OLD DOMINION UNIVERSITY

CS 495: Introduction to Web Science
Instructor: Micheal L. Nelson, Ph.D
Fall 2014 Thursdays 4:20pm – 7:10pm ECSB 2120

Assignment # 7
Joseph Elder UIN: 00844802

Honor Pledge
I pledge to support the Honor System of Old Dominion University. I will refrain from any form of academic dishonesty or deception, such as cheating or plagiarism. I am aware that as a member of the academic community it is my responsibility to turn in all suspected violations of the Honor Code. I will report to a hearing if summoned

November 3, 2014

Contents

1 Assignment 7

1.1 Question 1

1.1.1 The Problem

(10 points)

1. Using D3, create a graph of the Karate club before and after the split.

- Weight the edges with the data from: http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/zachary.dat

- Have the transition from before/after the split occur on a mouse click.

1.1.2 The Solution

The data file “zachary.dat” which contains the matrices which are representative of the karate club is downloaded from the Internet to be used for the dataset for this assignment. The python file “ConvertData.py” is written in python 2.7.6 and is used to generate data into an easily usable format called JSON. JSON stands for JavaScript Object Notation. Since D3 is a JavaScript library which will be used to handle the data and display a graph on an HTML website, using JSON as a data format is an appropriate choice. “ConvertData.py” writes the file “before.json”.

The file “before.json” will ultimately be used as input data for creating graphs which show connections, connection weight, and the factions for club members using D3. The JSON is composed of two lists, one containing the nodes of the graph the other contains the information about how nodes connect or link together. The nodes list contains two attributes name and group. The name attribute is generated by simply iterating through integers until there are 34 total, starting with the id of 0. For the before graph all club members will be in group 1 since there is only a single group in existence at this point in time.

The second list is generated by parsing the second matrix to find connections between nodes and the strength or weight of those connections. It contains three attributes for each object: source, target, and value. The source attribute is the selected node, the target attribute is a connection which that selected node has. The value attribute represents the weight of the connection between the source and target nodes.

The style of graph I have chosen to display this information is called a Force Directed Graph. This type of graph will be used for the before and after graphs. The file “before.html” displays the D3 graph of all nodes and their connections or edges before the split of the group. All nodes are a blue circle and all edges are a gray line as shown in figure 1. To explore the data of the graph in 'before.html' node labels will be displayed when hovering over a selected circle with the mouse cursor. In another version of the same graph 'beforeL.html' labels are displayed in close proximity to each node.

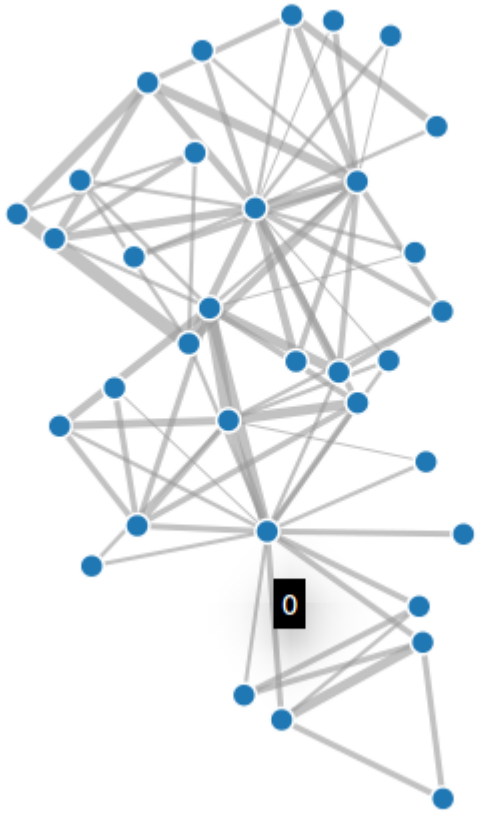

Figure 1: Graph displayed on 'before.html'.

