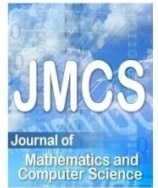Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com

A Graph Based Approach for Clustering Ensemble of Fuzzy Partitions

Mohammad Ahmadzadeh
Mazandaran University of Science and Technology
m.ahmadzadeh@ustmb.ac.ir

Zahra Azartash Golestan
Mazandaran University of Science and Technology
z.azartash@ustmb.ac.ir

Javad Vahidi
Iran University of Science and Technology
j.vahidi@iust.ac.ir

Babak Shirazi
Mazandaran University of Science and Technology
shirazi@ustmb.ac.ir

Abstract

Fuzzy clustering and Cluster Ensemble are important subjects in data mining. In recent years, fuzzy clustering algorithms have been growing rapidly, but fuzzy Clustering ensemble techniques have not grown much and most of them have been created by converting them to a fuzzy version of Consensus Function. In this paper, a fuzzy cluster ensemble method based on graph is introduced. Proposed approach uses membership matrixes obtained from multiple fuzzy partitions resulted by various fuzzy methods, and then creates fuzzy co-association matrixes for each partition which their entries present degree of correlation between related data points. Finally all of these matrixes summarize in another matrix called strength matrix and the final result is specified by an iterative decreasing process until one gets the desired number of clusters. Also a few data sets and some UCI datasets data set are used for evaluation of proposed methods. The proposed approach shows this could be more effective than base clustering algorithms same of FCM, K-means and spectral method and in comparison with various cluster ensemble methods, the proposed methods consist of results that are more reliable and less error rates than other methods.

1. Introduction

Data mining is used for knowledge discovery tasks; there are many methods for data mining tasks, such as association rule mining, classification and clustering are more important ones. It is this kind of classification when human recognizes two different things (as fruit) and separates them from each

other. Hence, classification is a supervised approach while clustering is same of classification with little difference in its approach, clustering is an unsupervised approach [1].

In recent years ensemble subjects are of interested so that Researchers agree on the composition of a number of grouped results leading to better quality in the final results versus results of one grouping method, and finally it causes cluster ensemble has become as a hot topic in data mining, cluster ensemble try to Consensus different partitions with most diversity from different clustering methods. Hence the weakness of any clustering methods is covered and high quality results with minimum distance from correct labeled are generated [2]. Of course in recent years, one new step is created to select better quality partitions for combination process, in order to gain high quality final result [3,4]. On the other hand, fuzzy clustering techniques has also grown rapidly because in crisp clustering each object belongs to one and only one cluster but also in fuzzy clustering each object can belong to more than one cluster that is measured by membership function , so accuracy in fuzzy clustering is more than in usual status [5].

Many algorithms are presented in fuzzy clustering until now that, all of them try to obtain better and more accurate results.
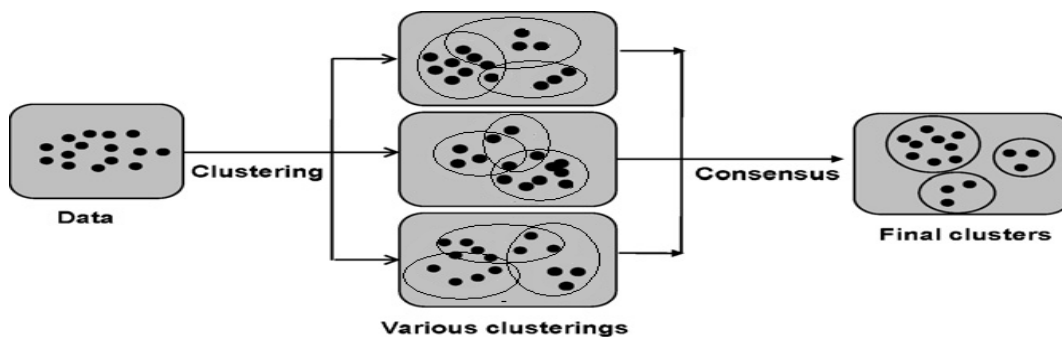

Figure1. Fuzzy clustering ensemble process

In case of fuzzy clustering ensemble (as shown in figure 1), fewer methods have been proposed and most of them are proposed by changing the main ensemble algorithms. In this paper, an approach is proposed to combine fuzzy partitions, which have the lowest error rate in the actual label, and is competitive with other clustering ensemble methods.

The rest of this paper is organized as follows: in section 2 different clustering techniques are reviewed. Section 3 proposes fuzzy clustering ensemble approach and its implementation on some datasets are described, in section 4 experimental results using some UCI's datasets are shown, and finally conclusions are in the last section.

