Frontiers of Network Science
Fall 2018

Class 14: Degree Correlations I
(Chapter 7 in Textbook)

Boleslaw Szymanski

based on slides by
Albert-László Barabási
and Roberta Sinatra

www.BarabasiLab.com

EXTENDED MODEL: Small-k cutoff

$$P(k) \sim (k+\kappa(p,q,m))^{-\gamma(p,q,m)} \qquad \gamma \in [1,\infty)$$

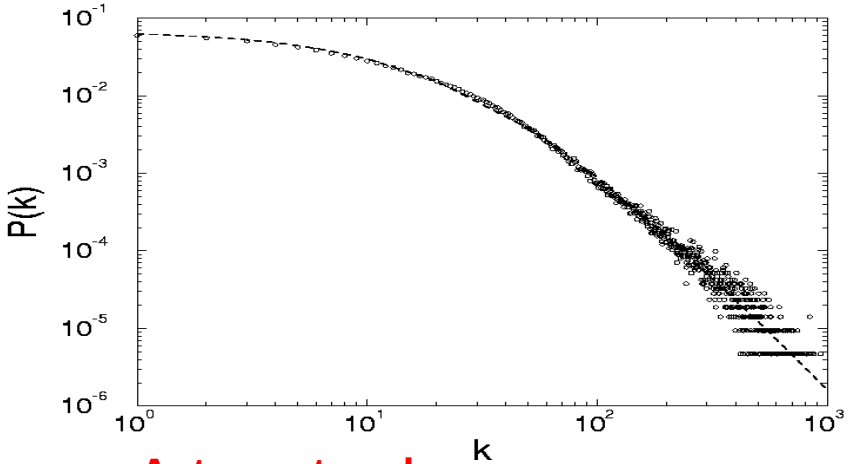
Extended Model

- prob. p : internal links
- prob. q : link deletion
- prob. 1-p-q : add node

→Predicts a small-k cutoff

→a correct model should predict all aspects of the degree distribution, not only the degree exponent.

→Degree exponent is a continuous function of *p,q*, *m*



Actor network

p=0.937

m=1

κ = 31.68

γ = 3.07

NONLINEAR PREFERENTIAL ATTACHMENT: MORE MODELS

- Non-linear preferential attachment:

$$\Pi(k) = \frac{k^{\alpha}}{\sum_i k_i^{\alpha}}$$

$\rightarrow$ P(k) does not follow a power law for $\alpha \neq 1$

$\Rightarrow$ $\alpha<1$: stretch-exponential $P(k) \approx exp\left(-(k/k_0)^{\beta}\right)$

$\Rightarrow$ $\alpha>1$: no-scaling ($\alpha>2$: “gelation”)

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

INITIAL ATTRACTIVENESS

BA model: *k=0* nodes cannot aquire links, as *Π(k=0)=0*
(the probability that a new node will attach to it is zero)

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

A - initial attractiveness

Initial attractiveness shifts the degree exponent:

$$\gamma_{in} = 2 + \frac{A}{m}$$

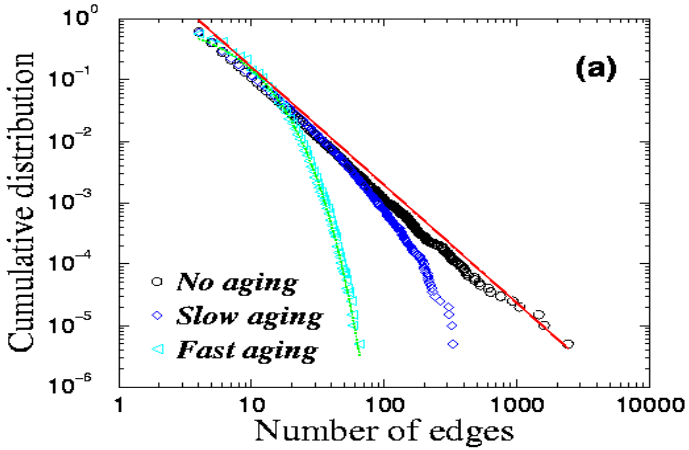
Note: the parameter A can be measured from real data, being the rate at which k=0 nodes acquire links, i.e. Π(k=0)=A

