Mini Project Assignment Questions

Network Science for the Web - DS 608

Aug-Dec 2019

Submission Details:

1) 15 Min presentation about the findings of all questions.
2) A report explaining the results. (Can also be a fairly well documented jupyter notebook code showing all plots and explanations).
3) Zip file of the code including:
    a) Code files
    b) Readme file with execution details of the code
    c) Text document containing detailed tasks performed by each team member

Presentation Details:

1. All team members should speak about different datasets
2. Keep it brief and stick to 15 presentation
3. Focus on reasoning of selection of dataset and interpretation from various measures
4. Presentation to be done as a part of mid-term examination

Questions:

1. Find the following Centrality Measures on relevant datasets and explain your plots and findings:
    a. Degree Centrality
    b. Find graph centres and radius (Eccentricity)
    c. Find graph medians (Closeness)
    d. Find nodes with max radiality
    e. Stress Centrality
    f. Betweenness
    g. Node with max flow vitality (Use a weighted dataset)
    h. Node with max closeness vitality
    i. Node with max stress vitality

2. Find the following Graph Properties on relevant datasets and explain your plots and findings:
    a. Weiner Index
    b. Max clique/s

c. Plex
d. Core

3. Find the optimal number of partitions of a graph using modularity and betweenness. (Code given as a part of iGraph tutorial)

4. Find the following Web Centralities on web datasets and explain your plots and findings:
   a. Page Rank
   b. HITS
   c. Compare Page Rank and HITS

5. Generate a graph with power law degree distribution. Plot and show that it is scale free at different levels. (Use Configuration Model from Barabasi's book -http://networksciencebook.com/chapter/7#introduction7)

6. Choose a centrality measure (like betweenness). Find the node with max centrality measure. Detect a max clique around the node with max centrality. Find the max cliques of the graph using inbuilt function (like answer of 2b). Compare and specify if the max clique around the node of max centrality is actually a max clique.

7. Estimate the following centralities and compare it with the exact centrality values (one can use random walk function for estimating centralities)
   a. Degree Centrality
   b. Betweenness
   c. Closeness

8. Find the following Feedback based Centrality on relevant datasets and explain your plots and findings:
   a. Bonacich Centrality
   b. Hubbell Centrality

9. Calculate and analyze DON based measure for dominance centrality (To be covered in the class to be held on 13 sept 2019) (Use datasets selected for question 1 and 8, and compare the results with other centrality measures)

Shortlisted Standard Datasets

Dataset Link and Number of nodes

Undirected UnWeighted : U -

http://konect.uni-koblenz.de/networks/contiguous-usa 49
http://konect.uni-koblenz.de/networks/dolphins 62
http://konect.uni-koblenz.de/networks/adjnoun_adjacency 112
http://konect.uni-koblenz.de/networks/arenas-jazz 198
http://konect.uni-koblenz.de/networks/arenas-email 1133
http://konect.uni-koblenz.de/networks/subelj_euroroad 1174
http://konect.uni-koblenz.de/networks/tntp-ChicagoRegional 1467
http://konect.uni-koblenz.de/networks/petster-friendships-hamster 1858
http://konect.uni-koblenz.de/networks/petster-hamster 2426
http://konect.uni-koblenz.de/networks/ego-facebook 2888
http://konect.uni-koblenz.de/networks/opsahl-powergrid 4941
http://konect.uni-koblenz.de/networks/arenas-pgp 10680
http://konect.uni-koblenz.de/networks/ca-AstroPh 18771
http://konect.uni-koblenz.de/networks/as-caida20071105 26475
http://konect.uni-koblenz.de/networks/filmtipset_friend 39199

Undirected Weighted : U +

http://konect.uni-koblenz.de/networks/moreno_train 64
http://konect.uni-koblenz.de/networks/moreno_lesmis 77

Directed UnWeighted : D -

http://konect.uni-koblenz.de/networks/maayan-foodweb 183 (self-loops)
http://konect.uni-koblenz.de/networks/moreno_innovation 241
http://konect.uni-koblenz.de/networks/moreno_blogs 1224 (self-loops)
http://konect.uni-koblenz.de/networks/maayan-faa 1226 (self-loops)
http://konect.uni-koblenz.de/networks/maayan-Stelzl 1706 (self-loops)
http://konect.uni-koblenz.de/networks/maayan-figeys 2239
http://konect.uni-koblenz.de/networks/opsahl-openflights 2939

http://konect.uni-koblenz.de/networks/dblp-cite 12591 (self-loop)
http://konect.uni-koblenz.de/networks/cfinder-google 15763 (self-loops)
http://konect.uni-koblenz.de/networks/subelj_cora 23166
http://konect.uni-koblenz.de/networks/ego-twitter 23370
http://konect.uni-koblenz.de/networks/ego-gplus 23628
http://konect.uni-koblenz.de/networks/cit-HepTh 27770 (self-loops)
http://konect.uni-koblenz.de/networks/linux 30837 (elf-loops)
http://konect.uni-koblenz.de/networks/cit-HepPh 34546 (self-loops)
http://konect.uni-koblenz.de/networks/p2p-Gnutella31 62586
http://konect.uni-koblenz.de/networks/soc-Epinions1 75879

Directed Weighted : D +

http://konect.uni-koblenz.de/networks/moreno_mac 62
http://konect.uni-koblenz.de/networks/moreno_highschool 70
http://konect.uni-koblenz.de/networks/foodweb-baydry 128
http://konect.uni-koblenz.de/networks/foodweb-baywet 128
http://konect.uni-koblenz.de/networks/moreno_oz 217
http://konect.uni-koblenz.de/networks/opsahl-usairport 1574
http://konect.uni-koblenz.de/networks/moreno_health 2539
http://konect.uni-koblenz.de/networks/advogato 6541 (self-loops)

More Datasets:
https://toreopsahl.com/datasets/
http://networksciencebook.com/translations/en/resources/data.html

Or you can choose any other relevant dataset having at least 50 nodes.