Redes y Sistemas Complejos

Cuarto Curso del Grado en Ingeniería Informática

Tema 5: Modelos de Redes

5.2. Mundos Pequeños

Oscar Cordón García

Dpto. Ciencias de la Computación e Inteligencia Artificial

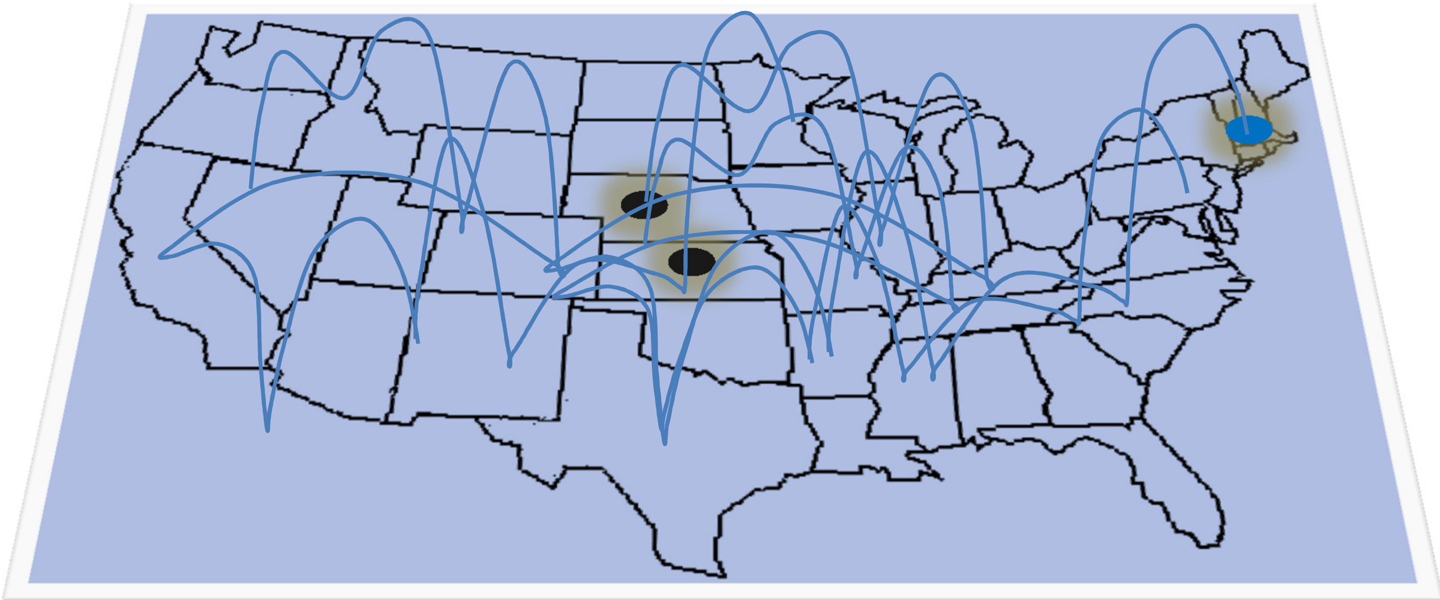
ocordon@decsai.ugr.es

EL VOLUMEN EXPLOSIVO DE LAS REDES COMPLEJAS

EXPERIMENTO DE MILGRAM:

Las primeras cartas en cadena… Orígenes: Omaha, Nebraska y Wichita, Kansas, EEUU. Destino: Boston y Sharon, Massachusetts, EEUU

Se enviaron 296 cartas. La primera llegó en pocos días, pasando sólo por 2 enlaces. Al final llegaron 64 con un máximo de 12 intermediarios



SEIS GRADOS DE SEPARACIÓN:

La media de intermediarios fue de entre 5.5 y 6

Travers y Milgram, Sociometry 32, 425 (1969)

INTERPRETACIÓN DEL EXPERIMENTO DE MILGRAM

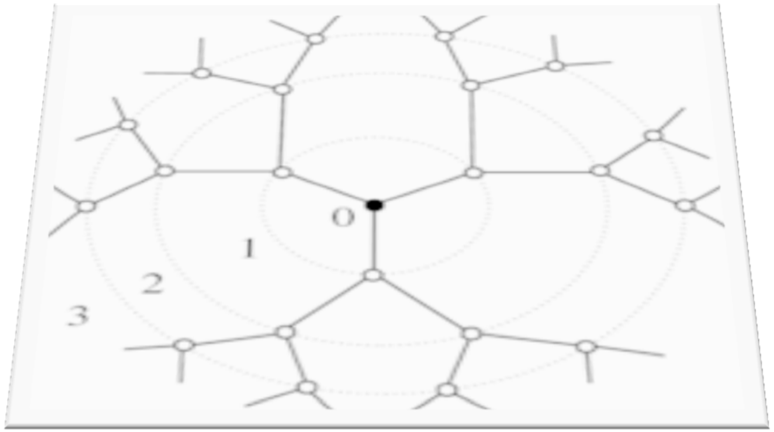
- ¿Es 6 un **resultado sorprendente** para el experimento?
    - ¿En los 60? ¿Hoy en día? ¿Por qué?

- Si la red social mundial fuera una red puramente aleatoria…
    - **Pool y Kochen (1978)**: Cada persona tiene unos 1000 amigos, variando en 500-1500
    - Mínimo ~ 500 elecciones para el primer enlace
    - ~ 500^2 = 250,000 vecinos potenciales para el segundo grado
    - ~ 500^3 = 125,000,000 vecinos potenciales para el tecer grado

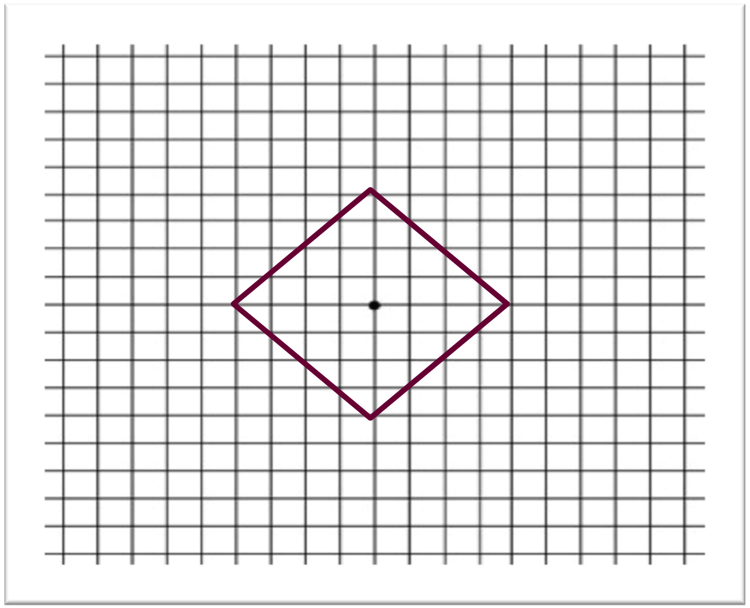
- ¿Y si las redes complejas fueran completamente "***cliquish***" (subgrafos totalmente conectados (clique): coeficiente medio de clustering $\rightarrow$ 1)?
    - Todos los amigos de mis amigos serían mis amigos
    - ¿Qué ocurriría?

EL VOLUMEN EXPONENCIAL DE LAS REDES COMPLEJAS (1)

El secreto que subyace al efecto de mundos pequeños se encuentra en el volumen de la red



$$S(d) = 4d$$

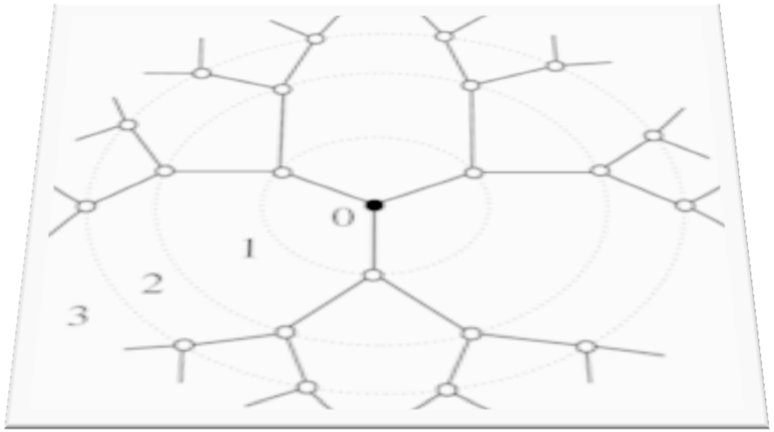


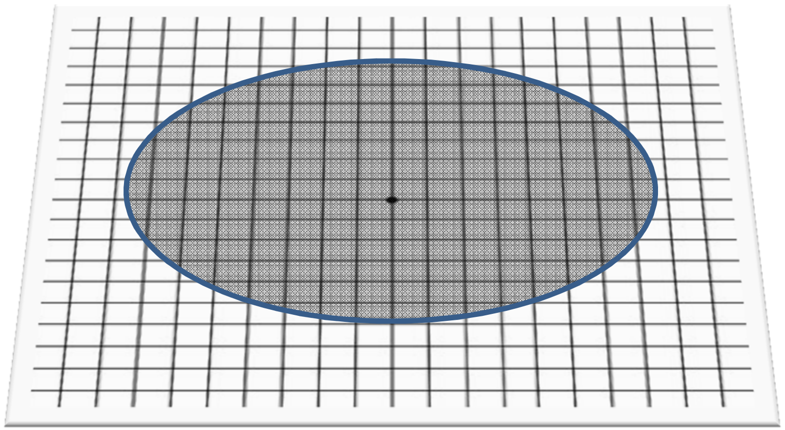
EL VOLUMEN EXPONENCIAL DE LAS REDES COMPLEJAS (2)

