DOMINANT-SET-BASED CONSENSUS FOR FUZZY C-MEANS CLUSTERING ENSEMBLE

PAN SU[1],TIANHUA CHEN[2], WEIFENG XU[1], XUQIANG SHAO[1], HONGTAO WANG[1], YITIAN ZHAO[3*]

[1]School of Control and Computer Engineering, North China Electric Power University, Baoding, China
[2]Department of Computer Science, School of Computing and Engineering, University of Huddersfield, Huddersfield, UK
[3*]Corresponding Author: Cixi Institute of Biomedical Engineering, Ningbo Institute of Industrial Technology, Chinese Academy of Sciences, Ningbo, China
E-MAIL: supan@ncepu.edu.cn, t.chen@hud.ac.uk, weifengxu@163.com, shaoxuqiang@ncepu.edu.cn, wanght@ncepu.edu.cn, yitian.zhao@nimte.ac.cn

Abstract:

Conventional clustering approaches partition a set of objects into a certain (some can automatically detect the number of clusters such as DBScan) number of clusters. During the partitioning process, the clusters of objects are produced where each object is assigned to one cluster. On the other hand, the dominant-set-based clustering provides a formalisation of clusters by sequentially searching for individual clusters in the set of objects. The resultant clusters do not necessarily form a partition of the set. With the popularity of clustering ensemble, graph-based consensus approaches have been proposed with promising results achieved, many of which are based on the partition of the graph. In this paper, a dominant-set-based consensus method for fuzzy-c-means-based clustering ensemble is proposed. Different from traditional graph-based consensus techniques, the graph generated by the fuzzy clusters are grouped on the basis of the extracted dominant sets. The proposed approach employs a similarity relation to denote the links between component clusters from which the final clusters of ensemble are derived with the extracted dominant set. The proposed method is tested on benchmark data sets against several alternative ensemble methods for fuzzy c-means. The results of experiment show that the proposed dominant-set-based clustering ensemble method generally achieves higher accuracy than its competitors.

Keywords:

Dominant set; Clustering ensemble; Consensus function; Fuzzy c-means; Clustering

1. Introduction

Clustering is one of the popular methods which is able to extract hidden structures from unlabelled and labelled data sets. In general, the goal of clustering is to assign objects to clusters where objects in the same cluster are similar to each other, and dissimilar to those in different clusters [1]. In the literature, many clustering algorithms have been proposed and applied to solve a wide range of problems for real-world applications [2, 3, 4, 5]. A number of the existing methods separate the set of objects into clusters which form a partition of the data set [6]. However, the one-class clustering, which attempts to find an individual cluster by locating a hidden structure or pattern in the data, has attracted attention with successful applications such as detecting outliers [7].

Cluster ensembles have shown to outperform standard clustering algorithms in terms of accuracy and robustness across different data collections [8]. Similar to the classifier ensemble [9] and feature selection ensemble [10], cluster ensemble combines multiple base clustering results into a single consolidated clustering. The performance of cluster ensembles generally depends on both the quality and the diversity of ensemble components. Consequently, two essential steps are commonly employed in the implementation of clustering ensemble, i.e., a) base clustering members generation and b) consensus function.

Recently, several works have been published in the literature to address the issue of consensus. These include: the voting-based scheme [11], feature-based approaches which are based on label-assignment matrix [12]; pairwise similarity-based approaches which create a pairwise similarity matrix amongst data points based on the base clusters [13]; graph-based approaches which employ graph representation of base clusters (or base clustering members) [14].

Although several graph-based consensus have been proposed for the development of clustering ensemble, most of the graph-based consensus methods are based on partitioning the graph of base clusters. Interesting departures from the partition-based clustering have been reported, such as the dominant set clustering which provides a formalisation of each individual cluster by considering the clustering process as a sequential search of dominant sets. Following this trend, a consensus approach based on dominant sets for constructing ensembles of fuzzy c-means is proposed in this paper, where a fuzzy graph representing the similarities between fuzzy base clusters is employed and the final result is generated by extracting dominant sets from the fuzzy graph. The proposed methods are tested on benchmark data sets against counterparts which utilise partition-based graph refinement. The results of experimentation shows that the proposed dominant-set-based consensus method outperforms its counterparts for fuzzy c-means ensemble in terms of accuracy.

