

Fuzzy Rough Clustering Analysis Algorithm Based on Attribute Reduction



Hao Ouyang, Zhi Wen Wang, Zhen Jin Huang, and Wei Ping Hu

Abstract Fuzzy clustering analysis algorithm has good ability to solve fuzzy problems. Rough clustering analysis has good ability to solve these problems its prior knowledge is uncertain. But in the real world, there are many problems that not only are fuzzy but also are rough and uncertain; the paper combines the idea of these two algorithms. In order to improve correction of clustering, it imports attributes reduction algorithm to get importance of each attribute, and dynamically changes attribute weight by the importance. The new algorithm firstly computes fuzzy membership degree of every object and then estimates the object that belongs to lower approximation or upper approximation of one cluster. In the analysis process, the paper provides a new way to get the cluster centers, combining fuzzy and rough theory. From experiments of four UCI data sets, it is proved that the new algorithm is better effective.

Keywords Clustering · Fuzzy theory · Rough set theory · Attribute reduction

1 Introduction

Clustering analysis is different from classifying analysis that is an unsupervised analysis process. Before the analysis, it does not know the characteristics of all attributes and the importance of all kinds of attributes for the clustering analysis process [1]. The paper [2] introduces the information entropy theory to analyze the attributes of data objects and obtains the importance of each attribute, so as to dynamically adjust the analysis process. The paper [3] introduces the idea of C4.5 algorithm, uses the information gain for the continuous attributes and category attributes, so as to help the category analysis. Although the above papers dynamically analyze the importance of each attribute, they fail to effectively combine the

H. Ouyang (✉) · Z. W. Wang · Z. J. Huang · W. P. Hu
School of Computer Science and Communication Engineering, Guangxi
University of Science and Technology, LiuZhou, Guangxi Province, PR China

fuzzy theory and the rough theory to deal with nonlinear problems. Fuzzy theory and rough set theory describe problems from the two different aspects: the fuzziness and the uncertainty. Therefore, a lot of papers combine fuzzy theory or rough set theory with clustering analysis algorithm to analyze problems. These papers [4–6] use fuzzy clustering algorithm to analyze medical image problem, social service problem, and carbonate fluid detection problem. These papers [7–9] use rough clustering algorithm to analyze fault detection in rotating machinery equipment, gene expression problems, and psychological anxiety problems of students. The paper [10] tries to combine the fuzzy theory and rough set theory and applies the algorithm to analyze hydrological engineering. In the process of combination, FCM algorithm is firstly adopted to obtain the mean points of fuzzy clustering, and then the rough set is calculated to obtain the upper and lower approximation of clusters. In the process, the two theories are not combined well, instead there are separated. In the paper [11], data object sets are preprocessed by using support vector machine, then introduces fuzzy clustering and rough set respectively, so the two steps are also independent. In the paper [12], fuzzy rough clustering is applied to the image analysis of brain MRI. In the process of analysis, the difference of membership degree is used to determine the rough boundary, but for calculating the mean point, only using the same method in the FCM algorithm. There are the same defect likely papers [10, 11].

In this paper, a new algorithm is provided. The reduction algorithm in rough set theory is used to obtain the importance of each attribute, dynamically adjust the weight values of each attribute in the iteration calculate process, so as to optimize the value of important attributes. In the clustering analysis, to combine the fuzzy and rough set theories, firstly it gets the fuzzy membership degree of each data object, and then gets fuzzy lower approximations and upper approximations of these clusters. The process effectively combines fuzzy and rough theory by calculating cluster mean points. This method more approximately describes the real world. The new algorithm presented in this paper is proved to have better analysis performance by several related experiments.

Chapter 2 introduces rough set theory and attribute reduction algorithm. This chapter introduces fuzzy clustering and rough clustering. In Chap. 4, fuzzy rough clustering based on reduction is introduced. In Chap. 5, a new algorithm is adopted to analyze multiple data sets. Chapter 6 is the conclusion and expectation.

2 Rough Set Theory and Reduction Algorithm

Rough set theory is a common method to analyze uncertain problems. In this theory, the most important concepts include lower approximation, upper approximation, bound and approximation quality.

