

Frontiers of Network Science 6250/4250

Assignment 1, Networks

Select three different real world networks for analysis. One network can be from your own project or research. Two other networks (or all three if you don't have your own network) should be taken from one of the following collections:

- SNAP repository (<http://snap.stanford.edu/>)
- Koblenz set (KS) (<http://konect.uni-koblenz.de/>)
- Network Repository (NR) (<http://www.networkrepository.com/>)
- Pajek datasets (<http://vlado.fmf.uni-lj.si/pub/networks/data/>)
- Mark Newman's Collection (<http://www-personal.umich.edu/~mejn/netdata/>)
- DIMACS Challenge Graphs (<http://www.dis.uniroma1.it/challenge9/download.shtml>)
- UF Sparse Matrix Collection (<http://www.cise.ufl.edu/research/sparse/matrices/index.html>)
- Laboratory for Web Algorithmics (<http://law.di.unimi.it/datasets.php>)

The network has to have at least 500 nodes and the number of edges has to be at least 2,000. We encourage you to select the networks that your computer is capable of processing, given your choice of tools, so no larger than 2,000 nodes and 10,000 edges.

Send us your list of three networks for approval, specifying where you obtained the networks (please, provide an URL) and its size to ensure that all students work on different networks. If it turns out that the network you wish to analyze has already been taken by someone else, we will ask you to select another network. So, the earlier you send us your selection, the better are your chances of getting the networks you want. Do not start working on the networks until you receive an OK from us.

You can use whatever tools (either your own or third party) you deem appropriate for the job. If you are using a third-party tool (e.g., Gephi, Neo4j, etc.), it has to be available free of charge (or at least a free of charge fully functional evaluation version should be available). Please document which tools you use for which task, including the URL of the tool Web site. If you are using your own tools (e.g., you are writing your own programs), please specify which ones and for which tasks and provide the source code and executable for ThinkPad running Microsoft Windows 10 along with the relevant instructions on how to run them. If you are using third-party frameworks or applications (e.g., Matlab, Microsoft, Excel, etc.) indicate which tool was used for which task and provide all user files (".m", ".xls", etc.) In your report, provide a brief justification for your choice of a particular tool for a particular task.