Modified direct clustering algorithm for group formation in cellular manufacturing

G C Potdar\(^1\), S U Sapkal\(^1\) and A S Shivade\(^1\)

\(^1\)Department of Mechanical Engineering, Walchand College of Engineering, Sangli, Maharashtra, India.

potdar.c.gaurav@gmail.com, sagarus1201@gmail.com, shivadeanand@yahoo.in

Abstract. In cellular manufacturing system formation of machine-part families is an important task. Cellular manufacturing systems deals with various clustering algorithms which are particularly relevant to the problem of machine component group formation. In this paper, case study from well-known gearbox manufacturing company is considered where there is scope for using cellular manufacturing system. Company is manufacturing the gearbox varying from single stage reduction to four stage reduction. A modified direct clustering algorithm is proposed and is applied after the initial incidence matrix derived from direct clustering algorithm. This algorithm overcomes the problem arises during application of direct clustering algorithm and gives the optimum solution. The algorithm is specifically designed to deal with large amount of data during realistic situations. The results of proposed algorithm are compared with well-known existing algorithms and it is found that the proposed algorithm gives the better solution than the existing algorithms under consideration.

1. Introduction

Group technology deals with the fragmentation of a manufacturing system into subsystems [4]. In order to enable this fragmentation various coding systems and classification have been developed. With the help of these codes, clustering of machine-component into subsystems is accomplished. Clustering analysis is one of the most frequently applied mathematical tools in group technology [6]. Cellular manufacturing is an application of GT to manufacturing. In cellular manufacturing, a few machines, usually dissimilar in function, are grouped into a cell. For mass production a dedicated cell is required which enable to process a family of parts [6].

An XYZ Pvt. Ltd. company which produces gearbox is being considered as a part of work. Company has made changes in their plant layout earlier but failed to meet desired objectives. Distances in between the processes is not feasible as far as plant layout is concern. To overcome these problems company has decided to use cellular manufacturing system. Presently, gearbox varying from single to four stage reduction is being manufactured in the plant. There are common processes and common components in between single to four stage reduction gearbox. Consider an example of single stage gearbox having 13 components and 23 processes. Some processes are common in those 13 components. So, clustering those common processes by applying various algorithms is the requisite task.

As described by Arumugam S [9] the input for a cellular manufacturing problem consists of , a set Y of p parts and a set X of m machines and an m × p matrix A = (a\(_{ij}\)) where a\(_{ij}\) = 1 or 0 according as the part \(p_i\) which is processed on the machine \(m_j\). In the context of similarity, it includes design attributes such as size, shape, etc. and/or manufacturing attributes like length, diameter, surface finish, tolerance, etc. Once if the parts are grouped, then machines, tools and equipment required to process these parts having similarity are grouped together to take the advantages of similarities in manufacturing.

