OPSEARCH (Jan–Mar 2014) 51(1):76–97
DOI 10.1007/s12597-013-0133-6

THEORETICAL ARTICLE

A novel approach for part family formation for reconfiguration manufacturing system

Ashutosh Gupta · P. K. Jain · Dinesh Kumar

Accepted: 11 March 2013 / Published online: 6 April 2013

Abstract The Reconfigurable Manufacturing Systems (RMS) is the next step in manufacturing, allowing the production of any quantity of highly customised and complex parts together with the benefits of mass production. In RMSs, parts are grouped into families, each of which requires a specific system configuration. Initially system is configured to produce the first family of parts. Once it is finished, the system is reconfigured in order to produce the second family, and so forth. The effectiveness of a RMS depends on the formation of the optimum set of part families addressing various reconfigurability issues. For this, a two-phase approach is developed where parts are first grouped into families and then families are sequenced, computing the required machines and modules configuration for each family. In the First phase, parts are grouped into families based on their common features. The correlation matrix is developed as operations sequence similarity coefficient matrix. Principal Component Analysis (PCA) is applied to find the eigenvalues and eigenvectors on the correlation similarity matrix. A scatter plot analysis as a cluster analysis is applied to make parts groups while maximizing correlation between parts as per operations sequence similarity. Agglomerative Hierarchical K-means algorithm improved the parts family formation using Euclidean distance resulting a set of part families. In the second phase, optimal selection and sequences of the resulted part families is achieved by using a Mixed Integer Linear Programming (MILP) model minimizing reconfigurability and under-utilization costs to get the minimum cost solution.

Keywords Reconfigurable manufacturing system · Principle Component Analysis · K-means algorithm · Part family selection · Cluster analysis · Sequencing · MILP

A. Gupta (✉) · P. K. Jain · D. Kumar
Mechanical and Industrial Engineering Department, Indian Institute of Technology Roorkee, Roorkee 247667, India
e-mail: ashuaec72@gmail.com

1 Introduction

Current market situation is characterized by globalization, new product requirements, rapidly changing demands, and a continuous improvement of the existing technology for manufacturing activities. Time reduction to introduce new products to the market with high quality and low cost is essential for enterprise survival in this new scenario. For achieving this, the manufacturing system must be able to yield different batch sizes from different part types, with the exact capacity and functionality in each case. Nowadays, the manufacturing industry faces the challenge of responding quickly to the dynamic requirements of the customers. The key factor for sustaining in this highly responsive environment is the ability of the companies to launch new products to the market, with high quality, low cost, and fast delivery. To answer this, Koren [1] define a RMS as "a manufacturing system designed at the outset for rapid changes in structure as well as in hardware and software components in order quickly to adjust production capacity and functionality within a part family in response to sudden changes in market or in regulatory requirements." In the same way, Xiaobo [2] consider a RMS as a manufacturing system configured to produce a family of parts that share some similarities. As these definitions state, the formation of part families is a central issue in RMS. The key attribute of part families is that all the components within a family require similar production systems and thus RMSs should have the exact capacity and functionality required to manufacture a part family, allowing cost effectiveness [3]. It is suggested that grouping parts into families in RMSs has a positive effect on the introduction of new products in market [4]. The systematic part family problem for reconfigurable manufacturing system was first taken by Galan [5] stated that formation of part family in RMS has to be based on some grouping criteria and developed a hierarchical clustering of parts. The hierarchical clustering techniques have been commonly used in cellular manufacturing for machine grouping and part family formation [6]. These Hierarchical clustering methods always yielded a dendrogram representing the nested family of parts and similarity levels at which families change. There are various similarity coefficients available in the literature [7] for part family formation. Based on the results of comparative study of similarity coefficient, it is difficult to find the omnipotent similarity coefficient to solve part family formation because most of the existing methods suffer from one or more drawbacks. Their major common drawbacks are the inflexibility and the limited industrial application due to the unavailability of commercial software programs supporting them [8]. So, a new similarity coefficient is needed to overcome these limitations and to develop a comprehensive approach for part family formation which supports the available commercial software. To process the part families, Initially RMS is configured to produce the first selected family and then it is reconfigured to effectively produce the following part family, and so forth. In each change of configuration, the manufacturing company incurs in a changeover cost, which depends on the current configuration and the destination configuration. The time required for reconfiguration is considered insignificant as the RMS should be designed for rapid reconfigurations [2]. In the case of joining all parts in a single family, the manufacturing system is composed of all the machines required for manufacturing the parts. In this situation the company does not incur changeover costs, but idle machines and/or machines whose functionality is not fully used exist. On the contrary case, selecting a family for each part, the company has to face the costs of system changeover, the number of idle machines is minimized and their functionalities and capacities are both fully utilized. So, there is a clearly need that a smaller number of families than the number of different parts may be selected. In this situation, although the change over cost is not avoided, the changes are few and consequently the cost is low. Moreover, the idle machines are fewer than in first case and the functionalities and capacities of the machines are not fully used but the utilization rate is higher if compared to the first case. Concluding, the key parameters to take into consideration for selecting the part families are changeover cost of the system, cost of idle machines, cost of under-utilization of the capacity of the machines. So, minimum cost solution is clearly needed at each level of dendrogram to select the optimum set part family.

