THE PRODUCTION DATA-BASED SIMILARITY COEFFICIENT VERSUS JACCARD'S SIMILARITY COEFFICIENT

Hamid Seifoddini and Manucher Djassemi

Department of Industrial and Systems Engineering
University of Wisconsin--Milwaukee
Milwaukee, Wisconsin

ABSTRACT

This paper compares the performance of Jaccard's similarity coefficient with the production data-based similarity coefficient. A number of machine-component charts taken from the literature or randomly generated are used to form machine-component groups. Then, the sum of intercellular and intracellular material handling costs for each machine-component group is calculated and used as a basis for performance evaluation of the two similarity coefficients.

INTRODUCTION

The machine-component grouping process is a basic step in the implementation of cellular manufacturing in which part-families are processed in dedicated machine cells each capable of processing one or more part-families. Cellular manufacturing improves productivity through reduction in setup times [2,5].

There are a number of different approaches to the machine-component grouping problem among them the similarity coefficient method is more effective in forming machine cells [3,4,6,7,12]. In this method, a measure of similarity (similarity coefficient) is defined between two machines (parts) and a clustering algorithm is used to group machines into machine cells [6]. The similarity coefficient between two machines is defined as the number of parts visiting both machines divided by the number of parts visiting either of the two machines [6,7].

This can be mathematically expressed as:

$$S_{ij} = \frac{\sum_{k=1}^{n} X_{ijk}}{\sum_{k=1}^{n} Y_{ijk}}$$

where,

n = the number parts

$$X_{ijk} = \begin{cases} 1 & \text{if part } k \text{ visits both machines } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

$$Y_{ijk} = \begin{cases} 1 & \text{if part } k \text{ visits at least one of machines } i \text{ and } j \\ 0 & \text{otherwise} \end{cases}$$

This definition is based on Jaccard's similarity coefficient [1,6].

One of the major problems with Jaccard's similarity coefficient is its limitation in incorporating various types of production data into the machine-component grouping process [8]. This similarity coefficient uses only the data in the machine-component chart which represent the machining requirements of parts in the product mix. As a result, important production data such as production volume, processing times, and sequence of operations are not considered in the development of cellular manufacturing systems.

A modified version of Jaccard's similarity coefficient called "the production data-based similarity coefficient" can overcome the above mentioned problem [8]. This similarity coefficient uses a weighting factor to reflect the effect of different types of production data including production volume.

The production data-based similarity coefficient can be mathematically expressed as:

$$S_{ij} = \frac{\sum_{k=1}^{n} N_k X_{ijk}}{\sum_{k=1}^{n} N_k Y_{ijk}}$$

where,

N_k = production volume for part type k

S_{ij}, n, X_{ijk}, Y_{ijk} as defined before.

The incorporation of production volume into the machine-component grouping process promises to generate more effective cellular manufacturing systems.

SOLUTION METHODOLOGY

The performances of the two similarity measures, Jaccard's similarity coefficient and the production-based similarity coefficient, are compared through the performance evaluation of the corresponding cellular manufacturing systems. Each of the two similarity coefficients will be employed in conjunction with a clustering algorithm to develop a cellular manufacturing system based on a given machine component chart. Then, a proper performance measure is employed to compare these cellular manufacturing systems which represent the similarity coefficients.

Several performance measures have been developed for the evaluation of cellular manufacturing systems, including the sum of intercellular and intracellular material handling costs, group efficiency, group efficacy, and group capability index [4]. Most of these measures, however, are inconsistent in determining the performance and generate less than perfect scores, even when a complete block diagonal form is formed [4].

Among the existing performance measures, the sum of intercellular and intracellular material handling costs is effective in the performance evaluatio cellular manufacturing systems [7]. is due to two major factors. First, material handling cost directly affects production cost. Therefore, any reduc in the sum of intercellular and in cellular material handling costs impr the performance of the cellular manufac ing system. Secondly, the reduction intercellular material handling cost achieved by placing machines with a l number of operations close to each ot Such an arrangement provides the basis the implementation of group tooling group scheduling which reduce the s costs and result in further reduction production costs.

The sum of intercellular and in cellular material handling cost is a f tion of the arrangement of machines machine cells which is directly affecte the type of similarity coefficient use form them. Since the new similarity mea uses the production volume as a weigh the calculation of the similarity c ficient between machines, the simila coefficient between two machines w process parts with high volume will be h As a result, these machines are more li to be assigned to the same machine cell reducing the number of intercellular m by replacing them with intracellular mo This will affect the sum of intercell and intracellular material handling c and make this sum an effective perform measure for the comparison of diffe similarity coefficients.

The intercellular material hand cost is a function of the number of ex tional parts, the number of intercell moves created by each exceptional part, travelling distances between machines, the unit transportation cost. A numbe computerized algorithms exists that capable of calculating the intercell material handling cost based on a subopt layout of machine cells. One such algor is CRAFT [11]. This algorithm can be used to calculate the intracellular mate handling cost.

The sum of intercellular and intra lular material handling costs is calcul for the two cellular manufacturing sys formed by using Jaccard's similarity c ficient and the production data-based s larity coefficient. The result is use compare the performance of each simila coefficient. To minimize the effect special situations, 10 different prob have been used. Furthermore, the produc volume of parts in the machine-compo charts is generated randomly.

