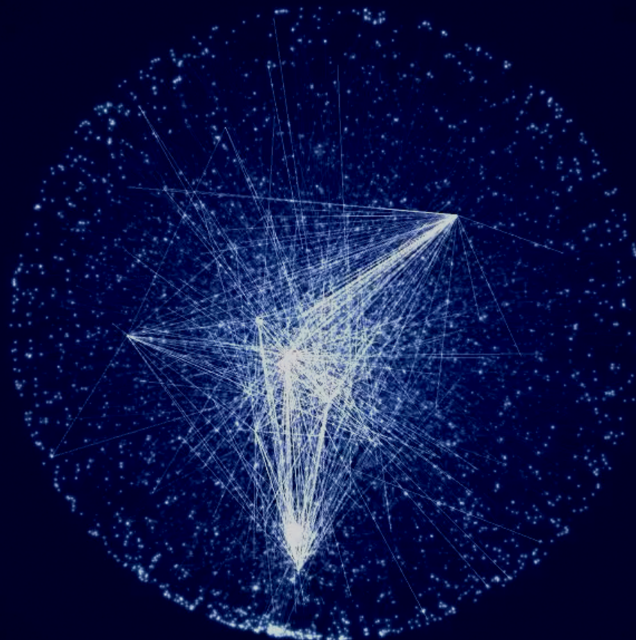
PARTIMOS EN BREVE

MUCHAS GRACIAS

El test comienza a las 11:55 en Canvas. Sección Evaluaciones

Network Science



Video: Protein-Protein Interactions
Credit: Mauro Martino

Dr. Cristian Candia

Director del Magister en Data Science y del Computational Research in Social Science Laboratory,
Instituto de Data Science, Centro de Inv. en Complejidad Social,
Facultades de Ingeniería y Gobierno
Universidad del Desarrollo, Chile

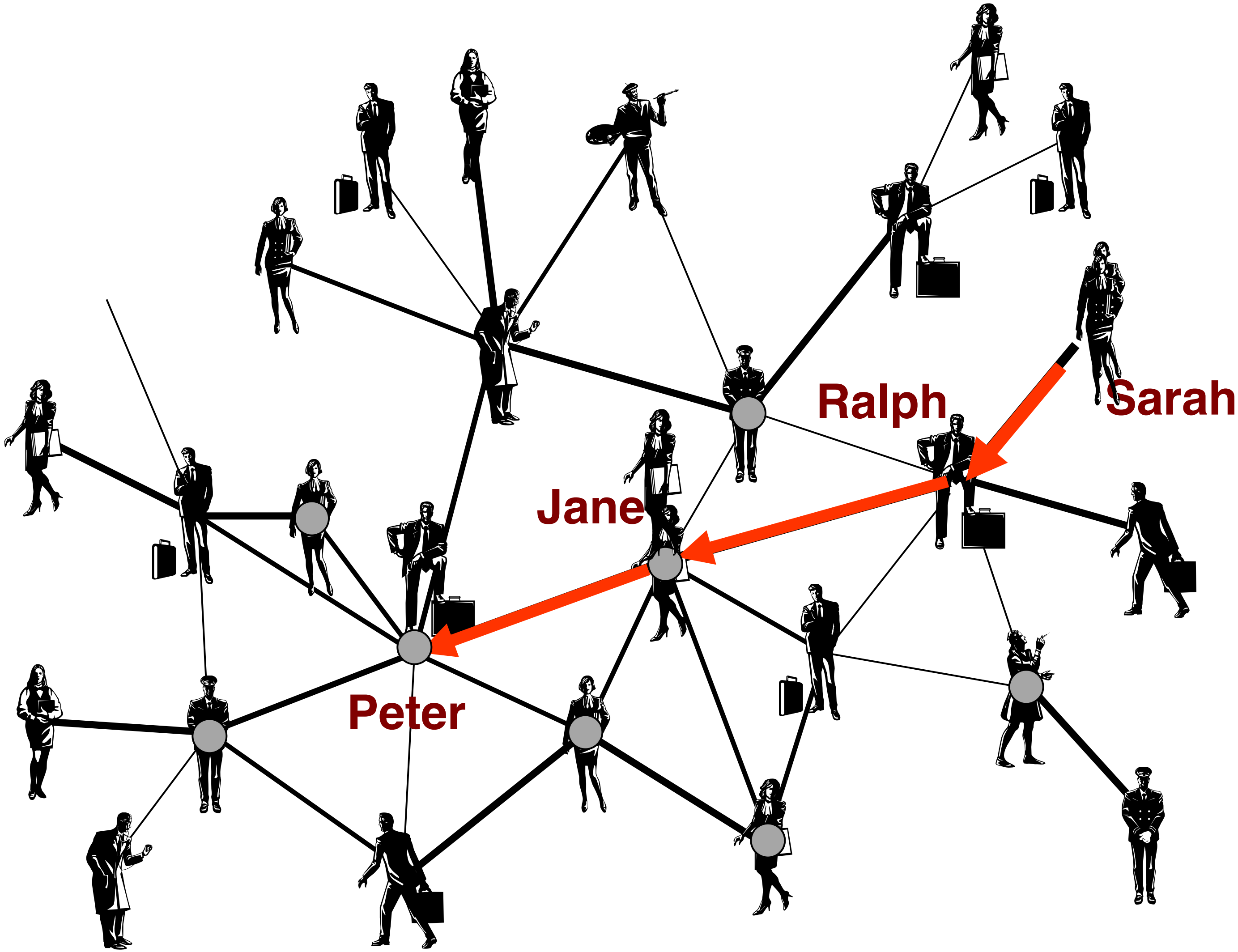
Académico Adjunto,
Northwestern University, United States

Modelos de Formación de Redes

Estas diapositivas se basan parcialmente en el curso del
Prof. Albert-László Barabási, de Northeastern University, con autorización.
El contenido ha sido traducido para su uso en este curso.

Mundo Pequeño

(Small World, modelo Watts-Strogatz)



Frigyes Karinthy, 1929
Stanley Milgram, 1967

1929: Frigyes Kartinthy

Frigyes Karinthy (1887-1938)
Hungarian Writer

1929: *Minden másképpen van* (Everything is Different)
Láncszemek (Chains)

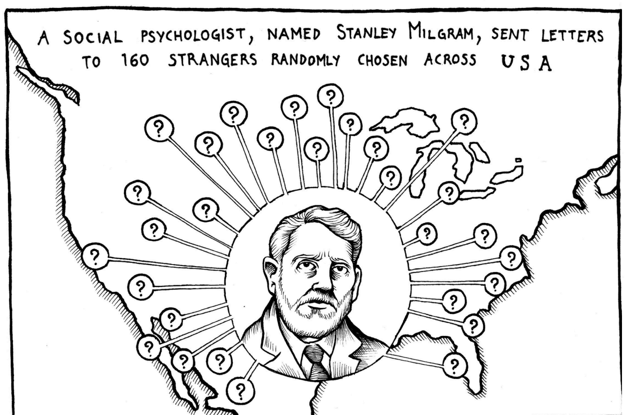
“Mire, Selma Lagerlöf acaba de ganar el Premio Nobel de Literatura, por lo que está obligada a conocer al Rey Gustav de Suecia, después de todo, él fue quien le entregó el Premio, como lo exige la tradición. El Rey Gustav, sin duda, es un apasionado jugador de tenis, que siempre participa en torneos internacionales. Se sabe que jugó con el Sr. Kehrling, a quien, por lo tanto, lo debe conocer con seguridad, y sucede que yo mismo conozco bastante bien al Sr. Kehrling ".

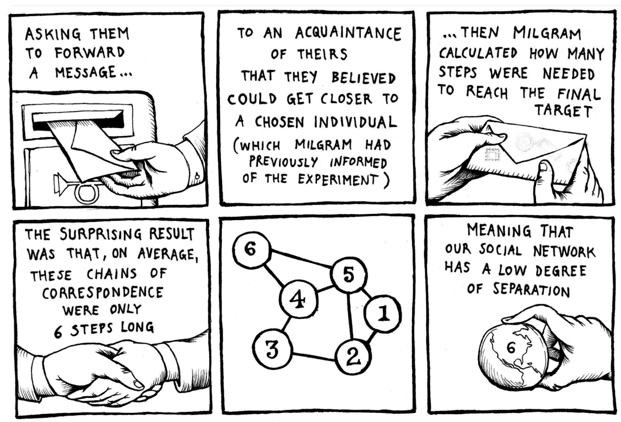
"El trabajador conoce al gerente de la tienda, quién conoce a Ford; Ford tiene una relación amistosa con el director general de Hearst Publications, que el año pasado se hizo muy amigo de Arpad Pasztor, alguien que no solo conozco, sino que es un buen amigo mío. Así que fácilmente podría pedirle que envíe un telegrama a través del director general para que le diga a Ford que debe hablar con el gerente para que le informe al trabajador de la tienda que arme un auto rápidamente para mí, ya que necesito uno."

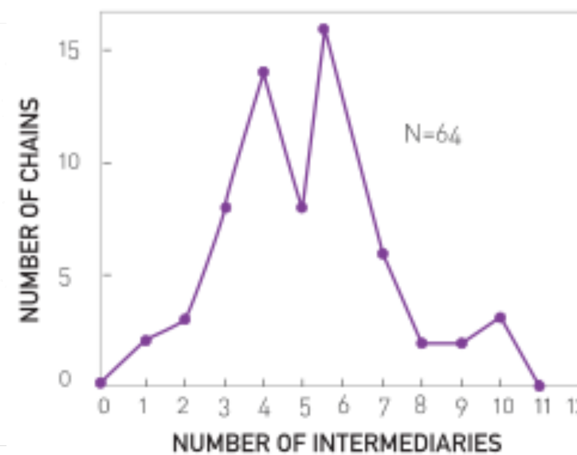
The Small World Experiment

—Se entregaron cartas a personas en Nebraska para que las enviaran a una persona objetivo en Boston.
—Se instruyó a las personas para que pasaran las cartas a alguien que conocían por su nombre de pila.
—Las cartas que llegaron a destino siguieron caminos de una longitud de alrededor de 6 enlaces
—Seis grados de separación, ¿Cómo puede ser eso?

1967

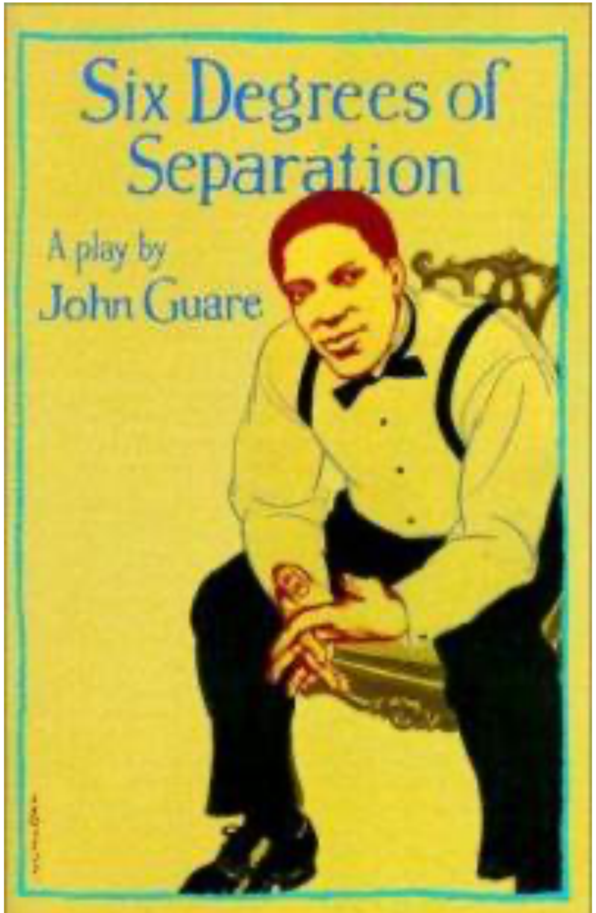






https://elpais.com/elpais/2016/12/15/ciencia/1481797356_294611.html#:~:text=En%201967%2C%20el%20psic%C3%B3logo%20estadounidense,veros%C3%ADmil%2C%20que%20a%20su%20vez

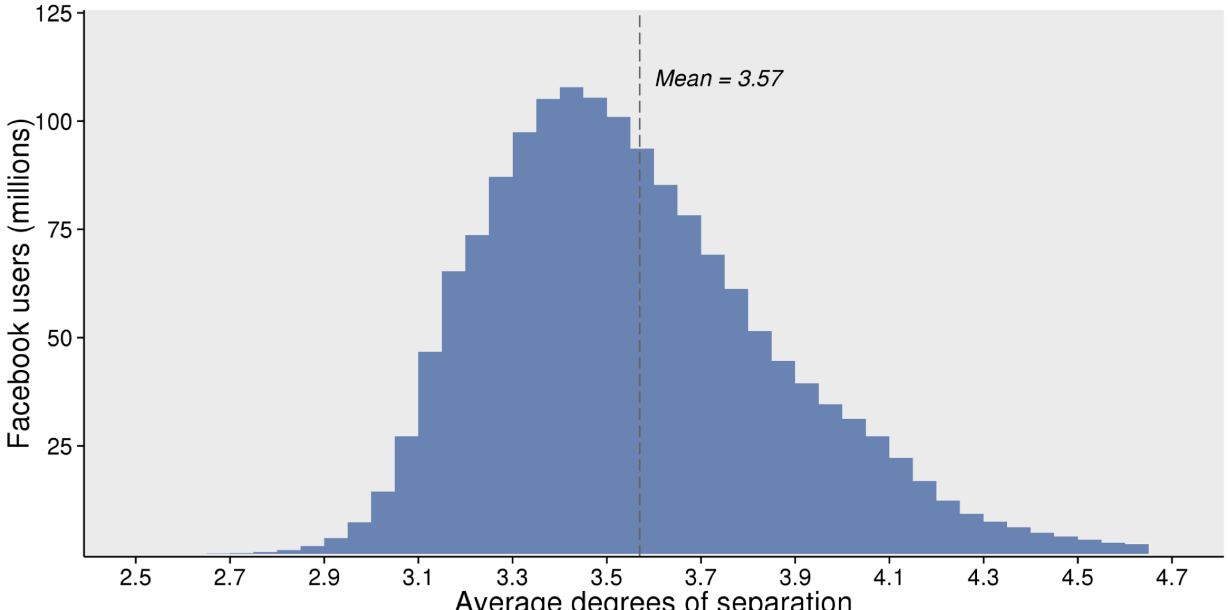
SEIS GRADOS | 1991: John Guare



"Todos en este planeta están separados solo por otras seis personas. Seis grados de separación. Entre nosotros y todos los demás en este planeta. El presidente de los Estados Unidos. Un gondolero en Venecia ... No son solo los grandes nombres. Es cualquiera. Un nativo en una selva tropical. Un Tierra del Fuegano. Un esquimal. Estoy atado a todos en este planeta por un rastro de seis personas. Es un pensamiento profundo. Cómo cada persona es una puerta nueva, abriéndose a otros mundos ".

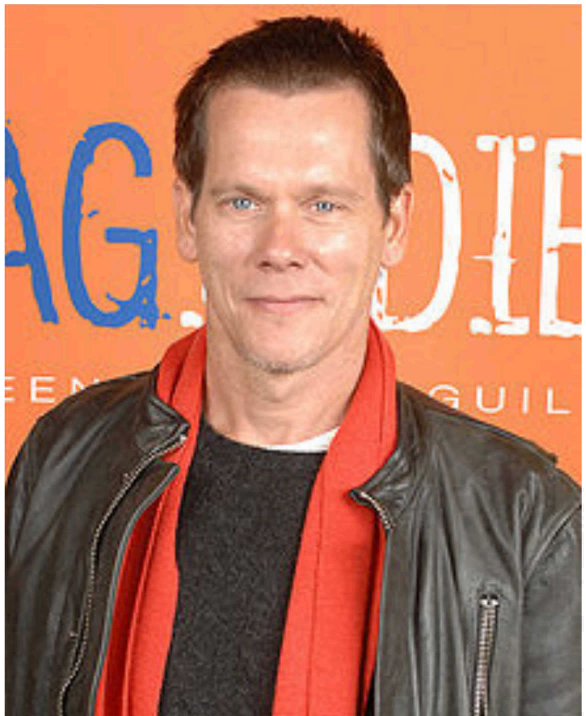
SEIS GRADOS | Experimento 2.0

—Informado en el blog de investigación de Facebook en 2016;
—La mayoría de las personas en Facebook tienen promedios de entre 2,9 y 4,2 grados de separación;
—La figura muestra los grados promedio estimados de separación entre todas las personas en Facebook. La persona promedio está conectada con todas las demás personas en un promedio de 3.57 pasos. La mayoría de personas tiene un promedio de entre 3 y 4 pasos;



https://research.fb.com/blog/2016/02/three-and-a-half-degrees-of-separation/

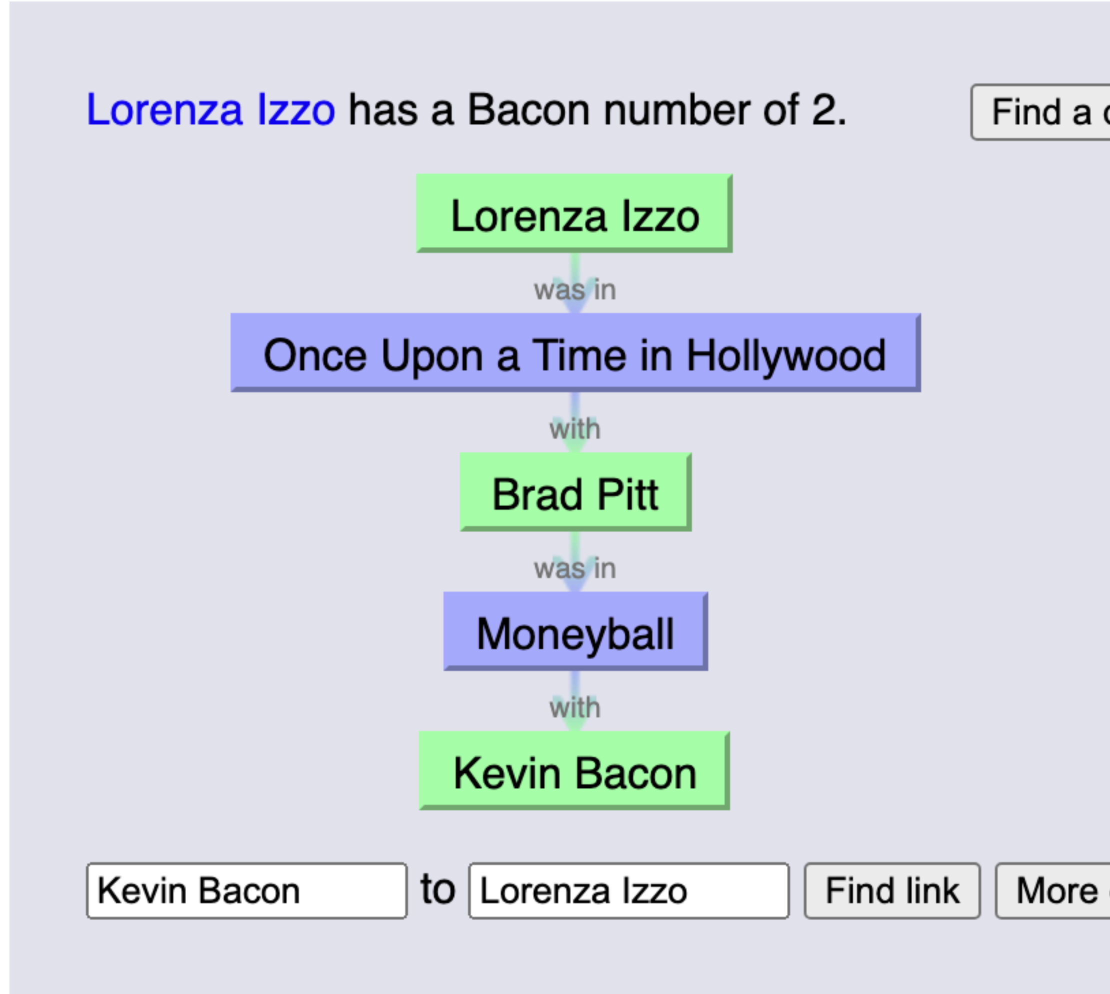
The Six Degrees of Kevin Bacon



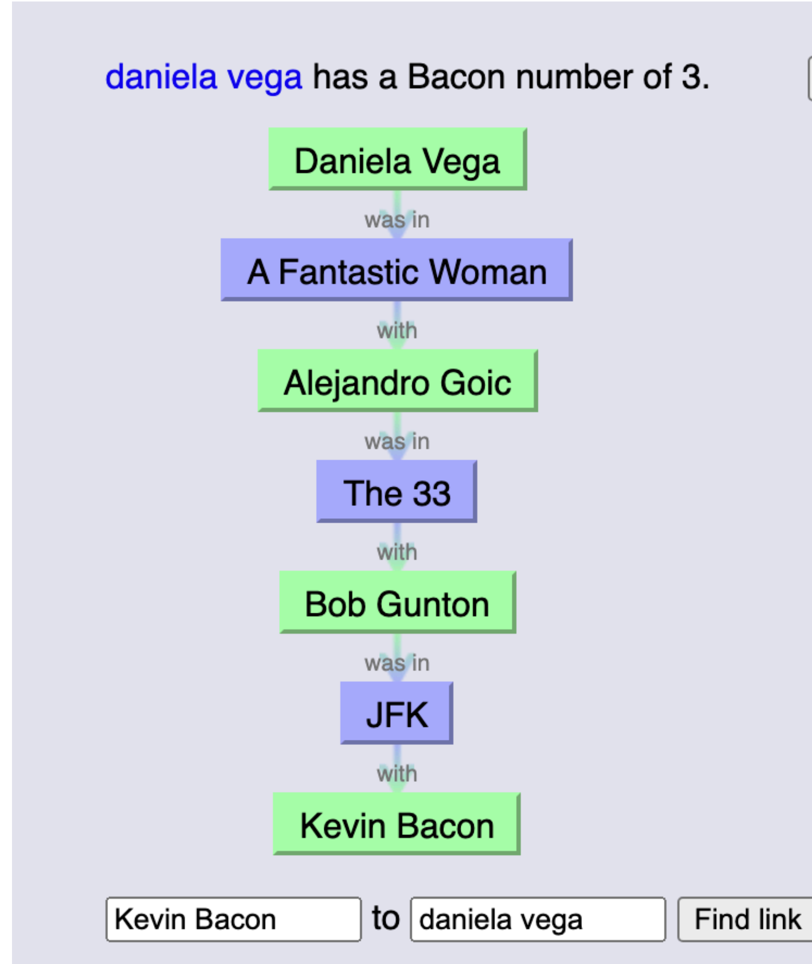
Paso 1. Nombra a un actor de cine

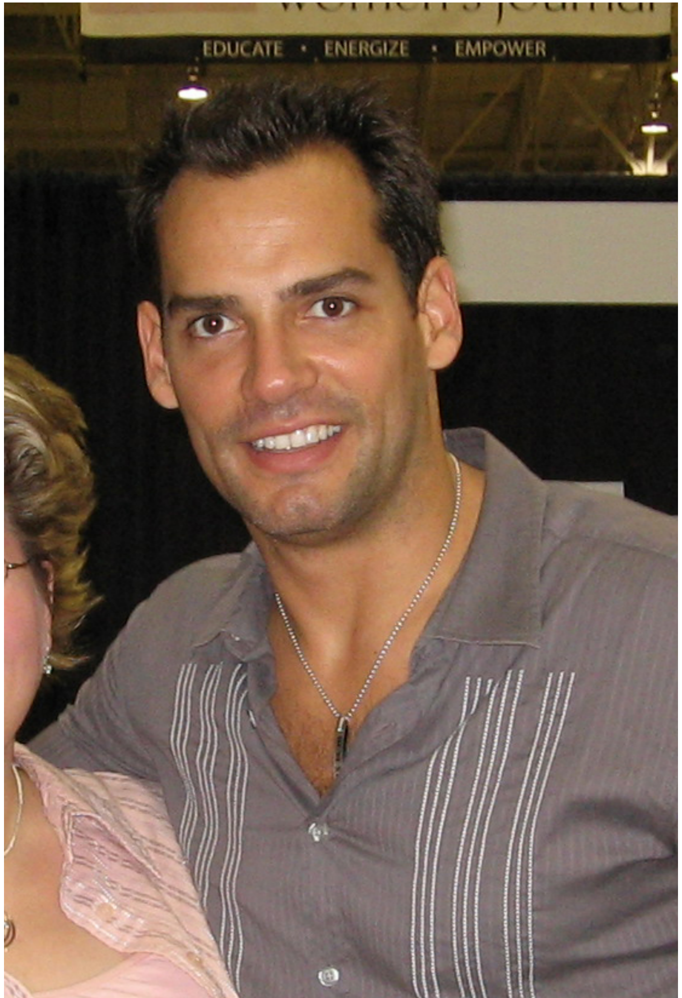
Paso 2. A través de los papeles en las películas, identifique una secuencia de actores que conecten al actor del Paso 1 con Kevin Bacon.

