Note On a Comparative Evaluation of Nine Well-Known Algorithms for Solving the Cell Formation Problem in Group Technology

PRAFULLA JOGLEKAR
La Salle University, Philadelphia, PA 19141. USA

Q B. CHUNG[1]
Villanova University, Villanova, PA 19085. USA

MADJID TAVANA
La Salle University, Philadelphia, PA 19141. USA

Abstract.   Over the last three decades, numerous algorithms have been proposed to solve the work-cell formation problem. For practicing manufacturing managers it would be nice to know as to which algorithm would be most effective and efficient for their specific situation. While several studies have attempted to fulfill this need, most have not resulted in any definitive recommendations and a better methodology of evaluation of cell formation algorithms is urgently needed. Prima facie, the methodology underlying Miltenburg and Zhang's (M&Z) (1991) evaluation of nine well-known cell formation algorithms seems very promising. The primary performance measure proposed by M&Z effectively captures the objectives of a good solution to a cell formation problem and is worthy of use in future studies. Unfortunately, a critical review of M&Z's methodology also reveals certain important flaws in M&Z's methodology. For example, M&Z may not have duplicated each algorithm precisely as the developer(s) of that algorithm intended. Second, M&Z's misrepresent Chandrasekharan and Rajagopalan's [C&R's] (1986) grouping efficiency measure. Third, M&Z's secondary performance measures lead them to unnecessarily ambivalent results. Fourth, several of M&Z's empirical conclusions can be theoretically deduced. It is hoped that future evaluations of cell formation algorithms will benefit from both the strengths and weaknesses of M&Z's work.

1.   Introduction

The basic notion of Group Technology (GT) is to decompose a manufacturing system into subsystems so that higher productivity can be achieved by exploiting the underlying commonality among different sub-processes of manufacturing. Cellular Manufacturing (CM) is one application of the GT philosophy. Using

---

[1] Requests for reprints should be sent to Q B. Chung, Villanova University, Villanova, PA 19085. USA.

CM, a firm's manufacturing system is organized into temporary work-cells (or simply, cells) that can exploit the advantages of a mass production system while maintaining the flexibility required by rapid changes in product mix and demand patterns. Each work-cell consists of a number of dissimilar machines grouped together to produce a set of parts with similar processing requirements. Major benefits of GT and CM include reduction in manufacturing lead times, in-process inventories, setup costs, and handling costs; and improvements in process planning and quality control (Boe & Cheng 1991; Burbidge 1992; Fazakerlay 1974; Gallagher & Knight 1984; Hyer 1984; Hyer & Wemmerlov 1984; Kusiak 1990; Pullen 1976; Suresh & Meredith 1985).

Temporary as the work-cells are, they must be rearranged every few months in response to the changing demand patterns and product mix in a manufacturing facility. Hence, an efficient algorithm for solving the cell formation (CF) is essential for companies that have adopted cellular manufacturing. Burbidge's production flow analysis (1963, 1975) was one of the earliest approaches to solving the CF problems. In his approach, a machine-part incidence matrix $[a_{ij}]$ is developed, which consists of "0" and "1" entries, where an entry of "1" indicates that machine i is used to process part j, while "0" indicates that machine i is not used to process part j. Typically, when an initial machine-part incidence matrix is constructed, clusters of machines processing similar parts are not clearly visible. Burbidge's method examined the similarity in the relationships between parts and machines and rearranged the initial incidence matrix to group together and identify clusters of highly compatible parts and machines. Following Burbidge's work, over the last three decades, numerous algorithms have been proposed to solve the CF problem. Several papers survey and classify the available algorithms (e.g., King and Nakoranchai, 1982; Wemmerlov and Hyer, 1986; Offodile, 1993; Vakharia and Wemmerlov, 1995; Mosier, Yelle, and Walker, 1997).

For practicing manufacturing managers, it would be nice to know as to which one of the available algorithms would be most effective and efficient for their specific situation. Tarsuslugil and Bloor (1979), Mosier and Taube (1985), Mosier (1989), Gupta (1991) and Shafer and Rogers (1993) represent some of the attempts at fulfilling this need. Unfortunately, none of these studies provide any definitive recommendations to practicing managers. Vakharia and Wemmerlov [V&W] (1995) provide a critical review of these studies. They identify the chosen data sets, partition levels, clustering techniques, and performance metrics as some of the reasons why "persistent and generalizable findings with respect to specific technique or index choices cannot be extracted from the studies …" in their review (V&W, 1995, p. 121). V&W then go on to conduct their own assessment of seven hierarchical clustering techniques combined with eight dissimilarity measures. Using 24 data sets and a large

number of partitions of data, they succeed in ruling out some techniques and measures as unsuitable for CF problems. V&W find that the performance of other techniques is sensitive to many factors, notably the underlying data and the stopping parameters. V&W's major conclusions are that "… more work is needed to link data structures to choice of clustering technique and dissimilarity measure," and "… we need to find measures and methods under which cell system solutions can be compared …" (V&W, 1995, p. 136). In short, V&W's study emphasizes the need for a better methodology of evaluating algorithms for solving the CF problem.