Burbidge J L [1] described that there are three stages in the analysis of the production flow in which the first stage is to determine the design of the department in the factory through the analysis of the factory flow, the second stage is to determine the group of machine components in the group analysis and the third stage is to determine the layout of each group of machine components in the line analysis. Burbidge J L [2] suggested a method for sorting components using punched card processes. This technique created a flow sequence based on the map, but it failed to define a grouping. King J R [3] reviewed the single cluster analysis method, the production flow analysis method, as well as the bond
energy method, and built a rank cluster method. Kusiak A [4] developed an integer programming model that offers a more convenient representation of clustering problems than a matrix model that enhances the performance of component families and machine cells. Ballakur A et al [5] developed a heuristic focused on cell utilization for the design of cell manufacturing systems. Machines are allocated to cells based on workload and cell size limitations. Components are 'assigned' to cells in such a way that most of their operations can be done within cells. Seifoddini H et al [6] analyzed the components of the machine cell formation cycle such as Similarity coefficients, clustering algorithms, and efficiency measurements. New Consistency metrics are implemented and a comparative analysis of three different similarity coefficients is performed - the Jaccard similarity coefficient, the weighted similarity coefficient, and the commonality value. Alhourani F [7] implemented similarity coefficient that incorporates multiple process routing along with operation sequence, production volumes, duplicate machines, and machine capacity. They also proposed a clustering algorithm for the formation of machine cells. The developed similarity coefficient showed greater sensitivity to intercellular movements and improved machine grouping. Telsang M et al [8] implemented cellular manufacturing system to reduce intercellular movements, transport waste, employees' wages to increase productivity. Arumugam S et al [9] developed an algorithm for the identification of bottleneck machines and bottleneck sections in a cellular manufacturing problem along with a cell partition. Mehta S et al [10] developed differential bond energy algorithm and results are compared to the classic bond energy algorithm. Dhayef D et al [11] developed branch and bound algorithm for machine cell and part family and for sequencing and scheduling groups for makespan calculation. The developed algorithm was evaluated by a case study consisting of four items processed on nine machines. Built algorithm can be used to decide the best job and minimize the order of the best group for any kind of problem and completion time (makespan). Chan H M et al [12] developed a direct clustering algorithm that shapes component families and machine groups for cellular manufacturing by iteratively restructurung the machine component matrix. The direct clustering algorithm consists of the matrix being concatenated, moving the rows with the leftmost assigned to the top and the columns with the topmost assigned cell to the left of the matrix.

2. Direct clustering algorithm

2.1. Introduction

Direct clustering algorithm is the method of finding machine and component groups by revamping the sequence in machines and components which are listed in the matrix. The initial incidence matrix is constructed by allocating '1' to the cells when a component goes through a process in a particular machine and '0' when it is not. While applying clustering algorithm it is desired to form a diagonal clustering of component and machines. Therefore, the matrix is sorted in two stages. Firstly, sorting is being concluded based on the counts of the number of entries of '1'; columns and rows are sorted in decreasing order of counts. The predicament emerges in second sorting where it postulates the sort as 'leftmost' rows allotments to the top and 'topmost' column allotment to the left. When it is attempted to iterate it for the second sorting, the diagonal element formed is not optimum and can be iterate further. Thus, the algorithm is unable to come to the final matrix which indicate diagonalized form of clustering.

The industrial data of 13 components and 23 processes is converged to 12 components and 12 processes by eliminating common allotments in all components (for example, inward inspection is common for all components so it can be eliminated) and the unique allotment for a single component (for example, inner diameter grinding is a unique process for planet gear and no other component undergone such process). It is not necessary that every time the number of components must be equal to number of processes.

2.2. The clustering algorithm

a. Count the number of '1' in each column and row. Regroup the machine component incidence matrix with rows and columns in decreasing order with virtue of counts of '1'.

b. Contemplating firstly with column number one of the matrix relocate the rows which have '1' as an entries in this column to the top of the matrix. Continual of this procedure with respect to the other column will give the new arrangement for the rows.

c. Is the current matrix being optimum diagonalized?
if yes, then go to 'f'
if no, then go to 'd'
d. Rerun the procedure mentioned in 'b' for rows, starting from row 1. This will give new arrangement for the columns.
e. Is the current matrix being optimum diagonalized?
   if yes, then go to 'f'
   if no, then go to 'b'
f. Stop

2.3. Application
The machine component incidence matrix with 12 machine and 12 components is explicated in Table 1. Table 1 illustrates the matrix after the first step which is used in direct clustering algorithm. A row numbering from 2 to 4 designate to components and column numbering from 6 to 7 designate to machine or process. The bottommost row and rightmost column designate the count of number of entries of '1' in each column and row respectively.

| S. No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--------|---|---|---|---|---|---|---|---|---|----|----|----|
| 1      | 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 11 | 7  | 4  |
| 2      | 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 9   |
| 3      | 10| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 8   |
| 4      | 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 8   |
| 5      | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 7   |
| 6      | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 7   |
| 7      | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 7   |
| 8      | 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 5   |
| 9      | 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 5   |
| 10     | 12| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 3   |
| 11     | 11| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 3   |
| 12     | 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1   | 3   |