2. Related work

In this section, base clustering methods and clustering ensemble methods are described in two parts:

2.1. Base clustering methods

There are many base clustering methods whether crisp or fuzzy clustering algorithms that are divided to four groups generally as hierarchical, density based, grid based and partitional. Among these, partitional and hierarchical clustering are the most popular.

In partitioning method with determined number of clusters, an initial partition is created. Then a repetitive process that attempts to improve the partitioning is used. In this way, objects are transmitted from one group to other groups. General criteria for good partitioning is that, objects are grouped in a

cluster be near or be linked together and in contrast objects in different clusters are away from each other or be as different as possible [6][7].

A hierarchical clustering method works by grouping data objects into a hierarchy or "tree" of clusters [8]. They merge smaller cluster into larger ones or split larger clusters into smaller ones. Partitioning Clustering methods is generally faster. But it requires the analyst to decide. Instances of these decisions are used to determining about the number of clusters or to select the minimum distance to be two elements in a cluster. Of course, there are many methods to determine optimal cluster number. Few methods are expressed in [9-12]. Hierarchical methods allow user to examine a wide range of solutions, obtained from a defined measure of similarity, in an efficient manner. However, hierarchical methods have a few shortcomings. For example, done allocation in one step, is not changeable in the later stages that can lead to bad and irreversible decisions. Thus, base clustering methods have advantages and limitations so that hybrid approaches are needed to cover their limitations and improve their accuracy.

2.2. Clustering ensembles

Since most base clustering methods emphasis on specific aspects of the data as a result, on the specific dataset are very efficient. Therefore, we require methods that these algorithms can use to combine the strengths of each, to produce efficient results.

Usually in the first step of cluster ensemble, several basic clustering is created that each of them emphasize on specific features of data. Using various clustering algorithms is the first and easiest way to create different and diverse results spread of a data set. Base clustering algorithms are most commonly used in cluster ensemble, including hierarchical clustering algorithm and partitioning clustering algorithms are classified. After the initial results (possible distributed) were produced, these results are combined usually by using a Consensus Function.

Many clustering ensemble methods have been proposed in recent years. The consensus problem has been faced by using several mathematical and computational tools. Methods based on Relabeling and Voting, Co-association Matrix, Graph and Hyper graph partitioning, Mirkin Distance, Information Theory, Finite Mixture Models, Genetic Algorithms, Locally Adaptive Clustering Algorithms (LAC), Kernel methods, Non -Negative Matrix Factorization (NMF) and Fuzzy techniques can be found. The two most important approaches are co-association methods [13] and hyper-graph methods [14]. Strehl and Ghosh proposed three efficient heuristics to partition the hyper-graphs in [15], including Cluster-based Similarity Partitioning Algorithm (CSPA), Hyper Graph Partitioning Algorithm (HPGA), and Meta-Clustering Algorithm (MCLA). Fred and Jain [13] adopted co-association approach and developed an evidence accumulation clustering (EAC) method based on single-link (SL) and average-link (AL) hierarchical clustering.

In recent years although fuzzy clustering algorithms have been growing rapidly, but fuzzy Clustering ensemble techniques have not grown much, most of them have been created by converting them to a fuzzy Consensus Function. Some of these methods, called sMCLA, sCSPA, sHBGF, which are the fuzzy version of MCLA, CSPA, HBGF [16].

This paper proposes a new approach to combine fuzzy partition with high performance.

3. Proposed method

The proposed clustering method is an unsupervised approach to combine multiple fuzzy clustering results obtained by various methods and algorithms to achieve reliable clustering results.

Implementation process of proposed approach contains four phases. In sequence advantages of this method and how it is expressed.

1. A dataset is clustered by multiple fuzzy clustering algorithms such as FCM, KFCM and KFAMC [17].
2. Fuzzy co- association matrix FC is created based on membership matrix.
3. All of association matrixes resulted from last step aggregate together and are put in a matrix by name of strength matrix.
4. Final result is obtained by graph based iterative decreasing process [18].

In sequence, each of them is explained in detail:

Phase1:

In the first stage, several fuzzy clustering methods such as FCM, KFCM , KFACM and etc are applied on the data sets and finally generate primary fuzzy partitions where the results of each algorithms are created with membership degree for each data point into different clusters and matrixes by name membership matrix are constructed through these membership degree for data points.

So we'll have a number of membership matrixes equal with number of applied algorithms that it will be represented by the letter U. In fuzzy clustering each data points can belong to multiple clusters while in crisp clustering each data point belong to a cluster or not.

