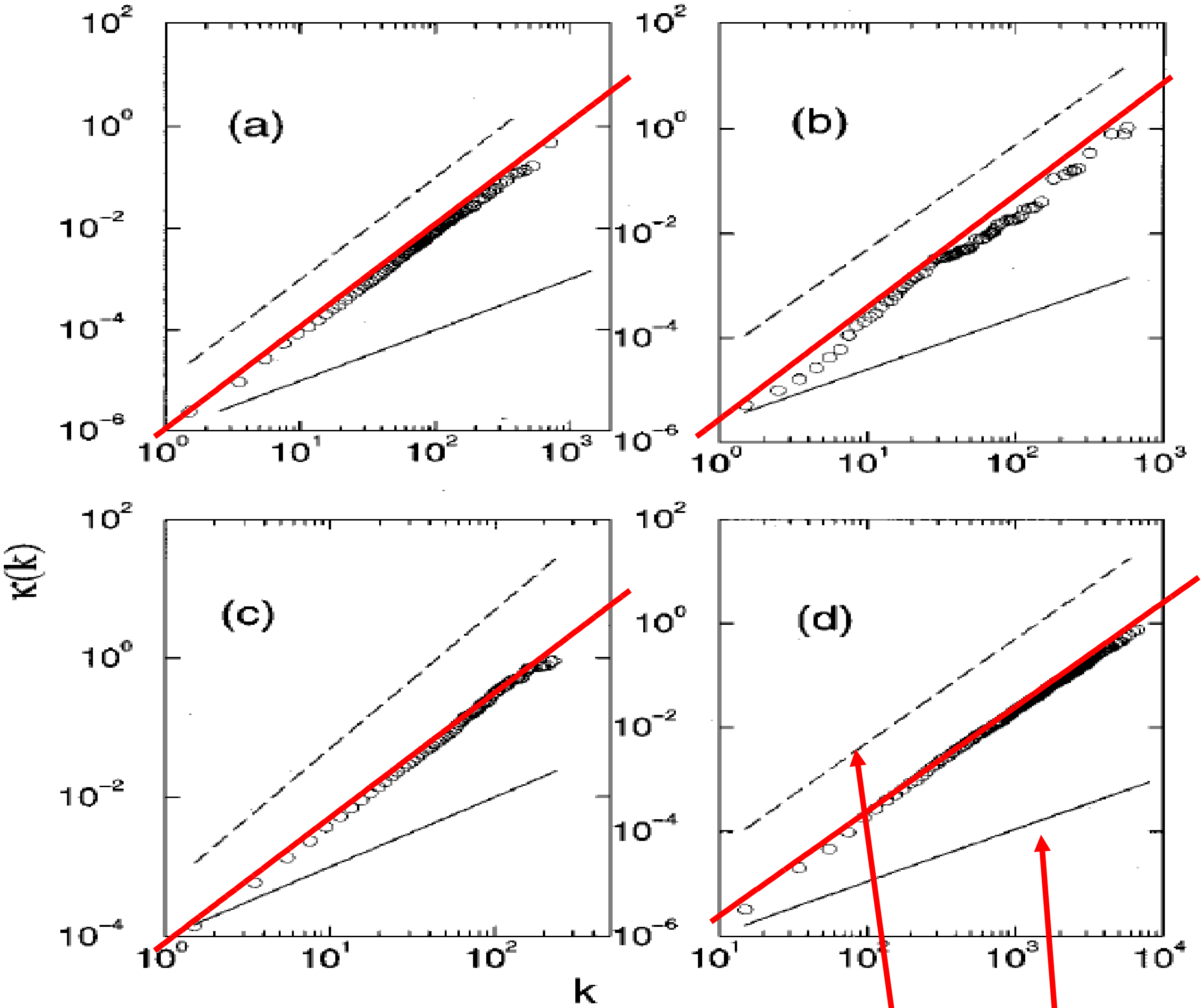


FIG. 26. Cumulative preferential attachment for (a) the citation network; (b) the Internet; (c) the neuroscience scientific collaboration network; (d) the actor collaboration network. In all panels the dashed line corresponds to linear preferential attachment, and the solid line to no preferential attachment. After Jeong, Néda, and Barabási (2001).

Multiples variaciones y efectos

1) nonlinear preferential attachments
2) otros modos de crecimiento
3) rewiring
4) edge removal
5) etc.