The Enhancement of the Overall Group Technology Efficacy using Clustering Algorithm for Cell Formation

Lan Xuan Phung*, Trung Kien Nguyen, Son Hoanh Truong

School of Mechanical Engineering, Hanoi University of Science and Technology,
Hanoi, 100000, Vietnam, lan.phungxuan@hust.edu.vn,
trung.nguyenkien@hust.edu.vn, son.truonghoanh@hust.edu.vn
*Corresponding author

Abstract: Cellular manufacturing is a principal application of group technology in which machine cells and part families are generated based on their similarity in the production process to minimize overall movement cost and maximize machine utilization by using complex mathematical programming procedures or computer tools with a lot of computational effort and time to solve problems. In this study, the clustering analysis based on a similarity coefficient is developed to efficiently solve cell formation problems in both single and multiple process routings. A novel similarity coefficient is developed to integrate operation sequence especially adjacent operation, processing time, production volume, machine capacity, and multi-visits to minimize the number of actual inter-cell moves and voids in machine cells. An improved clustering algorithm is proposed for grouping machines into cells and simultaneously determining the machine sequence in cells to reduce intra-cell moves as well as selecting the best process routing for each part. The practical effectiveness of the proposed method is demonstrated through computational experiments involving eighteen test instances, varying in scale from small to large problems. When compared to other complex methods, the proposed approach not only enhances overall group technology efficacy but also significantly reduces computational time, making it a highly promising and practical solution for addressing cellular manufacturing challenges.

Keywords: cell formation; overall group technology efficacy; similarity coefficient; clustering algorithm

---

1 Introduction

Lean production focuses on optimizing efficiency, reducing costs, minimizing waste and improving overall quality. Cellular manufacturing (CM) is one of the key principles used to achieve these objectives and plays a crucial role in lean production. [1, 2]. CM is based on group technology principles by separating machines into groups with similar characteristics and distributing parts into part families to achieve higher productivity and flexibility compared to traditional manufacturing [3, 4]. The principal problem in the CM system is cell formation to minimize the number of moves and voids in the cells, maximizing the utilization of machines, equipment, and labour in the production process. By grouping machines and processes in self-contained cells, CM reduces material handling and setup times. This results in a more streamlined and efficient production flow, minimizing the time required to move materials between workstations and improving overall process efficiency. For over three decades, various production-oriented methods have been introduced for cell formation problems such as mathematical programming, heuristic and (meta-) heuristic algorithm, graph partitioning, and most commonly hierarchical method. The mathematical programming methods focus on developing the model to maximize the total operations in each cell and/or to minimize the moves between cells [5, 6]. Some studies addressed heuristic algorithms based on flow-matrix to solve cell formation problems (CFP) and machine layout generation [7-9]. Another heuristic approach with two stages based on the similarity score was developed to produce the cellular facility layout [10]. Recently, (meta-) heuristic methods such as simulated annealing algorithm [11], genetic algorithm [12, 13], ant colony algorithm [14], etc. have been introduced as promising methods to obtain "acceptable" solutions with the optimization in "reasonable" computational time. Adaptive resonance theory, a class of neural networks, has also demonstrated the ability to solve the CFP [15, 16]. The graph partitioning approach represents a graph with nodes and arc weight defined as machines and similarity measure of machine pair, respectively to minimize inter-cell travels [17, 18]. In other work, the hybrid algorithm was developed to solve complex, multiple objective optimizations for the CFP [19]. The hierarchical method is the main approach to cluster analysis in which the similarity or distance function and hierarchy of clusters are determined [20, 21]. Besides time-saving and computation calculation, hierarchical clustering methods based on similarity coefficient (SC) are more flexible in integrating various production data into CFP such as operation sequence, production volume, processing time, machine capacity, etc. [22-24].

To address the problem in the hierarchical method, most researchers start from the machine-part matrix to get a transformed matrix in a more structured form as diagonal blocks by grouping machines into cells and parts into families to minimize the number of out-of-blocks referred to as exceptional elements. There are four steps in this method: production data collection, SC determination for each machine pair, clustering algorithm applying for cell formation based on SC, and part family generation. The most important production input data, machine-part matrix MP $[m_{ij}]$, is first determined. The binary matrix is commonly used as an MP matrix in simple CFP, where $m_{ij} = 1$, if part i proceeded on machine j and otherwise [14, 25]. In other work, a number presented for the operation sequence index is used instead of using the value "1" in the MP matrix [26-28]. Besides the MP matrix, processing time, and production volume, machine capability is also initial production data for the problem. Based on production input data, the SC is determined by different formulas in the literature. McAuley developed the original SC that considers only the number of machine pairs proceeding with both operations and/or at least one operation [25]. T. Gupta and H. I. Seifoddini incorporated processing time and production volume in the SC formula [26]. The operation sequence and multi-visit problem are also considered in the SC formula [27]. This study focuses on two specific operations: the first and last operations, primarily aimed at reducing the number of inter-cell moves. However, the adjacent operation, which plays a crucial role in diminishing duplicate or repeated inter-cell moves, has not been addressed.

In the next step, the clustering algorithms based on the SC are applied to separate and group machines into cells. There are some common algorithms for group machines to form machine cells (MC) such as the single linkage algorithm (SLINK) [25] and the average linkage algorithm (ALINK) [28]. Linear cell clustering (LCC) is a method employed to generate consistent machine groups by linearly comparing the similarity scores between two machines [20]. LCC is known as a fast method in computing, a simple algorithm, and an easy solution in programming. Finally, parts can be allocated to families corresponding with assigned machines to optimize the inter-cell moves and voids in MC. In practical works, each part may have alternative process routings that make CFP more complicated. Thus, the formation of machine cells (MCs), part families, and selection of best routing need to be considered in CFP to achieve the overall objectives [29-31]. Intra-cell moves are classified into two main types: forward moves, including in-sequence and by-pass movements, and backward moves. Due to the reverse material flow, the cost of backward moves is significantly higher than that of other moves. Thus, minimizing backward moves becomes a key factor in reducing intra-cell move cost. However, most previous clustering algorithms mainly focus on reducing the inter-cell moves regardless of paying attention to minimising the backward moves and the compactness in CFP for the multi-routing problem. In the proposed work, the operation sequence, multi-visits, multi-routings, processing time, production volume and machine capacity are integrated into a similarity coefficient in a unique model to solve CFP for both single-routing and multi-routing problems.

To evaluate the group technology performance, several measuring methods are used in the literature. Three types of evaluation performance measurements were used including global efficiency, group efficiency, and group technology efficiency [32]. These measurements are quite individual and insufficient to provide an overall evaluation of the effectiveness of cell formation. To solve this problem, Nair and Narendran developed bond efficiency incorporating both inter-cell moves and compactness [27]. Lee and Anh proposed group technology efficacy (GTE) as shown in Eq. 1 for the performance measurement of cell formation considering both actual inter-cell moves and cell compactness [33]. Based on Lee's GTE, S. Raja defined GTE considering the backward moves as presented in Eq. 2 [34]. In this measurement, the effect of inter-cell moves, and backward moves are equal although they are quite different in the actual production. Moreover, the definition of backward move in this study only covers the operations inside a cell containing the part while this move exists even for external operations of the part in a different cell.

$$Lee's\ GTE = \frac{1 - \frac{AIM}{PIM}}{\frac{NV}{NI}} \tag{1}$$

$$Raja's\ GTE = \frac{1 - \frac{AIM + BM}{PIM}}{\frac{NV}{NI}} \tag{2}$$

$$PIM = NO - NP$$

where AIM is number of actual inter-cell moves; BM is number of backward moves; PIM is number of possible internal moves; NV is the number of voids; NI is number of operation inside machine cells; NO is the number of operation outside machine cells; NP is the number of parts.

