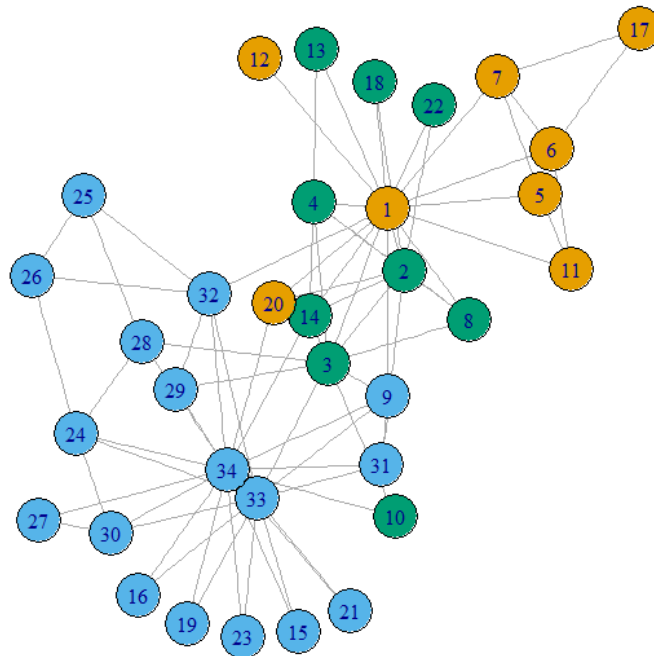


Final Report Assignment 5

Question 1:

In order to prove that a weighted graph of social interaction could be used to predict the result of the split, I created this test case from R of a karate <- make_graph("Zachary") to create a Zachary karate club data graph. The nodes, connections to nodes and degree of nodes are shown below for the data. I also plotted the graph to give the reader a better understanding of how the nodes actually are connected on the next page.