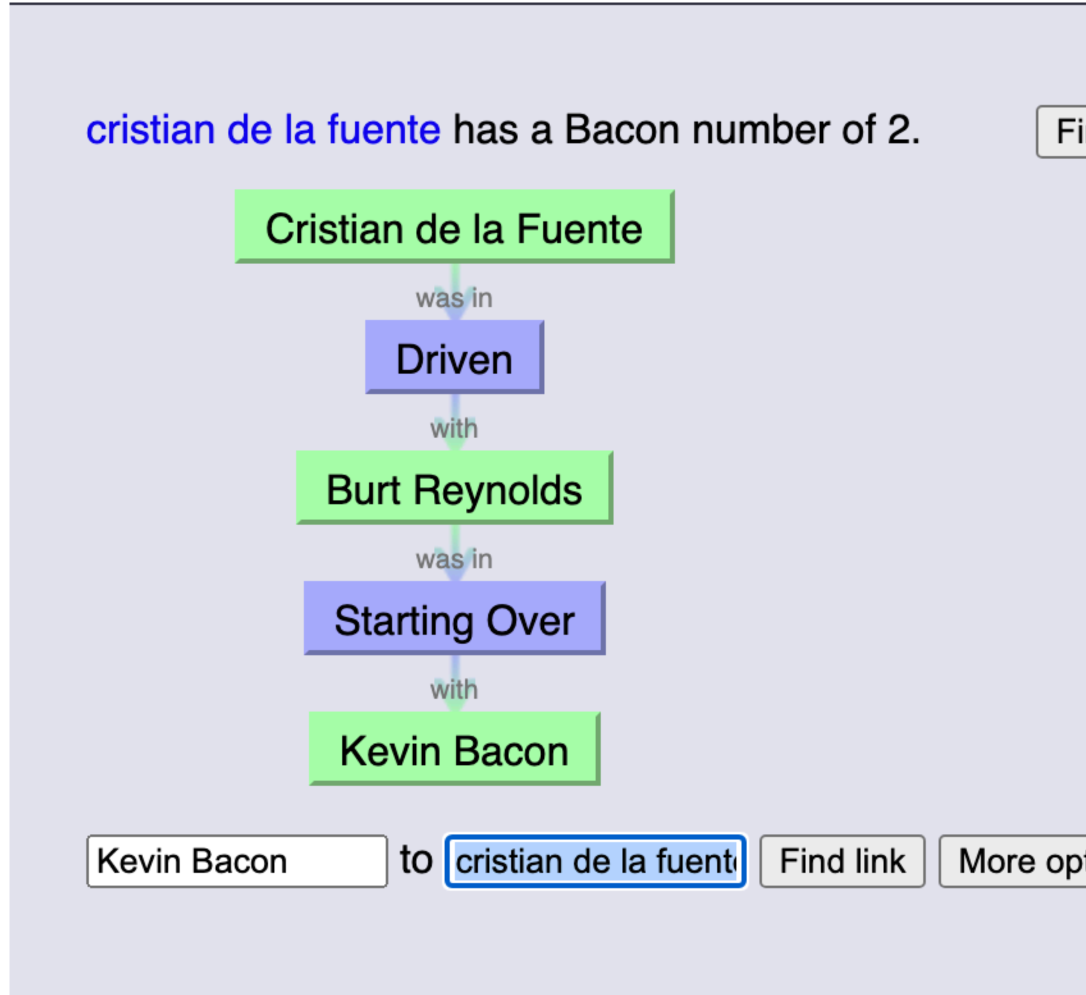
Paso 3. El ganador es la persona que encuentra el camino más corto

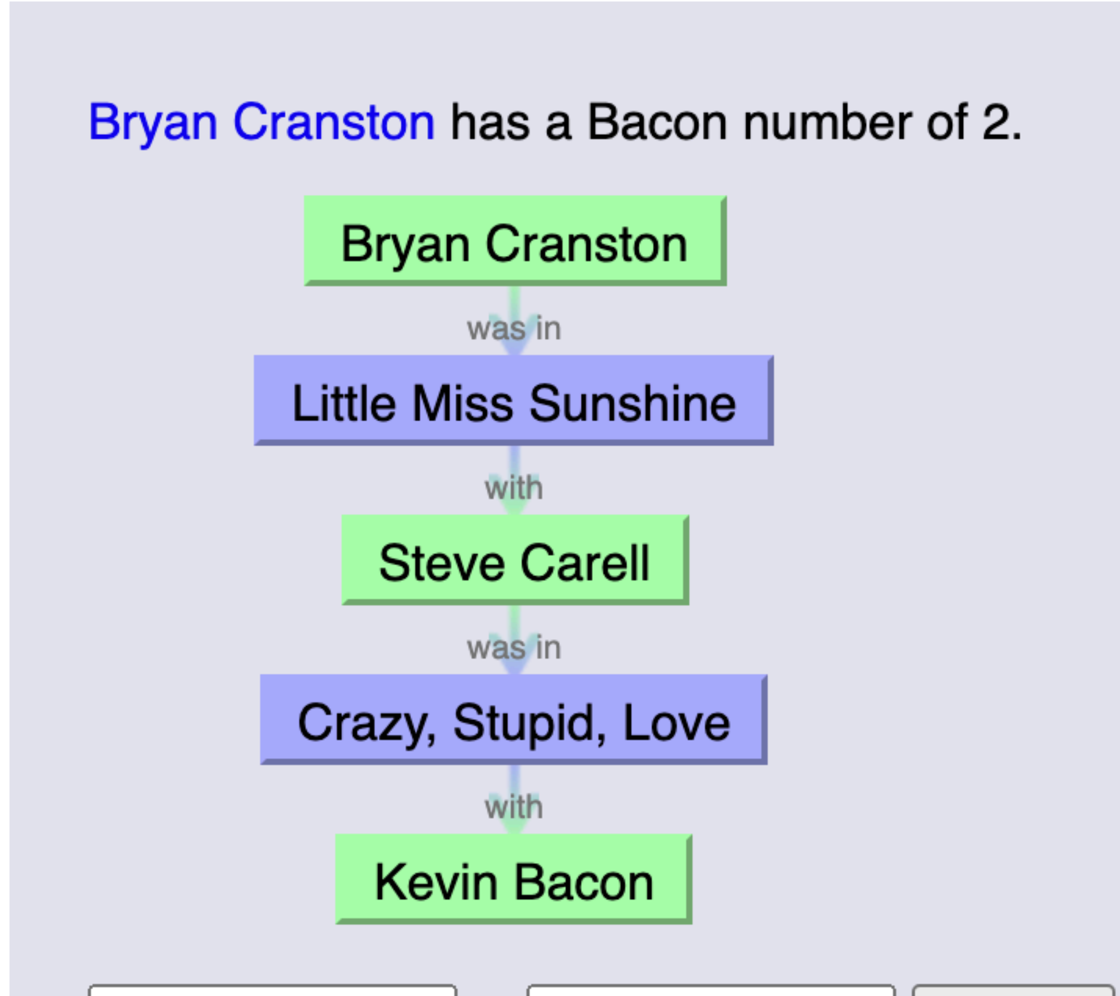


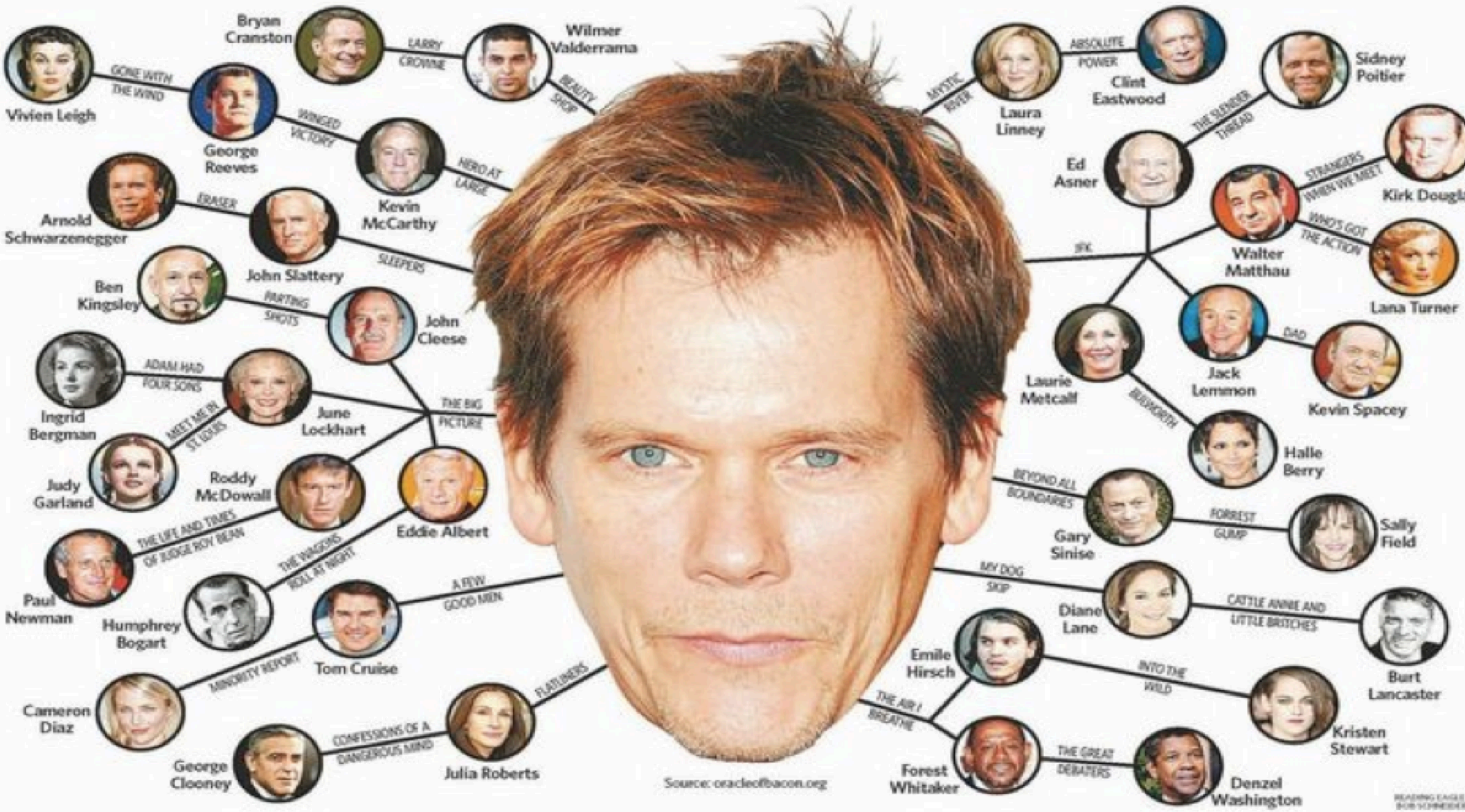
https://oracleofbacon.org


daniela vega has a Bacon number of 3.
Daniela Vega
was in
A Fantastic Woman
with
Alejandro Goic
was in
The 33
with
Bob Gunton
was in
JFK
with
Kevin Bacon
Kevin Bacon
to
daniela vega
Find link


EDUCATE • ENERGIZE • EMPOWER

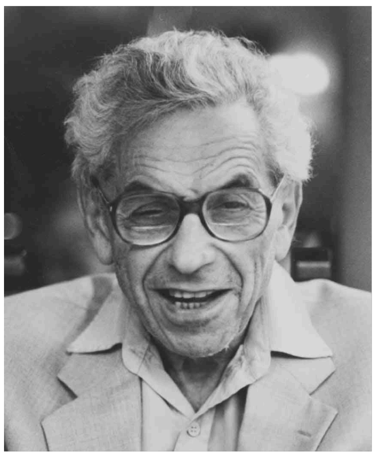

cristian de la fuente has a Bacon number of 2.
Cristian de la Fuente
was in
Driven
with
Burt Reynolds
was in
Starting Over
with
Kevin Bacon
Kevin Bacon
to
cristian de la fuent
Find link


Bryan Cranston has a Bacon number of 2.
Bryan Cranston
was in
Little Miss Sunshine
with
Steve Carell
was in
Crazy, Stupid, Love
with
Kevin Bacon

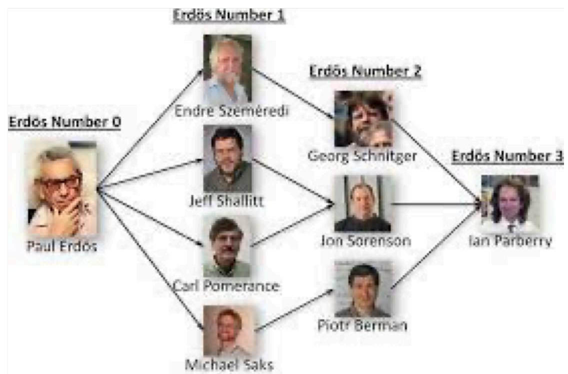


https://oracleofbacon.org

Erdos Number



Paul Erdős



Erdös number	0	---	1	person
Erdös number	1	---	504	people
Erdös number	2	---	6593	people
Erdös number	3	---	33605	people
Erdös number	4	---	83642	people
Erdös number	5	---	87760	people
Erdös number	6	---	40014	people
Erdös number	7	---	11591	people
Erdös number	8	---	3146	people
Erdös number	9	---	819	people
Erdös number	10	---	244	people
Erdös number	11	---	68	people
Erdös number	12	---	23	people
Erdös number	13	---	5	people

DATOS DE RED: REDES DE COLABORACIÓN DE CIENCIAS

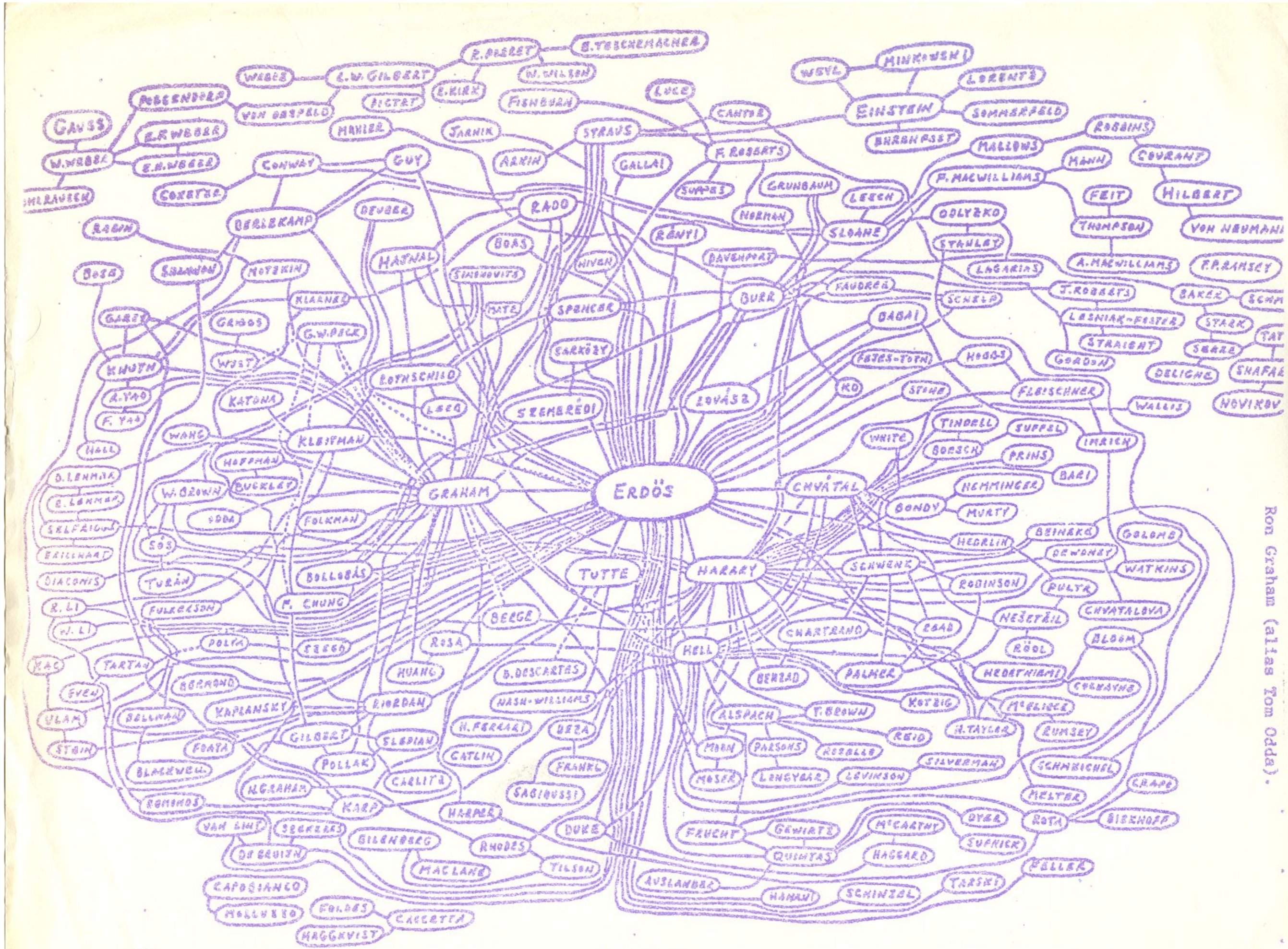


Figure 1
To appear in Topics in Graph Theory (F. Harary, ed.) New York Academy of Sciences (1979).

Erdos:
1,400 papers
507 coauthors

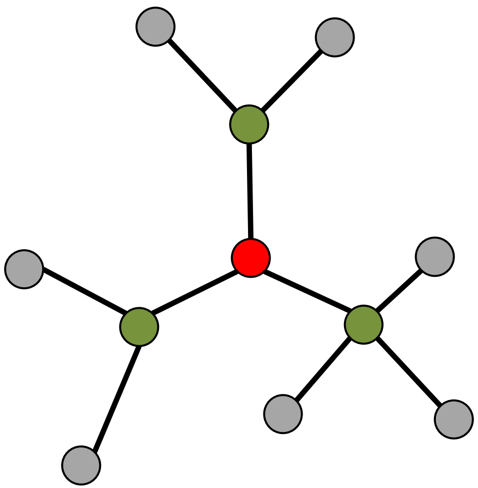
Einstein: EN=2
Paul Samuelson EN=5

….

ALB: EN= 3
CCV: EN=4

DISTANCIA EN GRAFOS ALEATORIOS

Los grafos aleatorios tienden a tener una topología en forma de árbol con grados de nodo casi constantes.



$\langle k \rangle$ nodes at distance one (d=1).

$\langle k \rangle^2$ nodes at distance two (d=2).

$\langle k \rangle^3$ nodes at distance three (d =3).

...

$\langle k \rangle^d$ nodes at distance d.

Podemos escribir la cantidad total de nodos de la red como:

$$N = 1 + \langle k \rangle + \langle k \rangle^2 + \ldots + \langle k \rangle^{d_{\max}} = \frac{\langle k \rangle^{d_{\max}+1} - 1}{\langle k \rangle - 1} \approx \langle k \rangle^{d_{\max}} \quad \Rightarrow \quad d_{\max} = \frac{\log N}{\log \langle k \rangle}$$

DISTANCIA EN GRAFOS ALEATORIOS

$$d_{\max} = \frac{\log N}{\log\langle k\rangle}$$

En la mayoría de las redes, esto ofrece una mejor aproximación a la distancia promedio entre dos nodos elegidos al azar, ⟨d⟩, que al diametro d_{max} .

$$< d >= \frac{\log N}{\log\langle k\rangle}$$

Llamaremos a este fenómeno de mundo pequeño la propiedad de que la **longitud promedio de la trayectoria o el diámetro depende logarítmicamente del tamaño del sistema**. Por lo tanto, "pequeño" significa que ⟨d⟩ es proporcional al logaritmo de N, en lugar de a N.

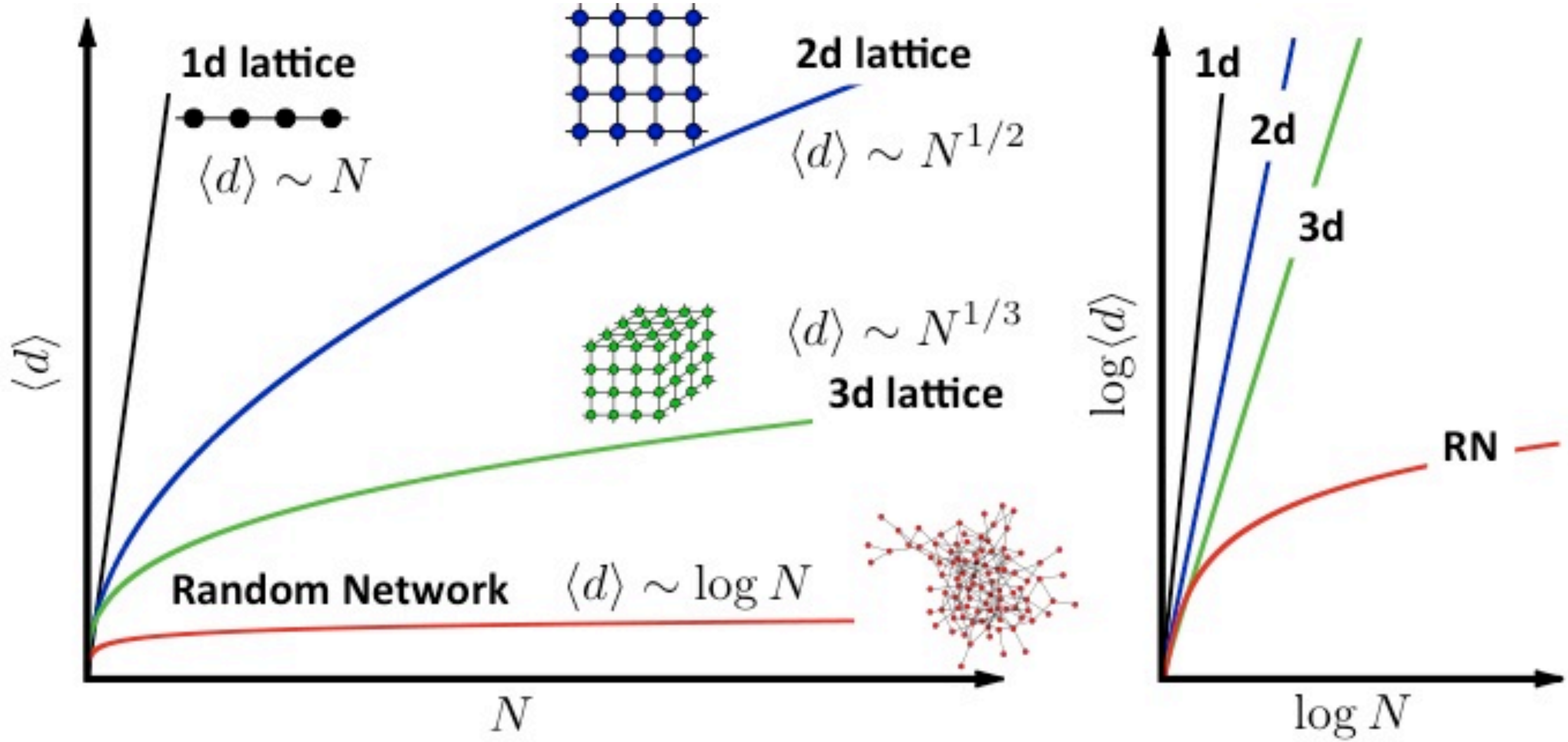
El término 1/log⟨k⟩ implica cuanto más densa sea la red, menor será la distancia entre los nodos.

