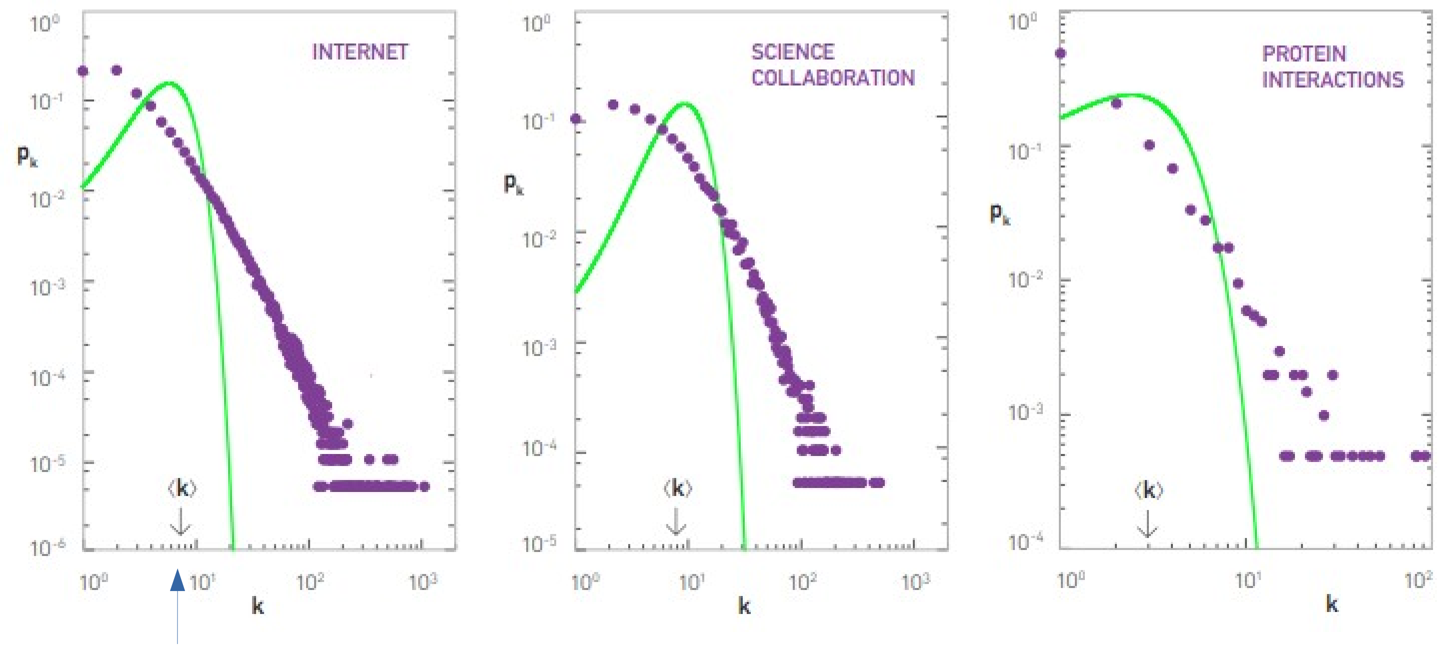
IIC-3641 Aprendizaje Automático basado en Grafos

https://github.com/marcelomendoza/IIC3641

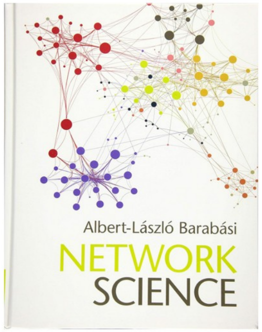
- RECAPITULACIÓN -

Las redes reales no son aleatorias (AF1)

Distribución de grado de redes reales



Grado promedio



Unidad 1: Modelos de grafos

Barabási, Albert-Lásló. Network Science. Cambridge University Press, 2016.

Pero son un objeto interesante de estudiar

Dinámica de redes aleatorias

En una red real el *p* no está fijo. Por ejemplo, en esta clase, antes de conocernos, la probabilidad de colaborar era muy baja. En la medida que nos conocemos, esto puede cambiar.
Si lo hacemos bien, el *p* debiera aumentar.

Veamos los casos extremos:

- Si $p = 0$, $\langle k \rangle = 0$. La componente conexa de mayor tamaño tiene $N_G = 1$.
- Si $p = 1$, $\langle k \rangle = N-1$. Entonces $N_G = N$.

Supongan que p aumenta gradualmente. ¿Qué creen que ocurre con el tamaño de la componente conexa más grande de la red?

a) Aumenta bruscamente, con tasa de crecimiento superior a *p*.
b) Aumenta gradualmente, con tasa de crecimiento proporcional a *p*.
c) Aumenta de forma no lineal a *p*.
d) Ni idea.

