TELKOMNIKA, Vol. 11, No. 4, April 2013, pp. 1781~ 1786
ISSN: 2302-4046

An Approach to Determining the Optimal Cell Number of Manufacturing Cell Formation

Jianwei Wang
College of Mechanical Engineering, Dalian University, Dalian, China
Corresponding author, e-mail: wangjw72@163.com

Abstract

An approach to determining the optimal cell number of manufacturing cell formation is presented. Firstly, the difference of weighting exponent, cluster center and metrics how to have an impact upon the clustering results and membership function are studied. Secondly, a method to determine the optimal m value is given. Two-order partial derivative of the objective function for FCM is calculated, and the variational weighting exponent m is obtained that can prevent the parameter from being the unique value and play an important role in the process of fuzzy clustering. Moreover, in order to avoid a single validity index can not assess correctly, partition coefficient (PC), classification entropy (CE), Fukuyama and Sugeno (FS) and Xie and Beni (XB) are considered as multi-performance indexes to evaluate the cluster validity, and then an optimal number c is chosen based on these validity measures. Finally, test exampls are given to illustrate the validity of the proposed approach.

Keywords: *cell formation, cell number, fuzzy c-mean, evaluation*

1. Introduction

Cellular manufacturing is a useful way to improve overall manufacturing performance. Group technology is used to increase the productivity for manufacturing high quality products and improving the flexibility of manufacturing systems. Cell formation is an important step in group technology. It is used in designing good cellular manufacturing systems. The key step in designing any cellular manufacturing system is the identification of part families and machine groups for the creation of cells that uses the similarities between parts in relation to the machines in their manufacture [1]. Cluster analysis is a method for clustering a data set into groups of similar individuals, its principle accords with the requirement of cell formation.

In cluster analysis, the fuzzy c-means (FCM) clustering algorithm is the best known and used method for cell formation problem. During the last two decades of research, a large number of cell formation methods based on FCM have been developed. Xu and Wang [2] first applied the fuzzy clustering to cell formation. Chu and Hayya [3] then improved its usage. Gindy et al. [4] considered optimal numbers of part families and machine groups using some validity indexes. Venugopal [5] gave a state-of-the-art review on the use of soft computing including fuzzy clustering. Moreover, Güngör and Arikan [6] applied fuzzy decision making in CF. However, it is necessary to pre-assume the cell number c in those FCM clustering algorithms, the cell number c is generally unknown. If cell number c is assigned an inaccurate value, it will cause invalid or worse cluster. Therefore, it is worthy how to determine the optimal cell number c of manufacturing cell formation.

In this paper, an approach to determining the optimal cell number of manufacturing cell formation is presented. Firstly, the influence factors of FCM algorithm and cluster validity are analysed. Secondly, based on the relationship fuzzy objective function with weighting exponent m, a novel method of choosing m in FCM is proposed. Taking into account the effect of clustering center subject to FCM, the objective function is modified by revising the constraint term based on simulated annealing to avoid the accordant cluster centers happening. The new measure style of fuzzy cluster is adopted to decrease the defect of Euclid distance in cell formation. Aiming at none of uniform performance index for evaluating the cluster validity, a synthetic performance indexes are adopted to assess the cluster validity and select the optimal cell number. Finally, set of test examples are given, the simulation results demonstrate the proposed approach is both effective and feasible.

Received January 10, 2013; Revised February 6, 2013; Accepted February 18, 2013

The rest of the paper is organized as follows: Sections 2 introduce the details of the proposed algorithm. The test examples are given and simulation results and discussions are presented in Section 3. Finally, conclusions are given.

2. An Approach to Determining the Optimal Cell Number of Manufacturing Cell Formation

2.1. FCM Clustering Algorithm

The FCM is an iterative algorithm using the necessary conditions for a minimum of the FCM objective function $J_m(\mu,v)$ [7]. It can be described as follows:

$$J_m(\mu,v)=\sum_{k=1}^{N}\sum_{i=1}^{c}(\mu_{ik})^m d(z_k,v_i) \tag{1}$$

where $\mu=\{\mu_1, \mu_2,\dots, \mu_c\}$. $v=\{v_1, v_2,\dots,v_c\}$ is the set of c cluster centers. μ_{ik} is the membership of the kth sample to the ith cluster center, its value is assigned in the interval [0,1] and $\sum_{i=1}^{c}\mu_{ik}=1$. m is the weighting exponent, $m\in[1,+\infty)$. $d(z_k,v_i)$ is the Euclidean distance between the sample z_k and the cluster center v_i, $d(z_k,v_i)=\left[\sum_{j=1}^{s}|z_{kj}-v_{ij}|^2\right]^{1/2}$. $Z=\{z_1,z_2,\cdots,z_N\}\subset R^s$ is the data set, N is the sample number of data set.

