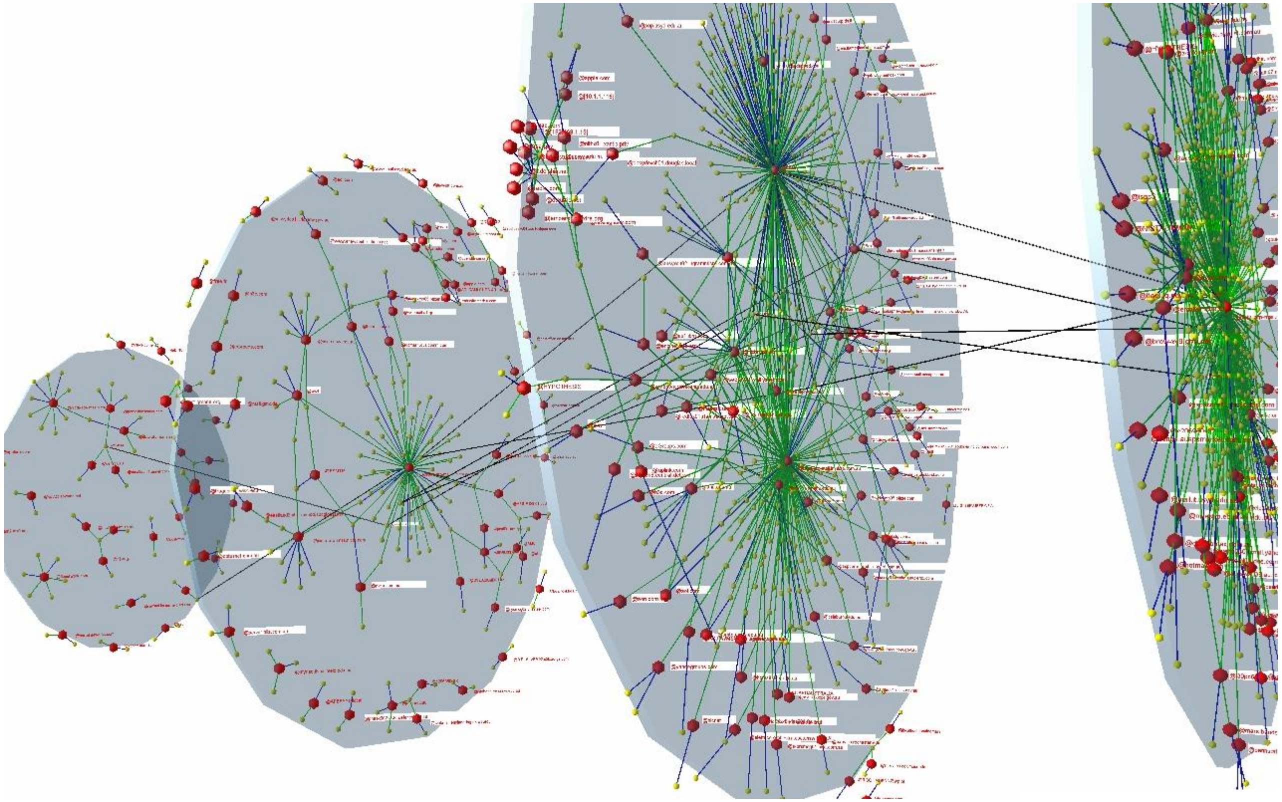
Dynamics and Evolution of Complex Networks



Dr. Michael Maragakis, UoM

Blagoevgrad 2013

European Territorial Cooperation Programme
Greece-Bulgaria 2007-2013
INVESTING IN OUR FUTURE

European Territorial Cooperation Progamme
Greece - Bulgaria 2007 - 2013
This Project is co-funded by the European Union (ERDF) and National Funds of Greece and Bulgaria

HELLENIC REPUBLIC
UNIVERSITY OF MACEDONIA
ECONOMIC & SOCIAL SCIENCES

ICoSCIS

What do we know up to know?

- Networks have a complex structure and their topology is highly irregular in most cases.
- Real world networks can be simulated.
- Real world networks are subject to attacks and/or errors.

Reminder: Small world networks

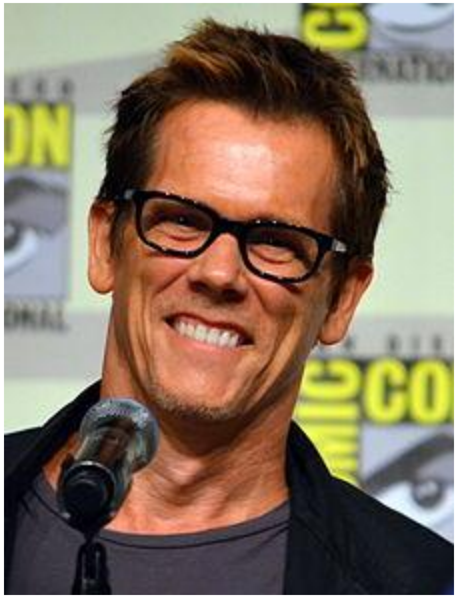
- Real networks have a small world character.
- This means that despite their often large size, there is a relatively short path between any two nodes.
- The distance of two nodes is the number of edges along the shortest path connecting them.

Reminder: Small world networks

- "Six degrees of separation" – Milgram's experiment
- The typical distance between any two nodes in small world networks scales as the logarithm of the number of nodes → random graphs
- Any Hollywood actor is on average connected to any other through just 3 connections (co-stars)!

Reminder: Small world networks

Kevin Bacon in 1994 commented that he had worked with everybody in Hollywood or someone who's worked with them. On April 7, 1994, a lengthy newsgroup thread headed "Kevin Bacon is the Center of the Universe" appeared.



Reminder: Small world networks

Six Degrees of Kevin Bacon game (based on the "six degrees of separation" concept)

That idea became a game:

- movie buffs try to find the shortest path between any actor and veteran Hollywood character actor Kevin Bacon.

It assumes that any actor involved in the Hollywood film industry can be linked through film roles to Kevin Bacon by 6 steps.

Reminder: Small world networks

The game requires a group of players to try to connect any such individual to Kevin Bacon as quickly as possible and in as few links as possible.

In 2007, Bacon started a charitable organization named SixDegrees.org.

