Complex Dynamical Networks:

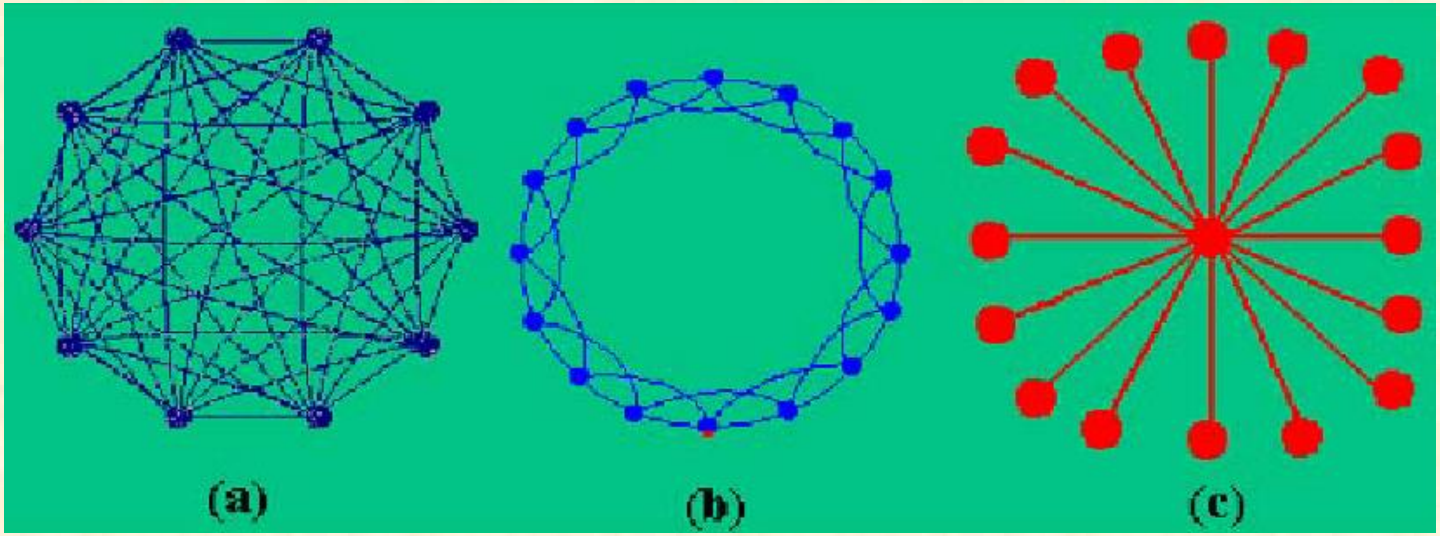
Lecture 2: Network Topologies

-- Basic Models and Properties

EE 6605

Instructor: G Ron Chen

Regular Networks

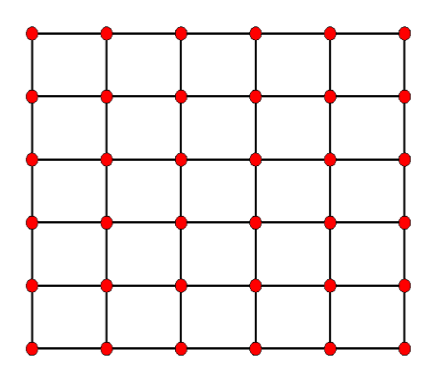


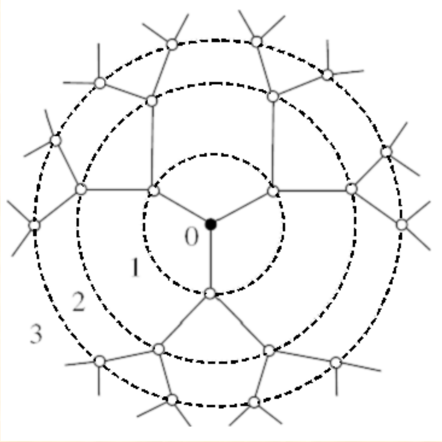
(a) **Fully-connected network**

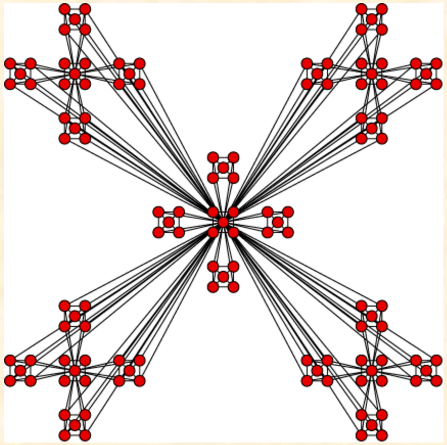
(b) **Ring-shaped coupled network**

(c) **Star-shape coupled network**

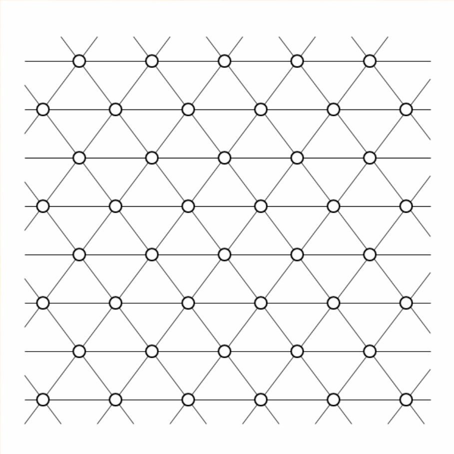
Regular Networks







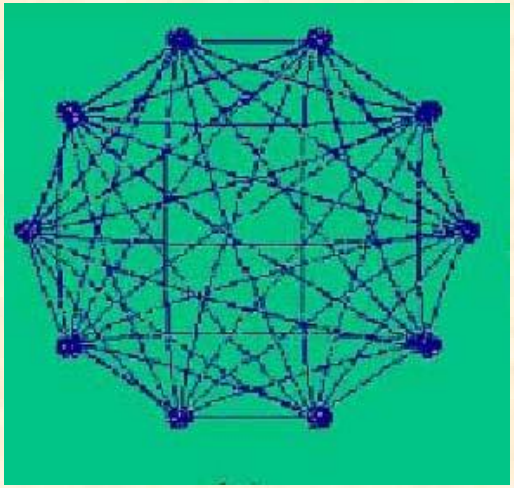
Fractal Network



Lattice

Basic Properties of Regular Networks:

(a) Fully-connected networks:



- **Average path length:**

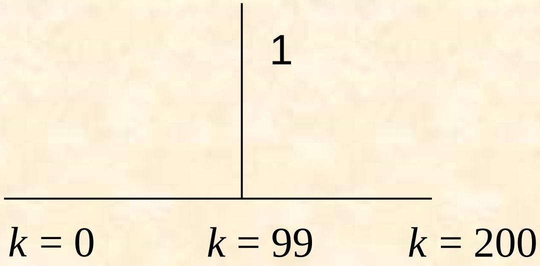
$$L_{full} = 1$$

- **Clustering coefficient:**

$$C_{full} = 1$$

- **Degree distribution:**

delta

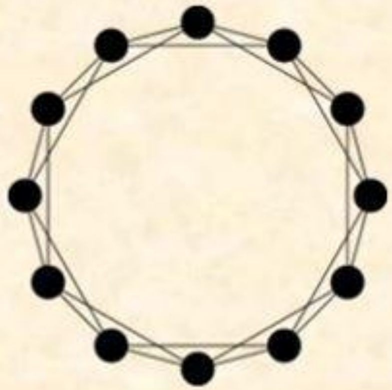


Example: $N = 100$

Basic Properties of Regular Networks:

(b) Ring-shaped coupled networks

- **Average Path Length:**



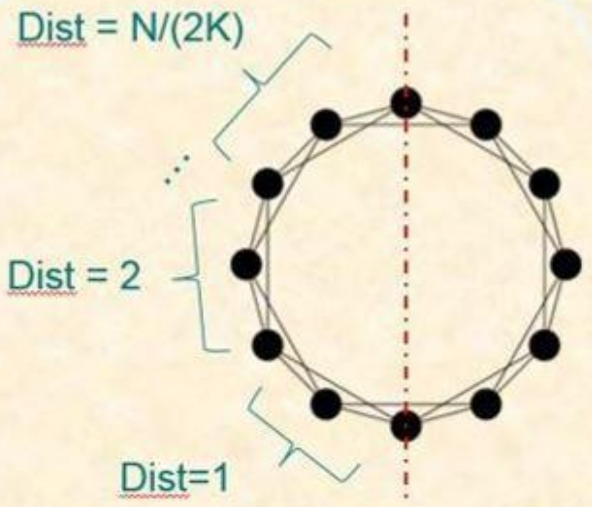
K – number of neighboring connections

$$L_{ring} \approx \frac{1}{N/2}\left[\frac{K}{2}\left(1+2+\cdots\frac{N}{2K}\right)\right] \approx \frac{N}{2K}$$

$K = \frac{N}{2}$ → fully connected → $L_{ring} = 1$

$L_{ring} \to \infty$ as $N \to \infty$

(K fixed)



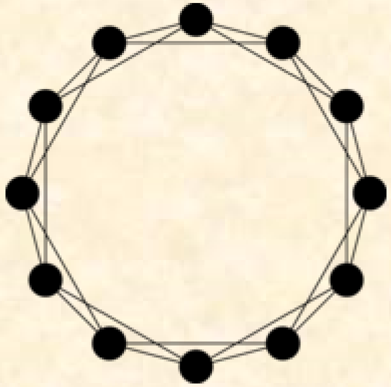
Note: This is an asymptotic formula: the bigger the N and K, the more accurate the formulas

Basic Properties of Regular Networks:

(b) Ring-shaped coupled networks

- **Clustering Coefficient:**

$$C_{ring} \approx \frac{3(K-1)}{2(2K-1)} \rightarrow \frac{3}{4} \quad (K \rightarrow \infty)$$



K =2

- K is the number of edges on one side
- Verified by mathematical induction
- It is a local formula, independent of N

Note: This is an asymptotic formula: the bigger the K (hence N), the more accurate the formulas

For a perfect ring:
$K = 1, \quad C_{ring} = 0$

Basic Properties of Regular Networks:

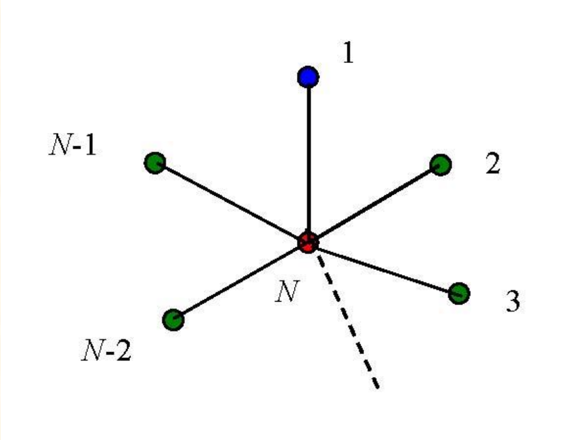
(c) Star-shaped coupled networks:

- **Average Path Length:**

$$L_{star} = 2 - \frac{2}{N} \to 2 \qquad (N \to \infty)$$

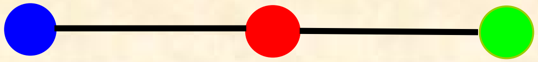
Verifying:

$$\frac{(N-1)\times 1 + [(N-2)+(N-3)+\cdots+2+1]\times 2}{(N-1)+[(N-2)+(N-3)+\cdots+2+1]} = 2 - \frac{2}{N}$$



- **Clustering Coefficient:** $C_{star} = 0$
- **Degree Distribution: delta-like**

Example: $N = 3$

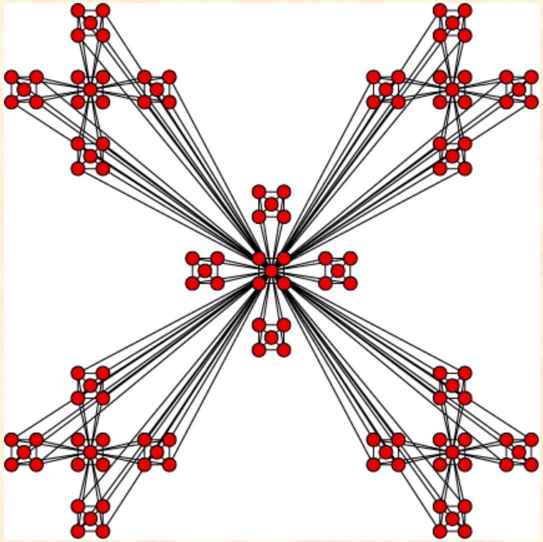


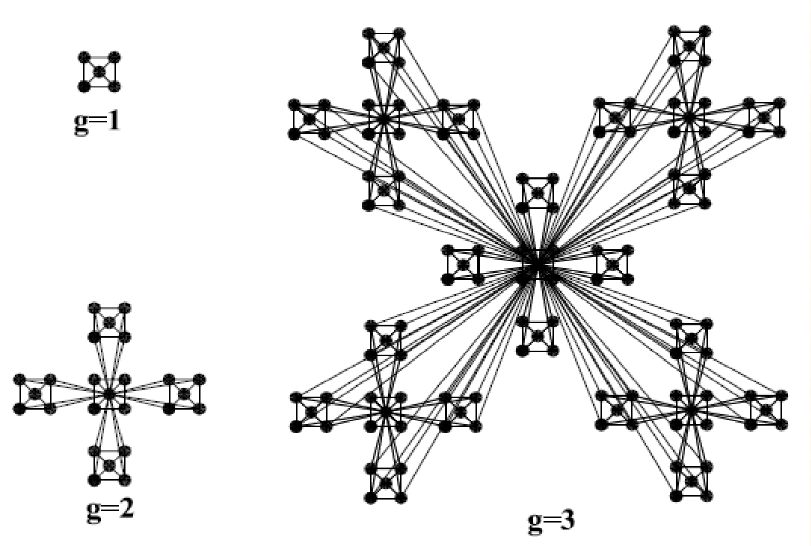
There are 3 different paths with lengths 1, 1, 2

$L = (1+1+2) / 3 = 4 / 3$ or $L = 2 - [2 / (N=3)] = 4 / 3$

Fractal Networks

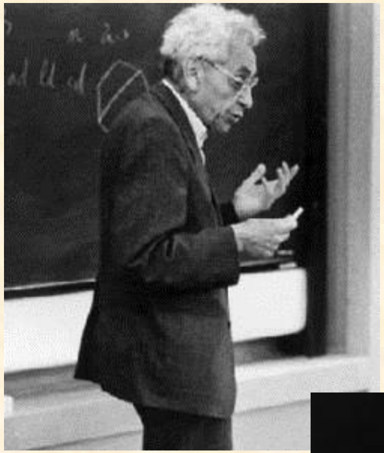
Example:

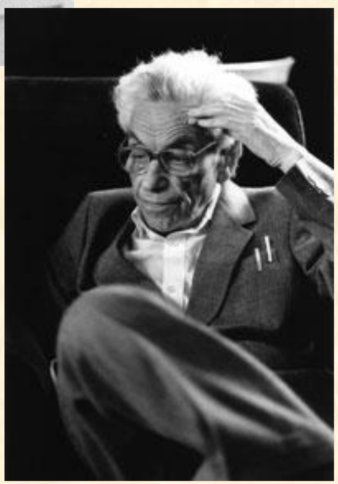


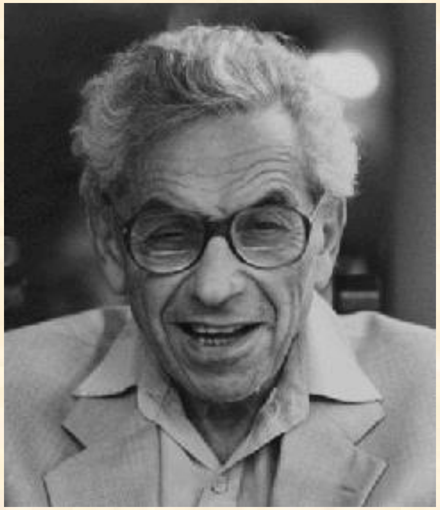


Step g =	1	2	3	N
Maximum degree $d_{\max}$	4	$4+4\times4$ $=20$	$4+4\times4+4\times4\times4$ $=84$	$4+4\times4+\cdots+4\times4\times4+\cdots+4^N$ $=\frac{4}{3}\left(4^N-1\right)$
Minimum degree $d_{\min}$	3	3	3	3
How many nodes with $d_{\min}$	4	4	$4+4\times4$ $=20$	$4+4\times4+\cdots+4\times4\times4+\cdots+4^{N-1}$ $=\frac{4}{3}\left(4^{N-1}-1\right) \quad (N\geq2)$

