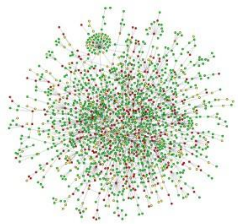
Escuela de Invierno
Dinámica de Redes Complejas
Aplicaciones Bioelectrónica y Bioinformática
Del 30 de Noviembre al 4 de Diciembre
Universidad de Guadalajara
Centro Universitario de los Lagos
Centro Científico y Cultural de la Región / UdeG

Applications of Complex Networks

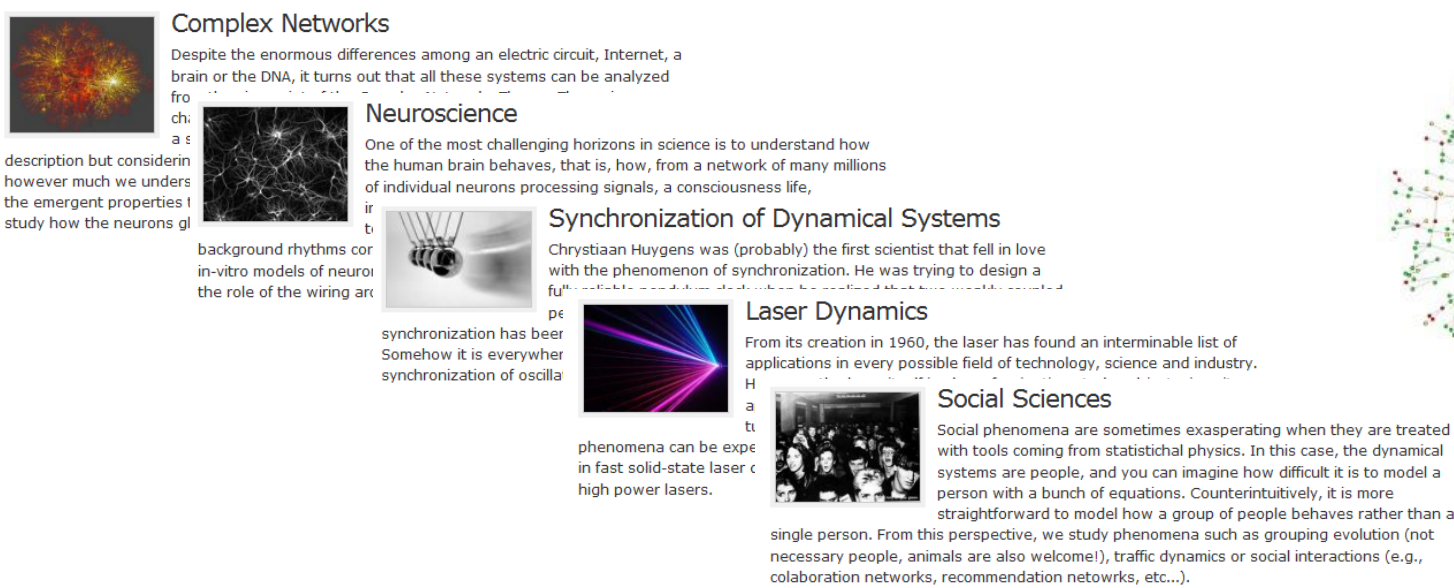
Javier M. Buldú
http://www.complex.etsit.urjc.es/jmbuldu
Complex Systems Group
Universidad Rey Juan Carlos (Madrid, Spain)

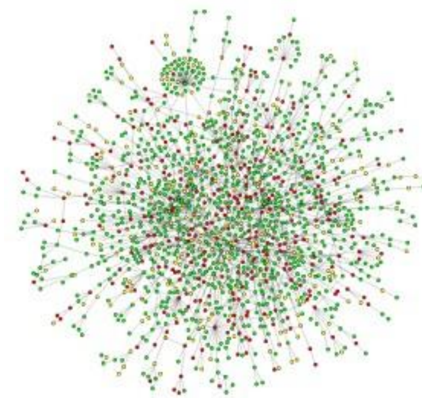


CSG

Universidad Rey Juan Carlos

http://www.complex.etsit.urjc.es/





CSG Universidad Rey Juan Carlos

OUTLINE OF THE COURSE

Universidad Rey Juan Carlos

0.- BIBLIOGRAPHY

Review Articles about Complex Networks:

- M. E. J. Newman, ***The structure and function of complex networks***, SIAM Reviews, 45(2): 167-256, 2003. (available at arXiv.org/cond-mat/0303516)
- S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, ***Complex networks: Structure and dynamics***, Phys. Rep. 424, 175 (2006).
- R. Albert and A.-L. Barabasi, *Statistical Mechanics of Complex Networks*, Rev. Mod. Phys., 74, 47-97 (2002). (available at arXiv.org/cond-mat/0106096)
- M. Newman, D. Watts, A.-L. Barabási, *The Structure and Dynamics of Networks*, Princeton University Press, 2006.
- J.A. Almendral: “Dynamics and Topology in Complex Networks”, Ph.D. Thesis. http://complex.escet.urjc.es/pdfs/almendral_tesis.pdf

Review articles about REAL complex Networks:

- Luciano da F.Costa et al., *Analyzing and Modeling Real-World Phenomena with Complex Networks: A Survey of Applications*. (available at http://arxiv.org/abs/0711.3199)
- Aaron Clauset et al., *Power-law distributions in empirical data*, arXiv:0706.1062v2.
- Chris Anderson, *The Long Tail*, Wired 12.20. October 2004.
- S. Bornholdt and S.G. Shuster, *Handbook of grahps and networks*, Wiley-VCH, Weinheim 2003.

CSG Universidad Rey Juan Carlos

0.- BIBLIOGRAPHY

Popular Science Books:

- Duncan J. Watts, *Six Degrees: The Science of a Connected Age*, W. W. Norton and Company. 2003.
- A.-L. Barabási, *Linked: The New Science of Networks*, Perseus, Cambridge, MA, 2002.
- Mark Buchanan, *Nexus: Small Worlds and the Groundbreaking Theory of Networks*, W. W. Norton and Company. 2003.

Complex Networks Databases:

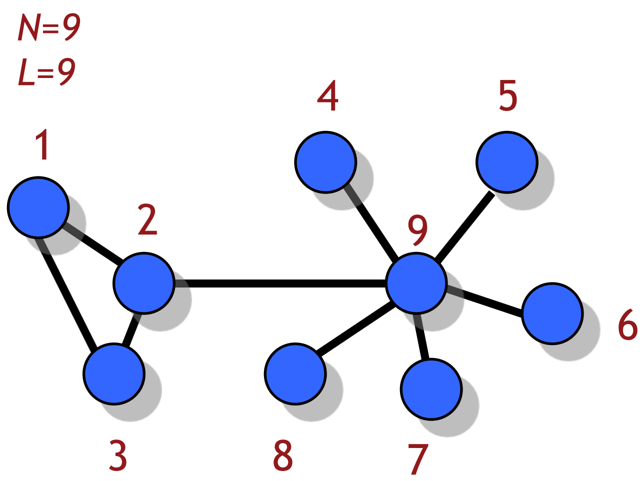
- Mark Newman, **University of Michigan**: http://www-personal.umich.edu/~mejn/netdata/
- Alberto L. Barabási, **University of Notre Dame**: http://www.nd.edu/~networks/resources.htm
- Alex Arenas, **Universitat Rovira y Virgili**: http://deim.urv.cat/~aarenas/data/welcome.htm
- **Indiana University** databases: http://iv.slis.indiana.edu/db/index.html

1.- INTRODUCTION TO COMPLEX NETWORKS

1.1.- What is a (complex) network?

1.1.- WHAT IS A (COMPLEX) NETWORK?

- A Network is a set of elements with connections between them



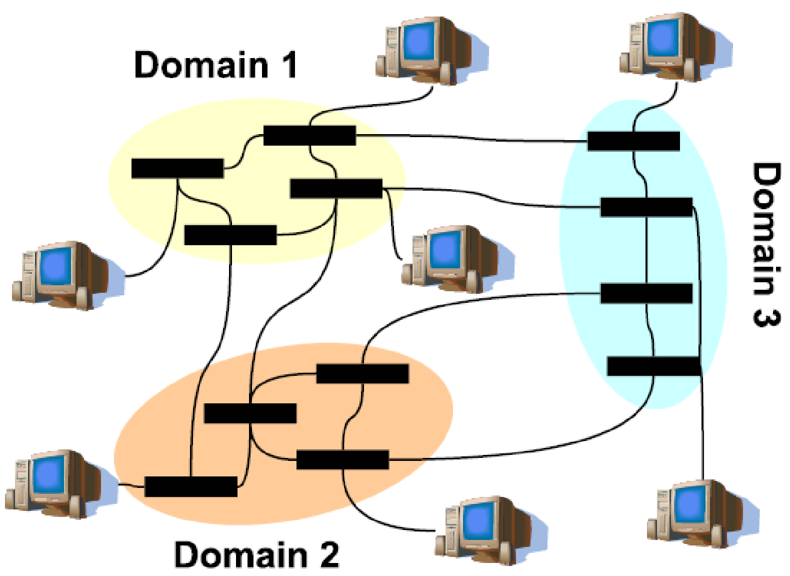
A network (graph) $G=(N,M)$ consists of a set of $N=\{n_1, n_2, \ldots, n_N\}$ nodes and a set of $L=\{l_1, l_2, \ldots, l_M\}$ links.

A graph is the mathematical abstraction of a network. Despite it is not rigorous, we will use both terms, graph and network, as synonyms.

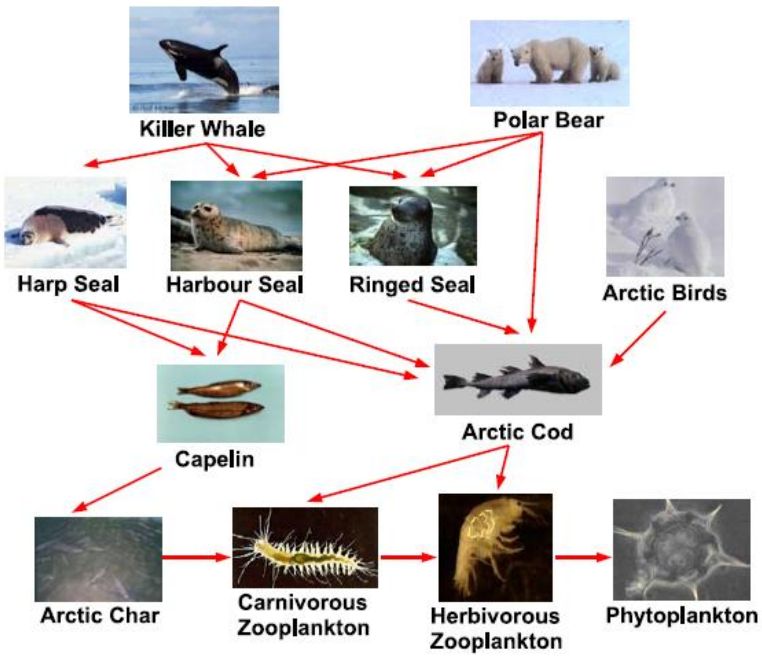
From this viewpoint, each element is represented by a site (physics), node (computer science), actor (sociology) or vertex (graph theory) and the interaction between two elements corresponds to a bond (physics), link (computer science), tie (sociology) or edge (graph theory).

1.1.- WHAT IS A (COMPLEX) NETWORK?

- Nodes and links may arise from completely different contexts:

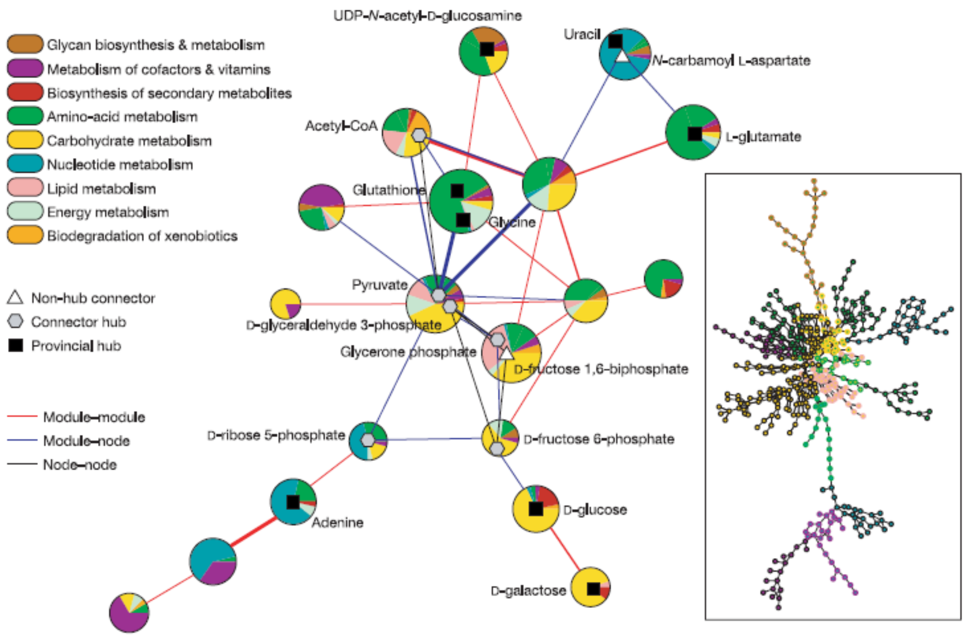


Schematic representation of a network of hosts and routers.