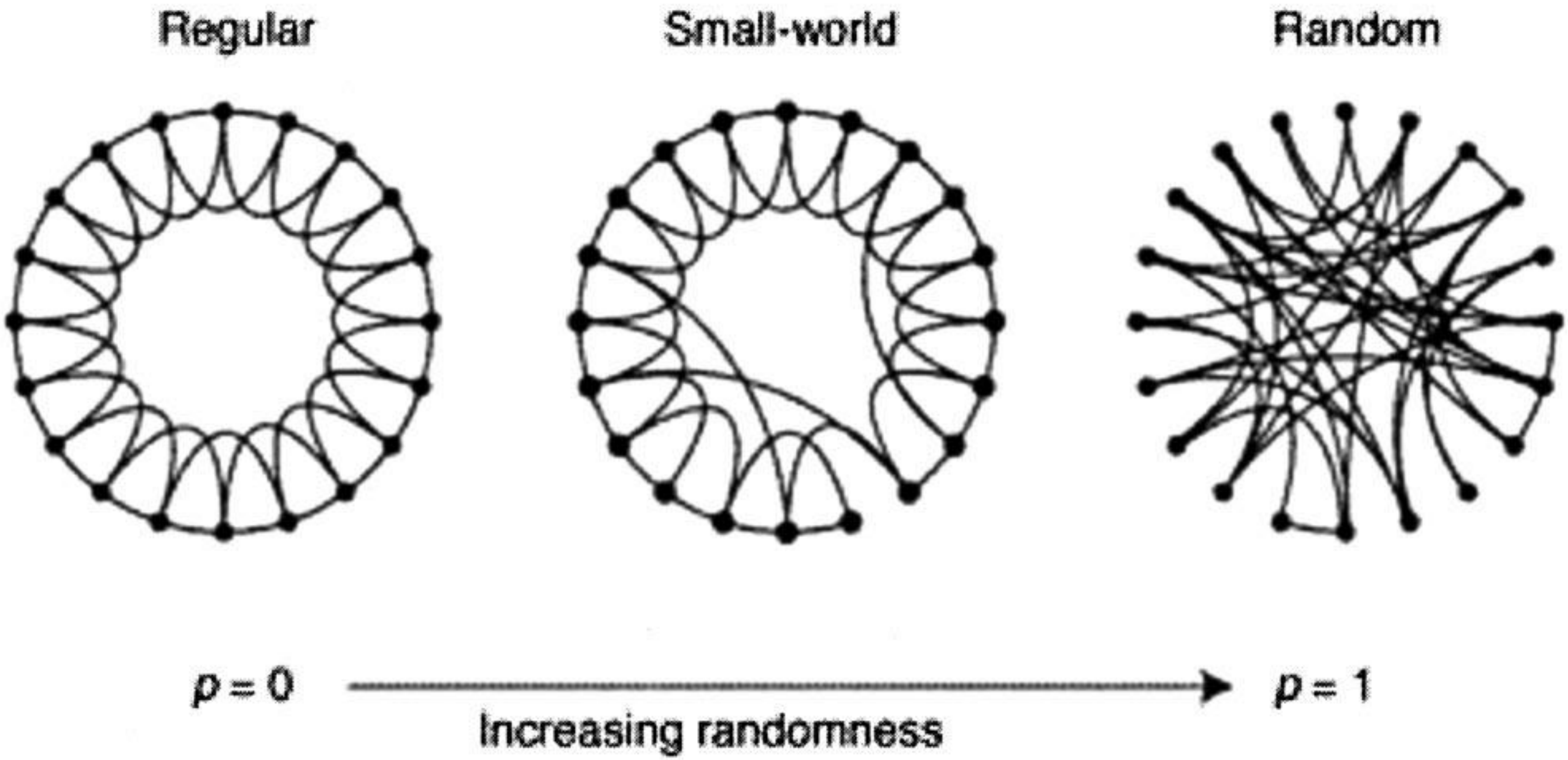
This organization helped charitable organizations do research on cancer.

Therefore, Kevin Bacon “cured cancer”.

Reminder: Watts – Strogatz model

- Watts and Strogatz is a one-parameter model that interpolates between an ordered finite dimensional lattice and a random graph.

Reminder: Watts – Strogatz model



Reminder: Watts – Strogatz model

- The algorithm behind the model is the following:
- *Start with order*: Start with a ring lattice with N nodes in which every node is connected to its first K neighbors (K/2 on either side). In order to have a sparse but connected network at all times, consider $N>>K>>\ln(N)>>1$.

Reminder: Watts – Strogatz model

- *Randomize*: Randomly rewire each edge with probability *p* such that no self-connections and duplicate edges exist.
- We introduce $pNK/2$ long-range edges which connect nodes that otherwise would be part of different neighborhoods.
- By varying *p* we monitor the transition between order (p=0) and randomness (p=1).

Reminder: Newman Watts variant

- Edges are added between randomly chosen pairs of sites.
- No edges are removed.
- It is easier to analyze than the original Watts-Strogatz model because no isolated clusters are formed.
- For sufficiently small p and large N it is equivalent to the original model.

Reminder: Barabasi Albert model

- The origin of the power-law degree distribution observed in networks was first addressed by Barabasi and Albert (1999).
- They argued that the scale-free nature of real networks is rooted in two generic mechanisms shared by many real networks.

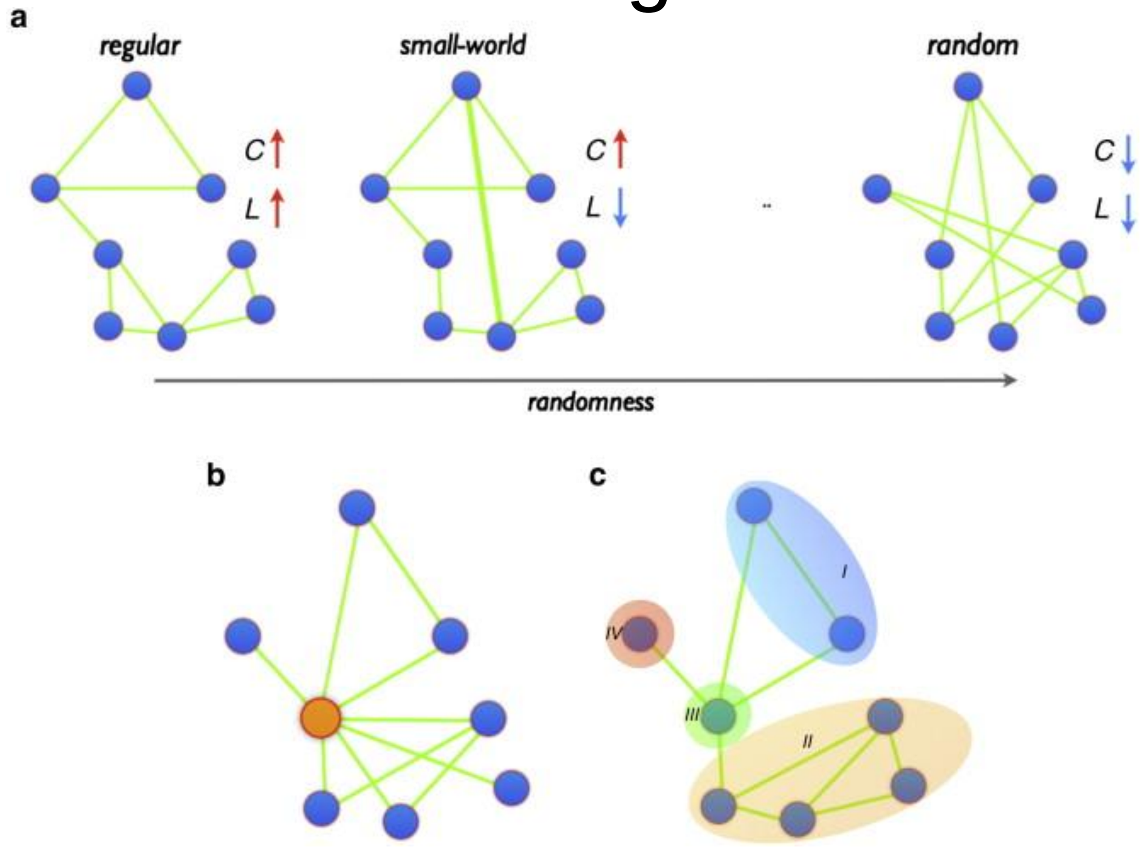
Reminder: Barabasi Albert model

- *Growth*: Starting with a small number (m_0) of nodes, add a new node with m ($<=m_0$) edges to different existing nodes.
- *Preferential attachment*: When choosing the nodes a new one connects to, assume that the probability Π to connect to node i depends on the degree k_i of node i, such that

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

Reminder: Barabasi Albert model

Watts – Strogatz model



Scale free   modular

What do we observe ourselves?

- Networks are not static.

R&D in Germany

World cup movie 1 & 2

Network topologies and evolution

They can change topology over time (cell phone, wireless device networks).

Kuramoto oscillators

What do we observe ourselves?

- They can grow over time (friendship, World Wide Web, sexual contact networks).

Evolution of “Green” companies relations, the so-called “GORT” cloud

- They can die-out over time (friendship, World Wide Web).

Sociopatterns 1 & 2

What do we observe ourselves?

- These changes can be:

a) really fast → wireless communication networks

snapshots at one point in time are meaningless in another,

b) really slow → power grid

such networks are easily considered as static.

So, networks change, evolve!

- We need to study the dynamics and the evolution of networks that change over time.
- The Barabasi-Albert (BA) model is a minimal model that captures the mechanisms responsible for the power-law degree distribution. Not very good for the dynamics and future of a network!

BA vs real networks

- Compared to real networks, the BA model has evident limitations: it predicts a power-law degree distribution with a fixed exponent, while the exponents measured for real networks vary between 1 and 3.