Phase 2:

In this phase, fuzzy co-association matrix is created based on membership matrix in the first phase.

Fuzzy co-association matrix is the fuzzy version of co-association matrix. Co-association matrix is applied in crisp methods and is an n*n matrix with each of its entries equal to 0 or 1. Each member of this matrix is identified by its i and j entry such that it expresses that i and j data points are in one cluster or not thus if two data points be in one cluster then their related entry equals with 1 and otherwise equals with 0.

The co-association matrix is a two-dimensional n*n matrix too, that for short is called FC. But fuzzy co-associations are calculated in the other way. For computing how much two data points are in a cluster, it needs a co-association matrix that be fit with fuzzy status.

$$0 \leq FC_{ij} \leq 1 \tag{1}$$

As was mentioned, the FC matrix is calculated based on membership matrix and corresponding to with each membership matrix, there is a FC matrix. Each entry of this matrix, for example fcij, expresses that i and j data points have how much unity and how much belong to one cluster in other words, they have how much similarity , of course similarity means being in one cluster and being close to each other that we call it amount of correlation.

For achieving amount of two data points correlation into one cluster, at first is intersected their membership degree into that cluster. The result of intersection is the minimum amount of their membership degree in that cluster and then this action is done for all of the clusters, finally these values are summed together and make up entry of FC matrix. Any how this amount be close to 0, it means that related data points are with less Possibility in one cluster in final result, and any how this amount be close to 1 , it means that related data points are more possible in one cluster in final result.

The below formula expresses, how fuzzy co-association matrixes are computed based on membership matrix that obtained in the phase ago:

$$\begin{cases} FC_{mn} = \sum_{i=1}^{C} u_{im} \cap u_{in} & m \neq n \\ FC_{mn} = 0 & m = n \end{cases} \tag{2}$$

If m=n, $FC_{mn} = 0$ and otherwise FC_{mn} is equal with union of intersection u_{im}, u_{in} where m, n are data points, in other words adjacent matrix A shows relationship of data points. So whatever the amount of adjacent matrix FC elements has much more values show x_i and x_j have more similarity and stability, and if FC_{mn} has low value, related data points have weak relation and less similarity and stability.

Phase 3:

In this phase of proposed approach all of association matrixes results from last step aggregate together and are putted in a matrix by name of strength matrix(S). Which edges with more weight indicate that their nodes have more similarity and stability status in most of the clustering methods. Because it expresses that in most of clustering, that nodes are in one cluster and so sum of the fuzzy co-association matrixes cause to increase related entry value, and it is computed by the formula as follows:

$$S = \sum_{i=1}^{P} FC_i \tag{3}$$

P is the number of all used methods and FCi is fuzzy co-association matrix.

Phase 4:

In this phase, final result is obtained by graph based iterative decreasing process. This result consists the most stable clusters between all of the partitions resulted of various algorithms.

The iterative decreasing process is implemented on strength matrix obtained from the phase ago. The process reduces entry values of strength matrix, amounts to 0.5 unit until achieving to desired cluster number. Reason of decreasing in each iteration is that , loose joints or in the other words, the weaker clusters get 0 value thus are broken and create stable clusters with remaining strength joints. Another reason of reducing 0.5 units is that, this value has better result than reducing 1 unit. If the reduced value is 1 unit then entries with 0.3 and 0.8 have no difference together while decreasing 0.5 unit makes 0 entry for 0.3 value but entry with 0.8 can remain its connection in next step.We can reduce this value to 0.2 but algorithm will take more running time. This step is related to the following formula:

$$S_{ij}^{new} = \begin{cases} S_{ij}^{previous} - 0.5 & S_{ij}^{previous} > 0 \\ 0 & otherwise \end{cases} \tag{4}$$

For better understanding, a numerical example is shown as follows:

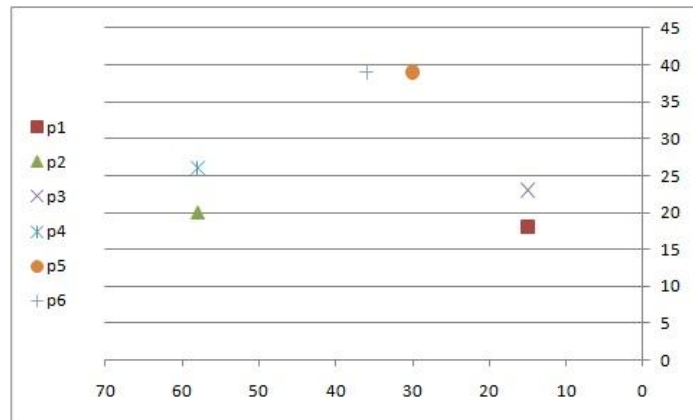$$X = \begin{bmatrix} 15 & 58 & 15 & 58 & 30 & 36 \\ 18 & 20 & 23 & 26 & 39 & 39 \end{bmatrix}$$