Definition 1 for Low approximation, Upper approximation, and Bound. For an information system $IS = (U, A, V, F)$, $B \subseteq A$, there exists equivalent relation

R_B , $\langle U, R_B \rangle$ is defined as approximate space. For any set X , $X \subseteq U$, the low approximation of X for $\langle U, R_B \rangle$ is defined as:

$$\underline{R}_B X = \{x \in U \mid [x]_B \subseteq X\} \quad (1)$$

The upper approximation of X for $\langle U, R_B \rangle$ is defined as:

$$\overline{R}_B X = \{x \in U \mid [x]_B \cap X \neq \emptyset\} \quad (2)$$

The bound of X for $\langle U, R_B \rangle$ is defined as:

$$BN_B = \overline{R}_B X - \underline{R}_B X \quad (3)$$

In formula (1), U is a finite object set for analysis, A is an attribute set for objects. V is a domain set for A , $f : U \times A \rightarrow V$ is an information function. By definition, $\underline{R}_B X$ is looked as the set composed by objects that certainly belong to X , $\overline{R}_B X$ is looked as the set composed by objects that probably belong to X , BN_B is looked as the set composed by objects that probably belong or not belong to X .

Definition 2, Approximate quality. $r_B(X)$ is the rough degree of set X for B , it is defined as:

$$r_B(X) = \frac{|\underline{R}_B X|}{|X|} \quad (4)$$

In formula (4), $|\bullet|$ is the number of elements contained in the set. If $r_B(X) = 1$, it means the bound does not exist, $BN_B = \emptyset$. If $r_B(X) < 1$, it means that X is rough referring to B .

In the analysis of actual data, there is a lot of redundant information, and the importance of each attribute is different. In order to deal with problems effectively, the attribute reduction of rough set theory can be used to find the importance of attributes and eliminate the irrelevant or redundant attributes, so as to improve the correctness of clustering.

Definition 3, Attribute importance. For the decision table $DT = (U, C \cup D)$, if $B \subseteq C$ and $a \in C - B$, $\text{sig}_r(a, B, D)$ is the importance of attribute a referring to attribute set B , $\text{sig}_r(a, B, D)$ is defined as:

$$\text{sig}_r(a, B, D) = r_{B \cup \{a\}}(D) - r_B(D) \quad (5)$$

In formula (5), $r_B(D)$ is the dependent degree of decision attribute D referring to conditional attribute.

According to formula (5), $\text{sig}_r(a, B, D)$ can be seen as the analysis contribution of attribute a in attribute set B . If the figure of $\text{sig}_r(a, B, D)$ is bigger, the attribute a is more important for decision. According to this reason, we can get an attribute reduction algorithm. The reduction algorithm starts from an empty attribute set, and then iteratively selects the attribute with the maximum importance. In order to avoid the influence of noise, the algorithm imports a stop threshold ε .

The reduction algorithm is described as follows:

Input: decision table $DT = (U, C \cup D)$, stop threshold ε

Output: set B that is a relative reduction of set C for set D

Steps:

```

Calculate the reliance of decision attributes on conditional attributes,  $r_C(D)$ ;
 $B \leftarrow \phi$ ;
while  $B \subset C$  do
  begin
    for every  $a \in C - B$  do
      Calculate  $\text{sig}_r(a, B, D)$ ;
      Select the attribute  $a$  having biggest  $\text{sig}_r(a, B, D)$ , if many attributes
meet this condition, then select the attribute  $a$  that has least number of combinations
with attribute set  $B$ ;
       $B \leftarrow B \cup \{a\}$ ;
      if  $r_C(D) - r_B(D) \leq \varepsilon$ 
        break;
      end if
    end for
  for every  $a \in B$  do
    if  $r_C(D) - r_{B - \{a\}}(D) \leq \varepsilon$ 
       $B \leftarrow B - \{a\}$ ;
      End if
    output  $B$ ;
  end for
end while

```

3 Fuzzy Clustering and Rough Clustering

Fuzzy Clustering

Fuzzy c-mean clustering (FCM) is using fuzzy theory into clustering analysis process. In this algorithm, membership degree is used to indicate the possibility of object belonging to a certain object cluster. The membership degree u_{ik} is described as how much degree of data object x_k belonged to the No. i cluster, $u_{ik} \in [0, 1]$, $\sum_{i=1}^C u_{ik} = 1$.