Real networks characteristics

Network	Size	$\langle k \rangle$	κ	γ_{out}	γ_{in}	ℓ_{real}	ℓ_{rand}	ℓ_{pow}	Reference	Nr.
WWW	325 729	4.51	900	2.45	2.1	11.2	8.32	4.77	Albert, Jeong, and Barabási 1999	1
WWW	4×10^7	7		2.38	2.1				Kumar *et al.*, 1999	2
WWW	2×10^8	7.5	4000	2.72	2.1	16	8.85	7.61	Broder *et al.*, 2000	3
WWW, site	260 000				1.94				Huberman and Adamic, 2000	4
Internet, domain*	3015–4389	3.42–3.76	30–40	2.1–2.2	2.1–2.2	4	6.3	5.2	Faloutsos, 1999	5
Internet, router*	3888	2.57	30	2.48	2.48	12.15	8.75	7.67	Faloutsos, 1999	6
Internet, router*	150 000	2.66	60	2.4	2.4	11	12.8	7.47	Govindan, 2000	7
Movie actors*	212 250	28.78	900	2.3	2.3	4.54	3.65	4.01	Barabási and Albert, 1999	8
Co-authors, SPIRES*	56 627	173	1100	1.2	1.2	4	2.12	1.95	Newman, 2001b	9
Co-authors, neuro.*	209 293	11.54	400	2.1	2.1	6	5.01	3.86	Barabási *et al.*, 2001	10
Co-authors, math.*	70 975	3.9	120	2.5	2.5	9.5	8.2	6.53	Barabási *et al.*, 2001	11
Sexual contacts*	2810			3.4	3.4				Liljeros *et al.*, 2001	12
Metabolic, *E. coli*	778	7.4	110	2.2	2.2	3.2	3.32	2.89	Jeong *et al.*, 2000	13
Protein, *S. cerev.**	1870	2.39		2.4	2.4				Jeong, Mason, *et al.*, 2001	14
Ythan estuary*	134	8.7	35	1.05	1.05	2.43	2.26	1.71	Montoya and Solé, 2000	14
Silwood Park*	154	4.75	27	1.13	1.13	3.4	3.23	2	Montoya and Solé, 2000	16
Citation	783 339	8.57			3				Redner, 1998	17
Phone call	53×10^6	3.16		2.1	2.1				Aiello *et al.*, 2000	18
Words, co-occurrence*	460 902	70.13		2.7	2.7				Ferrer i Cancho and Solé, 2001	19
Words, synonyms*	22 311	13.48		2.8	2.8				Yook *et al.*, 2001b	20

<k>→ average degree, κ → cutoff, γ_{out} → out degree, γ_{in} → in degree,
l_{real} → average path lengths of real networks,
l_{rand} → average path lengths from random-graph theory,
l_{pow} → average path lengths from power-law degree distribution

Network evolution questions

- Discrepancies between model and real networks led to increased interest in addressing questions on network evolution:

1. How can we change the scaling exponents?
2. Are there universality classes similar to those seen in critical phenomena, characterized by unique exponents?

Network evolution questions

3. How do microscopic processes, present in real networks, influence network topology?
4. Are there quantities (besides the degree distribution) that could help in classifying networks?

Network evolution questions

Results signal the emergence of a self-consistent theory of evolving networks, offering unprecedented insights into network evolution and topology.

Preferential attachment

A central ingredient of all models aiming to generate scale-free networks is preferential attachment.

Thus, we assume that the likelihood of receiving new edges increases with the node's degree.

The BA model assumes that the probability $\Pi(k)$ that a node attaches to node *i* is proportional to the degree k of node *i*.

Preferential attachment

This involves two hypotheses:

- first, that the probability $\Pi(k)$ depends on k, in contrast to random graphs in which $\Pi(k)=p$,
- and second, that the functional form of $\Pi(k)$ is linear in k.

Preferential attachment

The precise form of $\Pi(k)$ is an important question, as studies have shown that the degree distribution depends strongly on $\Pi(k)$.

To review these developments we start by discussing the empirical results on the functional form of $\Pi(k)$, followed by the theoretical work predicting the effect of $\Pi(k)$ on the network topology.

Measuring Π(k) for real networks

- The functional form of Π(*k*) can be determined for networks for which we know the time at which each node joined the network.
- Such data are available for the co-authorship network of researchers, the citation network of articles, the actor collaboration network, and the Internet at the domain level.

Measuring $\Pi(k)$ for real networks

- Consider the state of the network at a given time, and record the number of ‘‘old’’ nodes present in the network and their degrees.
- Measure the increase in the degree of the ‘‘old’’ nodes over a time interval ΔT, much shorter than the age of the network.

Measuring Π(k) for real networks

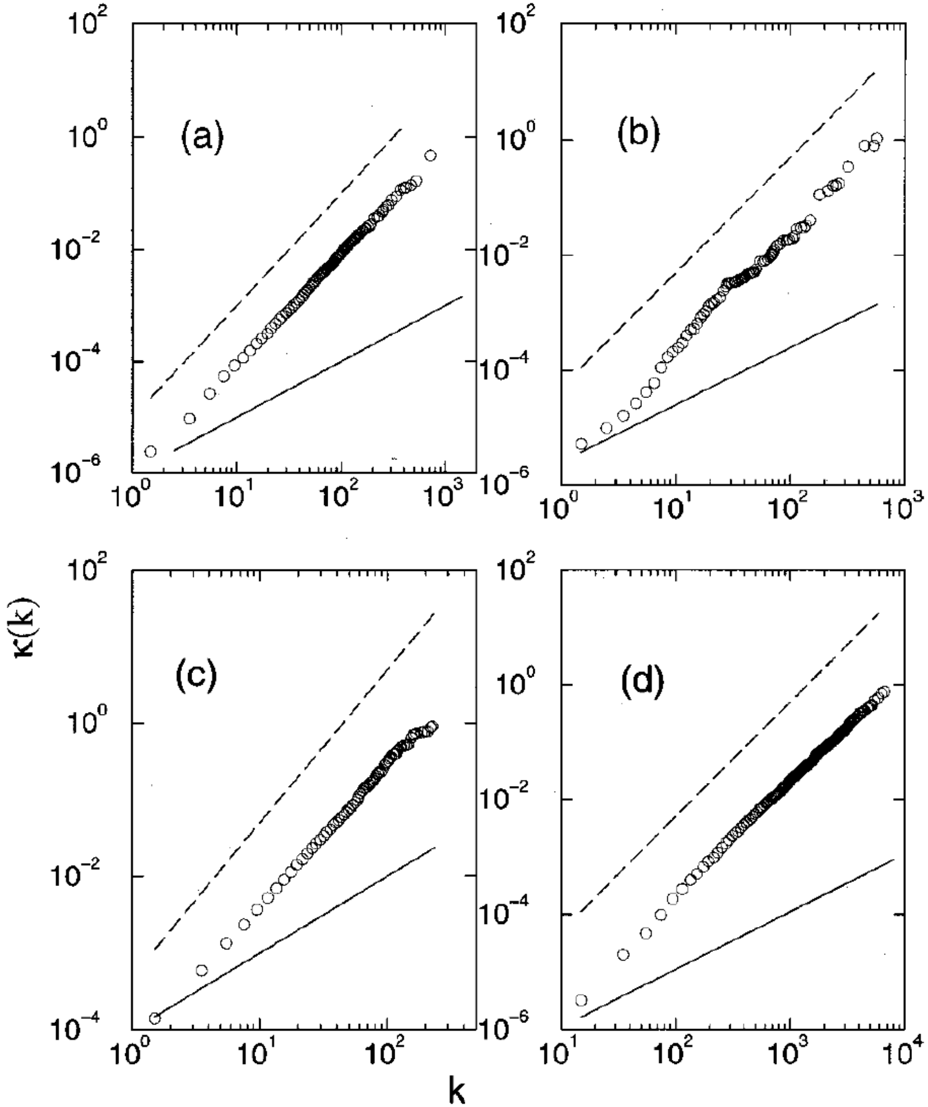
•Plotting the relative increase $\Delta k_i/\Delta k$ as a function of the earlier degree k_i for every node gives the $\Pi(k)$ function.

Δk → number of edges added in time ΔT.

