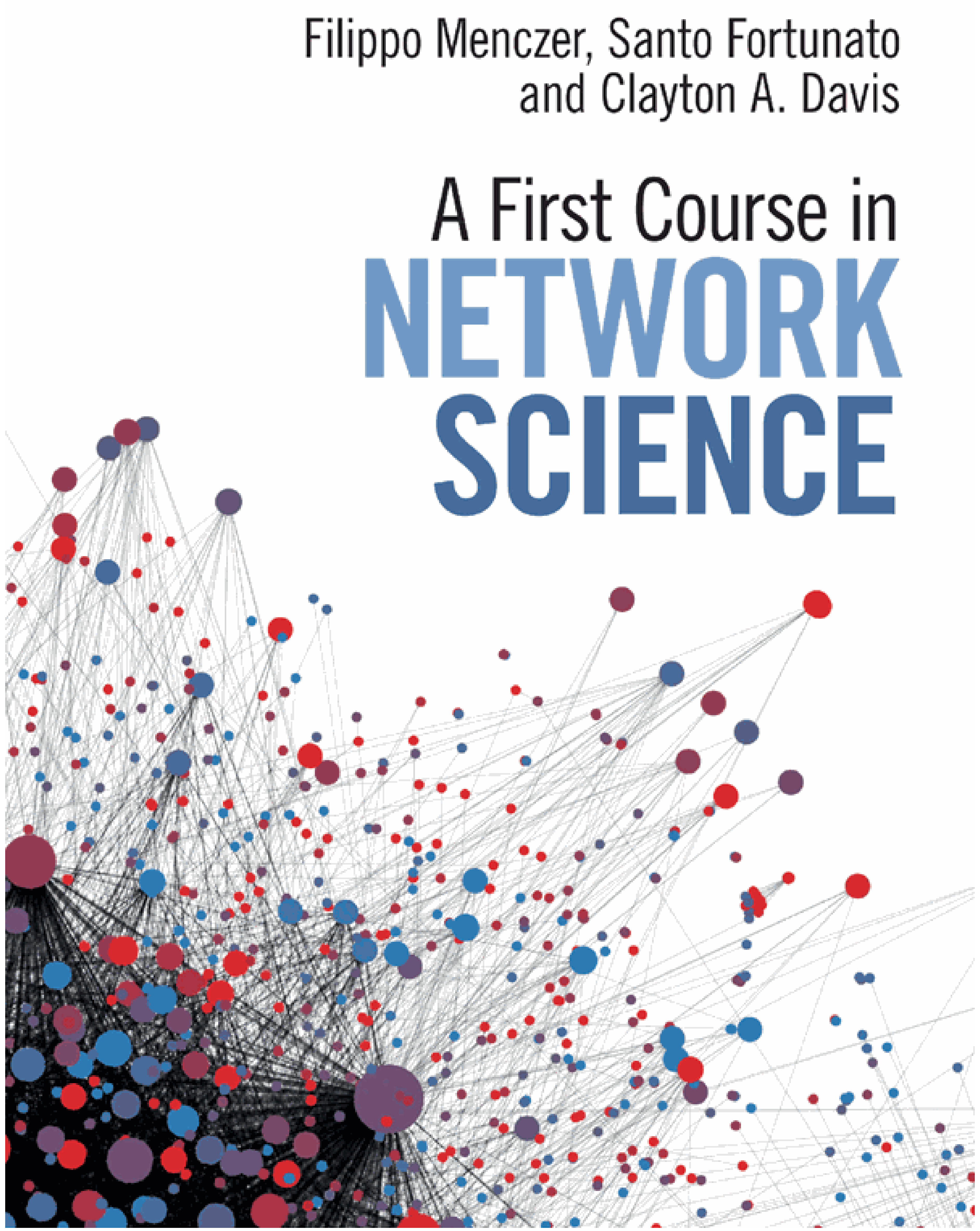
netherlands eScience center

First course in Network Science
Section 5.1

The content of this presentation is based on the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