The objective function of FCM algorithm is:

$$J_m(U, V) = \sum_{i=1}^C \sum_{k=1}^n (u_{ik})^m d_{ik}^2 \quad (6)$$

In formula (6), d_{ik} is described as the distance between the object x_k and the mean v_i , it is defined as follows:

$$d_{ik}^2 = \|x_k - v_i\|^2 \quad (7)$$

In formula (6), m is the fuzzy factor, $m \in [1, \infty]$. Usually, $m = 2$. Through many iterative calculations, the mean point is obtained by the following formula:

$$\forall \begin{matrix} 1 \leq i \leq c \\ 1 \leq k \leq n \end{matrix} u_{ik} = \begin{cases} (d_{ik})^{2/(1-m)} / \left[\sum_{j=1}^c (d_{jk})^{2/(1-m)} \right], & \text{if } v_i \neq x_k; \\ 1, & \text{if } v_i = x_k; \\ 0, & \text{if } v_l = x_k, l \neq i; \end{cases} \quad (8)$$

and

$$\forall_{1 \leq i \leq c} v_i = \sum_{k=1}^n (u_{ik})^m x_k / \sum_{k=1}^n (u_{ik})^m \quad (9)$$

Rough Clustering

Rough c-mean clustering (RCM) is using rough theory into clustering analysis process. In the algorithm, it is agreed that an object can only belong to the lower approximation of one cluster mostly. When calculating the distance between the data object and each mean point, if the difference between these distances is less than a certain threshold value, the object will be divided into the upper approximation of each corresponding cluster. Otherwise, it is divided into the lower approximation of nearest cluster.

According to the definition of lower approximation, upper approximation, and bound in formula (1)–(3), the new definition of mean point in RCM is obtained as follows:

$$\forall_{1 \leq i \leq c} v_i = \begin{cases} w_l \frac{\sum_{x_k \in \underline{C}_i} x_k}{|\underline{C}_i|} + w_{bn} \frac{\sum_{x_k \in C_i^{BN}} x_k}{|C_i^{BN}|}, & \text{if } \underline{C}_i \neq \phi \wedge C_i^{BN} \neq \phi; \\ \frac{\sum_{x_k \in \underline{C}_i} x_k}{|\underline{C}_i|}, & \text{if } \underline{C}_i \neq \phi \wedge C_i^{BN} = \phi; \\ \frac{\sum_{x_k \in C_i^{BN}} x_k}{|C_i^{BN}|}, & \text{if } \underline{C}_i = \phi \wedge C_i^{BN} \neq \phi; \end{cases} \quad (10)$$

In the above formula, w_l is the weight of lower approximation, w_{bn} is the weight of bound. Then $w_l + w_{bn} = 1$. If w_{bn} is larger, the prior knowledge is less complete, and the information set is rougher.

4 Fuzzy Rough Clustering Based on Attribute Reduction

Fuzzy Rough Clustering

In practical application scenarios, there are many problems often lacking prior knowledge that are fuzzy description. In this case, if only fuzzy clustering or rough clustering is used, it will inevitably lead to one-sidedness of the analysis process. In this paper, the fuzzy theory and rough set theory are applied to clustering analysis, and a new fuzzy clustering algorithm is designed. In the algorithm, two equivalent classes can be obtained for the set U : they are upper approximation and non-upper approximation; upper approximation can be divided continually into lower approximation and bound; this process is defined in formula (11). However, the definition of mean point is combined with the fuzzy theory and the rough theory. In the analysis, it is believed that the fuzzy degree of data objects belong to the lower approximation, boundary or non-upper approximation of each cluster is different, and the calculation method of fuzzy membership degree is shown in formula (12).

$$\begin{aligned} \bigcup_{1 \leq i \leq c} U &= \{x_k | x_k \in \overline{C}_i(X)\} \cup \{x_k | x_k \notin \overline{C}_i(X)\} \\ &= \{x_k | x_k \in \underline{C}_i\} \cup \{x_k | x_k \in C_i^{BN}\} \cup \{x_k | x_k \notin \overline{C}_i(X)\} \end{aligned} \quad (11)$$

$$\forall_{1 \leq i \leq c} v_i = \begin{cases} \left(\frac{\sum_{x_k \in \underline{C}_i} (u_{ik})^m x_k}{|\underline{C}_i|} + w_{bn} \frac{\sum_{x_k \in C_i^{BN}} (u_{ik})^m x_k}{|C_i^{BN}|} \right) / \sum_{k=1}^n (u_{ik})^m, \\ \text{if } \underline{C}_i \neq \phi \wedge C_i^{BN} \neq \phi; \\ \left(\frac{\sum_{x_k \in \underline{C}_i} (u_{ik})^m x_k}{|\underline{C}_i|} + \sum_{x_k \notin \overline{C}_i} (u_{ik})^m x_k \right) / \sum_{k=1}^n (u_{ik})^m, \\ \text{if } \underline{C}_i \neq \phi \wedge C_i^{BN} = \phi; \\ \left(\frac{\sum_{x_k \in C_i^{BN}} (u_{ik})^m x_k}{|C_i^{BN}|} + \sum_{x_k \notin \overline{C}_i} (u_{ik})^m x_k \right) / \sum_{k=1}^n (u_{ik})^m, \\ \text{if } \underline{C}_i = \phi \wedge C_i^{BN} \neq \phi; \end{cases} \quad (12)$$

