Complex and Social Networks

Laboratory 1

The Watts–Strogatz and Erdős–Rényi Models

Davide Volpi & Daniel Weroński Falcó

Academic Year 2025/2026

1 Introduction

The task of this first laboratory session is to plot the clustering coefficient and the average shortest-path as a function of the parameter p of the Watts–Strogatz (WS) model, and plot the average shortest-path length as a function of the network size of the Erdős–Rényi (ER) model.

In the first plot, in order to include both values — average shortest path and clustering coefficient — in the same plot, the clustering coefficient and the average shortest-path values will be normalised to be within the range [0, 1]. This can be achieved by dividing the values by the value obtained at the left-most point, that is, when $p = 0$.

In case of the second plot, we will experiment with appropriate values of p, which will depend on the parameter n.

2 Experiments and Results

2.1 Preconditions

To ensure the statistical reliability of our examples and to mitigate the influence of random fluctuations regarding the graph process, we limit both experiments to 15 samples in their respective x-spaces, logarithmically spaced across the chosen parameter range for efficient sampling. The experiment is repeated $n_{runs} = 7$ times and $n_{runs} = 10$ for every sample in the WS and ER experiments respectively. This choice represents a balance between obtaining a representative average and maintaining a total execution time within one day, given the computational constraints associated with large values of n (specifically, $n_{WS} = 2000$ and $n_{ER} = 15000$ in our experiments).

Additionally, we explicitly check that each generated graph is connected before computing metrics, as disconnected graphs can lead to undefined or misleading values for metrics such as the average shortest path length. This approach guarantees the validity and consistency of our measurements.

In the ER experiment, we have used a result from [1] which states that if

$$p > \frac{(1+\epsilon)\ln(n)}{n}, \tag{1}$$

where p is the rewiring probability, ϵ is a small arbitrary value ($\epsilon = 0.001$ in our case), and n is the graph's size, then the graph $G(n, p)$ will almost surely be connected. This allows choosing an appropriate p-value that will almost guarantee the graphs are connected.

2.2 Watts-Strogatz Results

The plot shows how the clustering coefficient and the average path length change as the rewiring probability p increases in the WS model.

For a small p, the clustering coefficient remains very high, indicating that the network is highly clustered. However, the average path length is also high, meaning nodes are not easily reachable from each other.

As p increases, the average path length rapidly drops, meaning that the network quickly gains shortcuts that connect distant parts of the network, reducing the distance between nodes. The clustering coefficient also decreases, albeit in a far slower fashion.

For large values of p, both clustering coefficient and the average path length are low, representing a behaviour of a random graph.

Theoretically, the "sweet spot" for achieving high clustering and small diameter is approximately $p \approx 0.01$. In our experimental results, with a lower amount of experimentation, we observe that this occurs between $0.001 < p < 0.01$, which is consistent with the theoretical expectations.

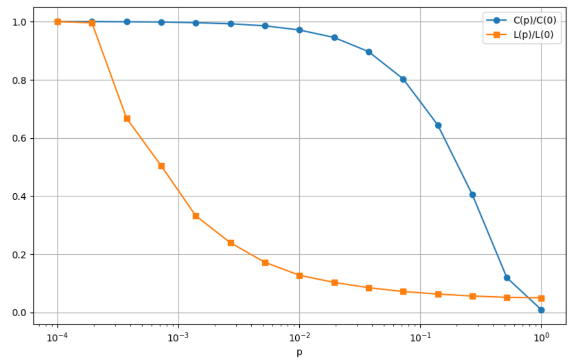

Figure 1: Clustering Coefficient and Average Path Length in WS Model

2.3 The Erdős-Rényi Model

As previously mentioned, 15 log-spaced samples of the number of nodes were taken. For each of these x-values, 10 connected graphs were taken in order to calculate the average, and be able to compute an error estimation. This was done using the standard error on the mean procedure as:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \tag{2}$$

where $\sigma_{\bar{x}}$ is the standard error on the mean, σ_x is the standard deviation of a sample and n is the sample size — in this case 10.

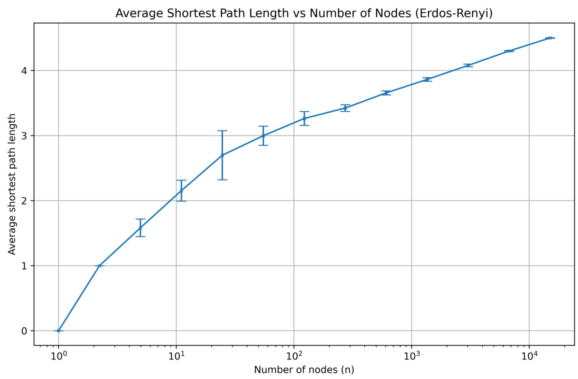

Figure 2: Average Shortest Path Length vs Number of Nodes in ER Model with errorbars showing the standard error on the mean $\sigma_{\bar{x}}$ as defined in eq. 2.

Figure 2 above shows a non-negligible decreasing gradient graph, characteristic of the logarithmic growth expected. However, the inherent maximum graph size limitations due to compute constraints mean that the logarithmic pattern is not as readily distinguishible as expected.

Furthermore, the error analysis shows the interesting detail that after a network size of $\sim$ 1000 the graphs have increased consistency and predictability when compared to the earlier ones (barring the first 3 data-points, where the graphs are too small to have substantial variance).

For these large graphs, different random instances produce very similar average shortest path lengths, leading to a smaller error and greater confidence in the average measurement.

3 Conclusion

The first experiment demonstrates the "small-world" effect: there is a range of p values where the network has both a high clustering coefficient and a low average path length. This matches the real-word networks, which are highly clustered with short paths between nodes.

The second experiment provided evidence that shows the average shortest path length grows logarithmically with the number of nodes in the network. This means that even as a random network grows to an very large size, the average distance between nodes increases very slowly.

This result reinforces the concept of "short paths" in random graphs and provides a baseline for understanding how efficiently information can traverse a network that lacks an explicitly organised structure.

References

[1] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.