

INTRODUCTION TO NETWORK SCIENCE

EXERCISE 1

Josip Hanak

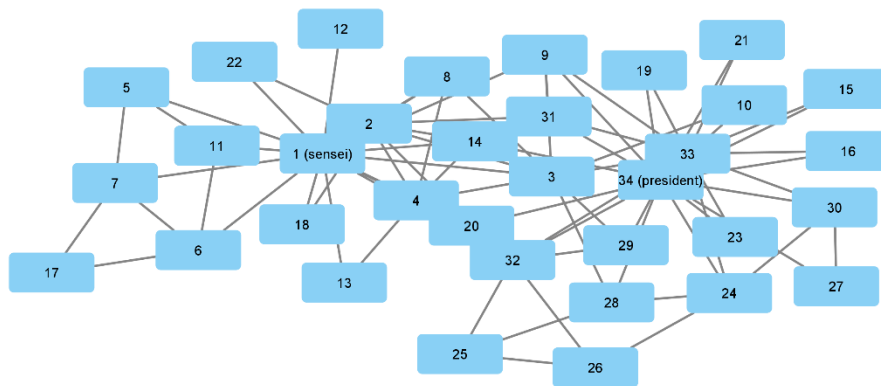
1.

In this assignment we are working with three graphs: starwars, us_companies_ownership and karate. Starwars graph represents the characters in the Star Wars saga and their relations between them. This graph is the second largest and has only one connected component, symbolizing that all the characters are connected. It has 110 vertices and 440 edges.

Us_companies_ownership is the largest graph. It represents the companies in the US their relations and origins. It is consisted of 718 components, 7172 vertices and 6721 edges. Interestingly it has more vertices than edges, symbolizing that a great percentage isn't affiliated with other companies.

Karate is the smallest graph representing a karate dojo organization. It is made up of 1 component (it isn't a franchise), 34 vertices and 78 edges. It has twice the amount of edges as to vertices representing a strong connectivity among the vertices.

1.1.1



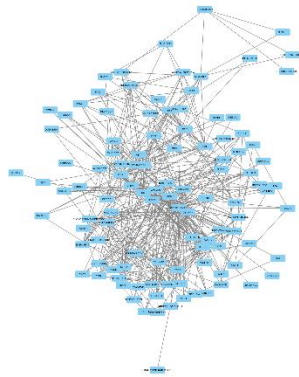
The two highest degree nodes are 34 (president) and 1 (sensei). I got that by analyzing the network as a node table (not edge table) and then sorting by number of degrees.

1.1.2

Force directed graph drawing is a algorithm for drawing graphs which represents the data in a more aesthetically pleasing way and it delivers more information visually. First all the nodes are shown connected by equal length edges and then the edges are adjusted depending on the forces that is present in the data. The forces may be represented through pushing (Coulombs law) or pulling (Springs or gravitational force).

