Agents for Clustering Techniques

K Sri Vijaya
IT Department
PVPSIT
Vijayawada, India

Amar Jukuntla
CSE Department
VFSTR University
Guntur, India

Kolluru Pavan Kumar
CSE Department
VFSTR University
Guntur, India

ABSTRACT – **A multi agent system framework is proposed for clustering mechanism and each single agent is treated as the single cluster. In this framework first initial clusters has to be formed and agents will be able to start improvement of the initial clusters with the help of the negotiation. This framework is used for finding the connected components in the network with the help of the spectral clustering algorithm. In this proposed method improves the clustering mechanism in the spectral clustering algorithm for finding the number of connected components in the network.**

1. Introduction

Clustering is an important mechanism for grouping of similar data types and many clustering algorithms are used in various applications among those clustering algorithms spectral clustering algorithm is most important algorithm and became famous in recent years. This spectral clustering algorithm is simple and efficient[1]. The performance of the algorithm is depending on the matrix.

Adjacency matrix is calculated with the help of the adjacency nodes in the network. The use of this matrix is to find the number of clusters in the network. The generation of matrix is critical for spectral cluster success.

There are many methods are proposed for improving the efficiency of the matrix and these methods are categories into two ways : supervised and unsupervised. In this supervised data labels are known and based on the known information clusters are formed, based on this class labels matrix is constructed. In unsupervised data labels are unknown, this proposed work belongs to this category[2-3].

Though many techniques have mentioned for obtaining the good matrix without considering the relationship between spectral clustering and random walk. In these studies they revealed how to optimize number of cuts in the cluster.

Fig 1. Clusters Formation

In this approach, we used multi agents for finding the number of connected components in the network and improving the efficiency of the clusters in the network. These multi agents are useful for formation of the efficient clusters in the networks with optimized cuts in the network[4-5].

2. Related work

ordering and recovery is then required to adequately utilize these
In this section we give brief overview of clustering algorithms and focus on number of cuts in the graph and its relation to spectral cluster.

The main goal of this clustering is to form different clusters based on the data points similarity. Based on the data points in the graph clusters are constructed, vertices are represented with data entry points and similarities between the data points are represented. In this paper A is represents the matrix and G represents for graph[7]. The spectral cluster algorithm as follows,

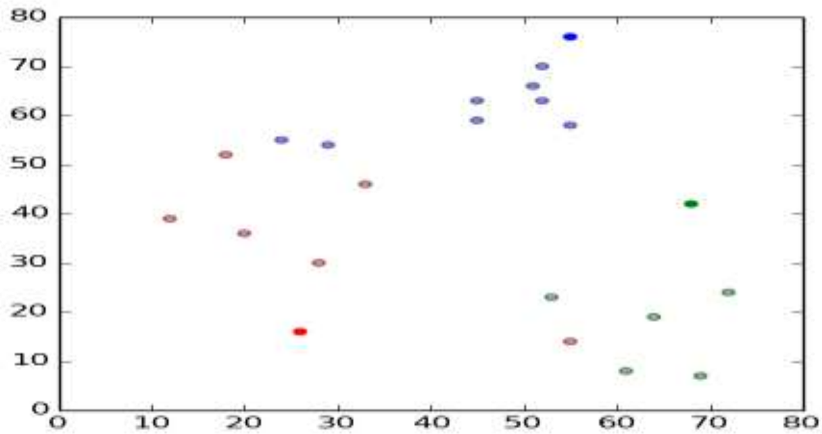

Fig 2. Initial Clusters formation

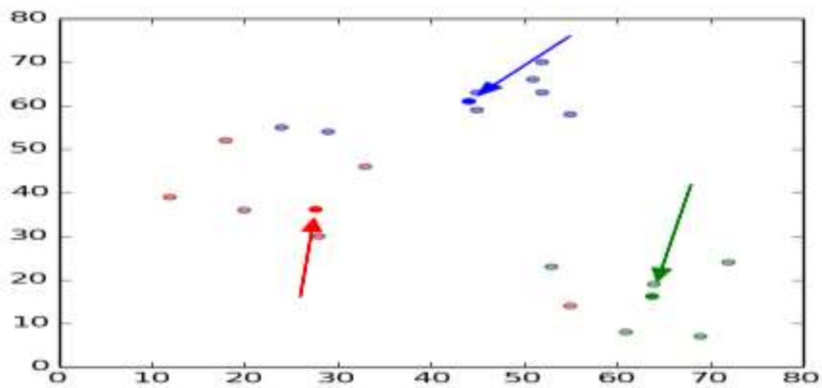

Fig 3. Centroids of clusters

Algorithm 1

Input: A set of data points, Set={s1 ,s2 ,s3……. Sn}, n=number of data point entries;
Output: Clusters
Step1: Compute similarity of the data points entry in matrix A.
Step 2: Compute Laplacian Matrix L= D-A
Step 3: Compute the first eigenvectors of Laplacian Matrix
Step 4: Compute the Eigenvalues of the matrix
Step 5: Form a cluster using k-means
Step 6: Clusters are formed using adjacency matrix.

