Networks
an introduction

Fakhteh Ghanbarnejad

TU Berlin

Related courses

Mark Newman

Albert-László Barabási

Constantine Dovrolis

Mehmet Gunes

Central European University, Dep. of Network and Data Science

Textbooks

Networks: An Introduction, Mark Newman (Oxford University Press, 2010)
Network Science, Albert-Laszlo Barabasi (Cambridge University Press, 2016)
Dynamical Processes on Complex Networks,
A. Barrat, M. Barthelemy, A. Vespignani (Cambridge Uni. Press, 2008)
Dynamical Systems on Networks: A Tutorial, M. Porter, J. Gleeson (Springer, 2016)
Lectures on Complex Networks, SN. Dorogovtsev (Oxford Uni. Press, 2010)
Evolution of Networks, SN. Dorogovtsev, JFF. Mendes (Oxford Uni. Press, 2003)
The Structure and Dynamics of Networks,
M. E. J. Newman, A.-L. Barabasi, D. J. Watts (Princeton University Press, 2006)
Complex and Adaptive Dynamical Systems, Claudius Gros (Springer, 2015)
Networks, Crowds, and Markets: Reasoning About a Highly Connected World,
David Easley and Jon Kleinberg (Cambridge University Press, 2010) .
The Structure of Complex Networks Theory and Applications, E. Estrada (Oxford Uni. Press,
Exploratory Social Network Analysis with Pajek,
Nooy, Wouter de, Andrej Mrvar, and Vladimir Batagelj (Cambridge University Press, 2011)
The Structure and Function of Complex Networks, M. E. J. Newman (SIAM Review, 2003)
Random Graph Dynamics, Rick Durrett (Cambridge University Press, 2007)

Textbooks

Random Graphs and Complex Networks, Remco van der Hofstad (pdf notes, 2013)
Complex Networks: Structure, Robustness and Function,
Reuven Cohen and Shlomo Havlin (Cambridge University Press, 2010)
Scale-Free Networks: Complex Webs in Nature and Technology,
G. Caldarelli (Oxford University Press, 2007)
Evolution and Structure of the Internet,
R. Pastor-Satorras and A. Vespignani (Cambridge University Press, 2007)
Social Network Analysis, J. Scott (Sage, 2000)
Social and Economic Networks, Matthew Jackson (Princeton University Press, 2008)
Social Network Analysis, S. Wasserman and K. Faust (Cambridge University Press, 1994)
Statistical analysis of network data, E. Kolaczyk (Springer, 2009)
Matrix Analysis and Applied Linear Algebra, C. D. Meyer (SIAM, Philadelphia, 2000)
Graph Spectra for Complex Networks, P. Van Mieghem (Cambridge Univ Press, 2011)
Introduction to Graph Theory, R. J. Wilson (Addison-Wesley, 1997)
Graph Theory, R. Diestel (Springer, 2000)
Six Degrees: The Science of a Connected Age, D. J. Watts (Norton, New York, 2003)

Events


NetSci®
International School and Conference on Network Science
11 - 15 June 2018 • Paris • France


COMPLEX NETWORKS
COMPLEX NETWORKS 2018
THE 7TH INTERNATIONAL CONFERENCE ON
COMPLEX NETWORKS AND THEIR APPLICATIONS
December 11 -13, 2018, Cambridge, United Kingdom

