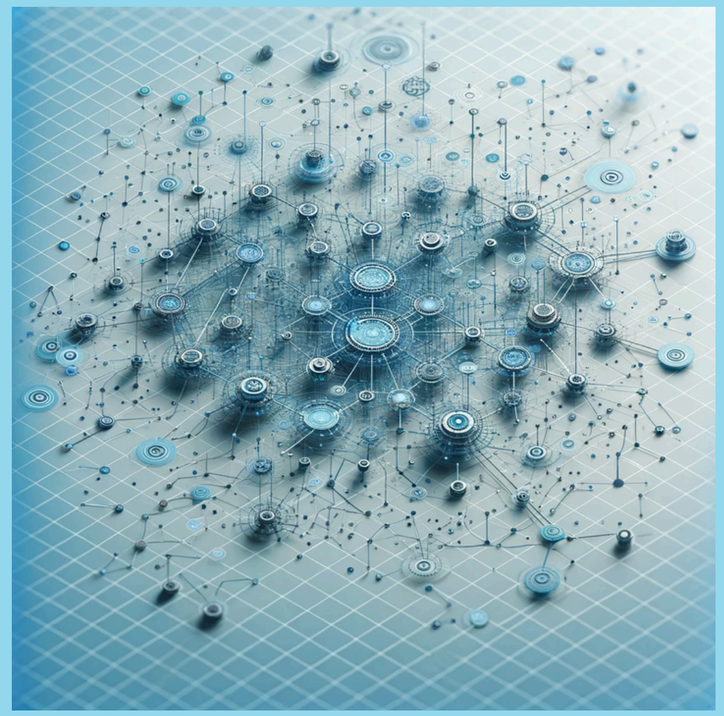


Universität Konstanz

UNIVERSITÄT KONSTANZ

Network Theory

Computational Modelling of Social Systems

Giordano De Marzo
Max Pellert

Recap

The Paretian World

Our world is dominated by extreme events that are not described by gaussians.

Power Law Probability Distributions

Power Law probability distributions well describe extreme events.

Zipf’s Law and Heaps’ Law

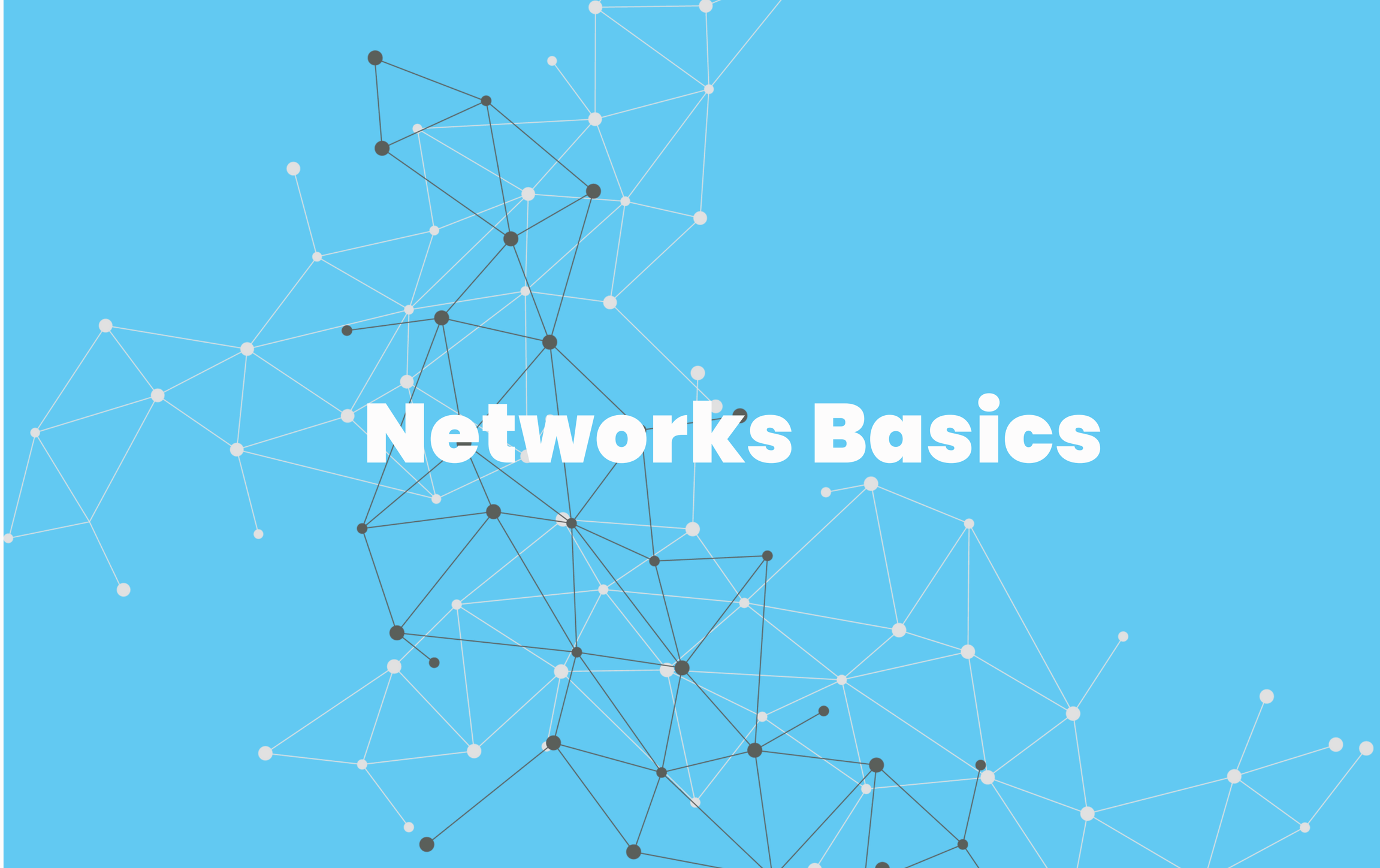
These two statistical regularities are different manifestations of power laws and are observed in most complex systems.

Modelling the Cryptocurrency Market

We can model innovation in the crypto market using the adjacent possible idea.

Outline

1. Networks Basics
2. Real World Networks
3. Network Formation Models


Network's Basics

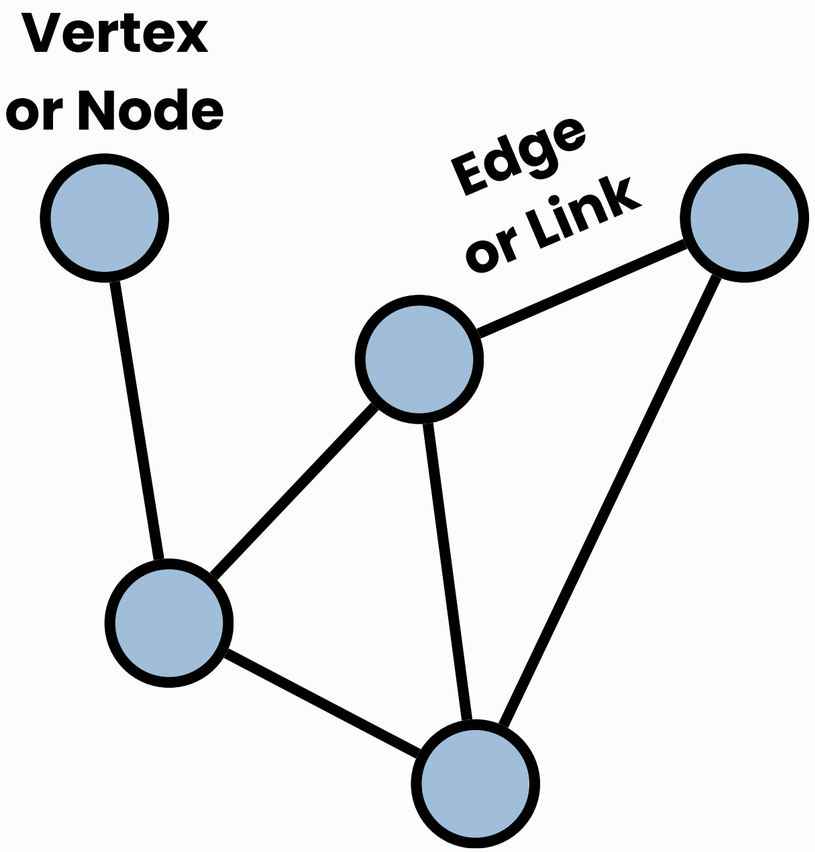
What is a Network?

A Network or Graph G(V, E) is a set of vertices or nodes V and and edges or links E

- nodes represent entities in the system (eg. people on a social network)
- edges represent connections among the nodes (eg. friendship in a social network)

We denote by

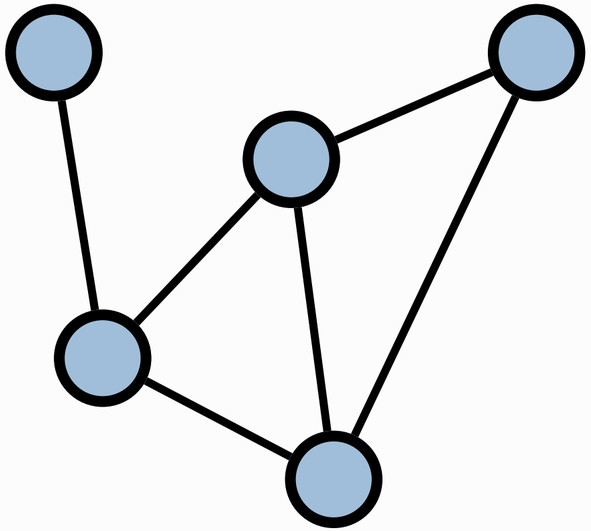
- N the number of nodes
- E the number of links



