Data Analytics - Assignment #2

Nishant Kumar (Sr.NO. - 21495)

September 2023

Community Detection

1 Community Detection using Fiedler Vector

1.1 Facebook Dataset

Identify two distinct communities within the Facebook dataset by computing the Fiedler vector of the normalized Laplacian matrix. Create a visualization that clearly displays the dolphin names, the network structure, and the two identified communities. Describe your approach, considering that the two communities may have different sizes.
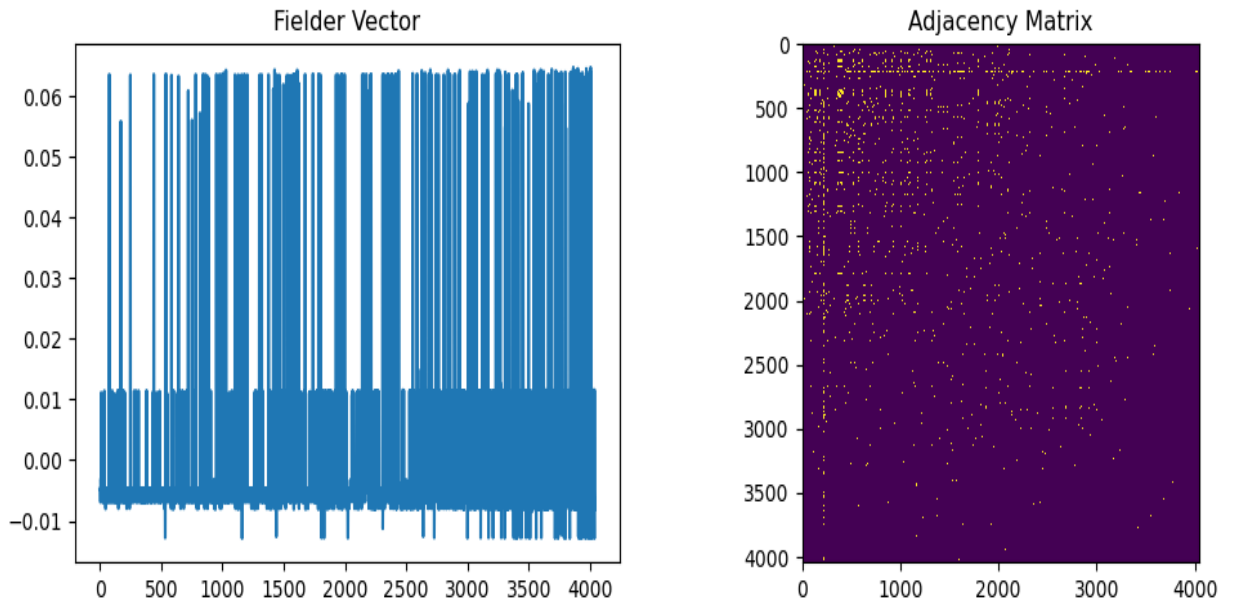
Approach:

- Read the dataset, clean it and make it numpy array nX2 i.e. node connectivity list
- That nX2 numpy array convert it into a graph using *networkx* package.
- Performs spectral decomposition to find the eigenvalues and eigenvectors, sorting them by eigenvalue.
- Now computer the Normalized Laplacian matrix Using the same package *networkx*.
- Extract EigenValues and EigenVectors corresponding to the second smallest eigenvalue {Fiedler Vector} and determine the graph partition based on its values.
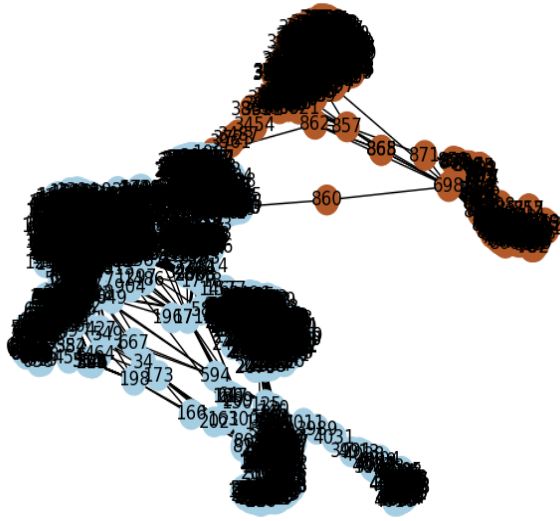- Finally, we plot the sorted Fiedler vector, the adjacency matrix, and the graph partition for visualization.
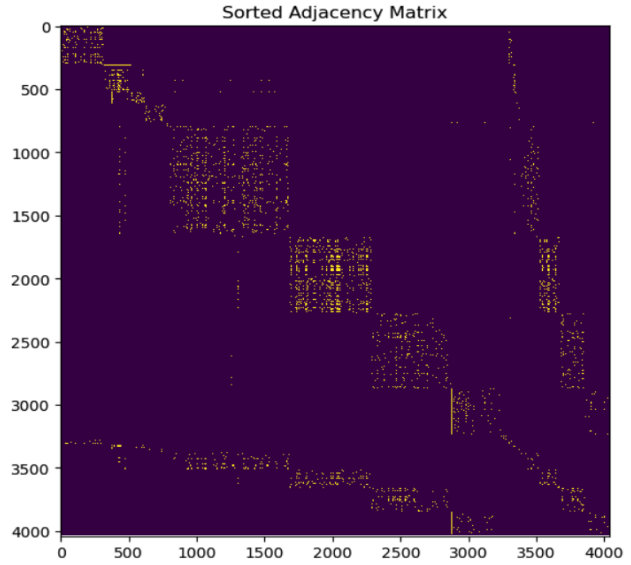
Figure 1: Grpah based on Fiedler vector

Figure 2: Sorted Adjacency matrix

1.2 Bitcoin Dataset

Identify two distinct communities within the Bitcoin dataset by computing the Fiedler vector of the normalized Laplacian matrix. Create a visualization that clearly displays the dolphin names, the network structure, and the two identified communities. Describe your approach, considering that the two communities may have different sizes. The BTC dataset contains 4 columns 'source', 'target', 'rating', and 'time' and we just need to work on the source and target to find out the community in the dataset. the approach is the same as the above that is used on the Facebook dataset, compute the eigenvalues and eigenvectors and then extract the second smallest eigenvector that corresponds to the second smallest eigenvalue, After extracting the Fiedler vector and adjacency matrix plot the graph of partitions.

