EFFECTIVE CLUSTERING METHOD FOR GROUP TECHNOLOGY PROBLEMS: A SHORT COMMUNICATION

Tamal Ghosh*, Pranab K Dan

Department of Industrial Engineering,
West Bengal University of Technology, Kolkata, India.
BF 142, Sector 1, Salt Lake City, Kolkata 700064, India. Tel. +91-33-2334-1014.
Email. tamal.31@gmail.com, danpk.wbut@gmail.com

ABSTRACT

Cellular manufacturing (CM) has acquired an upward interest of researchers in recent era. The most important problem in designing cellular manufacturing systems is the cell formation (CF), which constructs part families with similar processing requirements and designates machines into machine cells to optimize the production process. This article presents a short research report based on hybrid approach to the CF problem in CM. The proposed approach exploits Pearson's correlation coefficient and weighted average linkage clustering technique. Thereafter a modified part grouping method is introduced to improve the quality of solutions. The proposed technique is tested on 12 published datasets and the experimental results signify that the proposed method is quite effective for small to medium size problems.

Keywords
Group technology, cell formation, machine part grouping, hybrid clustering technique, correlation coefficient.

1. INTRODUCTION

Traditional manufacturing systems, namely job shops and flow shops are inefficient to control contemporary manufacturing firms which are under extreme pressures due to shorter product life-cycles, impulsive demands and varied customer needs [1]. Cellular manufacturing (CM) has materialized as a viable replacement to these which is the application of group technology (GT), a philosophy that utilizes similarities in product design and production processes. A primitive concern in CM is to determine the part families and machine cells, known as the cell formation (CF) problem which dissects the manufacturing systems into cells to reduce setup times, tool requirements and work-in-process inventories, improve product quality and productivity, shorten lead times, and enhance the overall control of operations [2]. The CF problem has long been identified as the most tricky problem in grasping the concept of CM, which begins with two fundamental tasks, (i) machine-cell formation, where similar machines are grouped and dedicated to manufacture part-families. (ii) part-family construction, where parts with similar design, features, attributes, shapes are grouped and manufactured within a cell. Many computational techniques are developed to design cells due to the NP-hard nature of the problem [3].
The CF problems are represented by machine-part incident matrix (MPIM), where elements are presented as 0/1. Parts are arranged in cloumns and machines are in rows (Fig. 1). A 0 indicates no operation and an 1 indicates an active operation.
Solution matrix is obtained as block diagonal structure (Fig. 4). An 1 outside the block known as an exceptional element (EE) and a 0 inside a cell known as 'void'. The objective is to minimize the EEs and voids [2].

	P1	P2	P3	P4	P5
M1	0	1	1	0	1
M2	1	0	0	1	1
M3	0	0	1	0	1

Fig. 1 Machine-part incidence matrix

Various techniques are developed to solve CF problems since decades. The similarity coefficient approach was implemented by McAuley [4]. CF procedure is based on rearranging rows and columns of the MPIM.. Some of the methods are Rank order clustering (ROC) [5], Bond energy algorithm [6] etc. Array based methods consider the rows and columns of the MPIM as binary patterns [7, 8, 9]. Graph Theoretic Approach states the machines as vertices and the similarity between machines as the weights on the arcs [10, 11].

2. METHODOLOGY

2.1. Similarity Coefficient

The similarity metric is utilized based on Pearson's corellation coefficient [12]. If sample X_k has n features, therefore it can be written as $X_k=\{x_{1k},x_{2k},\ldots,x_{nk}\}$. The formula used to calculate r is given as:

$$r = \frac{n\sum_{i=1}^{n} x_{ik}x_{im} - (\sum_{i=1}^{n} x_{ik})(\sum_{i=1}^{n} x_{im})}{\sqrt{[n\sum_{i=1}^{n} x_{ik}^2 - (\sum_{i=1}^{n} x_{ik})^2][n\sum_{i=1}^{n} x_{im}^2 - (\sum_{i=1}^{n} x_{im})^2]}} \quad (1)$$

It is used to calculate the similarity between two samples by using:

$$S = \frac{(r+1)}{2} \quad (2)$$

Further the distance between the samples is calculated using:

$$D = (1 - r) \quad (3)$$

and used to generate the distance matrix in Fig. 2 obtained from the example of Fig. 1.

	m1	m2	m3
m1	0		
m2	1.667	0	
m3	0.333	1.167	0

Fig. 2 Distance matrix of Fig. 1

2.2. Machine clustering technique

An Weighted Average Linkage Clustering (WALC) model is used in this study to form machine clusters [8]. It delivers informative descriptions and visualization of potential data clustering structures. WALC uses a recursive definition for the distance between two clusters, given as:

$$d(r,s) = \frac{(d(p,s) + d(q,s))}{2} \quad (4)$$

d(r,s)=distance between cluster *r* and *s*
d(p,s)=distance between cluster *p* and *s*
d(q,s)=distance between cluster *q* and *s*

A $(m\text{-}1)\times 3$ matrix is obtained using (4), where m is the number of machines. Columns contain cluster indices linked in pairs to form a binary tree. The leaf nodes are numbered from *1 to m*.

