Lê Thị Cẩm Bình
Khoa TT-TV

ĐẠI HỌC HẠ LONG

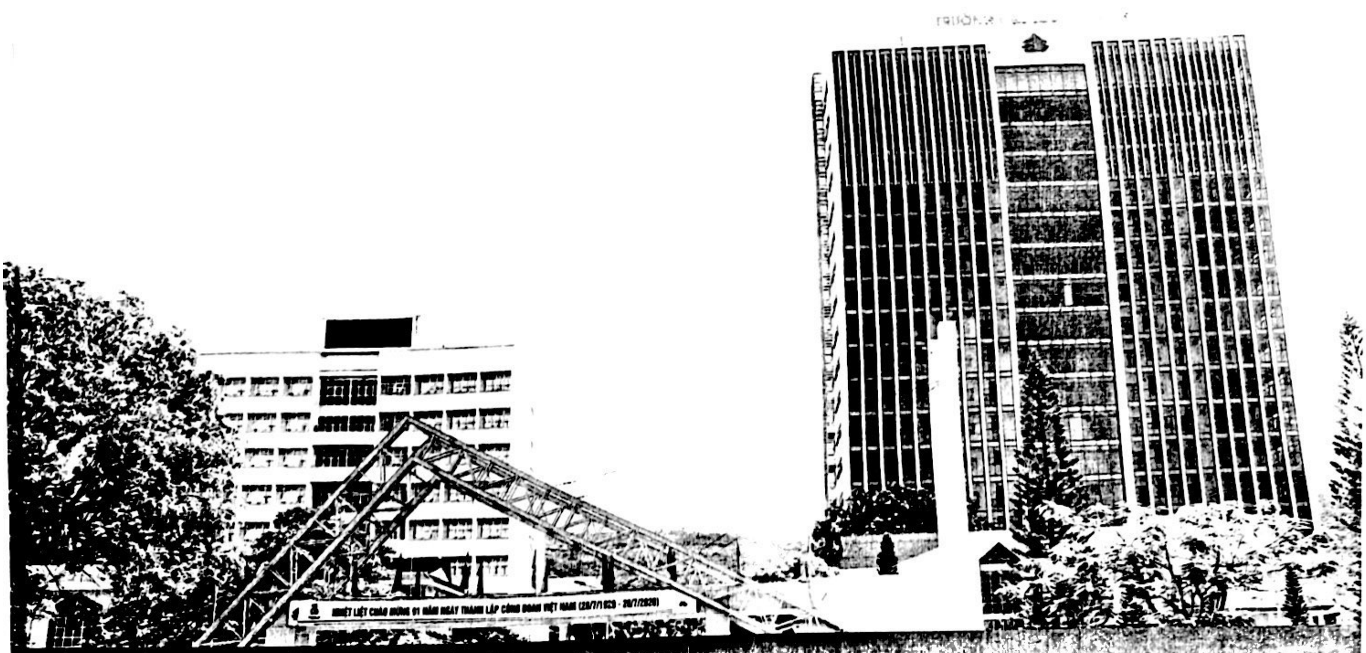
@

MỘT SỐ VẤN ĐỀ C...
...ỦA CÔNG NGHỆ ...
VÀ TRUYỀN THÔ...

Chủ đề: Smart City

Viết chung

TOÀN VĂN CÁC BÁO CÁO



NHÀ XUẤT BẢN KHOA HỌC VÀ KỸ THUẬT

Multi-view fuzzy co-clustering algorithm for high-dimensional data classification

Le Thi Cam Binh
Hanoi University of Culture
Hanoi, Vietnam
cambinhlt@gmail.com

Pham Van Nha
Academy of Military Science and Technology
Hanoi, Vietnam
nhapv@mod.gov.vn

***Abstract:* Multi-view clustering considers the diversity of different views and fuses these views to produce a more accurate and robust partition than single-view clustering. In this paper, we proposes a multi-view fuzzy co-clustering algorithm for high-dimensional data classification. We call MvFCoC algorithm. The proposed algorithm is demonstrated through experiments on benchmark data sets. The experimental results show that the clustering quality is better by evaluating using validity indexes in comparison with previous methods.**