In the formula (12), u_{ik} is got in formula (8).

When analyzing, the fuzzy rough clustering algorithm needs to comply with the following three rules:

- (a) A data object can only belong to the lower approximation of a cluster.
- (b) If the data object belongs to the lower approximation of a cluster, it must also belong to the upper approximation of the cluster.
- (c) When the difference between the maximum fuzzy membership of a cluster (it is u_{Max}) and the fuzzy membership of other cluster (it is u_i) is less than a certain threshold value, the data object can belong to the upper approximation of two or more clusters.

In this paper, the threshold in point C above is called the fuzzy rough threshold, which is denoted as δ .

Fuzzy Rough Clustering Based on Attribute Reduction

In the paper, a fuzzy rough clustering based on attribute reduction (FRCM_BR) is proposed. In the calculation process, the algorithm obtains the importance of each attribute by the reduction algorithm. Then it gives different weight value to each attribute according to the importance. At last, it makes corresponding adjustments for the distance between data object and cluster by different weight value. The new difference is calculated as shown in formula (13).

$$d_{ik}^2 = \sum_{p=1}^{\text{cond}} r_p (x_{kp} - v_{ip})^2 \quad (13)$$

where r_p is the weight of the No. p attribute of the object, cond is the count of the attributes, and

$$1 = \sum_{p=1}^{\text{cond}} r_p \quad (14)$$

Fuzzy rough theory is introduced in clustering analysis.

The fuzzy rough clustering algorithm is described as follows:

Input: Set U has n data objects; number of clusters (c); lower approximate weight (w_l); bound weight (w_{bn}); fuzzy rough threshold (δ); algorithm stop threshold (ε).

Output: c clusters

Steps:

Initialize the weight of each attribute of a data object: $r_p \leftarrow 1/n$;

Randomly generate the c mean points of each cluster,
 The number of iterations: $j \leftarrow 0$;
 do
 Running the reduction algorithm to get the weight of each attribute;
 Using formula (13) calculate the difference between data objects and each cluster;
 Using formula (8) get the membership degree of each object (u_{ik}) in the No. j iteration calculation;
 if $|u_{ik} - u_{mk}| \leq \delta$ && $i \neq m$
 $\underline{C}_i \leftarrow \underline{C}_i \cup \{x_k\}$;
 $\underline{C}_m \leftarrow \underline{C}_m \cup \{x_k\}$;
 else
 $\underline{C}_i \leftarrow \underline{C}_i \cup \{x_k\}$
 end if-else
 Using formula (12) get the mean points of each cluster again;
 The number of iterations: $j \leftarrow j + 1$;
 while $(\max \{|u_{ik}^{(j)} - u_{ik}^{(j-1)}|\} > \varepsilon)$;
 end do-while;

5 Experiments of FRCMbR Algorithm

Experimental Setting

The computer configuration for testing the algorithm: CPU is 2.93 GHz dual-core, and the memory is 4 GB. Software configuration: Window 7, Matlab2007. The experimental UCI data sets are Iris, Nursery, Ionosphere, and Isolet5, respectively. The characteristics of the data sets are shown in Table 1.

The definition of accuracy in the experiment is as follows:

$$\text{accurate}(U) = \frac{\sum_{k=1}^c \text{correct}_i}{|U|} \quad (15)$$

In formula (15), $|U|$ is the number of data sets, and correct_i is the number of data objects correctly divided into corresponding clusters.