In this present work, a two phased methodology has been developed i.e. to group an amount of parts during a time horizon into families and to configure the resources needed to each part family, and to find the best sequence for all part families trying to minimize reconfiguration and under-utilization cost.

In first phase, a approach based on new similarity coefficient is developed for the part family formation. To this effect, a method has been proposed using correlation as a similarity coefficient. Furthermore, Principal Component Analysis (PCA) is used to cluster the parts on the basis of operational sequence similarity followed by agglomerative hierarchical K-means for identifying part groups on the basis of operational similarity.

Principal Component Analysis (PCA) is the best known and oldest technique in multivariate analysis [9, 10]. Pearson [11] was first to introduce it to recast linear regression analysis into a new form. PCA is frequently used for the data set problems with some intrinsic complexity [12–15]. Lokesh and Jain [16] used PCA for concurrent part machine group formation problem in CMS. It is a quantitatively rigorous method for achieving the simplification. The method generates a new set of variables, called Principal Components (PCs). Each PC is a linear combination of the original variables. All the PCs are orthogonal to each other, so there is no redundant information. The number of PCs extracted in a PCA is equal to the number of observed variables being analyzed. However, in most analyses, only the first few components account for meaningful amount of variance, and hence those first few components are retained, interpreted, and used in subsequent analysis and rest are neglected. When the analysis is complete, the resulting components display varying degrees of correlation with the observed variables, but are completely uncorrelated with one another.

The K-means algorithm is a clustering technique. The K-means algorithm randomly selects K data points as initial cluster centroids. A centroid is an artificial point in the space which represents an average location of the particular cluster. K clusters are formed by assigning each data point to its nearest centroid. New virtual centroids are then calculated for each cluster. These processes are iterated until a predefined number of iteration is reached or the clusters did not change anymore. The major problem with the K-means algorithm is that its initial starting points are generated randomly and does not guarantee the unique clustering results [17]. Also, due to the non-hierarchical nature of the algorithm, a hierarchical relationship between the clusters is needed. This hierarchical relationship is effective to visualize and analyze the large data sets. The hierarchical technique is classified into agglomerative method and divisive method. The divisive method is the top down approach in which initially all the objects are included in a single cluster. Then, the single cluster is divided into sub-clusters until each object constitutes a cluster. An agglomerative method is the bottom up approach in which each object is assumed as a separate cluster and then they are clustered in succession until a single cluster which consists of the entire object set is formed.

In the second phase, selection and sequencing of the part families can be solved calculating the cost at each level in the hierarchy (dendrogram), and the level with the minimum cost will be selected. Thus, all the possible solutions are to be evaluated. Therefore, for the selection of the families, a model that includes the key parameters of cost is required. For effective working of the RMS, a family of parts is selected for being produced. When finishing, the following family is ready to be produced within a specifically-designed system. This process is repeated until the production of each part family. Only when the companies add or remove parts from their portfolios, or their demand are very different, the methodology may select different part families and different sequences of production. This problem is very close to the travelling salesmaen problem(TSP), which seeks to identify an itinerary that minimises the total distance travelled by a salesman who has to visit a certain number of cities once, leaving from one of them (base city) and returning to this one. Therefore, some similarities must be highlighted. First, cities in the TSP are the part families in RMS, Second, the goal in the TSP is to minimise the total distance travelled. In RMS, the goal is to minimise the total cost. Finally, in TSP the salesman has to arrive at the base city, and in RMS when the last family has been produced the system is reconfigured for the first one. So, the proposed model solves a TSP in each level and selects the level which presents the minimum cost. Therefore, the problem to solve is a TSP-multilevel. As the TSP is a NP complete problem, the selection and sequencing of part families is NP complete too.