The purpose of this article is to solve CFP by considering the actual inter-cell moves and actual backward moves and the compactness through the novel SC and modified clustering algorithm. In this study, both cell formation and machine sequence are solved simultaneously in consideration of the most important factors including operation sequence, processing time, machine capacity, production volume, multi-visits and multi-routings. The results contribute to increasing the overall GTE for both single routing and multi-routing problems.

2 Methodology

2.1 Proposed Similarity Coefficient

Besides considering the important factors such as operation sequence, processing time, production volume, and multi-visits, the proposed SC integrates the key factors including adjacent operation and the compactness in the calculation. The proposed SC between machine j and machine k is determined by S_{jk} as shown in Eq. 3.

$$S_{jk} = \frac{\sum_{i=1}^{n} A_{jk}\, w_i}{\sum_{i=1}^{n} A_{jk}\, w_i + \sum_{i=1}^{n} O_{jk}\, w_i + C_{jk}/N} \tag{3}$$

where A_{jk} refers to the total actual flow through machine j proceeded by part i which uses both machine j and machine k as calculated in Eq. 4; O_{jk} refers to the actual flows to or from machine j only (excluding machine k) made by the part i;

C_{jk} is the void possibility between machine j and machine k; N is the total routes of all parts; w_i is the production volume of part i.

$$A_{jk} = \sum_{\mathrm{r}=1}^{n_{ir}} \left(A_{irjk} + A_{irkj}\right)\left(T_{irjk} + Z_{irjk}\right) \tag{4}$$

$$A_{irjk} = \sum_{p=1}^{n_{irj}} \sum_{q=1}^{n_{irk}} a_{ij}^{pq}$$

$$\text{where } a_{ij}^{pq} = \begin{cases} 0 \; if \; b_{irj}^{p} = 0 \; or \; b_{irk}^{q} = 0 \\ 1 \; if \; b_{irj}^{p} = 1 \; or \; r_i \\ 2 \; otherwise \end{cases} \tag{5}$$

where A_{irjk} indicates the total actual flow through machine j proceeded by part i which uses both machine j and machine k in the route r as shown in Eq. 5 n_{ir} refers to the number of routes associated with part i; n_{irj} is the number of operations in which part i processes on machine j in the route r; a_{ij}^{pq} indicates the flows through machine j proceeded by part i which uses both machine j for time p and machine k for time q in the route r; b_{irj}^{p} indicates the operation index if part i moves through machine j for time p in the route r;

T_{irjk} is the proportion of minimal and maximal total processing time of part i in the route r spending on both machine j and k as shown in Eq. 6.

$$T_{irjk} = \frac{min\left(\sum_{\mathrm{r}=1}^{n_{ir}} \sum_{p=1}^{n_{irj}} \frac{t_{irj}^{p}}{C_j}, \sum_{\mathrm{r}=1}^{n_{ir}} \sum_{\mathrm{q}=1}^{n_{irk}} \frac{t_{irk}^{q}}{C_k}\right)}{max\left(\sum_{\mathrm{r}=1}^{n_{ir}} \sum_{p=1}^{n_{irj}} \frac{t_{irj}^{p}}{C_j}, \sum_{\mathrm{r}=1}^{n_{ir}} \sum_{\mathrm{q}=1}^{n_{irk}} \frac{t_{irk}^{q}}{C_k}\right)} \tag{6}$$

where t_{irj}^{p} is the processing time if the part i uses machine j for the time p in the route r; C_j is the machine capacity of machine j

Z_{irjk} is the proportion considering the adjacent operation between both machines j and k for part i in route r as shown in Eq. 7. It is the key factor to reduce the inter-cell moves.

$$Z_{irjk} = \frac{\sum_{p=1}^{n_{irj}} \sum_{q=1}^{n_{irk}} z_{irj}^{1\mathrm{pq}}}{\sum_{p=1}^{n_{ij}} \sum_{q=1}^{n_{ik}} z_{irj}^{2\mathrm{pq}}}$$

$$\text{where } z_{irj}^{1\mathrm{pq}} = \begin{cases} 1 \; if \; (both \; b_{irj}^{p} \; and \; b_{irk}^{q} \neq 0 \; and \; \left|b_{irj}^{p} - b_{irk}^{q}\right| = 1)\,) \\ 0 \; otherwise \end{cases} \tag{7}$$

$$z_{irj}^{2\mathrm{pq}} = \begin{cases} 1 \; if \; both \; b_{irj}^{p} \; and \; b_{irk}^{q} \neq 0 \\ 0 \; otherwise \end{cases}$$

where $z_{irj}^{1\mathrm{pq}}$ is the number of adjacent operations between machine j in the time p and machine k in the time q in the route r by part i. $z_{irj}^{2\mathrm{pq}}$ is the number of operations proceeded on both machine j in the time p and machine k in the time q in the route

r by part i. Only the actual flows to or from machine j (O_{jk}) is calculated by Eq. 8; where O_{irjk} is the actual flows to or from machine j (excluding machine k) made by the part i in the route r as shown in Eq. 9; o_{irj}^{pq} is the flow through machine j in the time p (excluding machine k in the time q) made by part i in the route r

$$O_{jk} = \sum_{r=1}^{n_{ir}} \left(O_{irjk} + O_{irkj}\right) \tag{8}$$

$$O_{irjk} = \sum_{p=1}^{n_{irj}} \sum_{q=1}^{n_{irk}} o_{irj}^{\mathrm{pq}}$$

$$\text{where } o_{irj}^{\mathrm{pq}} = \begin{cases} 0\ if\ b_{irj}^{p} = 0\ or\ b_{irk}^{q} \neq 0 \\ 1\ if\ b_{irj}^{p} = 1\ or\ r_i \\ 2\ otherwise \end{cases} \tag{9}$$

C_{jk} is defined by the absolute value of the difference between the total number of parts visiting machine j and the total number of parts visiting machine k in route r described in Eq. 10. This factor takes into account the impact of voids in cells.

$$C_{jk} = \left| \sum_{i=1}^{n} \sum_{\mathrm{r}=1}^{n_{ir}} c_{irj} - \sum_{i=1}^{n} \sum_{\mathrm{r}=1}^{n_{ir}} c_{irk} \right|$$

$$\text{where } c_{irj} = \begin{cases} 1\ if\ m_{irj} \neq 0 \\ 0\ otherwise \end{cases} \tag{10}$$

c_{irj} is the number of operations proceeded by part i by machine j in the route r

2.2 Improved Clustering Algorithm

The clustering algorithm is modified from the LCC algorithm by incorporating the sort algorithm, which facilitates the simultaneous identification of the appropriate machine positions within the cell during MC formation, and the selection algorithm, used for determining the best routing. The flowchart of modified clustering algorithm for grouping and generating machine cell is shown in Figure 1 and the algorithm includes the following steps:

Step 1: Acquire the matrices for machine-part ($MP[ir, j]$), process time ($PT[ir, j]$), production volume ($PV[i]$), and machine capacity ($C[j]$). Set value for the similarity coefficient threshold ($sThreshold$) for group merging considerations utilized at step 3b and weight factor (q) as the ratio between backward and inter-cell move cost used at step 6.

