2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Incremental Algorithms for Fuzzy Co-clustering of Very Large Cooccurrence Matrix

Katsuhiro Honda, Daiji Tanaka and Akira Notsu

***Abstract*— Handling very large data is an important issue in FCM-type clustering and several incremental algorithms have been proved to be useful in FCM clustering. In this paper, the incremental algorithms are extended to fuzzy co-clustering of cooccurrence matrices, whose goal is to simultaneously partition objects and items considering their cooccurrence information. Single pass and online approaches are applied to fuzzy clustering for categorical multivariate data (FCCM) and fuzzy CoDoK, which try to maximize the aggregation degrees of co-clusters adopting entropy-based and quadratic-based membership fuzzifications. Several experimental results demonstrate the applicability of the incremental approaches to fuzzy co-clustering algorithms.**

I. Introduction

Handling very large data is an important issue in various real world data analysis and many extended algorithms for applying fuzzy c-means (FCM) clustering [1] to large data have been proposed. Havens *et al.* [2] introduced a random sampling approach of *random sampling plus extension* FCM (rseFCM) and two incremental approaches of *single-pass* FCM (spFCM) [3] and *online* FCM (oFCM) [4], and extended the approaches to kernel fuzzy c-means (KFCM) [5]. rseFCM is a simple process of reducing the number of objects through random sampling but clustering quality might be significantly degraded. In the incremental approaches, the whole data set was randomly partitioned into data chunks and FCM clustering was performed with a chunk composed of a small subset at a time. The cluster structures extracted with each chunk were inherited in single-pass processes or merged in online (parallel) processes. In FCM clustering, each cluster is represented by its prototypical cluster center and the cluster centers are identified with additional *objects* in later operations.

In this paper, the sampling and incremental approaches are extended to fuzzy co-clustering models. Co-clustering achieves the task of simultaneous partitioning of objects and items considering their cooccurrence information. The data to be clustered is given by a cooccurrence matrix, whose elements represent the degree of coocurrence of object-item pairs, e.g., the number of appearance of keywords in documents and the frequencies of purchase of items by customers. Fuzzy clustering for categorical multivariate data (FCCM) [6] is an FCM-type fuzzy co-clustering algorithm, which simultaneously estimates two types of fuzzy memberships so that the degree of co-cluster aggregations of objects and items is maximized. The fuzzy memberships of objects are designed in a similar concept to FCM, in which they represent exclusive assignment of objects to clusters such that the sum of memberships w.r.t. clusters are constrained to be one. On the other hand, the fuzzy memberships of items are mostly responsible for representing the mutual typicalities in each cluster such that the sum of memberships w.r.t. items are constrained to be one in each cluster. In FCCM, the fuzzification of memberships is achieved by the entropy-based method [7], [8]. Fuzzy CoDoK [9] is another fuzzy co-clustering model, in which the fuzzification of memberships is achieved by the quadratic-based method [10]. The quadratic-based fuzzification method is useful for handling large data sets avoiding overflow in membership calculation while it needs a trick for deriving non-negative memberships.

Because FCCM and fuzzy CoDoK are prototype-less co-clustering algorithms, the conventional incremental approaches for FCM, in which prototypes are used as the messengers of cluster structures, cannot be directly applied to them. In this paper, item memberships are adopted for taking over the cluster characteristics in the incremental processes. Although item memberships cannot be identified with virtual objects unlike prototypical cluster centers in FCM, they have structural information of co-clusters. Then, they are utilized as additional virtual objects after normalization.

The remaining parts of this paper are organized as follows: Section II gives a brief review on the background of this research. The incremental procedures for fuzzy co-clustering of very large cooccurrence matrices are proposed in Section III. The applicability of the proposed procedures is demonstrated in Section IV and the summary conclusions are given in Section V.

II. Background

A. FCM and weighted FCM

FCM [1] is a basic method of fuzzy clustering, in which each cluster is represented by its prototypical cluster center. Assume that we have n objects, which are characterized by p attributes: $\boldsymbol{x}_i = (x_{i1}, \dots, x_{ip})^\top$, $i = 1, \dots, n$, and the goal is to partition the objects into C fuzzy clusters, whose cluster centers are $\boldsymbol{b}_c$, $c = 1, \dots, C$. Cluster assignments of object i are given by fuzzy memberships $u_{ci} \in [0, 1]$, $c = 1, \dots, C$ such that $\sum_{c=1}^{C} u_{ci} = 1$. The objective function to

Katsuhiro Honda, Daiji Tanaka and Akira Notsu are with the Graduate School of Engineering, Osaka Prefecture University, Osaka, Japan (email: {honda, notsu}@cs.osakafu-u.ac.jp).