DISTANCIAS EN GRÁFOS ALEATORIOS | comparados con datos reales

Network	N	L	‹k›	‹d›	d_{max}	ln*N*/ln‹*k*›
Internet	192,244	609,066	6.34	6.98	26	6.58
WWW	325,729	1,497,134	4.60	11.27	93	8.31
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile-Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	93,437	8.08	5.35	15	4.81
Actor Network	702,388	29,397,908	83.71	3.91	14	3.04
Citation Network	449,673	4,707,958	10.43	11.21	42	5.55
E. Coli Metabolism	1,039	5,802	5.58	2.98	8	4.04
Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Dadas las enormes diferencias en diámetro, tamaño y grado promedio, la aproximación es excelente.

¿Por qué sorprenden los pequeños mundos? Sorprendente en comparación con qué?



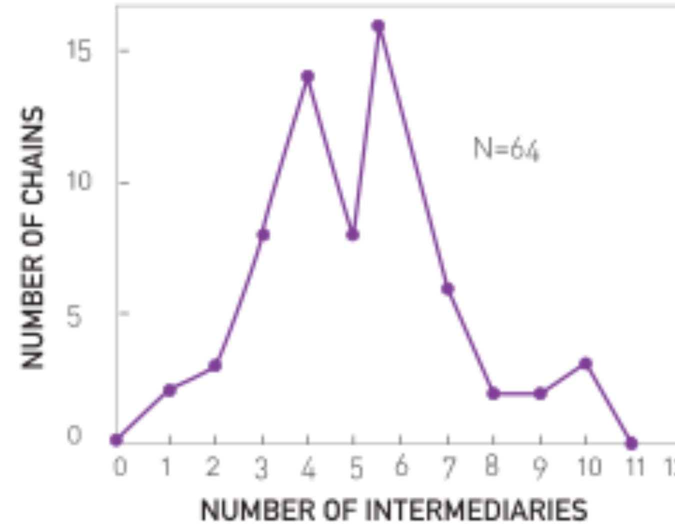
Tres, cuatro o seis grados?

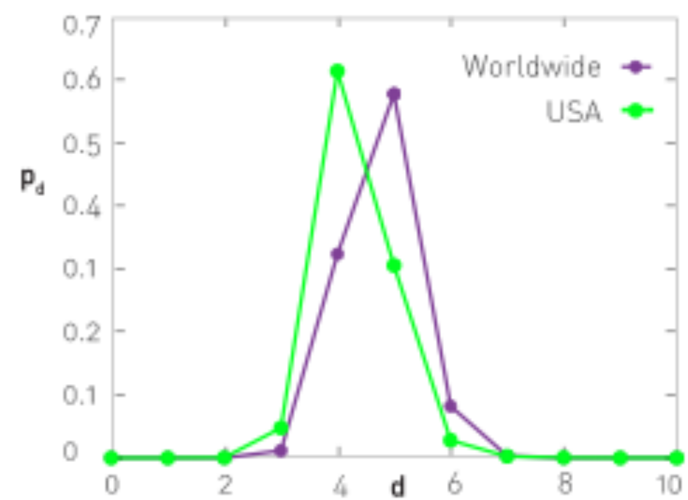
Para las redes sociales del mundo:

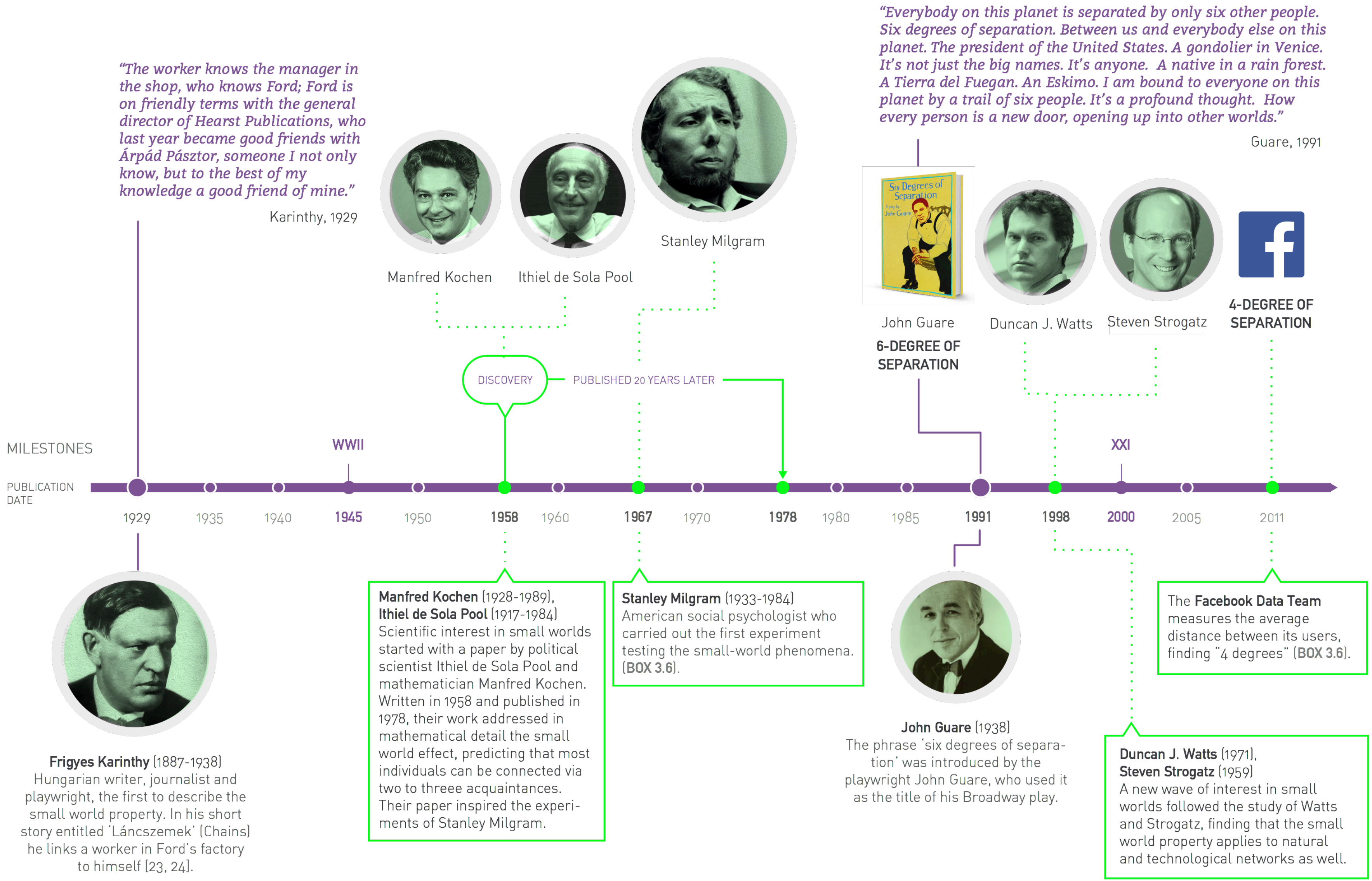
$\langle k \rangle \simeq 10^3$

$N \simeq 7 \times 10^9$ para la población mundial.

$$< d > = \frac{\ln(N)}{\ln\langle k \rangle} = 3.28$$





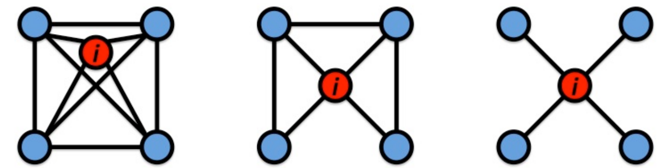


"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but to the best of my knowledge a good friend of mine."
Karinthy, 1929
Manfred Kochen
Ithiel de Sola Pool
Stanley Milgram
"Everybody on this planet is separated by only six other people. Six degrees of separation. Between us and everybody else on this planet. The president of the United States. A gondolier in Venice. It's not just the big names. It's anyone. A native in a rain forest. A Tierra del Fuegan. An Eskimo. I am bound to everyone on this planet by a trail of six people. It's a profound thought. How every person is a new door, opening up into other worlds."
Guare, 1991
Six Degrees of Separation
John Guare
6-DEGREE OF SEPARATION
Duncan J. Watts
Steven Strogatz
4-DEGREE OF SEPARATION
DISCOVERY
PUBLISHED 20 YEARS LATER
MILESTONES
WWII
XXI
PUBLICATION DATE
1929
1935
1940
1945
1950
1958
1960
1967
1970
1978
1980
1985
1991
1998
2000
2005
2011
Frigyes Karinthy (1887-1938)
Hungarian writer, journalist and playwright, the first to describe the small world property. In his short story entitled 'Láncszemek' (Chains) he links a worker in Ford's factory to himself [23, 24].
Manfred Kochen (1928-1989), **Ithiel de Sola Pool** (1917-1984)
Scientific interest in small worlds started with a paper by political scientist Ithiel de Sola Pool and mathematician Manfred Kochen. Written in 1958 and published in 1978, their work addressed in mathematical detail the small world effect, predicting that most individuals can be connected via two to threee acquaintances. Their paper inspired the experiments of Stanley Milgram.
Stanley Milgram (1933-1984)
American social psychologist who carried out the first experiment testing the small-world phenomena. (**BOX 3.6**).
John Guare (1938)
The phrase 'six degrees of separation' was introduced by the playwright John Guare, who used it as the title of his Broadway play.
Duncan J. Watts (1971), **Steven Strogatz** (1959)
A new wave of interest in small worlds followed the study of Watts and Strogatz, finding that the small world property applies to natural and technological networks as well.
The **Facebook Data Team** measures the average distance between its users, finding "4 degrees" (**BOX 3.6**).

Coeficiente de Agrupamiento (Clustering)

COEFICIENTE DE AGRUPAMIENTO

$$C_i \equiv \frac{2 < L_i >}{k_i(k_i - 1)}$$



Dado que los enlaces son independientes y tienen la misma probabilidad p,

$$< L_i > \cong p\frac{k_i(k_i - 1)}{2} \Rightarrow C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$

k_i=número de vecinos del nodo i

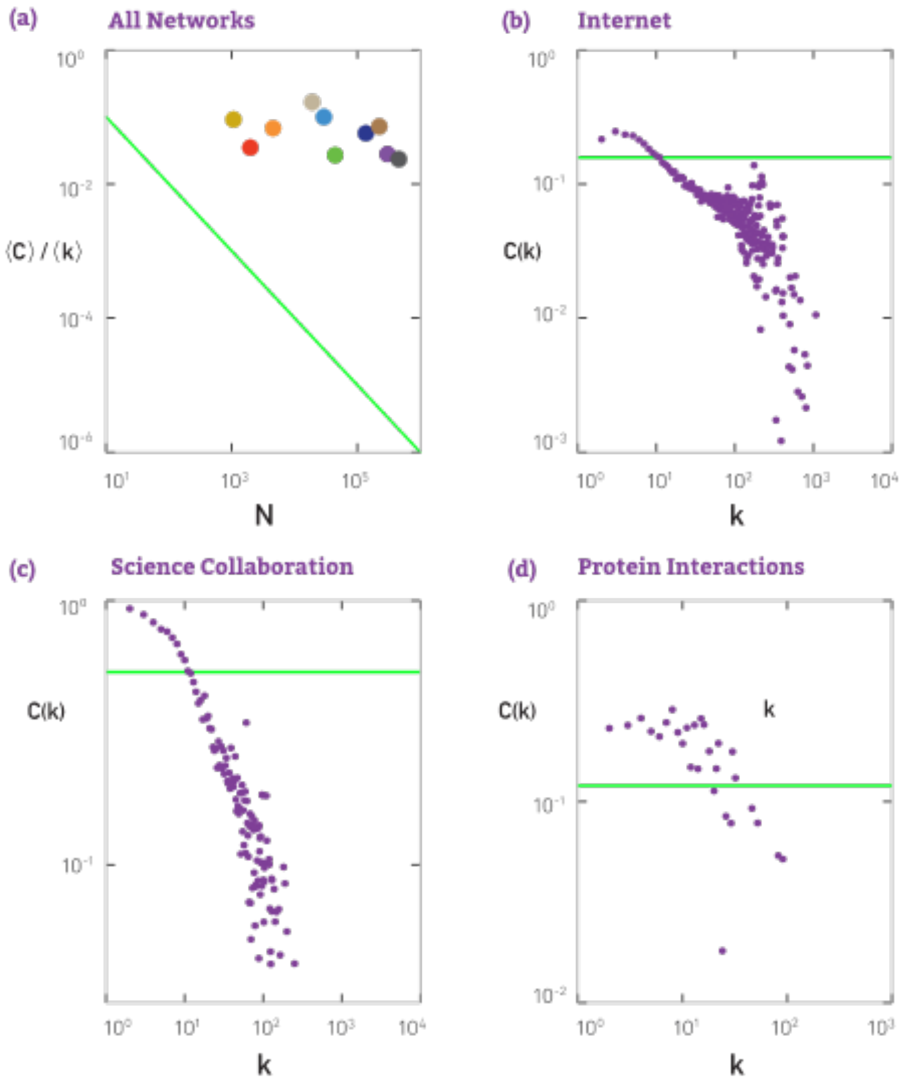
L_i=Número de enlaces que pasan por el nodo I

•El coeficiente de agrupamiento de los grafos aleatorios es pequeño.

•Para grados fijos, C disminuye con el tamaño del sistema N.

•C es independiente del grado k de nodos individuales.

COEFICIENTE DE AGRUPAMIENTO



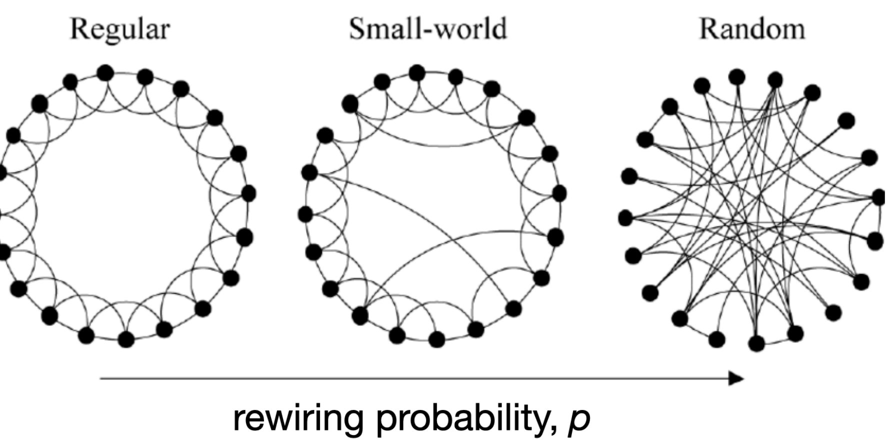
$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$

C disminuye con el tamaño del sistema N.

C es independiente del grado k de los nodos individuales.

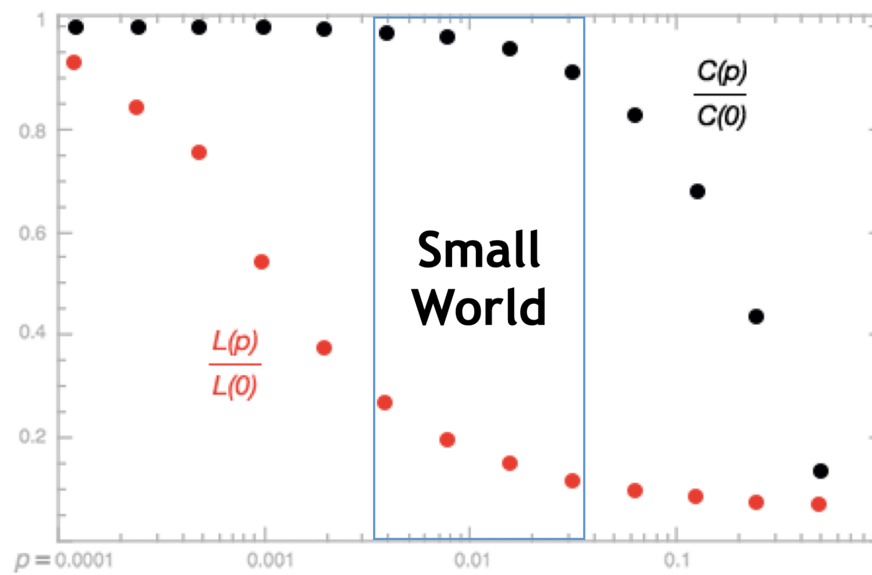
Small World Networks

Duncan J Watts Steven Strogatz



Watts & Strogatz algorithm

A partir de una red de anillo regular, destruya cada enlace con probabilidad p, creando un nuevo enlace a un nodo aleatorio.



Las redes reales no son aleatorias.

¿Las redes reales son como los grafos aleatorios?

A medida que se dispone de datos cuantitativos sobre redes reales, podemos compara su topología con las predicciones de la teoría de grafos aleatorios.

Tenga en cuenta que una vez que tenemos N y <k> para una red aleatoria, de ella podemos derivar cada propiedad medible. De hecho, tenemos:

Longitud media del camino:

$$< l_{rand} > \approx \frac{\log N}{\log \langle k \rangle}$$

Coeficiente de agrupamiento:

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$

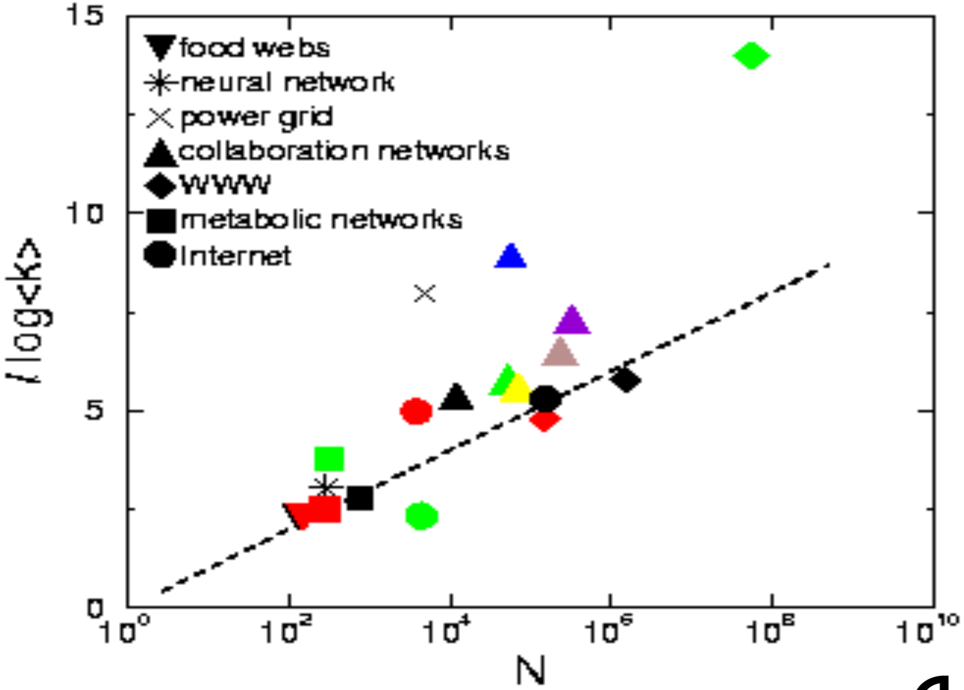
Distribución de Grados:

$$P(k) = e^{-<k>} \frac{< k >^k}{k!}$$

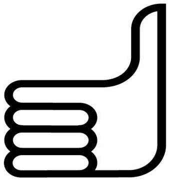
LONGITUD DE CAMINOS EN REDES REALES

Predicción:

$$< d >= \frac{\log N}{\log\langle k\rangle}$$



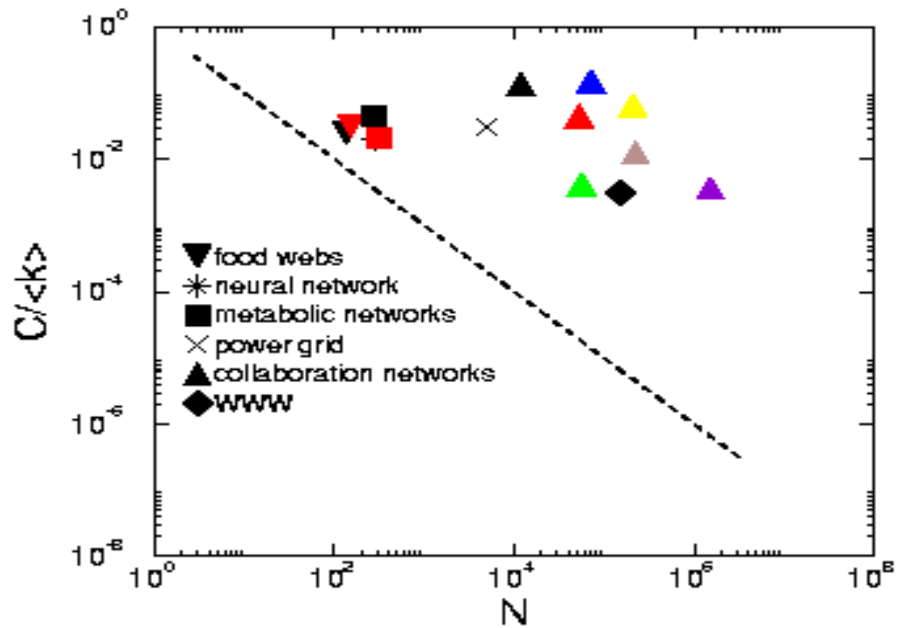
Las redes reales tienen distancias cortas como los grafos aleatorios.



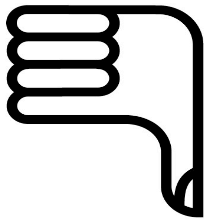
COEFICIENTE DE CLUSTERING

Predicción:

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$



C_{rand} Subestima en órdenes de magnitudes el coeficiente de agrupamiento de las redes reales.



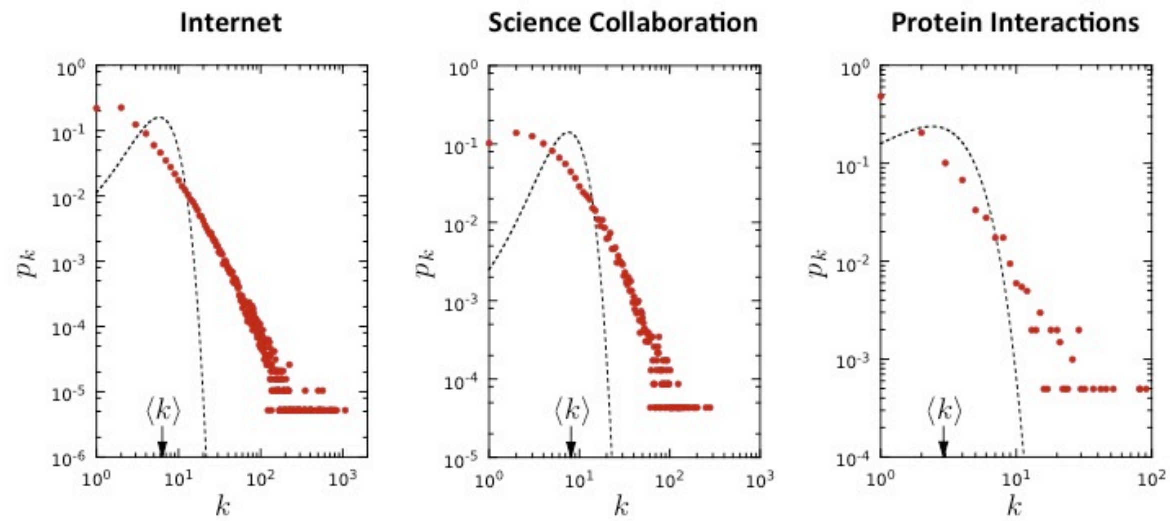
LA DISTRIBUCIÓN DE GRADO

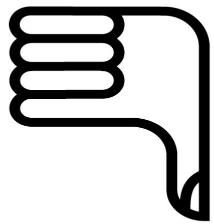
Predicción:

$$P(k) = e^{-<k>} \frac{<k>^k}{k!}$$

Data:

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





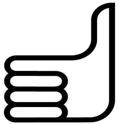
¿Las redes reales son como los grafos aleatorios?

A medida que se dispone de datos cuantitativos sobre redes reales, podemos compara su topología con las predicciones de la teoría de grafos aleatorios.