El secreto que subyace al efecto de mundos pequeños se encuentra en el volumen de la red

$$N(d) = \sum_{x=1}^{d} 4x = 2d(d+1) \sim d^2$$

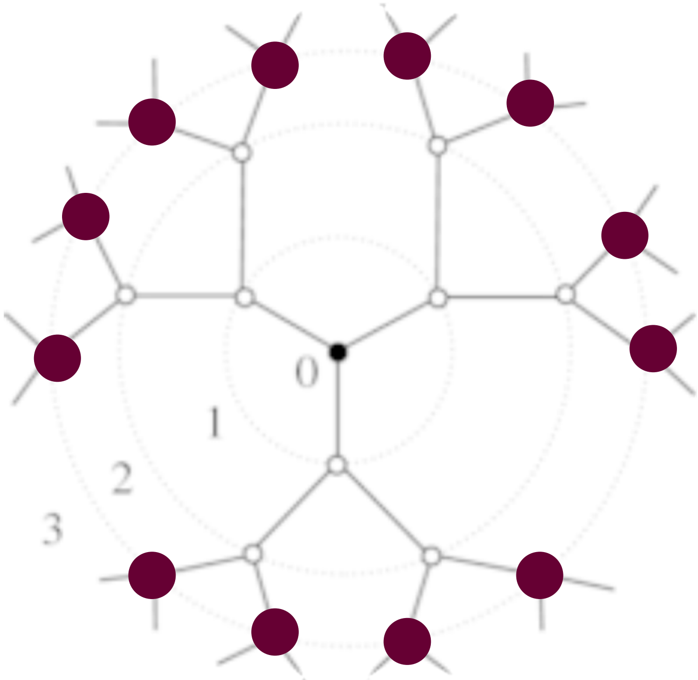
Crecimiento polinomial





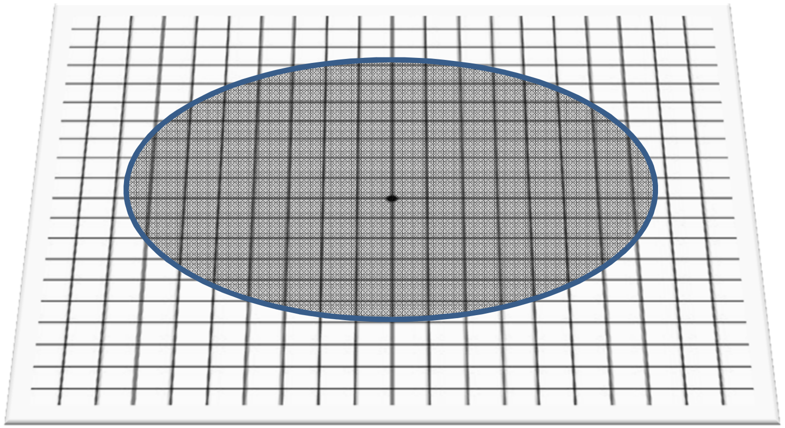
EL VOLUMEN EXPONENCIAL DE LAS REDES COMPLEJAS (3)

El secreto que subyace al efecto de mundos pequeños se encuentra en el volumen de la red



$$N(d) = \sum_{x=1}^{d} 4x = 2d(d+1) \sim d^2$$

Crecimiento polinomial



EL VOLUMEN EXPONENCIAL DE LAS REDES COMPLEJAS (4)

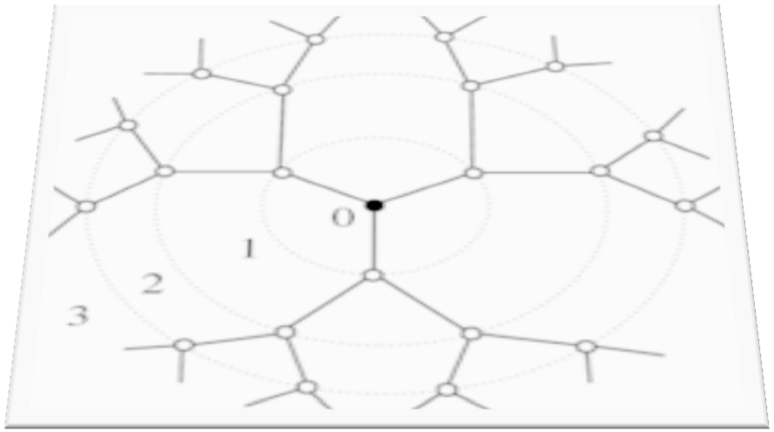
El secreto que subyace al efecto de mundos pequeños se encuentra en el volumen de la red

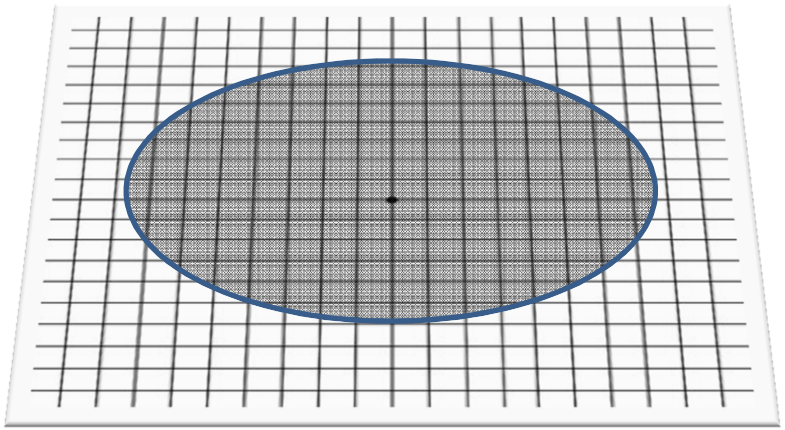
$$N(d) = \sum_{x=1}^{d} \langle k \rangle^{x} = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} \sim \langle k \rangle^{d}$$

$$N(d) = \sum_{x=1}^{d} 4x = 2d(d+1) \sim d^{2}$$

Crecimiento exponencial

Crecimiento polinomial



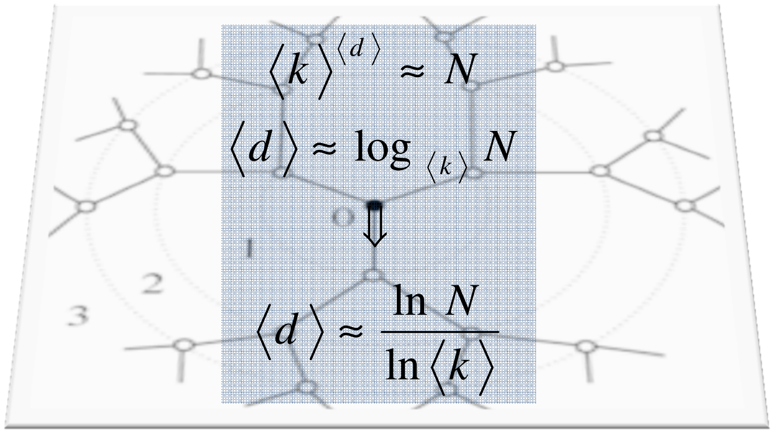


EL VOLUMEN EXPONENCIAL DE LAS REDES COMPLEJAS (5)

El secreto que subyace al efecto de mundos pequeños se encuentra en el volumen de la red

$$N(d) = \sum_{x=1}^{d} \langle k \rangle^{x} = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} \sim \langle k \rangle^{d}$$

Crecimiento exponencial



$$\langle k \rangle^{\langle d \rangle} \approx N$$

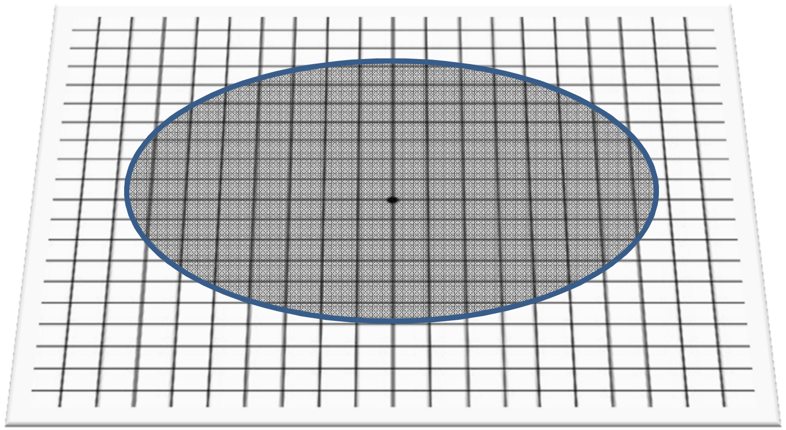
$$\langle d \rangle \approx \log_{\langle k \rangle} N$$