Hint. Recordar que:

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

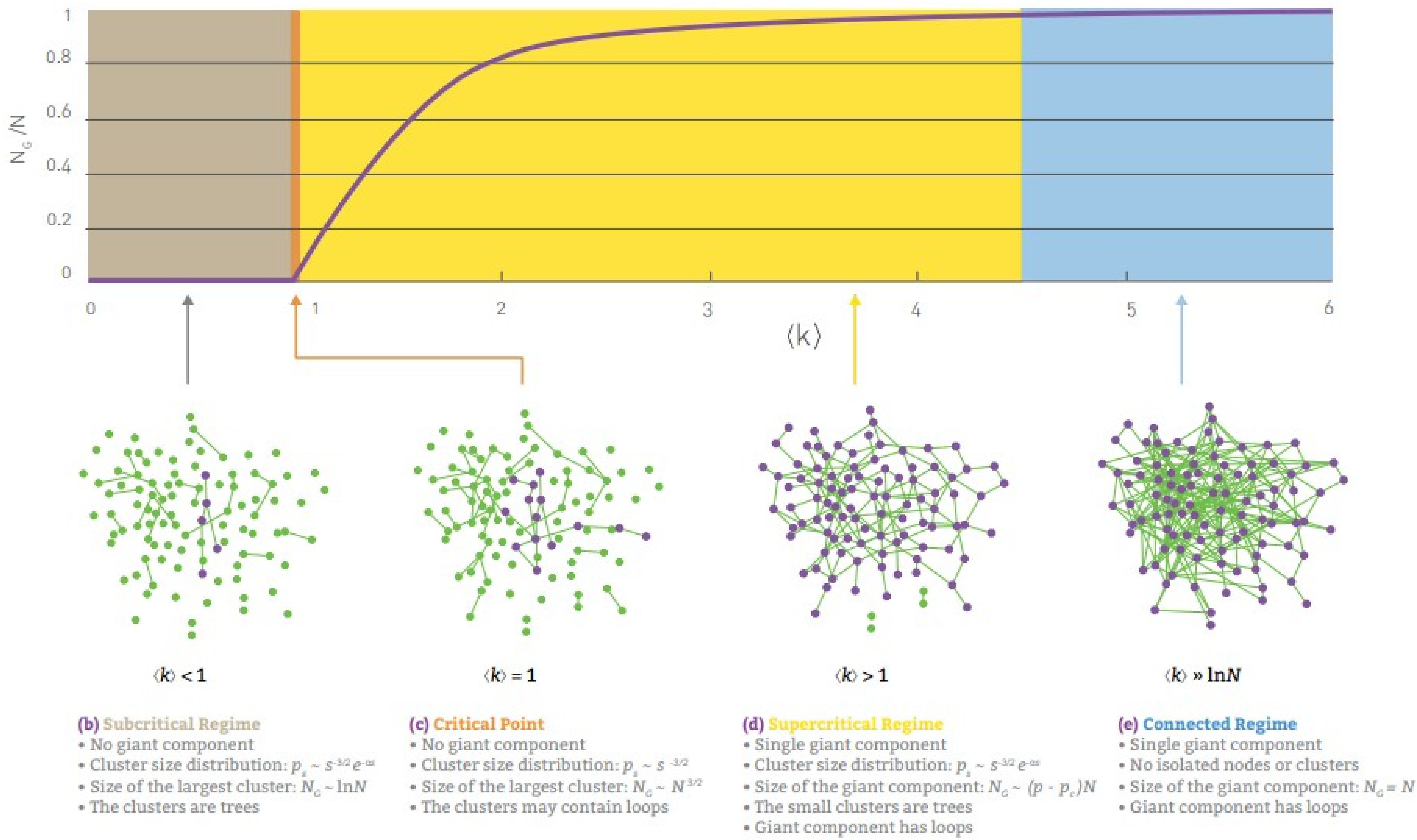
Fijo

Respondan aquí

https://app.wooclap.com/BKHKYC?from=status-bar

Dinámica de redes aleatorias

En realidad existe un valor crítico relacionado con <k>. Una vez que <k> excede el valor crítico, surge rápidamente una componente gigante. Erdös y Rényi mostraron que el valor crítico es <k> = 1.



P. Erdős and A. Rényi. On random graphs, I. Publicationes Mathematicae, 6:290-297, 1959

Dinámica de redes aleatorias

Notemos que como el valor crítico es <k> = 1, y dado que $\langle k\rangle = \frac{2\langle L\rangle}{N} = p(N-1).$ entonces:

$$p_c = \frac{1}{N-1} \approx \frac{1}{N},$$

es decir, en la medida que la red es más grande (en términos de N), el *p* crítico es menor.