Random Graph Theory





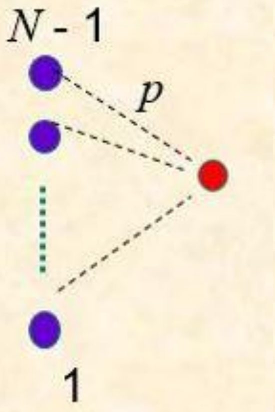


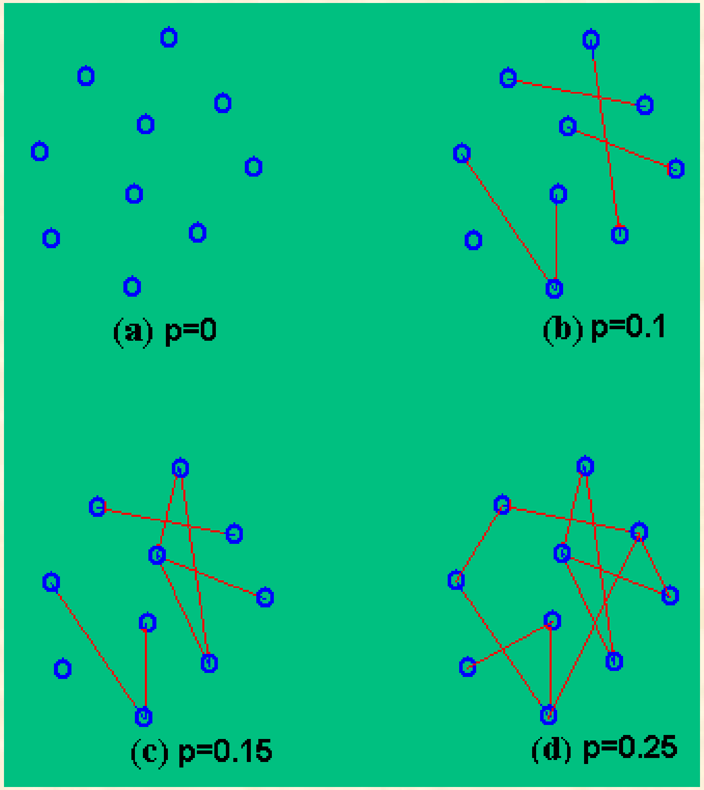
Paul Erdös
(1913-1996)

Alfred Rényi
(1921-1970)

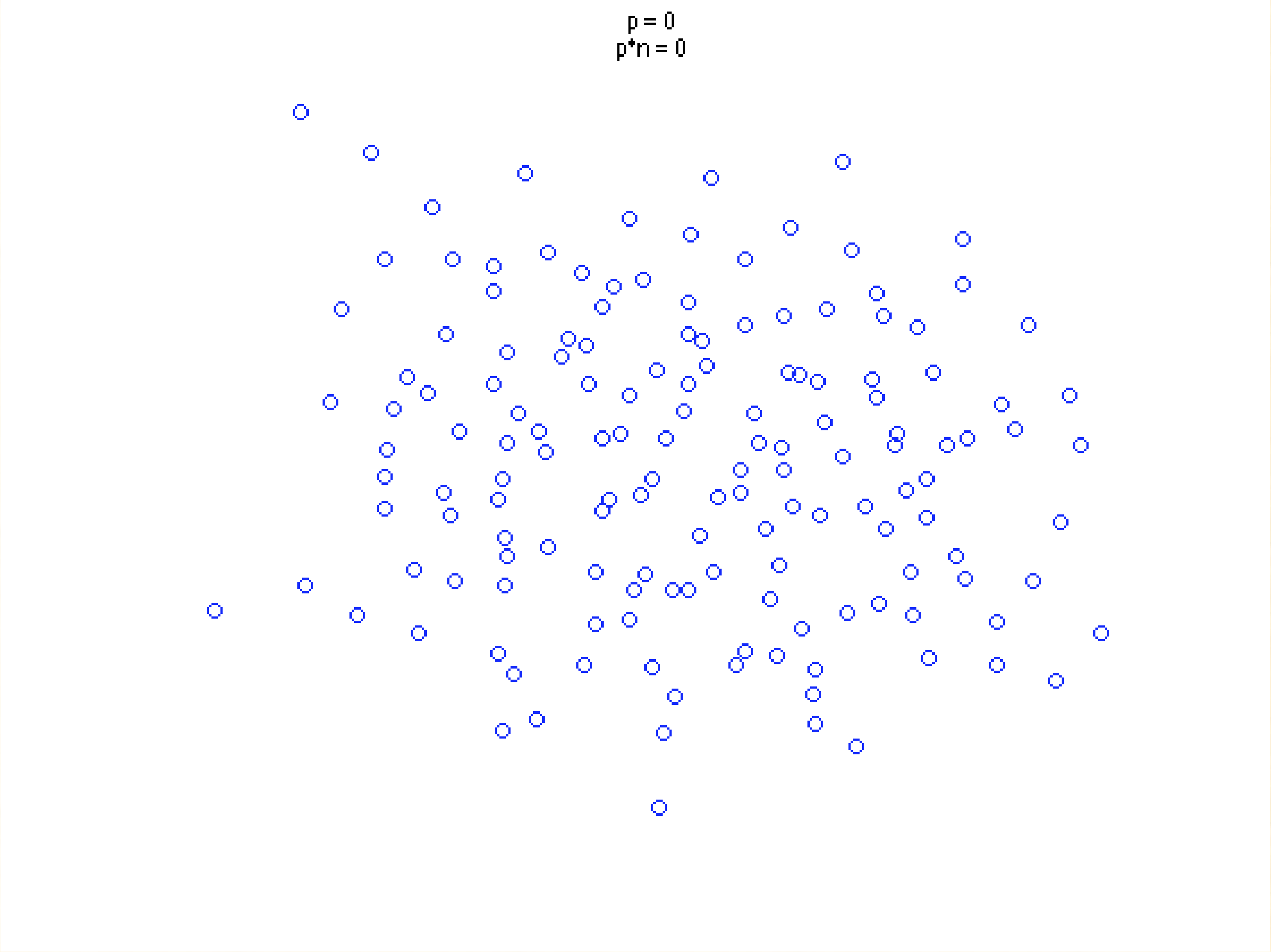
Erdös-Rényi Random Graph Theory

- Given N isolated nodes
- Add an edge between every possible pair of nodes with probability p
- Statistically, the expected number of edges is $pN(N-1)/2$



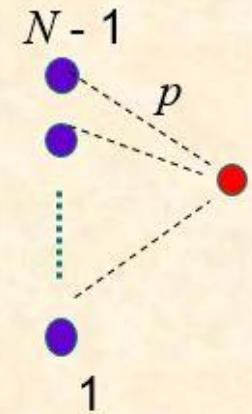


Example



Random-Graph Networks

Theorem:

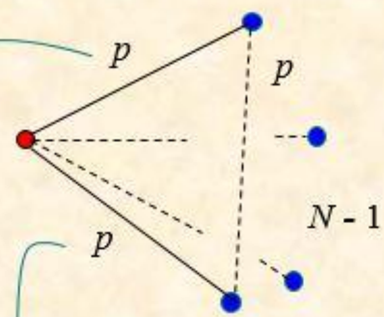


- **Average Degree:** $\langle k \rangle = p(N-1) \approx pN$
- **Clustering Coefficient:** $C = p = \langle k \rangle / N$
- **Degree Distribution:** $P(k) = \dfrac{\mu^k}{k!} e^{-\mu}$ (Poisson)
- **Average Path Length:** $L_{ER} \sim \ln N \,/\, \ln\langle k \rangle$

Verifying the Clustering Coefficient:

triangle

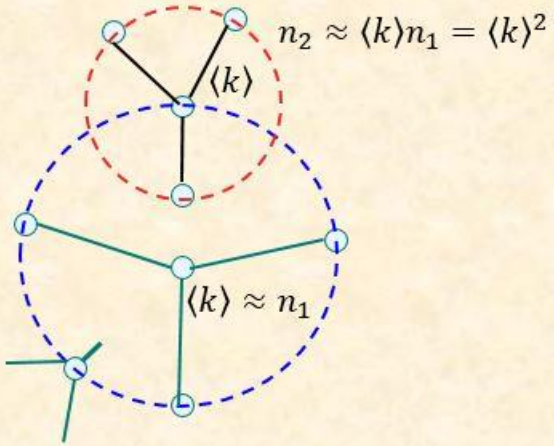
$$C = \frac{p \cdot p^2 \cdot \binom{N-1}{2}}{p^2 \cdot \binom{N-1}{2}} = p = \langle k \rangle / N$$



Triangular shape

Random-Graph Networks

Theorem:



- **Average Degree:**
- **Clustering Coefficient:**
- **Degree Distribution:**
- **Average Path Length:** $L \sim \ln N / \ln\langle k\rangle$

Verifying: For any node, it can connect to roughly $n_1 \sim \langle k\rangle$ nodes. Each next node can connect to roughly $n_2 \sim \langle k\rangle n_1 = \langle k\rangle^2$ nodes, and so on. Thus, because the average distance between each pair of nodes is L there are roughly $n = N \sim \langle k\rangle^L$ nodes in the network, yielding

$$L \sim \ln N / \ln\langle k\rangle$$

Poisson Distribution:

$$\boxed{P(k)=\frac{\mu^k}{k!}e^{-\mu}}$$

- The probability of a node connecting to k nodes among other $N-1$ nodes (but not connected to the rest $N-1-k$ nodes) is given by

$$P(k\,|\,N')=\binom{N'}{k}p^k(1-p)^{N'-k}$$

where $N'=N-1$

- Let $\mu=pN'$ (referred to as expectation value) and fix it as constant
- Thus, $P_\mu(k\,|\,N')=\dfrac{N'!}{k!(N'-k)!}\left(\dfrac{\mu}{N'}\right)^k\left(1-\dfrac{\mu}{N'}\right)^{N'-k}$

$$P(k)=\lim_{N'\to\infty}P_\mu(k\,|\,N')$$

$$=\lim_{N'\to\infty}\frac{N'(N'-1)\cdots(N'-k+1)}{N'^k}\frac{\mu^k}{k!}\left(1-\frac{\mu}{N'}\right)^{N'}\left(1-\frac{\mu}{N'}\right)^{-k}$$

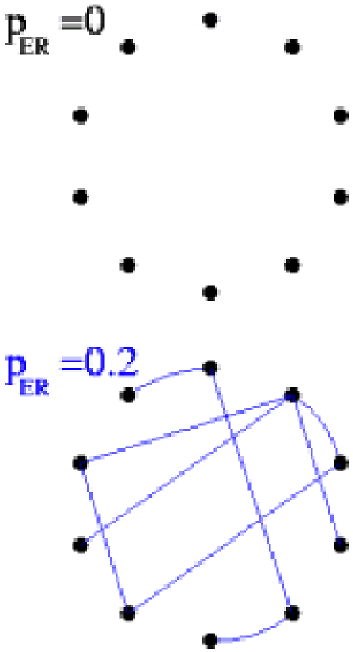
$$=1\cdot\frac{\mu^k}{k!}\cdot e^{-\mu}\cdot 1$$

ER Random Graph Models

Erdős-Rényi

(Publ. Math. Inst. Hung. Acd. Sci. **5**, 17 (1960))

N nodes, each pair of node is connected with probability p



Features:

- **Connectivity**
  node degree distribution - **Poisson**
- **Homogeneity**
  all nodes have about the same number of edges
- **Non-growing**

Random Graph and Poisson Degree Distribution

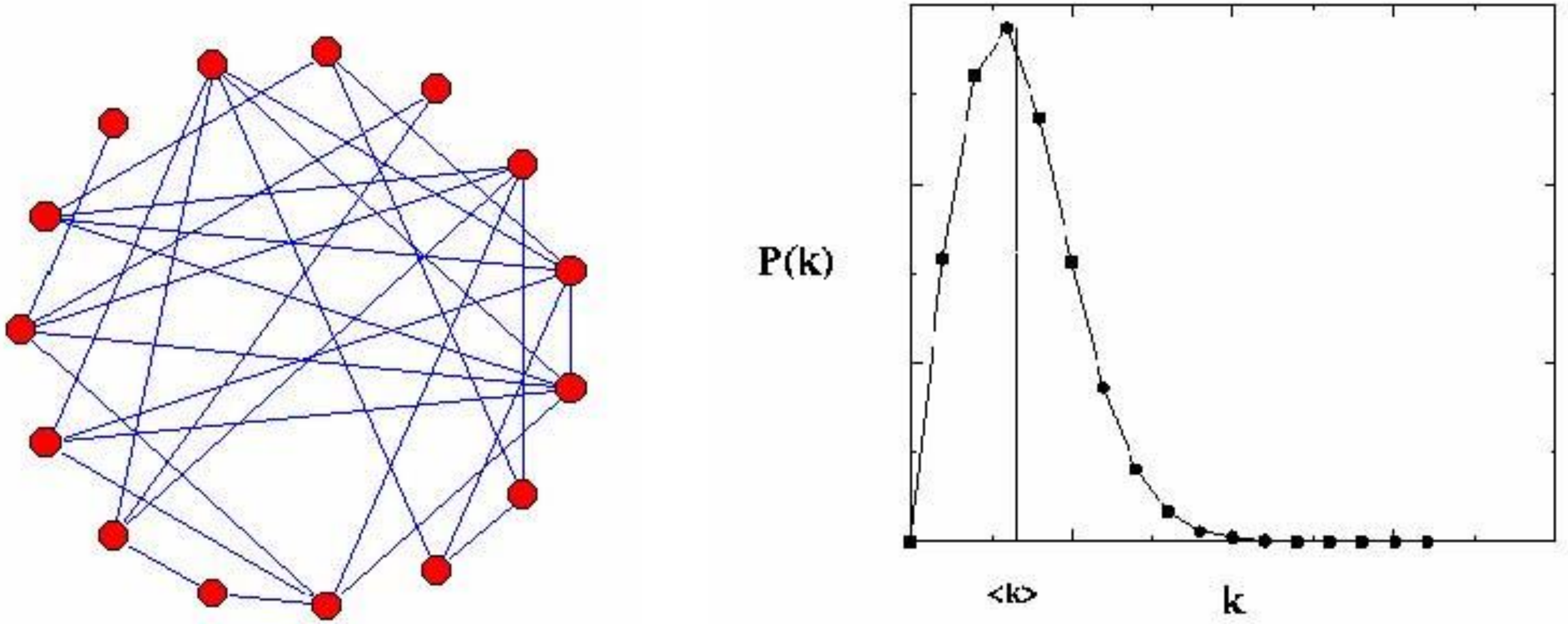


Illustration of Erdös-Rényi randon-graph network model

Small-World Networks

"Collective dynamics of 'small-world' networks"

--- Nature, 393: 440-442, 1998

D. J. Watts

S. H. Strogatz

(Cornell University)

A Historical Remark

CHAIN-LINKS
by
Frigyes Karinthy
(1929, *Everything is Different*)

We were arguing energetically about whether the world is actually evolving, headed in a particular direction, or whether the entire universe is just a returning rhythm's game, a renewal of eternity. "There has to be something of crucial importance," I said in the middle of debate. "I just don't quite know how to express it in a new way; I hate repeating myself:'

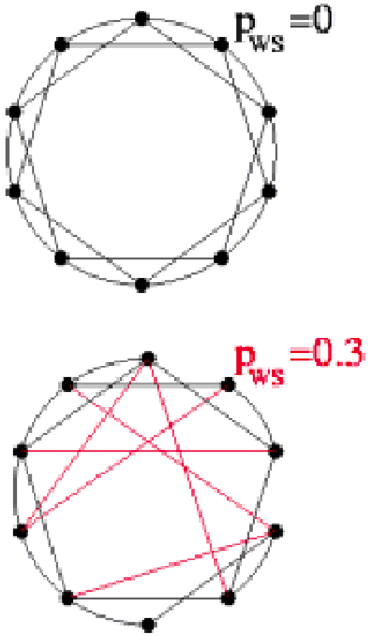
Let me put it this way: Planet Earth has never been as *tiny* as it is now. It shrunk - relatively speaking of course - due to the quickening pulse of both physical and verbal communication. This topic has come up before, but we had never framed it quite this way. We never talked about the fact that anyone on Earth, at my or anyone's will, can now learn in just a few minutes what I think or do, and what I want or what I would like to do. If I wanted to convince myself of the above fact: in couple of days I could be - *Hocus pocus!* - where I want to be.

A fascinating game grew out of this discussion. One of us suggested performing the following experiment to prove that the population of the Earth is closer together now than they have ever been before. We should select any person from the 1.5 billion inhabitants of the Earth - anyone, anywhere at all. He bet us that, using no more than *five* individuals, one of whom is a personal acquaintance, he could contact the selected individual using nothing except the network of personal acquaintances. For example, "Look, you know Mr. X.Y., please ask him to contact his friend Mr. Q.Z., whom he knows, and so forth."