***Keywords:* Multi-view; Multi-view fuzzy clustering; Fuzzy co-clustering; High-dimensional data.**

I. INTRODUCTION

Fuzzy co-clustering is an unsupervised learning technique. Fuzzy co-clustering was published in 2001 [1] with the aim to replace traditional clustering techniques for many-feature data clustering. Recently, some works have applied fuzzy co-clustering to deploy practical applications. Works in [2]-[5] have applied fuzzy co-clustering to categorize documents, where paragraphs are dealt with data objects and phrases or keywords be considered as features of data objects. The works in [6, 7] have applied fuzzy clustering to segment color images, where, the pixels are treated as data objects, and the three color components R, G, and B are considered as features. The works in [8]-[11] used fuzzy co-clustering to classify the multi-dimensional data and the multi-spectral image in applications related to the environment and the earth. In particular, the pixels are considered as data objects and the spectral components of the image are considered as features. In order to further improve the quality of clustering in handling uncertainty and overlapping clusters, we have integrated some advanced techniques of interval-valued fuzzy sets with co-clustering. We call it interval-valued fuzzy co-clustering [9]. In order to exploit the basic advantages of fuzzy co-clustering in many-feature data mining, we have applied fuzzy co-clustering to process the hyperspectral image data in remote sensing and medical applications [12, 13]. Where, hyperspectral pixels have not only three color components as color images but hundreds of spectral components as many-feature data. However, fuzzy co-clustering still exists the technical and application challenges. Furthermore, co-clustering were only concerned with single-view and limited-size datasets. These limitations have opened up valuable opportunities for us to develop fuzzy co-clustering.

With the rapid development of multimedia technology, multi-view data began to appear in large numbers, meaning that the same objects were described from different perspectives. Therefore, the applications of multi-view learning [14–16] in clustering problems produce many novel multi-view clustering algorithms for multi-view data.

The most primitive multi-view clustering is to simply stitch all data features and then use them for clustering. This obviously does not take advantage of the information complementarity between different views but also does not have any interpretability. Due to the simplicity of K-means, many researchers have tried to construct models for multi-view data based on its theory, including works in [17, 18]. Cleuziou et al.[19] developed a multi-view clustering approach by using the conventional FCM framework (Co-FKM algorithm), which has become the basis of the collaborative multi-view fuzzy clustering method.

In this paper, we are motivated by the advantages of fuzzy co-clustering in many-feature clustering and multi-view. We propose a the Multi-view Fuzzy Co-clustering algorithm (MvFCoC algorithm). In MvFCoC algorithm, the data is projected under different views, at the same time two weighting strategies are applied to the object and feature. Experiments on real-world datasets show the effectiveness of the proposed approach.

The remainder of this paper is organized as follows: The main works related to the paper are presented in Section II. Section III proposes the Multi-view Fuzzy Co-clustering algorithm. Section IV is the experimental results, and finally concludes some contributions of this paper in Section V.

II. RELATED WORK

In this section, we briefly presents the feature reduction based Co-FKM algorithm and fuzzy co-clustering algorithm.

A. Related notation

In this study, we consider a set of data $X = \{x_1, x_2, \ldots, x_i, \ldots, x_N\}$ with N samples, T views and C clusters ($2 \le C \le N$). Where $1 \le i \le N$. For $1 \le t \le T$ and $1 \le c \le C$ in the *t-th* view, D_t denotes the number of the features and $x_{ij,t}$ denotes the *j-th* feature value of the *i-th* sample in the *t-th* view with $j \in D_t$. In the fuzzy clustering framework, to cluster the sample set X into C classes, $u_{ci,t}$ denotes the fuzzy membership degree of x_i to cluster c in view t, $U = \{u_{ci,t}\}$ is called the object membership function matrix in t-th view. $v_{cj,t}$ denotes the fuzzy membership degree of x_{ij} to cluster c in view t, $V = \{v_{cj,t}\}_{c=\overline{1,C}, j=\overline{1,D}, t=\overline{1,T}}$ is called the feature membership function matrix in *t-th* view. $p_{c,t}$ denotes the center of cluster c in view t, $d_{ci,t}$ denotes the Euclidean distance between x_i and p_c in view t, $p_{cj,t}$ denotes the *j-th* feature value of *c-th* cluster center in *t-th* view.

