Modelos de Redes

Inteligencia Colectiva y Sistemas de Recomendación

Master Propio en Data Science y Big Data

Modelos de Redes

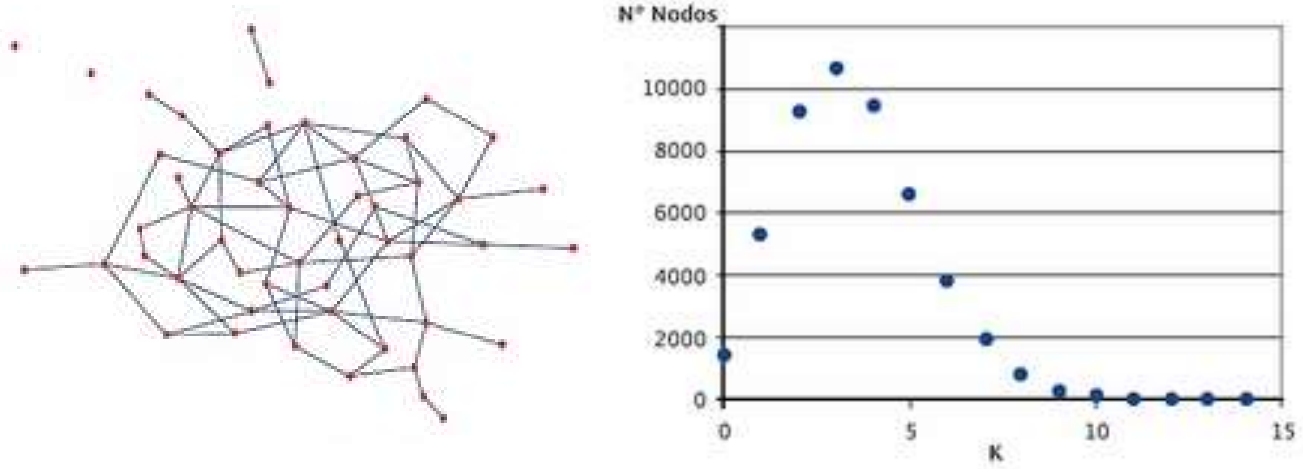
Una primer aproximación para clasificar las redes complejas de forma automática y general se puede realizar a partir de su *Distribución de Grados P(k)*.

Las más importantes son:

- Topología de Poisson.
- Topología Libre de Escala.

Topología de Poisson

En las redes de Poisson todos los nodos tienen un número similar de conexiones, es decir, las conexiones están distribuidas homogéneamente entre sus nodos.

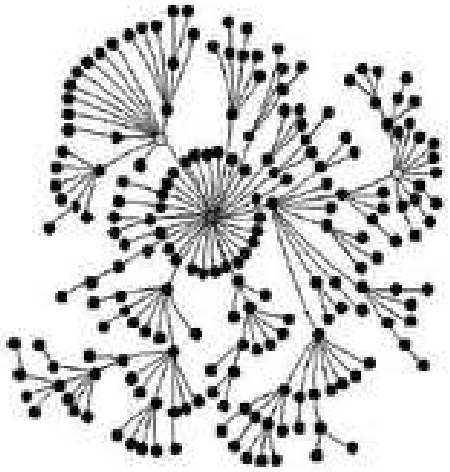


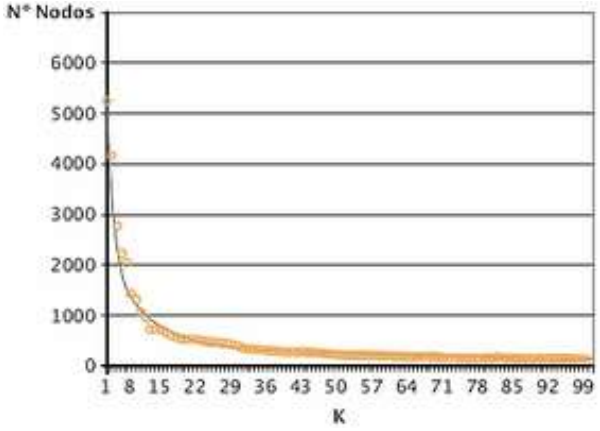
El valor medio de los grados es representativo del grado de los vértices.

