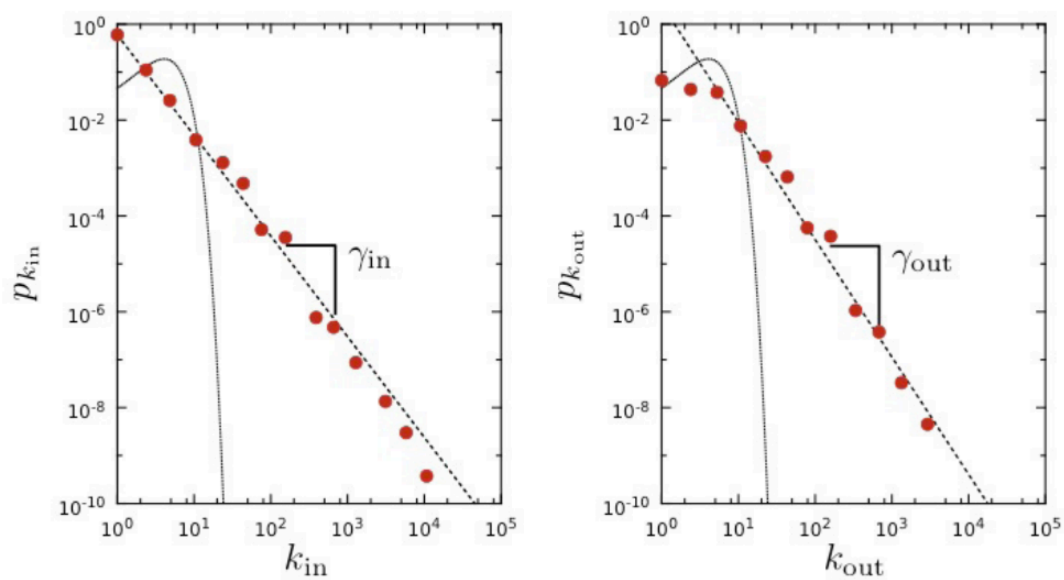
# 7-3- Preferential Attachment and the Barabasi-Albert Model

Complex Network Analysis Course
Reza Rezazadegan
Shiraz University, Spring 2025
https://dreamintelligent.com/complex-network-analysis-2025/

## Scale-Free Networks



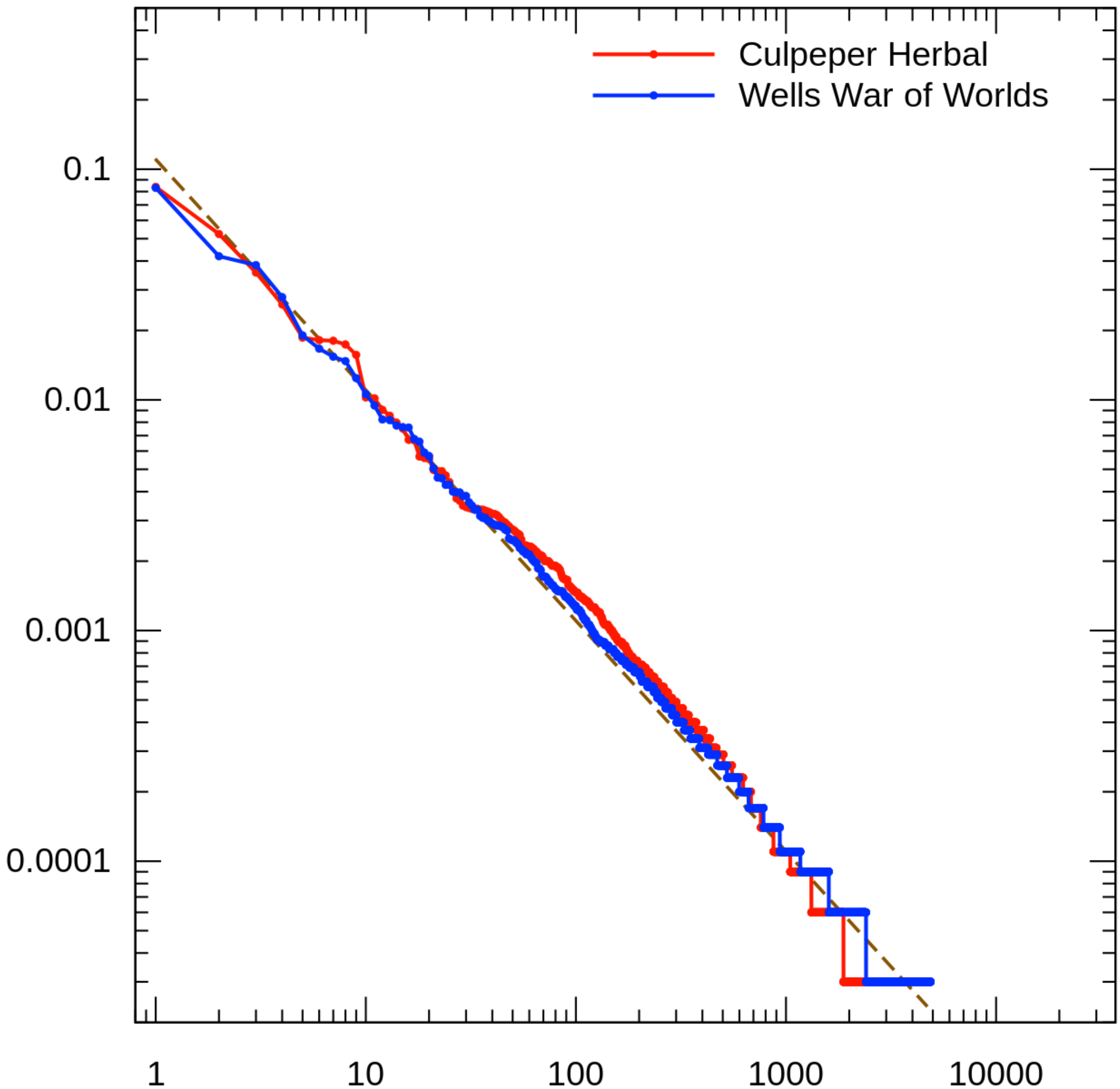

The distribution of in-degree and out-degree for the Web. The structure of the Web network can be obtained, or at least sampled, using a web crawler. Instead of the Poisson distribution of a random network (dotted curve), the data fits the power-law distribution which, in the log-log plot, takes the form of a line. Source: Barabasi, Network Science