B. Multi-view Fuzzy Clustering Algorithm

A multi-view fuzzy clustering algorithm was proposed in 2009 by Cleuziou et al. They call a collaborative multi-view fuzzy clustering algorithm (Co-KFM algorithm). In Co-FKM algorithm, each view has a specific partition and a penalty term is introduced to reduce the inconsistency between partitions from different views. The objective function is defined as the minimization of the distance between samples and cluster centers in each view while penalizing the disagreement between any pairs of views.

$$J_{Co-FKM} = \sum_{t=1}^{T}\sum_{c=1}^{C}\sum_{i=1}^{N} u_{ci,t} d_{ci,t} + \eta \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{c=1}^{C}\sum_{i=1}^{N} (u_{ci,t'} - u_{ci,t}) d_{ci,t'} \quad (1)$$

Where, m is a fuzzy coefficient, η is a parameter used to control the penalty associated with the disagreement.

C. Fuzzy Co-Clustering algorithm

Fuzzy co-clustering is a variant algorithm of traditional fuzzy clustering that was developed to replace fuzzy clustering in the complex many-feature data processing. The basic difference of fuzzy co-clustering compared to traditional fuzzy clustering is that fuzzy co-clustering considers data simultaneously by the weight of the data object through the object function membership and feature function membership. Meanwhile, fuzzy clustering only considers the weight of data objects through only object function membership. Therefore, the fuzzy co-clustering is suitable for more many-feature data.

Let T_u and T_v be weights that indicate fuzzy level. The objective function J_{FCoC} is expressed in the following form (3),

$$J_{FCoC} = \sum_{c=1}^{C}\sum_{i=1}^{N}\sum_{j=1}^{D} u_{ci} v_{cj} d_{cij} + T_U \sum_{c=1}^{C}\sum_{i=1}^{N} u_{ci} \log u_{ci} + T_V \sum_{c=1}^{C}\sum_{j=1}^{D} v_{cj} \log v_{cj} \quad (2)$$

Where, C is the number of clusters which is a known input parameter for each data set, N is the number of data objects and D is the number of features of data.

III. THE PROPOSED MULTI-VIEW FUZZY CO-CLUSTERING ALGORITHM

In this section, we present a MvFCoC algorithm. The objective function of the MvFCoC algorithm is modified as follows,

$$\begin{aligned} J_{MvFCoC} = & \sum_{t=1}^{T}\sum_{c=1}^{C}\sum_{n=1}^{N}\sum_{j=1}^{D} u_{ci,t} v_{cj,t} d_{cij,t} + \eta_1 \frac{1}{T-1} \sum_{t'=1,t'\neq t}^{T} \sum_{c=1}^{C}\sum_{n=1}^{N}\sum_{j=1}^{D} (u_{ci,t'} - u_{ci,t}) d_{cij,t} + \\ & \eta_2 \frac{1}{T-1} \sum_{t'=1,t'\neq t}^{T} \sum_{c=1}^{C}\sum_{n=1}^{N}\sum_{j=1}^{D} (v_{cj,t'} - v_{cj,t}) d_{cij,t} + T_u \sum_{t=1}^{T}\sum_{c=1}^{C}\sum_{i=1}^{N} u_{ci,t} \log u_{ci,t} + T_v \sum_{t=1}^{T}\sum_{c=1}^{C}\sum_{j=1}^{D} v_{cj,t} \log v_{cj,t} \end{aligned} \quad (3)$$

To get optimal clustering results, the objective function (3) is minimized subject to the following constraints (4),

$$\sum_{c=1}^{C} u_{ci,t} = 1, u_{ci,t} \in [0,1], \forall i = \overline{1,N}, \forall t = \overline{1,T} \qquad (4)$$
$$\sum_{j=1}^{D} v_{cj,t} = 1, v_{cj,t} \in [0,1], \forall c = \overline{1,C}, \forall t = \overline{1,T}$$

Where, η_1 and η_2 is parameters used to control the penalty associated with the disagreement. The terms $\eta_1 \frac{1}{T-1} \sum_{t'=1,t'\neq t}^{T} \sum_{c=1}^{C} \sum_{i=1}^{N} \left(u^m_{ci,t'} - u^m_{ci,t}\right) d_{ci,t}$ and $\eta_2 \frac{1}{T-1} \sum_{t'=1,t'\neq t}^{T} \sum_{c=1}^{C} \sum_{n=1}^{N} \sum_{j=1}^{D} \left(v_{cj,t'} - v_{cj,t}\right) d_{cij,t}$ are disagreement terms, which can be considered as the divergence between partitions from different views, i.e., the lower the value of $\left(u_{ci,t'} - u_{ci,t}\right)$, the lower the divergence between the object membership functions in views, the lower the value of $\left(v_{ci,t'} - v_{ci,t}\right)$, the lower the divergence between the feature membership functions in views.

