Economic Networks

Theory and Empirics

Giorgio Fagiolo
Laboratory of Economics and Management (LEM)
Sant'Anna School of Advanced Studies, Pisa, Italy

http://www.lem.sssup.it/fagiolo/
giorgio.fagiolo@sssup.it

Network Models

What's Next

- What is a network? Examples of networks
- Why networks are important for economists?
- Networks and graphs
- Measures and metrics on networks
- Distributions of metrics and measures in large networks
- Models of network formation
- Null statistical network models
- Economic applications

Empirical Findings vs. Theoretical Models

- Empirics of complex networks
  - ✓ Identifying a series of universal properties characterizing the topological architecture of networks in biology, computer science, sociology, economics, etc.
  - ✓ Why different networks exhibit similar structural properties?
- Universal properties
  - ✓ Small diameters and APLs
  - ✓ Large clustering
  - ✓ Bell-shaped or power-law degree distributions
- This lecture
  - ✓ Empirical findings
  - ✓ Simple graph-theoretic models able to explain these empirical properties

Stylized fact #1: It is a small world!

- Degrees of Separation
  - ✓ How many steps are required to get in touch with an arbitrarily far node?
  - ✓ Is it always possible to do that (i.e. is the graph connected)?
  - ✓ What is the average path length?

- Milgram's Experiment: Is it a small world?
  - ✓ How long is the shortest path connecting unacquainted people?
  - ✓ Letters gave to hundred of people in Wichita, KS and Omaha, NE, to be sent to XY, who works (but does not live) in Boston
  - ✓ Rule: Letter can be actually sent to a person whom you know personally. If you know XY then you can send the letter directly to him

- How many letters needed to get to XY?
  - ✓ On average, only 6 !!
  - ✓ Anyone in the planet is reachable in no more than 6 steps!

Stylized fact #1: It is a small world!

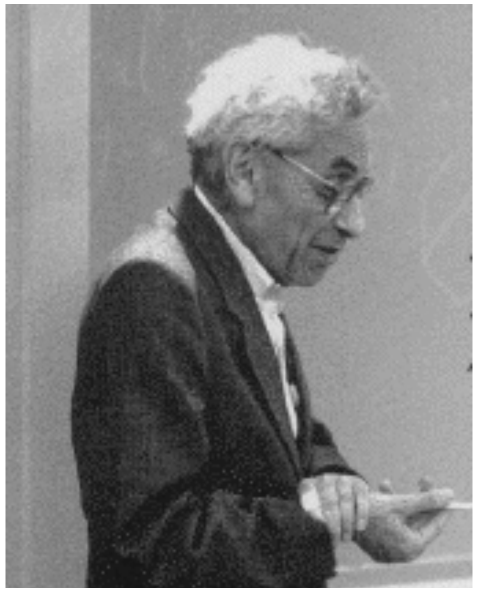
- Is it really so for all networks?
  - ✓ Pages in the WWW: 19 steps (over a few billions of nodes?)
  - ✓ Species in food webs and world trade network: 2 steps
  - ✓ Co-authorships: 4-6 steps
  - ✓ All other networks: between 2 and 14. They are really small worlds!!

- Does SW mean "easy to find"?
  - ✓ Not actually!
  - ✓ Small degrees of separation come from high node degree
  - ✓ You have to take the right path each time you travel trough a node! Search must be intelligent and not random.

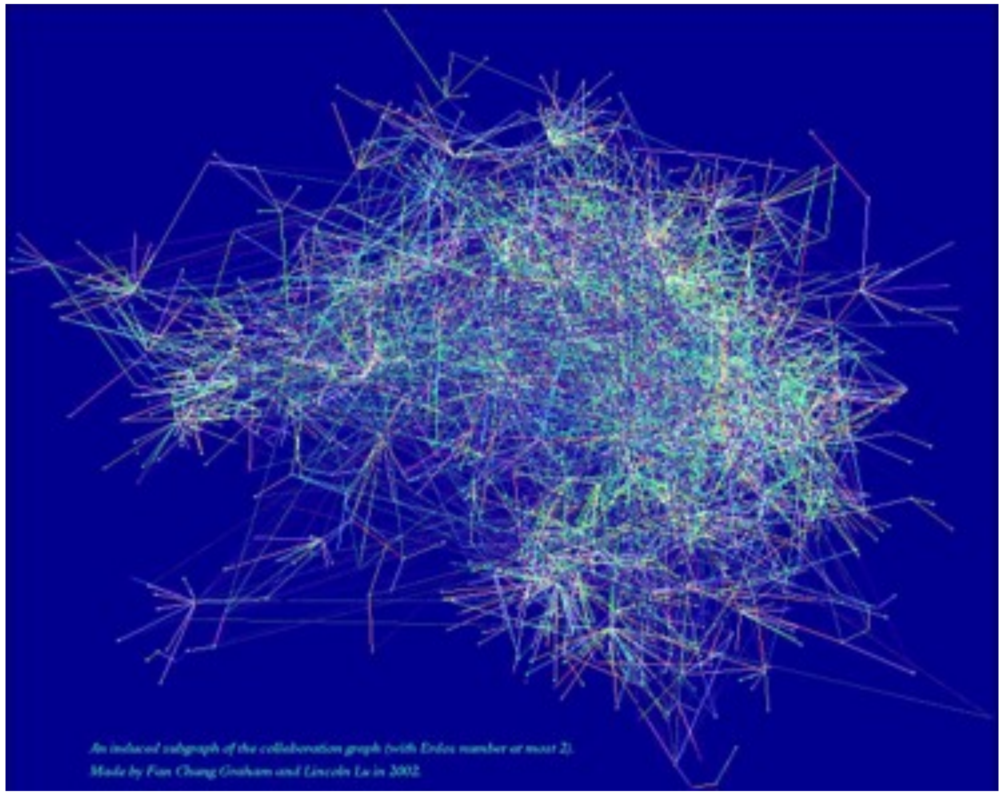
- Is there a model explaining SF #1?
  - ✓ Yes, the good-old Erdos-Renyi random graph model

The Erdos-Renyi Random Graph Model (1)

- A simple and beautiful model
  - ✓ Paul Erdos: an extremely prolific mathematical “pilgrim”
  - ✓ He wrote around 1,525 mathematical articles in his lifetime
  - ✓ The Erdos network and Erdos number



- Erdos was the founder of the field of random graphs
  - ✓ Hoffman, Paul (1998), The man who loved only numbers: the story of Paul Erdos and the search for mathematical truth, Hyperion
  - ✓ Bollobas (2001), Random Graphs, Cambridge University Press



The Erdos network with EN≤2

The Erdos-Renyi Random Graph Model (2)

- The G(N,p) model
  - ✓ Take N nodes initially not connected
  - ✓ Go through each of the $\binom{N}{2}$ possible links
  - ✓ Form a link with a probability $0<p\leq 1$ (iid)

- The G(N,m) model
  - ✓ Draw at random m links from all $\binom{N}{2}$ possible links

- Large-scale system properties
  - ✓ Take $N\to\infty$ as $N\cdot p(N)\to c$

ON THE EVOLUTION OF RANDOM GRAPHS

by

P. ERDÖS and A. RÉNYI

Institute of Mathematics
Hungarian Academy of Sciences, Hungary

1. *Definition of a random graph*

Let $E_{n,N}$ denote the set of all graphs having n given labelled vertices $V_1, V_2, \cdots, V_n$ and N edges. The graphs considered are supposed to be not oriented, without parallel edges and without slings (such graphs are sometimes called linear graphs). Thus a graph belonging to the set $E_{n,N}$ is obtained by choosing N out of the possible $\binom{n}{2}$ edges between the points $V_1, V_2, \cdots, V_n$, and therefore the number of elements of $E_{n,N}$ is equal to $\binom{\binom{n}{2}}{N}$. A random graph $\Gamma_{n,N}$ can be defined as an element of $E_{n,N}$ chosen at random, so that each of the elements of $E_{n,N}$ have the same probability to be chosen, namely $1/\binom{\binom{n}{2}}{N}$. There is however an other slightly different point of view, which has some advantages. *We may consider the formation of a random graph as a stochastic process* defined as follows: At time $t=1$ we choose one out of the $\binom{n}{2}$ possible edges connecting the points $V_1, V_2, \cdots, V_n$, each of these edges having the same probability to be chosen; let this edge be denoted by e_1. At time $t=2$ we choose one of the possible $\binom{n}{2}-1$ edges, different from e_1, all these being equiprobable. Continuing this process at time $t=k+1$ we choose one of the $\binom{n}{2}-k$ possible edges different from the edges $e_1, e_2, \cdots, e_k$ already chosen, each of the remaining edges being equiprobable, i.e. having the probability $1/\{\binom{n}{2}-k\}$. We denote by $\Gamma_{n,N}$ the graph consisting of the vertices $V_1, V_2, \cdots, V_n$ and the edges $e_1, e_2, \cdots, e_N$.

Paul Erdos and A. Renyi, On the evolution of random graphsMagyar Tud. Akad. Mat. Kut. Int. Kozl. 5 (1960) 17-61.

The Erdos-Renyi Random Graph Model (2)

- Density
  - ✓ G(N,p): Density is on average p, as on average we have p*(N-1)*N/2 links
  - ✓ G(N,m): Density is equal to 2*m/[(N-1)*N] in all instances

- Degree distribution
  - ✓ Probability that each node has degree k is a Binomial distribution (i.e. probability of getting k successes out of i.d.d. N-1 Bernoulli trials)

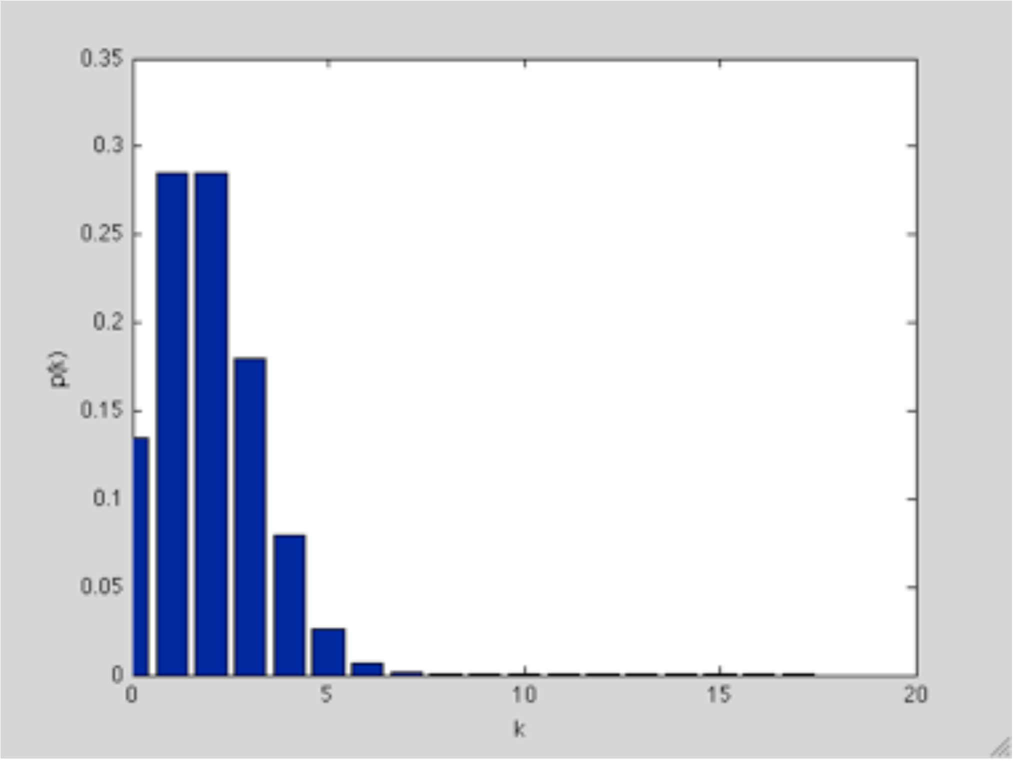
$$p(k) = \binom{N-1}{k} p^k (1-p)^{N-k-1}$$

  - ✓ For large N, this is also the fraction of nodes with degree k (due to the LLN)
  - ✓ If $N \to \infty$ and $N \cdot p(N) \to c$, we know that the Binomial distribution tends to a Poisson(c). Thus

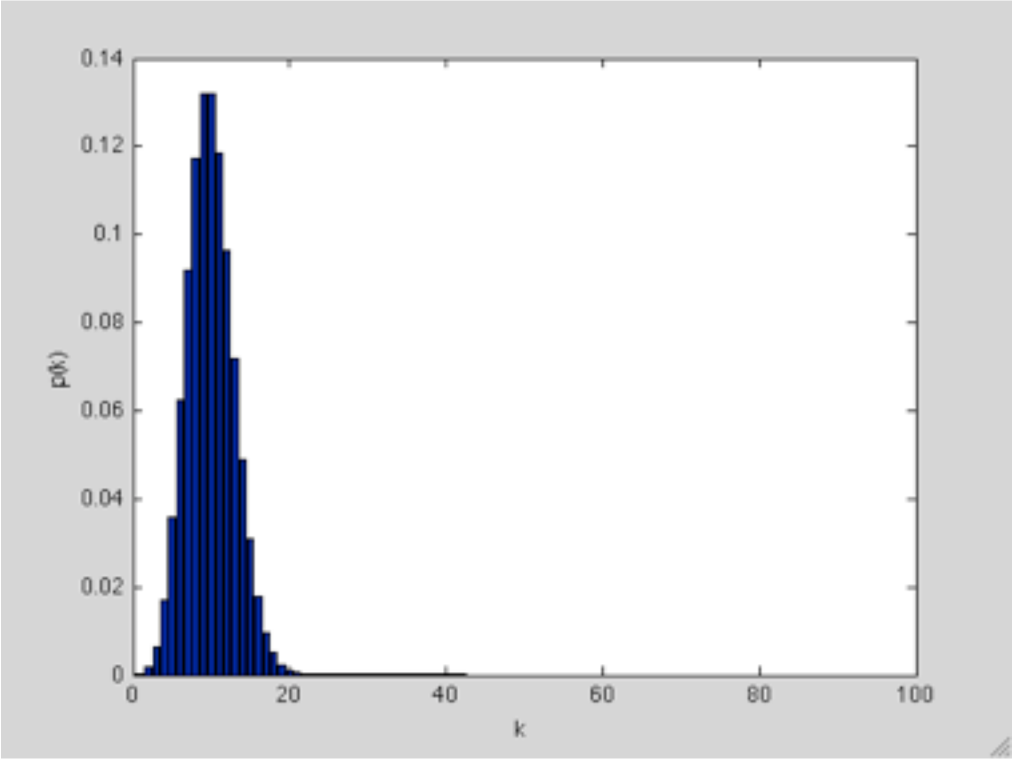
$$p(k) = \frac{e^{-c} c^k}{k!}$$

  - ✓ That's why ER random graphs are also called Poisson networks

The Erdos-Renyi Random Graph Model (2)



Degree Distribution for N=20, p=0.1



Degree Distribution for N=100, p=0.1

The Erdos-Renyi Random Graph Model (3)