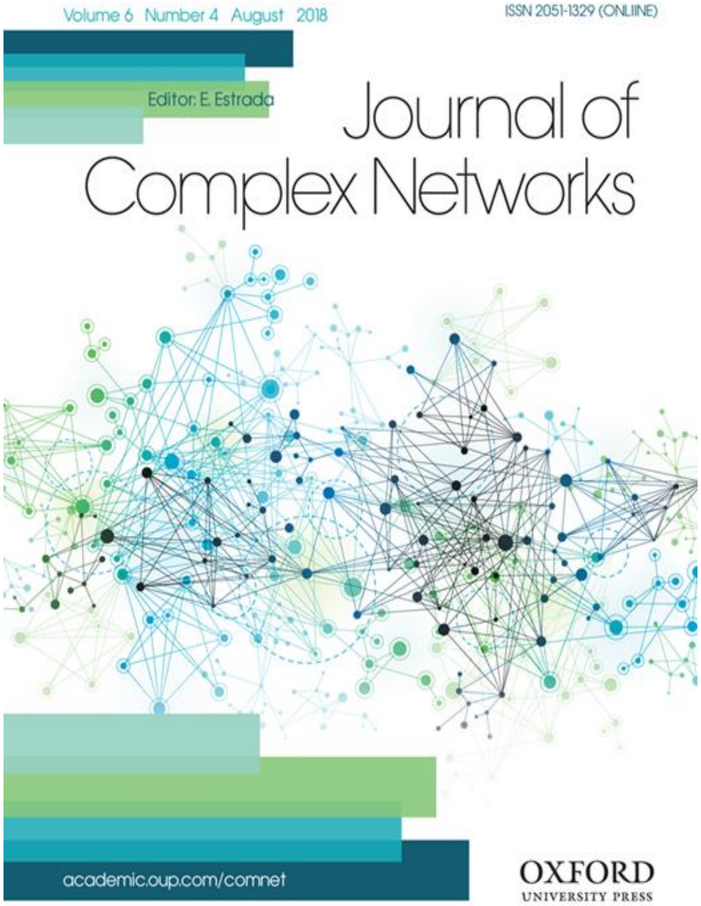

BURLINGTON VERMONT
MAY 27-31, 2019
NetSci®
INTERNATIONAL
CONFERENCE
AND SCHOOL
ON NETWORK
SCIENCE


NetSci X 2019
International School and Conference
on Network Science


Conference on Complex Networks
COMPLENET'19
TARRAGONA, SPAIN
18-21 MARCH 2019

Journals


VOL. 6 • 2018 • NO. 2
NETWORK
SCIENCE
CAMBRIDGE
UNIVERSITY PRESS


Volume 6 Number 4 August 2018
ISSN 2051-1329 (ONLIINE)
Editor: E. Estrada
Journal of
Complex Networks
academic.oup.com/comnet
OXFORD
UNIVERSITY PRESS

What is a network

Disconnected Nodes

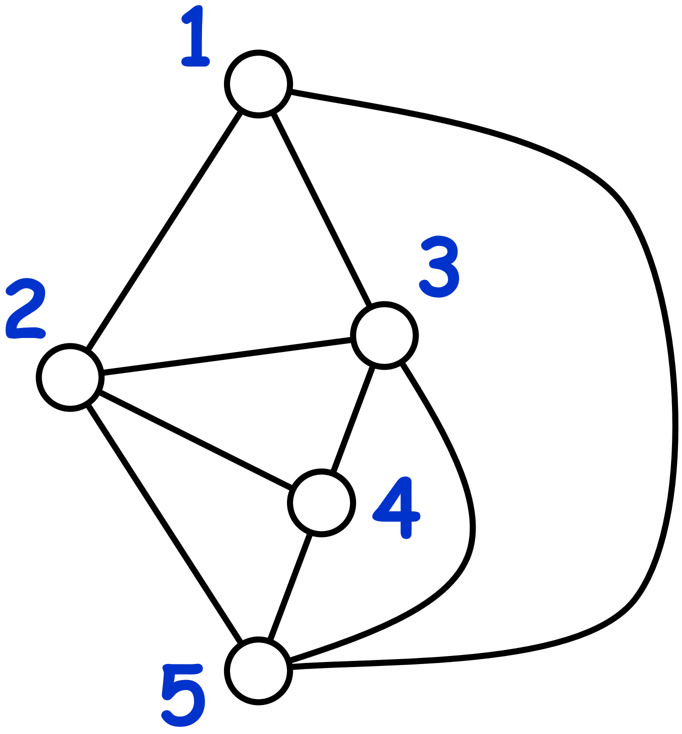
→

Networks

Interactions!

Networks

nodes (vertices) & links (edges)



Nodes = 1, 2, 3, 4, 5
Nodes may have states

Links =
(1,2), (1,3), (1,5),
(2,3), (2,4), (2,5),
(3,4), (3,5), (4,5)

Links may have directions and weights

Representation of a network

Adjacency matrix:

	A	B	C	D	E
1	1	1	1	1	0
2	1	1	1	0	1
3	0	1	1	1	0
4	0	0	1	0	1

undirected graph → symmetric

Adjacency list:

i->j: (i,j)

Networks are everywhere!

Your daily life:

- Food
- Energy
- Security
- Public health
- Social activities

Graph vs Network

Mathematical represenations

Size (Thermodynamics limit)

Dynamics On and Of Networks

Fundementals of network theory

Mathematics of networks

Measures and metrices

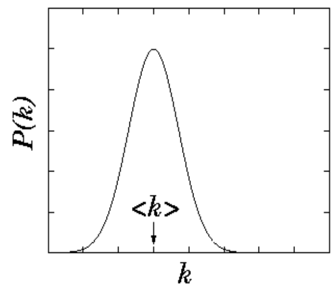
Topological properties

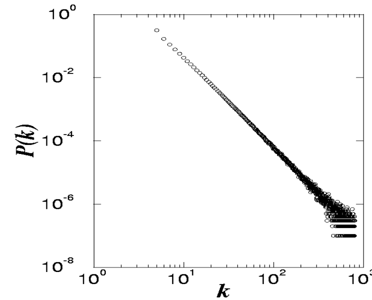
- # nodes
- # links
- <k>: Average degree (# of links per node)
- # connected components
- Connectivity measures
  - Degree distribution
  - Clustering coefficient
  - Motifs
  - Diameter
  - Assortativity
- Shortest paths
- Spectrum

Degree distribution
(first order of interactions)

- P(k) = # of nodes with degree k

a rough profile of how the connectivity is distributed within the network

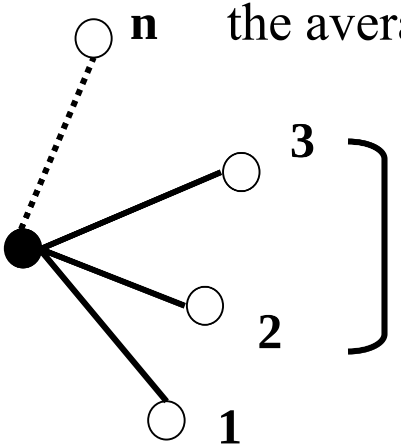




- total number of nodes?