Step 2: Calculate S_{jk} for each machine pair j and k and generate a similarity coefficient matrix. Arrange the SC in decending order.

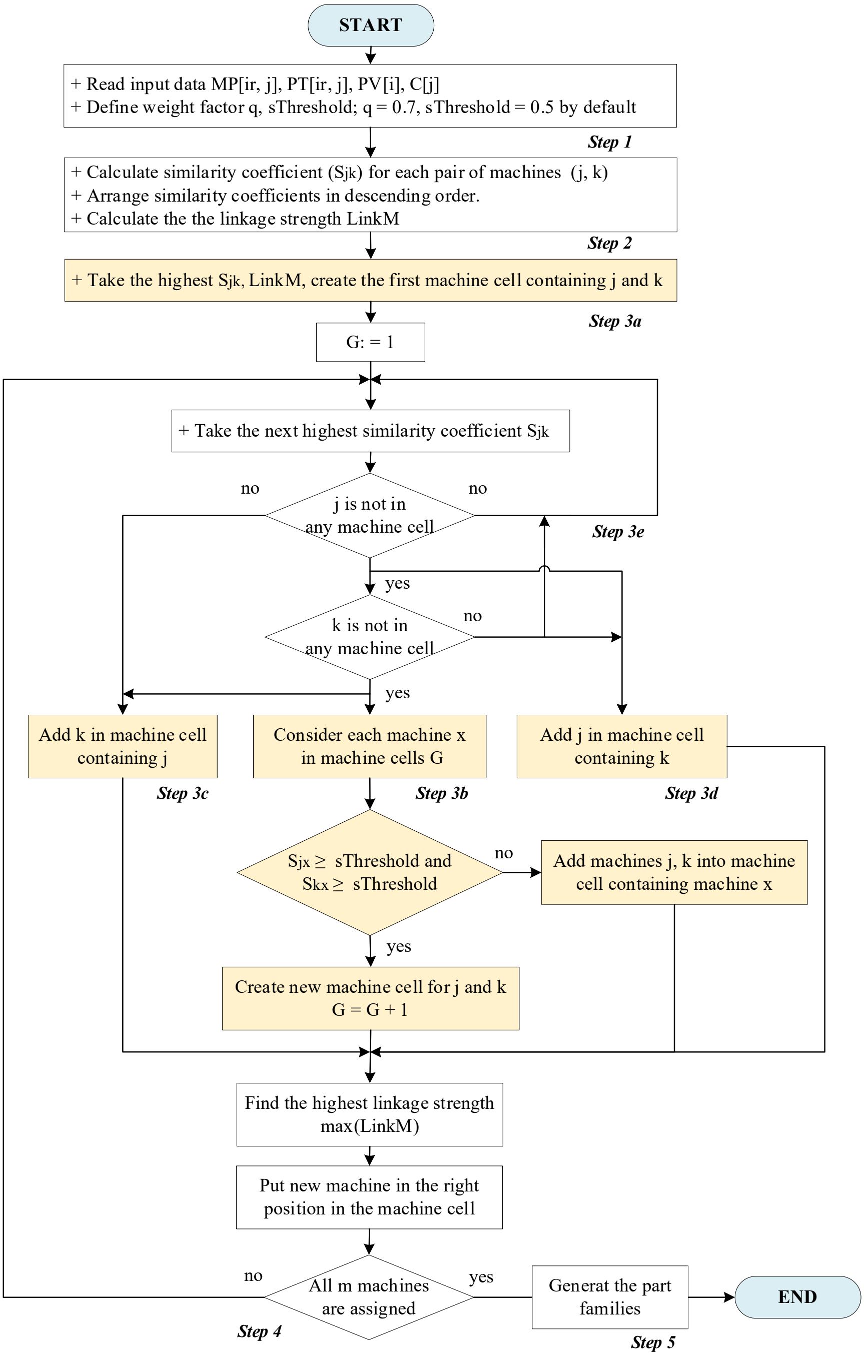


Figure 1

The flowchart of modified clustering algorithm for generating the machine cell

Step 3: Cluster machines into MC by considering the machine pair with the highest S_{jk}. During processing, the position of a machine in the assigned MC is also determined by calculating and comparing the linkage strength ($LinkM_{jk}$) between machines j and machine k. The order with the highest $LinkM_{jk}$ is given priority. With machines j and k in a cell, the linkage strength $LinkM_{jk}$ is calculated based on the total backward moves between them as shown in Eq. 11.

$$LinkM_{jk} = \sum_{i=1}^{n}\sum_{\mathrm{r}=1}^{n_{ir}}\sum_{p=1}^{n_{irj}}\sum_{q=1}^{n_{irk}} L_{jkpq}^{ir}$$

$$\text{where } L_{jkpq}^{ir} = \begin{cases} = 1\ if\ (b_{irk}^{q} - b_{irj}^{p} = 1\ and\ b_{irj}^{p} \neq 0\ and\ b_{irk}^{q} \neq 0)\ ) \\ = 0.25\ if\ (b_{irk}^{q} - b_{irj}^{p} \geq 2\ and\ b_{irj}^{p} \neq 0\ and\ b_{irk}^{q} \neq 0)\ ) \\ = 0\ otherwise \end{cases} \quad (11)$$

In this step, there are some cases as follows:

3a) If no MC is generated, the first MC is created containing machines j and k. The order in the first cell can be (j, k) or (k, j) depending on the comparison values of $LinkM_{jk}$ and $LinkM_{kj}$ as above-mentioned.

3b) If there is at least one generated MC and both machines j and k have not been assigned to any MC yet. Then, the merging group process is applied to determine whether both machines are assigned in a new MC or a pre-generated MC. Let x be any machine in generated MC. If both S_{jx} and S_{kx} are larger than $sThreshold$ and the number of machines in the merged cell is not larger than the maximum expected the number of machines in cells, machine j and k are merged into the assigned MC. Otherwise, a new MC is created for the machine pair.

3c) If machine j has already belonged to the cell, and machine k has not been assigned. Machine k is allocated to the cell that includes machine j. The position of machine k in MC depends on the value of $LinkM_{kx}$ and $LinkM_{xk}$ between k and each assigned machine (machine x) in the MC to find the position having the highest linkage strength value.

3d) If machine k has already belonged to the cell, and machine j has not been assigned. Machine j is assigned to the cell that includes machine k.