The necessary conditions for a minimum (μ,v) of $J_m(\mu,v)$ are the following update equations:

$$\mu_{ik}^{(l)}=\left(\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}\right)^{-1} \quad i,j=1,2,\dots,c,\ k=1,2,\dots,N. \tag{2}$$

$$v_i^{(l)}=\sum_{k=1}^{N}\left(\mu_{ik}^{(l)}\right)^m z_k\Big/\sum_{k=1}^{N}\left(\mu_{ik}^{(l)}\right)^m \quad i=1,2,\dots,c \tag{3}$$

2.1.1. The weighting Exponent *m* Influences on FCM Algorithm

The weighting exponent m is called the fuzzifier which can have an influence on the clustering performance of FCM. The best choice for m is probably in the interval [1.5, 2.5], whose mean and midpoint m=2, have often been the preferred choice for many users of FCM [8]. It is important to choose correctly m according to the different problems.

There is the implicit relationship between $J_m(\mu,v)$ and m, then

$$\frac{\partial J_m(\mu,v)}{\partial m}=\sum_{i=1}^{c}\sum_{k=1}^{n}(\mu_{ik})^m\cdot\lg(\mu_{ik})\cdot(d_{ik})^2=\sum_{i=1}^{c}\sum_{k=1}^{n}\left[\mu_{ik}\lg(\mu_{ik})\right]\left[\mu_{ik}{}^{m-1}(d_{ik})^2\right]<0 \tag{4}$$

From the equation, it can be found that $J_m(\mu,v)$ will monotonically decrease with the increasing m. However, the decreasing rate of $J_m(\mu,v)$ can be divided into two parts: a sharp drop and slow drop, then there is a inflection point between the two parts. The optimal weighting exponent is the value corresponding to the inflection point. m* can be calculated by the following equation:

$$m^*=\left\{m\,\middle|\,\frac{\partial}{\partial m}\left[\frac{\partial J_m(\mu,v)}{\partial m}\right]=0\right\} \tag{5}$$

2.1.2. The Fuzzy Distance $d(z_k,v_i)$ Influences on FCM Algorithm

During the cell formation, the Euclid distance is mostly adopted to determine $d(z_k,v_i)$. However, the same or different element number are taken into account firstly, the Euclid distance should not reflect the characteristic of cell formation problem. The distance function $d(z_k,v_i)$ between part z_k and cluster center v_i can be described as follows:

$$d_{ik}\left(z_k,v_i\right)=\left(0.5\sum_{j=1}^{s}\left|z_{kj}-v_{ij}\right|\right)\Big/\left(0.5+\sum_{j=1}^{s}z_{kj}v_{ij}\right) \quad (6)$$

where $\sum_{j=1}^{s}\left|z_{kj}-v_{ij}\right|$ is the number of used different machine between part z_k and cluster center v_i .

$\sum_{j=1}^{s}z_{kj}v_{ij}$ is the number of used same machine between part z_k and cluster center v_i .

2.1.3. The Cluster Center v_i Influences on FCM Algorithm

In FCM algorithm, the cluster center v_i should be keep the differentiation degrees and avoid the consistency. The philosophy of simulated annealing is referenced and the value of γ is revised constantly, then the influence degree of cluster center can be improved. In the initial stage, the value of γ is large to ensure the separation between clusters. In the final stage, the value of γ decreases to 0 to ensure the compactness between clusters.