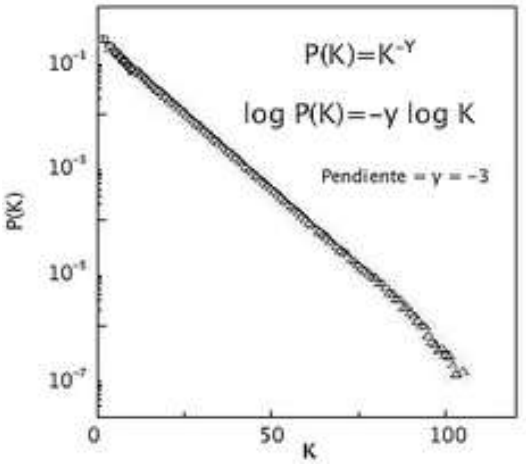
Topología Libre de Escala

Las redes Libres de Escala poseen alta heterogeneidad, al contener nodos con pocas, medias y muchas conexiones. El valor medio de la distribución no es representativo de la conectividad de la red. La mayoría de los vértices tienen baja conectividad y alta solo unos pocos: "muchos con poco, pocos con mucho"

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







Robustez de las topologías Libres de Escala

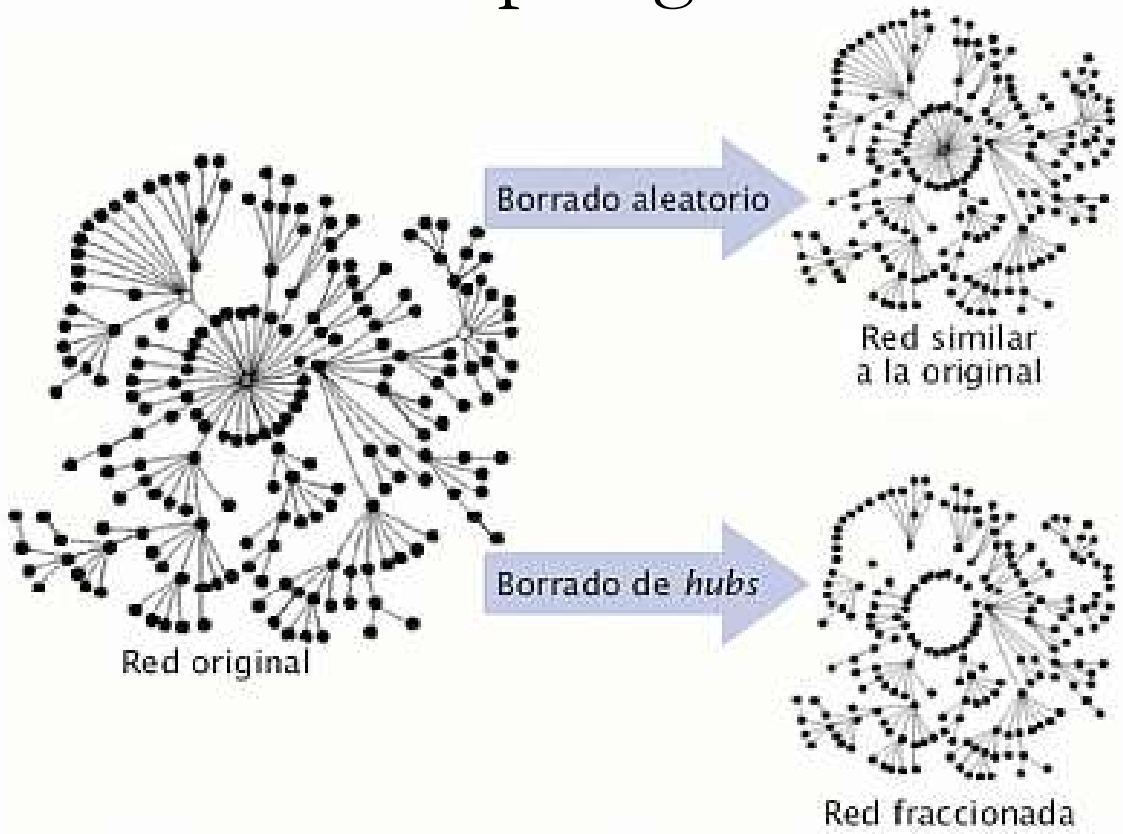


Fig 11 Ejemplo de un borrado aleatorio y un borrado selectivo de nodos en una red Libre de Escala. En un borrado aleatorio, como la mayoría de los nodos tienen pocas conexiones los efectos sobre la red son mínimos. En un borrado selectivo de hubs los efectos son dramáticos dejando muchos nodos no conectados y la red fraccionada.