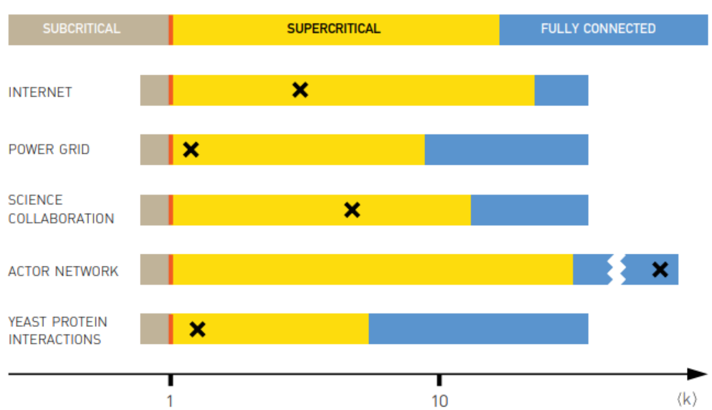
Dos propiedades de las redes aleatorias son muy relevantes para entender las redes reales:

1) El surgimiento de una componente gigante en una red aleatoria es un evento muy probable, dado que las condiciones que se requieren implican un *p* pequeño. Sólo una vez que *<k> > 1*, los nodos se organizarán como una red reconocible.

2) Si *<k> > ln N* todas las componentes de la red aleatoria son absorbidas por la componente gigante en una red conexa.

Por inspección se ha mostrado que la propiedad (1) se cumple en redes reales.

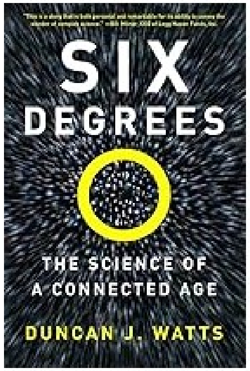
NETWORK	N	L	$\langle k\rangle$	lnN
Internet	192,244	609,066	6.34	12.17
Power Grid	4,941	6,594	2.67	8.51
Science Collaboration	23,133	94,439	8.08	10.05
Actor Network	702,388	29,397,908	83.71	13.46
Protein Interactions	2,018	2,930	2.90	7.61



P. Erdős and A. Rényi. On random graphs, I. Publicationes Mathematicae, 6:290-297, 1959

- SMALL WORLDS -

Fenómeno small worlds (seis grados de separación)



Fenómeno: Dos individuos cualesquiera en el mundo están conectados por un camino de a lo más largo seis en la red que indica quién conoce (en persona) a quién.

¿Por qué las distancias en redes reales son cortas?

Esto se debe a que la conectividad crece de forma exponencial en base $\langle k \rangle$.

En una red aleatoria, en promedio:

$\langle k \rangle$ nodos se encuentran a distancia 1.
$\langle k \rangle^2$ nodos se encuentran a distancia 2.
…
$\langle k \rangle^d$ nodos se encuentran a distancia d.

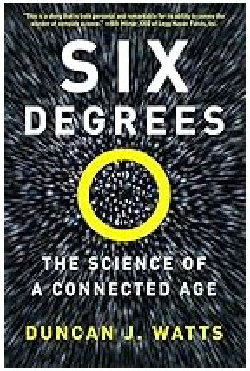
Es decir, el número esperado de nodos $N(d)$ hasta una distancia d está dado por:

$$N(d) \approx 1 + \langle k \rangle + \langle k \rangle^2 + \dots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1}.$$

Dado que $N(d)$ no debe exceder N, surge la noción de una d_{max} (diámetro). Podemos aproximar:

$$\langle k \rangle^{d_{max}} \approx N. \longrightarrow d_{max} \approx \frac{\ln N}{\ln \langle k \rangle}$$

Fenómeno small worlds (seis grados de separación)

