

CS 432/532: Web Science

Spring 2017

Assignment 5

Michael Micros

Instructor: Michael L. Nelson

Old Dominion University
Norfolk, Virginia

March 16, 2017

Problem 1: Predicting the Karate Club split(Zachary,1977)

The goal for this assignment was to determine whether it is possible to predict the members of the 2 groups that formed after the Karate Club split. Based on the data provided in the original paper, an undirected graph can be constructed with each edge representing interactions between members outside of the Karate Club. The idea is that based on these interactions of the members, we can predict which group each member will belong to after the split, or even if there is a dispute and a split imminent.

The data for the experiment was obtained from the UC Irvine Network Data Repository. The specific file is "karate.gml" and can be found at the following URI:

[http:// networkdata.ics.uci.edu/data/karate/](http://networkdata.ics.uci.edu/data/karate/)

The python program "part1.py" performs the prediction using igraph's built in algorithm "community_edge_betweenness", which calculates the betweenness of the edges. The argument "clusters" specifies how many communities we want after the split, and accordingly the edge or edges with the highest betweenness are cut resulting in the 2 or more factions.

In the code, f1 and f2 are the node IDs of the members) that sided with Mr. Hi and the officer respectively, according to the original paper.

```
3 import igraph
4 f1 = [1,2,3,4,5,6,7,8,9,11,12,13,14,17,18,20,22]
5 f2 = [10,15,16,19,21,23,24,25,26,27,28,29,30,31,32,33,34]
6
7 #load in graph
8 g = igraph.Graph.Read_GML("karate.gml")
9 layout=g.layout()
10
11 #plot original graph
12 igraph.plot(g, 'Karate club.pdf', layout = layout, vertex_label=g.vs["id"])
13
14 #plot factions
15 colorDict = {"f2": "blue", "f1": "red"}
16 for v in g.vs:
17     if v["id"] in f2:
18         v["color"] = "blue"
19     else:
20         v["color"] = "red"
21 igraph.plot(g, 'Karate club factions.pdf', layout = layout, vertex_label=g.vs["id"])
22
23 #plot predictions
24 pred = g.community_edge_betweenness(clusters=2, directed=False)
25 predictedSplit=pred.as_clustering()
26 igraph.plot(predictedSplit, "prediction.pdf", layout=layout)
```

Figure 1: "part1.py"