V&W's review seems to have missed Miltenburg and Zhang's [M&Z] (1991) evaluation of nine well-known CF algorithms. Prima facie, M&Z's methodology seems impressive. They code each one of the nine algorithms in C and run the code on a microcomputer to solve eight previously published problems and sixty randomly generated large problems. They use three different performance measures and provide a sophisticated analysis of variance of the many variables in their problem set. On each one of their performance measures, M&Z identify the algorithm that does the best. When they find that no algorithm performs better than all others on each one of their three performance measures, M&Z suggest a procedure for solving a CF problem using all three of their performance measures. Clearly, the amount of effort that M&Z have put in is very commendable, and M&Z's methodology is very promising.

In view of V&W's (1995) articulation of the need for a better methodology of evaluating algorithms for solving the CF problem and the prima facie promise of M&Z's (1991) work, we take a critical look at M&Z's methodology. We find that the primary performance measure proposed by M&Z effectively captures the objectives of a good solution to a CF problem and is useful for future studies. Unfortunately, we also find several flaws in M&Z's methodology. We show that:

(a)   M&Z may not have duplicated some of the algorithms precisely as the algorithm developers intended;

(b)   M&Z justify their primary performance measure ($M1$) on the basis of an incorrect appraisal of Chandrasekharan and Rajagopalan's [C&R's] (1986) grouping efficiency measure, η;

(c)   Instead of "enriching" the analysis, M&Z's secondary performance measures lead them to unnecessarily ambivalent results;

(d)   In view of their finding that none of the algorithms performs better than all other algorithms on each one of their performance measures, M&Z (1991) recommend their own procedure for solving a CF problem. We believe that M&Z's procedure may not work in most situations, and when it does work, it may be suboptimal.

(e)   Several of M&Z's empirical conclusions can be theoretically deduced; and finally,

(f)   One of M&Z's "findings" cannot be theoretically supported. It must be based on an incorrect analysis of their data.

It is hoped that that future evaluations of CF algorithms will benefit from both, the strengths and weaknesses of M&Z's work.

2.   Duplicating the Algorithms

Many analytical procedures have been developed for solving the CF problem. Wemmerlov and Hyer (1986) divide these procedures into four approaches, one of which "identifies part families and machine groups simultaneously." All of the nine algorithms evaluated by M&Z (1991) belong to this category. Since the algorithms are well known, and M&Z have done a reasonably good job of describing them, we shall not repeat those descriptions. To enhance the brevity of our presentation, we shall also refer to these algorithms by the abbreviations used by M&Z.

All CF algorithms begin with a given machine-part incidence matrix $A^{(i)} = \{a_{ij}\}$, where an element a_{ij} is either "1" if machine i is used to process part j, or "0" otherwise. They rearrange the rows and columns of $A^{(i)}$ to form a final matrix, $A^{(f)}$, where machines that process the same group of parts are close to each other, and parts that are processed by the same group of machines are close to one another. Typically, the non-zero entries in $A^{(f)}$ are clustered around the diagonal of the matrix. This diagonal arrangement facilitates the identification of the manufacturing cells, i.e., the simultaneous identification of part families and machine groups that would form specific work cells. However, as M&Z recognize, the nine algorithms belong to two different sets. Once $A^{(f)}$ is formed, algorithms in one set (e.g., ROC/ROC, or BEA) depend on manual methods to actually identify the cells, while algorithms in the other set (e.g., ISNC or SC-Seed) use computerized methods to identify the cells. Yet, M&Z describe their method of duplicating the algorithms as:

"In this research, each algorithm's solution to a cell formation problem was displayed as a final part/machine matrix in standard block diagonal form. The computer and researchers then interactively partitioned the matrix into cells so that the following measure, *M1*, was maximized." (M&Z, 1991, p. 51).