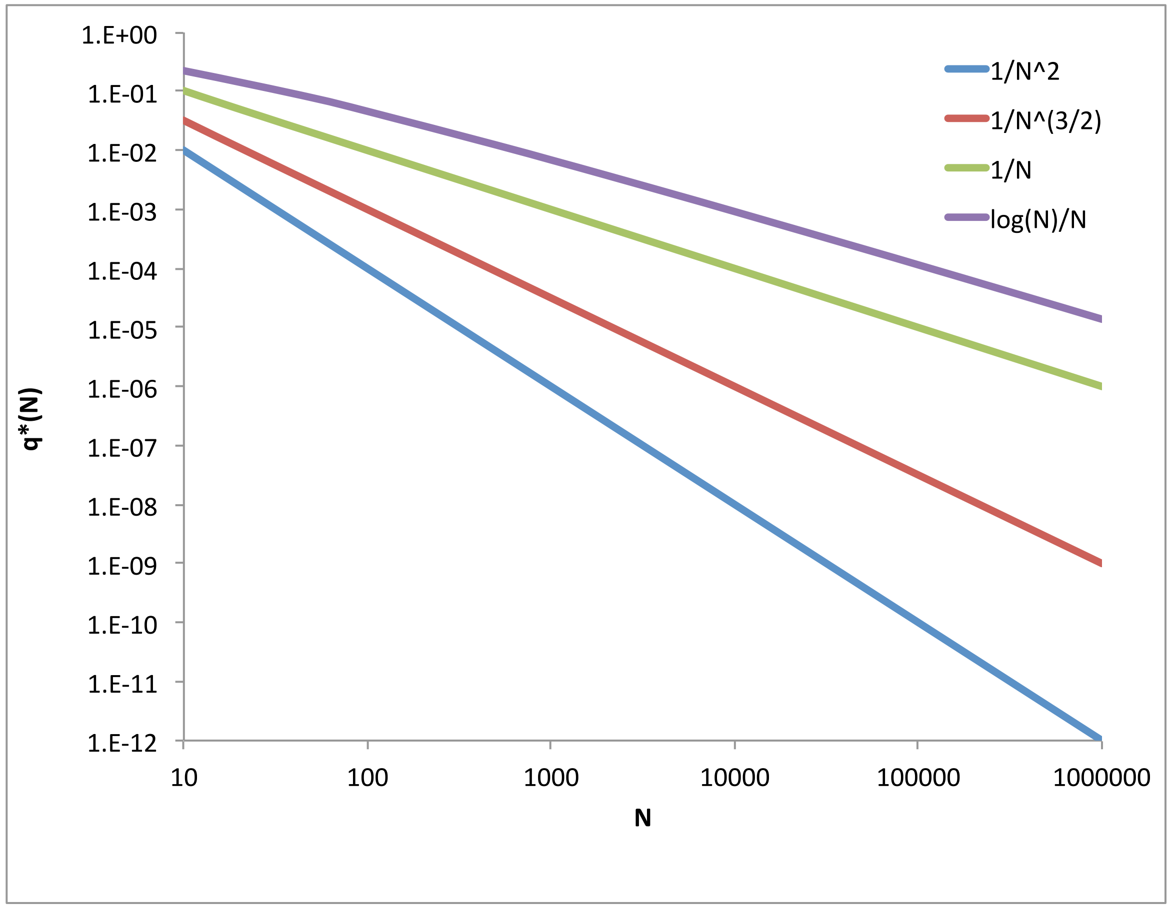
- Threshold functions
  - ✓ Let p=p(N) and let us focus on a given "property" of the ER graph G(N,p(N))
  - ✓ Examples of "properties": the network has no isolated nodes, the network displays cycles of order 3, the network is connected, etc.
  - ✓ Given a property P, a threshold function is a function q*(N) such that as N→∞

$$\frac{p(N)}{q^*(N)} \to \infty \Rightarrow \text{P holds with prob 1}$$

$$\frac{p(N)}{q^*(N)} \to 0 \Rightarrow \text{P holds with prob 0}$$

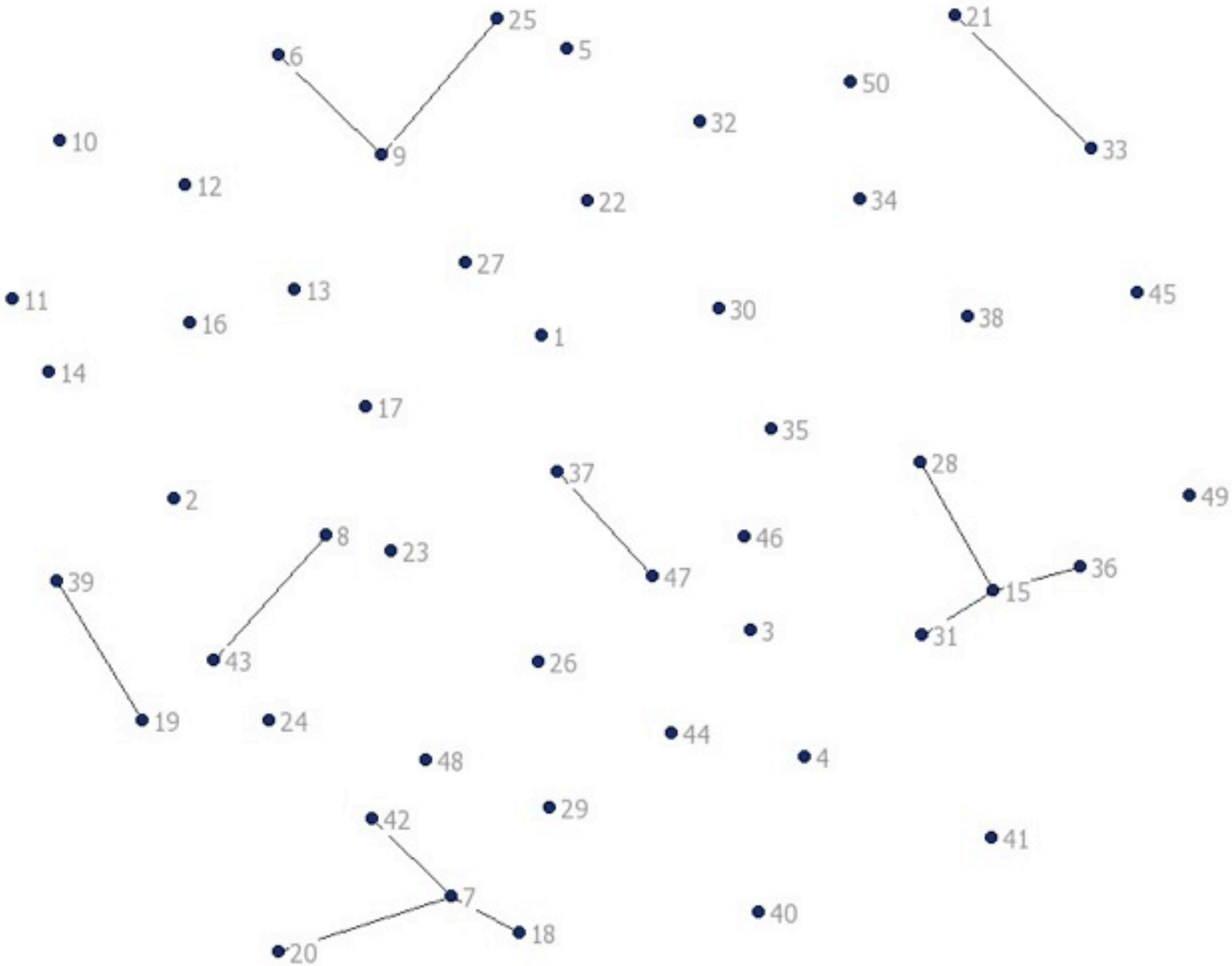
- Meaning of thresholds
  - ✓ As the number of nodes grows, and the probability p=p(N) of forming a link decreases with N, then if p(N) decreases slower than q*(N), we are going to observe the property P almost surely

The Erdos-Renyi Random Graph Model (4)



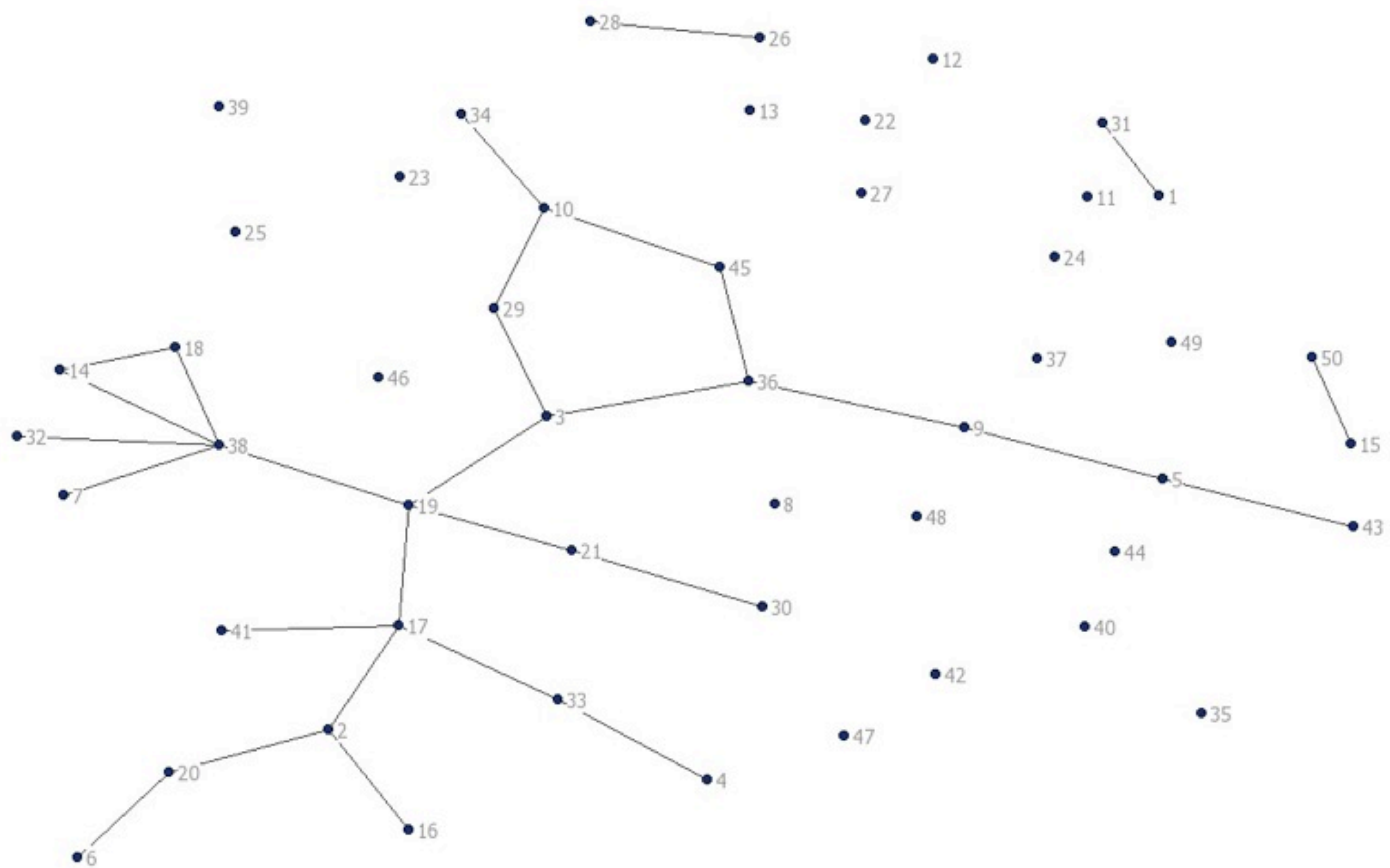
- $q^*(N)=1/N^2$
  - ✓ The network has at least one link
- $q^*(N)=1/N^{3/2}$
  - ✓ The network has at least a component with at least 3 nodes
- $q^*(N)=1/N$
  - ✓ The network has at least one cycle and contains a giant component (i.e. a unique largest component whose size grows as N and contains a non-trivial fraction of all nodes)
- $q^*(N)=\log(N)/N$
  - ✓ The network is connected

The Erdos-Renyi Random Graph Model (4)



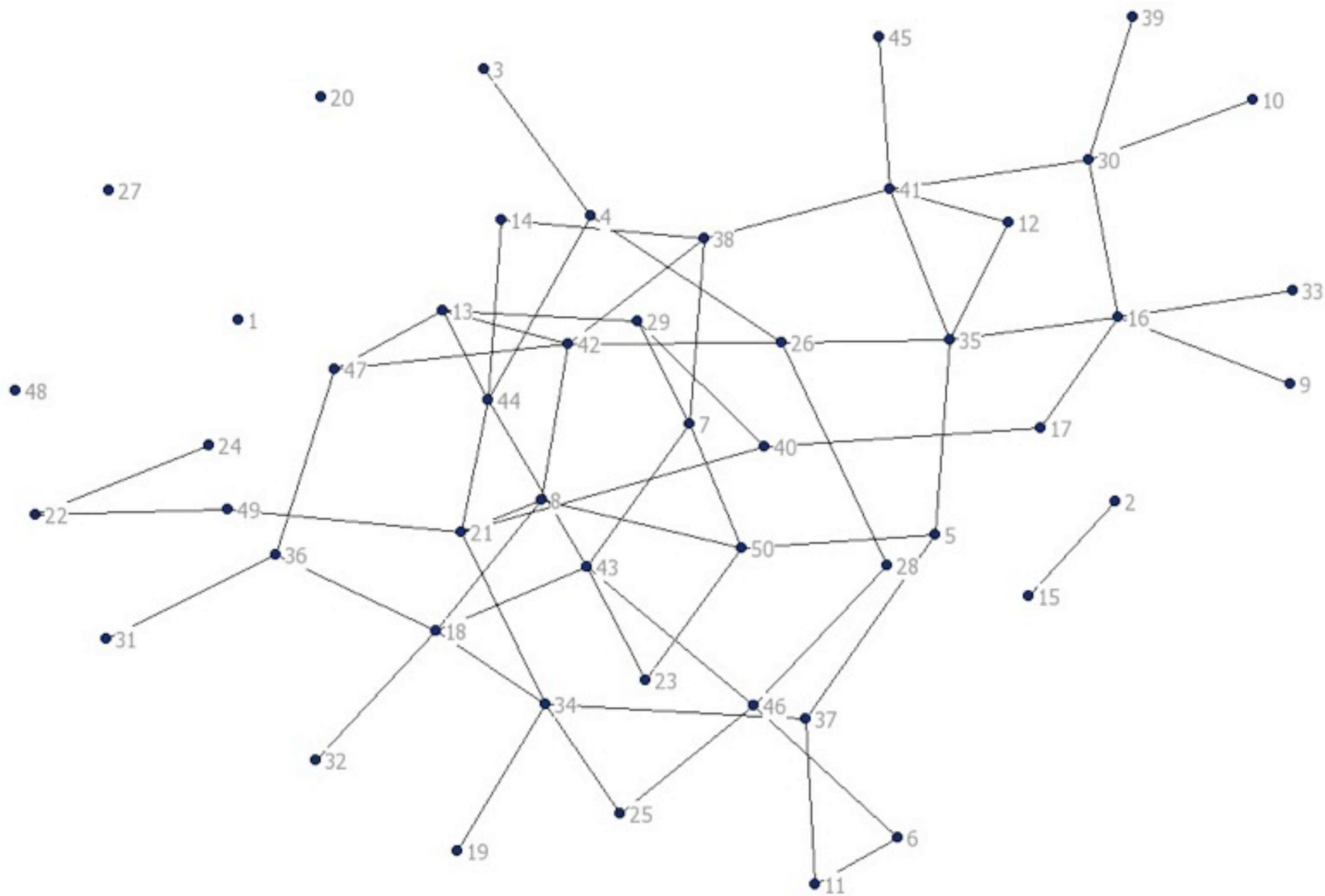
N=50, p=0.01, $N^{-3/2}<p<N^{-1}$: Components with more than 2 nodes emerge