Clustering coefficient (topological correlations)

- of a node:



the average probability for two of one's friends to be friends too

Higher probability to be connected

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

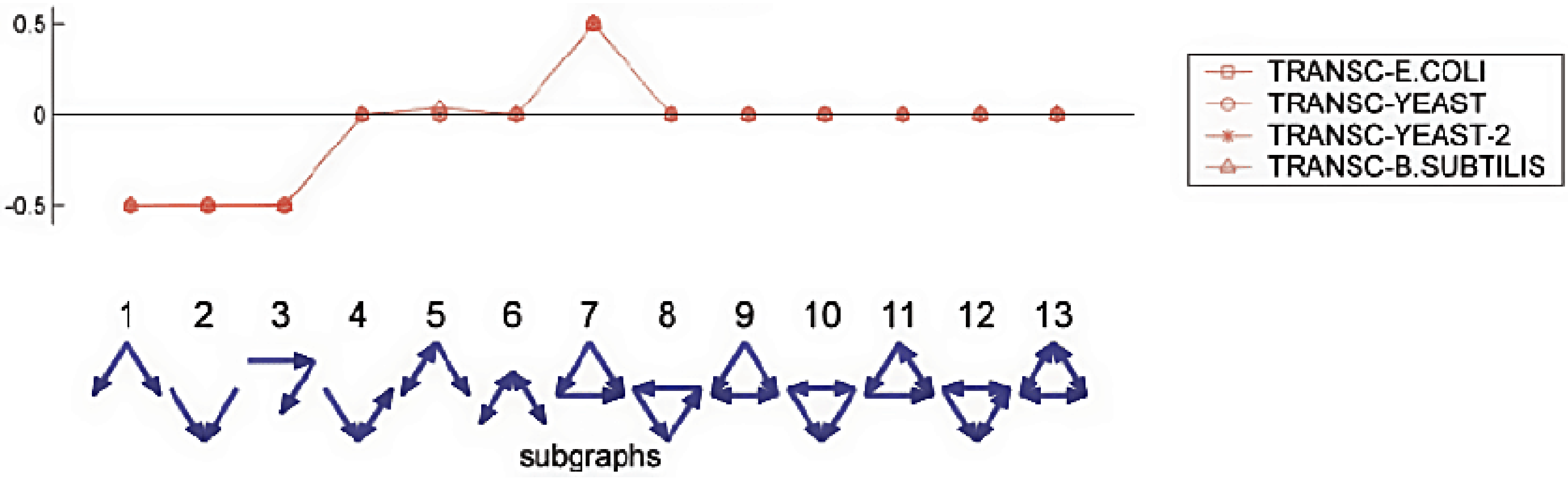
- of the network:

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i$$

Motifs

- Subgraphs of a particular type like: triangles, triples, ...
- To measure the frequency, we compare with how expected it is to see such patterns in a random network
- The significance profile (SP) of the network is a vector of those frequencies.

Motifs



- 13 possible directed connected graphs of 3 vertices, called a *triad* significance profile (TSP).

- 4 networks of different micro-organisms are shown to have very similar TSPs

Cycles
(higher order of interactions)

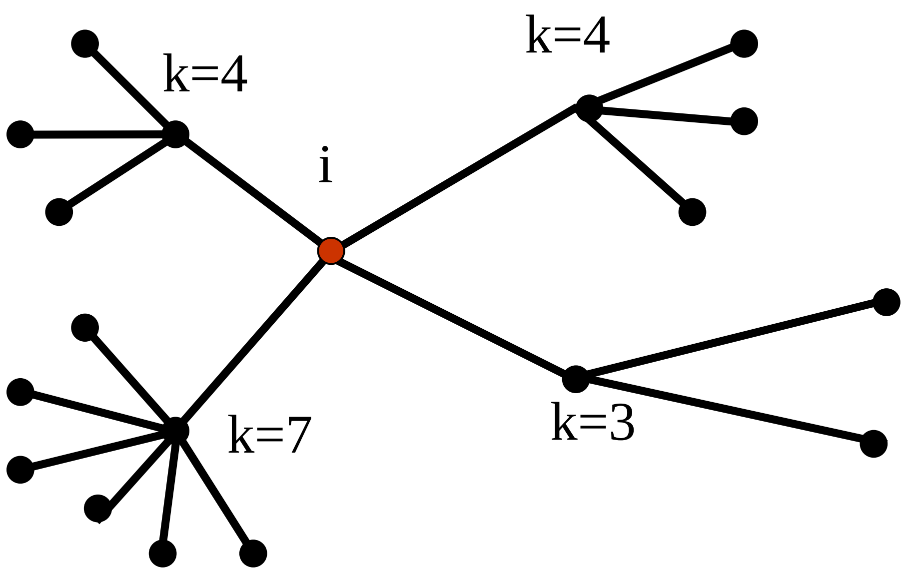
- Inhomogeneous evolution of subgraphs and cycles in complex networks (Vazquez, Oliveira, Barabasi. Phys. Rev E71, 2005).