To minimize the objective function J_{MvFCoC} with constraints are given by (4), we construct an objective function with Lagrange coefficients $\lambda_c \left(c = \overline{1,C}\right)$ for constraint $\sum_{c=1}^{C} u_{ci,t} = 1$, $\gamma_j \left(j = \overline{1,D}\right)$ for $\sum_{j=1}^{D} v_{cj,t} = 1$, we obtain:

$$J_{MvFCoC} = \sum_{t=1}^{T} \sum_{c=1}^{C} \sum_{n=1}^{N} u_{ci,t} v_{cj,t} d_{ci,t} + \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{c=1}^{C} \sum_{n=1}^{N} \left(u_{ci,t'} - u_{ci,t}\right) d_{ci,t} +$$
$$\eta_2 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{c=1}^{C} \sum_{n=1}^{N} \sum_{j=1}^{D} \left(v_{cj,t'} - v_{cj,t}\right) d_{cij,t} + T_u \sum_{t=1}^{T} \sum_{c=1}^{C} \sum_{i=1}^{N} u_{ci,t} \log u_{ci,t} + T_v \sum_{t=1}^{T} \sum_{c=1}^{C} \sum_{j=1}^{D} v_{cj,t} \log v_{cj,t} + \qquad (5)$$
$$\sum_{c=1}^{C} \lambda_c \left(u_{ci,t} - 1\right) + \sum_{j=1}^{D} \gamma_j \left(v_{cj,t} - 1\right)$$

The MvFCoC algorithm is resolved by the following steps. Firstly, we calculate the membership function U by fixing V and P, then minimizing the objective function (5) according to U, and taking derivatives of objective function with respect to the fuzzy object memberships and setting them to zero, we obtain,

$$\frac{\partial J_{MvFCoC}}{\partial u_{ci,t}} = \sum_{j=1}^{D} v_{cj,t} d_{ci,t} + \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{j=1}^{D} \left(u_{ci,t'} - 1\right) d_{ci,t} + T_u (\log u_{ci,t} + 1) + \lambda_c = 0 \qquad (6)$$

By some algebraic simplifications in Eq. (7), we obtain,

$$u_{ci,t} = \frac{e^{\frac{-\sum_{j=1}^{D} v_{cj,t} d_{cij,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{j=1}^{D} \left(u_{ci,t'} - 1\right) d_{ci,t}}{T_u} - 1}}{e^{\frac{\lambda_c}{T_u}}} \qquad (7)$$

Because of the constraint $\sum_{c=1}^{C} u_{ci,t} = 1$, the Lagrange multiplier λ_c are eliminated as

$$\sum_{c=1}^{C} u_{ci} = \sum_{c=1}^{C} \frac{e^{\frac{-\sum_{j=1}^{D} v_{cj,t} d_{cij,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{j=1}^{D} \left(u_{ci,t'} - 1\right) d_{ci,t}}{T_u} - 1}}{e^{\frac{\lambda_c}{T_u}}} = 1 \qquad (8)$$

$$\sum_{c=1}^{C} u_{ci} = \frac{\sum_{c=1}^{C} e^{\frac{-\sum_{j=1}^{D} v_{cj,t} d_{cij,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{j=1}^{D} \left(u_{ci,t'} - 1\right) d_{ci,t}}{T_u} - 1}}{e^{\frac{\lambda_c}{T_u}}} = 1 \qquad (9)$$

$$\Rightarrow e^{\frac{\lambda_c}{T_u}} = \sum_{c=1}^{C} e^{\frac{-\sum_{j=1}^{D} v_{cj,t} d_{cij,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t'\neq t)} \sum_{j=1}^{D} \left(u_{ci,t'} - 1\right) d_{ci,t}}{T_u} - 1} \qquad (10)$$

By using Eq. (8) in Eq. (10), the closed-form solution for the optimal object membership function is obtained as,

$$u_{ci,t} = \frac{e^{\frac{-\sum_{j=1}^{D} v_{cj,t} d_{cij,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t' \neq t)} \sum_{j=1}^{D} (u_{ci,t'} - 1) d_{ci,t}}{T_u} - 1}}{\sum_{k=1}^{C} e^{\frac{-\sum_{j=1}^{D} v_{kj,t} d_{ki,t} - \eta_1 \frac{1}{T-1} \sum_{t'=1}^{T(t' \neq t)} \sum_{j=1}^{D} (u_{ki,t'} - 1) d_{ki,t}}{T_u} - 1}} \quad (11)$$