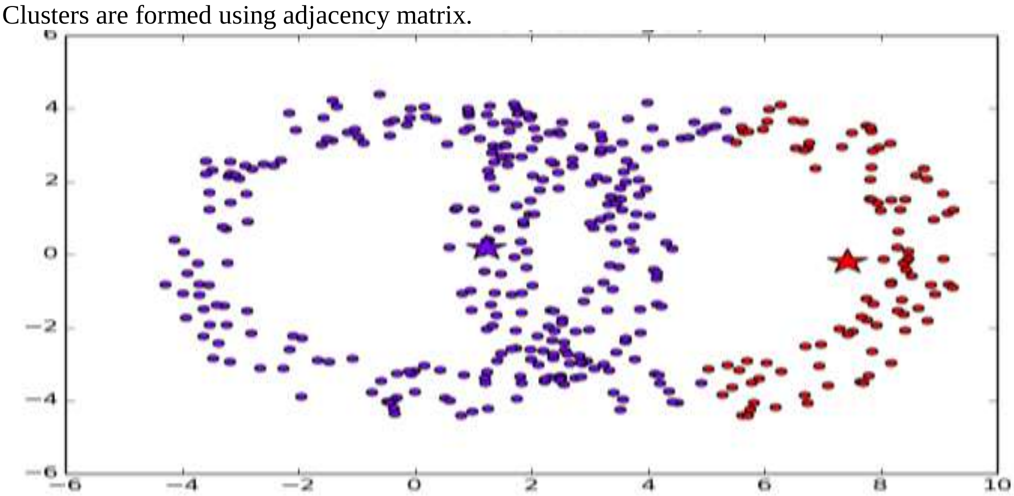

Fig 4. Spectral Cluster formation

Number of cuts will be identified, after formation of clusters. For formation of clusters k-means algorithm is used in the spectral cluster. Based on the number of clusters agents are identified. In initial stage of cluster formation agent will be generated and start communicating with other clusters based on the number of cut in the cluster[6,9]. These cuts will help for finding the number of connected components in the network. Agent communication and agent formation in the clusters will be discussed in the next section.

3. Proposed Framework

Proposed framework is presented in the following flowchart. In this, first traditional spectral clustering algorithm is used for formation of the clusters and agents are used for finding the connected components in the network through agent communication.