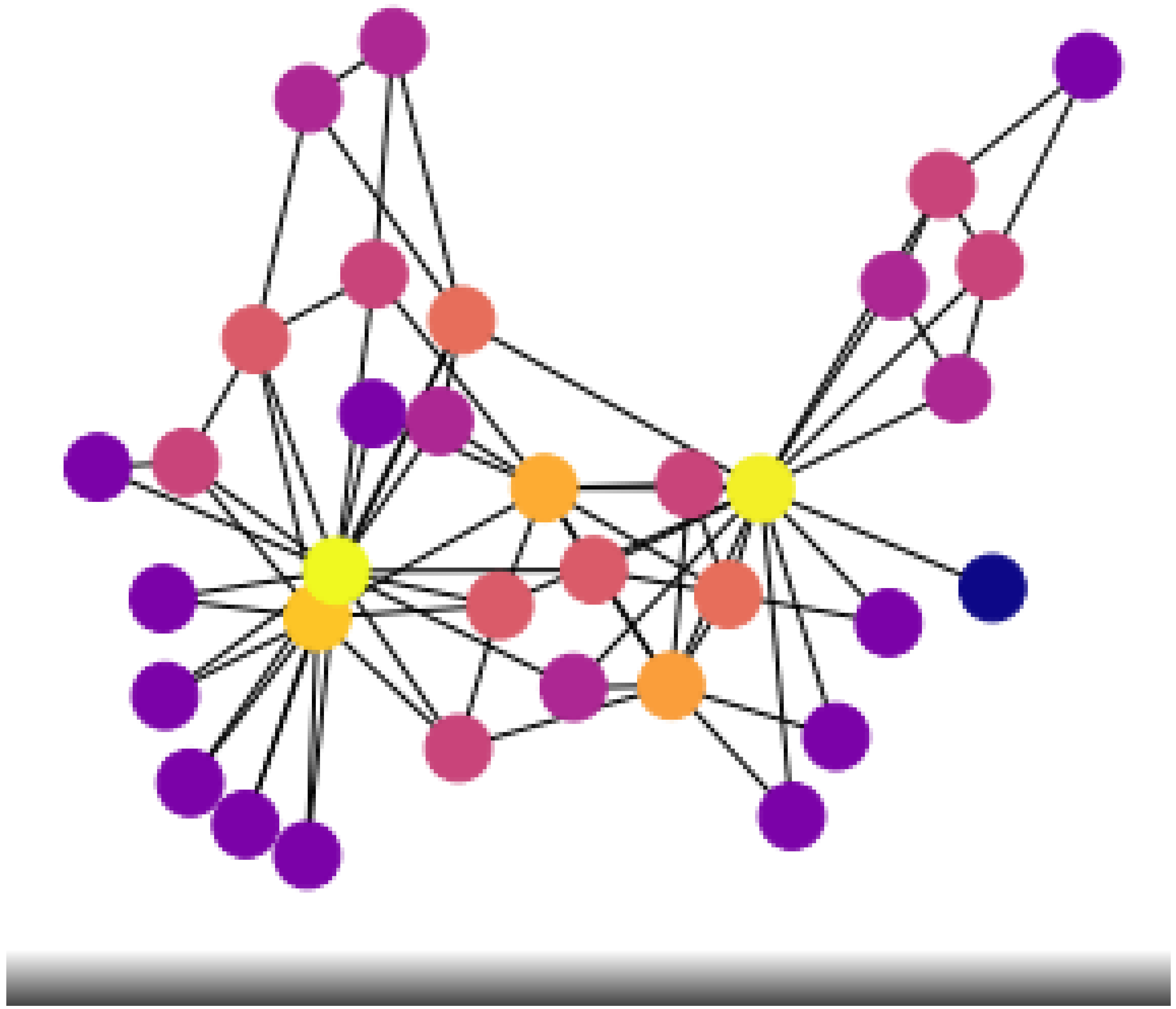
Sarah Alidoost
Analytics SIG, 12/04/2021


