2017

Assignment 5

CS 432

MIRANDA SMITH

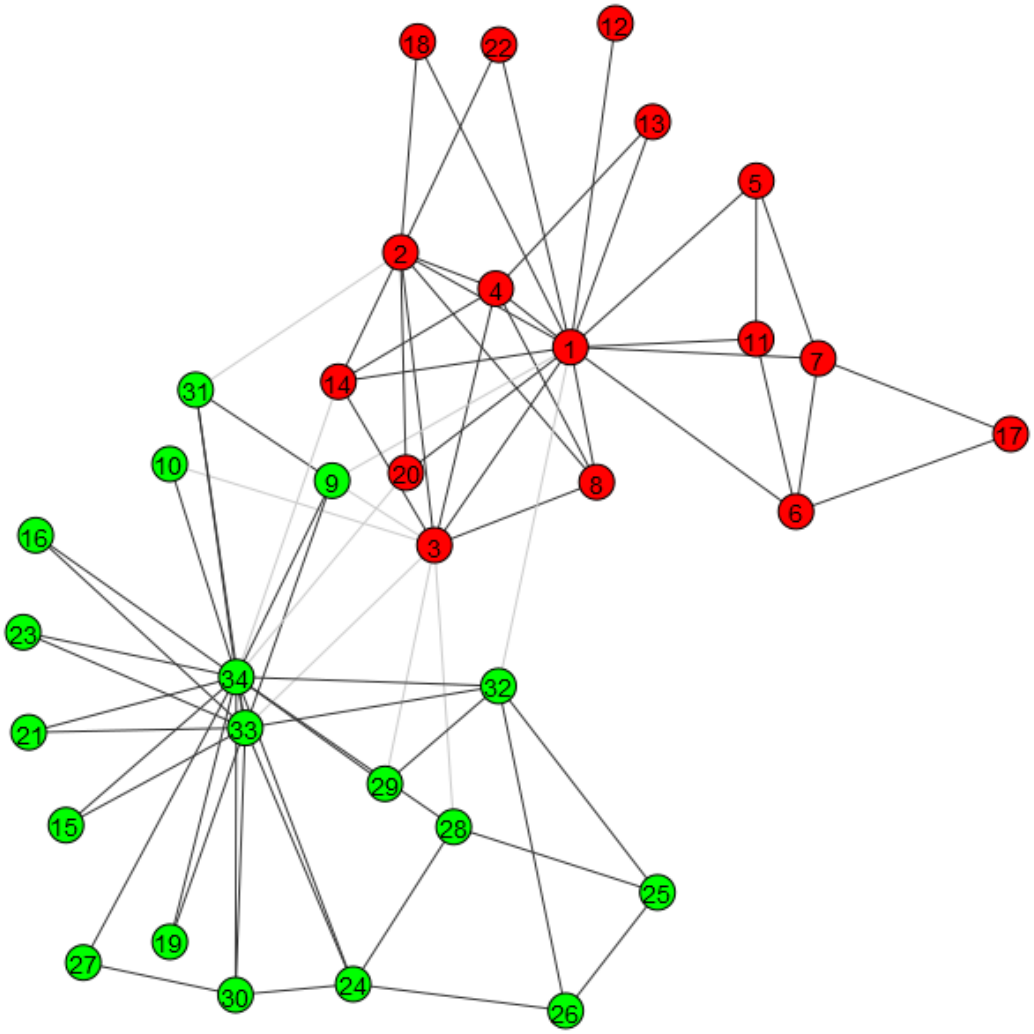
Question 1

We know the result of the Karate Club (Zachary, 1977) split. Prove or disprove that the result of split could have been predicted by the weighted graph of social interactions. How well does the mathematical model represent reality?

I wrote a python program that makes use of igraph to get the karate club graph and model the split. I ran it with two community detection functions that are implemented in igraph, community_leading_eigenvector and community_edge_betweenness because those where the two that allowed the final number of clusters to be specified. The documentation does not indicate this however, it is only found out through looking at the source code of igraph. I chose community_leading_eigenvector because it had one more node correct than community_edge_betweenness. The eigenvector function is an implementation of Newman's eigenvector community structure detection.

```
#!/usr/bin/env python
"""
Zachary's Karate Club graph

Data file from:
http://vlado.fmf.uni-lj.si/pub/networks/data/Ucinet/UciData.htm

Reference:
Zachary W. (1977).
An information flow model for conflict and fission in small groups.
Journal of Anthropological Research, 33, 452-473.
"""
import igraph

groups = open("karateGroups.txt", "w")

g = igraph.Graph.Read_GML("karate.gml")

#labels done in this weird way so that they wouldn't be floats
g.vs["label"] = g.vs["id"]
numbers = list(map(int, g.vs["label"]))
g.vs["label"] = list(map(str, numbers))
#igraph.plot(g)
clusters = g.community_leading_eigenvector(clusters=2)

#dendrogram = g.community_edge_betweenness(clusters=2)
#clusters = dendrogram.as_clustering()

igraph.plot(clusters, "karateClubModeledGraph.png")

membership = clusters.membership

groups.write("ID Group\n")
for name , membership in zip(g.vs["label"], membership):
    groups.write(name + " " + str(membership) + "\n")

groups.close()
```

The output of the previous program is this graph showing the two communities color coded and the removed edges greyed out.



A text file is also outputted ‘karateGroups.txt’ to numerically list which node ID belongs to which group. Comparing that to ‘karateGroupsFromData.txt’ which is the identically formatted observed data from the incident shows that there is only one discrepancy. Node 9 in the model is in group 1 (green), while the actual data shows that node 9 is in group 0 (red).

References:

Data from the Karate Club Split.

http://aris.ss.uci.edu/~lin/76.pdf

Used for help with the programming semantics.

http://stackoverflow.com/questions/25254151/using-igraph-in-python-for-community-detection-and-writing-community-number-for

http://igraph.org/python/doc/igraph.GraphBase-class.html