# Network models

Properties common to many large-scale networks, independently of their origin and function:

1. The degree and betweenness distribution are decreasing functions, usually power-laws. scale - free
2. The distances scale logarithmically with the network size

$$l \approx \frac{\log N}{\log\langle k\rangle}$$

small world

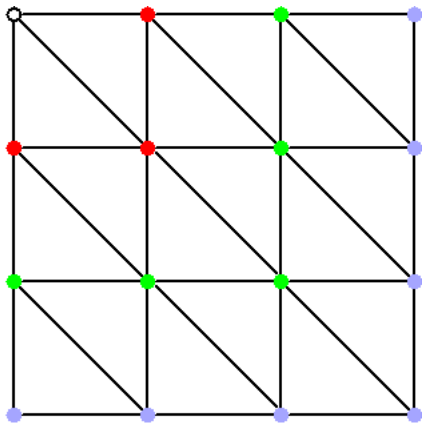
3. The clustering coefficient does not seem to depend on the network size, and is larger than the clustering coefficient of comparable random graphs

There are two model families proposed to explain these properties: Small world network models and scale-free network models.

# Benchmark 1: regular lattices

One-dimensional lattice: $\boldsymbol{l \approx N, \;\; k = const, C = const}$

Two-dimensional lattice:



$$\boldsymbol{l \approx L \approx N^{1/2}}$$

$$k = 6 = \text{const. for inside nodes}$$

$$C = \frac{6}{15} = \text{const. for inside nodes}$$

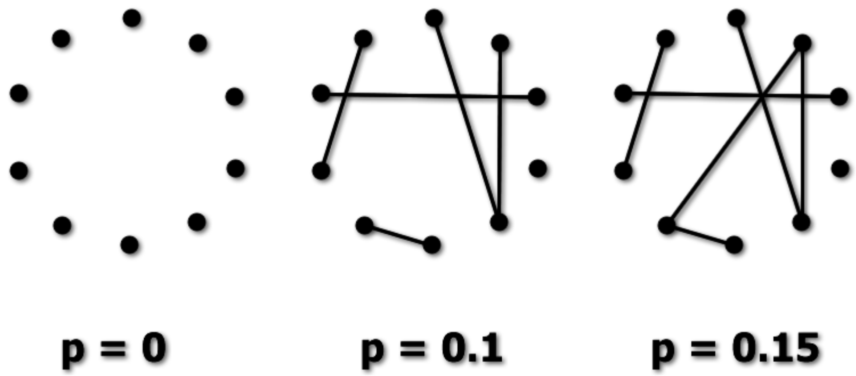
D-dimensional lattice:

The average path-length varies as $\boldsymbol{l \approx N^{1/D}}$

Constant degree (coordination number), constant clustering coefficient.

# Benchmark 2: random graph theory

Erdös-Rényi algorithm - Publ. Math. Debrecen 6, 290 (1959)




- fixed node number $N$
- connecting pairs of nodes with probability $p$

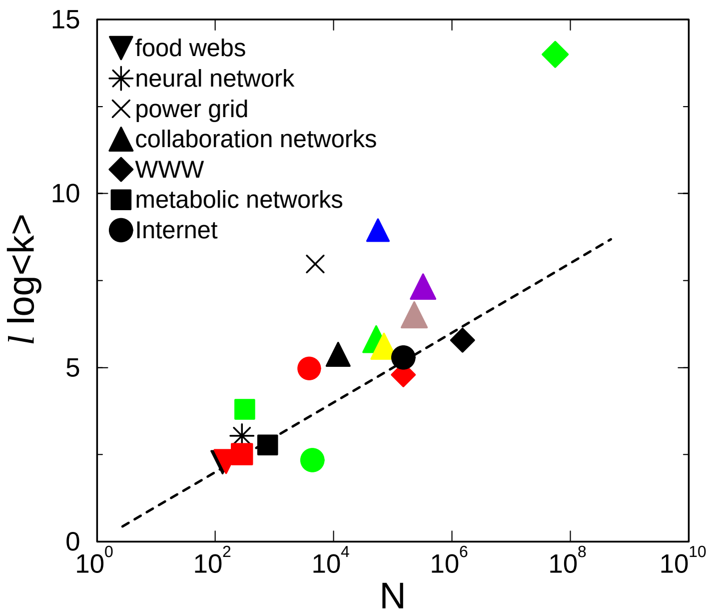
Expected degree distribution: $P_{rand}(k) \cong C_{N-1}^{k} p^{k} (1-p)^{N-1-k}$

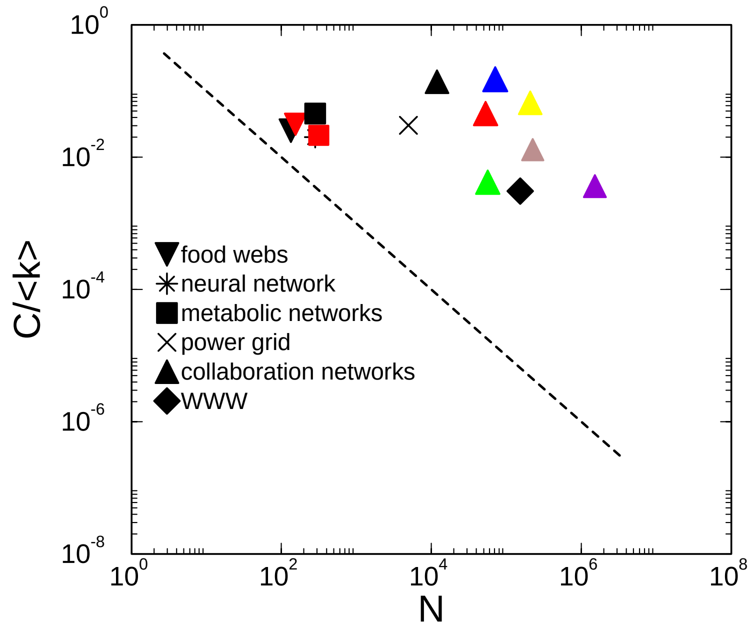
Expected path length: $l_{rand} \approx \frac{\log N}{\log\langle k\rangle}$

Expected clustering coefficient: $C_{rand} = p = \frac{\langle k\rangle}{N}$

# Path length and order in real networks

$$l_{rand} = \frac{\log N}{\log\langle k \rangle} \qquad C_{rand} = \frac{\langle k \rangle}{N}$$



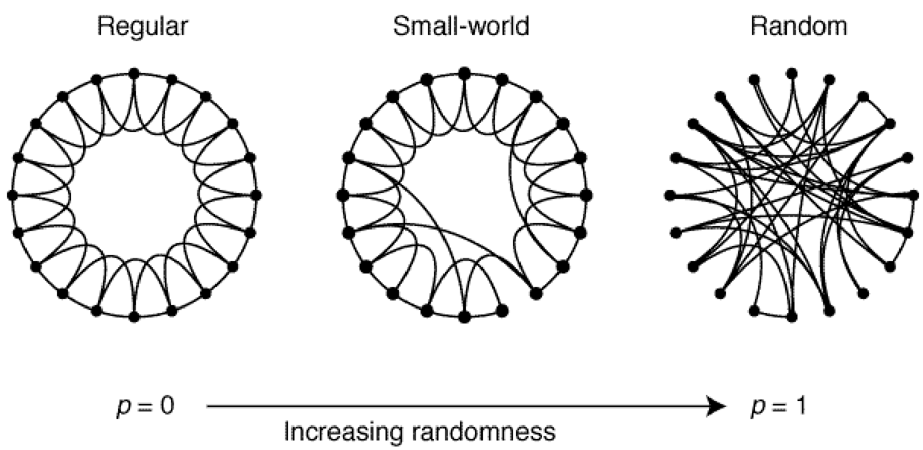




Real networks have short distances like random graphs yet show signs of local order.