Filippo Menczer, Santo Fortunato
and Clayton A. Davis
A First Course in
NETWORK
SCIENCE

5 Network models / introduction

Common properties of real networks:

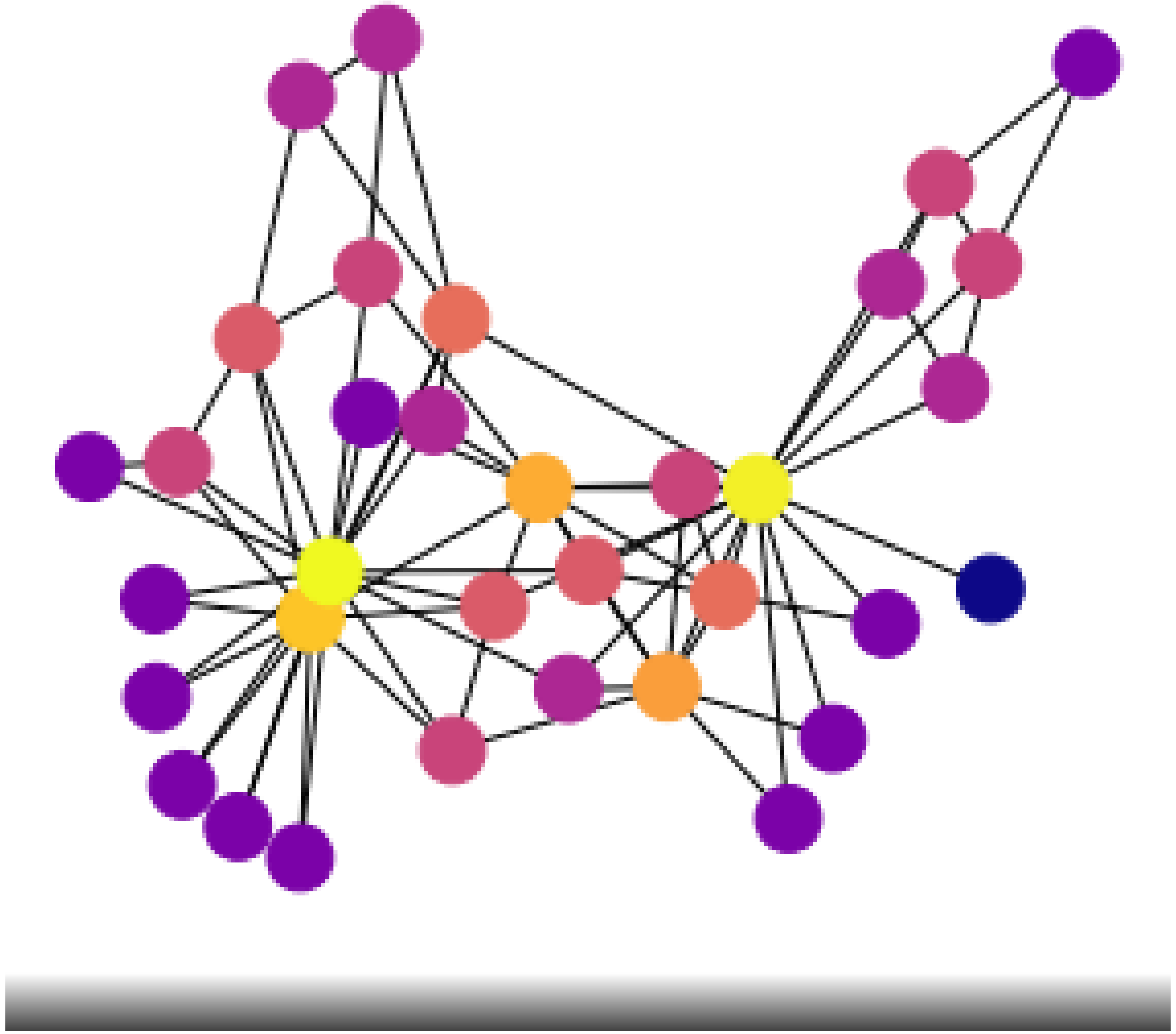
- Short paths,
- Many triangles,
- Degree and weights.

