Complex and Social Networks - Lab 1

Irene Simó Muñoz and Sara Montese

23 September 2023

1 Introduction

Graph theory is a fundamental mathematical framework used to model and understand various real-world phenomena, ranging from social networks to biological systems and transportation networks. Within the realm of graph theory, random graphs serve as essential models to study the properties and behavior of networks with a degree of randomness. This lab report explores the characteristics of random graphs using R, a powerful programming language for statistical computing and data visualization.
Random graphs hold a pivotal position in the field of network science. They allow us to investigate network properties under different degrees of randomness, offering insights into the behavior of complex systems where connections are not predetermined.
The primary objective of this lab report is twofold: first, to generate random graphs with controlled parameters using R (in particular the library *igraph*), and second, to explore and visualize key characteristics of these graphs. By doing so, we aim to gain a deeper understanding of how randomness impacts the structural properties of networks and how these properties align with theoretical expectations.

2 Clustering Coefficient and Average Shortest-Path in WS model

In this section, we present the foundation and the results of our analysis, which focuses on understanding how the parameter p - the probability of rewiring edges - in the Watts-Strogatz (WS) model influences two key network characteristics:

- *Clustering Coefficient* (or transitivity): often denoted as C, it measures the extent to which nodes in a network tend to cluster together. It quantifies the likelihood that neighbors of a given node are also connected to each other.
- *Average shortest-path length*: it quantifies the average number of edges that must be traversed to reach one node from another in the network. It provides insights into the network's overall navigability and efficiency of information transfer.

2.1 Watts–Strogatz Model

The Watts–Strogatz model is a random graph generation model that produces graphs with small-world properties, including short average path lengths and high clustering. An example can be seen in Fig. 1.
The network on the left represents a ring lattice with circular boundary conditions. Starting from this configuration connections are randomly rewired with a given rewiring probability p. For $p = 0$ (no rewiring), the network retains its regular lattice topology with a high clustering coefficient. This is because in a regular lattice, nodes are highly interconnected, and triangles (groups of three nodes with all possible edges) are abundant.
Conversely, as we increase the value of p, the network's regularity decreases, and it transitions towards a more random structure. This transition is evident as the clustering coefficient decreases.
At $p = 1$, the network becomes a random graph, and the clustering coefficient drops to a minimal value.

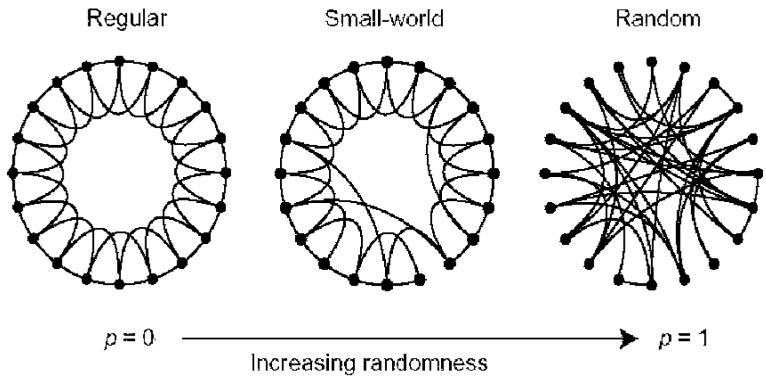

Figure 1: An example of increasing the randomness in the network [1]

2.2 Results

In this section, we illustrate our work with regards to the behaviour of the Clustering Coefficient and the Average Shortest-Path in WS model, as functions of p.
Using the function *sample_smallworld* of the package *igraph*, we generate a small-world network with 1000 nodes, where each node is initially connected to its 4 nearest neighbors.The p parameter, that determines the probability of rewiring edges to create a small-world structure, varies from 0 to 1.
To take into account the randomness of the model, to have a more robust result we compute the metrics 20 times for each value of p. Fig. 2 illustrates the Clustering Coefficient (C) and the Average Shortest-Path (L) as functions p of the WS model. Notice that the metrics are normalized by their values for the regular network (C(0), L(0)). Lower