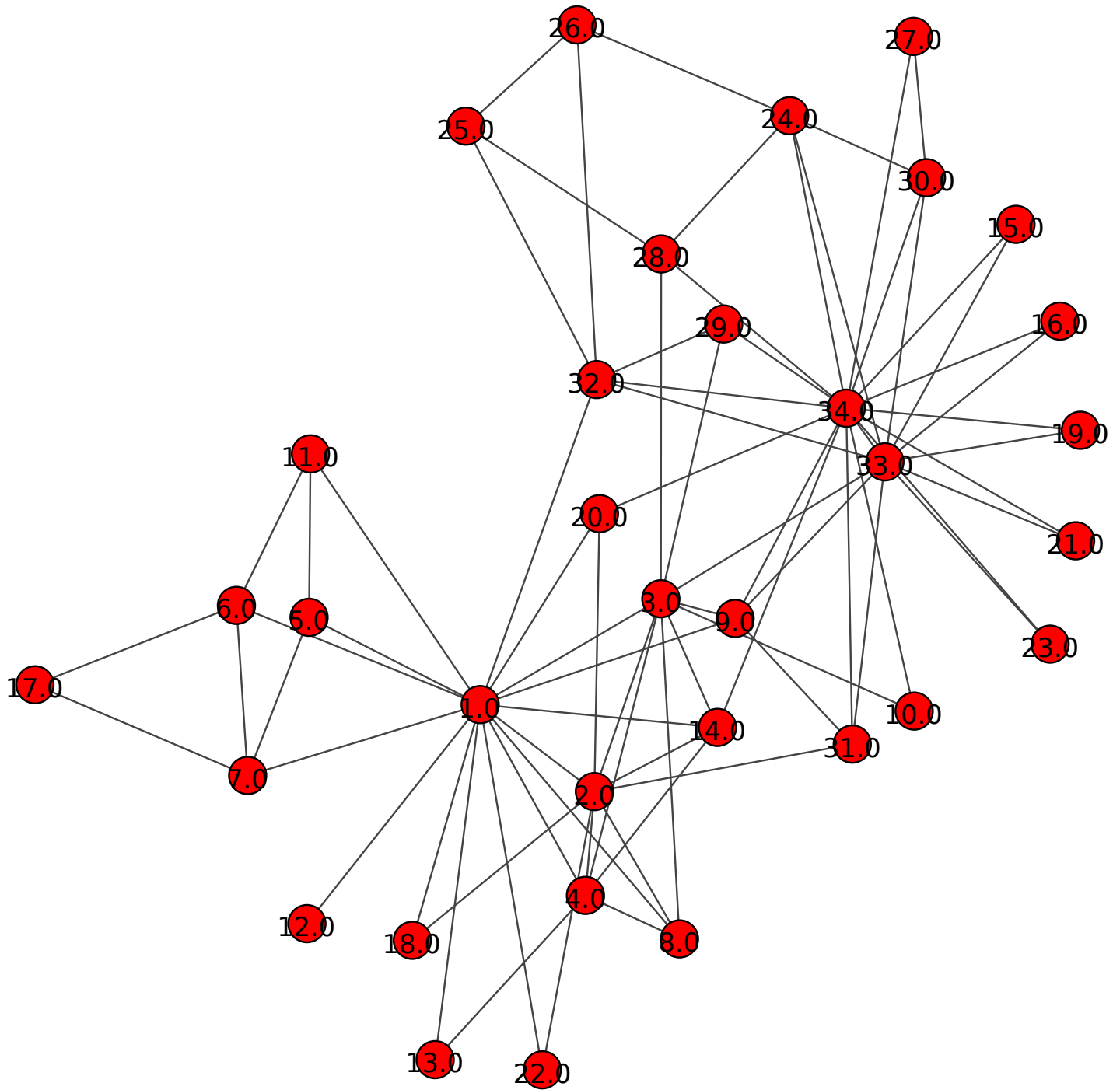


Figure 2: Graph of the 34 Karate Club members from the original paper by Zachary

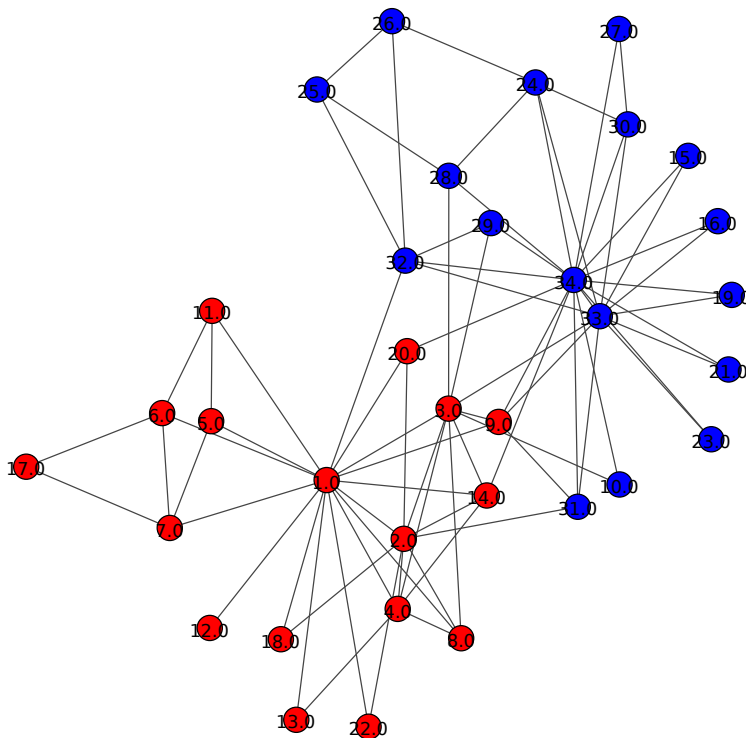


Figure 3: The 2 groups that formed after the split in the original paper.

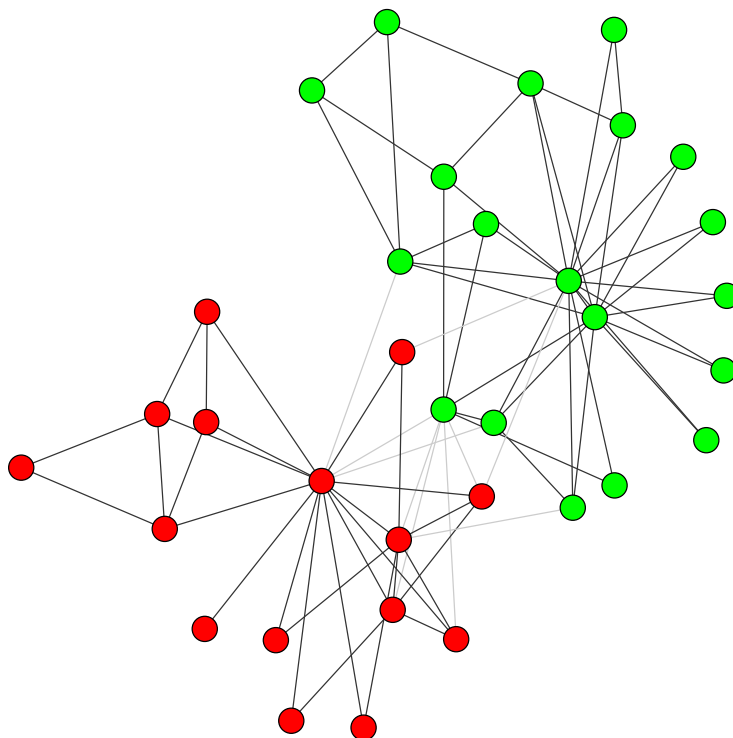


Figure 4: The prediction generated by the python script

By comparing Figures 3 and 4 it is obvious that the program predicted the final groups right with the exception of members 3 and 9, who had interactions with both groups and it is difficult to say with any certainty to which faction they belonged. That is an accuracy of 0.94, which is acceptable.