3e) If both machine j and machine k have already belonged to the same cell. Therefore, the machine pair can be ignored and go to the next step.

3f) If machines j and k have been allocated to two distinct cells. This information can be reserved for future processes such as merging two MCs with specific conditions.

Step 4: Repeat step 3 with the next highest S_{jk} until all m machines have been assigned to MCs.

Step 5: Assign parts corresponding to MCs to generate part families. The following sub-steps are required to assign parts into part families corresponding to the generated MCs. For special cases, further process such as merging cells is applied.

5a) For each part i with route r and each machine cell, determine NV_{ir} representing the total number of machines that are not visited by part i, and NO_{ir} which denotes the sum of operations of part i outside this cell.

5b) Part i with route r is assigned to the part families corresponding to the machine cell where the sum NVO_{ir} is minimal as calculated by Eq. 12.

$$NVO_{ir} = NV_{ir} + NO_{ir} \tag{12}$$

5c) Repeat step 5a until all parts are allocated in part cells.

For parts with multi-routings, the following processes in step 6 are proceeded to select the best routing.

Step 6: Select routings for each part

6a) For each part i in the route r, determine $SAIB_{ir}$ that is calculated by the weighted sum of the total number of actual inter-cell moves (AIM_{ir}) and the total number of actual backward moves (ABM_{ir}) as shown at Eq. 13.

$$SAIB_{ir} = AIM_{ir} + q.ABM_{ir} \tag{13}$$

6b) Among the various routings, the route with smallest $SAIB_{ir}$ is chosen. In the case of the same $SAIB_{ir}$, two additional sub-conditions are employed. The first sub-condition involves evaluating the ratio of voids for route r, denoted as $RNVI_{ir}$ and calculated using Eq. 14. The second sub-condition pertains to the processing time objective, measured as the total processing time in route r (ST_{ir}) and computed using Eq. 15. Depending on the selected objective, one of two sub-conditions is compared across various routings to determine the optimal routing, which yields the smallest value.

$$RNVI_{ir} = \frac{NV_{ir}}{NI_{ir}} \tag{14}$$

$$ST_{ir} = \sum_{j=0}^{m} \sum_{p=1}^{n_{irj}} t_{irj}^{p} \tag{15}$$

where NV_{ir} is number of voids for each part i in the route r; NI_{ir} is number of operations inside machine cells for part i in the route r

6c) Repeat Step 6b until all parts can choose the best routing.

Step *7*: Stop

2.3 Modified Group Technology Efficacy (MGTE)

The modified group technology efficacy ($MGTE$) is introduced to integrate the actual backward moves (including external operations) and the weigh factor q. In practical production, the inter-cell move is the main concern and has the highest effects on the travelling cost in CFP. Thus, the inter-cell move should have a stronger effect than other factors and q should be added in $MGTE$ calculation. Depending on the pratical inter-cell and intra-cell move cost, the q value can change to meet specific production conditions but is not fixed for all cases. Using the proposed $MGTE$, the overall group technology efficacy can be evaluated as shown in Eq. 16.

$$MGTE = \frac{1 - \frac{AIM + q.ABM}{PIM}}{1 + \frac{NV}{NI}} \tag{16}$$

3 Illustrative Examples

To explain the calculation procedure of the proposed method, example 1 was generated with six parts and five machines for single routing, incorporating multi-visits, Subsequently, example 2 utilized production data featuring seven parts, ten machines, and fourteen alternative routings.

3.1 Example of Single Routing and Multi-Visits (Example 1)

Random production data is generated with a machine-part matrix of 6x5, indicating the operation sequence, and the production volume is shown in Table 1. In this example, all parts have the single routing. The SC calculation in step 2 is applied for all machine pairs, and the results are shown in Table 2. At step 3a, S_{14} is determined as the highest value (0.8703), so machines 1 and 4 should be grouped in the first machine cell. The order of machines 1 and 4 is determined by comparing the linkage strength calculation of the machine pair. Because $LinkM_{14} = 2$ and $LinkM_{41} = 1$, machine 1 should have been in front of machine to obtain the highest linkage strength value. S_{23} is the second-highest value (0.8481), so machines 2 and 3 should have been grouped in the second machine cell. The linkage strength between machines 3 and 2 is higher than that between machines 2 and 3. Hence, machine 3 should have been in front of machine 2. A similar process is executed for the other machine pairs, and two MCs are finally generated, including (1, 4) and (3, 2, 5) at the end of step 4. The parts then are assigned to the MC in the step 5. For each part, NV_{ir} and NO_{ir} for each machine cell are calculated to assign the part to the specific MC with the smallest sum of NV_{ir} and NO_{ir}. Finally, two part families are generated, including (3, 5, 6) and (1, 2, 4) according to two machine cells.

Table 1

The production data for Example 01

Part *i*	Machine					PV
	M1	M2	M3	M4	M5	
P1	3	1	2			160
P2		4	1, 3		2	310
P3	2, 4	1		3		280
P4		1		3	2	265
P5				2	1	80
P6	1			2		150

Table 2

Similarity coefficient matrix for Example 01

Machine	Machine				
	M1	M2	M3	M4	M5
M1	1	0.5892	0.2941	**0.8703**	
M2	0.5892	1	**0.8481**	0.3948	0.7611
M3	0.2941	**0.8481**	1		**0.7777**
M4	**0.8703**	0.3948		1	0.5406
M5		0.7611	**0.7777**	0.5406	1

3.2 Example of Multi-Routings and Selected Objective (Example 2)

The second example with a test instance size of 7x10x14 uses a production data sample from the existing literature, including 7 parts and 10 machines and 14 alternative routings [8]. The processing time for all operations is assumed to be the same. According to the calculation of the SC matrix, the machine pairs (8, 3), (6, 4), (9, 6), and (7, 10) have highest SC value. Therefore, the machine pair (8, 3) should be assigned to the first MC and the machine pair (6, 4) assigned to the second MC since the merging condition at step 3d is not satisfied.

The next highest SC is for machine pair (9, 6). Machine 6 has already been assigned to the second MC. Then machine 9 is assigned to the second MC. The machine pair (7, 10) has not been assigned to any MCs. The merging condition of this pair is satisfied because both SC values for machine pairs (7, 8) and (10, 8) are higher than the $sThreshold$ value (0.5). Hence, they are added to the first MC with the machine pair (8, 3). The process continued until all machines are assigned to MCs. At the end of step 4, two MCs corresponding with two-part families for all routings are generated, as shown in Table 3. During step 6, $SAIB_{ir}$ for all parts with alternative routings are calculated as Eq. 13. The routing with lowest $SAIB_{ir}$ is selected. For parts 3, 5 and 7, two alternative routings have the same $SAIB_{ir}$. Thus, the $RNVI_{ir}$

and ST_{ir} are calculated corresponding with two objectives: the compactness and processing time. To obtain the MC compactness for the case (7x10x14$^{(C)}$), the smallest $RNVI_{ir}$ is more important than ST_{ir}. Two-part families are determined: (P1-R1, P2-R3, P3-R2, P5-R1) and (P4-R1, P6-R1, P7-R2). To achieve the processing time objective for the case (7x10x14$^{(T)}$), the smallest ST_{ir} has a higher priority and two families are identified: (P1-R1, P2-R3, P3-R2, P7-R1) and (P6-R1, P4-R1, P5-R2). Two machine cells with machine layout for case 7x10x14$^{(C)}$ and examples of material flow for P2-R3 and P7-R2 are shown on Figure 2.