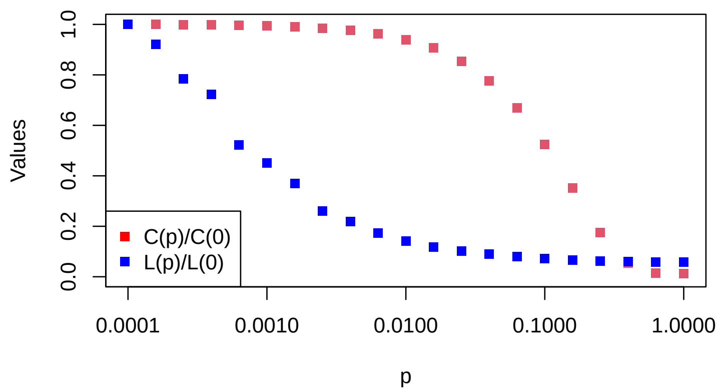

Figure 2: Clustering Coefficient (C) and the Average Shortest-Path (L) as functions p in WS model

values of p lead to networks with high clustering but longer average shortest-path lengths, resembling regular lattices. Conversely, higher values of p result in networks with lower clustering but shorter average shortest-path lengths, displaying small-world characteristics.
As p increases and the network becomes more random, the average shortest-path length decreases due to the introduction of more random edges that create shortcuts in the network. This is indicative of the emergence of long-distance connections in the network, which is a characteristic feature of small-world networks.
It can be noticed that there is a range of p where the path length has been reduced substantially, while the clustering coefficient is still close to the p=0 value. Within this range, networks exhibit small-world attributes. For example, when $p \approx 0.01$, the local clustering of the graph is still high but the path lengths in the graph becomes very short.

These findings align with the theoretical understanding of the WS model, where the parameter p controls the balance between regularity and randomness in network structure. This analysis enhances our comprehension of how small-world properties emerge in real-world networks, where both local clustering and global connectivity are essential.

3 Average Shortest-Path Length in ER model

In this section, we present the results of our analysis which focuses on understanding how network size influences the average shortest-path length in the Erdős-Rényi (ER) model. We conducted this analysis while ensuring network connectivity by dynamically adjusting the probability parameter p based on the graph size.

3.1 ER Model with Connectivity

The Erdős-Rényi (ER) model is a fundamental random graph model that generates networks with a specified number of nodes (n) and a probability parameter (p) governing the likelihood of edge creation between nodes. While the ER model provides a versatile framework for studying random networks, it does not guarantee network connectivity, especially for small values of p. To address this, we have used the result from [2] stating that if:

$$p > \frac{(1+\varepsilon)log(n)}{n}$$

then a graph will almost surely be connected.
To take into account the randomness of the model and have a more robust result, we iterate 20 times for each value of $p(n)$ and n and take the average.

Considering that for large values of the network size the computing time is high, we restrict our study to values that range in [0,16000].

3.2 Results

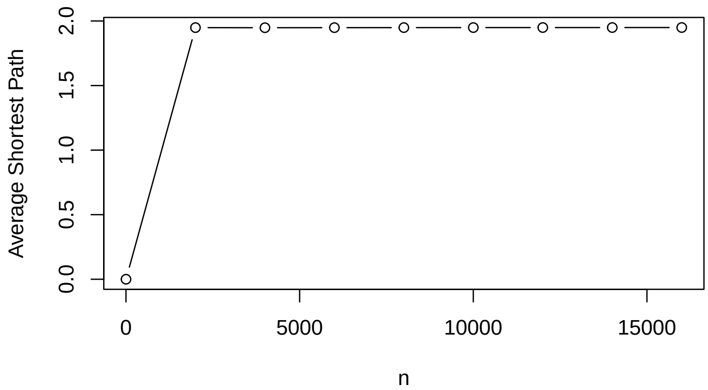

Figure 3: Average Shortest Path as a function of n in ER model

In Fig. 3, we observe distinct patterns in the average shortest-path length as a function of network size (n) and the probability of edge creation (p) in the ER model. As the network size (n) increases, we notice that the average shortest-path length converges to a stable value.
This observation suggests that beyond a certain network density, further increasing the probability of edge creation has diminishing returns in terms of reducing average shortest-path lengths. Consequently, in large networks, most nodes can be reached from every other node through a relatively small number of intermediate steps.

References

[1] S. D. . J. Xu, Z., “Small-world characteristics on transportation networks: a perspective from network autocorrelation,” 2007.
[2] P. Erdős and A. Rényi, “On the evolution of random graphs,” 1960.