This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology, Japan, through a Grant-in-Aid for Scientific Research (#23500283, #26330281).

be minimized is defined as the sum of within-cluster errors:

$$L_{fcm} = \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}^{m} ||\boldsymbol{x}_i - \boldsymbol{b}_c||^2. \quad (1)$$

m ($m > 1$) is an exponential weight for membership fuzzification and the model is reduced to the crisp k-Means [11] if $m = 1$. The clustering algorithm is an iterative process of membership estimation and cluster center updating.

Although the conventional FCM considers that each object is equally important in the clustering solution, it is also the case that objects have different relative importance. In such cases, FCM clustering should be performed considering the relative importance, which is given by α_i for object i. Introducing the weight α_i, the objective function of weighted FCM (wFCM) is defined as:

$$L_{wfcm} = \sum_{c=1}^{C}\sum_{i=1}^{n} \alpha_i u_{ci}^{m} ||\boldsymbol{x}_i - \boldsymbol{b}_c||^2. \quad (2)$$

In wFCM, the objects with larger α_i have higher responsibility in cluster center estimation while fuzzy memberships are estimated in the same manner with FCM.

B. Sampling and Incremental approaches in FCM

In handling very large data sets composed of a huge number of objects, it is difficult to process all objects in the batch algorithm. In order to reduce the computational overload, the FCM algorithm should be applied to a subset randomly extracted from the whole data sets.

Random sampling plus extension FCM (rseFCM) is a simple approach, in which the conventional FCM algorithm is performed with only a subset randomly extracted objects, and then, the derived cluster centers are extended for remaining objects in membership estimation. Although rseFCM has its computational efficiency, the cluster quality may be significantly degraded because a large part of data objects are removed from cluster structure estimation.

In order to utilize whole data objects with low computational efforts, several incremental algorithms have been proposed, in which FCM algorithms are performed with a chunk composed of a small subset at a time. The whole data set is randomly partitioned into data chunks and the cluster structures extracted with each chunk are inherited in single-pass processes or merged in online (parallel) processes. In FCM clustering, each cluster is represented by its prototypical cluster center and the cluster centers are identified with additional *objects* in later operations.

In *single-pass* FCM (spFCM) [3], FCM or wFCM are sequentially performed in a single pass, where the previous cluster centers are added into the next chunk as virtual objects considering their responsibility weights of how many objects they are supported by. Then, the cluster centers derived from the final chunk can be representative cluster centers of whole data set while the fuzzy memberships should be re-calculated for all objects using the final cluster centers.

Online FCM (oFCM) [4] is a two stage algorithm. In the first stage, FCM is separately performed with every data chunk in parallel. Then, in the second stage, the cluster centers estimated from each data chunk are gathered into a new virtual data set, and wFCM is applied considering their responsibility weights.

In these incremental algorithms, cluster centers play an important role of the messengers of cluster structures and are identified with *virtual objects* because they are also defined in the same data space with vector observations.

In the remaining parts of this paper, the incremental approaches are modified for performing prototype-less co-clustering of very large data sets.

III. Fuzzy Co-clustering and Incremental Algorithms

Assume that we have a cooccurrence matrix $R = \{r_{ij}\}$ on objects $i = 1, \ldots, n$ and items $j = 1, \ldots, p$, in which r_{ij} represents the degree of cooccurrence of item j with object i. The goal of co-clustering is to simultaneously partition objects and items by estimating two types of fuzzy memberships. The fuzzy memberships of objects u_{ci} are designed in a similar concept to FCM, in which they represent exclusive assignment of objects to clusters such that $\sum_{c=1}^{C} u_{ci} = 1$. On the other hand, in order to avoid trivial solutions, the fuzzy memberships of items w_{cj} are mostly responsible for representing the mutual typicalities in each cluster such that $\sum_{j=1}^{p} w_{cj} = 1$.