The outline of the paper is as follows: The proposed model is presented in the Section 2 followed by the methodology for grouping parts into families and mathematical model for selection and sequencing the part families are presented in Section 3. Afterwards, one numerical example gives illustration of the proposed methodology in phased manner in Section 4. Lastly, conclusions are drawn in Section 5.

2 Proposed model

Different types and quantities of parts must be manufactured within a certain time horizon in a Reconfigurable Manufacturing System, using Reconfigurable Machine Tools and their available Modules. A Reconfigurable System is configured with the necessary RMT and modules to manufacture a family of similar products at the same time. Once a family is manufactured, the system is reconfigured for manufacturing the following family effectively. In each change, the system incurs in a reconfigurable cost, which depends on the current configuration and the destination configuration. The proposed model possesses the following features as described below:

1. Manufacturer receives orders for Q different part types for manufacturing in the next reconfiguration cycle. The order receipt is closed at a certain pre-defined time limit before the completion of running reconfiguration cycle to provide sufficient time for the reconfiguration exercise. Out of the Q part types, only those P (≤Q) part types are accepted which satisfy two conditions as per company policy: (i) the total order d_i of all the customers for a particular part type P_i must exceed a pre-decided minimum quantity, Dmin(i) (i.e. $d_i \geq Dmin(i)$) and (ii) the anticipated execution time of all the orders must not exceed a predefined maximum allowable reconfiguration cycle span.
2. The operations required to manufacture a part and their sequences are stated on a process plan. The plan so selected could be used for the manufacturing of full demand volume of the concerned part type. Therefore, it is being proposed that the system planner will have the selected process plan in hand for formation of part families.
3. Each part, P_i is manufactured in ordered quantity, d_i using various operations by one of the various alternative operation sequences. It has been proposed to give equal priority to all alternative operation sequences and to select only one operation sequence from the set of alternatives for the manufacturing of full demand volume of the concerned part.
4. These part types are divided into k part families so that 1≤k≤P. Therefore, in extreme cases, either all the part types will fall in the same family or each family will be consisting of only one part type. The key attribute of a part family is that all the part types within a family require same operations and hence same production resources.
5. A part is associated to a family based on the number of common operations that it shares with other members of the family. A reconfigurable machine tool has to remain idle when a member of the family does not require an operation. A higher value of similarity coefficient implies a higher number of operations are common and therefore less machine idleness. Therefore, it is proposed to choose operations based on the highest similarity coefficient. It implies that operation sequences correlation will be beneficial to use in the clustering procedure and accordingly part family will be selected.
6. Selection and sequences of the product families with the minimum operational costs (that is, reconfiguration costs and the costs of under-utilization of resources).

Figure 1 shows a schematic representation of the proposed model approach, with P_i (say P_1, P_2, P_3…P_{11}) types of parts to be manufactured within a certain time horizon. Based on the result of the first phase (Part Families Formation) k families are formed (say {P_2-P_6-P_9}, {P_1-P_3-P_{11}-P_7}, {P_4-P_5}, {P_8-P_{10}}). Each part family require a different configuration of the RMS. These similar parts allow a reduction of Setups, Lead-times, Work-In-Process and Material Handling, thus increasing productivity. For each part family the system is reconfigured and corresponding total cost are

