M2 – NSD (Practical Work 2 - Sessions 4,5)

Graph models

Lionel Tabourier, Fabien Tarissan

During these PW sessions, we aim at generating and describing the properties of graph models seen during the courses. We then compare them to real-world graphs properties. It is therefore necessary to know how to make the measurements corresponding to the first 3 PW sessions.

In the following (and in future practical works), we represent networks by graphs. We denote a graph $G = (V, E)$, where V is the set of vertices (also called nodes) of the graph and numbered from 0 to $|V| - 1$ and E is the set of links (also called edges).

1 Erdös-Rényi random graphs

Exercise 1 — ***Generation.*** Create a program which, given two integers n and m generates an Erdös-Rényi random graph with n nodes and m edges and writes the list of edges in a file. Multiple edges are allowed, but not loops.

Exercise 2 — ***Characteristics.*** Using the previously defined programs,

1. Generate an Erdös-Rényi graph with $n = 7236$ nodes and $m = 22270$ edges.
2. Compute the most important features of this graph : number of connected components, size of the largest connected component, number of isolated nodes, average clustering coefficient, average distance between nodes in the largest connected component.
3. Display the degree distribution of this graph using (for example) GNUPLOT.
4. Comment these results.

Exercise 3 — ***Comparison to* ppi-droso *drosophila network.*** We use the protein-protein interaction network of the drosophila seen during the first PW sessions (named PPI-DROSO in the following). Delete the loops (if it's not already done).

Compare the measurements made on PPI-DROSO network during the first PW sessions with those made on the Erdös-Rényi graph. Comment.

2 Random graphs with fixed degree distribution

Exercise 4 — ***Generation using direct method*** We first generate a random graphs with a fixed degree distribution using the direct method, that is the configuration model.

1. Create a program which, given a degree list, creates a random graph with a similar degree list using the configuration model seen during the course.
2. Extract the degree list of PPI-DROSO to generate a random graph with a similar degree distribution.

Exercise 5 — ***Generation using switching method*** Now, we implement the switching method.

1. Create a program which, given a real graph, achieves P random switches of links ends. Take good care that switches do not create any loop or multiple edge.

2. Starting from the PPI-DROSO graph of the drosophila, realize $P = 10^6$ permutations, and writes in a file the list of edges obtained.
3. Add to the code a module that allows to measure the clustering of the graph every 10^4 switches, and to write in a file the value of the clustering throughout the process.
4. Plot the corresponding values obtained with PPI-DROSO graph. Comment the results.

Exercise 6 — *Comparison to* ppi-droso *drosophila network.* We know have two random graphs with the same degree distribution as PPI-DROSO.

1. Compare these two graphs using the important features previously cited : are they identical ? If not, what is different from one to the other ?
2. Compare the measurements made on PPI-DROSO network during the first PW sessions with those made on the graph generated using the configuration model. Comment.

3 Barabási-Albert random graphs

Exercise 7 — *Barabási-Albert graph generation.* Create a program generating a graph according to Barabási-Albert method such that :

1. the graph is built by adding n nodes to the initial graph, the initial graph will be given as an input of the program,
2. any node is connected to m' nodes of the existing graph according to the preferential attachment rule, m' is an input parameter of the program too.

Exercise 8 — *Comparison to* ppi-droso *drosophila network.* Use the previous algorithm to generate a scale-free Barabási-Albert network with $n = 7236$ nodes, to compare with PPI-DROSO graph, according to the following indications :

1. use as an initial graph the 4 nodes graph containing the following edges :
   0 1
   1 2
   2 3
2. set the value of m in such a way that the density is of the same order as PPI-DROSO density.

Compare the structure of the graph obtained to PPI-DROSO structure. Comment.

4 Watts-Strogatz random graphs

Exercise 9 — *Comparison to* ppi-droso *drosophila network.* Open Question.

Suppose that you want to compare PPI-DROSO structure to a Watts-Strogatz network with comparable size and density.

1. Describe how you would proceed. (Meaning how would you choose the characteristics to generate the graph ?)
2. Implement the solution that you proposed and then discuss the results.