Figure 2. An artificial dataset for proposed approach implementation

At first a dataset X is clustered with specified cluster number by three fuzzy clustering algorithms as FCM, KFCM and KFACM [17], and then the adjacency matrix is created based on membership matrix.

$$U1 = \begin{bmatrix} 0.9869 & 0.0042 & 0.9233 & 0.0069 & 0.1789 & 0.0018 \\ 0.0090 & 0.0084 & 0.0563 & 0.0186 & 0.6249 & 0.9962 \\ 0.0041 & 0.9874 & 0.0203 & 0.9745 & 0.1962 & 0.0020 \end{bmatrix}$$

$$U2 = \begin{bmatrix} 0.0090 & 0.0084 & 0.0563 & 0.0186 & 0.1789 & 0.9962 \\ 0.0041 & 0.9874 & 0.0203 & 0.9745 & 0.1962 & 0.0020 \\ 0.9869 & 0.0042 & 0.9233 & 0.0069 & 0.6249 & 0.0018 \end{bmatrix}$$

$$U3 = \begin{bmatrix} 0.0405 & 0.0212 & 0.0059 & 0.0047 & 0.9932 & 0.4100 \\ 0.9499 & 0.0124 & 0.9921 & 0.0033 & 0.0046 & 0.3803 \\ 0.0096 & 0.9663 & 0.0020 & 0.9920 & 0.0022 & 0.2097 \end{bmatrix}$$

In which the degree of similarity or in other words amount in common in a cluster is calculated by the below formula.

$$\begin{cases} FC_{mn} = \displaystyle\sum_{i=1}^{C} u_{im} \cap u_{in} & m \neq n \\ FC_{mn} = 0 & m = n \end{cases}$$

Suppose the data set is partitioned into three clusters (C=3) and the membership matrix is resulted in this phase. In phase 2, these fuzzy partitions must be combined together to obtain the final result. At first, adjacent matrixes are organized by membership matrixes as follows:

$$FC1 = \begin{bmatrix} 0 & 0.0167 & 0.9364 & 0.0200 & 0.1920 & 0.0128 \\ 0.0167 & 0 & 0.0329 & 0.9871 & 0.2088 & 0.0122 \\ 0.9364 & 0.0329 & 0 & 0.0458 & 0.2555 & 0.0601 \\ 0.0200 & 0.9871 & 0.0458 & 0 & 0.2217 & 0.0224 \\ 0.1920 & 0.2088 & 0.2555 & 0.2217 & 0 & 0.6287 \\ 0.0128 & 0.0122 & 0.0601 & 0.0224 & 0.6287 & 0 \end{bmatrix}$$

$$FC2 = \begin{bmatrix} 0 & 0.0167 & 0.9364 & 0.0200 & 0.6380 & 0.0128 \\ 0.0167 & 0 & 0.0329 & 0.9871 & 0.2088 & 0.0122 \\ 0.9364 & 0.0329 & 0 & 0.0458 & 0.7015 & 0.0601 \\ 0.0200 & 0.9871 & 0.0458 & 0 & 0.2217 & 0.0224 \\ 0.6380 & 0.2088 & 0.7015 & 0.2217 & 0 & 0.1827 \\ 0.0128 & 0.0122 & 0.0601 & 0.0224 & 0.1827 & 0 \end{bmatrix}$$

$$FC3 = \begin{bmatrix} 0 & 0.0432 & 0.9578 & 0.0176 & 0.0473 & 0.4304 \\ 0.0432 & 0 & 0.0203 & 0.9743 & 0.0280 & 0.2433 \\ 0.9578 & 0.0203 & 0 & 0.0100 & 0.0125 & 0.3882 \\ 0.0176 & 0.9743 & 0.0100 & 0 & 0.0102 & 0.2177 \\ 0.0473 & 0.0280 & 0.0125 & 0.0102 & 0 & 0.4168 \\ 0.4304 & 0.2433 & 0.3882 & 0.2177 & 0.4168 & 0 \end{bmatrix}$$

$$\mathbf{fc_{3,4}} = \sum_{i=1}^{C} u_{i3} \cap u_{i4} = u_{1,3} \cap u_{1,4} + u_{2,3} \cap u_{2,4} + u_{3,3} \cap u_{3,4} = 0.9233 \cap 0.0069 +$$

$$0.0563 \cap 0.0186 + 0.0203 \cap 0.9745 = 0.0069 + 0.0186 + 0.0203 = \mathbf{0.0458}$$