- Degree-dependent intervertex separation in complex networks (Dorogovtsev, Mendes, Oliveira. Phys Rev. E73 2006)

Assortativity
(topological correlations)

$$k_{nn,i} = \frac{1}{k_i} \Sigma_j \, a_{ij} k_j$$

$$k_{nn}(k) = \frac{\Sigma_i \, \delta(k_i - k) k_{nn,i}}{\Sigma_i \, \delta(k_i - k)}$$



k_i=4

$k_{nn,i}$= (3+4+4+7)/4=4.5

Assortativity

- Assortative behaviour: growing $k_{nn}(k)$
  - **Example: social networks**
  - **Large sites are connected with large sites**

- Disassortative behaviour: decreasing $k_{nn}(k)$
  - **Example: internet**
  - **Large sites connected with small sites, hierarchical structure**

Distance between two nodes

- Number of links that make up the path between two points
- “Geodesic” = shortest path
- 6 degree?

Diameter

- In mathematics: Maximum of shortest path lengths between pairs of nodes
- In recent network theory: Average shortest path lengths
- Characterizes how globally the network connected is:

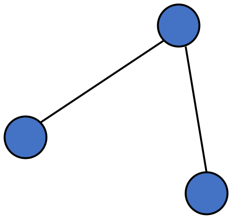
a small diameter = well connected globally

Directedness

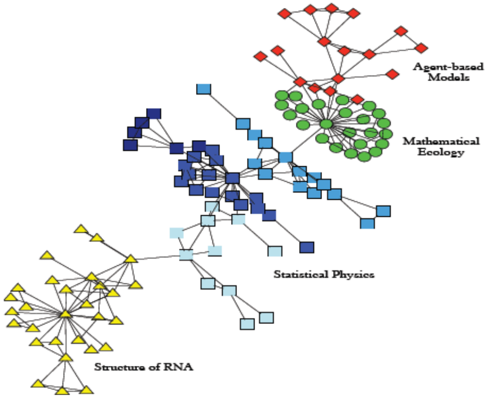
- The flow of resources depends on direction of the degree:
    - In-degree
    - Out-degree

- Careful definition of magnitudes like clustering

Communities



Dynamics on Different Structure Scales:



Microscopic ? Macroscopic

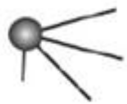
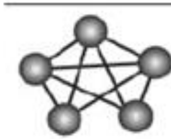
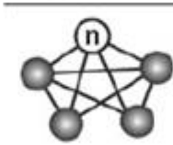
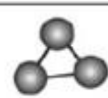
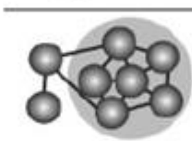
Nodes
Motif
Clusters
…

what is a Meso (midterm) pattern?

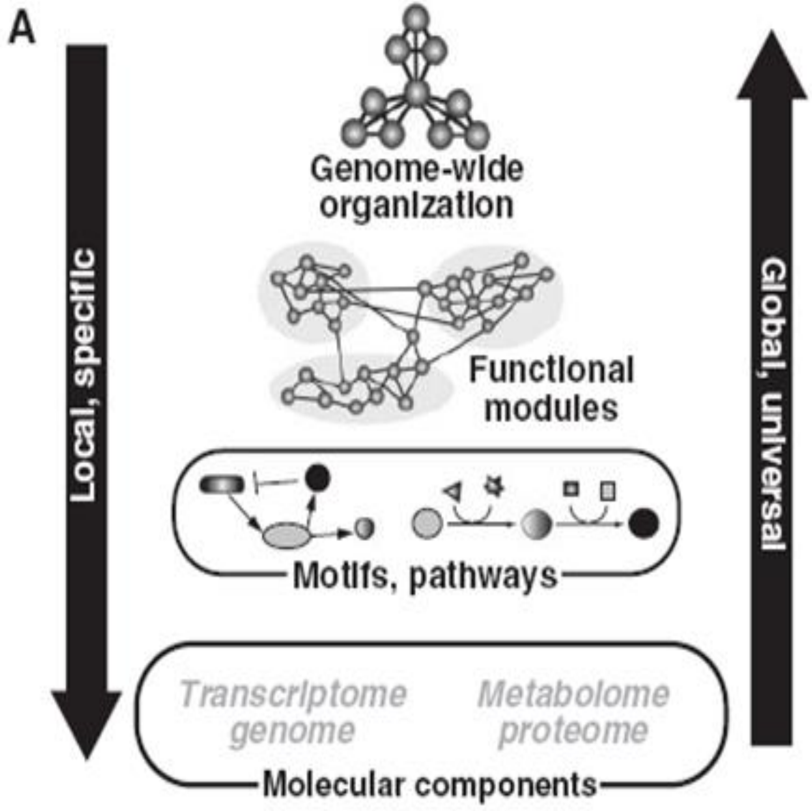
Network

Communities
(how to measure: multi-scale structure?)

Structure Description

	Node
	Link
	Degree: number of links made by a node.
	Network density: a ratio of the observed number of links to all possible links.
	Clustering coefficient: clustering coefficient of node (n) is a ratio of observed number of links between n' s neighbors to number of all posiible links between n' s neighbors.
	Motif: statistically over-represented sub-graphs.
	Module: a group of nodes that linked more densely within the group.

