UPPSALA UNIVERSITY
DEPARTMENT OF MATHEMATICS
Fiona Skerman

Resit version
Summer 2020

Lab 4: Network Properties.

The deadline for this sheet is midnight Monday 31st of August.

Please submit hand-ins on Studentportalen. All code should be included. Please feel free to submit videos illustrating your results where appropriate, via Studentportalen or uploaded elsewhere. You may work in groups of size 1-4, and only one group member needs to submit the assignment. State clearly the members of the group.

4a. Robustness of networks to edge deletion

For this exercise all sections are to be done both for an example real world network and a random networks model of your choice. For this question it is enough to simulate the network and/or random deletion process multiple times (≥ 10) and take the average value.

What is the size of the maximum connected component after α proportion of edges have been deleted. By repeatedly simulating the model while systematically changing α plot the average size of the maximum connected component in the graph as a function of α. The networks should be large enough that it makes sense to delete α proportion of edges for several values of α.

1. If edges deleted are chosen randomly
2. Each step pick two edges xy and uv randomly. Let c_x (respectively c_y) be the size (number of vertices) of the connected component containing x (respectively y) after edge xy is deleted. Similarly c_u and c_v are sizes of the connected components containing u and v after edge uv deleted. Delete edge xy if $c_x \geq c_y$; otherwise delete uv.
3. Another edge deletion method of your choosing. Describe your chosen algorithm clearly.

(4 points)

4b. (Global) clustering coefficient of networks

In this exercise you will consider the global clustering coefficient: called clustering coefficient in lecture slides and notes, but be careful not to confuse it with the local clustering coefficient. The global clustering coefficient of  is 1, of  is 3/4, of  is 6/10, of  is 0, and of  is 3/11. (The python package networkx, calls this transitivity.).

1. For a graph on 5 or 6 vertices show a worked example of how to calculate the global clustering coefficient - where the value is in the interval $[0.1, 0.9]$ **(2 points)**.

2. Calculate the global clustering coefficient for an example real world network **(1 points)**.

4c. Random graphs as null models for networks

For this exercise you may consider either the (global) clustering coefficient, the maximum modularity or any other non-trivial measure of the network as the statistic of interest.

Find a real network examples and calculate the statistic (clustering coefficient, max modularity etc) of those networks. In the case of maximum modularity you can use a heuristic method such as Louvain algorithm or Leiden algorithm which outputs a modularity and a score, but is not necessarily the maximum modularity (it would be too slow to test all possible partitions).

1. **Configuration Model as a null model.** Simulate a graph with the configuration model which has the same degree sequence as your example network. (You may choose to generate a multigraph with same degree sequence, or sample repeatedly keeping only simple graphs). By repeated simulations, (≥ 30), record the values of your statistic on the random graph. Illustrate this using a box and whisker plot and show on the same plot the value of your statistic on the real network.

   Was your test statistic on the real-network greater then 95% of the values for the test statistic which you recorded for the random graphs? **(3 points)**.

[1] If you submitted Lab 4 during the course, ensure that you use a different example real world network in 4a-c, a different random network model in 4a and a different test statistic in 4c.