If a network's degree distribution follows the the power-law distribution, $p(k) \sim k^{-\gamma}$, then we call it a *scale free network*. The constant $\gamma$ is called the *exponent* of the network. "Scale" here is in the statistical sense and refers to the second moment (or equivalently standard deviation) of the degree distribution, as we will see. Put differently, if we scale the independent variable, $k \to ck$, then $p(k)$ scales accordingly. For the Web, as a directed network we have $\gamma_{in} = 2.1, \gamma_{out} = 2.45$. For real scale-free networks we have $2 < \gamma < 3$.

In the power-law distribution, we have $p(k) = Ck^{-\gamma}$ where $C, \gamma > 0$ are constants. Since the probabilities must add up to one, we have $C = \frac{1}{\sum_k k^{-\gamma}}$.
Note that for for $k = 0$ this diverges, thus isolated nodes have to be taken into account separately.

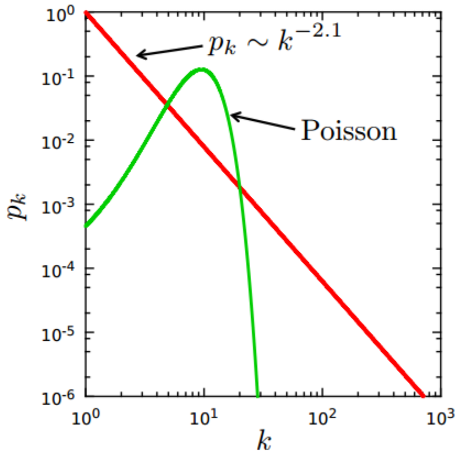
Scale-free networks can be approximately studied using the continuous power-law distribution by assuming that degree can take any value. In this regime, $p(k) = Ck^{-\gamma}$ and it follows from $\int_{k_{min}}^{\infty} p(k)dk = 1$ that $p(k) = (\gamma - 1)\, k_{min}^{k-1}\, k^{-\gamma}$.

Vilfred Pareto discovered in 19th century that the income distribution in Italy followed a power law.
Zipf's law states that the frequency of a word in a corpus is about twice the frequency of the next word. In other words, word frequency follows a power law, too.




The frequency of words as a function of their frequency rank, in two different English books.

## Hubs in scale-free networks




Comparison of Poisson and power-law distributions, both with mean equal to 10.

For the Web with average degree of 4.1, the probability of the existence of a node with degree 100 is $10^{-30}$ in the Poisson distribution and $10^{-4}$ in the power-law distribution.
This means that hubs are much less likely in random networks, compared to scale-free networks.

# Estimating the degree of the largest hub

"Largest" here means that there is at most one node with degree higher than $k_{max}$, i.e.

$$\int_{k_{max}}^{\infty} p(k)dk = 1/N.$$

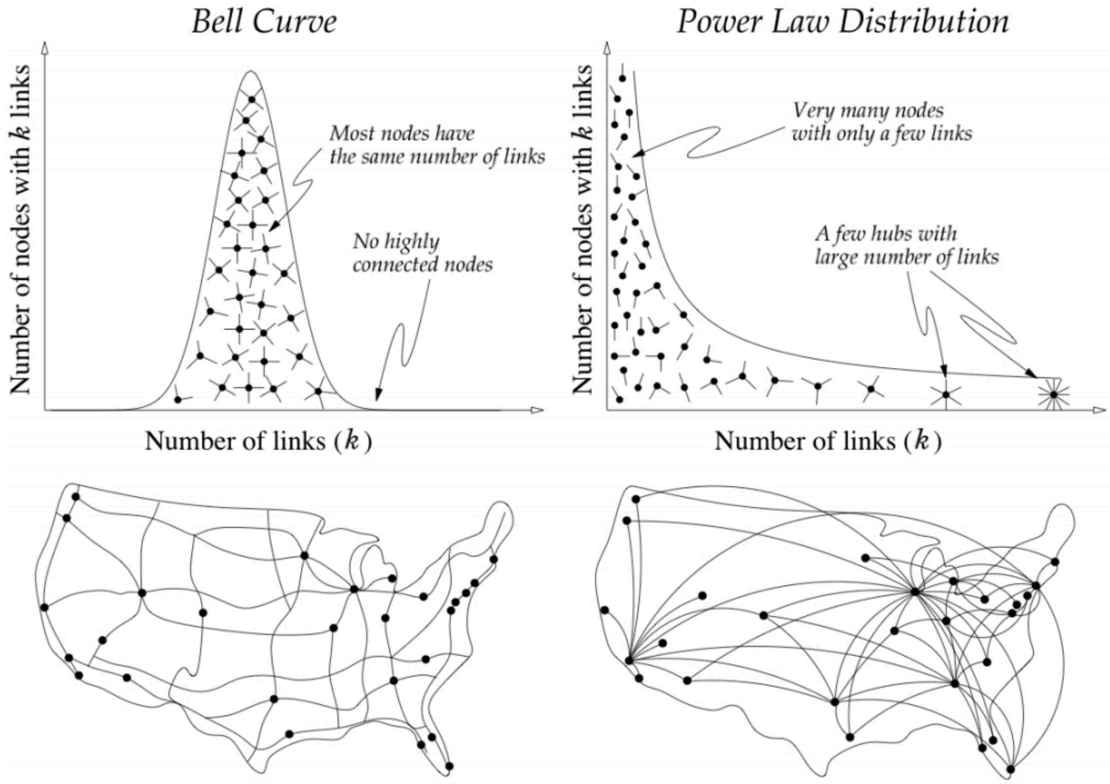
**Exercise:** show that if degree distribution is given by the exponential distribution $p(k) = Ce^{-\lambda k}$ (as an approximation for the Poisson distribution), then

