# Complex Networks

# Project 2: Centrality, degree-degree correlation and community structure in networks

**Send the solutions to: projetosicmc@gmail.com**

1 – For the following networks:

a) E-road network (http://konect.cc/networks/subelj_euroroad),

b) Hamsterster friendships (http://konect.uni-koblenz.de/networks/petster-friendships-hamster).

c) C. elegans neural network (http://www-personal.umich.edu/~mejn/netdata/celegansneural.zip)

d) US airport network (http://toreopsahl.com/datasets/#usairports)

Construct a correlation matrix between the centrality measures: (i) degree, (ii) k-core, (iii) closeness centrality, (iv) betweenness centrality, (v) eigenvector centrality, (vi) pagerank, (vi) random walk accessibility, (v) communicability centrality.

Discuss the highest correlations and interpret the results.

2 – Choose four dataset of cities (from OSMX or any other dataset) and compare the cities in terms of the centrality measures. That is, construct the histogram of (i) degree, (ii) closeness centrality and (iii) betweenness centrality. Discuss which city is easier to navigate in terms of these distributions.

3 – For the following networks:

a) Human protein network (http://konect.cc/networks/maayan-vidal)

b) C. elegans protein network 2007 (http://interactome.dfci.harvard.edu/C_elegans/index.php?page=download)

c) C. elegans protein network 2004 (http://interactome.dfci.harvard.edu/C_elegans/index.php?page=download)

d) Western US power grid (http://www-personal.umich.edu/~mejn/netdata/power.zip)

e) R dependency network (http://www.mas.ncl.ac.uk/~ncsg3/blog/dependencies.csv)

Construct the histogram of the centrality measures and compare them. Use the measures (i) degree, (ii) eigenvector centrality, (iii) pagerank and (v) closeness

centrality. Are there similarities between networks of the same type, such as among technological or biological networks?

4 – Consider the networks in the last exercise. For each network, consider the measures:
(i) degree, (ii) k-core, (iii) closeness centrality, (iv) betweenness centrality, (v) eigenvector centrality, (vi) pagerank, (vii) random walk accessibility, (viii) communicability centrality.
And for each measure, calculate:
(i) average, (ii) standard deviation, (iii) second moment, (iv) Shannon entropy. Thus, each network will be represented by a feature vector with 8*4 = 32 elements. Project the networks into a two dimensional space by using Principal Component Analysis (PCA). Discuss the similarities of the networks in terms of their features.

5 – For the following networks:

a) E-road network (http://konect.cc/networks/subelj_euroroad),

b) C. elegans neural network

(http://www-personal.umich.edu/~mejn/netdata/celegansneural.zip)

c) US airport network (http://toreopsahl.com/datasets/#usairports)

d) Human protein network (http://konect.cc/networks/maayan-vidal)

Obtain the scatterplot of knn(k) X k and calculate the Pearson correlation coefficient between these two measures. Compare it with the assortativity coefficient. Comment the results.

6 – Construct the Givan-Newman benchmark (N=128, kin+kout = 16 and 4 communities). You can use:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.community_generators.LFR_benchmark_graph.html
Compare the following methods:
Fastgreedy:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.modularity_max.greedy_modularity_communities.html#networkx.algorithms.community.modularity_max.greedy_modularity_communities
Label propagation:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.label_propagation.label_propagation_communities.html#networkx.algorithms.community.label_propagation.label_propagation_communities
Givan Newman:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.centrality.girvan_newman.html#networkx.algorithms.community.centrality.girvan_newman

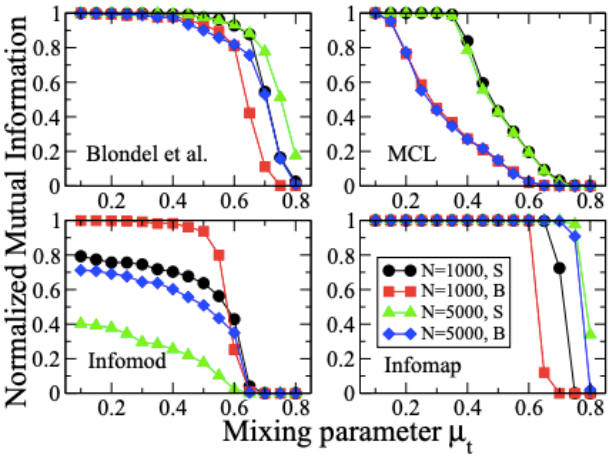
Louvain: https://github.com/taynaud/python-louvain
Use the normalized mutual information as a measure of accuracy.

7 –Consider the methods for community identification:
Netcarto: https://amaral.northwestern.edu/resources/software/netcarto
Informap: https://www.mapequation.org/code.html
Fastgreedy:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.modularity_max.greedy_modularity_communities.html#networkx.algorithms.community.modularity_max.greedy_modularity_communities
Label propagation:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.label_propagation.label_propagation_communities.html#networkx.algorithms.community.label_propagation.label_propagation_communities
Givan Newman:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.centrality.girvan_newman.html#networkx.algorithms.community.centrality.girvan_newman
Louvain: https://github.com/taynaud/python-louvain

Show the partition of the Zachary karate club for these networks.

8 – Consider the Fortunato benchmark:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.community_generators.LFR_benchmark_graph.html
Compare the following methods. Consider the Normalized Mutual Information (NMI) measure.
Netcarto: https://amaral.northwestern.edu/resources/software/netcarto
Informap: https://www.mapequation.org/code.html
Fastgreedy:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.modularity_max.greedy_modularity_communities.html#networkx.algorithms.community.modularity_max.greedy_modularity_communities
Label propagation:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.label_propagation.label_propagation_communities.html#networkx.algorithms.community.label_propagation.label_propagation_communities
Givan Newman:
https://networkx.github.io/documentation/stable/reference/algorithms/generated/networkx.algorithms.community.centrality.girvan_newman.html#networkx.algorithms.community.centrality.girvan_newman
Louvain: https://github.com/taynaud/python-louvain

Show the comparison as bellow for networks with N=1000 nodes.




9 – Consider the following networks:

a) E-road network (http://konect.cc/networks/subelj_euroroad),

b) C. elegans neural network

(http://www-personal.umich.edu/~mejn/netdata/celegansneural.zip)

c) US airport network (http://toreopsahl.com/datasets/#usairports)

d) Human protein network (http://konect.cc/networks/maayan-vidal)

Construct a table showing the following properties of these networks: N, Average degree, assortativity coefficient, average shortest path length, modularity. In the case of modularity, show the values for the following methods: (i) Fastgreedy, (ii) Label propagation, (iii) Givan Newman, (iv) Louvain. a) Compare the results in terms of the modularity. (b) Are there some relation between assortativity and other network measures?

10 – Elaborate an exercise and solve it.