Then all of association matrixes resulted from last step aggregate together and are put in a matrix by name of strength matrix. The edges with more weight indicate that their nodes have more similarity and stability status in most of the clustering methods.

$$S = \sum_{i=1}^{P} FC_i = \begin{bmatrix} 0 & 0.0766 & 2.8306 & 0.0576 & 0.8773 & 0.4560 \\ 0.0766 & 0 & 0.0861 & 2.9485 & 0.4456 & 0.2677 \\ 2.8306 & 0.0861 & 0 & 0.1016 & 0.9695 & 0.5084 \\ 0.0576 & 2.9485 & 0.1016 & 0 & 0.4536 & 0.2625 \\ 0.8773 & 0.4456 & 0.9695 & 0.4536 & 0 & 1.2282 \\ 0.4560 & 0.2677 & 0.5084 & 0.2625 & 1.2282 & 0 \end{bmatrix}$$

$$S_{ij}^{new} = \begin{cases} S_{ij}^{previous} - 0.5 & S_{ij}^{previous} > 0 \\ 0 & otherwise \end{cases}$$

$$
\begin{bmatrix}
0 & 0.0766 & 2.8306 & 0.0576 & 0.8773 & 0.4560 \\
0.0766 & 0 & 0.0861 & 2.9485 & 0.4456 & 0.2677 \\
2.8306 & 0.0861 & 0 & 0.1016 & 0.9695 & 0.5084 \\
0.0576 & 2.9485 & 0.1016 & 0 & 0.4536 & 0.2625 \\
0.8773 & 0.4456 & 0.9695 & 0.4536 & 0 & 1.2282 \\
0.4560 & 0.2677 & 0.5084 & 0.2625 & 1.2282 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 2.3306 & 0 & 0.3773 & 0 \\
0 & 0 & 0 & 2.4485 & 0 & 0 \\
2.3306 & 0 & 0 & 0 & 0.4695 & 0.0084 \\
0 & 2.44 & 0 & 0 & 0 & 0 \\
0.3773 & 0 & 0.4695 & 0 & 0 & 0.7282 \\
0 & 0 & 0.0084 & 0 & 0.7282 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 1.8306 & 0 & 0 & 0 \\
0 & 0 & 0 & 1.9485 & 0 & 0 \\
1.8306 & 0 & 0 & 0 & 0 & 0 \\
0 & 1.9485 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.2282 \\
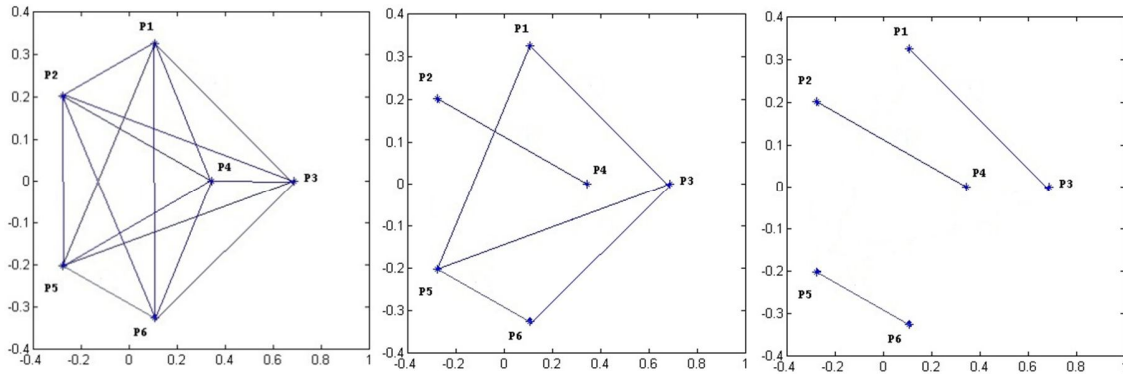0 & 0 & 0 & 0 & 0.2282 & 0
\end{bmatrix}
$$

Figure 3. Iterative decreasing process until achieving to desired cluster number

Then begin to decrease the strength matrix until you achieve the desired cluster number because loose joints or, in other words, the weaker clusters are broken and create stable clusters with remaining strength joints.