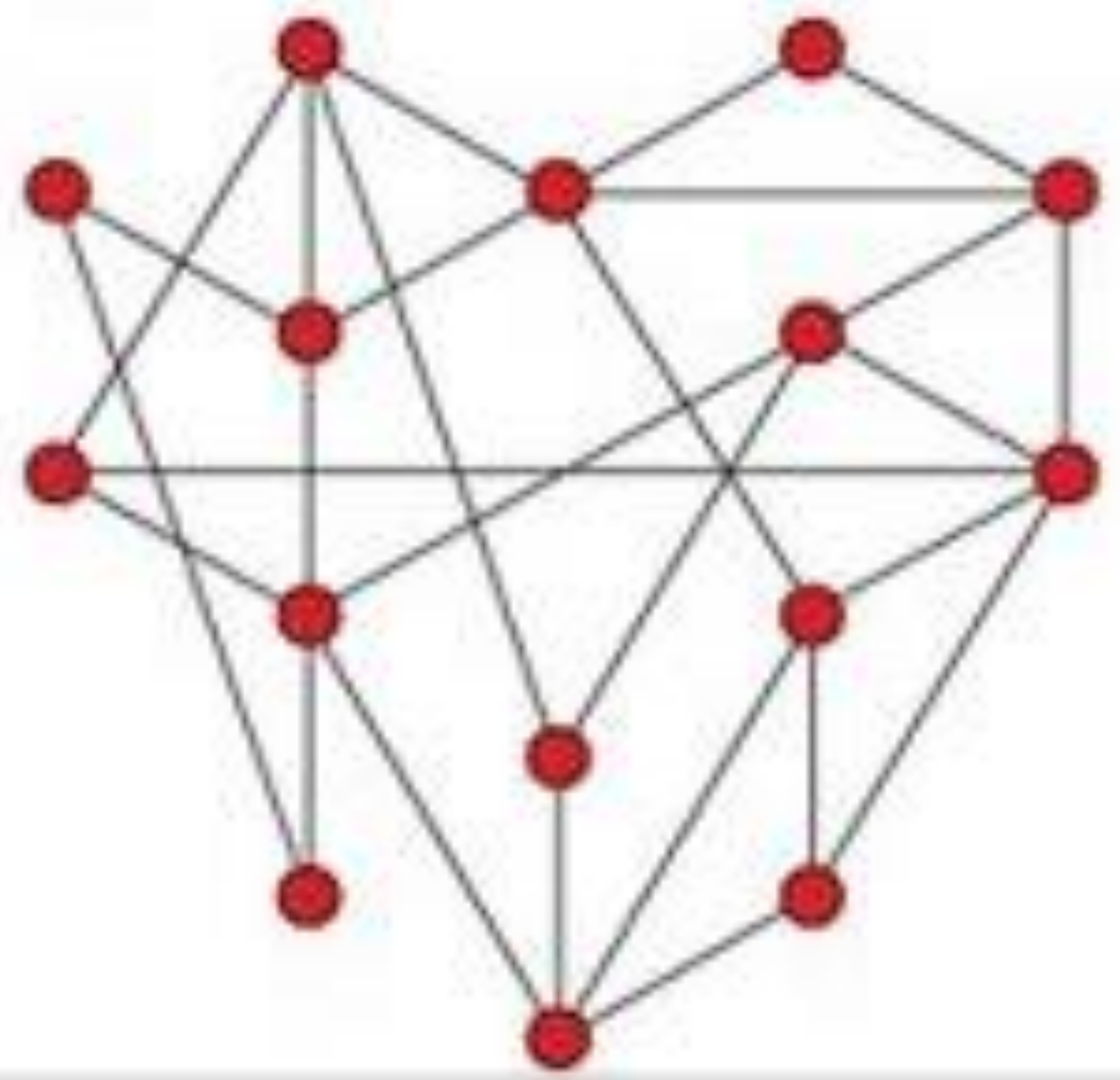


5 Network models / introduction

Where such properties come from:

- How do nodes choose their Neighbours
- How are hubs generated?
- How are triangles formed?

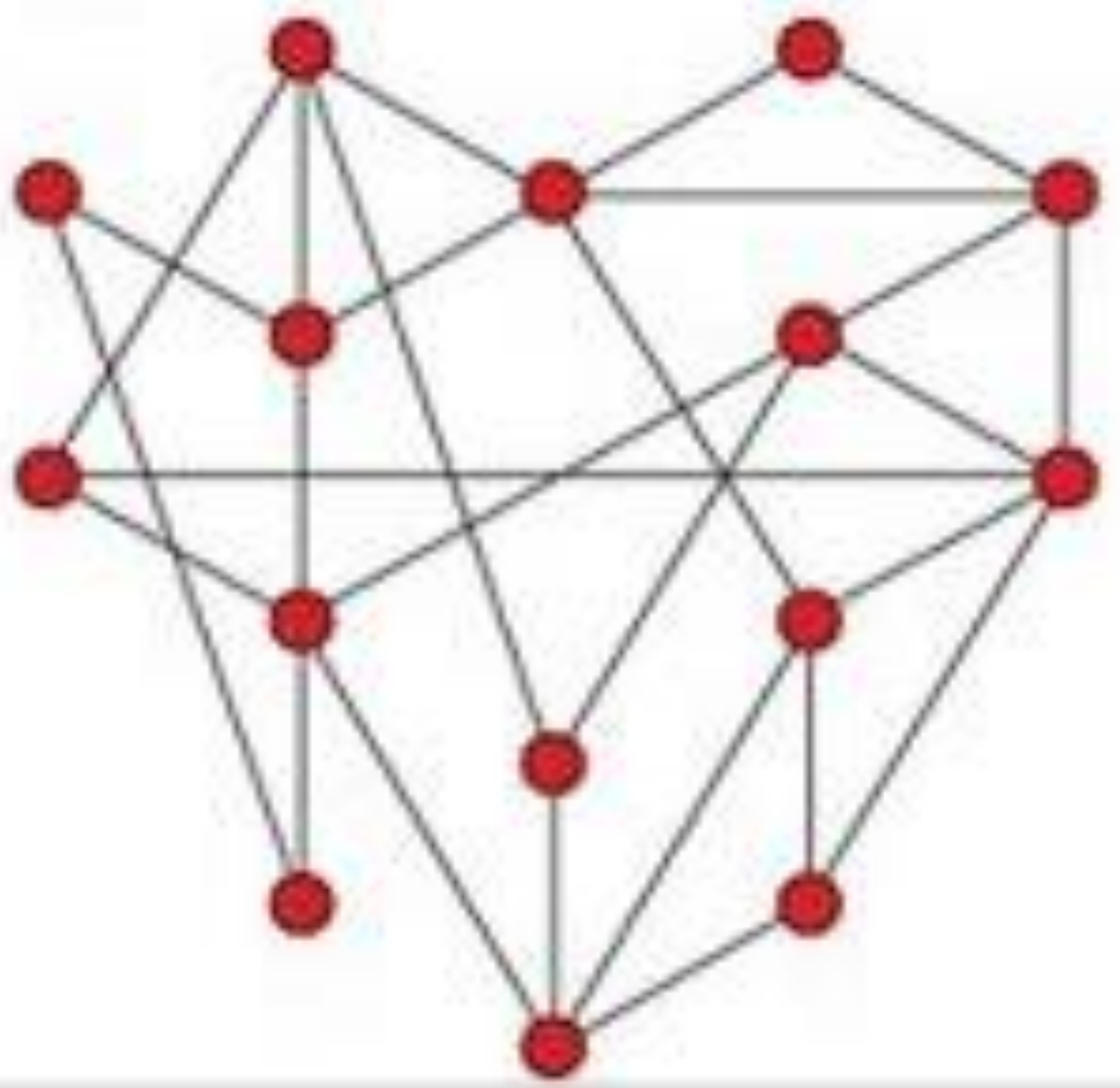




5.1 Random Networks / introduction

Random network: One of the classis models in network science

- 5.1.1 Density
- 5.1.2 Degree distribution
- 5.1.3 Short paths
- 5.1.4 Clustering coefficients



5.1 Random Networks

Egalitarian approach: to place links between randomly selected pairs of nodes.

A network built in this way is called a **random** (Erdos-Renyi) network.

Gilbert model:

- number of nodes *N*,
- a link probability *p*.