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

GROWTH CONSTRAINTS AND AGING CAUSE CUTOFFS

- Finite lifetime to acquire new edges

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

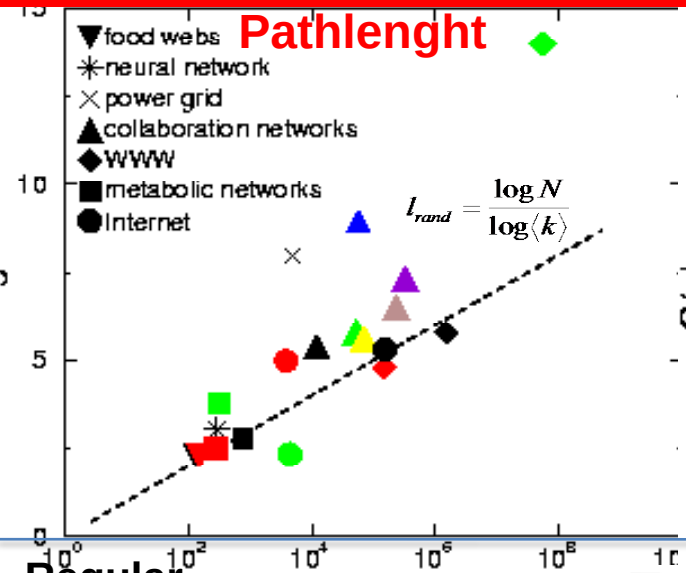


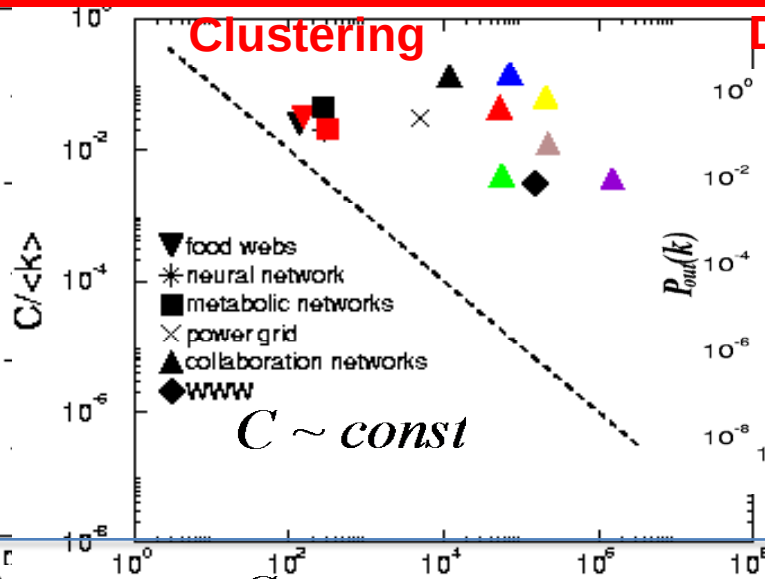
- Gradual aging: $\Pi(k_i) \propto k_i (t - t_i)^{-\nu}$

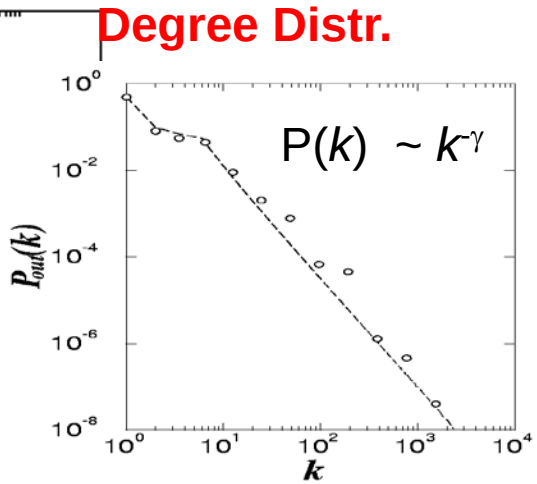
γ increases with ν

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

THE LAST PROBLEM: HIGH, SYSTEM-SIZE INDEPENDENT C(N)





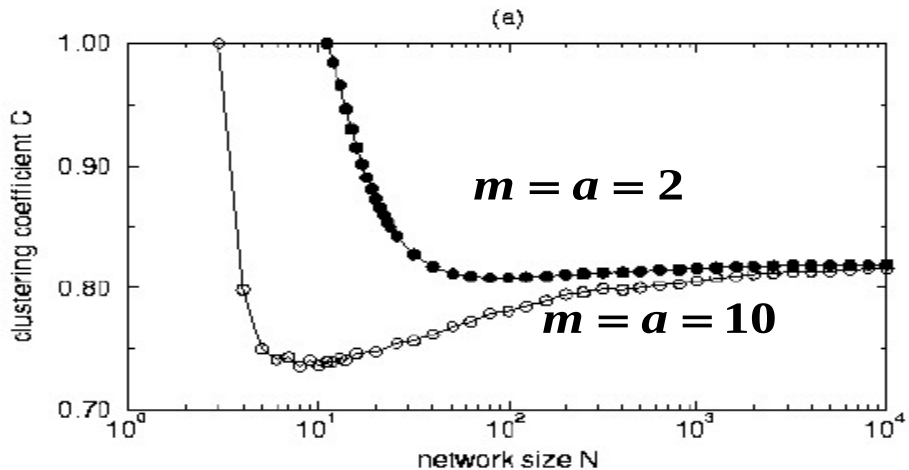


Regular network	$l \approx N^{1/D}$	$C \sim const$	$P(k)=\delta(k-k_d)$
Erdos-Renyi	$l_{rand} \approx \frac{\log N}{\log \langle k \rangle}$	$C_{rand} = p = \frac{\langle k \rangle}{N}$	$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$
Watts-Strogatz	$l_{rand} \approx \frac{\log N}{\log \langle k \rangle}$	$C \sim const$	*Exponential*
Barabasi-Albert	$l \approx \frac{\ln N}{\ln \ln N}$	$C \sim \frac{(\ln N)^2}{N}$	$P(k) \sim k^{-\gamma}$