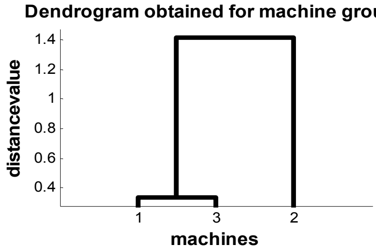

Fig. 3 Dendrogram of problem in Fig. 1

Leaf nodes are the singleton clusters from which all higher clusters are built. The dendrogram can be obtained with a tree of potential solutions (Fig. 3), states the construction of two clusters. Cluster 1 contains machines 1 and 3 and cluster 2 contains machine 2.

2.3. Part Family Formation

A modified part grouping technique is adopted in this study inspired from Zolfaghari and Liang [13], identifies a machine cell which processes the part for a maximum number of operations than any other machine cell. Therefore parts are assigned to the cells which further form tangible part families using membership value:

$$D_{cj} = \frac{m_{cj}}{k_c} \times \frac{m_{cj}}{n_j} \times \frac{1}{v} \qquad (5)$$

D_{cj} = Membership index of part j to cell c
m_{cj} = Number of machines in cell c which process part j
k_c = Total number of machines in cell c
n_j = Total number of machines required by part j
v = Total number of voids

Tab. 1 membership index values for parts

	Membership Index Values	
Parts	Cell 1	Cell 2
P1	0	**1**
P2	**0.5**	0
P3	**1**	0
P4	0	**1**
P5	**0.667**	0.333

Parts are assigned to cells with largest membership index value. The computed membership index values are depicted in Table 1. The above analysis illustrates the solution, parts 1, 3, 5 are grouped into family 1 and parts 2, 4 are grouped into family 2. The final block diagonal matrix is obtained in Fig. 4. The flow-chart of the proposed technique is presented in Fig. 5.

	P3	P5	P2	P1	P4
M1	1	1	1	0	0
M3	1	1	0	0	0
M2	0	1	0	1	1

Fig. 4 Final block diagonal matrix

2.4. Performance Measure

In this study grouping efficacy [14] measure is used as the evaluation criterion to test the goodness of the solutions, stated as,

$$\tau = \frac{E - E_e}{E + E_v} \qquad (6)$$

Where
E = Total number of 1s in MPIM
E_e = Total number of EEs
E_v = Total number of voids

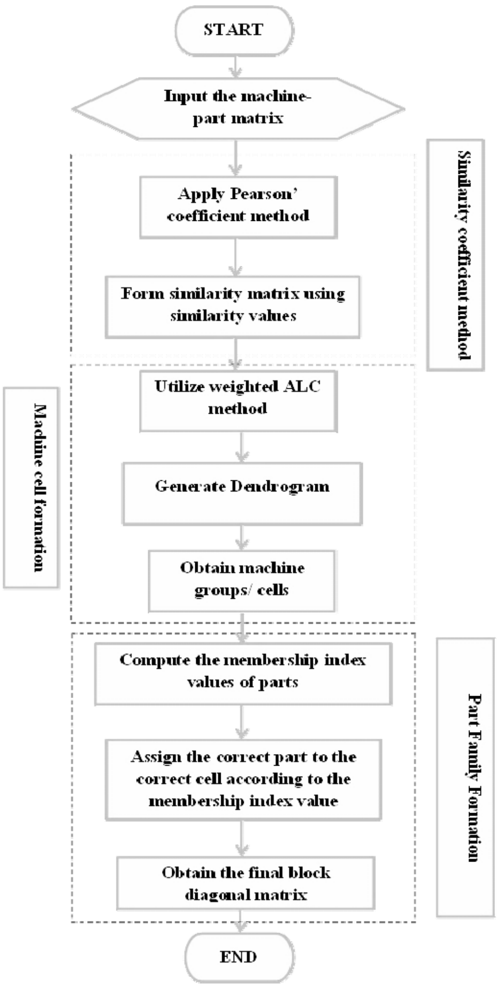

Fig. 5 Flow chart of methodology

The final machine-part block diagonal structure is achieved as,
Cell 1= {Machine 1, 3 || Parts 2, 3, 5}; Cell 2= {Machine 2 || Parts 1, 4}; EEs = 1; Voids = 1.
Grouping efficacy measure = (8-1)/(8+1) = 7/9 =0.7778 Which is 77.78%.

3. EXPERIMENTS

The proposed technique is simulated on Matlab 7.0 and a PIV computer, tested on 12 datasets [2, 15], compared with the best results obtained from the literature [2, 15] and given in Table 2. For the problems solved with proposed technique, the grouping efficacy value is better or equal in all instances with negligible computational time (< 10 seconds). This indicates that proposed technique is very efficient and less complex because of its simplicity in simulation. Improvement curve is shown in Fig. 6. The proposed technique outperforms the standard techniques in 6 instances, and equal in 6 instances, i.e. 50% improved results achieved in terms of solution quality, time and space complexities.