Simplified representation of the Arctic food web

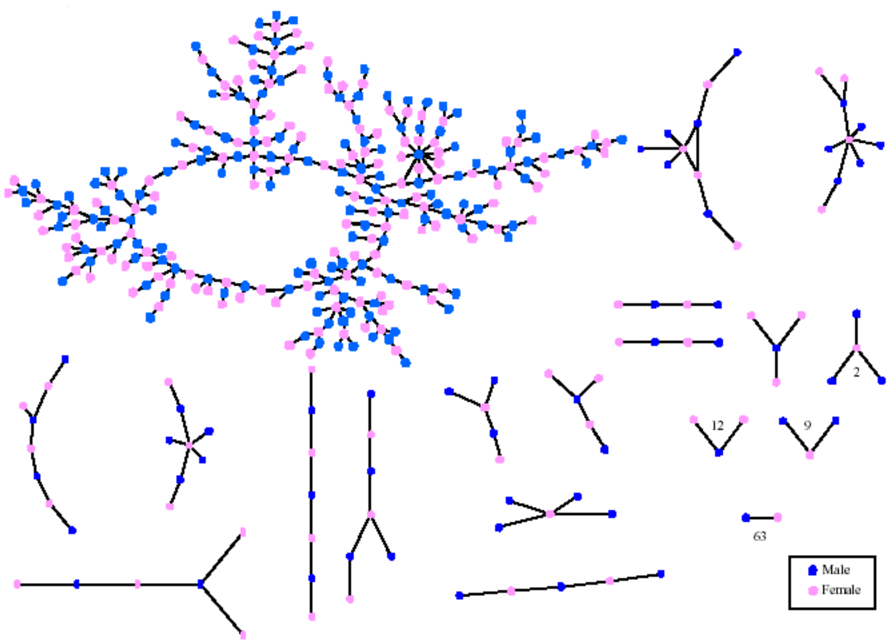
1.1.- WHAT IS A (COMPLEX) NETWORK?



Metabolic network of the *E. Coli*.
From Guimerà et al., Nature, 433, 895, 2005

Madrid Power Grid.
From http://www.ree.es

1.1.- WHAT IS A (COMPLEX) NETWORK?

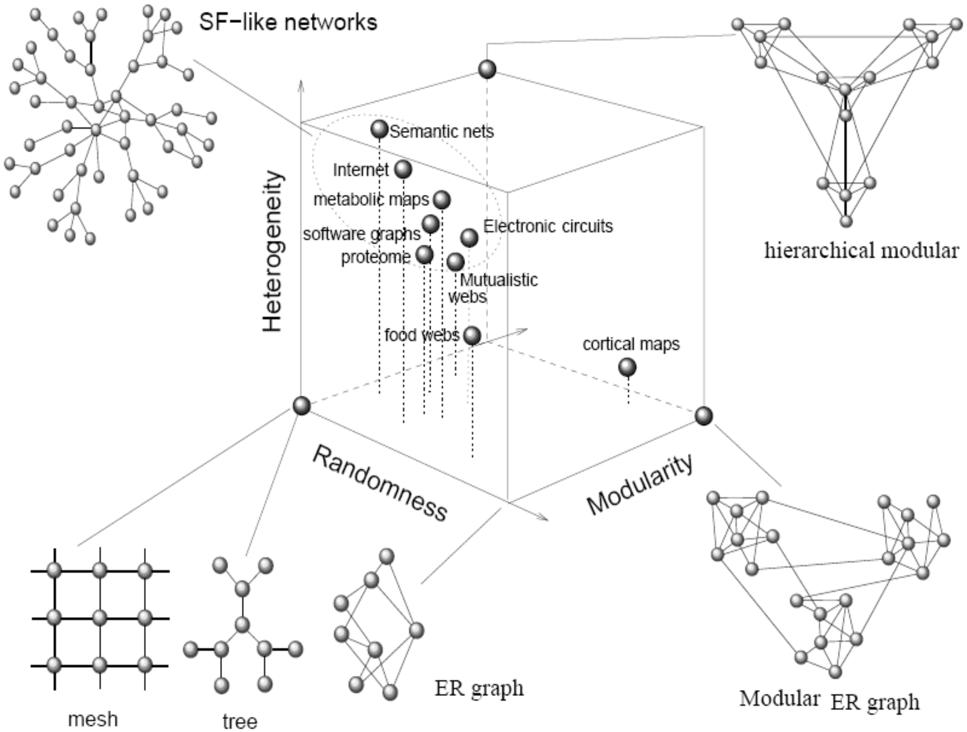


Structure of romantic and sexual contact at Jefferson High School
From P.S. Bearman et al., AJS, 110, 44 (2004)

Neuron network

1.1.- WHAT IS A (COMPLEX) NETWORK?

❑ A **Complex Network** is a network with non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor purely random.



From: R.V. Solé and S. Valverde, Lecture Notes in Physics, **650**, 189, 2004

1.2.- Types of networks

1.2.- TYPES OF NETWORKS

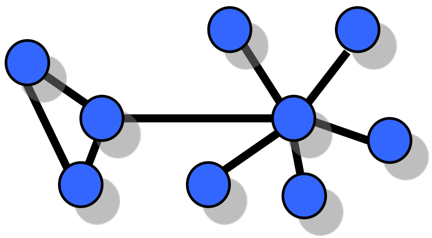
- **There exist different classifications of networks:**
  - According to the direction of the links: **directed or undirected**.
  - According to the kind of interaction: **weighted or unweighted**.
  - According to the differences between nodes: **bipartite** or not.
  - According to the evolution of their topology: **static or evolving.**
  - According to the dynamics of the nodes: **with/without dynamics**.
  - …

more at: M.E.J. Newman,SIAM Reviews, 45, 167 (2003)
S. Boccaletti et al., Phys. Rep., 424, 175 (2006)

❑ Directed and undirected networks:

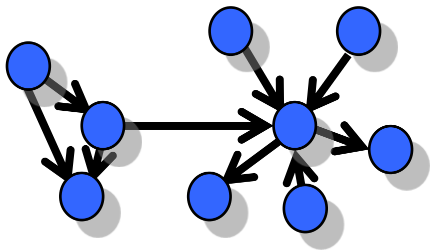
The relationship between nodes may be symmetric (undirected networks) or asymmetric (directed networks).

Undirected network



Examples: router network, power grid.
collaboration networks, etc...

Directed network (digraph)



Examples: internet, food webs,
e-mail/telephone networks, etc...

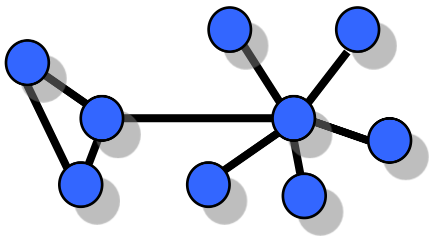
The direction of the links is crucial in dynamical processes ocurring in the network, such as information spreading, synchronization or network robustness.

1.2.- TYPES OF NETWORKS

❑ Weighted and unweighted networks:

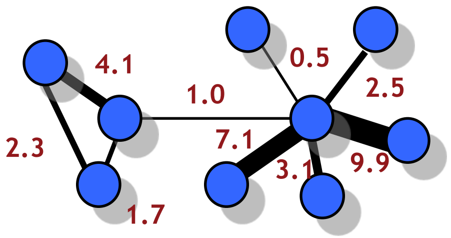
The capacity or intensity of the relationship between nodes may be heterogeneous (weighted networks).

Unweighted network



Examples: citation network, internet, etc...

Weighted network

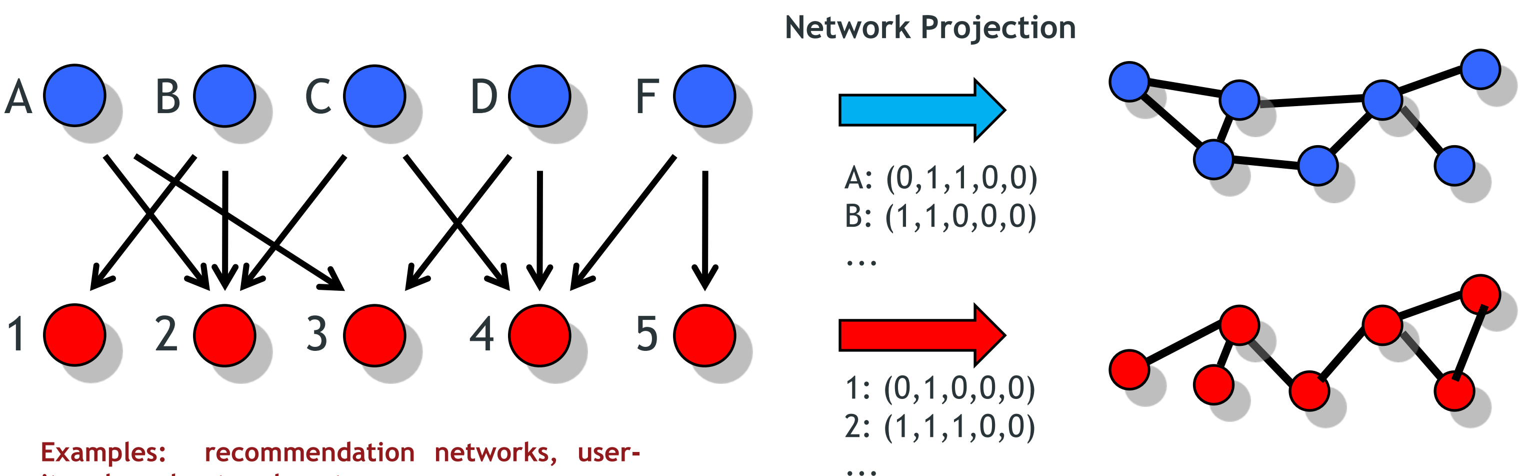


Examples: e-mail/telephone networks, food webs, power grid, colaboration network, etc...

Again, the weight of the links is crucial in dynamical processes ocurring in the network, such as information spreading, synchronization or network robustness.

❑ Bipartite networks:

Networks with two (or more) kind of nodes and links joining ONLY nodes of unlike type.

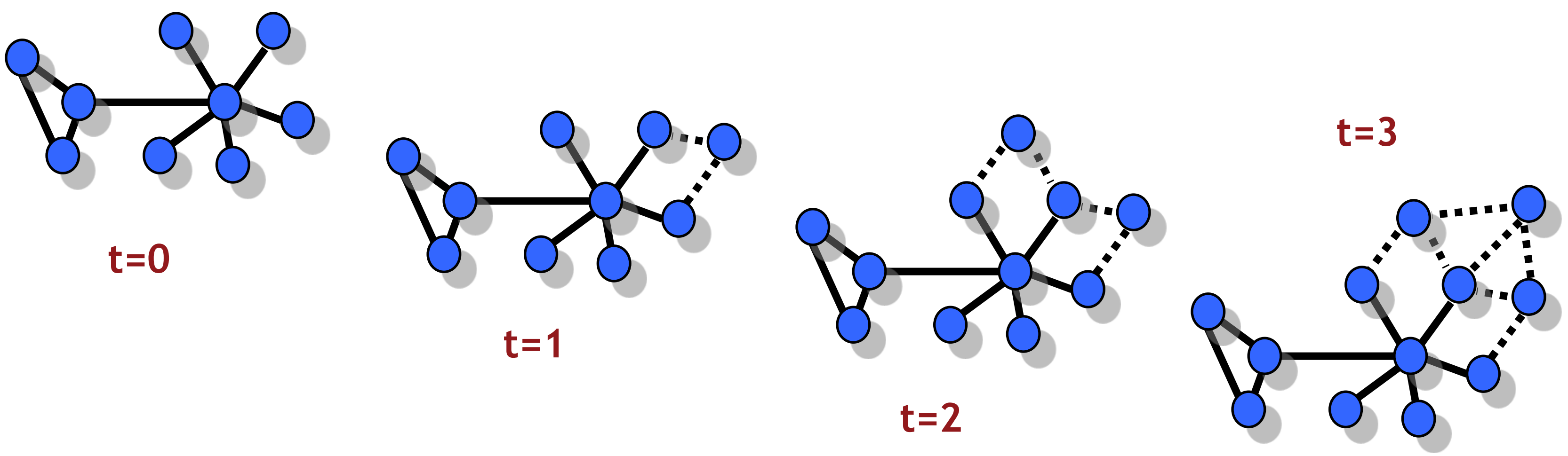


Examples: recommendation networks, user-item based networks, etc...