In other words, the nine algorithms are treated as if they all belonged to the first set. Consequently, M&Z may not have duplicated the algorithms in the second set precisely as the developers of those algorithms intended. First, since the algorithms in the second set are designed to let the computer identify the precise composition of the cells, there is no need for M&Z's approach of letting the

computer and researcher to "then interactively partition the matrix." Second, regardless of the criterion sought by the original algorithm (such as the maximization of h sought by the ISNC algorithm, or the maximization of bond energy sought by the BEA algorithm), M&Z's method of duplicating an algorithm is to maximize *M1*. Lastly, as shown in the next section, M&Z misrepresent the correct formula for h. Thus, even if they meant to duplicate the ISNC algorithm exactly, they would have maximized an incorrect criterion. Note also that M&Z make no attempt to show that the solution they attribute to an algorithm is identical to the one intended by the algorithm's developer(s).

3.   The Performance Measures

M&Z consider three measures of the goodness of a solution to a CF problem, *M1*, the primary measure, and *M2* and *M3*, two secondary measures. M&Z explain that a good CF solution seeks to attain the following two objectives: (1) within a cell, each machine is visited by many parts, and (2) few parts require processing on machines in other than their own cells. Ever since C&R's work (1986), these objectives have been accepted as the right objectives in the CF literature. Let us examine M&Z's measures one at a time to see if they capture the objectives of a good CF solution.

3.1. The Primary Measure, *M1*

M&Z define their primary measure in the following terms:
Suppose that the final machine-part matrix, $A^{(f)}$, of a CF problem is partitioned into k submatrices $\{D_r / r = 1, 2, ..., k\}$ where the submatrix D_r consisting of machines M_r and parts C_r is the r-th machine-part cell. Then *M1*, the primary grouping measure, is given by

$$M1 = \zeta_g = \zeta_u - \zeta_m, \quad -1 \leq \zeta_g \leq 1, \tag{1}$$

where ζ_u, a measure of the usage of the machines by parts in their respective cells, is given by

$$\zeta_u = \left(\sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right) \Bigg/ \left(\sum_r |M_r||C_r| \right), \quad 0 \leq \zeta_u \leq 1, \tag{2}$$

and ζ_m, a measure of movement of parts between cells, is given by

$$\zeta_m = 1 - \left(\sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right) \bigg/ \left(\sum_{i,j} a_{ij} \right), \quad 0 \le \zeta_m \le 1. \tag{3}$$

As M&Z point out, large values of ζ_u occur when each part in a cell uses most of the machines in that cell, and small values of ζ_m occur when there are few parts that require processing by machines outside their own cells. Hence, large values of ζ_u and small values of ζ_m are preferred. Consequently, large values of *M1* are also preferred. It is clear that *M1* does capture reasonably well the two objectives of cell formation. However, as shown below, M&Z justify *M1* on the basis of an incorrect appraisal of η. A formal definition of η given by C&R (1986) is

$$\eta = q\eta_1 + (1-q)\eta_2 \tag{4}$$

where η_1, C&R's usage measure is identical to M&Z's ζ_u, and η_2, C&R's measure for the *lack* of movement of parts between cells, is defined as

$$\eta_2 = 1 - \left(\sum_{\substack{i \in M \\ j \in N}} a_{ij} - \sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right) \bigg/ \left(MN - \sum_r |M_r| |C_r| \right) \tag{5}$$

while q is a weighting factor ($0 \le q \le 1$) that allows the analyst to alter the emphasis between η_1 and η_2, depending on the specific requirements of a problem. C&R (1986) recommend the use of $q = 0.5$ as the default value for a general problem.

It should be clear that C&R's η also captures the two objectives of cell formation reasonably well. In addition, compared to *M1*, η can be seen as a more flexible way of capturing those objectives because it allows for a problem-specific value of q. Yet, M&Z incorrectly claim that η is deficient in several ways. M&Z's mistake lies in a misrepresentation of η_2. Instead of Equation (5) above, M&Z (1991, p. 53) assume that η_2 is given by Equation (6) below.

$$\eta_2 = 1 - \left(\sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right) \bigg/ \left(MN - \sum_r |M_r| |C_r| \right) \tag{6}$$