A. FCCM

Oh *et al.* [6] proposed the FCM-type co-clustering model, which is called FCCM, by modifying the FCM algorithm for handling cooccurrence information, where the cluster aggregation degree of each cluster is maximized:

$$\begin{aligned} L_{fccm} \quad = \quad & \sum_{c=1}^{C}\sum_{i=1}^{n}\sum_{j=1}^{p} u_{ci} w_{cj} r_{ij} \\ & -\lambda_u \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci} \log u_{ci} \\ & -\lambda_w \sum_{c=1}^{C}\sum_{j=1}^{p} w_{cj} \log w_{cj}. \end{aligned} \quad (3)$$

In FCCM, the entropy-based fuzzification method [7], [8] were adopted instead of the standard approach in FCM because the exponential weight m in FCM can work only in the minimization framework of positive objective functions. λ_u and λ_w play a similar role to m, where larger λ brings fuzzier partitions while small λ brings crisp partitions.

The clustering algorithm is an iterative process of updating u_{ci} and w_{cj} using the following rules:

$$u_{ci} = \frac{\exp\left(\lambda_u^{-1} \sum_{j=1}^{p} w_{cj} r_{ij}\right)}{\sum_{\ell=1}^{C} \exp\left(\lambda_u^{-1} \sum_{j=1}^{p} w_{\ell j} r_{ij}\right)}, \quad (4)$$

and

$$w_{cj} = \frac{\exp\left(\lambda_w^{-1}\sum_{i=1}^{n} u_{ci} r_{ij}\right)}{\sum_{\ell=1}^{p}\exp\left(\lambda_w^{-1}\sum_{i=1}^{n} u_{ci} r_{i\ell}\right)}. \quad (5)$$

B. Fuzzy CoDoK

In the entropy-based FCCM algorithm using the exponential-type component function, membership calculation may be suffered from overflow when the numbers of objects and/or items are large. Fuzzy CoDoK [9] is a modified algorithm of FCCM, in which the fuzzification of memberships is achieved by the quadratic-based method [10].

$$\begin{aligned} L_{codok} &= \sum_{c=1}^{C}\sum_{i=1}^{n}\sum_{j=1}^{p} u_{ci} w_{cj} r_{ij} \\ &\quad -T_u \sum_{c=1}^{C}\sum_{i=1}^{n} u_{ci}^2 \\ &\quad -T_w \sum_{c=1}^{C}\sum_{j=1}^{p} w_{cj}^2. \end{aligned} \quad (6)$$

T_u and T_w play a similar role to λ_u and λ_w of the entropy-based method, and tune the degree of fuzziness, where a larger T brings fuzzier partitions. Considering the optimality condition $\partial L_{codok}/\partial u_{ci} = 0$ and $\partial L_{codok}/\partial w_{cj} = 0$, the updating rules for u_{ci} and w_{cj} are derived as follows:

$$u_{ci} = \frac{1}{C} + \frac{1}{2T_u}\left(\sum_{j=1}^{p} w_{cj} r_{ij} - \frac{1}{C}\sum_{\ell=1}^{C}\sum_{j=1}^{p} w_{\ell j} r_{ij}\right), \quad (7)$$

$$w_{cj} = \frac{1}{p} + \frac{1}{2T_w}\left(\sum_{i=1}^{n} u_{ci} r_{ij} - \frac{1}{p}\sum_{\ell=1}^{p}\sum_{i=1}^{n} u_{ci} r_{i\ell}\right). \quad (8)$$

Here, we must note that the updating rules of Eqs.(7) and (8) can derive negative memberships while u_{ci} and w_{cj} must be non-negative. In a simple strategy, negative memberships are replaced with *zero-memberships* and the remaining positive memberships are renormalized so as to sum to one.

C. Sampling and Incremental Algorithms for Fuzzy Co-clustering

In this paper, rseFCM, spFCM and oFCM are modified for applying to fuzzy co-clustering tasks.

1) Random Sample and Extend Approach: rseFCM is extended to fuzzy co-clustering algorithms of rseFCCM (or rseFuzzy CoDoK). In this approach, $n_s (n_s < n)$ objects are randomly selected from n objects and an $n_s \times p$ matrix R_s is constructed. Then, the conventional FCCM (or fuzzy CoDoK) is performed with R_s. After convergence, the derived item memberships w_{cj} are extended to the remaining objects for calculating their memberships u_{ci}.