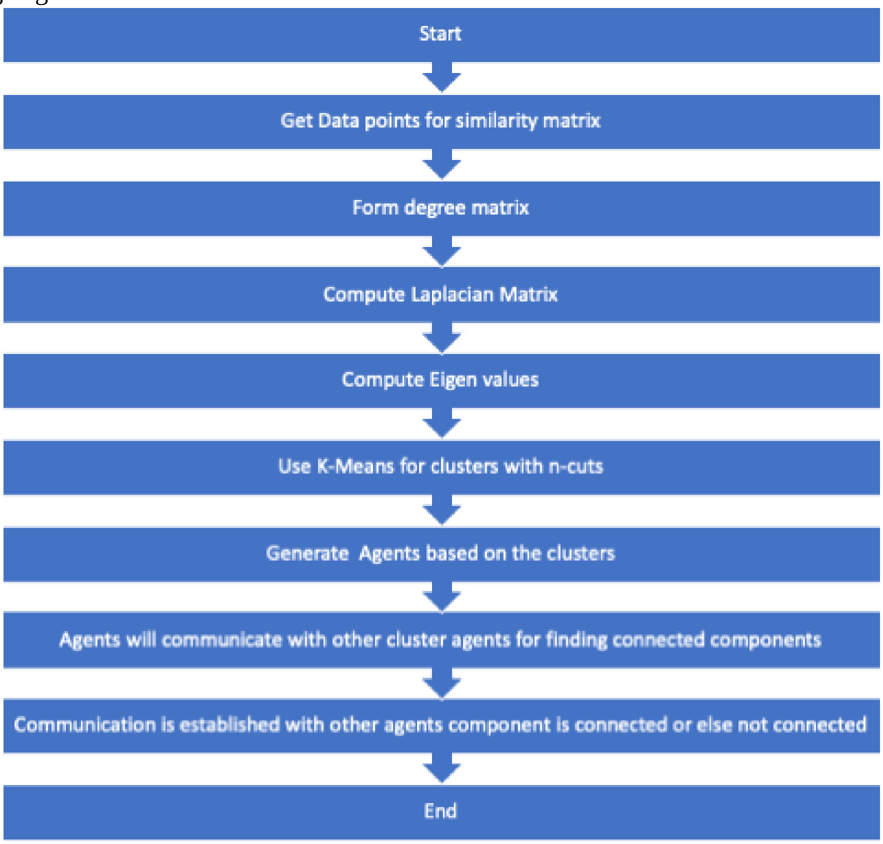

Fig 5. Proposed framework

After formation of initial clusters Agents are generated and start communicating with other agents which are present in the another cluster. These agents will communicate with connected components in the network. These cuts will helpful for finding the communication between the agents. Once communication is established between the clusters. Agents will start communicating and searches for connected components in the network. These will take care of the remaining cluster formation with other agents communications.

Algorithm 2

Step1: Compute similarity of the data points entry in matrix A.
Step 2: Compute Laplacian Matrix L= D-A
Step 3: Compute the first eigenvectors of Laplacian Matrix
Step 4: Compute the Eigenvalues of the matrix
Step 5: After computation of eigen values and eigen vectors, cuts in the network are identified.
Step 6: After identification of cuts in the network generate agents in the clusters.
Step 7: After agents identification is done, agents will start communicating with other agents which are present in the another group.
Step 8: Agents will start communicating with other
Form a cluster using k-means
Step 9: Once communication is done agents will start identification of connected components in the network.
Step 10: Stop the procedure if agents identified same set of connected components or else repeat the procedure.

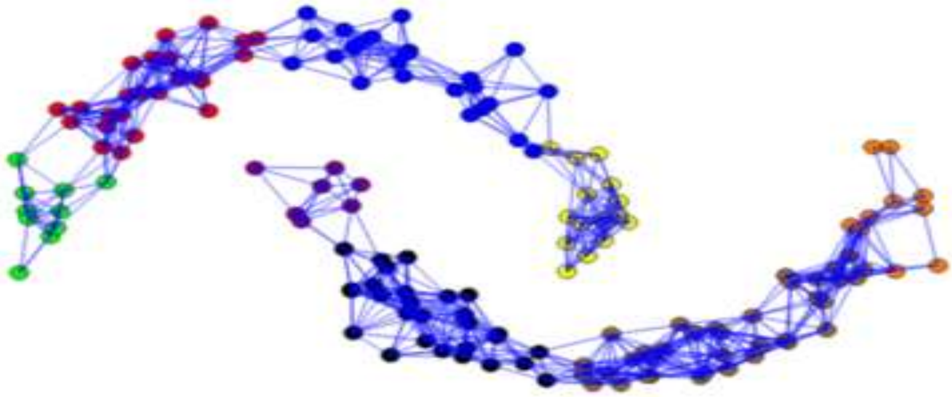

Fig 6. Clusters Formation without connected components

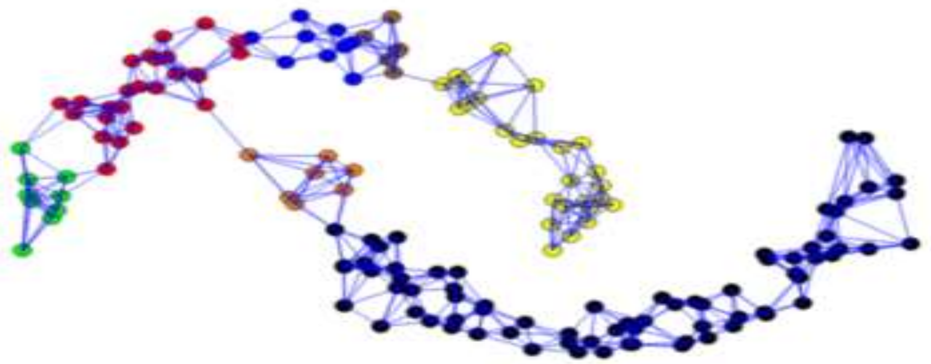

Fig 7. Clusters Formation with connected components

Spectral Clustering algorithm uses graph with weights to divide the network into clusters based on the sizes. Where G is an undirected with vertex set V={v1,…,vn} and edge at E={(vi,vj):vi,vj∈ V}. For each vi we define the degree of vi∈ V as

Distance Di=$\sum_{j=1}^{n}$ wij; Where wij is weight of the vertices.