The Erdos-Renyi Random Graph Model (4)



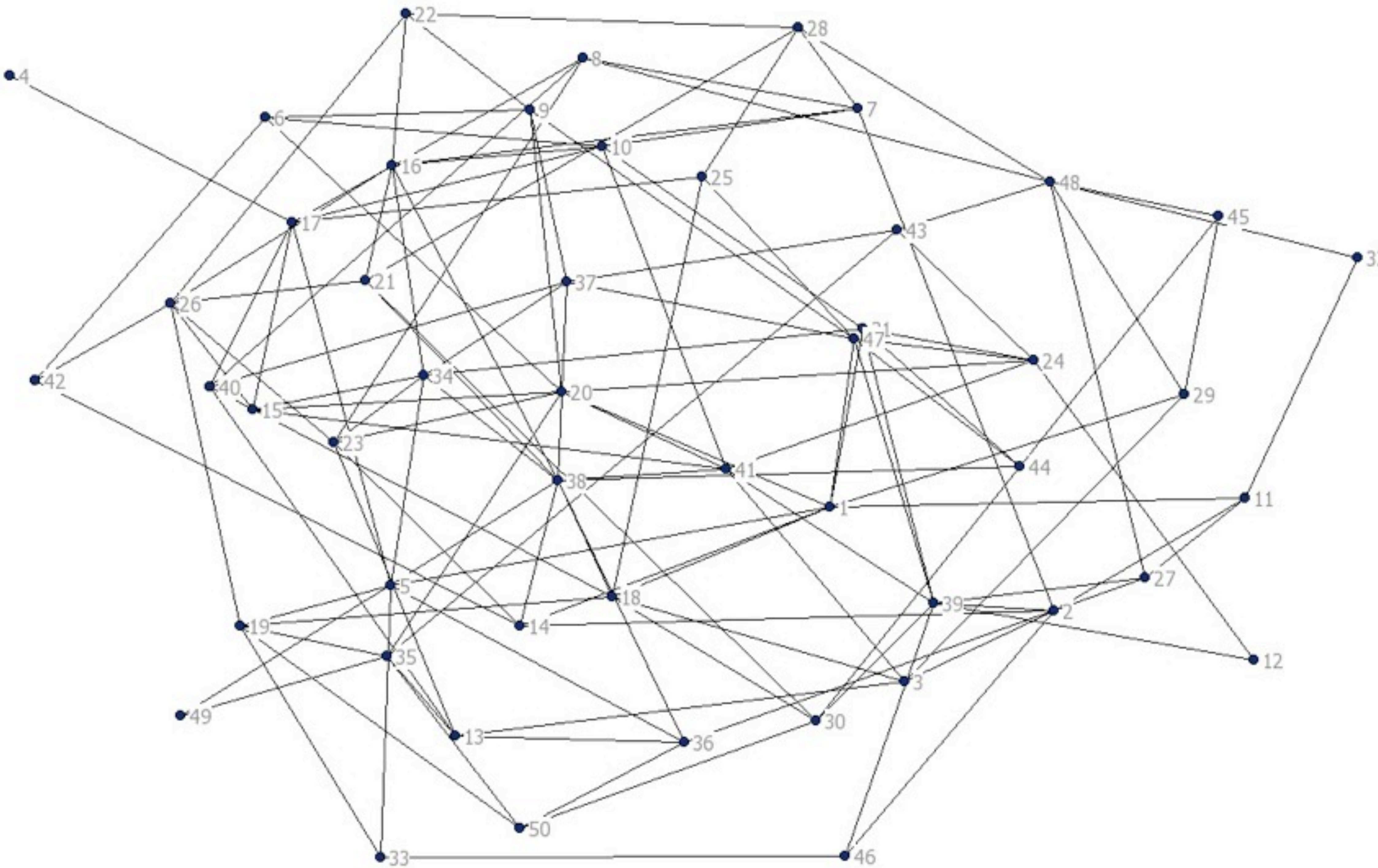
N=50, p=0.03, $N^{-1}<p<\log(N)/N$: Cycles emerge

The Erdos-Renyi Random Graph Model (4)



N=50, p=0.05, $N^{-1}<p<\log(N)/N$: The giant component emerges

The Erdos-Renyi Random Graph Model (4)



N=50, p=0.1, p>log(N)/N: The network becomes connected

The Erdos-Renyi Random Graph Model (6)

- Diameter Estimation: Are ER graphs small worlds?
  - ✓ Suppose p(N) as $N \rightarrow \infty$ is such that:

$$\frac{Np(N)}{\log(N)} \rightarrow \infty$$

  - ✓ This means the the graph is connected a.s.; define by $z=p(N-1)$ the average degree of the network

- What is the maximum path length between any two nodes?
  - ✓ Start at any given node i and estimate the average cardinality of the set of its direct neighbors V_1: it is easy to see that $|V_1|=z$
  - ✓ Iterate the argument and compute the magnitude of the neighbors of neighbors of node i, etc.: we approximately have $|V_k|=z^k$
  - ✓ But $|V_k| \leq N$, thus $z^k \leq N$ and $k \leq \log(N)/\log(z)$. Therefore, k=length of path from i to any other neighbor cannot exceed a function that grows as log(N)
  - ✓ Note that: $\log(1000) \cong 7$, $\log(10^6) \cong 14$: very small numbers!

The Erdos-Renyi Random Graph Model (7)

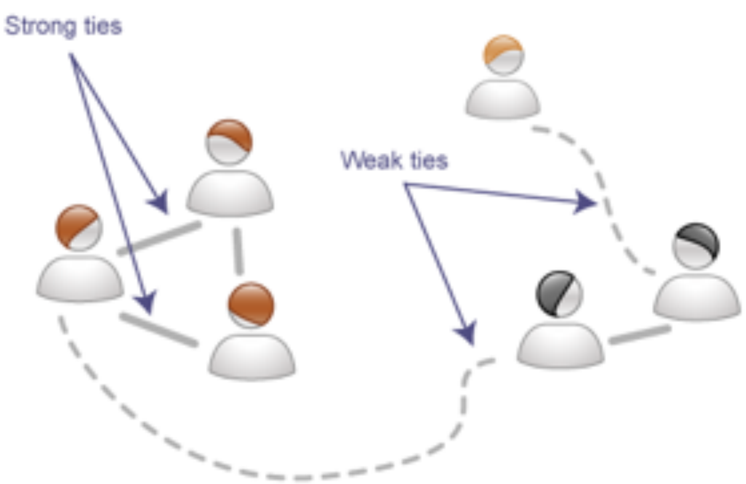
- ER Graphs are Small Worlds
  - ✓ Diameter growth as log(N) as in many real-world graphs

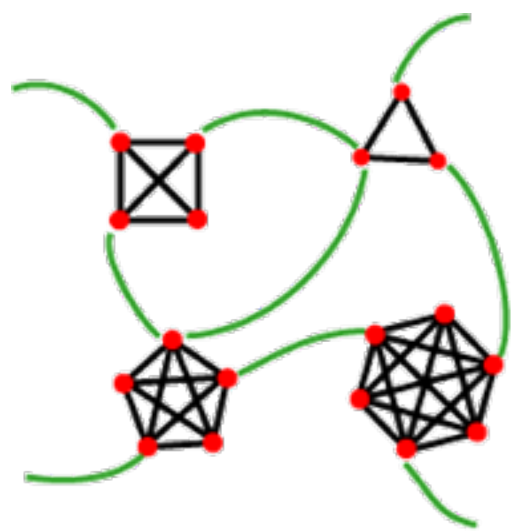
- The ER Random Graph Model: Discussion
  - ✓ Number of nodes is constant
  - ✓ Link formation is an i.i.d. process
  - ✓ All nodes have a "characteristic" scale: average number of partners, deviations from it are exponentially rare
  - ✓ If I take any two partners of a given node, the probability that they are friends is equal to p for all agents, clustering is very small (=density)

- The ER Random Graph Model and Empirical Evidence
  - ✓ Link formation is not an i.i.d. process
  - ✓ Networks grow over time (adding nodes and links)
  - ✓ Degree distribution may not be Poisson
  - ✓ Clustering may be very high

Stylized fact #2: It is a clustered world!

- Granovetter (1971): The strength of weak ties
  - ✓ How do you find a job?
  - ✓ Agents are strongly connected to a small circle of "friends"
  - ✓ They occasionally hold "weak ties" to other people belonging to different social groups
  - ✓ Statistically, info on job openings are gathered more through weak than strong ties





- Empirical evidence on clustering rejects random-graph model
  - ✓ Ex: in co-authorship network, CC 10000 times larger than predicted (density=10E-5)
  - ✓ How can one explain TOGETHER low average path length (SF #1) and high clustering (SF #2) observed in real-world networks?

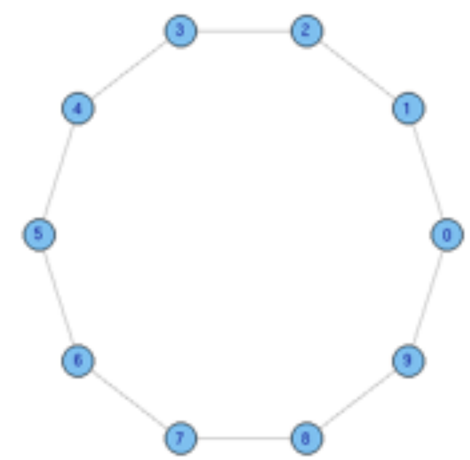
The Watts-Strogatz Model (1)

- The basic ideas behind the model
  - ✓ Goal: Building a random-graph model delivering small-world with high clustering coefficient
  - ✓ Two extremes as far as clustering is concerned
    1. ER Random Graphs: Clustering is equal to link probability (p), if p is small (to match real-world network densities, which are often sparse) so is clustering
    2. Regular Lattices: Clustering is very high, as lattices mimic geographical networks, where any two neighbors of mine are themselves neighbors with a high probability because they are close in geographical space
  - ✓ Two extremes as far as APL is concerned
    1. ER Random Graphs: APL is small (similar argument used for diameter)
    2. Regular Lattices: APL is high, as it takes a lot to move from one node to another distant one in the lattice (there are no shortcuts in geographical space)

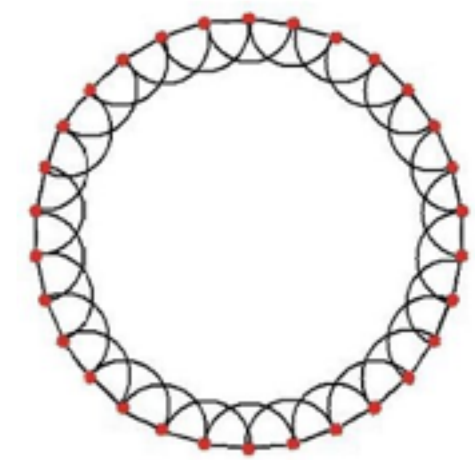
Example: Ring network

N=Number of nodes

z=2*r=node degree



r=1, z=2



r=2, z=4

$$D = \frac{N-1}{2r}$$

$$APL = \frac{1}{2} + \frac{N-1}{4r}$$

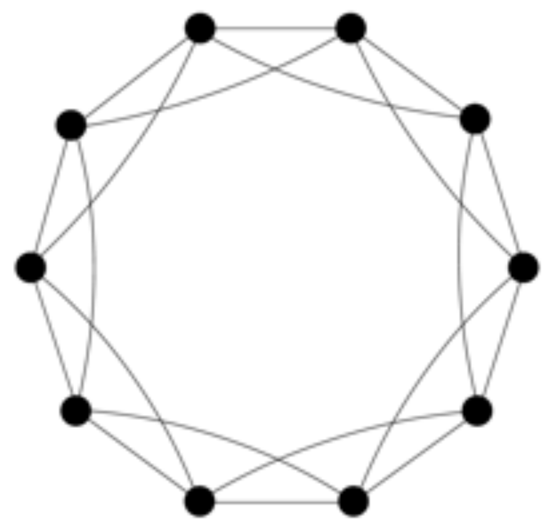
$$CC = \frac{3(r-1)}{4(r-\frac{1}{2})}$$

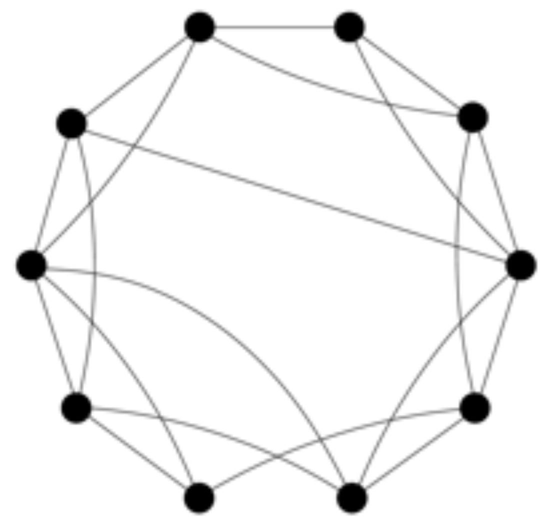
r>1

The Watts-Strogatz Model (2)

- Perturbing the ring structure
  - ✓ What if we add random links between distant locations and delete existing local links?
  - ✓ This may have four consequences
    1. N and L remain constant
    2. Node degree is perturbed in such a way not to introduce any bias: we should expect unimodal distributions
    3. Local link removal should destroy clustering, but if perturbations are not too strong clustering coefficients should stay relatively large
    4. Randomly adding links between distant nodes should dramatically lower APL by short-cuts (weak ties) between otherwise very clustered but disconnected islands
  - ✓ Is all that true?
    1. A more formal description of the model
    2. Analysis of model behavior

Perturbing a Ring network
N=Number of nodes
z=2*r=node degree

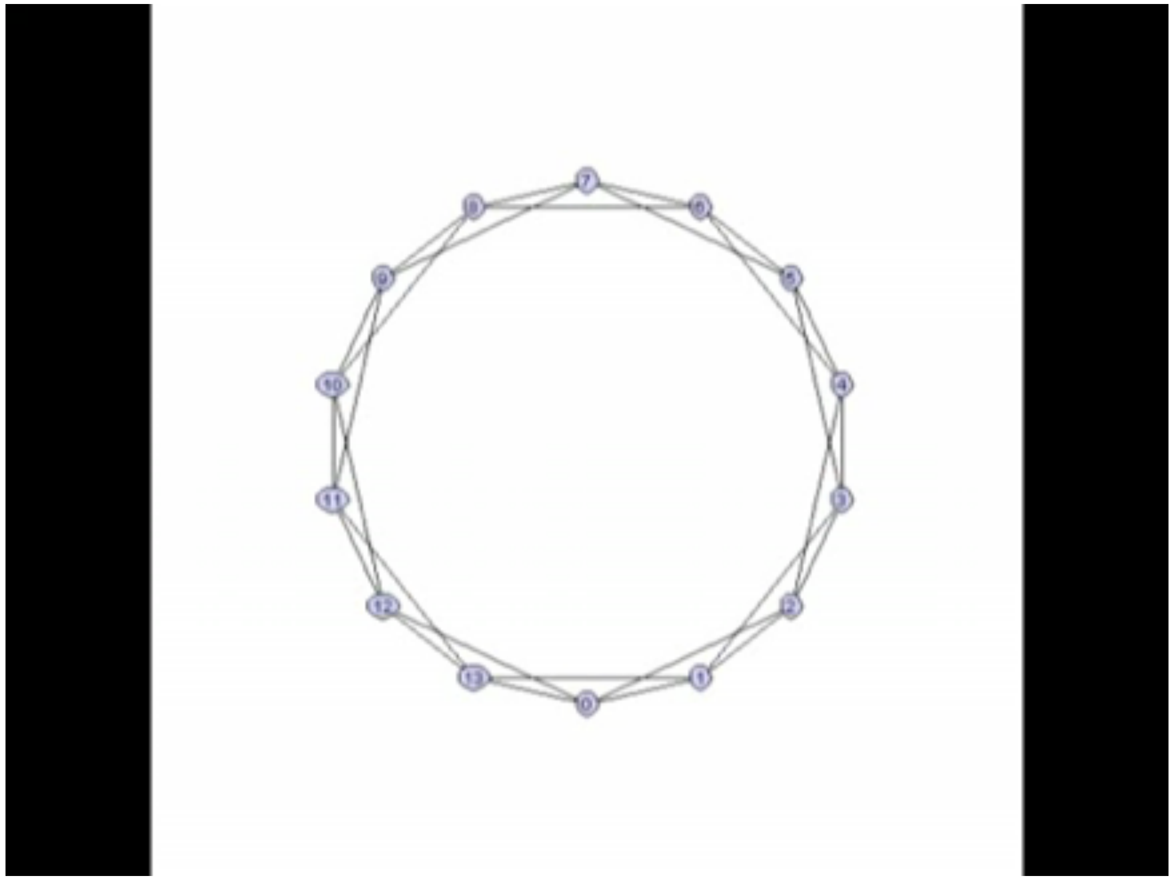




The Watts-Strogatz Model (3)