Table 1 Characterization of four UCI sets

Data set	Number of samples	Attribute dimension	Number of clusters	Instructions
Iris	150	4	3	Data on Iris flowers
Nursery	12,960	8	3	Nursery data
Ionosphere	351	34	2	Radar acquisition data
Isolet5	1559	617	26	Alphabetic predictive data

Table 2 Comparison of correction for Iris set

Algorithm	Number of samples	Correct samples	Wrong samples	Accuracy (%)
K-Means	150	105	45	70.0
FCM	150	130	20	82.7
RCM	150	129	21	86.0
FRCM	150	132	18	88.0
FRCMbR	150	139	11	92.7

Table 3 Comparison of correction for Nursery set

Algorithm	Number of samples	Correct samples	Wrong samples	Accuracy (%)
K-Means	12,960	8451	4509	65.2
FCM	12,960	8935	4025	68.9
RCM	12,960	9051	3909	69.8
FRCM	12,960	10,455	4170	80.7
FRCMbR	12,960	10,561	2399	81.4

Table 4 Comparison of correction for Ionosphere set

Algorithm	Number of samples	Correct samples	Wrong samples	Accuracy (%)
K-Means	351	207	144	59.0
FCM	351	234	117	66.7
RCM	351	241	110	68.7
FRCM	351	249	102	70.9
FRCMbR	351	292	59	83.2

Table 5 Comparison of correction for Isolet5 set

Algorithm	Number of samples	Correct samples	Wrong samples	Accuracy (%)
K-Means	1559	894	665	57.3
FCM	1559	1032	527	61.2
RCM	1559	1025	534	65.7
FRCM	1559	1098	461	70.4
FRCMbR	1559	1286	273	82.4

Comparison of the Accuracy of Each Algorithm

Tables 2, 3, 4, and 5 shows the accuracy of different algorithm after analysis of the above four data sets. For the four experiments of FRCMbR algorithm, the stooping threshold $\epsilon = 0.00001$, the lower approximate weight $w_l = 0.9$, and the bound weight $w_{bn} = 0.1$.

According to the analysis results in Tables 2, 3, 4, and 5, the traditional FRCM algorithm combines the fuzzy theory with the rough theory, which makes the analysis more robust. Compared with FCM and RCM, it is more capable of analyzing nonlinear problems, so its analysis results are more accurate. The FRCMbR algorithm proposed in this paper is superior to the other four algorithms

because it not only recombines the fuzzy and rough theory, but also conforms to the concept of fuzzy roughness and is close to the real world. In the process of analysis, the weights of each attribute are dynamically adjusted according to the analysis results, so as to optimize the analysis results. From the perspective of Iris and Nursery data sets, if the attribute dimension of the data set is small, the FRCMbR algorithm is superior to the FRCM algorithm, but its advantages are not particularly obvious. When the attribute dimension increases, the FRCMbR algorithm's analytical capability gradually shows that it is superior to the FRCM algorithm in a word, FRCMbR algorithm has two advantages:

- (a) FRCMbR algorithm combines the characteristics of fuzzy theory and rough theory, making it better able to deal with nonlinear problems, especially when the background knowledge is not clear, fuzzy, and incomplete, it has more advantages.
- (b) FRCMbR algorithm introduces reduction algorithm, which can dynamically adjust the weight of attributes according to the actual importance in the analysis, so as to strengthen the major attributes on final analysis results.

Parameter Experiments of FRCMbR Algorithm

The important parameters to be adjusted in FRCMbR analysis include: lower approximate weight w_l , bound weight w_{bn} , fuzzy rough threshold δ , algorithm stop threshold ε . Where the algorithm stop threshold ε generally selects a minimum value approaching 0. The approximate weight w_l , bound weight w_{bn} , and fuzzy rough threshold δ are directly related to rough calculation. Therefore, this paper conducts experiments on these three parameters to further verify the validity of FRCMbR algorithm. Among the three parameters, w_l and w_{bn} are correlated, $w_l + w_{bn} = 1$, so when analyzing these two parameters, only one of them needs to be set, and the other one can be determined.

In order to test the w_l and w_{bn} , it sets $\delta = 0.1$. The algorithm deals with data set Iris, Nursery, Ionosphere, and Isolet5, respectively, 20 times each for each data set, and then average the results. These test results are shown in Fig. 1.

As can be seen in Fig. 1, when it deals with Iris data set, it sets $w_{bn} = 0.1$, the analysis accuracy 92.7% is the highest; when it deals with Nursery data set, it sets $w_{bn} = 0.15$, the analysis accuracy 82.7% is the highest; when it deals with Ionosphere data set, it sets $w_{bn} = 0.15$, the analysis accuracy 85.9% is the highest; when it deals with Isolet5 data set, it sets $w_{bn} = 0.2$, the analysis accuracy 88.4% is the highest. Conclusion can be drawn from the above analysis:

- (a) When w_{bn} is within the [0.1, 0.2] range, the accuracy is highest. This is because only a small part of the data object that clusters in the whole problem domain is uncertain, but the overall information is uncertain. The impact of these uncertain data objects on the overall analysis results can be small.