Tenga en cuenta que una vez que tenemos N y <k> para una red aleatoria, de ella podemos derivar cada propiedad medible. De hecho, tenemos:

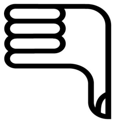
<u>Longitud media del camino:</u>

$$< l_{rand} > \approx \frac{\log N}{\log \langle k \rangle}$$



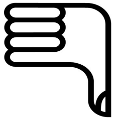
<u>Coeficiente de agrupamiento:</u>

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i - 1)} = p = \frac{\langle k \rangle}{N}.$$



<u>Distribución de Grados:</u>

$$P(k) = e^{-<k>} \frac{< k >^k}{k!}$$



¿ES EL MODELO ALEATORIO RELEVANTE PARA LOS SISTEMAS REALES?

(B) Lo más importante: debemos preguntarnos, ¿son las redes reales aleatorias?

La respuesta simplemente es no

No hay ninguna red en la naturaleza que sepamos que sería descrita por el modelo de red aleatorio.

SI ES INCORRECTO E IRRELEVANTE, ¿POR QUÉ DEDICAMOS A UNA CLASE COMPLETA?

- **Es el modelo de referencia para el resto del curso**

- Nos ayudará a calcular muchas cantidades, que luego se pueden comparar con los datos reales, entendiendo hasta qué punto es una propiedad particular el resultado de algún proceso aleatorio.

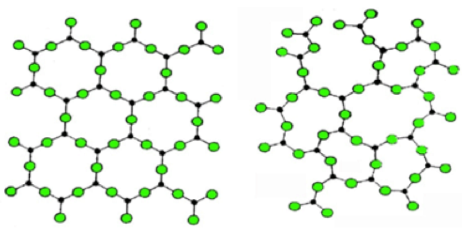
- Patrones en redes reales, que son compartidos por una gran cantidad de redes reales, pero que se desvían de las predicciones del modelo de red aleatorio.

- Para identificarlos, debemos entender cómo se vería una propiedad en particular si está totalmente controlada por procesos aleatorios.

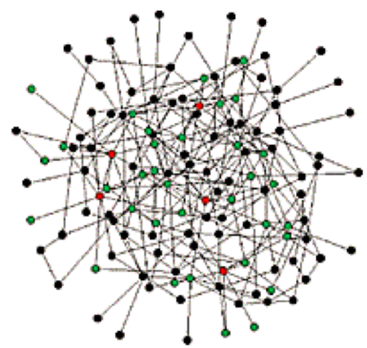
Si bien es EQUIVOCADO e IRRELEVANTE, ¡resultará extremadamente ÚTIL!

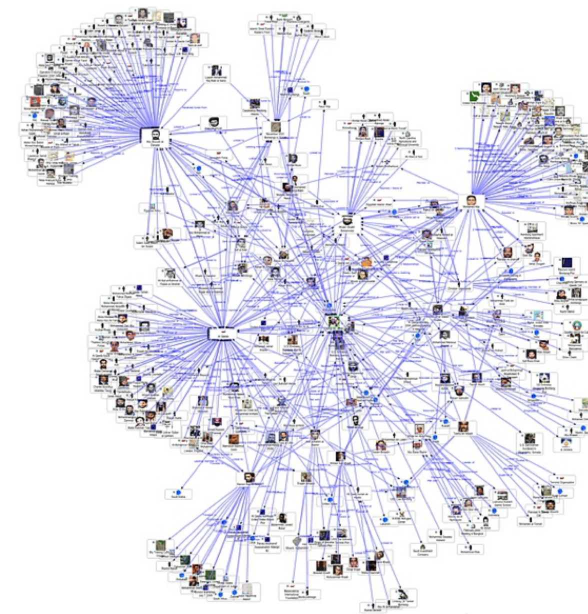
Scale-free Networks

Desde sistemas físicos a sociales



Crystalline silica   Vitreous silica



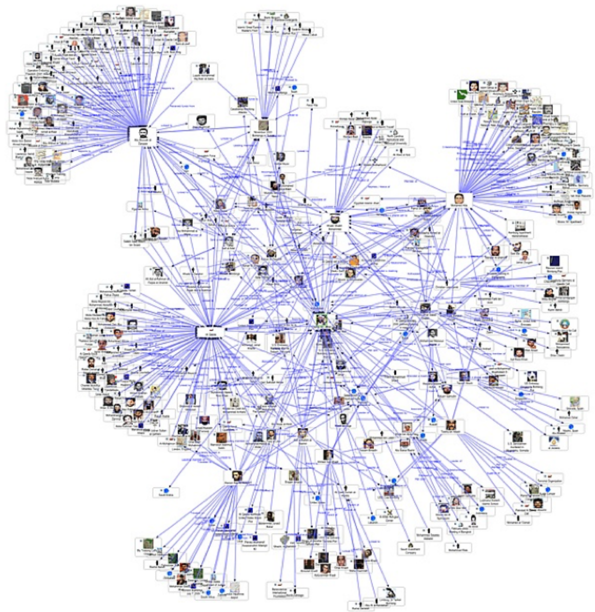


increasing network heterogeneity/complexity

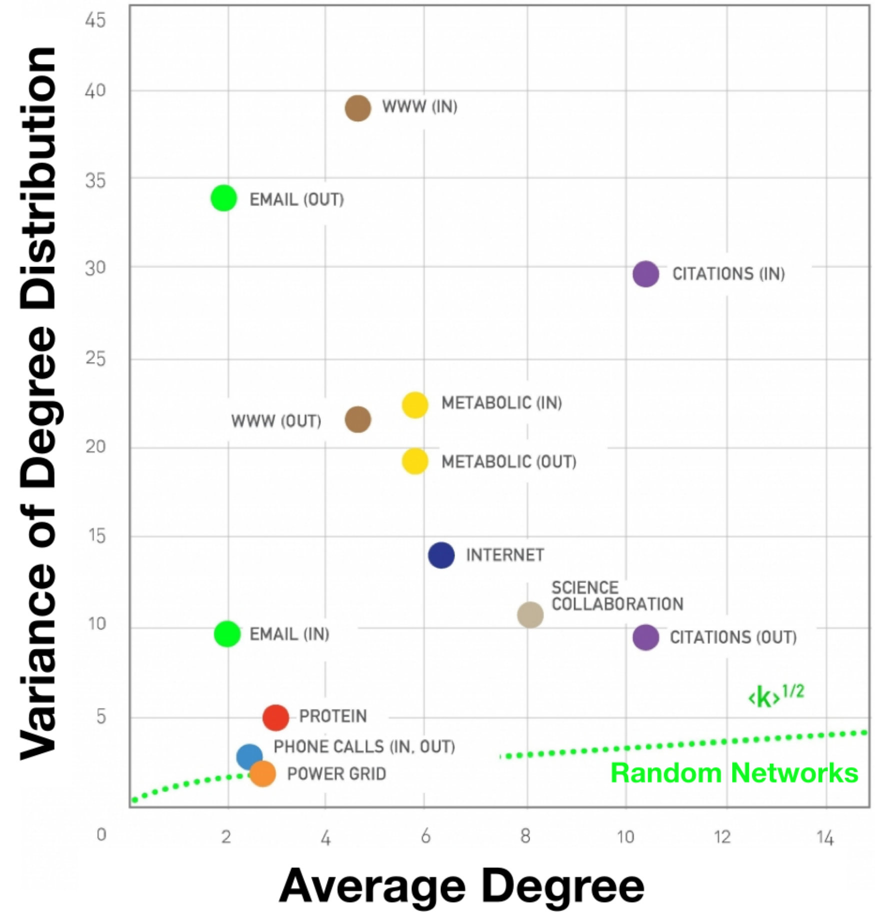
regular, lattices
physical systems

random structures
mathematics

complex structures
social systems



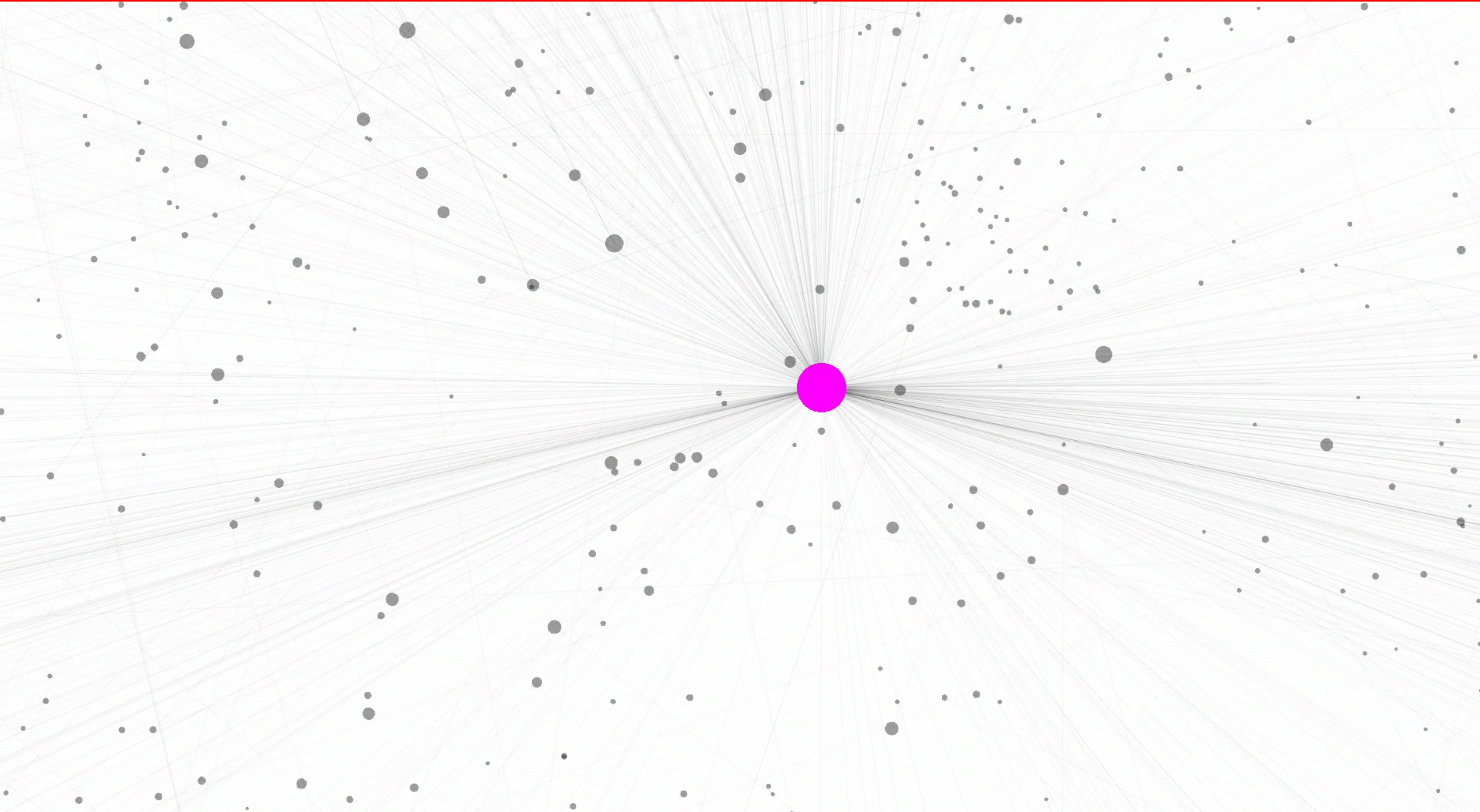
Las redes reales no son random son libre de escala!



Leyes de potencia y redes libre de escala

El nombre libre de escala captura la falta de una escala interna, una consecuencia del hecho de que los nodos con grados muy diferentes coexisten en la misma red. Esta característica distingue las redes libre de escala de las grillas, en las que todos los nodos tienen exactamente el mismo grado ($\sigma = 0$), o de las redes aleatorias, cuyos grados varían en un rango estrecho ($\sigma = \langle k \rangle^{1/2}$)

WORLD WIDE WEB



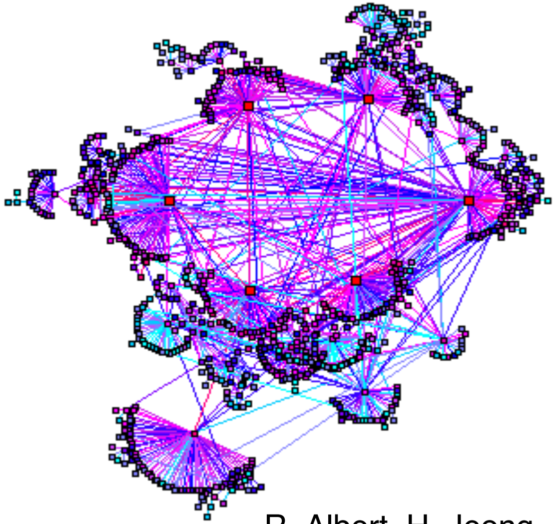
R. Albert, H. Jeong, A-L Barabasi, *Nature*, 401 130 (1999).

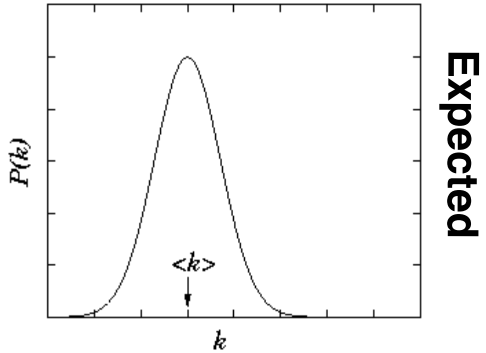
WORLD WIDE WEB

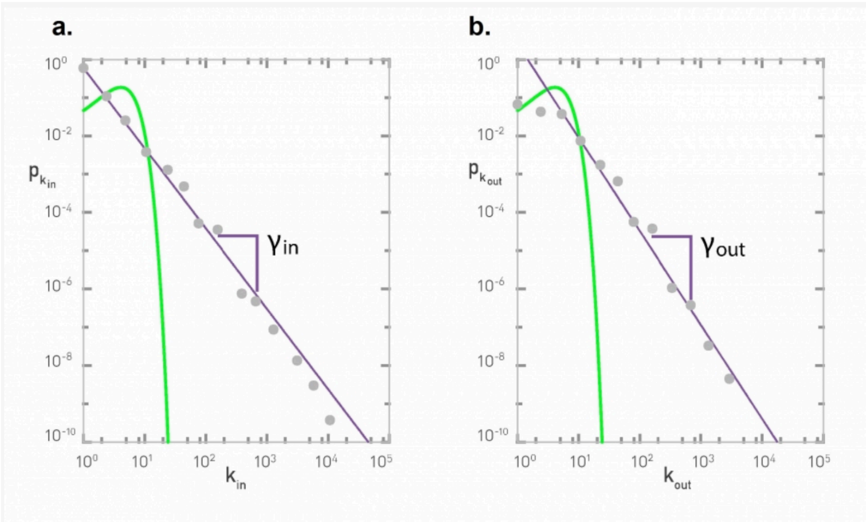
Nodos: **documentos WWW**
Enlaces: **links URL**

Más de 3 mil millones de documentos.

ROBOT: recopila todas las URL encontradas en un documento y las sigue de manera recursiva

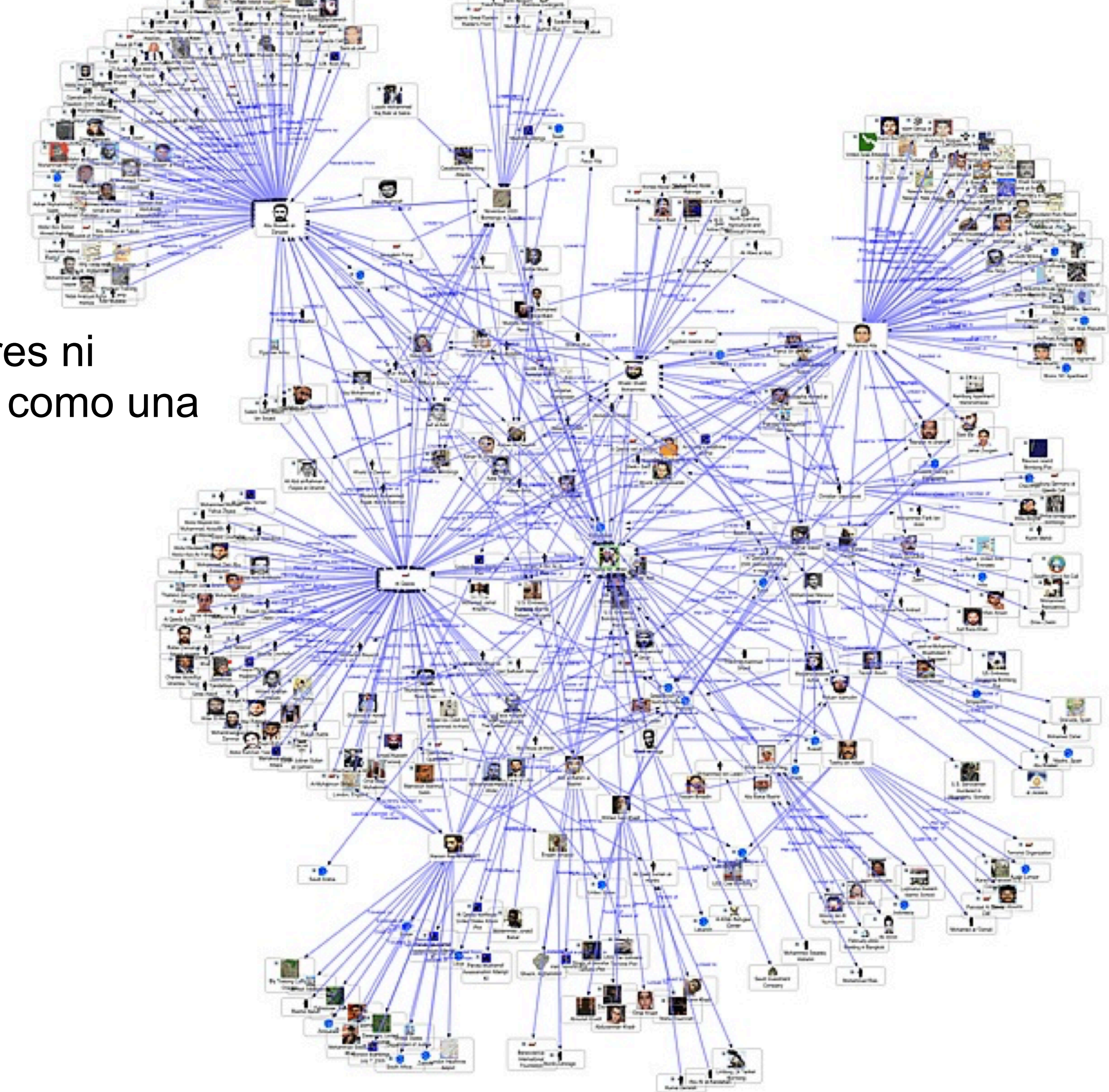






R. Albert, H. Jeong, A-L Barabasi, *Nature*, 401 130 (1999).

Las redes reales no son regulares ni aleatorias, ¡son major descritas como una red libre de escala!



- El nombre sin escala captura la falta de una escala interna, una consecuencia del hecho de que nodos con grados muy diferentes coexisten en la misma red. Esta característica distingue las redes sin escala de las redes, en las que todos los nodos tienen exactamente el mismo grado ($\sigma = 0$), o de las redes aleatorias, cuyos grados varían en un rango estrecho ($\sigma = \langle k \rangle^{1/2}$)

Formalismo discreto vs. continuo.

Formalismo discreto

Como los grados de nodo son siempre enteros positivos, el formalismo discreto captura la probabilidad de que un nodo tenga exactamente k enlaces:

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

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

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

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

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

Formalismo continuo

En los cálculos analíticos a menudo es conveniente suponer que los grados pueden tomar cualquier valor real positivo:

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

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

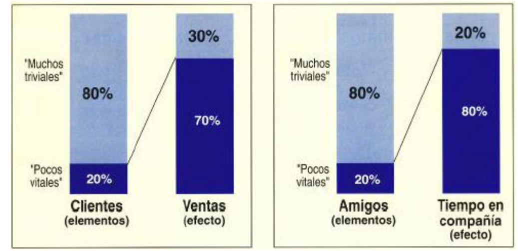
$$C = \frac{1}{\int_{k_{\min}}^{\infty} k^{-\gamma} dk} = (\gamma - 1)k_{\min}^{\gamma-1}$$

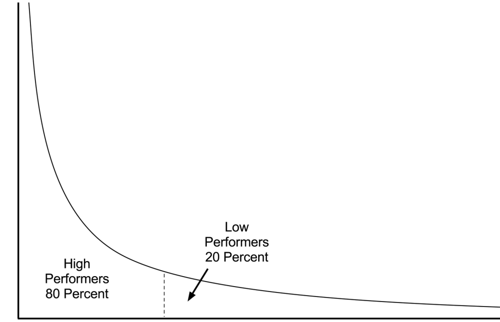
$$p(k) = (\gamma - 1)k_{\min}^{\gamma-1} k^{-\gamma} .$$

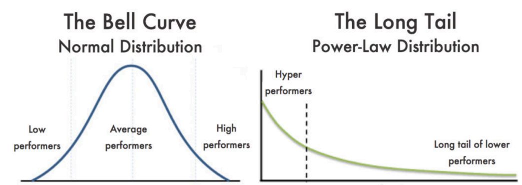
INTERPRETATION: p_k $\int_{k_1}^{k_2} p(k)dk$

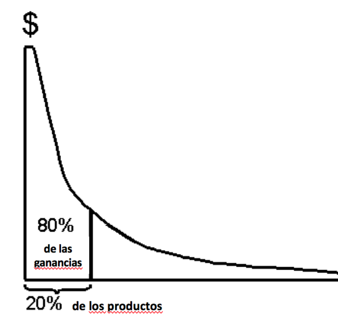
REGLA 80/20

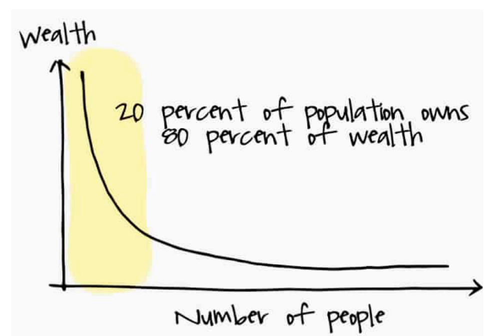
- **Vilfredo Federico Damaso Pareto (1848 – 1923),** Economista, politólogo y filósofo italiano, que contribuyó de manera importante a nuestra comprensión de la distribución del ingreso y al análisis de las elecciones individuales. Algunos de sus principios fundamentales llevan su nombre, como la eficiencia de Pareto, la distribución de Pareto (otro nombre para una distribución de ley de poder), el principio de Pareto (o ley 80/20).

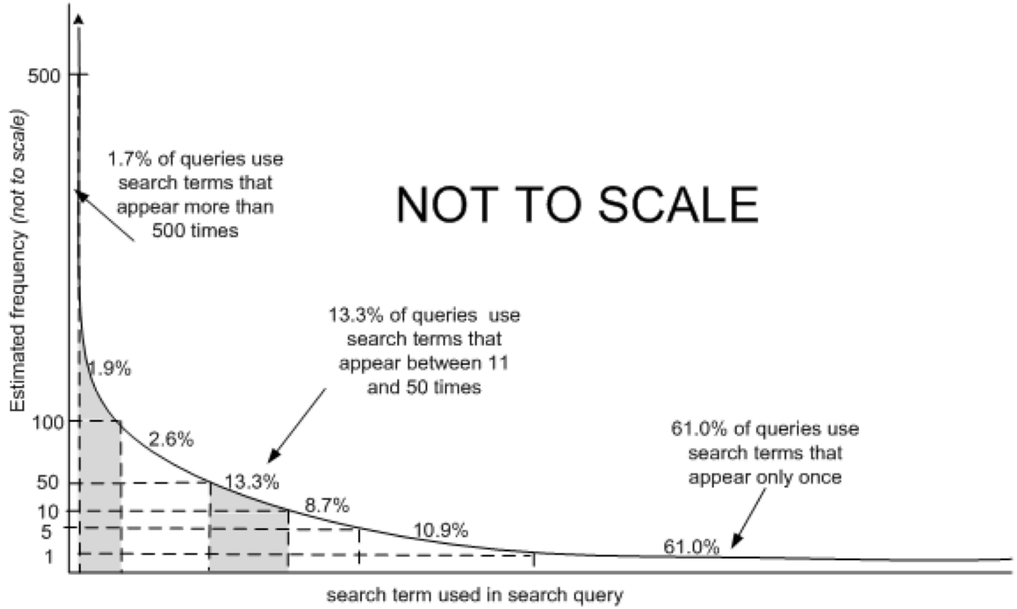


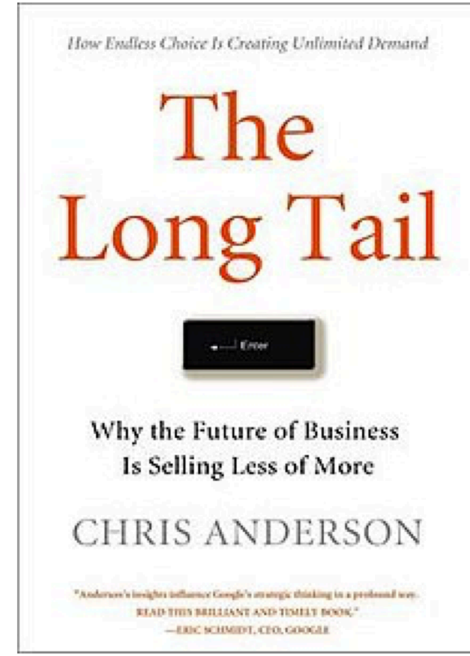












¿Qué sucede cuando desaparecen los cuellos de botella que se interponen entre la oferta y la demanda en nuestra cultura y todo está al alcance de todos?