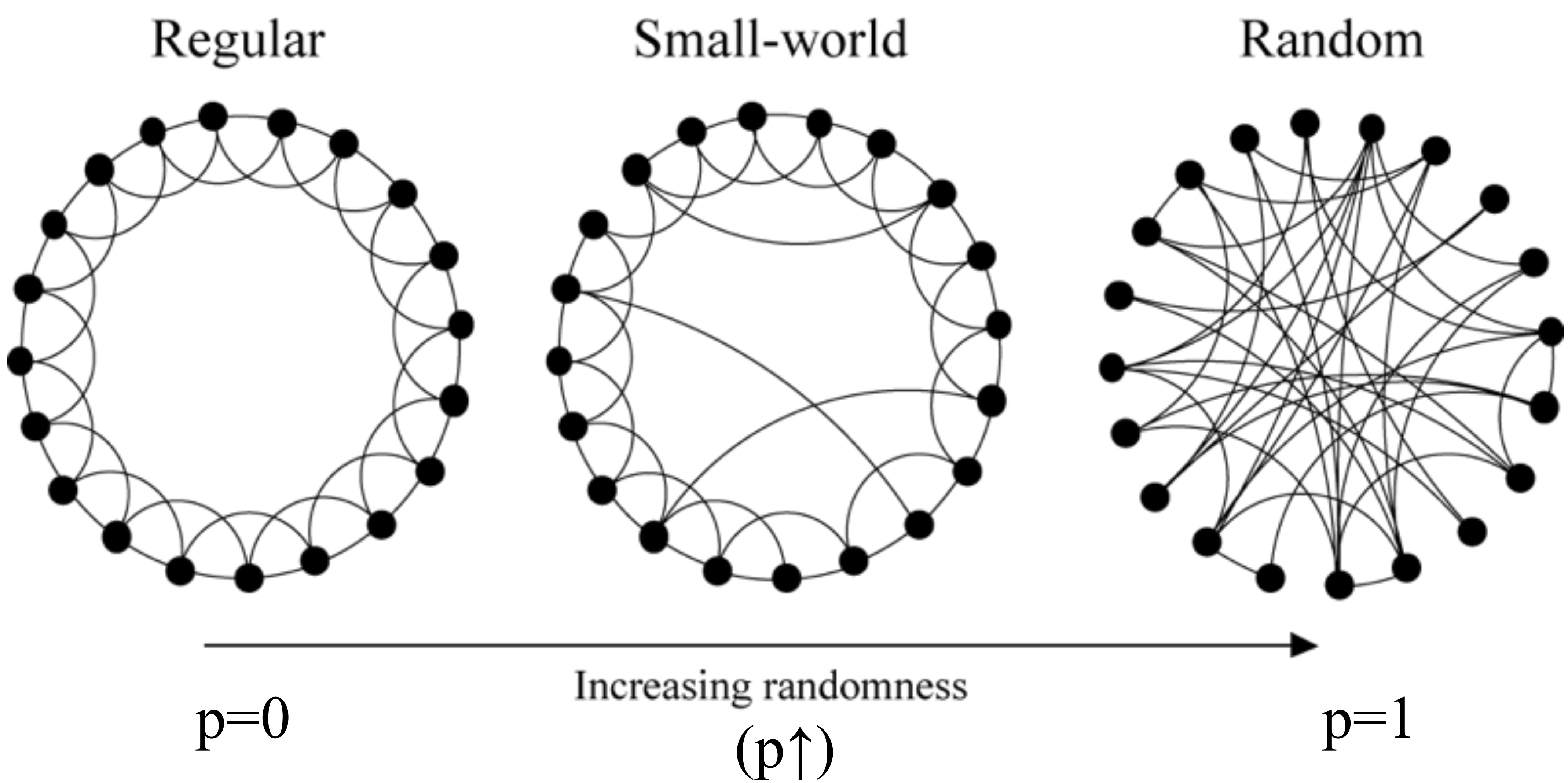
- The Model
  - ✓ Start from N nodes located on a ring
  - ✓ Each node is linked to r neighbors on the left and r neighbors on the right
  - ✓ Thus the network is regular and all nodes have degree k=2r
  - ✓ Define $0 \leq p \leq 1$ as the rewiring probability

- Dynamics
  1. Start from node 1 and proceed clockwise through all subsequent nodes
  2. For each node i, consider its links with nodes at distance 1
  3. For any link of node i in this set, with probability p delete the link and rewire node i with another one from the list of nodes that are not currently linked to i
  4. When we arrive at node N, start again with node 1 and consider all its links at distance 2 and repeat from 3
  5. After r laps (after having considered links with nodes at distance r), stop the process



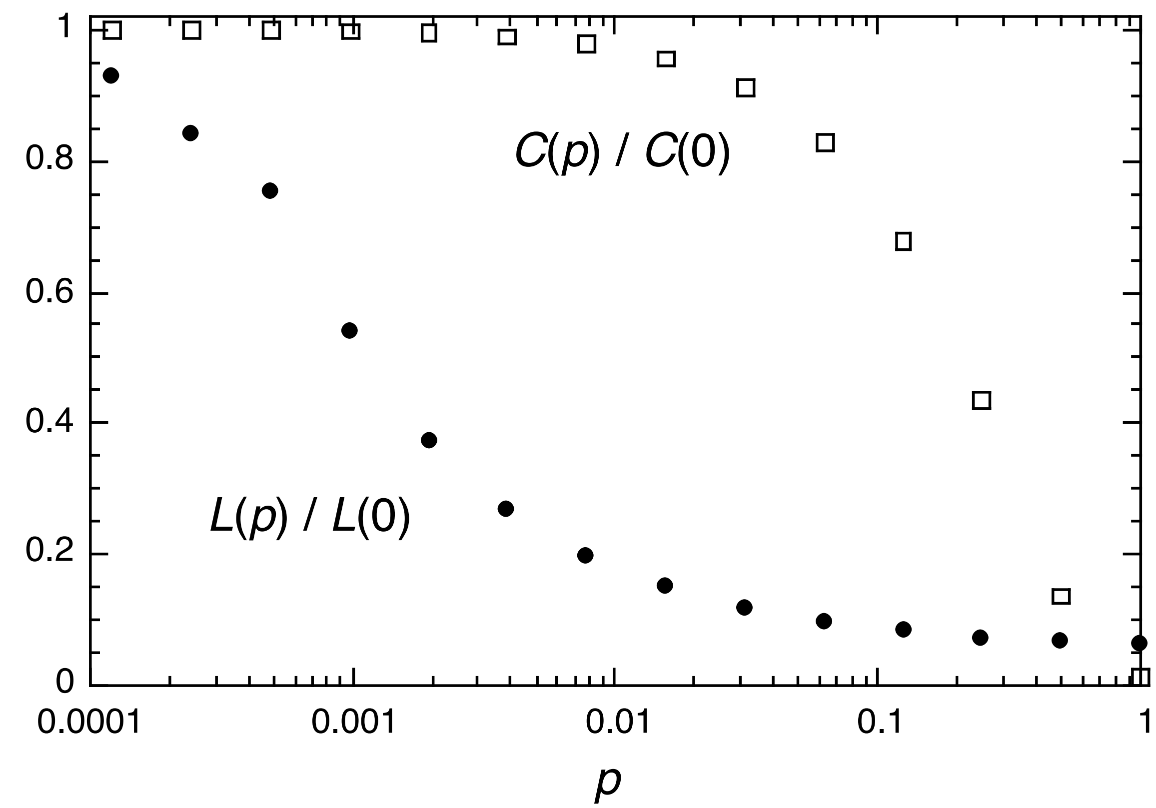
The Watts-Strogatz Model (4)

- The Model: Some observations
  - ✓ If p=0 then the resulting graph is the initial lattice (no rewiring takes place). Thus the degree distribution is degenerate (peaked on k=2r)
  - ✓ If p=1 then the model is equivalent to a ER random lattice, as all links are rewired a.s.: therefore the resulting graph is fully random and the degree distribution is Poissonian (for large N)
  - ✓ But what happens to CC and APL when we tune p from 0 to 1?



The Watts-Strogatz Model (5)

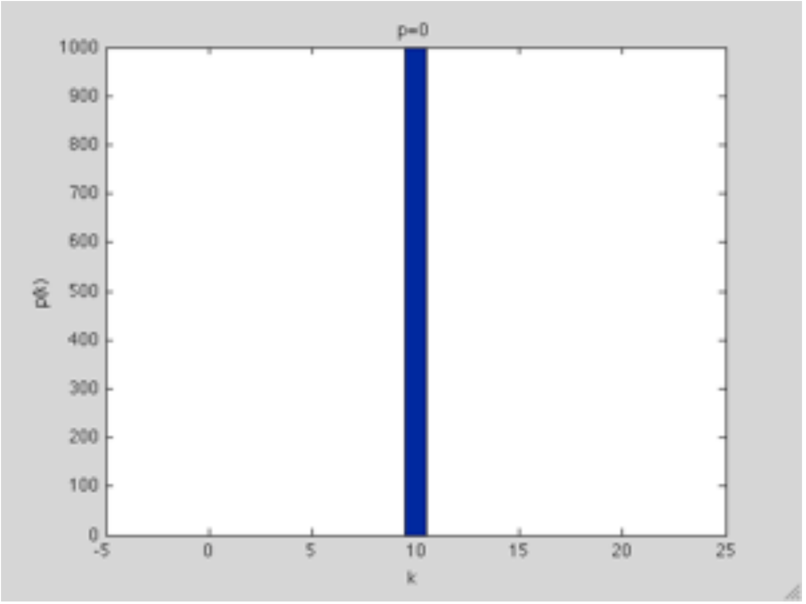
- If p=0 then we observe a high clustering and high APL; if p=1 a small clustering and APL
- As p increases, there is a sudden decrease in APL: this means that even a few rewirings introduce shortcuts that dramatically decrease the distance between any two nodes: the graph becomes a small world very soon
- Clustering keeps instead very high, as link removals do not alter too much the local structure. This means that the small-world property is not perceived as a local phenomenon
- There emerges a large intermediate region of rewiring probabilities where the resulting graph exhibits both small APL and large clustering, in line with what is observed in reality

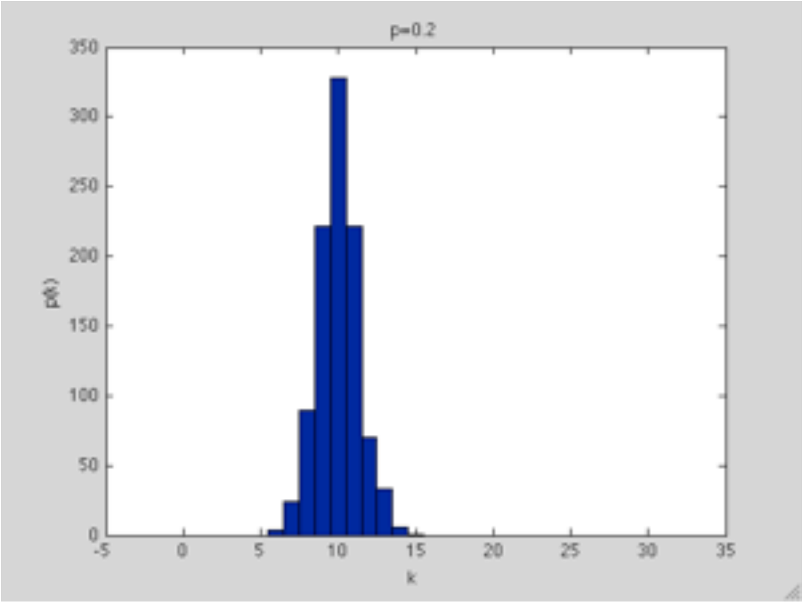


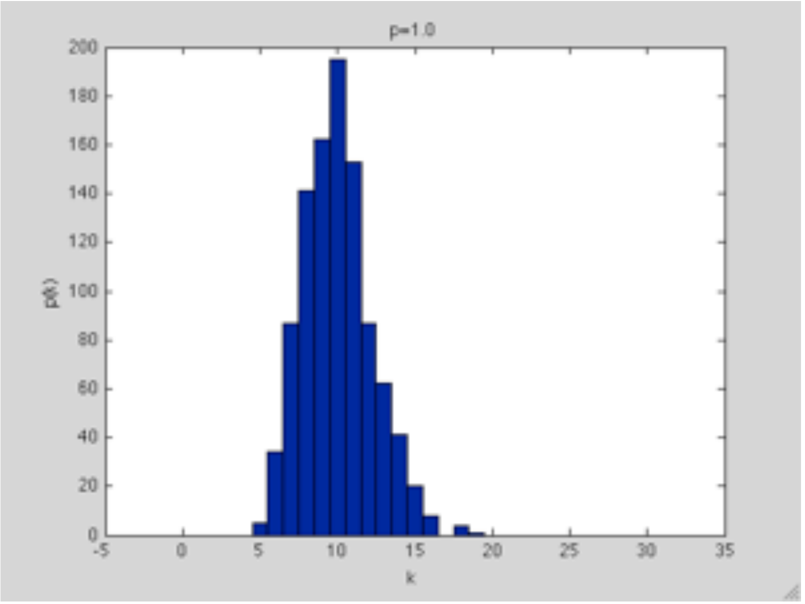
- N=1000, r=5 (N,r chosen in such a way not to have for any p a disconnected graph)
- C is the node-average of clustering coefficients, L is average node path-length
- Values of C and L for p=0, i.e. in the ring, that is C(0) and L(0), are use no normalize all other values
- Points in the graphs obtained averaging out over a large number of replications.
- Source: D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. Nature 393, 440-442 (1998).

The Watts-Strogatz Model (6)

- The model introduces about p·r·N non lattice links, but in order not to have disconnected graphs we need N>>2r>>log(N)>>1
- The degree distribution can be worked out analytically but simulations can help in understanding how it looks like as p increases







- Mean(k) is always close to initial one (p=0), i.e. 2*r
- Distribution tends to a Poissonian one as p grows

The Watts-Strogatz Model (7)

- Is this the end of our journey? The WS model explains
  - ✓ Low average path length (Milgram)
  - ✓ High clustering (Granovetter)
  - ✓ But: Degree distribution still has exponentially-decaying tails:
  - ✓ Characteristic scale: almost all nodes have same number of partners, because links are rewired at random!
  - ✓ Deviations are very rare events: almost impossible to find a 50 or 100mt tall guy wandering around in our streets…

- Are all real-world degree distributions like that?
  - ✓ Not at all… see previous lectures! Barabasi et al. (1999)
  - ✓ WWW network: many low-degree networks coexisting with not-so-rare hubs holding many links
  - ✓ Comparison: national highway network (most nodes have the same number of links) vs. air-traffic network (a few hubs coexist with many small airports)

Stylized fact #3: It is a scale-free world!