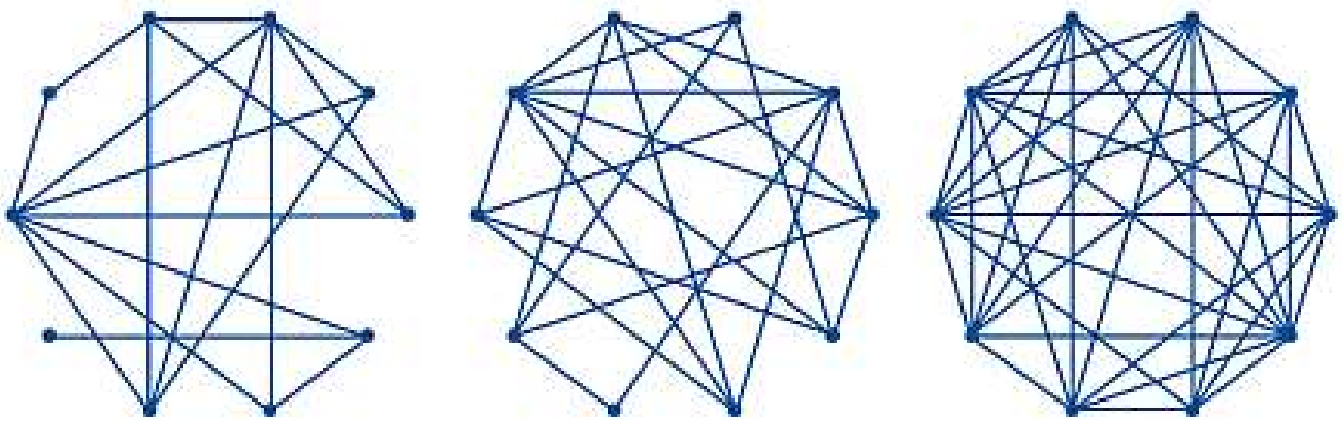
Modelos de construcción de Redes

- Dos procedimientos principales:
  - Modelos Estadísticos
    - Modelo de Grafos aleatorios
    - Modelo de Wattz-Strogatz
  - Modelos Dinámicos
    - Enlace Preferencial
    - Duplicación

Modelo de Grafos Aleatorios

Iniciador de la Teoría de Redes Complejas: Erdos-Renyi (50's).

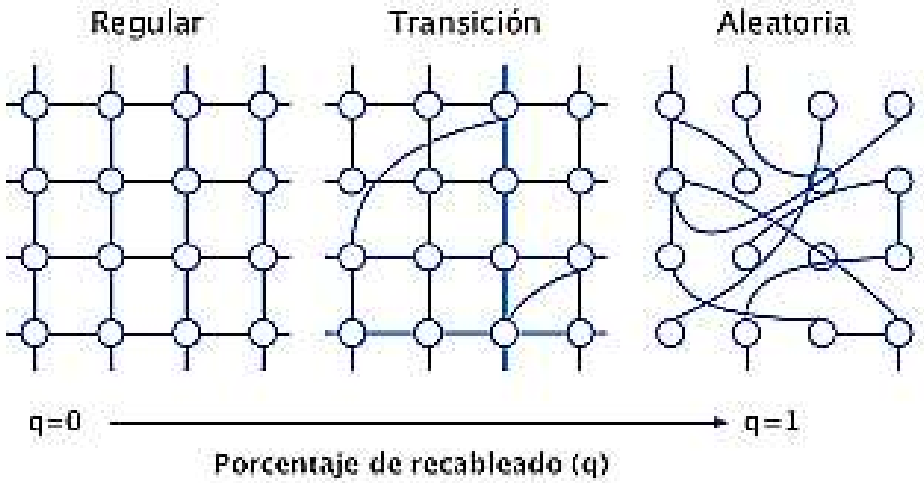
- Construye la red enlazando nodos elegidos al azar según determinada probabilidad