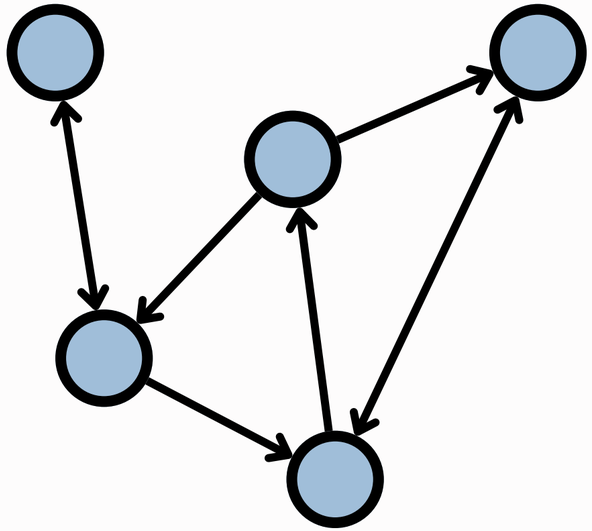
Network Types

There are three main types of graphs

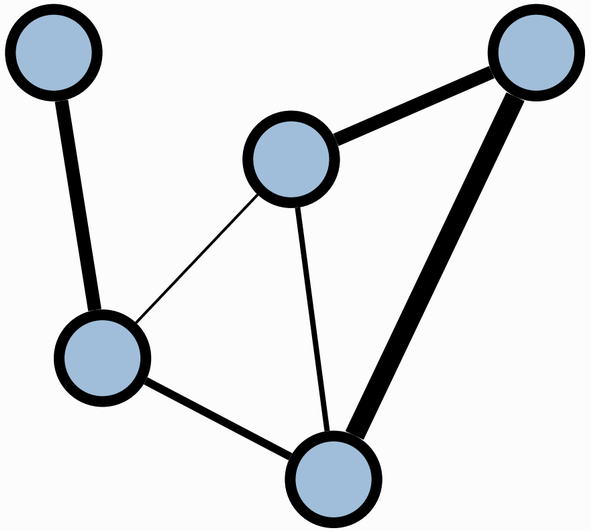
Undirected
Links are bidirectional
E.g. Facebook



Directed
Links are directional
E.g. Twitter



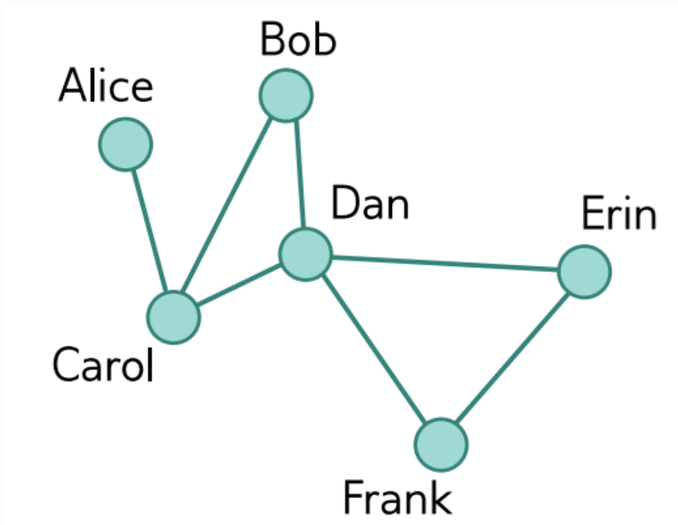
Weighted
Links are weighted
E.g. Road Network



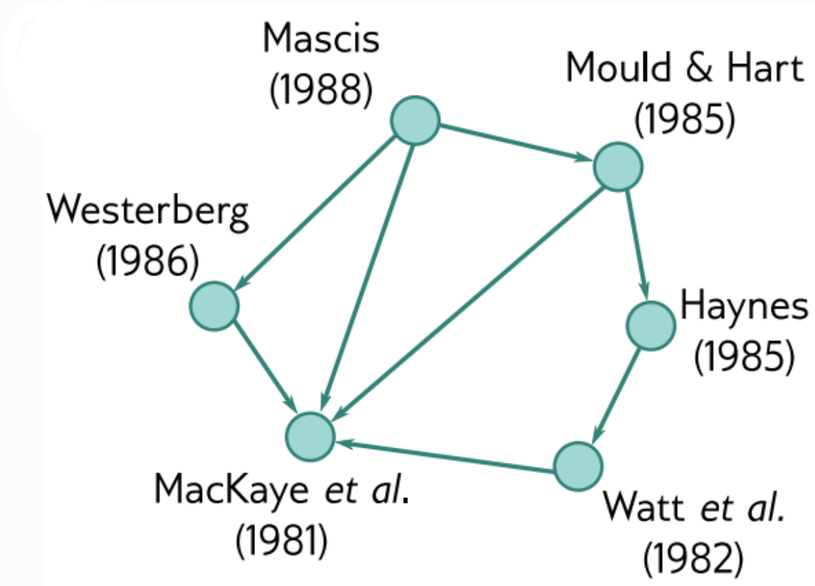
Examples of Networks

Graphs are everywhere, many systems can be described using this formalism

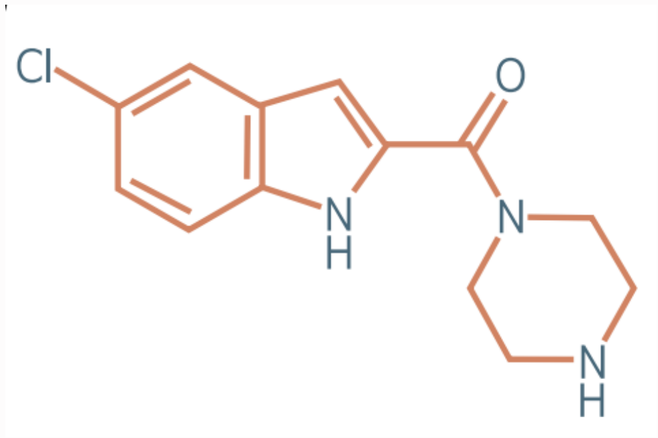
Undirected
Friendship Networks



Directed
Citation Networks



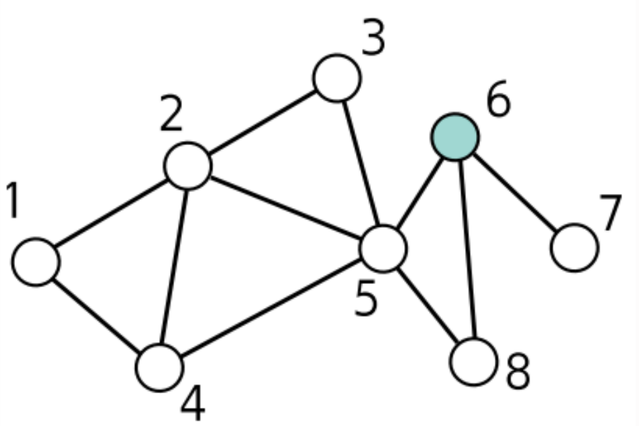
Weighted
Molecules



The Adjacency Matrix

A network is mathematically represented by its adjacency matrix A

- it's an NxN matrix
- the element A_{ij} of the matrix is different from zero if there is a link going from i to j
  - $A_{ij}=1$ in unweighted graphs
  - $A_{ij}=w_{ij}$ in weighted graphs
- if the graphs is undirected the adjacency matrix is simmetric $A_{ij}=A_{ji}$
- elements on the diagonal (self-loops) are typically null $A_{ii}=0$



$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

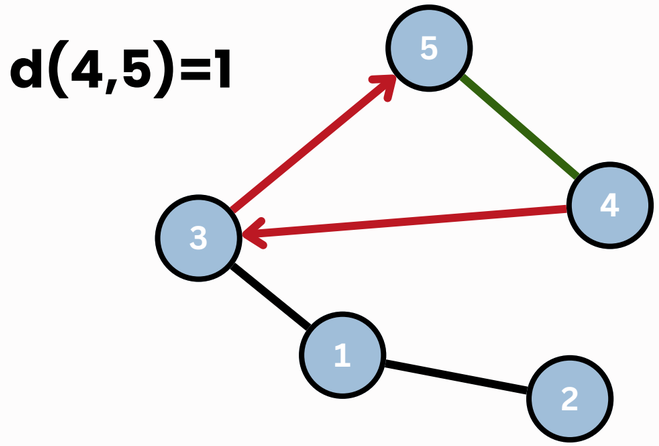
Distance on a Network

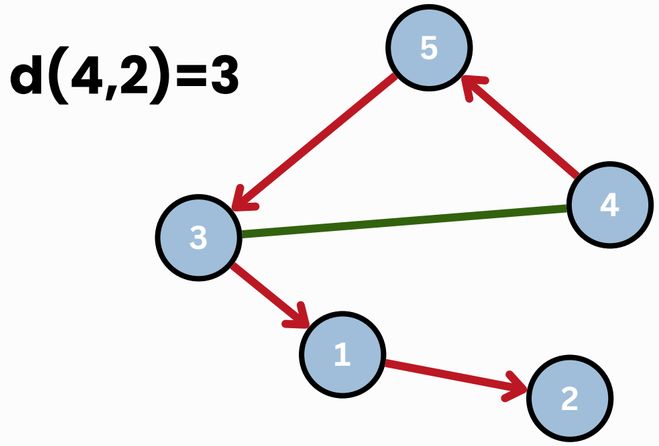
A path on a network is a sequence of links connecting a series of nodes

- the length of a path is the number of links it contains
- the distance d(i,j) between two nodes i, j defined as the length of the shortest path connecting them

We can find the number of paths between two nodes using the adjacency matrix

- A_{ij} gives the number of length 1 paths from i to j
- $(\mathbf{A}^2)_{ij}$ gives the number of length 2 paths from i to j





Diameter of a Network

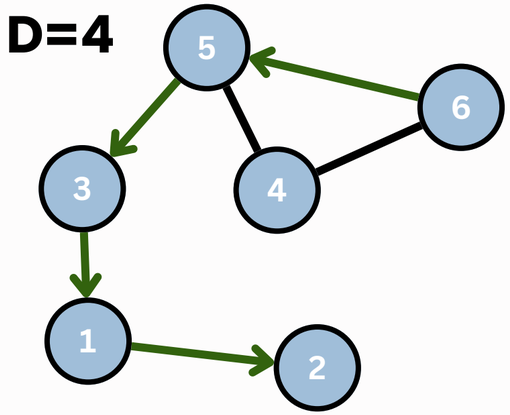
As we already explained a network is different from a typical metrical space like a lattice. However also in this case we can define the dimension or diameter of a network

