October 2016, Volume 18, Issue 5

International Journal of Fuzzy Systems

ISSN: 1562-2479 (Print) 2199-3211 (Online)

Description

The International Journal of Fuzzy Systems (IJFS) is an official journal of Taiwan Fuzzy Systems Association (TFSA) and is published quarterly. IJFS will consider high quality papers that deal with the theory, design, and application of fuzzy systems, soft computing systems, grey systems, and extension theory systems ranging from hardware to software. Survey and expository submissions are also welcome.

Browse Volumes & Issues

Impact Factor	Available
2.198	2015 - 2017
Volumes	**Issues**
3	13

Latest Articles

October 2016, Volume 18, Issue 5, pp 894–903

International Journal of Fuzzy Systems

A Novel Multiple Fuzzy Clustering Method Based on Internal Clustering Validation Measures with Gradient Descent

Authors

Le Hoang Son & Pham Van Hai

*Corresponding to the author: Pham Van Hai haipv@soict.hust.edu.vn

DOI: 10.1007/s40815-015-0117-1

Cite this article as:

Son, L.H. & Van Hai, P. Int. J. Fuzzy Syst. (2016) 18: 894. doi:10.1007/s40815-015-0117-1

Abstract

In this paper, we propose a novel multiple fuzzy clustering method based on internal clustering validation measures with gradient descent. Firstly, some single fuzzy clustering algorithms such as Fuzzy C-Means, Kernel Fuzzy C-Means and Gustafson–Kessel are used to construct similarity matrixes for each partition. Secondly, those similarity matrixes are aggregated into a final one by means of the direct sum of weighted vectors where the values of weights are determined by internal clustering validation measures. Finally, final membership matrix is calculated by minimizing the sum of square errors through the gradient descent method. The proposed approach has been validated in terms of clustering quality on UCI Machine Learning Repository datasets. Experimental results show that the proposed approach's performance is better than those of other ensemble and standalone methods.

Keywords: Fuzzy clustering ensemble Gradient descent Internal clustering validation Similarity matrix

Int. J. Fuzzy Syst. (2016) 18(5):894–903
DOI 10.1007/s40815-015-0117-1

A Novel Multiple Fuzzy Clustering Method Based on Internal Clustering Validation Measures with Gradient Descent

Le Hoang Son[1,2] · Pham Van Hai[3]

Received: 5 July 2015 / Revised: 5 October 2015 / Accepted: 14 November 2015 / Published online: 1 December 2015

Abstract In this paper, we propose a novel multiple fuzzy clustering method based on internal clustering validation measures with gradient descent. Firstly, some single fuzzy clustering algorithms such as Fuzzy C-Means, Kernel Fuzzy C-Means and Gustafson–Kessel are used to construct similarity matrixes for each partition. Secondly, those similarity matrixes are aggregated into a final one by means of the direct sum of weighted vectors where the values of weights are determined by internal clustering validation measures. Finally, final membership matrix is calculated by minimizing the sum of square errors through the gradient descent method. The proposed approach has been validated in terms of clustering quality on UCI Machine Learning Repository datasets. Experimental results show that the proposed approach's performance is better than those of other ensemble and standalone methods.

Keywords Fuzzy clustering ensemble · Gradient descent · Internal clustering validation · Similarity matrix

✉ Pham Van Hai
haipv@soict.hust.edu.vn

Le Hoang Son
lehoangson@tdt.edu.vn

1 Division of Data Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam

2 Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

3 School of Information and Communication Technology, Hanoi University of Science and Technology, Hanoi, Vietnam

1 Introduction

Clustering involves in homogeneous classes so that items in the same class can be similar in different classes [29]. Additionally, clustering can be interpreted as a form of data compression, where a large number of samples are converted into small ones [19]. However, depending on data and applications, various types of similarity measures may be used to identify classes and control how the clusters are formed. Some examples are named such as distance, connectivity, and intensity [12]. In non-fuzzy or hard clustering, a dataset is divided into crisp clusters, where each data point belongs to exactly one cluster [8]. In fuzzy clustering, data points belong to one or more clusters, and associated with each point are membership grades indicating the degree to which data points belong to different clusters [8]. *Fuzzy clustering* techniques are commonly used in a variety of pattern recognition problems, knowledge discovery, risk assessment, geo-demographic analysis, compression, medical diagnosis, and financial forecast [6, 18–22, 25].

Many fuzzy clustering approaches have investigated a group of individual objects in a population such as Fuzzy C-Means (FCM) [5], Gustafson–Kessel (GK) [15], and Kernel Fuzzy C-Means (KFCM) [29]. Nonetheless, those algorithms faced the problems regarding local minima [12], thus degrading clustering quality. Recent efforts for dealing with this problem have aimed to use various clustering solutions derived by different clustering schemes so that the global solutions, by any aggregation method, could have better quality than those obtained in a single scheme [24]. This strategy is called fuzzy clustering ensemble or *multiple fuzzy clustering* [2].

The same manners regarding multiple fuzzy clustering considered mathematical and computational tools such as re-labeling voting, co-association matrices, graph and hyper-graph partitioning, information theory, and finite mixture models to improve the clustering solutions. The following list summarizes principal ideas of the typical methods:

- Ahmadzadeh et al. [1] presented a graph-based approach for clustering ensemble of fuzzy partitions. However, it outputs hard clustering results due to effects of the partition graph;
- Strehl & Ghosh [24] used co-association and hyper-graph methods to design new fuzzy clustering ensemble algorithms such as *Similarity Partitioning Algorithm* (CSPA), *Meta-Clustering Algorithm* (MCLA), and *Hyper-Graph Partitioning Algorithm* (HGPA). In CSPA, similarity among two data points is directly proportional to the number of constituent clusterings of the ensemble. HGPA seeks to directly partition the hyper-graph where all vertices and hyper-edges are weighted equally. MCLA solves the cluster correspondence problem by grouping the clusters identified in the individual clusterings and uses voting to place data points into the final consensus clusters [8];
- Fern & Brodley [9] proposed *Hybrid Bipartite Graph Formulation* (HBGF) to model instances and clusters simultaneously in a bipartite graph, in which individual data points and clusters of constituent clusterings are both vertices [8];
- Vega-Pons & Ruiz-Shulcloper [27] extended the methods of Strehl & Ghosh in fuzzy environment. Specifically, *soft CSPA* (sCSPA) extends CSPA using fuzzy partitions to create a similarity matrix, which is then partitioned by a graph partitioning algorithm—METIS to produce a desired number of clusters. *Soft MCLA* (sMCLA) extends MCLA by accepting soft clustering as an input. *Soft HBGF* (sHBGF) extends HBGF by representing the ensemble as a bipartite graph with clusters and instances as nodes, and edges between the instances and the clusters they belong to, and then using METIS to partition;
- Other works on multiple fuzzy clustering algorithms can be seen in [3, 4, 7, 8, 10, 11, 13, 14, 16, 17, 23]. Among them, the Information-Theoretic K-Means (ITK) algorithm [8] is an efficient multiple fuzzy clustering algorithm which uses Kullback–Leibler (KL) divergence as a distance measure between two instances to solve the soft ensemble. Other works were either developed on different metrics or had worse performance than the methods above so could not be discussed herein.

To the best of our knowledge, the methods of Vega-Pons & Ruiz-Shulcloper (sCSPA, sMCLA, and sHBGF) [27] and of De Oliveira & Pedrycz (ITK) [8] achieved reliable performance over other algorithms working in the same environment including metrics and criteria. However, a major drawback of those algorithms is that they were deployed on the graph-based approach, which could degrade their performances in the case that the number of data points is large. When processing large data points, a vector quantization algorithm is usually used to create representatives which are further regarded as vertices of a graph. Then those vertices are classified instead of the original data points, thus making a trade-off between computational complexity and performance of the algorithm. Nevertheless, the algorithms in [8, 27] considered the original data points for clustering without using representatives. Therefore, finding feasible solutions in such a large graph is inappropriate in terms of computational complexity. Another approach for multiple fuzzy clustering is indeed necessary in this situation.

In this paper, we propose a novel multiple fuzzy clustering method based on internal clustering validation measures with gradient descent. Specifically, main ideas of the new method are highlighted as follows.

(1) *Firstly*, some single fuzzy clustering algorithms such as FCM [5], GK [15], and KFCM [29] are used to create similarity matrixes $S^{(i)}$, $i = 1, 2, 3$ for each partition;

(2) *Secondly*, the similarity matrixes are aggregated into a final one by means of the direct sum of weighted vectors where the values of weights are determined by internal clustering validation measures;

(3) *Thirdly*, the final membership matrix is calculated by the minimization of the sum of square errors through the gradient descent method. An iterative strategy to find this matrix is presented.

From the mentioned ideas, the contributions and novelty (unique features) of this work are demonstrated as follows:

(1) first attempt to combine internal clustering validation measures and gradient descent in order to make an efficient multiple fuzzy clustering algorithm in terms of clustering quality and computational time;

(2) proposes a new aggregation method of single clustering solutions based on the direct sum of weighted vectors where the values of weights are determined by internal clustering validation measures. In the previous works [8, 27], the authors used complicated aggregation operators like entropy function and neglected the determination of weights. The new method remedies these by choosing weights that give the best performance among all and using them in a more concise and simpler aggregation way;

(3) takes into account the error of aggregation in separation process of the final similarity matrix into the membership matrix by means of gradient

descent. The previous works also ignored the error and used the similarity matrix to specify which clusters the data elements belong to. This is inaccurate since in fuzzy clustering the membership matrix is the most popular way to create fuzzy clusters. The new method has solved the existing problem by proposing the model to determine the final membership matrix;

(4) all the new contributions are expressed in a novel multiple fuzzy clustering method described in Sect. 2 of this paper. This algorithm would obtain a more reliable and higher accuracy than other single and multiple clustering methods. The proposed framework will be evaluated on the benchmark datasets of UCI Machine Learning Repository [26] in terms of clustering quality [28]. The findings would suggest the efficiency and advantages of the proposed approach for the clustering problem.

The significance, advantages, and practical implication of this study can be demonstrated as follows. *Firstly*, the proposed algorithm enriches the knowledge of deploying clustering algorithms in ensemble environments for the sake of clustering quality and computational complexity. *Secondly*, it can be applied to various pattern recognition problems that require high accuracy of results such as land classification, medical segmentation, emergency response systems, and so on [6, 18–22, 25]. *Thirdly*, the new method is straightforward and easy for implementation. *Fourthly*, the theoretical contribution of this paper could expand a minor research direction regarding multiple fuzzy clustering algorithms based on internal clustering validation measures with gradient descent.

Despite having above advantages, some drawbacks of the new method are listed as follows. *Firstly*, the new method did not show how many single clustering solutions are enough in the sense that they guarantee both final clustering quality and computational complexity. *Secondly*, the order of single solutions was not taken into account. It is obvious that only several single solutions contribute greatly to the values of final results. Indeed, an order of such solutions should be put into design of the algorithm. *Thirdly*, more numbers of internal clustering validation measures should be used to give more accurate weights. *Finally*, it is necessary to specify equivalent clusters in the single clustering algorithms for aggregation procedure. Those limitations are our further researches for this theme.

The structure of this paper is organized as follows: Sect. 2 presents the proposed method accompanied with the implementation. Experimental results on the benchmark datasets of UCI Machine Learning Repository are described in Sect. 3. Finally, Sect. 4 gives the conclusions and delineates further works.

2 The Proposed Method

In this section, we firstly present the main ideas of the new algorithm in Sect. 2.1. Pseudo-code of the algorithm is given in Sect. 2.2.

2.1 The Main Ideas

In this section, we elaborate the summary of ideas in Sect. 1 in detail.