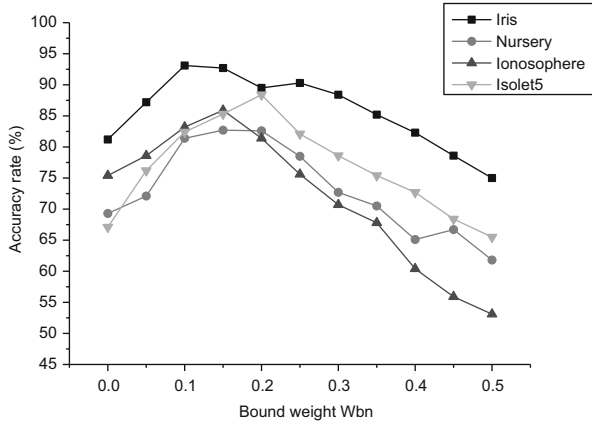
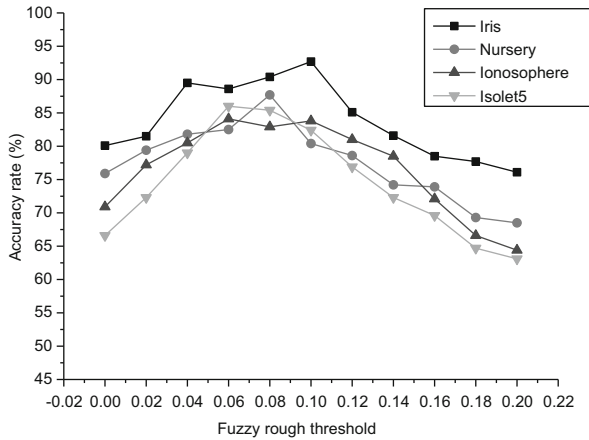


Fig. 1 Influence of boundary weight for correction

Fig. 2 Influence of fuzzy rough threshold for correction



- (b) At that time $w_{bn} = 0$, FRCM δ R algorithm is focused only on the data objects determined by prior knowledge, completely ignoring the bound objects in rough set. And the algorithm is reduced to a fuzzy clustering, that only dealt with lower approximation of each cluster.
- (c) At that time $w_{bn} > 0.25$, the analytical capability of FRCM δ R algorithm is gradually declined. This is because, in the process, the algorithm will only pay more attention to these objects whose prior knowledge is uncertain, but easing to ignore the objects of the low approximation.

In the fuzzy rough threshold δ test, this paper sets the boundary weight $w_{bn} = 0.1$, and the algorithm tests the four data sets, respectively. The test results are shown in Fig. 2.

As can be seen from Fig. 2 when $\delta = 0.1$ in the test for Iris data set, the highest accuracy rate is 92.7%; when $\delta = 0.08$ in the test for Nursery data set, the highest

accuracy rate is 87.7%; when $\delta = 0.06$ in the test for Ionosphere data set, the highest accuracy rate is 84.1%; when $\delta = 0.06$ in the test for Isolet5 data set, the highest accuracy rate is 86.0%. The conclusion is drawn from the above analysis:

- (a) When the δ value range within the $[0.06, 0.1]$, the accuracy is the highest. This is because most of the data is not fuzzy and uncertainty, only the data of bound that is fuzzy.
- (b) Comparing with the bound weight w_{bn} , the fuzzy rough threshold δ is a smaller values The reason is that the δ is a difference of fuzzy membership degrees for the data object belonging to different clusters. (e.g., $|u_{ik} - u_{ij}|$ is the difference of fuzzy membership degree for the No. i data object belonging to the No. k or the No. j cluster). The fuzzy membership degree value is small and it is in $[0, 1]$, so the δ is more smaller.
- (c) At that time $\delta = 0$, the FRCM_bR algorithm is no longer capable of rough analysis, and the algorithm is degraded to fuzzy clustering algorithm.
- (d) At that time $\delta > 0.1$, the analysis capability of FRCM_bR algorithm is also declining. This is because as δ increases, the algorithm divides more data objects into the bound, believing that the whole analysis object is more rough, which is inconsistent with the actual situation.