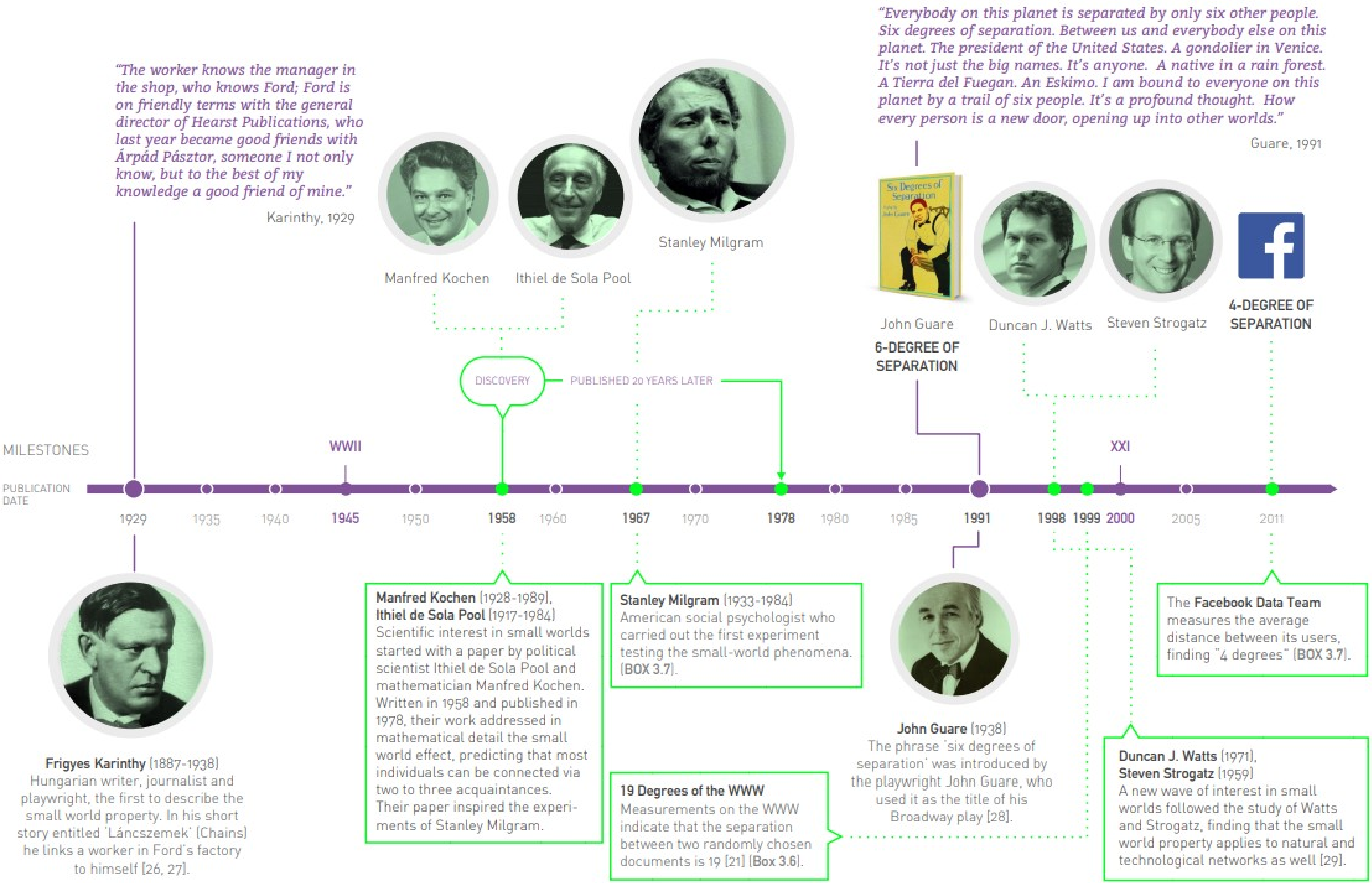


En redes reales ocurre esto:

NETWORK	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\frac{\ln N}{\ln \langle k \rangle}$
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	93,439	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

El diámetro es mayor que la predicción de seis grados de separación. **Sin embargo, si seleccionamos dos nodos al azar, en promedio la distancia entre ellos está bien aproximada por los grados de separación.**

Fenómeno small worlds (seis grados de separación)



Modelo Watts-Strogatz

El modelo Watts-Strogatz es una extensión del modelo Erdös-Rényi motivado por dos observacione

a) En redes reales la distancia promedio entre dos nodos depende logarítmicamente de N.

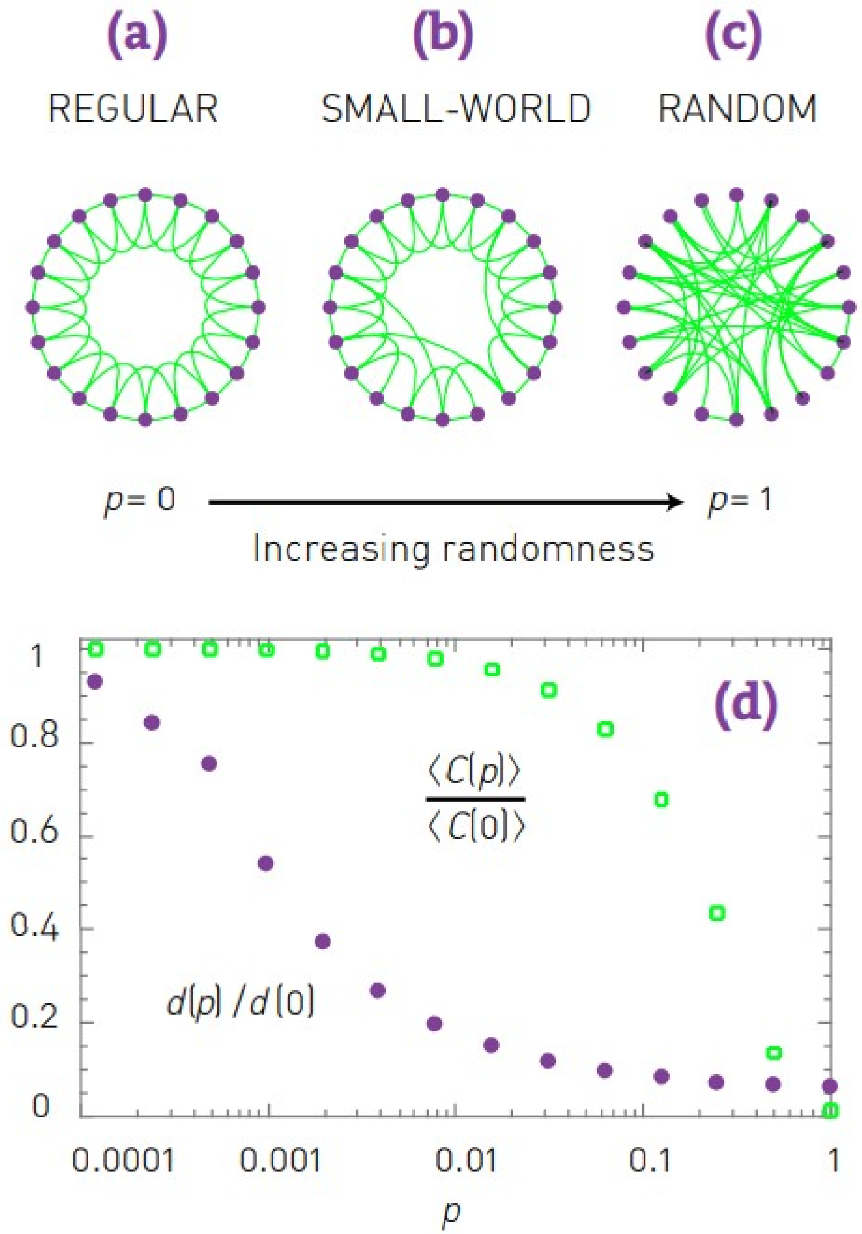
b) El coeficiente de clustering promedio de una red real es mucho mayor que el de una red aleatoria (AF

