netherlands eScience center

Network science book chapter 3.1-3.2

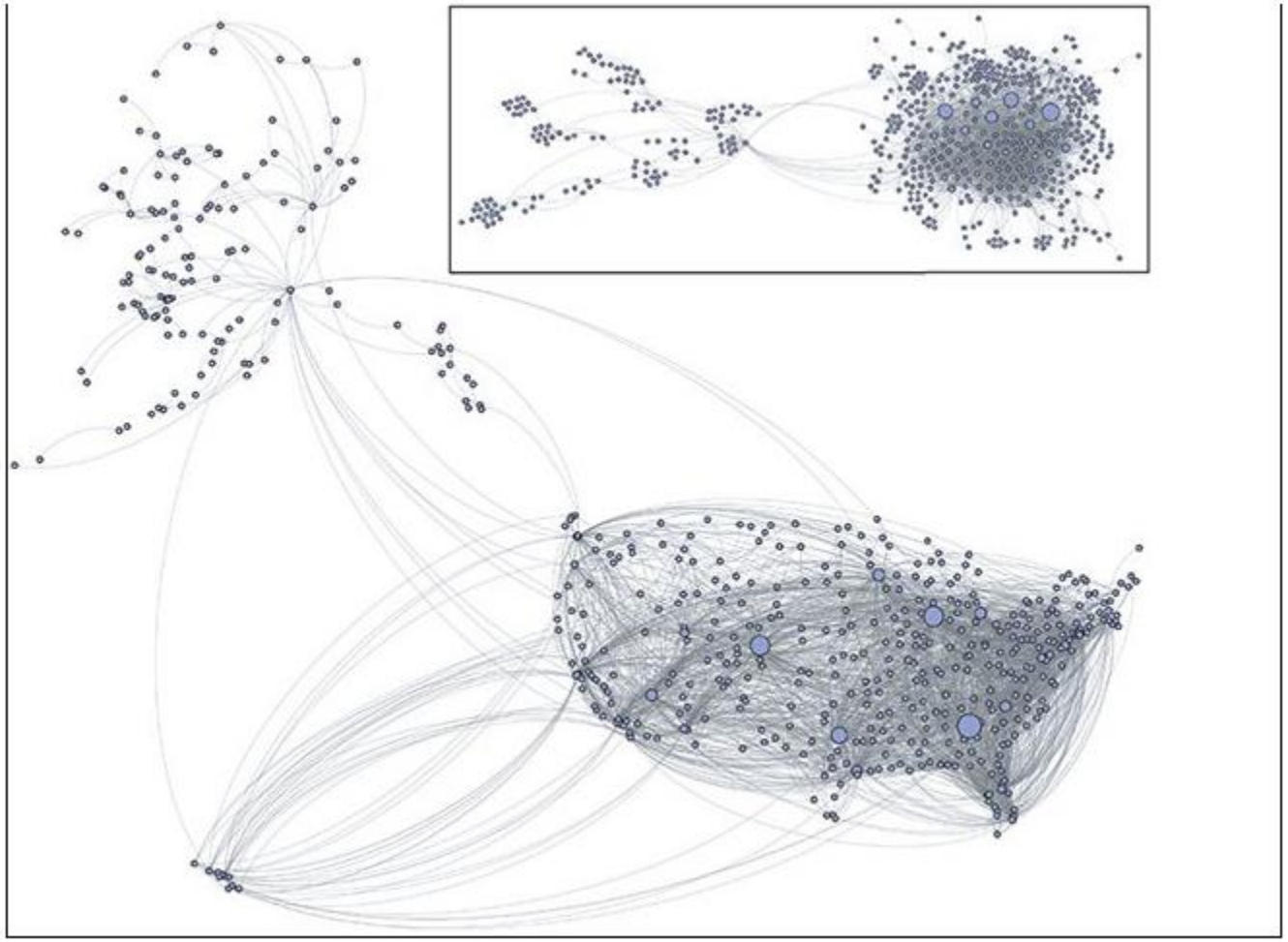
Dafne van Kuppevelt

Analytics SIG, 2021-2-15

All illustrations are copied from the book:

A First Course in Network Science - Clayton A. Davis, Santo Fortunato, Filippo Menczer

Hubs



netherlands eScience center

Centrality

- Importance of specific **node** or link in network
- Many different centrality measures, including:
  - Degree
    nr of links of the node
  - Closeness
  - Betweenness