The paper is outlined as follows: Section II introduces the preliminaries of fuzzy clustering ensemble. Section III introduces dominant sets and presents its applications to consensus functions for creating ensembles of fuzzy clusters. Section IV describes the evaluation and discussion of the proposed method based on experimentation. Finally, Section V draws conclusion of the paper and makes suggestions for further works.

2. Preliminaries of Fuzzy Clustering Ensemble

Many methods [15, 16, 17] have been successfully developed in the framework of fuzzy set theory, among which, fuzzy c-means allows an object belonging to differen clusters to various degrees, overcoming boolean boundaries that are often not natural or even counterintuitive. Each cluster in a fuzzy partition $\widetilde{\pi}$ is a fuzzy set $\widetilde{C}_k, k = 1, \cdots, K$ where $\widetilde{C}_k(x_t) \in [0, 1]$ represents the degree of a data point $x_t \in X$ belonging to the corresponding fuzzy cluster. Usually, this degree is normalised with all the clusters in a partition to satisfy that $\sum_{k=1}^{K} \widetilde{C}_k(x_t) = 1$.

Formally, a fuzzy (or soft) cluster ensemble can be described as follows [18]. Let $X = \{x_1, \cdots, x_N\}$ be a set of N data points and $\Pi = \{\widetilde{\pi}_1, \cdots, \widetilde{\pi}_m, \cdots, \widetilde{\pi}_M\}$ be M fuzzy ensemble members. Each ensemble member returns a set of fuzzy clusters $\widetilde{\pi}_m = \{\widetilde{C}_1^m, \cdots, \widetilde{C}_k^m, \cdots, \widetilde{C}_{K_m}^m\}$, where K_m is the number of fuzzy clusters constructed by that member. The fuzzy clusters generated by all ensemble members together form a set of fuzzy base clusters for the ensemble: $\{\widetilde{C}_1, \cdots, \widetilde{C}_n\} = \bigcup_{m=1}^{M} \widetilde{\pi}_m$, where $n = \sum_{m=1}^{M} K_m$. For each $x_t \in X$ and each ensemble member $\widetilde{\pi}_m \in \Pi$, $\widetilde{C}_k^m(x_t) \in [0, 1]$ denotes the degree of which the data point x_t belongs to the fuzzy cluster $\widetilde{C}_k^m$. An example of the so-called instance-cluster matrix of a fuzzy cluster ensemble is shown in Table. 1. The task of a fuzzy cluster ensemble is: for a given dataset X, find a new partition π^* which summarises the information embedded in the whole cluster ensemble Π. Such a cluster ensemble technique does not specify whether the final clusterings should be crisp or fuzzy.

The two procedures: a) base clustering members generation and b) consensus function (indicated previously) are also employed in the implementation of fuzzy clustering ensemble approaches. First, diverse component clustering members are generated by using fuzzy clustering algorithms. Second, in order to generate the final clustering result, a consensus functions is employed based on the fuzzy base clusters generated by component clustering members. The implementation of the described fuzzy clustering ensemble is shown in Figure 1.

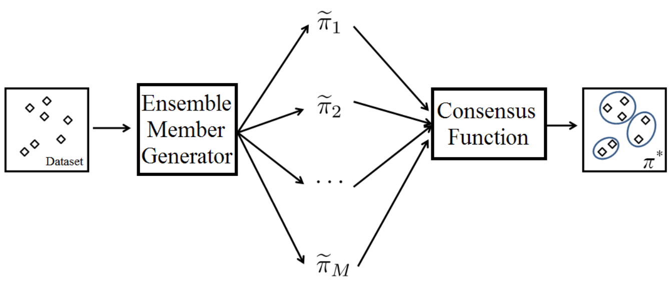

FIGURE 1. Fuzzy clustering ensemble

In the clustering ensemble, a consensus function can be represented by a map from a set of component ensemble members to forming one final result of the ensemble $f : \Pi \rightarrow \pi$. A number of consensus functions are based on the ensemble-information matrix which directly concludes the results of component clustering members. If a hard-boundary clustering algorithm (such as k-means) is utilised in the generation of base clusters, the membership of an object belonging to a cluster is either 1 or 0. Alternatively, in fuzzy clustering ensemble, the membership of an object belonging to a fuzzy cluster is in $[0, 1]$. An example of fuzzy ensemble-information matrix is illustrated in Tables 1.

Based on the fuzzy ensemble-information matrix, a graph whose nodes are fuzzy base clusters can be extracted. The edges amongst the nodes can be weighted by the similarities between fuzzy base clusters. After that, graph partition methods can then be employed to obtain a clustering ensemble output based on the graph.