Despite being bipartite, it is possible to project the network.

❑ Static or evolving networks:

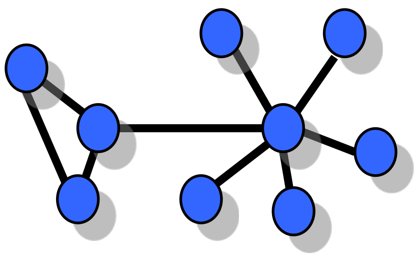
Networks do not appear suddenly. We have to know if the network that we are studying is static (its structure is stationary) or if it is still evolving

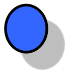


Two fundamental questions are addressed when working with evolving networks: what are the rules governing the evolution? What consequences have the rules on the final topology?

❑ Networks of dynamical systems:

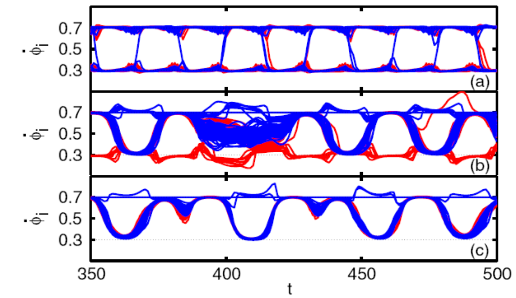
Nodes are dynamical systems whose dynamics is influenced through the matrix of connections.





Nodes are (coupled) dynamical systems
(periodic oscilators, excitable systems, chaotic oscilators, bistable systems, …)

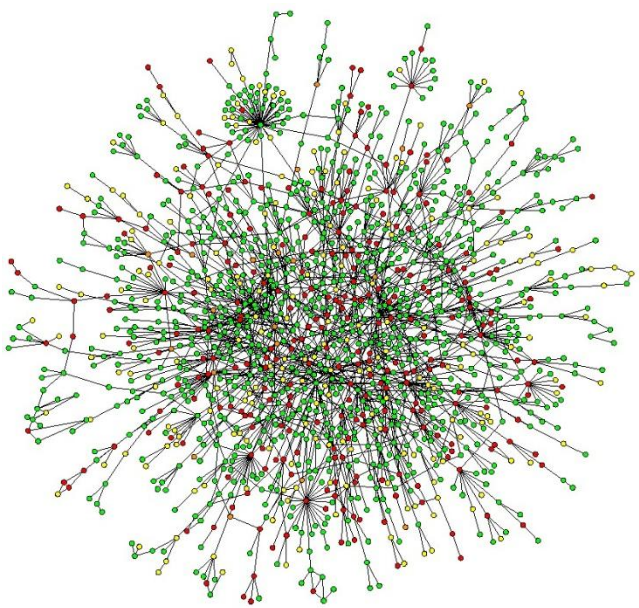
$$\dot{\phi}_i = \begin{cases} \omega_i + \frac{d}{(k_i+k_{p_i})} \sum_{j=1}^{N} a_{ij} \sin(\phi_j - \phi_i) \\ + \frac{d_p k_{p_i}}{(k_i+k_{p_i})} \sin(\phi_{p_i} - \phi_i), \end{cases}$$

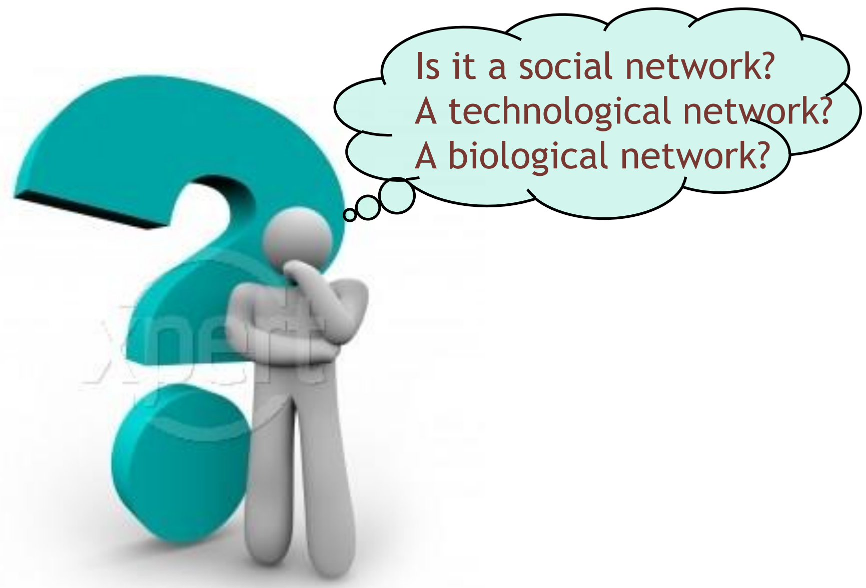


In this case, we have to study the influence of the topology in the dynamical processes occurring in the network (synchronization, stochastic processes, etc..) … … and vice-versa!

1.2.- TYPES OF NETWORKS

- Despite the different types of networks, which in turn are obtained from completely different interacting systems (people, neurons, proteins, routers,...) we will see that they share some universal properties



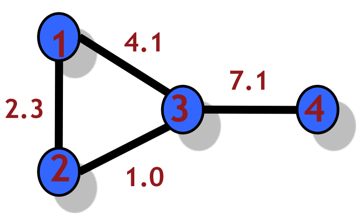


1.3.- Basic concepts about networks

Universidad Rey Juan Carlos

❑ Adjacency, Weights and Laplacian Matrix:

All the former networks can be described using a matricial formalism. Given a set of *N* nodes with *M* conections between them:



Weights Matrix (*W*):

Entries of the matrix are the weights w_{ij} (i,j=1, ..., N) of the connections

$$\begin{pmatrix} 0.0 & 2.3 & 4.1 & 0.0 \\ 2.3 & 0.0 & 1.0 & 0.0 \\ 4.1 & 1.0 & 0.0 & 7.1 \\ 0.0 & 0.0 & 7.1 & 0.0 \end{pmatrix}$$

Adjacency Matrix (*A*):

$a_{ij}=1$ if there exists a link between i and j, and $a_{ij}=0$ otherwise

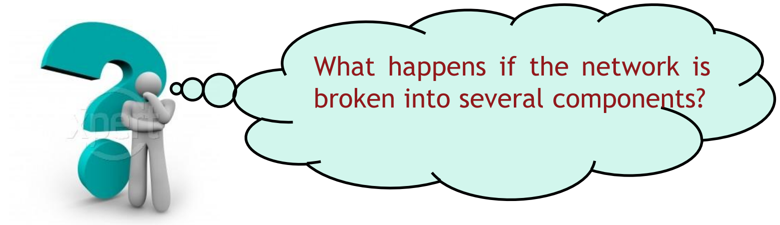
$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Laplacian Matrix (*L*):

The Laplacian matrix is defined as $L=K-A$, where K is a diagonal matrix of elements $k_{ii}=\sum a_{ij}$. Thus, it has a zero-row sum.

$$\begin{pmatrix} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -3 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Matrices will be symmetric if networks are undirected.



❑ Shortest path, average path length and diameter:

Shortest path (d_{ij}):

The shortest path d_{ij} between nodes *i* and *j* corresponds to the minimal distance (or weight) between all paths that connect *i* and *j*

Average path length (*l*):

The average path length *l* is the average shortest path between all nodes in the network:

$$\ell = \langle d_{ij} \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$$

when the network is not connected it is usefull to define the “harmonic mean”

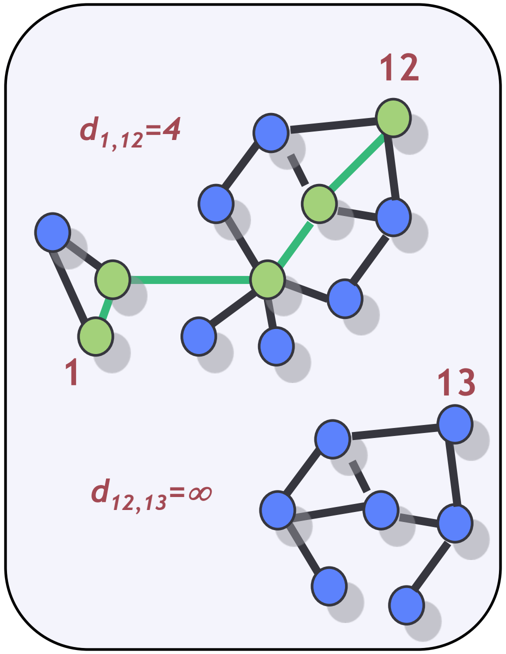
$$\ell = \frac{1}{\langle d_{ij}^{-1} \rangle} = \left(\frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \right)^{-1}$$

Diameter(*D*):

The maximum between all shortest paths $D=max(d_{ij})$

Component:

The set of nodes reachable from a given node.



1.3.- BASIC CONCEPTS ABOUT NETWORKS

Is the node with the highest degree, the one with highest strength?

and, which node has the highest betweennes?

❑ Degree, strength and betweenness:

Degree (k_i):

The degree k_i of a node i is the number of connections of the node

Strength (s_i):

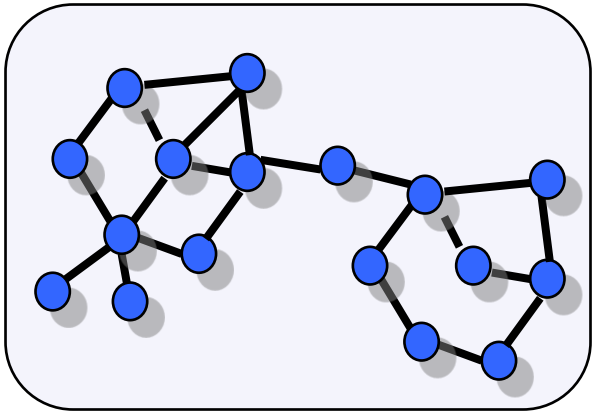
The strength s_i of a node i is the sum of the weigths of the connections to that node $s_i = \sum w_{ij}$

Betweenness (b_i):

The betweeennes of a node i (or a link) accounts for the number of shortest paths passing through that node (or link).

$$b_i = \sum_{j,k \in \mathcal{N}, j \neq k} \frac{n_{jk}(i)}{n_{jk}}$$

where n_{jk} is the number of shortest paths connecting j and k, and $n_{jk}(i)$ are those shortest paths between j and k that pass through i.

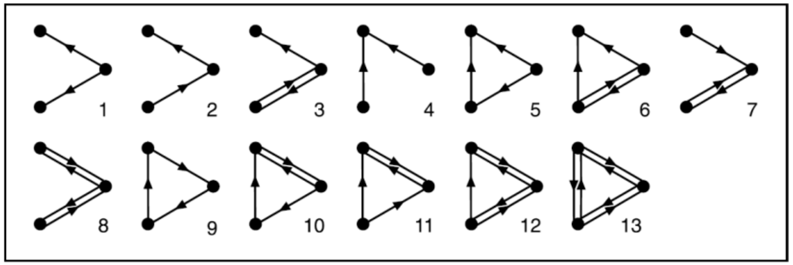


CSG Universidad Rey Juan Carlos

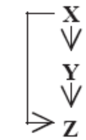
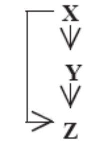
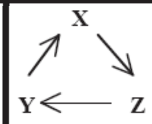
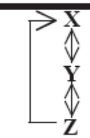
Network Motifs:

Network motifs are patterns (sub-graphs) that recur within a network much more often than expected at random.