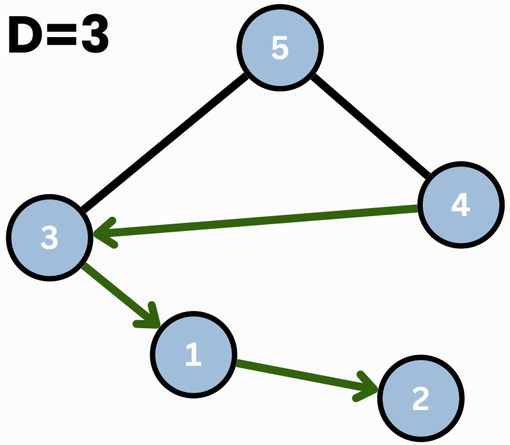
- we take all possible paths between any pair of nodes in the network
- the diameter is defined as the largest of this path

$$D = \max[d_{ij}]$$

Another useful measure is the average path length L

$$L = \frac{1}{N(N-1)} \sum_{i,j} d_{ij}$$





Degree Distribution

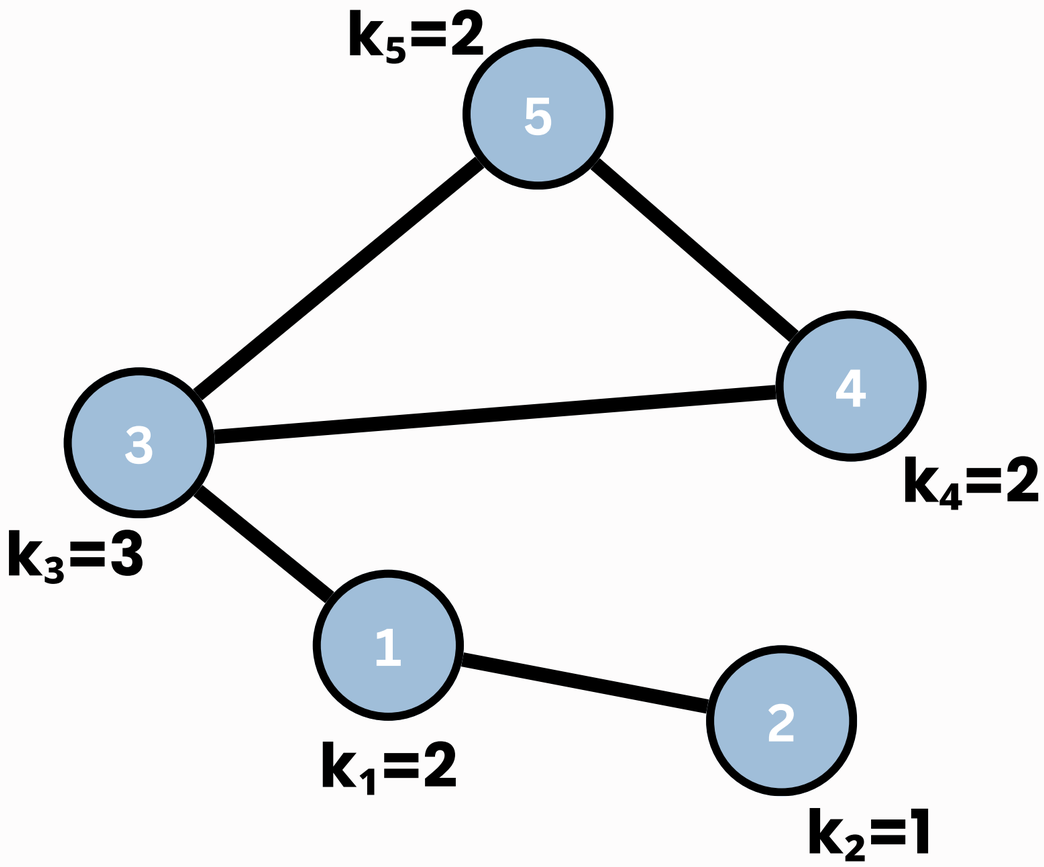
We define the degree of a node as the number of links this node has. In the case of directed network node have 2 degrees

- in-degree (number of incoming links)
- out-degree (number of outgoing links)

The degree k_i of node i can be computed using the adjacency matrix as

$$k_i = \sum_j A_{ij}$$

Finally we denote by P(k) the degree (probability) distribution, that gives the probability of observing a giving degree.



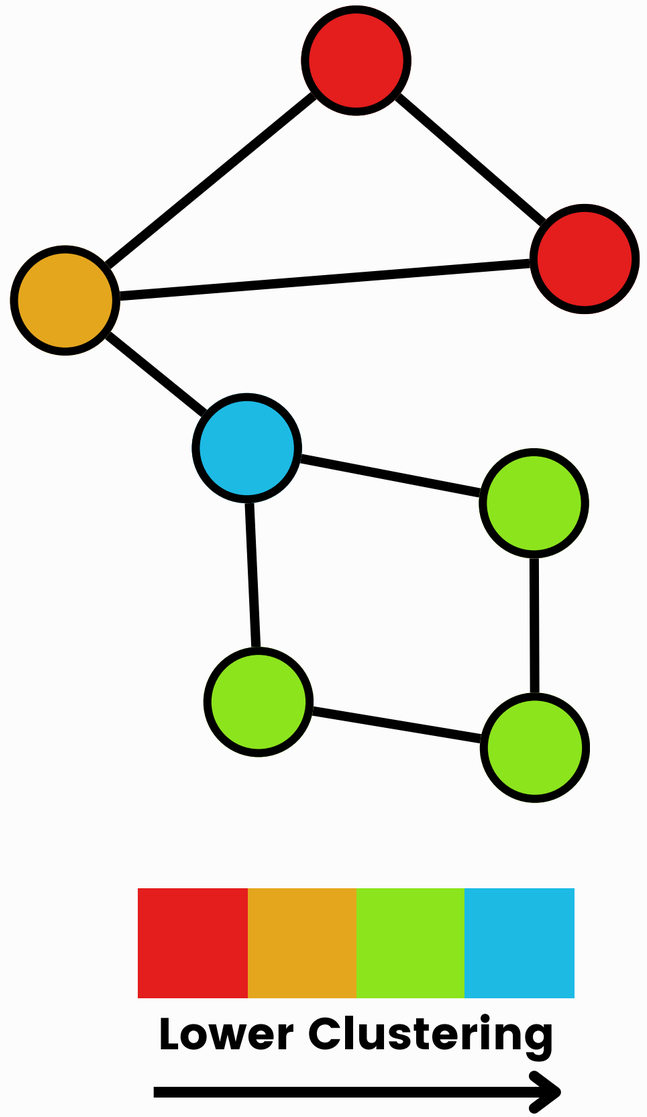
Clustering Coefficient

The clustering coefficient measures the likelihood that the friends of a node are also friends with each other

- **Local Clustering Coefficient:** For a single node, it is the ratio of the number of actual connections between its neighbors to the number of possible connections between those neighbors.

$$C_i = \frac{2 \cdot t_i}{k_i(k_i - 1)}$$

- **Global Clustering Coefficient:** it is the average of the local clustering coefficients of all nodes in the network.

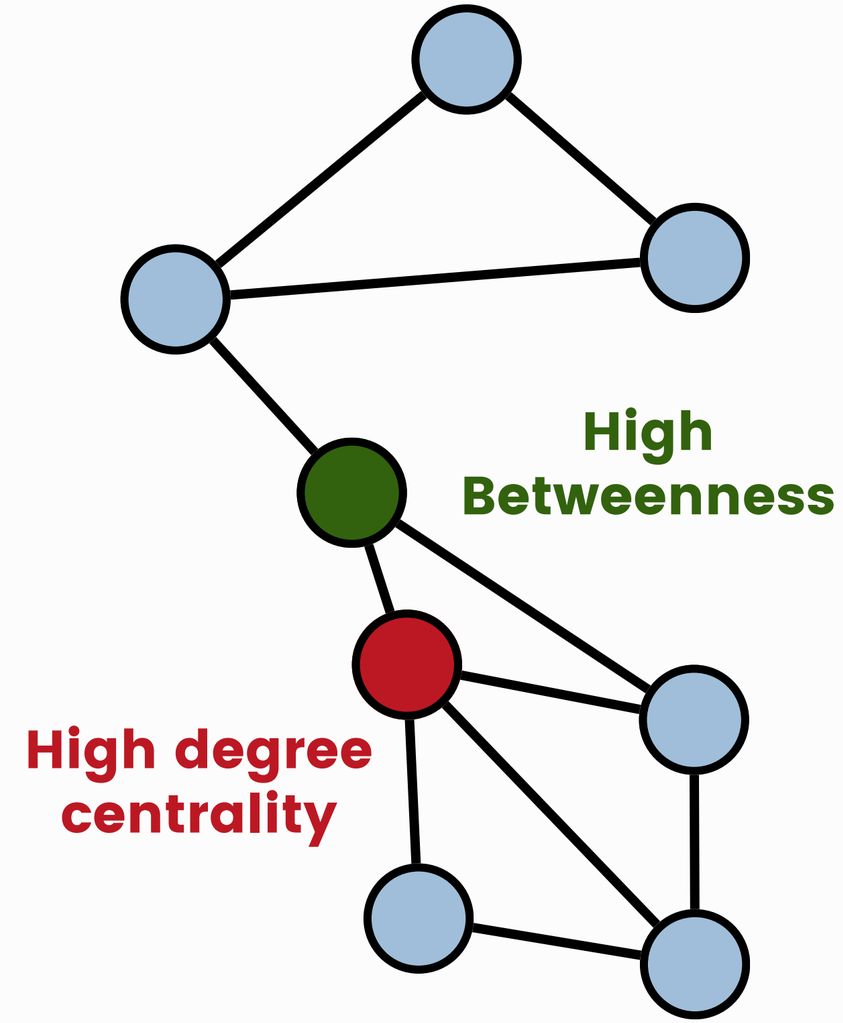


Centrality Measures

Centrality measures help identify the most important nodes within a network, based on various criteria.

- **Degree Centrality.** Nodes with higher degree centrality are more connected and can quickly interact with many others.
- **Betweenness Centrality.** Measures the extent to which a node lies on the shortest paths between other nodes.
- **Closeness Centrality.** Measures how close a node is to all other nodes in the network.

Each centrality measure captures a different property, there is no right centrality!



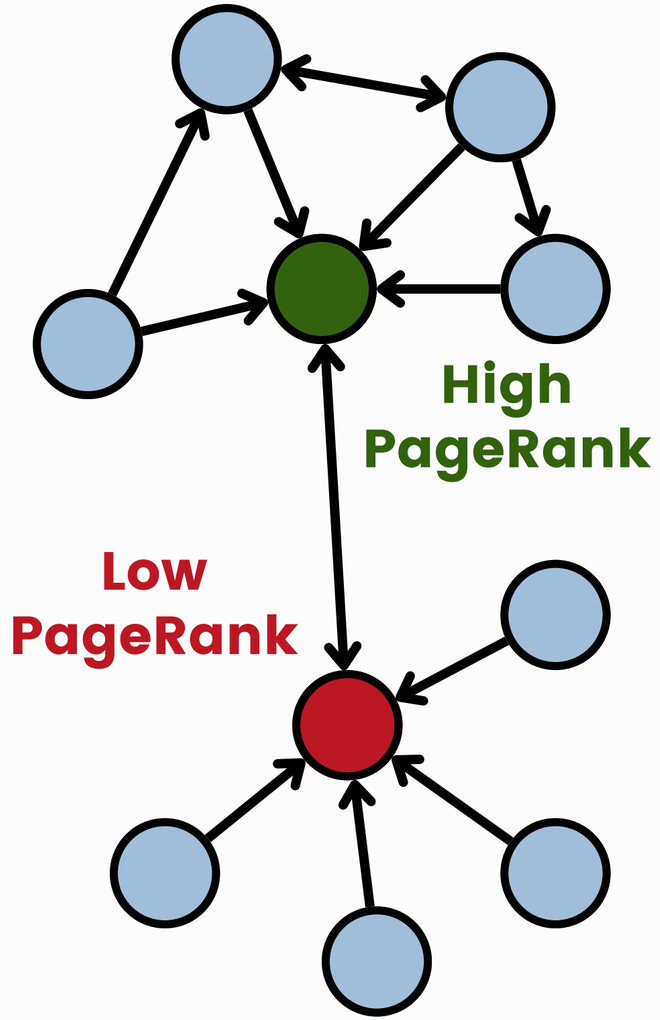
The PageRank

PageRank is an algorithm used to measure the importance of nodes within a network, originally developed by Google to rank web pages in search results.