TABLE 1. Example of fuzzy ensemble-information matrix

	$\widetilde{\pi}_1$		$\widetilde{\pi}_2$		$\widetilde{\pi}_3$	
	$\widetilde{C}_1^1$	$\widetilde{C}_2^1$	$\widetilde{C}_1^2$	$\widetilde{C}_2^2$	$\widetilde{C}_1^3$	$\widetilde{C}_2^3$
x_1	0.6	0.4	0.6	0.4	0.6	0.4
x_2	0.8	0.2	0.8	0.2	0.8	0.2
x_3	0.5	0.5	0.9	0.1	0.8	0.2
x_4	0.7	0.3	0.2	0.8	0.8	0.2
x_5	0.2	0.8	0.4	0.6	0.6	0.4
x_6	0.4	0.6	0.6	0.4	0.1	0.9
x_7	0.0	1.0	0.7	0.3	0.1	0.9

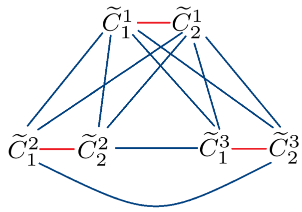

FIGURE 2. Graph generated from ensemble members of Table 1

3. Dominant-set-based Consensus for Fuzzy C-means Ensemble

In the graph-based clustering ensemble, the weights of edges are usually defined by the similarity amongst base clusters. Therefore, the base cluster members are usually generated from a dataset with all objects, so that the different base clusters in a cluster ensemble may share common objects. These shared objects create a linkage between a pair of base clusters and hence, it is possible to evaluate the similarity between them by measuring how much they are overlapped to each other. In the following of this paper, it is assumed that for a dataset with N objects $X = \{x_1, \cdots, x_N\}$, each fuzzy ensemble member $\widetilde{\pi}_m = \{\widetilde{C}_1^m, \cdots, \widetilde{C}_k^m, \cdots, \widetilde{C}_{K_m}^m\}$ satisfies $\sum_{k=1}^{K_m} \widetilde{C}_k^m(x_i) = 1$ for $i = 1, \cdots, N$.

Formally, let $\mathbb{C} = \bigcup_{m=1}^{M} \widetilde{\pi}_m = \{\widetilde{C}_1, \cdots, \widetilde{C}_n\}, n = \sum_{m=1}^{M} K_m$ be a set of fuzzy base clusters, a fuzzy graph $< \mathbb{C}, \widetilde{L} >$ is defined on them where $\widetilde{L}$ is a fuzzy set of edges. The membership function $\mathbb{C} \times \mathbb{C} \rightarrow [0, 1]$ of $\widetilde{L}$ is defined as:

$$\widetilde{L}(\widetilde{C}_i, \widetilde{C}_j) = \frac{\sum_{l=1}^{N} \min(\widetilde{C}_i(x_l), \widetilde{C}_j(x_l))}{\sum_{l=1}^{N} \max(\widetilde{C}_i(x_l), \widetilde{C}_j(x_l))}, \text{if } i \neq j \quad (1)$$

where $\widetilde{C}_i(x_t)$ indicates the the degree of a data point x_t belonging to a fuzzy cluster $\widetilde{C}_i$. For all $i = j$, $\widetilde{L}(\widetilde{C}_i, \widetilde{C}_j) = 0$, i.e., $< \mathbb{C}, \widetilde{L} >$ is a self-loop-free graph. It is worth noticing that $\widetilde{L}(\widetilde{C}_i, \widetilde{C}_j) = \widetilde{L}(\widetilde{C}_j, \widetilde{C}_i)$. The degree assigned to the link connecting fuzzy clusters $\widetilde{C}_i$ and $\widetilde{C}_j$ is thus defined in accordance with the proportion of their overlapping degree on all data points in X. For instance, the example illustrated in Table 1 can be represented as the graph shown in Fig. 2.