Observe that, in the above expression, M&Z incorrectly substitute

$$\left(\sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right) \text{ in place of C\&R's } \left(\sum_{\substack{i \in M \\ j \in N}} a_{ij} - \sum_r \sum_{\substack{i \in M_r \\ j \in C_r}} a_{ij} \right).$$
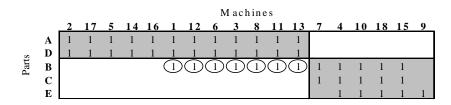
As a consequence of this misrepresentation, M&Z claim that η_2 does not satisfy the condition $0 \leq \eta_2 \leq 1$, when, in fact, it does. M&Z suggest that one reason they prefer $M1$ to η is that η_2 does not have the property that $\eta_2 = 0$ when there is no inter-group movement. However, in C&R's methodology, η_2 represent the *lack* of inter-group movement. As such, as can be seen from (5), the correct formula for η_2, when there is no inter-group movement, η_2 attains its largest possible value of 1. Note that in C&R's methodology, the final goodness measure, η, is calculated by taking a weighted average of η_1 and η_2, whereas in M&Z's methodology, $M1$ is calculated by subtracting ζ_m from ζ_u. In short, M&Z's justification of $M1$ is based on an incorrect appraisal of η.
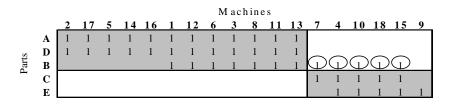
Although the calculus of $M1$ and η is different, and numerically, in most situations, η (which is always between 0 and 1) is larger than $M1$ (which can range between -1 to +1), because both $M1$ and η capture the two objectives of a good CF solution reasonably well, their values are likely to be highly correlated. However, there may be a reason to consider whether $M1$ still represents an improvement over η. After all, Seifoddini (1989), Boe and Cheng (1991), and Ng (1996) have all pointed out that, particularly at the default value of $q = 0.5$, η does not put adequate emphasis on a reduction of inter-cell movements. At the first sight, a comparison of η_2 and ζ_m formulae suggests that $M1$ may have overcome this problem. In η_2, the penalty for inter-group movement is calculated by the ratio of inter-group movements to all ungrouped entries in a matrix, whereas in ζ_m, the penalty for inter-group movement is calculated by the ratio of inter-group movements to all the "1" entries in a matrix. Thus, particularly in matrices with small densities, ζ_m is likely to represent a greater penalty for inter-group movement than η_2 does.

However, consider the two solutions (See Figure 1) to a CF problem cited by Seifoddini (1989). Both solutions use the same rearranged matrix and both give a two-cell solution. Solution **A** in Figure 1 requires seven inter-cell movements while Solution **B** requires only five inter-cell movements. On the basis of the criterion of minimum inter-cell movements, Solution **B** should be preferred to Solution **A**. Yet, as summarized in Part (c) of Figure 1, with $q = 0.5$, the η (=0.8914) for Solution **A** is larger than the η (=0.8676) for Solution **B**. Thus, as Seifoddini (1989), Boe and Cheng (1991), and Ng (1996) have suggested, with $q = 0.5$, η under-emphasizes inter-cell movements. M&Z's $M1$ measure does not remedy this situation. $M1$ (=0.7764) for Solution **A** is also larger than $M1$ (=0.7455) for Solution **B**.

On the other hand, $M1$ does represent an improvement over η in one situation. M&Z (p.51, Figure 2) present a 10-machine, 8-part problem with twenty-six "1" entries. M&Z show that this problem can have a 4-cell solution, or 3-cell solution, or 2-cell solution, or 1-cell solution. We summarize the relevant

Machines

Parts	2	17	5	14	16	1	12	6	3	8	11	13	7	4	10	18	15	9
A	1	1	1	1	1	1	1	1	1	1	1	1						
D	1	1	1	1	1	1	1	1	1	1	1	1						
B						(1)	(1)	(1)	(1)	(1)	(1)	(1)	1	1	1	1	1	
C													1	1	1	1	1	
E														1	1	1	1	1

(a) Solution **A**

Machines

Parts	2	17	5	14	16	1	12	6	3	8	11	13	7	4	10	18	15	9
A	1	1	1	1	1	1	1	1	1	1	1	1						
D	1	1	1	1	1	1	1	1	1	1	1	1						
B						1	1	1	1	1	1	1	(1)	(1)	(1)	(1)	(1)	
C													1	1	1	1	1	
E														1	1	1	1	1

(b) Solution **B**

	Solution **A**	Solution **B**
Number of cells	2	2
Inter-cell movements	7	5
η (at $q = 0.5$)	0.8914	0.8676
M1	0.7764	0.7455
M2	3.4511	3.4511
M3	1.3958	1.3958

(c) Comparison of two solutions

Note: The circled **1**'s in (a) and (b) represent the entries requiring inter-cell movements.

Figure 1. **Two CF solutions for a given final matrix**

properties of these four solutions in Table 1. Applying the two criteria of maximum within-cell usage and minimal inter-cell movements, the 3-cell solution should be judged the best. As Table 1 shows, *M1* (=0.812) for the 3-cell solution is largest of the *M1* values for the four solutions, while η (=0.933) for

the 4-cell solution is largest of the η values for the four solutions. Thus, in this case, *M1* seems to be superior to η. On the other hand, note that, in general, the values of *M1* and η are highly correlated.

