THE ITERATIVE FIEDLER VS. THE GENERALIZED FIEDLER CLUSTERING METHODS

KEIVAN HASSANI MONFARED

ABSTRACT. Here we explore the difference between two clustering methods of simple graphs using an eigenvector corresponding to the second smallest eigenvalue of the Laplacian of the graph: the iterative method and the generalized one. The numerical experiments show that the iterative method on average performs better in most cases.

1. INTRODUCTION

The **iterative Fiedler clustering method** is as follows:
(1) Start with all the vertices in one cluster.
(2) For each cluster:
  (a) consider this cluster as a standalone graph,
  (b) divide the cluster into two clusters by separating the positive and negative components of the second smallest eigenvector of the Laplacian.
(3) Repeat until each cluster consists of only one vertex.

One consideration with this method is which cluster shall be separated first. To address this, we compute the normalized algebraic connectivity of each cluster and choose the one with the lowest algebraic connectivity to be separated first.

Another consideration with this method is that for subgraphs the second smallest eigenvector of the Laplacian does not take into account the external forces that pull vertices together (push away from each other, in the case of negative edges), that is, it ignores the edges outside the subgraph.

One suggested workaround for both of these concerns is to use more eigenvectors of the Laplacian of the whole graph to cluster in one step. This gives rise to a new method for clustering.

The **generalized Fiedler clustering method** is as follows:
(1) Compute k largest eigenvectors of the Laplacian of the graph ($k \leq \log_2(n)$, where n is the number of vertices of the graph).
(2) Put the vertices with the same sign patterns in the same cluster.

Two families of graph were generated; first, graphs with 4 clusters on 78 vertices of comparable sizes 20, 15, 25, and 18; second, graphs with 4 clusters of sizes 100, 12, 8, and 15. The density of edges inside each cluster is 90%, and we let the density of the edges between clusters to vary from 0% to 99%, with increments of 1%. For example, and example when the inter density is a 10% is shown in Figure 1.
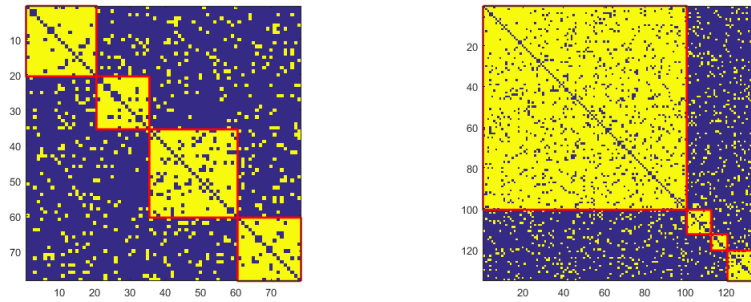
FIGURE 1. The adjacency matrices of two graphs with four clusters (Left:) of comparable sizes, and (Right:) of non-comparable sizes. Both graphs have 90% intra density and 10% inter density.

2. METHODS

For each density of inter edges, we generated 1000 random graphs, and applied two methods of clustering on them. The adjusted Rand index (ARI) against our known clustering, and the Girvan-Newman modularity (GNM) for each clustering was computed and the average of them is recorded.