Example: all 13 types of three-node connected subgraphs:



Each network motif can carry out specific information-processing functions

Network	Nodes	Edges	N_{real}	$N_{rand} \pm SD$	Z score
Gene regulation (transcription)					**Feed-forward loop**
E. coli	424	519	40	7 ± 3	10
*S. cerevisiae**	685	1,052	70	11 ± 4	14
Neurons					**Feed-forward loop**
C. elegans†	252	509	125	90 ± 10	3.7
Electronic circuits (digital fractional multipliers)					**Three-node feedback loop**
s208	122	189	10	1 ± 1	9
s420	252	399	20	1 ± 1	18
s838‡	512	819	40	1 ± 1	38
World Wide Web					**Feedback with two mutual dyads**
nd.edu§	325,729	1.46e6	1.1e5	2e3 ± 1e2	800

Figures from: Milo et al., Science, 298, 824 (2002)

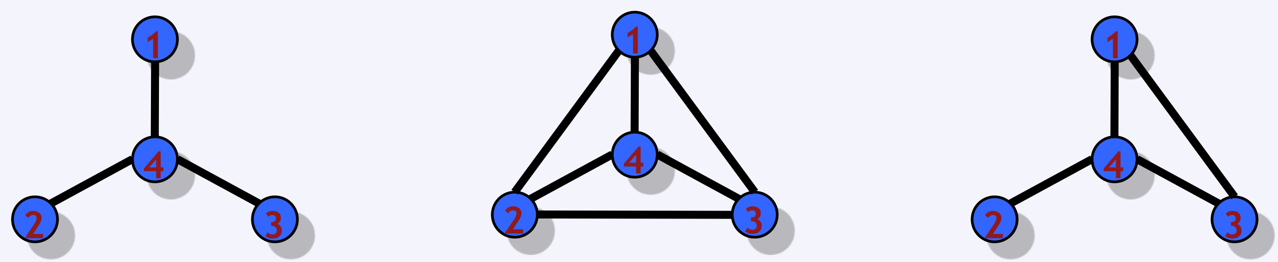
Universidad Rey Juan Carlos

❑ Clustering coefficient:

The clustering coefficient *C* accounts for the number of triangles in the network. Specifically, C_i is the ratio between the number of links *E* connecting the nearest neighbors of *i* and the total number of possible links between these neighbors.

$$C_i = \frac{2E}{k_i(k_i - 1)}$$

The clustering coefficient of the network *C* is the average of C_i over all nodes.



$C_{1,2,3,4} = \{0,0,0,0\}$
$C=0$

$C_{1,2,3,4} = \{1,1,1,1\}$
$C=1$

$C_{1,2,3,4} = \{1,0,1,1/3\}$
$C=7/12$

Local and Global Efficiency:

The efficiency overcomes the divergence of the shortest paths if the graph is disconnected

Global Efficency (E):

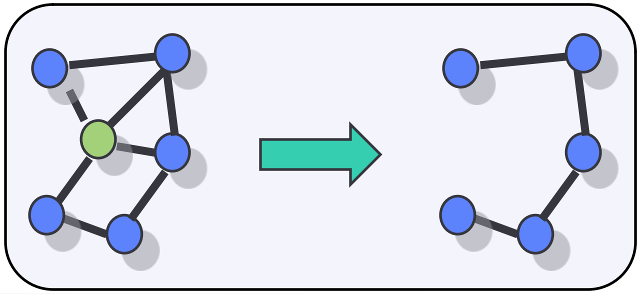
The global efficiency is the harmonic mean of the geodesic paths between all nodes of the network:

$$E = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} \frac{1}{d_{ij}}$$

Local Efficency (E_i):

The local efficiency E_i of a node i, measures the shortest path length between the subset G_i of neighbors of the node i, when i is not present.

$$E_{\text{loc}} = \frac{1}{N} \sum_{i \in \mathcal{N}} E(G_i)$$



The local efficiency is related, somehow, with the clustering cofficient

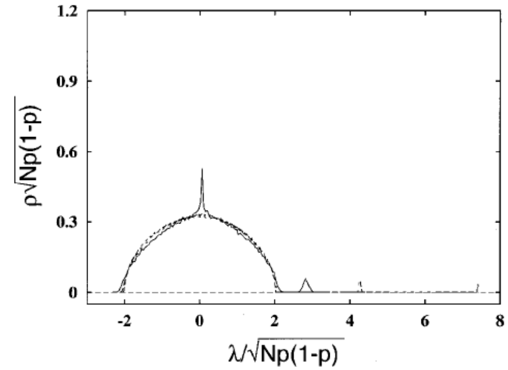
Graph Spectrum:

The spectrum of a graph is the set of eigenvalues of its adjacency (or Laplacian) matrix *A*. A graph $G_{N,M}$, has *N* eigenvalues $\mu_i=(\mu_1, \mu_2,\ldots, \mu_N)$ and *N* associated eigenvectors $v_i=(v_1,v_2,\ldots, v_N)$.

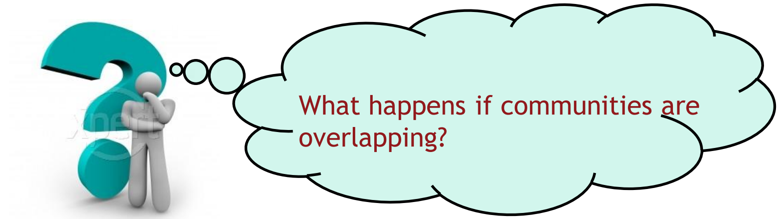
The eigenvalues and associated eigenvectors of a graph are intimately related to important topological features such as the diameter, the number of cycles, information transmission and the connectivity properties of the graph.

Spectral density:

$$\rho(\mu) = \frac{1}{N}\sum_{i=1}^{N}\delta(\mu-\mu_i)$$



Rescaled spectral density of three random graphs having p=0.05 and size N=100, N=300, and N=1000.The isolated peak corresponds to the principal eigenvalue.



1.3.- BASIC CONCEPTS ABOUT NETWORKS

❑ Community Structure (I):

Given a graph $G_{N,M}$, a community is a subgraph $G'_{N',M'}$ *whose nodes are thightly connected (or at least, more connected than in a random equivalent network).*

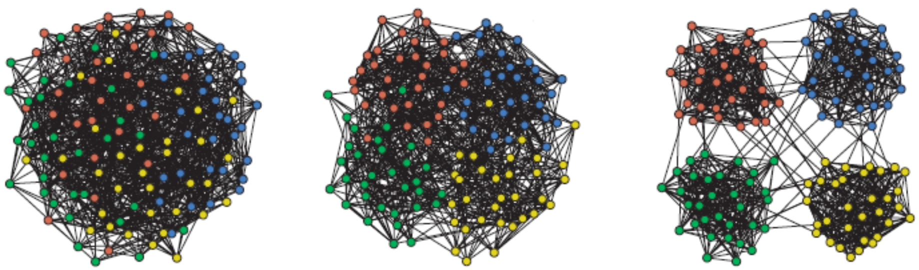


Figure from: Guimerà et al., Nature, 433, 895(2005)

Zachary Karate Club

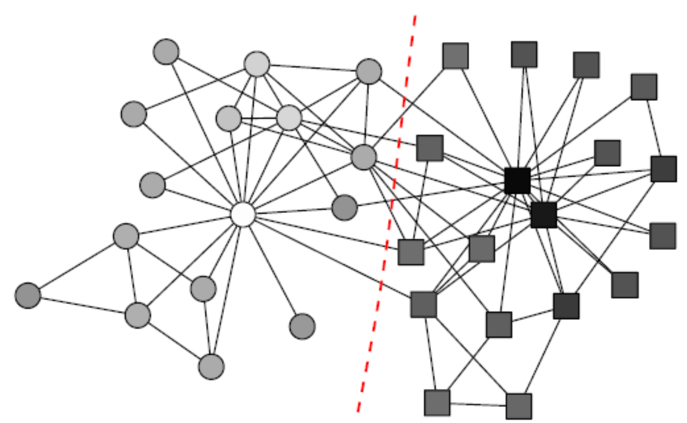


Figure from: M. E. J. Newman, Proc. Natl. Acad. Sci. USA 103, 8577 (2006)

❑ Community Structure (II):

Several algorithms have been proposed in order to split a sparse network into communities:

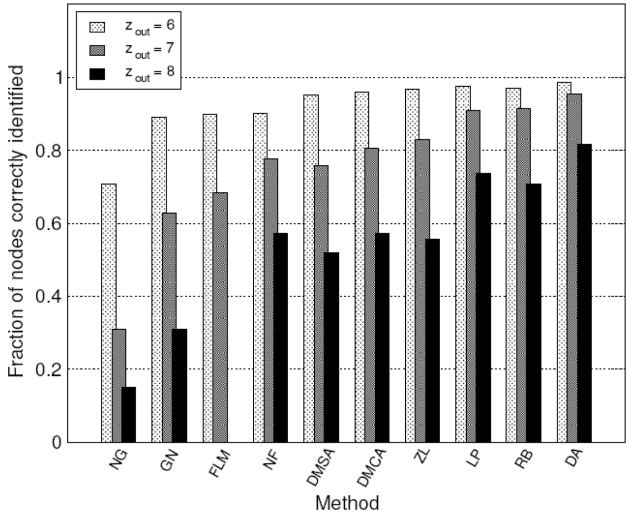


Figure from: L. Danon et al., World Scientific, 93-113 (2007)

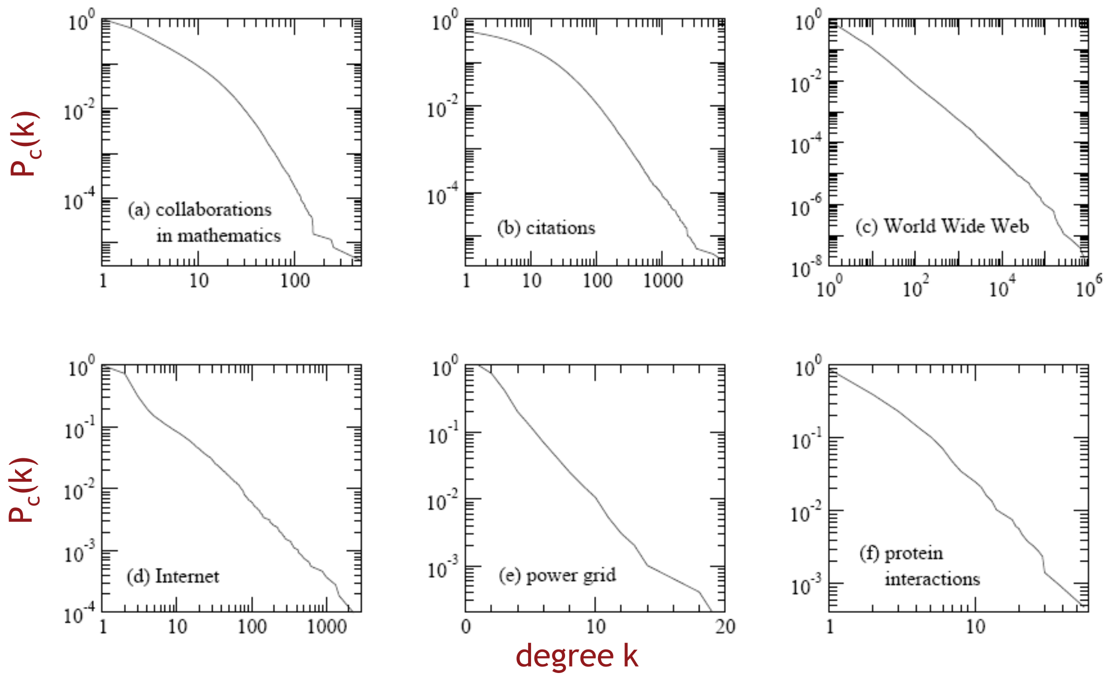
Modularity *M* is and objective measure in order to evaluate community division:

$$M \equiv \sum_{s=1}^{N_M} \left[\frac{l_s}{L} - \left(\frac{d_s}{2L} \right)^2 \right]$$

where N_M is the number of modules, L is the number of links in the network, l_s is the number of links between nodes in module s, and d_s is the sum of the degrees of the nodes in module s.

❑ Degree Distribution (I):

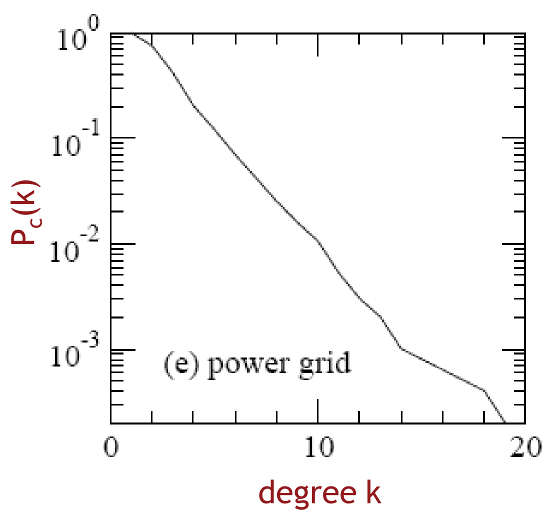
The [cumulative] degree distribution [$P_c(k)$] *p(k)* accounts for the fraction of nodes in the network with a degree [higher than] equal to *k*.