As it is shown in Fig. 2, an arbitrary pair of fuzzy base clusters have an edge linked to each other, and the membership value of a given edge represents the similarity between the corresponding two base clusters. In crisp clustering ensemble, base clusters within the same ensemble member usually do not have common objects with each other and the weights of those edges between the clusters within the same ensemble member are zero. Therefore, additional calculation may be desired to retrieve the similarities amongst clusters within an ensemble member. Take the connected-triple as an example, edges cross ensemble members are employed to estimate the similarities within ensemble members [19]. However, in fuzzy clustering ensemble, even the similarities amongst fuzzy base clusters within the same ensemble member (indicated as the red edges in Fig. 2) are possible to be of non-zero values. In other words, by using Eqn. (1), given that the number of ensemble members and number of base clusters in each ensemble member are fixed, the graph of a fuzzy clustering ensemble is more dense than that of a crisp clustering ensemble in terms of edges.

Traditional graph-based consensus methods are usually based on partitioning the set of base clusters $\mathbb{C}$, which implies that all the base clusters have to be assigned to one set in the partition. However, an ensemble member may generate a poor result and its corresponding base clusters are very dissimilar to other base clusters. In this case, it makes little sense to force all base clusters to belong to one set of the partition, which might result in poor ensemble results. Alternatively, the dominant set clustering provides a formalisation of coherent subgraph (named dominant set) individually and considering the clustering process as a sequential search of such coherent subgraphs.

The concept of dominant set arises from the study of graph theory, by which a continuous formulation of the maximum clique problem is defined. The nodes to be clustered are represented as an undirected graph with weighted edges $G = (V, E, \omega)$. The edge set $E \subseteq V \times V$ indicates all the possible connections. $\omega : E \rightarrow \mathbb{R}$ is the positive weight function. In the context of graph-based fuzzy clustering ensemble, $\mathbb{C} = \{\widetilde{C}_1, \cdots, \widetilde{C}_n\}$ is V and the membership function of $\widetilde{L}$ is ω. The symmetric matrix $A = (a_{ij})$ is used to represent the graph G with weighted adjacency matrix. This non-negative adjacency matrix is defined as: $a_{ij} = \widetilde{L}(\widetilde{C}_i, \widetilde{C}_j)$.

In general, the weights of edges within the dominant set of a edge-weighted graph should be large, representing high internal homogeneity or similarity. By contrast, the weights of edges linked to the ones from external dominant set will be small. The assignment of the edge-weights can be analysed based on the above perspectives. Let $\mathbb{S} \subseteq \mathbb{C}$ be a nonempty subset of nodes, $\widetilde{C}_i \in \mathbb{C}$ and $\widetilde{C}_j \in \mathbb{S}$. The relative similarity between $\widetilde{C}_i$ and $\widetilde{C}_j$ with respect to the average similarity between $\widetilde{C}_j$ and its neighbours in $\mathbb{S}$ can be defined as:

$$\phi_{\mathbb{S}}(i,j) = a_{ij} - \frac{1}{|\mathbb{S}|} \sum_{\widetilde{C}_k \in \mathbb{S}} a_{jk}. \tag{2}$$

It can be observed that $\phi_{\mathbb{S}}(i,j)$ can be either positive or negative. The weight of $\widetilde{C}_i$ with regard to $\mathbb{S}$ is assigned as:

$$W_{\mathbb{S}}(i) = \begin{cases} 1 & \text{if } |\mathbb{S}| = 1 \\ \sum\limits_{\widetilde{C}_j \in \mathbb{S}\setminus\{\widetilde{C}_i\}} \phi_{\mathbb{S}\setminus\{\widetilde{C}_i\}}(i,j) W_{\mathbb{S}\setminus\{\widetilde{C}_i\}}(j) & \text{otherwise.} \end{cases} \tag{3}$$

where $\mathbb{S} \setminus \{\widetilde{C}_i\}$ indicates the the nodes set $\mathbb{S}$ excluding the node $\widetilde{C}_i$, and $W_{\mathbb{S}}(i)$ demonstrates the similarity between node $\widetilde{C}_i$ and the nodes of $\mathbb{S} \setminus \{\widetilde{C}_i\}$ with respect to the mutual similarity amongst the nodes in $\mathbb{S} \setminus \{\widetilde{C}_i\}$. Finally, the total weight of $\mathbb{S}$ is calculated by $W(\mathbb{S}) = \sum_{\widetilde{C}_i \in \mathbb{S}} W_{\mathbb{S}}(i)$.

