Practice Session 01: Cytoscape Basics

Tània Pazos Puig
tania.pazos01@estudiant.upf.edu

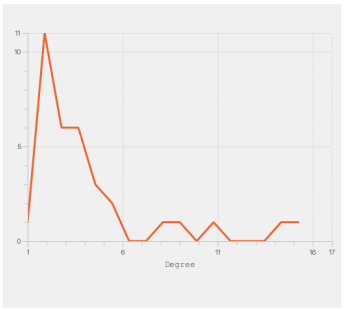
September 28, 2023

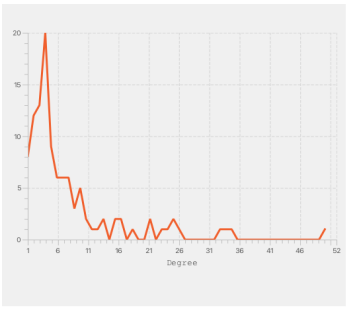
1. Networks description

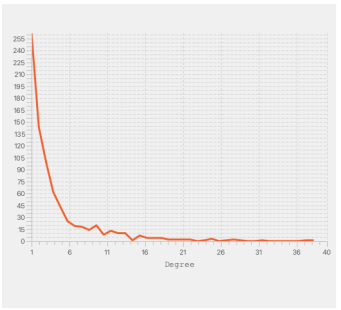
The first network, Zachary's Karate Club, depicts 34 disciples of a karate club as well as the relations between them and other individuals outside the club.
The second network represents all characters that perform together in a particular scene of a Star Wars movie, each edge representing a scene where they appear together.
Lastly, the third network is a bipartite graph that illustrates a burrow of tortoises in the Nevada desert.

	Zachary's Karate Club	Star Wars characters	Tortoises
Nodes	34	110	787
Edges	78	444	1713

Figure 1: Table of nodes and edges of each of the three networks.







Figures 2,3,4: Node distributions of Zachary's Karate Club, Star Wars characters and tortoises graphs respectively.

2. *Karate Club* graph

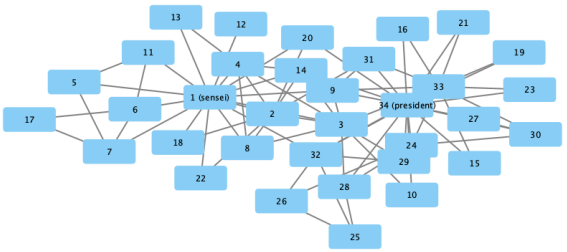


Figure 5: Zachary's Karate Club network. Node 34 (president) has degree 17, node 1 (sensei) has degree 16, and node 33 has degree 12.

Force directed graph drawing algorithms provide a simple, visual and organised representation of a graph. The graph that results from such algorithms is depicted in two or three dimensions and is often symmetric, with as few crossing edges as possible.

After analyzing the node table of the Karate Club graph, it can be noticed that the three nodes with largest closeness centrality are node 1 (sensei) with closeness centrality 0.5689655172413793, node 3 with closeness centrality 0.5593220338983051, and node 34 (president) with closeness centrality 0.55.

Regarding the betweenness centrality, node 1 (sensei) has betweennes centrality 0.4376352813852814, node 34 (president) has betweennes centrality 0.304074975949976, and node 33 has betweennes centrality 0.14524711399711396.

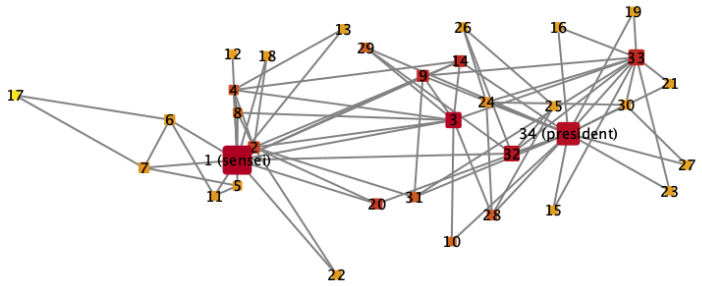


Figure 6: Karate Club graph. The size of its nodes is directly proportional to the betweenness centrality and the redness of the nodes is directly proportional to the closeness centrality.

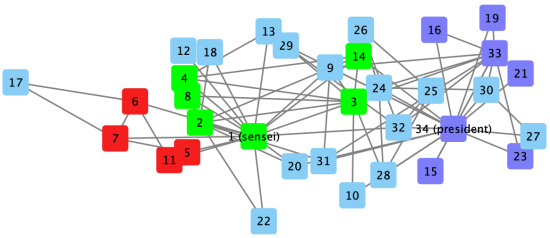


Figure 7: MCODE Cluster algorithm of the Karate Club graph.