2) Single-pass Approach: spFCM is extended to fuzzy co-clustering algorithms of spFCCM (or spFuzzy CoDoK). In this approach, n objects are first partitioned into s disjoint subsets, each of which consists of n_s objects, and $n_s \times p$ submatrices $R_1, \ldots, R_s$ are constructed. In the single pass process, FCCM (or fuzzy CoDoK) is sequentially applied to $R_1, \ldots, R_s$ inheriting the previous co-cluster structures.

In spFCM, previous cluster centers were added to next chunks as the messengers of the previous cluster structures. However, in fuzzy co-clustering, we do not have such prototypes, which are estimated in the same data space with object observations. In this paper, the applicability of item memberships w_{cj} is considered because they play a role for characterizing each co-cluster. Introducing the virtual objects, a submatrix R_k is modified to an $(n_s + C) \times p$ matrix $R_k^* = \{r_{ij}^*\}$, whose $(n_s + c)$th row is given by the item memberships of cluster c. Then, FCCM (or fuzzy CoDoK) is sequentially applied considering their responsibility weights.

When each object is associated with its responsibility weight α_i, weighted FCCM is implemented by modifying the updating rule for item memberships of Eq. (5) as follows:

$$w_{cj} = \frac{\exp\left(\lambda_w^{-1}\sum_{i=1}^{n} \alpha_i u_{ci} r_{ij}\right)}{\sum_{\ell=1}^{p}\exp\left(\lambda_w^{-1}\sum_{i=1}^{n} \alpha_i u_{ci} r_{i\ell}\right)}. \quad (9)$$

In a same manner, weighted fuzzy CoDoK modifies Eq. (8) as follows:

$$w_{cj} = \frac{1}{p} + \frac{1}{2T_w}\left(\sum_{i=1}^{n} \alpha_i u_{ci} r_{ij} - \frac{1}{p}\sum_{\ell=1}^{p}\sum_{i=1}^{n} \alpha_i u_{ci} r_{i\ell}\right). \quad (10)$$

It can be regarded in these formula as if we have α_i object i. On the other hand, object memberships are still updated by Eqs. (4) and (7).

In the submatrix R_k^*, objects $i = 1, \ldots, n_s$ are weighted as $\alpha_i = 1$ while the additional virtual objects $i = n_s + 1, \ldots, n_s + C$ are weighted considering the number of objects, which support the item memberships in the previous subprocess, as follows:

$$\alpha_{n_s+c} = \sum_{i=1}^{n_s+C} \tilde{u}_{ci} \times \tilde{\alpha}_i, \quad (11)$$

where $\tilde{u}_{ci}$ and $\tilde{\alpha}_i$ are the parameters of the previous subprocess.

Here, in contrast to FCM, we should note that item memberships $\boldsymbol{w}_c = (w_{c1}, \ldots, w_{cp})^\top$, $c = 1, \ldots, C$ are not directly comparative with a row element of the cooccurrence matrix $\boldsymbol{r}_{\cdot} = (r_{\cdot 1}, \ldots, r_{\cdot p})$ because $\boldsymbol{w}_c$ are normalized so that they do not represent cooccurrence degree of object-item pairs but represent relative typicalities among items, i.e., w_{cj} often have much smaller values than r_{ij} because of the sum-to-one condition. In the proposed algorithm of spFCCM (or spFuzzy CoDoK), before introducing $\boldsymbol{w}_c$ into R_k $(k > 1)$, the elements of $\boldsymbol{w}_c$ are normalized so as to have same range with original cooccurrence information r_{ij} as follows:

$$w^*_{cj} = r^{\min}_{ij} + \frac{w_{cj} - w^{\min}_{cj}}{w^{\max}_{cj} - w^{\min}_{cj}}(r^{\max}_{ij} - r^{\min}_{ij}) \qquad (12)$$

Additionally, it is obvious that $\sum_{i=1}^{n} \alpha_i u_{ci} r_{ij}$ and $\sum_{i=1}^{n} \alpha_i u_{ci} r_{ij} - \frac{1}{p}\sum_{\ell=1}^{p}\sum_{i=1}^{n} \alpha_i u_{ci} r_{i\ell}$ are proportional to the number of objects and the effects of λ_w and T_w are decreased as we use more virtual objects [12]. Then, the fuzzy degrees of λ_w and T_w should also be proportionally increased corresponding to the sum of objects weights as follows:

$$\lambda^*_w(T^*_w) = \lambda_w(T_w) \times \frac{\sum_{i=1}^{n_s+C} \alpha_i}{n_s}. \qquad (13)$$

After the sth subprocess with R_s, the final item memberships $\boldsymbol{w}_c$, $c = 1, \ldots, C$ are then extended to the objects included in $R_1, \ldots, R_{s-1}$ for calculating their final object memberships.