Table 3

Cell formation for all multi-routings

Part	Route	Machine										$SAIB_{i}$	$RNVI_{i}$	ST_{ir}
		M3	M5	M7	M8	M10	M1	M2	M4	M6	M9			
P1	**R1**		**3**	**4**		**5**	**1**		**2**			**1**		
P2	**R3**	**2**	**3**	**4**	**5**	**6**		**1**				**1**		
P3	R2	1	2		3	5					4	2	**0.25**	**5**
P5	R1	1		3	4	5				2		2	**0.25**	*5*
P6	**R1**			**3**		**4**		**1**	**2**			**1**		
P7	R1	1			2	3						0	*0.67*	**3**
P1	R2	2			4			1		3	5	4		
P2	R1		4			6	1	2	3	5		3		
P2	R2	2					1		3	4	5	2		
P3	R1			3	4		1		2		5	2	*0.67*	*5*
P4	**R1**							**1**	**2**	**3**	**4**	**0**		
P5	R2			3			1	2			4	**2**	*0.67*	**4**
P6	R2				2		1				3	2		
P7	R2							1	2	3	4	0	**0.25**	*4*

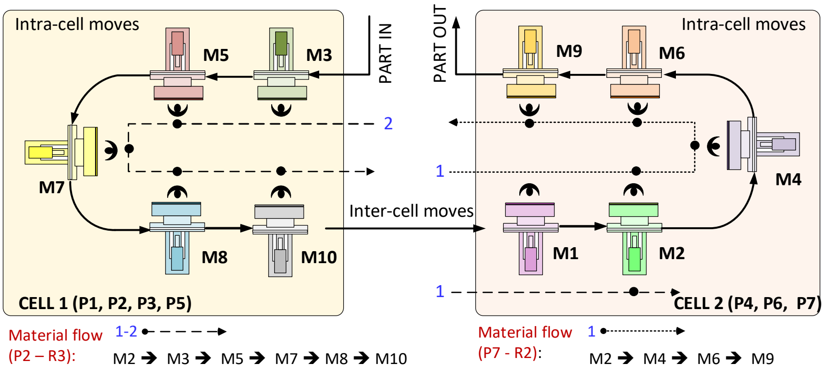
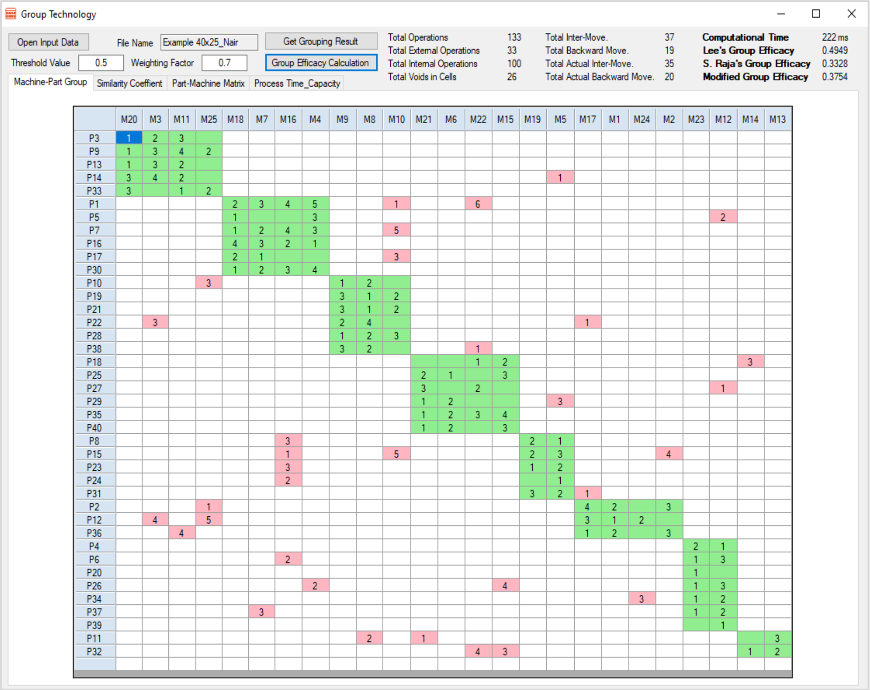


Figure 2

Cell formation and machine layout (7x10x14$^{(C)}$) and material flows for P2 - R3 and P7 - R2

4 Comparison Results and Discussion

The proposed method used seventeen common problems from the literature from small size (5x4) to large size (51x20) for the evaluation and comparison. Three kinds of GTE values including Lee's GTE [33], Raja's GTE [34] and proposed MGTE are calculated for the comparison. A developed software built by Visual C# allows to quickly calculate and display the group technology results based on the proposed method. Figure 3 shows the final solution of case 40x25 in the developed application.



File Name: Example 40x25_Nair
Threshold Value: 0.5
Weighting Factor: 0.7

Total Operations 133
Total External Operations 33
Total Internal Operations 100
Total Voids in Cells 26
Total Inter-Move. 37
Total Backward Move. 19
Total Actual Inter-Move. 35
Total Actual Backward Move. 20
Computational Time 222 ms
Lee's Group Efficacy 0.4949
S. Raja's Group Efficacy 0.3328
Modified Group Efficacy 0.3754

	M20	M3	M11	M25	M18	M7	M16	M4	M9	M8	M10	M21	M6	M22	M15	M19	M5	M17	M1	M24	M2	M23	M12	M14	M13
P3	1	2	3																						
P9	1	3	4	2																					
P13	1	3	2																						
P14	3	4	2														1								
P33	3		1	2																					
P1					2	3	4	5			1			6											
P5					1			3															2		
P7					1	2	4	3			5														
P16					4	3	2	1																	
P17					2	1					3														
P30					1	2	3	4																	
P10				3					1	2															
P19									3	1	2														
P21									3	1	2														
P22		3							2	4								1							
P28									1	2	3														
P38									3	2				1											
P18														1	2									3	
P25												2	1		3										
P27												3		2									1		
P29												1	2				3								
P35												1	2	3	4										
P40												1	2		3										
P8							3									2	1								
P15							1				5					2	3				4				
P23							3									1	2								
P24							2										1								
P31																3	2	1							
P2				1														4	2		3				
P12		4		5														3	1	2					
P36			4															1	2		3				
P4																						2	1		
P6							2															1	3		
P20																						1			
P26								2							4							1	3		
P34																				3		1	2		
P37						3																1	2		
P39																							1		
P11										2		1													3
P32														4	3									1	2

Figure 3

The result of the production data size of 40x25 in the developed application

Table 4 displays the computational results obtained by the proposed method for each test instance, along with a comparison to the best results achieved by other approaches in the literature. The proposed method generally outperforms previous studies in terms of AIM and ABM, except for the 35x18 and 43x16 examples. Although these two cases exhibit a higher number of AIM, they demonstrate superior compactness of machine cells with significantly lower NV. Furthermore, for the 40x25 example, the proposed method yields machine cells with one additional ABM, while maintaining four fewer AIM compared to the literature's methods. Consequently, the proposed method consistently achieves minimal overall moves while ensuring higher compactness. The results demonstrate that the proposed method produces solutions with lower weighted overall moves and improved compactness across most test instances.