- **Basic Idea:** Nodes with more incoming links (especially from important nodes) have higher PageRank values.
- **Recursive Nature:** PageRank of a node is influenced by the PageRank of nodes linking to it.

The PageRank is computed exploiting an iterative algorithm

$$PR(i) = \frac{1-d}{N} + \sum_j A_{ji} \frac{PR(j)}{k_j^{out}}$$



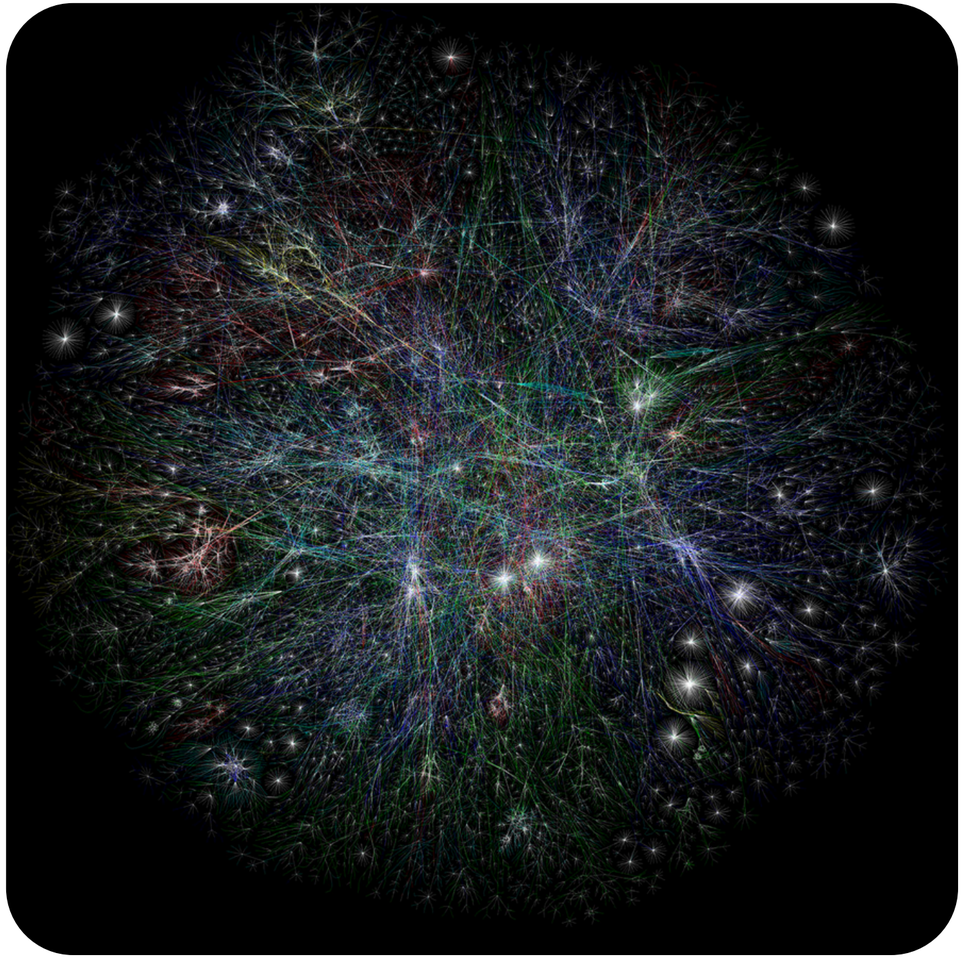

Real World Networks

Networks are Everywhere!

Networks are very versatile and many systems can be described using this mathematical formalism. Some examples include

- internet
- online social networks
- powerline networks
- airline networks
- food webs

The idea is that whenever there are “things” interacting or “talking” pairwise we can always describe the system as a network.

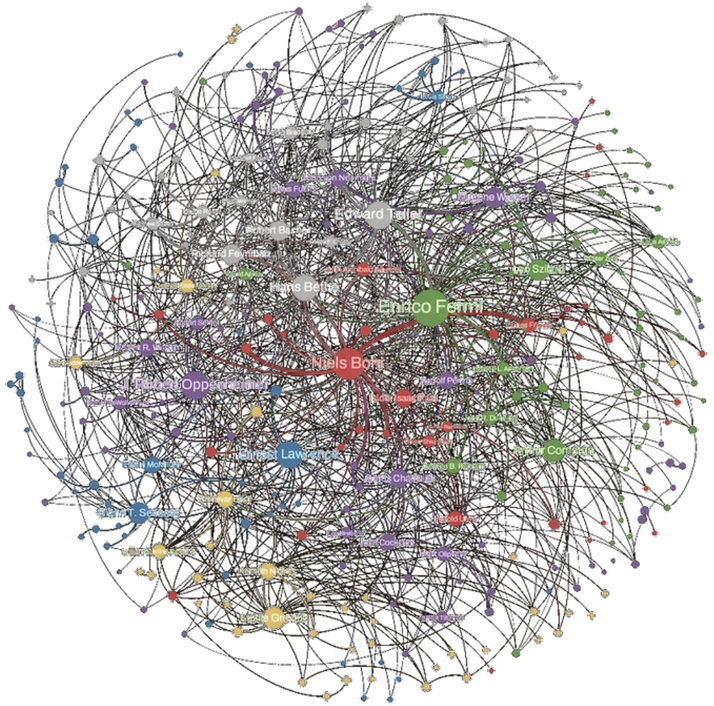


Examples from the Social Domain

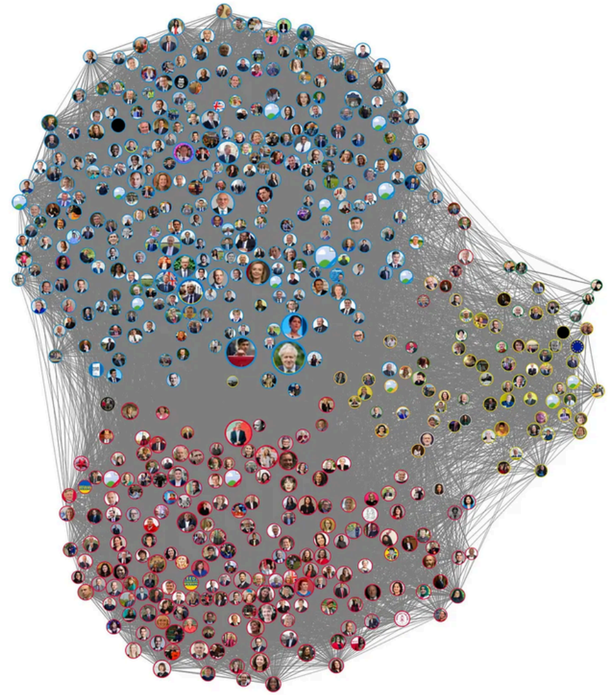
Many social systems spontaneously form networks

Board of Directors Network

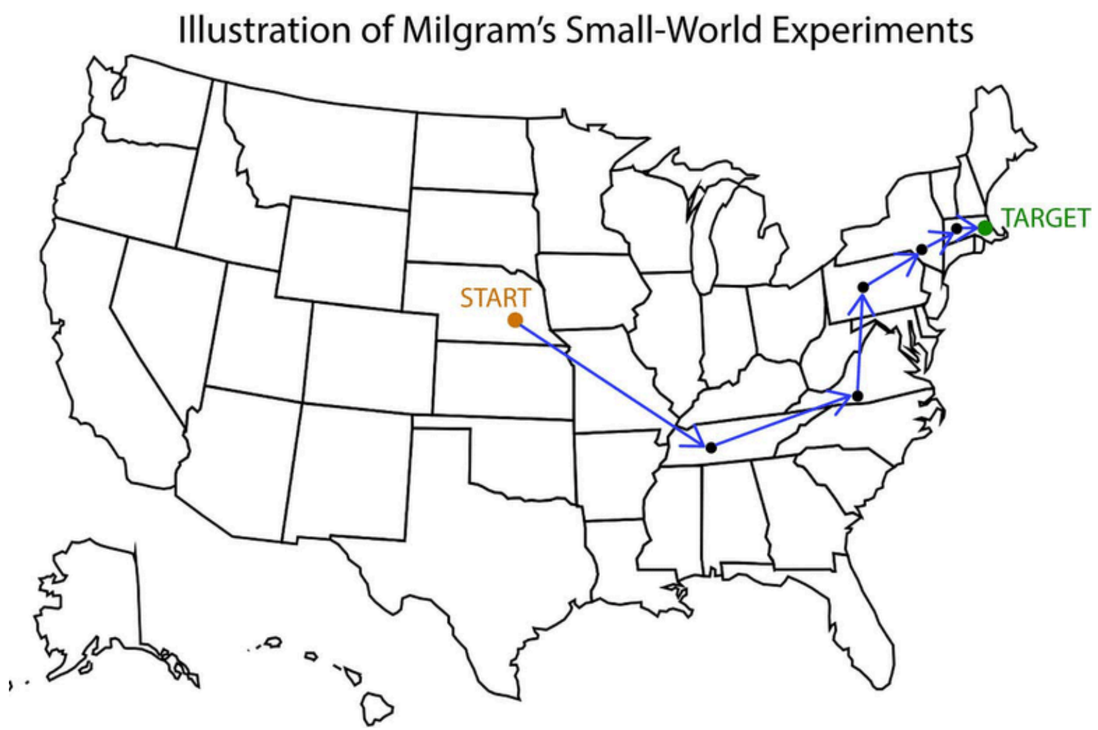
Project Manhattan Collaboration Network



House of Commons MP Twitter Network



Milgram's Experiment



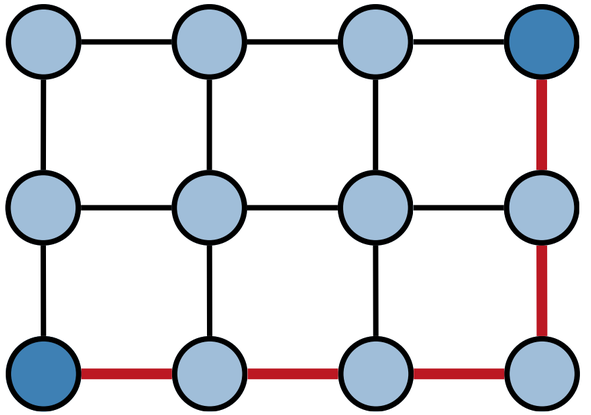
In 1967 Milgram measured the average path length in social networks

- **Participants:** Randomly selected individuals in Omaha, Nebraska
- **Task:** Send a package to a stockbroker in Boston
- **Method:** Each participant mailed a packet to a friend they thought was socially closer to the target. The process was repeated until the packet reached the stockbroker or the chain ended.