- Barabasi et al. (1999): Scale-free networks
  - ✓ Many real-world networks display a power-law degree distribution

$$p(k) = c_1 k^{-\gamma-1} \Rightarrow \log[p(k)] = c_2 - (\gamma+1)\log(k)$$

  - ✓ Scale-free: there is no characteristic scale (mean is meaningless)
  - ✓ Hierarchical structure and importance of hubs in taking the network together: hubs guarantee that average path length is small

- Scale or degree exponent (γ)
  - ✓ Measures the heaviness of the tails (likelihood of hubs)
  - ✓ It is a Pareto distribution (80-20 rule): a downward sloping line in a log-log paper

Stylized fact #3: It is a scale-free world!

- Examples: Widespread evidence
  - ✓ In-coming or out-coming links in web pages; actor network in Hollywood; airline traffic; co-authorship and Erdos number; citations
  - ✓ Molecules interacting within a cell, and many others

- Power laws are ubiquitous in natural and social systems
  - ✓ Bell-shaped curves emerge in disordered random systems
  - ✓ Power laws are the signature of complex behavior and self-organization
  - ✓ Most economic variables typically follow (quasi) power laws: firms size, wealth and income, size of economic fluctuations, etc.

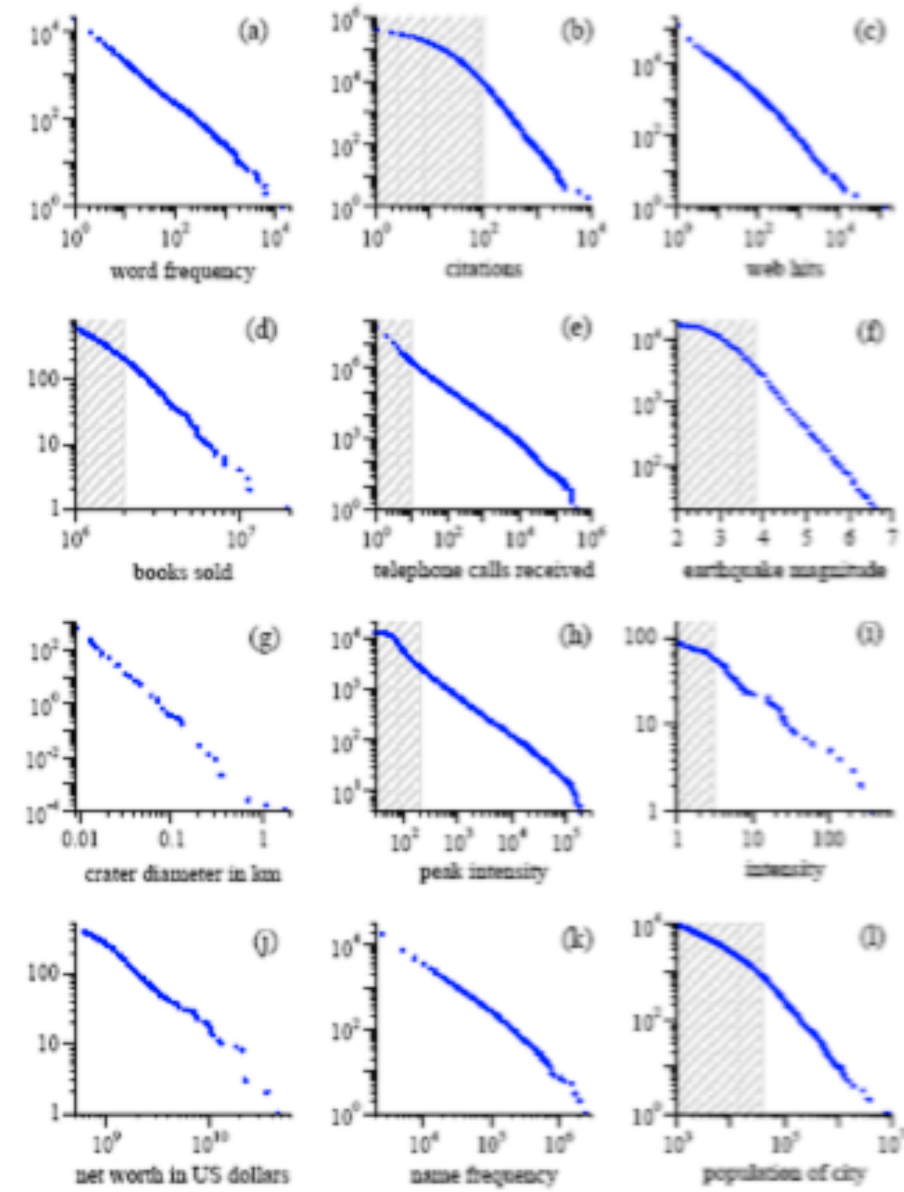


FIG. 4 Cumulative distributions or "rank/frequency plots" of twelve quantities reputed to follow power laws. The distributions were computed as described in Appendix A. Data in the shaded regions were excluded from the calculations of the exponents in Table I. Source references for the data are given in the text. (a) Numbers of occurrences of words in the novel *Moby Dick* by Hermann Melville. (b) Numbers of citations to scientific papers published in 1981, from time of publication until June 1997. (c) Numbers of hits on web sites by 60 000 users of the America Online Internet service for the day of 1 December 1997. (d) Numbers of copies of bestselling books sold in the US between 1895 and 1965. (e) Number of calls received by AT&T telephone customers in the US for a single day. (f) Magnitude of earthquakes in California between January 1910 and May 1992. Magnitude is proportional to the logarithm of the maximum amplitude of the earthquake, and hence the distribution obeys a power law even though the horizontal axis is linear. (g) Diameter of craters on the moon. Vertical axis is measured per square kilometre. (h) Peak gamma-ray intensity of solar flares in counts per second, measured from Earth orbit between February 1980 and November 1989. (i) Intensity of wars from 1816 to 1980, measured as battle deaths per 10 000 of the population of the participating countries. (j) Aggregate net worth in dollars of the richest individuals in the US in October 2003. (k) Frequency of occurrence of family names in the US in the year 1990. (l) Populations of US cities in the year 2000.

The Barabasi-Albert Model (1)

- Goal: A simple model of network evolution that
  - ✓ Reproduces Watts-Strogatz features
  - ✓ Allows for equilibrium power-law degree distributions
  - ✓ See Barabasi and Albert (1999), Science, 286.

- Two essential features
  - ✓ Growth: The network grows over time through the successive arrival of new nodes that upon entry link to some preexisting nodes
  - ✓ Preferential attachment: New nodes stochastically choose upon entry existing ones to form links with a bias in favor of highly-connected ones

- Consequences
  - ✓ Growth: Size of nodes is not fixed as in ER or WS models. Node degree may attain different magnitudes as the network continues to grow
  - ✓ Preferential attachment: Rich-get-richer process where nodes become more connected and this in turn induces higher probability to get even more connections

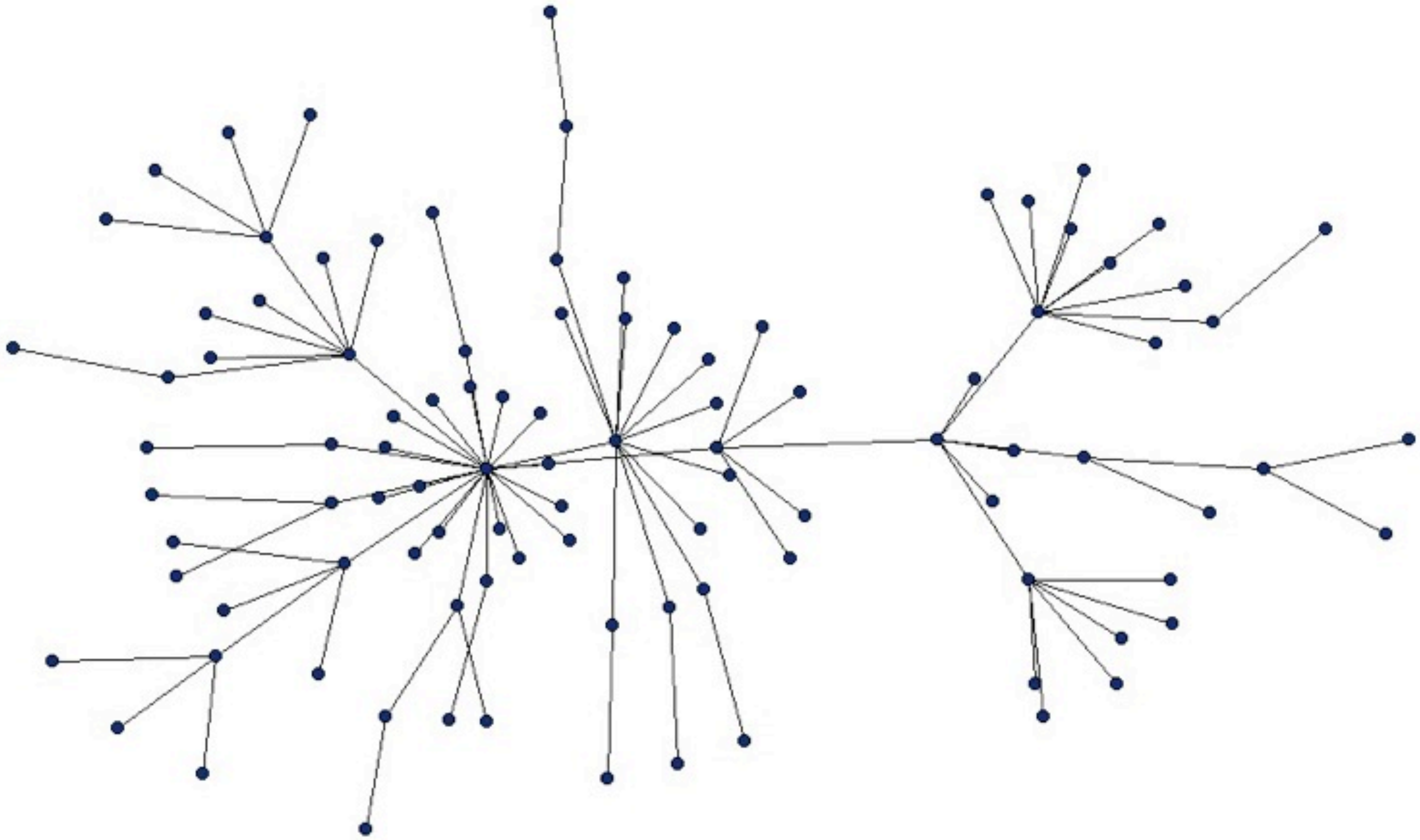
The Barabasi-Albert Model (2)

- The Model
  - ✓ Start at t=0 with N_0 nodes connected in a complete undirected graph
  - ✓ At any t=1,2,... add a new node
  - ✓ The new node chooses m distinct nodes among the pre-existing ones, each with probability proportional to its current degree and creates a link with each of them (m new links)
- More formally
  - ✓ At t=0 there are N_0 nodes and $L_0 = N_0(N_0-1)/2$ links
  - ✓ At time t>0, before entry, there will be $N_t = N_0+t$ nodes and $L_t = L_0+mt$ links; let $k_{i,t}$ be the degree of node i at time t
  - ✓ The entrant node will form m new links; the probability that a new link is formed with a pre-existing node i is

$$p_{i,t} = \frac{k_{i,t}}{\sum_{j=1}^{N_t} k_{j,t}} = \frac{k_{i,t}}{2 * L_t}$$

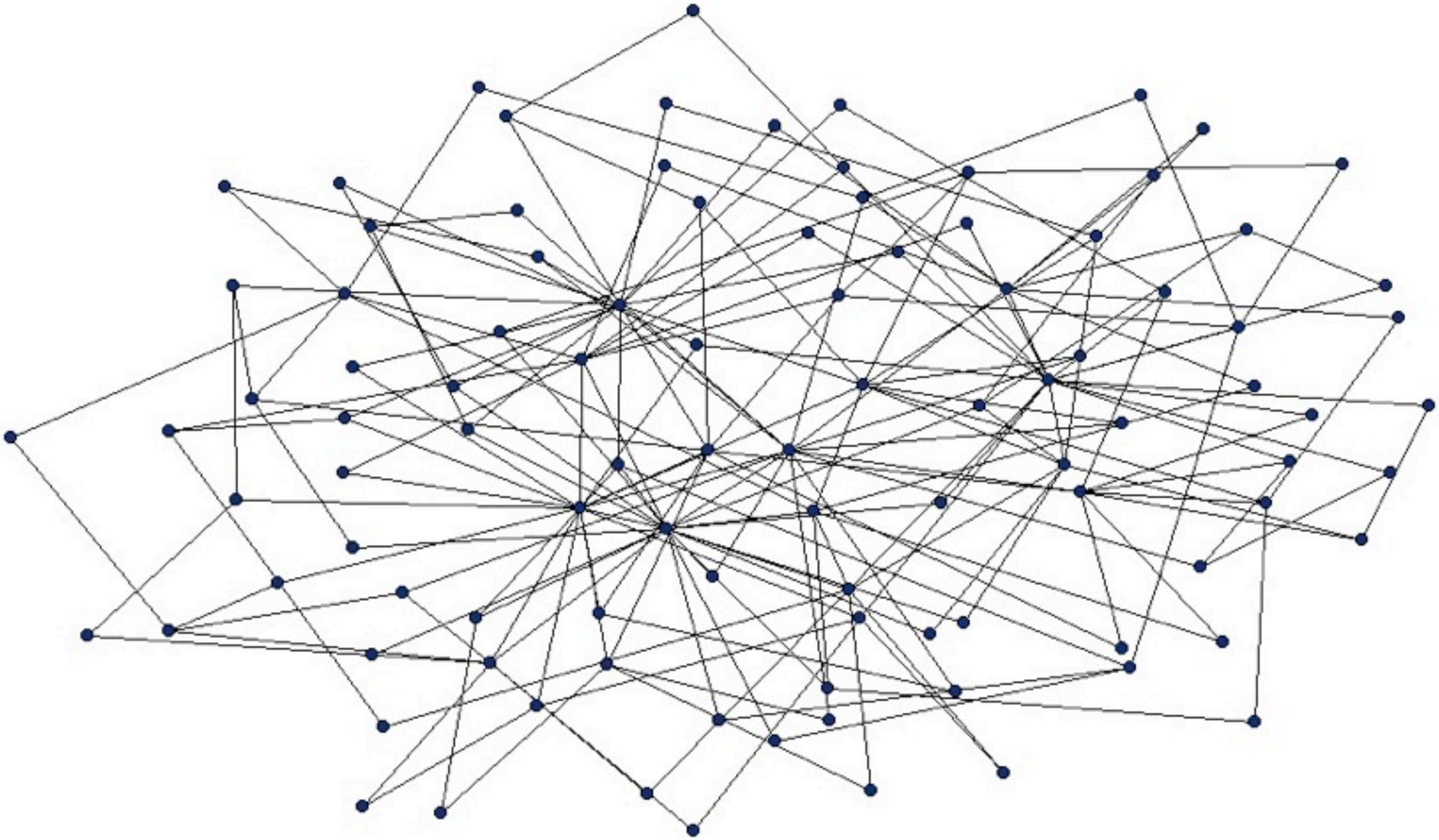
  - ✓ Implementation: We consider m steps; at the beginning of each step s, probabilities are computed on the remaining sample of N-s+1 nodes; after each step the node just chosen is removed from the list of available nodes. NB: Probabilities depend on the order of sampling (it is not a weighted random sampling w/o replacement!)