1.2

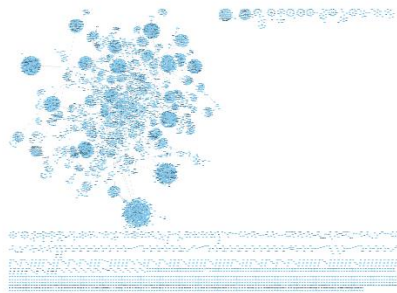
C3P0



1.2.1 As shown in the picture C3P0 has a degree larger than 22 which is 35.

1.2.2 The only two characters with larger degrees are Obi-Wan(36) and Darth Vader(52)

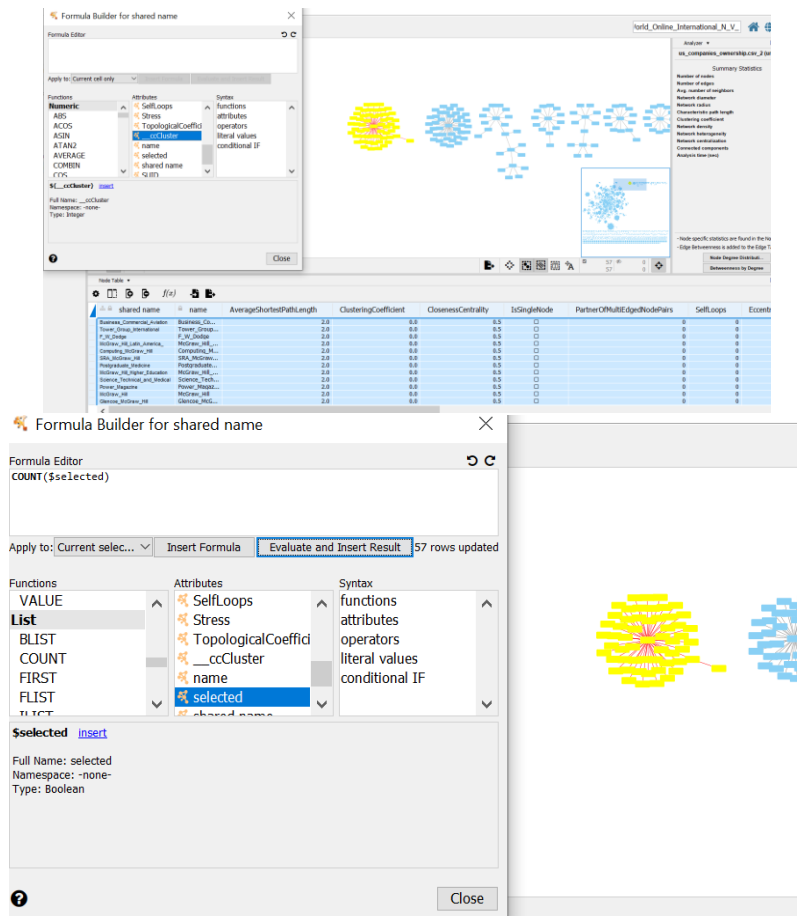
1.3



A connected component is a set of vertices that are linked to each other by a set of paths (each node is reachable to another by a path). That said the us_companies_ownership contains 718 connected components in total. I tried an automatic path for analyzing components with cytoscape. First I tried the Tools > Analyze Network, selecting a connected component with CTRL + drag over vertices in a connected component but the Analyze Network option hadn't changed. So then I did research and found an article (<https://www.biostars.org/p/454441/>) that recommended using an app ClusterMaker (I did this before reading the 3rd and 4th assignment) but even with that I hadn't made any progress. So all the observations are made by my eye.

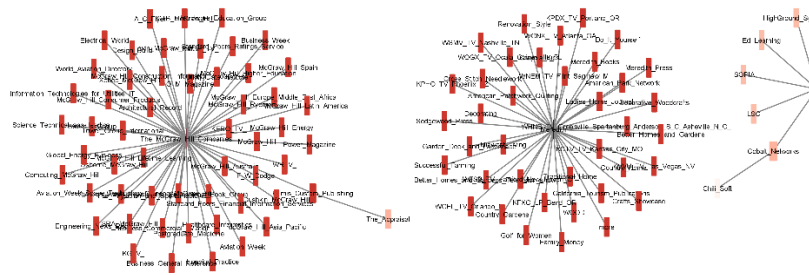
Edit:

When selecting a component with CTRL + drag there I saw the option f(x) in the table on the bottom of the screen that might have the option for gathering more data through functions.



With this method I calculated that the second largest connected component is 57 and the second largest one is 39

1.3.1

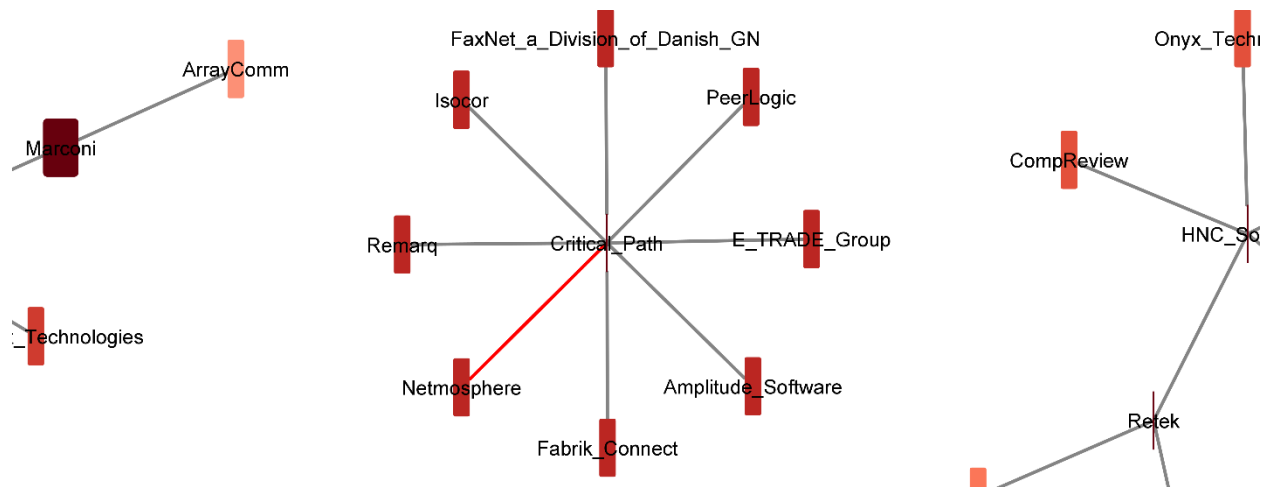


Second and third largest connected component

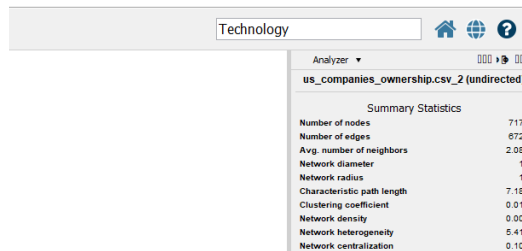
1.3.2

I did not see a cycle in this graph. Perhaps there is but the layout of the vertices are such that a cycle is difficult to notice. I presume there is an algorithm for finding a cycle – perhaps through the Cytoscape.js and importing a DFS algorithm for cycle detection.