$$\Downarrow$$

$$\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle}$$

$$N(d) = \sum_{x=1}^{d} 4x = 2d(d+1) \sim d^{2}$$

Crecimiento polinomial

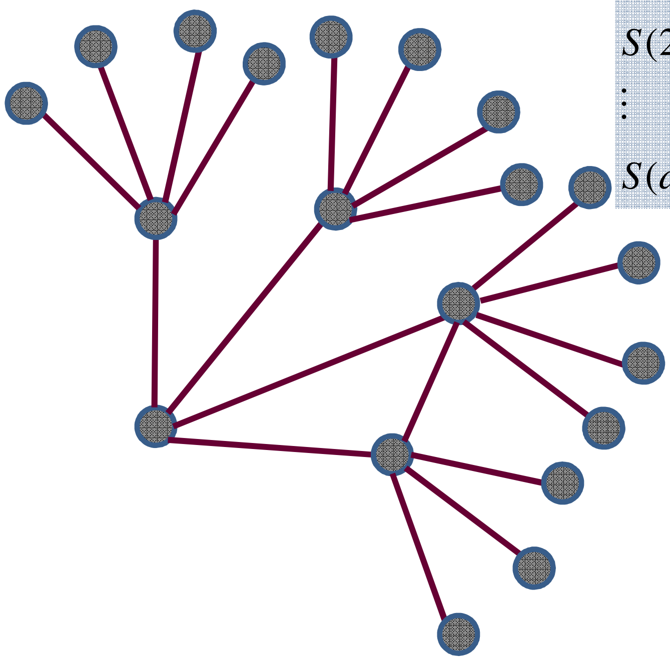


LAS REDES COMPLEJAS NO SON ÁRBOLES (1)

Crecimiento exponencial: $$\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle}$$

El clustering inhibe la propiedad de mundos pequeños, reduce el volumen exponencial

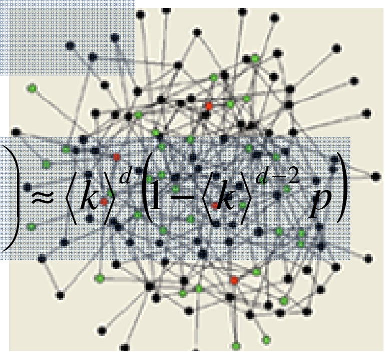
Algunos de los vecinos de tus vecinos también son tus vecinos



$$S(1) = \langle k \rangle$$
$$S(2) < \langle k \rangle^2$$
$$\vdots$$
$$S(d) < \langle k \rangle^d$$

$$S(0) = 1$$
$$S(1) = \langle k \rangle$$
$$S(2) = \langle k \rangle^2 \left(\frac{N - \langle k \rangle - 1}{N} \right) \approx \langle k \rangle^2 (1 - p)$$
$$S(3) = \langle k \rangle S(2) \left(\frac{N - \langle k \rangle^2 (1 - p) - \langle k \rangle}{N} \right) \approx \langle k \rangle^3 (1 - \langle k \rangle p)$$
$$\vdots$$

$$S(d) = \langle k \rangle S(d-1) \left(1 - \frac{S(d-1) + S(d-2)}{N} \right) \approx \langle k \rangle^d \left(1 - \langle k \rangle^{d-2} p \right)$$

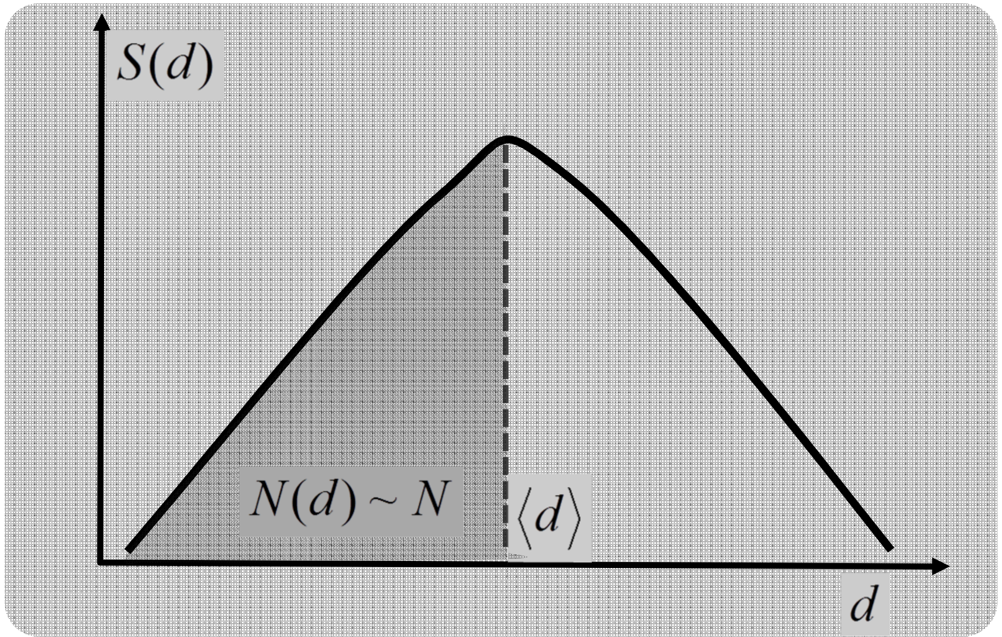


LAS REDES COMPLEJAS NO SON ÁRBOLES (2)

Crecimiento exponencial: $\langle d \rangle \approx \dfrac{\ln N}{\ln \langle k \rangle}$

El crecimiento exponencial continúa mientras $N(d) < N$

$$S(d) \approx \langle k \rangle^{d} \left(1 - \langle k \rangle^{d-2} p\right) = \langle k \rangle^{d} \left(1 - \frac{\langle k \rangle^{d-1}}{N}\right)$$

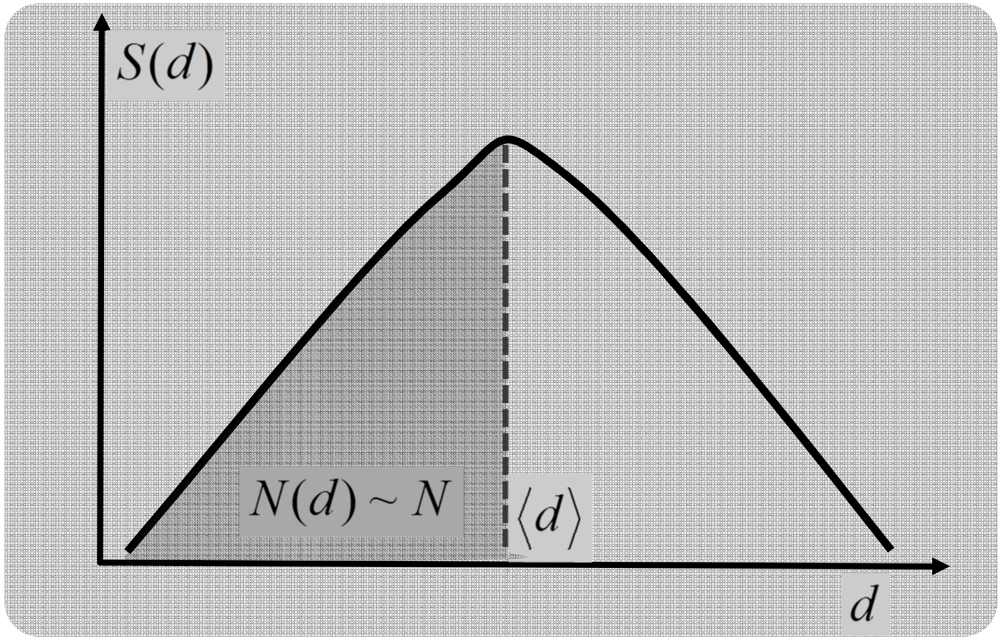


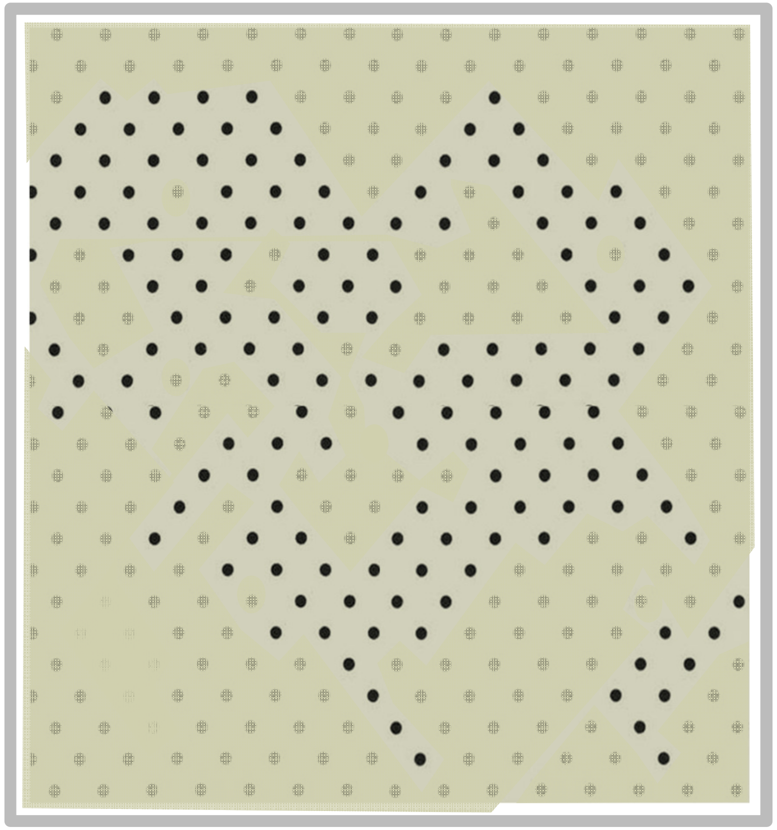
LAS REDES COMPLEJAS NO SON ÁRBOLES (3)