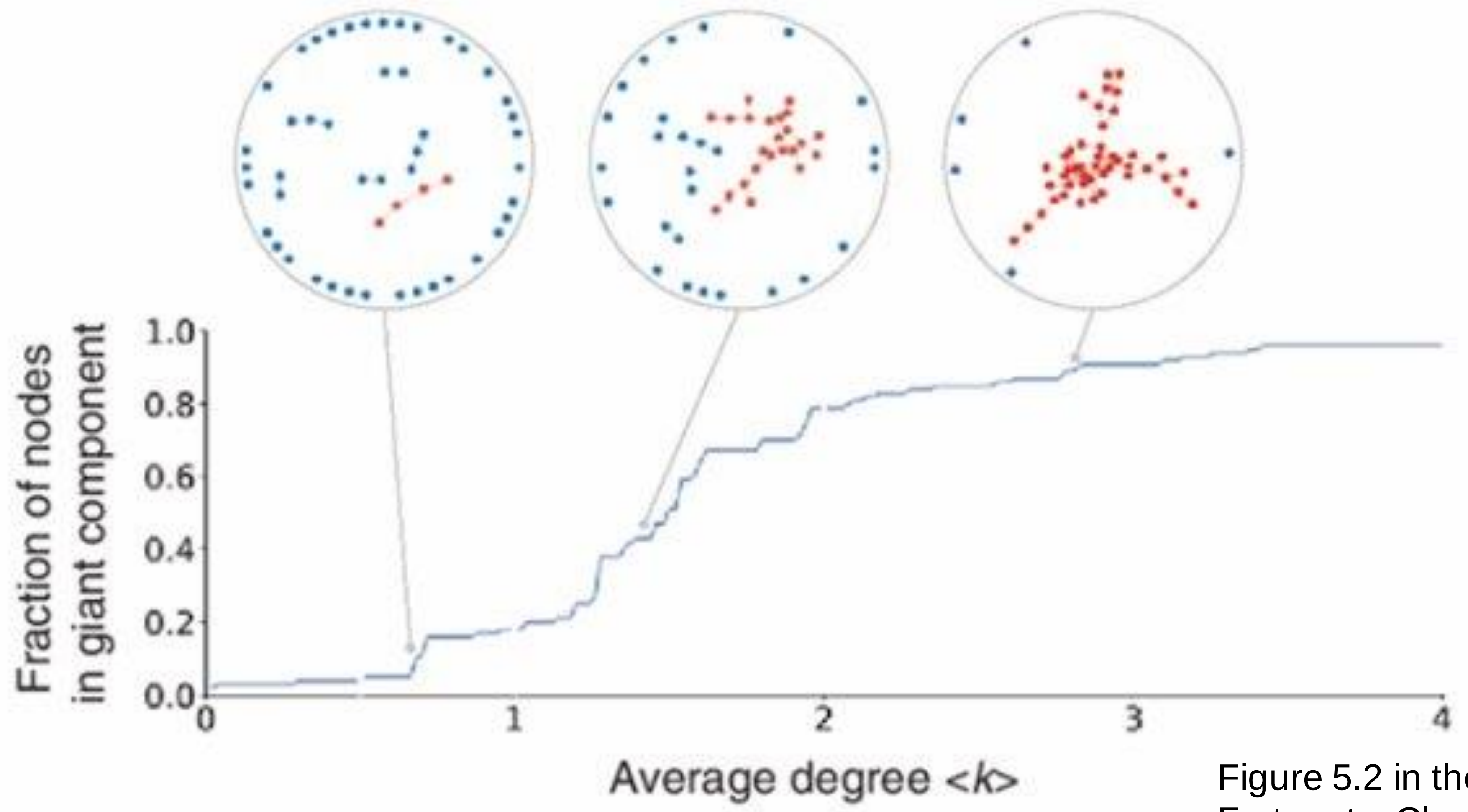


Figure 5.2 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

5.1 Random Networks

A giant component is formed when <*k*> = 1 i.e. each node has one neighbor on average.

The giant component is very small before <*k*>=1 and grows rapidly with the average degree.

5.1.1 Density

- The expected number of links in a random network: $$\langle L \rangle = p\binom{N}{2} = \frac{pN(N-1)}{2}.$$
- The expected average degree <*k*>: $$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1).$$
- The expected density is: $$\langle d \rangle = p.$$

Real networks are usually sparse (very small density).

For a random network to be a good model, the link probability should be close to zero.

5.1.2 Degree distribution

The probability that a node has k neighbors is given by a **binomial distribution**:

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

In many **real sparse** networks, the binomial distribution is approximated by a **bell-shaped distribution** with mean $\langle k \rangle$ and variance $\langle k \rangle$.

The **average degree** is a good statistical descriptor of the distribution.

- Figure **a**: degree distribution in a random network.
- Figure **b**: the degree distribution of the world flight network.
- Figure **c**: comparison between **a** and **b** in logarithmic scale.