In Table 1, row 6, 4,10,9 and 2 clustering with column 2,5,12 and 8. Continual of iteration through the rows and column it is observed that these formed clustered get disturbed and they get deviated from their desired results of forming machine component family through their diagonal arrangement. The iteration can be done by relocating leftmost row to the top and topmost column to the left. As this iteration process does not follow a proper path for clustering, it becomes difficult to determine optimum diagonal elements. As this diagonal element can further be optimized which can determined by using modified direct clustering algorithm. Modified direct clustering algorithm enables to allocate possible optimum cell arrangement by assigning weights based on the priority of the positioning in the diagonal form.

3. Modified direct clustering algorithm

3.1. Introduction
As discussed earlier, modified direct clustering algorithm is applied after an initial incidence matrix formed in direct clustering algorithm. The aim of clustering is to diagonally form machine component family matrix in three clusters of $4 \times 4$ (only for $12 \times 12$ data). First cluster consist of the elements allocated at upper left most part of the matrix (row and column from s no.1 to 4 in table 1), followed by
second cluster which is at the middle part of the matrix (row and column from s no. 5 to 8) and third one which is at lower right most part of the matrix (row and column s no. 9 to 12).

The matrix has undergone the first sorting which is based on the counts of entries '1' in the cell. As discussed in King J R [4], weights are the most important parameter in the clustering. Weights can be given as per the element’s configurations required. Here, it is required to cluster the elements in diagonal form in three clusters as mentioned above. Therefore, it is essential to give the priority to first four cells of each column and row as $4^3$ and remaining eight cells as 4. Repeat this procedure with respect to middle four values and last four values of each row and column. Now the priority for three clusters of 12 rows and three clusters for 12 columns is set. Sum up those priorities for each row and column and apply the equation 1 for finding out the weights. Seifoddini H [6] worked on similarity coefficient algorithm in which Jaccard’s coefficient is applied. Referring to the Jaccard's coefficient the formula is applied as mention in equation 1.

$$W_i = \frac{\text{Sum of priorities of first cluster for } i^{th} \text{ row}}{768} \tag{1}$$

$W_i$ is the weight for $i^{th}$ row and 768 is the maximum sum of the priority for a row or a column ($4^3 \times 12$). Similarly, equation 1 is applied for other row and columns.

3.2. The proposed clustering algorithm

a. Rearrange the rows and columns to form it as initial incidence matrix as in direct clustering algorithm (based on counts of number of entries of '1' in decreasing order)

b. Formation of the first cluster in which s no. 1 to 4 of rows and columns allotted the priority of $4^3$ and for remaining eight elements priority of 4.

c. Sum up the priority for each row and column for constructing the first cluster and apply the equation 1 individually.

d. Sort according to weights in decreasing order and then relocate the rows and columns in decreasing order of their weights.

e. Identify the diagonal elements (while identifying do not repeat the element number in a cluster).

f. Is the current matrix being optimum diagonalized?
   if yes then go to 'g'
   if no then go to 'e'

g. Repeat the procedure 'b', 'c' and 'd' for second cluster and third cluster.

h. Is the current matrix being optimum diagonalized?
   if yes then go to 'k'
   if no then go to 'e'

i. Check for other diagonal elements by rearranging rows and columns (keep on eliminating allotted or overlapped elements)

j. Is the current matrix being optimum diagonalized?
   if yes then go to 'k'
   if no then go to 'e'

k. Stop

The steps in the proposed algorithm are applied on $12 \times 12$ incidence matrix. The machine component incidence matrix using modified direct clustering algorithm is illustrated in table 2.