6 Conclusion

Fuzzy theory and rough set theory helps to understand the real world from different angles. When combined with K-Means algorithm, the two theories can be selected according to different scenarios. Fuzzy K-Means is better at the ambiguous and paradoxical problems, while rough K-Means is better at the problems of incomplete prior knowledge or incomplete prerequisite knowledge. The FRCM_bR algorithm proposed in this paper combines fuzzy theory and rough set theory and deals with problems from two aspects at the same time. FRCM_bR algorithm also introduces the reduction algorithm in rough set theory, in order to obtain the important degree of each attribute, then, dynamically adjust their weight, and further optimize the important attribute. The FRCM_bR algorithm is verified through several experiments on UCI data sets. However, in the process of algorithm calculation, the parameter setting will affect the final analysis results. Therefore, it is necessary to further improve the algorithm in future research to reduce the influence of experience factors on the algorithm.

Acknowledgments Young Teachers' Basic Ability Improving Project of Guangxi Education Hall under Grant No. 2018KY0321; The National Natural Science Foundation of China (61462008, 61751213, 61866004); The Key projects of Guangxi Natural Science Foundation (2018GXNSFDA294001, 2018GXNSFDA281009); The Natural Science Foundation of Guangxi (2017GXNSFAA198365); 2015 Innovation Team Project of Guangxi University of Science and Technology (gxkjdx201504); Scientific Research and Technology Development Project of Liuzhou (2016C050205); Guangxi Colleges and Universities Key Laboratory of Intelligent Processing of

Computer Images and Graphics under Grant No. GIIP201508; Young Teachers' Basic Ability Improving Project of Guangxi Education Hall under Grant No. KY2016YB252; Natural Science Foundation of Guangxi University of Science and Technology under Grant No. 174523.

References

1. Han, J.-w., Kamber, M., & Pei, J. (2012). *Data mining: Concept and techniques* (3rd ed.). Beijing: China Machine Press.
2. Ouyang, H., Dai, X.-S., Wang, Z.-w., & Wang, M. (2015). Rough K-prototypes clustering algorithm based on entropy. *Computer Engineering and Design*, 36(5), 1239–1243.
3. Ouyang, H., Wang, Z.-w., Dai, X.-S., & Liu, Z.-q. (2015). A fuzzy K-prototypes clustering algorithm based on information gain. *Computer Engineering & Science*, 37(5), 1009–1014.
4. Kannan, S. R., Devi, R., Ramathilagam, S., & Takezawa, K. (2013). Effective FCM noise clustering algorithms in medical images. *Computers in Biology and Medicine*, 43(2), 73–83.
5. Ghodousi, M., Alesheikh, A. A., & Saeidian, B. (2016). Analyzing public participant data to evaluate citizen satisfaction and to prioritize their needs via K-means, FCM and ICA. *Cities*, 55(6), 70–81.
6. Liu, L., Sun, S. Z., Yu, H., Yue, X., & Zhang, D. (2016). A modified Fuzzy C-Means (FCM) clustering algorithm and its application on carbonate fluid identification. *Journal of Applied Geophysics*, 129(6), 28–35.
7. Pacheco, F., Cerrada, M., Sánchez, R.-V., Cabrera, D., Li, C., & de Oliveira, J. V. (2017). Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery. *Expert Systems with Applications*, 71(4), 69–86.
8. Nayak, R. K., Mishra, D., Shaw, K., & Mishra, S. (2012). Rough set based attribute clustering for sample classification of gene expression data. *Procedia Engineering*, 38, 1788–1792.
9. Yanto, I. T. R., Vitasari, P., Herawan, T., & Deris, M. M. (2012). Applying variable precision rough set model for clustering student suffering study's anxiety. *Expert Systems with Applications*, 39(1), 452–459.
10. Wang, H., Feng, Q., Lin, X., & Zeng, W. (2009). Development and application of ergodicity model with FRCM and FLAR for hydrological process. *Science China: Technological Sciences*, 52(2), 379–386.
11. Saltos, R., & Weber, R. (2016). A rough–fuzzy approach for support vector clustering. *Information Sciences*, 339(4), 353–368.
12. Dubey, Y. K., Mushrif, M. M., & Mitra, K. (2016). Segmentation of brain MR images using rough set based intuitionistic fuzzy clustering. *Biocybernetics and Biomedical Engineering*, 36(2), 413–426.