2.1.1 Clustering by Single Algorithms

The problem is formulated as follows. Consider a dataset X consisting of N data points in r dimensions. Let us divide the dataset into C clusters with some pre-defined parameters such as the fuzzifier m and the maximal number of iteration steps max *Step*. The first step of the new algorithm is using some single fuzzy clustering algorithms such as FCM [5], GK [15], and KFCM [29] to create various clustering solutions. Details of activities of those clustering algorithms are mentioned in the equivalent papers so that they are not discussed herein.

2.1.2 Synthesis of the Single Clustering Results

After getting the single clustering solutions, we aggregate them into the final one as follows. Consider the Euclidean distance among two data points in the view of the multiple clustering scheme as follows:

$$d_{ij}^{(q)} = d^{(q)}\left(X_i, X_j\right) = \left(\sum_{l=1}^{C(q)} \left(u_{il}^{(q)} - u_{jl}^{(q)}\right)^2\right)^{1/2}, \tag{1}$$

$i, j = \overline{1, N}; i \neq j,$

where $U_{il}^{(q)}$ is the membership degree of data point ith to cluster lth $(i = \overline{1, N},\ l = \overline{1, C(q)})$ in the qth clustering results. The value of $C(q)$ could be different for various clustering results, but in our case $C(q) = C,\ \forall q = 1, 2, 3$. The membership matrix for each clustering result satisfies the constraints (2):

$$\begin{cases} u_{kj}^{(q)} \in [0, 1] \\ \sum_{j=1}^{C(q)} u_{kj}^{(q)} = 1 \\ k = \overline{1, N}; j = \overline{1, C(q)} \end{cases}. \tag{2}$$

The similarity matrix $S^{(q)}$ for the qth clustering results $(\forall q = 1, 2, 3)$ is calculated as

$$S_{ij}^{(q)} = e^{-\left(d_{ij}^{(q)}\right)^2}. \tag{3}$$

The final similarity matrix is aggregated by direct sum of weighted vectors.

$$S = F\left(S^{(1)}, S^{(2)}, S^{(3)}\right) = \sum_{q=1}^{3} w_q \times S^{(q)}, \tag{4}$$

where w_q is the weight of the similarity matrix $S^{(q)}$ satisfying

$$\sum_{q=1}^{3} w_q = 1. \tag{5}$$

2.1.3 Finding Appropriate Weights

According to Eq. (4), weights of the similarity matrix must be determined in order to calculate the final similarity matrix. Our idea for this task is to use some internal clustering validation measures such as Dunn's index (DI) and the partition coefficient (PC) [28] to create those weights. Definitions of the measures are as follows:

- DI:

$$V_{DI} = \min_{1 \le i \le C} \min_{\substack{1 \le j \le C \\ i \ne j}} \frac{\delta(C_i, C_j)}{\max\limits_{1 \le k \le C} \Delta_k} \tag{6}$$

$$\delta(C_i, C_j) = \min\{d(X_i, X_j) | X_i \in C_i, X_j \in C_j\}. \tag{7}$$

In those equations, $\delta(C_i, C_j)$ is the distance between clusters C_i and C_j and Δ_k is the average distance from cluster elements to the center of cluster kth. The larger value of DI index means a better clustering result.

- PC:

$$V_{PC} = \frac{1}{N} \sum_{k=1}^{N} \sum_{j=1}^{C} (u_{kj})^2. \tag{8}$$

The larger value of PC index means a better clustering result.

From Eqs. (4–8), the following formula is used to generate the weights:

$$w_q^h = \frac{V_h^{(q)}}{\sum_{q=1}^{3} V_h^{(q)}}, \tag{9}$$

$$w'_q = \left(\sum_{h=1}^{2} w_q^h\right)/2, \tag{10}$$

$$w_q = \frac{w'_q}{\sum_{q=1}^{3} w'_q}, \tag{11}$$

where $V_h^{(q)}$ is the value of validation measure hth ($h = 1$(DI) or 2 (PC)) for the qth clustering results ($\forall q = 1, 2, 3$). Using the internal clustering validation measures, the final similarity matrix leads to the clustering results having the best performance among them.

2.1.4 Determining Final Results

Now, we have the final similarity matrix S. In order to determine the final membership matrix from S, it is necessary to solve the equation:

$$S_{kl} = \sum_{j=1}^{C} u_{kj} u_{lj} + \varepsilon_{kl}, \tag{12}$$

where ε_{kl} is an error between two data points X_k and X_l.

The gradient descent method is applied to solve Eq. (12) by minimizing the following sum of squares error:

$$\varphi^2 = \frac{\sum_{k=1}^{N} \sum_{l=1}^{N} \left(S_{kl} - \sum_{j=1}^{C} u_{kj} u_{lj}\right)^2}{\sum_{k=1}^{N} \sum_{l=1}^{N} \left(S_{kl} - \overline{S}\right)^2} \to \min. \tag{13}$$

Reducing (13), we have the following problem:

$$J = \sum_{k=1}^{N} \sum_{l=1}^{N} \left(S_{kl} - \alpha \sum_{j=1}^{C} u_{kj} u_{lj}\right)^2 \to \min. \tag{14}$$

Taking the derivative of J with respect to α, we obtain

$$\alpha = \frac{\sum_{k=1}^{C} \sum_{l=1}^{C} S_{kl} \sum_{j=1}^{C} u_{kj} u_{lj}}{\sum_{k=1}^{N} \sum_{l=1}^{N} \left(\sum_{j=1}^{C} u_{kj} u_{lj}\right)^2}. \tag{15}$$

The descent vectors are determined as follows:

$$\frac{\partial J}{\partial u_{kj}} = -2\alpha \sum_{\substack{l=1 \\ l \ne k}}^{N} u_{lj} \left(S_{kl} - \alpha \sum_{j=1}^{C} u_{kj} u_{lj}\right). \tag{16}$$

From (15–16), the following method is used to find the final solution U.