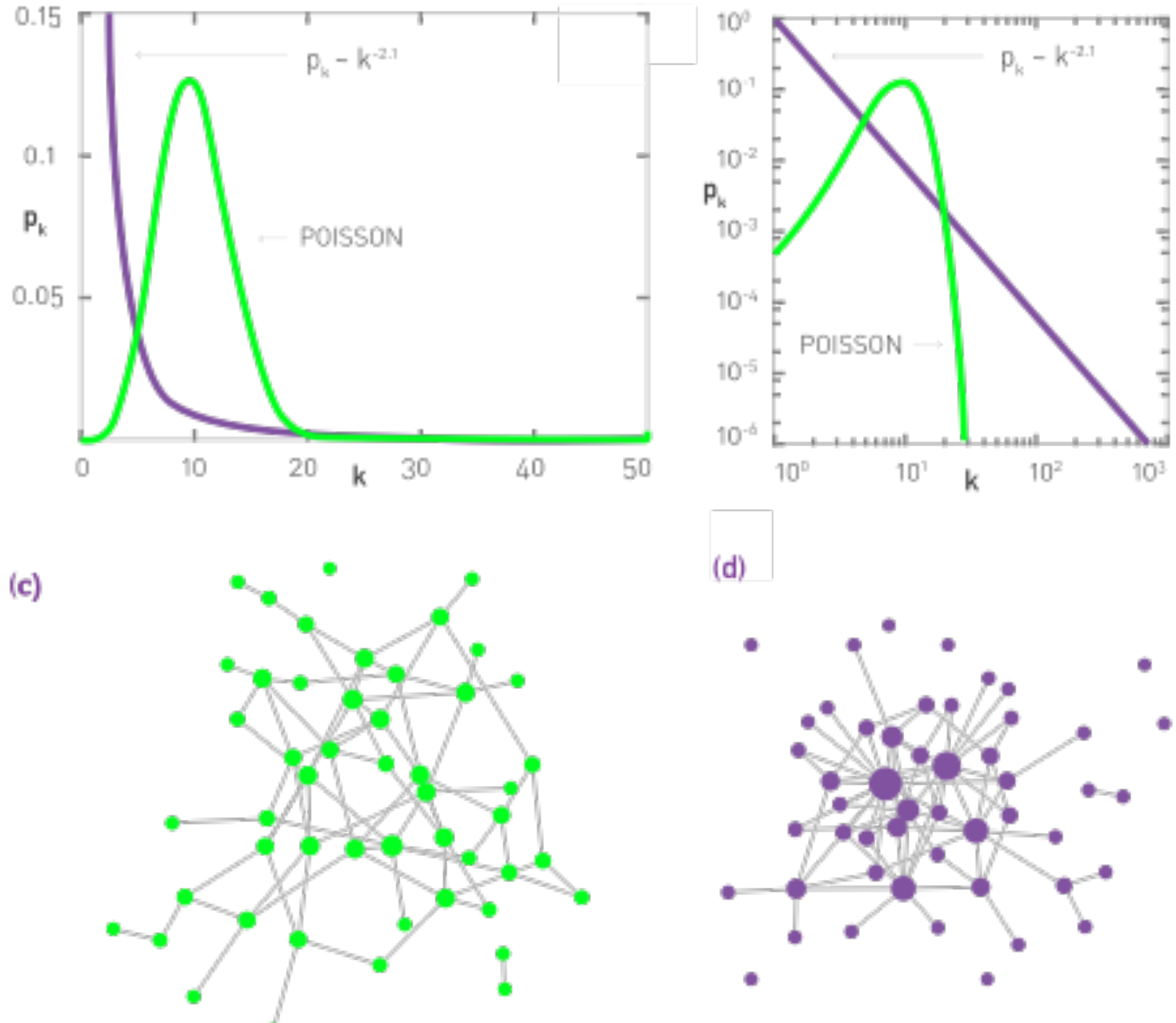
"The Long Tail" es una nueva y poderosa fuerza en nuestra economía: el surgimiento del nicho. A medida que el costo de llegar a los consumidores se reduce drásticamente, nuestros mercados están pasando de un modelo único de atractivo masivo a uno de variedad ilimitada para gustos únicos. Desde los estantes de los supermercados hasta las agencias de publicidad, la capacidad de ofrecer una gran variedad de opciones lo está cambiando todo y nos obliga a repensar dónde se encuentran nuestros mercados y cómo llegar a ellos. La selección ilimitada está revelando verdades sobre lo que quieren los consumidores y cómo quieren obtenerlo, desde DVD en Netflix hasta canciones en iTunes y publicidad en Google.

Sin embargo, esto no es solo una virtud de los mercados en línea; es un ejemplo de un modelo económico completamente nuevo para los negocios, uno que recién comienza a mostrar su poder. **Después de un siglo de obsesionarnos con los pocos productos que se encuentran a la cabeza de la curva de demanda, la nueva economía de la distribución nos permite centrarnos en los muchos más productos que se encuentran en la cola, que colectivamente pueden crear un nuevo mercado tan grande como el que tenemos. ya saben.**

The Long Tail se trata realmente de la economía de la abundancia. Las nuevas eficiencias en la distribución, la fabricación y el marketing esencialmente están restableciendo la definición de lo que es comercialmente viable en todos los ámbitos. Si el siglo XX fue de hits, el XXI será igualmente de nichos.

Hubs

La diferencia entre una ley de potencia y una distribución exponencial.



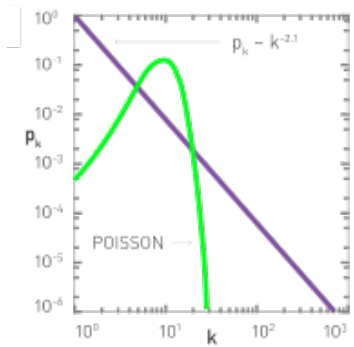
La diferencia entre una ley de potencia y una distribución exponencial.

Usemos la WWW para ilustrar las propiedades del régimen de alto-k.
La probabilidad de tener un nodo con k ~ 100 es

•*Cercana a* $p_{100} \simeq 10^{-30}$ *en una distribución de Poisson*

•*Cercana a* $p_{100} \simeq 10^{-4}$ *si* p_k *sigue una ley de potencia*

•*En consecuencia, si la WWW fuera una red aleatoria, de acuerdo con la predicción de Poisson, esperaríamos* 10^{-18} *nodos de grado k>100 , o ninguno.*
•*Para una distribución de grado de ley de potencia, esperamos alrededor de* $N_{k>100} = 10^9$ *nodos de grado k>100*




(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
Number of links (k)
(d)
Chicago
Boston
Los Angeles

El tamaño del hub más grande.

Todas las redes reales son finitas → exploremos sus consecuencias.

→ Tenemos un grado máximo esperado, k_{max}

Estimando k_{max}

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

Por qué: la probabilidad de tener un nodo mayor que k_{max} no debe exceder la probabilidad de tener un nodo, es decir, una fracción 1 / N de todos los nodos

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

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

El tamaño del hub más grande.

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

Red libre de escala:

$$\gamma = 2.1 \qquad k_{max} \simeq 85{,}000$$

Red libre aleatoria (hubs prohibidos):

$$k_{max} \simeq 13$$

Red real WWW: k_{max}=10,721

Redes libre de escala finitas

Grado máximo esperado, k_{max}

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

•k_{max}, Aumenta con el tamaño de la red.

→ Cuanto más grande es un sistema, más grande es su **hub** más grande.

•Para γ>2 k_{max} aumenta mas lento que N

àel **hub** más grande contendrá una fracción decreciente de enlaces a medida que N aumente.

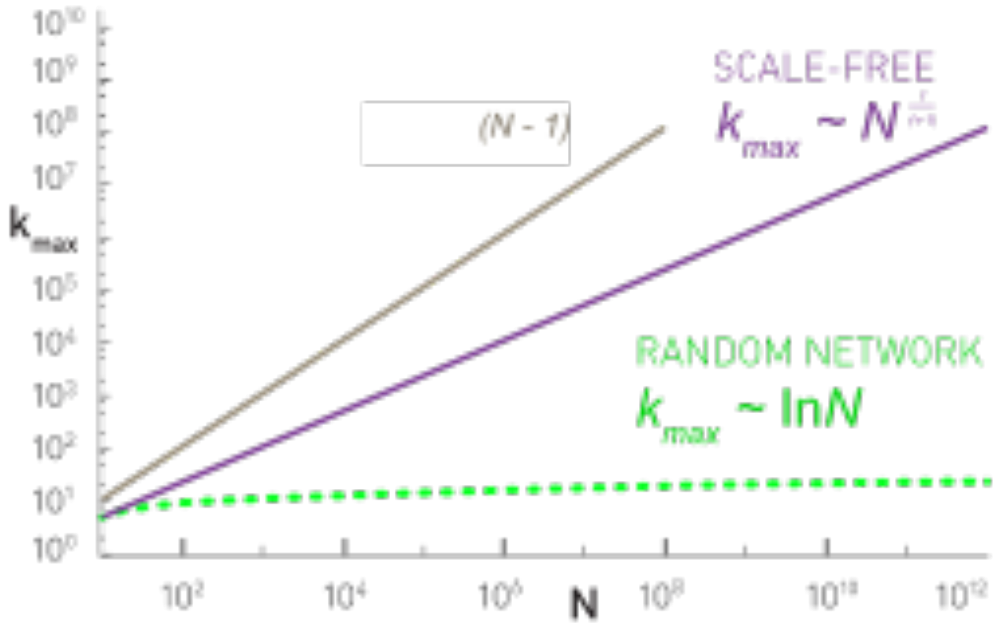
•Para γ=2 k_{max}~N.

→ El tamaño del centro más grande es O(N) (lineal con N)

•Para γ<2 k_{max} aumenta más rápido que N: fenómenos de condensación

→ el **hub** más grande tomará una fracción creciente de enlaces. ¡Anomalía!

El tamaño del hub más grande.



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

El significado de la propiedad libre de escala

Redes libre de escala: Definición

Definición :

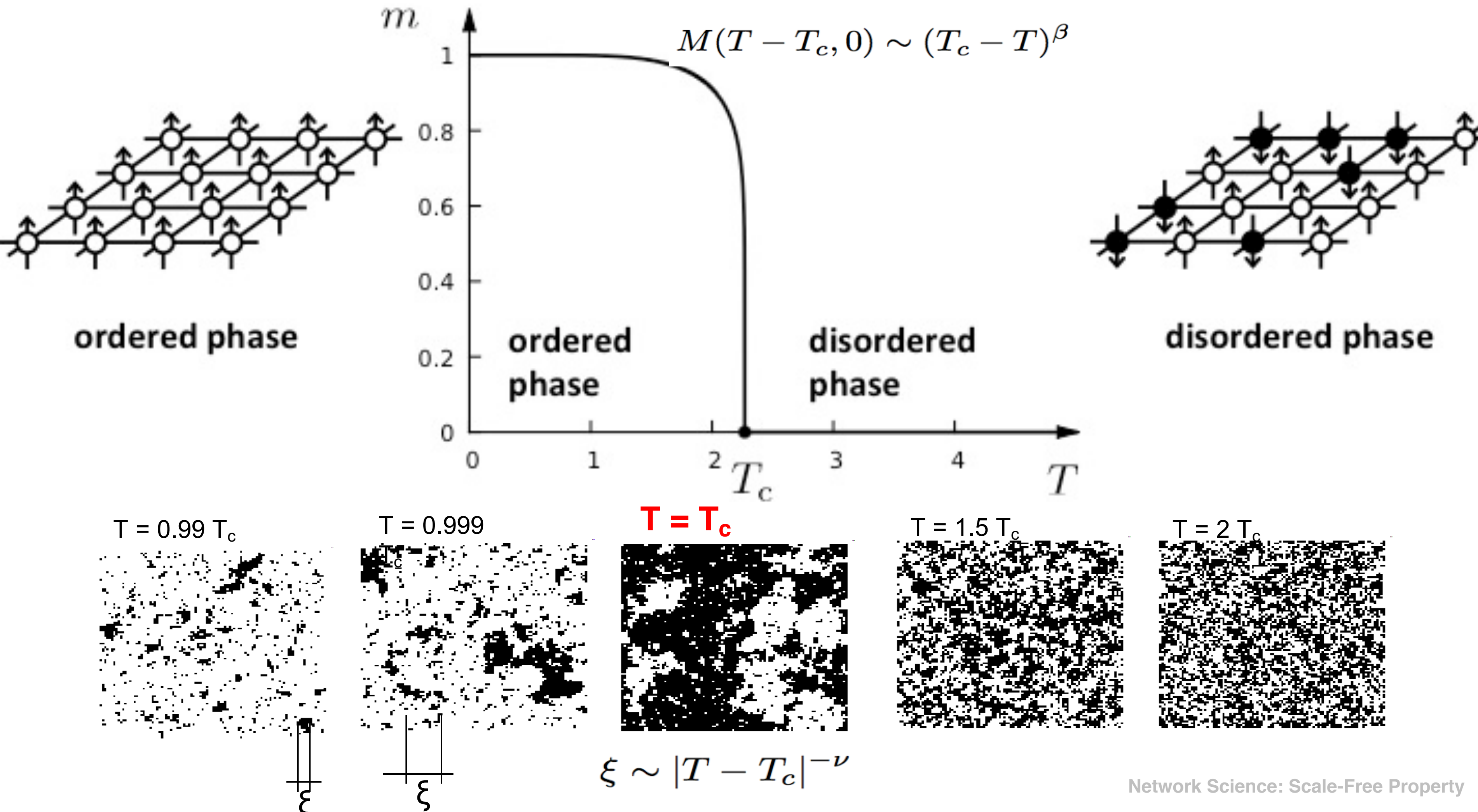
Las redes con una ley de potencia en su distribución de grado se llaman 'Redes libre de escala'

¿De dónde viene el nombre?

Fenómenos críticos e invariancia de escala.

(un desvío)

Transiciones de fase en sistemas complejos I: Magnetismo



Comportamiento libre de escalas en el espacio.

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

a T = Tc:

la longitud de correlación diverge

Las fluctuaciones emergen en todas las escalas:

comportamiento libre de escala

FENOMENOS CRITICOS

- La longitud de correlación diverge en el punto crítico: ¡todo el sistema está correlacionado!

- Invariancia de escala: no hay una escala característica para la fluctuación (comportamiento libre de escala).

- Universalidad: los exponentes son independientes de los detalles del sistema.

Divergencias en distribuciones sin escala.

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

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

Exponente debe ser negativo

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

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

If m-γ+1>0, La integral diverge.

Para un γ fijo, esto significa que todos los momentos con m>γ-1 divergen.

DIVERGENCIA DE LOS MOMENTOS SUPERIORES

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

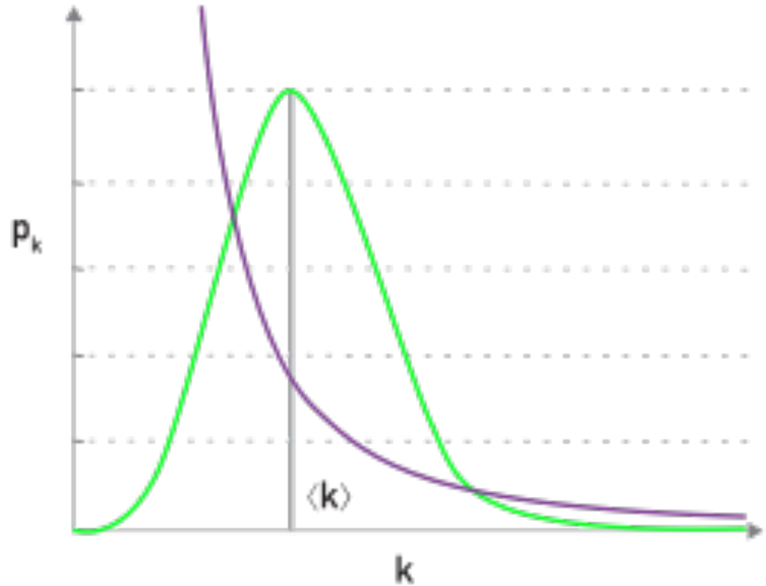
Para un γ fijo, esto significa que todos los momentos con $m>\gamma-1$ divergen.

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

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*-

Muchos exponentes de grado son más pequeños que 3

→ $<k^2>$ diverge en el límite $N \to \infty$



Random Network

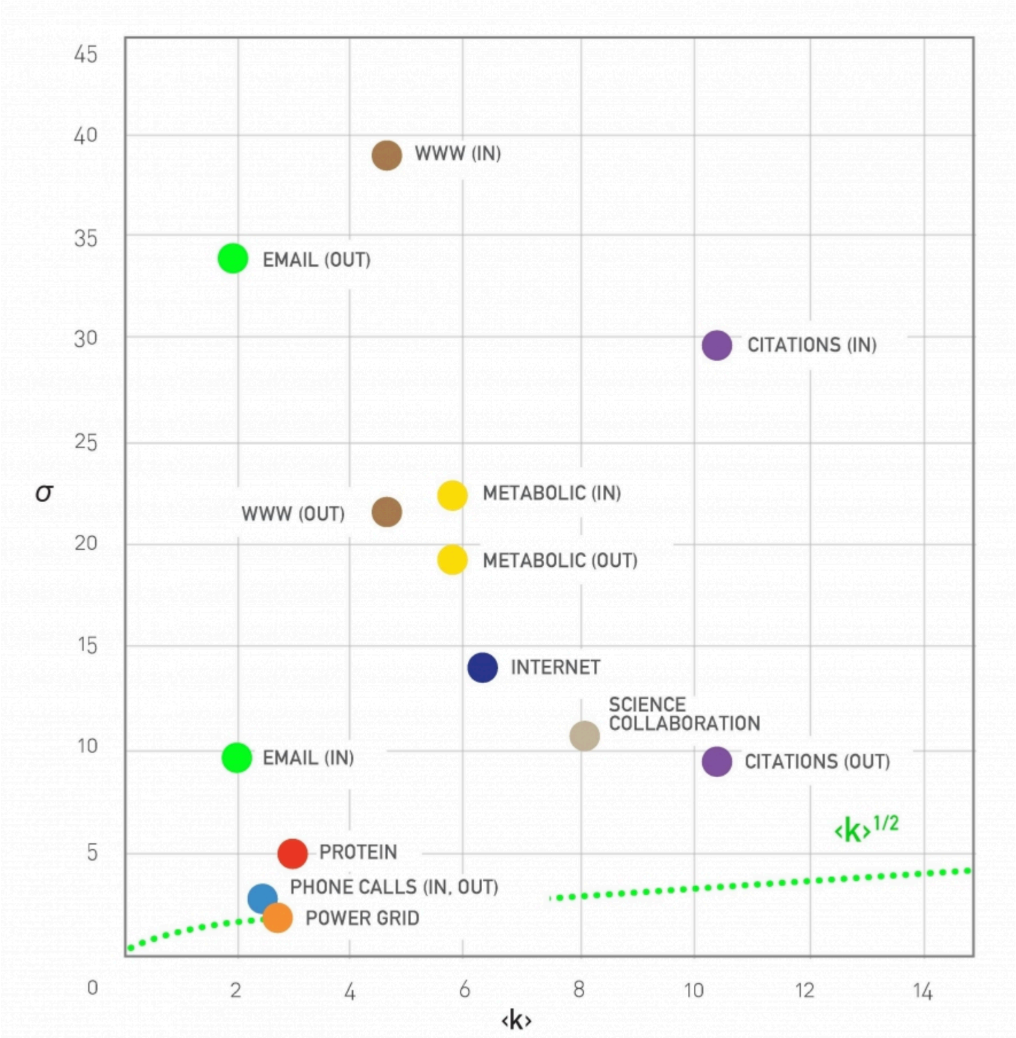
Nodo elegido al azar: $k = \langle k \rangle \pm \langle k \rangle^{1/2}$

Scale: $\langle k \rangle$

Scale-Free Network

Nodo elegido al azar: $k = \langle k \rangle \pm \infty$

Scale: none

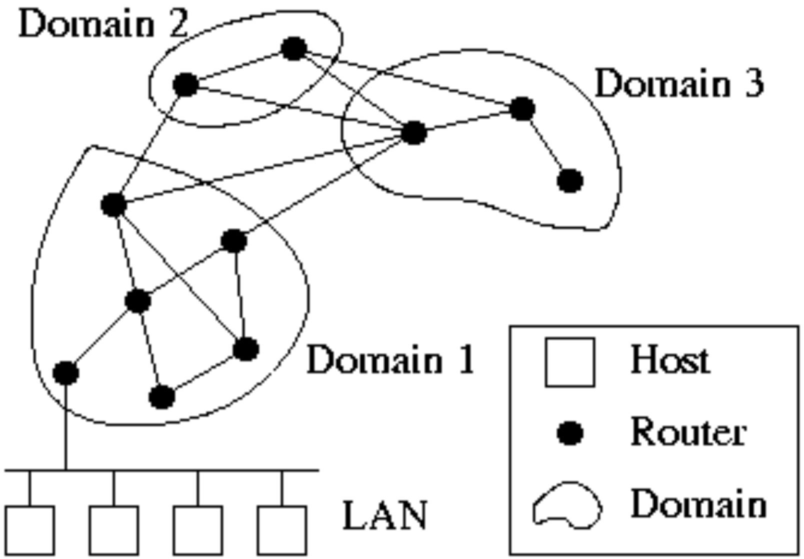


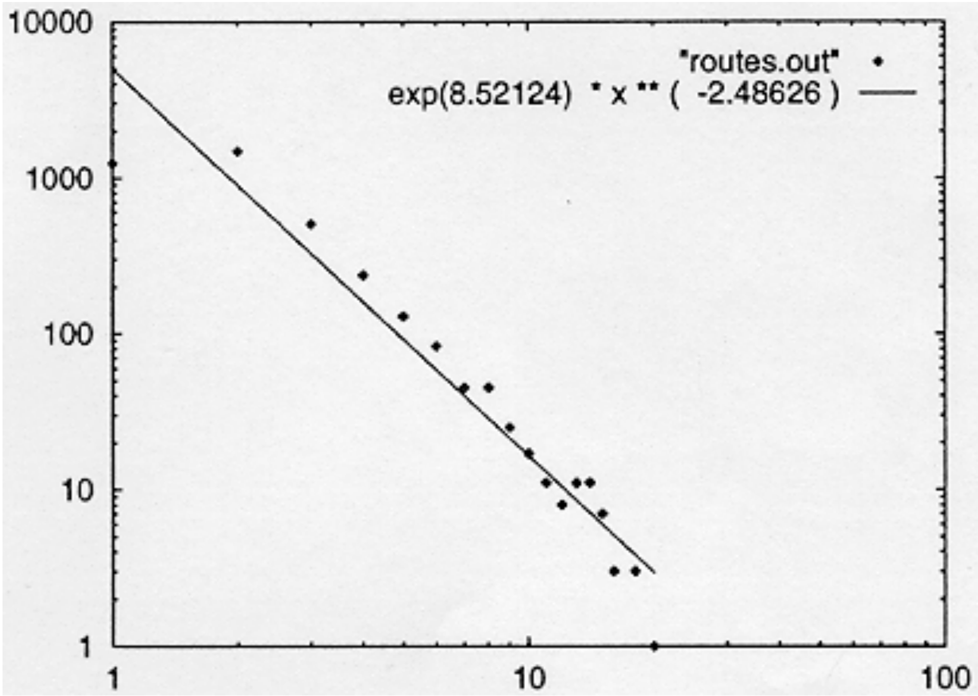
$$k = \langle k \rangle \pm \sigma_k$$

Universalidad

Backbone de internet

Nodos: computadores, routers
Enlaces: líneas físicas





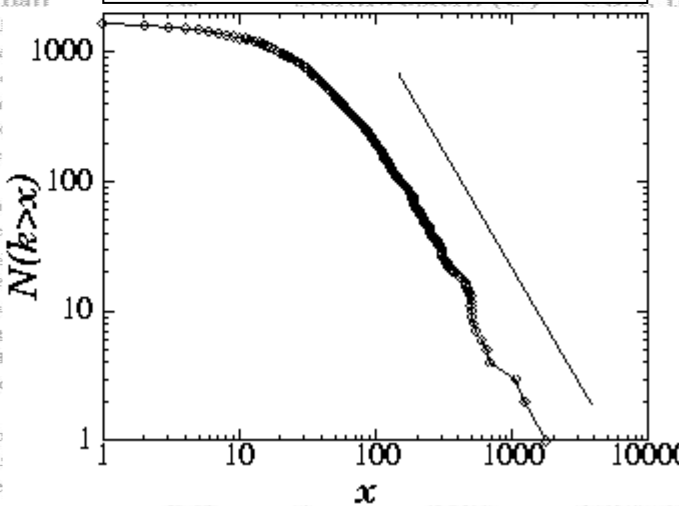
(Faloutsos, Faloutsos and Faloutsos, 1999)