Closeness centrality

Inverse of the sum of distances of node to all other nodes:

$$g_i = \frac{1}{\sum_{j \neq i} l_{ij}}$$

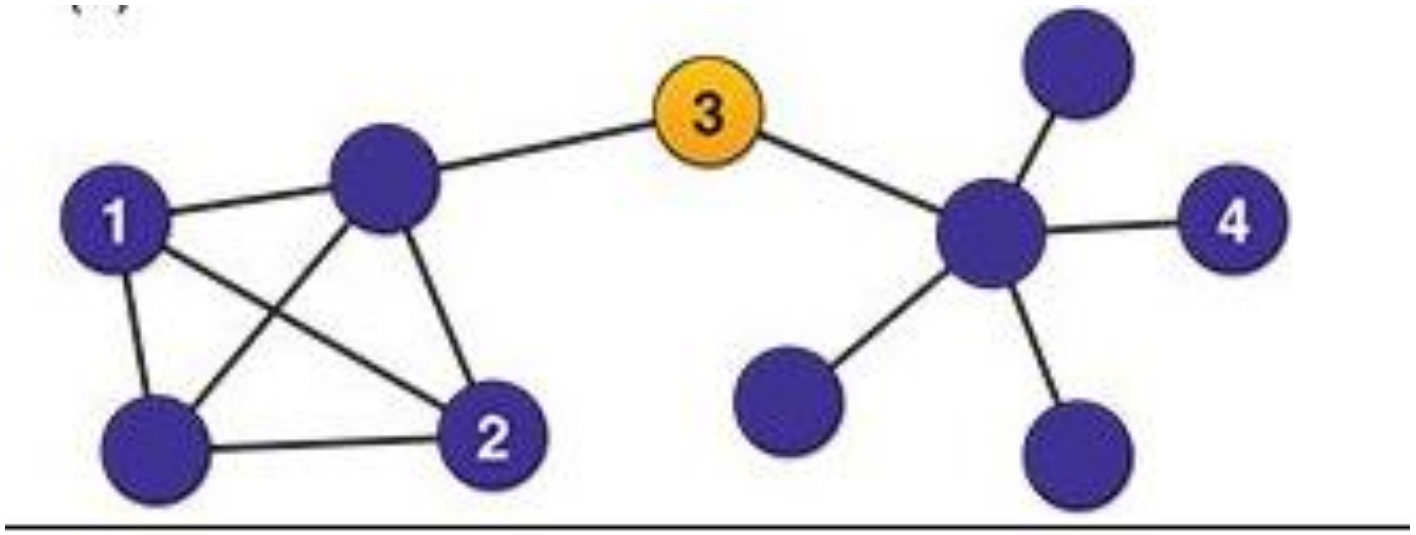
with l_{ij} the distance between *i* and *j*

In networkx:

```
nx.closeness_centrality(G, node)
```

Betweenness centrality

- How many shortest paths go through this node?
- Especially relevant in *diffusion processes*



Betweenness centrality (2)

Definition:

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

Where σ_{hj} is the total number of shortest path between *h* and *j* and $\sigma_{hj}(i)$ the number of these that pass through *i*

(similar: edge betweenness)

Maximum betweenness depends on network size, so can also *normalize.*

Exercise

E3.2: Undirected network with 100 nodes and 200 links. What is the average degree of nodes?