In the similar way to U, to find the optimal fuzzy feature memberships V, taking derivatives of objective function with respect to the fuzzy feature memberships and setting them to zero, we obtain,

$$\frac{\partial J_{MvFCoC}}{\partial v_{cj,t}} = \sum_{c=1}^{C} u_{ci,t} d_{cij,t} + \eta_2 \frac{1}{T-1} \sum_{t'=1}^{T(t' \neq t)} \sum_{j=1}^{D} \left(v_{cj,t'} - 1\right) d_{cij,t} + T_v (\log v_{cj,t} + 1) + \gamma_j = 0 \quad (12)$$

By some algebraic simplifications in Eq. (12), we reach,

$$v_{cj,t} = \frac{e^{\frac{-\sum_{j=1}^{D} u_{ci,t} d_{ci,t} - \eta_2 \frac{1}{T-1} \sum_{t'=1}^{T(t' \neq t)} \sum_{j=1}^{D} (v_{cj,t'} - 1) d_{cij,t}}{T_v} - 1}}{\sum_{k=1}^{C} e^{\frac{-\sum_{j=1}^{D} u_{ki,t} d_{ki,t} - \eta_2 \frac{1}{T-1} \sum_{t'=1}^{T(t' \neq t)} \sum_{j=1}^{D} (v_{kj,t'} - 1) d_{kij,t}}{T_v} - 1}} \quad (13)$$

Before finding the cluster centroids P the square of Euclidean distance $\|x_{ij,t} - p_{cj,t}\|^2$ is defined as $\|x_{ij,t} - p_{cj,t}\|^2 = (x_{ij,t} - p_{cj,t})^2 = x_{ij,t}^2 - 2x_{ij,t} p_{cj,t} + p_{cj,t}^2$, we obtain,

$$\frac{\partial J_{FRFCoC}}{\partial p_{cj,t}} = v_{cj,t} \sum_{i=1}^{N} u_{ci,t} x_{ij,t} - v_{cj,t} p_{cj,t} \sum_{i=1}^{N} u_{ci,t} = 0 \quad (14)$$

By some algebraic simplifications in Eq. (14), we reach,

$$p_{cj,t} = \frac{\sum_{i=1}^{N} u_{ci,t} x_{ij,t}}{\sum_{i=1}^{N} u_{ci,t}} \quad (15)$$

In MvFCoC algorithm, the idea of fuzzy co-clustering ensemble is adopted to combine individual view fuzzy partitions $u_{ci,t}$, $v_{cj,t}$ and obtain the global clustering result $\bar{u}_{ci}$, $\bar{v}_{cj}$. The consensus function is defined as the geometric mean of $u_{ci,t}$, $v_{cj,t}$ for each view and expressed as follows:

$$\bar{u}_{ci} = \sqrt[T]{\prod_{t=1}^{T} u_{ci,t}} \quad (16)$$

$$\bar{v}_{cj} = \sqrt[T]{\prod_{t=1}^{T} v_{cj,t}} \quad (17)$$

MvFCoC improved the performance of multi-view clustering, and MvFCoC considered that each view and each feature contributed equally to clustering, which may decrease the clustering performance when the views and features had different importance.

The MvFCoC algorithm diagram consists of the learning processes of membership function matrixs U, V that are shown as Algorithm 1.

Algorithm 1. MvFCoC algorithm

Input: T data sets $X_t = \{x_{i,t}, x_{i,t} \in R^{D_t}\}$, $i = \overline{1, N_t}$, the number of clusters C.

Output: Clustering result.

1. Initialize parameters T_u, T_v, η_1, η_2, ε , η_1, η_2, ε, the maximum number of iterations τ_{max}.

2. Initialize $u_{ci,t}$ satisfying Eq. (4).

3. τ=1.

4. REPEAT

5. Update $p_{cj,t}$ using (13).

6. Calculate $d_{ci,t}$ and d_{cij} using Eq. (1).

7. Update $v_{cj,t}$ using (11).

8. Update $u_{ci,t}$ using (9).

9. Update $\bar{u}_{ci}$ using (14).

10. Update $\bar{v}_{cj}$ using (15).

11. τ=τ+1.