Table 4

Comparison of the proposed method with other methods in detail

Test No.	Size PxMxR	Best result/**Proposed Method**						Method	Best source
		NC	NO	NI	NV	AIM	ABM		
1	6x5x8	**-/2**	**-/4**	**-/14**	**-/3**	**-/4**	**-/3**	-	-
2	5x4x10	2/**2**	0/**0**	9/**9**	1/**1**	0/**0**	2/**2**	Mathematical model	[35]
3	9x9	2/**2**	4/**5**	29/**28**	20/**14**	4/**5**	1/**0**	Similarity coefficient	[24]
4	12x10	3/**3**	4/**5**	34/**33**	7/**8**	8/**5**	9/**3**	Two-mode similarity coefficient	[36]
5.1	7x10x14$^{(C)}$	**-/2**	**-/7**	**-/26**	**-/9**	**-/7**	**-/0**	-	-
5.2	7x10x14$^{(T)}$	**-/2**	**-/7**	**-/24**	**-/11**	**-/7**	**-/0**	-	-
5.3	7x10x14	2/**3**	8/**8**	22/**23**	12/**11**	11/**9**	0/**0**	Heuristic algorithm	[8]
6	18x10	3/**3**	6/**6**	50/**50**	13/**14**	7/**7**	14/**12**	Similarity coefficient	[21]
7	19x12	3/**3**	26/**20**	53/**59**	16/**30**	28/**22**	7/**6**	Simulated annealing	[11]
8.1	12x12x20$^{(C)}$	3/**3**	15/**15**	31/**32**	30/**16**	14/**18**	12/**4**	Similarity coefficient	[31]
8.2	12x12x20$^{(T)}$	3/**3**	15/**15**	31/**29**	30/**19**	14/**18**	12/**5**	Similarity coefficient	[31]
9	20x8	3/**3**	9/**9**	52/**52**	0/**0**	16/**16**	8/**8**	Flow matrix	[34]
10.1	8x9x20$^{(C)}$	2/**2**	2/**2**	26/**26**	14/**10**	2/**2**	0/**0**	Tabu search algorithm	[29]
10.2	8x9x20$^{(T)}$	2/**2**	2/**2**	26/**24**	14/**11**	2/**2**	0/**0**		
11	10x10x24	3/**3**	2/**2**	30/**30**	3/**3**	2/**2**	1/**1**	Tabu search algorithm	[37]
12	16x10x32	2/**2**	5/**5**	66/**67**	17/**17**	6/**6**	18/**11**	Similarity coefficient	[31]
13	20x20	5/**5**	14/**14**	65/**65**	18/**18**	18/**18**	21/**11**	Two-mode similarity coefficient	[36]
14	35x18	4/**5**	49/**44**	118/**123**	91/**21**	54/**60**	29/**27**	Genetic algorithm	[13]
15	40x25	8/**8**	33/**33**	100/**100**	23/**26**	39/**35**	19/**20**	Flow matrix	[34]
16	43x16	4/**5**	28/**32**	119/**115**	107/**53**	37/**44**	19/**20**	Similarity coefficient	[31]
17	45x20	4/**5**	31/**30**	129/**130**	65/**66**	41/**41**	41/**22**	Genetic algorithm	[13]
18	51x20	5/**5**	42/**30**	138/**150**	65/**84**	46/**37**	27/**29**	TOPSIS	[38]
PxMxR: *Part number* x *Machine number* x *Route number*; *(C): Compactness objective; (T): Processing time objective; NC: Number of cells*									

Figure 4 presents the comparison of the proposed method's results against other approaches, specifically for the 40x25 test instance, using three GTE measures. The consideration of adjacent operation in the SC formula can significantly reduce the total number of AIM resulting in the best Lee's GTE in comparison with other methods in the literature. Raja's GTE and proposed MGTE integrated ABM in measures are also higher in the proposed method than in other previous approaches. Table 5 presents a comparative analysis of the proposed method against the best solutions from the literature, focusing on Lee's GTE, Raja's GTE, and the proposed MGTE measures across eighteen instances. The comparison results consistently demonstrate the superiority of the proposed method over other approaches in terms of Raja's GTE and the proposed MGTE. Notably, in instances 18x10 and 45x20, the proposed method achieves smaller Lee's GTE but higher Raja's GTE than other methods. This is attributed to the proposed method yielding a final solution with one more NV while maintaining the same AIM and fewer ABM. In instances 35x18 and 43x16, the proposed method exhibits a higher number of AIM due to a higher NC. However, the substantial reduction in NV across all machine cells leads to significantly higher GTE in all measure types compared to existing methods.

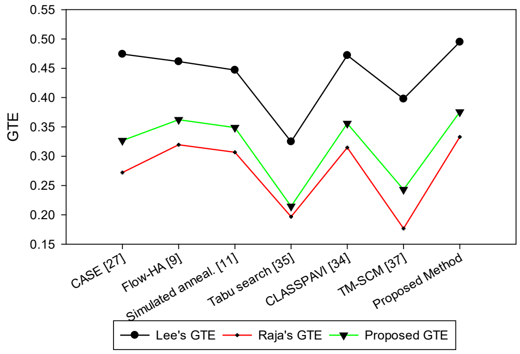


Figure 4
Comparison of the GTEs obtained by the proposed method over existing methods for problem 40x25

Furthermore, in order to assess the efficacy of the proposed method when dealing with a large-scale problem, we employ the production data 100x40 using a binary machine-part matrix, as introduced by Gonçalves [12], where 100 parts and 40 machines are involved. The original production data is also modified to change into an operation sequence-based machine-part matrix to evaluate the performance of the proposed method. The CPU time to solve these problems is less than three seconds. It indicates the merit of the proposed method to apply to big-size problems in a short computational time while still obtaining the optimal overall GTE. For all remaining evaluated instances, the computational time is significantly short in comparison with other methods in the literature due to the simple programming method [11, 13, 29, 37].

Figure 5 shows the CPU time comparison results between the proposed method and other algorithms for popular instances. The comparison results emphasize the significance of incorporating adjacent operations in the SC calculation, resulting in decreased AIM and NV in machine cells. The effectiveness of the sorting algorithm in determining machine positions during clustering leads to reduced ABM. Selecting the best routing for each part, based on both overall moves and machine cell compactness, allows for achieving optimal overall GTE in multi-routing problems within the context of CFP.