# Small-world networks

Real networks resemble both regular lattices and random graphs – perhaps they are in between.

Watts-Strogatz model - D. Watts, S. Strogatz, Nature 393, 440 (1998)

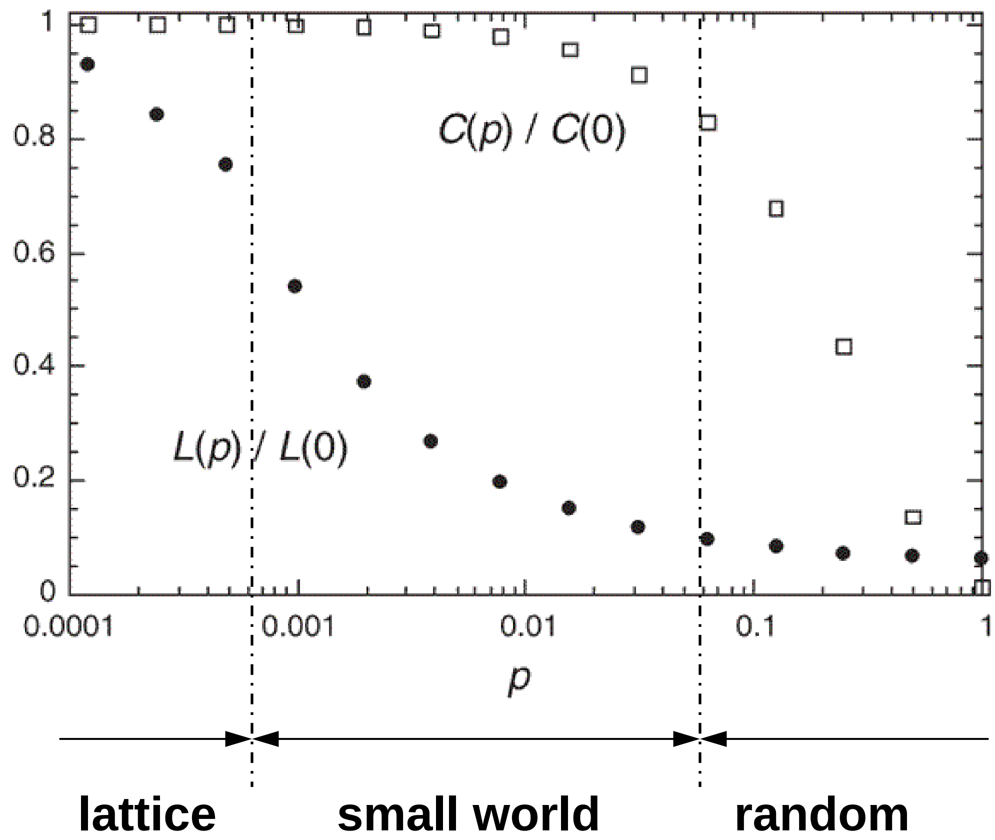



- lattice with $K$ neighbors
- rewire edges with probability $p$

$$l = \frac{N}{2K}, \; C = \frac{3(K-2)}{4(K-1)} \quad \Longrightarrow \quad l \approx \frac{\log N}{\log K}, \; C \approx \frac{K}{N}$$

Is there a regime with small *l* and large C?

# Transition from a lattice to a small world




There is a broad interval of p over which $C(p) \cong C(0)$ but $l(p) \cong l(1)$

# The onset of the small-world behavior depends on the system size

$$l(N,p) \approx \frac{N^{1/d}}{K} f(pKN)$$

*d* is the dimension of the lattice

$$f(u) = \begin{cases} const \text{ if } u << 1 \\ \ln u / u \text{ if } u >> 1 \end{cases}$$

lattice - like

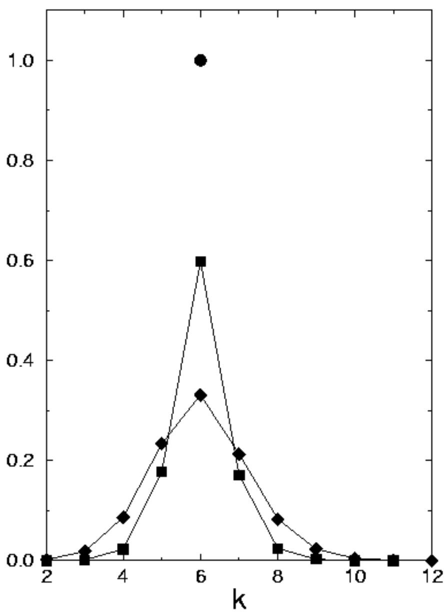
random graph - like

The transition point depends on the rewiring probability, the size of the network and the average degree.

$$C(p) = C(0)(1-p)^3$$

These results cannot be directly compared to most real networks because the rewiring probability *p* is not known.

# Degree distribution of a small-world network




Rewiring does not change the average degree, but modifies the degree distribution.

$$\langle k \rangle = K$$

*P(k)* depends on the rewiring parameter *p,* but is always centered around <*k*>.

Degree distribution similar to that of a random graph, with exponentially small probability for very highly connected nodes.

Ex. 1

A variant of the Watts-Strogatz model adds random edges to a regular lattice. Start with a 1D lattice where every node has degree K . For each existing edge of a node, add an edge with a probability *p*. The endpoint of the edge is selected randomly from all other nodes. How many edges do you expect the graph will have after edge addition?