Small-World Networks

Watts-Strogatz

(Nature **393**, 440 (1998))



N nodes forms a regular lattice. With probability p, each edge is rewired randomly

Features:

- **Connectivity Poisson distribution**
- **Homogeneous nature: each node has roughly the same number of edges**
- **Small average path length but large clustering coefficient**
- **Not growing**

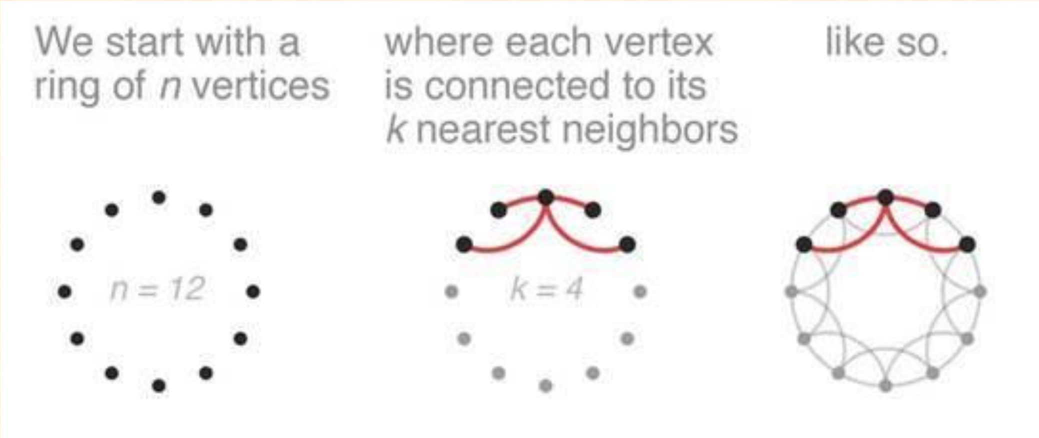
WS Small-World Network Model

Algorithm (rewiring):

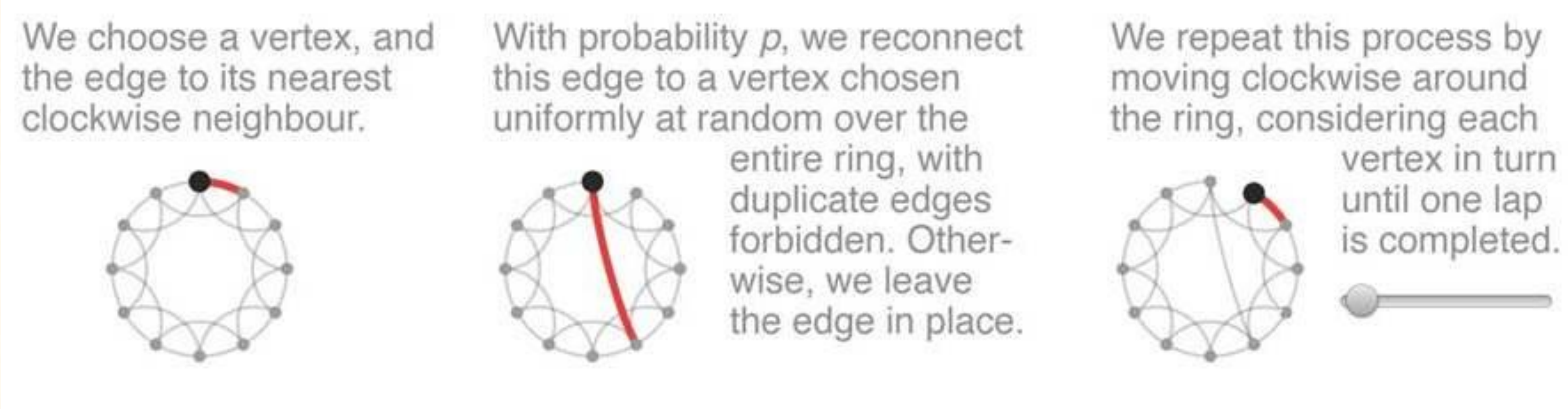
- Start from a ring-shaped network with N nodes, in which each node is connected to its $2K$ neighbors, where K is a (usually small) positive integer.

- Go around the ring-shaped network clockwise (or counterclockwise). Pick a node and operate on its connections to the K neighbors one by one: an edge connecting this node to its neighbor will be kept unchanged; another end of the edge will be disconnected with probability p and then be re-connected to a randomly-picked node from the network (ignore self-loops and multiple edges).

More precisely …



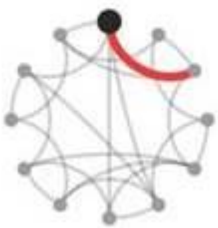
WS Small-World Network formation



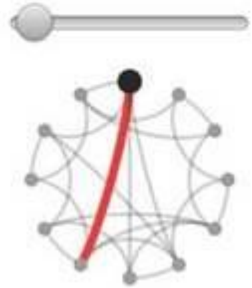
D.J. Watts and S.H. Strogatz:
Collective dynamics of 'small-world' networks *Nature* 393, 440-442 (4 June 1998)

Continuing …

Next, we consider the edges that connect vertices to their second-nearest neighbours clockwise.



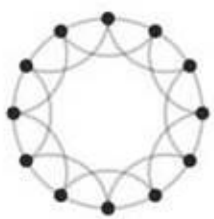
As before, we randomly rewire each of these edges with probability p.



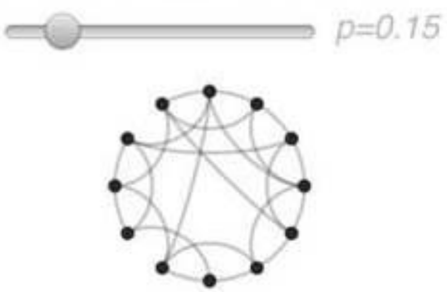
We continue this process, circulating around the ring and proceeding outward to more distant neighbours after each lap, until each original edge has been considered once.

As there are $nk/2$ edges in the entire graph, the rewiring process stops after $k/2$ laps.

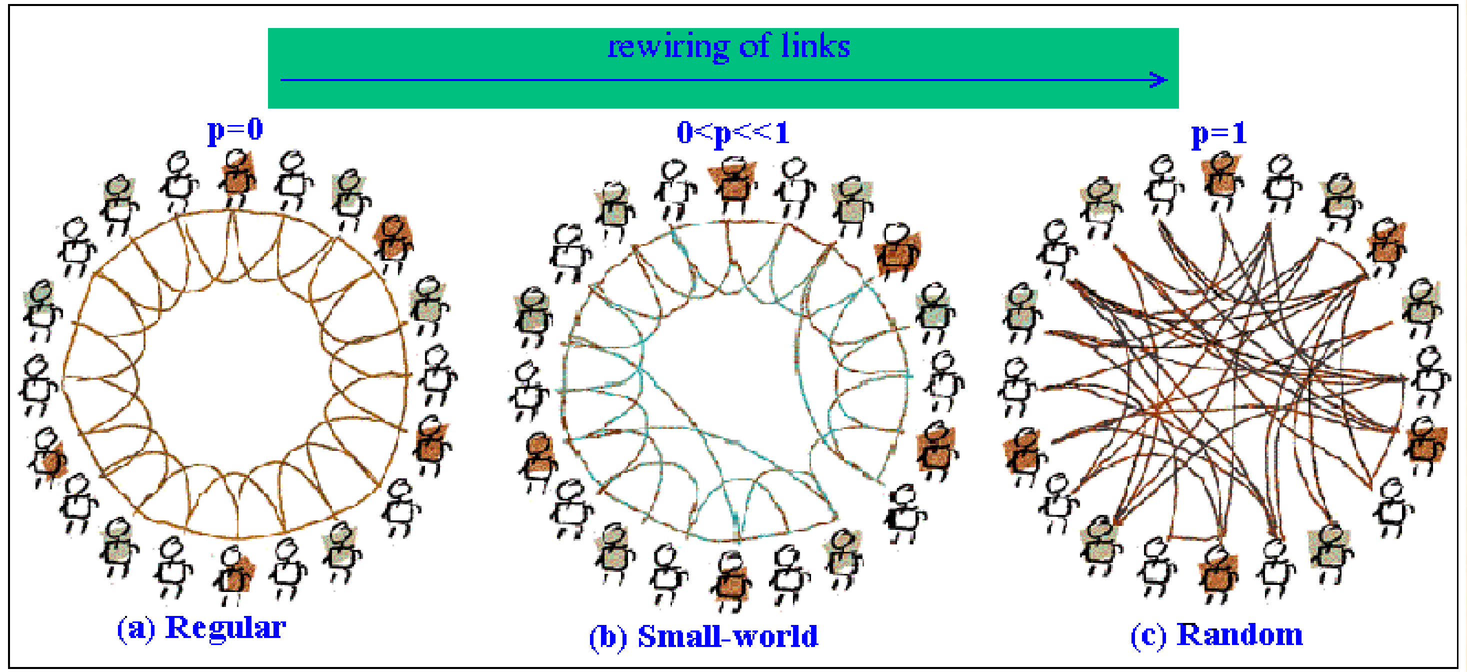
For $p = 0$, the ring is unchanged.

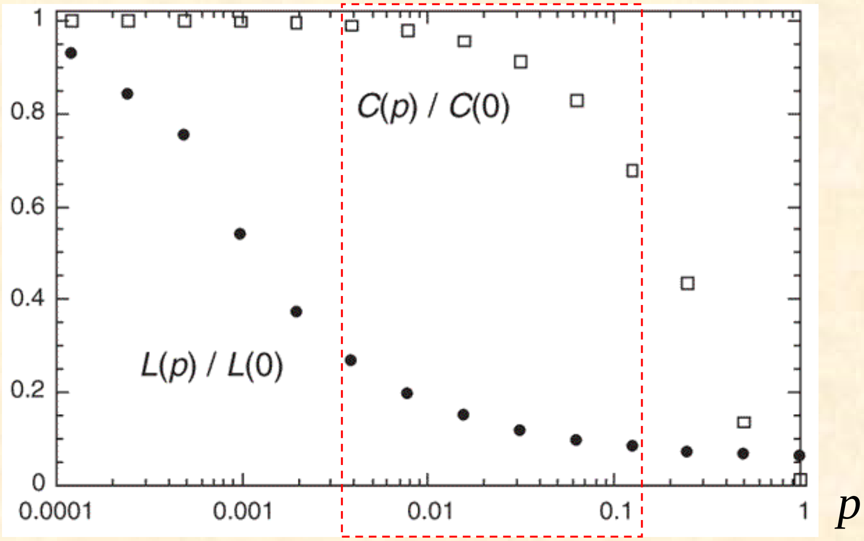


As p increases, the graph becomes increasingly disordered.



At $p = 1$, all edges are re-wired randomly.





$0.05 < p < 0.1$

Small-world features:

- **Large clustering coefficient C**
- **Small average path-length L**

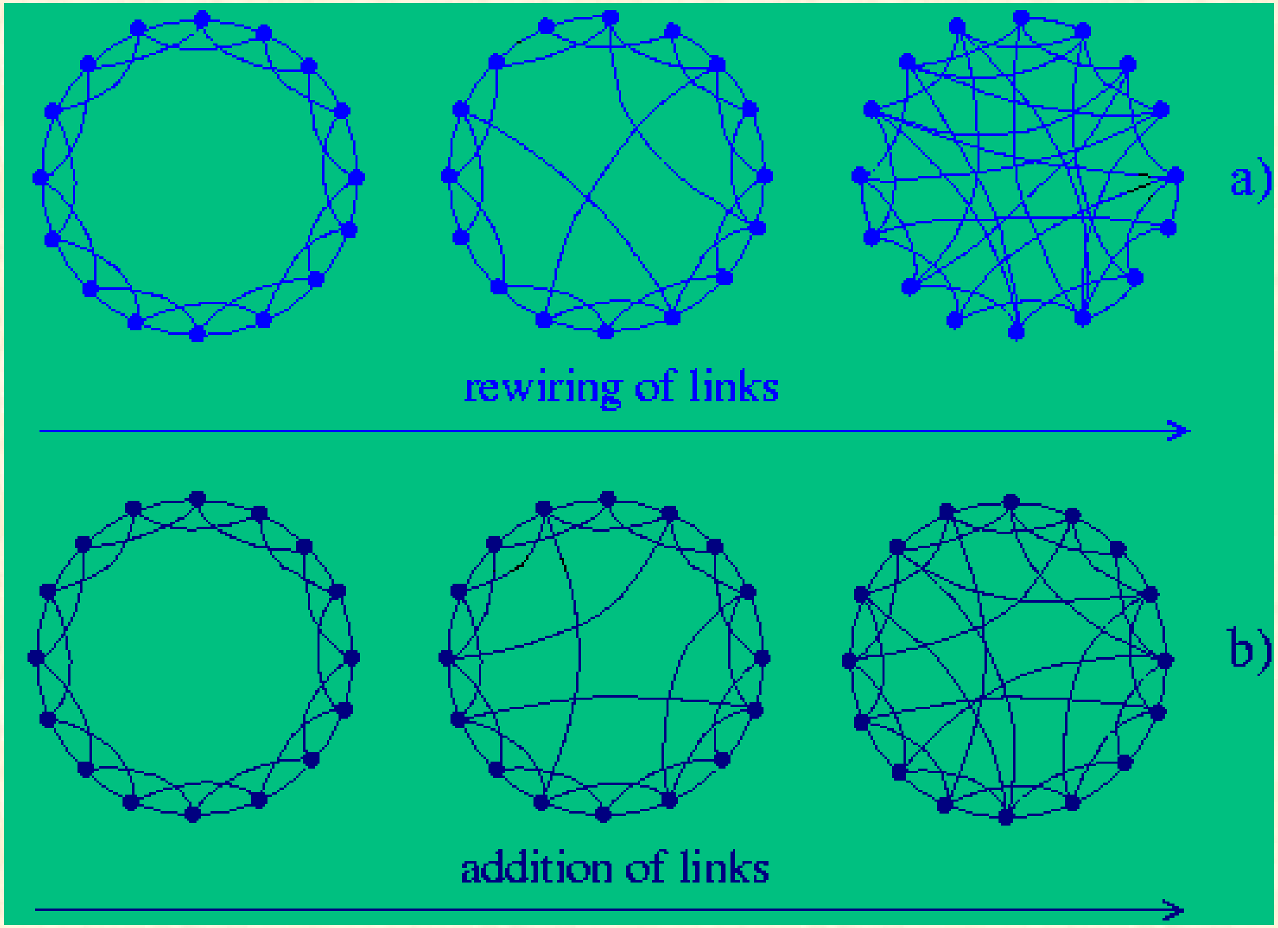
NW Small-World Network Model

Algorithm (adding edges):

- Start from a ring-shaped network with N nodes, in which each node is connected to its $2K$ neighbors, where K is a (usually small) positive integer.
- With a (small) probability p, add an edge to each pair of nodes in the network.

Mark E J Newman
Univ of Michigan

Two Small-World Network Models



a) **WS** small-world network model   b) **NW** small-world network model

Small-World Networks

Theorem: For large enough $\boldsymbol{N}$,

Clustering Coefficient of the WS small-world network model:

$$C(p) = \frac{3(K-2)}{4(K-1)}(1-p)^3$$

Clustering Coefficient of the NW small-world network model:

$$C(p) = \frac{3(K-2)}{4(K-1)+4Kp(p+2)}$$

M.E.J. Newman, The structure and function of networks, Computer Physics Communications, 147: 40-45, 2002

M.E.J. Newman, The structure and function of complex networks, SIAM Review, 45: 167-256, 2003.

Small-World Networks

Theorem:

Average Path Length of WS small-world network model:

$$L(p) = \frac{2N}{K} f(NKp/2)$$

$$f(x) = \begin{cases} c & x << 1 \\ \ln x / x & x >> 1 \end{cases}$$

(c — constant)

Average Path Length of NW small-world network model:

$$L(p) = \frac{2N}{K} f(NKp/2)$$

$$f(x) \approx \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \sqrt{\frac{x}{x+2}}$$

M.E.J. Newman and D.J. Watts, Randomization group analysis of the small-world network model, Phys. Lett. A, 253: 341-346, 1999.
A. Barrat and M. Weigt, On the properties of small world networks, European Phys. Journal B, 13: 547-560, 2000

Small-World Networks

Theorem:

Degree Distribution of WS small-world network model:

$$P(k) = \sum_{i=0}^{\min\{k-K/2,K/2\}} \binom{K/2}{i} (1-p)^i \, p^{(K/2)-i} \, \frac{(pK/2)^{k-i-K/2}}{(k-N-K/2)!} e^{-pK/2} \qquad k \geq K/2$$

$$P(k) = 0 \qquad k < K/2$$

Degree Distribution of NW small-world network model:

$$P(k) = \binom{N}{k-K} \left(\frac{Kp}{N}\right)^{k-K} \left(1 - \frac{Kp}{N}\right)^{N-k+K} \qquad k \geq K$$

$$P(k) = 0 \qquad k < K$$

M.E.J. Newman, The structure and function of networks, Computer Physics Communications, 147: 40-45, 2002

M.E.J. Newman, The structure and function of complex networks, SIAM Review, 45: 167-256, 2003.