In short, while M&Z seem to have justified *M1* on incorrect grounds, and *M1* may not be a perfect measure, it does capture the two CF objectives effectively. Furthermore, in some cases, it produces results that are superior to those produced by η. Hence, *M1* is worthy of use and further development in future studies.

3.2. The Secondary Measures, *M2* and *M3*

Even though *M1* effectively captures the two objectives of cell formation, M&Z introduce two secondary measures of performance, *M2* and *M3*, "to enrich the analysis" (1991, p. 55). We do not see how these secondary measures enrich M&Z's "clustering measure," *M2*, is given by the average Euclidean distance of a non-zero a_{ij} from the diagonal of the final matrix. Smaller values of *M2* are preferred to larger values. From the definition of *M2*, it is clear that, once the final matrix, $A^{(f)}$, is formed, *M2* is fully determined. Any alternative cell formations of that $A^{(f)}$, in terms of the number of cells formed or the specific machine-part membership of each cell, has no effect on the value of *M2*. Thus, *M2* can add nothing to *M1*'s evaluation of how well an algorithm performs in terms of the two objectives of cell formation.

M&Z's other secondary measure, *M3*, is based on the premise underlying the Bond Energy Algorithm (BEA), namely that, the non-zero elements of a machine-part matrix should be located close to each other. We will not reproduce the formula here. Succinctly, *M3* represents the average closeness of a non-zero a_{ij} from the other non-zero a_{ij}s. Thus, larger values of *M3* are preferred to smaller values. As in the case of *M2*, given the definition of *M3*, it follows that, once the final matrix, $A^{(f)}$, is formed, *M3* is fully determined, and alternative cell formations of the same $A^{(f)}$ do not affect the value of *M3*. Thus, *M3* also can add nothing to *M1*'s evaluation of how well an algorithm performs in terms of the two objectives of cell formation.

Consistent with the above argument, note that, in M&Z's Table 2 (1991, p. 54), the *M2* and *M3* values change only between the initial and the final matrix, but not among the one, or two, or three-cell solutions of the final matrix. In short, M&Z's secondary performance measures add no new insights. Indeed, instead of helping definitive conclusions, the inclusion of these secondary measures leads M&Z to arrive at such an ambivalent conclusion as, "No solution algorithm was found to be better than all other algorithms on all performance measures for both the randomly generated data set and the literature data set." (M&Z, 1991, p. 68).

Had they employed *M1* as the only measure, based on their data, M&Z would have reported simply that the ISNC algorithm was significantly better than any other algorithm. Such a definitive conclusion would have been considerably more valuable than the conclusions M&Z actually report.

Table 1. Comparison of η and ζ for the four solutions in M&Z's Example (M&Z, 1991, p.51, Figure 2)

	Number of Cells in the Solution			
	1	**2**	**3**	**4**
m	10	10	10	10
n	8	8	8	8
m X *n*	80	80	80	80
Sum of all non-zero entries in the matrix	26	26	26	26
Sum of all *m* X *n* in cells	80	42	27	25
Sum of all non-zero entries in cells	26	25	24	23
Number of inter-cell moves	0	1	2	3
$\eta_1 = \zeta_u$	0.325	0.595	0.889	0.920
η_2	1.000	0.974	0.962	0.945
η	0.663	0.784	0.926	**0.933**
ζ_m	0.000	0.038	0.077	0.115
ζ *(M1)*	0.325	0.557	**0.812**	0.805

Note: The 1-cell solution and the 2-cell solution do not fulfill the first objective of a good CF solution. I.e., their $\eta_1 = \zeta u$ values are unacceptable. The 3-cell solution and the 4-cell solution both accomplish the first objective reasonably well. However, because the 3-cell solution has only 2 inter-cell movements while the 4-cell solution has 3 inter-cell movements, the 3-cell solution should be preferred. η fails insofar as its value for the 4-cell solution is larger. *M1* succeeds in correctly identifying the more desirable solution.

4.   M&Z's Suggested Procedure for Solving a CF Problem

In view of their finding that none of the algorithms performs better than all other algorithms on each one of their performance measures, M&Z (1991) recommend their own procedure for solving a CF problem using all three of their performance measures. They describe this procedure as:
"A solution algorithm produces a solution which is displayed as a final part/machine matrix in standard block diagonal form. Values for the clustering measure and the bond energy measure are computed. If both values are sufficiently good, the user (or computer) proceeds to partition the final

part/machine matrix into cells in such a way that the value of the grouping measure is maximized. If either the clustering or the bond energy measure is poor, then another algorithm is selected, or selected parts and/or bottleneck machines are removed, selected routings are changed, etc., and the process is repeated." (M&Z, 1991, p. 66).

