Social Networks

RWTH Aachen, Summer Semester 2019

Home Assignment 1
Handed out April 11, 2019
Due April 25, 2019, 23:59

For the following tasks use the network provided in the file jazz.txt (edgelist format)

In this home assignment we will use data from the network of collaborations of jazz bands described in *Community Structure in Jazz*, P.Gleiser and L. Danon, Adv. Complex Syst.6, 565 (2003). In case you are interested in the dataset you are welcome to read the paper: https://arxiv.org/abs/cond-mat/0307434

Create a Jupyter notebook that loads the network file and extracts the following information using the library NetworkX. You may use built-in NetworkX functions unless otherwise noted. **Python 3.x is required. Libraries can be used.**
The notebook must follow the following structure (you can use the template HA1-template.ipynb provided)

1   Basic statistics
      1.1 Number of nodes
      1.2 Number of links
      1.3 What is the density of the network? Is it sparse? Justify your answer using text in a markdown cell.
      1.4 Node id and degree of the node with the highest degree

2   Paths
      2.1 Which is the shortest path from nodes 1 to 10: Give the length and the node IDs it traverses
      2.2 Diameter of the network
      2.3 The average distance between nodes
      2.4 Print the nodes with the 10 highest betweenness centrality scores

3   Plotting
      3.1 Plot the network using a spring layout. For your own enjoyment: try to plot the network with different layouts.
      3.2 Plot in green colour the nodes and links of the shortest path in 2.1
      3.3 Plot the degree distribution
      3.4 Plot the network with node size proportional to closeness centrality.

4   Additional questions.
      4.1 How many connected components are there in the network

4.2 Which the average clustering coefficient of the network

4.3 What is the number of triangles in the network

4.4 Challenge: For each edge calculate the number of neighbors its constituent nodes have in common, sometimes called the embeddedness of an edge. For example: the friendship edge of Tobias and Johannes has embeddedness of 2 if Tobias and Johannes have two friends in common. Remove all edges with embeddedness less than or equal to 2 and plot the resulting network. For your own enjoyment: think about why the network changed the way it did.

Grading

Task 1. Basic statistics	25%
Task 2. Paths	25%
Task 3. Plotting	25%
Task 4. Additional questions	25%

Useful documentation/Notes

NetworkX library for Python http://networkx.github.io/
Anaconda Python distribution https://anaconda.org/anaconda/python

Provided files
Network jazz in edgelist format provided in the file jazz.txt
Submission notebook template HA1-template.ipynb

Submission
This assignment is due April 25th, 23:59. Submit your .ipynb file to the course Moodle.

The due date is a soft deadline. That is, you can still hand in your submission until the hard deadline at 11:59 (am!) on the following day, but your points will be scaled to 66%, i.e., you will lose one third of your points. Submissions after 12:00 noon the following day will be ignored.