Overview

The task, that was given was to create three networks by using different models – $G_{n,p}$; Small World and Preferential Attachment. Afterwards, we had to interpret and compare the results, as shown in the lecture. The differences between the graphs are obvious and there are possible explanations, that I've provided for each of them.

$G_{n,p}$ Model

This is the simplest method to construct a stochastic graph. It requires to have the number of vertices (n) and the probability to connect two vertices (p). The input parameters given in this case were 500 vertices and p=0.015. I have started the research by counting the number of the edges and comparing with the mean number of edges m. The numbers, as expected are pretty close to each other and this is mainly because the whole method is completely random and the attachment of new vertices with edges is with the same probability and it doesn't change. Afterwards I checked the minimum and maximum degree **– 1** and **16.** By looking at these values and more specifically the minimum degree, we can easily determine the vertex and edge connectivity, even without using the provided function. The edge connectivity refers to how many edges have to be cut in order to disconnect the graph, and in this case if we remove only **1** vertex, then the graph will become disconnected, because there are vertices with only 1 edge. Also the vertex connectivity is **one**. The diameter of the graph is **6,** which means, that the longest shortest path between two different vertices goes through 6 edges. The observed and the approximated average path length (apl) are again close, because of the randomness property of using this model. The transitivity of the graph is very low, compared to the other two observed models. This is because in the $G_{n,p}$ Model, the transitivity depends on the probability (p). Finally, I've plotted the degree distribution and the cumulative degree distribution. We can observe that as expected there is a binomial distribution of the degrees and they are not concentrated on one point, but are rather equally distributed. This is again due to the random allocation of the edges.