3. Results

3.1. **Clusters with almost equal sizes.** Since the meaning of cluster fades away as the density of inter edges increases, we have included figures of a range of densities for better intuition in Figure 2.
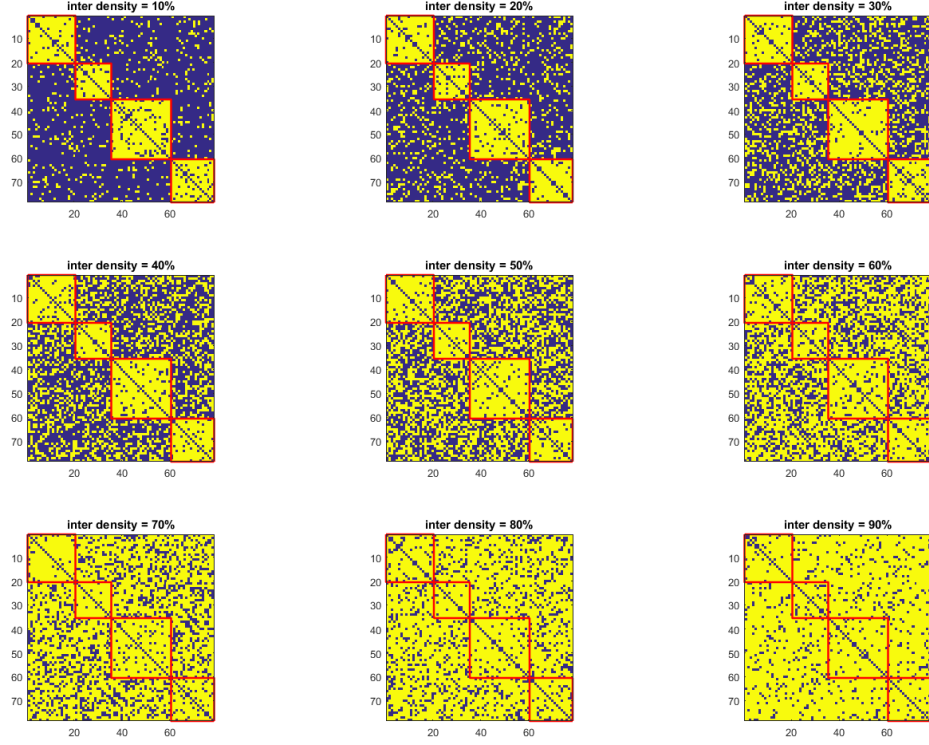

FIGURE 2. Sample graphs with four clusters of comparable sizes.
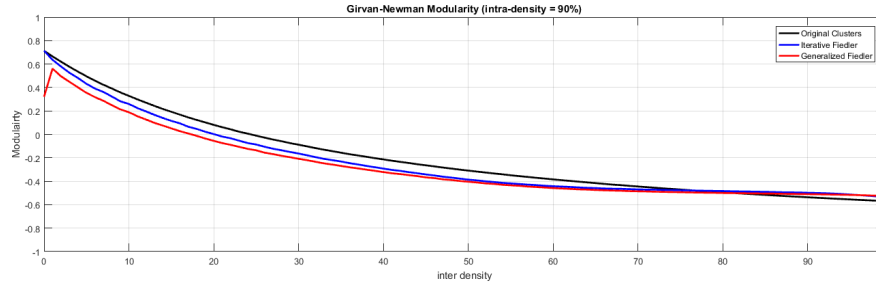
And here are the results:

FIGURE 3. Averages of Girvan-Newman modularity of graphs with four clusters of comparable sizes for each inter density for 1000 runs each.
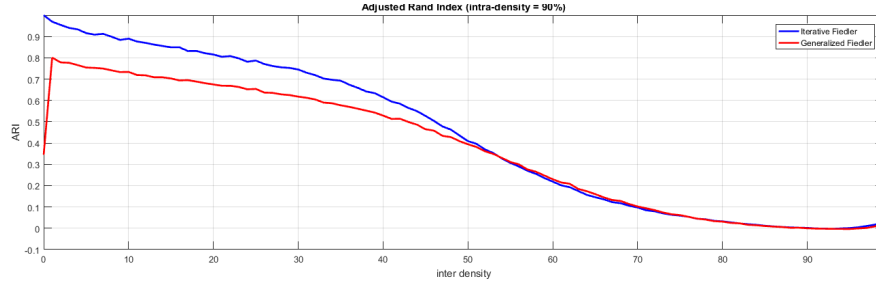
FIGURE 4. Averages of adjusted Rand index of graphs with four clusters of comparable sizes for each inter density for 1000 runs each.