M&Z do not indicate whether they have actually experimented with this procedure and whether it has proved to be practical. However, we suspect that requiring both M2 and M3 to be sufficiently good may eliminate a majority of algorithms from further consideration. This is because while *small* values of *M2* and *large* values of *M3* are desirable, large values of *M3* (the bond energy measure) often result in large values of *M2* (the clustering measure). As M&Z themselves note, "Where the final part/machine matrix is in standard block diagonal form, a high bond energy measure leads to a high grouping measure (and a high clustering measure)." (M&Z, 1991, p. 55).

Let us consider the implications of this correlation between the values of *M2* and *M3* for M&Z's procedure. While M&Z do not indicate as to what values of *M2* and *M3* are "sufficiently good" or what are "poor" in implementing their procedure, consider a simple approach to determine such good and poor values. Since in M&Z's Table 6, the mean *M2* values for the various algorithms range from 14.397 to 18.671 (the mid point of the range being 16.534), and since smaller values of *M2* are preferred, a simple rule may be that *M2* values of 16.534 *or less* are "sufficiently good," and values above 16.534 are "poor." Similarly, since in M&Z's Table 7, the mean *M3* values for the various algorithms range from 0.4794 to 0.7626 (the mid point of the range being 0.6209), and since larger values of *M3* are preferred, a simple rule may be that *M3* values of 0.6209 *or greater* are "sufficiently good," and values below 0.6209 are "poor."

Now, note from M&Z's Table 6 that when the BEA algorithm is used, 95% of the problems have an *M2* score ranging between 17.14 to 17.86. Since these values are "very poor" by our interpretation of M&Z's procedure, BEA would not be an acceptable algorithm in these 95% problems. At the same time, from M&Z's Table 7, we see that, except for BEA, every algorithm's 95% interval estimate of *M3* is substantially below our desired value for *M3*, the largest interval being 0.505 to 0.527 for the ISNC algorithm. In other words, if M&Z's procedure is followed, in the cases of a large majority of problems, none of the eight algorithms included in Tables 6 and 7 would score "sufficiently good" on both the *M2* and the *M3* criteria. Thus, none would advance to the level of partitioning the final machine-part matrix into cells!

Of course, this need not happen if the "sufficiently good" *M2* and *M3* values are defined at points that are substantially below the midpoints of their ranges. Then

the question is whether *M2* and *M3* really serve as good screening devices for eliminating the not-so-good algorithms.

Finally, in M&Z's procedure, if the very first attempted algorithm results in a final matrix with acceptable *M2* and *M3* values, then that matrix is used to maximize *M1* and no other algorithm is considered. This procedure is suboptimal to one where the final machine/part matrices generated by several different algorithms (regardless of their *M2* and *M3* values) are advanced to the step of partitioning to maximize *M1*, and then the solution with the best *M1* value is selected for implementation.

5.   Statistical Analysis and Empirical Findings

The amount of effort M&Z put into their analysis is overwhelming. They code nine different algorithms, and use each algorithm to solve eight problems from the open literature and sixty randomly generated large problems. They evaluate each solution on three different performance measures. Finally, M&Z conduct a statistically sound analysis of variance of the major variables in their data sets. Unfortunately, several of M&Z's empirical conclusions could have been theoretically deduced. At the same time, one of M&Z's conclusions simply cannot be supported theoretically.

First, some examples of conclusions that can be theoretically deduced:

M&Z (1991, p. 60) conclude that, in terms of their primary measure, *M1*, the ISNC algorithm is better than all other algorithms. Note that the ISNC algorithm maximizes C&R's (1986) η. As discussed earlier, the underlying calculus of *M1* and η is somewhat different and for a given solution. Typically, η is numerically larger than *M1*. Although there are circumstances where the two measures provide different rankings for alternative solutions to a problem, overall, both *M1* and η capture the two goals of CF formation reasonably well and consequently, these measures are highly correlated. Thus, it follows that an algorithm that maximizes η would do very well on the *M1* measure.

The most important piece of information from M&Z's work is that on the basis of the *M1* measure, the ISNC algorithm is *significantly* better than all other algorithms. Unfortunately, the importance of this piece of information is lost because of M&Z's use of their secondary performance measures.

M&Z (1991, p. 61) find that, in terms of the clustering measure, *M2*, the ROC/ROC algorithm is better than all other algorithms. Again, this conclusion can be theoretically derived. The fact is that *M2* measures the average distance of a non-zero entry from the diagonal, with the understanding that smaller distance is preferred. Naturally then, the ROC/ROC algorithm would score the best on the

M2 measure, since it is designed precisely to minimize the average distance of the non-zero entries from the diagonal of a matrix.