A MODEL WITH HIGH CLUSTERING COEFFICIENT

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

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



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

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

C→ C* when N→∞

K. Klemm and V. Eguiluz, Phys. Rev. E 65, 036123 (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 $a_\infty\to\infty$ $\gamma\to\infty$ if $a_\infty\to 0$	Krapivsky, Redner, and Leyvraz, 2000
Initial attractiveness $\Pi(k_i)\sim A+k_i$	$\gamma=2$ if $A=0$ $\gamma\to\infty$ 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 $\theta\to 1$ $\gamma\to 2$ if $\theta\to 0$	Dorogovtsev and Mendes, 2001a
Internal edges with probab. p	$\gamma=2$ if $q=\dfrac{1-p+m}{1+2m}$	
Rewiring of edges with probab. q	$\gamma\to\infty$ if $p,q,m\to 0$	Albert and Barabási, 2000
c internal edges or removal of c edges	$\gamma\to 2$ if $c\to\infty$ $\gamma\to\infty$ if $c\to -1$	Dorogovtsev and Mendes, 2000c
Gradual aging $\Pi(k_i)\sim k_i(t-t_i)^{-\nu}$	$\gamma\to 2$ if $\nu\to-\infty$ $\gamma\to\infty$ if $\nu\to 1$	Dorogovtsev and Mendes, 2000b
Multiplicative node fitness $\Pi_i\sim\eta_i k_i$	$P(k)\sim\dfrac{k^{-1-C}}{\ln(k)}$	Bianconi and Barabási, 2001a
Edge inheritance	$P(k_{in})=\dfrac{d}{k_{in}^{\sqrt{2}}}\ln(ak_{in})$	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 $\Pi(k_i,k_j)\propto(k_i^{in}+\lambda)(k_j^{out}+\mu)$	$\gamma_{in}=2+p\lambda$ $\gamma_{out}=1+(1-p)^{-1}+\mu p/(1-p)$	Krapivsky, Rodgers, and Redner, 2001

Section 11: Summary

Number of Nodes

$N = t$

Number of Links

$N = mt$

Average Degree

$\langle k \rangle = 2m$

Degree Dynamics

$k_i(t) = m\,(t/t_i)^{\beta}$

Dynamical Exponent

$\beta = 1/2$

Degree Distribution

$p_k \sim k^{\gamma}$

Degree Exponent

$\gamma = 3$

Average Distance

$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient

$\langle C \rangle \sim (\ln N)^2 / N$

The network grows, but the degree distribution is stationary.

Section 11: Summary

Number of Nodes
$N = t$

Number of Links
$N = mt$

Average Degree
$\langle k \rangle = 2m$

Degree Dynamics
$k_i(t) = m\,(t/t_i)^{\beta}$

Dynamical Exponent
$\beta = 1/2$

Degree Distribution
$p_k \sim k^{\gamma}$

Degree Exponent
$\gamma = 3$

Average Distance
$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient
$\langle C \rangle \sim (\ln N)^2 / N$

Consequently, the modeling philosophy behind the model is simple: *to understand the topology of a complex system, we need to describe how it came into being.*

The network grows, but the degree distribution is stationary.

Section 11: Summary

Number of Nodes
$N = t$

Number of Links
$N = mt$

Average Degree
$\langle k \rangle = 2m$

Degree Dynamics
$k_i(t) = m\,(t/t_i)^{\beta}$

Dynamical Exponent
$\beta = 1/2$

Degree Distribution
$p_k \sim k^{\gamma}$

Degree Exponent
$\gamma = 3$

Average Distance
$\langle d \rangle \sim \log N / \log \log N$

Clustering Coefficient
$\langle C \rangle \sim (\ln N)^2 / N$

- The model predicts γ=3 while the degree exponent of real networks varies between 2 and 5 (Table 4.2).

- Many networks, like the WWW or citation networks, are directed, while the model generates undirected networks.

- Many processes observed in networks, from linking to already existing nodes to the disappearance of links and nodes, are absent from the model.

- The model does not allow us to distinguish between nodes based on some intrinsic characteristics, like the novelty of a research paper or the utility of a webpage.