Table 2 comparison of the results

#	size	Grouping Efficacy Value		cell	EE	void
		Best result in literature	proposed method			
1	5×7	62.5	**69.56****	2	6	2
2	5×7	73.68	73.68	2	2	3
3	5×18	79.59	79.59	2	7	3
4	6×8	76.92	76.92	2	2	4
5	7×11	70.37	70.37	3	2	6
6	7×11	53.13	**59.26****	4	7	4
7	8×10	66.67	**68.96****	4	7	2
8	8×20	85.24	**86.67****	3	8	0
9	9×9	73.53	**74.28****	3	6	3
10	15×12	86.67	86.67	4	0	6
11	20×8	83.87*	82.25	3	9	1
12	25×10	63.93	**70.27****	4	9	13

** inconsistent result shown in [5], actual computed value is 82.25*
*** better result shown in boldface*

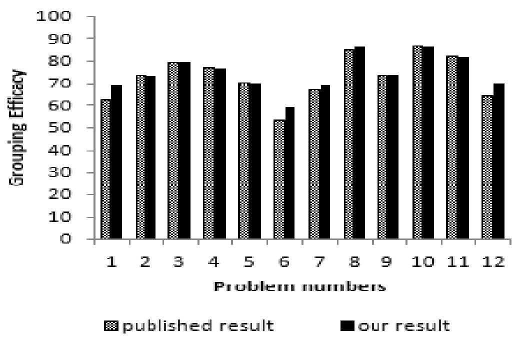

Fig. 6 improvement shown by proposed method

4. CONCLUSION

This short communication portrays an effective clustering methodology that blends Pearson's correlation coefficient and WALC technique with modified part grouping method. Obtained results validate that the proposed technique outperforms the published methods, enhancing the clustering results. The proposed method obtains better quality solutions by consuming lesser CPU time. It is also shown that the technique performs at least as well as, and often better than the available algorithms for the CF on all problems tested. Therefore it is verified as a promising method in aforestated area. This research work is yet under development phase.

REFERENCES

[1] Burbidge, J.L. (1963). "Production flow Analysis". *Production Engineer*, 42: 742-752.

[2] Unler, A. and Gungor, Z. (2009). "Applying K-harmonic means clustering to the part-machine classification problem". *Expert Systems with Applications*, 36: 1179–1194.

[3] Ghosh, T., Sengupta, S., Chattopadhyay, M. and Dan, P.K. (2011). "Meta-heuristics in cellular manufacturing: A state-of-the-art review". *International Journal of Industrial Engineering Computations*, 2: 87-122.

[4] McAuley, J. (1972). "Machine grouping for efficient production". *Production Engineer*, 51: 53-57.

[5] King, J.R. (1980). "Machine-component grouping in production flow analysis: an approach using a rank order-clustering algorithm". *International Journal of Production Research*, 18: 213-232.

[6] McCornick, W. T., Schweitzer, J.P.J. and White, T.W. (1972). "Problem decomposition and data reorganization by a clustering technique", *Operations Research*, 20: 993-1009.

[7] King, J.R. and Nakornchai, V. (1982). "Machine-component group formation in group technology: review and extension", *International Journal of Production Research*, 20: 117-133.

[8] Chandrasekharan, M.P. and Rajagopalan, R. (1986). An ideal seed non-hierarchical clustering algorithm for cellular manufacturing, *International Journal of Production Research*, 24: 451-464.

[9] Chan, H.M. and Milner, D.A. (1982). "Direct Clustering algorithm for group formation in Cellular Manufacturing". *Journal of Manufacturing Systems*, 1: 65-75.

[10] Rajagopalan, R. and Batra, J.L. (1975). "Design of Cellular production system – A graph theoretic approach". *International Journal of Production Research*, 13: 567-579.

[11] Ballakur, A. and Steudel, H.J. (1987). "A with-in cell utilization based heuristic for designing cellular manufacturing systems", *International Journal of Production Research*, 25: 639-655.

[12] Benesty, J., Chen, J., Huang, Y. and Cohen, I. (2009) Pearson Correlation Coefficient. In: Cohen, I., Huang, Y., Chen, J. and Benesty, J. Eds. *Noise Reduction in Speech Processing, Springer Topics in Signal Processing*, Vol.2, p.1-4.

[13] Zolfaghari, S. and Liang, M. (2003). "A new genetic algorithm for the machine/part grouping problem involving processing times and lot sizes". *Computers and Industrial engineering*, 45: 713-731.

[14] Kumar, S.C. and Chandrasekharan, M.P. (1990). "Grouping efficacy: a quantitative criterion for goodness of block diagonal forms of binary matrices in group technology". *International Journal of Production Research*, 28: 233-243.

[15] James, T.L., Brown, E.C. and Keeling, K.B. (2007). "A hybrid grouping genetic algorithm for the cell formation problem". *Computers & Operations Research*, 34: 2059–2079.