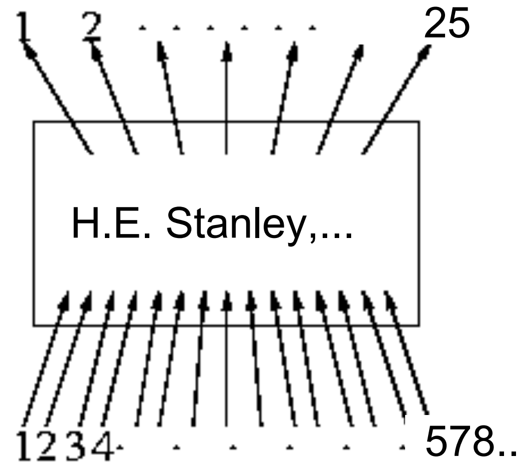
INDICE DE CITACIONES EN CIENCIA

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

Nodos: papers
Enlaces: citaciones

1736 PRL papers (1988)





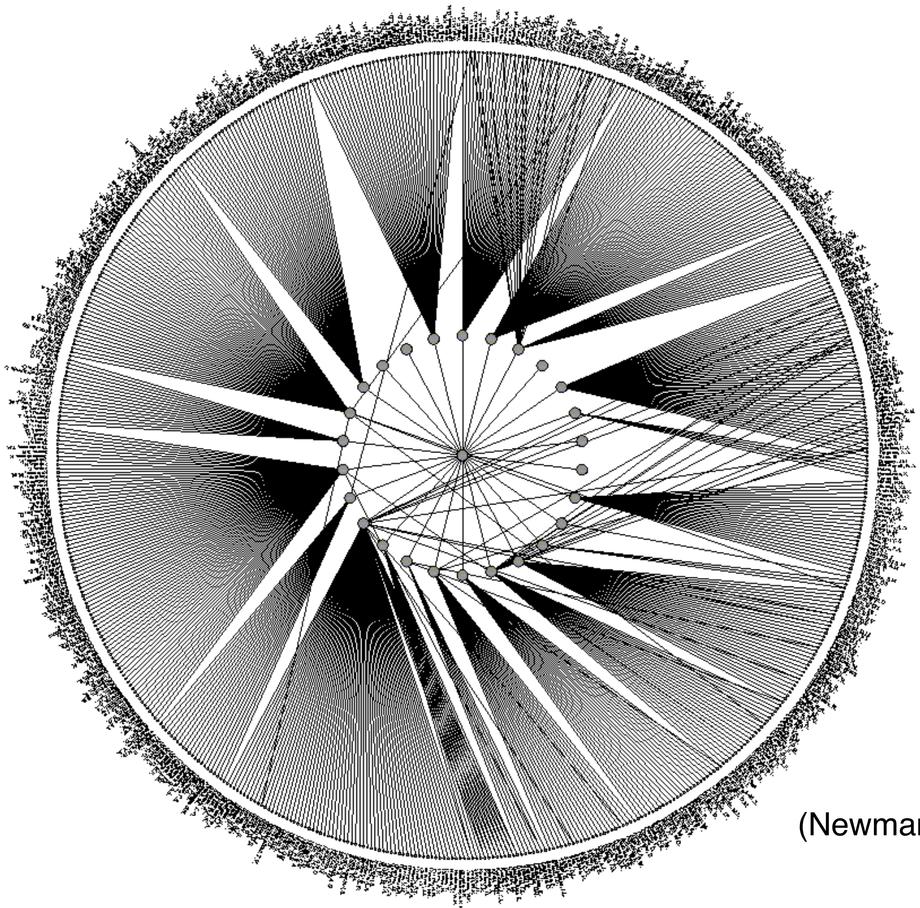
$P(k) \sim k^{-\gamma}$
($\gamma = 3$)

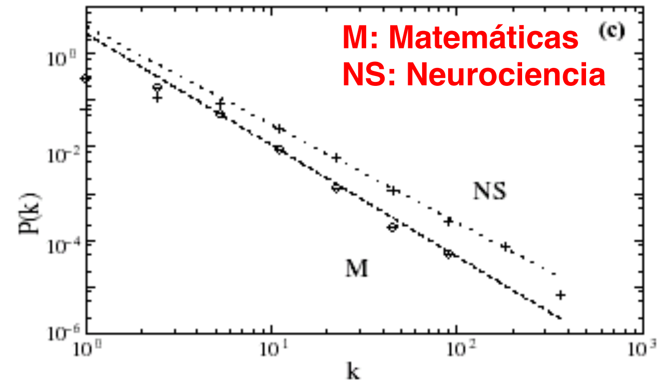
(S. Redner, 1998)

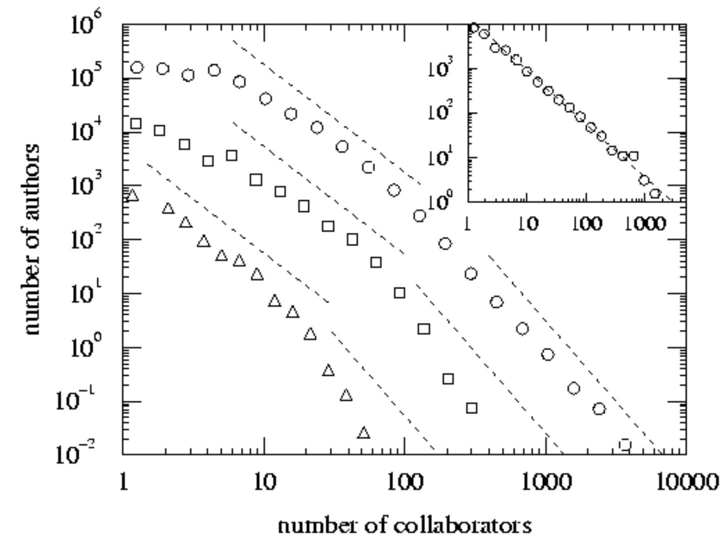
* citation total may be skewed because of multiple authors with the same name

CO-AUTORIA EN CIENCIA

Nodos: Científicos (autores)
Links: Publicaciones conjuntas







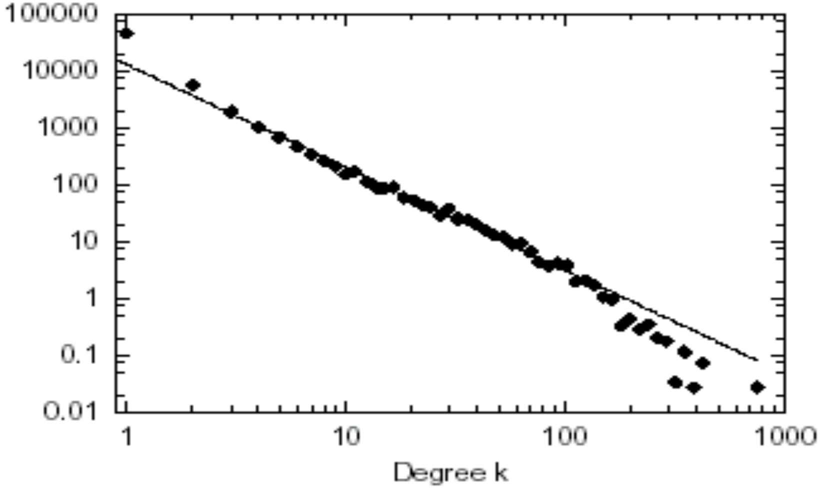
(Newman, 2000, Barabasi et al 2001)

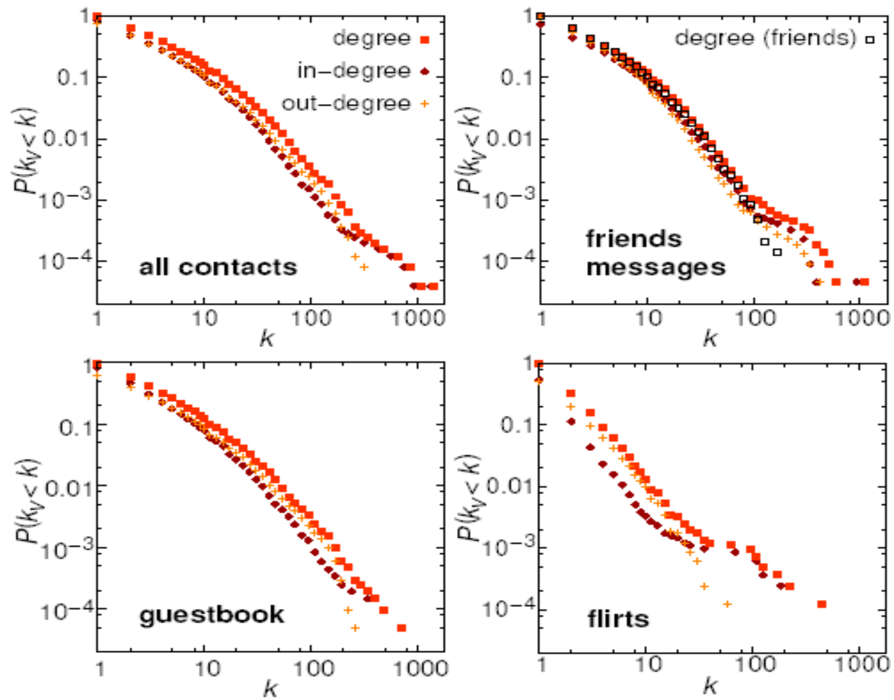
COMUNIDADES ONLINE

Nodos: Usuarios online
Enlaces: contactos email

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

Archivos log de Kiel University
112 días, N=59,912 nodos



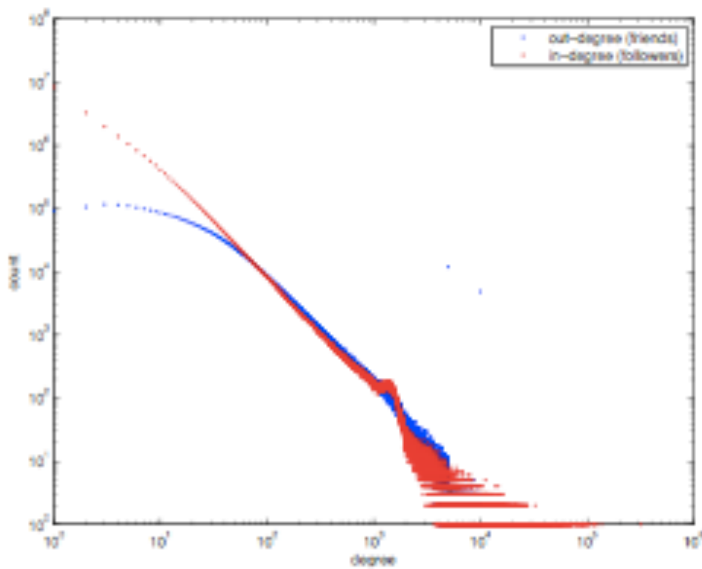


Ebel, Mielsch, Bornholdtz, PRE 2002.

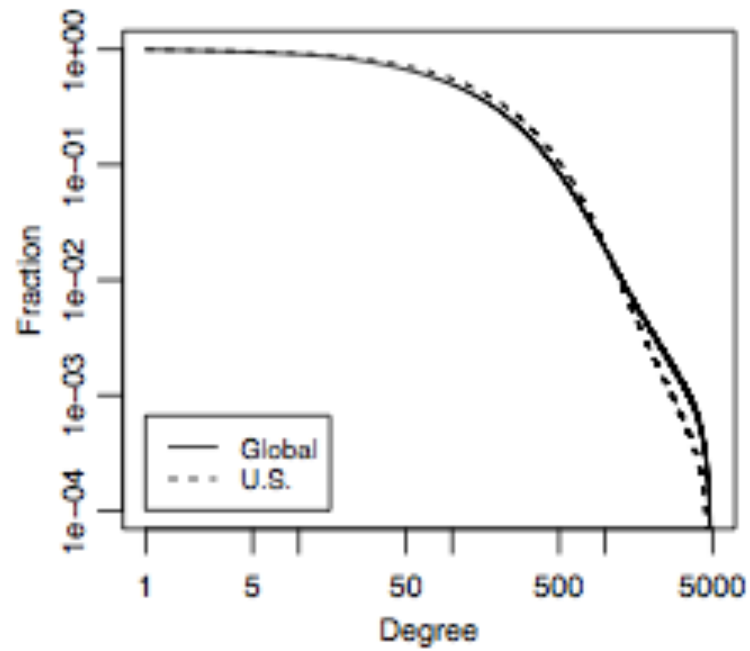
Holme, Edling, Liljeros, 2002.

COMUNIDADES ONLINE

Twitter:

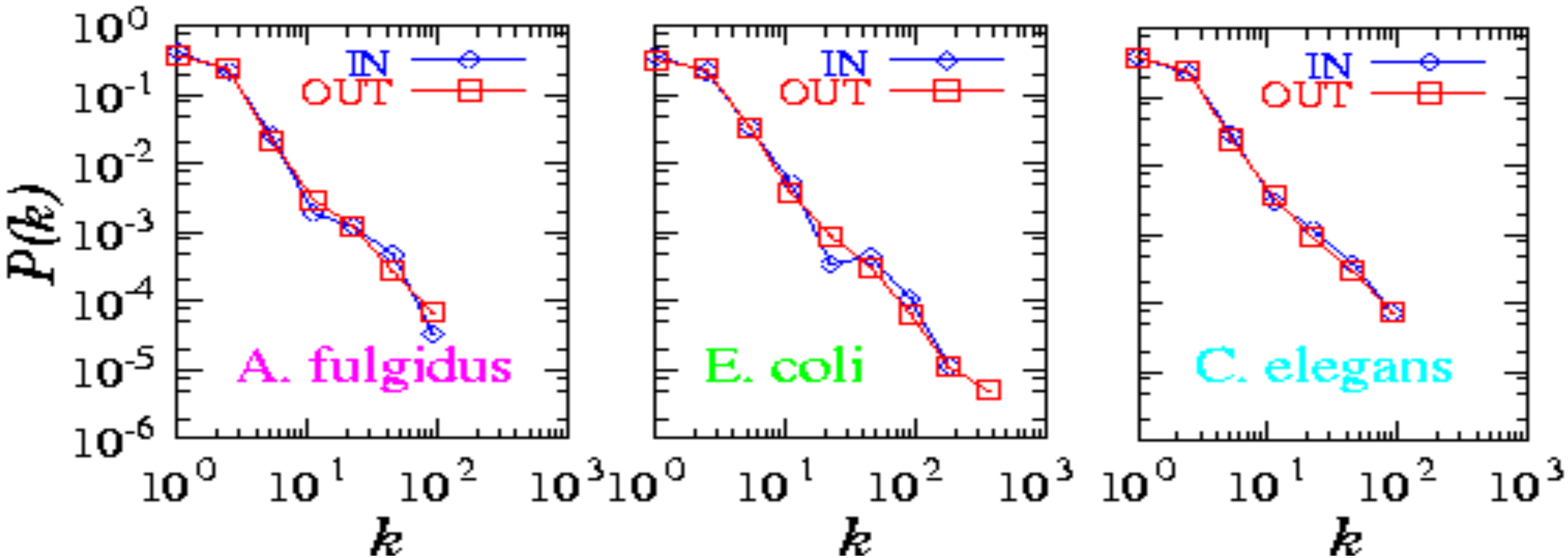


Facebook



Brian Karrer, Lars Backstrom, Cameron Marlowm 2011

RED METABOLICA



Archaea Bacteria Eukaryotes

¡Los organismos de los tres dominios de la vida son libres de escala!

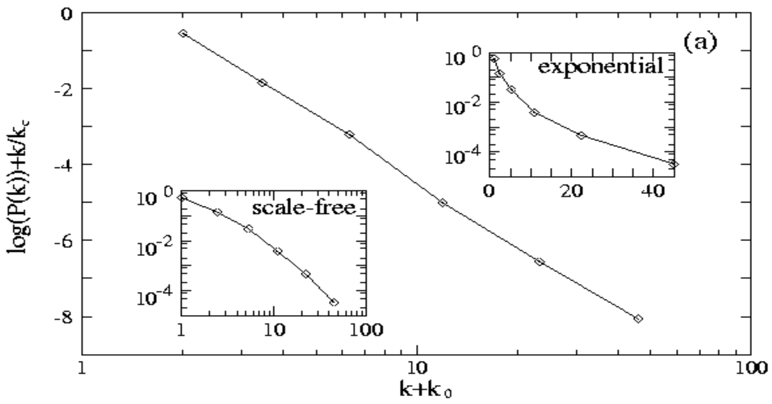
$$P_{in}(k) \approx k^{-2.2}$$

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

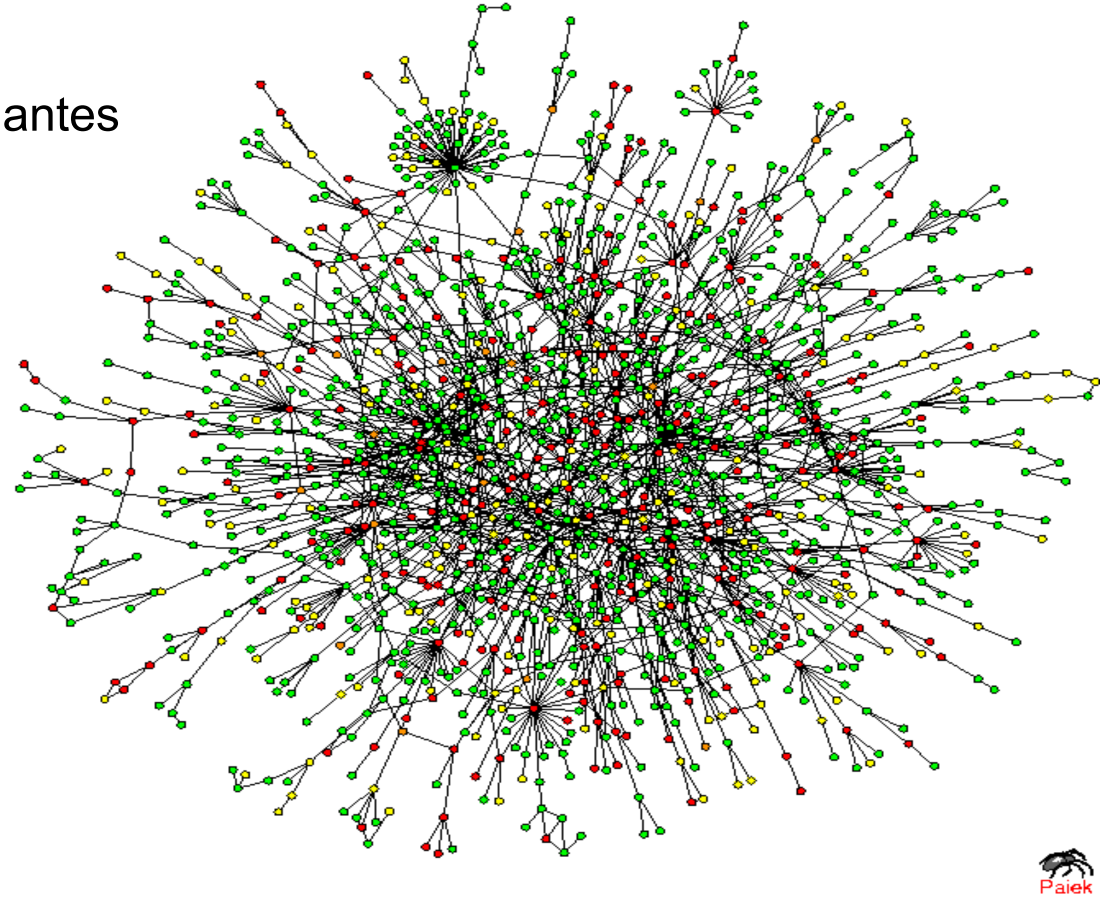
H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabasi, *Nature*, 407 651 (2000)

TOPOLOGÍA DE LA RED DE PROTEÍNAS

Nodos: proteínas
Enlaces: interacciones físicas vinculantes



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

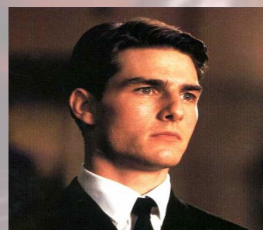


H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Nature 411, 41-42 (2001

RED DE ACTORS

Nodos: actores
Enlaces: película conjunta

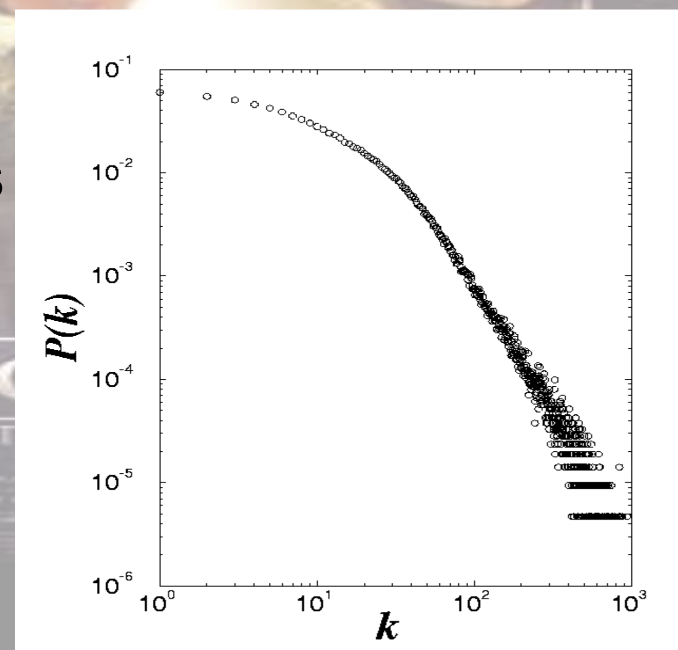
IMDb Internet Movie Database



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

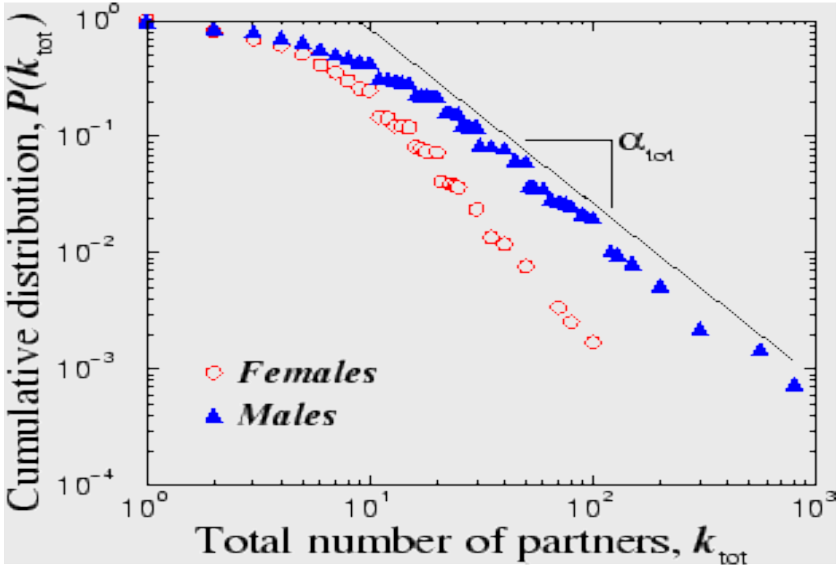
N = 212,250 actores
$\langle k \rangle = 28.78$