12. UNTIL $(\|\bar{U}(\tau)\| - \|\bar{U}(\tau+1)\| < \varepsilon)$ **or** $(\tau > \tau_{max})$

IV. EXPERIMENT RESULTS

In this section, we evaluate our proposed algorithm on the real-world data. We performed experiments on Windows 7 of HP Elitebook 8560W, Core i7-2670QM, 8 GB RAM, NVIDIA Quadro 2000M, and C#.Net development environment. Experiments are conducted on four real-world benchmark datasets. The clustering results are evaluated by comparing the obtained label of each instance with the provided label by the dataset. Two metrics, the accuracy and the recall index and precision index are used to measure the clustering performance [20,21].

Table 1. Summary of the benchmark data sets

Dataset	Instances	Views	Clusters	D_1	D_2	D_3	D_4	D_5	D_6
100leaves	1600	3	100	64	64	64	-	-	-
Handwritten digit (HW)	2000	6	10	216	76	64	6	240	47
Image segmentation (IS)	2310	2	7	19	19	-	-	-	-
PEMS-SF	440	3	7	138672	4789	8900	-	-	-

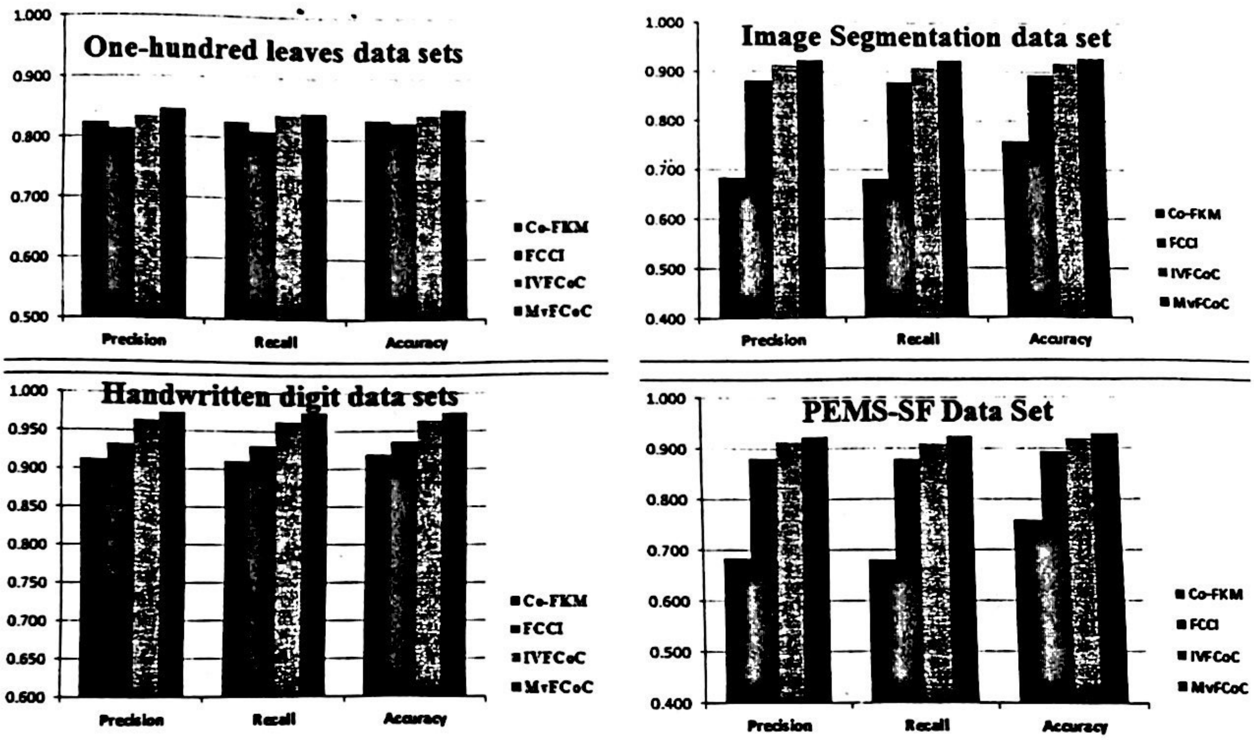


Figure 1. Clustering performance comparison of each algorithm in terms of Precíion, Recall and Accuracy on four real-world datasets.