Navigable Small-World Networks

International Congress of Mathematics (ICM)

22-28 August 2006, Madrid, Spain

Jon M Kleinberg (Cornell Univ.) received the Nevanlinna Prize for Applied Mathematics

He gave a 45-minute talk -
"Complex Networks and Decentralized Search Algorithms"

J M Kleinberg, "Navigation in a small world," *Nature*, 2000

J M Kleinberg, The small-world phenomenon: An algorithmic perspective. *Proc. 32nd ACM Symposium on Theory of Computing*, 2000: 163-170

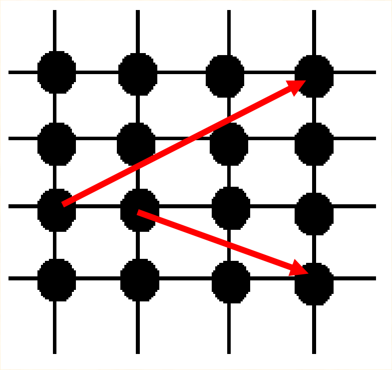
Kleinberg's Navigable Networks

Problem –

To deliver a message from a source to a destination on a network, in as few steps as possible

Suppose that

1. Given a lattice network
2. Add m long-range edges anywhere in the network



Kleinberg's Navigable Networks

Assumption: The probability of a new edge connecting a node u to any other node v in the given lattice:

$$P(u,v) = \beta d_{uv}^{-\alpha}$$

where d_{uv} is the distance between u and v with parameters $\alpha \geq 0, \beta > 0$ satisfying $\beta \sum_{(u,v)} d_{u,v}^{-\alpha} = 1$ as a probability measure.

When $\alpha = 0$, the new edge-addition probability is a constant, so every node has the same probability to receive the new edge; in this case, the model reduces to the original NW small-word network model.

Kleinberg's Navigable Networks

A network is *navigable* if it requires at most $poly(\log N)$ time-steps to deliver a message from any source to any destination on the network.

Kleinberg Theorem:

For $\alpha = 2$, the lattice is navigable in $O(\log N)^2$ time-steps at a rate depending on m, but for all $\alpha < 0$ it is not navigable.

For $0 \leq \alpha < 2$, it takes at least $O(N^{(2-\alpha)/3})$ time-steps to deliver a message from any source to any target.

For $\alpha > 2$, it will take at least $O(N^{(\alpha-2)/(\alpha-1)})$ time-steps.

For the Kleinberg problem and solution, formulated in the original NW ring-shaped network setting, see:

M E J Newman: ***Networks: An Introduction***. Oxford, UK, 2010: 713-718

Some More **Concepts**

Degree Correlations

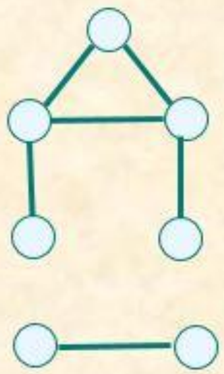
Weighted Networks

Evolving Networks

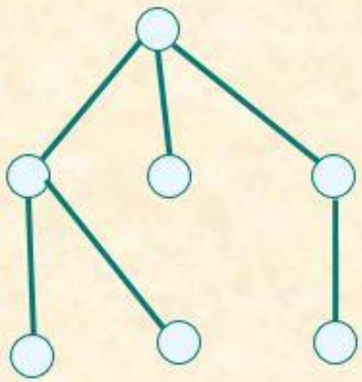
Degree Correlations

$$\langle k \rangle = \sum_{k=0}^{N-1} kP(k)$$

$P(k)$ = probability of a randomly picked node has degree k



$$\langle k \rangle = \frac{12}{7}$$



$$\langle k \rangle = \frac{12}{7}$$

Verifying:

$$\langle k \rangle = 1 \cdot \frac{4}{7} + 2 \cdot \frac{1}{7} + 3 \cdot \frac{2}{7} = \frac{12}{7}$$

→ Average degree alone is not sufficient to characterize a network

Degree Correlations

Degree Correlation (second-order degree distribution)

between two nodes of degree k and k' respectively

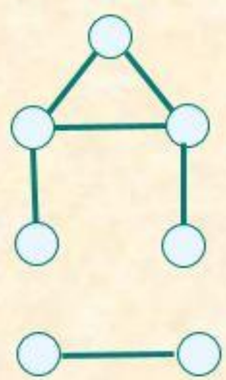
$$P(k,k') = \frac{a_{k,k'} \cdot m(k,k')}{2M}$$

$M =$ total number of edges in the network

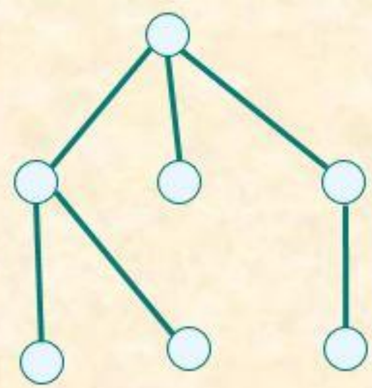
$m(k,k') =$ total number of edges between nodes of degree k and k'

$a_{k,k'} = 2$ (if $k = k'$) or $= 1$ (otherwise)

Example:



$$P(1,1) = \frac{2\times 1}{2\times 6} = \frac{1}{6}$$



$$P(1,1) = \frac{2\times 0}{2\times 6} = 0$$

Degree Correlations

Define the probability of a randomly picked node whose randomly picked neighbor has degree k as

$$P_N(k) = \sum_{k'=0}^{N-1} P(k,k') \quad \text{with } P(k,k') \text{ defined above}$$

Relationship:

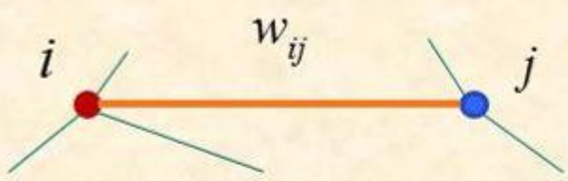
$$P(k) = \frac{\langle k \rangle}{k} \sum_{k'=0}^{N-1} P(k,k') = \frac{\langle k \rangle}{k} P_N(k)$$

Theorem:

A network is degree uncorrelated if and only if

$$P(k,k') = P_N(k) P_N(k') \quad \text{for all } k, k'$$

Weighted Undirected Networks



Static Networks:

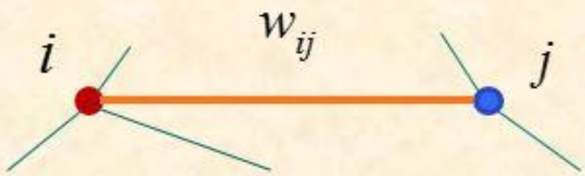
Weight of edge e_{ij} $\qquad w_{ij} = \sqrt{k_i k_j}$

Example:

$$w_{ij} = \sqrt{4 \times 3} = \sqrt{12} = 2\sqrt{3}$$

Weighted Undirected Networks

Dynamic Networks:



Weight of Edge e_{ij}: $$w_{ij} = \sum_{h=1}^{H} \delta_i^h \delta_j^h \quad w_{ii} = 0 \quad w_{ij} = w_{ji}$$

i, j: nodes

h : index of action

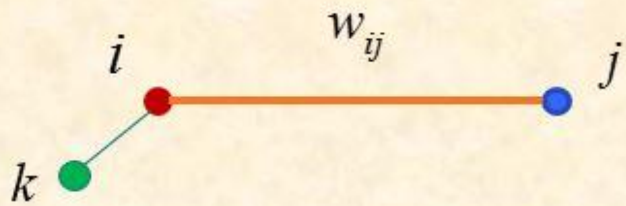
$\delta_i^h = 1$, if node i has action h $(h = 1, \ldots, H)$

$= 0$, otherwise

Total coupling strength of node i: $S_i = \sum_k w_{ik}$

Weighted Undirected Networks

Example:



Nodes: i, j, k

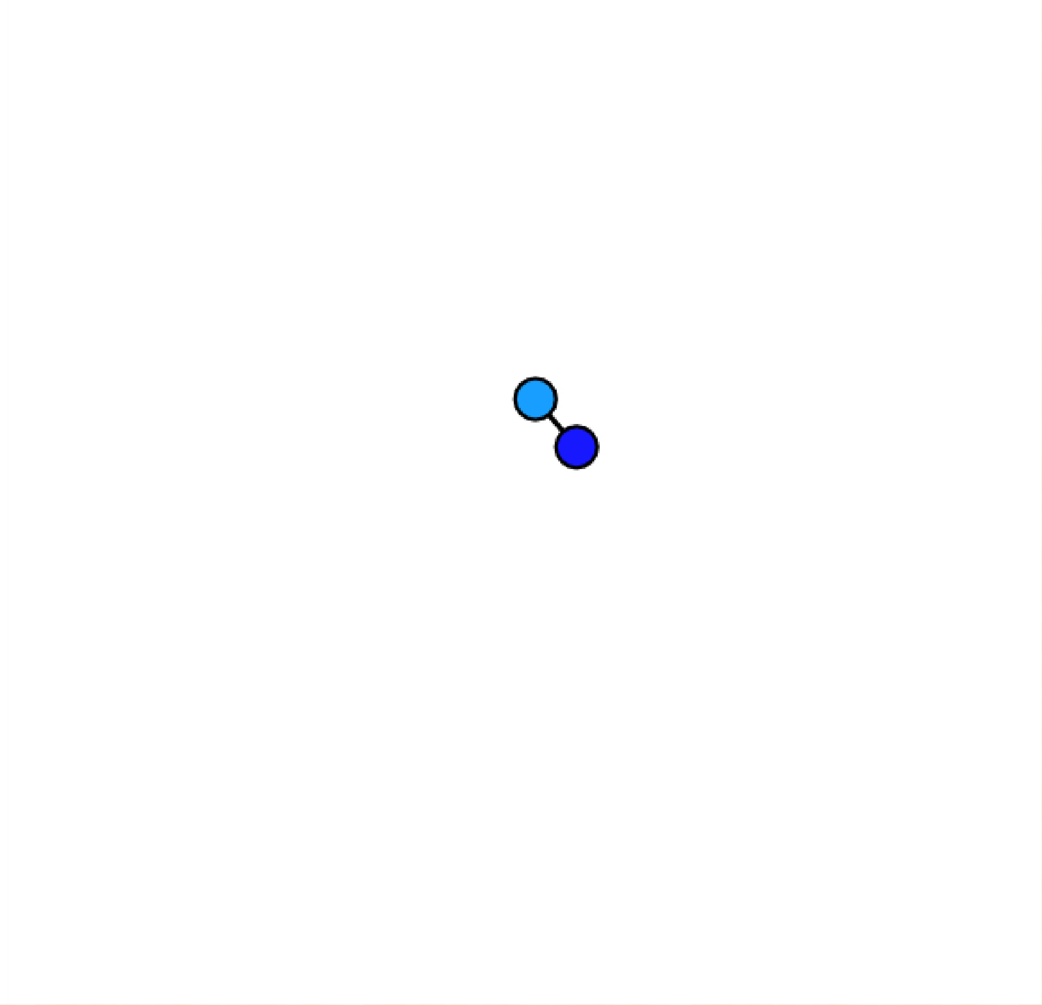
Actions: Send and Receive E-mails

i sent 2 e-mails to j: indexed $h = 1, 2$ (so, $H = 2$)
j received the first e-mail: indexed $h = 1$ but it rejected the second e-mail
i sent 1 e-mail to k (the second e-mail): indexed $h = 2$ (so, $H = 2$)
k received the (second) e-mail from i

→ $\delta_i^1 = 1,\ \delta_i^2 = 1;\ \ \delta_j^1 = 1,\ \delta_j^2 = 0;\ \ \delta_k^1 = 0,\ \delta_k^2 = 1$

→ $w_{ij} = \sum_{h=1}^{H} \delta_i^h \delta_j^h = 1\times 1 + 1\times 0 = 1 \quad w_{ik} = \sum_{h=1}^{H} \delta_i^h \delta_k^h = 1\times 0 + 1\times 1 = 1$

Evolving Networks



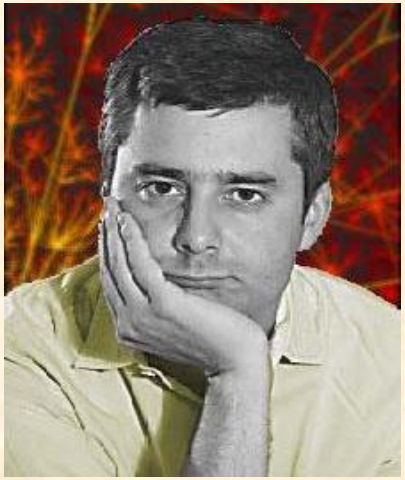
Reference

BREAK

10 minutes

Scale-Free Networks

"Emergence of scaling in random networks"
Science 286: 509 (1999)



A.-L. Barabási

R. Albert

(Norte Dame University)

BA Scale-Free Network Model

Start with a connected network having $m_0 \geq 1$ nodes

(i) Add new nodes:

Add 1 new node into the network:

This node is connected to m ($m \leq m_0$) existing nodes simultaneously

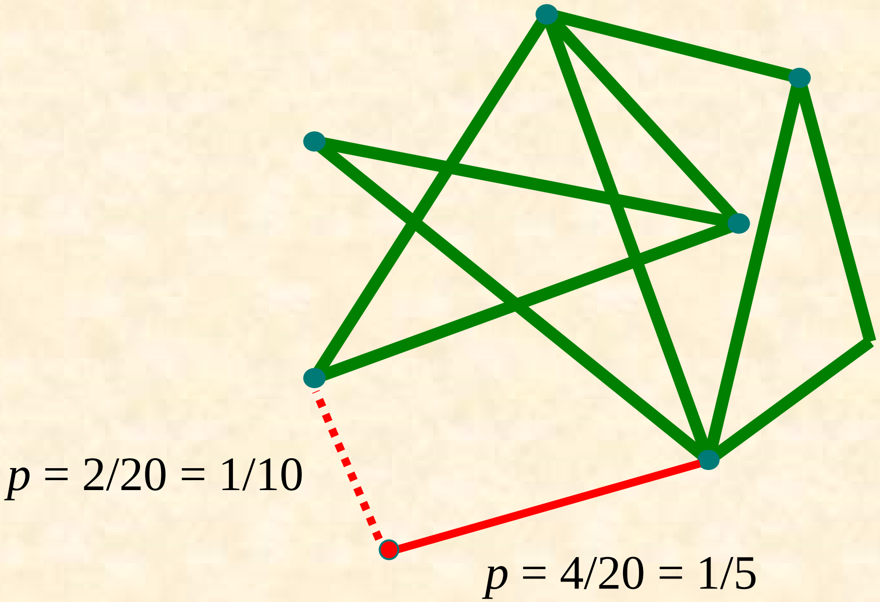
(ii) Add new edges:

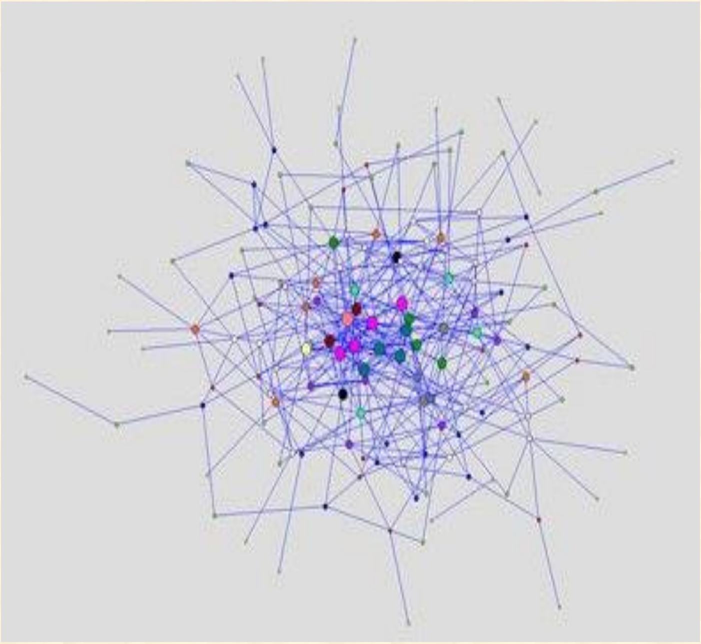
The way to add the m new edges into the network:

Every existing node is to be chosen with probability

$$\Pi(k_i) = \frac{k_i}{\sum_l k_l}$$

Scale-Free Network Modeling





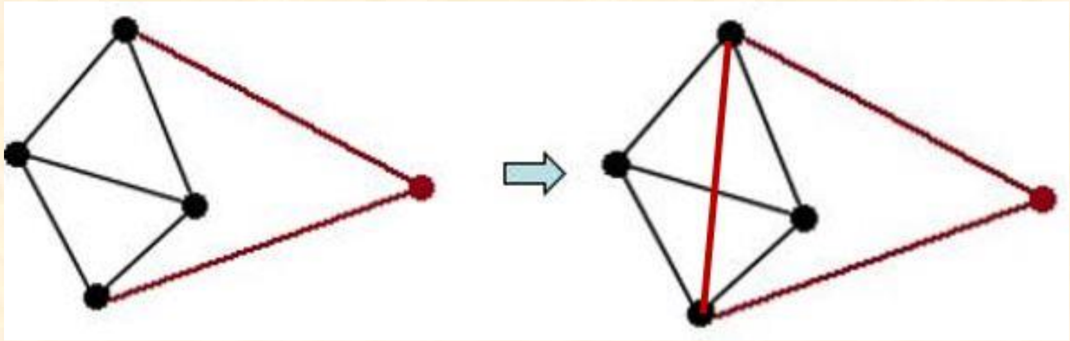
Henneberg Network

Step 1. Start with a small connected network.