Similarly, as M&Z (1991, p. 62) themselves recognize, it is no surprise that BEA outperforms any other algorithm in terms of *M3*, the bond energy measure. BEA is designed to maximize bond energy.

Of more substantial concern is the fact that one of M&Z's empirical conclusions simply cannot be supported theoretically. M&Z (1991, p. 60) conclude that, for their data set, "the number of machines has a larger effect on the performance measures than does the number of parts." This just cannot happen in any CF problem situation. All of the CF formation algorithms studied by M&Z basically rearrange the rows and columns of a $m \times n$ matrix, and it does not matter whether the machines are represented as rows and parts as columns, or machines are represented as columns and parts as rows. Thus, any solution to an m-machines, n-parts problem has an exactly equivalent solution to an m-parts, n-machines problem.

In explaining the reasoning behind their conclusion, M&Z says,

"… in part 2 of Table 5, 95% interval estimates are computed for, among other things the mean grouping measure for problems having 50 machines and for problems having 50 parts. Since the interval for problems having 50 parts is higher than the interval for problems having 50 machines, (recall that high grouping measure is preferred) and the intervals do not overlap, we conclude that problems having 50 machines are more difficult than problems having 50 parts." (M&Z, 1991, p. 60)

However, this non-overlap of intervals is logically impossible. Note that in M&Z's data set, of the 240 problems involving 50 machines, 120 problems involve 50 parts. These 120 problems are identical to 120 of the 240 problems involving 50 parts. Thus, at least for these 120 problems, the mean grouping measures should be identical. Now, a 95% interval of 50-machine data (consisting of 240 problems) must include at least 90% of these 120 problems. Similarly a 95% interval of 50-part data (consisting of 240 problems) must also include at least 90% of these 120 problems. Yet, in their Table 5, M&Z report that the 95% interval estimate of the 50-machine data ranges from 0.198 to 0.203 while the 95% interval estimate of the 50-part data ranges from 0.219 to 0.225. The only possible explanation for the reported non-overlap of the intervals is that M&Z must have inadvertently made some computational error. M&Z's conclusion on this matter simply cannot be supported on theoretical grounds.

6.   Conclusion

With the proliferation of algorithms to solve the work-cell formation problem, practicing manufacturing managers need assistance in identifying an algorithm that would be most effective and efficient for their specific situation. Miltenburg and Zhang's (M&Z) (1991) work on the evaluation of nine well-known cell formation algorithms represents an important contribution towards that goal. The primary performance measure proposed by M&Z effectively captures the objectives of a good solution to a cell formation problem and is worthy of use in future studies. In this note we have also pointed out some of the shortcomings of M&Z's methodology. It is hoped that that future evaluations of cell formation algorithms will benefit from both, the strengths and weaknesses of M&Z's work.

References

1.   Boe, W. J. & Cheng, C. H. (1991). A Close Neighbour Algorithm for Designing Cellular Manufacturing Systems. *International Journal of Production Research,* vol. 29, no. 10, pp. 2097-2116.

2.   Burbidge, J. L. (1963). Production Flow Analysis. *Production Engineer*, vol. 42, no. 12, pp. 742-752.

3.   Burbidge, J. L. (1975). *The Introduction of Group Technology*. New York: Halsyer Press and John Wiley.

4.   Burbidge, J. L. (1992). Production Flow Analysis for Planning Group Technology. *Journal of Operations Management*, vol. 10, no. 1, pp. 5-27.

5.   Chandrasekharan, M. P. & Rajagopalan, R. (1986). An Ideal Seed Non-hierarchical Clustering Algorithm for Cellular Manufacturing. *International Journal of Production Research*, vol. 24, no. 2, pp. 451-464.

6.   Fazakerlay, G. M. (1974). Group Technology: Social Benefits and Social Problems. *The Production Engineer*. October 1974, pp. 383-386.

7.   Gallagher, C. C. & Knight, W. A. (1986). *Group Technology: Production Methods in Manufacturing*. New York: Halsted Press.

8.   Gupta, T. (1991). Clustering Algorithms for the Design of Cellular Manufacturing Systems - An Analysis of their Performance. *Computers in Industrial Engineering.* vol., 20, no. 4, pp. 461-468.

9.   Hyer, N. L. (1984). The potential of Group Technology for U.S. Manufacturing. *Journal of Operations Management*. Vol. 4, no. 3, pp. 183-202.