Crecimiento exponencial: $\langle d \rangle \approx \dfrac{\ln N}{\ln \langle k \rangle}$

El crecimiento exponencial continúa mientras $N(d) < N$

$$S(d) \approx \langle k \rangle^d \left(1 - \langle k \rangle^{d-2} p\right) = \langle k \rangle^d \left(1 - \frac{\langle k \rangle^{d-1}}{N}\right)$$



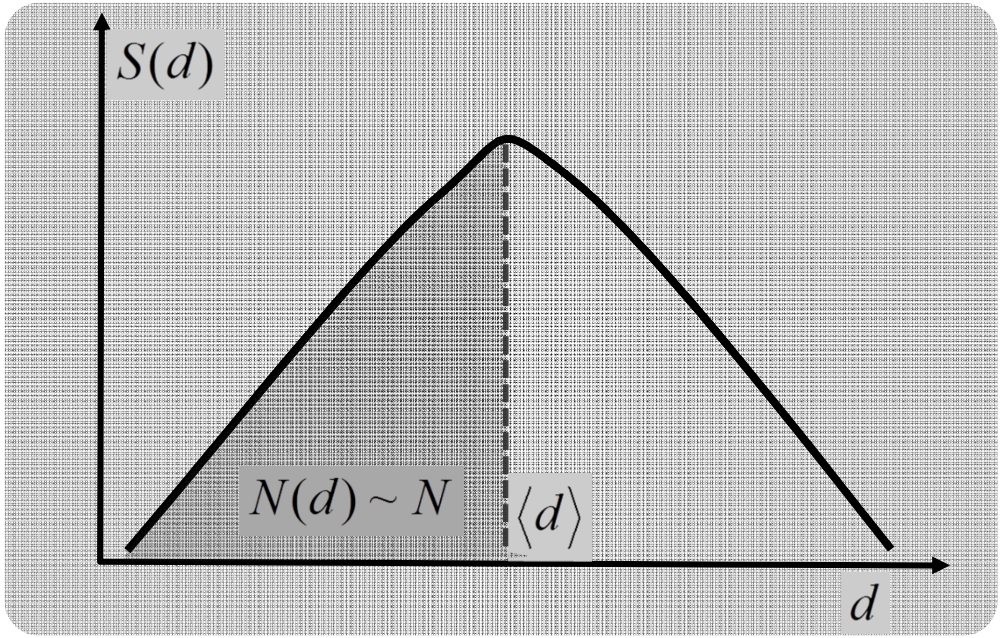


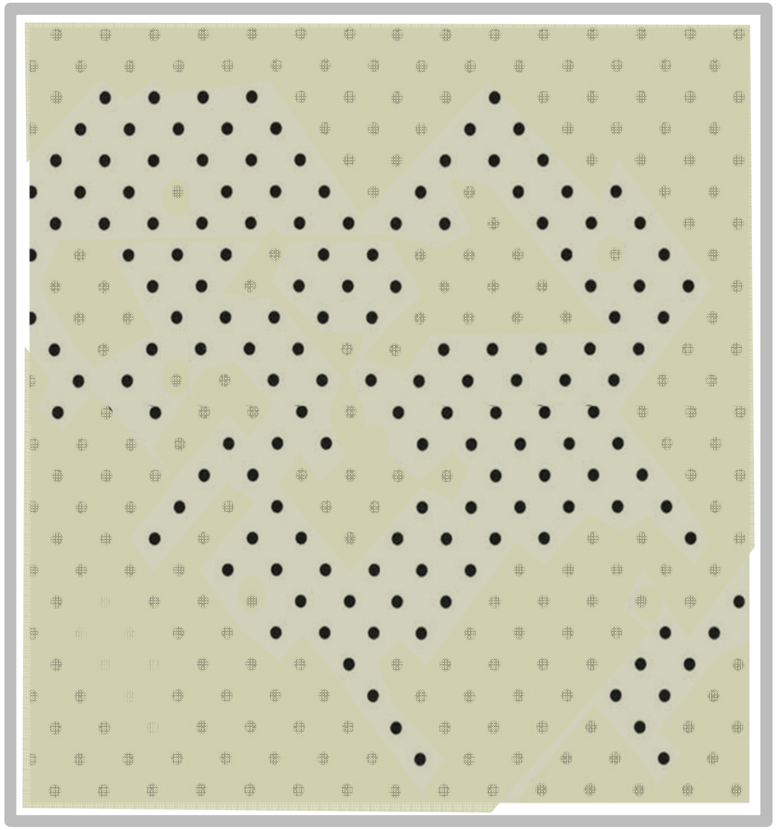
LAS REDES COMPLEJAS NO SON ÁRBOLES (4)

Crecimiento exponencial: $\langle d \rangle \approx \dfrac{\ln N}{\ln \langle k \rangle}$

El crecimiento exponencial continúa mientras $d \leq \langle d \rangle$

$$S(d) \approx \langle k \rangle^d \left(1 - \langle k \rangle^{d-2} p\right) = \langle k \rangle^d \left(1 - \frac{\langle k \rangle^{d-1}}{N}\right)$$





CLUSTERING versus ALEATORIEDAD

Concepto

Una red puede ser un mundo pequeño en tanto en cuanto el clustering pueda ignorarse

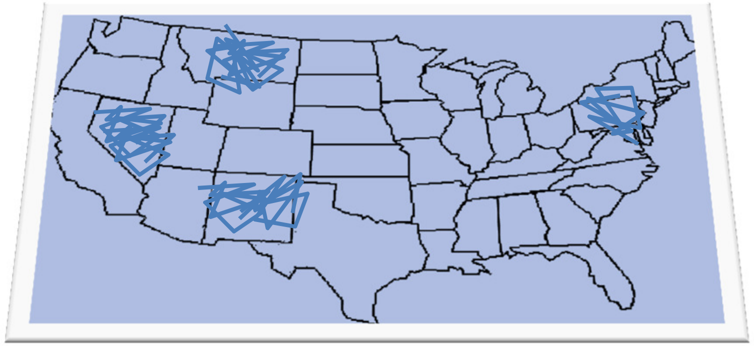
¿Dónde se localizaría la red social?

Red clusterizada

Red aleatoria

CLUSTERING versus ALEATORIEDAD

Interpretación del Clustering (1)