On average, 6 steps are needed to reach the target.

The Small World Property

Milgram’s experiment shows that despite networks can be huge, often the path connecting any two elements in the network can be surprisingly short. This phenomenon is often summarized by the popular notion of "six degrees of separation".

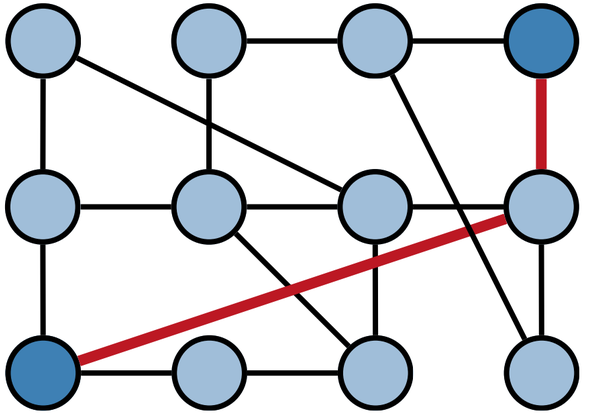


This property is mathematically expressed in terms of the average path length L and it is called small world property

$$L \sim \log(N)$$

Note that this is not true for lattices, for instance in D=2

$$L \sim \sqrt{N}$$



High Clustering

Many real world networks are characterized by two apparently opposite properties:

- high clustering C (nodes tend to form triangles)
- small world (the average path length L is small)

These two properties are apparently in conflict since

- random networks have small L but also small C (no local structure)
- lattices have high C, but also high L (too much local structure)

Network	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
C. elegans	2.65	2.25	0.28	0.05

Scale-free Networks

Many real world networks are characterized by a power law distribution of degrees. We call such graphs scale-free networks. In a scale-free network there are many nodes with few connections, but also few nodes with an enormous number of links.

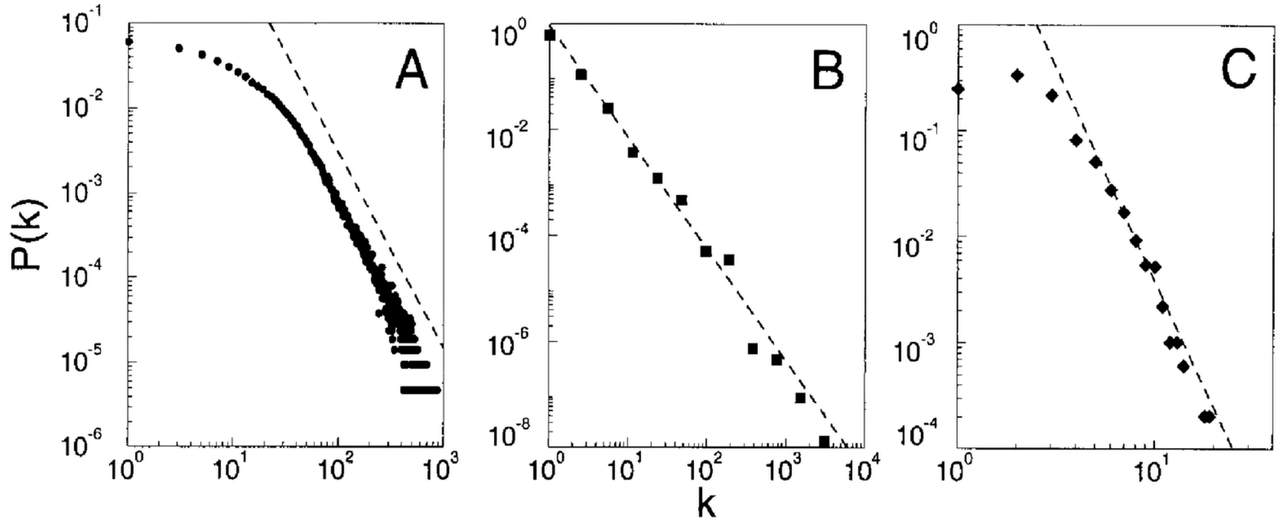


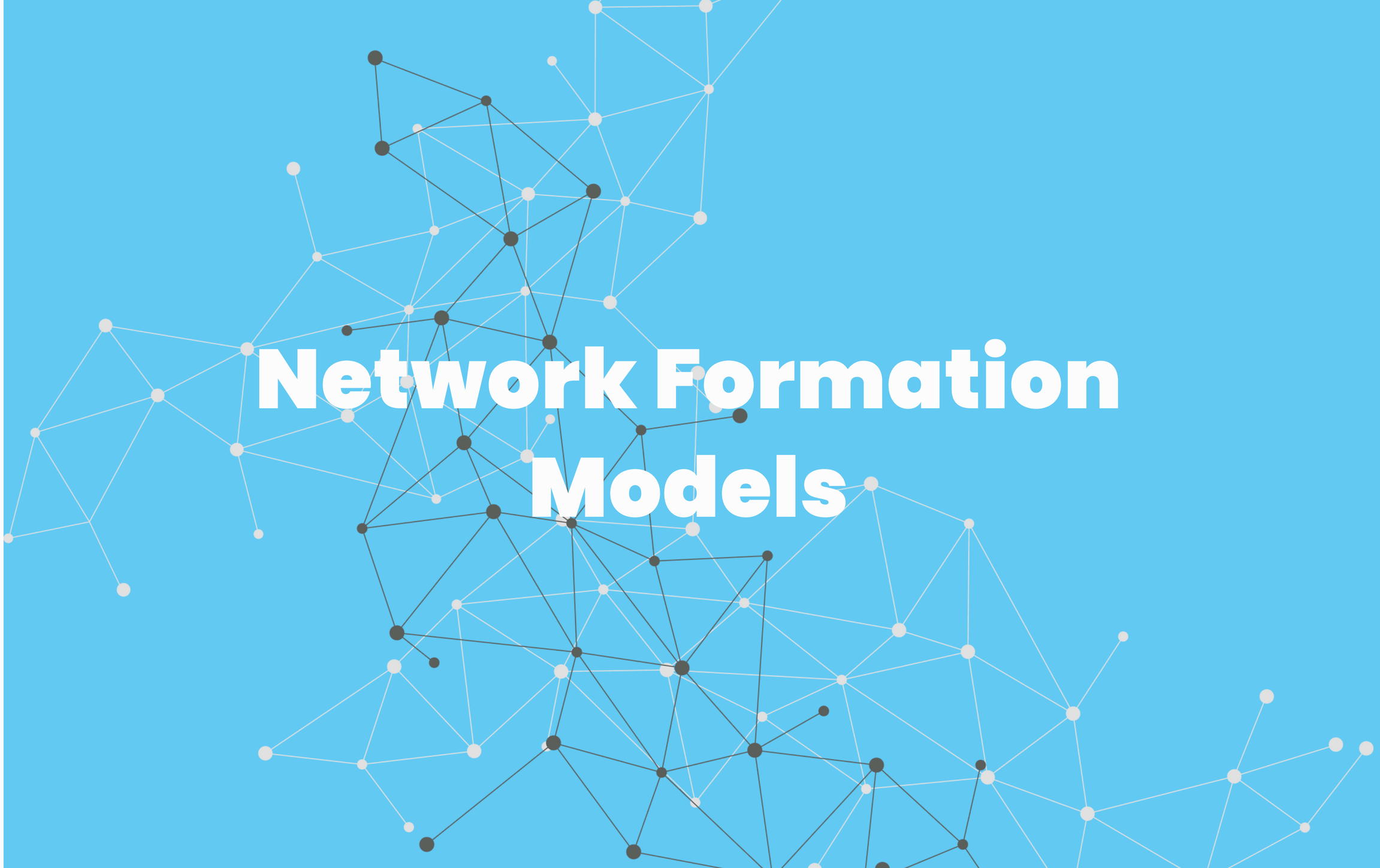
Fig. 1. The distribution function of connectivities for various large networks. (**A**) Actor collaboration graph with $N = 212{,}250$ vertices and average connectivity $\langle k \rangle = 28.78$. (**B**) WWW, $N = 325{,}729$, $\langle k \rangle = 5.46$ (*6*). (**C**) Power grid data, $N = 4941$, $\langle k \rangle = 2.67$. The dashed lines have slopes (A) $\gamma_{actor} = 2.3$, (B) $\gamma_{www} = 2.1$ and (C) $\gamma_{power} = 4$.

Research Questions

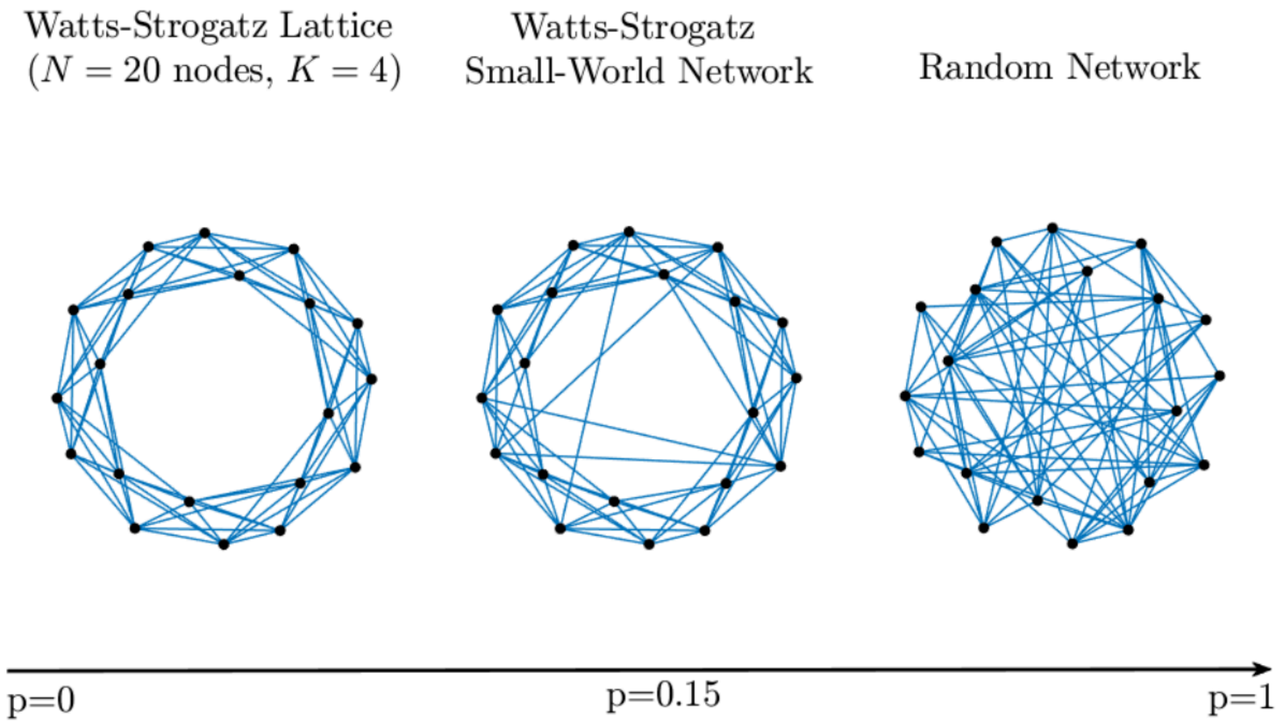
We saw that real world networks are characterized by 3 main properties, two of them often being observed together. We want to understand what are the factors and the mechanisms that are making these features emerge. In particular:

- how can we have both high clustering and the small world property?
- which are the generative mechanisms producing these features?
- how can we get a scale-free network from a simple model?
- can we get realistic networks only using local mechanisms?