ANALYSIS OF RESULTS

The machine-component charts with 18 machines and 24 parts in Fig. 1 will be used to demonstrate the procedure for performance evaluation. The production volumes for parts 1 to 24 are 78, 91, 76, 139, 97, 69, 61, 115, 16, 120, 78, 91, 71, 67, 82, 61, 8, 75, 87, 31, 93, 60, 128, and 49. When Jaccard's similarity coefficient is used, the machine-component groups are as presented in Fig. 2. The sum of intercellular and intracellular material handling costs for this solution is $7,314.

The machine component groups when the production data-based similarity coefficient is used are given in Fig. 3. The sum of intercellular and intracellular material handling costs for this solution is $6,919. This reduction in material handling costs indicates that the incorporation of production volume in the machine-component grouping process improves the productivity of cellular manufacturing systems.

The results from 10 different problems show that in six cases the application of the production data-based similarity

Components

Machines	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1		1			1	1						1												
2																1				1				
3			1		1	1		1	1	1					1		1						1	1
4		1			1	1		1	1			1	1		1		1					1		
5				1																				
6				1																				
7								1	1					1		1			1			1		
8										1													1	
9						1	1						1	1				1	1		1	1		
10	1						1						1		1		1	1						
11							1											1			1	1		
12																							1	
13			1																					1
14	1		1	1						1										1			1	1
15		1		1	1			1	1	1	1	1			1	1	1						1	1
16		1										1							1					
17		1			1			1	1			1			1		1							
18																		1		1	1			

Figure 1. Initial machine component chart

Components

Machines	1	3	5	6	8	9	11	15	16	17	24	20	4	10	23	7	13	14	18	19	21	22	2	12
4			1	1	1	1		1		1							1					1	1	1
17			1		1	1		1		1													1	1
15			1		1	1	1	1	1	1	1		1	1	1								1	1
3		1	1	1	1	1		1		1	1			1	1									
14	1	1									1	1	1	1										
13		1									1													
2							1		1			1												
10	1	1						1		1				1		1	1		1					
5													1											
6													1											
8														1	1									
12														1										
9				1										1	1	1	1	1	1	1	1	1		
11														1		1	1		1		1	1		
18												1							1		1			
7					1	1												1		1		1		
1			1	1																			1	1
16																				1			1	1

Figure 2. Machine-component groups based on Jaccard's similarity coefficient

Components

Machines	2	5	6	8	9	12	15	17	19	4	1	7	13	14	18	20	21	22	3	10	23	24	11	16
4	1	1	1	1	1	1	1	1					1					1						
17	1	1	1	1	1	1	1	1																
1	1	1	1			1																		
16	1					1			1															
7				1	1				1					1				1						
5										1														
6										1														
11												1			1		1	1		1				
18															1	1	1							
9			1						1			1	1	1	1		1	1		1	1			
10							1	1			1	1	1		1				1	1				
8																				1	1			
12																				1				
14										1	1					1			1	1	1	1		
13																			1			1		
3		1	1	1	1		1	1												1	1	1		1
15	1	1		1	1	1	1	1		1										1	1	1	1	1
2																1							1	1

Figure 3. Machine-component groups based on the production data-based similarity coefficient

coefficient improved the performance of the corresponding cellular manufacturing system. In the remaining cases the results were the same.

CONCLUSION

Ten different machine-component grouping problems were used to compare Jaccard's similarity coefficient with the production data-based similarity coefficient. The results indicate that in most cases the sum of intercellular and intracellular material handling costs decreases when the production data-based similarity coefficient is used to form machine-component groups.

REFERENCES

1. Anderberg, M. R., Cluster Analysis for Applications, Academic Press, New York, 1973.
2. Burbidge, J. L., "Production Flow Analysis," Seminar on the First Steps to Group Technology, Birneihill Institute, East Kulbride, Glasgow, 1970.
3. Carrie, A. S., "Numerical Taxonomy Applied to Group Technology and Plant Layout," Int. J. Prod. Res., April 1973, pp. 399-416.
4. Hsu, C. P., "Similarity Coefficient Approaches to Machine-Component Cell Formation in Cellular Manufacturing: A Comparative Study," Unpublished M.S. Thesis, Industrial and Systems Engineering Department, University of Wisconsin-Milwaukee, 1990.
5. King, J. R., "Machine-Component Grouping in Production Flow Analysis: An Approach Using a Rank Order Clustering Algorithm," Int. J. Prod. Res., Vol. 18, March 1980, pp. 213-237.
6. McAuley, J. "Machine Grouping for Efficient Production," The Production Engineer, Vol. 52, pp. 53-57, February, 1972.
7. Seifoddini, H. and Wolfe, P. M., "Application of the Similarity Coefficient Method in Group Technology," IIE Transactions, September 1986, Vol. 18, No. 3, pp. 271-277.
8. Seifoddini, H., "Incorporation of the Production Volume in Machine Cells Formation in Group Technology Applications," Proceedings of the IXth ICPR, Oct. 1987, pp. 2348-2356.
9. Srinivasan, G., Narendran, T. and Mahadevan, B., "An Assignment Model for the Part-Families Problem in Group Technology," Int. J. Prod. Res., Vol. 28, 1990, pp. 145-152.
10. Sokal, R. R. and Sneath, P. H. A., Principles of Numerical Taxonomy, 1963, Freeman, San Francisco, California.
11. Tompkins, J. A. and Moor, M. J., "Computer-Aided Layout: A User's Guide," Publication No. 1 in the Monograph Series, Facilities Planning and Design Division, AIIE Inc., 1980.
12. Vakharia, A. J. and Wemmerlov, U., "Designing a Cellular Manufacturing System: A Material Flow Approach Based on Operations Sequences," Institute of Industrial Engineers (IIE) Transactions, Vol. 22, No. 1, 1990.