Graph based on Adjacency Matrix BTC

Sorted Adjacency Matrix BTC
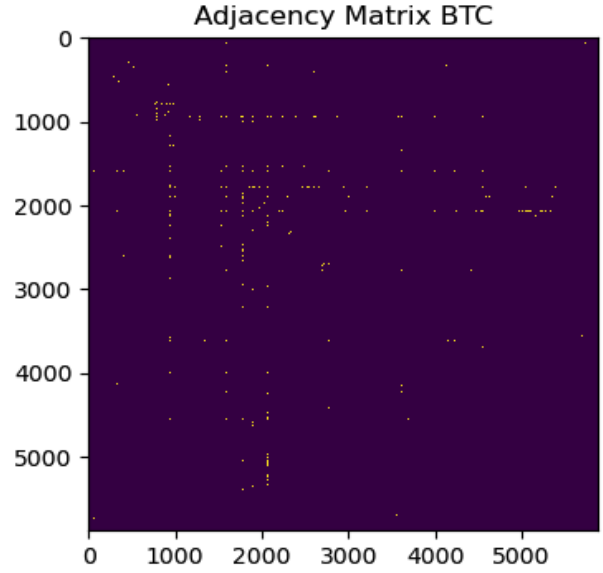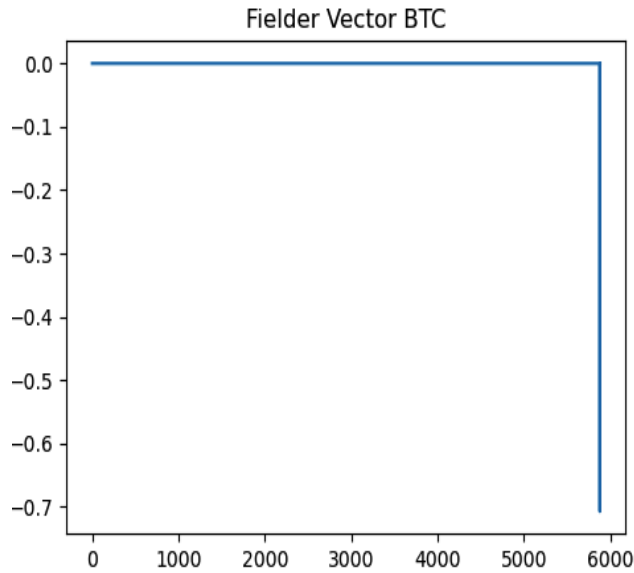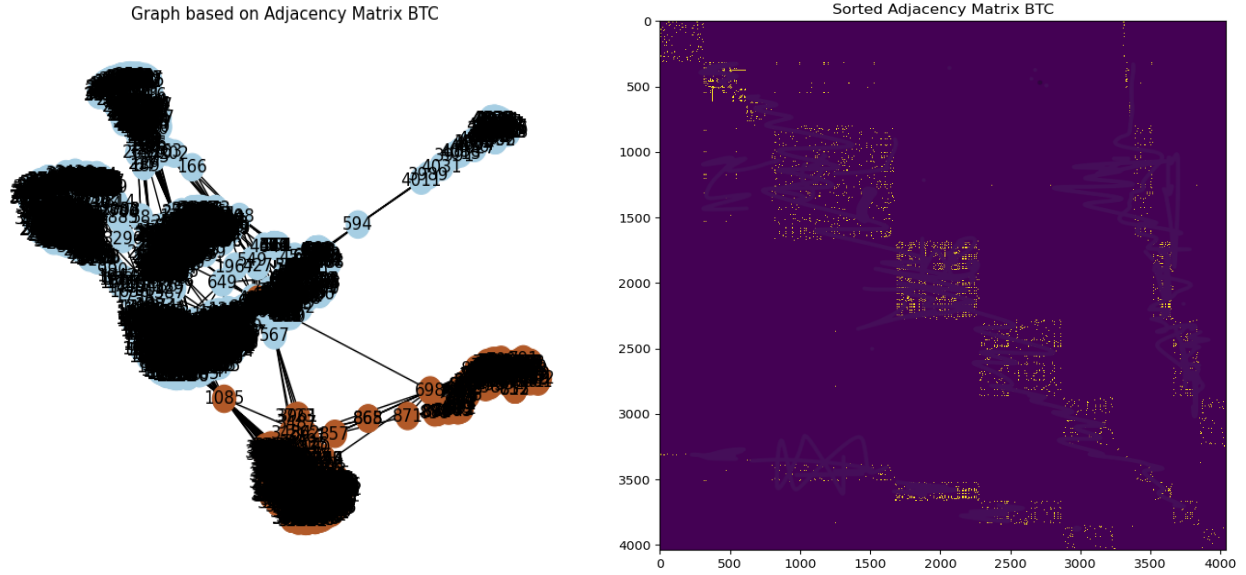
2 Community Detection using Louvain Method

The Louvain community detection algorithm from scratch is a complex task, and it involves several steps, including modularity optimization and community merging.

- We start by initializing each node as its own community.

- We iteratively consider moving nodes to their neighbours' communities if it results in a modularity gain.

- The algorithm continues until no further community merging is possible

- Finally, we visualize the graph with nodes colored according to their communities.

2.1 Facebook Dataset

Louvain Community Detection (One Iteration)