El modelo Watts-Strogatz interpola entre un lattice regular (que tiene un alto coeficiente de clustering) y una red aleatoria (que tiene bajo coeficiente de clustering pero exhibe el fenómeno small worlds).

Modelo Watts-Strogatz:

a) Comenzamos con un anillo de *N* nodos.
b) Con probabilidad *p* cada enlace es reconectado al azar con otro nodo.

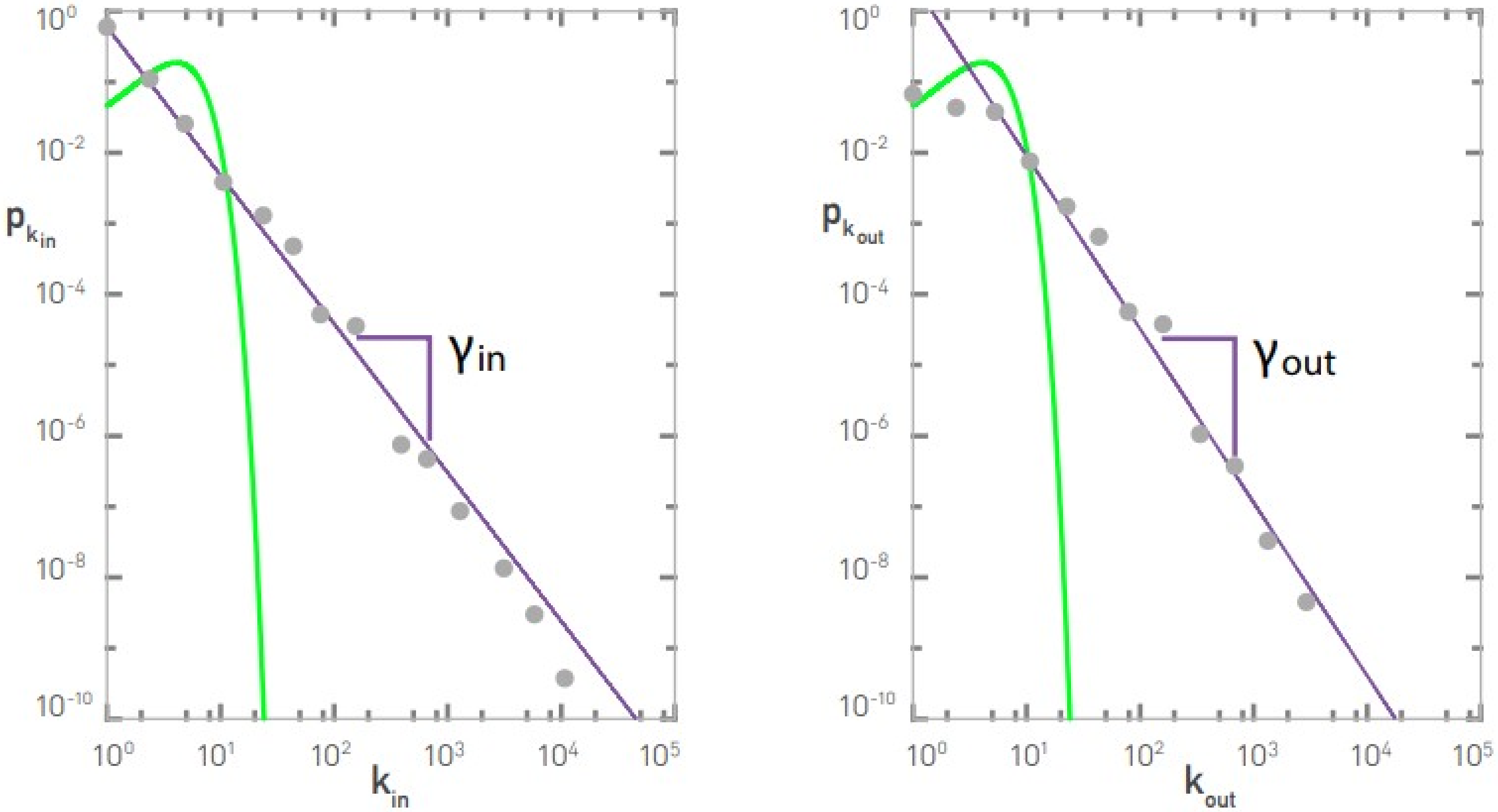
Notamos que si *p=1* la red se transforma en una red aleatoria. Se observa que *0.001 < p < 0.1* permite que existan redes con distancias pequeñas y alto coeficiente de clustering.