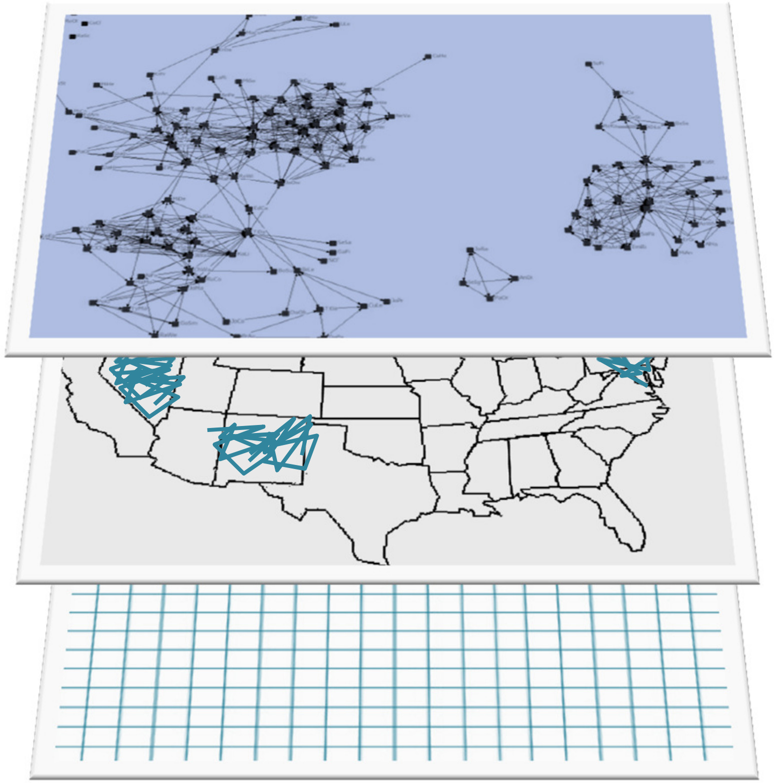


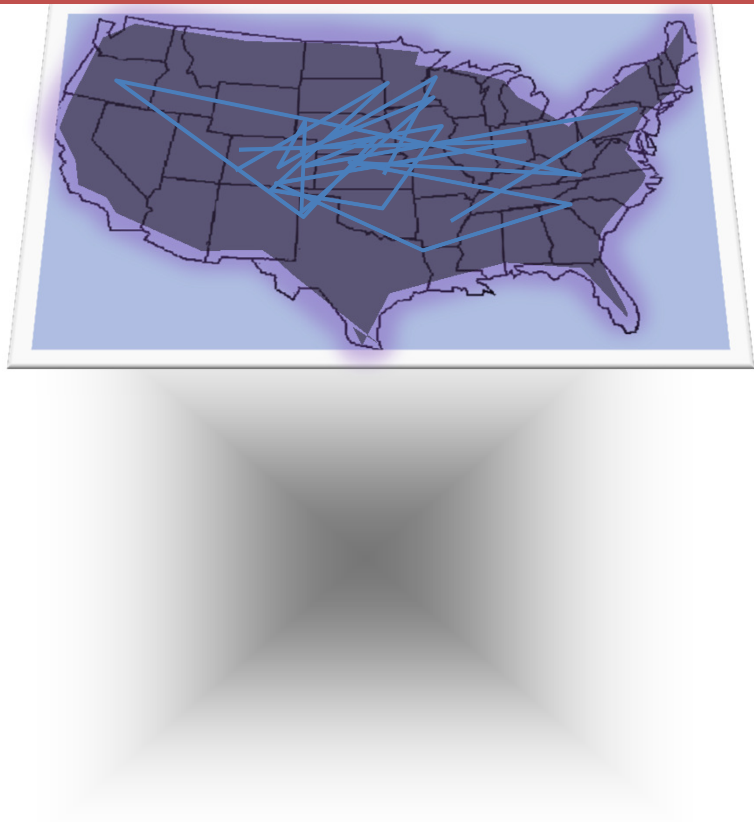
Localmente estructurada

Puramente aleatoria

CLUSTERING versus ALEATORIEDAD | **Interpretación del Clustering (2)**

El Clustering implica localidad ⟷ La aleatoriedad permite atajos



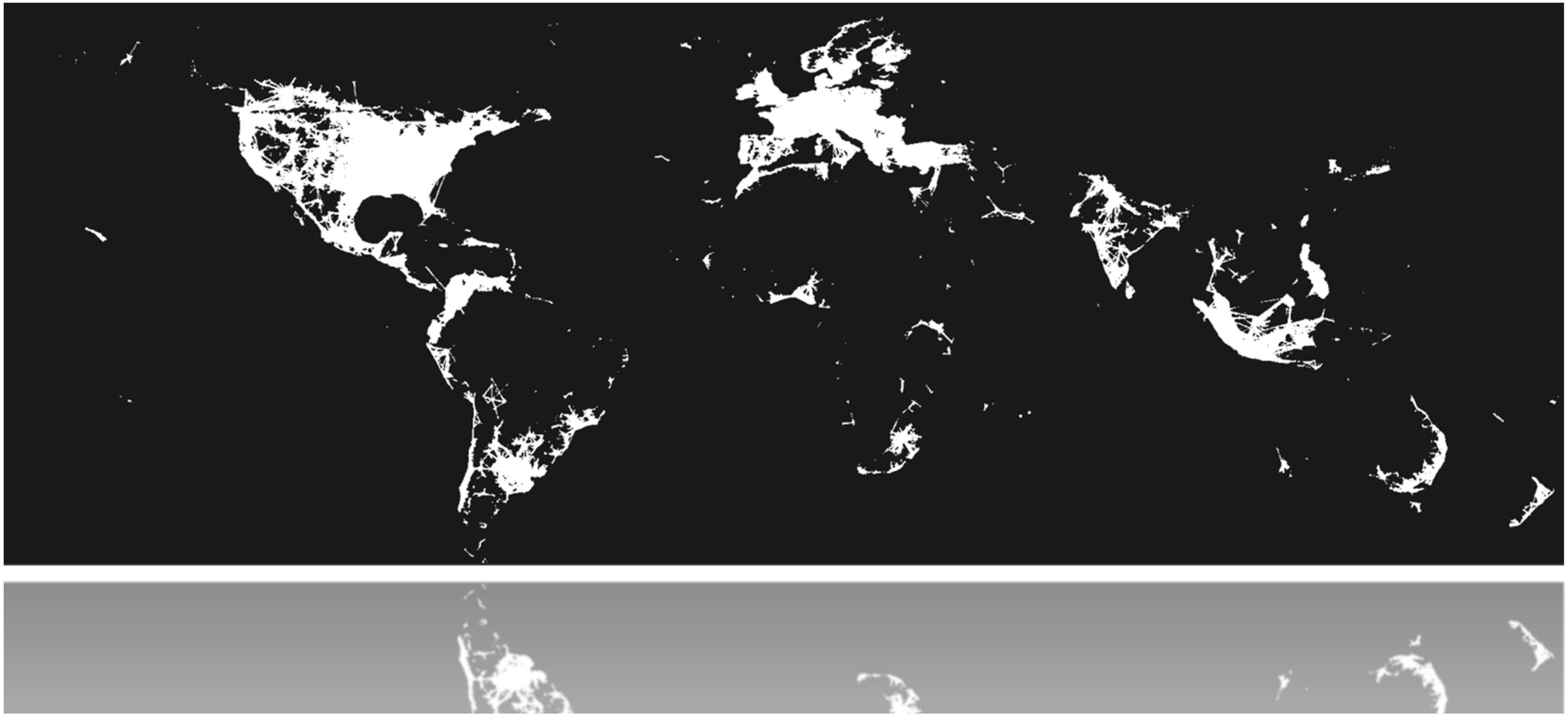


Localmente estructurada

Puramente aleatoria

CLUSTERING versus ALEATORIEDAD — ¿Son reconciliables? (1)

¿Puede una red con una estructura local muy fuerte ser a la vez un mundo pequeño?



CLUSTERING versus ALEATORIEDAD ¿Son reconciliables? (2)

- Fenómeno de mundo pequeño = distancias medias pequeñas:

$$\langle d \rangle_{red} \approx \ln(N)$$

- Clustering: $\langle C \rangle_{red} \gg \langle C \rangle_{aleatoria}$

Ejemplos de redes reales con ese comportamiento:

- Red neuronal del C. elegans
- Redes semánticas de los idiomas
- Red de actores de Hollywood
- Redes alimentarias
- Red eléctrica de EEUU

EL MODELO DE WATTS Y STROGATZ

MODELO DE WATTS Y STROGATZ (1)

La solución para reconciliar mundos pequeños y clustering es **mezclar estructura y aleatoriedad**

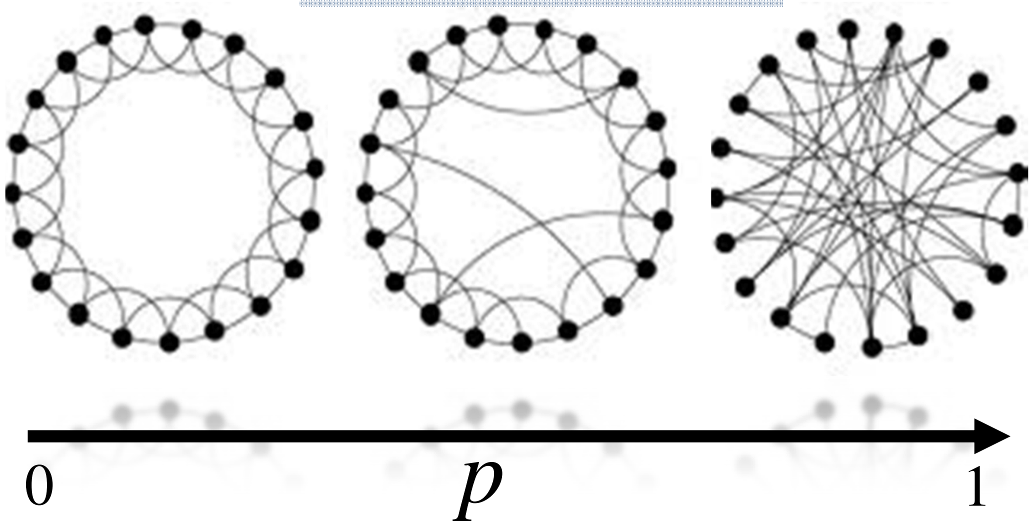
Modelo Watts-Strogatz para la generación de mundos pequeños:

1. Construir una red de retículo en anillo con *N* nodos, cada uno con <*k*> vecinos, con *L*=N · <*k*>/2 enlaces (*N* >> <*k*> >> ln *N*)

2. Reasignar cada enlace con probabilidad *p* (en sentido horario) → <*k*>/2 vueltas para cubrir los N · <*k*>/2 enlaces

No permitir auto-enlaces ni enlaces repetidos (múltiples)

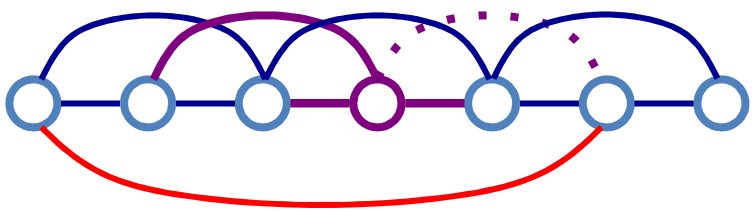
$N = 20 \quad ; \quad \langle k \rangle = 4$



D.J. Watts y S.H. Strogatz. Collective dynamics of 'small-world' networks. Nature 393: 440-442 (1998)