❑ Degree Distribution (II):

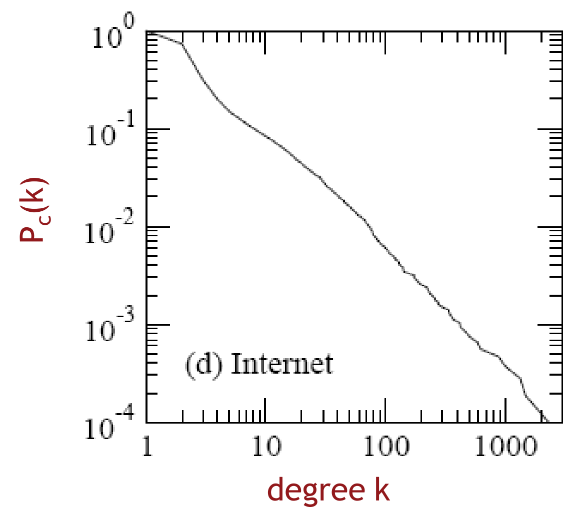
Two types of degree distribution appear more frequently in real networks :

Exponential decay: $P_c(k) \sim e^{-\alpha k}$

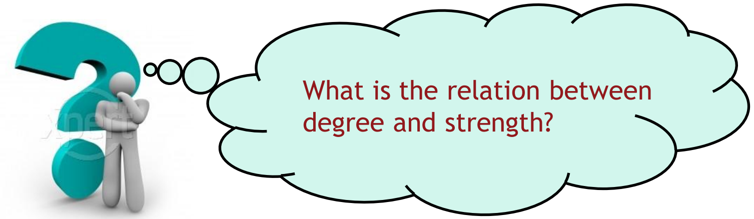


Typical in random networks

Power-law decay: $P_c(k) \sim k^{-\gamma}$



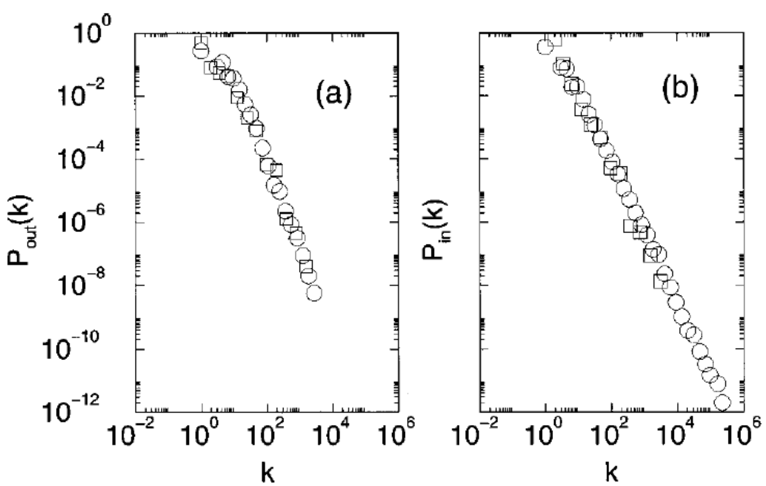
Networks with power-law decay are called scale-free networks.



❑ Degree Distribution (III):

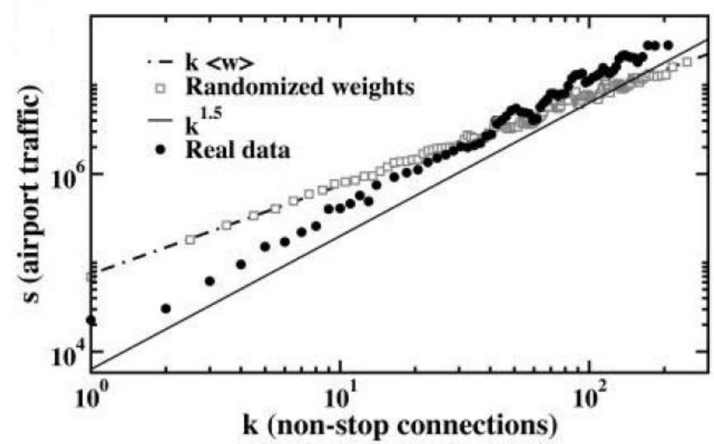
Other related distributions are:

In/out degree distributions
(directed networks)



In/out degree distributions of WWW (from two different samples: 325.729 and 200.000.000 nodes). From R. Albert et al., Rev. Mod. Phys. 74, 47 (2002).

Strength distribution
(weighted networks)

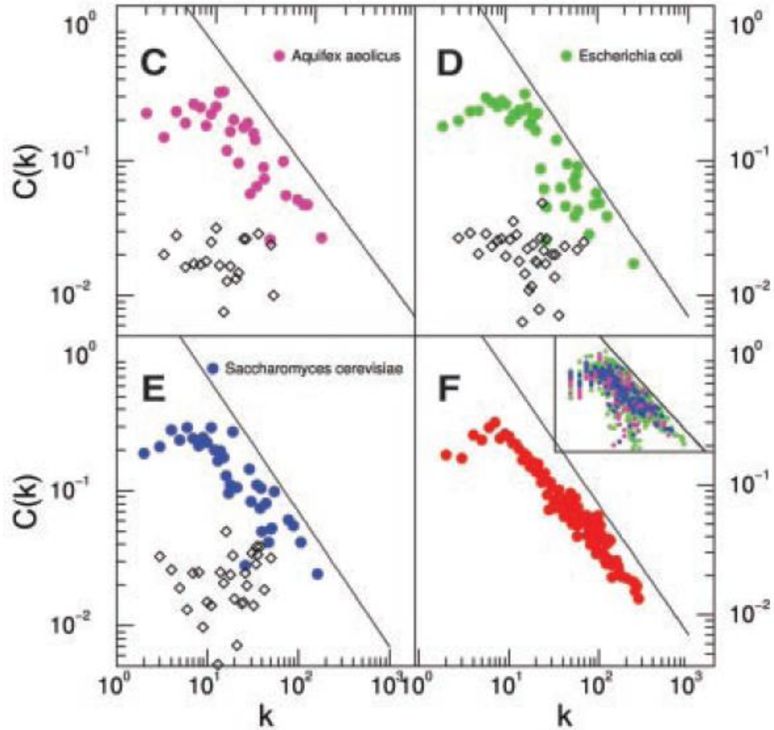


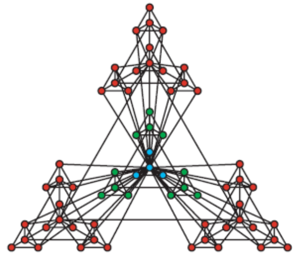
Strength distribution of the International Air Transportation Network (www.iata.org). From A. Barrat et al., PNAS, 101, 3747 (2004).

❑ Clustering Distribution C(k):

The clustering distribution has been related with the modularity and hierarchy of the network:

Figure: Clustering distribution in three organisms: Aquidex aeolicus (archaea) (C), Escherichia coli (bacterium) (D), and Saccharomyces cerevisiae (eukaryote) (E). (F) The C(k) curves averaged over all 43 organisms is shown, and the inset displays all 43 species together. Lines correspond to $C(k) \sim k^{-1}$, and diamonds represent the C(k) value expected for an equivalent scale-free network, indicating the absence of scaling



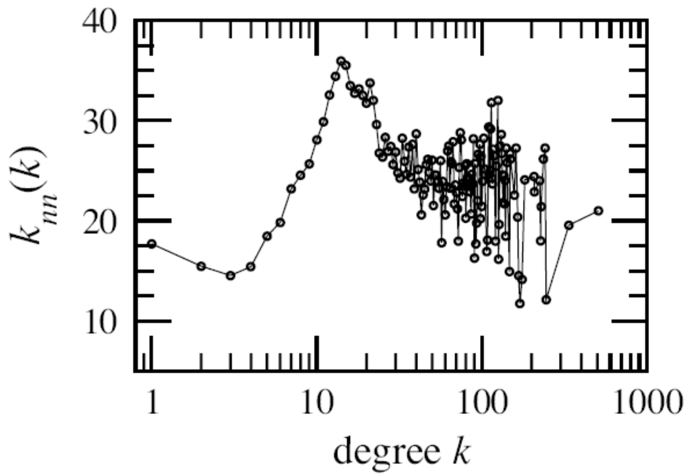


From E. Ravasz et al., Science, 297, 1551 (2002).

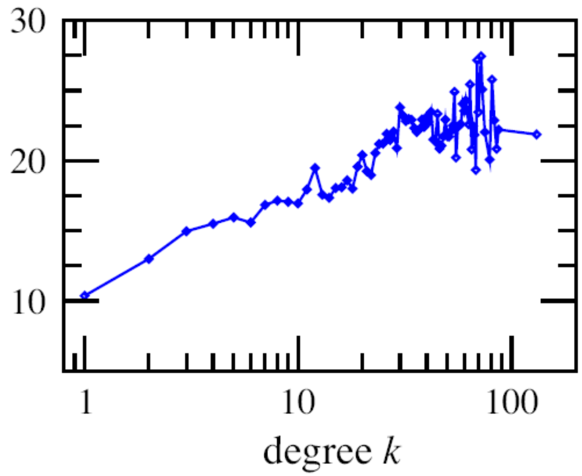
❑ Nearest neighbor degree $k_{nn}(k)$ and assortativity

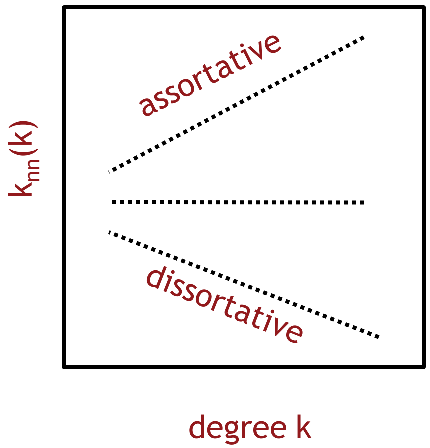
The $k_{nn}(k)$ distribution measures the degree of the nearest neighbors. It is an indicator of the **assortativity** of the network.

Music collaboration network



Music similarity network





Colaboration and similarity network obtained from a music database (AllMusic Guide).
From J. Park et al., IJBC, 17, 2281 (2007).

1.4.- Brief historical background

1.4.- BRIEF HISTORICAL BRACKGROUND

❑ Leonard Euler (Basel 1707 – St. Petersburg 1783)

Some revealing data about Leo:

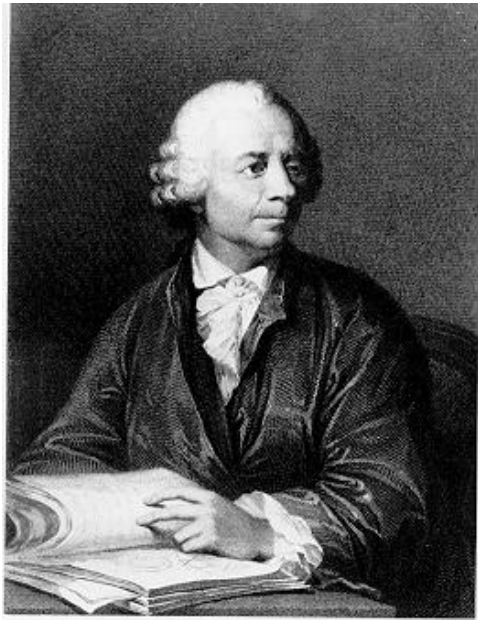
❑ Euler worked in almost all areas of mathematics: geometry, calculus, trigonometry, algebra, and number theory, as well as continuum physics, lunar theory and other areas of physics.

❑ Large number of topics of physics and mathematics are named in his honour (e.g., Eulers's function, Euler's Equation or Euler's formula).

❑ All his work is collected in *Opera Omnia*, which consists of 886 books.

❑ With one eye from 1738 and completely blind from 1766!