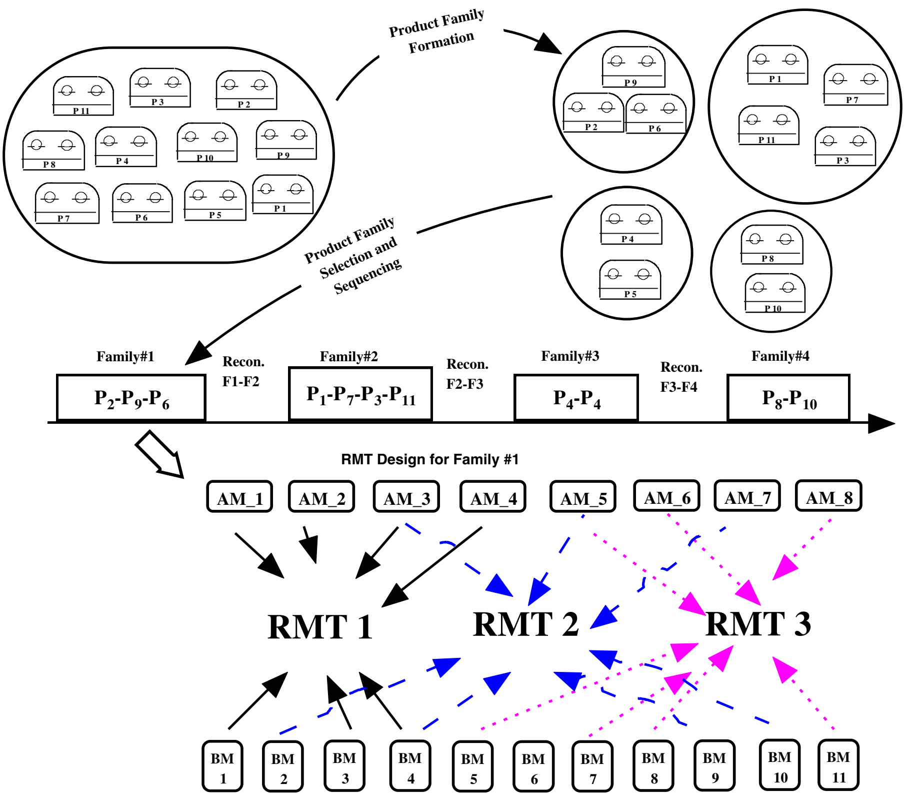

Fig. 1 Schematic representation of the proposed approach

calculated and minimum cost solution for sequence of part families obtained in the second phase.

3 Methodology

3.1 Formation of part family

In the proposed RMS model, it has been assumed that all the parts belonging to a family are manufactured on a system configuration that provides all the resources required for their manufacturing. All the parts are manufactured sequentially (one after the other) on the same system and orders are accepted as per a predefined policy which ensures that the manufacturing time of all the parts does not exceed the cycle span. Therefore, demand data will not influence the decision of part family formation. It implies that for a part to be a member of a family, commonality of operations is a sufficient condition. Thus, the operations required to manufacture a part are the most relevant attributes for the considered RMS model, consistent with the objective of minimising the number and hence, cost of reconfigurations. Therefore, binary part-operation incidence matrix (POIM) is the sufficient input for this problem. In this section, a method based on PCA and AHKCA is developed to form part family on the basis of the input received.

The objective is to cluster the parts into k part families based on operations similarity. The proposed methodology consists of three stages as shown in Fig. 2.

3.1.1 Development of similarity coefficient matrix

The first phase starts with building a similarity coefficient matrix. The initial part-operation incidence matrix as shown in Eq. (1) is a binary matrix in which rows represent the operations and columns stand for parts. This matrix looks like the transpose of the classical part-machine incidence matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \tag{1}$$

Where a_{ij} =1 if part j requires operation i and a_{ij}=0 otherwise.

Let P_j^A is a binary row vector of matrix A, such that $P_j^A = \left[a_{1j}, a_{2j}, \ldots, a_{nj}\right]$.