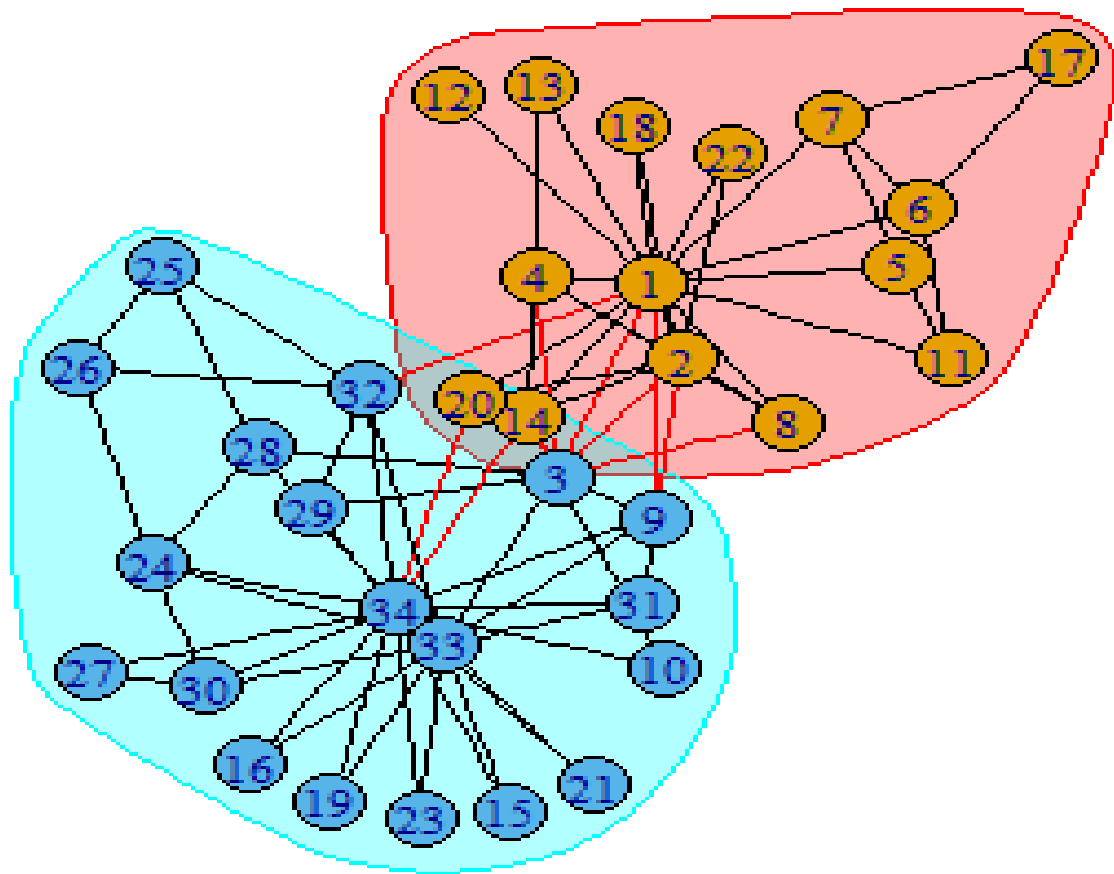
| Number of Node | Node Connected to | Degree |
|----------------|---|--------|
| 1 | 2,3,4,5,6,7,8,9,11,12,13,14,18,20,22,32 | 16 |
| 2 | 1,3,4,8,14,18,20,22,31 | 9 |
| 3 | 1,2,4,8,9,10,14,28,29,33 | 10 |
| 4 | 1,2,3,8,13,14 | 6 |
| 5 | 1,7,11 | 3 |
| 6 | 1,7,11,17 | 4 |
| 7 | 1,5,6,17 | 4 |
| 8 | 1,2,3,4 | 4 |
| 9 | 1,3,31,33,34 | 5 |
| 10 | 3,33 | 2 |
| 11 | 1,5,6 | 3 |
| 12 | 1 | 1 |
| 13 | 1,4 | 2 |
| 14 | 1,2,3,4,34 | 5 |
| 15 | 33,34 | 2 |
| 16 | 33,34 | 2 |
| 17 | 6,7 | 2 |
| 18 | 1,2 | 2 |
| 19 | 33,34 | 2 |
| 20 | 1,2,34 | 3 |
| 21 | 1,2 | 2 |
| 22 | 33,34 | 2 |
| 23 | 26,28,30,33,34 | 5 |
| 24 | 26,28,32 | 3 |
| 25 | 24,25,32 | 3 |
| 26 | 30,34 | 2 |
| 27 | 3,24,25,34 | 4 |
| 28 | 30,34 | 2 |
| 29 | 3,32,34 | 3 |
| 30 | 24,27,33,34 | 4 |
| 31 | 2,9,33,34 | 4 |
| 32 | 1,25,26,29,33,34 | 6 |
| 33 | 3,9,15,16,19,21,23,24,30,31,32,34 | 12 |
| 34 | 9,10,14,15,16,19,20,23,23,24,27,28,29,30,31,32,33 | 17 |



Since the data was created, I used the Girvan–Newman algorithm which is used to help determine the community structure of the graph. This process is summarized as:

1. The betweenness of all existing edges in the network is calculated first.
2. The edge with the highest betweenness is removed.
3. The betweenness of all edges affected by the removal is recalculated.
4. Steps 2 and 3 are repeated until no edges remain.