Network Formation Models



The Watts-Strogatz Model



The Watts-Strogatz Model is one of the most simple models

- start with a ring with connections only to near nodes (on both sides)
- rewire each link with probability p

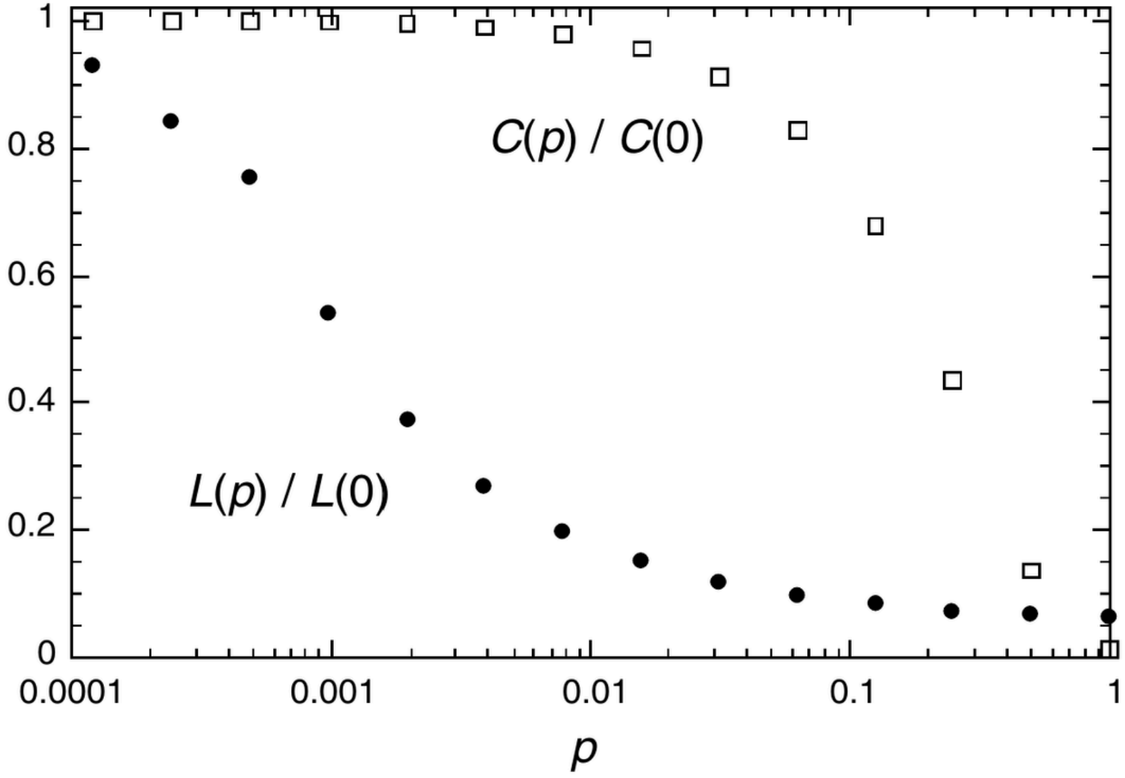
For p=0 we have a regular ring network (similar to a lattice), while for p=1 we have a random network. What happens in between?

Properties of the Model

The Watts-Strogatz Model interpolates between a regular graph and a random graph. For intermediate values of p we observe:

- high clustering (inherited from the initial regular graph)
- low average path length (deriving from the rewiring)

In practice the few random connection we are adding make it much easier to move around the network.



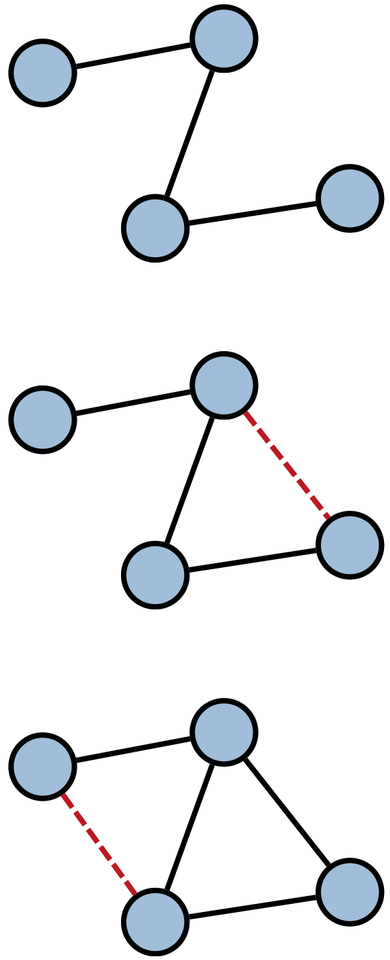
Triadic Closure Mechanism

The Watts-Strogatz Model reproduces real networks properties, however it is not very realistic:

- in real life we don’t know much about the full network, we tend to link more with close people (i.e. friends of friends)

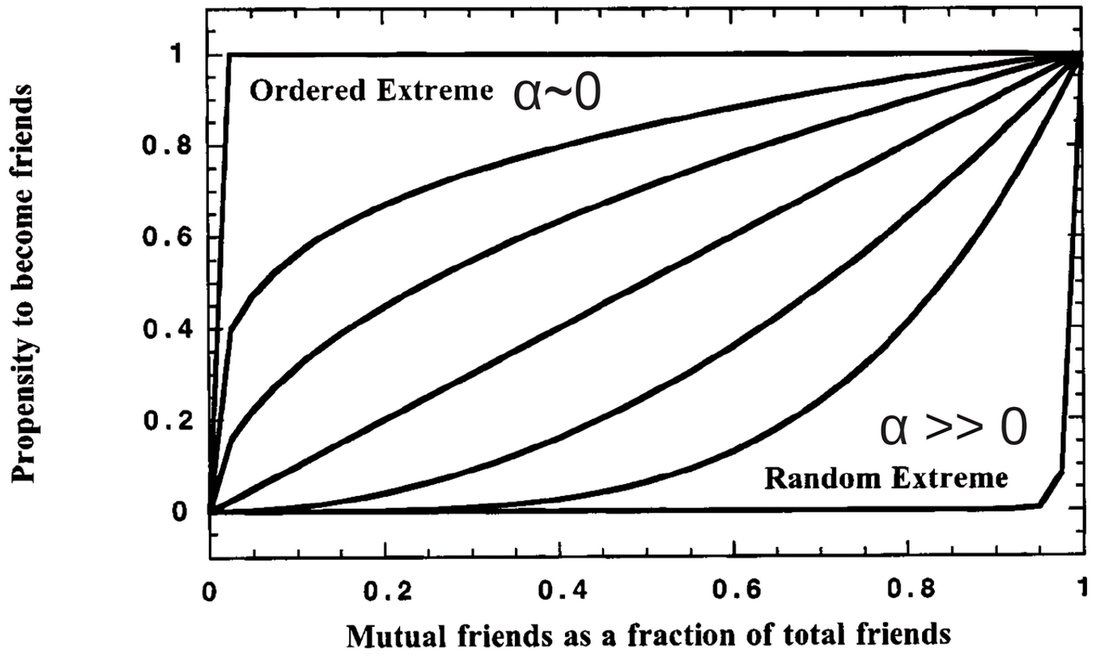
We can achieve similar networks without relying on the random rewiring. Instead we can perform a rewiring based on triadic closure

- the idea is that nodes having a “common friend” are more likely to link
- we always start with a regular ring
- we add new links with a probability that depends on the number of shared friends



Propensity to Triadic Closure

$$R_{i,j} = \begin{cases} 1 & m_{i,j} \geq k \\ \left[\frac{m_{i,j}}{k}\right]^{\alpha}(1-p)+p & k > m_{i,j} > 0 \\ p & m_{i,j} = 0 \end{cases}$$



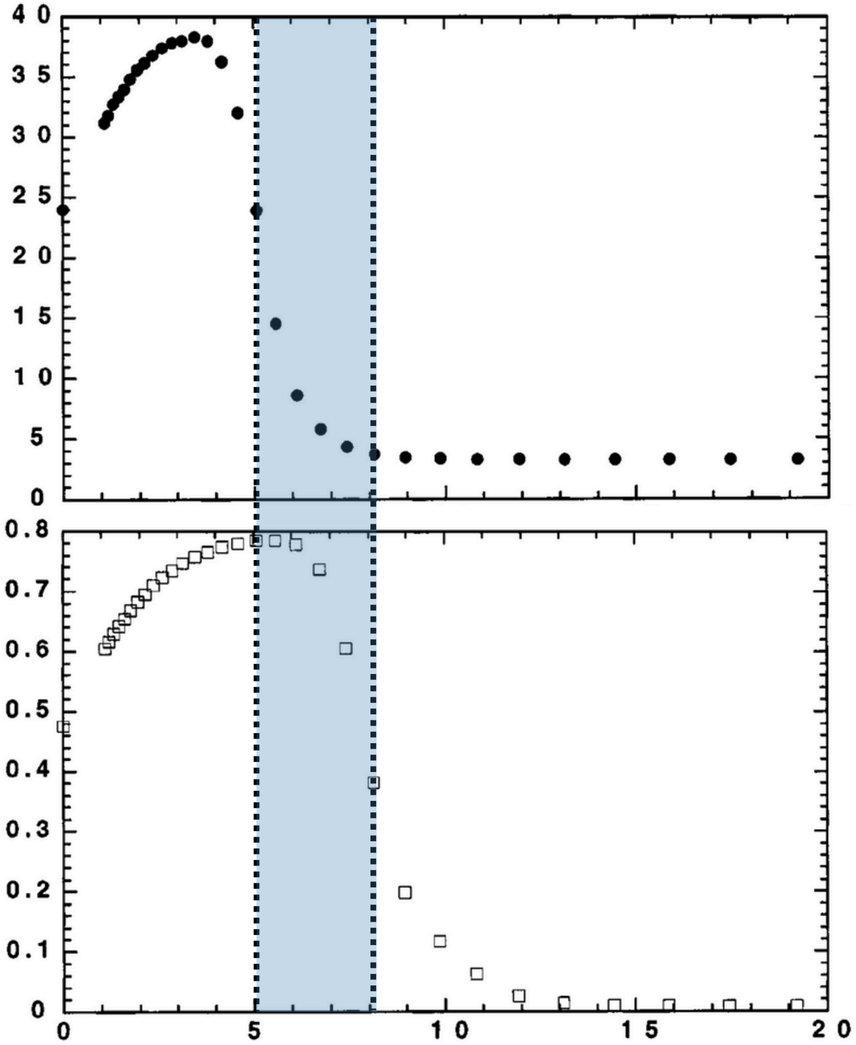
The model works as it follows:

- Start with a ring of n nodes
- For each pair of nodes (in random order):
  - Calculate number of shared friends $m_{i,j}$
  - Calculate propensity to connect $R_{i,j}$ based on $m_{i,j}$
  - Connect them with probability $R_{i,j}$
- p gives the probability to connect even in absence of mutual friends
- α sets the relevance of the common friend mechanism

Properties of the Model

Similarly to the Watts-Strogatz Model, we observe a sweet spot (in α) for which the model produces networks with both low path length and high clustering

- this is much more realistic than the Watts-Strogatz model
- the rewiring process is based on local characteristics of the network
- the process resembles what we humans tend to do in real life



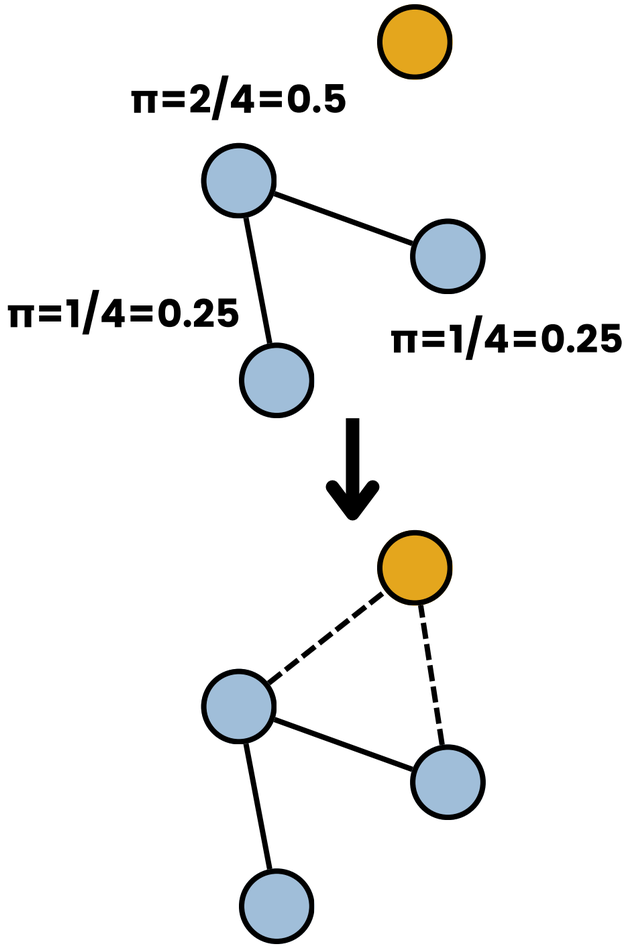
The Barabasi-Albert Model

The other property of real networks we want to explain is their power law degree distribution. The Barabasi-Albert model is a simple network growth process showing that scale-free networks can emerge from a simple mechanism

- we start with an initial network
- at each time step we add a new node
- this new node links to m existing nodes
- the linking probability π_i to link to node i is proportional to the node's degree

$$\pi_i = \pi(k_i) = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2mN}$$

https://sarah37.github.io/barabasialbert/



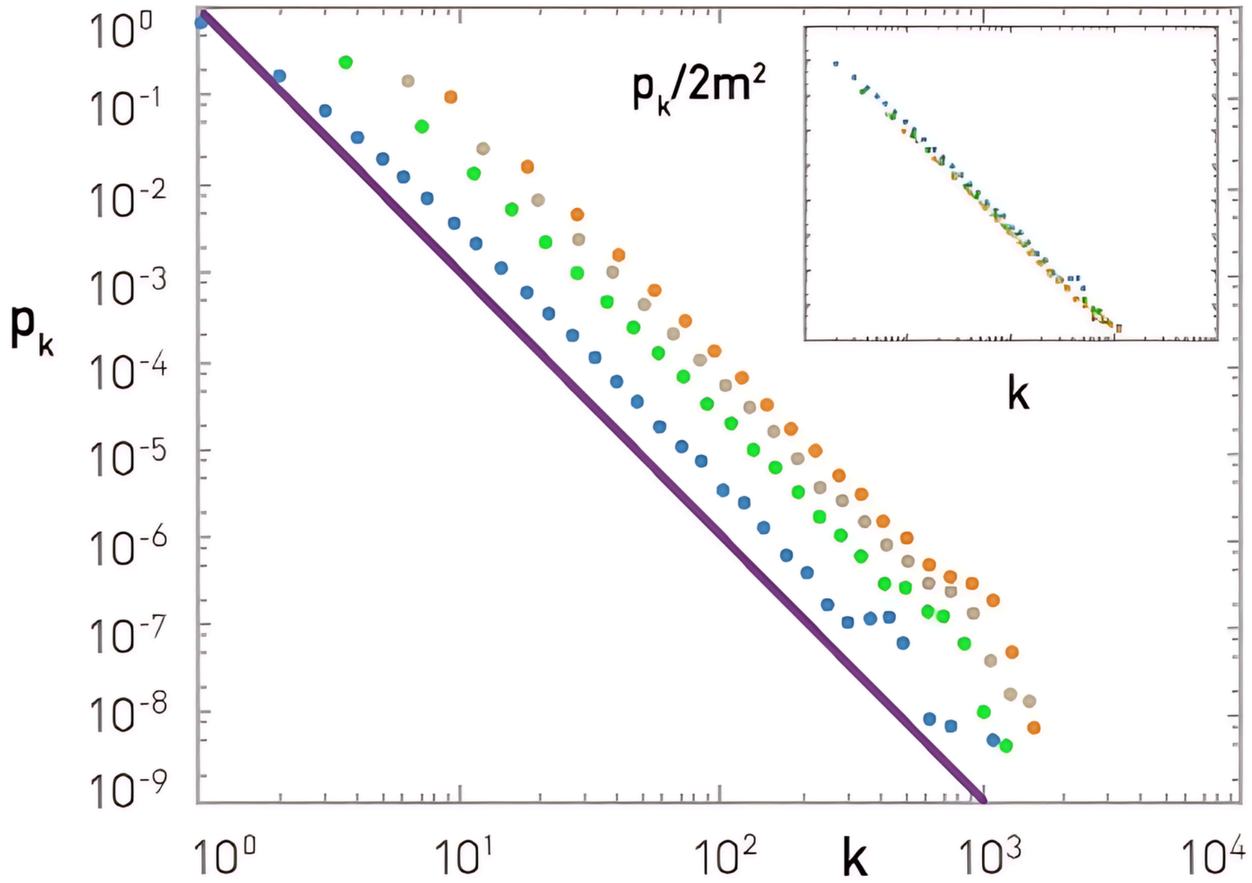
Scale-Free Degree Distribution

The Barabasi-Albert model generates scale free networks

- the power law exponent is independent of
  - the number of links m
  - the initial network
- the model asymptotically produces a degree distribution with exponent -3