Ex. 2

How do you expect the degree distribution will look like after edge addition? Will it be symmetrical or not?

$$k_i = K + m_i + n_i$$

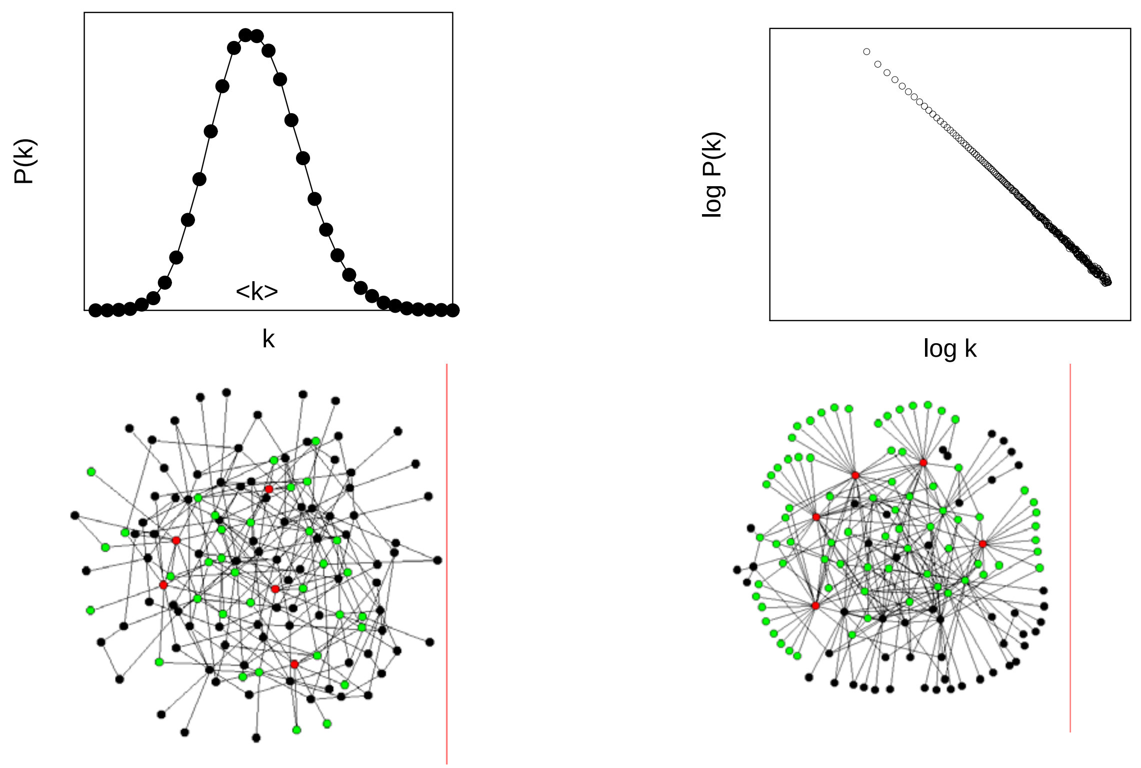
$m_i$: starting point, $n_i$: endpoint

$$P(m) \cong C^{m}_{K/2} p^{m} (1-p)^{K/2-m}$$

$$P(n) \cong C^{n}_{pNK/2} \left(\frac{1}{N}\right)^{n} \left(1-\frac{1}{N}\right)^{pNK/2-n}$$

Minimum: K, peak: K+pK

# The scale-free degree distribution indicates a heterogeneous topology




New models are needed to reproduce this feature.

# We need to uncover the mechanisms responsible for the scale-free *P(k)*

- random graphs
- small-world networks
- scale-free random graphs

**Static** (#of nodes fixed)

Real networks continuously change

- random graphs
- small-world networks

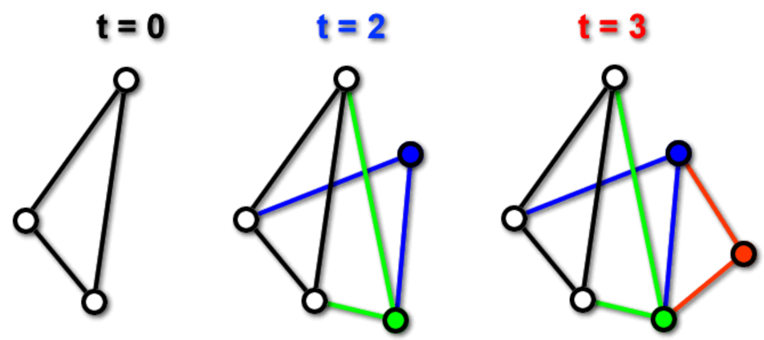
**Homogeneous**

Scale-free degree distribution - the nodes are not equivalent

We need to model the evolution of networks, not just their topology.

# A simple model of network assembly and evolution (BA model)

Start with a small seed of $m_0$ nodes and $m_0(m_0-1)/2$ edges.



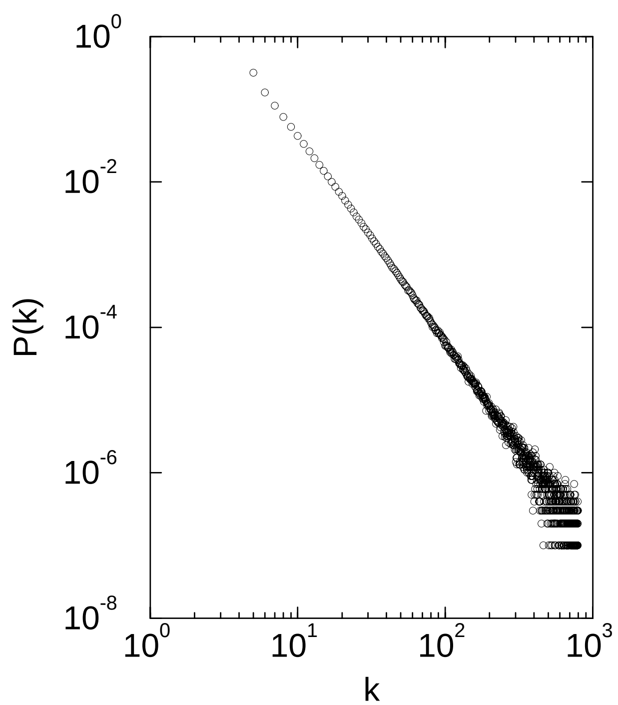

- **growth**: a node and *m* edges added at every step
- **preferential attachment**: $\Pi(k_i) = \dfrac{k_i}{\Sigma_j k_j}$

Price, J. Amer. Soc. Inform. Sci. 27, 292 (1976)
Barabási and Albert, Science 286, 509 (2000)

# General properties of the network

- nr. of nodes: $N = m_0 + t$
- nr. of edges: $E = \frac{m_0(m_0 - 1)}{2} + m\,t$
- average degree: $\langle k \rangle = \frac{2E}{N} \rightarrow 2m$
- degree distribution:

$$P(k) \xrightarrow[t \to \infty]{} Ak^{-3}$$




Although the network grows, the degree distribution becomes stationary.