$$k_{max} = k_{min} + \frac{\log N}{\lambda}.$$

**Exercise:** Show that for a scale-free network:

$$k_{max} \sim k_{min} N^{1/(\gamma-1)}.$$

Thus, the difference between $k_{min}$ and $k_{max}$ is much higher in a scale-free network.




Comparison of a random network (left) and a scale-free network (right). The top part is the degree distribution. Source: Barabasi: Network Science

# Moments of the degree distribution (optional)

Remember that the *moments* of a distribution $P$ are given by $E[P^n]$ for $n \geq 1$. For $n = 1$, this gives the mean of the distribution and we have $\sigma^2 = E[k^2] - E[k]^2$ and thus, standard deviation can be expressed in terms of the second moment.
For the Poisson distribution with mean equal to $\lambda$, the standard deviation equals $\sqrt{\lambda}$.
For the power-law distribution we have

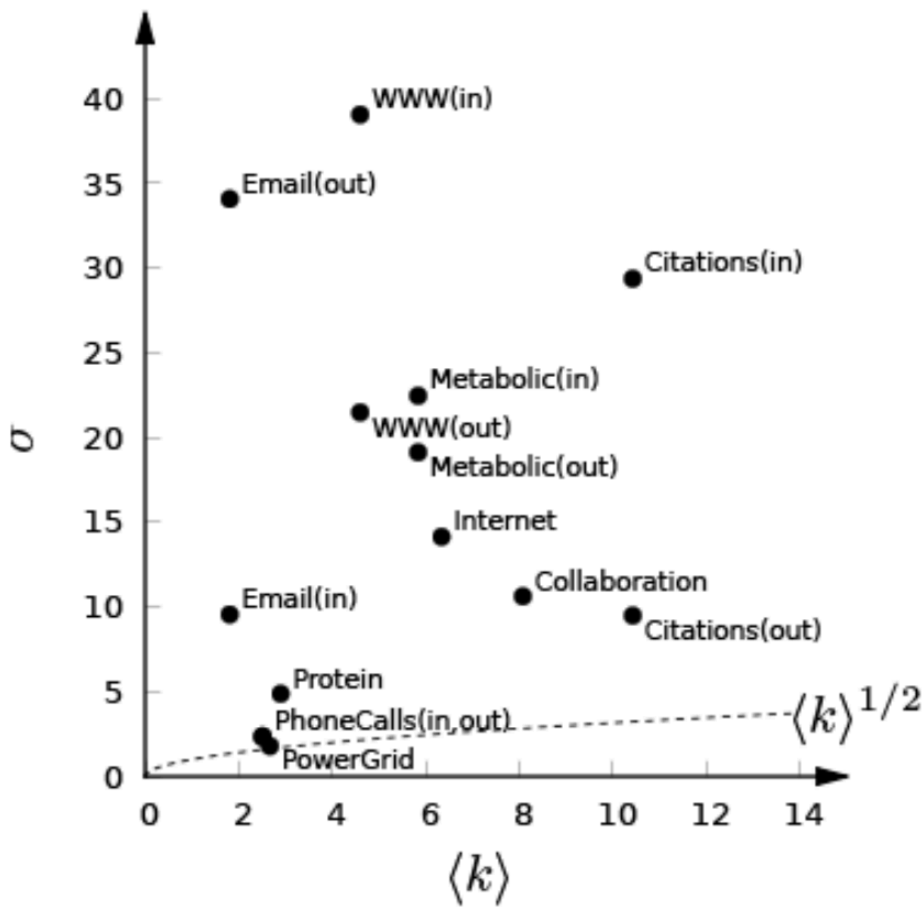
$$E[k^n] = \int_{k_{min}}^{k_{max}} k^n k^{-\gamma} dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n - \gamma + 1}.$$

We have $E[k] = C \frac{k_{max}^{2-\gamma} - k_{min}^{2-\gamma}}{2-\gamma}$ if $\gamma \neq 2$ and $C(\log k_{max} - \log k_{min})$ otherwise. It follows that if $\gamma \neq 2$ then $E[k] \sim N^{\frac{2-\gamma}{\gamma-1}}$.