Hierarchical Description: Module ⇔ Function



What is community on complex networks?

- Groups of nodes, having more internal than external connections between them:
  - Share common properties
  - and/or play similar roles within the graph.

- People are still improving the definition and also the methods to detect the true communities in real world.

Community detection

Traditional Methods

Graph partitioning: dividing the vertices in g groups of predefined size

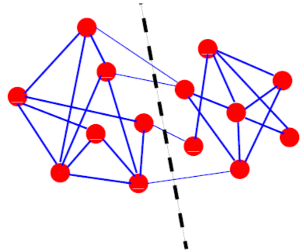


FIG. Graph partitioning. The dashed line shows the solution of the minimum bisection problem for the graph illustrated, i. e. the partition in two groups of equal size with minimal number of edges running between the groups. Reprinted figure with permission from Ref. (Fortunato and Castellano, 2009). ©2009 by Springer.

Hierarchical clustering: definition of a similarity measure between vertices

Partitional clustering: separate the points in k clusters such to maximize or minimize a given cost function based on distances between points.

Spectral Clustering: eigenvectors of matrix Adjacent or Laplace.

Divisive algorithms

The algorithm of Girvan and Newman: according to the values of measures of edge centrality, estimating the importance of edges according to some property or process running on the graph

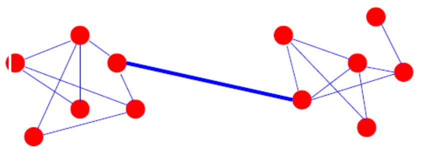


FIG. Edge betweenness is highest for edges connecting communities. In the figure, the edge in the middle has a much higher betweenness than all other edges, because all shortest paths connecting vertices of the two communities run through it. Reprinted figure with permission from Ref. (Fortunato and Castellano, 2009). ©2009 by Springer.

Modularity-based methods

Modularity optimization

Modifications of modularity

Limits of Modularity

Spectral algorithms: Use the eigenvalue and eigenvectors

Spectrum

- Set of eigenvalues of the adjacency matrix
- Spectral density (density of eigenvalues)

$$\rho(\lambda) = \frac{1}{N} \sum_{j=1}^{N} \delta\left(\lambda - \lambda_j\right)$$

Relation with graph topology

- k-th moment

$$M_k = \frac{1}{N}\sum_{j=1}^{N}\left(\lambda_j\right)^k = \frac{1}{N}Tr\left(A\right)^k = \frac{1}{N}\sum_{i_1,..,i_k} A_{i_1,i_2}A_{i_2,i_3}...A_{i_k,i_1}$$

- N*M = number of loops of the graph that return to their starting node after k steps
- k=3 related to clustering

- A symmetric and real => eigenvalues are real and the largest is not degenerate
- Largest eigenvalue: shows the density of links
- Second largest: related to the conductance of the graph as a set of resistances

Network models

Erdös-Rényi

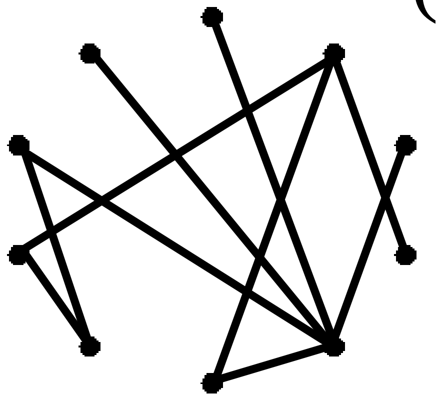
Small World (WS)

Scale Free (AB)

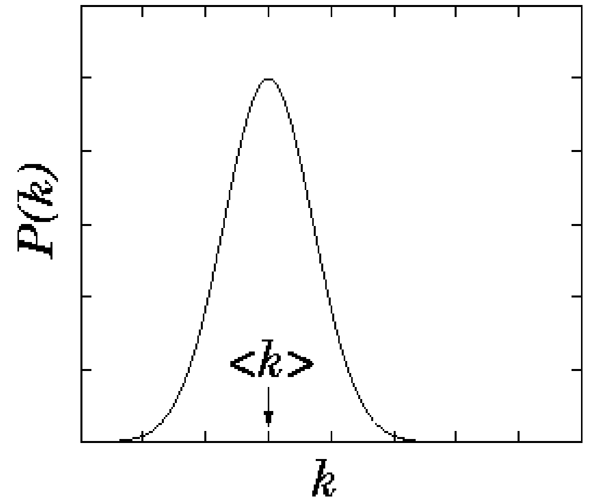
Erdös-Rényi model (1960)

- N nodes
- Links with probability p are connected
- Static random graphs

(p=O(1/N))