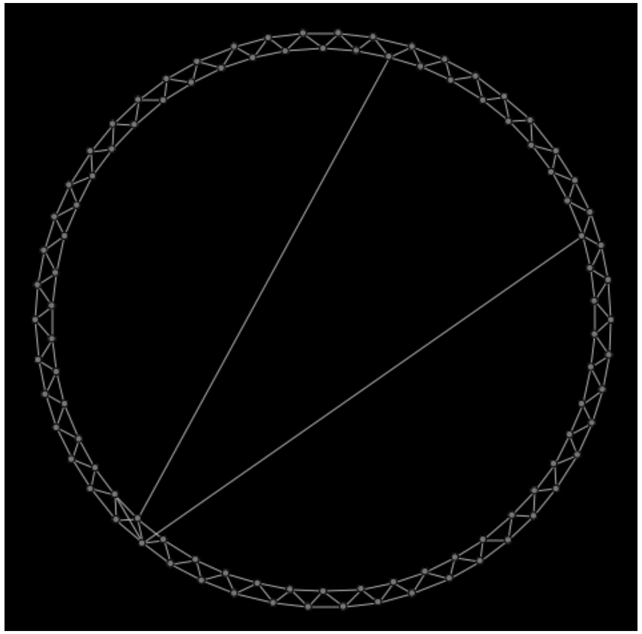
MODELO DE WATTS Y STROGATZ (2)

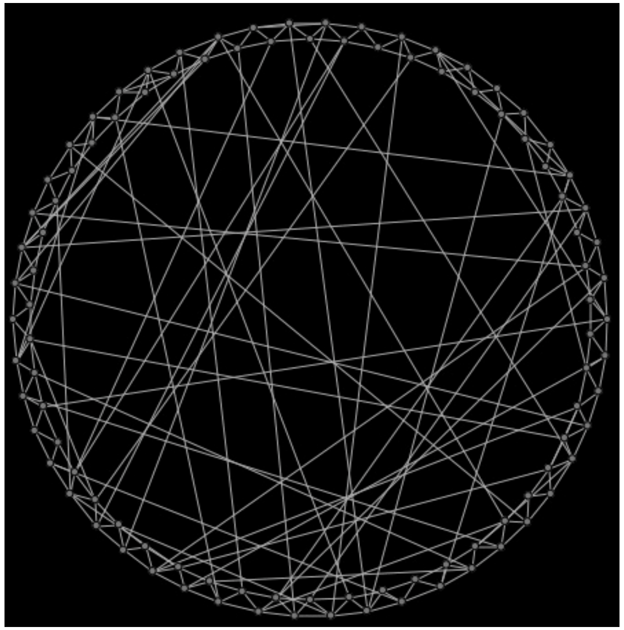
- Cada nodo tiene <k> ≥ 4 vecinos (estructura local fuerte)
- Cada nodo tiene <k>/2 vecinos a cada lado
- Modelo ajustable: se puede variar la probabilidad $0 \leq p \leq 1$ de reasignar un nodo
- Con un p pequeño se mantiene una red retículo regular
- Con un p grande se transforma en una red totalmente aleatoria



MODELO DE WATTS Y STROGATZ (3) | Pregunta

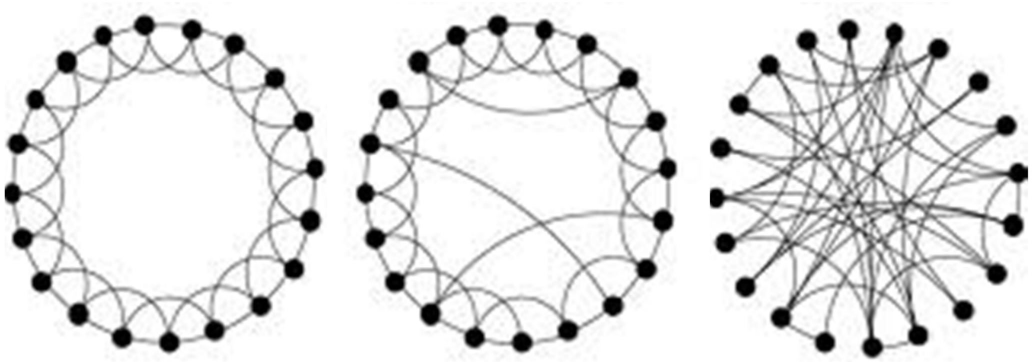
¿Cuál de las dos salidas siguientes del modelo corresponde a una mayor probabilidad de reasignación de enlaces?





MODELO DE WATTS Y STROGATZ (4)

La solución para reconciliar mundos pequeños y clustering es **mezclar estructura y aleatoriedad**



Mundo **grande** fuertemente clusterizado: <*d*> es función de *N*

$$\langle d \rangle_{retículo} = \frac{N}{2\langle k \rangle}$$

$$\langle C \rangle_{retículo} = \frac{3}{4}$$

$$\langle d \rangle_{aleatoria} = \frac{\ln N}{\ln \langle k \rangle}$$

$$\langle C \rangle_{aleatoria} = \frac{\langle k \rangle}{N}$$

Mundo **pequeño** débilmente clusterizado: <*d*> es función de ln *N*

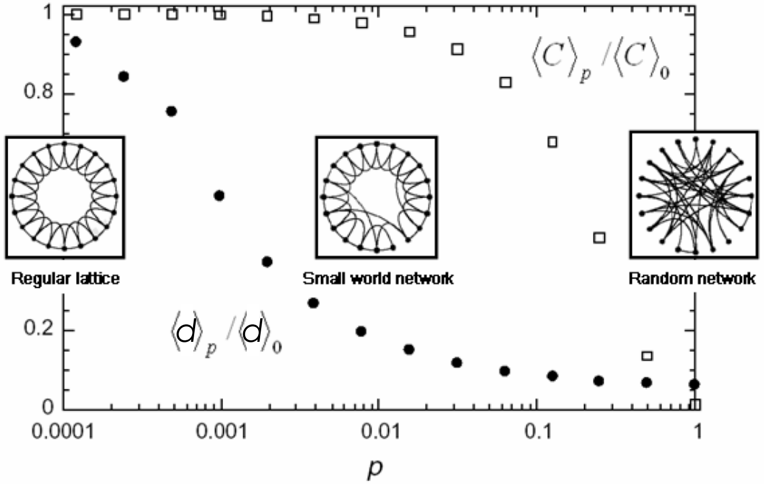
$0 \qquad p \qquad 1$

MODELO DE WATTS Y STROGATZ (5)

¿Qué ocurre en la zona intermedia?

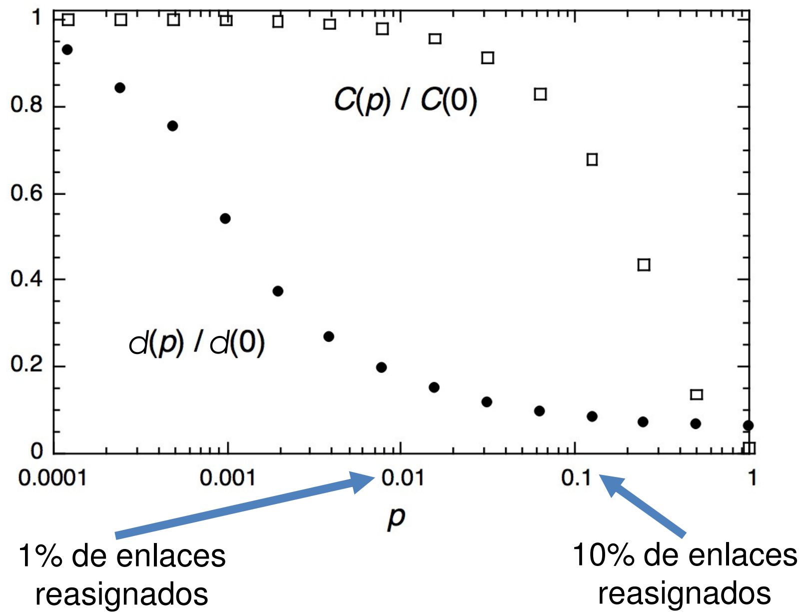
El algoritmo introduce $p \cdot N \cdot \langle k \rangle/2$ enlaces no regulares. Mediante simulación numérica, podemos observar que:

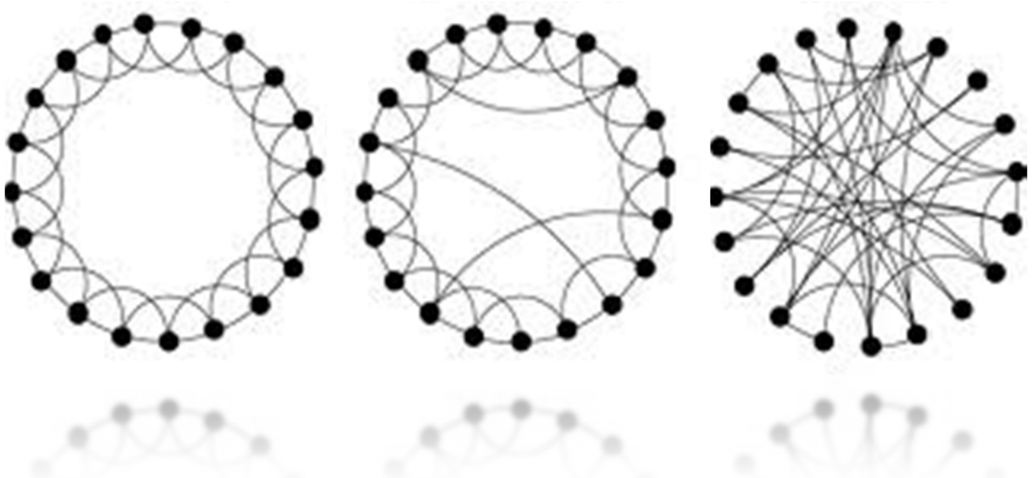
- Hay una reducción muy rápida de la distancia media $\langle d \rangle$ por la aparición de los enlaces "atajo"
- Hay una reducción muy suave del coeficiente de clustering $\langle C \rangle$