Initial matrix A is further standardized by using a suitable method of standardization [18]. In this work, the general standardization is used and applied to the initial incidence matrix. The standardization process is expressed as follows:

$$P_j^B = \frac{P_j^A - E_j}{\sigma_j} \tag{2}$$

Where, E_j is the average of row vector P_j^A and P_j^B is the row vector of the standardized matrix B. E_j can be expressed as:

$$E_j = \frac{\sum_{k=1}^{n} a_{kj}}{n} \tag{3}$$

Here a_{kj} is the element of initial incidence matrix; n is the number of elements in a row vector. Similarly,

$$\sigma_j^2 = \text{variance} = \frac{1}{n}\sum_{k=1}^{n}\left(a_{kj} - E_j\right)^2 \tag{4}$$

To simplify the Eq. 4 further, Huyghens Koning theorem is applied to yield

$$\sigma_j^2 = \sqrt{E_j - \left(E_j\right)^2} \tag{5}$$

Once the standardized matrix is formed, the proposed similarity coefficient is based on the simple correlation matrix of the standard incidence matrix. The correlation matrix S is defined as follows:

$$S = \frac{1}{n}B^T B \tag{6}$$

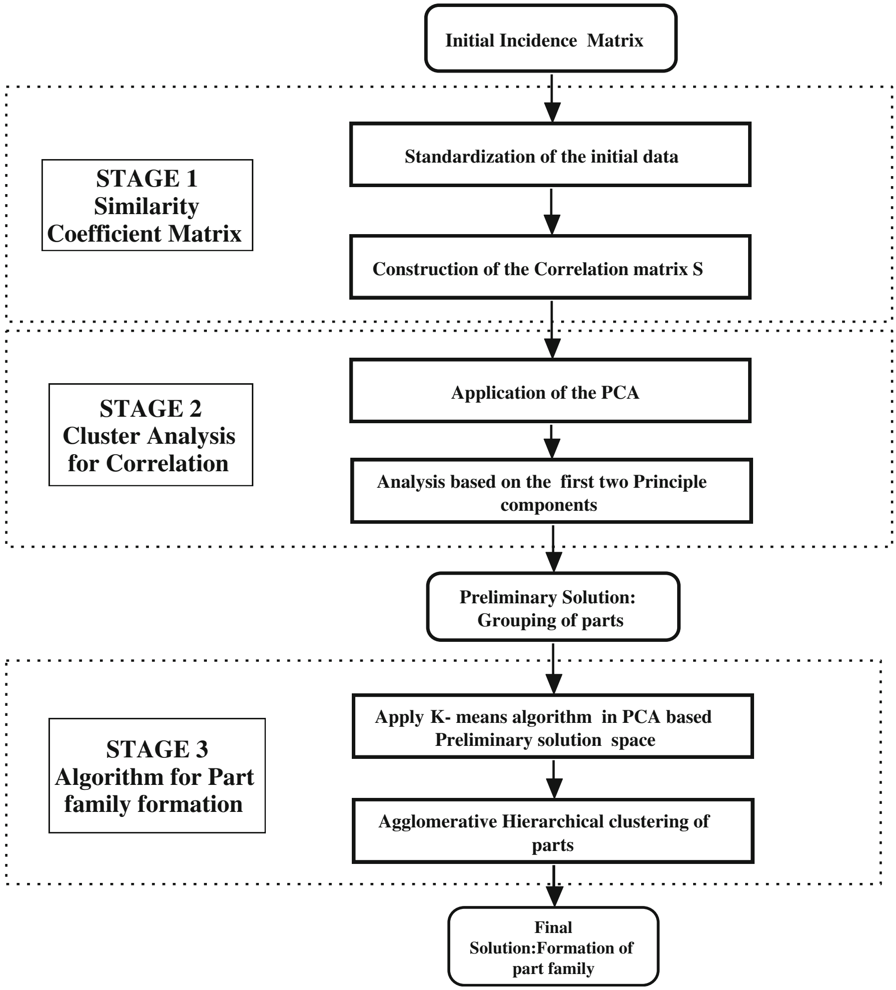

Fig. 2 Methodology for part family formation

S is the square matrix in which elements are given by:

$$\mathrm{S_{ii}} = 1 \quad \text{and} \quad \mathrm{S_{ij}} = \frac{1}{n}\sum_{k=1}^{n} b_{ik}b_{jk} \tag{7}$$

3.1.2 Cluster analysis for correlation

In the second phase of the proposed approach, the part family formation is done on the basis of operations similarity using Principle Component Analysis. Principle Component Analysis is a dimension reduction technique which attempts to model the total variance of the original data set, via new uncorrelated variables called Principal Components. PCA consists of determining a small number of principal components that recover as much variability in the data as possible. These components are linear combinations of the original variables and account for the total variance of the original data. The first principle component is a single axis in space. When each observation is projected on that axis, resultant is a new variable. The variance of this variable is maximum among all possible choices of the first axis. The second principle component is another axis in space, perpendicular to the first one. Projecting the observations on this axis generates another new variable. The variance of this new variable is again maximum among all possible choices of this second axis. The full set of principal components is as large as the original set of variables. However, the sum of the variances of the first few principal components is usually 80 % or more of the total variance of the original data [19].