Algorithm description:

As shown in below algorithm, at first fuzzy co-association matrix is constructed based on membership matrix (line7) and then strength matrix is generated that related to phase 3 (line12). Finally strength matrixes are used as input for next stage. BFS_TRAVELSAL is breath first search on the matrix (line16) that is used to obtain cluster number by getSubGraphNumber() method (line23) and then in each stage strength matrix S is decreased 0.5 unit until you achieve the desired cluster number.

```
for m=1:6
for n=1:6
if m==n
fc(m,n)=0
else
for i=1:3
fc(m,n)=fc(m,n)+ min(U(i,m),U(i,n))
end
end
end
end
for i=1:3
S=S+FCi
Input: a strengh matrix S
N=0
Gnew[0]=BFS_TRAVELSAL();
S^previous = S
```

While(cluster-number > 3)
{
N=N+1;

$$S^{new} = S^{previous} - 0.5;$$

Gsubgraphs[n] = BFS_TRAVELSAL(Snew);
cluster-number = Gsubgraphs[n].getSubGraphNumber();
$S^{previous} = S$new;
}
Return Gsubgraphs [n];
Output: a set of sub- graphs- Gsubgraphs

4. Experimental results

In this section, our proposed approach is implemented on a few artificial datasets and widely used dataset as iris, which is available on the [19].

4.1. Clustering results

Results of proposed approach implementation and cluster number distribution until getting specific cluster number C are shown in fig 4,5, 6 and 7. Datasets are portioned by various fuzzy methods with specified cluster number and the strength matrix is calculated based on fuzzy co-association matrixes, and then iterative decreasing process is begun until you achieve the slightly cluster number.

For example, clustering results and cluster number distribution of Dataset1 and Datast2 until getting 5 and 3 clusters are shown in fig 5.
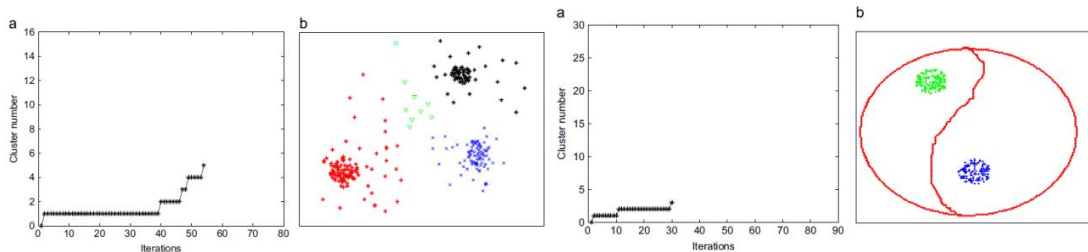

Figure 4. implementation of proposed approach on Dataset1 and Dataset2 with their cluster number distribution

4.2 Comparisons with base clustering algorithms

Clustering results of Datasets3 and Dataset4 on base algorithms as k-means and FCM compared with proposed methods in figure 5 and figure 6 show that direct implementation of k-means and FCM are sensitive to outliers and noisy data. The proposed approach can overcome this limitation.
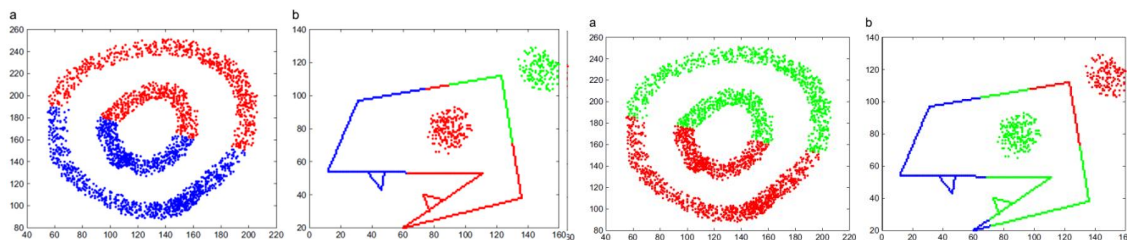