Definition 1 *[20] An non-empty subset of nodes $\mathbb{S}, \mathbb{S} \subseteq \mathbb{C}$ such that $W(S) > 0$ for any non-empty $S \subseteq \mathbb{S}$ is said to be a dominant set if:*

$$W_{\mathbb{S}}(i) > 0, \textit{for all } \widetilde{C}_i \in \mathbb{S}; \tag{4}$$

and

$$W_{\mathbb{S} \cup \{\widetilde{C}_i\}}(i) < 0, \textit{for all } \widetilde{C}_i \notin \mathbb{S}. \tag{5}$$

Dominant sets can be identified by local solutions of program:

$$\begin{array}{ll} \text{maximize} & f(\mathbf{z}) = \mathbf{z}^\top A \mathbf{z} \\ \text{subject to} & \mathbf{z} \in \Delta \end{array} \tag{6}$$

where

$$\Delta = \left\{ \mathbf{z} \in \mathbb{R}^{|\mathbb{C}|} : \sum_{i=1}^{|\mathbb{C}|} z_i = 1 \text{ and } z_i \geq 0 \text{ for all } i = 1, \cdots, |\mathbb{C}| \right\}.$$

A strict local solution $\mathbf{z}^*$ of Eqn. (6) indicates a dominant set $\mathbb{S}$ of G, where $z_i > 0$ means that the according node $\widetilde{C}_i \in \mathbb{S}$. An effective optimization approach for solving Eqn. (6) is given by the so-called *replicator dynamics*:

$$z_i^{(t+1)} = z_i^{(t)} \frac{(A\mathbf{z}^{(t)})_i}{\mathbf{z}^{(t)\top} A \mathbf{z}^{(t)}}, \tag{7}$$

where $i = 1, 2, \cdots, |\mathbb{S}|$. It has been proven that for any initialization of $\mathbf{z} \in \Delta$, its trajectory will remains in Δ with the increase of iteration t. With the increasing of t in Eqn. (7), the objective function $f(\mathbf{z})$ in Eqn. (6) is either strictly increasing or constant. In practice, the stopping criteria of the dynamic system can be set as a maximal number of iteration t or a minimal increment of $f(\mathbf{z})$.

For the solution of replicator dynamics, only one cluster can be detected for each dynamic system and different initialisations will result in different dominant sets. A peeling-off strategy has been proposed in [20], which iteratively extracts a dominant set $\mathbb{S}$ each time by using Eqn. (7) and repeats the process in the new set of nodes $\mathbb{C} = \mathbb{C} \setminus \mathbb{S}$. In the task of fuzzy clustering ensemble, the maximum number of peeling-off dominant sets is set to the number of output clusters in the final result π^*. Given there are D dominant sets extracted from $< \mathbb{C}, \widetilde{L} >$, the membership of an object $x_i, i = 1, \cdots, N$ belongs to the dominant set $\mathbb{S}_d, d = 1, \cdots, D$ is calculated as:

$$\mathbb{S}_d(x_i) = \frac{\sum_{\widetilde{C}_j \in \mathbb{S}_d} \widetilde{C}_j(x_i)/|\mathbb{S}_d|}{\sum_{l=1}^{D} \sum_{\widetilde{C}_j \in \mathbb{S}_l} \widetilde{C}_j(x_i)/|\mathbb{S}_l|}. \tag{8}$$

Different from that all the base clusters have to be assigned to one set in the partition, an base cluster which is very dissimilar to others can be excluded from all D dominant sets by using the proposed method. Therefore, the final result of clustering ensemble can be improved by ignoring those poor base clusters in the dominant-set-based consensus function.

4. Experimentation and Evaluation

In this experiment, the proposed method is tested on seven datasets downloaded from the UCI benchmark repository [21]. The true labels of instances in these datasets are known but are not explicitly used in the fuzzy clustering ensemble process. The performance of the proposed method is assessed in terms of accuracy as the ground truth of each dataset is known. The summary of these datasets is shown in Table 2.

TABLE 2. Summary of datasets used

Datasets	Instances	Attributes	Classes
Iris	150	4	3
Wine	178	13	3
Parkinsons	195	22	2
Glass (Identification)	214	9	6
Ecoli	336	7	8
Ionosphere	351	34	2
(Pima Indians) Diabetes	768	8	2

The fuzzy c-means clustering algorithm is used to generate the base clustering members. For the proposed algorithm, the number of ensemble members is set to the number of attributes in each dataset and the number of base clusters in each ensemble member is set to $K_m = 12$. Both the fuzzy c-means and replicator dynamics are initialised randomly in each run. The result of the proposed method is compared with four clustering ensemble approaches $FLink$, FCO, $FCTS$, and CTS (using a fixed number $K_m = \lceil\sqrt{N}\rceil$) [22]. The number of final clusters of ensemble (i.e., the number of dominant sets D) is set to the number of true classes on each dataset. The decay factor (DC) of CTS is set to 0.5 according to the suggestion of [19]. It is worth noticing that by setting the $K_m = 12$ for the proposed algorithm, the size of its generated graph is smaller than those of its counterparts, which facilities the extraction of dominant sets.