As $N \to \infty$, $k_{max}$ goes to infinity. This means that in this limit, all moments for $n > \gamma - 1$ diverge. Since for real networks $\gamma$ is between 2 and 3, this means that the second and higher moments diverge. In other words, the scale (standard deviation) of the degree distribution is infinite, for a large network. For the Poisson distribution we have $\sigma = \sqrt{E[k]}$.

| NETWORK | NL | | $\langle k \rangle$ $\langle k_{in} \rangle = \langle k_{out} \rangle$ | $\sigma_{in}$ | $\sigma_{out}$ | $\sigma$ | $\gamma_{in}$ | $\gamma_{out}$ | $\gamma$ |
|---|---|---|---|---|---|---|---|---|---|
| Internet | 192,244 | 609,066 | 6.34 | - | - | 14.14 | - | - | 3.42* |
| WWW | 325,729 | 1,497,134 | 4.60 | 39.05 | 21.48 | - | 2.31 | 2.00 | - |
| Power Grid | 4,941 | 6,594 | 2.67 | - | - | 1.79 | - | - | Exp. |
| Mobile Phone Calls | 36,595 | 91,826 | 2.51 | 2.39 | 2.32 | - | 4.69* | 5.01* | - |
| Email | 57,194 | 103,731 | 1.81 | 9.56 | 34.07 | - | 3.43* | 2.03 | - |
| Science Collaboration | 23,133 | 93,439 | 8.08 | - | - | 10.63 | - | - | 3.35 |
| Actor Network | 702,388 | 29,397,908 | 83.71 | - | - | 200.86 | - | - | 2.12 |
| Citation Network | 449,673 | 4,689,479 | 10.43 | 29.37 | 9.49 | - | 3.03** | 4.00 | - |
| E. Coli Metabolism | 1,039 | 5,802 | 5.58 | 22.46 | 19.12 | - | 2.43 | 2.90 | - |
| Yeast Protein Interactions | 2,018 | 2,930 | 2.90 | - | - | 4.88 | - | - | 2.89* |

The mean and standard deviation of degree for some real networks. * indicates that data fits a power law distribution while ** indicates fit for a finite second momentum. Source: Barabasi, Network Science

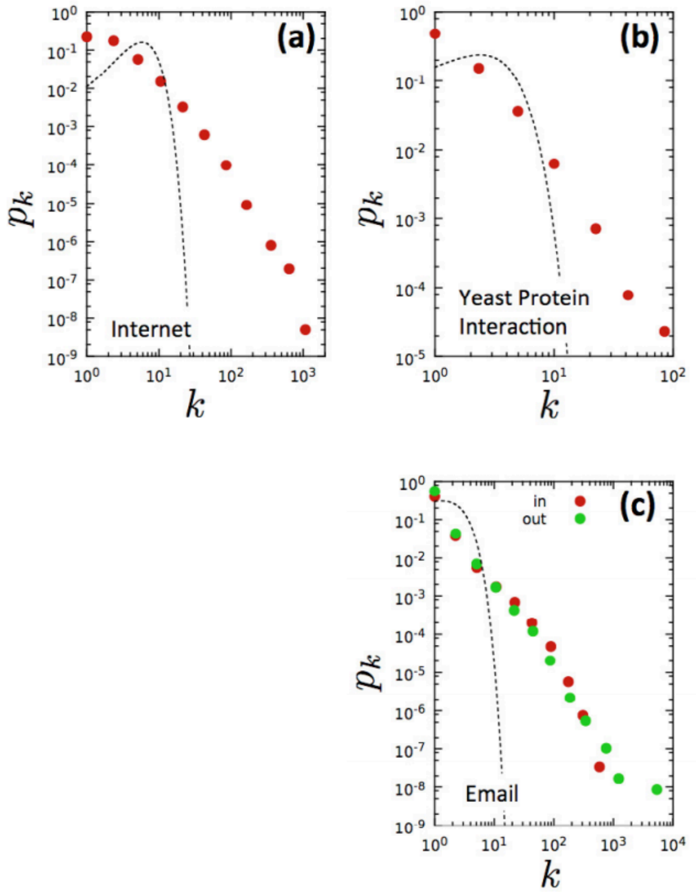



The plot of the standard deviation of degree vs average degree for some real networks. Source: Barabasi, Network Science

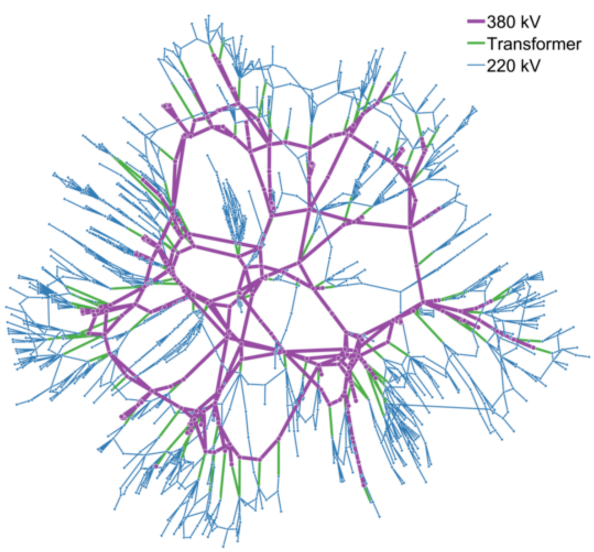
# Universality of the scale-free property

It was shown by de Solla Price in 1965 that the number of citations scientific publications receive follows a power-law distribution.

The scale-free property was first established for the Web in 1999. It was later shown to hold for other networks such as yeast protein interaction network and email communication network.




However not all networks are scale-free, e.g. power grid and the regular graphs arising in chemistry. If a system has a limit on the number of links a node can have, the network will not be scale-free.




The structure of a power gird. Source: Wikimedia Commons

A few "crossover" disributions are used in network theory as well, such as the following:

- Power-law with exponential cutoff:

$$P(k) = Ck^{-\gamma}e^{-\lambda k}$$

- Log-normal distribution

$$P(k) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(\log(k)-\mu)^2/2\sigma}$$

In plotting the degree distribution of a real network, we use a log-log plot, and we must use logarithmic binning instead of linear binning. This means that the size of the bins grows exponentially with $k$, to make sure that each bin has a comparable number of nodes.