The first component extracted in principal component analysis accounts for a maximal amount of total variance in the observed variables. Under typical conditions, this means that the first component is correlated with at least some of the observed variables. The second component extracted is having two important characteristics. First, this component accounts for a maximal amount of variance in the data set that was not accounted for by the first component. This means that the second component is correlated with some of the observed variables that did not display strong correlations with first component. The second characteristic of the second component is that it is uncorrelated with the first component. This means that the correlation between first and second principle components is zero [20].

The remaining principle components that are extracted in the analysis display the same two characteristics mentioned above. That is, each principle component accounts for a maximal amount of variance in the observed variables that was not accounted for by the preceding components, and is uncorrelated with all of the preceding components. Principal components analysis proceeds in this fashion, with each new component accounting for progressively smaller and smaller amounts of variance.

Thus, the study of principal components is considered as putting into terms the usual developments of eigenvalues and eigenvectors for positive semi-definite matrices. The eigenvector equation where the terms $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ are real, non-negative roots of the determinant polynomial of degree P is given as:

$$\det(S - \lambda_i I) = 0; \qquad i \in < 1, m > \tag{8}$$

Let $\{F_1, F_2, \ldots, F_m\}$ be corresponding eigenvectors. When PCA was performed on the mean centered data, a model with the first and the second principal components was usually obtained. This model explains the procedure to determine the principle component in the data.

$$\text{where,} \quad \mathrm{PC} = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^{m} \lambda_k} = \frac{\lambda_1 + \lambda_2}{m} \tag{9}$$

In this application of PCA, the objective is to cluster parts into families. As part-operation matrix is binary in nature, two principal components are enough to analyse correlation between elements (i.e. parts).

3.1.3 Algorithm for part family formation

The objective of the third phase is to assign parts into families after the preliminary grouping as done in the second phase of the proposed approach. An Agglomerative Hierarchical K-means Clustering Algorithm is used for this purpose. The algorithm for assigning the parts (P_i) into parts families is given below:

For each part k=1 to P_i do

Step 1: Take n points (Initial seeds) in 2-dimensional plane, where n is the number of variables. (Initial seeds are taken based on score values (coordinate) of each part on first two principle components. Each seed represents a part which contains associated part operations.)

Step 2: Compute Euclidean distance for each part.

$$(P_k,\ P_i) = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{10}$$

Where, x_i and y_i are the co-ordinates of part P_i on two principal components axis.

Step 3: Since the objective is to group parts with minimum distance, part under consideration (say P_i) is assigned to a family (say P_k) on the basis of least smallest distance to the part family P_k.

Step 4: Draw the dendrogram of the sequence of family formation.

The iteration continues until all parts (P_i) are assigned to part families.

3.2 Part families selection and sequencing

To select the dendogram level and the sequencing of the corresponding families a minimum costs criterion is proposed. The different operating costs for manufacturing the parts in a RMS is defined as

- Reconfiguration costs: costs due to changing the configuration of the RMS (machines and modules) when changing the part family to manufacture. These costs include cost of adding or removing from the system a module k from a machine m (α_{mk});Cost of removing from the system a machine m currently not required (β_m); Cost of adding a required machine m (γ_m).
- Underutilization costs: costs for not using at the same level all the resources assigned to the system for the manufacturing of all the parts in a family. These costs include cost per unit time of underutilization of module k in a machine m (ε_{mk}).