MODELO DE WATTS Y STROGATZ (6)

La solución para reconciliar mundos pequeños y clustering es **mezclar estructura y aleatoriedad**



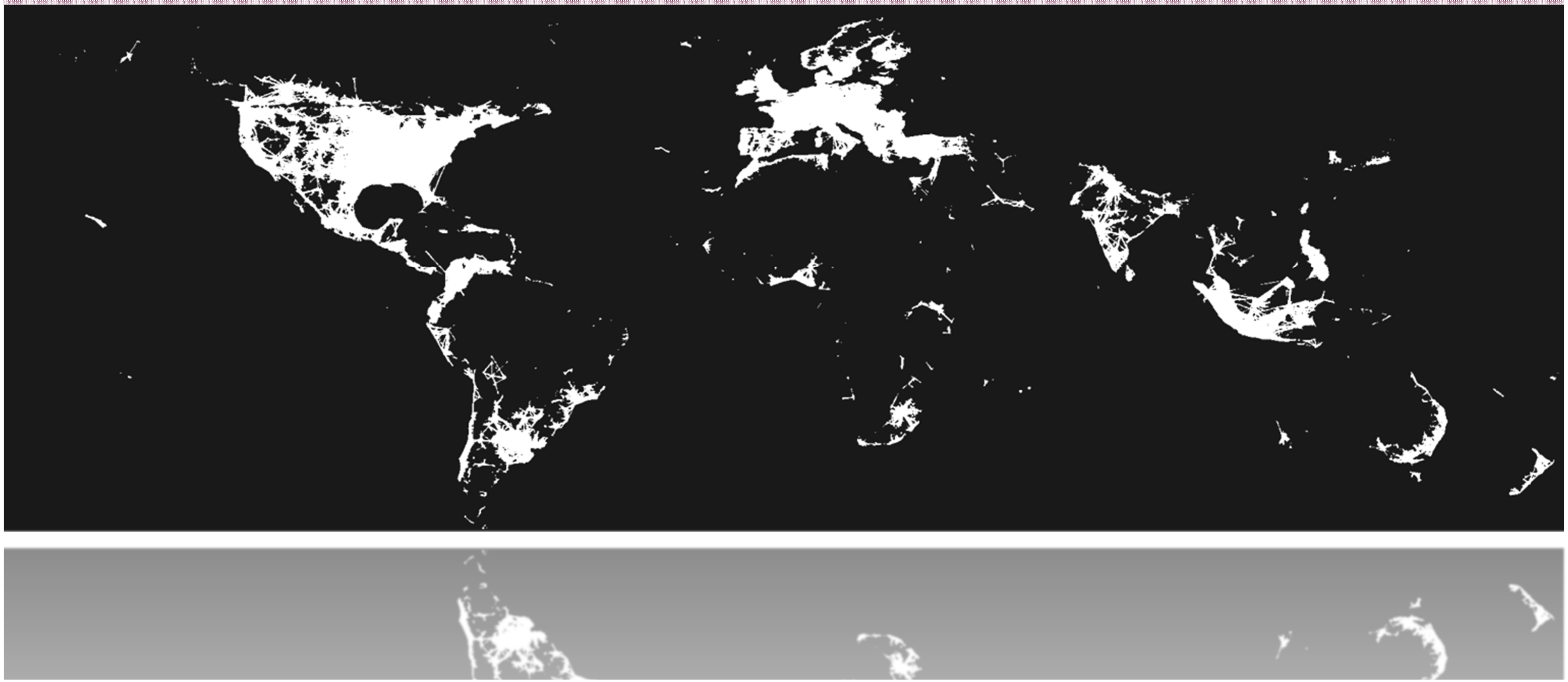


Modelo Watts-Strogatz:
Para eliminar el clustering hace falta una alta aleatoriedad pero para debilitar la localidad basta una poca

MODELO DE WATTS Y STROGATZ (7)

¿Puede una red con una estructura local muy fuerte ser a la vez un mundo pequeño?

¡SI! Basta con unos pocos enlaces aleatorios



MODELO DE WATTS Y STROGATZ (8)

Red de Facebook

¿Puede una red con una estructura local muy fuerte ser a la vez un mundo pequeño?
¡SI! Basta con unos pocos enlaces aleatorios

MODELO DE WATTS Y STROGATZ (9)

Otras redes reales

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}
WWW, site level, undir.	153 127	35.21	3.1	3.35	0.1078	0.00023
Internet, domain level	3015–6209	3.52–4.11	3.7–3.76	6.36–6.18	0.18–0.3	0.001
Movie actors	225 226	61	3.65	2.99	0.79	0.00027
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}
Neurosci. co-authorship	209 293	11.5	6	5.01	0.76	5.5×10^{-5}
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006
Power grid	4941	2.67	18.7	12.4	0.08	0.005
C. Elegans	282	14	2.65	2.25	0.28	0.05

Albert and Barabási, Reviews of Modern Physics **74**,47 (2002)

MODELO DE WATTS Y STROGATZ (10) — Otras redes reales (2)

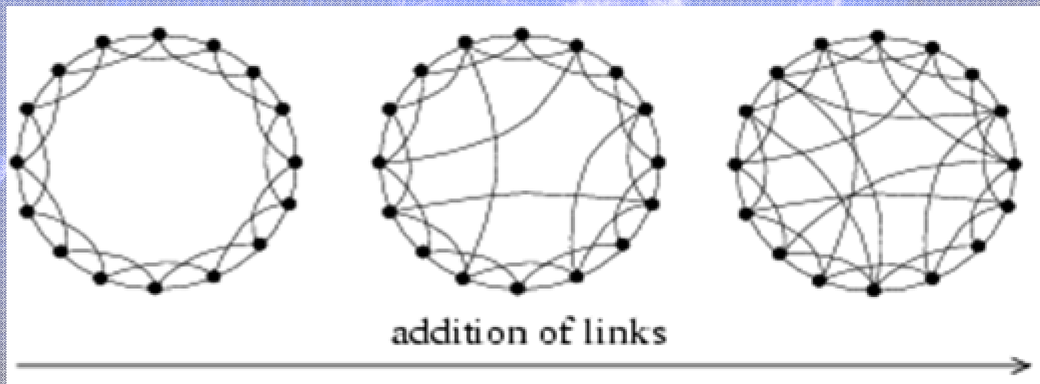
Mapa de colaboraciones científicas

MODELO DE WATTS Y STROGATZ (11) — Conclusiones (1)

¿Puede una red con una estructura local muy fuerte ser a la vez un mundo pequeño?
¡SI! Basta con unos pocos enlaces aleatorios

Modelo Watts-Strogatz:

- **Alternativamente se puede considerar otra variante del modelo que añade enlaces aleatorios con probabilidad p, manteniendo el retículo inicial**



MODELO DE WATTS Y STROGATZ (12) | Conclusiones (2)

¿Puede una red con una estructura local muy fuerte ser a la vez un mundo pequeño?
¡SI! Basta con unos pocos enlaces aleatorios

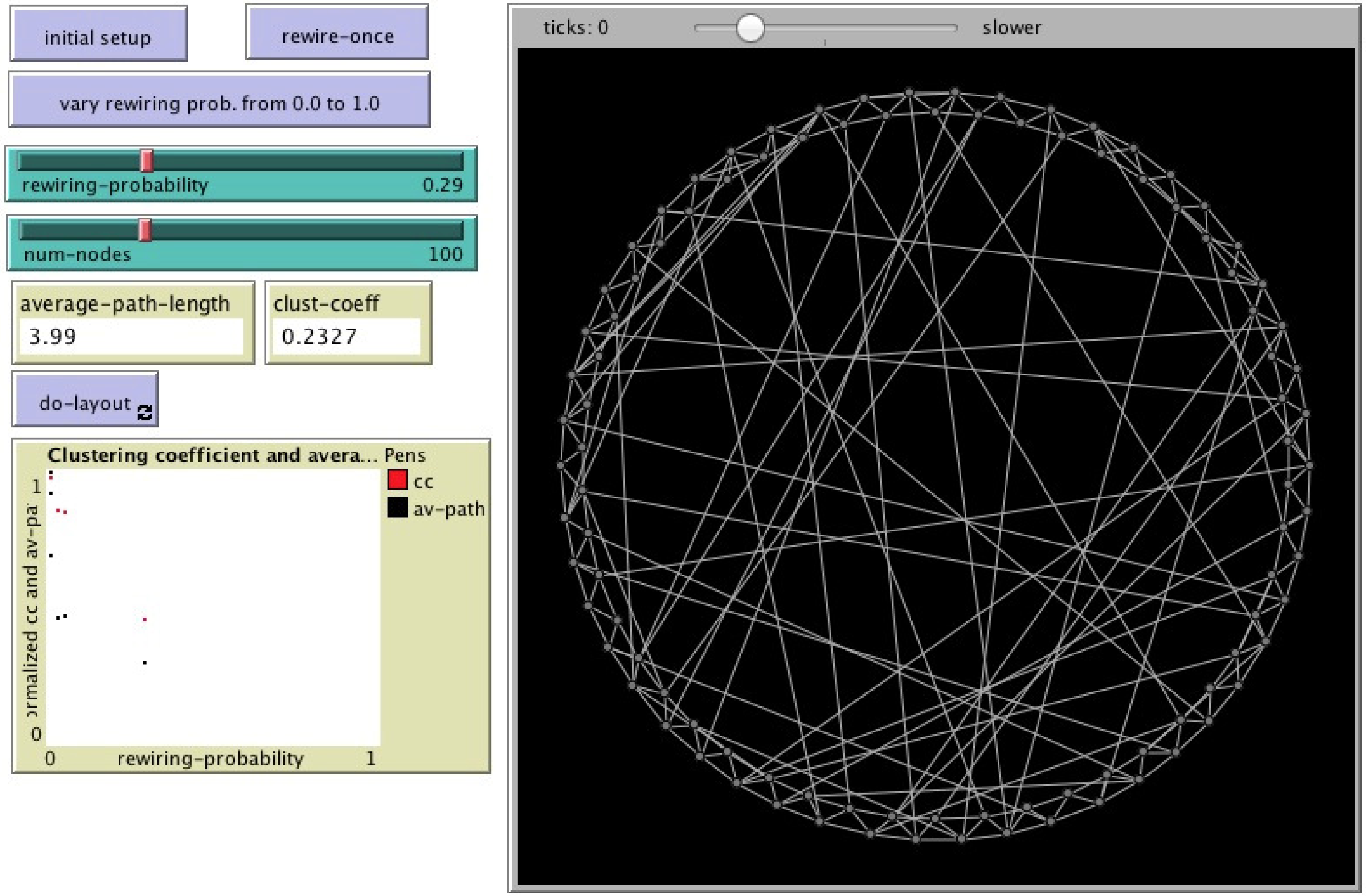
Modelo Watts-Strogatz:

