

Assignment 3

In this assignment you will analyze a real dataset from an online social network. In particular you will use the social graph of *Brightkite* [1] and you are required to provide a report with the following:

- 1) A plot of the probability degree distribution (30 pts)
- 2) A plot of the average local clustering coefficient as a function of the node degree (30 pts)
- 3) The following statistics (40pts):
 - a. Average node degree
 - b. Average local clustering coefficient
 - c. Theoretical clustering coefficient (that is considering a random graph with the same degree distribution)
 - d. Entropy of the node degree distribution (see Assignment 2)
 - e. Assortativity coefficient for the graph (consider the degree of a node as the characteristic under consideration)
- 4) **Bonus question (20 pts):** Make a *nice* visualization of the graph, by assigning different colors to nodes that belong to different *communities*. At a high level, a community is a set of vertices of the graph such that the members of this set are *tightly* connected, while they have loose connections to vertices outside the community. You can find more information about communities and how you can detect them in your textbook (chapters 11.2-11.11). For this task you can perform community detection using any known algorithm or even devise your own heuristic.

For the first 3 questions you have to write your own code - in any programming language you prefer - which you also need to submit with your report. For the bonus part, you can use any graph analysis package (e.g., gephi, iGraph for R and Python, Pajek etc.). However, you will need to provide details on what algorithm you used, what package you used etc. Furthermore, if you pick to use an existing graph analysis package for the first 3 questions, you will automatically lose 30 points from your total. For instance, if you answer all questions correctly, your sum will be 120, but you will actually get 90.

[1] Available online: <http://snap.stanford.edu/data/loc-brightkite.html>