Table 5

The GTE comparison between other approaches and the proposed method

Test No.	Size PxMxR	Results from the literature			Proposed method			CPU (ms)
		Lee's GTE	Raja's GTE	Proposed MGTE	Lee's GTE	Raja's GTE	Proposed MGTE	
1	6x5	-	-	-	**0.5490**	**0.3431**	**0.4049**	3
3	5x4x10	0.9	0.45	0.585	**0.9**	**0.45**	**0.585**	1
2	9x9	0.4932	0.4685	0.4759	**0.5277**	**0.5277**	**0.5277**	8
4	12x10	0.5741	0.2870	0.3731	**0.6500**	**0.5572**	**0.5850**	12
5	7x10x14$^{(C)}$	-	-	-	**0.5428**	**0.5428**	**0.5428**	15
	7x10x14$^{(T)}$	-	-	-	**0.4857**	**0.4857**	**0.4857**	16
	7x10x14	0.3376	0.3376	0.3376	**0.4637**	**0.4637**	**0.4637**	16
6	18x10	0.6475	0.3551	0.4423	*0.6373*	**0.3906**	**0.4646**	18
7	19x12	0.4097	0.3200	0.3469	**0.4198**	**0.3646**	**0.3734**	31
8	12x12x20	0.3139	0.1200	0.1734	**0.3238**	**0.2476**	**0.2704**	20
9	20x8	0.6097	0.4146	0.4731	**0.6097**	**0.4146**	**0.4731**	10
10	8x9x20	0.6213	0.6213	0.6213	**0.6500**	**0.6500**	**0.6500**	17
11	10x10x24	0.8264	0.7851	0.7975	**0.8264**	**0.7851**	**0.7975**	22
12	16x10x32	0.7084	0.4482	0.5263	**0.7121**	**0.5554**	**0.6024**	35
13	20x20	0.5442	0.2655	0.3491	**0.5442**	**0.3982**	**0.4420**	128
14	35x18	0.3600	0.2501	0.2831	**0.4659**	**0.2912**	**0.3436**	122
15	40x25	0.4721	0.3122	0.3558	**0.4949**	**0.3328**	**0.3754**	222
16	43x16	0.3392	0.2430	0.2719	**0.4032**	**0.2667**	**0.3076**	149
17	45x20	0.4279	0.1908	0.2619	*0.4267*	**0.2999**	**0.3379**	148
18	51x20	0.4374	0.2951	0.3378	**0.4571**	**0.3131**	**0.3562**	321

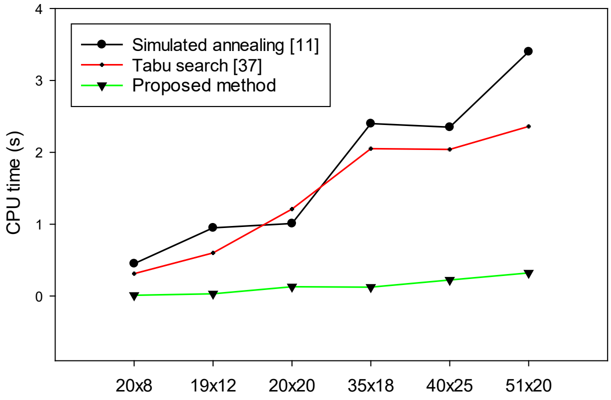


Figure 5

CPU time comparison for popular instances

Conclusions

This paper presents a novel similarity coefficient and an improved clustering algorithm to address machine cell formation and machine sequence generation simultaneously. The proposed method integrates realistic production data, such as operation sequence, production volume, processing time, machine capacity, multi-visits, and multi-routings. The modified group technology efficacy is utilized to evaluate the overall performance of the solutions for practical problems. Comparative analyses with existing methods using eighteen problems reveal the following conclusions:

- The proposed method outperforms other approaches in reducing weighted overall moves and voids in machine cells. It consistently achieves higher overall group technology efficacy for most test instances.

- The proposed method demonstrates significant time savings, with most test instances solved in less than 0.4 seconds, even big-size problem from the literature in under 3 seconds.

- The effectiveness of the proposed approach makes it a promising method for simultaneously addressing cell formation and machine sequence in complex problems within the realm of cellular manufacturing.

References

[1] Baysan, S., O. Kabadurmus, E. Cevikcan, S. I. Satoglu, and M. B. Durmusoglu, *A simulation-based methodology for the analysis of the effect of lean tools on energy efficiency: An application in power distribution industry*. Journal of Cleaner Production, 2019, **211** pp. 895-908, DOI: 10.1016/j.jclepro.2018.11.217

[2] Belhadi, A., F. E. Touriki, and S. El Fezazi, *Benefits of adopting lean production on green performance of SMEs: a case study*. Production Planning & Control, 2018, **29**(11): pp. 873-894, DOI: 10.1080/09537287.2018.1490971

[3] Pigler Cs., F.-V. Á., Abonyi J., *Scalable co-Clustering using Crossing Minimization - Application to Production Flow Analysis*. Acta Polytechnica Hungarica, 2016, **13**(2): pp. 209-228

[4] Gidwani, V. S. a. B. D., *Cellular manufacturing system practices in manufacturing industries: a pilot study*. International Journal of Indian Culture and Business Management, 2022, **25**(4): pp. 550-569, DOI: doi.org/10.1504/IJICBM.2022.122763

[5] Mahdavi, I., B. Javadi, K. Fallah-Alipour, and J. Slomp, *Designing a new mathematical model for cellular manufacturing system based on cell utilization*. Appl. Math. Comput., 2007, **190**(1): pp. 662-670 DOI: 10.1016/j.amc.2007.01.060

[6] Tóth, T., S. Radeleczki, L. Veres, and A. Körei, *A new mathematical approach to supporting group technology*. Eur. J. Ind. Eng., 2014, **8**(5): pp. 716-737 DOI: 10.1504/EJIE.2014.065734

[7] Mahdavi, I., B. Shirazi, and M. M. Paydar, *A flow matrix-based heuristic algorithm for cell formation and layout design in cellular manufacturing system*. Int J Adv Manuf Technol, 2008, **39**(9-10): pp. 943-953 DOI: 10.1007/s00170-007-1274-7

[8] Shashikumar, S., R. D. Raut, V. S. Narwane, B. B. Gardas, B. E. Narkhede, and A. Awasthi, *A novel approach to determine the cell formation using heuristics approach*. OPSEARCH, 2019, **56**(3): pp. 628-656 DOI: 10.1007/s12597-019-00381-4

[9] Mahdavi, I. and B. Mahadevan, *CLASS: An algorithm for cellular manufacturing system and layout design using sequence data*. Robot Comput Integr Manuf, 2008, **24**(3): pp. 488-497 DOI: 10.1016/j.rcim.2007.07.011

[10] Kumar, R. and S. P. Singh, *A similarity score-based two-phase heuristic approach to solve the dynamic cellular facility layout for manufacturing systems*. Eng. Optim., 2017, **49**(11): pp. 1848-1867 DOI: 10.1080/0305215X.2016.1274205

[11] Paydar, M. M., I. Mahdavi, I. Sharafuddin, and M. Solimanpur, *Applying simulated annealing for designing cellular manufacturing systems using MDmTSP*. Comput. Ind. Eng., 2010, **59**(4): pp. 929-936 DOI: 10.1016/j.cie.2010.09.003