n=500 p=0.015

$\overline{m} = \frac{n(n-1)}{2}p$ = 1871.25 |E| = 1888

$\overline{\delta}$ = (n-1)*p = 7.485

Apl(expected) = ln(n)n($\overline{\delta}$) = 3.0874 Apl(actual) = 3.298397

$\tau(G_{n,p})$ = 0.01272804

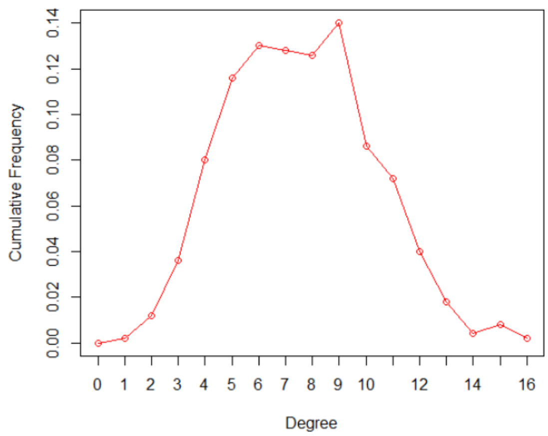


Fig 1.1 – $G_{n,p}$ Graph – Degree Distribution

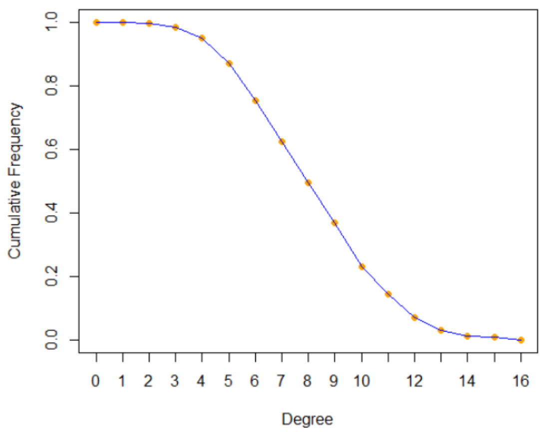


Fig 1.2 - $G_{n,p}$ Graph – Cumulative Degree Distribution

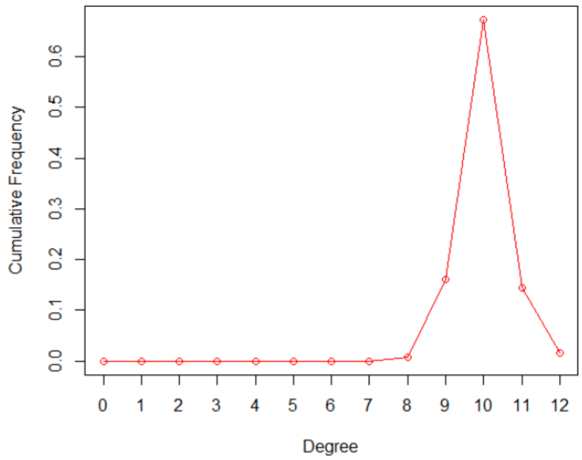
Small World Graph

As mentioned above, the $G_{n,p}$ graphs don't provide good clustering and the small world graphs address this issue. They are built by first creating a cycle graph and then depending on which method is chosen removing an edge and connecting it to a different random vertex(Watts-Strogatz), or by just adding the new edge without removing one(Newman-Watts). This is done to improve the clustering. I have calculated the same coefficients and variables as above, but we can see that the values differ. The actual and the expected number of edges is equal. There is a big difference in the expected and the observed average path length and this is because of the addition of new edges. This can be fixed by increasing n. The diameter here is **9,** which is greater than the previous graph model, because here the vertices are connected to their neighbors more, than with random vertices on the other side of the network. But the most notable improvement here is the transitivity, which is in fact allowing us to think of the small world property. Regarding the degree distribution, as shown in fig 2.1 and 2.2, we can observe that there are no hubs and that most of the vectors have degree of 10, which is 2x5, the number of edges in a cluster. We can also note, that the method used was Watts-Strogatz, because there are vectors with a degree of less than 10, which means some edges were removed. Minimum degree is 8, Maximum = 12.

$n = 500$ $p = 0.025$ $x = 5$

expected Edges = $\frac{1}{2}(n_x k_x p) = 2500$ $|E| = 2500$

apl(expected) = $\ln(n_x k_x p)/k^2 p$ = 1,931   apl(observed) = 4.646333



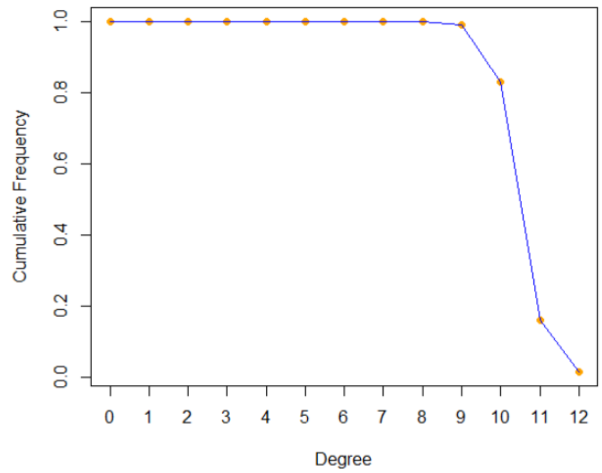
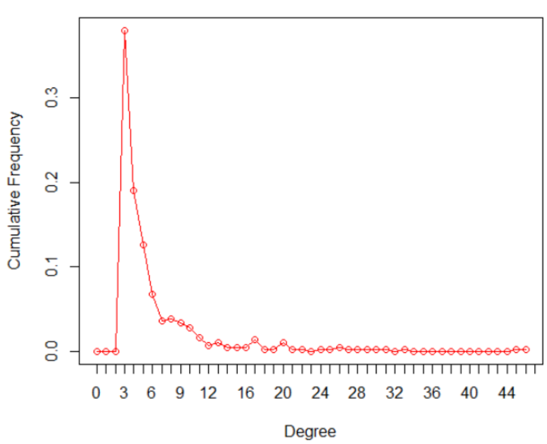


Fig. 2.1 – Degree distribution – Small World **Fig 2.2 – Cumul. Degree Distr. Small World**

Preferential Attachment Graph

The last graph model that we had to examine was the preferential attachment one. It assumes that in the real world, there is a different pattern of the networks. For example – the distribution of wealth (Pareto 80/20 principle) which means that the rich have big chances to get richer and the poor to stay poor. This is called heavy-tailed distribution and we can see it in Fig. 3.1 and 3.2. The way that this model works is that the probability (p) when adding a new vertex (m) to the network depends on the degree of the other vertices in the graph. The bigger the degree-the bigger the probability. We can see this again in the figures provided. There is a lot of low-degree vertices (concentrated at the minimum **(3))** and few with higher degree, with the highest one being **46**. The diameter (**6**) and the apl(**3.346076**) are relatively low, because of the hubs that are present. The transitivity (**0.02614424**) is rather low compared to the Small World graph and we cannot observe the small world property properly here. This is again, because of the hubs and the vertices with low number of edges. Vertex and edge connectivity again correspond to the value of the vertex with the lowest degree (**3**).



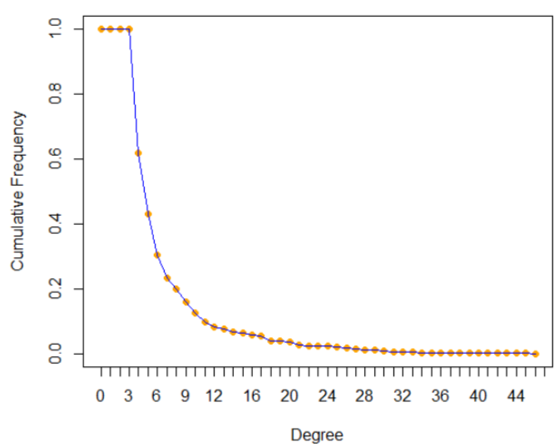


Fig 3.1. Degree Distribution – PAG **Fig 3.2 Cumulative Deg. Distribution - PAG**