- Proporciona conocimiento sobre la interrelación entre el clustering y la topología de mundos pequeños
- Captura la esencia estructural de muchas redes reales
- Tiene en cuenta el alto coeficiente de clustering observado en las redes reales

- **No representa una distribución de los grados realista (hubs)**
- **Los enlaces largos son menos frecuentes que los cortos**

MODELO DE WATTS Y STROGATZ (13) — Modelo Netlogo

http://www.ladamic.com/netlearn/NetLogo4/SmallWorldWS.html



RETOMANDO EL EXPERIMENTO DE MILGRAM

EXPERIMENTO DE MILGRAM: Actualización (1)

Experimento basado en e-mails:

Dodds, Muhamad, Watts (2003). Science 301

- 18 objetivos
- 13 países distintos

- Más de 60,000 participantes
- 24,163 cadenas de mensajes
- 384 alcanzaron el objetivo

Distancia media = 4.0

Fuente de la imagen: NASA
http://visibleearth.nasa.gov/view_rec.php?id=2429

EXPERIMENTO DE MILGRAM: Actualización (2)

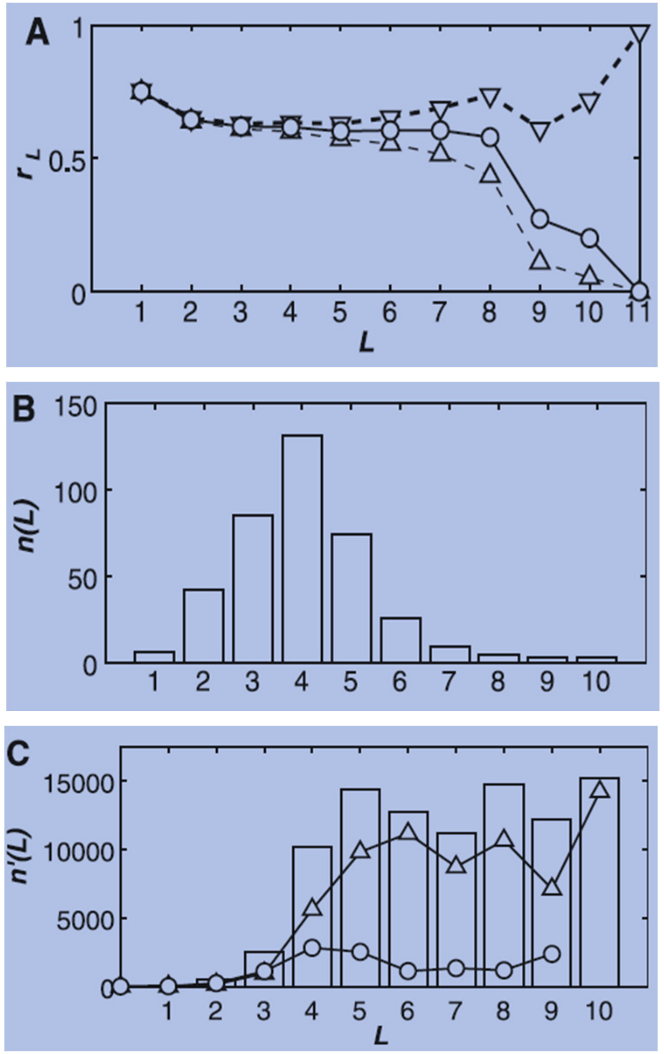
Experimento basado en e-mails:

Dodds, Muhamad, Watts (2003). Science 301

- 18 objetivos
- 13 países distintos

- Más de 60,000 participantes
- 24,163 cadenas de mensajes
- 384 alcanzaron el objetivo

Distancia media = 4.0



EXPERIMENTO DE MILGRAM: Actualización (3)

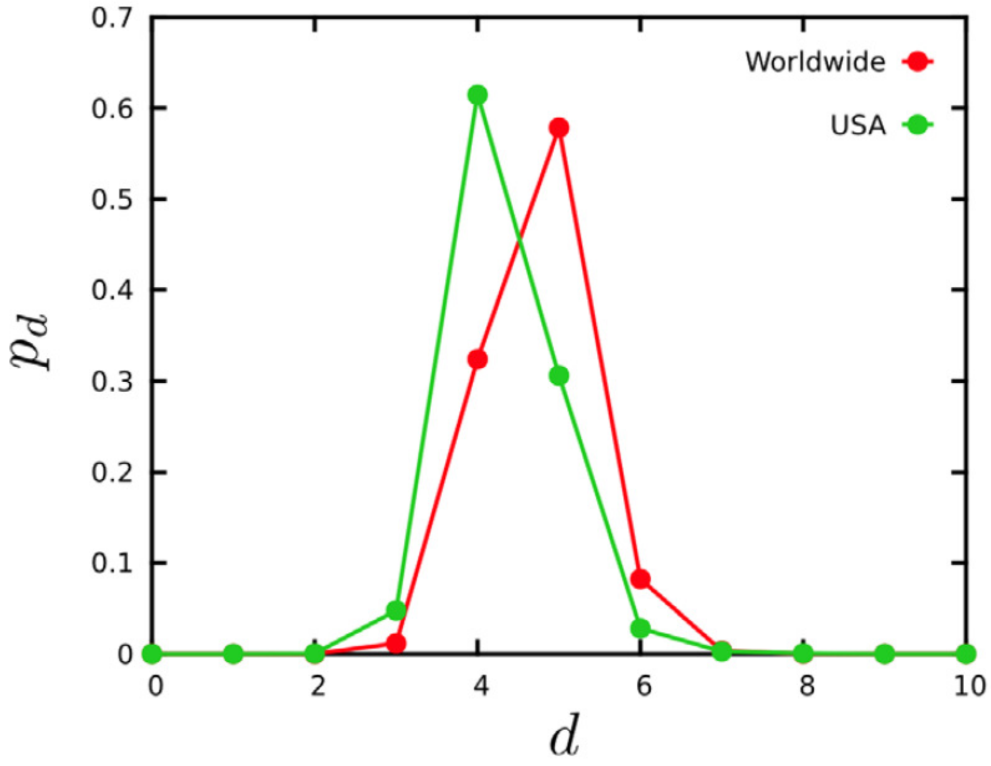
Experimento basado en Facebook:

Backstrom, L., Boldi, P., Rosa, M., Ugander, J. & Vigna, S. (2011). Four degrees of separation. CoRR, abs/1111.4570

El experimento de Milgram no disponía de un mapa adecuado de la red social mundial. Hoy en día, FB es una buena aproximación

Mapa FB Mayo 2011: 721 millones de usuarios activos, 68 billones de relaciones: **distancia media = 4.74**

Este valor más cercano a la distancia media teórica: 3.90



MUNDOS PEQUEÑOS: Red social mundial

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

En sociología, una persona cualquiera conoce directamente a otras mil, *k* ~1000

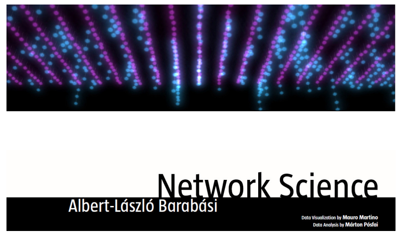
La población mundial está actualmente en torno a los 7,000 millones ($7 \cdot 10^9$)

$$\langle d \rangle = \frac{\ln 7 \times 10^9}{\ln(10^3)} = 3.28.$$

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



- Curso on-line "Social Network Analysis" de Lada Adamic, Coursera, Universidad de Michigan: https://www.coursera.org/course/sna