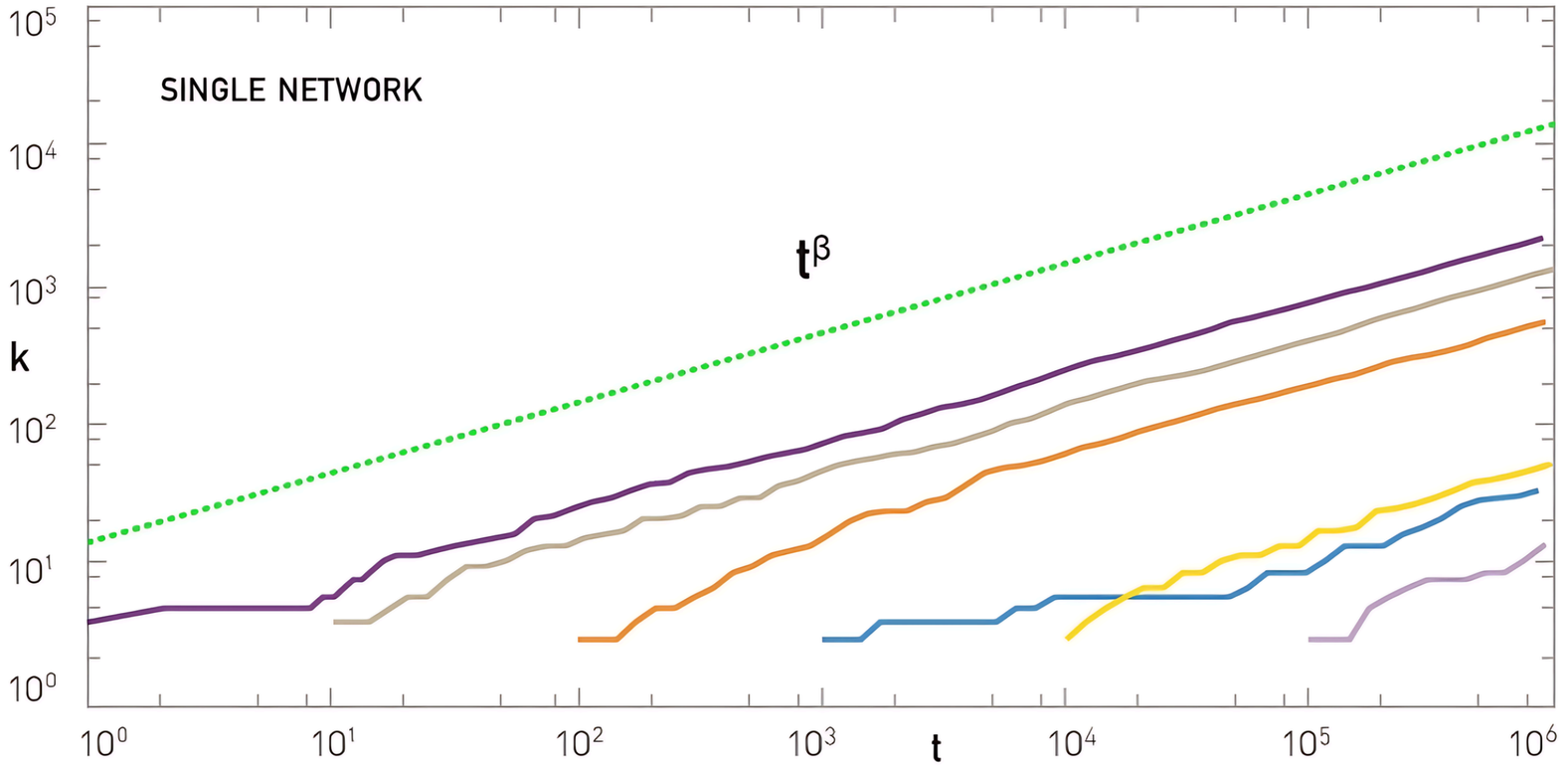
$$P(k) = \frac{2m^2}{k^3}$$

- small modifications allow to get any exponent >2

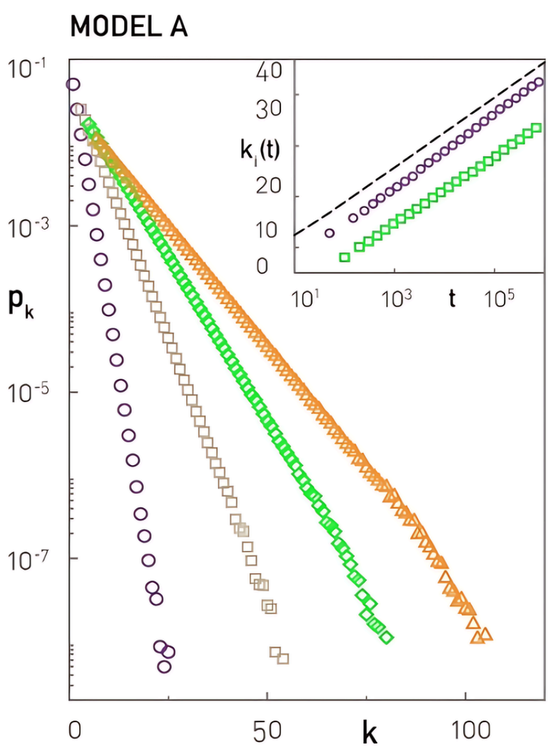


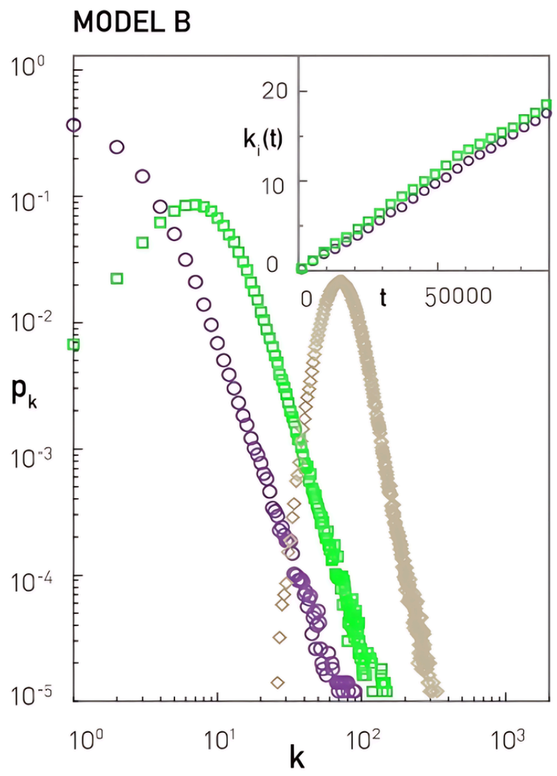
Rich-get-Richer Effect

In the Barabasi-Alber model, older nodes have an advantage over younger nodes. This is called Rich-get-Richer effect (or cumulative advantage)



Necessary Conditions





What are the necessary ingredients to get a scale-free networks? Are both growth and (linear) preferential attachment crucial?

- without (linear) preferential attachment we get random networks (exponential degree distribution)
- without growth (no new nodes) the distribution never reaches a stationary state and peaks on a specific value (depending on N)

The Vertex Copy Model

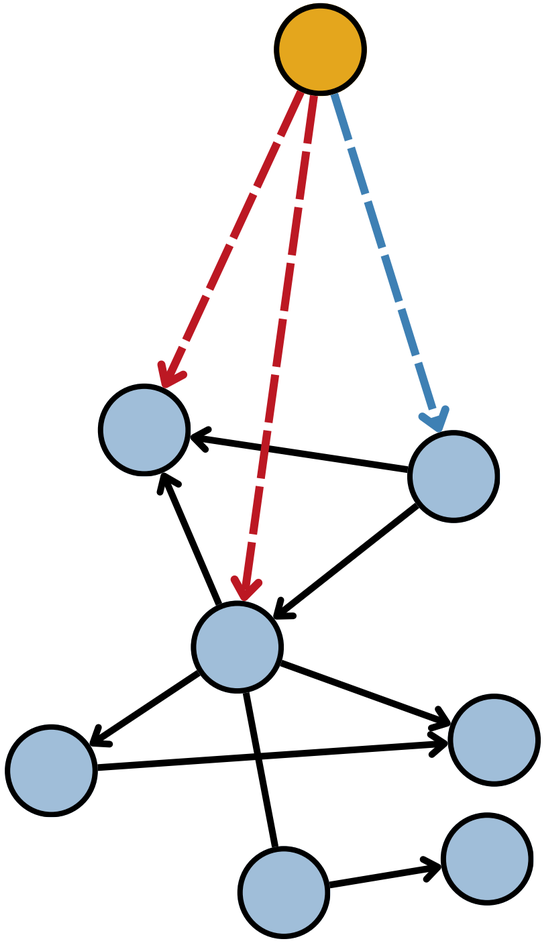
The Barabasi-Albert model is not very realistic:

- in order to compute the linking probability we have to know all degrees
- in most situation we can only observe a very limited portion of a network

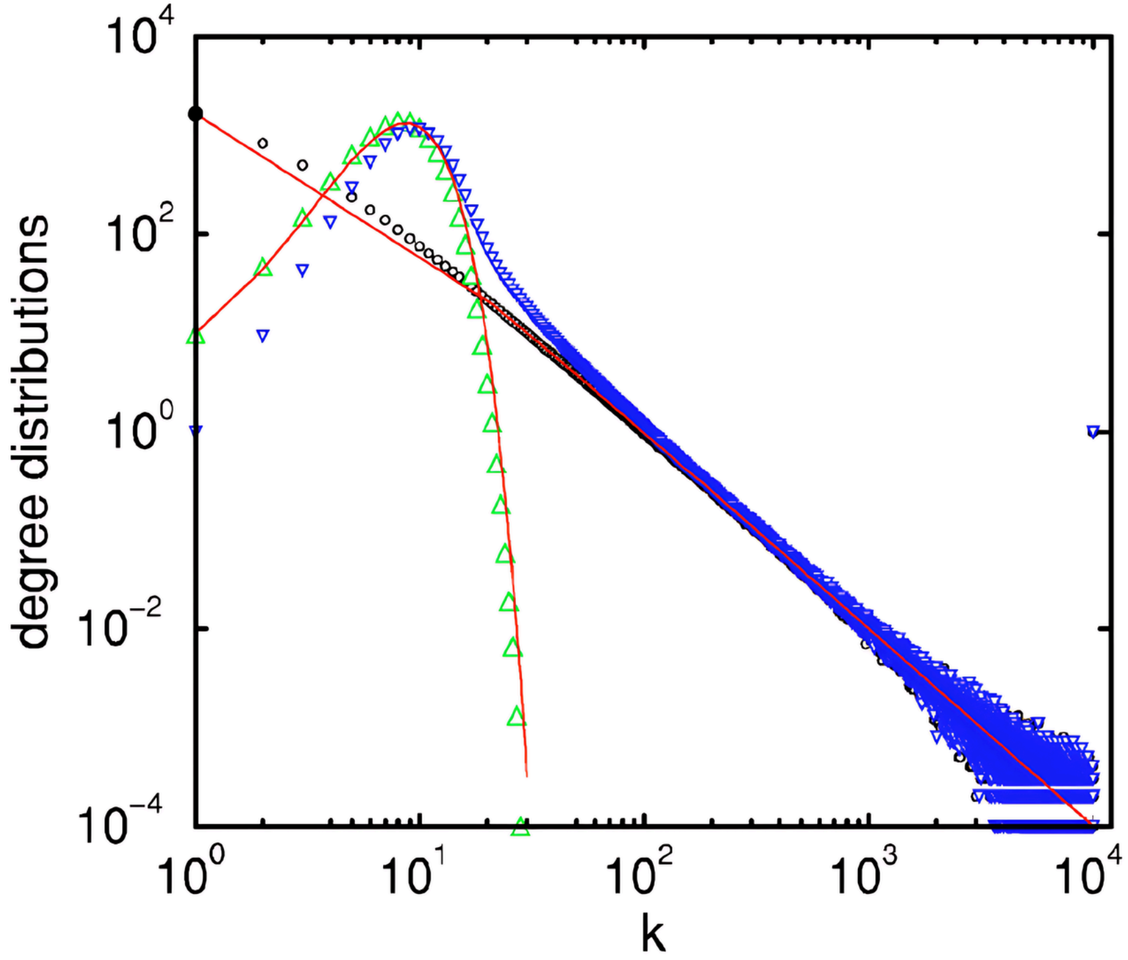
In the Vertex Copy model these limitations are overcame by exploiting a more local mechanism

- at each time step a new node is added
- this node links to a random node (blue arrow)
- it then copied all the connections of the node it has linked to (red arrows)

In this way we only need to know the local structure around a node.



Degree Distribution



The Vertex Copy model produces directed networks. The relevant property to look at is the in-degree (incoming connections)

- the out-degree distribution is peaked
- the in-degree distribution is a power law with exponent -2

This implies that it is possible to obtain scale free networks even if only local information is used. The edge copy mechanism is creating a sort of proxy of the linear preferential attachment.

Conclusions

Networks Basics

Many systems can be described in terms of networks. We introduced the main properties of networks (degree, clustering, diameter, centrality measures).

Real World Networks

Real world networks are characterized by the small world property but also by a high global clustering coefficient.

Network Formation Models

We saw that real networks' properties can be obtained using simple and local growth mechanisms.

Quiz

- What are some networks that we use everyday?
- Do you have any real life example of the small world property?
- Is there any flaw in Milgram’s experiment?
- Do you know any scale free network?
- What are the implausible assumptions of the Watts-Strogatz model?
- What about the Barabasi Albert model?
- Which model better describes how an online social network work?