Part III: Networks

Michael Small

Complex Systems Group
School of Mathematics and Statistics
The University of Western Australia

Mineral Resources
Commonwealth Scientific and Industrial Research Organisation
Australia

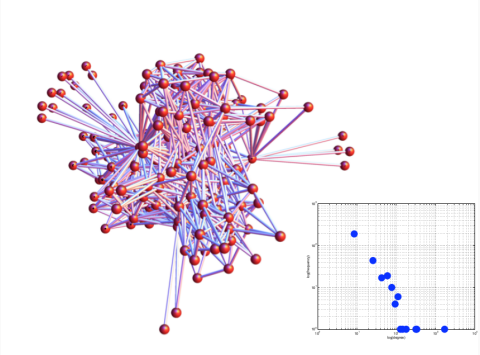
Discussion

Where do you have networks in your business?

Network Quantification

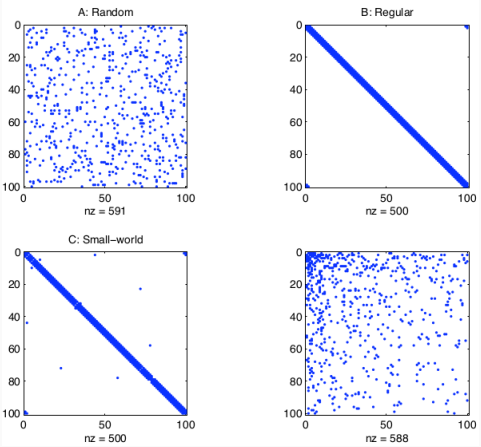
Quantifying Topology of Networks

- Degree sequence
- Degree histogram
- Path-length
- Clustering
- Assortativity
- Betweenness Centrality
- Eigenvalue Centralitiy
- Modularity and Communities
- Hubs and richclubs
- Robustness and fragility
- Motifs and superfamily

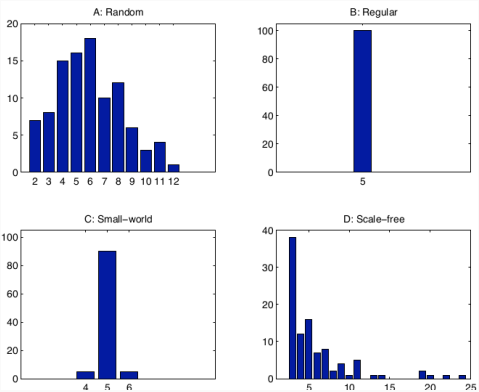


Adjacency matrix

Let A be a binary $N \times N$ matrix with $a_{ij} = (A)_{ij} := 1$ iff there is a link from node-i to node-j (otherwise $a_{ij} = 0$). Generalise to weighted networks with $a_{ij} = w_{ij}$ is the weight, and directed networks with an asymmetric A.
The graph Laplacian L is defined so that $L = A$ except for the diagonal: $(L)_{ii} = -\sum_{j \neq i} a_{ij}$.