Computation of Laplacian matrix L= D-W ;

Cut between clusters=$\sum_{i=C1,j=c2}$ Wij

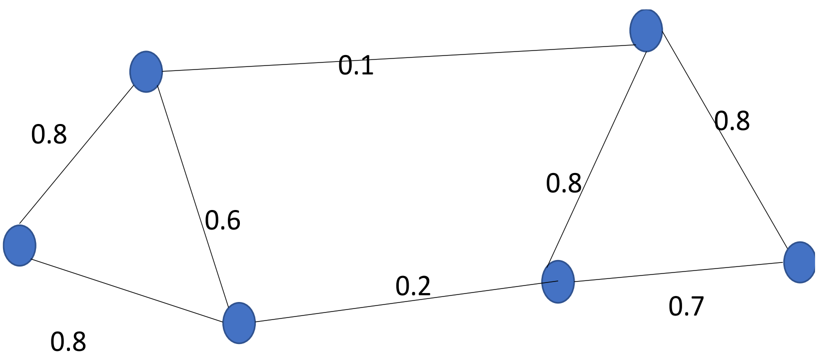

Fig 8. Example for connected graph

$$W=\begin{bmatrix} 0 & 0.8 & 0.6 & 0 & 0.1 & 0 \\ 0.8 & 0 & 0 & 0 & 0 & 0 \\ 0.6 & 0.8 & 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0.8 & 0.7 \\ 0.1 & 0 & 0 & 0.8 & 0 & 0.8 \\ 0 & 0 & 0 & 0.7 & 0.8 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} 1.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.5 \end{bmatrix}$$

$$L = \begin{bmatrix} 1.5 & -0.8 & -0.6 & 0 & -0.1 & 0 \\ -0.8 & 1.6 & -0.8 & 0 & 0 & 0 \\ -0.6 & -0.8 & 1.6 & -0.2 & 0 & 0 \\ 0 & 0 & -0.2 & 1.7 & -0.8 & -0.7 \\ -0.1 & 0 & 0 & -0.8 & 1.7 & -0.8 \\ 0 & 0 & 0 & -0.7 & -0.8 & 1.5 \end{bmatrix}$$

$$\text{Eigenvectors of L} = \begin{bmatrix} 0.0 \\ 0.4 \\ 2.2 \\ 2.3 \\ 2.5 \\ 3.0 \end{bmatrix}$$

Clusters ={0.2,0.2,0.2}, {-0.4,-0.7,-0.7}

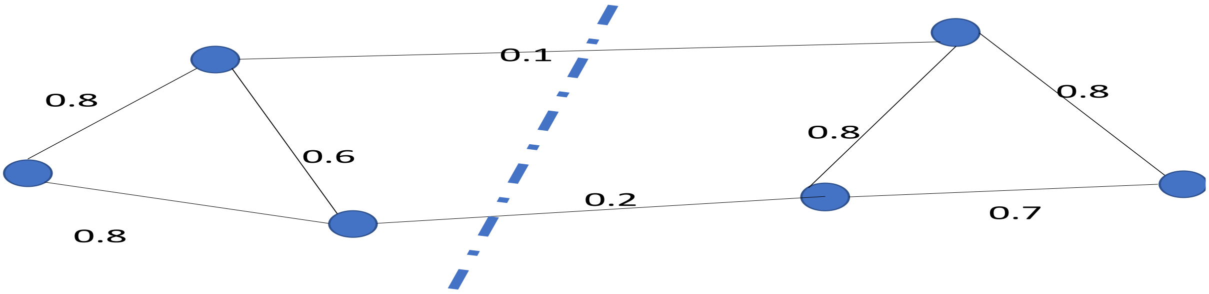

Fig 9. Example for partition graph

$$\text{Final normalized Laplacian matrix}= \begin{bmatrix} 1.0 & -0.52 & -0.39 & 0 & -0.06 & 0 \\ -0.52 & 1.0 & -0.50 & 0 & 0 & 0 \\ -0.39 & -0.50 & 1.0 & -0.12 & 0 & 0 \\ 0 & 0 & -0.12 & 1.0 & -0.47 & -0.44 \\ -0.1 & 0 & 0 & -0.8 & 1.7 & -0.50 \\ 0 & 0 & 0 & -0.44 & -0.50 & 1.0 \end{bmatrix}$$