Poisson distribution



Asymptotic behavior

Lattice

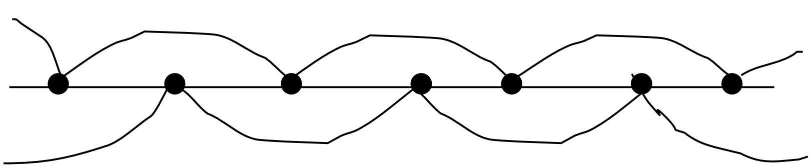
$$L(N) = N^{1/d}$$

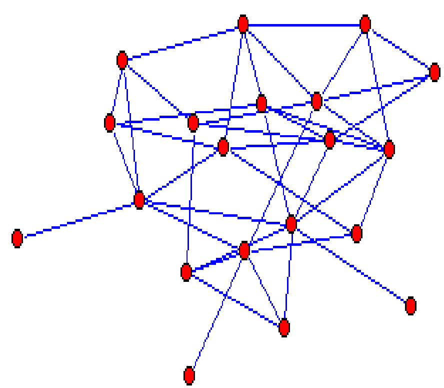
$$C(N) \approx const.$$

Random graph

$$L(N) = \log N$$

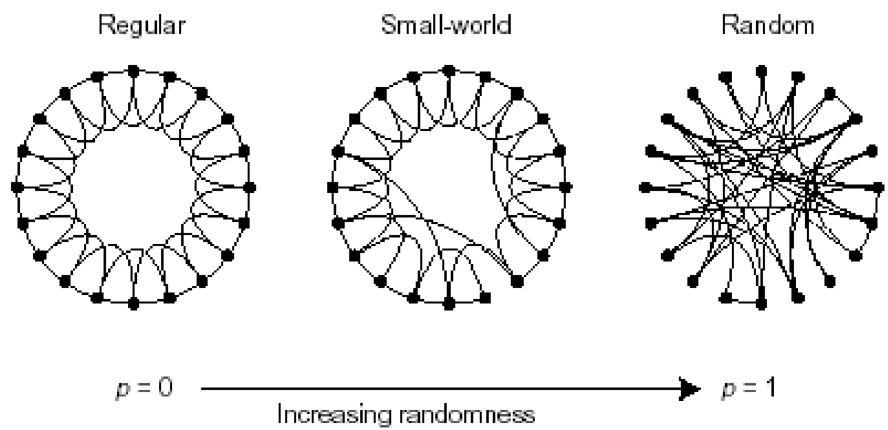
$$C(N) \approx N^{-1}$$



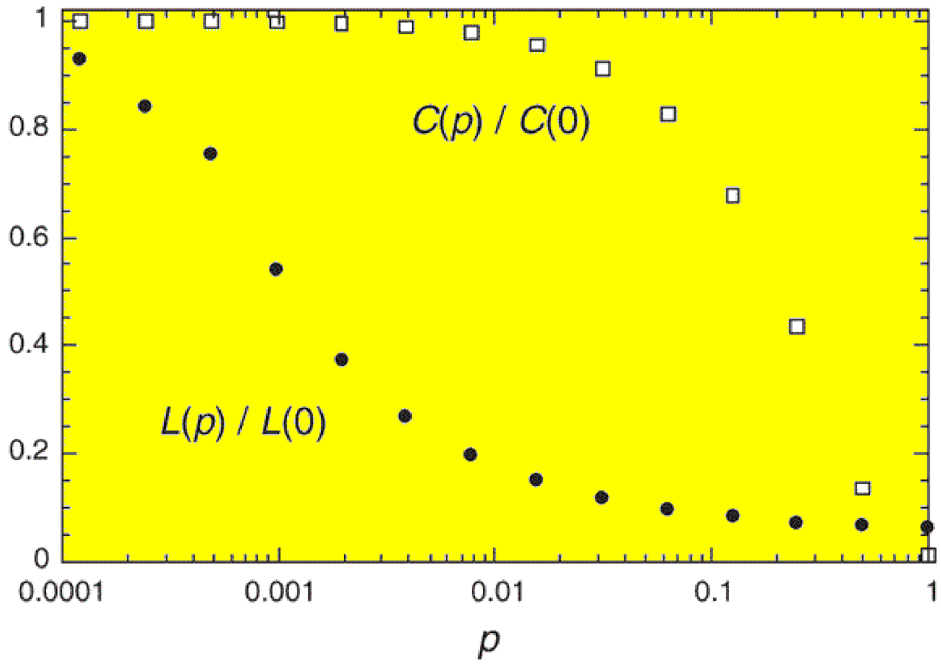


Watts-Strogatz model (1998)

- Nodes are initially arranged in a circle
- Each node is connected to k nearest neighbors
- Then links are rewired randomly with probability p