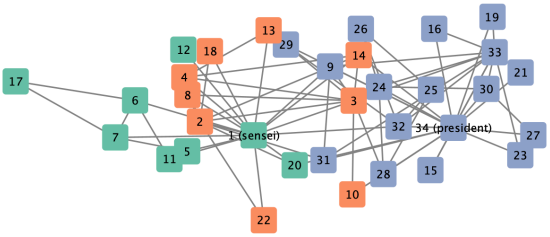
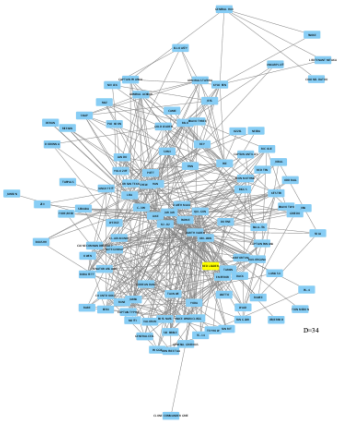
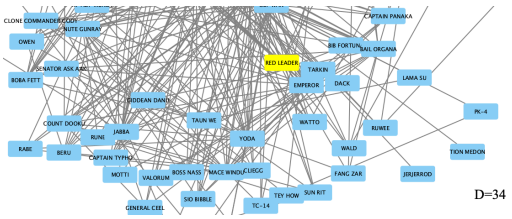


Figure 8: Community Clustering (GLay) algorithm of the Karate Club graph.

In both representations it can be seen that there are two main nodes around which the clusters are formed: node 1 (sensei) and node 34 (president).
On the one hand, the MCODE representation indicates that the green and red clusters are clearly attached to the sensei while the purple cluster is attached to the president; the light blue cluster is rather a “grey area”: some of the nodes are attached to the president and others to the sensei. On the other hand, the GLay representation suggests that the green cluster is linked to the sensei whereas the purple cluster is linked to the president. Again, there is a “not-so-clear” orange cluster with some nodes connected to the close contacts of the sensei and other nodes near the social circle of the president.
As a result, it can be foreseen that the karate club would split into two main groups -formed by more than one cluster in the attached representations.

3. *Star Wars* graph





Figures 9 and 10: Star Wars character network. The yellow node represents a character that appears together with other characters of the graph in more than 25 scenes, in particular in 34.

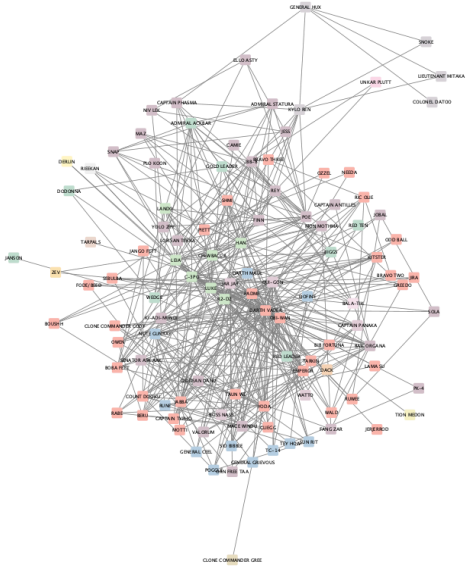


Figure 11: Star Wars network where each color represents a cluster. Considering the fact that the array source of the cluster is called “scenes”, it can be deduced that clusters are formed by characters that have the most scenes in common, that is, that often appear together in the same scene. For instance, a cluster is formed by C-3PO and R2-D2 -androids that are friends-, and Luke -who buys C-3PO-. Thus, it seems logical that they appear together in most scenes.

4. *Tortoises* graph

Figure 12: Tortoises network, formed by 104 connected components.

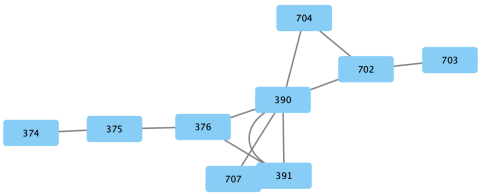


Figure 13: Second largest connected component -9 nodes- of the tortoises graph.

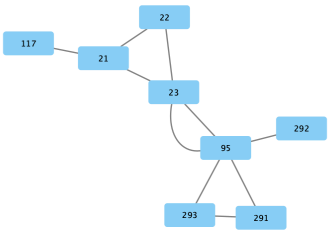


Figure 14: Third largest connected component -8 nodes- of the tortoises graph.

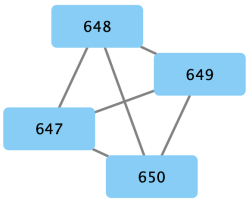


Figure 15: Largest connected component in the tortoise graph that is a *clique*, with 4 nodes.

Figure 16: Largest connected component in the tortoise graph that is a *line graph*, with 5 nodes.

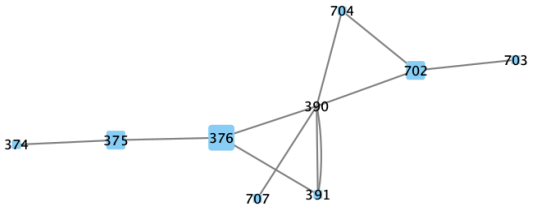


Figure 17: Close up of one connected component of the tortoises network after having set the size of the nodes larger for those nodes with high betweenness. Note that node 702 -with betweenness 0.25- is larger than node 703 -with betweenness 0.0-.

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