We can reduce fluctuations (noise) in data by plotting the cumulative distribution

$$\boldsymbol{\kappa}(k) = \sum_{k_i=0}^{k} \Pi(k_i)$$

Cumulative preferential attachment



Cumulative preferential attachment for:
(a) citation network;
(b) Internet;
(c) neuroscience scientific collaboration network;
(d) actor collaboration network.
Dashed line corresponds to linear preferential attachment.
Solid line to no preferential attachment.

Cumulative preferential attachment

- The obtained Π(*k*) supports the existence of preferential attachment. Furthermore, it appears that in each case Π(*k*) follows a power law, i.e.,

$$\Pi(k) \sim k^{\alpha}$$

Cumulative preferential attachment

- In some cases, such as the Internet, the citation network, Medline, and the Los Alamos archive we have $\alpha \approx 1$, i.e., $\Pi(k)$ depends linearly on k as assumed in the BA model.
- For other networks the dependence is sublinear, with $\alpha = 0.8 \pm 0.1$ for the neuroscience co-authorship and the actor collaboration networks.

Nonlinear probability Π(*k*)

- We have two distinct cases:

1. *Sublinear case* ($\alpha<1$)

The degree distribution can be expressed by a series and the result is a stretch exponential in which a new term arises whenever α decreases below $1/l$, where l is an arbitrary positive integer.

$$P(k)=\frac{\mu}{k^{\alpha}}\prod_{j=1}^{k}\left(1+\frac{\mu}{j^{\alpha}}\right)^{-1}$$

Nonlinear probability $\Pi(k)$

2.Sublinear case (α>1)

There is no analytic solution.

- If α>2, A “winner takes all” phenomenon arises. Almost all nodes have a single edge that connects them to the entire network.
- If 1.5<α<2, the number of nodes with two edges grows as $t^{2-\alpha}$, while the number of nodes with more than two edges is finite. The other edges belong to the gel node.

Nonlinear preferential attachment

- Analytical calculations demonstrate that the scale-free nature is destroyed for nonlinear preferential attachment.
- It remains scale free only when the preferential attachment is asymptotically linear.
- In this case the rate equation leads to $P(k) \sim k^{-\gamma}$ with $\gamma = 1 + \mu/\alpha_{\infty}$.
- It can be tuned to any value from 2 to ∞.

Initial attractiveness

- A general feature of $\Pi(k)$ in real networks is that $\Pi(0)\neq 0$, i.e., there is a nonzero probability that a new node attaches to an isolated node.
- Thus, in general $\Pi(k)$ has the form $\Pi(k)=A+k^{\alpha}$,

  where A is the initial attractiveness of the node i.

Initial attractiveness

- if A=0, a node that has k=0 can never increase its connectivity.
- However, in real networks every node has a finite chance to be ‘‘discovered’’ and linked to, even if it has no edges to start with.

  Thus, A describes the likelihood that an isolated node will be discovered, such as a new article’s being cited the first time.

Initial attractiveness

Analytically solved model:

- at every time step a new node is added, followed by the addition of m directed edges pointing from any node in the network to preferentially chosen nodes.
- The probability that a node will receive an incoming edge is proportional to the sum of an initial attractiveness and the number of incoming edges, i.e., $\Pi(k_{in})=A+k_{in}$.

Initial attractiveness

- Calculations indicate that the degree distribution follows $\Pi(k) \sim k^{-\gamma}$ with $\gamma = 2 + A/m$.
- Consequently, the scale-free nature of the degree distribution remains even with initial attractiveness; only the degree exponent changes.

Growth

- In the BA model the number of nodes and edges increases linearly in time. Consequently the average degree (<*k*>) of the network is constant.
- What happens on the network dynamics and its topology when we have a nonlinear growth rate?

Nonlinear growth rate results

- The ability of networks to follow different growth patterns is supported by several recent measurements.

1. Internet <*k*> in November 1997 was 3.42.

   It increased to 3.96 by December of 1998. (In 13 months!!!)

2. Similarly, the World Wide Web has increased its <*k*> from 7.22 to 7.86 in just five months in 1999.

Nonlinear growth rate results

3. <*k*> of scientific coauthorship has continuously increased over an eight-year period.
4. Comparison of metabolic networks of organisms of different sizes indicates that the <*k*> of the substrates increases approximately linearly with the number of substrates involved in the metabolism.

Nonlinear growth rate results

- The increase of the <*k*> indicates that in many real systems the number of edges increases faster than the number of nodes, supporting the presence of a phenomenon called ***accelerated growth***.

Nonlinear growth rate results

- Opposite effects can also happen. $<k>$ can be reduced with a growing network.
- A Facebook sample of 5.8 Million users on 2007 had $<k>=5.73$.

  In 2011 Facebook said that the 723 Million users (it had then) have $<k>=4.74$.
- Also, 99.91% of Facebook users were interconnected, forming a large connected component.

Example

- LinkedIn (as a service) operates based exactly on this idea.
- It tells you how many steps you are away from a person you wish to communicate with.

Example

- i.e. Last week I was connected to people from AUBG only (SWU teachers do not have a Linkedin account). One of them is connected to a Tanya Tancheva, so if you know her, you are my 3rd degree connection! If you know someone who knows her you are my 4th degree connection!

Example

- This week I may connect to Ivan Trenchev.
- The site also encourages you to pass messages to people in your network via the people in your 1st-degree connections list, who in turn pass it to their 1st-degree connections.
- If you become connected to them then the networks $\langle k \rangle$ is reduced.

Analytical results

- Since I see you are not into analytical results, I will skip those….

Analytical results

Just kidding!

Now comes the interesting (who am I kidding?) part!

Analytical results

- An analytic study of the effect of accelerated growth on the degree distribution, generalized the directed model with asymptotically linear preferential attachment.
- At every step a new node is added to the network, receiving *n* incoming edges from random nodes in the system.

Analytical results

Additionally, $c_0 t^\theta$ new edges are distributed, each directed from a random node to one with high *in* degree, with asymptotically linear preferential attachment $P(k_{in}) \sim A + k_{in}$.

Results showed accelerated growth, controlled by the exponent θ. The scale-free nature of the degree distribution was not affected (only different degree exponent)

$$\gamma = 1 + \frac{1}{1+\theta}$$

Analytical results

While this model is based on a directed network, other models use an undirected model motivated by measurements on the evolution of the co-authorship network.

In the undirected model new nodes connect to the system with a constant rate to *b* *existing* nodes with preferential attachment

$$P_i = b \frac{k_i}{\sum_j k_j}$$

Analytical results

Additionally, at every time step a linearly increasing number of edges (constituting a fraction *α* of the nodes that are present in the network) are distributed between the nodes. The probability that an edge is added between nodes *i* and *j* being

$$P_{ij} = \frac{k_i k_j}{\sum_{s,l}' k_s k_l} N(t) a$$

Analytical results

$N(t)$ is the number of nodes in the system and the summation goes over all nonequal values of s and l.

As a result of these two processes the $\langle k \rangle$ increases linearly in time, following $\langle k \rangle = \alpha t + 2b$, in line with real co-author network data.

Analytical results

Continuum theory predicts that the time-dependent degree distribution displays a crossover at a critical degree,

$$k_c = \sqrt{b^2 t}(2 + 2at/b)^{3/2}$$

for $k = k_c$, $P(k)$ follows a power law with exponent $\gamma = 1.5$ and for $k \gg k_c$ the exponent is $\gamma = 3$.

Analytical results

This explains the fast-decaying tail of the degree distributions measured, and indicates that as time increases the scaling behavior with γ=1.5 becomes increasingly visible.