Input	Data X and the number of clusters—C
Output	The membership matrix U
Algorithm	
1	Initialize $U(0)$. Set the step number: $p = 0$
2	Set $p = p + 1$ and calculate $\alpha(p)$ by Eq. (15)
3	Calculate $\frac{\partial J}{\partial u_{kj}(p-1)}$ by Eq. (16) and find the optimal solution with respect to the direction of the descent vector using one-dimensional direct search
4	Update rule: $u_{kj}(p) = u_{kj}(p-1) - \lambda \frac{\partial J}{\partial u_{kj}(p-1)}$ with $\lambda > 0$ being the step size
5	If $\|U(p) - U(p-1)\| < \varepsilon$ then stop; otherwise go to Step 2

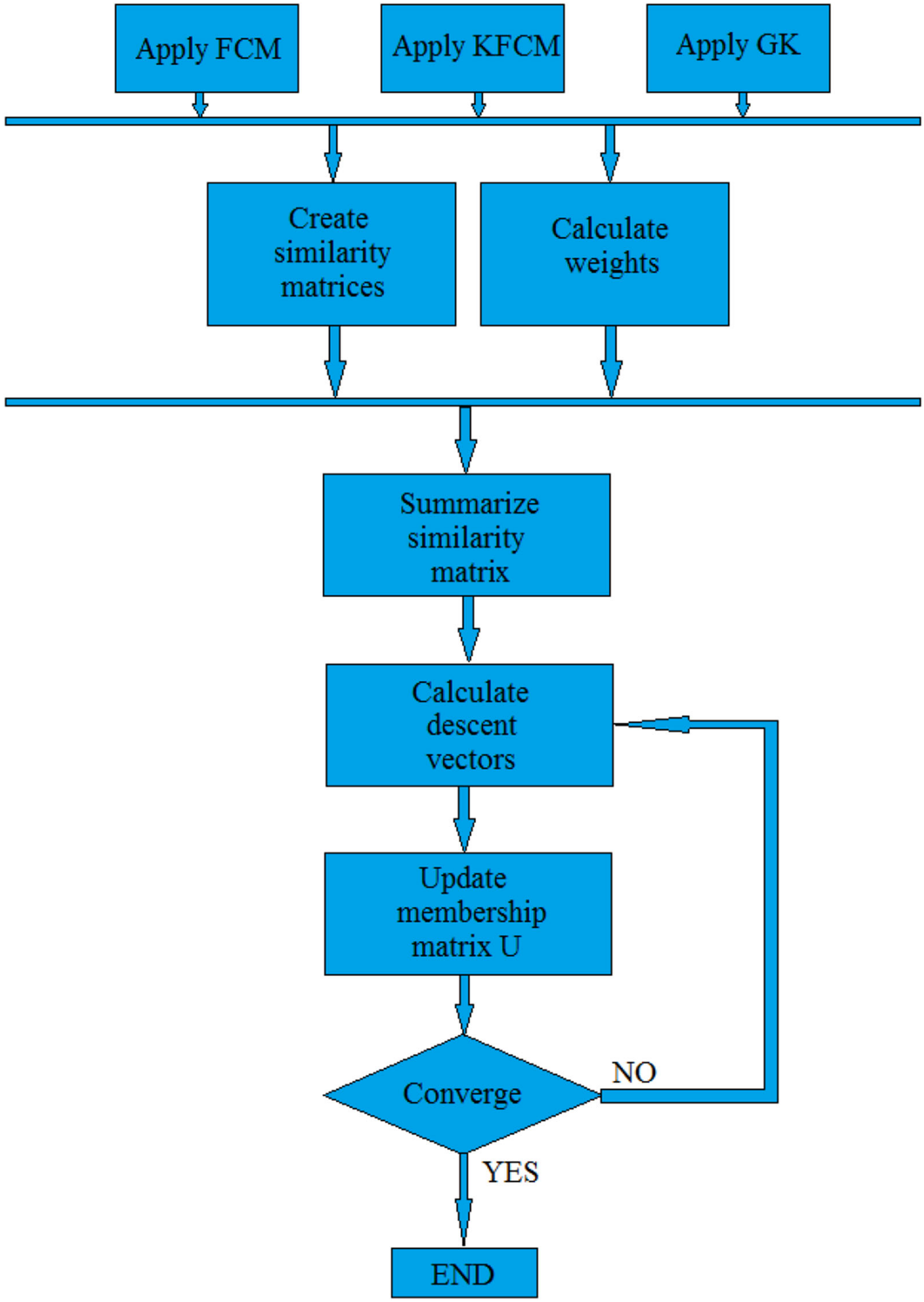

Fig. 1 The flowchart of the algorithm

It is straightforward to determine the clusters and cluster centers from the membership matrix computed by the iteration scheme above.

2.2 The Pseudo-Code

The following pseudo-code shows the activities of the algorithm where the construction of single solutions by

Table 1 UCI dataset

Name	Size	Number of instances	Number of attributes	Number of classes	Number of instances of each class
Iris	4.4 K	150	4	3	50,50,50
Wine	11 K	178	13	3	59,71,48
Glass	12 K	214	10	6	70,76,17,13,9,29
Seeds	9.1 K	210	7	3	70,70,70
Knowledge Modeling	57 K	258	5	4	24,88,83,63
Haberman's Survival	3 K	306	3	2	225,81

Table 2 Comparison by the PC index

	GK	FCM	KFCM	Proposed method
Iris	0.3351	0.8009	0.7069	0.9813
Wine	0.3357	0.5792	0.4606	0.5776
Glass	0.1713	0.7069	0.6602	0.7792
Seeds	0.3356	0.7365	0.6186	0.9887
Knowledge modeling	0.2509	0.5482	0.4379	0.5339
Haberman's survival	0.5076	0.7393	0.6517	0.9249

Table 3 Comparison by the DI index

Method	GK	FCM	KFCM	Proposed method
Iris	0.0659	0.1077	0.1064	0.3066
Wine	0.0041	0.0118	0.0220	0.0232
Glass	0.0212	0.03	0.0299	0.0302
Seeds	0.0403	0.0462	0.0467	0.0470
Knowledge modeling	0.0737	0.0739	0.0711	0.074
Haberman's survival	0.0210	0.0272	0.0262	0.0273

FCM, KFCM, and GK algorithms is shown in lines 2, 5, and 8. The final similarity matrix is constructed on line 14 using internal clustering validation measures. A flow chart of the algorithm is shown in Fig. 1.