Figure 5. Implementation of FCM and K-Means algorithm on Dataset3 and Dataset4
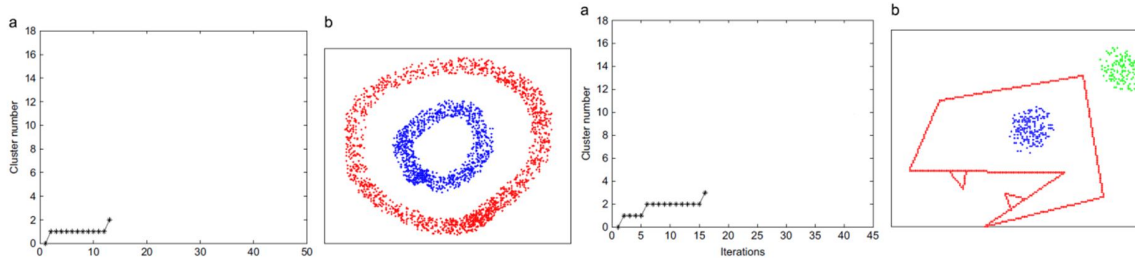
Figure 6. Implementation of proposed approach on Dataset3 and Dataset4

In comparison with the spectral method, same of proposed method, the spectral method is supported by graph theory. The difference between spectral methods and the proposed method is that spectral method uses graph to represent data relationship but the proposed approach uses graph to aggregate results of various clustering algorithms and then decreases them to cut weak connections and obtain more reliable and quality final clustering result.

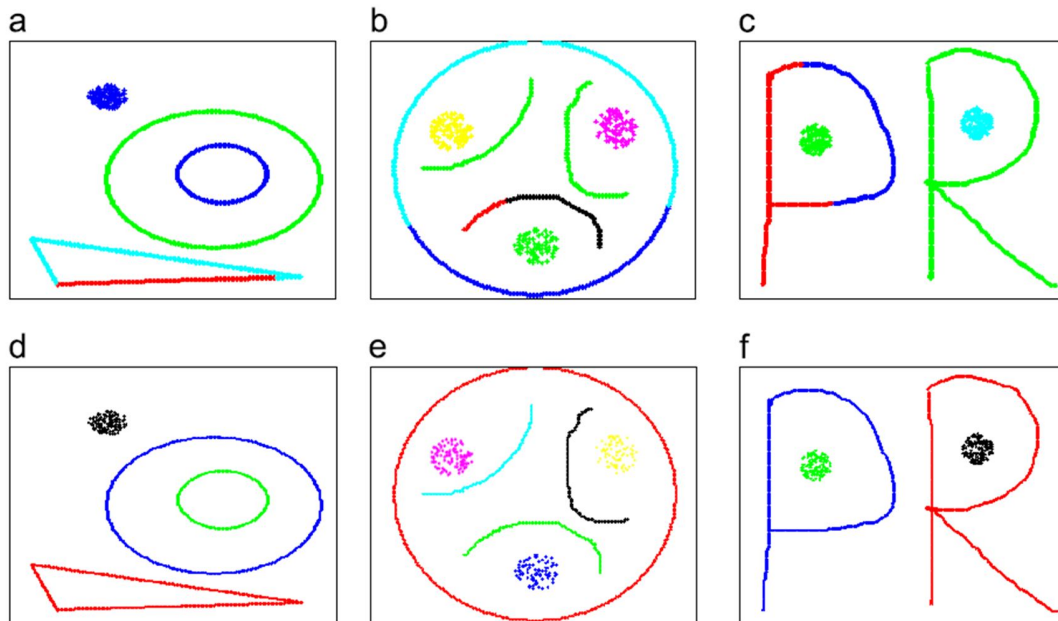In this section the comparison between spectral method and proposed method is shown in

Figure 7 Implementation of spectral methods in fig a, b and c[21] and implementation of proposed method in fig d , e and f

Figure 7 states that proposed method results better than spectral method, on the other hand, proposed method is adaptable with different clustering methods.

4.3 Comparisons with cluster ensemble methods

In this section, the proposed method is compared with clustering ensemble methods, since the proposed approach has similarity with clustering ensemble methods, thus it must be interesting when comparing these together.

The two most important approaches are co-association methods [13] and hyper-graph methods [14]. Strehl and Ghosh proposed three efficient heuristics to partition the hyper-graphs in [15], including

Cluster- based Similarity Partitioning Algorithm (CSPA), Hyper Graph Partitioning Algorithm (HPGA), and Meta-Clustering Algorithm (MCLA). Fred and Jain[13] adopted co-association approach and developed an evidence accumulation clustering (EAC) method based on single-link (SL) and average-link (AL) hierarchical clustering.

Some UCI datasets are used to compare proposed method with a number of said cluster ensemble methods in above.

Table 1 error rate of ensemble methods on some UCI datasets dataset

Method	WKF	GWKF	CSPA	HPGA	EAC-SL	EAC-AL	Our Proposal
Iris	10.6	10.8	13.3	37.3	11.1	11.1	9.8
Wine	28	30	31	48	37	28.7	28
Glass	51	51	57	60	59.8	49.1	47

As shown in table 1, the error rate of various methods is computed by comparing the results of methods with actual label of the data sets.