To compare the effects of different clustering algorithms on multi-view clustering, MvFCoC is fed with other three clustering algorithms: The multi-vew fuzzy clustering algorithm [19], the fuzzy co-clustering - FCCI algorithm [7] and the interval-valued fuzzy co-clustering - IVFCoC algorithm [9].

Table 2. Clustering performance comparison of each algorithm in terms of Precíion, Recall and Accuracy on four real-world datasets

Data sets	Algorithms	Pre	Rec	Acc	Time
100leaves	Co-FKM	0.824	0.827	0.832	47
	FCCI	0.815	0.812	0.828	25
	IVFCoC	0.835	0.838	0.842	38
	MvFCoC	**0.848**	**0.843**	**0.852**	**15**
HW	Co-FKM	0.913	0.912	0.920	35
	FCCI	0.933	0.932	0.937	20
	IVFCoC	0.964	0.964	0.965	32
	MvFCoC	**0.974**	**0.974**	**0.975**	**16**
IS	Co-FKM	0.684	0.682	0.759	15
	FCCI	0.882	0.881	0.894	2
	IVFCoC	0.913	0.909	0.920	10
	MvFCoC	**0.924**	**0.923**	**0.929**	**3**
PEMS-SF	Co-FKM	0.737	0.727	0.788	1415
	FCCI	0.946	0.946	0.949	1066
	IVFCoC	0.965	0.964	0.965	1528
	MvFCoC	**0.980**	**0.980**	**0.981**	**921**

Tables 2 and Fig. 1 show the comparison results which is quantified by Precision, Recall and Accuracy. Where the values in bold indicate the best performance results among the four algorithms. In Table II, time is measured in seconds. From Tables II and Fig. 1, we observe that MvFCoC algorithm obtained Precision, Recall and Accuracy indexes higher than Co-KFM, FCoC and IVFCoC algorithms. This means that the MvFCoC algorithm achieves higher classification accuracy than Co-KFM, FCoC and IVFCoC algorithms. In addition, the classification time of MvFCoC algorithm is lower than Co-KFM, FCoC and IVFCoC algorithms.

To clarify the effectiveness of the MvFC algorithm, we have a look at a few comparisons below. For the Co-KFM algorithm, the MvFCoC algorithm is based on the original FCoC algorithm which is dedicated to processing data with many features. Therefore, the algorithm MvFCoC will classify more accurately and faster than the Co-KFM algorithm. For the FCoC algorithm which is the original algorithm of the MvFCoC algorithm. There, the FCoC algorithm processes each data view and then aggregates their results into the general result. The MvFCoc algorithm concurrently processes views, then interacts between views using Eq. (16) and Eq. (17). Therefore, the algorithm MvFCoC will get more accurate and faster classification results. For IVFCoC algorithm which has higher computational complexity than the FCoC algorithm and the same operating mechanism as the FCoC algorithm. Therefore, the IVFCoC algorithm will process the data more slowly and less accurately than the MvFCoC algorithm.

V. CONCLUSION

In this paper, we developed a multi-view fuzzy co-clustering algorithm built upon a combination of a fuzzy co-clustering algorithm and multi-view for calculating the weight of features and objects in different views. MvFCoC algorithm is guided by the new objective function with the feature-weighted entropy and the object-weighted entropy and an unsupervise learning schema. Experimental results have proved that the proposed algorithm is more effective than some previously proposed algorithms in some considered cases.

This paper just experimented and presented some results on small datasets that have not yet exploited the real potential of multi-view and fuzzy co-clustering in big data processing. In the future, we will focus on experiments on larger datasets and discuss in-depth the working mechanism of the proposed algorithms.

REFERENCES

[1] C.H. Oh, K. Honda, H. Ichihashi, Fuzzy clustering for categorical multivariate data, Congress and 20th NAFIPS International Conference, pp. 2154-2159, 2001.

[2] Y. Yan, L. Chen, W. C. Tjhi, Fuzzy semi-supervised co-clustering for text documents, Fuzzy Sets and Systems, Vol. 215, pp. 74-89, 2013.

[3] W. C. Tjhi, L. Chen, Possibilistic fuzzy co-clustering of large document collections, Pattern Recognition, Vol. 40(12), pp. 3452-3466, 2007.

[4] C. Laclau, M. Nadif, Hard and fuzzy diagonal co-clustering for document-term partitioning, Neurocomputing, Vol. 19312, pp. 133-147, 2016.