$P(k) \sim k^{-\gamma}$
$\gamma = 2.3$



SWEDISH SE-WEB

Nodos: Personas (Mujeres; Hombres)
Enlaces: Relaciones sexuales



4781 Suizos; 18-74;
59% tasa de respuesta.

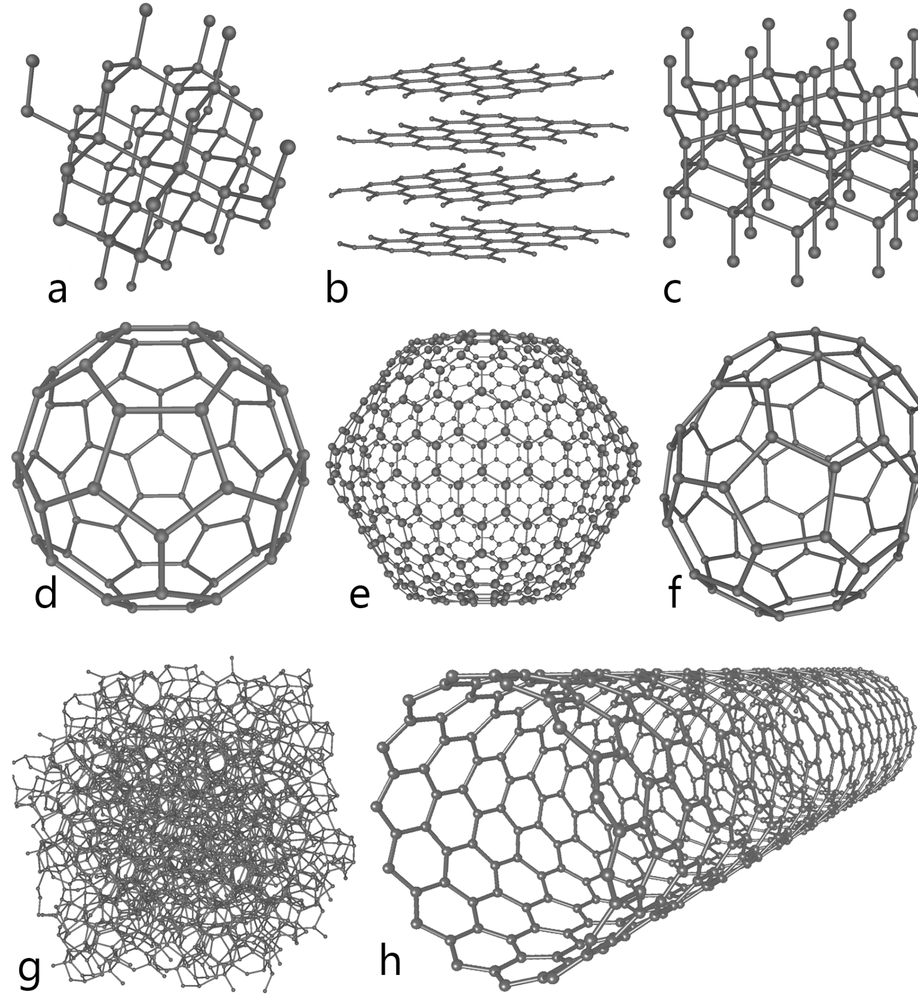
Liljeros et al. Nature 2001

No todas las redes pueden ser descritas como libres de escala

- *Redes que aparecen en la ciencia de los materiales, como la red que describe los enlaces entre los átomos en materiales cristalinos o amorfos, donde cada nodo tiene exactamente el mismo grado.*

- *La red neuronal del gusano C.elegans.*

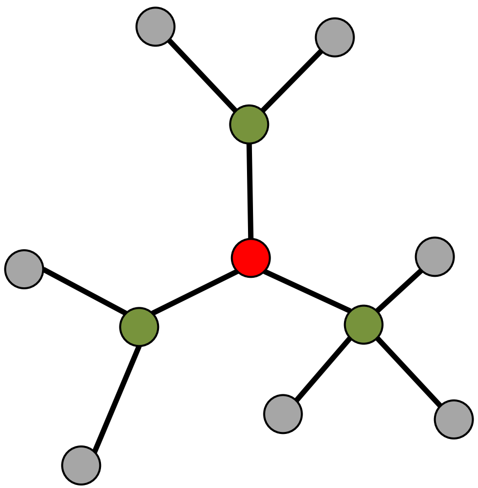
- *La red eléctrica, compuesta por generadores e interruptores conectados por líneas de transmisión.*



Propiedad ultra-pequeña

DISTANCIAS EN GRAFOS ALEATORIOS

Los grafos aleatorios tienden a tener una topología en forma de árbol con grados de nodo casi constantes.



- nr de primeros vecinos $N_1 \cong \langle k \rangle$
- nr de segundos vecinos: $N_2 \cong \langle k \rangle^2$
- nr de vecinos a distancia d: $N_d \cong \langle k \rangle^d$
- estimar la distancia máxima:

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

COMPORTAMIENTO DE MUNDO PEQUEÑO EN REDES SIN ESCALA

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

Ultra Small World

Small World

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

El tamaño del hub más grande es de orden O (N). La mayoría de los nodos se pueden conectar dentro de dos enlaces (la gran mayoria estan conectados al hub), por lo que la longitud promedio de la ruta será independiente del tamaño del sistema.

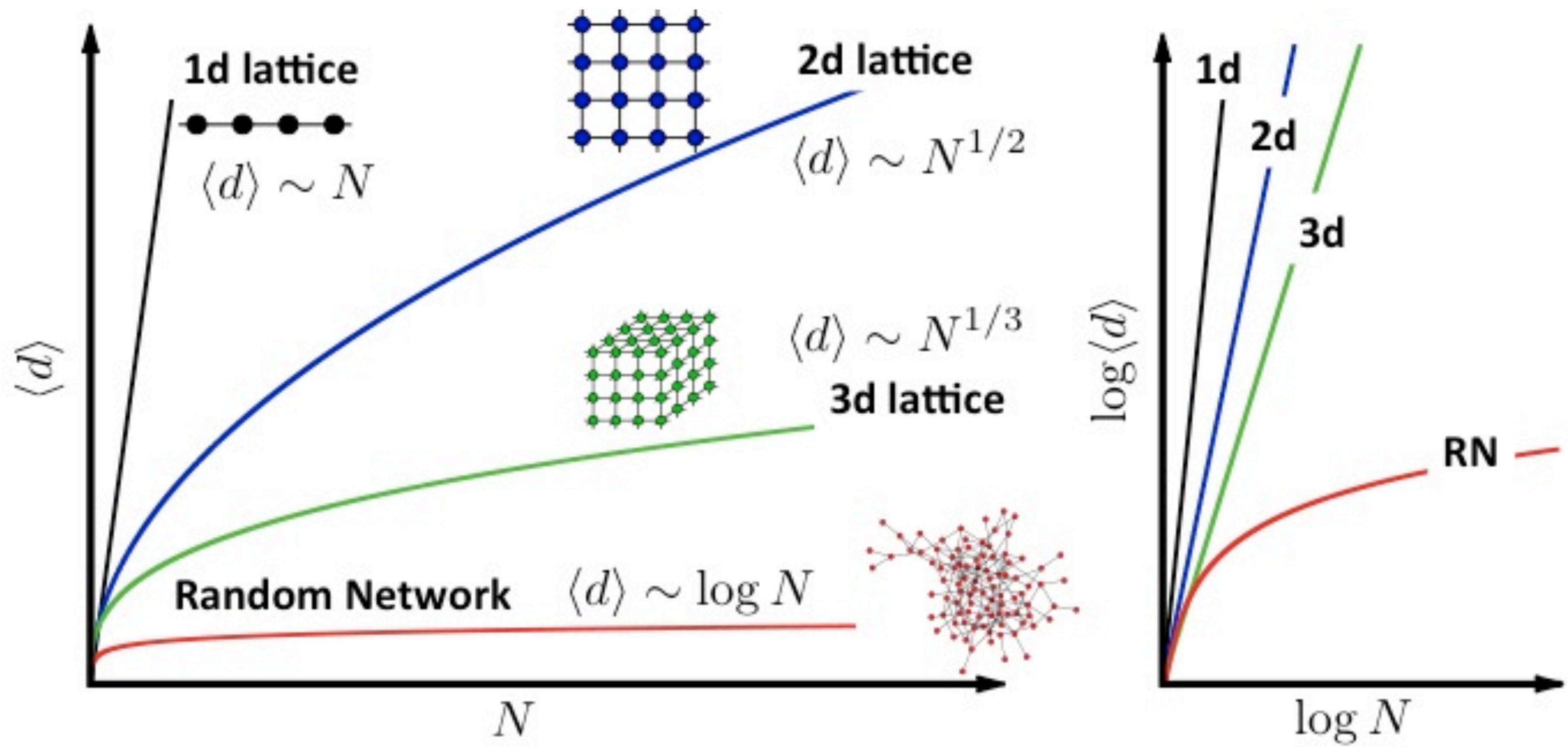
La longitud media de la ruta aumenta más lentamente que logarítmicamente. En una red aleatoria, todos los nodos tienen un grado comparable, por lo que la mayoría de las rutas tendrán una longitud comparable. En una red libre de escala, la gran mayoría de la ruta pasa por unos pocos hubs de alto grado, lo que reduce las distancias entre los nodos.

Algunos modelos clave producen $\gamma = 3$, por lo que el resultado es de particular importancia para ellos. Esto fue derivado por primera vez por Bollobas et al. para el diámetro de la red en el contexto de un modelo dinámico, pero también se aplica a la longitud de ruta promedio.

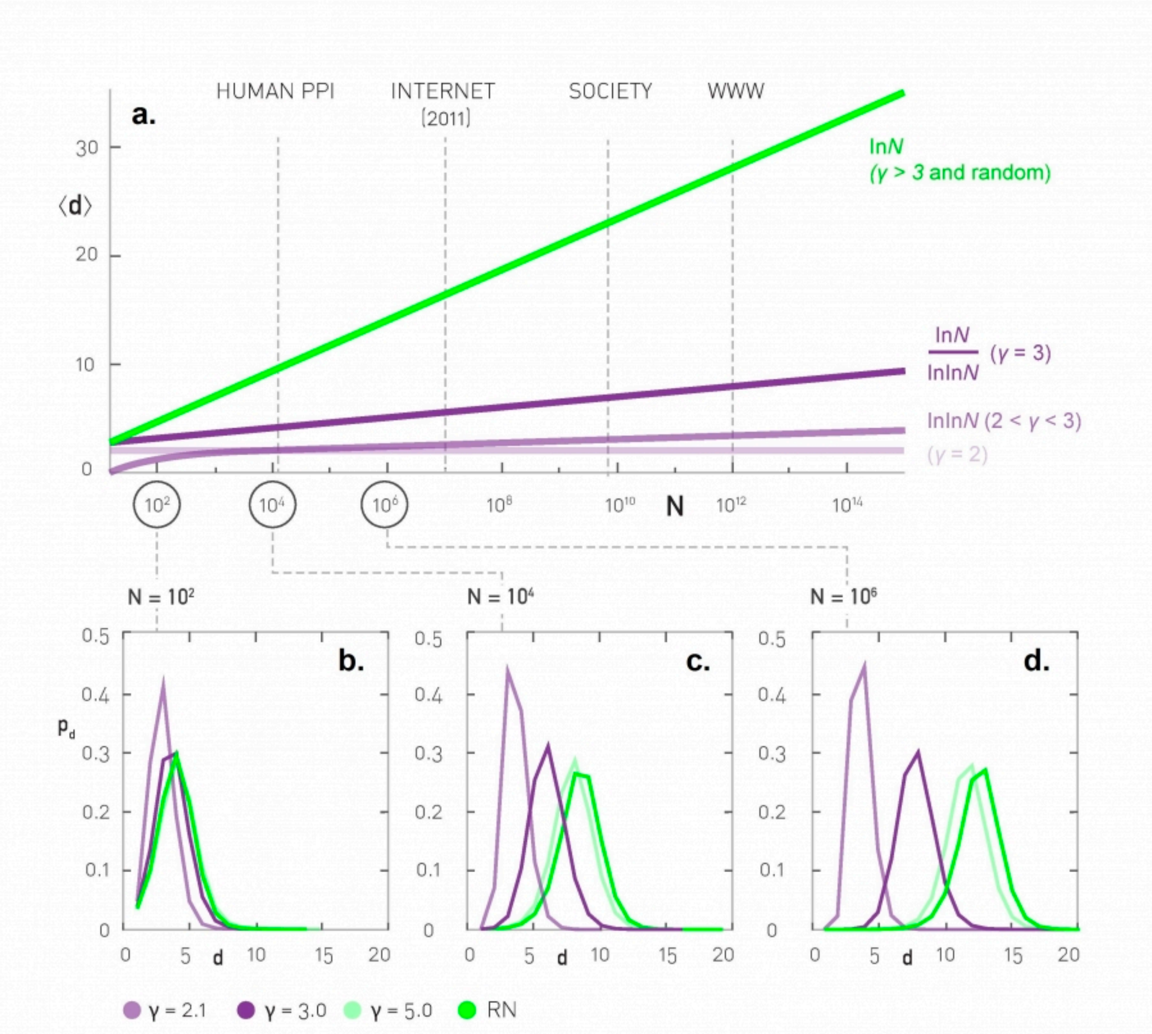
El segundo momento de la distribución es finito, por lo que de muchas maneras la red se comporta como una red aleatoria. Por lo tanto, la longitud promedio de la ruta sigue el resultado que obtuvimos anteriormente para el modelo de red aleatorio.

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

¿Por qué sorprenden los pequeños mundos? Sorprendente en comparación con qué?

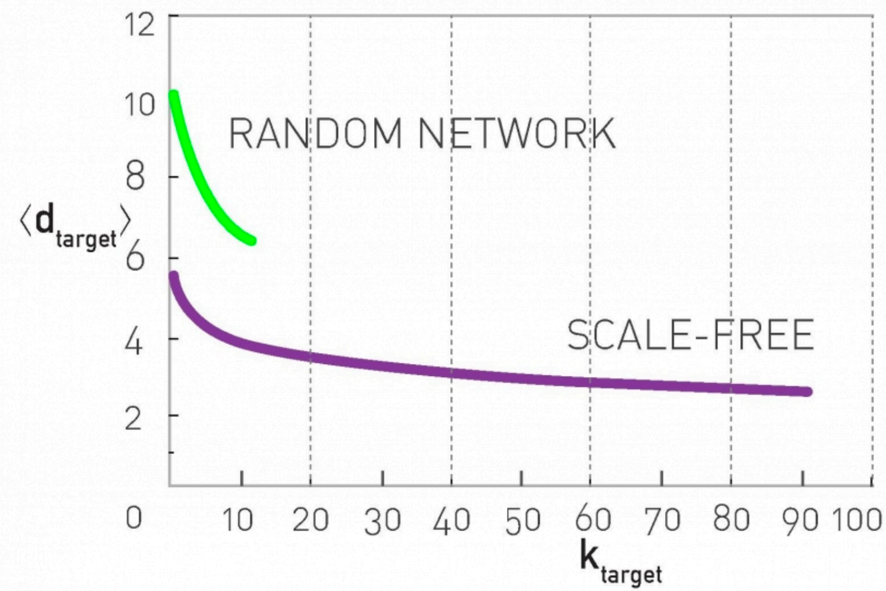


COMPORTAMIENTO DE MUNDO PEQUEÑO EN REDES SIN ESCALA



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

”Siempre es más fácil encontrar a alguien que conozca a una figura famosa o popular que a una persona insignificante.”
(Frigyes Karinthy, 1929)



Para generar redes libres de escala, Laszlo Barabási y Réka Albert sugirieron el algoritmo

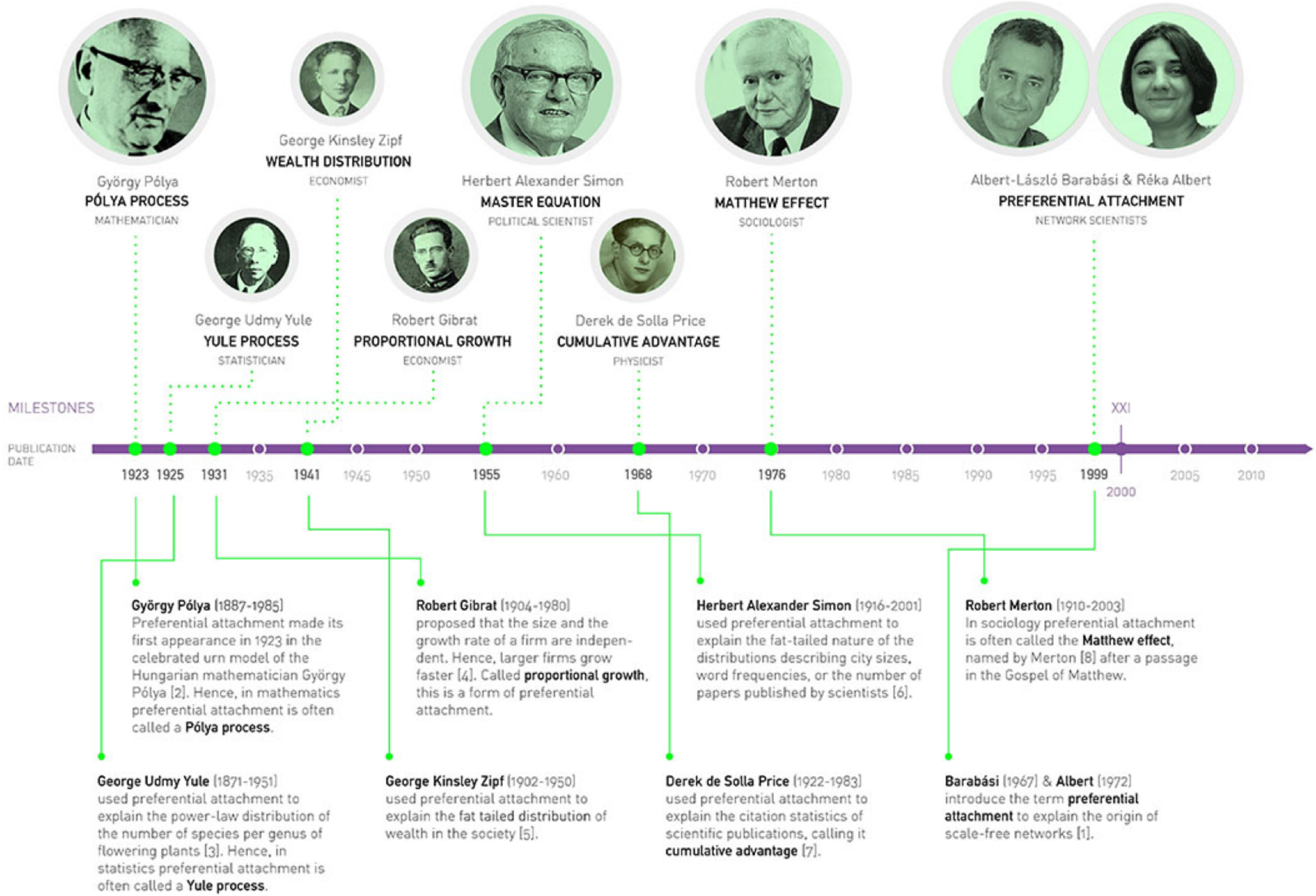
Barabási-Albert model of Growth and Preferential Attachment

Growth and Preferential Attachment

Crecimiento significa que tendremos una red en la que el número de nodos crecerá iterativamente en el tiempo hasta que alcancemos el tamaño de población deseado.

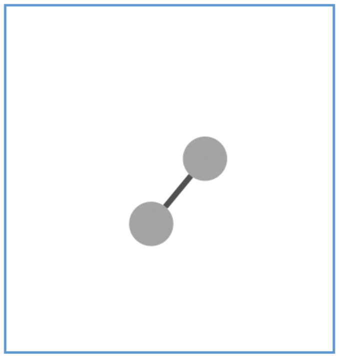
Los nuevos nodos se unirán a los nodos existentes, pero lo harán preferentemente a los más conectados.

Preferential Attachment



Growth and Preferential Attachment

Step 0



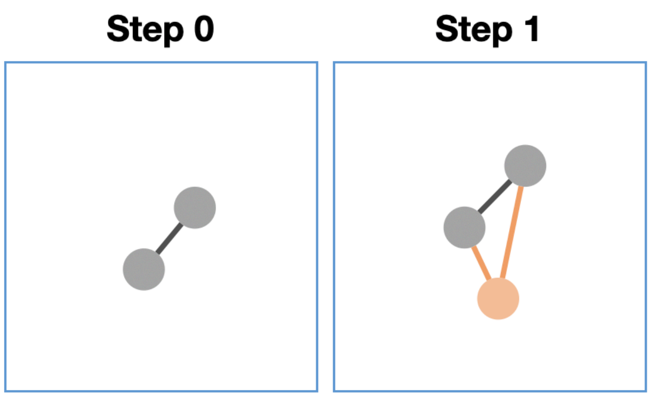
El algoritmo consta de los siguientes pasos:

- Comenzando con m0 nodos, los enlaces se eligen arbitrariamente, siempre que cada nodo tenga al menos un enlace.
- Para las N-m0 iteraciones, agregue un nodo nuevo que se conecte con m (< m0) enlaces a nodos preexistentes con probabilidad:

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

la probabilidad de que un enlace del nuevo nodo se conecte al nodo i depende del grado ki como

Growth and Preferential Attachment



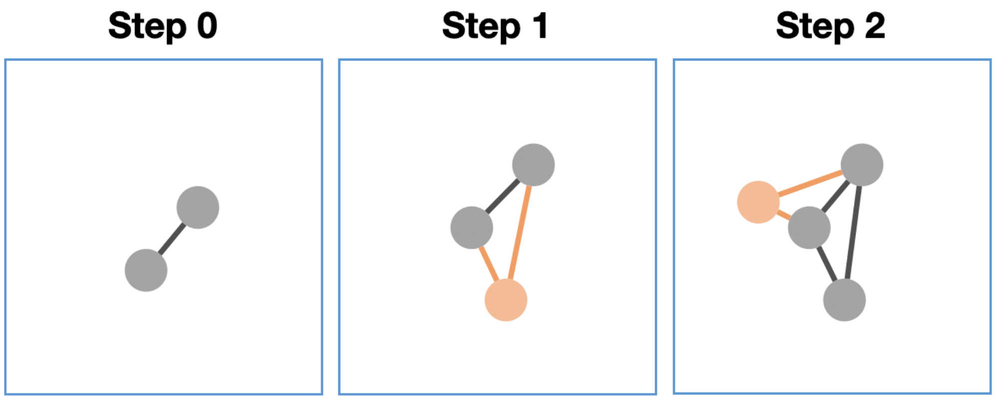
El algoritmo consta de los siguientes pasos:

- Comenzando con m0 nodos, los enlaces se eligen arbitrariamente, siempre que cada nodo tenga al menos un enlace.
- Para las N-m0 iteraciones, agregue un nodo nuevo que se conecte con m (< m0) enlaces a nodos preexistentes con probabilidad:

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

la probabilidad de que un enlace del nuevo nodo se conecte al nodo i depende del grado ki como