Growth constraints

- For many real networks the nodes have

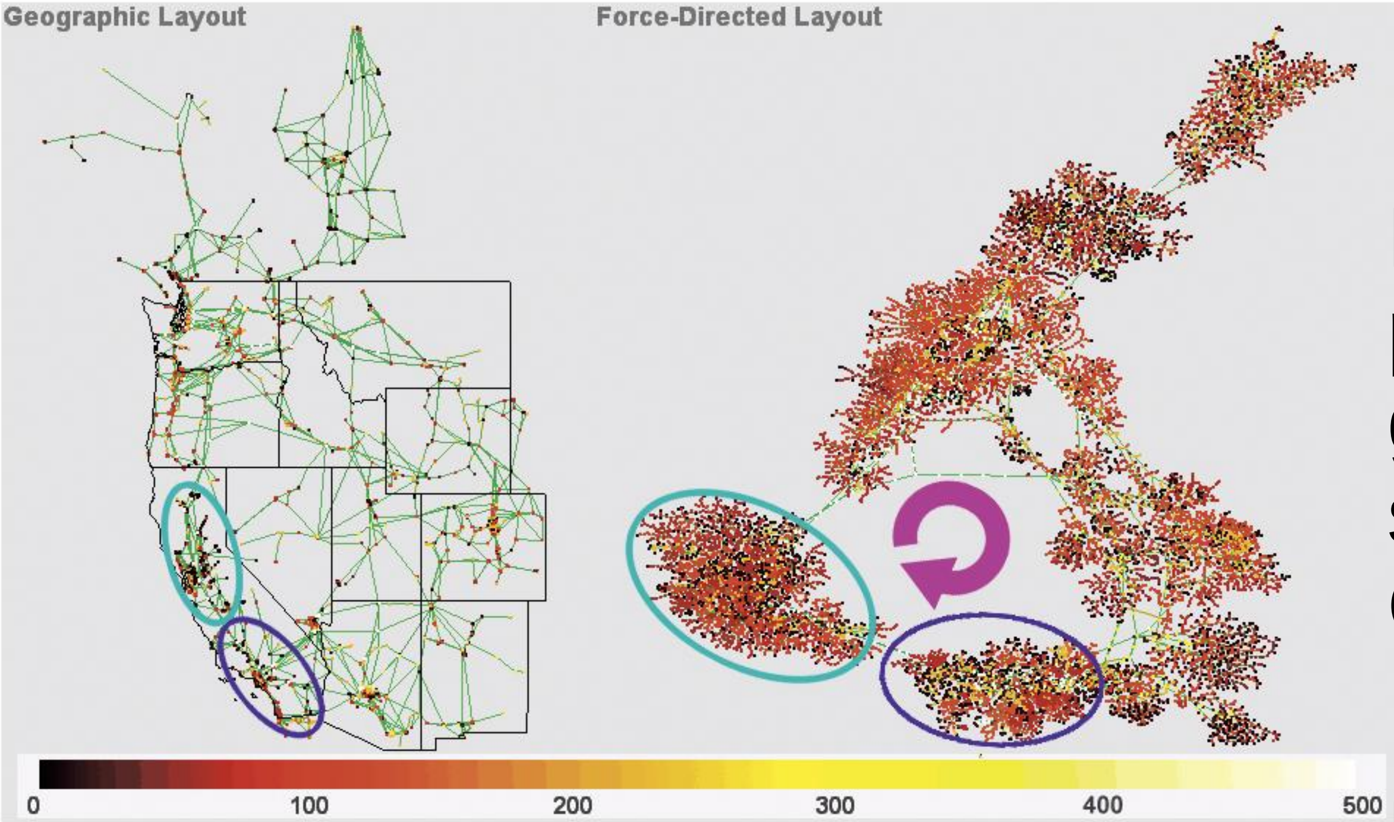
a. a finite lifetime. For example, in social networks – people-nodes, or friendships-connections die out
b. a finite edge capacity. Internet routers or nodes in the electrical power grid.

- The extend to which such constraints affect the degree distribution has been studied.

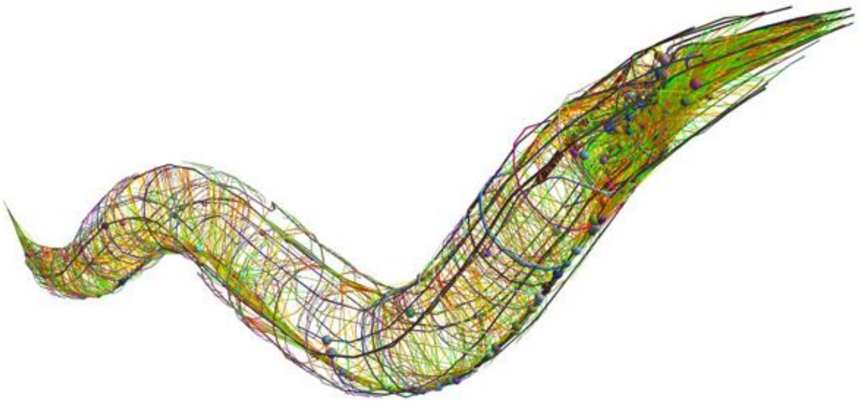
Aging and cost

- While several networks show deviations from the power-law behavior, they are far from being random networks.
- For example, the degree distribution of the
  - electric power grid of southern California
  - neural network of the worm *C. elegans*

  is more consistent with a single-scale exponential distribution.



Electric power grid of southern California



Neural network of the worm *C. elegans*

Aging and cost

- Other networks (extended actor collaboration network) have a degree distribution where power-law scaling is followed by exponential cutoff for large k.
- Constraints limit the addition of new edges.

Aging and cost

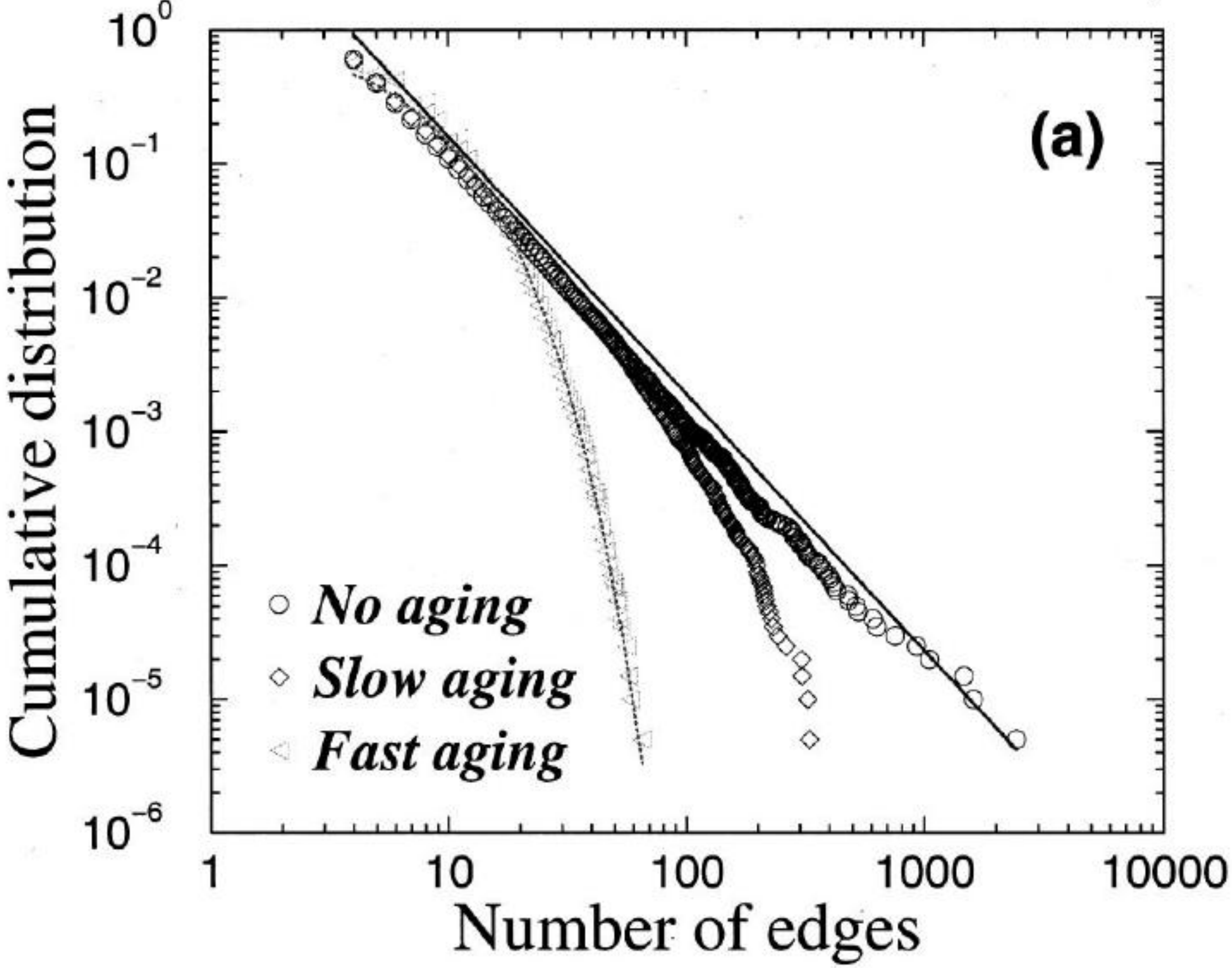
For example

- the actors have a finite active period during which they collect new edges (retirement)
- the electrical power grid or other neural networks, have constraints driven by economic, physical, or evolutionary reasons (too expensive to build new lines – “evolutionary” expensive to do the same).