[5] T. Kondo, Y. Kanzawa, Comparison of fuzzy co-clustering methods in collaborative filtering-based recommender system, Springer: Lecture Notes in Computer Science, pp. 103-116, 2017.

[6] Y. Liu, S. Wu, Z. Liu, H. Chao, A fuzzy co-clustering algorithm for biomedical data, PLoS One, Vol. 12(4), 2017: doi:10.1371/journal.pone.0176536.

[7] M. Hanmandlua, O.P. Verma, S.S., V.K. Madasu, Color segmentation by fuzzy co-clustering of chrominance color features, Neurocomputing, Vol. 120, pp. 235-249, 2013.

[8] K. Kummamuru, A. Dhawale, R. Krishnapuram, Fuzzy Co-clustering of Documents and Keywords, FUZZ-IEEE 2003, Vol. 2, pp. 772-777, 2003.

[9] V.N. Pham, N.T. Long, W. Pedrycz, Interval-valued fuzzy set approach to fuzzy co-clustering for data classification, Knowledge-Based Systems, Vol. 107, pp. 1-13, 2016.

[10] W. C. Tjhi, L. Chen,A heuristic-based fuzzy co-clustering algorithm for categorization of high-dimensional data, Fuzzy Sets and Systems, Vol. 159, pp. 371-389, 2008.

[11] Y. Liu , J. Chen, H. Chao, A Fuzzy Co-Clustering Algorithm via Modularity Maximization, Mathematical Problems in Engineering, 2018: https://doi.org/10.1155/2018/3757580.

[12] V.N. Pham, L.T. Ngo, D.T. Nguyen, Feature-Reduction Fuzzy Co-Clustering algorithm for hyperspectral image clustering, FUZZ-IEEE 2017, pp. 1-6, 2017.

[13] V.N. Pham, L.T. Pham, T.D. Nguyen, L.T. Ngo, A new cluster tendency assessment method for fuzzy co-clustering in hyperspectral image analysis, Neurocomputing, Vol. 30713, pp. 213-226, 2018.

[14] C. Xu, D. Tao, C. Xu, A survey on multi-view learning, 2013, ArXiv: Learning.

[15] S. Sun, A survey of multi-view machine learning, Neural Comput. Appl. 23 (2013) 2031–2038.

[16] J. Zhao, X. Xie, X. Xu, S. Sun, Multi-view learning overview: recent progress and new challenges, Inf. Fusion 38 (2017) 43–54.

[17] D. Arthur, S. Vassilvitskii, k-means++: The advantages of careful seeding, in: Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 2007, pp. 1027–1035.

[18] X. Shen, W. Liu, I. Tsang, F. Shen, Q.S. Sun, Compressed k-means for large-scale clustering, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2017, pp. 2527–2533.

[19] G. Cleuziou, M. Exbrayat, L. Martin, and J. H. Sublemontier, CoFKM: A centralized method for multipleview clustering, in Proc. 9th IEEE Int. Conf. Data Mining, Miami, FL, USA, 2009, pp. 752–757.

[20] D. Cai, X. He, J. Han, Document clustering using locality preserving indexing, IEEE Trans. Knowl. Data Eng, Vol. 17(12), 2005, pp. 1624–1637.

[21] D.L. Olson, D. Delen, Advanced Data Mining Techniques, Springer ISBN 3-540-76916-1, 1st edition, 2008, page 138.

[22] C. Mallah, J. Cope, J. Orwell, Plant leaf classification using probabilistic integration of shape, texture and margin features, Proceedings of the IASTED International Conference Signal Processing, Pattern Recognition and Applications, 2013, pp. 279–286.

[23] M. van Breukelen, R.P.W. Duin, D.M.J. Tax, J.E. den Hartog, Handwritten digit recognition by combined classifiers, Kybernetika 34 (4) (1998) 381–386.

[24] [24] C. Blake and C. J. Merz, UCI repository of machine learning databases, http://archive.ics.uci.edu/ml/index.php, 1998.

[25] M. Cuturi, Fast Global Alignment Kernels, Proceedings of the Intern. Conference on Machine Learning 2011.

[26] I.T. Jolliffe, Principal Component Analysis, second edition (Springer), 2002.

[27] E. Bingham, H. Mannila, Random projection in dimensionality reduction: Applications to image and text data, Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2001, pp. 245–250.