1. U_1 = *FCM(data,number_of_cluster);*
2. S_1 = *createSimilarity*(U_1);
3. V_1 = *validity*(U_1,*measure_name*);
4. U_2 = *KFCM(data,number_of_cluster);*
5. S_2 = *createSimilarity*(U_2);
6. V_2 = *validity*(U_2,*measure_name*);
7. U_3 = *GK(data,number_of_cluster);*
8. S_3 = *createSimilarity*(U_3);
9. V_3 = *validity*(U_3,*measure_name*);
10. *SumV* = $V_1 + V_2 + V_3$;
11. $w_1 = V_1$/*SumV*;
12. $w_2 = V_2$/*SumV*;
13. $w_3 = V_3$/*SumV*;
14. $S = w-_1{*}S_1 + w_2{*}S_2 + w_3{*}S_3$;
15. $U = U_1$;
16. *ax = calculatealpha(U,S);*
17. *do {*
18. *Uold = U;*
19. *for i = 1:n do*
    a. *for j = i:n do*
    b. *for k = 1:K do*
    c. *Q(i,j) = Q(i,j) + U(i,k)*U(j,k);*
    d. *endfor;*
    e. *Q(j,i) = Q(i,j);*
    f . *endfor;*
20. *endfor;*
21. *for a = 1:n do*
    a. *for l = 1:K do*
    b. *for i = 1:n do*
    c. *if (a ! = i) then*
    d. *G(a,l) = G(a,l)-2*ax*(S(a,i)-ax*Q(a,i))*U(i,l);*
    e. *endif;*
    f. *endfor;*
    g. *U(a,l) = U(a,l)-step_size*G(a,l);*
    h. *endfor;*
22. *endfor;*
23. *ax = calculatealpha(U,S);*
24. *} while(max(abs(U-Uold)) > threshold);*