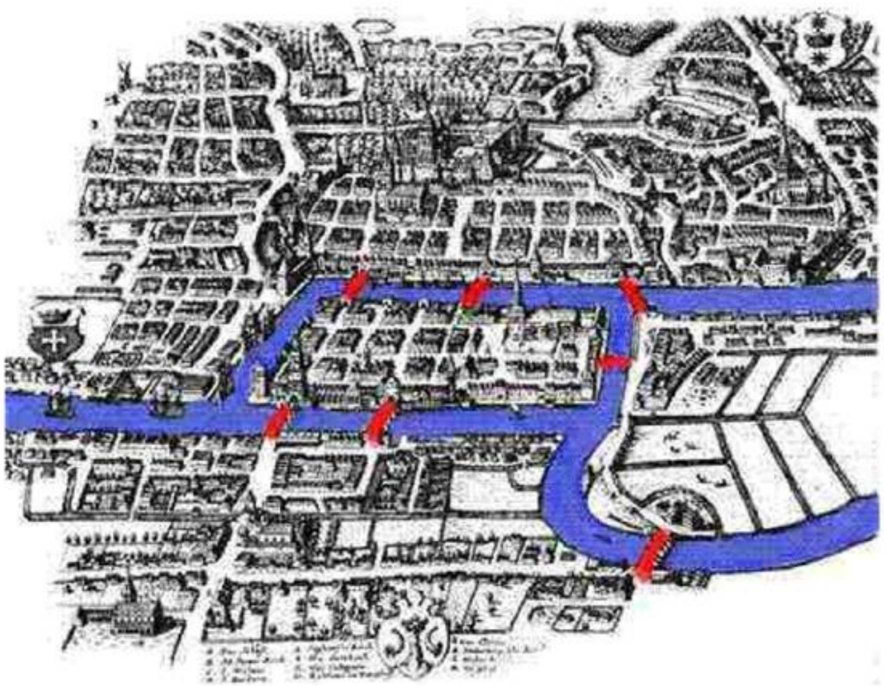
❑ And the most atonishing data: all of that with **13 children!**

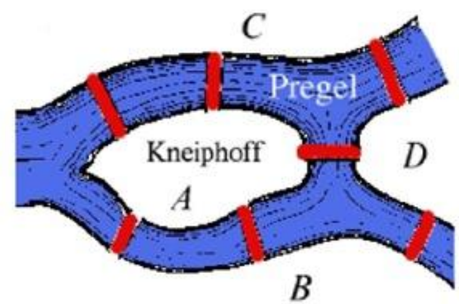


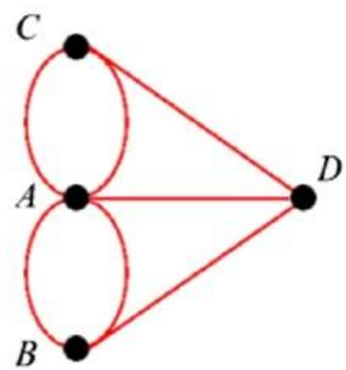
CSG Universidad Rey Juan Carlos

❑ Euler, the father of graph theory:

The seven bridges of Konïgsberg and the origin of graph theory:
Is it possible to cross the seven bridges only once?







Euler's Solution:

N_0 = Number of nodes with odd degree

1.- If $N_0>2$, no solution.

2.- If $N_0=2$, only one solution starting from one of the odd nodes.

1.- If $N_0 < 2$, there are solutions starting from any node.

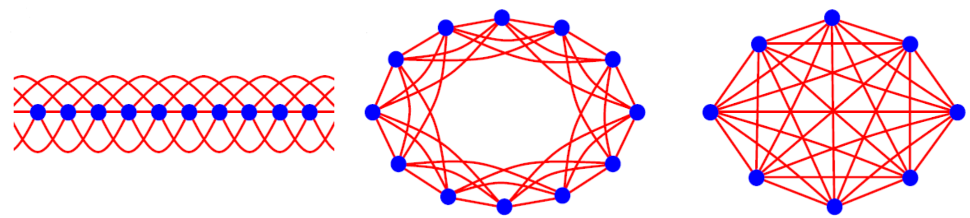
❑ Regular Graphs

- ❑ After the death of Euler, graph theory received many contributions from mathematicians such as Hamilton, Kirchhoff or Cayley.

- ❑ The core of graph theory focused on the study of regular graphs:

Regular graph: a graph where all nodes have the same degree.

Lattice: a regular network where all nodes are coupled to its nearest neighbor.



N = number of nodes
K = degree
C = clustering coefficient
d = dimension of the lattice
l = average path length

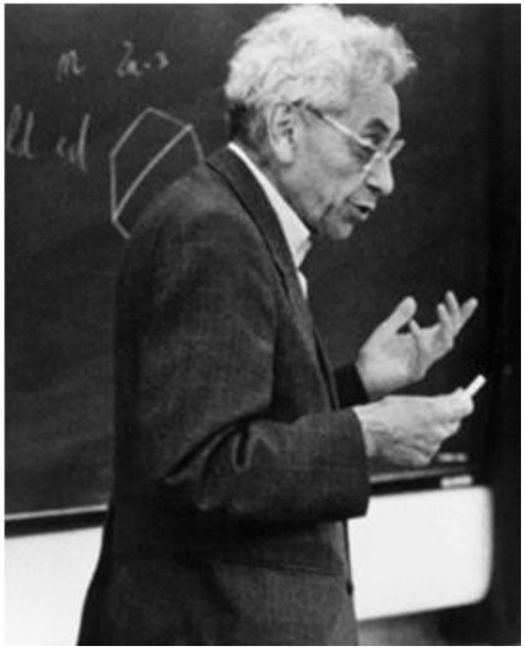
$$C = \frac{3(K-2d)}{4(K-d)}$$

(if $K < 2N/3$)

$$\ell \sim \sqrt[d]{\frac{N}{K}}$$

❑ Paul Erdös (Budapest 1913 – Warsaw 1996)

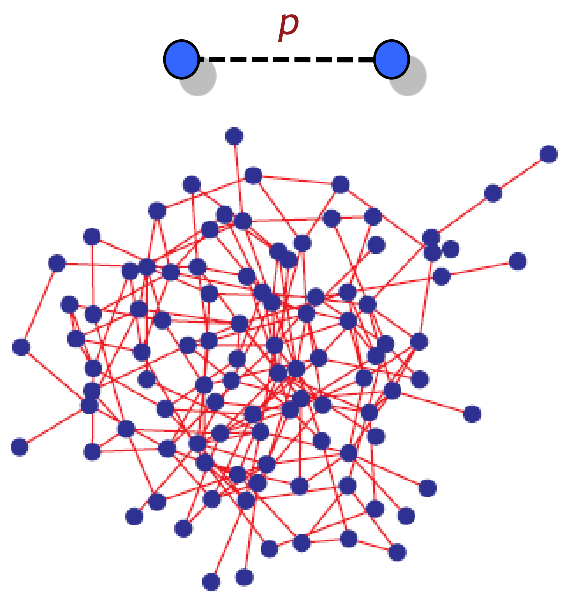
Some revealing data about Paul:



- ❑ Seminal contributions in combinatorics, graph theory, number theory, classical analysis, approximation theory, set theory, and probability theory.
- ❑ Paul wrote 1475 papers and collaborated with 511 scientists.
- ❑ Excentric person, he had an special vocabulary (children="epsilons", women="bosses", U.S="samland", etc...)
- ❑ Paul offered small prizes for solutions to unresolved problems (from 25$ to some thousands), and there are still open problems!
- ❑ "You don't have to believe in God, but you should believe in The Book." (he recognized that he took amphetamines)

Paul Erdös and Alfred Rényi

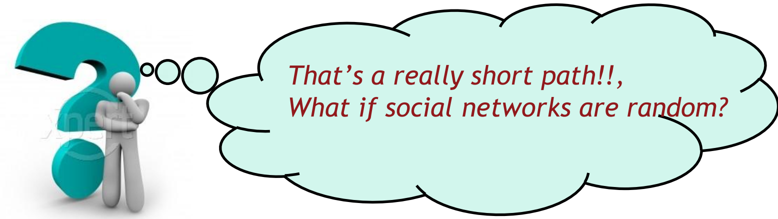
They worked on the analysis of social networks by finding analogies with the so-called ***random graphs***, in which the existence of a link between a pair of nodes has a probability *p*.



N = number of nodes
<*k*> = mean degree
<L> = number of random connections
p = probability of connection between two nodes

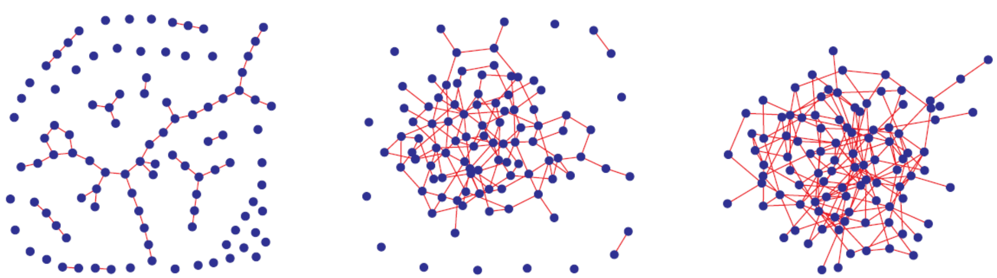
Mean degree of the network → $\langle k \rangle = p(N-1) \cong pN$

Number of random connections → $\langle L \rangle = \frac{1}{2}\, pN(N-1) \cong \frac{1}{2} \langle k \rangle N$



❑ Emergence of a giant component

When propability *p* crosses a critical value p_c, there emerge a giant component that contains and extensive fraction of the nodes in the network



p increases →

Critical probability (N→ ∞):

$$p_c \sim \frac{\ln N}{N}$$

Critical mean degree:

$$\langle k \rangle_c \sim \ln N$$

Clustering coefficient:

$$\overline{C} = p \cong \frac{\langle k \rangle}{N} \ll 1$$

N=1000
<k>=2
C ~ 0.002

Average shortest path:

$$\ell \sim \frac{\ln N}{\ln \langle k \rangle}$$

N=1000000
<k>=5
l ~ 8.6

❑ Stanley Milgram (New York 1933 – New York 1984)

Stanley Milgram was an American social psychologist most notable for his controversial studies on the obedience to authority.

Some Stanley's famous experiments:

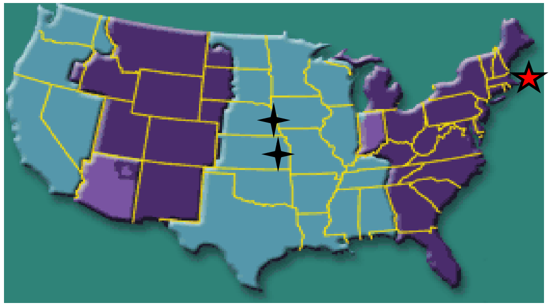
- ❑ The Milgram experiment 18

- ❑ The lost-letter experiment

- ❑ The small-world experiment

❑ The small-world experiment

A group of people from Omaha (Nebraska) and Wichita (Kansas) was asked to send a letter to an unknown person in Boston (Massachussetts).

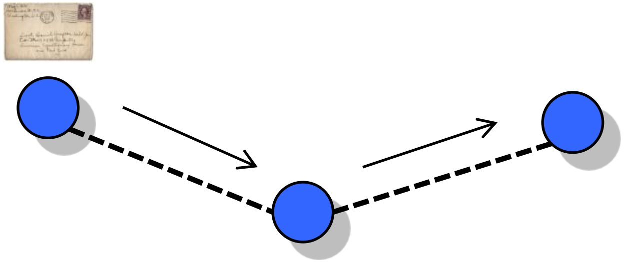


Basic Rule of the experiment:

- ❑ People should forward the letter to a person that they consider closer to the target person

Results of one experiment (in fact, there where several!):

- ❑ 232 out of 296 letters never reached the target
- ❑ 64 letters reached the target (with paths from 2 to 10)
- ❑ The average path length was …. 5.2 (steps)



- It's a small world! (que pequeño es el mundo!)

or in other words:

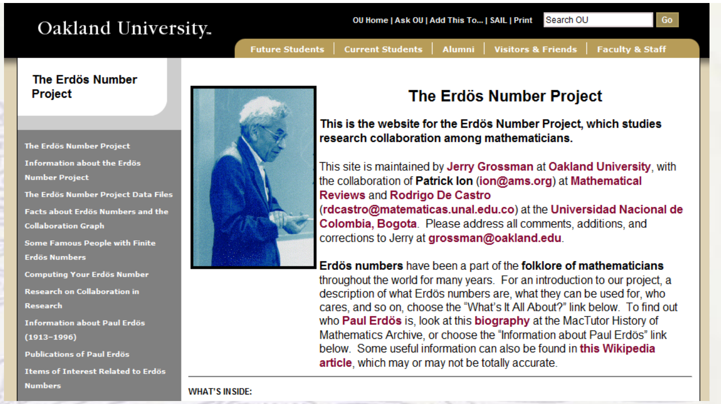
$d_{ij} << N$

1.4.- BRIEF HISTORICAL BRACKGROUND

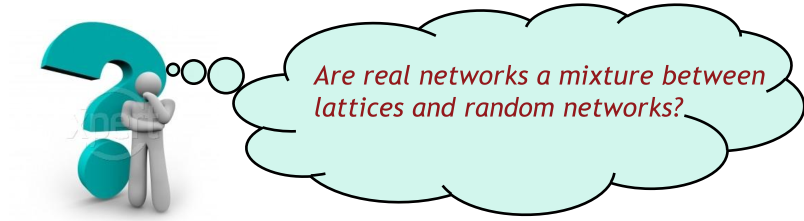
- **Let's go back to Erdös:**

  You can measure the distance with Paul Erdös. (http://www.oakland.edu/enp/)

  - Mean Erdös number: ~5
  - Largest Erdös number: ~13



Max von Laue	1914	4
Albert Einstein	1921	2
Niels Bohr	1922	5
Louis de Broglie	1929	5
Werner Heisenberg	1932	4
Paul A. Dirac	1933	4
Erwin Schrödinger	1933	8
Enrico Fermi	1938	3
Ernest O. Lawrence	1939	6
Otto Stern	1943	3
Isidor I. Rabi	1944	4
Wolfgang Pauli	1945	3
Frits Zernike	1953	6
Max Born	1954	3
Willis E. Lamb	1955	3



1.4.- BRIEF HISTORICAL BRACKGROUND

❑ It's a small world everywhere!