# Analytical determination of P(k)

The degree of "old" nodes increases by acquiring new edges. The probability of an old node with degree $k_i$ receiving a new edge is

$$m\Pi(k_i) = m\frac{k_i}{\sum_j k_j} \approx \frac{k_i}{2t}$$

Degree increase:

$$\Delta k_i = \begin{cases} 1 & \text{with prob. } \dfrac{k_i}{2t} \\ 0 & \text{otherwise} \end{cases}$$

Choices:

follow the increase in the number of nodes with degree $k_i$ (rate equation approach)

follow the increase in time in $k_i$ (continuum theory)

# Rate equation approach

Change in average number of nodes with degree *k*

k-1 → k          k → k+1

$$\frac{dN_k}{dt} = m\frac{(k-1)N_{k-1}(t) - kN_k(t)}{\sum_k kN_k(t)} + \delta_{k,m}$$

← first node

number of edges of new node

normalization

$$P(k) = N_k(t)/N = \lim_{t\to\infty} N_k(t)/t$$

Plug in: $$P(k) = m\frac{(k-1)P(k-1)\cancel{t} - kP(k)\cancel{t}}{\sum_k kP(k)\cancel{t}} + \delta_{k,m}$$

<k>

# Degree distribution

The rate equation leads to a recursive relationship between P(k) and P(k-1)

$$P(k)=\frac{(k-1)P(k-1)-kP(k)}{2}+\delta_{k,m}$$

$$P(k)=\begin{cases}\dfrac{k-1}{k+2}P(k-1) & for\ k>m\\[2ex] \dfrac{2}{m+2} & for\ k=m\end{cases}$$

Leads to $$P(k)=\frac{2m(m+1)}{k(k+1)(k+2)}\approx k^{-3}$$

Stationary power law with an exponent $\gamma = 3$

P. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. 85, 4629 (2000)

Ex. 1

Start from a seed of two nodes connected by an edge. At each step, add a new node, and connect it by a single edge to a preexisting node.

Let the probability of selection be directly proportional with the degree of the “old” node. (Is there an easy way to do this?)

Continue growing the graph until you reach 15 nodes. Describe the graph (average degree, degree distribution,
clustering coefficient, connectivity, maximum distance).

Ex. 2

How will the properties of the graph change if at each step a new node and two new edges are added?

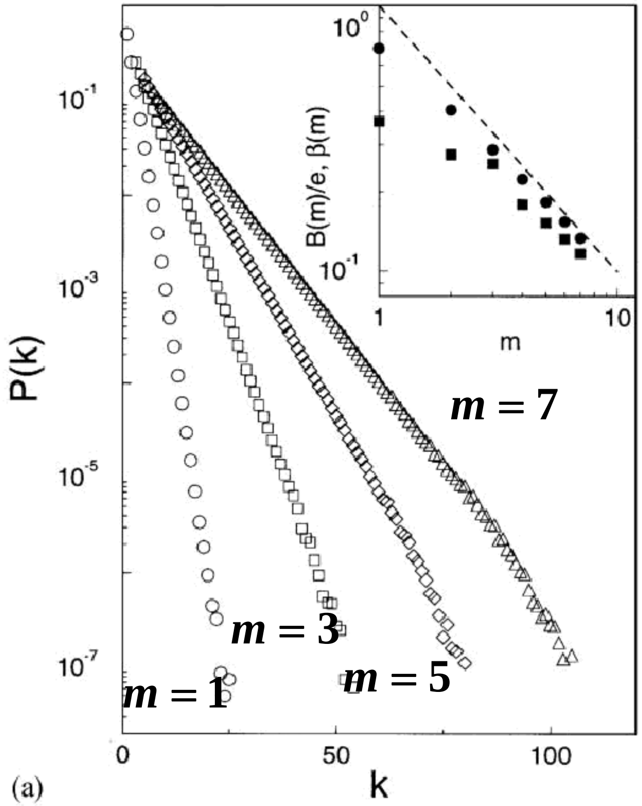
# Model A

| growth | ~~preferential attachment~~ |
|---|---|

$\Pi(k_i)$ : uniform

$$\frac{dk_i}{dt} = A\Pi(k_i) = \frac{m}{m_0 + t - 1}$$

$$P(k) = \frac{e}{m}\exp\left(-\frac{k}{m}\right)$$




Characteristic degree scale: m

# Model B

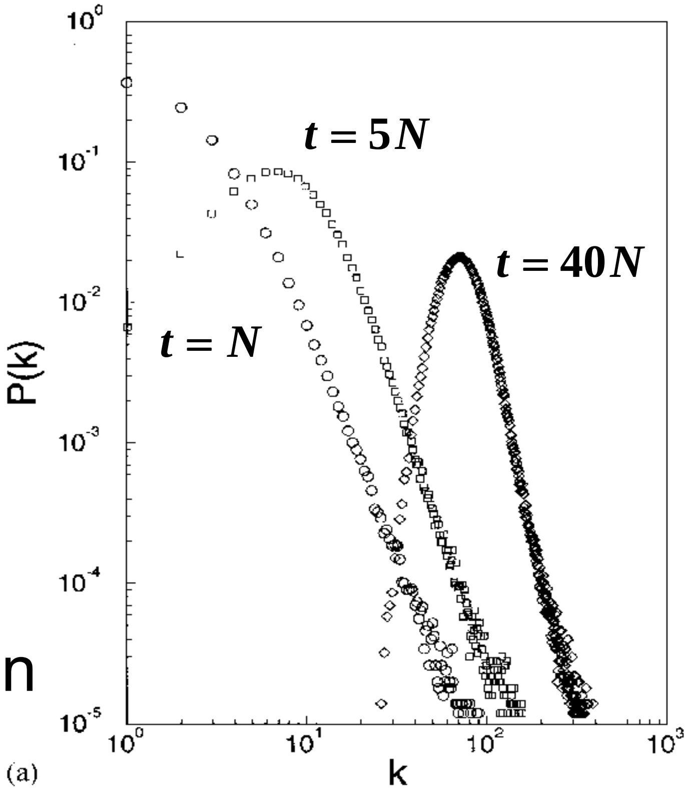
~~growth~~ preferential attachment

Fixed N, edges connect a randomly selected node with a preferentially selected node

$$\frac{dk_i}{dt} = A\Pi(k_i) + \frac{1}{N} = \frac{N}{N-1}\frac{k_i}{2t} + \frac{1}{N}$$

$$\langle k \rangle = \frac{2mt}{N}$$

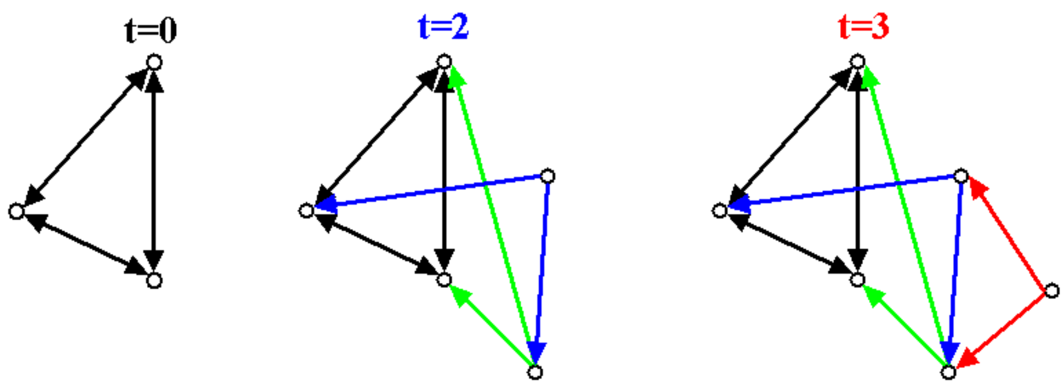
P(*k*) : power law (initially) ⇒ Gaussian