Aging

- In order to explain these deviations from a pure power law aging and cost or capacity constraints must be incorporated.
- Suggested models follow growth and preferential attachment, but when a node reaches a certain age (aging) they do not allow new edges to connect.

Aging

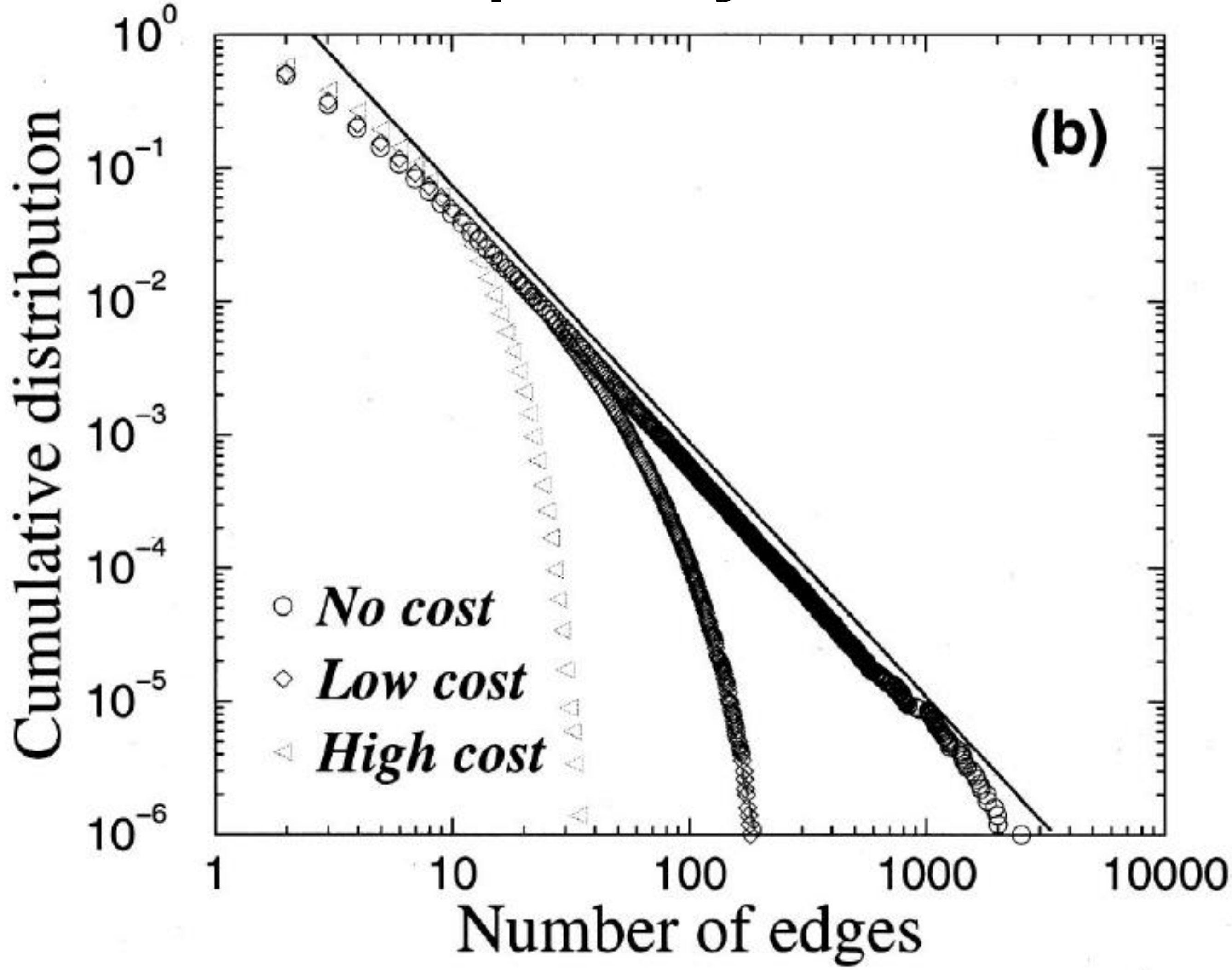


(a)
Cumulative distribution
Number of edges
No aging
Slow aging
Fast aging
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
1
10
100
1000
10000

Capacity cost

- Similar models as before with one main difference in the constraints produce similar results.
- Instead of aging they take into account capacity costs. If a node has more than a critical number of edges they do not allow new edges to connect.

Capacity cost



Aging and cost results

In both cases numerical simulations indicate that while for small *k* the degree distribution still follows a power law, for large *k* an exponential cutoff develops.

Local events

- The BA model has only one mechanism for network growth: addition of new nodes connecting to existing nodes.
- In real systems, a series of microscopic events shape the network evolution, i.e. addition or rewiring of new edges or removal of nodes or edges.

Example time – wake up!

In simple terms the Facebook analogue is:

- Your connections increase when you connect to existing nodes you had not connected before (by simply finding them as being friends of a friend that you happen to also know).
- They also reduce if you remove a link to someone or if they decide to stop using Facebook (*as if this happens*…).

Local events

- Several models exist to investigate the effect of selected processes on the scale-free nature of the degree distribution, offering a more realistic description of various real networks.
- Any local change in network topology can be obtained through a combination of four elementary processes: addition or removal of a node or an edge.

Local events

- But in reality these events come jointly.
- For example, the rewiring of an edge is a combination of an edge removal and the addition of another edge originating from one (common with before) node to another one.

Internal edges and rewiring

- Several models exist that incorporate new edges between existing nodes and rewire current ones.
- Starting with m_0 isolated nodes, at each time step we perform one of the following three operations:

Internal edges and rewiring

i. With probability p we add $m(m \leq m_0)$ new edges.

One end of a new edge is selected randomly, the other with probability given by

$$\Pi(k_i) = \frac{k_i + 1}{\sum_j (k_j + 1)}$$

Internal edges and rewiring

ii. With probability q we rewire m edges.

We randomly select a node i and remove an edge l_{ij} connected to it, replacing it with a new edge $l_{ij'}$ that connects i with node j' chosen with probability $\Pi(k'_j)$ given by the previous equation.

Internal edges and rewiring

iii. With probability 1-*p*-*q* we add a new node.

The new node has m new edges that are connected to nodes *i* already present in the system with probability $\Pi(k_i)$.

Internal edges and rewiring

- The growth rate of the degree of a node *i* is given by $$\frac{\partial k_i}{\partial t} = (p-q)m\frac{1}{N} + m\frac{k_i+1}{\sum_j (k_j+1)}$$
- The first term on the right-hand side corresponds to random selection of node *i* as a starting point of a new edge (with probability *p*) or as end point from which an edge is disconnected (with probability *q*).

Internal edges and rewiring

- The growth rate of the degree of a node *i* is given by $$\frac{\partial k_i}{\partial t} = (p-q)m\frac{1}{N} + m\frac{k_i+1}{\sum_j (k_j+1)}$$
- The second term corresponds to the selection of node *i* as an end point of an edge with the preferential attachment present in all three of the possible processes.