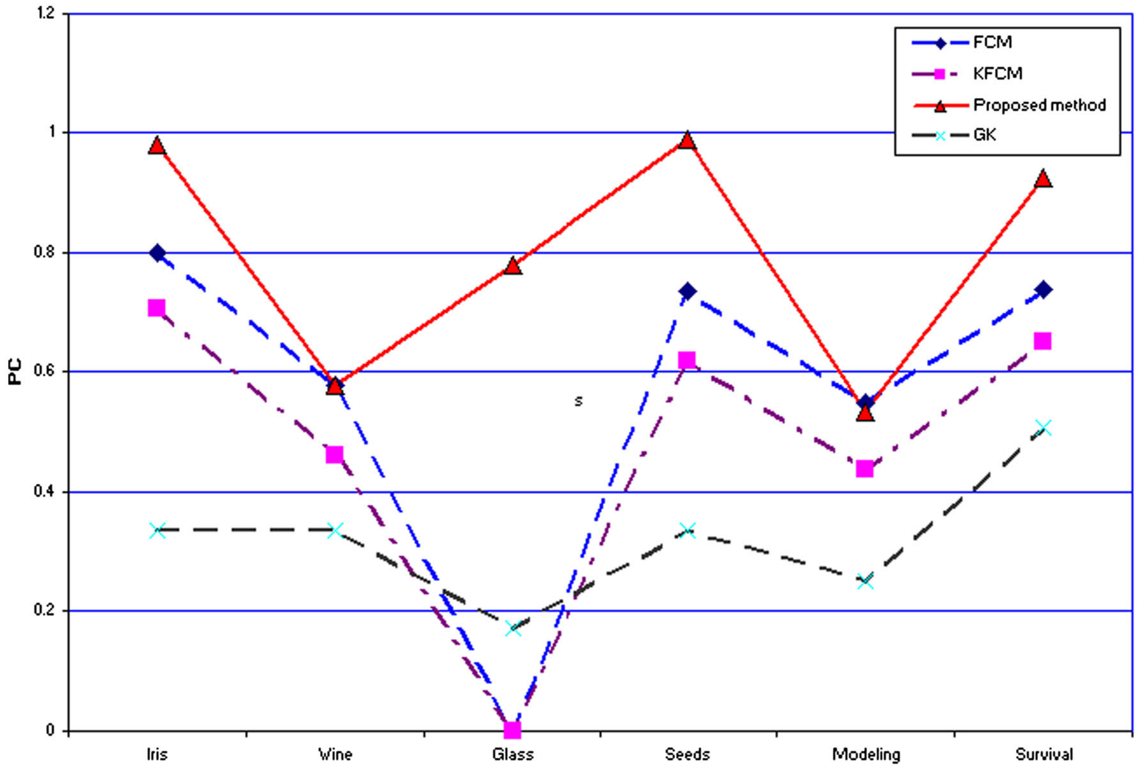

Fig. 2 PC values of algorithms

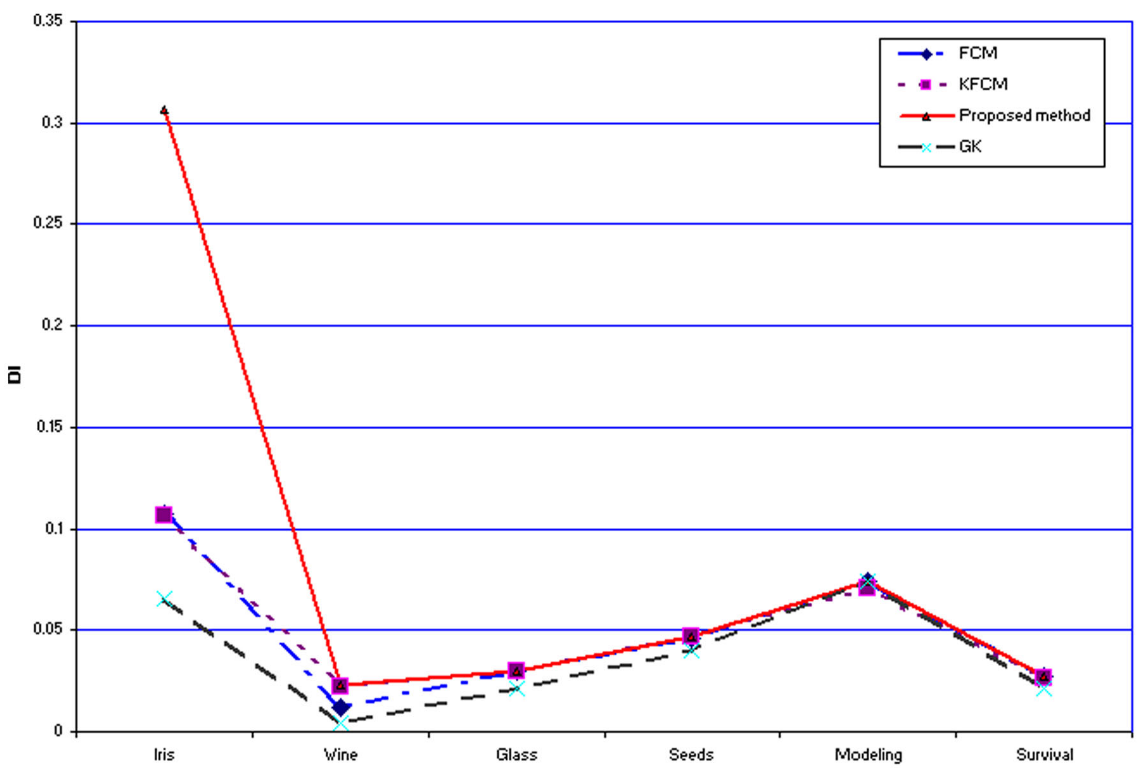

Fig. 3 DI values of algorithms

Table 4 Comparison of multiple fuzzy clustering algorithms (%)

Method	ITK	sCSPA	sMCLA	sHBGF	Proposed method
Iris	60	74	66	70	80
Wine	71	79	70	73	81
Glass	49	49	54	46	53
Seeds	72	83	77	79	86
Knowledge Modeling	78	85	83	83	85
Haberman's Survival	79	92	90	92	92

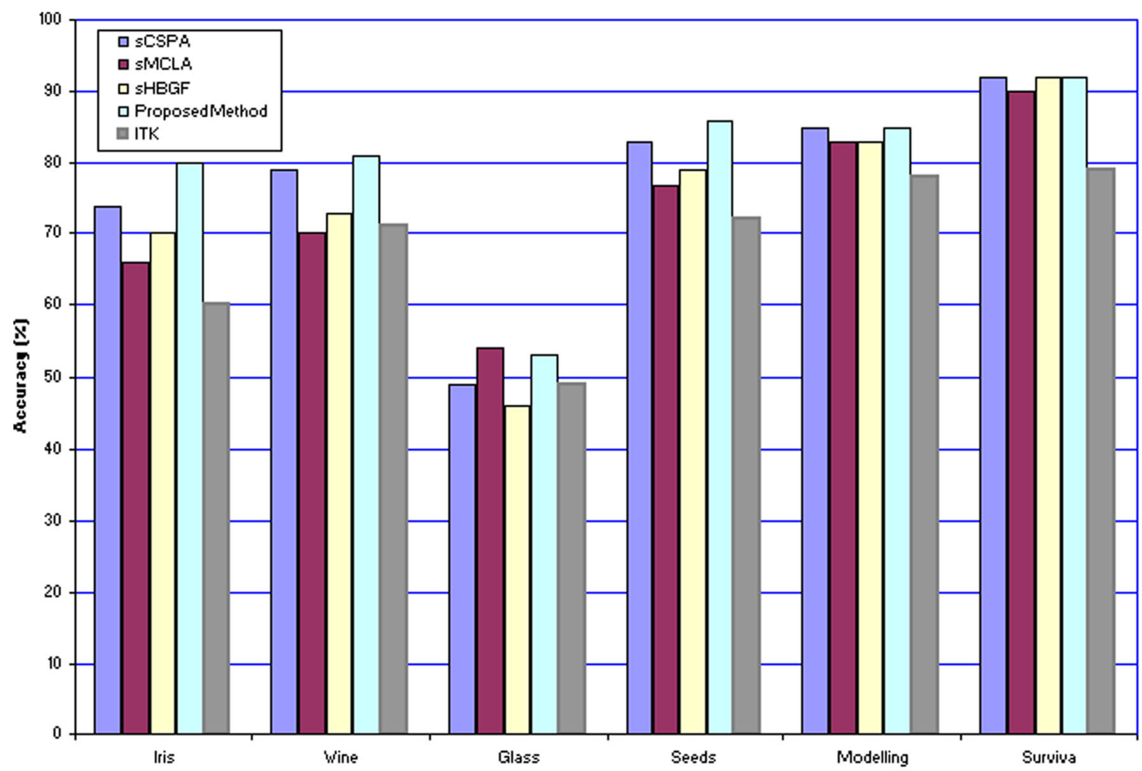

Fig. 4 Accuracies of algorithms (%)

Table 5 Comparison of computational time (sec)

Method	GK	FCM	KFCM	ITK	sCSPA	sMCLA	sHBGF	Proposed method
Iris	24.6	14.7	11.1	65.6	105.1	331.7	73.2	63
Wine	23.7	13.2	8.1	86.3	97.6	175.8	482.4	56.3
Glass	56.1	29.4	18	106.9	93	455.7	89.1	129.4
Seeds	26.7	15.9	9.6	75.7	161.1	66.5	231.7	65.3
Knowledge Modeling	40.8	23.1	15	106.7	99.6	99.7	283.5	98.6
Haberman's Survival	31.8	16.2	14.1	90.8	80.5	83.6	156.7	77.6

3 Results and Discussion

We have implemented the proposed algorithm in addition to the single methods—FCM [5], GK [15], and KFCM [29], and the multiple fuzzy clustering algorithms, namely sCSPA, sMCLA, and sHBGF [27], and the Information-Theoretic K-Means (ITK) algorithm [8] in Matlab 2014 and executed them on a PC VAIO laptop with Core i5 processor. The experimental results are taken as the average values after 20 runs. The experimental datasets are taken from benchmark UCI Machine Learning Repository [26] including Iris, Wine, Glass, Seeds, User Knowledge Modeling, and Haberman's Survival (Table 1). The validity measures are DI, PC, and the accuracy.