10. Hyer, N. L. & Wemmerlov, U. (1984). Group Technology and Productivity. *Harvard Business Review*, July-August 1984, pp. 140-149.

11. King, J. & Nakoranchai V. (1982). Machine Component Group Formation in Group Technology: Review and Extension. *International Journal of Production Research.* Vol. 20, no. 2, pp. 117-133.

12. Kusiak, A. (1990) *Intelligent Manufacturing Systems*, Englewood Cliffs, New Jersey: Prentice Hall.

13. Miltenburg, J. & Zhang, W. (1991). A Comparative Evaluation of Nine Well-Known Algorithms for Solving the Cell Formation Problem in Group Technology. *Journal of Operations Management*, vol. 10, no. 1, pp. 44-72.

14. Mosier, C. T. (1989). An Experiment Investigating the Application of Clustering Procedures and Similarity Coefficients to the GT Machine Cell Formation Problem. *International Journal of Production Research.* Vol. 27, no. 10, pp. 1811-1835.

15. Mosier, C. & Taube, L. (1985). Weighted Similarity Measure Heuristics for Group Technology Machine Clustering Problem. *Omega*, vol. 13, no. 6, pp. 577-583.

16. Mosier, C. T., Yelle, J. & Walker, G. (1997). Survey of Similarity Coefficient Based Methods as Applied to the Group Technology Configuration Problem. *Omega*, vol. 25, no. 1, pp. 65-71.

17. Ng, S. M. (1996). On the Characterization and Measure of Machine Cells in Group Technology. *Operations Research*, vol. 44, No. 5, pp. 735-744.

18. Offodile, O. F. (1993). Machine Grouping in Cellular Manufacturing. *Omega*, vol. 21, no. 1, pp. 35-52.

19. Pullen, R. D. (1976). A Survey of Cellular Manufacturing Cells. *The Production Engineer*, vol. 55, no. 9, pp. 451-454.

20. Seifoddini, H. (1989). A Note on the Similarity Coefficient Method and the Problem of Improper Machine Assignment in Group Technology Applications. *International Journal of Production Research*, vol. 27, no. 2, pp. 1161-1165.

21. Shafer, S. M. & Rogers, D. F. (1993). Similarity and Distance Measures for Cellular Manufacturing. Part II An Extension and Comparison. *International Journal of Operations Research*, vol. 31, no. 6, pp. 1315-1326.

22. Suresh, N. C. & Meredith, J. R. (1985). Achieving Factory Automation through Group Technology. *Journal of Operations Management*, vol. 5, no. 2, pp. 151-167.

23. Tarsuslugil, M. & Bloor, J. (1979). The Use of Similarity Coefficients and Cluster Analysis in Production Flow Analysis. *Proceedings of 20th International Machine Tool Design and Research Conference*, Birmingham, UK. September, pp. 525-532.

24. Vakharia A. J. & Wemmerlov, U. (1995). A Comparative Investigation of Hierarchical clustering Techniques and Dissimilarity Measures Applied to Cell Formation Problem. *Journal of Operations Management*, vol. 13, pp. 117-138.

25. Wemmerlov, U. & Hyer, N. L. (1986). Procedures for the Part Family/Machine
     Group Identification Problem in Cellular Manufacturing. *Journal of Operations
     Management*, vol. 6, no. 2, pp. 125-167.

Special Issue on
Singular Boundary Value Problems for Ordinary Differential Equations

Call for Papers

The purpose of this special issue is to study singular boundary value problems arising in differential equations and dynamical systems. Survey articles dealing with interactions between different fields, applications, and approaches of boundary value problems and singular problems are welcome.

This Special Issue will focus on any type of singularities that appear in the study of boundary value problems. It includes:

- Theory and methods
- Mathematical Models
- Engineering applications
- Biological applications
- Medical Applications
- Finance applications
- Numerical and simulation applications

Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://www.hindawi.com/journals/bvp/guidelines.html. Authors should follow the Boundary Value Problems manuscript format described at the journal site http://www.hindawi.com/journals/bvp/. Articles published in this Special Issue shall be subject to a reduced Article Processing Charge of €200 per article. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at http://mts.hindawi.com/ according to the following timetable:

Manuscript Due	May 1, 2009
First Round of Reviews	August 1, 2009
Publication Date	November 1, 2009

Lead Guest Editor

Juan J. Nieto, Departamento de Análisis Matemático, Facultad de Matemáticas, Universidad de Santiago de Compostela, Santiago de Compostela 15782, Spain; juanjose.nieto.roig@usc.es

Guest Editor

Donal O'Regan, Department of Mathematics, National University of Ireland, Galway, Ireland; donal.oregan@nuigalway.ie