2.1.4. Sub Bab 2

$$J_m(\mu,v)=\sum_{k=1}^{N}\sum_{i=1}^{c}\left[\left(\mu_{ik}\right)^m d\left(z_k,v_i\right)-\frac{\gamma}{c}\sum_{t=1}^{c}d(v_i,v_t)\right] \quad (7)$$

where " i,k, $\mu_{ik}\in(0,1)$, $\sum_{i=1}^{c}\mu_{ik}=1$, i=1,2,…,c, k=1,2,…,N. The weighting exponent m*, fuzzy distance $d(z_k,v_i)$, cluster center v_i and membership function μ_{ik} are shown as follows:

$$m^*=\arg\left\{\min_{\forall m}\left\{\frac{\partial J\left(U^*,V^*\right)}{\partial m}\right\}\right\} \quad (8)$$

$$d\left(z_k,v_i\right)=\left(0.5\sum_{j=1}^{s}\left|z_{kj}-v_{ij}\right|\right)\Big/\left(0.5+\sum_{j=1}^{s}z_{kj}v_{ij}\right) \quad (9)$$

$$v_i=\left(\sum_{k=1}^{N}\left(\mu_{ik}\right)^m z_k\right)\Big/\sum_{k=1}^{N}\left(\mu_{ik}\right)^m \quad , \quad i=1,2,\ldots,c \quad (10)$$

$$\mu_{ik}=\left(\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}\right)^{-1} \quad (11)$$

2.2. Cluster Validity for Fuzzy Clustering

Whether does the FCM algorithm accurately represent the structure of the data set? There are four most cited validity indexes shown as follows:

(1) Partition coefficient (PC) [9]:

$$\max_{2\le c\le c_{max}}\{PC(c)\}=\max_{2\le c\le c_{max}}\left\{\frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_{ik})^2\right\} \quad (12)$$

where $1/c\le PC(c)\le 1$.

(2) Classification entropy (CE) [7]:

$$\min_{2\le c\le c_{max}}\{CE(c)\}=\min_{2\le c\le c_{max}}\left\{-\frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N}\mu_{ik}\log_2(\mu_{ik})\right\} \quad (13)$$

where $0 \le CE(c) \le \log_2 c$.

(3) Fukuyama and Sugeno (FS) [10]:

$$\min_{2\le c\le c_{max}} \{FS(c)\} = \min_{2\le c\le c_{max}} \left\{ \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{m} \|z_k - v_i\|^2 - \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{m} \|v_i - \bar{v}\|^2 \right\} \quad (14)$$

where $\bar{v} = \frac{1}{c}\sum_{i=1}^{c} v_i$.

(4) Xie and Beni (XB) [11]:

$$\min_{2\le c\le c_{max}} \{XB(c)\} = \min_{2\le c\le c_{max}} \left\{ \left(\sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{m} \|z_k - v_i\|^2 \right) \Big/ \left(N\min_{i,j} \|v_i - v_j\| \right) \right\} \quad (15)$$

Note that since no single validity index is the best, a better way of using validity indexes to solve the cluster validity problem is to consider all information proposed by all selected indexes, and then make an optimal decision. The four validity indexes are looked as a synthetic performance indexes to assess the cluster validity and choose the optimal c.

3. Case Study

To demonstrate the performance of the proposed method, four instances of Reference [8] is adopted and used the same data set. The initial machine-part matrix of four instances owns different data scales 5×7, 10×15, 24×40 and 40×100 respectively. The proposed approach and four performance indexes are employed for simulating test. The parameters for test cases were set as follows: ε=0.001, c_{min}=2. The performance indexes for test case are shown in Tables 1-4. The optimal cell number c is signed with gray for every test case shown in Tables 1-4. The validity indexes with different cluster numbers c for test case are shown in Figures 1-4. Table 5 shows the solutions of optimal cell number c obtained by the proposed method in this paper and Reference [8].

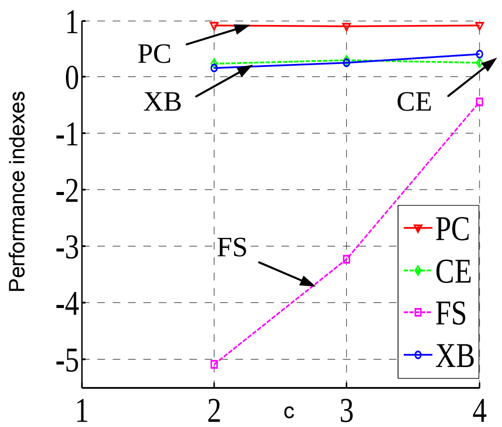

Figure 1. Validity Indexes with Different Cluster Numbers c (5×7)