- REDES SCALE-FREE -

La propiedad scale-free

La distribución de grado de una red dirigida real como la WWW se ve así (muestra 1999):



Esto sugiere que la distribución de grado puede ser aproximada por una ley de potencia:

$$p_k \sim k^{-\gamma}.$$

Para que la distribución cumpla con $\int_{k_{min}}^{\infty} p(k)dk = 1$ se llega a: $\boxed{p(k) = (\gamma - 1)k_{min}^{\gamma-1}\, k^{-\gamma}}$

La propiedad scale-free

Una red scale-free es una red cuya distribución de grado sigue una ley de potencia.

Si una red es dirigida, la propiedad scale-free aplica tanto a la distribución de grado de entrada y de salida.

Las redes scale-free permiten el surgimiento de nodos hubs, los cuales ocurren en redes reales. Para saber cual es el grado de hub más grande una red, k_{max}, lo que llamamos el corte de la distribución de grado (en cola larga), veamos que ocurre con la distribución de grado de una exponencial:

$$p(k) = Ce^{-\lambda k}.$$

Aplicando la condición de normalización:

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

encontramos que $C = \lambda e^{\lambda k_{min}}$. Dado que en una red se espera encontrar un hub que cumpla con la condición de máxima conectividad (óptimo global), la probabilidad de observar ese nodo por medio de muestreo es 1/N, es decir:

$$\int_{k_{max}}^{\infty} p(k)dk = \frac{1}{N}.$$

La propiedad scale-free

Reemplazando *C* en *p(k)* e integrando, obtenemos k_{max}:

$$k_{\max} = k_{\min} + \frac{\ln N}{\lambda}.$$

Se observa que la dependencia de $k_{\max}$ de *N* es una dependencia logarítmica (sublineal). En redes aleatorias esta dependencia es aún más lenta.

Si lo calculamos para una red scale-free pasa lo siguiente:

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

Luego al aplicar la condición de corte:

$$\int_{k_{max}}^{\infty} p(k)dk = \frac{1}{N}.$$

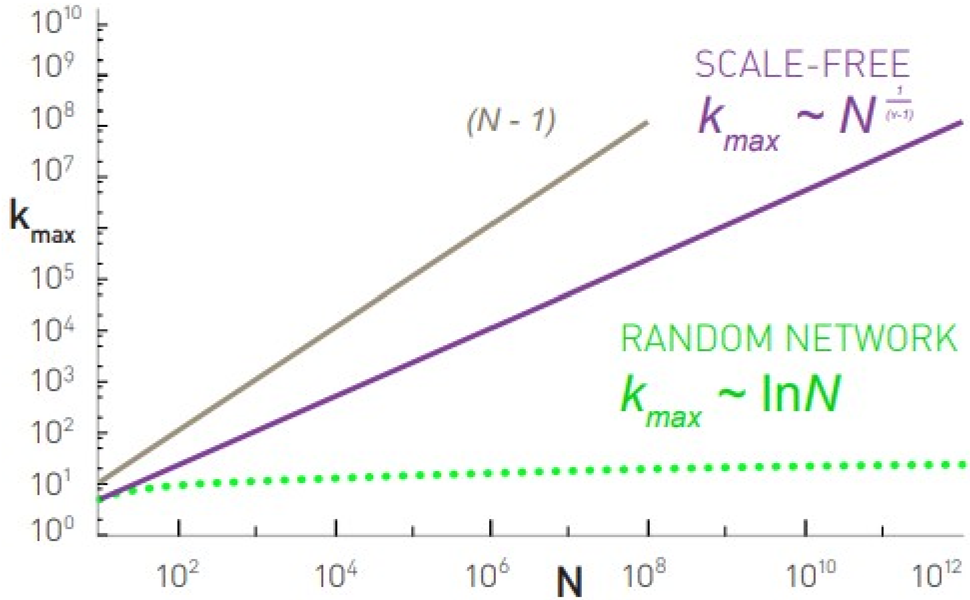
Llegamos a:

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

Es decir, mientras más grande sea la red (en términos de N) mayor será la conectividad de su hub más grande.