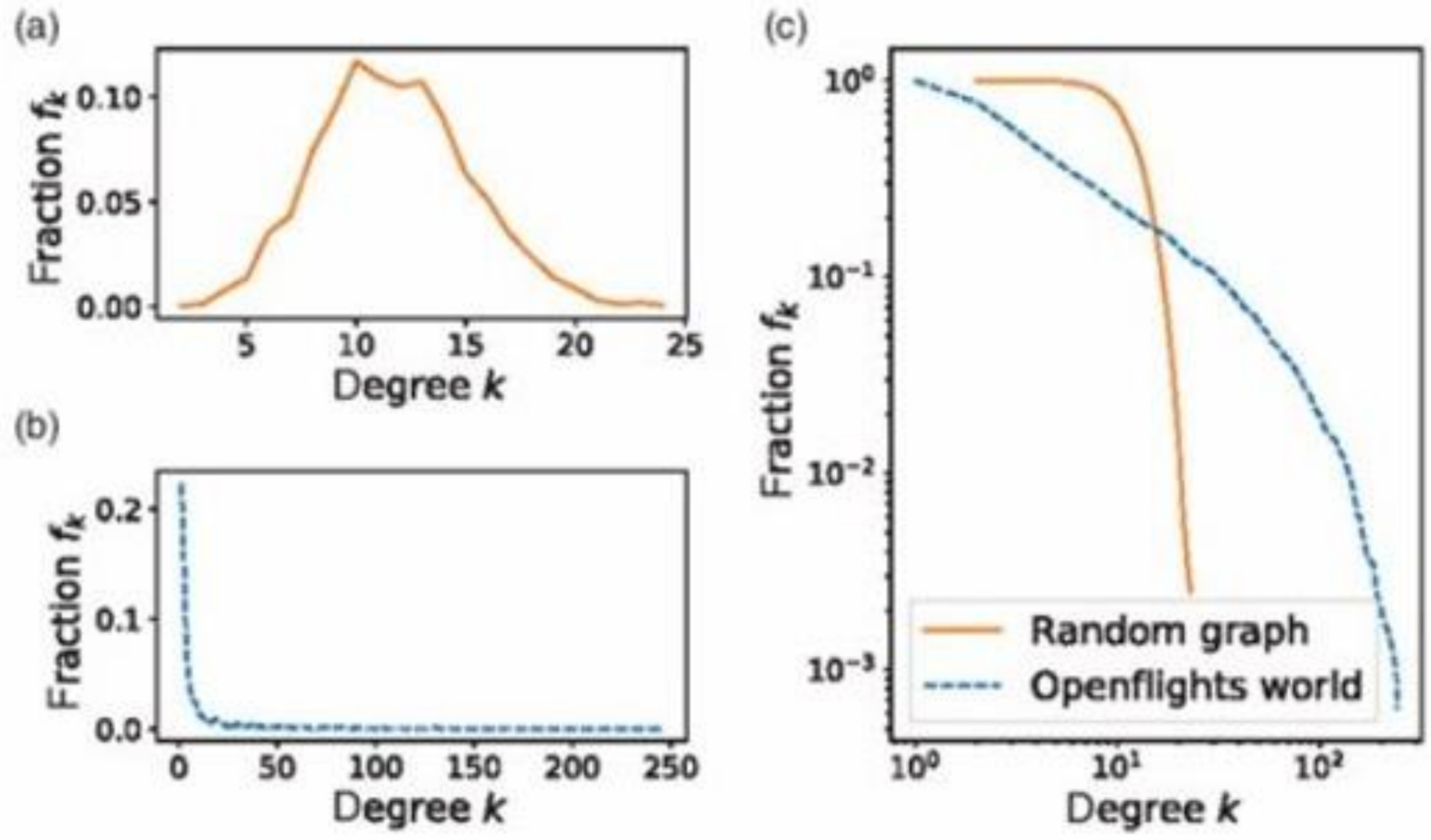


Figure 5.3 in the book: A First Course in NETWORK SCIENCE. Filippo Menczer, Santo Fortunato, Clayton A. Davis, ISBN: 9781108471138, Cambridge University Press.

The random network model does **not** provide a good description for real-world flight networks.

5.1.3 Short paths

Assume that nodes all have degree k=10.

If we start from a node within ℓ=1 step, there are k=10 nodes attach to it.

Each of them has k-1 neighbors and so on.

How far from a node do we have to go to reach all other nodes?

The diameter is:

$$\ell_{max} = \log_k N = \frac{\log N}{\log k}.$$

The fact that ℓ_{max} in a random network is small compared to the N, means that the network has short paths.

Example: the world's netwrok of social contacts

Imagine that it is a random network.

If we take k=150, the average number of regular contacts that humans can maintain.

At distance 5, the number of reachable people is 150^5 approx. 75 billion, a factor of 10 larger than the world population.

So, in principle, we could reach any individual in five steps or less that is compatible with the **small-world** experiment (section 2.7 Six Degrees of Separation).

5.1.4 Clustering coefficient

- **Clustering coefficient** of a node: the fraction of the node's pairs of neighbors that are connected to each other.
- In a random network, *p is* the probability that a pair of neighbors of a node is connected.
- In a **real sparse** network, *p* and thus the clustering coefficient is very small.
- In contrast, we know that real social networks have a high clustering coefficient.

If we want to model a real network with a high coefficient, we need a model with some specific rules.

Summary:

- Random network is one of the classis models in network science.
- Real networks are usually sparse (very small density).
- In many real sparse networks, the distribution is approximated by a bell-shaped distribution.
- A random network has short paths.
- A random network cannot model a real network with a high coefficient.