Therefore, as it is shown, the proposed method consists of lower error rates than others ensemble methods.

5. Conclusion

In this paper, a fuzzy cluster ensemble method based on graph is introduced. Proposed approach uses membership matrixes obtained from multiple fuzzy partitions resulted by various fuzzy methods, and then creates fuzzy co-association matrixes for each partition which their entries present degree of correlation between related data points.Finally, all of these matrixes are summarized in another matrix called strength matrix and the final result is specified by an iterative decreasing process.

 In addition, a few data sets and some UCI datasets data set are used for evaluation of proposed methods. The proposed approach is shown to be more effective than base clustering algorithms same of FCM, K-means and spectral method and in comparison with various cluster ensemble methods, the proposed methods consist of both reliable results and lower error rates.

6. References

[1] A.K. Jain, M.N. Murty, P.J. Flynn, Data clustering: a review, ACM Computing Surveys 31 **(1999)** 264–323.

[2] H. Alizadeh, Clustering Ensemble Based on a Subset of Primary Results, Master of Science Thesis, Iran University of Science and Technology **(2009)**.

[3] X. Z. Fern and W. Lin, Cluster ensemble selection, Stat. Anal. Data Min. 1(3) **(2008)** 128_141.

[4] Y. Hong, S. Kwong, H. Wang and Q. Ren, Resembling-based selective clustering ensembles, Pattern Recognition Letter 30(3) **(2009)** 298_305.

[5] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, **(1981)**.

[6] F. Hoppner, F. Klawonn, R. Kruse, T. Runkler, Fuzzy Cluster Analysis, Wiley Press, Chichester, **(1999)**.

[7] A.Fred and A.K.Jain, Learning Pairwise Similarity for Data Clustering, In Proc. Of the 18th Int. Conf. on Pattern Recognition (ICPR'06) , **(2006)**.

[8] J.Han, M.Kamber and J.Pei , Data Mining: Concepts and Techniques, Third Edition Elsevier Inc (2012).

[9] R.J.G.B. Campello, Generalized external indexes for comparing data partitions with overlapping Categories, Pattern Recognition Letters 31 **(2010)** 966–975.

[10] R.J.G.B. Campello, A fuzzy extension of the Rand index and other related indexes for clustering and Classification assessment, Pattern Recognition Letters 28 **(2007)** 833–841.

[11] F.Kovaˊ cs, C.Legaˊ ny, A.Babos, Cluster Validity Measurement Techniques, in: Proceedings of the fifth WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and DataBases, Madrid, Spain, **(2006)**, pp. 388–393.

[12] Das A.K. and Sil J., Cluster Validation using Splitting and Merging Technique, in proc. of Int. Conf. on Computational Intelligence and Multimedia Applications, ICCIMA**(2007).**

[13] A.L.N. Fred, A.K. Jain, Combining multiple clusterings using evidence accumulation, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 **(2005)** 835–850.

[14] A. Strehl, J. Ghosh, Cluster ensembles: a knowledge reuse framework for combining multiple partitions, Journal of Machine Learning Research 3 **(2002)** 583–617.

[15] S. Vega-Pons, J. Correa-Morris, J. Ruiz-Shulcloper, Weighted cluster ensemble using a kernel consensus function, in: J. Ruiz-Shulcloper, W.G. Kropatsch (Eds.), CIARP 2008, 5197, Lecture Notes in Computer Science, **(2008)**, pp. 195–202.

[16] S. Vega-Pons and J. Ruiz-Shulcloper, A survey of clustering ensemble algorithms, International Journal of Pattern Recognition and Artificial Intelligence Vol. 25, No. 3 **(2011)** 337_372.

[17] J.Liu, M.Xu, Kernelized fuzzy attribute C-means clustering algorithm, Fuzzy Sets and Systems 159 **(2008)** 2428 – 2445.

[18] P.Y. Mok, H.Q.Huang, Y.L.Kwok, J.S.Au A robust adaptive clustering analysis method for automatic Identification of clusters, Pattern Recognition 45 **(2012)** 3017–3033.

[19]<http://archive.ics.uci.edu>.

[20] M. Filippone, F. Camastra, F. Masulli, S. Rovetta, A survey of kernel and spectral methods for clustering, Pattern Recognition 41 **(2008)** 176–190.

[21] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) **(2000)** 888–905.