(a)

# BA algorithm with directed edges

New edges are directed from the new to the old nodes




$$k_i^{out} = m \quad for\ i > m_0$$

$k_i^{in}$ varies

$$\langle k^{in} \rangle = m$$

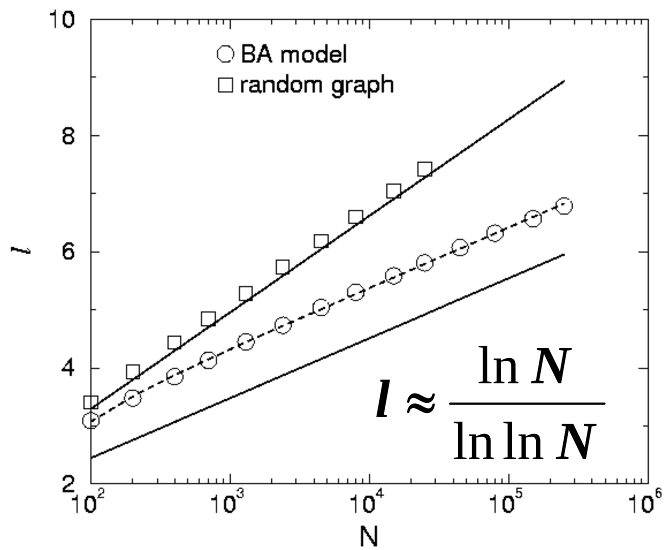
$$\frac{dk_i^{in}}{dt} = m\Pi(k_i^{in}) = m\frac{k_i^{in}}{\sum_j k_j^{in}} = \frac{k_i^{in}}{t}$$

$$P_{in}(k) \sim k^{-2}$$

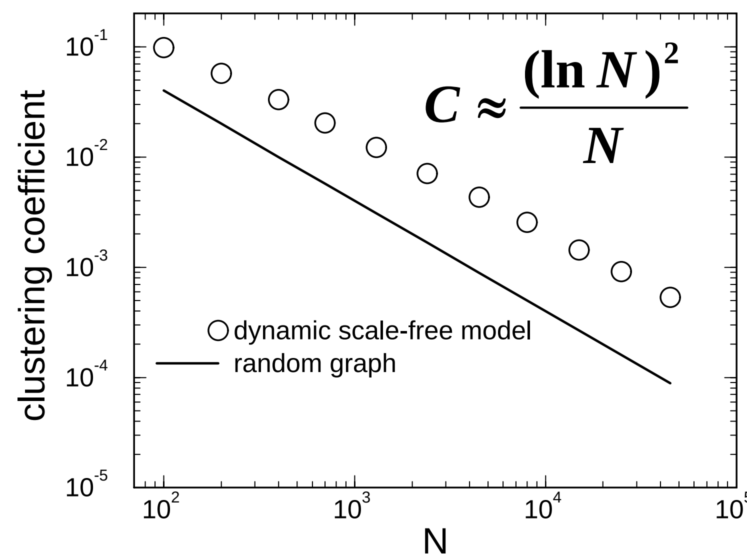
The degree exponent of the directed scale-free network is *2* !

# How do the other network measures compare with real networks?

Average path length




Clustering coefficient




Average distances smaller in the BA model than in equivalent random graphs. but not as small as in scale-free random graphs.

Cohen et al, in Handbooks of Graphs and Networks (2003)

Clustering coefficient decreases with network size.

B. Bollobás and O. Riordan, in Handbooks of Graphs and Networks (2003)

# Evolving network models

The scale-free model is only a minimal model.

Makes the simplest assumptions:

- linear growth $\langle k \rangle = 2m$
- proportional preferential attachment $\Pi(k_i) \propto k_i$

Does not capture

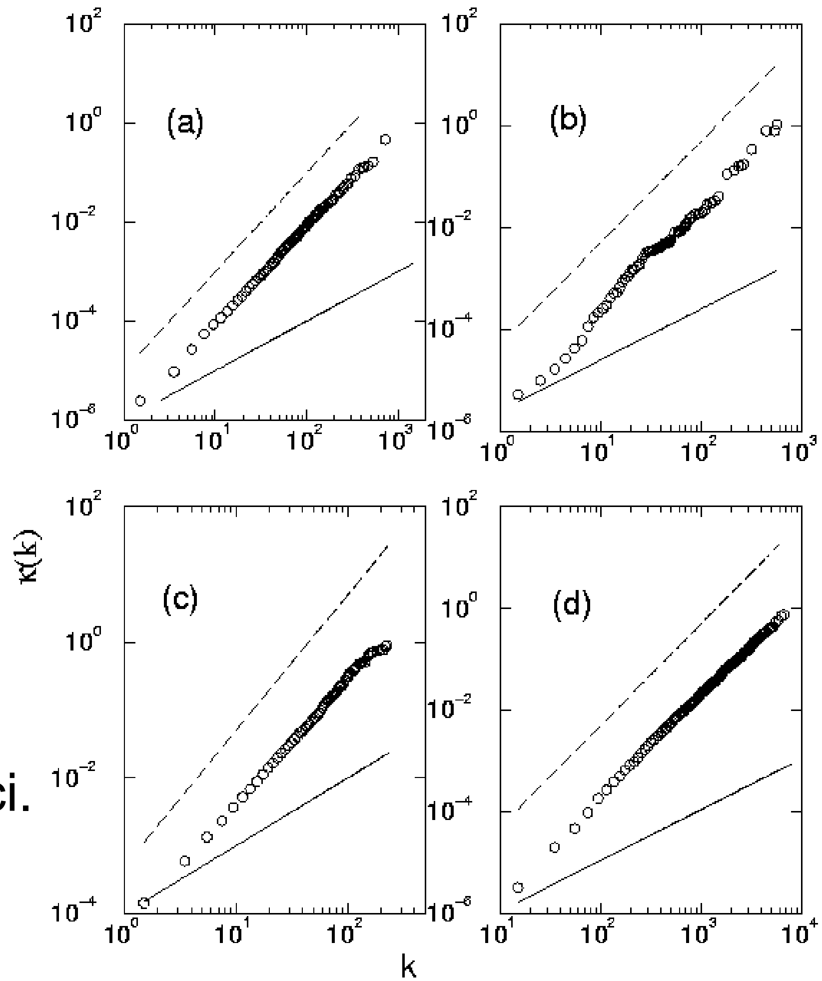
variations in the shape of the degree distribution
variations in the exponent of the power-law region
the size-independent clustering coefficient

Hypothesis: the basic mechanisms need to be augmented by the incorporation of

addition of edges without new nodes
edge rewiring, removal
node removal
constraints or optimization principles

# Preferential attachment in real networks

citation




Internet

$$\kappa(k) = \sum_{K<k} \Pi(K)$$

—— no pref. attach

- - - - - linear pref. attach

$$\Pi(k) \approx A + k^{\alpha},\ \alpha \leq 1$$

neurosci.

actor collab.