Internal edges and rewiring

- The degree distribution has the generalized power-law form (for $q>q_{max}$→ too many edges rewired)

$$P(k) \propto [k+\kappa(p,q,m)]^{-\gamma(p,q,m)}$$

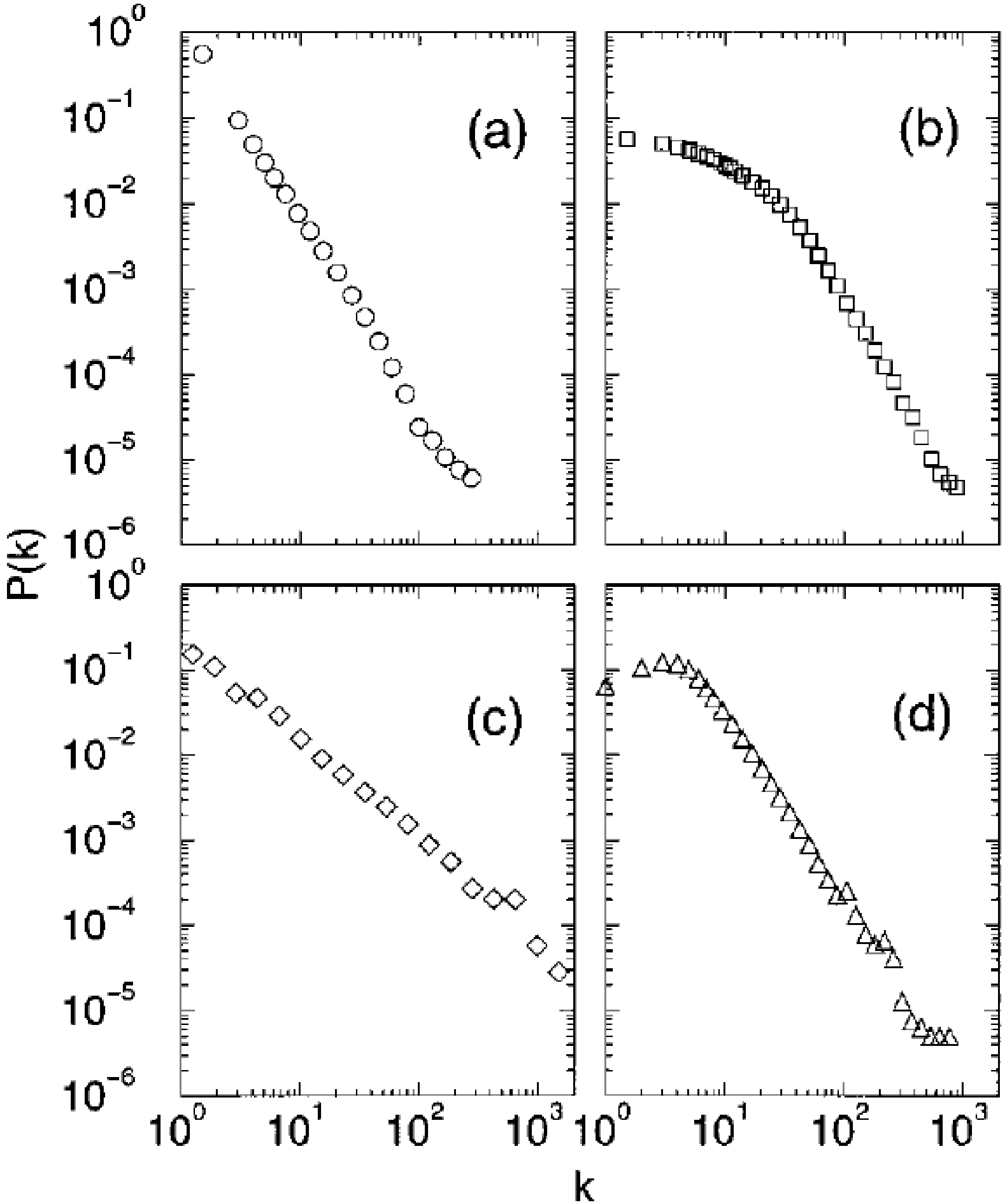
- For $q>q_{max}$ numerical simulations indicate that $P(k)$ approaches an exponential and the above equation is not valid.

Internal edges and rewiring

- While a power-law tail is present in any point of the scale-free regime, for small *k* the probability saturates at $P[\kappa(p,q,m)]$ (there is a kind of a plateau), a feature seen in real networks:

➢ Movie actor collaboration networks (b)

➢ Coauthorship network in neuroscience (d)



The degree distribution of several real networks:

(a) Internet at the router level.

(b) movie actor collaboration network.

(c) co-authorship network of high-energy physicists.

(d) co-authorship network of neuroscientists.

Internal edges and edge removal

- Let us consider a class of undirected models in which new edges are added between old sites and existing edges can be removed.
- In the first variant of the model, called a developing network, *c* new edges are introduced at every time step.

Internal edges and edge removal

- These edges connect two unconnected nodes *i* and *j* with a probability proportional to the product of their degrees [as in slide 54], an assumption confirmed by empirical measurements on the co-authorship network.
- This is explained by the fact that new researchers are added as young people write their first papers.

Internal edges and edge removal

- It is assumed that c can be tuned continuously, such that $c>0$ for a developing and $c<0$ for a decaying network.
- A two term rate of change of the node degrees can be calculated. One term corresponds to the linear preferential attachment and the other to the addition of c new edges.

Competition in evolving networks

- The BA model assumes that all nodes increase their degree following a power-law time dependence with the same dynamic exponent β=1/2.

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}$$

➢ Thus, the oldest nodes have the highest number of edges, since they had the longest lifetime to accumulate them.

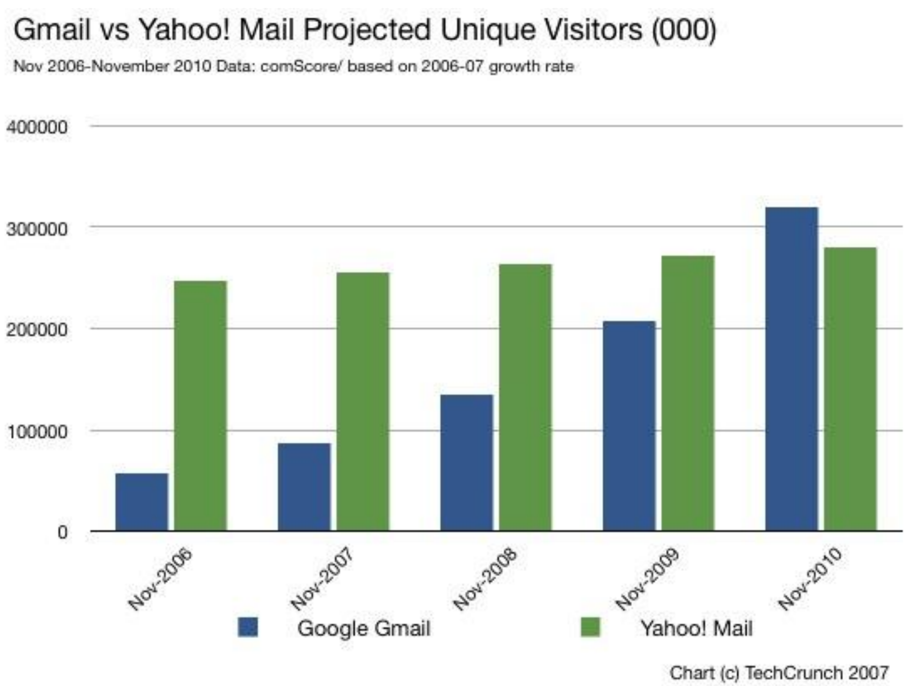
Competition in evolving networks

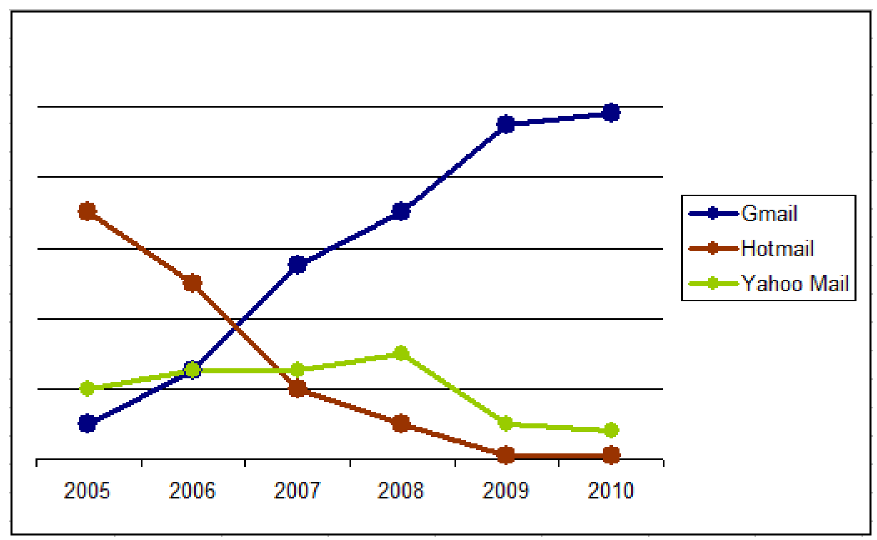
However, numerous examples indicate that in real networks a node's degree and growth rate do not depend on age alone.

