**ME274 Spring 2015 Report**

**Layout:**

**Based on our situation, we decided to apply for ForceAtlas 2 algorithm.**

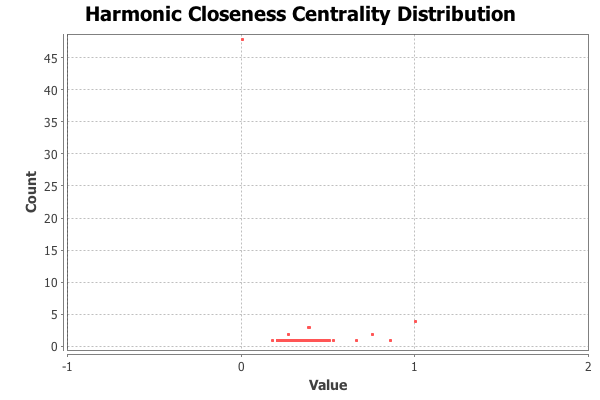
* **Modularity**

I can’t run modularity on my laptop due to a software issue ;(

* **Network Diameter:**

**Diameter: 9 The distance from graph center to the furthest node.**

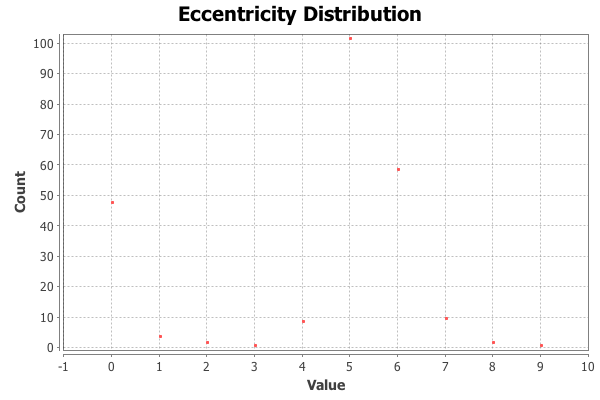
**Closeness centrality**

****

Under Network Diameter, closeness centrality measures the efficiency of a vertex in spreading information to other vertices. The larger the closeness centrality a vertex has, the shorter the average distance from this node to others. From the graph, it is not difficult to find that a few students have a closeness centrality above 0. Most students are just at point 0. Only one guy has a value 1 centrality.

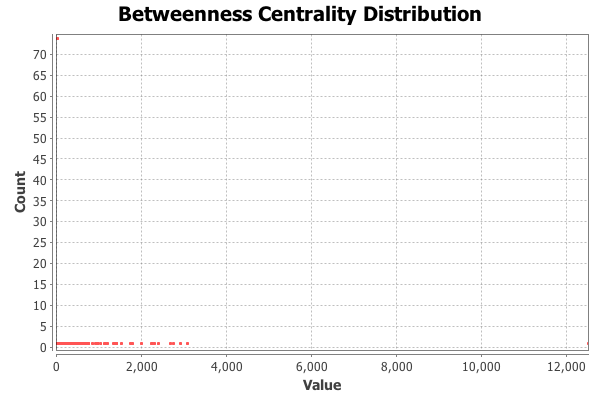
Closeness centrality is a very interesting statistic because it is inverse to the average path distance. The algorithm is based on the an eigenvector formula: λv = Av. Then we could apply: (A- λI) \* v = 0. Then v is the null space of (A-λI). If a node A links to another node B, which has a high degree, the eigenvector of A, v, has a high score and then A has a higher closeness. This is because B has a number of edges as A has a effect to B. Therefore, we say the efficiency of A of spreading information is high.

**Eccentricity Distribution**

****

Under Network Diameter, eccentricity distribution showed us the distance between a node and the node that is furthest from it. From the graph, it is shown that actually quite a few students(like 100) replied some messages under one’s post but not directly replied the poster.