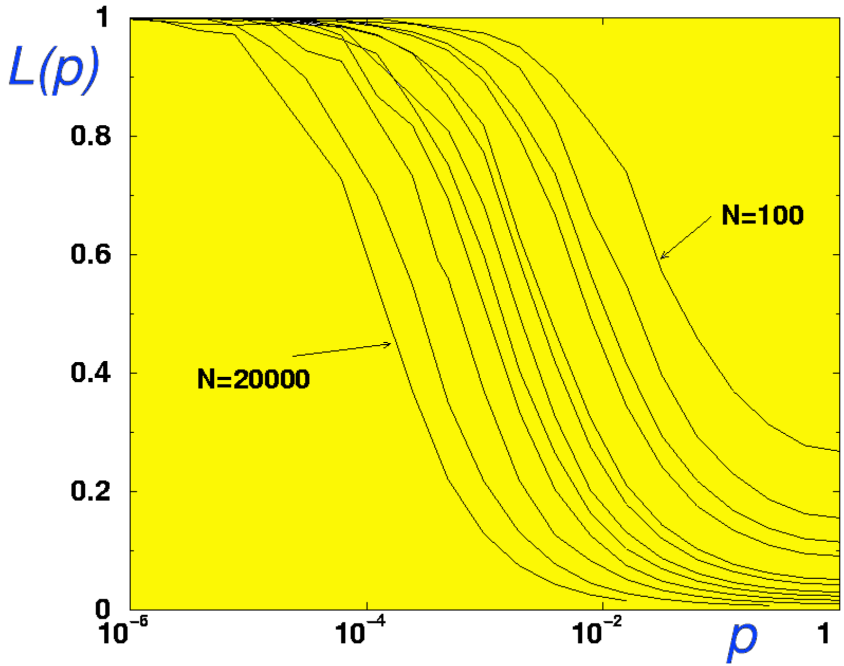


Small-world networks



Large clustering coeff.
Short typical path

Size-dependence



p >> 1/N => Small-world

Amaral & Barthélemy *Phys Rev Lett* **83**, 3180 (1999)

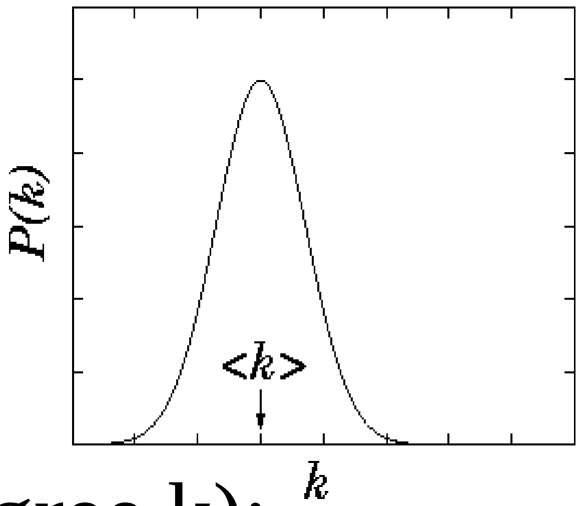
Newman & Watts, *Phys Lett A* **263**, 341 (1999)

Barrat & Weigt, *Eur Phys J B* **13**, 547 (2000)

Is that all we need?

NO, because...

ER & WS Networks are
homogeneous graphs
(small fluctuations of the degree k):



While

Emperical Networks

Technological

Social

Networks of information

Biological

Emperical Networks

Supercritcal = connected $<k> > 1$

Small word

Sparse

Not random

No homogounous degree dist.

Clustering

Two important observations:

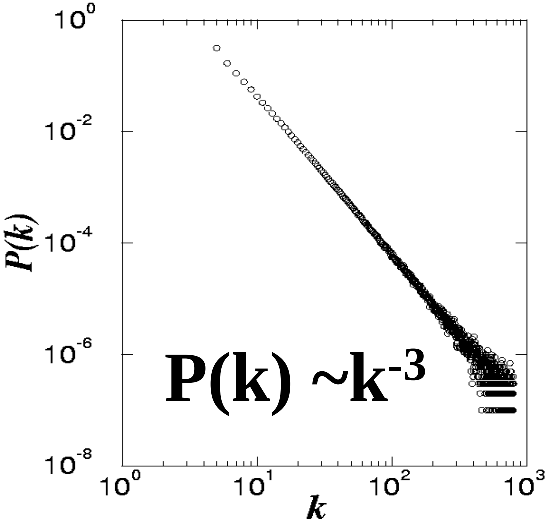
- The number of nodes (N) is NOT fixed.

- The attachment is NOT uniform

Barabasi-Albert model (1999)

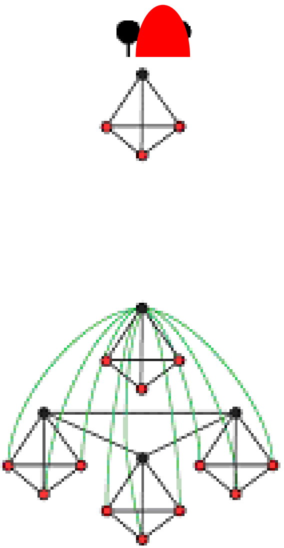
- Growth with time
- Preferential attachments (rich get richer)