# Consequences of nonlinear preferential attachment

$$\Pi(k) \approx A + k^{\alpha},\ \alpha \leq 1$$

$A$ - initial attractiveness

1. Sublinear preferential attachment leads to a stretch-exponential degree distribution.

$$P(k) \approx exp\left(-(k/k_0)^{\beta}\right)$$

P. Krapivsky, S. Redner, F. Leyvraz, Phys. Rev. Lett. 85, 4629 (2000)

2. Initial attractiveness only shifts the degree exponent.

$$\gamma_{in} = 2 + \frac{A}{m}$$ directed network, starting point is 2

Dorogovtsev, Mendes, Samukhin, Phys. Rev. Lett. 85, 4633 (2000)

# Mechanisms for preferential attachment

1. Copying mechanism
   directed network
   select a node and an edge of this node
   attach to the endpoint of this edge

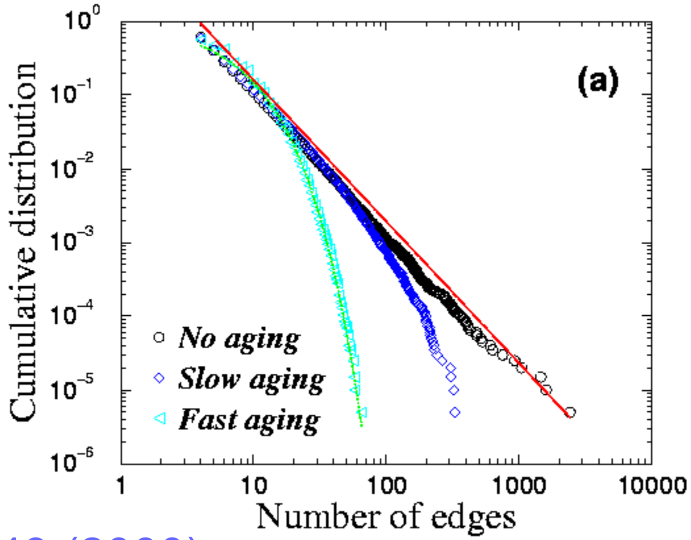
2. Walking on a network
   directed network
   the new node connects to a node, then to every
   first, second, … neighbor of this node

3. Attaching to edges
   select an edge
   attach to both endpoints of this edge

4. Node duplication
   duplicate a node with all its edges
   randomly prune edges of new node

# Growth constraints and aging cause cutoffs

- Finite lifetime to acquire new edges




L. A. N. Amaral et al., PNAS 97, 11149 (2000)

- Gradual aging: $\prod(k_i) \propto k_i (t - t_i)^{-\nu}$

$\gamma$ increases with $\nu$

S. N. Dorogovtsev and J. F. F. Mendes, Phys. Rev. E 62, 1842 (2000)

# Additional processes also change the degree exponent

- *mp* new edges
- *mq* edges rewired

$$P(k) \approx (k + k_0)^{-\gamma} \qquad k_0, \gamma = f(p, q, m)$$

R. Albert, A.-L. Barabási, Phys. Rev. Lett 85, 5234 (2000)

- *c* edges added or removed

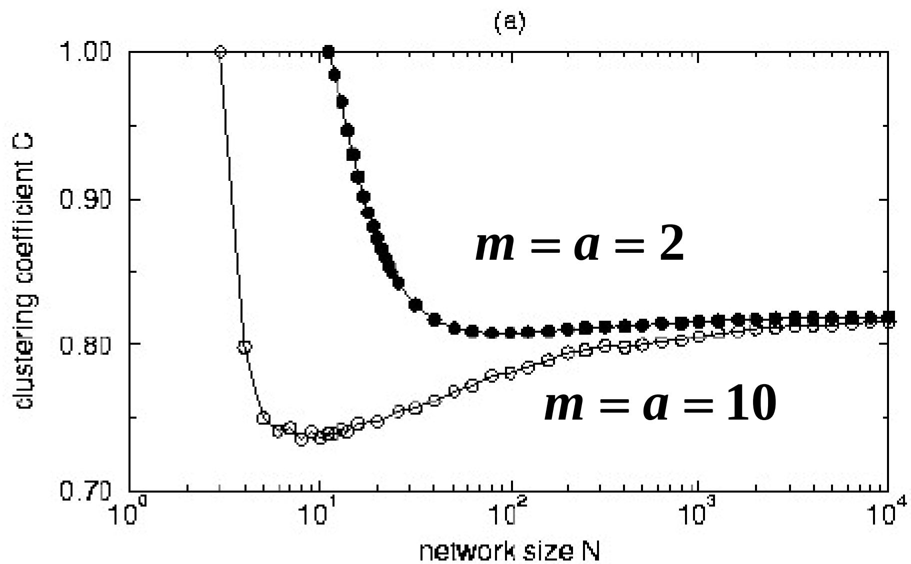
$$\gamma_{in} = 2 + \frac{1}{1 + 2c}$$

S. N. Dorogovtsev, J. F. F. Mendes, Europhys. Lett. 52, 33 (2000)

# A model with high clustering coefficient

- Each node of the network can be either active or inactive.
- There are only *m* active nodes in the network at any instance.

1. Start with *m* active, completely connected nodes
2. At each timestep add a new node (active) that connects to *m* active nodes.
3. Deactivate one active node $P_d(k_i) \propto (a + k_j)^{-1}$



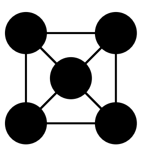

$$\Pi(k) \approx a + k$$

$$P(k) \approx k^{-2-a/m}$$

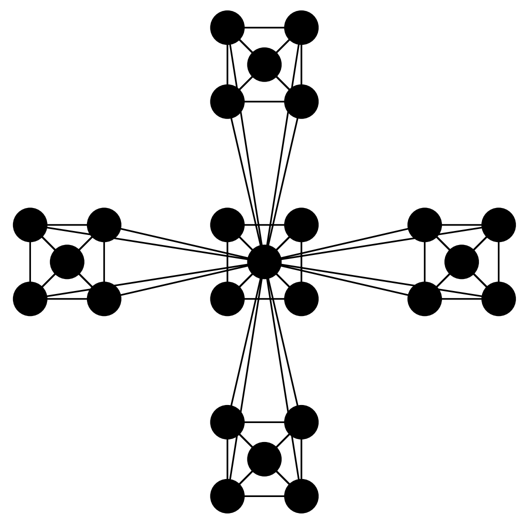
K. Klemm and V. Eguiluz, Phys. Rev. E 65, 036123 (2002)

# A deterministic scale-free model

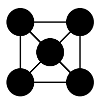
- Start with a completely connected graph with five nodes (one "central" , four peripheral
- Make four copies of the graph,  keep the original in the center. Connect the four peripheral nodes of each copy to the central node of the original.
- Make four copies of the graph, again connect peripheral nodes to the central node.



