Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 17 (2014) 44 – 49

www.elsevier.com/locate/procedia

Variety Management in Manufacturing. Proceedings of the 47th CIRP Conference on Manufacturing Systems

Generation of block diagonal forms using hierarchical clustering for cell formation problems

Simon Li[a,]*, Houman Mehrabadi[b]

[a]*University of Calgary, Calgary, Canada*
[b]*Concordia University, Montreal, Canada*

* Corresponding author. Tel.: +1-403-220-5599; fax: +1-403-282-8406. *E-mail address:* simoli@ucalgary.ca

Abstract

While cell formation (CF) problems have been studied for few decades, the purpose of this paper is to advance the solution technique using one classical approach, hierarchical cluster analysis (HCA). In the application of HCA, one technical challenge is to cluster both machines and parts simultaneously. In this paper, this challenge is addressed by quantifying the coupling between machines and parts in the clustering process. One feature of the proposed method is to generate block diagonal forms that show some intermediate sorting of machines and parts without specifying the structural criteria (e.g., the number of cells). Consequently, engineers can specify the structural criteria after inspecting the block diagonal forms instead of specifying them at the beginning. Some numerical examples from literature are used to examine and verify the proposed method.

Selection and peer-review under responsibility of the International Scientific Committee of "The 47th CIRP Conference on Manufacturing Systems" in the person of the Conference Chair Professor Hoda ElMaraghy"

Keywords: cell formation problem; hierarchical clustering; simultaneous clustering

1. Introduction

Cell formation (CF) problems have been widely studied for few decades, and relevant review papers have been published at different times to reflect the significance of this topic [1, 2, 3, 4]. This paper adapts the approach of cluster analysis to tackle the CP problems. In contrast to the optimization approach (e.g., genetic algorithm), the algorithms by cluster analysis are relatively simple, and their applications have been found in the early CF literature (e.g., [5]). Yin and Yasuda [3] have emphasized that the similarity coefficient method (SCM, rooted in cluster analysis) is a more flexible approach to solve CF problems. One specific argument is that SCM consists of several tractable solution phases so that SCM can be easily adapted for various CF problems.

In this context, the purpose of this paper is to propose a new method based on cluster analysis. On the one hand, the new method applies two traditional techniques: machine-part incidence matrix for cell representation and hierarchical clustering for group formation. On the other hand, the new method has two specific features that are intended to contribute to the CF community.

Firstly, using a machine-part incidence matrix as an input, it generates a block diagonal form (BDF) as intermediate clustering results before suggesting machine groups and part families. The notion of BDF can be found in the early literature, and BDFs are helpful for practitioners to discern the patterns for cell formation. The techniques to generate BDFs include the close neighbor algorithm [6], the steepest descent pairwise interchange (SDPI) algorithm [7], the integrated fuzzy clustering method [8], and the evolutionary algorithm [9]. Yet, the problem of generating BDFs is equivalent to a travelling salesman problem (TSP) that is also NP-complete [10]. One new feature of the proposed method is about using hierarchical clustering for generating BDFs.

Selection and peer-review under responsibility of the International Scientific Committee of "The 47th CIRP Conference on Manufacturing Systems" in the person of the Conference Chair Professor Hoda ElMaraghy"
doi:10.1016/j.procir.2014.01.143

Secondly, the proposed method does not require the number of cells at the beginning of the algorithmic execution. In other words, we can specify the number and size of cells after obtaining a BDF. This feature echoes some arguments that manufacturing cells should be naturally identified without knowing the number of cells in advance [11, 12]. Figure 1 shows the basic workflow of the proposed method with the highlights of specific features.