From the above costs coefficients the reconfiguration costs ($RC_{FF'}$) of changing from the parts of family F to the parts of family F′, both families belonging to the same dendogram level is computed as the total costs of adding and removing machines and modules.

$$RC_{FF'} = \sum_{m=1}^{M} \beta_m \cdot \Delta\left(\sum_{i\in F}\sum_{k=1}^{K} OPT_{imk}, \sum_{i\in F'}\sum_{k=1}^{K} OPT_{imk}\right) + \sum_{m=1}^{M} \gamma_m \cdot \Delta'\left(\sum_{i\in F}\sum_{k=1}^{K} OPT_{imk}, \sum_{i\in F'}\sum_{k=1}^{K} OPT\right) + \sum_{m=1}^{M}\sum_{k=1}^{K} \alpha_{mk} \cdot \left(\Delta\left(\sum_{i\in F} OPT_{imk}, \sum_{i\in F'} OPT_{imk}\right) + \Delta'\left(\sum_{i\in F} OPT_{imk}, \sum_{i\in F'} OPT_{imk}\right)\right) \tag{11}$$

where the indicator functions $\Delta$and Δ' are defined as: $\Delta(a,b)=1$ if $a>0$ and $b=0$; $\Delta'(a,b)=1$ if a=0 and $b>0$.

Also, the resource underutilization costs (UC_F) associated to a family F can be computed as the costs of not maximally using all the resources allocated to the family. Thus, compared with the resource with maximum utilization, all the other resources incur costs proportional to the difference in utilization levels, i.e.

$$UC_F = \sum_{m=1}^{M}\sum_{k=1}^{K} \varepsilon_{mk} \cdot \left(\underset{\forall(m',\mathrm{k},)}{Max}\left(\sum_{i\in F'} OPT_{im'\mathrm{k}'}\right) - \left(\sum_{i\in F} OPT_{imk}\right)\right) \tag{12}$$

From the given data the processing time (OPT_{imk}) of each product i on each resource (i.e. on each combination of machine m and module k) is computed as:

$$OPT_{imk} = d_i \cdot \sum_{j=1}^{P(i)} t_{ij} \cdot \delta_{o(j,i)mk} \tag{13}$$

Given the reconfiguration and underutilization costs computed above the problem of finding the optimal dendogram families and their sequencing corresponds to solving a Travelling Salesman Problem (TSP) for each dendogram level and selecting the level with minimum cost. There exist in the literature different MILP formulations of TSP [21].

Nomenclature

i, j	Indexes of the product families to manufacture
k	Index of the K available auxiliary modules
m	Index of the M available machine tools (RMT)
d_i	Demand of product i along time horizon considered
o	Index of the operation types
o(j,i)	Operation type of manufacture task j of product i
t_{ij}	Duration of manufacture task j of product i
	$\delta_{omk} = \begin{cases} 1 & \text{if operation O requires module k in machine m} \\ 0 & \text{otherwise} \end{cases}$
l	Index of the dendogram levels
L	Number of dendogram levels (level 1=each product forms a family; level L=all products belong to the same, single family)
F_l	Set of product families at dendogram level l, l =1,…,L
N_l	Number of product families at dendogram level l, i.e., $N_l=\lvert F_l\rvert$
R_{ijl}	Reconfiguration cost of changing from family i to family j both belonging to dendogram level $(i \in F_l \quad \mathrm{j} \in F_l \quad \mathrm{j} \neq i \quad l = 1, \ldots, L-1)$

H_{il} Resource underutilization cost of family i of dendogram level l $(i \in F_l \quad l = 1, \ldots, L)$

Variables

- T_{ijl}=1, if products of family i in dendogram level l are manufactured just before those of family j both belonging to dendogram level l $(i \in F_l \quad j \in F_l \quad j \neq i \quad l = 1, \ldots, L)$
- K_l=1, if the product families of dendogram level l are selected $(\ l = 1, \ldots, L)$
- U_{il} ≥0, ancillary variables used to prevent family assignment cycles in each dendogram level(sub tour) $(i \in F_l \quad l = 1, \ldots, L-5)$

Objective function The objective function is the minimization of the sum of reconfiguration costs between families plus underutilization of resources assigned to each family.