# The role of the degree exponent and the small world property

Remember that

$$E[k^n] = C\frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n-\gamma+1}$$

if $n \neq \gamma - 1$ and $C(\log(k_{max}) - \log(k_{min}))$ otherwise. This means that if $n \geq \gamma - 1$, $E[k^n]$ is divergent as $N \to \infty$.

## The anomalous case $\gamma \leq 2$

It follows from $k_{max} \sim k_{min}N^{1/(\gamma-1)}$ that if $\gamma \leq 2$ then $k_{max}$ grows as a polynomial of $N$ (linearly when $\gamma = 2$). This means that if $\gamma < 2$, the degree of the largest hub grows faster than network size. Thus, scale-free networks with $\gamma < 2$ do not exist, unless multiple edges between a pair of nodes are allowed. In both cases $E[k]$ goes to infinity as $N \to \infty$.

When $\gamma = 2$, $k_{max}$ grows linearly with $N$ which means the network has a single hub with the rest of the nodes attached to it. In this case, average distance does not depend on $N$.

## Ultra-small regime $2 < \gamma < 3$

Several real networks demonstrate scale-free property with $\gamma$ in this range. Remember that for a random network $E[d]$ was proportional to $\frac{\log N}{\log E[k]}$. For scale-free networks, the existence of hubs makes the average distance to be even smaller. It was shown that $E[d] \sim \frac{\log \log N}{\log(\gamma-1)}$.
In this regime, as mentioned above, $E[k]$ is finite but $E[k^2] \to \infty$ as $N \to \infty$. This means that

average degree is finite but its standard deviation goes to infinity with $N$.
We have $k_{max} \sim N^{1/(\gamma-1)}$, thus $k_{max}/N \sim N^{-\alpha}$ where $0 < \alpha < 1$. This means that the largest hub is directly connected to a sizeable portion of the network.

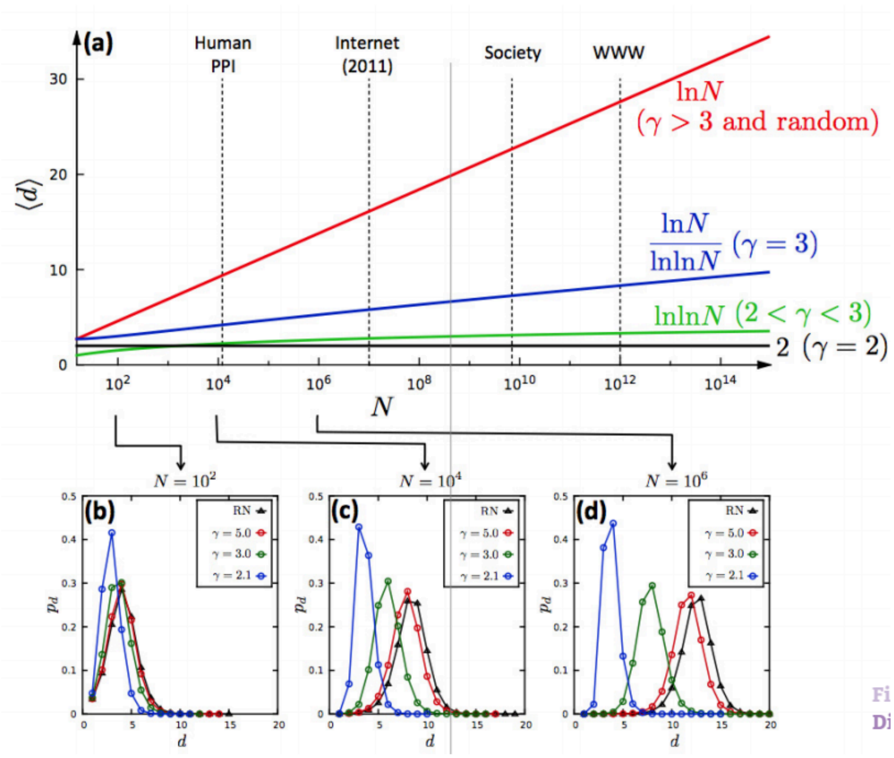

- B. Bollobás and O. Riordan. The Diameter of a Scale-Free Random Graph. Combinatorica, 24: 5-34, 2004.
- R. Cohen and S. Havlin. Scale free networks are ultrasmall, Phys. Rev. Lett. 90, 058701, 2003.


## Critical point $\gamma = 3$

In this case
$E[d] \sim \frac{\log N}{\log \log N}$

## The case $\gamma > 3$

The first and second momentum are finite and $E[d] \sim \log N$. Hence, the network behaves similar to a random network.




Other topics:
Goodness of fit for the degree distribution and the approximation of $\gamma$

# The Barabas-Albert Method

In this section we move from network topology to the evolution of the complex system it represents. We want to see how scale-free networks are formed.

The Barabasi-Albert model explains the evolution of networks using two concepts:

- **Growth:** most real networks are growing constantly. New web pages are created, new email and social media accounts are created and even new genes are formed. Also new patents are granted and new research papers are published every day.
- **Preferential attachment:** new nodes tend to link to nodes with a high degree. (This goes against the Erdős–Rényi model, in which new links are formed completely randomly.)