The result of accuracies shown in Table 3 is achieved by using a fixed number of base clusters in each ensemble member for dominant-set-based consensus method (DS) and its counterparts. The best result on each dataset is highlighted (in boldface) and each value in Table 3 is an average calculated from 50 random runs. The results show that the use of dominant-set-based consensus leads to five best accuracies out of the seven tested datasets, by building fuzzy c-means ensembles. This demonstrates the effectiveness of proposed approaches. However, the averaged accuracy of DS over the seven datasets is not better than that of $FLink$. This is mainly due to the poor performance of DS with $K_m = 12$ on the iris dataset.

TABLE 3. Comparison of accuracy

	$FLink$	FCO	$FCTS$	CTS	DS
Iris	86.36	**87.60**	80.97	71.35	66.93
Wine	94.51	91.58	94.45	80.75	**94.78**
Parkinsons	**81.92**	81.54	**81.92**	76.18	78.72
Glass	48.25	45.37	48.31	52.60	**59.25**
Ecoli	79.53	76.15	79.90	75.29	**80.54**
Ionosphere	64.10	64.10	64.10	64.10	**68.52**
Diabetes	66.64	66.87	66.63	65.82	**67.23**
Means	**74.4728**	73.3157	73.7547	69.4409	73.7085

Additionally, the performance of the proposed approach is evaluated with respect to the number of base clusters in each ensemble member. The number of fuzzy base clusters K_m in each ensemble member is set from 2 to 15 in each run (with an increment step of 1). Figure 3 shows the trend of accuracy (Y-axis) against the changing of fuzzy base cluster number in each ensemble member (X-axis). Each point in Fig. 3 is an average of values from 50 runs.

The performance of using dominant-set-based consensus is comparatively stable when the K_m (and consequently $|\mathbb{C}|$) is

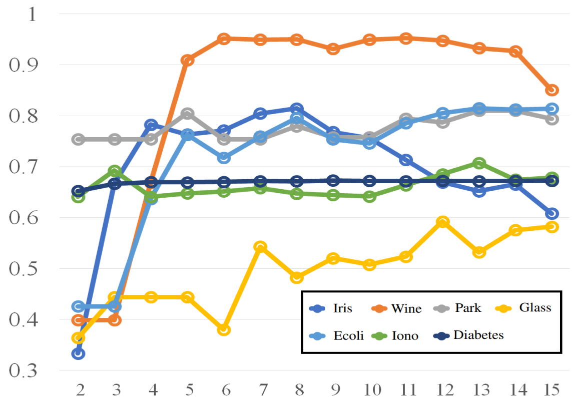

FIGURE 3. Trend of accuracy change against K_m

increased, as the performance generally does not decrease for six out of seven datasets. With the increase of K_m. However, the accuracies achieved on the iris dataset is dropped quickly when K_m is above eight. This shows that although extracting dominant sets one by one can help clustering ensemble to ignore those poor base clusters for building ensemble, it may lead to ignore useful base clusters if the value of K_m is set too high comparing with the number of dominant sets D.

5. Conclusions

This paper has introduced the application of dominant sets as a consensus function for fuzzy c-means clustering ensemble. The link between a pair of fuzzy base clusters is defined and a fuzzy graph is generated to represent the similarities amongst fuzzy base clusters. The proposed approach takes the advantage of dominant sets which are able to sequentially search stable structures in the graph of base clusters. Results of experiments over seven UCI datasets show that the proposed method generally outperforms the conventional fuzzy clustering ensemble methods.

The present work also enlightens ideas for further works. For example, the proposed method can be tested over more benchmark datasets to demonstrate its effectiveness and it would be interesting to investigate the performance of the proposed approach against the number of ensemble members as well as the diversity of base clustering algorithms.

Acknowledgements

This paper is supported by the Fundamental Research Funds for the Central Universities, China, under grant number: 2016MS118, the Natural Science Foundation of Hebei

Province under grant number: F2016502069, and the National Natural Science Foundation of China under grant number: 61502168.