By the way, in the experiments shown in the next section, the initial item memberships w_{cj} in each subprocess were given by the resulted memberships of the previous subprocess for computational efficiencies.

3) On-line Approach: oFCM is extended to fuzzy co-clustering algorithms of oFCCM (or oFuzzy CoDoK). In this approach, n objects are also partitioned into s disjoint subsets, each of which consists of n_s objects, and $n_s \times p$ submatrices $R_1, \ldots, R_s$ are constructed. In the on-line process, FCCM (or fuzzy CoDoK) is first separately applied to $R_1, \ldots, R_s$ in parallel. Then, the item memberships are merged into a new virtual $(s \times C) \times p$ matrix $\tilde{R}$, which is available in the second stage for applying weighted FCCM (or weighted fuzzy CoDoK). The final item memberships are extended to the whole original objects and the final object memberships are calculated.

In the same manner with spFCCM (or spFuzzy CoDoK), the item memberships derived from the kth submatrix $\boldsymbol{w}^k_c = (w^k_{c1}, \ldots, w^k_{cp})$ should be normalized so as to have same range with r_{ij} before the second stage. The responsibility weight on item membership vector $\boldsymbol{w}^k_c$ is also given by

$$\alpha^k_c = \sum_{i=1}^{n_s} u^k_{ci}, \qquad (14)$$

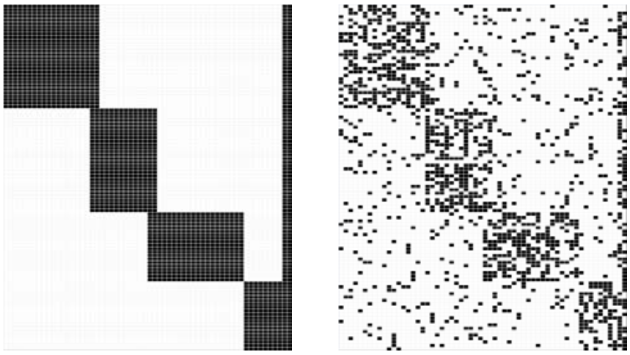
where u^k_{ci} are the object memberships derived from submatrix R_k. Considering the responsibility weights, weighted FCCM (or weighted fuzzy CoDoK) should also tune the fuzzification degree.

D. Complexity

Next, the time and space complexities are theoretically estimated. The updating rules for object memberships (Eqs.(4) and (7)) implies that the time complexity on a single u_{ci} is summarized as $O(tCm)$, where t is the number of iteration. On the other hand, from Eqs.(5) and (8), the time complexity on a single w_{cj} is $O(tmn)$. Then, the total time complexities are given as Table I, which indicates that the theoretical time complexities of spFCCM (spFuzzy CoDoK) and oFCCM (oFuzzy CoDoK) are equivalent to that of FCCM while rseFCCM (rseFuzzy CoDoK) can reduce the time complexity using a fewer number of objects only. However, as is also noted in [2], the incremental approaches needs much smaller number of iterations than the conventional whole data approaches in practice.

In a same manner, the space complexities can also be estimated as shown in Table I. The space complexity of both the sampling and incremental approaches are smaller than the whole data cases, and is proportional to the number of objects in each chunk (n/s) because we need keep parameters only on the current objects in each subprocess.

Then, the proposed sampling and incremental approaches can reduce computational time and memory requirement.