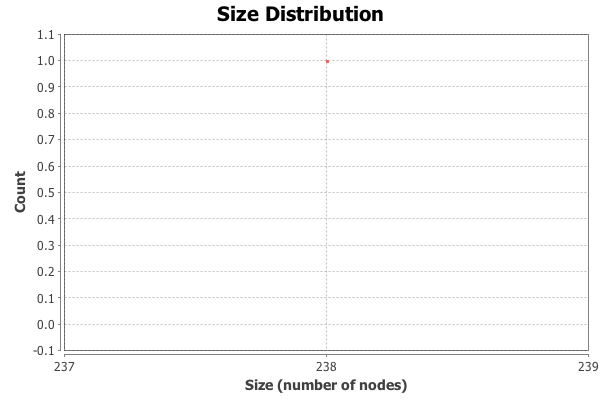
**Betweenness centrality**

****

Under Network diameter, betweenness centrality measures how often a node appears on the shortest paths between nodes in the network. From the graph we can see that except one node (the teacher) has a very high value all others are close to 0.

The betweenness centrality didn’t give us too much useful information. This is because in our situation most students just reply others message directly without passing others. That is why we can’t get too much from the graph.

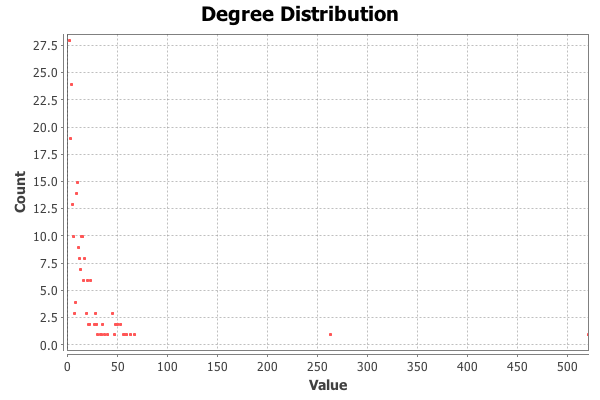
* **Connected Components**

****

A connected component is that any two vertices are connected to each other by paths, and is connected to no additional vertices. It is an important topological invariant of a graph and it is straightforward to compute by connections in linear time. From the report we can see that the number of weakly connected components is 1 and strongly connected components is 78.

Connected component implies how continuousness of a network graph. Since I used a directed graph, strongly connected components implies that there are 78 nodes reachable from every other vertex. Weakly connected components infers that there is on e subgraph which has the maximal number of paths.

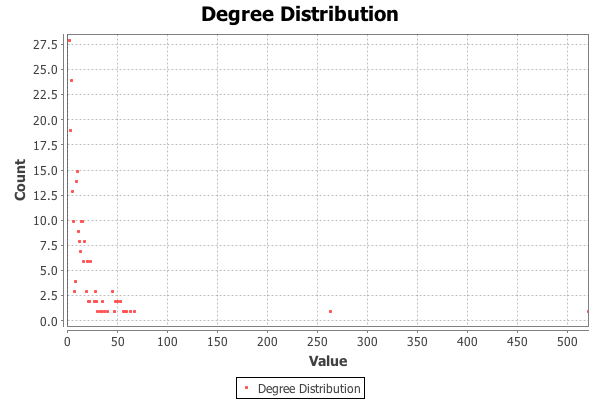
* **Average degree**

****

The average degree of a node shows the number of edges that are adjacent to this node. From the result we found that the average degree is 15.311. However, the results on Degree Distribution graph show that almost all students talked to others within 50 times, and just one node has a count of over 500.

The average degree here helps us understand the average engagement in this class. However, as what we can see from the graph, the teacher is very involved but some students never involved. In this case, the average degree is not persuasive enough to represent the extent of engagement. So I would recommend the average weighted degree in the following below.

* **Average weighted degree**

****

Average weighted degree is quite similar with average degree. It is based on the number of edges on a node, but it is doing the sum weight of edges. It is hard to say a weighted degree of a node. Instead, average weighted degree is more like the mass center of the overall edges. In another way, think about the difference between median and mean. Average degree is the mean statistic, average weighted degree is kind of like the median. According to this formula, the weighted degree is 7.655 and the graph has the same feature as that of average degree.

The average weighted degree is more helpful compared with average degree because we are kind of finding the center mass of edges. In our case, the teach is very involved but students are not. After weighted, the score represents the sum of every engagement. This averts the effect of a special point; The instructor had a super high degree as he posts a lot but other students didn’t.

* **Graph density**

Graph density measures how close the graph is to complete. For our directed parameter, our density is 0.032.

Graph density is not helpful because there are too many posted blogs without any response. That’s why we get a very low density graph.