Nomenclature

a_{ij}	entry of incidence matrix
m	total number of machines
m_i	ith machine
n	total number of parts
n_c	number of cells
n_e	total number of operations
n_{in}	total number of voids inside the cells
n_{out}	total number of operations outside the cells
p_j	jth part
r_{ij}	weighted similarity value between ith and jth objects
R_{ss2}	similarity coefficient of Sokal Sneath 2
$R_{two\text{-}mode}$	similarity coefficient of two-mode
w_a	weight for similarity value between machines
w_b	weight for similarity value between parts
w_c	weight for similarity value between a machine and a part
μ	grouping efficacy

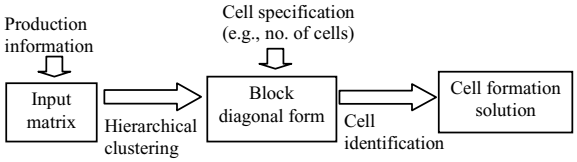

Fig. 1. Workflow of the proposed CF method.

2. Background: cell formation problem

In a cell formation (CF) problem, there exists a set of machines (m_1, m_2…) and parts (p_1, p_2…). The basic inquiry of the CF problem is to identify machine groups (i.e., subsets of m_i) and part families (i.e., subsets of p_j) in order to optimize some group efficacy measure. Based on the manufacturing requirements, it is specified which machines are required to make a part. In a matrix representation, a matrix's rows represent machines, and a matrix's columns represent parts. The dependency information is captured in a matrix entry a_{ij}, which is defined as follows.

$$a_{ij} = \begin{cases} 1 & \text{if machine } i \text{ is required to make part } j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In this problem, we want to form machine groups and part families so that the inter-cell elements can be minimized. Figure 2 illustrates the solution process. In the matrix format, the machines and parts are brought close to each other if they belong to the same block. For example, Figure 2 shows the formation of two blocks, where the first one consists of {m_2, m_3, p_2, p_3, p_5}. Possible inter-cell elements are shown as any nonzero matrix entries that relate two different blocks. Notably, some researchers have extended the binary relationships to non-binary ones, indicating the weights of the relationships between machines and parts (e.g., production data-based matrix in [13]). In this situation, the purpose of the cell formation problem is to minimize the weights of all inter-cell elements.

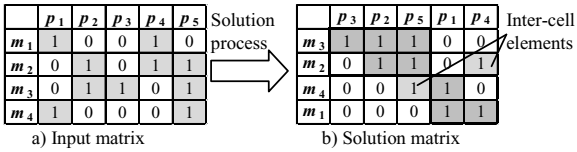

Fig. 2. Solution process.

To evaluate the solution quality, the group efficacy (denoted as μ) proposed in [14] is used in this paper, and its formulation is provided as follows.

$$\mu = \frac{n_e - n_{out}}{n_e + n_{in}} \quad (2)$$

where n_e, n_{out}, n_{in} are the total numbers of all operations, operations outside the cells, and voids inside the cells, respectively. In a "perfect" solution, there should be no operations outsides the cells (i.e., $n_{out} = 0$) and no voids inside the cells (i.e., $n_{in} = 0$), leading to $\mu = 1$. As an illustration, the grouping efficacy of the solution in Figure 2b is equal to (10-2)/(10+2) = 0.667.

Notably, the measure of CF solution quality is not a trivial topic as the goodness measure can be viewed from different aspects. Interested readers may find the surveys in [15, 16]. In this paper, the group efficacy is applied due to its generality and popularity in the field.

3. Proposed method

The proposed method consists of three major steps, namely, similarity analysis, sorting analysis and cell identification. These three steps will be discussed in the following sub-sections, and the simple example will Figure 2a will be used for illustration.

3.1. Similarity analysis

In this paper, two machines are said "similar" if they are required by some common parts. Likewise, two parts are said "similar" if they are manufactured by some common machines. In choosing the similarity coefficients, we have referenced [16] and examined three coefficients that generally yield good performance: Jaccard, Sorenson and Sokal Sneath 2. After our experimental study, we choose the coefficient "Sokal Sneath 2" since it can generally produce the solutions in favor of the group efficacy defined in Section 2. Using the matrix definition in Equation (1), this paper uses the max and min operators to formulate the coefficient (denoted as R_{ss2}). Particularly, the min operator counts the number of 1-1 matches, and the max operator counts the total number of 1-1 and 1-0 matches. The formulations (3) and (4) are provided for two machines and two parts, respectively.