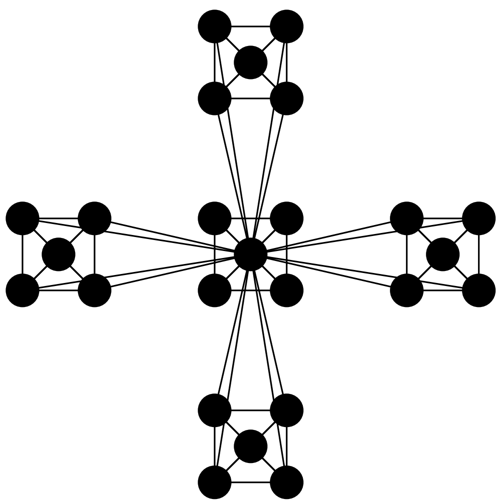
5-clique



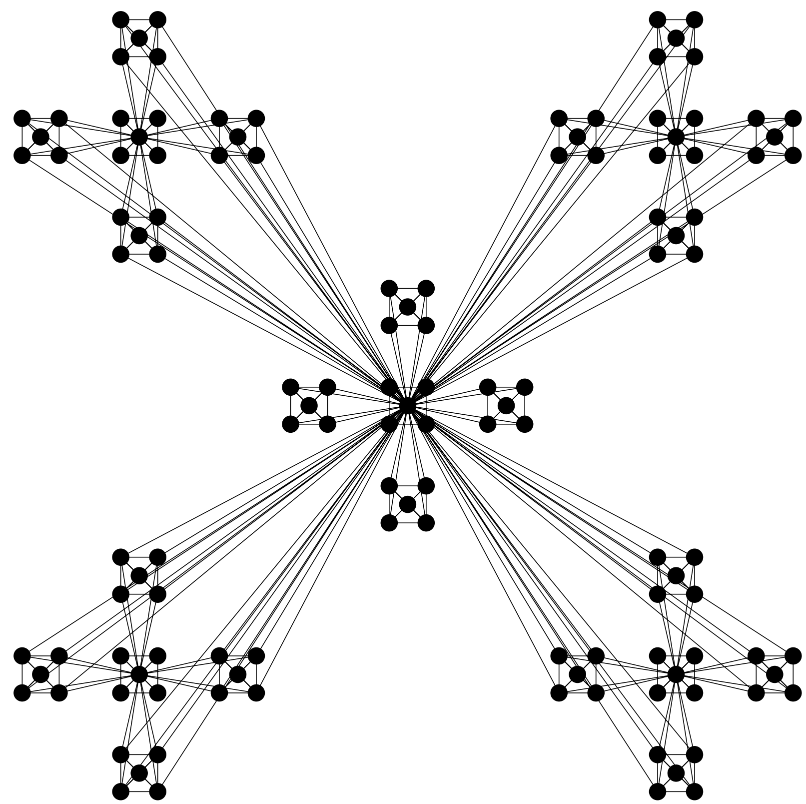
# A deterministic scale-free model



5-clique



connect peripheries
to central node



E. Ravasz, A.-L. Barabasi, Phys Rev E 67, 026112 (2003)

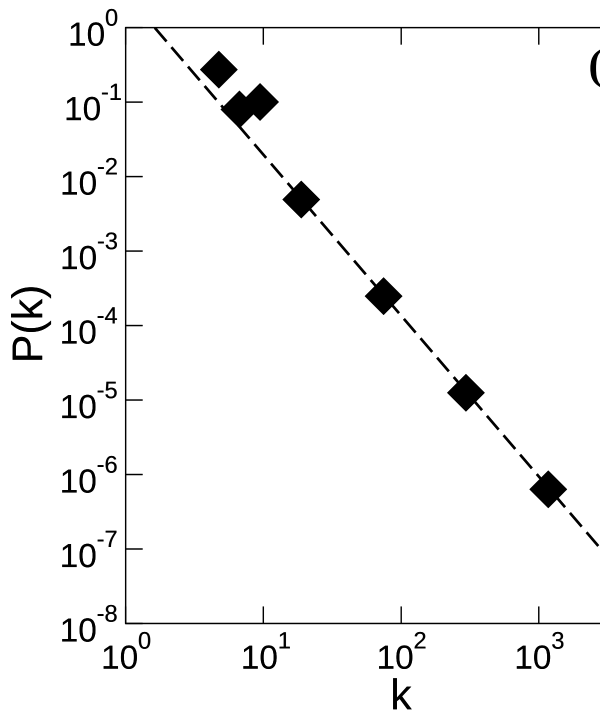
Ex. 1
How does the number of nodes increase as a function of time steps?

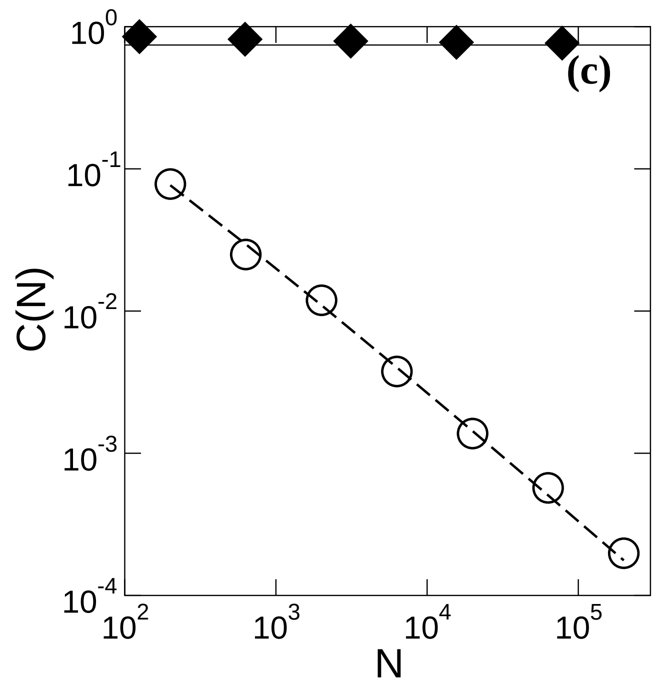
Ex. 2
How does the degree of the central node increase in time?

Ex. 3
How does the number of edges increase as a function of time steps?

Ex. 4
Can you identify the highest degree nodes?