4. Results and discussion
In modified direct clustering algorithm, there are two stages of sorting which are used after the first sorting from direct clustering algorithm. The second and third sorting from the proposed algorithm concentrates more on diagonal configuration of the matrix. In second sorting, priorities are divided in clusters of three (s no. 1 to 4, s no. 4 to 8 and s no. 9 to 12 for both rows and columns one by one). Priorities are assigned only to cells which has entry as ‘1’. Firstly, the priority of \(4^3\) is given for the first cluster (s no. 1 to 4) and for remaining it is 4. As mentioned by King J R [4], the priorities should be given in terms of exponent. For example, if 4, 3 and 2 are the priority for first cluster, second cluster and third cluster respectively. Then the purpose of diagonal configuration of the matrix will not be solved because if number of counts of ‘1’ increased in second and third cluster then weights (which is the part of third stage) obtained will tend towards second and third cluster. This will form element in row number s no. 1 to 4 and column number s no. 5 to 8. In third sorting, the weights are given based on the priority for machine and component and they are arranged in the descending order as shown in table 2. From table 2 it is observed that the elements are diagonalized and there is no scope for the further optimization. After finding out the clusters, elimination of overlapping elements becomes essential as it may affect in bottlenecks. The remaining rows and columns are rearranged manually and configured diagonally as mentioned in the table 2. Thus, the optimum solution is achieved and it is considered as final solution.

Contemplating from table 2 that machines 6,4,10 and 9 forming a cluster with components 2,5,12 and 8. On the other hand, machines 2 and 1 forming a cluster with components 6 and 3. Similarly 3-10, 8-9 and 5-1 are also the effective combinations. Further clustering is not possible as it may lead to the construction of bottlenecks.

5. Conclusion
The optimum solution for system component clustering is not feasible in the case of a direct clustering algorithm, because there is no efficient way to perform iterations. This algorithm is then modified by constructing a machine-component family with a diagonal structure and called as a modified direct clustering algorithm. A comparative evaluation of the machine component incidence matrix is carried out and it is found that direct clustering algorithms are less efficient when it comes to realistic problems as mentioned in earlier research. Moreover, it is found that modified direct clustering algorithms can handle large data and practical problems as they deal with priority by using the initial incidence matrix. Bottlenecks can also be found effectively by a modified direct cluster algorithm.
References

[1] Burbidge J L 1963 Production flow analysis The Production Engineer 43 p 742.
[2] Burbidge J L 1969 Production Flow Analysis Int. Centre for Advanced Technical and Vocational Training
[3] King J R 1980 Machine-component grouping in production flow analysis: an approach using a rank order clustering algorithm Int. J. of Production Research 18(2) pp 213-232
[4] Kusiak A 1987 The generalized group technology concept. Int. J. on Production Research 25(4) pp 561-569
[5] Ballakur A and Steudels H J 1987 A within-cell utilization based heuristic for designing cellular manufacturing systems Int. J. of Production Research 25(5) pp 639-666
[6] Seifoddini H and Hsu C 1991 Comparative Study of Similarity Coefficients and Clustering Algorithms in Cellular Manufacturing J. of Manufacturing Systems 13(2)
[7] Alhourani F 2013 Clustering algorithm for solving group technology problem with multiple process routings Computers & Industrial Engineering 6 pp 781–790
[8] Bharamgude P and Telsang M 2014 Design, analysis and performance simulation of cellular manufacturing system - a case study Int. J. of Mechanical And Production Engineering 2(6) pp 2320-2092
[9] Arumugam S and Saral J 2015 Identification of Bottleneck Elements in Cellular Manufacturing Problem Int. Conf. on Graph Theory and Information Security 74 pp 181 – 185
[10] Mehta S Agarwal P Shrivastava P and Barlawala J 2018 Differential bond energy algorithm for optimal vertical fragmentation of distributed databases J. of Computer and Information Sciences
[11] Dhayef D and Abdulkhafour A 2018 Developing branch and bound algorithm for cell formation and group scheduling Engineering and Technology Journal 36 pp 555-563
[12] Chan H M and Milner D A Direct Clustering Algorithm for Group Formation in Cellular Manufacture J. of Manufacturing Systems 1(1)

Acknowledgement
The authors are grateful to the KavitsuTransmission Pvt. Ltd., Satara for providing the data and necessary facility.