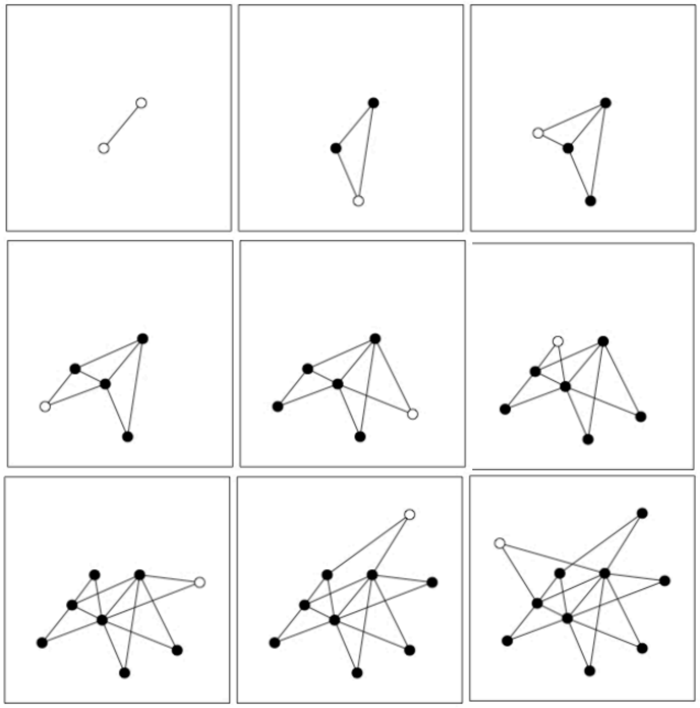
> "For everyone who has will be given more, and he will have an abundance. Whoever does not have, even what he has will be taken from him." Gospel of Mathew

Barabasi and Albert showed that these two concepts can be used to generate scale-free networks.

- The process starts with $m_0$ random nodes, with $L_0$ links spread randomly between them.
- At each step a new node is added, with $m \leq m_0$ links connecting it to the other (randomly chosen) nodes.
- The probability of the new node connecting to the node $i$ with degree $k_i$ equals

$$\pi(k_i) = \frac{k_i}{\sum_j k_j}.$$

After $t$ steps, the network has $N = m_0 + t$ nodes and $L_0 + tm$ links.



A simulation of the Barabasi-Albert process with $m_0 = m = 2$. Source: Barabasi Network Science

**Exercise:** using Python and Networkx, run a simulation of the Barabasi-Albert model with $m_0 = 2, m = 2$ for 10000 iterations. Inspect how close is the degree distribution of the resulting network to a power-law.

# Degree Dynamics

Note that at each time step, the degree of the node $i$ is expected to increase by $m\pi(k_i)$. Thus, considering time to be continuous, we have

$$\frac{\partial k_i}{\partial t} = m\frac{k_i}{\sum_j k_j} = m\frac{k_i}{2mt} = \frac{k_i}{2t}.$$

Thus,

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

where $\beta = 1/2$ and $t_i$ is the time node $i$ was introduced to the network.
For older nodes $t_i$ is smaller and therefore, their degree increases at a faster rate.

Note also that the growth in a node's degree is sublinear in time ($\sim t^{1/2}$), whereas in the random graph model, growth is linear in time (if new nodes are added at a constant rate). This is because the probability of link formation between the new node and existing nodes is the same for all the nodes. In the Barabasi-Albert model, newer nodes in the network face more competition for new links compared to the older ones.

To derive the degree distribution of this network, note that when a new node is added, it can affect the number $N(k,t) = Np_t(k)$ of nodes of degree $k$ by either:

- linking to a node of degree $k-1$ and thus increasing $N(k,t)$ by 1, or
- linking to a node of degree $k$ and consequently decreasing $N(k,t)$ by 1.

For the second case, the number of nodes of degree $k$ that the new node can connect to equals:

$$\pi_t(k) \times m \times N \cdot p_t(k) = \frac{k}{2tm} \times m \times tp_t(k) = \frac{k}{2}p_t(k).$$

The first equality is because $N = t$. In a similar way, the number degree $k-1$ nodes that the new node can connect to is $\frac{k-1}{2}p_t(k-1)$. Therefore, from the above argument we conclude that the difference between the number of nodes of degree $k$, between time $t$ and $t+1$ is:

$$(N+1)p_{t+1}(k) - Np_t(k) = \frac{k-1}{2}p_t(k-1) - \frac{k}{2}p_t(k).$$

Note that the minimum degree of the nodes is $m$, because each new node is connected to $m$ other nodes at birth. For $k = m$, the above equation has to be adjusted to

$$(N+1)p_{t+1}(m) = Np_t(m) + 1 - \frac{k}{2}p_t(m).$$

The 1 is for the new node. We are interested in the limiting stationary distribution $p(k) = \lim_{t\to\infty p_t(k)}$. In this limit, if we move the first term on the right hand side of the above equations to left hand side, we get $p(k) = \frac{k-1}{k+2}p(k-1)$ for $k > m$ and $p_m = 2/(2+m)$. Using induction we can show that

$$p(k) = \frac{2m(m+1)}{k(k+1)(k+2)}.$$

For large $k$ we have $p(k) \sim k^{-3}$ and thus, this is approximately a scale-free distribution with $\gamma = 3$ .

**Exercise:** using the equation for $k_i(t)$ and assuming that time and degree distributions are continuous, give an alternative derivation of the degree distribution. Hint: consider $p(k_i(t) < k)$.

# Measuring preferential attachment in real networks

We saw that the Barabasi-Albert method results in scale-free networks. Now we want to see whether preferential attachment happens in real networks.
We want to test two hypotheses:

- The probability of a new node connecting to an existing node is proportional to latter's degree $k$, thus it can be denoted by $\pi(k)$.
- The probability $\pi(k)$ depends linearly on $k$.