In the following Tables 2 and 3, we describe the comparison of the proposed method and the single methods—GK, FCM, and KFCM on the UCI datasets in terms of clustering quality. Tables 2 and 3 show the results evaluated by the PC and DI, respectively.

It is obvious that the values of the proposed method are larger than those of the other algorithms. In order to understand the performances of the algorithms, we show the variations of index values on various datasets (Figs. 2, 3). It is clear from the figures that the proposed method is effective on the dataset having nearly equal distribution of data points in clusters like IRIS. Furthermore, the PC and DI values of the proposed method are better than those of GK, FCM, and KFCM.

As shown in Table 4, we compare the performance of multiple fuzzy clustering algorithms in terms of accuracy. Each clustering result is compared with the accurate instances in clusters as in Table 1, and the accuracy is made by taking the maximal percentage of the number of corrected results overall in a cluster. From the illustration in Fig. 4, it is clearly shown that the proposed method has better accuracy than the other ones.

Table 5 demonstrates the comparison of computational time of all algorithms. It has been clearly observed that the proposed algorithm is slower than the single clustering algorithm, namely GK, FCM, and KFCM, but mostly faster than other multiple fuzzy clustering schemes. Combining the results in Tables 4 and 5, we recognize that the clustering quality of the proposed work is approximate to (better than) those of other multiple schemes, while its computational time is smaller than those of the others. This shows the advantages of the new method in comparison with the relevant ones.

4 Conclusions

In this paper, a novel multiple fuzzy clustering method based on internal clustering validation measures with gradient descent was proposed. The new algorithm uses the membership matrices obtained by various single fuzzy clustering models and creates similarity matrixes for each partition. These matrixes are then summarized in a global similarity matrix through the internal clustering validation measures. Final clustering results are specified by the gradient descent method of fuzzy clustering based on the similarity matrix. The experiments on the UCI datasets showed that the proposed approach is more effective than the single fuzzy clustering algorithms such as GK, FCM, and KFCM in the same manner as relevant multiple fuzzy clustering algorithms, namely ITK, sCSPA, sMCLA, and sHBGF in terms of reliable results and lower error rates. Computational time of the algorithms was also investigated.

There are several directions for future research such as using other validation measures to summarize similarity matrix or taking other models such as the Kernel-based ones. Some drawbacks of the proposed works stated in Introduction will be focused in the upcoming research works.

Acknowledgments The authors are greatly indebted to the Editor-in-chief, Prof. Shun-Feng Su, and anonymous reviewers for their comments and valuable suggestions that improved the quality and clarity of this paper. A special thanks to Mr. Pham Van Viet, HUST, for coding of algorithms in the revision stage. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2014.01.

References

1. Ahmadzadeh, M., Golestan, Z.A., Vahidi, J., Shirazi, B.: A graph based approach for clustering ensemble of fuzzy partitions. J. Math. Comput. Sci. **6**, 154–165 (2013)
2. Alizadeh, H., Minaei-Bidgoli, B., Parvin, H.: To improve the quality of cluster ensembles by selecting a subset of base clusters. J. Exp. Theor. Artif. Intell. **26**(1), 127–150 (2014)
3. Avogadri, R., & Valentini, G. (2008). Ensemble clustering with a fuzzy approach. In Supervised and Unsupervised Ensemble Methods and their Applications (pp. 49–69). Springer, Berlin
4. Avogadri, R., Valentini, G.: Fuzzy ensemble clustering based on random projections for DNA microarray data analysis. Artif. Intell. Med. **45**(2), 173–183 (2009)
5. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2), 191–203 (1984)
6. Chen, J., Zhao, S., Wang, H.: Risk analysis of flood disaster based on fuzzy clustering method. Energy Procedia. **5**, 1915–1919 (2011)
7. Chen, L., Chen, C.P., Lu, M.: A multiple-kernel fuzzy C-means algorithm for image segmentation. IEEE Trans. Syst. Man Cybern. B Cybern. **41**(5), 1263–1274 (2011)
8. De Oliveira, J.V., Pedrycz, W. (eds.): Advances in fuzzy clustering and its applications. Wiley, New York (2007)
9. Fern, X. Z., Brodley, C. E. (2004). Solving cluster ensemble problems by bipartite graph partitioning. Proceedings of the 21st ACM International Conference on Machine Learning, p. 36
10. He, M., Cai, W.J., Li, S.Y.: Multiple fuzzy model-based temperature predictive control for HVAC systems. Inf. Sci. **169**(1), 155–174 (2005)
11. Huang, H.C., Chuang, Y.Y., Chen, C.S.: Multiple kernel fuzzy clustering. IEEE Trans. Fuzzy Syst. **20**(1), 120–134 (2012)
12. Jayaram, B., Klawonn, F. (2013). Can fuzzy clustering avoid local minima and undesired partitions?. In Computational Intelligence in Intelligent Data Analysis, pp. 31–44. Springer, Berlin