(a) noiseless (R_0) (b) noisy (R_1)

Fig. 1. Artificial cooccurrence matrices

IV. Experimental Results

Comparative experiments were performed with an artificially generated cooccurrence matrix. Figure 1-(a) shows the ideal (noiseless) data matrix R_0 composed of 100 objects and 60 items, which includes four co-clusters. Black and white cells indicate $r_{ij} = 1$ and $r_{ij} = 0$, respectively. 10 noisy subsets, each of which is composed of 100 objects and 60 items, were generated by replacing randomly selected element with its counterpart. Then, the 10 subsets $R_1, \ldots, R_{10}$ were gathered into a 1000×60 cooccurrence matrix R. For example, the first subset R_1 is shown in Fig. 1-(b). In this experiment, FCCM (fuzzy CoDoK) algorithms were applied with $C = 4$ and $\lambda_u = 0.001$ ($T_u = 0.001$). λ_w (T_w) was set as $\lambda_w = 100.0$ ($T_w = 2000.0$) for 100 objects (single subset) cases, and was increased in proportion to the number of objects because the degree of fuzziness of item memberships is dependent on the number of objects. Each algorithm was performed 50 times with different initialization and the average of 50 trials is evaluated.

A. Comparison of Partition Quality

Tables II and III compares the quality of item memberships derived by sampling approaches, where the partition quality is evaluated by the similarity with the ideal (noiseless) result. The tables present the correlation coefficients between the ideal item memberships given from R_0 and those given

TABLE I
COMPARISON OF TIME AND SPACE COMPLEXITY

Algorithm	Time	Space
FCCM(Fuzzy CoDoK)	$O(tCmn(C+m))$	$O((C+m)n)$
rseFCCM(rseFuzzy CoDoK)	$O(tCmn(C+m)/s)$	$O((C+m)(n/s))$
spFCCM(spFuzzy CoDoK)	$O(tCmn(C+m))$	$O((C+m)(n/s))$
oFCCM(oFuzzy CoDoK)	$O(tCmn(C+m))$	$O((C+m)(n/s)+Cs)$

TABLE II
COMPARISON OF ITEM MEMBERSHIP QUALITIES WITH FCCM (CORRELATION COEFFICIENT WITH NOISELESS RESULT: MEAN VALUES AND VARIANCES)

	correlation coefficient: mean value (variance)
FCCM(whole data)	0.914 (0.0136)
rseFCCM	0.794 (0.0167)
spFCCM	0.882 (0.0076)
oFCCM	0.831 (0.0285)

TABLE III
COMPARISON OF ITEM MEMBERSHIP QUALITIES WITH FUZZY CODOK (CORRELATION COEFFICIENT WITH NOISELESS RESULT: MEAN VALUES AND VARIANCES)

	correlation coefficient: mean value (variance)
Fuzzy CoDoK(whole data)	0.917 (0.0172)
rseFuzzy CoDoK	0.809 (0.0180)
spFuzzy CoDoK	0.881 (0.0152)
oFuzzy CoDoK	0.857 (0.0240)

from the noisy very large data R. Means and variances were calculated from the results of 50 trials.

The tables indicate that both FCCM and fuzzy CoDoK derived similar performances and the result of rseFCCM (rse-Fuzzy CoDoK) is inferior to other ones, i.e., a single random sampling process is not enough to reconstruct the intrinsic cluster structure. The results of spFCCM (spFuzzy CoDoK) and oFCCM (oFuzzy CoDoK) are almost comparative and are superior to rseFCCM (rseFuzzyCoDoK) although they are slightly inferior to the whole data case. The variance in spFCCM (spFuzzy CoDoK) is smaller than that in oFCCM (oFuzzy CoDoK), i.e., spFCCM (spFuzzy CoDoK) is more stable than oFCCM (oFuzzy CoDoK). In spFCCM (spFuzzy CoDoK), the varieties of initial partitions may be gradually summarized in its sequential processes. On the other hand, in the parallel first stage of oFCCM (oFuzzy CoDoK), some subprocess can converge into local minima because they use only small chunks, and then, the errors may be cumulated in the second stage. This result implies that each data chunk should include enough amount of cooccurrence information for applying oFCCM (oFuzzy CoDoK).