La propiedad scale-free

Las redes scale-free permiten el surgimiento de hubs de mayor conectividad que las redes aleatorias.



Las redes scale-free derivan su nombre de la propiedad del segundo momento de la distribución. Si la distribución fuera Gaussiana, el corte que determina el k_{max} de la red depende de la varianza. En una red scale-free el segundo momento es infinito si el factor de escala es < 3, es decir, las fluctuaciones en torno de $\langle k \rangle$ pueden ser arbitrariamente grandes.

Por ejemplo, para la WWW, $\langle k \rangle = 4.6$. Dado que el factor de escala es 2.1, el segundo momento es infinito y por tanto el grado de un nodo muestreado al azar es $k = 4.60 \pm \infty$ si $N \to \infty$.

La propiedad scale-free

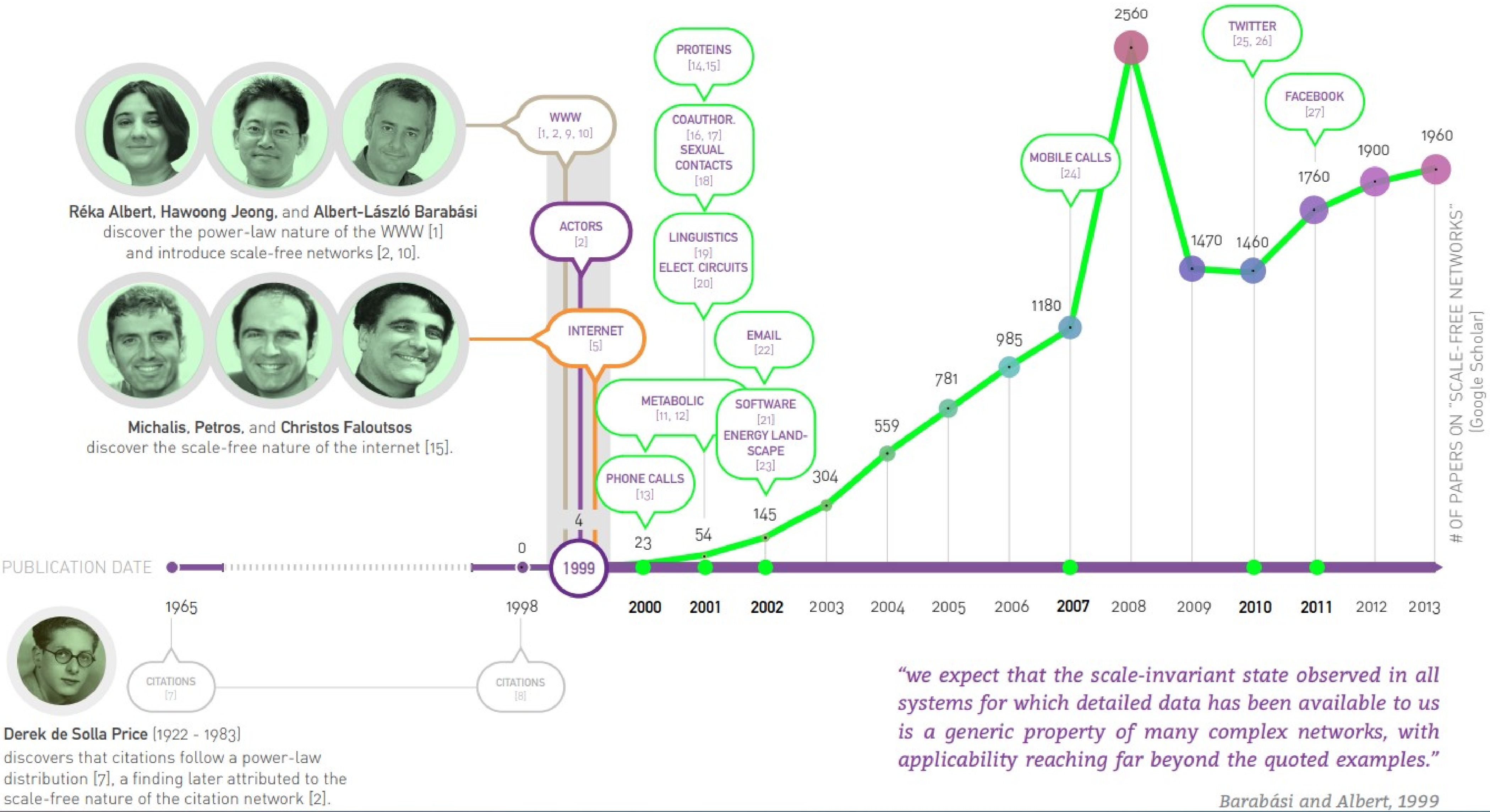
En la práctica el segundo momento no diverge en redes reales ya que *N* no tiende a infinito. Pero si ayuda a explicar el surgimiento de los hubs de gran conectividad.