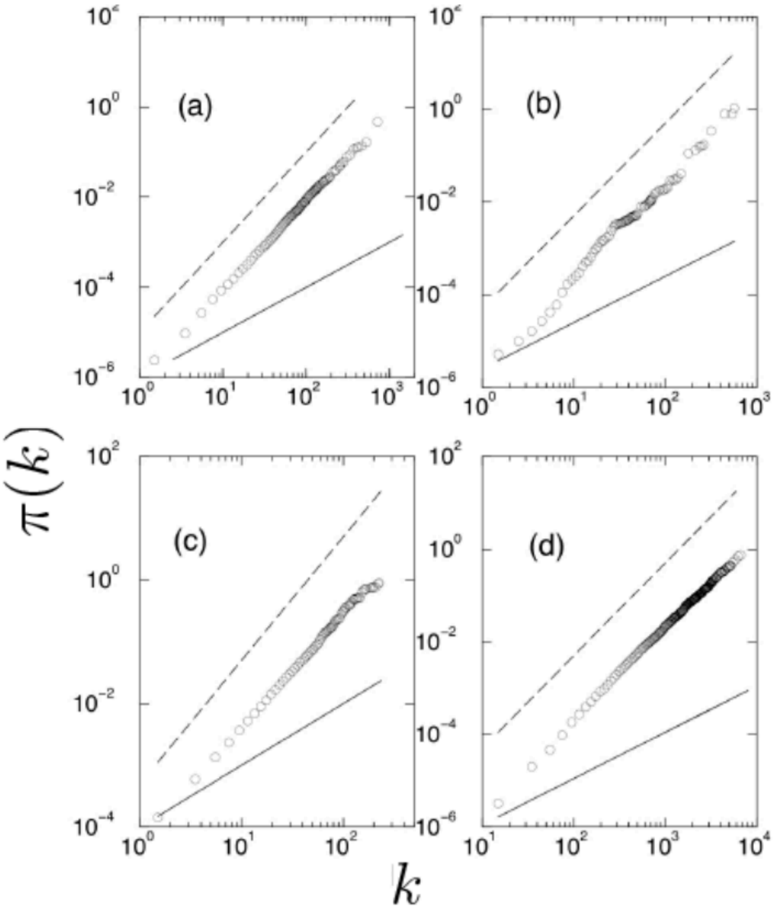
Imagine we have a network whose evolution we know over a time period. Actually it is sufficient to know the network map at two different times. If $i$ is a node whose degree $k_i$ is changed by the amount $\Delta k_i$ over this time period $\Delta t$ then we expect

$$\frac{\Delta k_i}{\Delta t} \sim \pi(k_i).$$

Since the $\frac{\Delta k_i}{\Delta t}$ curve is noisy, we consider the cumulative function

$$\Pi(k) = \sum_{k_i \leq k} \pi(k_i).$$

If there is no preferential attachment then $\pi(k_i)$ would be constant and thus $\Pi(k) \sim k$. When P.A. is present, we have $\Pi(k) \sim k^2$.




The plot of $\Pi(k)$ vs $k$ for a) citation network, b) internet, c) collaboration network, d) actor network. Source: Barabasi, Network Science

From the figure we conclude that in real networks $\pi(k) \sim k^{\alpha}$ is sublinear in $k$.

# Nonlinear preferential attachment (optional)

Inspired by the above discussion, we consider the B.A. model with $\pi(k) \sim k^{\alpha}$. When $\alpha = 0$ we get the random graph model and when $\alpha = 1$, we get the B.A. model.

## Sublinear case $0 < \alpha < 1$

In this case we have

$$\pi(k) = \frac{k^{\alpha}}{\mu(\alpha, t)}$$

where $\mu(\alpha, t) = E[k^{\alpha}] = \sum_k k^{\alpha} p(k, t)$. We denote $\mu(\alpha) = \lim_{t\to\infty} \mu(\alpha, t)$. We have $\mu(0) = 1$ and $\mu(1) = E[k] = 2L/N = 2mt/N \sim 2m$.
For $k > m$ we have:

$$p(k,t+1) = \frac{m(k-1)^\alpha}{\mu(\alpha,t)}p(k-1,t) - \frac{mk^\alpha}{\mu(\alpha,t)}p(k,t).$$

The first term on the right hand side is $m$ times the probability that the degree of a randomly chosen node is increased to $k$, when a new node is added. The second term is $m$ times the probability of such a node to stop being of degree $k$ when a new node is added.
When $k = m$, since there are no nodes of degree $m-1$, and the newly added node is of degree $m$, we have

$$p(m,t+1) = -\frac{m}{\mu(\alpha,t)}m^\alpha p(m,t) + 1/N$$

The last term indicates that the new node has degree $m$. Taking the limits when $t \to \infty$ we get $p_k = \frac{m}{\mu}((k-1)^\alpha p_{k-1} - k^\alpha p_k)$, and therefore, $p_k = \frac{(k-1)^\alpha}{\mu/m + k^\alpha}p_{k-1}$.