13. Jeng, J.T., Chuang, C.C., Tseng, C.C., Juan, C.J.: Robust interval competitive agglomeration clustering algorithm with outliers. Int. J. Fuzzy Syst. **12**(3), 227–236 (2010)
14. Lee, S.H., Pedrycz, W., Sohn, G.: Design of similarity and dissimilarity measures for fuzzy sets on the basis of distance measure. Int. J. Fuzzy Syst. **11**(2), 67–72 (2009)
15. Lesot, M. J., & Kruse, R. (2006). Gustafson-Kessel-like clustering algorithm based on typicality degrees. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, *IPMU*, pp. 1300–1307
16. Miyamoto, S.: Different objective functions in fuzzy c-means algorithms and kernel-based clustering. Int. J. Fuzzy Syst. **13**(2), 89–97 (2011)
17. Shieh, H.L.: A hybrid fuzzy clustering method with a robust validity index. Int. J. Fuzzy Syst. **16**(1), 39–45 (2014)
18. Son, L.H., Cuong, B.C., Lanzi, P.L., Thong, N.T.: A novel intuitionistic fuzzy clustering method for geo-demographic analysis. Expert Syst. Appl. **39**(10), 9848–9859 (2012)
19. Son, L.H., Linh, N.D., Long, H.V.: A lossless DEM compression for fast retrieval method using fuzzy clustering and MANFIS neural network. Eng. Appl. Artif. Intell. **29**, 33–42 (2014)
20. Son, L.H., Thong, N.T.: Intuitionistic fuzzy recommender systems: an effective tool for medical diagnosis. Knowledge-Based Syst. **74**, 133–150 (2015)
21. Son, L.H.: A novel kernel fuzzy clustering algorithm for geo-demographic analysis. Inf. Sci. **317**, 202–223 (2015)
22. Son, L.H.: DPFCM: a novel distributed picture fuzzy clustering method on picture fuzzy sets. Expert Syst. Appl. **42**(1), 51–66 (2015)
23. Srivastava, V., Tripathi, B.K., Pathak, V.K.: Evolutionary fuzzy clustering and functional modular neural network-based human recognition. Neural Comput. Appl. **22**(1), 411–419 (2013)
24. Strehl, A., Ghosh, J.: Cluster ensembles—a knowledge reuse framework for combining multiple partitions. J. Mach. Learn. Res. **3**, 583–617 (2003)
25. Thong, P. H., Son, L. H. (2014). A new approach to multi-variables fuzzy forecasting using picture fuzzy clustering and picture fuzzy rules interpolation method. Proceeding of 6th International Conference on Knowledge and Systems Engineering, pp. 679–690
26. UC Irvine (2015). UCI Machine Learning Repository. Available at: http://archive.ics.uci.edu/ml
27. Vega-Pons, S., Ruiz-Shulcloper, J.: A survey of clustering ensemble algorithms. Int. J. Pattern Recognit. Artif. Intell. **25**(03), 337–372 (2011)
28. Vendramin, L., Campello, R.J., Hruschka, E.R.: Relative clustering validity criteria: a comparative overview. Stat. Anal. Data Mining: The ASA Data Sci. J. **3**(4), 209–235 (2010)
29. Zhang, D., & Chen, S. Fuzzy clustering using kernel method. 2002 International Conference on Control and Automation, 2002. ICCA

Le Hoang Son obtained Ph.D. degree on Mathematics – Informatics at VNU University of Science, Vietnam National University (VNU). He has been working as a researcher and lecturer at the Center for High Performance Computing, VNU University of Science, Vietnam National University, since 2007. His major field includes Soft Computing, Fuzzy Clustering, Recommender Systems, Geographic Information Systems (GIS), and Particle Swarm Optimization. He is a member of IACSIT, member of Center for Applied Research in e-Health (eCARE), and also an associate editor of the International Journal of Engineering and Technology (IJET). Dr. Son served as a reviewer for various international journals and conferences such as PACIS 2010, ICMET 2011, ICCTD 2011, KSE 2013, BAFI 2014, NICS 2014 & 2015, ACIIDS 2015 & 2016, ICNSC15, GIS-2015, FAIR 2015, International Journal of Computer and Electrical Engineering, Imaging Science Journal, International Journal of Intelligent Systems Technologies and Applications, IEEE Transactions on Fuzzy Systems, Expert Systems with Applications, International Journal of Electrical Power and Energy Systems, Neural Computing and Applications, International Journal of Fuzzy System Applications, Intelligent Data Analysis, Computer Methods and Programs in Biomedicine, World Journal of Modeling and Simulation, Knowledge-Based Systems, and Engineering Applications of Artificial Intelligence. He gave a number of invited talks at many conferences such as 2015 National Fundamental and Applied IT Research (FAIR 15'), 2015 National conference of Vietnam Society for Applications of Mathematics (VietSam15'), 2015 Conference on Developing Applications in Virtual Reality, GIS and Mobile technologies, and International Conference on Mathematical Education Vietnam 2015 (ICME Vietnam 2015). Dr. Son has got 47 publications in prestigious journals and conferences including 17 SCI/SCIE papers and undertaken more than 20 major joint international and national research projects. He has published 2 books on mobile and GIS applications. So far, he has awarded the "2014 VNU Research Award for Young Scientists."

Pham Van Hai obtained a Masters Degree (M.Sc.) in computer science at Asian Institute of Technology. Following the Masters Degree, he has been awarded Doctor of Engineering degree (Ph.D.) in computer intelligence and information science by Ritsumeikan University, Japan, visiting researcher in Computational Methods and Machine learning at Lamar University (USA). He has been currently jointed research projects at Soft Intelligence Lab, Ritsumeikan University, and worked as a lecturer–assistant professor at School of Information Technology and Communication, Hanoi University of Science and Technology. He has been Master Supervisor and Ph.D. advisor. He has published 70 international/national journals and conference publications and joined with international projects, and has served as the editorial board of a number of international journals. He is interested in dynamic systems, intelligent hybrid systems, intelligent decision support systems, data mining, complex systems, context-aware systems, artificial intelligence, knowledge-based systems for the domain of health care, telemedicine, and automatic intelligent system dialogs between doctor and patient.