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	93,439	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*	2.9 0*	-
Protein Interactions	2,018	2,930	2.9 0	-	-	32.3	-	-	2.89*

Observemos que la exponente de la distribución de grado no es menor que 1.

La propiedad scale-free

Muchas redes son scale-free.



La dependencia entre <d> y el exponente de la distribución de grado

Cohen & Cavlin (2003) encontraron una relación de dependencia entre la distancia promedio y el exponente de la distribución de grado:

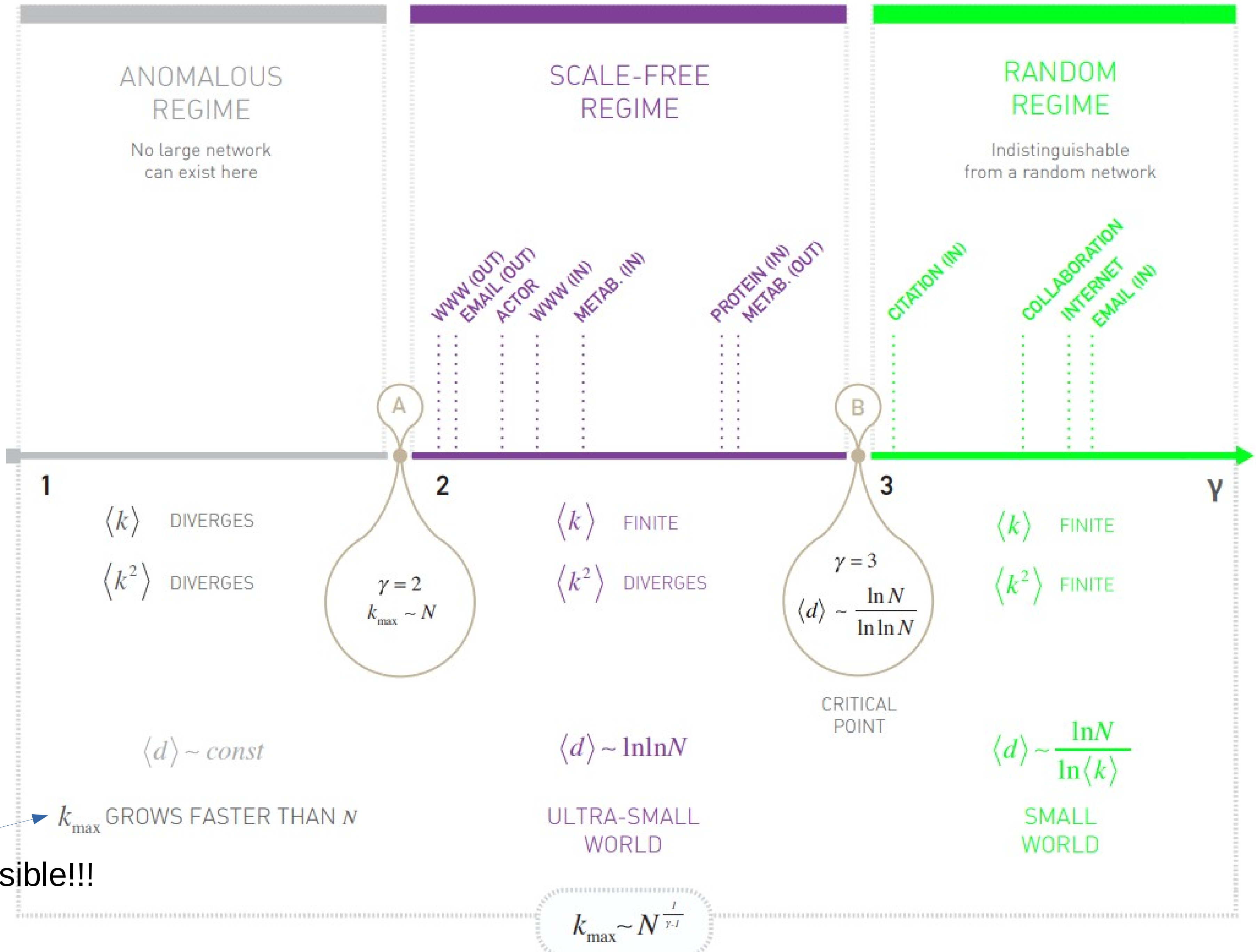
$$\langle d \rangle \sim \begin{cases} \text{const.} & \gamma = 2 \\ \ln \ln N & 2 < \gamma < 3 \qquad \longleftarrow \quad \text{Ultra small-world} \\ \dfrac{\ln N}{\ln \ln N} & \gamma = 3 \\ \ln N & \gamma > 3 \end{cases}$$

Esta relación define regímenes (o fases) de las redes scale-free.

- Ultra small-world: la distancia promedio crece más lento que en una red aleatoria.

R. Cohen and S. Havlin. Scale free networks are ultrasmall. Phys. Rev. Lett. 90, 058701, 2003.

Las fases de una red scale-free



Imposible!!!