**Exercise:** using induction show that

$$p_k = \frac{\mu(\alpha)k^\alpha}{m}\prod_{j=m}^{k}(1+\frac{\mu(\alpha)}{mj^\alpha})^{-1}$$

**Exercise:** by taking the logarithm of both sides of the above equation and approximating the resulting sum by an integral, show that for $1/2 < \alpha < 1$ we have:

$$p(k) \sim k^{-\alpha}\exp\left(-\frac{\mu(\alpha)}{m(1-\alpha)}k^{1-\alpha}\right)$$

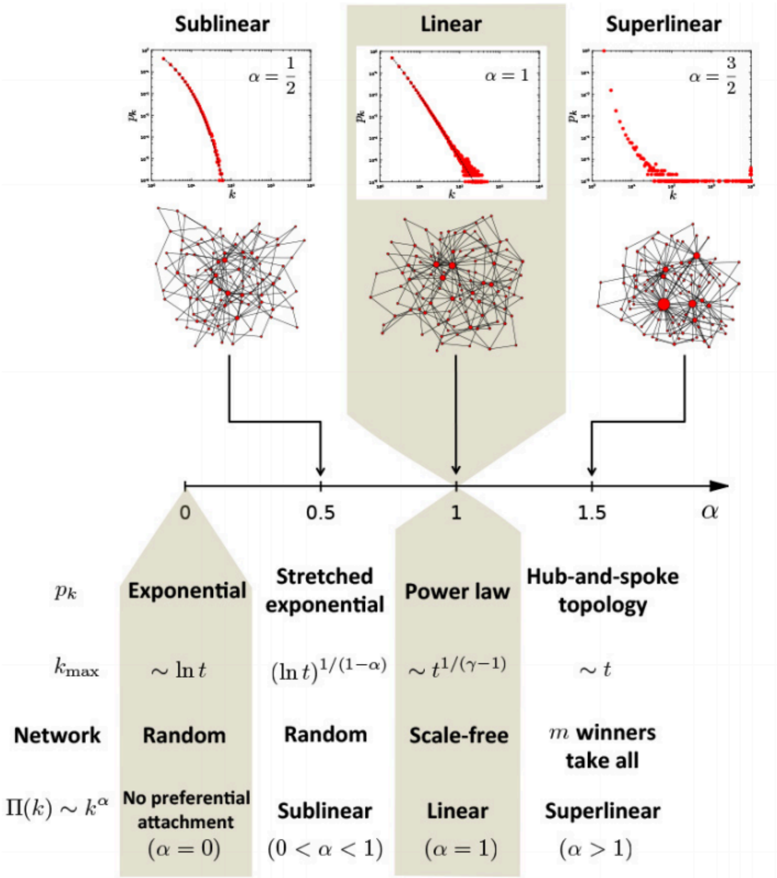
In this case degree distribution is a stretched exponential.
We see that sublinear P.A. limits the size and the number of the hubs. We also have

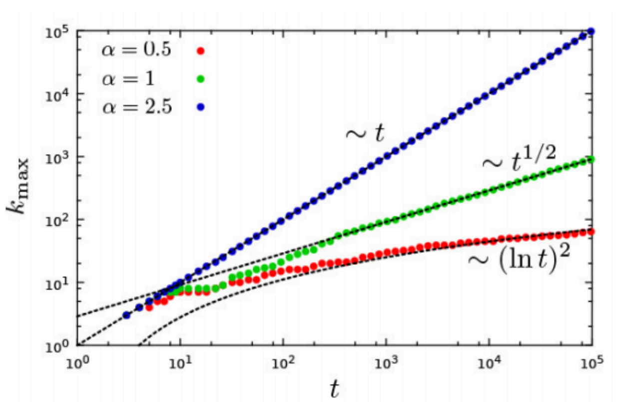
$$k_{max} \sim (\log t)^{1/(1-\alpha)}.$$

This can be responsible for deviations from a pure power lows seen in real networks.

# Superlinear case $\alpha > 1$

In this case, P.A. is amplified and we have $k_{max} \sim t \sim N$. We don't have a stationary distribution in this case.




The growth of the hubs. Source: Barabasi, Network Science.

# The Origins of Preferential Attachment

There are two theories for the origin of preferential attachment.

## Local or random mechanisms

The idea is that a new node chooses a *link* that is already in the network and tries to copy it, instead of choosing a node to connect to. At each step we choose an existing *link* in the network and connect the new node to one of the link's nodes.

In the **Copying Model**, for a new node $v$, a random target node $u$ is chosen, then

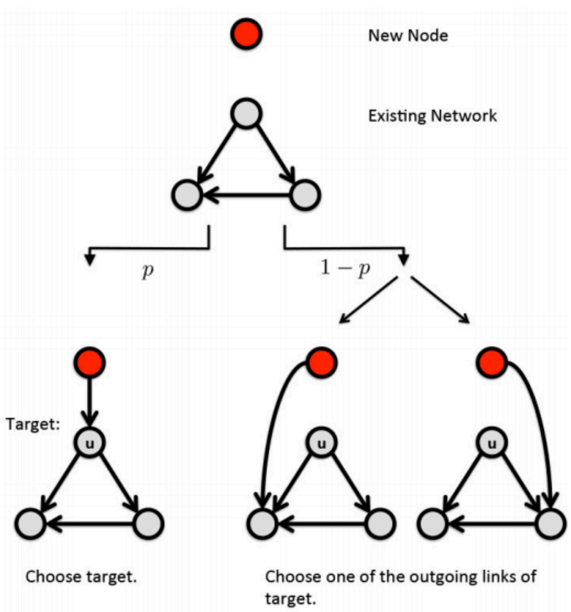
- with probability $p$ it links to $u$ and
- with probability $1-p$ it links to one of the (randomly chosen) nodes that $u$ links to.

Note that the second step is equivalent to choosing a random link from the network.

Copying model can be observed in real networks. For example:

- A person with a lot of friends is likely to be introduced to new friends by his/her current friends.
- A researcher is likely to cite the publications that the papers he/she has read refer to.

The probability of the new node $v$ connecting to a node of degree $k$ is $p \times p(k) + (1-p)\frac{k}{2L}$.




Copying model. Source: Barabasi, Network Science

- Kumar, et al., The Web as a graph. Proceedings of the 19th Symposium on principles of database systems, 2000.