2.3 Complex Networks

2.3.1 Overview

The first work on graph theory was carried out by Leonard Euler when he considered the Seven Bridges of Königsberg problem (1736), when the citizens of the city wondered whether it was possible to cross all the bridges in the city only once. His key insight was to conceptualize the land masses in the problem as "vertices" and the bridges as "edges" but it was not until probabilistic methods were applied to these graphs that something resembling modern network theory was established (Erdős and Rényi, 1959). However, while the Erdős-Rényi model proved to be effective for learning about topological features of networks, it did not manage to encapsulate many attributes that had been noted in real-world networks. For example, its edges follow a Poisson distribution instead of a power-law distribution and there is little local clustering among the nodes. It was not until the late 90's that two realistic methods for network creation were discovered. The first, the Watts-Strogatz model was able to account for the "Small-world" property (Milgram, 1967), meaning that the graph has a low characteristic path length but also a high local clustering coefficient (Watts and Strogatz, 1998). The second, the Barabási–Albert model captured the other missing attribute, vertex connectivities following a scale-free power-law distribution (Barabási and Albert, 1999).

2.3.2 Important Concepts

Network

A network can be considered as two sets. A set of nodes, N, and a set of links, L, between the nodes. N is the size of the network and each node is labeled with a number, i, from 1 to N. The degree, k, of a node specifies the amount of links it has, k_i therefore represents the degree of the i^{th} node.

Characteristic Path Length

The path length, d, between two nodes is the shortest distance between them. The characteristic path length, ℓ or CPL, is an average of these path lengths. It is of interest to us as it can be used to quantify global efficiency and parallelism of information propagation.

Clustering Coefficient

The local clustering coefficient, C_i, is a measure of the density of links between the neighbours of the node.

$$C_i = \frac{2L_i}{k_i\left(k_i - 1\right)} \qquad \textbf{(1)}$$

We take the average, C, of the local clustering coefficients to measure the overall amount of clustering in the network. It can be used to quantify the level of local connectivity and thus local efficiency of information propagation.

Erdős-Rényi Model

This model is a randomly connected graph, first formulated in 1959 (Erdős-Rényi, 1959). To initialise such a graph, we specify N and L, randomly selecting pairs of disconnected vertices and adding an edge until the graph satisfies the given definition. The Erdős-Rényi model usually has a low characteristic path length, however, the average clustering coefficient is usually much too low to be able to accurately model the level of modularity and clustering found in real-world networks.

Watts-Strogatz Model

To create a WS network, we must first specify how many nodes we wish the graph to have, the average degree of each node, and a special value called Beta, β. The β component is a probability that determines to what extent we move from a regular lattice towards a classic Erdős-Rényi graph.
These values must satisfy two conditions: $N \gg K \gg \ln N \gg 1$ and $0 \leq \beta \leq 1$.

We then create a regular ring lattice, with K/2 links on each side of it. Such that, there is an edge between nodes i and j iff

$$0 < |i - j|\, mod(N - 1 - \frac{k}{2}) \leq \frac{k}{2} \qquad \textbf{(2)}$$

We then go through each node, i, redistributing links where $i < j$ with probability β to another node, k, that must satisfy $k \neq i$ and no link already exists between i and k.

This has the effect that if the β component was set to 0 then we would still have a regular lattice but if the β component is set to 1 we effectively have an Erdős-Rényi graph. Thus the Watts-Strogatz small world model can be seen as a graph that is, characteristically, between a regular lattice and a random graph with its variance and average clustering coefficient in between the two, where the extent to which this is the case is governed by the β component.