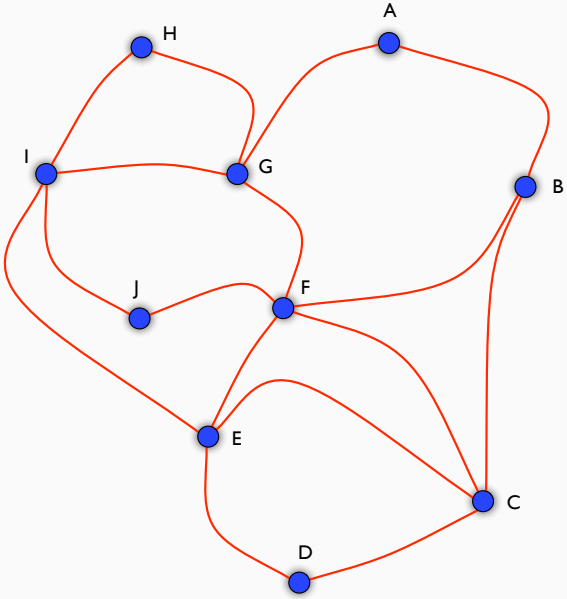


Degree distribution

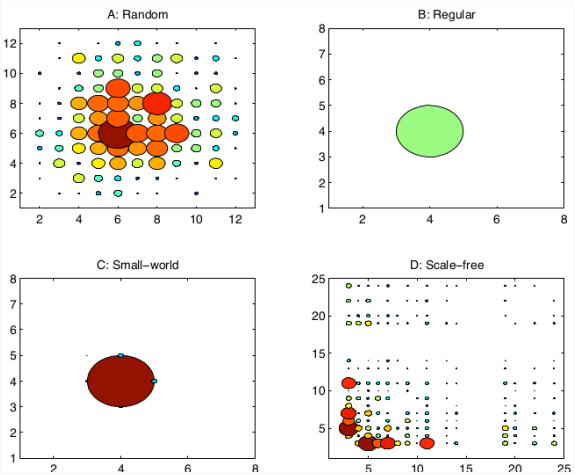


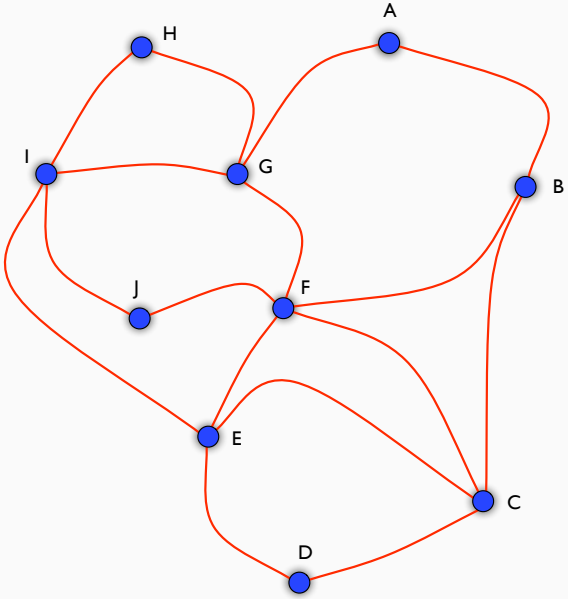
Compute the number of nodes n_k (or the probability p_k) with degree k.

- *Path-length:* The path-length is the shortest path (number of edges traversed) between two nodes.
- *Diameter:* The maximum path-length

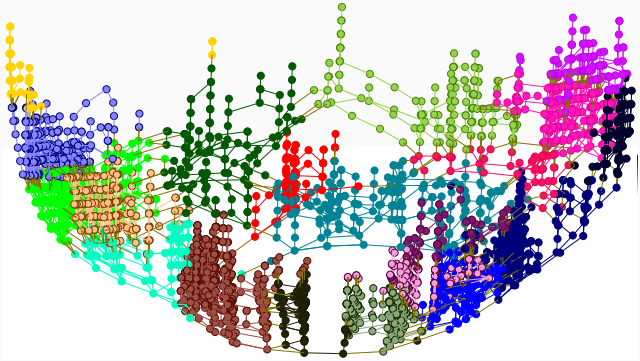


- *Clustering:* The number of triangles — the probability of neighbours being neighbours.
- *Assortativity:* The linear (Pearson) correlation between pairs of nodes with a give property





- *Betweenness Centrality:* The number (fraction) of shorter paths passing through a given node
- *Eigenvalue Centrality:* (AKA: Google's PageRank) Eigenvalue decomposition of the Laplacian



- *Communities* should have more links between members than between communities.
- *Modularity* Q measures this:

$$Q = \frac{1}{2m} \sum_{i,j} \left(a_{ij} - \frac{k_i k_j}{2m} \right) \delta(i,j).$$

- *Hubs* are nodes with the highest degree
- *Rich-club* is the connection between the hub nodes and the tendency of hub nodes to be connected to one another
- *Giant component:* The property that most of the nodes are connected (directly or indirectly) to one another.
- *Robustness:* The ability of a network to maintain it's giant component even after random removal of a relatively large number of edges (or, equivalently, nodes).
- *Fragility:* The corresponding inability of a network to maintain the rich club under targeted removal of edges (or nodes)
- *Motifs:* The structure of interconnection in sub-graphs of particular size k.
- *Super-familiy:* The relative frequency of all such sub-graphs for fixed k.

Pythonification

Discussion

Why would *you* need to know any of this stuff for your business?

Discussion

Why would *you* need to know any of this stuff for your business?

Exercise

Everything we've discussed here is implemented and easily usable within the `networkx` package. Refer to the notbook

References

Sources

- C. Braham and M. Small. “Complex networks untangle competitive advantage in Australian football” *Chaos* **28** (2018) 053105.
- Barabasi, Network Science, http://networksciencebook.com/
- https://www.python-course.eu/networkx.php
- M. Small. “Dynamics of biological systems” (network chapter).

michael.small@uwa.edu.au
michael.small@csiro.au