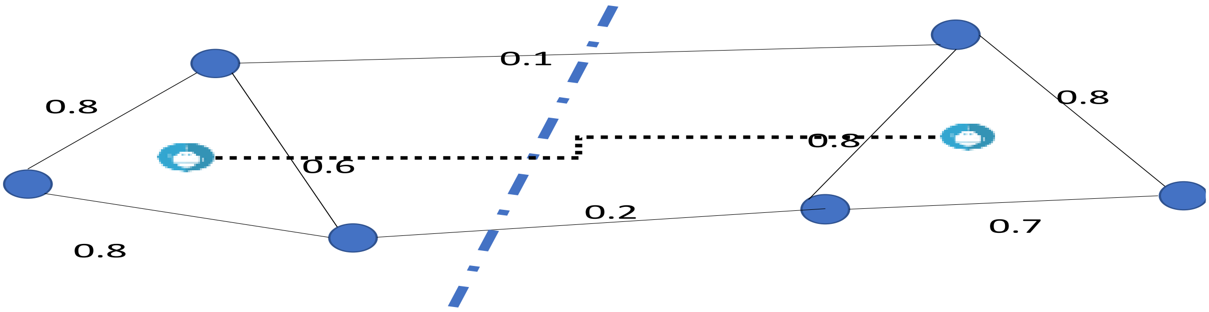

Fig 10. Connected graph with agents

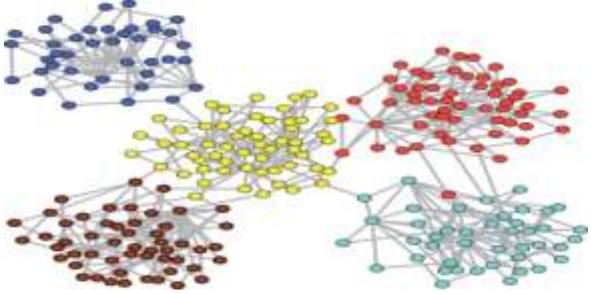

Fig 11. Clusters with n number of clusters

4. Conclusions

In this paper, we explained clustering techniques in unsupervised learning and developed a solution with spectral clustering by using agents. These clustering algorithm gives effective solution with the help of the agents. In this algorithm we adopted spectral clustering mechanism with agents. These agents will communicate with other agents which are available in the another cluster. These agents will communicate and establish a communication and find the connected components in the graphs and it gives the efficient solution in clustering mechanism. This algorithm can be improved with efficient machine learning techniques and connected components can be optimized with efficient solution.

5. References

1. D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial Bee Colony (ABC) algorithm", Applied Soft Computing, vol. 11, no. 1, (2011), pp. 652-657.
2. W. Y. Chen, Y. Song, H. Bai, C. J. Lin and E. Y. Chang, "Parallel spectral clustering in distributed systems", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 33 no. 3, (2011), pp. 568-586.
3. Y. Kim, K. Shim, M. S. Kim and S. Lee, "DBCURE-MR: an efficient density-based clustering algorithm for large data using MapReduce", Information Systems, vol. 42, (2014), pp. 15-35.
4. D. Marghescu, P. Sarlin and S. Liu, "Early-warning analysis for currency crises in emerging markets: A revisit with fuzzy clustering", Intelligent Systems in Accounting, Finance and Management, vol. 17, no. 3-4, (2010), pp. 143-165.
5. U. Kaymak and M. Setnes, "Extended fuzzy clustering algorithms", ERIM Report Series Reference No. ERS-2001-51-LIS, (2000).
6. T. Niknam and B. Amiri, "An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis", Applied Soft Computing, vol. 10, no. 1, (2010), pp. 183-197.
7. C. Alpert, A. Kahng, and S. Yao. Spectral partitioning: The more eigenvectors, the better. Discrete Applied Math, 90:3- 26, 1999.
8. N. Christianini, J. Shawe-Taylor, and J. Kandola. Spectral kernel methods for clustering. In Neural Information Processing Systems 14, 2002. [3] F. Chung. Spectral Graph Theory. Number 92 in CBMS Regional Conference Series in Mathematics. American Mathematical Society, 1997.
9. R. Kannan, S. Vempala, and A. Yetta. On clusterings- good, bad and spectral. In Proceedings of the 41st Annual Symposium on Foundations of Computer Science, 2000.
10. J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. In Perceptual Organization for Artificial Vision Systems. Kluwer, 2000.