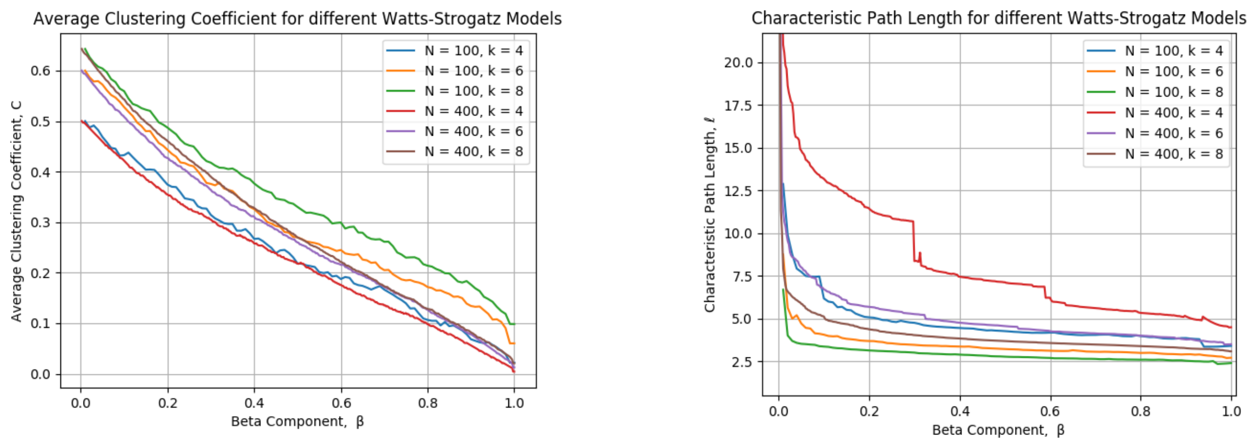

Figure 1: The Clustering Coefficient and Characteristic Path Length for various Watts-Strogatz networks during link redistribution.

Conceptually, we can imagine the regular ring lattice to be ordered and the process of redistribution to be increasing the level of disorder of the network. For certain values of β, we encounter networks with the "small-world" property, which is of particular relevance to the network of the brain. This is our first indication that there exists an area of interest between order and disorder.

Barabási–Albert Model

The creation of a Barabási–Albert model is considerably different to the two methods discussed so far. Barabási noticed that many real-world networks such as the Internet or the power grid are not initially created with all the nodes and links distributed. Instead, they rely on the concepts of growth and preferential attachment during their formation. Growth refers to the fact that in many networks, N grows over time. Preferential attachment is related to the phenomena that not all nodes are necessarily created equal. While later models, such as the Bianconi-Barabási model (Bianconi and Barabási, 2001), incorporate the idea of individual node fitness. We will take preferential attachment to mean that the more links a node has, the more likely it is that a link will be made between it and a new node. Thus, we have a positive feedback cycle that gives us the power law scaling that has been found in the brain (Kello et al., 2010). This is our first indication that relatively simple self-organising phenomenon can produce realistic features.

To create such a model, we initialise a network with m_0 nodes. We then attach one node to the network at a time, with m ($m \leq m_0$) links. $\Pi(k_i)$ is the probability that a link between a new node and node i, given as a function of the degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N} k_j} \quad \textbf{(3)}$$

Thus, the probability of the link is a function of the proportion of degree k_i to the sum of all degrees at that time step.

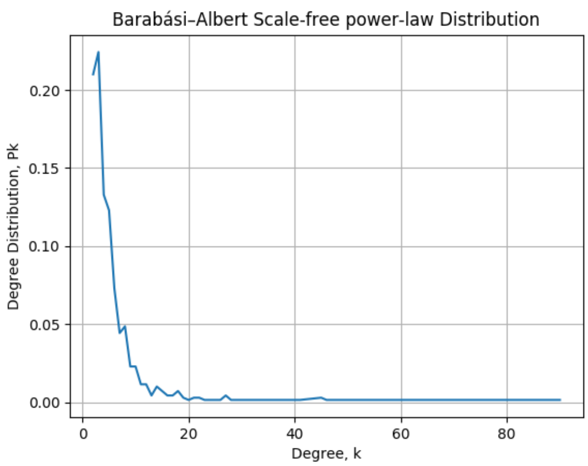

Figure 2: The degree distribution of a BA network with K=3. Note that there are many nodes with a low degree and a few that have a high degree.

These networks do not have an even distribution of links, certain nodes tend to have a much higher degree than others. Such nodes are known as “hubs” and have a role to play in information propagation and the robustness of real-world networks, as they can effectively communicate to and from a large number of nodes (Barabási, 2016).