- On the WWW some documents acquire a large number of edges in a very short time through good content and/or marketing.
- Some research papers acquire many more citations than their peers.

Example case

Launched as a beta version on AprilFools Day (1/4/2004).





Competition in evolving networks

How can such a phenomenon be explained?

Two models have tried with relatively good success to account for such cases.

1. Fitness model
2. Edge inheritance

Fitness model

- Real networks have a competitive aspect, as each node has an intrinsic ability to compete for edges at the expense of other nodes.
- They propose a model in which each node is assigned a fitness parameter η_i which does not change in time.

Fitness model

- Thus, at every time step a new node j with a fitness η_j is added to the system, where η_j is chosen from a distribution $\rho(\eta)$.
- Each new node connects with m edges to nodes present in the network
- The probability of connecting to a node i is proportional to the degree and the fitness of node i

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

Fitness model

- This generalized preferential attachment (compare with slide 14) ensures that even a relatively young node with a few edges can acquire edges at a high rate if it has a high fitness parameter. The rate of change of the degree of node *i* is

$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_k \eta_j k_j}$$

Fitness model

- Assuming that the time evolution of k_i with a fitness-dependent dynamic exponent $\beta(\eta)$, is given by

$$k_{\eta_i}(t, t_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}$$

- Thus, nodes with higher fitness increase their degree faster than those with lower fitness. The fitness model allows for late but fit nodes to take a central role in the network topology.

$$\beta(\eta) = \frac{\eta}{C}$$

Facebook example

- If there was no fitness then someone creating a Facebook account today would never be able to surpass the number of connections I have (I have a Facebook account since Oct 2007).
- I have practically stopped using Facebook for about 3-4 years.

Facebook example

- I have very low fitness in the expansion of Facebook.
- I do not even accept new friendship requests since I do not login!
- An active (fit) account from you can make more connections in a very low amount of time, although it is newer.

Edge inheritance

- A different mechanism that gives individuality to the new nodes exists.
- It builds on an evolving directed network algorithm, and assumes that the degree of the new nodes is not constant but depends on the state of the network at the time the new node is added to the system.

Edge inheritance

- Every new node is assumed to be an “heir” of a randomly chosen old node, and it inherits a fraction c of the old node’s incoming edges (i.e., a fraction c of the nodes that point to the parent node will also point to the heir).
- The parameter c is assumed to be distributed with a probability density $h(c)$.

Edge inheritance

- The time-dependent degree distribution for uniformly distributed c indicates that the fraction of nodes with no incoming edges increases and tends to 1 asymptotically.

Other types of network growth

- When a researcher enters a new scientific field, he is usually aware of only a few important papers and follows the references included therein to find other relevant articles.
- This process is continued recursively both ways (up and down), such that a manuscript will contain references to papers discovered this way.

Walking on a network

- This process is called “walking on a network”.
- It resembles a random walk in a network.
- The difference is that we can create “clones” of our walkers in each node and depending on the probability to follow this edge we will move to the next node or not.

Walking on a network

- A network algorithm that does this can be stated in the following way:
  - We start with an isolated node.
  - At every time step a new node is added with a directed edge to a random node.
  - We then follow edges starting from this node to another with probability p.

Walking on a network

- This last step is repeated starting from the nodes to which connections were established, until no new target node is found.
- High probability is practically a breadth-first search algorithm. If $p \rightarrow 1$ this means we will practically find ALL nodes in a network.
- Low probability means halting quite fast.

Walking on a network

- In the special case of $p=1$ one can see that nodes of high degree will be more likely to acquire new incoming edges, leading to a preferential attachment $\Pi(k)=(1+k)/N$.
- Consequently, the degree distribution follows a power law with $\gamma=2$.

Walking on a network

- If p varies between 0 and 1, simulations indicate a phase transition:

➢ for $p<p_c\approx 0.4$ the degree distribution decays exponentially,

➢ for $p>p_c$ it has a power-law tail with γ very close to 2 (the value of p=1).

- Thus, while the model does not explicitly include preferential attachment, the mechanism responsible for creation of the edges induces one.

Attaching to edges

- Perhaps the simplest model of a scale-free network without explicit preferential attachment is the following: at every time step a new node connects to both ends of a randomly selected edge.

Attaching to edges

- Consequently, the probability that a node will receive a new edge is directly proportional to its degree; in other words, this model has exactly the same preferential attachment as the BA model.
- It readily follows that the degree distribution has the same asymptotic form as the BA model, i.e., $P(k) \sim k^{-3}$.

Percolation in networks

- In grids and other organized lattices, in any dimension larger than one, a percolation phase transition occurs. The usual percolation model assumes that sites (nodes) or bonds (links) in the lattice are occupied with some probability (or density), p, and unoccupied with probability $q=1-p$.

Percolation in networks

- The system is considered percolating if there is a path from one side of the lattice to the other, passing only through occupied links and nodes.
- When such a path exists, the component or cluster of sites that spans the network from side to side is called the spanning cluster or the infinite cluster.

Percolation in networks

- The percolation phase transition occurs at some critical density p_c that depends on the type and dimensionality of the lattice.
- In networks no notion of side exists. However, the ideas of percolation theory can still be applied to obtain useful results.

Percolation in networks

- Main difference with lattices:
- ➢The condition for percolation is no longer the spanning property, but rather the property of having a component (cluster) containing $O(N)$ nodes, where N is the total original number of nodes in the network.
- Such a component, if it exists, is termed the **giant component**.

Percolation in networks

- The condition of the existence of a giant component above the percolation threshold and its absence below the threshold also applies to lattices, and therefore can be considered as more general than the spanning property.

Percolation in networks

- An interesting property of percolation, called **universality**, is that the behavior at the critical point (and near it) depends only on the dimensionality, and not on the microscopic connection details of the lattice.
- There is a set of “critical exponents” that are the same for all *d*-dimensional lattices, even if they are square, triangular, or hexagonal, site or bond percolation.

Percolation in networks

- A different set will be obtained for another dimension of a lattice.
- Above some critical *d* (d_c=6 for percolation in *d*-dimensional lattices), called **upper critical dimension**, the critical behavior remains the same, regardless of *d*.
- In high dimensions loops are insignificant, allowing us to find the critical exponents for high dimensions, by an "*infinite-dimensional*" or "*mean-field*" approach.

Percolation in networks

- Percolation on ER networks or on Cayley trees has the same critical exponents as for lattices above the upper critical dimension, since no spatial constraints are imposed on these networks.
- In fact, the heterogeneity of the degrees may still affect the critical behavior even above the critical dimension.

Percolation in networks

- In such cases each node has a typical number of neighbors, whereas in SFN the variation of node degrees is very large.
- The results are still "*mean field*" or "*infinite dimensional*" for the insignificance of the loops.
- However, results that differ from the standard mean-field percolation solution are obtained.

Percolation in networks

- For scale-free networks with $\gamma > 4$, the critical exponents are same as for ER and lattices in $d \geq d_c = 6$. γ is large enough and the translational symmetry (approximately) exists.
- For $\gamma < 4$, however, topology is different, thus, critical exponents are different.
- As a result, scale-free networks are like a generalization of ER networks.