Exercise

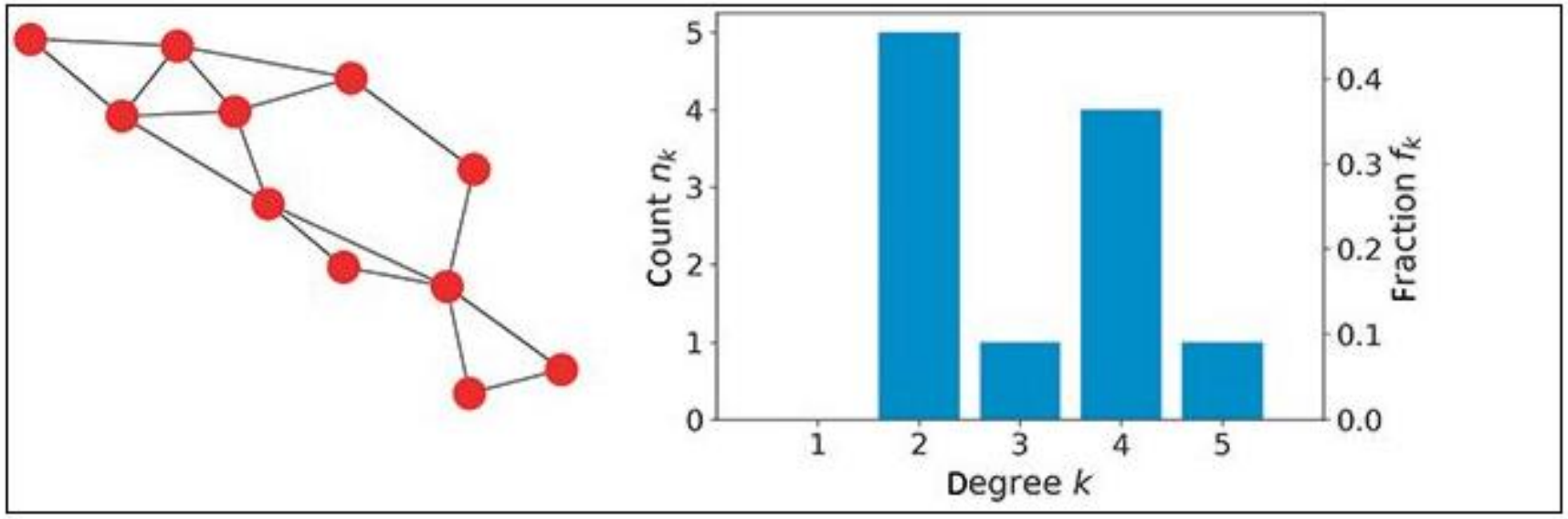
E3.2: Undirected network with 100 nodes and 200 links. What is the average degree of nodes?

Answer: 2 * 200 / 100 = 4

Note that each link contributes on both its sides to the total degree.

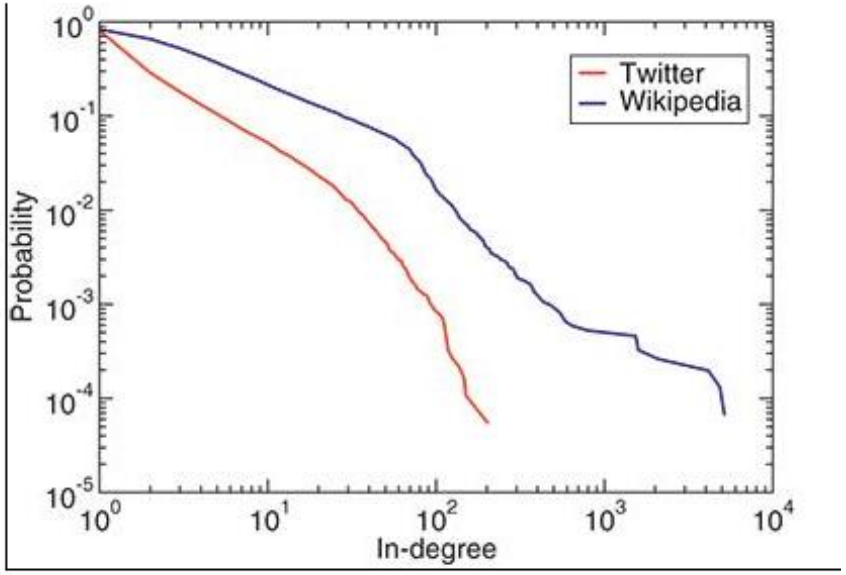
Centrality distributions

- In large networks, the *distribution* of centrality values is more interesting than individual nodes / edges
- Can draw conclusions on homogeneity / heterogeneity of network and network structure



Degree distribution of real networks

- Many real networks show heavy tail distribution



- Heterogeneity parameter:

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle^2}$$

Where $\langle k^2 \rangle$ is the average squared degree and $\langle k \rangle$ is the average degree of the network.

Exercise

E3.3: Consider a network formed by 250 students in a dormitory. The links in th network represent roommate relationships. In this dorm, the rooms are mostly double occupancy with a few triples and quads.

1. Is this graph connected?
2. What is the mode (most frequent value) of the node degree distribution?
3. How many nodes are in the largest clique?
4. Would you expect this graph to have any hubs?

Exercises with networkx

See the book for more exercises with networkx.

And the Tutorial