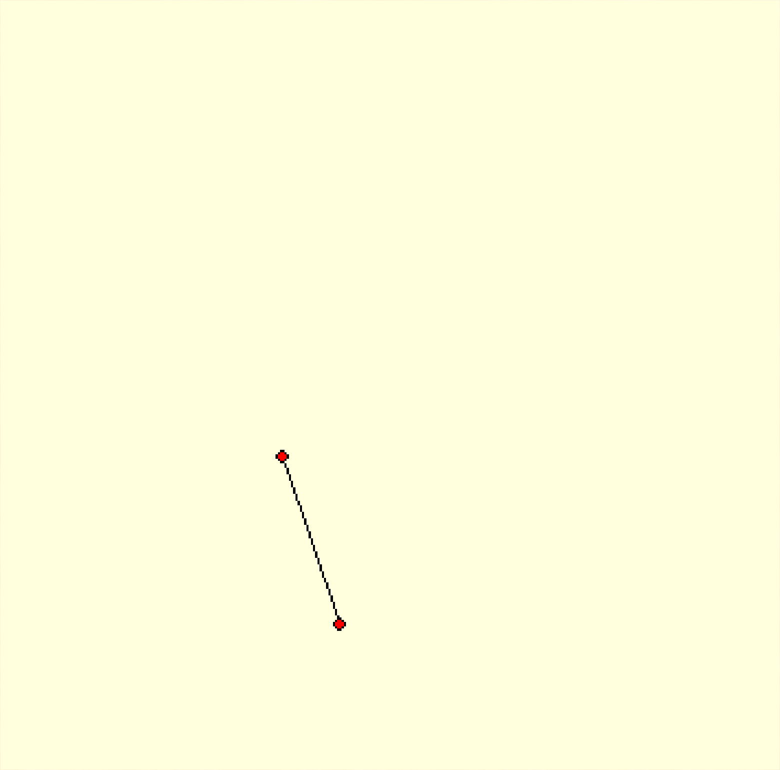
Step 2. Node addition: Add a new node to the existing network.
Connect it to any pair of existing nodes.

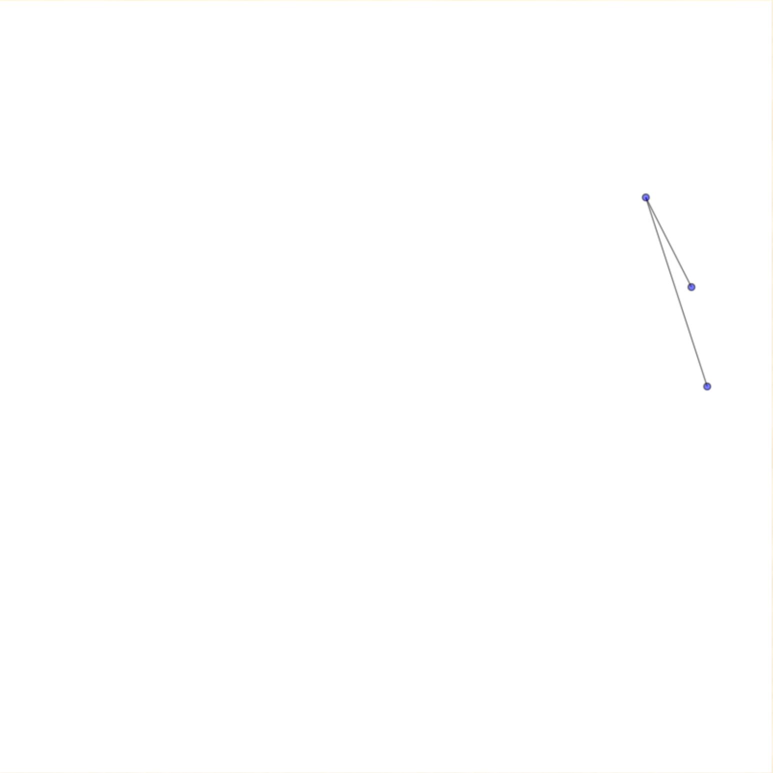
Step 3. Edge addition: If the pair of old nodes were connected,
then do nothing; otherwise, connect them with a new edge.

Step 4. Repeat Steps 2 and 3 till the network has the desired size.



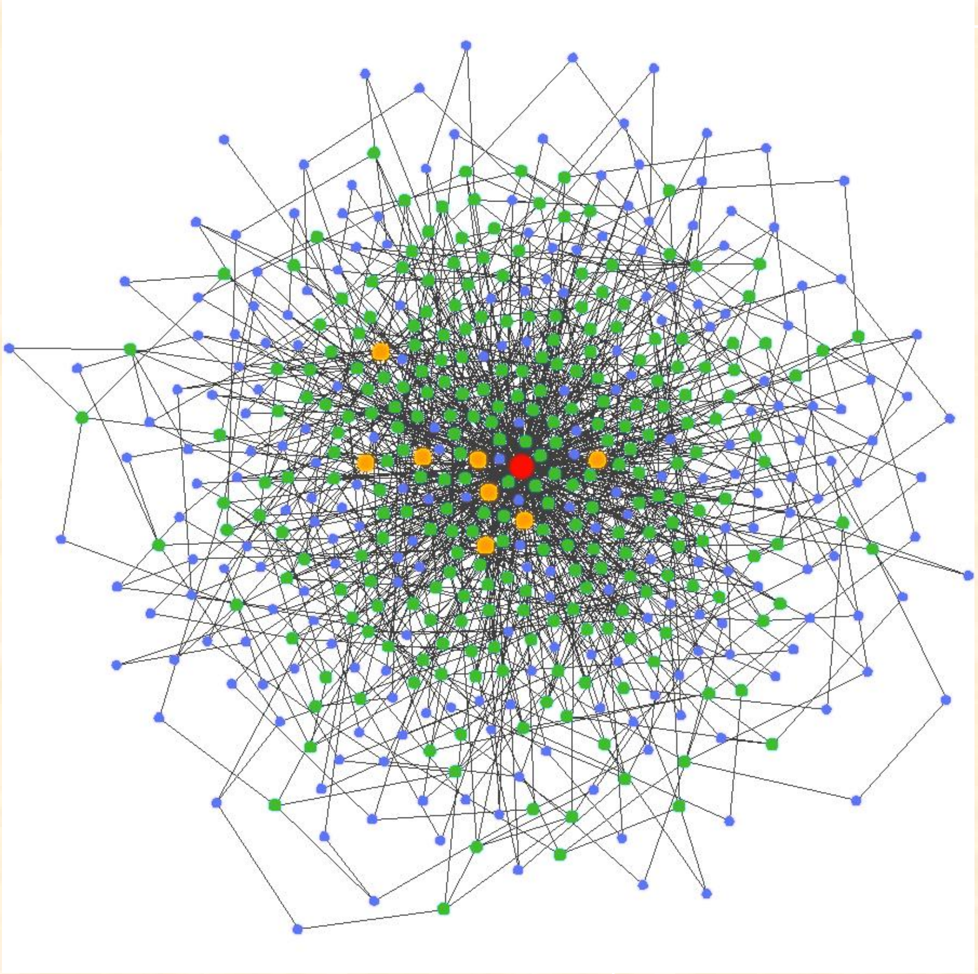
Facebook is Scale-Free Network





Different ways to displace

Scale-Free Networks

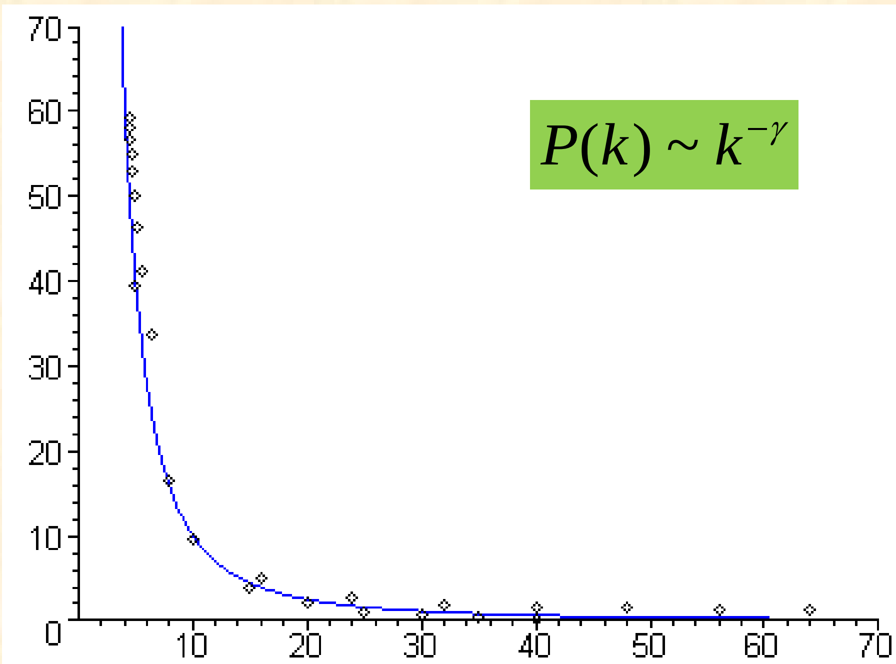


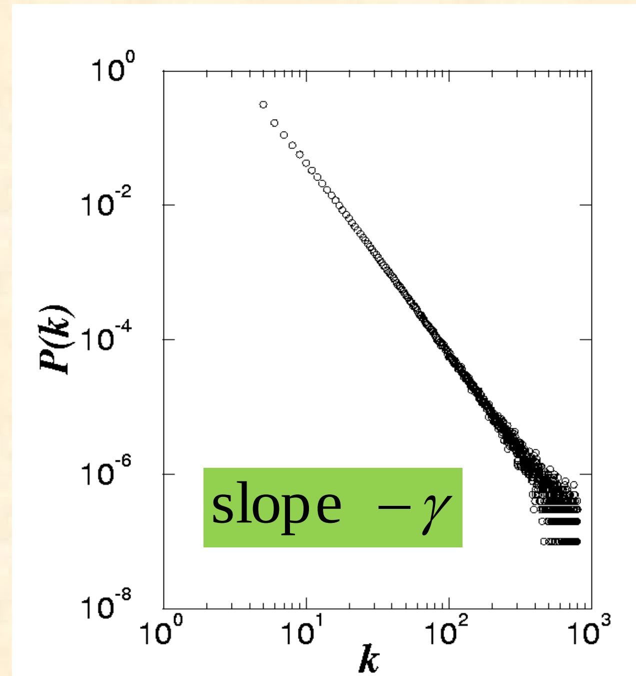
Features:

- **Connectivity:**
  in power-law form
- **Heterogeneous nature:**
  very few nodes have many links but most nodes have very few links
- **Growing**

(Hawoong Jeong)