References

[1] A. Jain, M. Murty, and P. Flynn, “Data clustering: a review,” *ACM Computing Surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[2] C.-L. Liu, W.-H. Hsaio, T.-H. Chang, and H.-H. Li, “Clustering data with partial background information,” *International Journal of Machine Learning and Cybernetics*, Feb 2018. [Online]. Available: https://doi.org/10.1007/s13042-018-0790-0

[3] P. Su, C. Shang, T. Chen, and Q. Shen, “Exploiting data reliability and fuzzy clustering for journal ranking,” *IEEE Transactions on Fuzzy Systems*, vol. 25, no. 5, pp. 1306–1319, 2017.

[4] Z. Li, C. Shang, and Q. Shen, “Fuzzy-clustering embedded regression for predicting student academic performance,” in *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*. IEEE, 2016, pp. 344–351.

[5] P. Su, C. Shang, and Q. Shen, “Link-based approach for bibliometric journal ranking,” *Soft Computing*, vol. 17, no. 12, pp. 2399–2410, 2013.

[6] S. R. Bulò and M. Pelillo, “Dominant-set clustering: A review,” *European Journal of Operational Research*, vol. 262, no. 1, pp. 1–13, 2017.

[7] G. Gupta and J. Ghosh, “Robust one-class clustering using hybrid global and local search,” in *Proceedings of the 22nd international conference on Machine learning*. ACM, 2005, pp. 273–280.

[8] T. Boongoen and N. Iam-On, “Cluster ensembles: A survey of approaches with recent extensions and applications,” *Computer Science Review*, vol. 28, pp. 1 – 25, 2018.

[9] R. Diao, F. Chao, T. Peng, N. Snooke, and Q. Shen, “Feature selection inspired classifier ensemble reduction,” *Cybernetics, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.

[10] Q. Shen, R. Diao, and P. Su, “Feature selection ensemble,” in *Proceedings of the Alan Turing centenary conference*, 2012.

[11] X. Sevillano, F. Alías, and J. C. Socoró, “Positional and confidence voting-based consensus functions for fuzzy cluster ensembles,” *Fuzzy Sets and Systems*, vol. 193, pp. 1–32, 2012.

[12] Z. Yu, L. Li, H.-S. Wong, J. You, G. Han, Y. Gao, and G. Yu, “Probabilistic cluster structure ensemble,” *Information Sciences*, vol. 267, pp. 16–34, 2014.

[13] D. Yan, A. Chen, and M. I. Jordan, “Cluster forests,” *Computational Statistics & Data Analysis*, vol. 66, pp. 178–192, 2013.

[14] N. Iam-On, T. Boongoen, S. Garrett, and C. Price, “A link-based approach to the cluster ensemble problem,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 12, pp. 2396–2409, 2011.

[15] P. Su, Q. Shen, T. Chen, and C. Shang, “Ordered weighted aggregation of fuzzy similarity relations and its application to detecting water treatment plant malfunction,” *Engineering Applications of Artificial Intelligence*, vol. 66, pp. 17–29, 2017.

[16] T. Chen, C. Shang, P. Su, and Q. Shen, “Induction of accurate and interpretable fuzzy rules from preliminary crisp representation,” *Knowledge-Based Systems*, vol. 146, pp. 152–166, 2018.

[17] T. Chen, Q. Shen, P. Su, and C. Shang, “Fuzzy rule weight modification with particle swarm optimisation,” *Soft Computing*, vol. 20, no. 8, pp. 2923–2937, 2016.

[18] K. Punera and J. Ghosh, “Consensus-based ensembles of soft clusterings,” *Applied Artificial Intelligence*, vol. 22, no. 7-8, pp. 780–810, 2008.

[19] N. Iam-On, T. Boongoen, and S. Garrett, “Refining pairwise similarity matrix for cluster ensemble problem with cluster relations,” in *Discovery Science*. Springer, 2008, pp. 222–233.

[20] M. Pavan and M. Pelillo, “Dominant sets and pairwise clustering,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 1, pp. 167–172, 2007.

[21] K. Bache and M. Lichman, “UCI machine learning repository,” 2013. [Online]. Available: http://archive.ics.uci.edu/ml

[22] P. Su, C. Shang, and Q. Shen, “Link-based pairwise similarity matrix approach for fuzzy c-means clustering ensemble,” in *Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1538–1544.