Grafos Aleatorios con distinta probabilidad de conexión entre pares de nodos (A) p=0.3, (B) p=0.5 y (C) p=0.8

Modelo de Wattz-Strogatz

(1998) Transforma un grafo regular en una red aleatoria al recablear enlaces añadiendo o moviendo los ya existentes.



Recableado de una malla regular hasta llegar a una aleatoria modelo de Watts y Strogatz. En una red regular existe un 0% de enlaces recacableados en forma aleatoria mientras que en una red aleatoria un 100% de los enlaces son recableados aleatoriamente. Entre ambos extremos se ubican una serie de topologías de transición.

Características de los Mod. Est.

- Permiten construir redes con homogeneidad en el conexionado.
- **Modelo de Grafos Aleatorios:** tienen la propiedad **Small World** y bajo **Coeficiente de Clustering**.
- **Modelo WS:** poseen característica **Small World** y alto **Coeficiente de Clustering**.
- Estos modelos no logran reproducir las características de los sistemas complejos.

Modelos Dinámicos

- Consideran las redes como sistemas con interacciones que varían en el tiempo según determinadas leyes.
- Se llaman también **Modelos de Crecimiento** o de **Evolución** ya que imitan los procesos de crecimiento mediante la adición gradual de nodos o enlaces.
- Logran reproducir la heterogeneidad en el conexionado, la **Distribución de Grados**, el **Coeficiente de Clustering** *y* el efecto **Small World** observados en sistemas complejos.

Modelo de Enlace Preferencial

Asume que la conexión de los nuevos nodos añadidos al sistema está regulada por la cantidad de conexiones de los ya presentes. Es decir, los nuevos elementos se unirán con mayor probabilidad a los más conectados ya ubicados en la red: "el rico se vuelve más rico".(Barabási y Albert, 1999)

Fig. 14. Modelo de Enlace Preferencial. Los nodos más conectados tiene mayor probabilidad de atrapar a los recién llegados. Las flechas indican estos enlaces.

Modelo de Duplicación

Asume que el origen de los nuevos elementos añadidos a la red es interno. Los nodos se duplican y se unen a los ya existentes según determinada probabilidad
(Pastor Santorras et al, 2003).