The small-world property has been reported in a large number of real networks of different origin.

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rnd}	C	C_{rnd}
1. Movie actors	225 226	61.0	3.65	2.99	0.79	0.00027
2. Power grid	4 941	2.67	18.7	12.4	0.08	0.00054
3. WWW site level (undir.)	153 127	35.2	3.10	3.35	0.11	0.00023
4. Words (co–ocurrence)	460 902	70.1	2.67	3.03	0.44	0.00015
5. LANL co–authorship	52 909	9.70	5.90	4.79	0.43	0.00018
6. MEDLINE co–authorship	1 520 251	18.1	4.60	4.91	0.07	0.00001
7. Math. co–authorship	70 975	3.90	9.50	8.21	0.59	0.00005

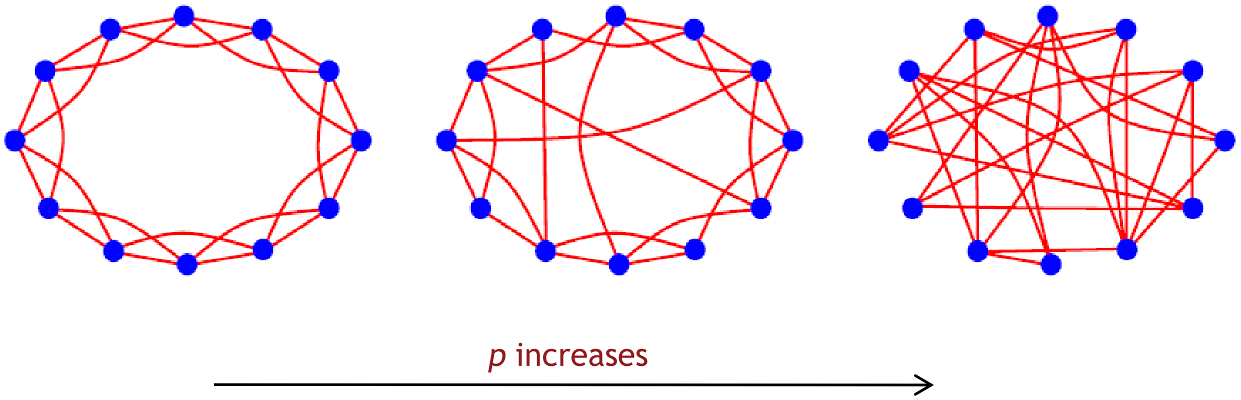
Average path length and clustering coefficient of some real networks. We compare the values in the real network with those of equivalent random networks

The average path length is similar in random networks (where $l \sim \ln N$) but the clustering coefficient is some orders of magnitude higher (and closer to the clustering coefficient of a lattice!).

❑ The Watts-Strogatz model (I)

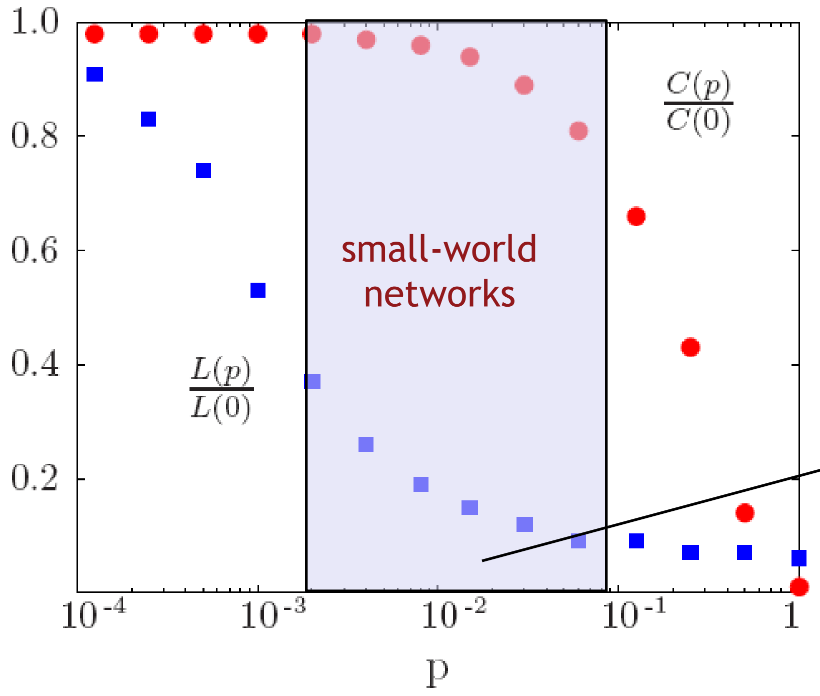
Watts and Strogatz (PRL 1998) proposed a network model that conciliated the high clustering and short average path length of real networks

Starting from a regular ring, a certain (random) rewiring is introduced with a probability *p*



❑ The Watts-Strogatz model (II)

Small-world networks are characterized by a low average shortest path and high clustering

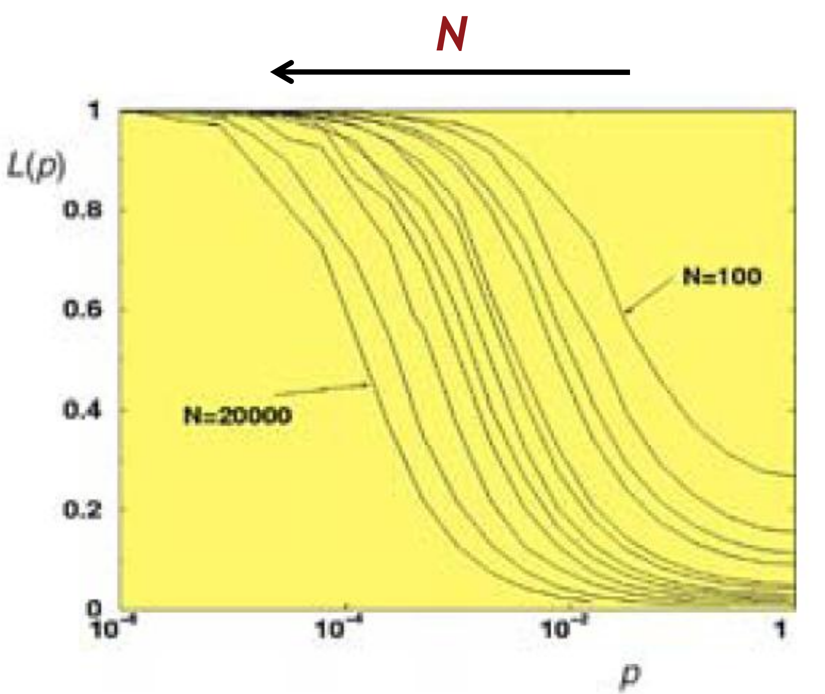


A low number of "shortcuts" reduces the distance between nodes without modifying the local properties

CSG Universidad Rey Juan Carlos

❑ The Watts-Strogatz model (II)

The larger the network, the higher probability to be small-wolrd.



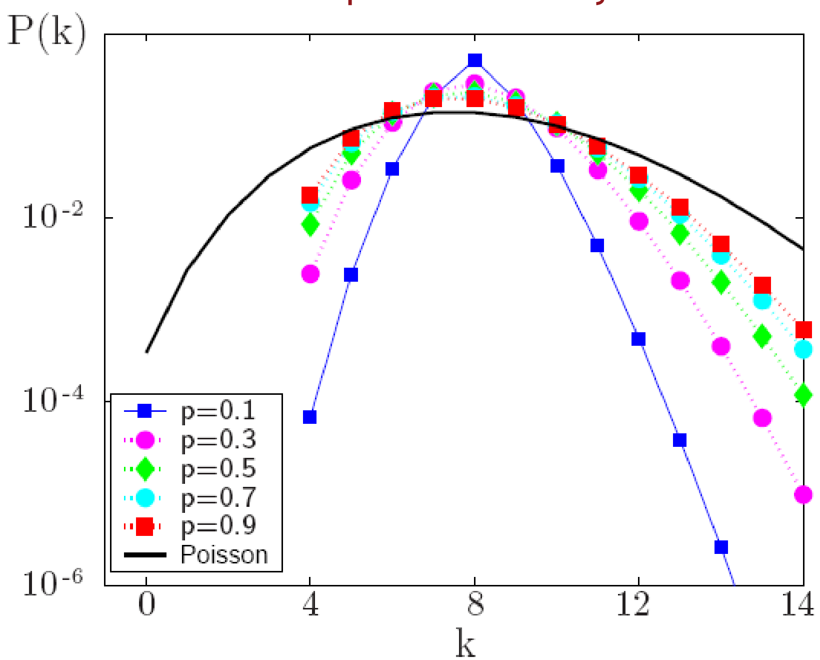
The rewiring of the links in order to entre the small world-region goes with:

$$p \sim 1/N$$

Figure from Barthelemy, PRL, 82,3180 (1999)

❑ The Watts-Strogatz model (III)

The probability degree distribution p(k) of WS small-world networks shows a pronounced peak around <k> and exponential decay

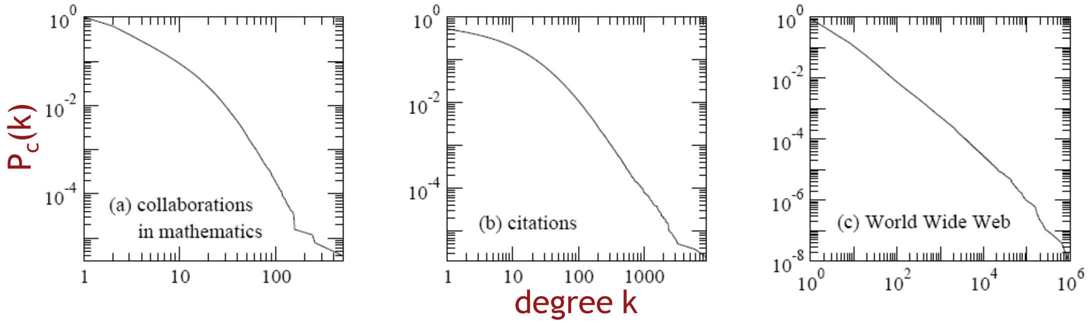


Degree distribution of the WS model for <*k*>=*8* and different rewiring probabilities

Networks obtained with the WS model are *"exponential networks"*

Scale-free networks (I)

Unfortunatelly (or luckily!) many real networks are not exponential. On the contrary, they have a power-law decay (i.e., $P(k) \sim k^{-\gamma}$).



- Scale-free networks have power law decays $P(k) \sim k^{-\gamma}$
- Power laws are relatively slow decreasing functions (the probability of having highly connected nodes is much higher than in exponential networks).
- A power-law distribution has no peak at its average value (no characteristic scale).