Scale-Free Network Model

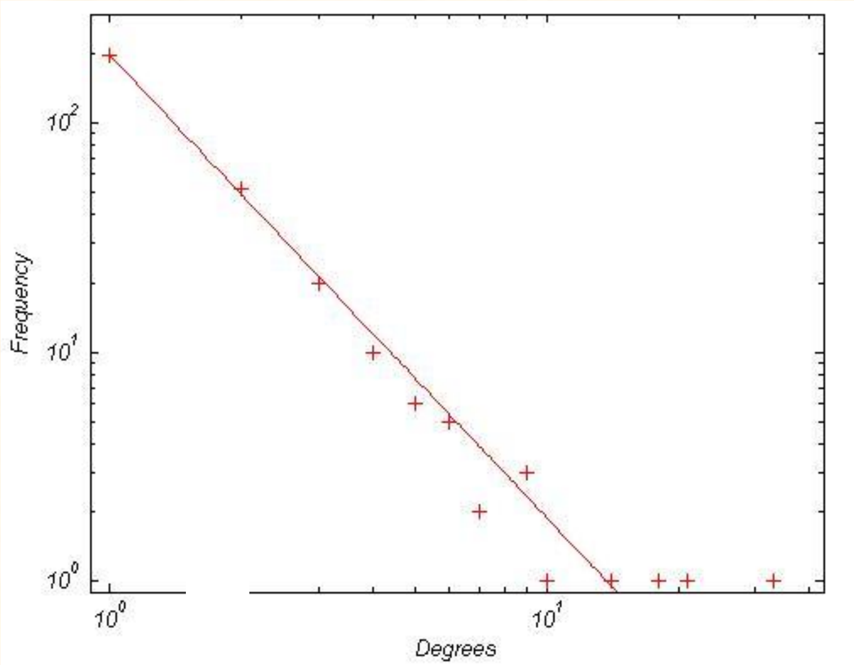


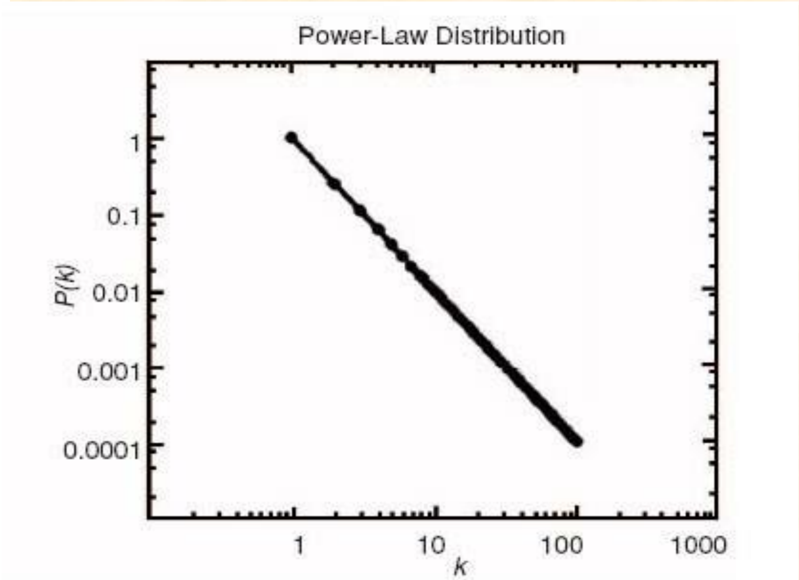


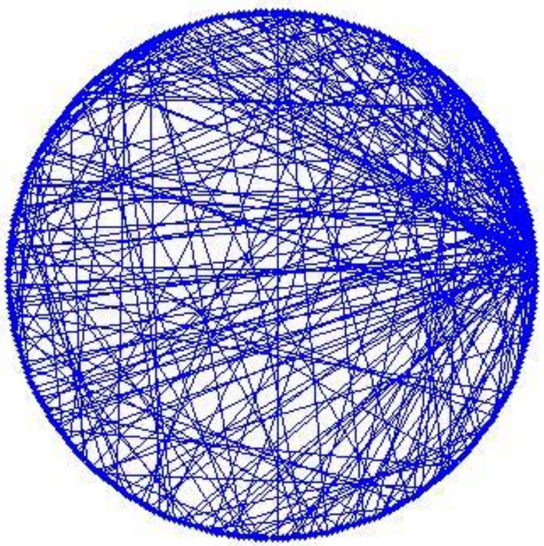
Power-Law

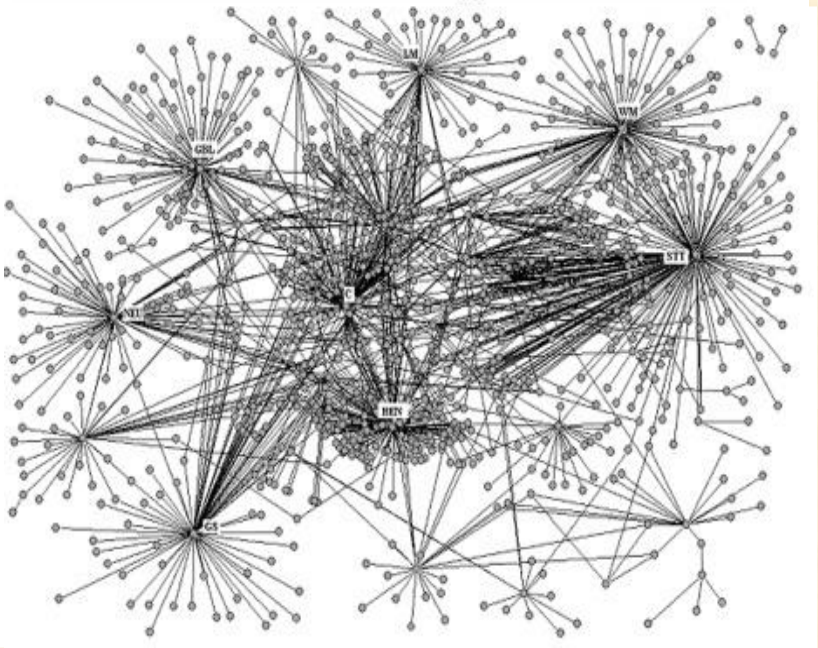
Log-Log Plot

Scale-free: It is independent of the network size



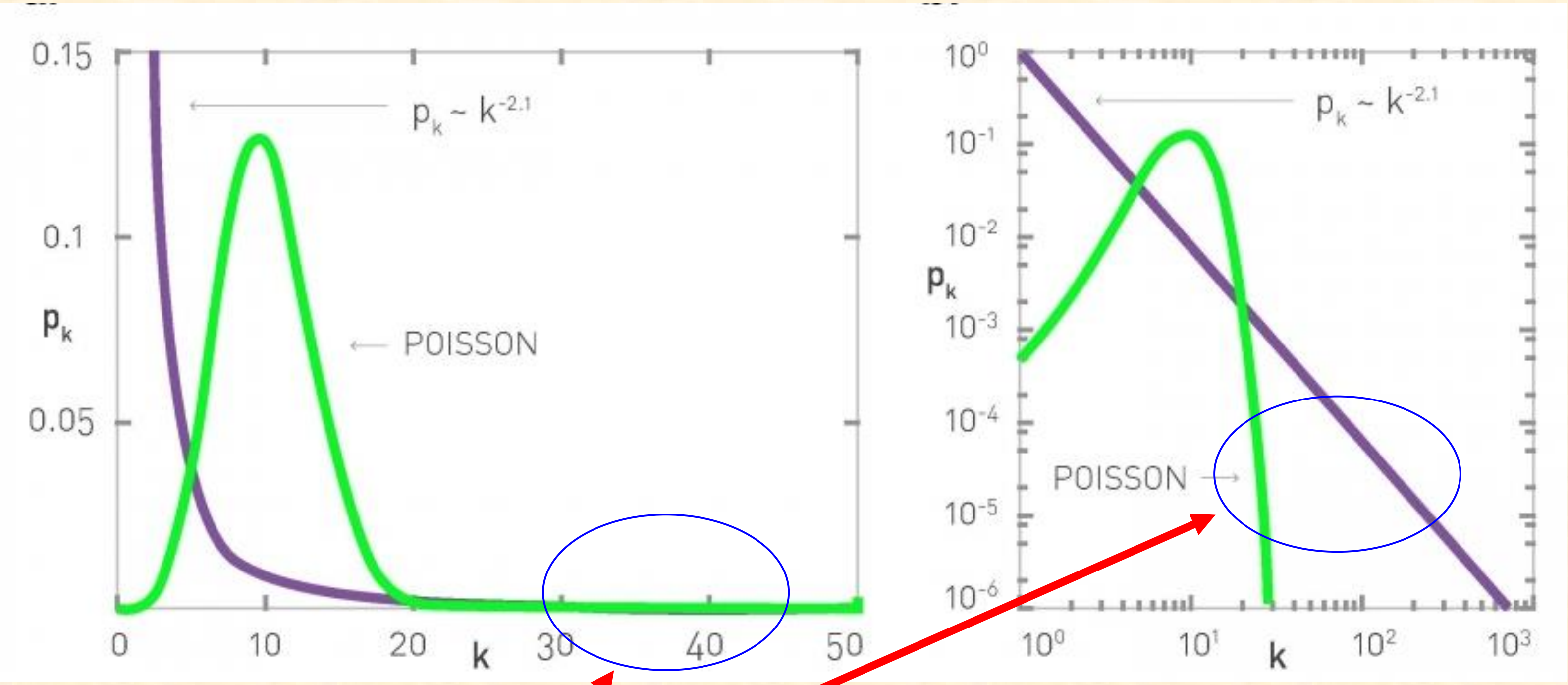






Different ways to displace

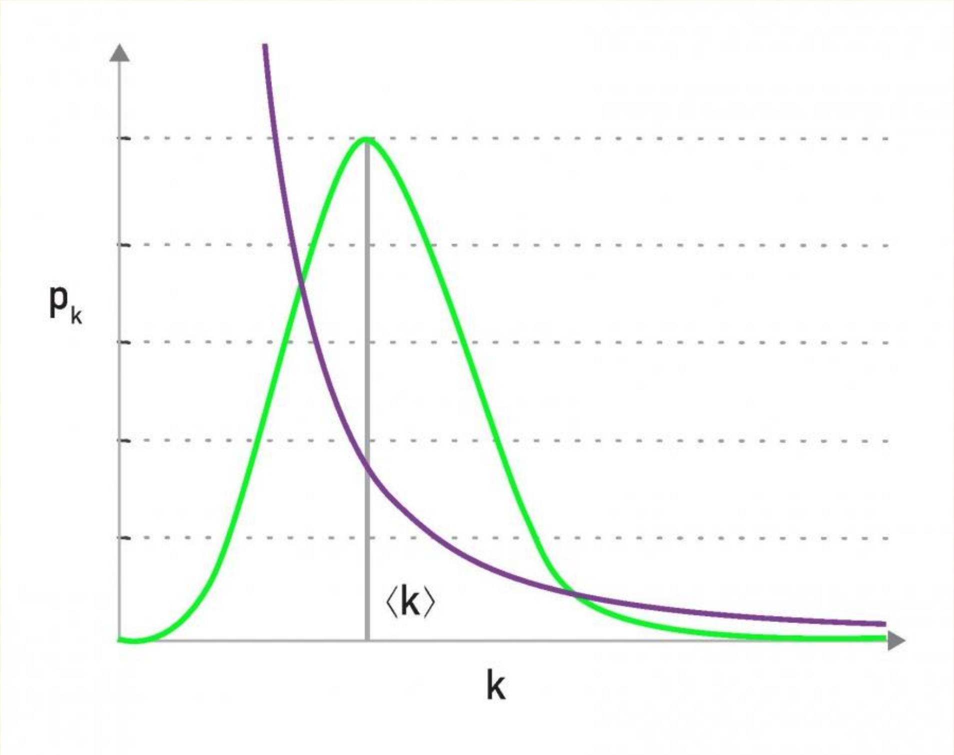
Why Log-Log Plot ?



(A.-L. Barabasi)

To see subtle details

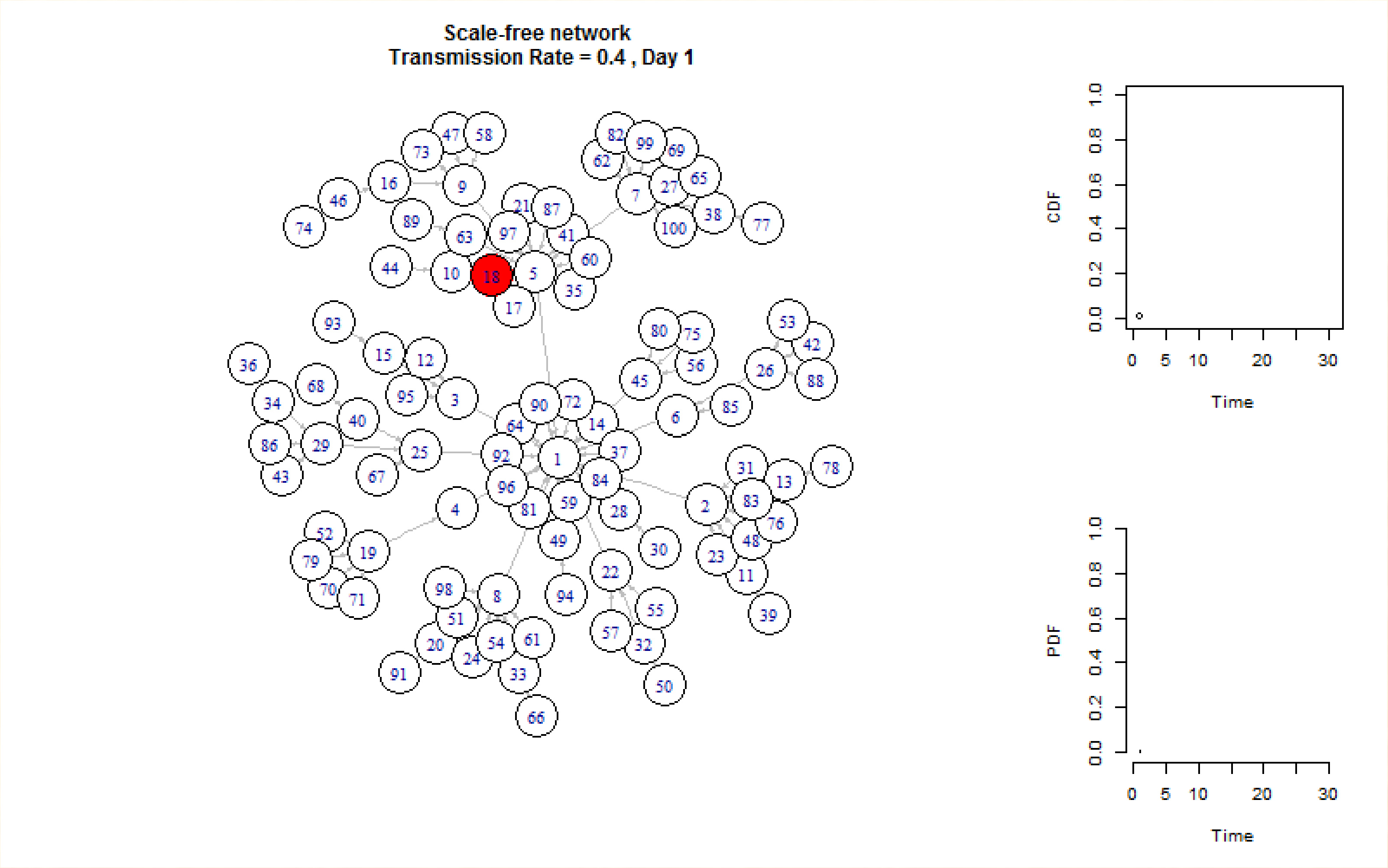
Comparison



Random Networks: $k = \langle k \rangle \pm \sqrt{\langle k \rangle}$ → A randomly picked node has degree around $\langle k \rangle$

Scale-free Networks: $k = \langle k \rangle \pm \infty$ → A randomly picked node is unlikely with degree $\langle k \rangle$

Scale-Free Network



How to derive the power law ? – A heuristic

Let i be a node, which was added to the network at time t_i

Suppose this node has degree k_i when it is being picked up at time $t \geq t_i$

Imaging that the node degree k_i is a continuous variable, so the rate of change of k_i is proportional to k_i Therefore,

$$\frac{\partial k_i}{\partial t} = ck_i = ak_i / \sum_j k_j = a\Pi(k_i) \quad (a := c\sum_j k_j \text{ is parameter})$$

Since the new node brings m edges in, the change of its degree is m, one has $a = m$ Also, ignoring initial nodes, $\sum_j k_j = 2mt$

Hence, $$\frac{\partial k_i}{\partial t} = \frac{mk_i}{2mt} = \frac{k_i}{2t} \qquad \left(\text{Recall: } \int \frac{dk}{k} = \int \frac{dt}{t} \rightarrow \ln(k) = \ln(t) + C\right)$$

Solving this equation, with the initial condition that node i was added to the network at t_i with connectivity $k_i(t_i) = m$, yields

$k_i(t) = m\sqrt{\frac{t}{t_i}}$ namely $t_i = \frac{m^2 t}{k_i^2}$

How to derive the power law ? – A heuristic 2/2

On the other hand, if $k_i < k$ then $t_i = \frac{m^2 t}{k_i^2} > \frac{m^2 t}{k^2}$ hence

$$P(k_i(t) < k) = P\left(t_i > \frac{m^2 t}{k^2}\right)$$

Assuming the time intervals $\{t_i\}$ are equally distributed, one has

$P(t_i) = \frac{1}{t}$ for all t_i on time interval $[0, t]$, so

$$P\left(t_i > \frac{m^2 t}{k^2}\right) = 1 - P\left(t_i \leq \frac{m^2 t}{k^2}\right) = 1 - \frac{m^2 t}{k^2} \cdot P(t_i) = 1 - \frac{m^2 t}{k^2}\frac{1}{t} = 1 - \frac{m^2}{k^2}$$

Consequently

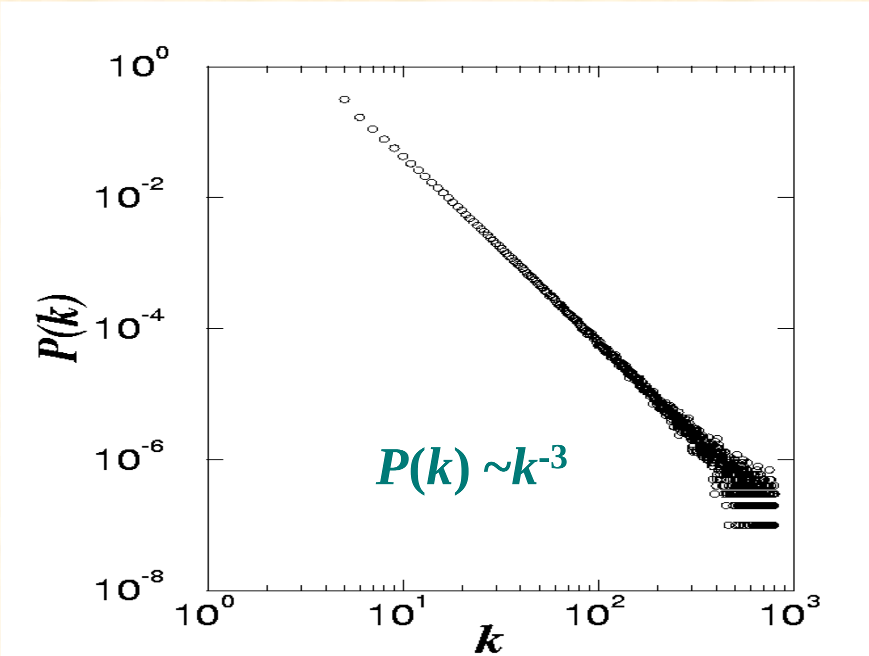
$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{\partial P\left(t_i > \frac{m^2 t}{k^2}\right)}{\partial k} = \frac{\partial\left(1 - \frac{m^2}{k^2}\right)}{\partial k} = 2m^2 k^{-3}$$

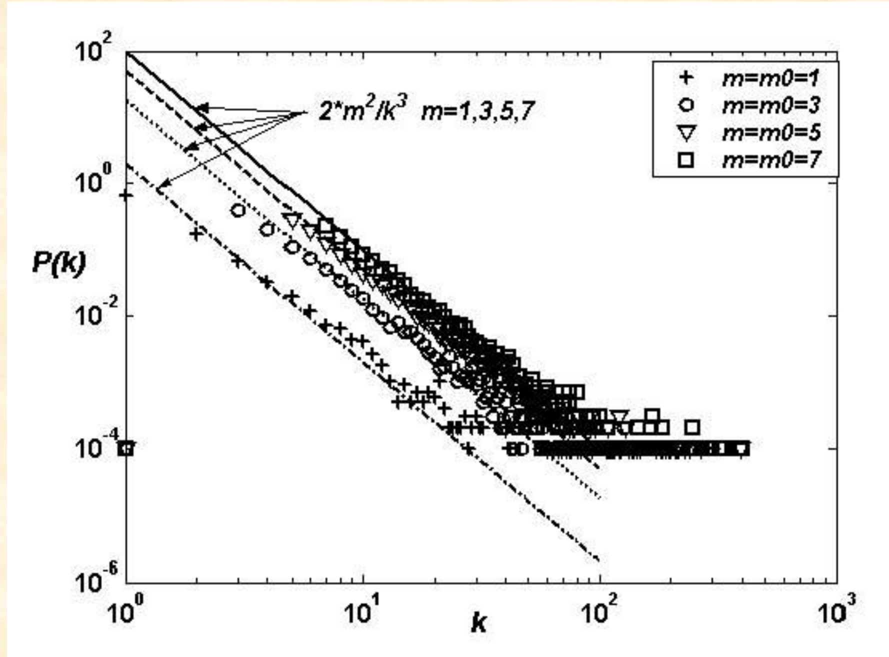
This is a power law of the form $b \cdot k^{-\gamma}$ with $b = 2m^2$, $\gamma = 3$

This power-law is a probability after normalized by choosing b

BA Model

$$P(k) = 2m^2 k^{-3}$$





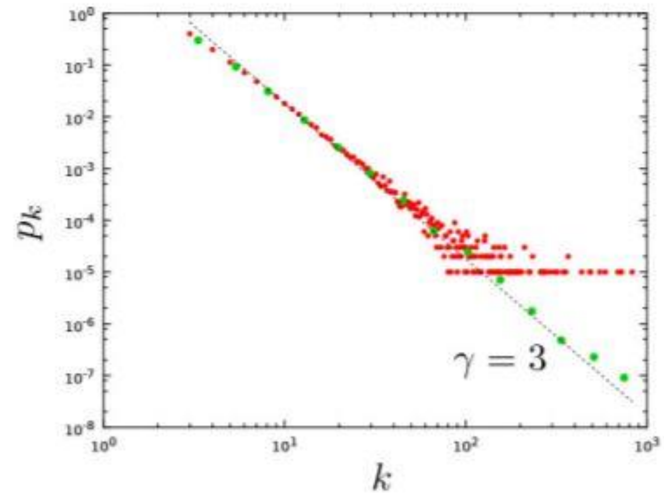
A.-L. Barabási, R. Albert, *Science* 286, 509 (1999)

Log-Binned Curves

Data binning (or bucketing) is a data pre-processing technique used to reduce the effects of minor observation errors.

The original data values which fall in a given small interval, a bin, are replaced by a value representative of that interval, often the central value.

It is a form of quantization.