The Barabasi-Albert Model (2)



Simulating the BA model. N=100, m=1: Even when N is small, hubs start to emerge.
The resulting graph is a tree as m=1

The Barabasi-Albert Model (2)



Simulating the BA model. N=100, m=2.
Note the richer structure and the presence of loops (clustering is not zero)

The Barabasi-Albert Model (3)

- Simulating the evolution of the BA model
  - ✓ An example for m=1
  - ✓ http://oldweb.ct.infn.it/cactus/applets/Preferential%20Attachment.html

- The model reproduces power-law degree distributions. Why?
  - ✓ Growth (G) and preferential attachment (PA) are both necessary
  - ✓ Assume G without preferential attachment (new links are put at random). It may be proven (Vega-Redondo, 2007, p. 67) that the limit degree distribution is geometric, i.e. $p(k)=2^k$, i.e. a skewed but narrow distribution with a characteristic scale exhibiting a sharp decay for high degrees (hubs are low-probability events as in Poisson networks)
  - ✓ Assume PA but no G (at each t one node is picked at random to establish m new links with the other N-1 nodes according to preferential attachment. Clearly, in the long run a complete network arises (no multi links)
  - ✓ When both G and PA are assumed, then power-law degree distributions do emerge thanks to a rich-get-richer process

The Barabasi-Albert Model (3)

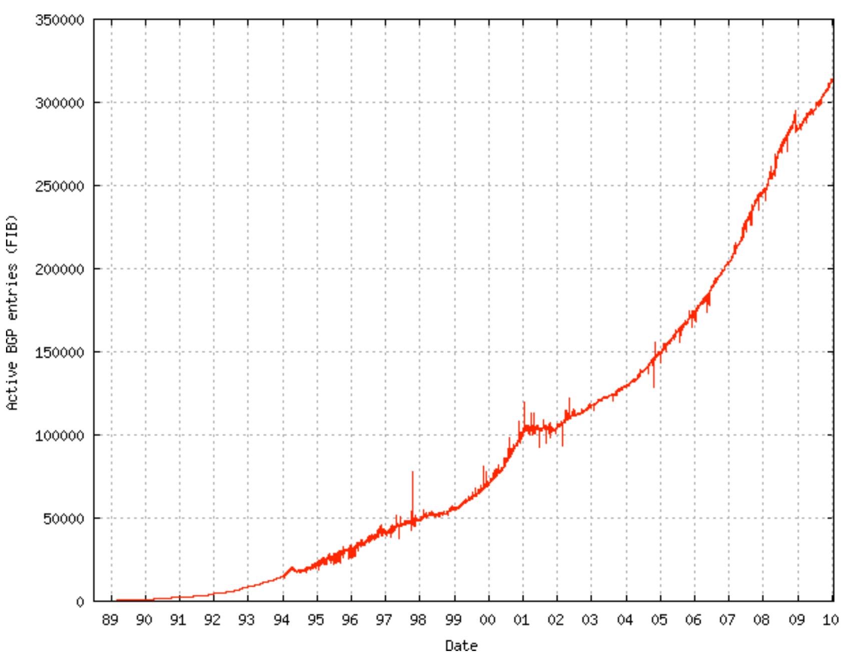
- Is the growth assumption justifiable?
  - ✓ Many real-world networks grow in size (order) over time (but others do not!)

Actor network

Number of movies in IMDB

Herr II, Bruce W., Ke, Weimao, Hardy, Elisha, and Börner, Katy. (2007) Movies and Actors: Mapping the Internet Movie Database. In Conference Proceedings of 11th Annual Information Visualization International Conference (IV 2007), Zurich, Switzerland, July 4-6, pp. 465-469.

Internet



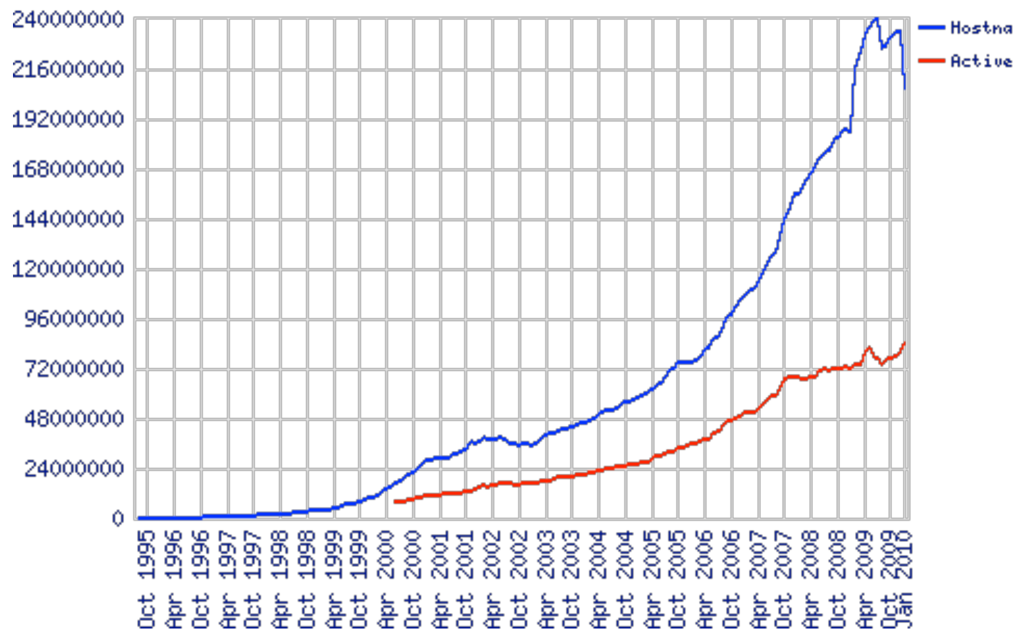
Growth of the Internet routing table

http://www.trainsignaltraining.com/ccna-ipv6

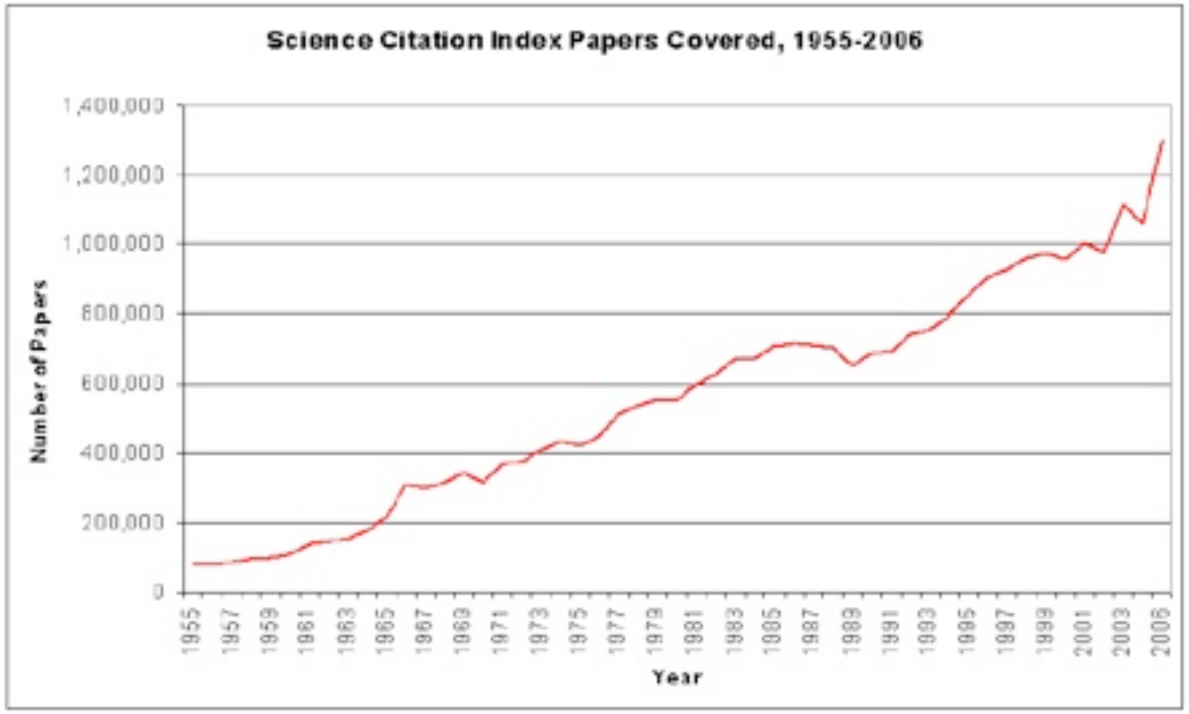
The Barabasi-Albert Model (3)

- Is the growth assumption justifiable?
  - ✓ Many real-world networks grow in size (order) over time (but others do not!)

WWW



Scientific Publications



http://website101.com/define-ecommerce-web-terms-definitions/

http://www.kk.org/thetechnium/archives/2008/10/the_expansion_o.php

Properties of BA Networks (1)

- The degree distribution converges, as $t \to \infty$ to a power-law with exponent 3, i.e. $P(K<k)=A*k^{-3}$
- Informal proof: Consider k as a continuous variable and note that new vertices enter at a constant rate; thus the variation of k with respect to t for any given new node is equal to the constant change in connectivity in one time step (m) times the probability of forming a link (degree/sum of all degrees); suppose for simplicity $m_0=0$

$$\frac{\partial k}{\partial t} = m\frac{k}{2mt}$$

- This is a differential equation. We need to solve it for k(t). The general form is dk/dt=f(t)h(k). The solution is:

$$\int h^{-1}(k)\,\mathrm{d}k = \int f(t)\,\mathrm{d}t$$

Properties of BA Networks (1)

- Replacing h(k)=k and f(t)=1/2t and solving one gets

$$\log(k) = \frac{1}{2}\log(t) + C_0 \Rightarrow k(t) = C_1 t^{\frac{1}{2}}$$

- To compute C_1, suppose that the node has entered at time t*; its initial degree was therefore m; thus replacing k(t*)=m one gets

$$C_1 = m(t^*)^{-\frac{1}{2}} \Rightarrow k(t) = m(\frac{t}{t^*})^{\frac{1}{2}}$$

- Let us then compute p(K<k). We get:

$$P(K < k) = P(m(\frac{t}{t^*})^{\frac{1}{2}} < k) = P(t^* > \frac{m^2 t}{k^2}) = 1 - P(t^* < \frac{m^2 t}{k^2})$$

Properties of BA Networks (1)

- Note that t* (entry time) is a random variable that is uniform in time as nodes enter at a constant rate. At time t there are N_0+t nodes. Therefore t* is U(0, N_0+t) and its CDF reads F(a)=P(t*<a)=a/(N_0+t)

$$P(K < k) = 1 - P(t^* < \frac{m^2 t}{k^2}) = 1 - F(\frac{m^2 t}{k^2}) = 1 - \frac{m^2 t}{k^2} \frac{1}{N_0 + t}$$

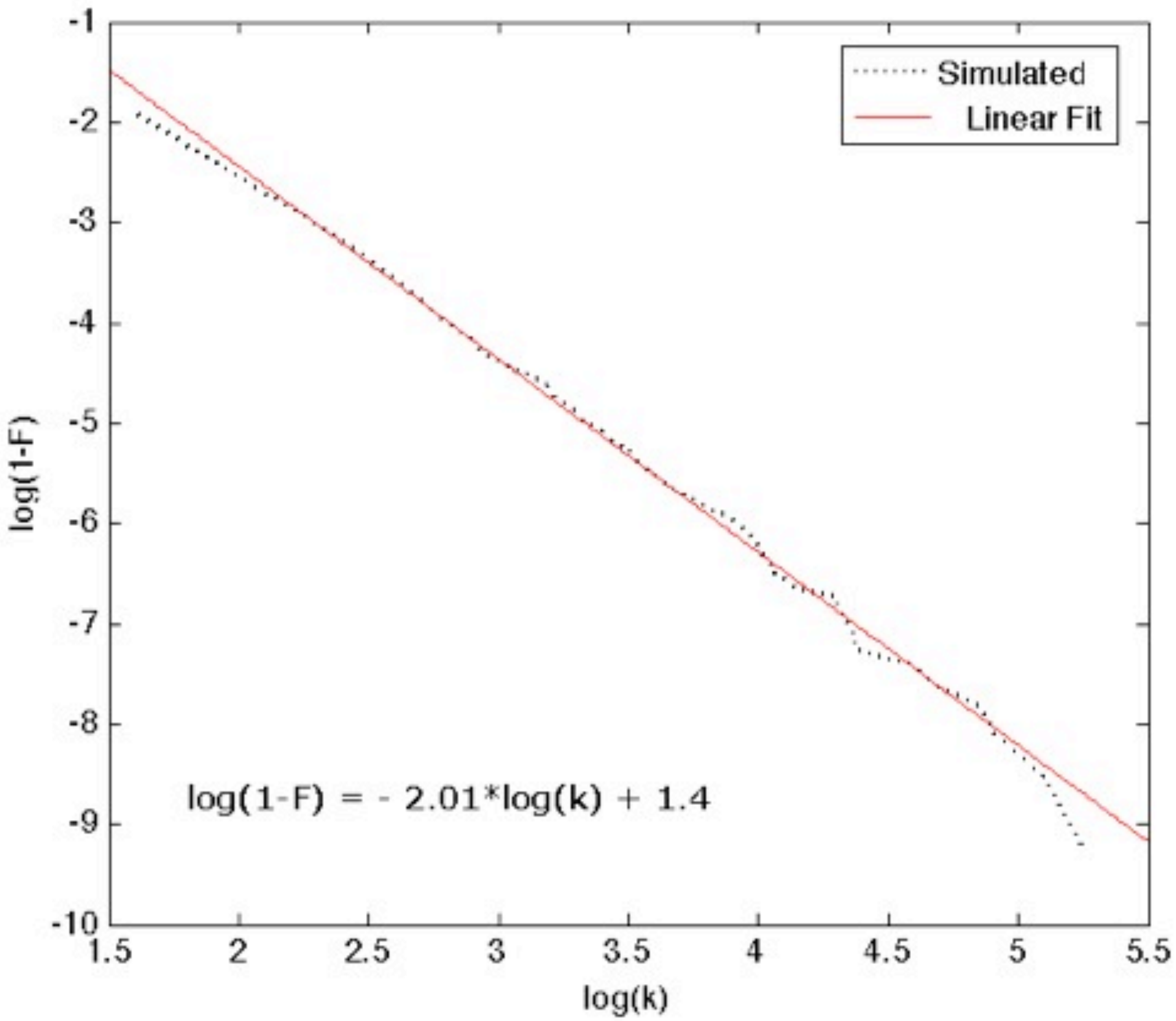
- Therefore the density is:

$$f_K(k) = \frac{\partial P(k)}{\partial k} = \frac{2m^2 t}{N_0 + t} \frac{1}{k^3}$$

That is: The BA model yields a power-law degree distribution with density exponent equal to 3, independently of m

This means that the limit degree distribution is a Pareto with parameter 2 (and therefore it does not admit a finite variance)

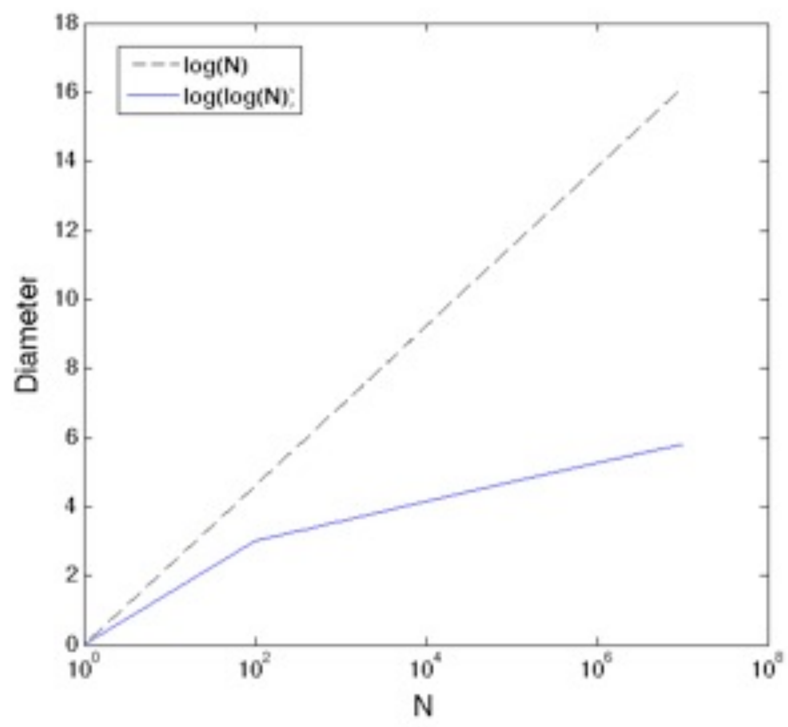
Pareto Degree Distribution in the BA Model



Simulating the BA model. N=10000, m=2.
Linear fit of the degree distribution using a rank-size plot. Not that the estimated slope is approximately equal to -2, i.e. the prediction of the BA model as N tends to infinity. Notice that the slope of the line log(1-F)=a+b*log(k) is not the MLE for alpha for the Pareto distribution.