TABLE IV
COMPARISON OF CONVERGENT SPEED WITH FCCM (AVERAGES OF TOTAL TIME (SEC.) AND ITERATION PER CHUNK)

	time (sec.)	iteration
FCCM(whole data)	0.029	8.3
rseFCCM	0.004	9.1
spFCCM	0.011	2.7
oFCCM	0.017	4.3 (1st phase) 4.2 (2nd phase)

TABLE V
COMPARISON OF CONVERGENT SPEED WITH FUZZY CODOK (AVERAGES OF TOTAL TIME (SEC.) AND ITERATION PER CHUNK)

	time (sec.)	iteration
Fuzzy CoDoK(whole data)	0.034	10.2
rseFuzzy CoDoK	0.004	10.8
spFuzzy CoDoK	0.013	2.9
oFuzzy CoDoK	0.021	5.1 (1st phase) 8.0 (2nd phase)

B. *Comparison of Convergent Speed*

Next, convergent speed of the algorithms are compared. Tables IV and V compare the averages of the total operation times (sec.) and the iteration needed for convergence in each FCCM (fuzzy CoDoK) process per chunk. Although the theoretical computational costs of spFCCM (spFuzzy CoDoK) and oFCCM (oFuzzy CoDoK) are equivalent to the conventional ones, the practical cost could be reduced because of the support of fewer iterations. It may be because each incremental process can achive a fast convergence adopting the previous local optima as the current initial partition. Then, it is expected that the proposed method contribute to reduction of computational costs from the view points of both speed and space.

V. CONCLUSIONS

In this paper, the applicability of sampling and incremental approaches for handling very large data in FCM to co-clustering was discussed. Although the fuzzy co-clustering models of FCCM and fuzzy CoDoK are prototype-less clustering algorithms, incremental algorithms could be constructed by introducing item memberships as additional virtual objects after normalization. Several experimental results demonstrated that the proposed incremental algorithms contribute to extracting robust co-cluster structures without significant computational costs.

Possible future works include application to real world cooccurrence data such as purchase history data and document-keyword relational data. The automatic mechanism for tuning the fuzzy degrees should also be developed in future works.

REFERENCES

[1] J. C. Bezdek: *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, 1981.

[2] T. C. Havens, J. C. Bezdek, C. Leckie, L. O. Hall and M. Palaniswami, "Fuzzy *c*-means algorithms for very large data," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 6, pp. 1130-1146, 2012.

[3] P. Hore, L. Hall and D. Goldgof, "Single pass fuzzy *c*-means," *Proc. IEEE Int. Conf. Fuzzy Syst.*, London, U.K., pp. 1-7, 2007.

[4] P. Hore, L. Hall, D. Goldgof, Y. Gu and A. Maudsley, "A scalable framework for segmenting magnetic resonance images," *J. Signal Process Syst.*, vol. 54, no. 1-3, pp. 183-203, 2009.

[5] S. Miyamoto and D. Suizu, "Fuzzy *c*-means clustering using kernel functions in support vector machines," *J. Advanced Computational Intelligence and Intelligent Informatics*, vol.7, no.1, pp.25-30, 2003.

[6] C.-H. Oh, K. Honda and H. Ichihashi, "Fuzzy clustering for categorical multivariate data," *Proc. of Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, pp. 2154-2159, 2001.

[7] S. Miyamoto and M. Mukaidono, "Fuzzy *c*-means as a regularization and maximum entropy approach," *Proc. of the 7th International Fuzzy Systems Association World Congress*, vol. 2, pp. 86-92, 1997.

[8] S. Miyamoto, H. Ichihashi and K. Honda: *Algorithms for Fuzzy Clustering*, Springer, 2008.

[9] K. Kummamuru, A. Dhawale and R. Krishnapuram, "Fuzzy co-clustering of documents and keywords," *Proc. 2003 IEEE Int'l Conf. Fuzzy Systems*, vol. 2, pp. 772-777, 2003.

[10] S. Miyamoto and K. Umayahara, "Fuzzy clustering by quadratic regularization," *Proc. 1998 IEEE Int. Conf. Fuzzy Systems and IEEE World Congr. Computational Intelligence*, vol. 2, pp. 1394-1399, 1998.

[11] J. B. MacQueen, "Some methods of classification and analysis of multivariate observations," *Proc. of 5th Berkeley Symposium on Math. Stat. and Prob.*, pp. 281-297, 1967.

[12] K. Honda, A. Notsu and C.-H. Oh, "Handling very large cooccurrence matrices in fuzzy co-clustering by sampling approaches," *Proc. of 14th International Symposium on Advanced Intelligent Systems*, pp. 1720-1728, 2013.