# Properties of the model

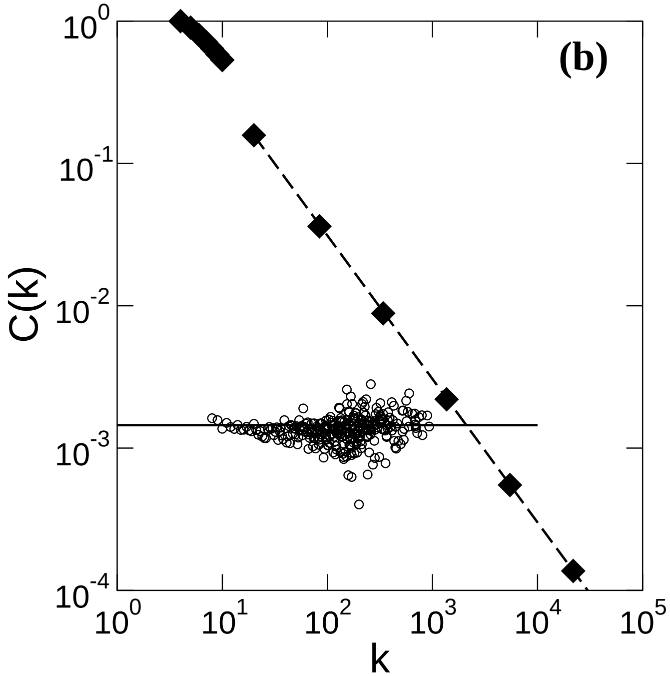


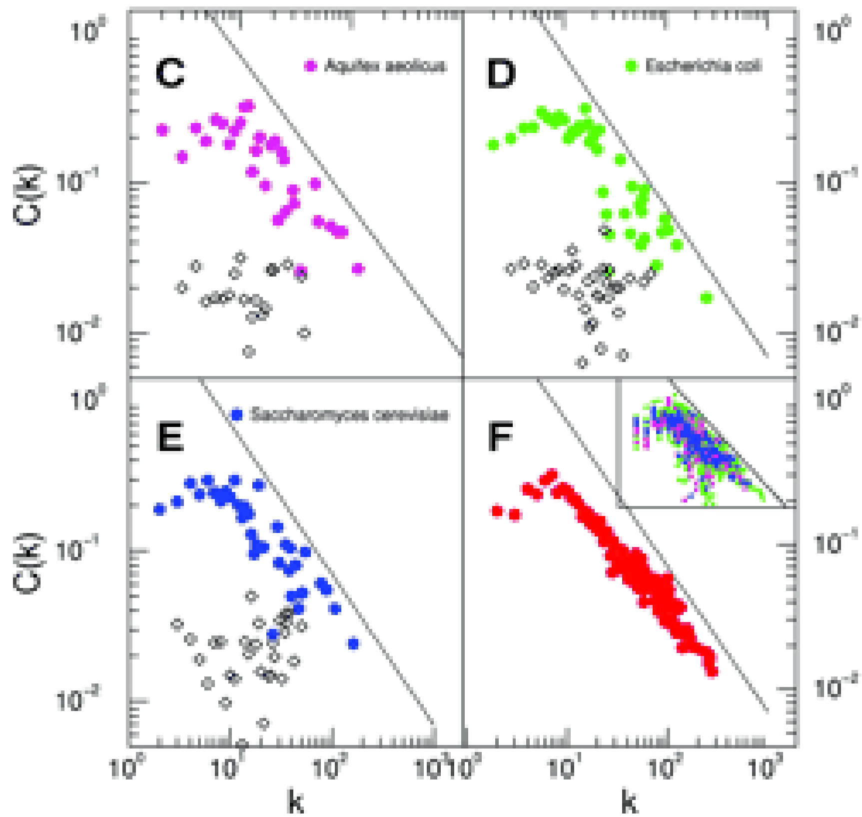




Degree distribution $P(k) \propto k^{-1-\ln 4/\ln 3}$

Clustering coefficient $C \approx 0.6$ independent of network size

# Hierarchical structure




Average clustering coefficient of nodes with degree k

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

Also observed in various cellular networks – sign of hierarchical, modular architecture

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

| | | |
|---|---|---|
| Linear growth, linear pref. attachment | $\gamma=3$ | Barabási and Albert, 1999 |
| Nonlinear preferential attachment $\Pi(k_i)\sim k_i^{\alpha}$ | no scaling for $\alpha\neq 1$ | Krapivsky, Redner, and Leyvraz, 2000 |
| Asymptotically linear pref. attachment $\Pi(k_i)\sim a_\infty k_i$ as $k_i\to\infty$ | $\gamma\to 2$ if $a_\infty\to\infty$<br>$\gamma\to\infty$ if $a_\infty\to 0$ | Krapivsky, Redner, and Leyvraz, 2000 |
| Initial attractiveness $\Pi(k_i)\sim A+k_i$ | $\gamma=2$ if $A=0$<br>$\gamma\to\infty$ if $A\to\infty$ | Dorogovtsev, Mendes, and Samukhin, 2000a, 2000b |
| Accelerating growth $\langle k\rangle\sim t^{\theta}$ constant initial attractiveness | $\gamma=1.5$ if $\theta\to 1$<br>$\gamma\to 2$ if $\theta\to 0$ | Dorogovtsev and Mendes, 2001a |
| Internal edges with probab. $p$<br>Rewiring of edges with probab. $q$ | $\gamma=2$ if $q=\frac{1-p+m}{1+2m}$<br>$\gamma\to\infty$ if $p,q,m\to 0$ | Albert and Barabási, 2000 |
| $c$ internal edges or removal of $c$ edges | $\gamma\to 2$ if $c\to\infty$<br>$\gamma\to\infty$ if $c\to -1$ | Dorogovtsev and Mendes, 2000c |
| Gradual aging $\Pi(k_i)\sim k_i(t-t_i)^{-\nu}$ | $\gamma\to 2$ if $\nu\to -\infty$<br>$\gamma\to\infty$ if $\nu\to 1$ | Dorogovtsev and Mendes, 2000b |
| Multiplicative node fitness $\Pi_i\sim\eta_i k_i$ | $P(k)\sim\frac{k^{-1-C}}{\ln(k)}$ | Bianconi and Barabási, 2001a |
| Edge inheritance | $P(k_{in})=\frac{d}{k_{in}^{\sqrt{2}}}\ln(ak_{in})$ | Dorogovtsev, Mendes, and Samukhin, 2000c |
| Copying with probab. $p$ | $\gamma=(2-p)/(1-p)$ | Kumar *et al.*, 2000a, 2000b |
| Redirection with probab. $r$ | $\gamma=1+1/r$ | Krapivsky and Redner, 2001 |
| Walking with probab. $p$ | $\gamma\simeq 2$ for $p>p_c$ | Vázquez, 2000 |
| Attaching to edges | $\gamma=3$ | Dorogovtsev, Mendes, and Samukhin, 2001a |
| $p$ directed internal edges $\Pi(k_i,k_j)\propto(k_i^{in}+\lambda)(k_j^{out}+\mu)$ | $\gamma_{in}=2+p\lambda$<br>$\gamma_{out}=1+(1-p)^{-1}+\mu p/(1-p)$ | Krapivsky, Rodgers, and Redner, 2001 |

# Lessons learned from evolving network models

1. There is no universal exponent characterizing all networks.

2. The origins of the preferential attachment might be system-dependent.

3. It is generally true that networks evolve.

4. Modeling real networks:
    - identify the processes that play a role
    - measure their frequency from real data
    - develop dynamical models that capture these processes