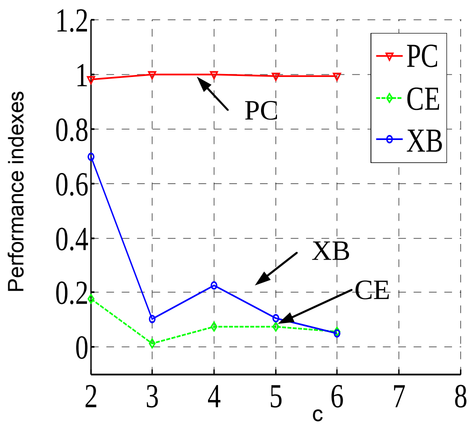

Figure 2. Validity Indexes with Different Cluster Numbers c (10×15)

Table 1. Synthetic Performance Indexes for Test Case (5×7)

c	Synthetic performance indexes			
	PC(c)	CE(c)	FS(c)	XB(c)
2	0.9188	0.2323	-5.0955	0.1635
3	0.8954	0.3060	-3.2208	0.2454
4	0.9155	0.2588	-0.4343	0.4016

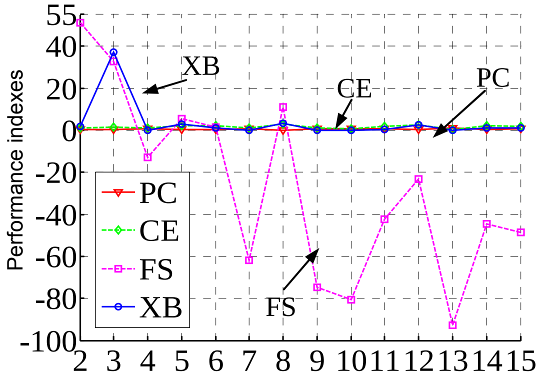

Figure 3. Validity Indexes with Different Cluster Numbers c (24×40)

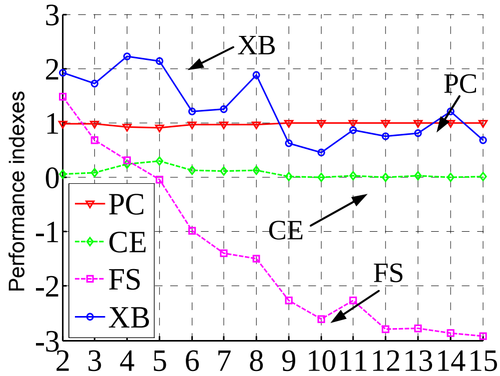

Figure 4. Validity Indexes with Different Cluster Numbers c (40×100)

Table 2. Synthetic Performance Indexes for Test Case (10×15)

c	Synthetic performance indexes			
	PC(c)	CE(c)	FS(c)	XB(c)
2	0.9809	0.1752	2.7661	0.6984
3	0.9997	0.0130	-25.9917	0.1023
4	0.9996	0.0759	-28.0633	0.2249
5	0.9954	0.0758	-29.6039	0.1042
6	0.9954	0.0563	-31.6841	0.0498
7	NaN	NaN	NaN	NaN
8	NaN	NaN	NaN	NaN

Table 3. Synthetic Performance Indexes for Test Case (24×40)

c	Synthetic performance indexes			
	PC(c)	CE(c)	FS(c)	XB(c)
2	0.5000	1.0000	51.2776	1.82e+2
3	0.3333	1.5850	32.8266	3.73e+3
4	0.5676	1.2215	-12.9215	0.8657
5	0.3057	2.0124	5.5703	3.03e+2
6	0.2756	2.2390	1.3559	1.16e+2
7	0.6895	1.0437	-62.0475	0.2796
8	0.1251	2.9995	11.1527	3.17e+2
9	0.7480	0.9104	-74.8708	0.1732
10	0.7814	0.8124	-80.7597	0.1316
11	0.5134	1.8330	-42.3003	4.15e+1
12	0.3171	2.6495	-23.1775	2.53e+2
13	0.8875	0.4481	-92.8612	0.0560
14	0.5010	2.0237	-44.5946	1.10e+2
15	0.5362	1.9192	-48.4852	1.23e+2