3.2. **Clusters with sizes far away from each other.** When the sizes of clusters are different from each other however, the methods do not behave in the same way. In particular, the modularity cannot reach the theoretical maximum and overall stays low.
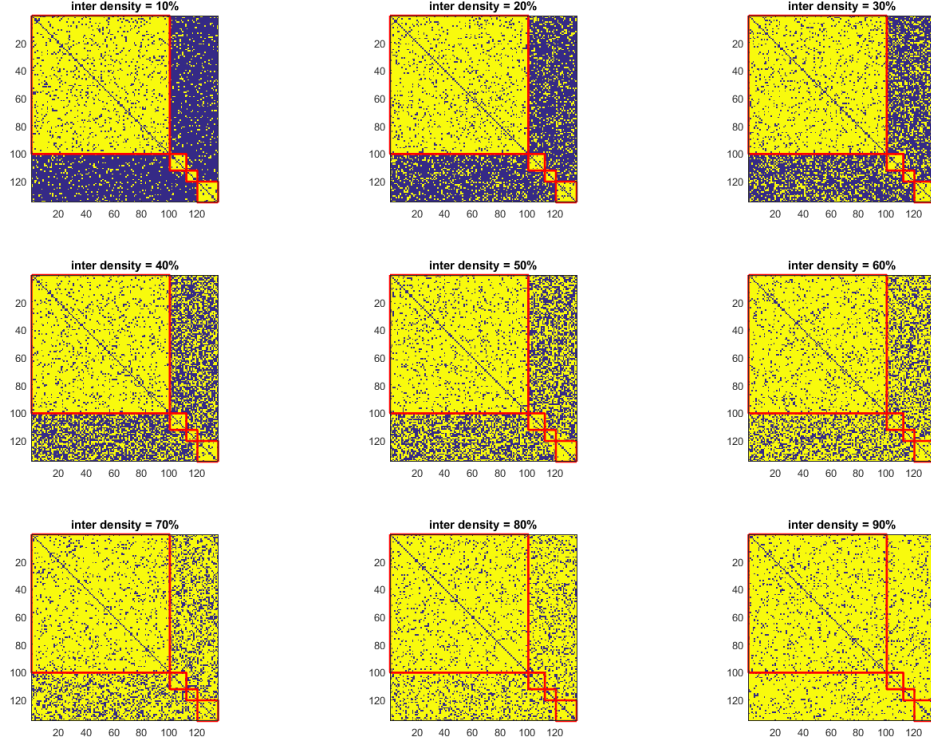

FIGURE 5. Sample graphs with four clusters of far away sizes.
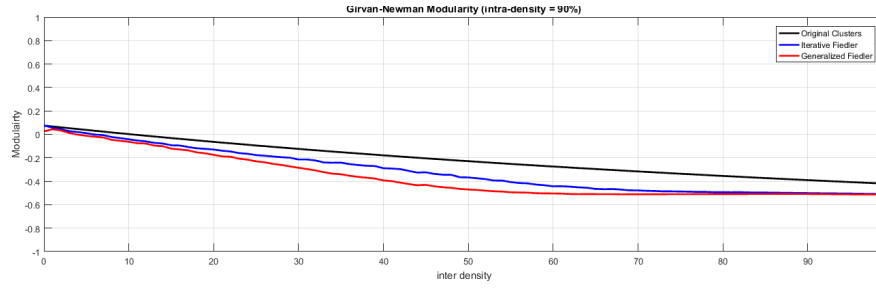
And here are the results:

FIGURE 6. Averages of Girvan-Newman modularity of graphs with four clusters of far away sizes for each inter density for 1000 runs each.
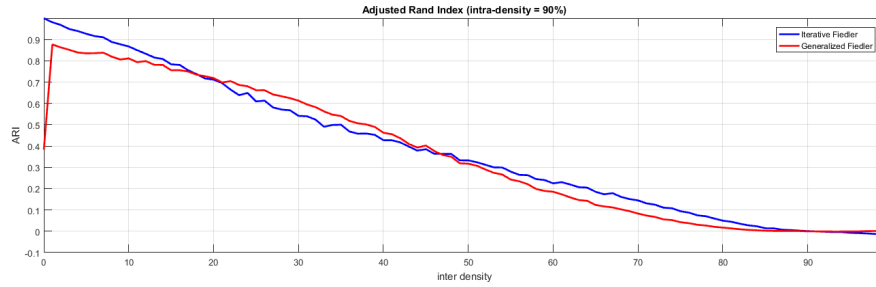

FIGURE 7. Averages of adjusted Rand index of graphs with four clusters of far away sizes for each inter density for 1000 runs each.

REFERENCES

[1] M. Fiedler. "Algebraic connectivity of Graphs", Czechoslovak Mathematical Journal 23(98) (1973), 298–305.

[2] Keivan Hassani Monfared, Kris Vasudevan, Jordan S. Farrell, G. Campbell Teskey, "Community structure detection and evaluation during the pre- and post-ictal hippocampal depth recordings", arXiv:1804.01568 [cs.SI].

K1MONFARED@GMAIL.COM