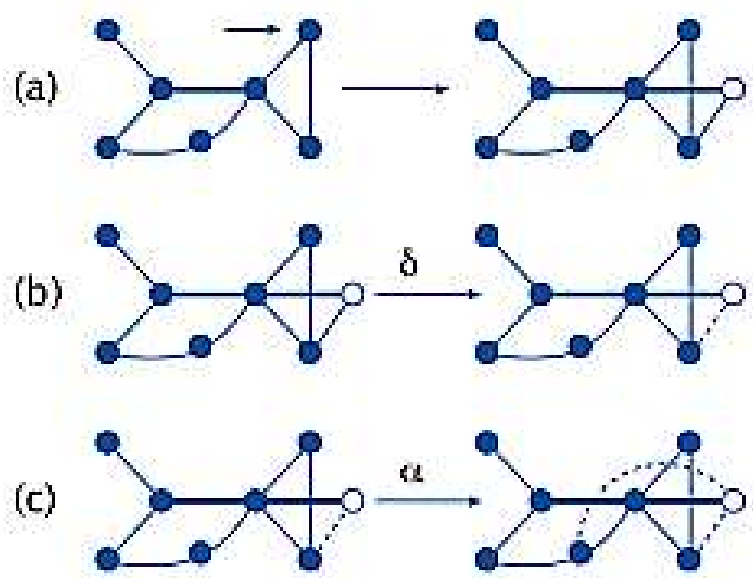
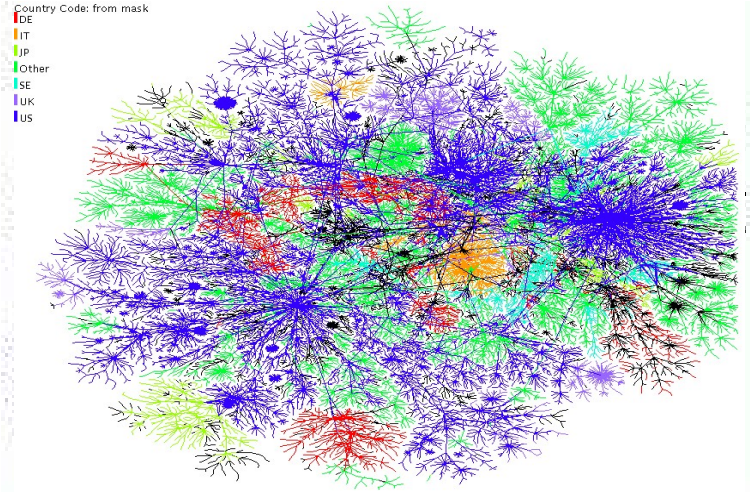
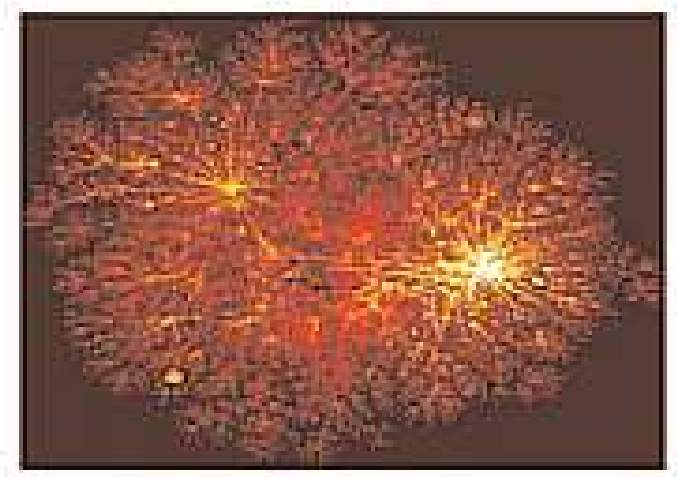


Fig. 15. Modelo de Duplicación en tres etapas: a) copia del nodo señalado con una flecha, b) eliminación de conexión según una probabilidad ‰ y c) generación de nueva conexión según una probabilidad ·

Redes Naturales vs. Redes Artificiales

- Las redes naturales evolucionan por adaptación al entorno.
- A diferencia de éstas, el origen de una red artificial, como Internet, está basado en un diseño humano inteligente.
- Sin embargo, presentan características topológicas análogas.





Comparativa de algunas redes

Red	Tipo	Nodos	Aristas	k	l	e	C1	C2
Social								
film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78
company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88
math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34
physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56
biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60
telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1		
email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16
email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13
student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001
sexual contacts	undirected	2 810				3.2		
Information								
WWW nd.edu	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29
WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7		
citation network	directed	783 339	6 716 198	8.57		3.0/–		
Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.13	0.15
word co-occurrence	undirected	460 902	17 000 000	70.13		2.7		0.44

k: grado medio, l: distancia media,
e: exponente libre escala,
C1-2: Coef. Clustering

Comparativa de algunas redes

Red	Tipo	Nodos	Aristas	k	l	exp	C1	C2
Technological								
Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.035	0.39
power grid	undirected	4 941	6 594	2.67	18.99	–	0.10	0.080
train routes	undirected	587	19 603	66.79	2.16	–		0.69
software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.070	0.082
software classes	directed	1 377	2 213	1.61	1.51	–	0.033	0.012
electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.010	0.030
peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.012	0.011
Biological								
metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.090	0.67
protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.072	0.071
marine food web	directed	135	598	4.43	2.05	–	0.16	0.23
freshwater food web	directed	92	997	10.84	1.90	–	0.20	0.087
neural network	directed	307	2 359	7.68	3.97	–	0.18	0.28

k: grado medio, l: distancia media,
e: exponente libre escala,
C1-2: Coef. Clustering