

# Data Analytics Assignment2-Report

## Community-Detection

Alice Laguri  
21275

September 6, 2023

## Implementation

There are two popular methods for detecting community in a graph based dataset:

- **The spectral decomposition algorithm** uses the eigenvalue and eigenvectors of a network's adjacency matrix to find communities.

- **The Louvain algorithm** is a greedy algorithm that iteratively merges nodes into communities based on their connectivity using. The algorithm starts by assigning each node to its own community. Then, it repeatedly merges the two communities that are most connected to each other. *The algorithm stops when no more communities can be merged.*

The facebook\_combined and sgn\_bitcoin datasets were used. There were no duplicates in the data. Hence, dimension of datasets are (88234, 2) and (71184, 4).

## Results

Number of community in facebook dataset using spectralDecomposition Algorithm: 8

Number of community in facebook dataset using Louvain Algorithm: 99

Number of community in bitcoin dataset using spectralDecomposition Algorithm: 8

Number of community in bitcoin dataset using Louvain Algorithm: 479

This assignment focuses on answering following questions by experiment:

- What would be your stopping criterion?  
The stopping criteria used in spectral decomposition is by ratio cut i.e.,  $\text{conductance} = \text{cut} / \min(v_0, v_1)$ . when conduction is 0.2, Graph partition is returned.
- How would you pick the best decomposition of nodes into communities? Using **modularity gain**, which is a measure of how well the nodes in a community are connected to each other compared to the rest of the graph. The demerge operation is performed when the modularity gain of merging two communities is negative and the merge operation is performed when the modularity gain of merging two communities is positive.
- What was the running time of the Spectral decomposition algorithm versus the Louvain algorithm on the data sets you were given?

Running time of the Spectral decomposition algorithm: 1.2 min for facebook dataset, 1.7 min for bitcoin dataset  
Running time of the the Louvain algorithm algorithm: 37 sec for facebook dataset, 30 sec for bitcoin dataset  
Total run time(including plotting time): 5 min 34.8s

## Conclusion

**In your opinion which algorithm gave rise to better communities, why?**

The spectral decomposition algorithm is a global algorithm, it takes into account the entire network when finding communities. The Louvain algorithm is a local algorithm, it only considers the connections between the nodes that are being merged.

In my opinion, the spectral decomposition algorithm is more accurate than the Louvain algorithm, but it is also more computationally expensive. The Louvain algorithm is less accurate, but it is much faster.

## Plots

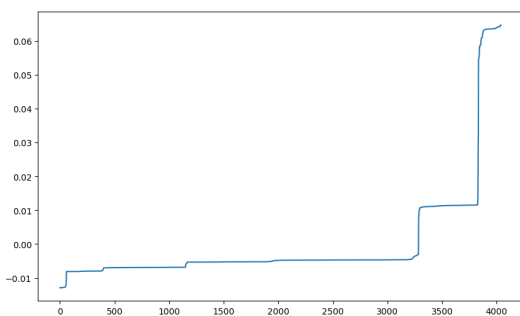


Figure 1: q1)sorted Fiedler vector

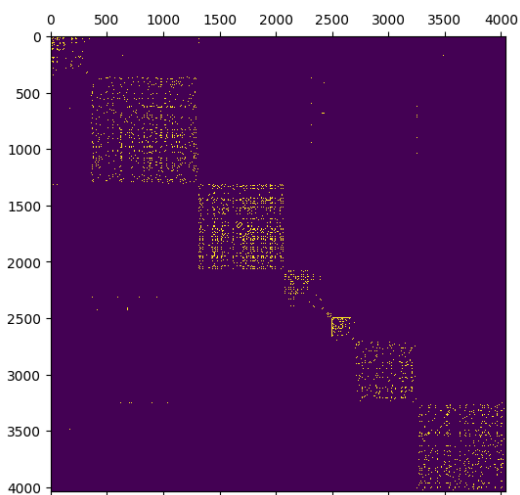


Figure 2: q1)associated adjacency matrix

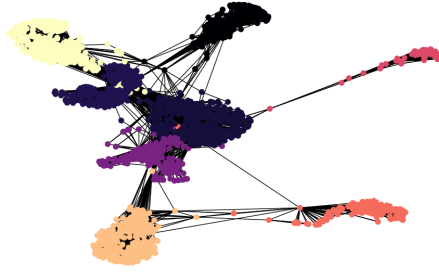


Figure 3: q1)the graph partition

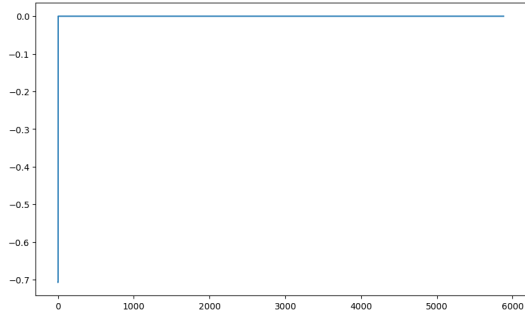


Figure 4: q3)sorted Fiedler vector

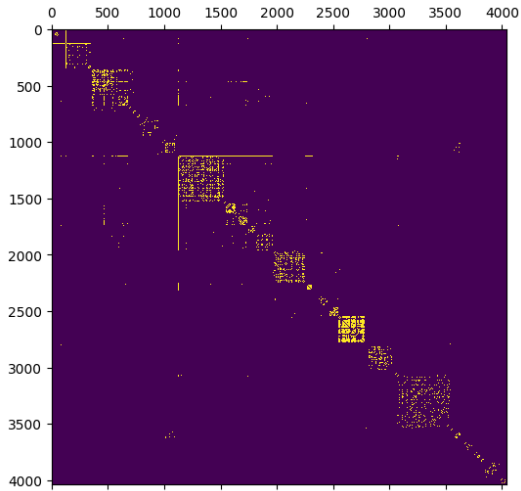


Figure 5: q3)associated adjacency matrix

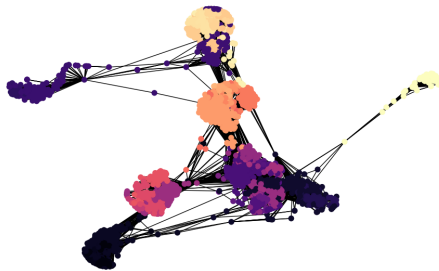


Figure 6: q3)the graph partition