[12] Gonçalves, J. F. and M. G. C. Resende, *An evolutionary algorithm for manufacturing cell formation*. Comput. Ind. Eng., 2004, **47**(2-3): pp. 247-273 DOI: 10.1016/j.cie.2004.07.003

[13] Boulif, M. and K. Atif, *A new branch-&-bound-enhanced genetic algorithm for the manufacturing cell formation problem*. Comput. Oper. Res., 2006, **33**(8): pp. 2219-2245 DOI: 10.1016/j.cor.2005.02.005

[14] Agrawal, A. K., P. Bhardwaj, and V. Srivastava, *Ant colony optimization for group technology applications*. Int J Adv Manuf Technol, 2011, **55**(5-8): pp. 783-795 DOI: 10.1007/s00170-010-3097-1

[15] Sudhakara Pandian, R. and S. S. Mahapatra, *Manufacturing cell formation with production data using neural networks*. Comput. Ind. Eng., 2009, **56**(4): pp. 1340-1347 DOI: 10.1016/j.cie.2008.08.003

[16] Park, S. and N. C. Suresh, *Performance of Fuzzy ART neural network and hierarchical clustering for part-machine grouping based on operation sequences*. Int. J. Prod. Res., 2003, **41**(14): pp. 3185-3216 DOI: 10.1080/0020754031000110277

[17] Selim, H. M., *Manufacturing cell formation problem: A graph partitioning approach*. Ind. Manag. Data Syst., 2002, **102**(5-6): pp. 341-352 DOI: 10.1108/02635570210432046

[18] Saral, J., S. Arumugam, I. Venkat, and A. Somasundaram, *Cellular manufacturing problem - A graph theoretic approach*. J. Adv. Mech. Des. Syst. Manuf., 2019, **13**(3): pp. JAMDSM0061-JAMDSM0061 DOI: 10.1299/jamdsm.2019jamdsm0061

[19] Danilovic, M. and O. Ilic, *A novel hybrid algorithm for manufacturing cell formation problem*. Expert Syst. Appl., 2019, **135** pp. 327-350 DOI: 10.1016/j.eswa.2019.06.019

[20] Seifoddini, H. and C. P. Hsu, *Comparative study of similarity coefficients and clustering algorithms in cellular manufacturing*. J. Manuf. Syst., 1994, **13**(2): pp. 119-127 DOI: 10.1016/0278-6125(94)90027-2

[21] Prabhakaran, G., T. N. Janakiraman, and M. Sachithanandam, *Manufacturing data-based combined dissimilarity coefficient for machine cell formation*. Int J Adv Manuf Technol, 2002, **19**(12): pp. 889-897 DOI: 10.1007/s001700200101

[22] Wu, L. and S. Suzuki, *Cell formation design with improved similarity coefficient method and decomposed mathematical model*. Int J Adv Manuf Technol, 2015, **79**(5-8): pp. 1335-1352 DOI: 10.1007/s00170-015-6931-7

[23] Goyal, K. K., P. K. Jain, and M. Jain, *A comprehensive approach to operation sequence similarity based part family formation in the reconfigurable manufacturing system*. Int. J. Prod. Res., 2013, **51**(6): pp. 1762-1776 DOI: 10.1080/00207543.2012.701771

[24] Wu, L., L. Li, L. Tan, B. Niu, R. Wang, and Y. Feng, *Improved similarity coefficient and clustering algorithm for cell formation in cellular manufacturing systems*. Eng. Optim., 2020, **52**(11): pp. 1923-1939 DOI: 10.1080/0305215X.2019.1692204

[25] McAuley, J., *Machine grouping for efficient production*. Production Engineer, 1972, **51**(2): pp. 53-53 DOI: 10.1049/tpe.1972.0006

[26] Gupta, T. and H. I. Seifoddini, *Production data based similarity coefficient for machine-component grouping decisions in the design of a cellular manufacturing system*. Int. J. Prod. Res., 1990, **28**(7): pp. 1247-1269 DOI: 10.1080/00207549008942791

[27] Nair, G. J. and T. T. Narendran, *CASE: A clustering algorithm for cell formation with sequence data*. Int. J. Prod. Res., 1998, **36**(1): pp. 157-180 DOI: 10.1080/002075498193985

[28] Seifoddini, H. and P. M. Wolfe, *Application of the similarity coefficient method in group technology*. IEE Trans (Ins. Ind. Engr.), 1986, **18**(3): pp. 271-277 DOI: 10.1080/07408178608974704

[29] Chung, S.-H., T.-H. Wu, and C.-C. Chang, *An efficient tabu search algorithm to the cell formation problem with alternative routings and machine reliability considerations*. Computers & Industrial Engineering, 2011, **60**(1): pp. DOI: 10.1016/j.cie.2010.08.016

[30] Alhourani, F., *Cellular manufacturing system design considering machines reliability and parts alternative process routings*. Int. J. Prod. Res., 2016, **54**(3): pp. 846-863 DOI: 10.1080/00207543.2015.1083626

[31] Alhourani, F., *Clustering algorithm for solving group technology problem with multiple process routings*. Comput. Ind. Eng., 2013, **66**(4): pp. 781-790 DOI: 10.1016/j.cie.2013.09.002

[32] Harhalakis, G., R. Nagi, and J. M. Proth, *An efficient heuristic in manufacturing cell formation for group technology applications*. Int. J. Prod. Res., 1990, **28**(1): pp. 185-198 DOI: 10.1080/00207549008942692

[33] Lee, K. and K. I. Ahn, *GT efficacy: A performance measure for cell formation with sequence data*. Int. J. Prod. Res., 2013, **51**(20): pp. 6070-6081 DOI: 10.1080/00207543.2013.794317

[34] Raja, S. and V. Anbumalar, *An effective methodology for cell formation and intra-cell machine layout design in cellular manufacturing system using parts visit data and operation sequence data*. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2016, **38**(3): pp. 869-882 DOI: 10.1007/s40430-014-0280-6

[35] Yin, Y., K. Yasuda, and L. Hu, *Formation of manufacturing cells based on material flows*. Int J Adv Manuf Technol, 2005, **27**(1-2): pp. 159-165 DOI: 10.1007/s00170-004-2143-2

[36] Kong, T., K. Seong, K. Song, and K. Lee, *Two-mode modularity clustering of parts and activities for cell formation problems*. Comput. Oper. Res., 2018, **100**pp. 77-88 DOI: 10.1016/j.cor.2018.06.018

[37] Chang, C. C., T. H. Wu, and C. W. Wu, *An efficient approach to determine cell formation, cell layout and intracellular machine sequence in cellular manufacturing systems*. Comput. Ind. Eng., 2013, **66**(2): pp. 438-450 DOI: 10.1016/j.cie.2013.07.009

[38] Ahi, A., M. B. Aryanezhad, B. Ashtiani, and A. Makui, *A novel approach to determine cell formation, intracellular machine layout and cell layout in the CMS problem based on TOPSIS method*. Comput. Oper. Res., 2009, **36**(5): pp. 1478-1496 DOI: 10.1016/j.cor.2008.02.012