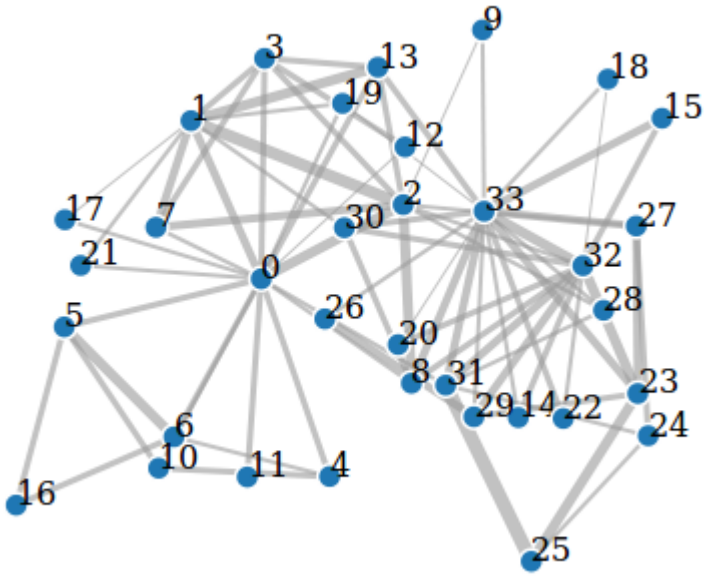

Figure 2: Graph displayed on 'beforeL.html'

For both graphs the edge weight can be seen by the relative thickness of the gray line representing each edge, where thicker lines equal stronger connections. When either 'before' graph is clicked the respective 'after' graph will be displayed by loading another web page. The files “after.html” and “afterL.html” show the karate club similarly to the previous two except after the division in the group. The data for these files comes from 'after.json' a file which is based on 'before.json' where the only difference is the group attribute in the JSON designates which faction that member belongs to after the fission in the group.

This group assignment was done manually for the sake of simplicity since the data set is relatively small. The group assignment was done by referencing the original study written by Zachary. For the graphs showing the club after the division the two factions can be seen by the different colored nodes on the graph. Note that all graphs generate slightly different each time and have a dynamic nature, while the data will always be the same the graph can be slightly changed visually by refreshing the page. All of these web pages can be found on github and additionally on my ODU CS website at www.cs.odu.edu/~jelder/cs495/before.html.

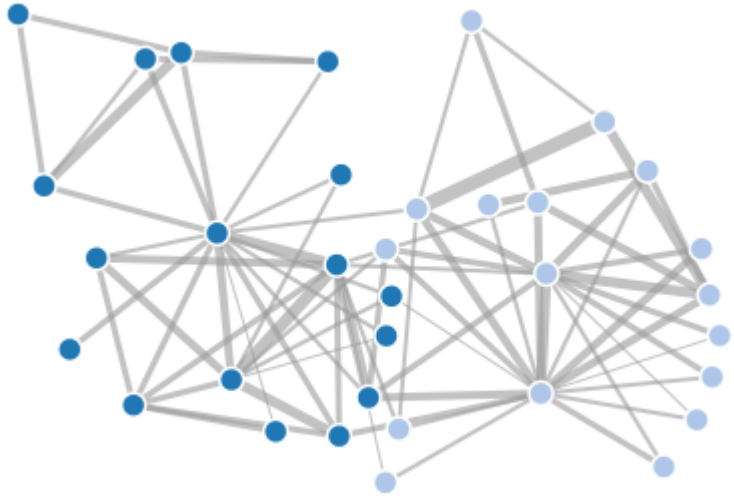

Figure 3: Graph displayed on 'after.html'

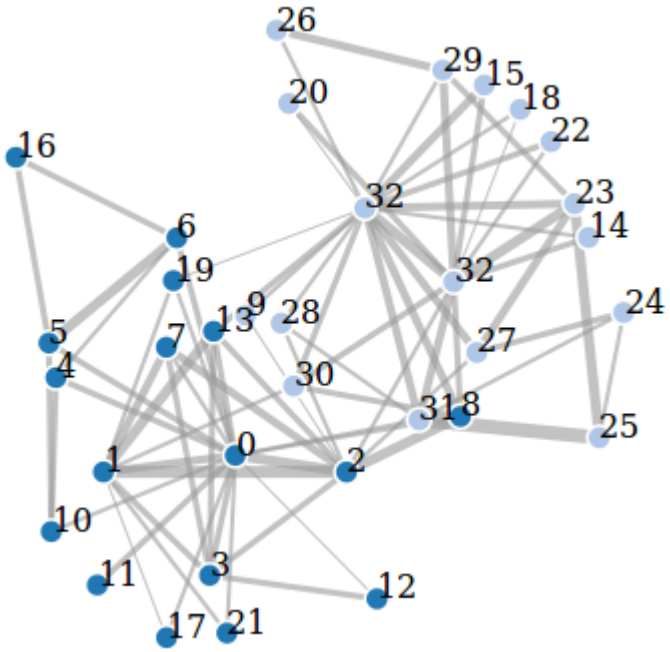

Figure 4: Graph displayed on 'afterL.html'

References

1. https://docs.python.org/2/
2. http://d3js.org/
3. http://vlado.fmf.uni-lj.si/pub/networks/data/ucinet/zachary.dat
4. http://aris.ss.uci.edu/~lin/76.pdf