Table 4. Synthetic Performance Indexes for Test Case (40×100)

c	Synthetic performance indexes			
	PC(c)	CE(c)	FS(c)	XB(c)
2	0.9879	0.0504	1.48e+02	1.9185
3	0.9799	0.0843	6.87 e+01	1.7285
4	0.9222	0.2403	3.14 e+01	2.2233
5	0.9085	0.3001	-3.7551	2.1446
6	0.9692	0.1237	-9.81 e+01	1.2114
7	0.9716	0.1097	-1.40 e+02	1.2571
8	0.9683	0.1231	-1.49 e+02	1.8824
9	0.9966	0.0182	-2.28 e+02	0.6216
10	0.9994	0.0041	-2.61 e+02	0.4600
11	0.9932	0.0336	-2.27 e+02	0.8729
12	0.9993	0.0044	-2.80 e+02	0.7497
13	0.9938	0.0228	-2.78 e+02	0.8178
14	0.9993	0.0045	-2.87 e+02	1.2127
15	0.9950	0.0196	-2.92 e+02	0.6904

Table 5. Simulated Data and Cluster Number

	Example	5×7	10×15	24×40	40×100
c	Reference [8]	2	3	7	10
	Proposed method	2	3	13	10

As can be seen in Tables 1-5 and Figures 1-4, the optimal cell number obtained by the proposed approach is in accord with that given in Reference [8]. Moreover, there are still three or four extremums of performance indexes corresponding with the optimal cell number, thus it can be proved that the presented method is effective and feasible. However, the fact is existing that some simulation results of the proposed algorithm are accord with the others, it reflects the fact that the complexity of cluster problems and the differentia of performance indexes. Therefore, it can be concluded that the approach for determining the optimal cell number of manufacturing cell formation is available and robust.

4. Conclusion

Taking into account the characteristics of cell formation problem and analysing the difference of weighting exponent, cluster center and metrics how to have an impact upon the clustering results and membership function, an approach to determining the optimal cell number of manufacturing cell formation is proposed. FCM algorithm is adopted to calculate the membership and cluster center of parts in the designated range of cell number. In order to assess correctly the clustering performance, four synthetic performance indexes are employed to select the optimal cell number. Finally, a set of test examples are given, the simulation results demonstrate the proposed approach is both effective and feasible.

Acknowledgment

The project is partly supported by National Nature and Science Foundation (No.51275061,50975033), Science & Technology Research Foundation of the Educational Department of Liaoning Province (No.LS2010006,LT2010006), Science & Technology Research Foundation of Liaoning Provincial (No.2008219013) and Doctoral Starting Foundation of Dalian University.

References

[1] Hung WL, Yang MS, Lee ES. Cell formation using fuzzy relational clustering algorithm. *Mathematical and Computer Modelling*. 2011; 53: 1776-1787.
[2] Xu H, Wang HP. Part family formation for GT applications based on fuzzy mathematics. *International Journal of Production Research*. 198; 927(9): 1637-1651.
[3] Chu CH, Hayya JC. A fuzzy clustering approach to manufacturing cell formation. *International Journal of Production Research*. 1991; 29: 1475-1487.
[4] Gindy NG, Ratchev TM, Case K. Component grouping for GT applications–a fuzzy clustering approach with validity measure. *International Journal of Production Research*. 1995; 33(9): 2493-2509.
[5] Venugopa V. Soft-computing-based approaches to the group technology problem: a state-of-the-art review. *International Journal of Production Research*. 1999; 37: 3335-3357.
[6] Güngör Z, Arikan F. Application of fuzzy decision making in part-machine grouping. *International Journal of Production Economics*. 2000; 63: 181-193.
[7] Bezdek JC. *Pattern Recognition with Fuzzy Objective Algorithms*. New York: Plenum Press. 1981.
[8] Pal NR, Bezdek JC. On cluster validity for the fuzzy c-means model. *IEEE Transaction on Fuzzy Systems*. 1995; 3(3): 370-379.
[9] Bezdek JC. Cluster validity with fuzzy sets. *Journal of Cybernetics*.1974; 3: 8-74.
[10] Fukuyama Y, Sugeno M. *A new method of choosing the number of clusters for the fuzzy c-means method.* Proceedings of Fifth Fuzzy Systems, Symposium. 1989: 247-250.
[11] Xie XL, Beni G. A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991; 13(8): 841-847.
[12] Chandrasekharan MP, Rajagopalan R. ZODIAC-An algorithm for concurrent formation of part families and machine cells. *International Journal of Production Research*. 1987; 25(6): 835-850.