The degree distribution

The degree distribution of a network generated by the Barabási-Albert model. The plot shows p_k for a single network of size N=100,000 and m=3. It shows both the linearly-binned (red symbols) as well as the log-binned version (green symbols) of p_k. The straight line is added to guide the eye and has slope γ=3, corresponding to the resulting network's degree distribution.

Comparison 1

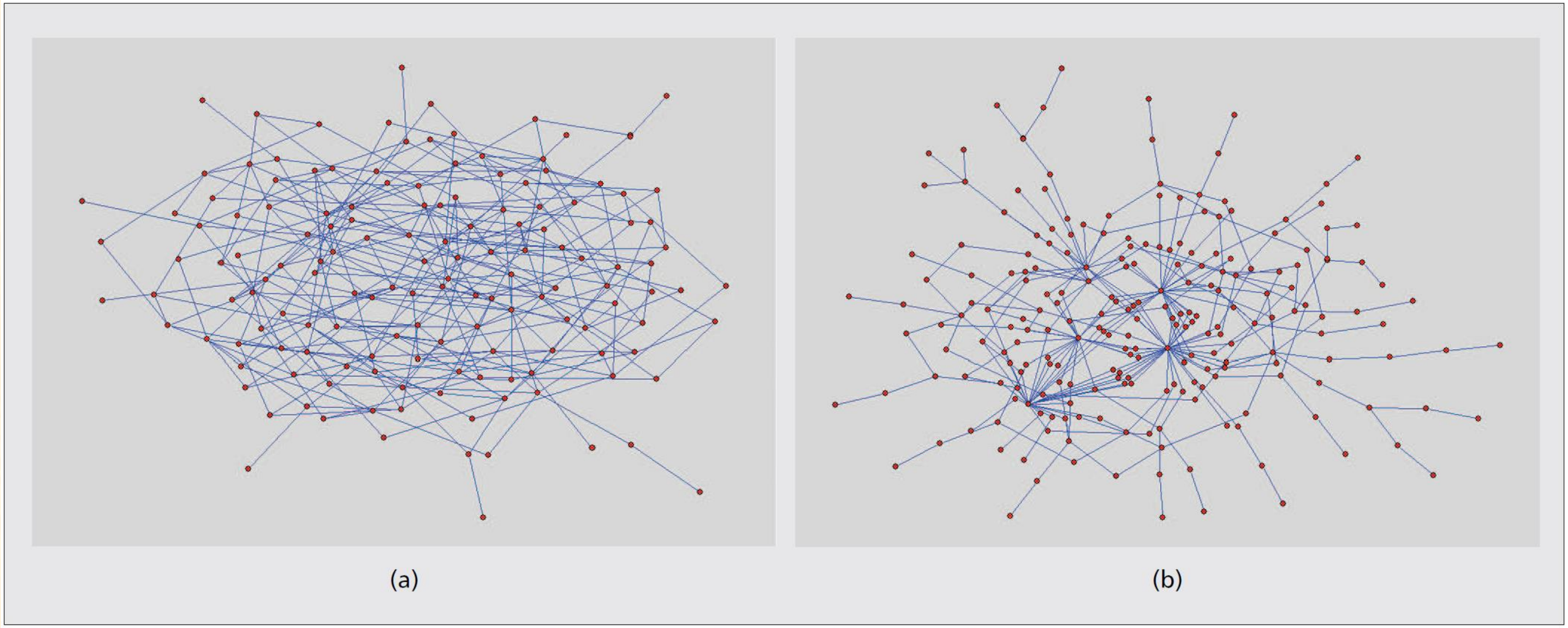


Figure 1. *Examples showing the differences between a random network and a scale-free network: a) a 200-node Erdos-Rényi random network; b) a 200-node scale-free network.*

Comparison 2

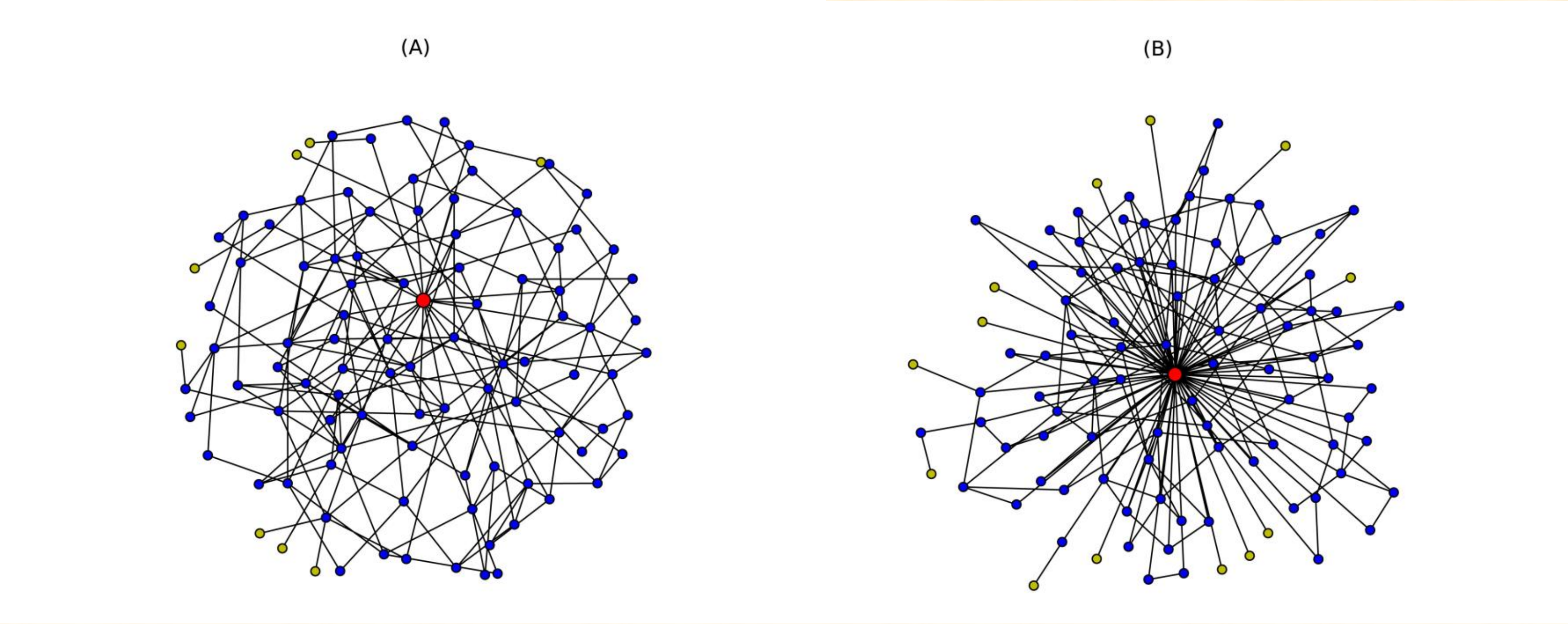
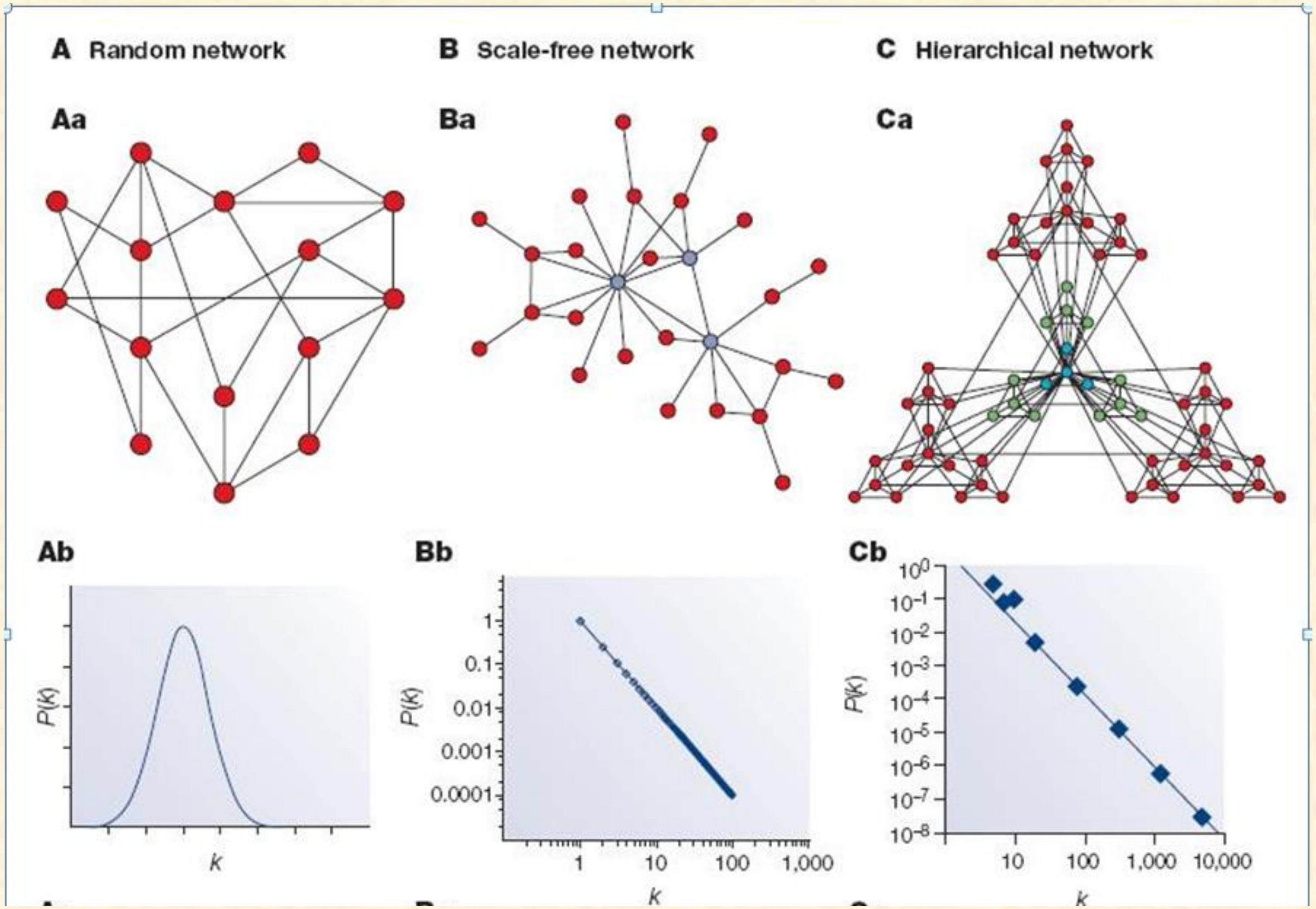


Figure 2. Homogeneous vs Heterogeneous networks, both with a Center

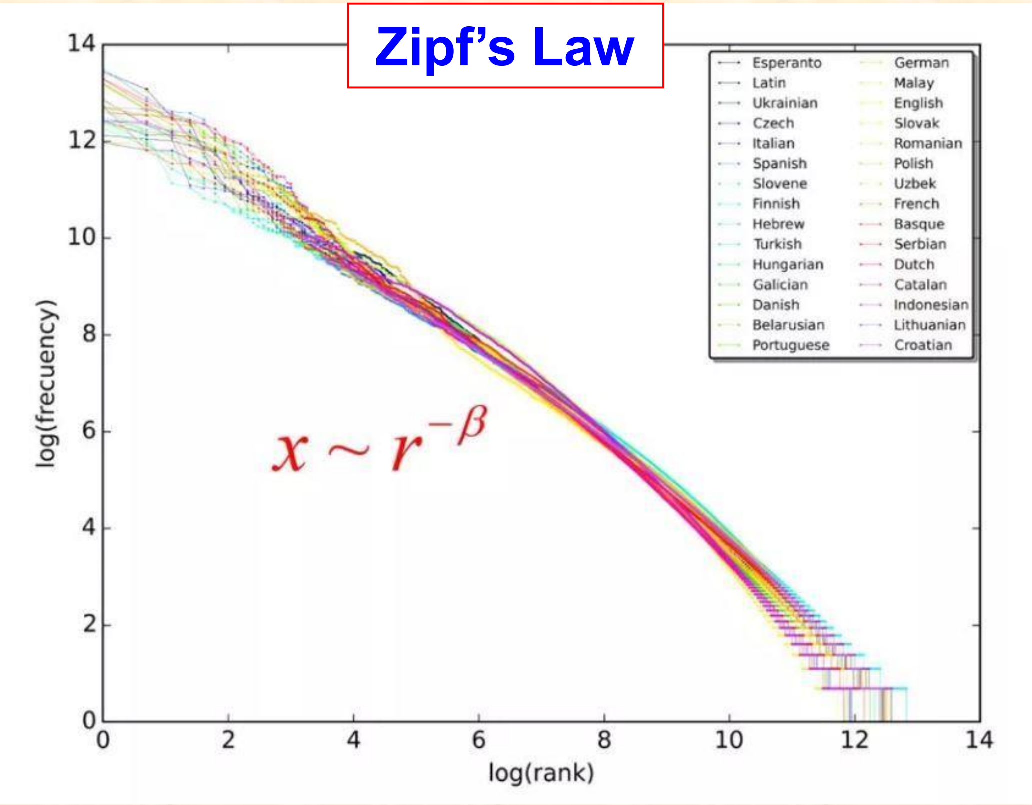
They have the same power-law distribution

Comparison 3



A Historical Remark (1)

Prof George Zipf from Harvard University studied the frequencies of words In English (and languages) in 1932-1935



Few words occur very frequently;
Most words occur only occasionally

A Historical Remark (2)

In 1965, Dr. Derek John de Solla Price (1922–1983), Professor of Science History at Yale University, consider the following network:

(o) **Initialization:** Start from $m_0 \geq 1$ isolated nodes

(i) **Growth:** At every step, add one new node to the existing network, which brings m $(m \leq m_0)$ directed edges, each pointing to one existing node; thus, the existing node receives one more "in-degree" (namely, its in-degree is increased by 1)

(ii) **Preferential Attachment:** For every existing node i with in-degree k_i^{in}, the above new node will connect to it according to the following probability:

$$\Pi_i = \frac{k_i^{in} + a}{\sum_j (k_j^{in} + a)} \qquad (a > 0 \text{ is a constant parameter})$$

It was then proved that the node-degree distribution of the above Price model is given by $P(k) \sim (k^{in})^{-\gamma}$ with $\gamma = 2 + a/m$.

Nevertheless, the BA model is much more influential in recent years

Power-Law Distributions: Scale-Free

Consider a probability distribution function $f(x)$.

If, for any given constant a, there is a constant b such that the following "**scale-free**" property holds:

$$f(a\,x) = b\,f(x)$$

then, with the assumption that $f(1)\,f'(1) \neq 0$, the function $f(x)$ is uniquely determined by a **power-law**:

$$f(x) = f(1)\,x^{-\gamma} \qquad \gamma = -f'(1)/f(1)$$

BA Model:

$$P(k) = 2m^2 k^{-3}$$

Note: The reverse may not be true

Power-Law Distributions: Scale-Free

Proof. In $f(a\,x) = b\,f(x)$, let $x = 1$. Then, $f(a) = b f(1)$, so $b = f(a) / f(1)$, giving

$$f(ax) = \frac{f(a)\,f(x)}{f(1)}$$

Since this equality holds for arbitrary a and x, one may also consider a as a variable and take a derivative w.r.t. a :

$$\frac{df(ax)}{d(ax)}\frac{d(ax)}{da} = \frac{f(x)}{f(1)}\frac{df(a)}{da}$$

Letting $a = 1$ gives

$$x\frac{df(x)}{d(x)} = \frac{f'(1)}{f(1)}\,f(x)$$

Recall: $\int \frac{df}{f} = \int \frac{dx}{x} \rightarrow \ln(f) = \ln(x) + C$

which has a unique solution

$$f(x) = f(1)\,x^{-\gamma} \qquad \gamma = -f'(1)/\,f(1)$$

BA Scale-Free Network Model

Theorem:

Average Path Length: $$L \sim \frac{\ln N}{\ln \ln N}$$

Clustering Coefficient: $$C = \frac{m^2(m+1)^2}{4(m-1)}\left(\ln\left(\frac{m+1}{m}\right) - \frac{1}{m+1}\right)\frac{(\ln t)^2}{t}$$

Degree Distribution: $$P(k) = 2m^2 k^{-3}$$

B. Bollobas and O. Riordan, Mathematical results on scale-free random graphs, In S. Bornholdt and H.G. Schuster (eds.), Handbook of Graphs and Networks: From the Gerome to the Internet, Wiley-VCH, 2003, pp. 1-34.

R. Cohen and S. Havlin, Scale-free networks are ultrasmall, Phys. Rev. Lett., 86: 3682-3685, 2003.

Extended BA (EBA) Model

The BA Model can describe scale-free networks having a power law with $r = 3$, but many real-world networks do <u>not</u> satisfy $r = 3$

EBA model (Albert and Barabasi, 2000)

Start with a connected network with $m_0 \geq 0$ nodes.