Properties of BA Networks (2)

- Does the BA model generate limit networks with small diameters? In other words, are BA networks small worlds?
  - Yes, the diameter can be shown to scale as log(N)/log(log(N)), see Bollobas and Riordan (2003), even slower than in a random graph



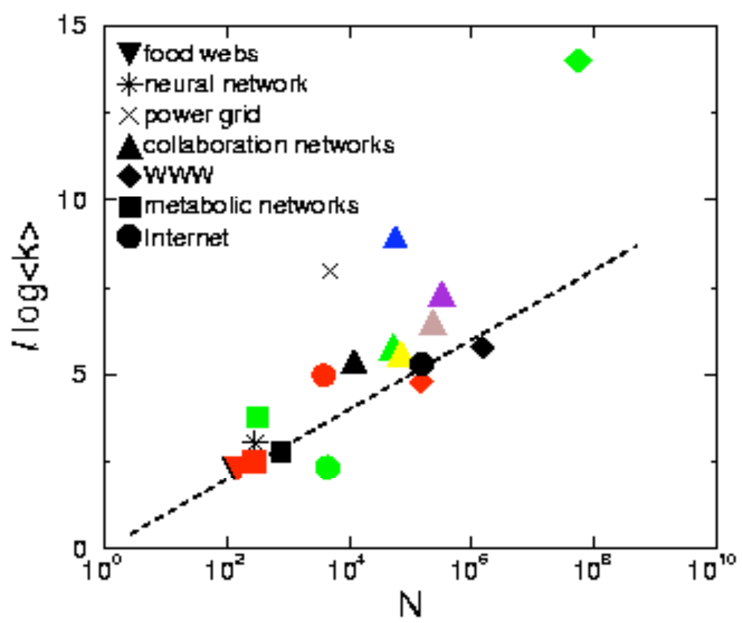
- Does the BA model generate limit networks with large clustering? In other words, are BA networks similar to WS graphs?
  - Almost, the CC is 5 times larger than that of random graphs and in general scales lower than in random graphs as N increases (order $(\log N)^2/N$ vs. $1/N$)

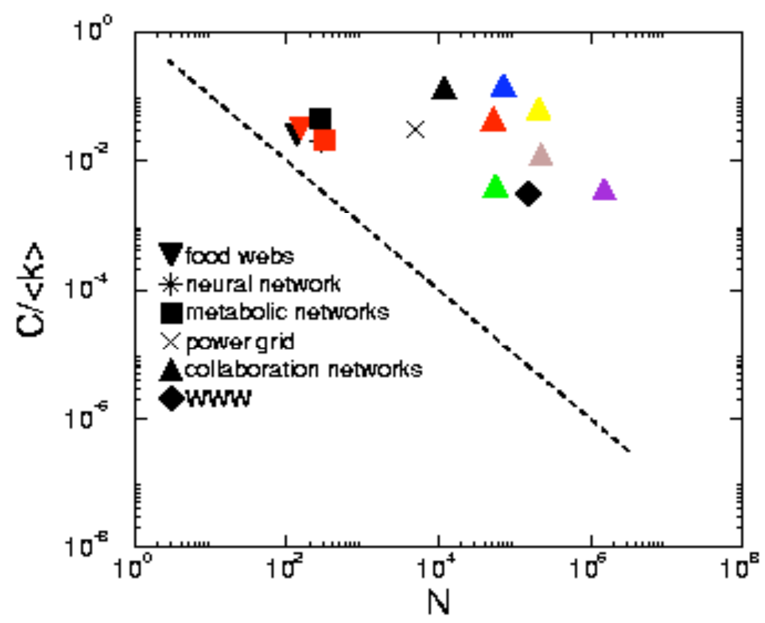
- Therefore the BA model generates limit graphs that are structurally similar to many observed networks... but
  - Is network growth an acceptable assumption in reality? What about preferential attachment?

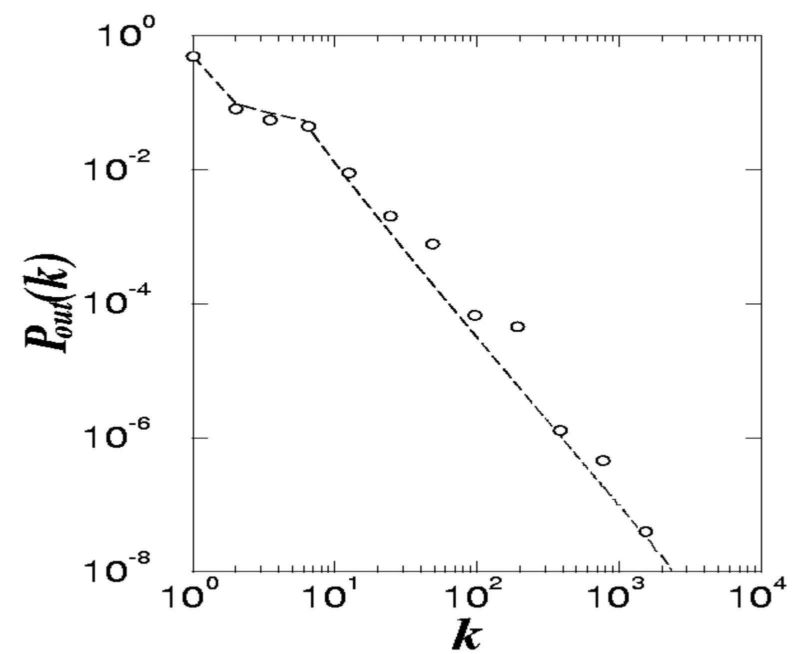
Network Models: Conclusions

- Robustness of networks to attacks
  - What if a randomly-targeted node is removed due to an attack (viruses, terrorists)?
  - What are the consequences in terms of connectivity?
  - Comparing Poisson random networks, WS and BA networks

- Summary of empirical findings about networks







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

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

$$P(k) \sim k^{\gamma}$$

Network Models: Conclusions

	Features	Average path length	Clustering	Degree distribution
Empirical Evidence	Stylized Facts	Low	High	Sometimes: Bell-shaped Most often: Power-law
Theoretical Models	Regular networks	Very high	Very High	All nodes have the same degree
	Random networks	Low	Low	Bell-shaped with exponentially-decaying tails
	Watts-Strogatz	Low	High	Bell-shaped with exponentially-decaying tails
	Barabasi et al.	Low	Medium-High	Power-law

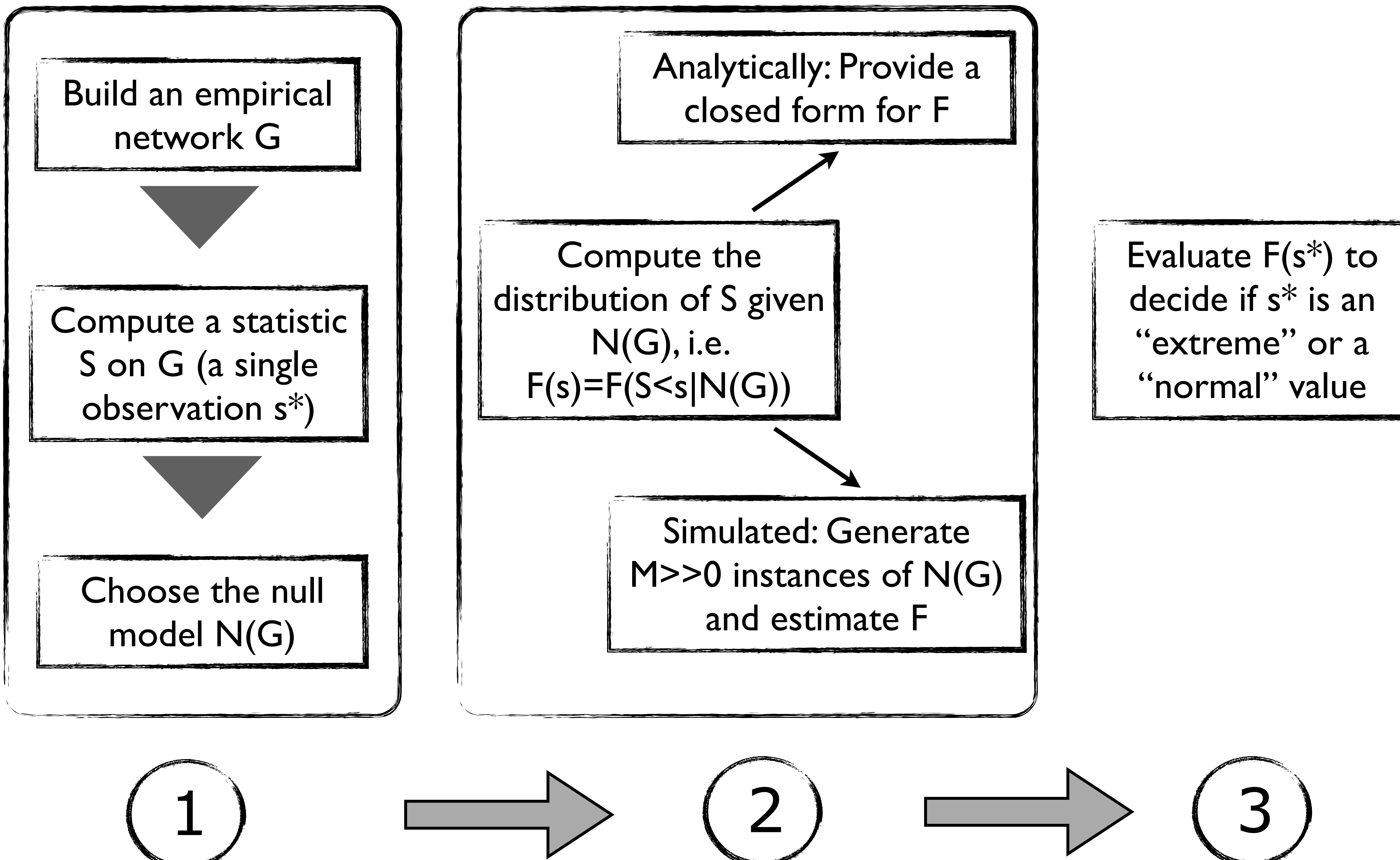
What's Next

- What is a network? Examples of networks
- Why networks are important for economists?
- Networks and graphs
- Measures and metrics on networks
- Distributions of metrics and measures in large networks
- Models of network formation
- Null statistical network models
- Economic applications

Null Statistical Network Models

- Inference on empirical network properties
  - ✓ Suppose we have observed a network G and we have computed a set of interesting network statistics on G, say $s_1(G),\dots,s_k(G)$
  - ✓ These may be: the clustering coefficient or the correlation between ND and ANND, or between NS and ANNS
- Problem: how can we say something about whether these observed values are large or small?
  - ✓ We need statistical benchmarks (null models) to assess the distribution of any given statistics given the null model at hand
  - ✓ Many ways to do it: one must choose the most appropriate null model, i.e. decide which properties of the observed graph we want to preserve
  - ✓ The null model generates maximally random graphs satisfying the selected constraints (preserved properties)
- Example: Maximally random graphs given
  - ✓ Binary graphs: Density only, degree distribution, degree sequence, etc.
  - ✓ Weighted graphs: Weight distribution, binary topology, etc.

Design of the Experiment



Binary Networks: Example 1

- Preserving density only
  - ✓ Compute density d on observed graph G with N nodes and L links
  - ✓ Generate a Poisson random graph with density d

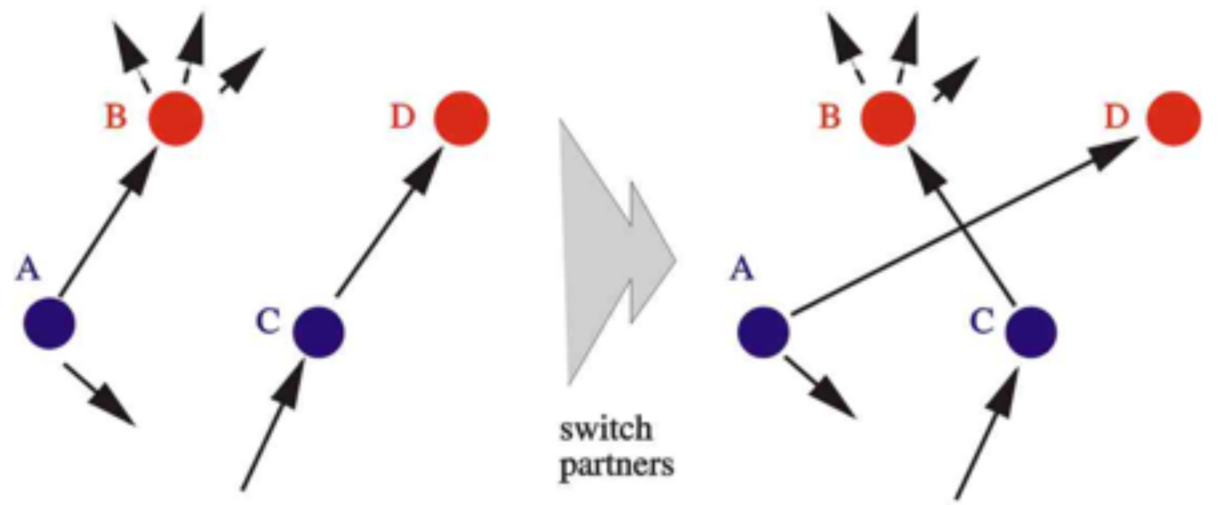
- Two alternatives
  - ✓ Average or exact density: using G(N,p) or G(N,m) models, where p=d and m=L (number of links in G)

- What is not preserved
  - ✓ Actual links, degree sequence and distribution all change