Growth and Preferential Attachment



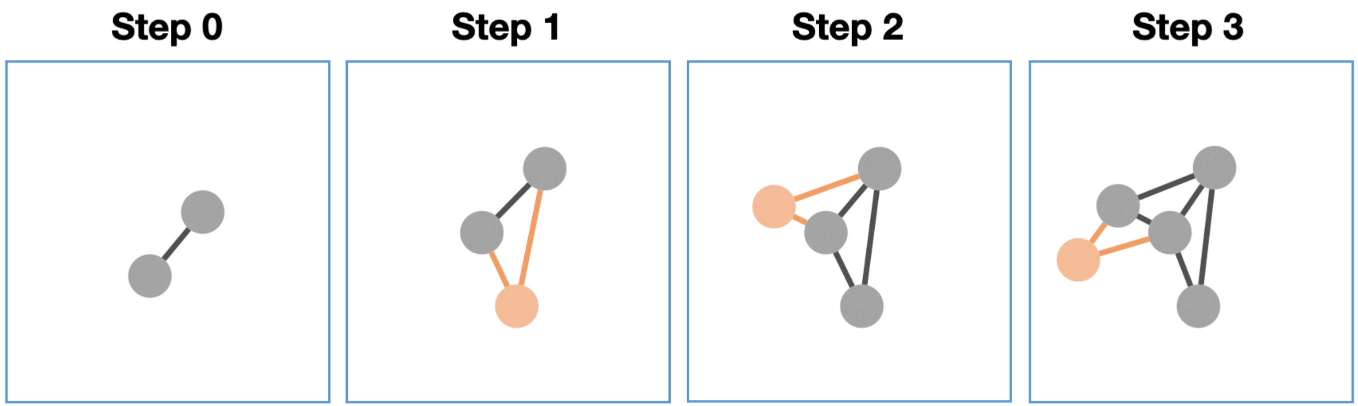
El algoritmo consta de los siguientes pasos:

- Comenzando con m0 nodos, los enlaces se eligen arbitrariamente, siempre que cada nodo tenga al menos un enlace.
- Para las N-m0 iteraciones, agregue un nodo nuevo que se conecte con m (< m0) enlaces a nodos preexistentes con probabilidad:

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

la probabilidad de que un enlace del nuevo nodo se conecte al nodo i depende del grado ki como

Growth and Preferential Attachment



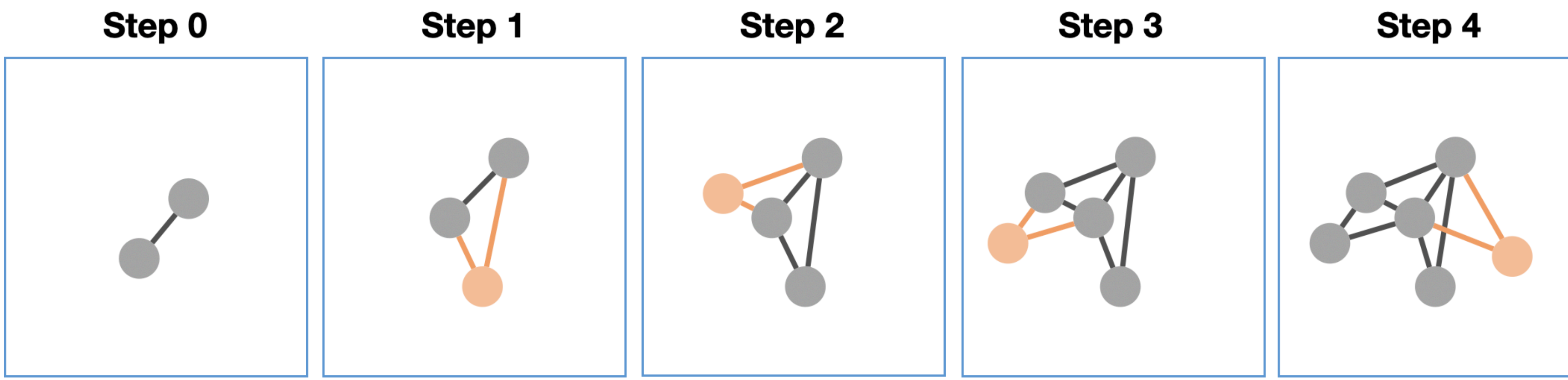
El algoritmo consta de los siguientes pasos:

- Comenzando con m0 nodos, los enlaces se eligen arbitrariamente, siempre que cada nodo tenga al menos un enlace.
- Para las N-m0 iteraciones, agregue un nodo nuevo que se conecte con m (< m0) enlaces a nodos preexistentes con probabilidad:

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

la probabilidad de que un enlace del nuevo nodo se conecte al nodo i depende del grado ki como

Growth and Preferential Attachment



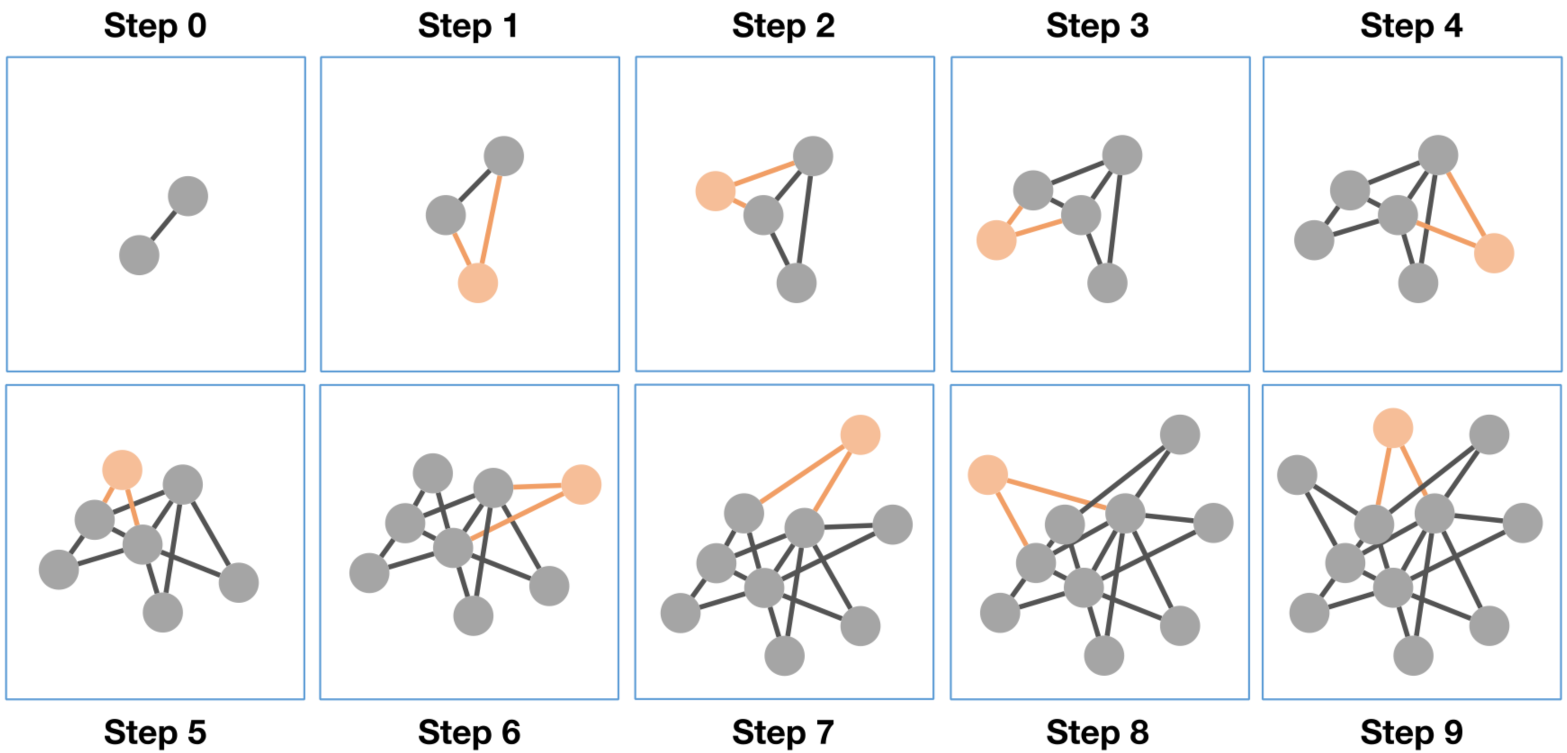
El algoritmo consta de los siguientes pasos:

- Comenzando con m0 nodos, los enlaces se eligen arbitrariamente, siempre que cada nodo tenga al menos un enlace.
- Para las N-m0 iteraciones, agregue un nodo nuevo que se conecte con m (< m0) enlaces a nodos preexistentes con probabilidad:

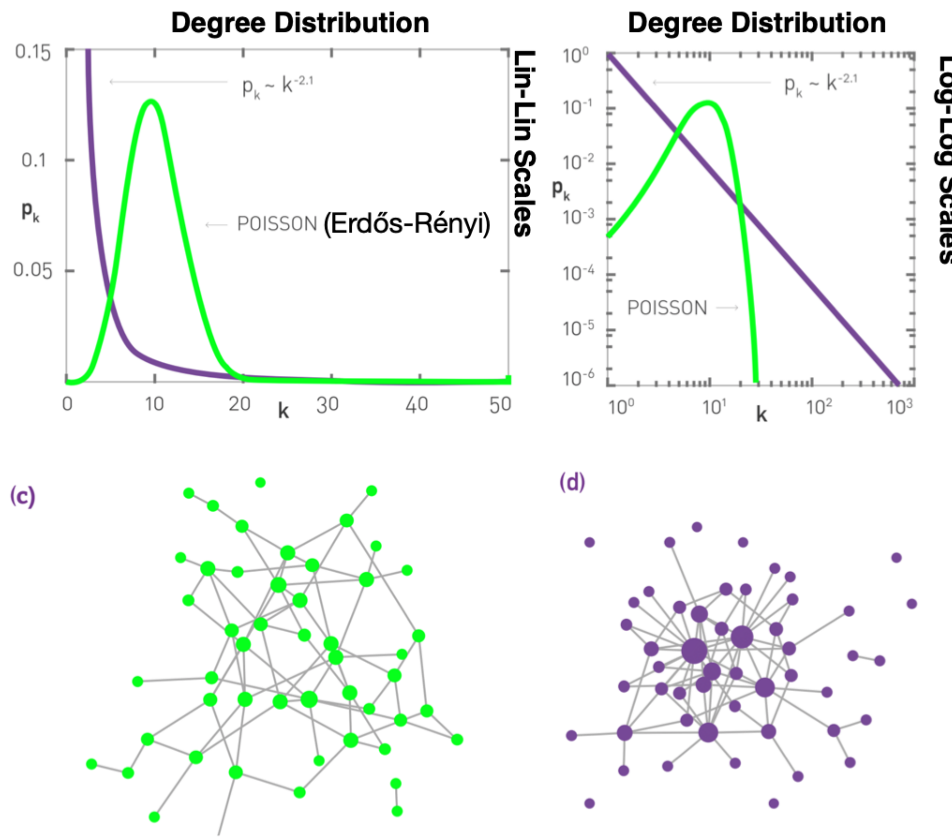
$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

la probabilidad de que un enlace del nuevo nodo se conecte al nodo i depende del grado ki como

Growth and Preferential Attachment



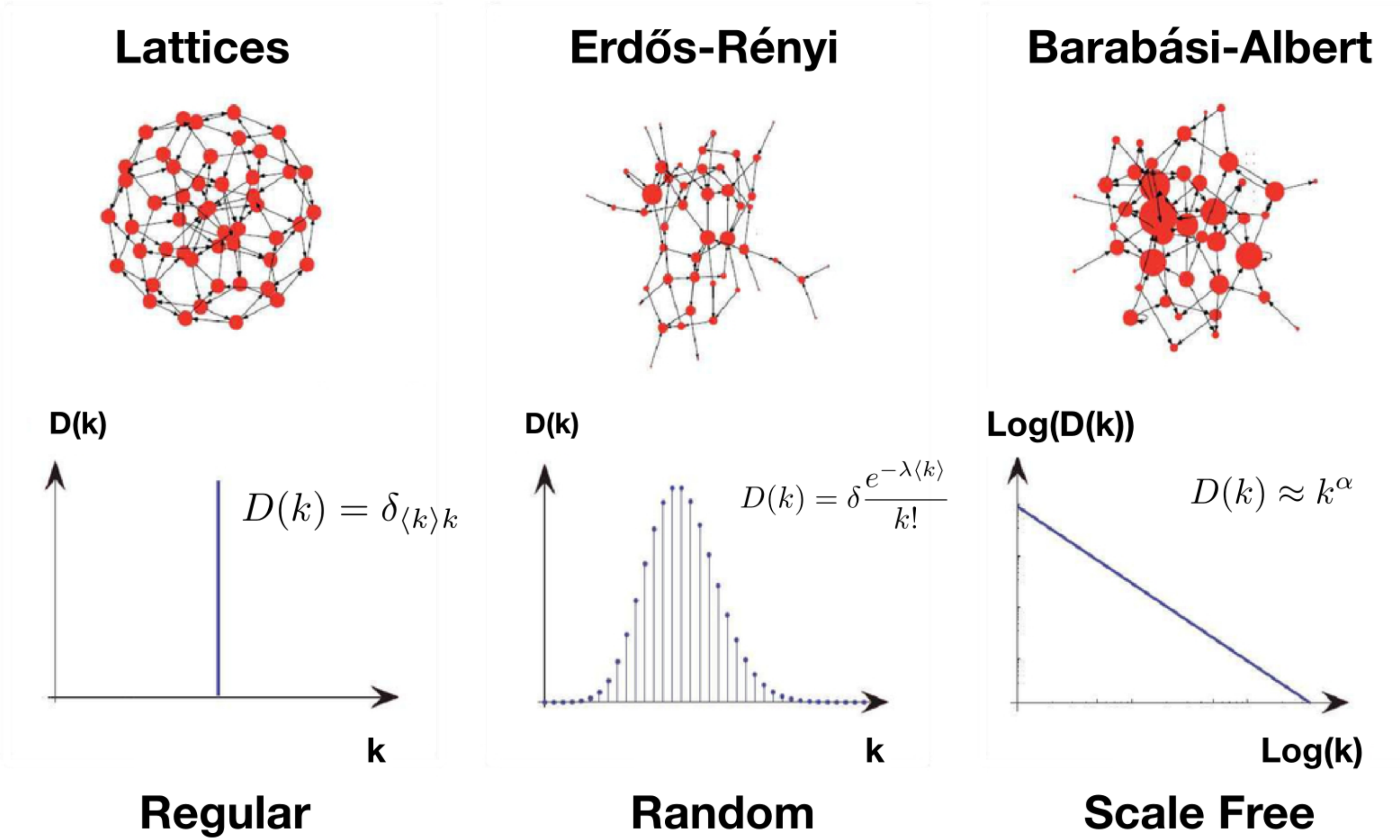
Scale-Free and Random Networks



(Erdős-Rényi)

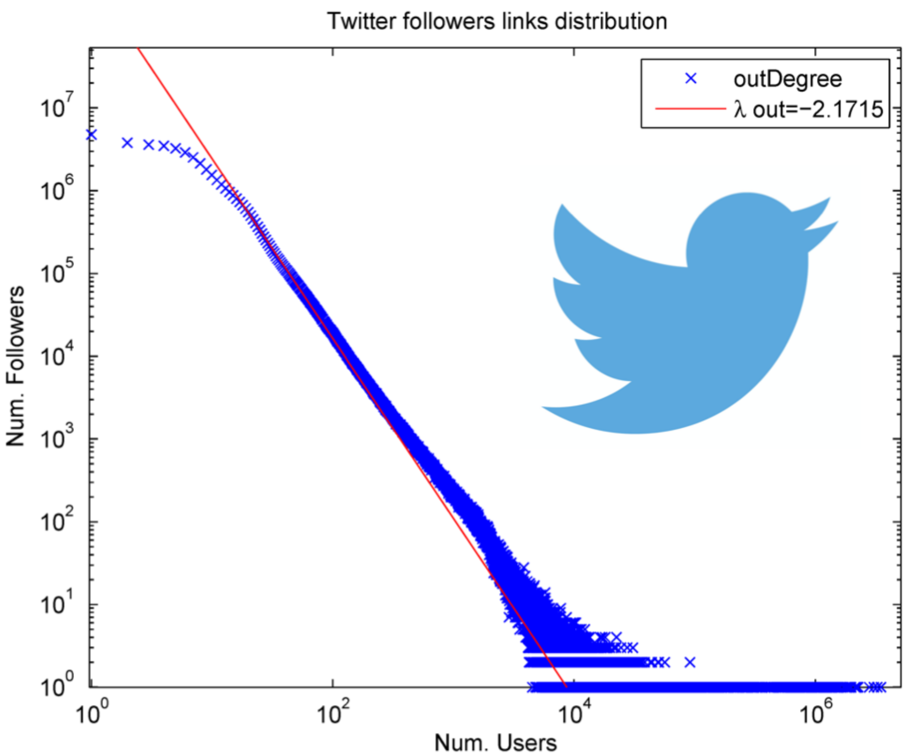
Scale-Free

¡Las redes sin escala tienen una distribución de grados que sigue una ley de potencia!



Lattices
Erdős-Rényi
Barabási-Albert
D(k)
$D(k) = \delta_{\langle k \rangle k}$
k
D(k)
$D(k) = \delta \frac{e^{-\lambda \langle k \rangle}}{k!}$
k
Log(D(k))
$D(k) \approx k^{\alpha}$
Log(k)
Regular
Random
Scale Free

Scale Free Networks



- La ubicuidad de la propiedad sin escala sugiere algún mecanismo subyacente fundamental;

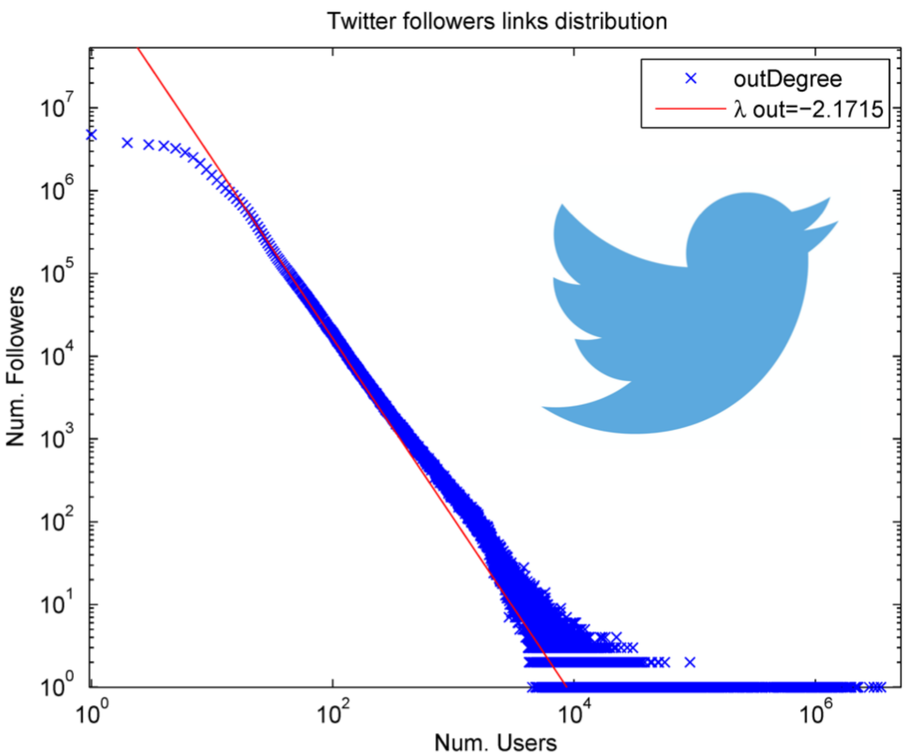
- Crecimiento y apego preferencial:
  - Crecimiento: se agregan nodos a la red con el tiempo;
  - Apego preferencial: estos nodos prefieren vincularse a nodos altamente conectados;

- El modelo simple de Barabási-Albert muestra que el crecimiento y el apego preferencial conducen a redes sin escala;

- Los ricos se hacen más ricos: los nodos altamente conectados tienden a seguir adquiriendo más conexiones

- Resultado claro del apego preferencial
  - En última instancia, conduce a una propiedad sin escala
  - Ventaja del primero: los nodos que se agregan antes tienen grados más altos

¿Es el modelo B.A. un buen modelo?

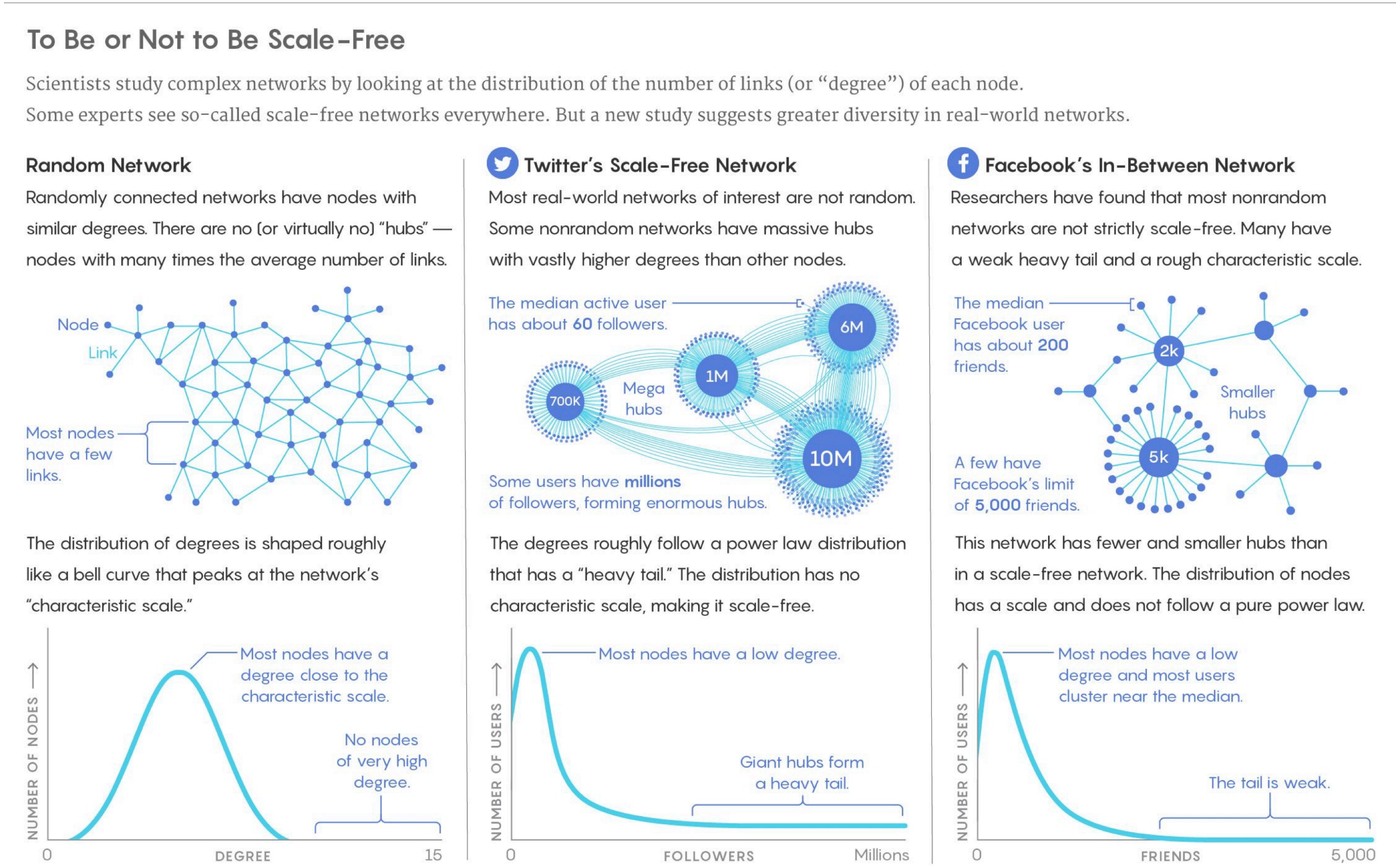


- Captura la intuición de que las redes reales crecen y que los nuevos nodos prefieren conectarse a nodos altamente conectados;
- Predice el comportamiento sin escala observado en redes reales;
- Que la preferencia sea directamente proporcional al número de enlaces parece simplista;
- La predicción de que los primeros nodos siempre están más conectados parece extraña;

Posibles extensiones:

- apego preferencial sublineal/superlineal
- métricas de fitness para nodos
- eliminación de nodos (edad fuera de la red)
- modelo mejor vinculado con la dinámica específica en cuestión.

“There is no general theory of networks”



https://www.quantamagazine.org/scant-evidence-of-power-laws-found-in-real-world-networks-20180215/?utm_content=buffer34d54&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Preguntas básicas

- Explica cómo se alcanza la propiedad de mundo pequeño en el modelo Watts-Strogatz.
- Qué es lo que explica el modelo Barabási-Albert en ciencia de redes?
- Compara y contrasta las distribuciones de grado de los tres modelos. De qué manera difieren y como se relacionan con la estructura de la red?