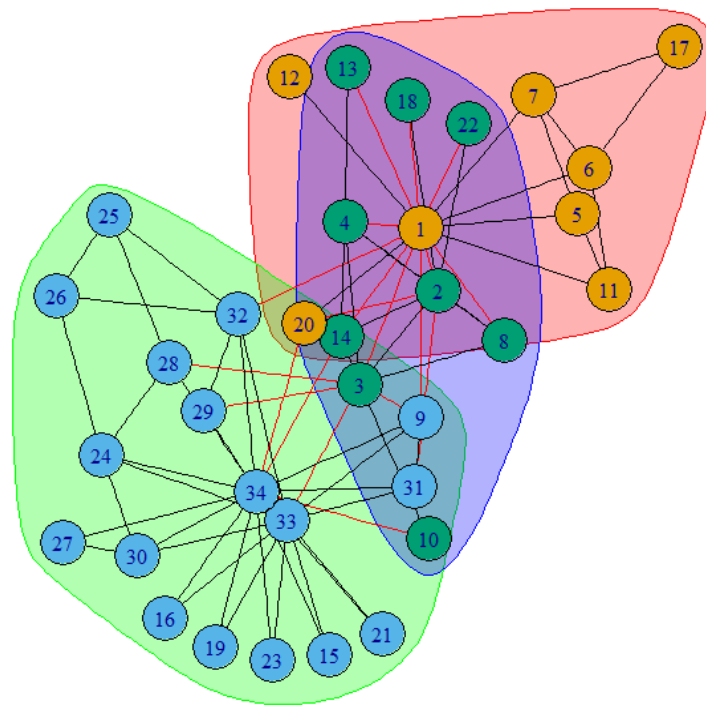
It calculates the weights for each of the pair of edges based on the shortest path between a pair of nodes. This process is defined in R as `edge.betweenness.community`. It is an iterative process that will continue until some sort of community structure is shown. I plotted the result of the community structure below based off of the example data.



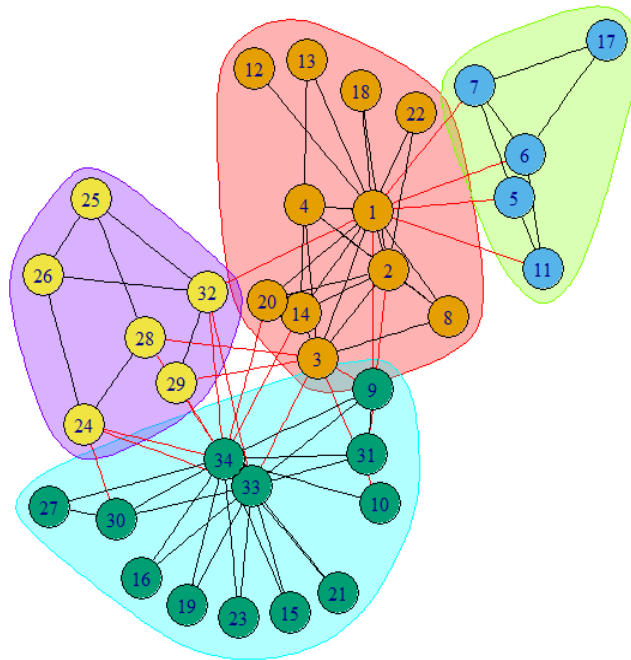
The Girvan-Newman Algorithm can be used as a way to define weights for the graph in order to predict the Zachary karate club split. This algorithm with its use of the weights for helping to remove edges and finding communities and structure of similarities together could have been used to predict the split and how it could have split. I believe that this model from the Girvan-Newman Algorithm reflects reality accurately. This is because it determines this centers of communities that are separate from each other just like cities and even people are. There are always central authorities that have more knowledge or even more relationships between other people who would either have separate ideals like in the Karate club or they are a part of the same group. This model helps to define these leaders and separate the communities based off of these connections which makes it an accurate representation.

Question 2:

In order to calculate the group of three, I used the fast_greedy_community method which tries to optimize a modularity by initially computing that every node belongs to its own community. It then slowly builds up to larger communities by merging neighboring nodes in order to optimize itself locally. This helped me to break up the groups in 3 communities as shown below:



In order to get a group of 4, I used the spinglass_community method which is an approach that involves statistical physics. The way it works is that each node can be in any amount of spins and interacts with another node to make a pair if they are in the same spin state. It continuously finds pairs that prefer to be in the same spin state and those who do not. This method helped me to figure out what would happen if they were to split in 4 separate groups as shown in the graph below.



In order to get 5 separate communities, I used the walktrap_community method which is based on random walks. It is based off an idea that if someone before random walks then they will be more inclined to stay within there communities since there are only minimal amount of edges that could possibly lead into a new community. It performs these short random walks and separates the communities based on the results of the walks.