- Extension to digraphs
  - ✓ Generating Poisson random digraphs in such a way to preserve either the total number of links (in- and out- density change) or the exact number of in and out links in the observed graph

Binary Networks: Example 2

- Degree-preserving random rewiring (Maslov & Sneppen, 2002)
  - ✓ Preserve exactly the degree sequence of the observed binary network



A pair of directed edges A->B and C->D is randomly selected. These edges are then rewired in such a way that A becomes connected to D, while C to B, provided that none of these edges already exist in the network, in which case the rewiring step is aborted and a new pair of edges is selected. Note that the above rewiring algorithm conserves both the in- and out-degree of each individual node (and degrees if the graph is undirected)

- Problems
  - ✓ To get a single instance we need many such rewirings (at least 4L): this takes time

Weighted Networks

- Random reshuffling preserving density only
  - ✓ Generate Poisson random graphs with exact observed density d
  - ✓ Randomly reshuffle existing (positive) link weights across the randomly generated instance of binary topology using G(N,m)
- Maslov-Sneppen for WUNs and WDNs
  - ✓ Maslov-Sneppen rewiring algorithm works perfectly also for WUNs and WDNs
  - ✓ Just move the link-weight together with the link that is rewired
  - ✓ This does not preserves strength (in/out/tot) sequence (check it)
- Preserving weight distribution and binary topology
  - ✓ Reshuffle weights among existing binary links
  - ✓ This preserves link-weight distributions and binary structure (A)
  - ✓ Therefore also the degree sequence is preserved exactly

Alternative Null Models (1)

- Configuration model
  - ✓ An alternative algorithm to generate random (binary undirected) networks with a given degree sequence $\{k_1,\ldots,k_N\}$
  - ✓ Suppose that $\{k_1,\ldots,k_N\}$ is graphic, i.e. it is a feasible degree sequence of a graph (e.g. sum of all degrees is even). This is automatically satisfied if the sequence comes from an empirically-observed graph

- Algorithm
  - ✓ Construct a sequence where node i is listed k_i times for all i

$$\underbrace{1,1,\ldots,1}_{k_1 \text{ times}} \quad \underbrace{2,2,\ldots,2}_{k_2 \text{ times}} \quad \cdots \quad \underbrace{N,N,\ldots,N}_{k_N \text{ times}}$$

  - ✓ Randomly pick any two elements from the list and form a link between the nodes corresponding to those entries.
  - ✓ Delete those entries from the list and repeat until we get to the end (note: if the sum of degrees were odd we will remain with a single node)

Alternative Null Models (2)

- Output of the configuration model
  - ✓ A random graph where the degree sequence is preserved
  - ✓ Problems: multiple self-loops and multi-edges are not ruled out
  - ✓ Therefore the configuration model generates multi graphs

- Ways out
  - ✓ Delete multi-edges and all self loops: this destroys degree sequence but if multi edges are not that frequent the resulting degree distribution is close to the observed one
  - ✓ Employ null models preserving degree sequence only on average

- Expected-degree models (and beyond)
  - ✓ Chung-Lu (2002)
  - ✓ Squartini, Garlaschelli (2011) , Squartini, Fagiolo, Garlaschelli (2011a,b)

Alternative Null Models (3)

- Chung-Lu model
  - ✓ Start with the observed degree sequence $\{k_1,\ldots,k_N\}$ and an empty graph
  - ✓ Go through each pair of nodes and form a link with probability

$$\frac{k_i k_j}{\sum_h k_h} = \frac{k_i k_j}{2L}$$

- Notice
  - ✓ Self-loops are still allowed with probability $k_i^2/2L$ but no multiple edges between different nodes
  - ✓ To have well-defined probabilities it must be that $\max\{k_i\}<\text{Sqrt}(2L)$. It can be checked that if the degree sequence is very broad this condition is not satisfied and the ratios above are larger than one. This means that the observed degree sequence cannot be replicated on average.
  - ✓ What is the probability of connecting i and j in the configuration model? Prove that is equal to $k_i k_j/(2L-1)$, i.e. equal to the one in Chung-Lu model for large L

Economic Interpretation of Null Models

- Null models provide a statistical benchmark to compare observed network statistics
  - ✓ Almost no economics in them
  - ✓ Why are they useful in economics?

- Null statistical models in economics
  - ✓ Suppose a given observation is in line with what predicted by a given null statistical model. Then that value of the statistics does not require additional economic explanations. It can be simply the outcome of randomness. If we provide an economic model that reproduces that observation then that model could not be selected against the null random model.
  - ✓ Suppose instead a given observation is found to be an extreme value for the null model at hand. Then the null model must be rejected because it cannot explain that observation. We need to find an explanation elsewhere, probably in the economic realm.

Null Models: The Case of ITN

- **Two levels**
  - Null models of the ITN
  - Economic models of the ITN

- **Null models of the ITN**
  - Can observed properties be replicated by a null random network model that only preserves some local (1^{st}-order) statistics?
  - What is (if any) the minimal amount of information about the ITN needed to reproduce all its properties using an otherwise random model?
  - Can one discriminate between statistically relevant and irrelevant properties?

- **Economic models of the ITN**
  - Standard Int'l Trade Models: Gravity Model (GM)
  - Economics-Inspired Stochastic Models of Network Formation

Null Models: The Case of ITN

- **Main Idea**
  - Given observed network, define a set of local properties of the network (constraints) that must be preserved (density, degree or strength sequence, etc.)
  - Characterize the ensemble of all networks that preserve on average these constraints but are otherwise purely random
  - Obtain expected value and standard deviation of higher-order network statistics (assortativity, clustering, centrality, etc.) over the ensemble
  - Compare observed vs. expected values
- **Application to the ITN**
  - We study null models where we keep fixed either (in/out) degree or strength sequences and we check higher order statistical network properties (disassortativity, clustering)
  - By product: Are standard (local) international-trade statistics sufficient for explaining higher-order network properties?

Null Models: The Case of ITN

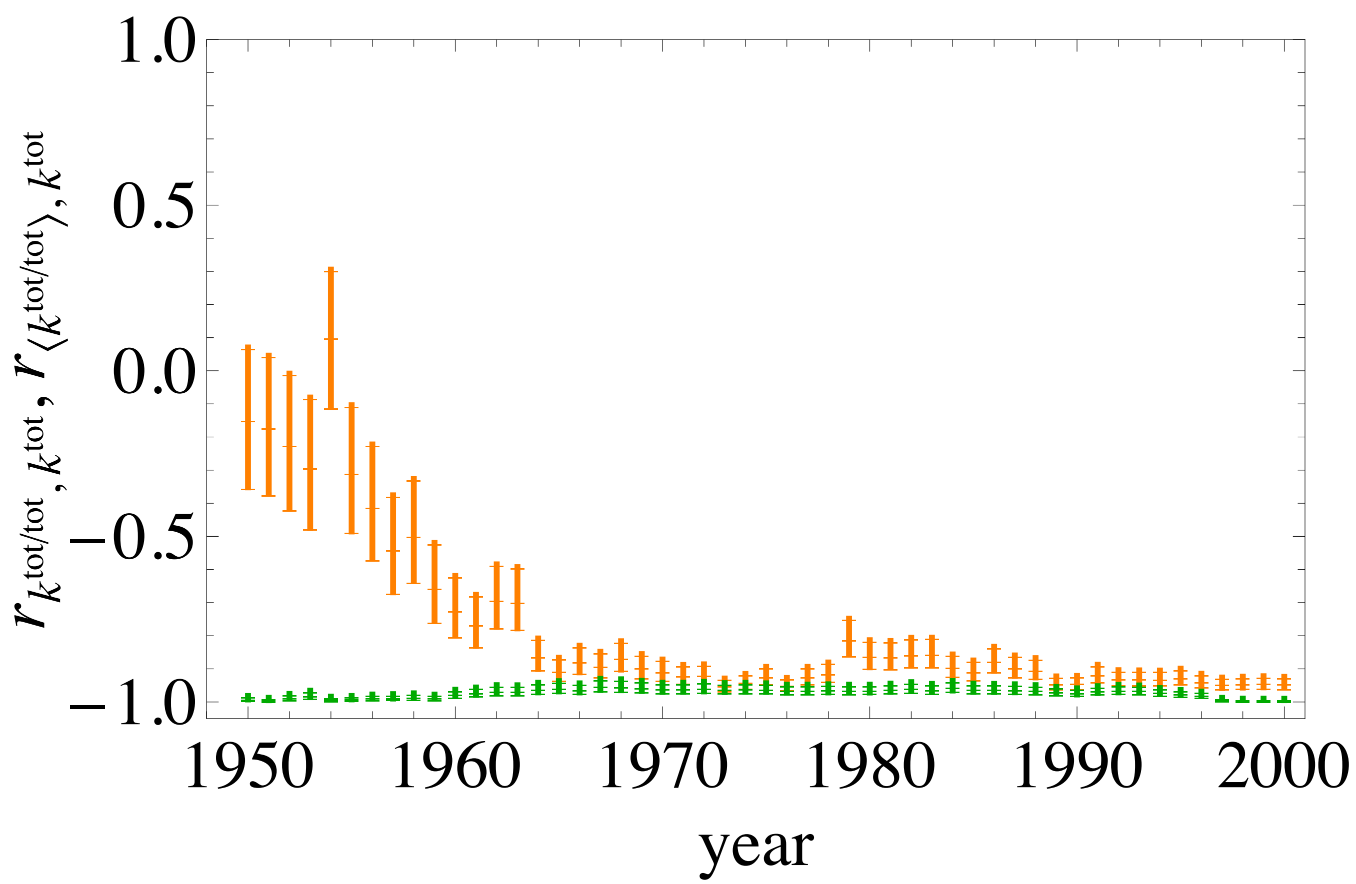
- **Features (Squartini & Garlaschelli, 2010)**
  - Fit to observed network the probability $P(G)$ of a random graph satisfying a list of local constraints (inferred from observed network)
  - Fully analytical method: no random variant must be generated
  - Works for directed/undirected, binary/weighted, sparse/dense networks
  - Expected properties computed in same time as empirical ones

- **A 3-Step Method**
  - Find the graph probability distribution $P(G; \overrightarrow{\theta})$ that maximizes graph entropy subject to constraints
  - Use observed data to estimate via ML free parameters $\overrightarrow{\theta}$ in the graph probability distribution obtained above
  - Use ML estimates of free parameters $\overrightarrow{\theta^*}$ to compute expected values and standard deviations of higher-order network statistics $X(G)$

$$E(X|\overrightarrow{\theta^*}) = \sum_G P(G|\overrightarrow{\theta^*})X(G)$$

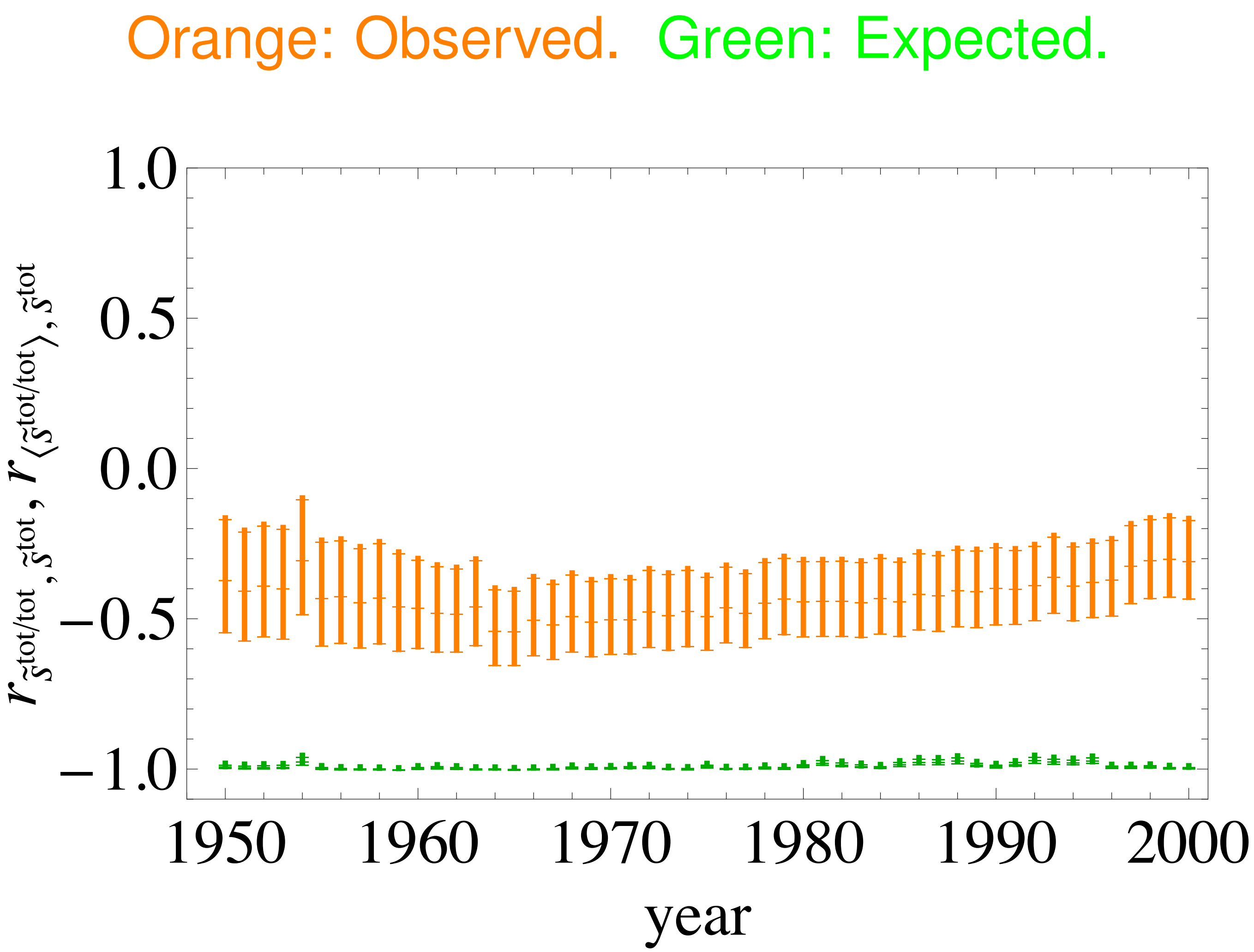
Null Models: The Case of ITN

Orange: Observed. Green: Expected.



- Contraint: Degree sequence
- Null model always predicts strong disassortativity
- ITN is strongly disassortative only after 1965
- Null model well predicts disassortativity (when it is a robust network feature)

Null Models: The Case of ITN



- Contraint: Strength sequence
- Null model always predicts extreme weighted disassortativity
- Weighted (weak) disassortativity patterns (arising consistently from 1950 to 2000) cannot be replicated

Null Models: The Case of ITN

- **General Results**
  - Binary ITN: Degrees are sufficient to reproduce all higher-order statistics
  - Weighted ITN: Strengths are not sufficient to reproduce higher-order statistics
- **Implications for network analysis**
  - Binary ITN: disassortativity and clustering patterns do not convey any interesting information
  - Weighted ITN: higher-order statistics convey fresh information, which is not already contained in strength sequences
- **Implications for international-trade empirics**
  - A weighted-network analysis brings value added wrt standard (local) int'l-trade statistics
  - Degree sequences are maximally informative: trade models should focus on explaining new-link formation and degrees (in addition to trade flows)

What’s Next

- What is a network? Examples of networks
- Why networks are important for economists?
- Networks and graphs
- Measures and metrics on networks
- Distributions of metrics and measures in large networks
- Models of network formation
- Null statistical network models
- Economic applications