- While the Barabási-Albert model is occasionally used as a model of the Internet or the cell, in reality it is not designed to capture the details of any particular real network. It is a minimal, proof of principle model whose main purpose is to capture the basic mechanisms responsible for the emergence of the scale-free property. Therefore, if we want to understand the evolution of systems like the Internet, the cell or the WWW, we need to incorporate the important details that contribute to the time evolution of these systems, like the directed nature of the WWW, the possibility of internal links and node and link removal.

LESSONS LEARNED: evolving network models

1. There is no universal exponent characterizing all networks.

2. Growth and preferential attachment are responsible for the emergence of the scale-free property.
3. The origins of the preferential attachment is system-dependent.
4. Modeling real networks:
    - identify the microscopic processes that take place in the system
    - measure their frequency from real data
    - develop dynamical models that capture these processes.

5. If the model is correct, it should correctly predict not only the degree exponent, but both small and large k-cutoffs.

LESSONS LEARNED: evolving network models

Philosophical change in network modeling:

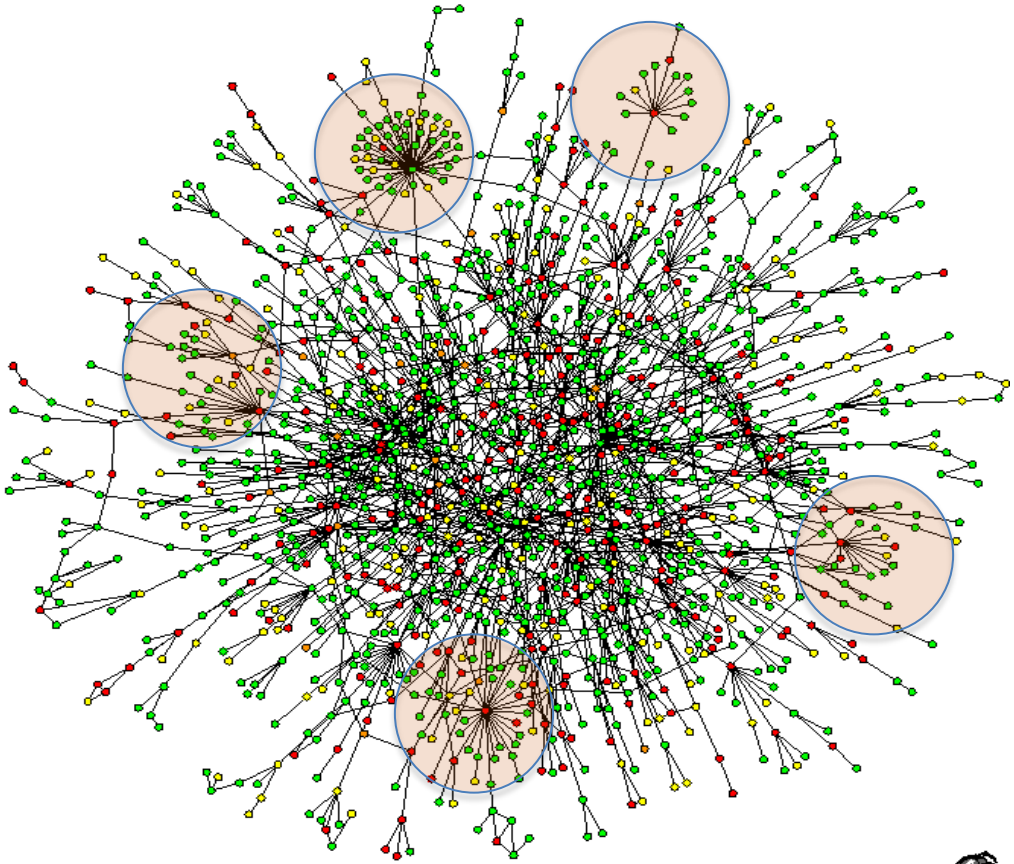
ER, WS models are static models – the role of the network modeler it to cleverly place the links between a fixed number of nodes to that the network topology mimic the networks seen in real systems.

BA and evolving network models are dynamical models: they aim to reproduce how the network was built and evolved.
Thus their goal is to capture the network dynamics, not the structure.
→ as a byproduct, you get the topology correctly

TOPOLOGY OF THE PROTEIN NETWORK

Nodes: proteins
Links: physical interactions (binding)

Puzzling pattern:
Hubs tend to link to small degree nodes.
Why is this puzzling?



In a random network, the probability that a node with degree *k* links to a node with degree *k'* is:

$$p_{kk'} = \frac{kk'}{2L}$$

k≅50, k'=13, N=1,458, L=1746

$p_{50,13} = 0.15$
$p_{2,1} = 0.0004$

Yet, we see many links between degree 2 and 1 links, and no links between the hubs.

H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Nature 411, 41-42 (2001)