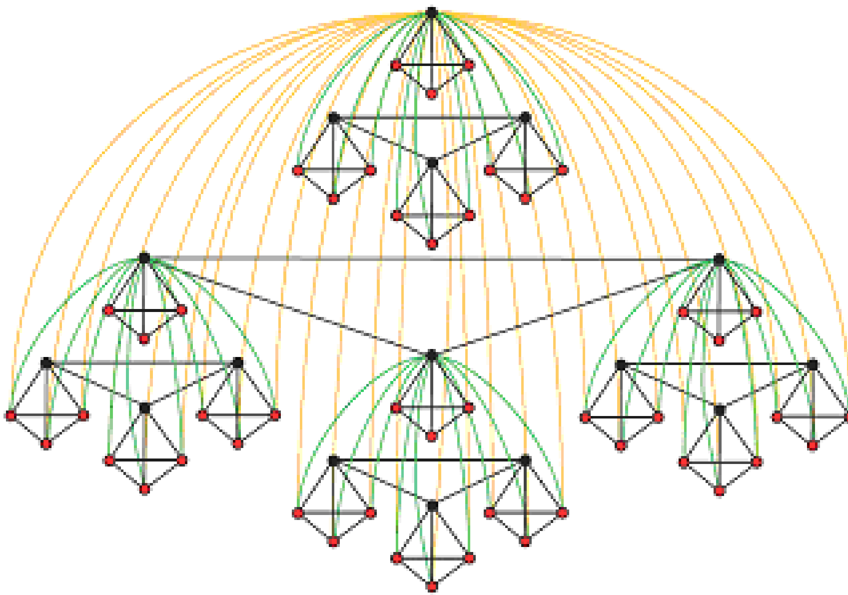
$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$



Scale-free

Scale-free networks





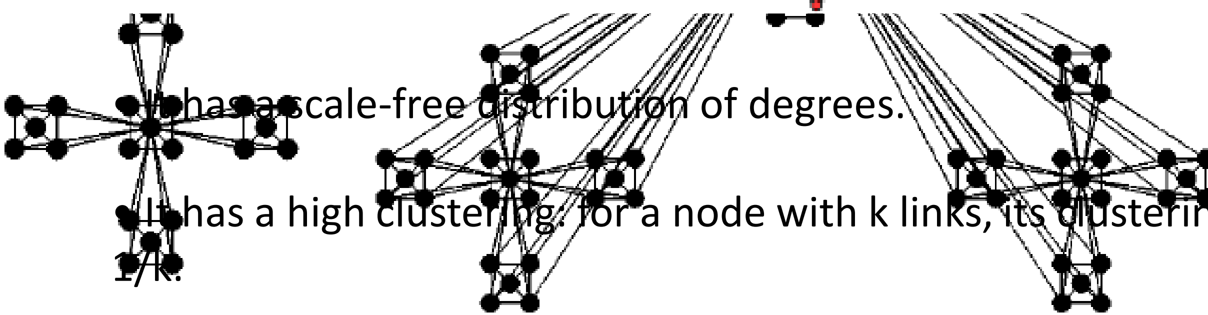
Start with a complete graph of N vertices (a dense group).

This construction was generalized in the family of graphs Hn,k.

Make N – 1 copies.

Start with a group of n nodes and iterate k steps (*i.e.* build k levels).

Link the root (red) to all vertices but the copies of the root.



- It has a scale-free distribution of degrees.
- It has a high clustering: for a node with k links, its clustering is C(k) ~ 1/k.

Repeat the process.

(b) n=1, N=25

(c) n=2, N=125

- Its diameter is 2k – 1 (this network is <u>not</u> small-world).

More SF models

- **Generalized BA model**

  Non-linear preferential attachment : $\Pi(k) \sim k^{\alpha}$

  (Redner et al. 2000)

  Initial attractiveness : $\Pi(k) \sim A+k^{\alpha}$

  (Mendes et al. 2000)

  Rewiring

  (Albert et al. 2000)

- **Highly clustered**

  (Dorogovtsev et al. 2001)

  (Eguiluz & Klemm 2002)

- **Fitness Model**

  (Bianconi et al. 2001)

$$\Pi(k_i) \cong \frac{\eta_i\, k_i}{\sum_j \eta_j\, k_j}$$

- **Multiplicative noise**

  (Huberman & Adamic 1999)

(....)

Tools for characterizing the various models

Connectivity distribution P(k)

=>Homogeneous vs. Heterogenous (SF)

Clustering

Assortativity

...

=>Compare with real-world networks

Tools
Visualization, Measures, ...

NetworkX: A Python package for studying the structure, dynamics & functions of networks
Gephi: An open visualization and exploration software
iGraph: A software package for creating and manipulating undirected and directed graphs
Pajek: A simple network visualization tool allowing to interactively manipulate the network
NetLogo: A multi-agent programmable modeling environment
Graphviz: A simple network visualization tool available for a variety of platforms
GUESS: An exploratory data analysis and visualization tool
JUNG: A Java Universal Network/Graph Framework
SNAP: Stanford Network Analysis Platform
UCINET: A social network visualization and analysis tool
IVC: InfoVis Cyberinfrastructure is a collection of data analysis & visualization algorithms
graph-tool: A python module to help with statistical analysis
NetSciDraw: A tool to draw your data

Data Resources

Sociopatterns
Penelope
Konect
InfoVis data
UCINET data
Netviki data
Harvard dataverse

Newman data
Alon data
Arenas data
Pozo data
Cx-Nets data
Reality Commons

Consequences of the topological heterogeneity

- Robustness and vulnerability
- Propagation of epidemics

Processes on networks

Percolation

Random walks

Epidemics

Synchronization

Evolutionary Game Theory

...

More

Networks of networks

Multilayer networks

Multiplex

Temporal netwoks

Temporal-Spatial networks

...

Approximations

Homogeneous Mean Field Approximations

Heterogeneous Mean Field Appr.

Pair Approximations

...