1.3.3



I found a large company in the Technology sector by typing Technology in the Node search bar.



This query produced all the Nodes connected to Technology (I think). Then I ordered them by degree and selected a node with a degree of 77 (it has 77 neighbors).






The_Business_Of_Kagan_World_Media

I am not 100% sure that this is connected to Technology (I couldn't find the `_Technology` node) but I suppose this is the right action. It is a star like subgraph, all the surrounding nodes are connected to the main one and they are not connected between each other.

3.1

Sensei, 3 and president have the largest closeness centrality (sorted the table after analyzing).

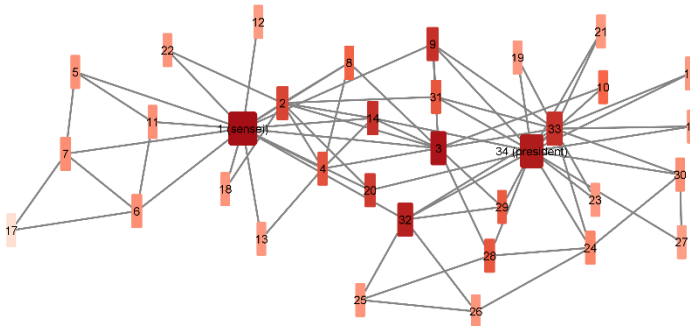
 shared name	 name	 label	AverageShortestPathLength	ClusteringCoefficient	ClosenessCentrality ▾	IsSi
1 (sensei)	1 (sensei)	1 (sensei)	1.7575757575757576	0.15	0.5689655172413793	
3	3	3	1.7878787878787878	0.24444444444444444	0.5593220338983051	
34 (president)	34 (president)	34 (president)	1.8181818181818181	0.11029411764705882	0.55	

3.2

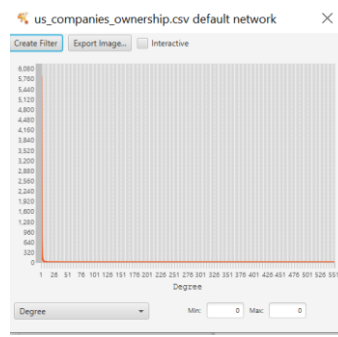
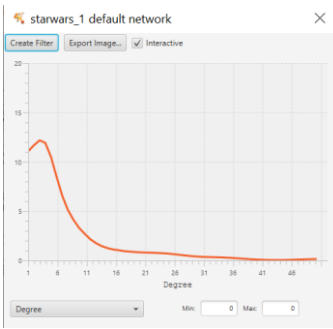
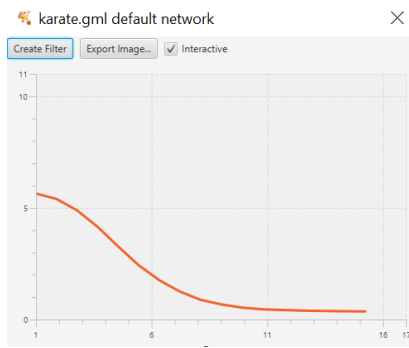
Sensei, 34 and 33 have the largest betweenness centrality (same algorithm as the task before).

1 (sensei)	1 (sensei)	1 (sensei)
34 (president)	34 (president)	34 (president)
33	33	33

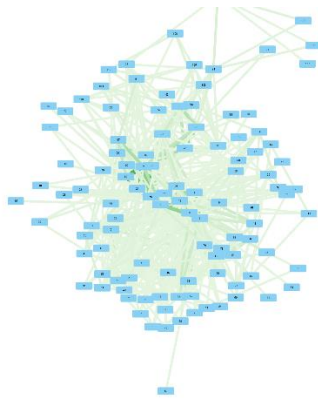
3.3



3.2

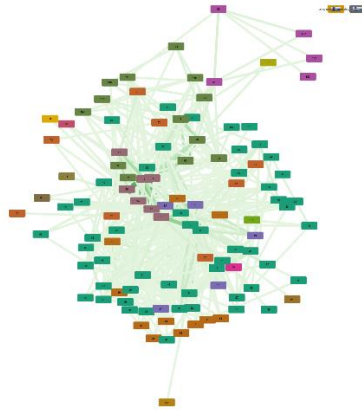


3.3



4.2

4.2.1



4.2.2

Cytoscape calculated 15 types of clusters and attributed a colour to them. The largest cluster contains about 50% of all vertices and the second cluster contains about 25-30%, then the clusters spiral down to 1 vertex in size. It probably suggest groups which vertices belong to, there was an example in theory class that some students sided with the president and the other with the sensei maybe that is the case?

4.2.3

I hereby declare that all of the text, tables, and figures in this report were produced by myself.