

DSO 510: Homework 3

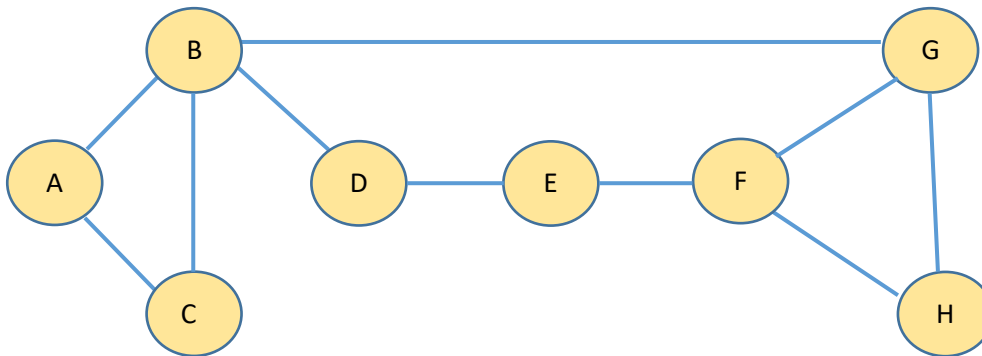
Due: Wednesday, December 1 @ midnight

General Instructions

You may work with a classmate or individually. If you choose to work with a partner, **only one member should submit the assignment**. Make sure that both members' names are visible on the word document submission. There is a submission portal on Blackboard under the assignments/homework tab.

You must submit a document with your screenshots and explanations.

Exercise 1. [2 points] Consider the following undirected network of 8 nodes. All edges have weight 1.



Tasks: compute the following metrics. For full credit, you need to show your detailed work (not just a number – discuss the paths).

- (a) Compute the **closeness centrality** of node F. Show your work, don't just provide a number.
- (b) Compute the **betweenness centrality** of node D.

Exercise 2. [4 points]

You will need to use **Gephi** for this part of the assignment.

Tasks: Use the employees.gephi file to answer the following questions. This file contains a network of employees and their connections. Each node is an employee and each edge is a connection. It is an undirected graph with equal weights.

- (a) The employees in the network all work in the same company. Employees in the same department are much more likely to have connections with each other than employees in other departments. Using the modularity tool (under Statistics), identify how many departments exist in the company. Identify an employee (other than employee 46) that is in the same department as employee 46. Explain your answer and show screenshots to validate.
- (b) The company is concerned that the departments are disjointed and don't have enough cross-departmental connections. To outline the importance of these cross-departmental connections, the company has set an objective to reward an employee that does a good job of connecting departments (specifically those that are bridges between departments). Management has identified employees number 25, 35, and 60 as potential candidates. How many connections does each of the mentioned employees (25, 35, and 60) have? Which would you recommend given the stated objective? Explain your answer and show screenshots to validate.

Exercise 3. [4 points] – Facebook Network Analysis

In this problem you will visualize and analyze an anonymized Facebook friends' network. Work with the undirected form of the network. The file has been renamed "FB.edges" (source <http://snap.stanford.edu/data/index.html> - ego-Facebook database - file 3437.edges).

- (a) Download the graph file "FB.edges" and open it in Gephi. Choose Undirected in the beginning when you open the graph. Paste a screenshot of the same.
- (b) Go to "Window" tab (top left) and select "Context" - how many Facebook friends does the network include?
- (c) Apply "Force Atlas 2" layout first. Then, once the network unfolds enough, stop it and apply "Noverlap" layout to make sure nodes do not overlap. Add the screenshot of the result.
- (d) Compute the degree, betweenness and closeness centrality statistics for the entire network. What is the average degree?

- (e) Rescale the nodes based on their betweenness centrality and provide a screenshot of the new graph (Assign min size of a node as 10 and the max size as 60)
- (f) Rerun the layout “Noverlap” so that there is no node overlap due to rescaling and provide a screenshot of the graph.
- (g) Go back to “Statistics” tab, run “Modularity” and show.
- (h) Recolor nodes based on their Modularity class. What does it reveal about the network structure?
- (i) Suppose we are interested only in the most socially active members of the network. Using filters, configure the network to display only the nodes that have at least 22 ties (degree at least 22). Provide a screenshot of the degree filter applied.
- (j) Add labels to the nodes. Make them proportional to the node’s size and adjust their size to be clearly visible. Add the screenshot of your current graph.
- (k) Rescale the nodes’ size using the degree value as the node’s size. Adjust the labels to be proportional to the new node sizes. Did your graph change? Attach a screenshot of the graph. Comment on your new graph.
- (l) Suppose you would like to increase the publicity about your company’s products. You want the most people to know about your product as fast as possible. You have a free sample that you can give away to any member of the network. You are choosing between members 3566, 3948, and 3933. Which would you choose? Explain your answer and show screenshots to validate.