(i) Add new nodes (incremental growth):

With probability p, one new node is added into the network

(ii) Re-wiring:

With probability q, m ($m \leq m_0$) edges are rewired

(iii) Add new edges (preferential attachment):

m ($m \leq m_0$) new edges are added into the network with probability

$$\Pi(k_i) = \frac{k_i}{\sum_l k_l}$$

EBA Model

In this model, $0 < p \le 1$ and $0 \le q < 1 - p$

If $q < \min(1-p, (1-p+m)/(1+2m)),$ then the degree distribution of nodes follows a shifted power-law:

$$P(k) \sim (k + A(p,q,m) + 1)^{-\gamma}$$

where $\gamma = 1 + B$ (typically, $2 < \gamma < 3$), and

$$A(p,q,m) = (p-q)\left(\frac{2m(1-q)}{1-p-q} + 1\right)$$

$$B(p,q,m) = \frac{2m(1-q) + 1 - p - q}{m}$$

Fitness Model

Recall: During the growth of the BA model, the degree of a node is changing by

$$k_i(t) = \sqrt{\frac{t}{t_i}}$$

where $k_i(t)$ is the degree of node i at time t and t_i is the instant at which node i is being added into the network.

This implies that the older a node, the higher its degree

In real life, however, this may not be true. For example:

- A teenager can have more social connections than an elder man
- A new website in WWW can receive more links than an old one
- ……

Fitness Model

1. Growth: Start from a small network of size $m_0 \geq 1$ and introduce one new node to the existing network each time, and with probability $\rho(\eta)$ this node is given a fitness value, $0 < \eta_i < 1$

2. Preferential Attachment: The new node is connected to $1 \leq m \leq m_0$ existing nodes, each is according to the probability

$$\Pi_i = \frac{\eta_i k_i}{\sum_{j=1}^{n} \eta_j k_j}$$

where $n = m_0 + t - 1$ is the total number of existing nodes at the $(t\text{-}1)st$ step of the process.

Basic Properties: Similar to the BA model

Main difference: Preferential attachment is proportional to both (1) node degree; (2) fitness value

Bianconi and Barabási, Phys. Rev. Lett. (2001)

A Simple Model
with power-law degree distribution

1. Start with nothing: no nodes, no links

2. At each time, with probability p, one new node is added to the existing network

3. For every pair of unconnected nodes, with probability $q = 1 - p$, one edge is added to connect them

Theorem: The degree distribution of the above network model has a power-law form, $P(k) \sim k^{-\gamma}$, where $\gamma = 1 + q^{-1}$

Corollary: If $1/2 < q < 1$ then $2 < \gamma < 3$

Aiello, Chung and Lu (2002)

Scale-Free Networks Debate 1/2

Criticism

SIAM Review

< Previous Article

Volume 51, Issue 4

Abstract | References | PDF | Cited By

SIAM Rev., 51(4), 661–703. (43 pages)

Power-Law Distributions in Empirical Data

Aaron Clauset, **Cosma Rohilla Shalizi**, **and** **M. E. J. Newman**

https://doi.org/10.1137/070710111

Power-law distributions occur in many situations of scientific interest and have significant consequences for our understanding of natural and man-made phenomena. Unfortunately, the detection and characterization of power laws is complicated by the large fluctuations that occur in the tail of the distribution—the part of the distribution representing large but rare events—and by the difficulty of identifying the range over which power-law behavior holds. Commonly used methods for analyzing power-law data, such as least-squares fitting, can produce substantially inaccurate estimates of parameters for power-law distributions, and even in cases where such methods return accurate answers they are still unsatisfactory because they give no indication of whether the data obey a power law at all. Here we present a principled statistical framework for discerning and quantifying power-law behavior in empirical data. Our approach combines maximum-likelihood fitting methods with goodness-of-fit tests based on the Kolmogorov–Smirnov (KS) statistic and likelihood ratios. We evaluate the effectiveness of the approach with tests on synthetic data and give critical comparisons to previous approaches. We also apply the proposed methods to twenty-four real-world data sets from a range of different disciplines, each of which has been conjectured to follow a power-law distribution. In some cases we find these conjectures to be consistent with the data, while in others the power law is ruled out.

Rebut

Love is All You Need

Clauset's fruitless search for scale-free networks

by Albert-László Barabási, March 6, 2018

In the past weeks, I have received several requests to address the merits of the Anna D. Broido and Aaron Clauset (BC) preprint [1] and their fruitless search for scale-free networks in nature. The preprint's central claim is deceptively simple: It starts from the textbook definition of a scale-free network as a network with a power law in the degree distribution [2]. It then proceeds to fit a power law to 927 networks, reporting that only 4% are scale-free. The author's conclusion that 'scale-free networks are rare,' is turned into the title of the preprint, helping it to get maximal attention. It worked—*Quanta* magazine accepted its conclusions without reservations. After the *Atlantic* carried the article, the un-refereed preprint received a degree of media exposure that the original discovery of scale-free networks never enjoyed.

While I saw the conceptual problems with the manuscript, I was convinced that the paper must be technically proficient. Yet, once I did dig into it, it was a real ride. If you have the patience to get to the end of this commentary, you will see where it fails at the conceptual level. But, we will learn that it also fails, repeatedly, at the technical level.

2018

Scale-Free Networks Debate 2/2

Criticism

Rebut

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-08746-5 OPEN

Scale-free networks are rare

Anna D. Broido[1] & Aaron Clauset[2,3,4]

Real-world networks are often claimed to be scale free, meaning that the fraction of nodes with degree k follows a power law $k^{-\alpha}$, a pattern with broad implications for the structure and dynamics of complex systems. However, the universality of scale-free networks remains controversial. Here, we organize different definitions of scale-free networks and construct a severe test of their empirical prevalence using state-of-the-art statistical tools applied to nearly 1000 social, biological, technological, transportation, and information networks. Across these networks, we find robust evidence that strongly scale-free structure is empirically rare, while for most networks, log-normal distributions fit the data as well or better than power laws. Furthermore, social networks are at best weakly scale free, while a handful of technological and biological networks appear strongly scale free. These findings highlight the structural diversity of real-world networks and the need for new theoretical explanations of these non-scale-free patterns.

PHYSICAL REVIEW RESEARCH **1**, 033034 (2019)

Scale-free networks well done

Ivan Voitalov,[1,2,*] Pim van der Hoorn,[1,2,3,†] Remco van der Hofstad,[3,‡] and Dmitri Krioukov[1,2,4,5,§]

[1]*Department of Physics, Northeastern University, Boston, Massachusetts 02115, USA*

[2]*Network Science Institute, Northeastern University, Boston, Massachusetts 02115, USA*

nt of Mathematics and Computer Science, Eindhoven University of Technology, Postbus 513, 5600 MB Eindhoven, Netherlands

[4]*Department of Mathematics, Northeastern University, Boston, Massachusetts 02115, USA*

[5]*Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02115, USA*

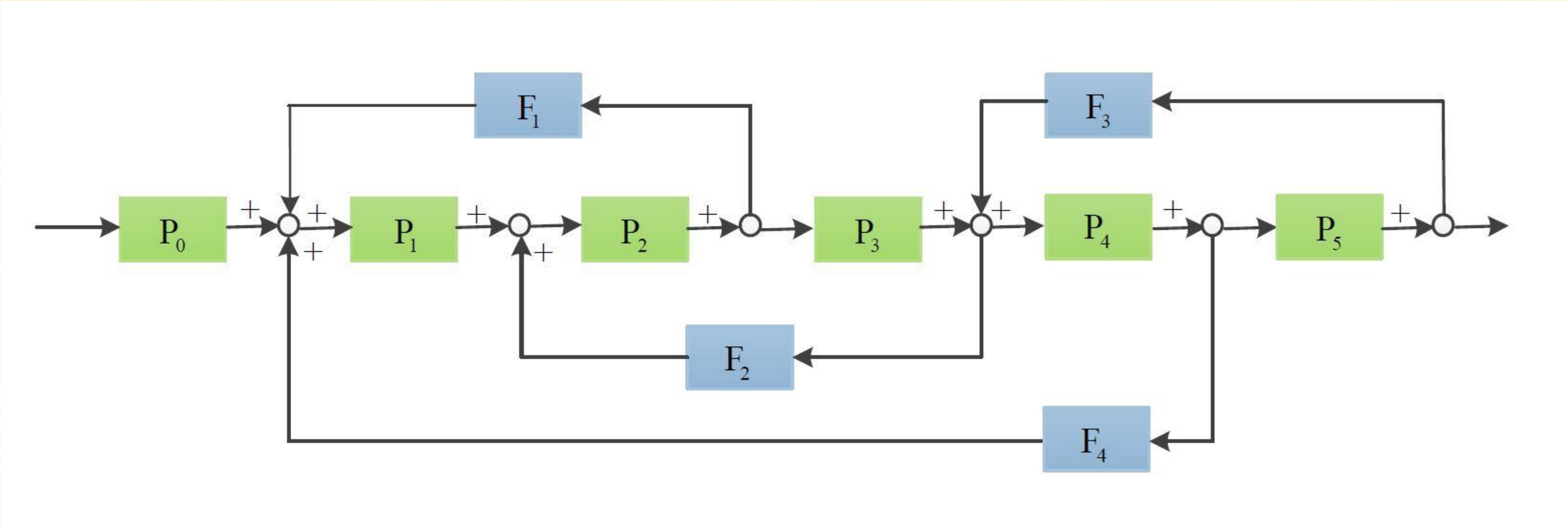
(Received 7 November 2018; revised manuscript received 26 May 2019; published 18 October 2019)

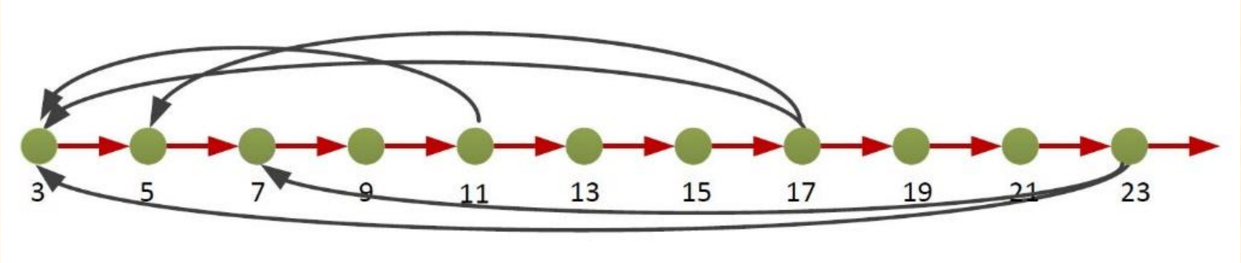
We bring rigor to the vibrant activity of detecting power laws in empirical degree distributions in real-world networks. We first provide a rigorous definition of power-law distributions, equivalent to the definition of regularly varying distributions that are widely used in statistics and other fields. This definition allows the distribution to deviate from a pure power law arbitrarily but without affecting the power-law tail exponent. We then identify three estimators of these exponents that are proven to be statistically consistent—that is, converging to the true value of the exponent for any regularly varying distribution—and that satisfy some additional niceness requirements. In contrast to estimators that are currently popular in network science, the estimators considered here are based on fundamental results in extreme value theory, and so are the proofs of their consistency. Finally, we apply these estimators to a representative collection of synthetic and real-world data. According to their estimates, real-world scale-free networks are definitely not as rare as one would conclude based on the popular but unrealistic assumption that real-world data come from power laws of pristine purity, void of noise, and deviations.

2019

Some Other Models

Assembly-Line Automation



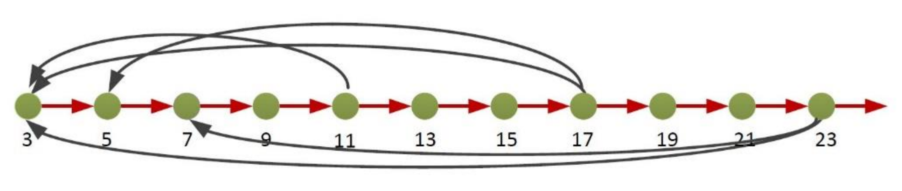


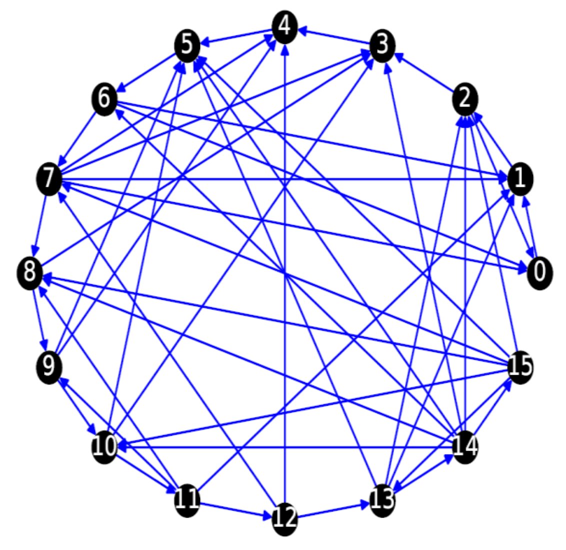
Snapback Network

Step 1. Start with a directed chain of N nodes

Step 2. From the root, move ahead node by node till the tip:

At every node, with a probability q connect back to each previous node, once and once only

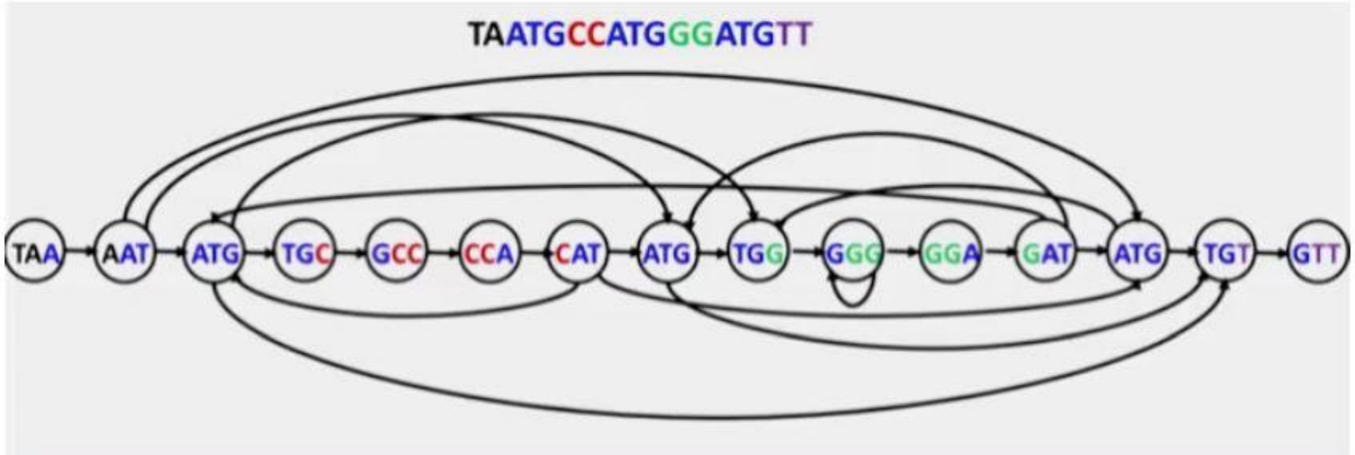


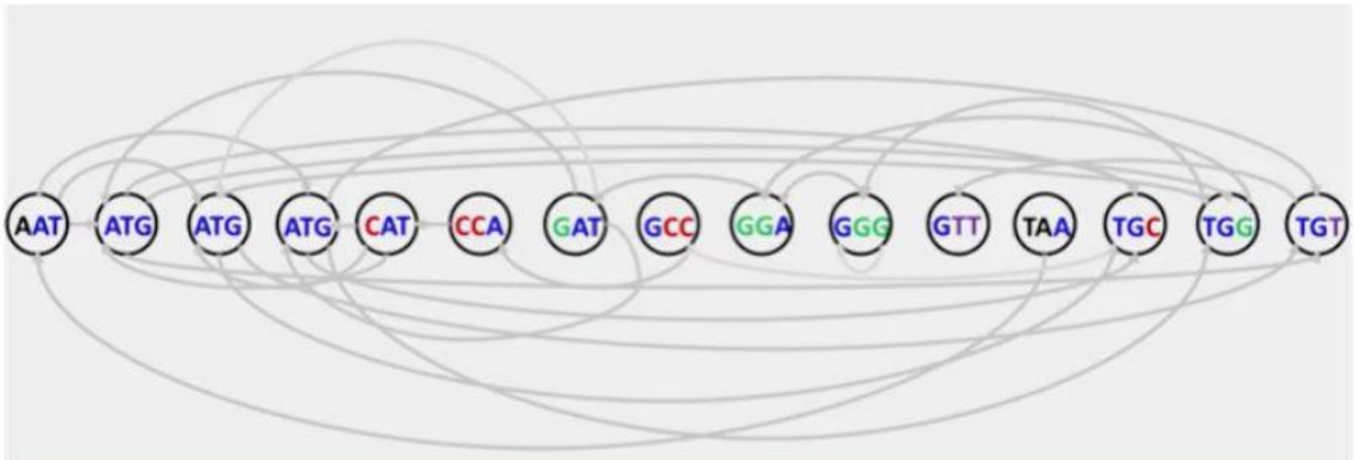


Note: The direction indicates how to operate the snapback connection

The model can be undirected if all the directions are finally removed

Genome





Nodes are arranged from left to right in lexicographic order.

End