Assignment 6

community detection

Deepti Gupta

M21CS005

In This assignment We need to do community detection by using three algorithms on the LFR benchmark graph.

Three algorithms are as follows:

- Griven Newman algorithm
- Label-propagation algorithm (LPA)
- label propagation algorithm based on node importance (NI-LPA)

LFR benchmark graph is constructed→

parameters: For this experiment, n is 1001 (1000 < n < 2000) nodes are considered for comparing the community detection algorithm. Minimum 20 communities are considered.

Graph information:

- Undirected Graph: YES
- Number of Nodes:1002
- Number of Edges: 2239
- Average Degree:7.15

LFR benchmark graph: The greater the radius has higher the degree of the respective node.

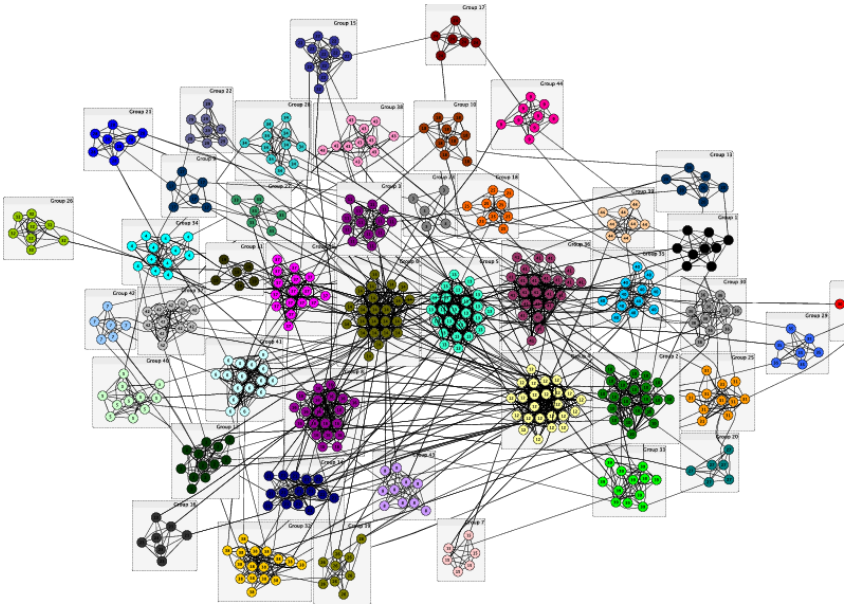

Fig: LFR benchmark graph

I.Community detection using Griven Newman algorithm: In This experiment compares the community with default and community by Girven Newman algorithm using Normalized mutual information.

NMI (Y, C) = 2 * I (Y ; C) / [H(Y) + H(C)] where Y= class labels, C = cluster labels, H(.) = Entropy, I(Y;C) = Mutual Information b/w Y and C.

- **Result:** NMI of Grivman Newman algorithm 0.5216719531242248

2. Community detection using Label Propagation algorithm: The Label Propagation algorithm is based on the label updating process on each t iteration. It gives the labels to unlabeled data using subsets of data.

- **Result**: NMI of label propagation algorithm0.8788076431861692

3. Community detection using Label Propagation algorithm using Node Importance:

NI-LPA is better than the LPA model. The performance of the NI-LPA model is significantly improved in terms of the modularity of the community.

Steps:

- Initialize all the level
- Calculating jaccard distance of every pair of nodes.
- Calculating k shell index value of each node.
- Calculating node importance score.
- Based on the node importance, update label of the current node.

- **Result:** NMI of Label propagation algorithm based on node importance 0.8863848614268359

Conclusion

The Griven Newman algorithm gives better results for a minimum number of communities in the graph. and For more community, node importance in graph Label Propagation algorithm based on the importance of nodes is better than traditional Label propagation algorithm.

References

Wang, T., Chen, S., Wang, X., & Wang, J. (2020). Label propagation algorithm based on node importance. *Physica A: Statistical Mechanics and its Applications, 551,* 124137.