Universidad Rey Juan Carlos

❑ Scale-free networks (II)

NETWORK	SIZE	γ_{in}/γ_{out}
1. Movie actors [57]	212 250	2.3
2. WWW [59]	$2 \cdot 10^8$	2.7/2.1
3. Internet, router [60]	260 000	—/1.94
4. Words (co–ocurrence) [13]	460 902	2.7
5. Neuro. co–authorship [61]	209 293	2.1
6. SPIRES co–authorship [48]	56 627	1.2
7. E-mail messages [62]	59 912	1.5/2.0
8. Metabollic network [63]	778	2.2

Real networks with scale-free structure. From Almendral, PhD. Thesis

Interestingly, the exponent of the power laws range from 1.2 to 3, with the majority between 2 and 3.

❑ The Barabási-Albert model (I)

They introduce a model in order to explain the origin of the power-law dstributions of real networks. A network is constructed from scratch following two fundamental rules:

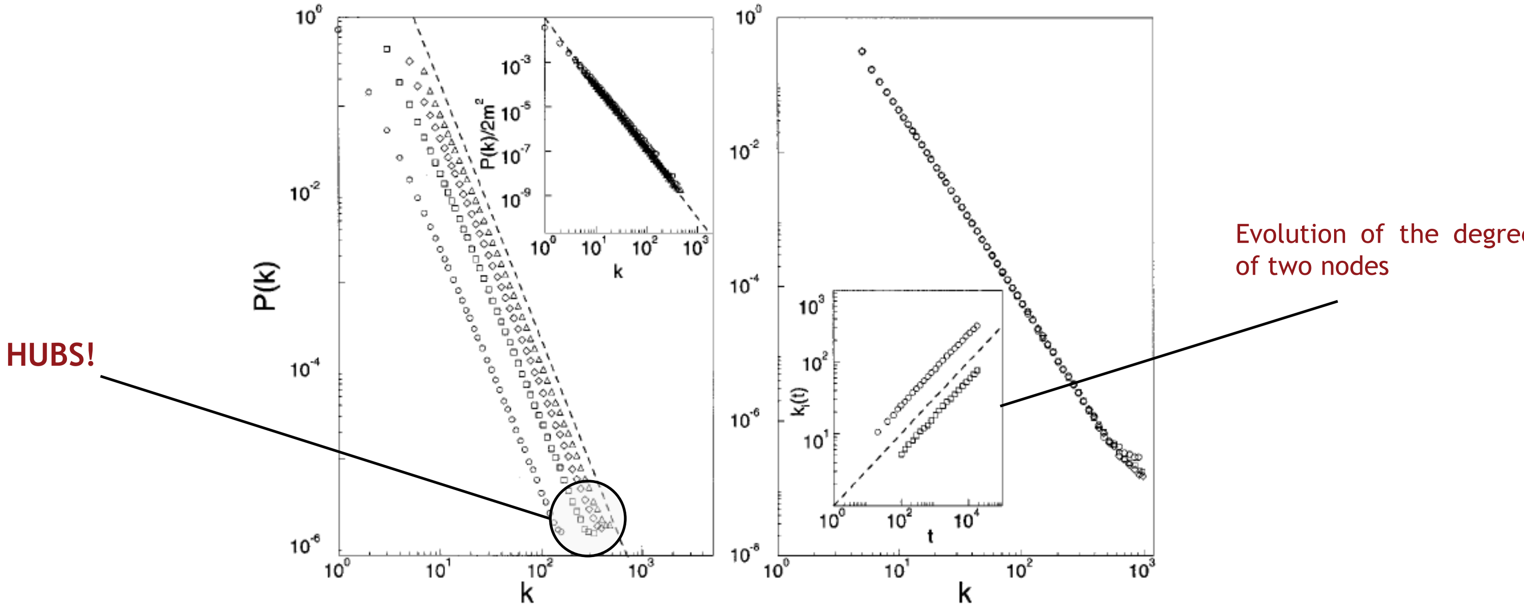
❑ *Growth. From an initial number of nodes N_0, new nodes are attached to the existing ones at discrete time steps. Thus, the number of nodes increases with time $N(t)= N_0+t$ and also the number of links $L(t)= mt$ (being m the number of links of each new node)*

❑ *Preferential attachment. The nodes to which the new node is attached are chosen following a preference function:*

$$p_i = \frac{k_i}{\sum_{j=1}^{N(t)} k_j}$$

❑ The Barabási-Albert model (II)

The BA model shows a power law decay independent of the number of links or the system size (with an exponent γ=3)

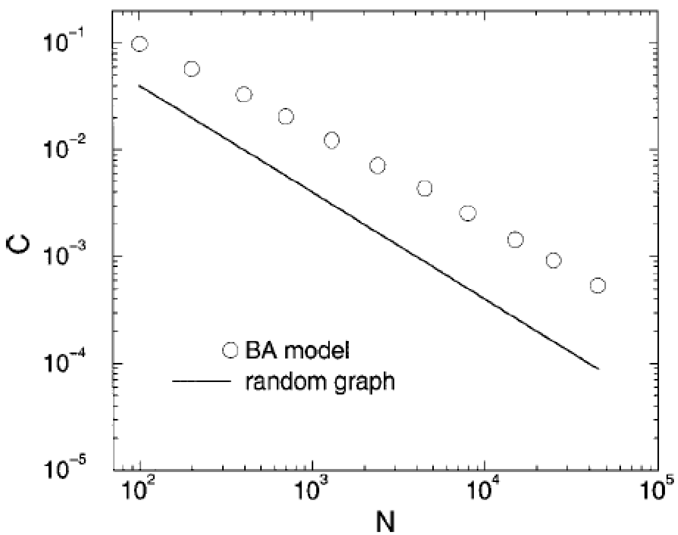


(Left) Degree distribution of the B-A model, with $N=m_0+t=300000$ and m_0=1,3,5,7. The dashed lines correspond to $P(k)=k^{-2.9}$. (Right) P(k) for m_0=5 and different systems size: m=100000, 150000 and 200000. From R. Albert et al., Rev. Mod. Phys. 74, 47(2002).

Are networks obtained with the BA model small world networks?

❑ The Barabási-Albert model (III)

As in random networks, the clustering coefficient obtained with the BA model is low



Clustering coefficient *C* of the network as a function of the system size *N*. From R. Albert et al., Rev. Mod. Phys. 74, 47(2002)

	network	type	n	m	C	ℓ
social	film actors	undirected	449 913	25 516 482	0.78	3.48
	company directors	undirected	7 673	55 392	0.88	4.60
	math coauthorship	undirected	253 339	496 489	0.34	7.57
	physics coauthorship	undirected	52 909	245 300	0.56	6.19
	biology coauthorship	undirected	1 520 251	11 803 064	0.60	4.92
	telephone call graph	undirected	47 000 000	80 000 000		
	email messages	directed	59 912	86 300	0.16	4.95
	email address books	directed	16 881	57 029	0.13	5.22
	student relationships	undirected	573	477	0.001	16.01
	sexual contacts	undirected	2 810			
technological	Internet	undirected	10 697	31 992	0.39	3.31
	power grid	undirected	4 941	6 594	0.080	18.99
	train routes	undirected	587	19 603	0.69	2.16
	software packages	directed	1 439	1 723	0.082	2.42
	software classes	directed	1 377	2 213	0.012	1.51
	electronic circuits	undirected	24 097	53 248	0.030	11.05
	peer-to-peer network	undirected	880	1 296	0.011	4.28
biological	metabolic network	undirected	765	3 686	0.67	2.56
	protein interactions	undirected	2 115	2 240	0.071	6.80
	marine food web	directed	135	598	0.23	2.05
	freshwater food web	directed	92	997	0.087	1.90

Clustering coefficient *C* and average path length of some real networks. From Newman, SIAM Rev, 45, 167 (2003)

❑ The Barabási-Albert model (IV)

Attractiveness, aging, capacity, … can modify the scale free behaviour of the BA model.

The Dorogovtsev-Mendes -Samukhin model

$$\prod_{j\to i} = \frac{k_i + k_0}{\sum_l (k_l + k_0)}$$

k_0= initial attractiveness
($-m < k_0 < \infty$)
m= number of new links

$$\gamma = 3 + k_0/m$$

$(2 < \gamma < \infty)$

Dorogovtsev et al.,
PRL 85 4633 (2000)

The Kaprivsky et al. model

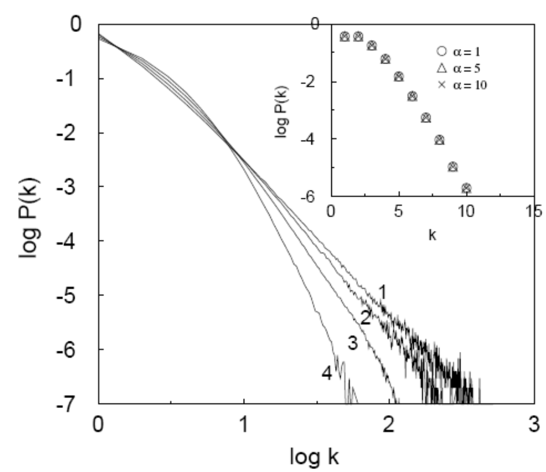
$$\prod_{j\to i} = \frac{k_i^\alpha}{\sum_l k_l^\alpha}$$

$\alpha < 1$: streched exponential decay
$\alpha > 1$: a single node dominates

Krapivsky et al.,
PRL, 4629 85 (2000)

The Dorogovtsev-Mendes model

Probability of linking depends on $\tau^{-\alpha}$
(being τ the age of the node)

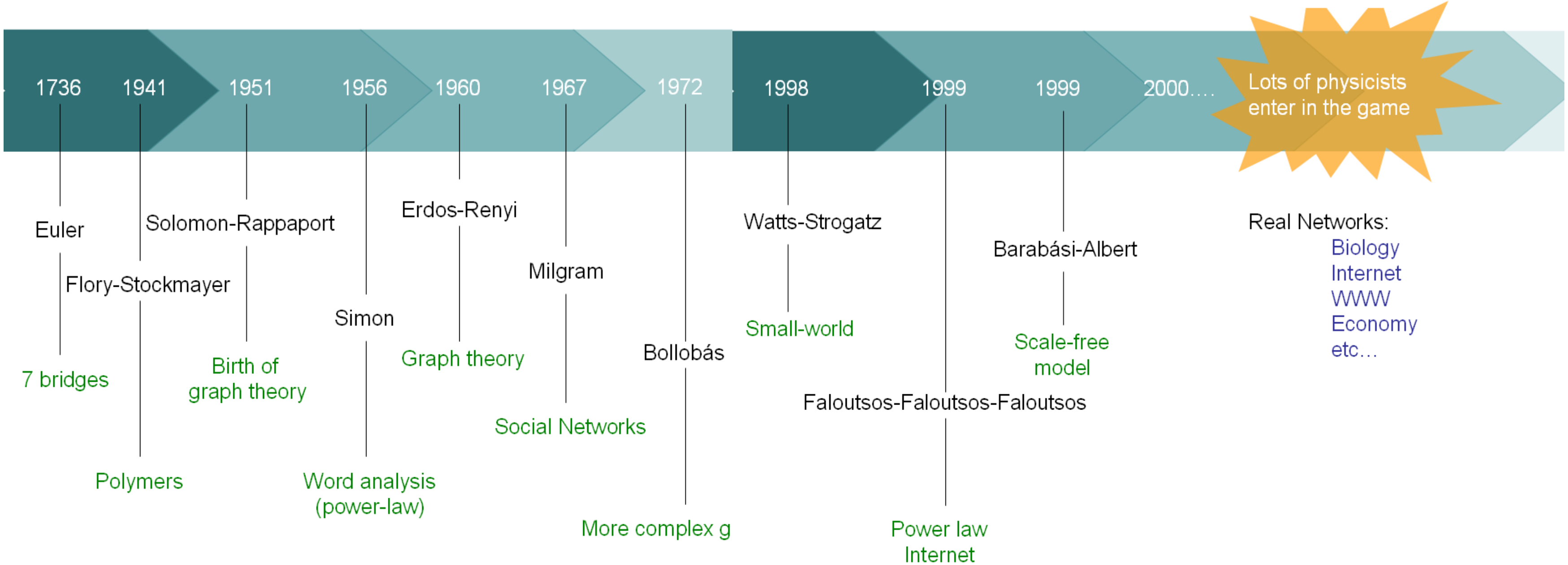


Probability distribution for serveral aging exponents:
1) 0.2, 2) 0.25, 3) 0.5 and 4) 0.75. $\alpha>1$ exponential decay.
From PRE62, 1842 (2000)

Universidad Rey Juan Carlos

1.4.- BRIEF HISTORICAL BRACKGROUND

- Complex Networks time line:



(from J.F.F. Mendes presentation)

1.4.- BRIEF HISTORICAL BRACKGROUND

All of that is nice, but does it have applications in real networks?

Thanks for your attention

mañana más, pero no mejor!