MILP Model

$$Min \sum_{l=1}^{L-1} \sum_{i \in F_l} \sum_{\substack{j=F_l \\ j \neq i}} R_{ijl} T_{ijl} + \sum_{l=1}^{L} K_l \sum_{i \in F_l} H_{il}$$

subject to

$$(0) : \sum_{l=1}^{L} K_l = 1$$

$$(1) : \sum_{i \in F_l} T_{ijl} = K_l$$

$$(2) : \sum_{j \in F_l} T_{ijl} - \sum_{j \in F_l} T_{jil} = 0 \qquad \forall i \in F_l; l = 1, \ldots, L-1$$

$$(3) : \sum_{\substack{j \in F_l \\ j \neq 1}} T_{1jl} = K_l \qquad \forall l = 1, \ldots, L-1$$

$$(4) : N_l \cdot T_{ijl} + U_{il} - U_{jl} \leq N_l - 1 \qquad \forall i \in F_l - 1; \quad j \in F_l - 1; \quad j \neq i; \quad l = 1, \ldots, L-5$$

$$(5) : T_{ijl} = [0, 1] \qquad \forall i \in F_l; \quad j \in F_l; \quad l = 1, \ldots, L-1$$

$$(6) : K_l = [0, 1] \qquad \forall l = 1, \ldots, L$$

$$(7) : U_{il} \geq 0 \qquad \forall i \in F_l - 1; \quad l = 1, \ldots, L-5$$

Constraints

0 One and only one dendogram level is to be selected

1 The families corresponding to the selected dendogram level will be sequentially processed once in each processing cycle. The families corresponding to the other dendogram levels will not be considered

2 For each family of the selected dendogram level there must be exactly one predecessor family and another successor family both of the selected level

3 Since the different families selected are sequenced cyclically, the first one can be any of them

4 Anti-cycle constraints

5 Integrality (variables T and K) and non-negativity (variables U) constraints

4 Numerical illustration

An example case has been taken to demonstrate the proposed methodology where seven operations are required for manufacturing of parts in next reconfiguration cycle. Table 1 shows the operation sequences, the total demand, d_i and the minimum acceptable demand size Dmin(i) for each part type. Table 2 shows the operations capability of the RMT resources in terms of processing time. Table 3 shows the resources (RMT) available for the manufacturing. The RMT are configured by two base module and maximum of three auxiliary module for the example problem.

From the above available data, POIM is constructed as shown in Table 4. Any order $d_i <$ Dmin(i) is not considered, hence P_{12} is not considered here. So, the POIM (Table 4) has total 11 parts and 7 operations

Now, applying the Eqs. 2, 3 and 5 to the part operation incidence matrix given above, yield standardized matrix B, which represents Standard Sequence Part Operation Matrix (SSPOM) in the case. The following procedure is adopted to determine various elements of SSPOM.

Say, for part P_1,

$$E_1 = 2/7 = 0.286$$
$$\sigma_1 = \sqrt{0.286 - (0.286)^2} = 0.452$$

The member coefficient between P_1 and OP_1 (i.e. b_{11}) is calculated as follows:

$$b_{11} = \frac{1 - 0.286}{0.452} = 1.580$$

Similarly, $b_{21} = \frac{1-0.286}{0.452} = 1.580$ and $b_{31} = \frac{0-0.286}{0.452} = -0.633$

The same procedure is repeated for other elements. Finally, SSPOM is obtained as shown in Table 5.

Table 1 Parts demand data and operations sequence with operation time

Part no.	Dmin(i)	Demand (di)	Operation sequence with operations time
P1	40	60	$OP_1(140)$ -$OP_2(115)$
P2	100	150	$OP_2(150)$ -$OP_3(135)$
P3	100	250	$OP_1(85)$- $OP_5(100)$-$OP_6(65)$
P4	50	80	$OP_4(70)$- $OP_5(65)$ -$OP_6(55)$ – $OP_7(60)$
P5	50	100	$OP_4(75)$ -$OP_5(85)$- $OP_6(95)$
P6	100	150	$OP_2(150)$-$OP_3(135)$
P7	100	120	$OP_1(90)$- $OP_6(110)$
P8	120	180	$OP_6(145)$-$OP_7(105)$
P9	70	110	$OP_3(160)$
P10	90	160	$OP_5(95)$-$OP_6(75)$ -$OP_7(80)$
P11	100	200	$OP_1(125)$- $OP_5(130)$
P12	120	80	$OP_6(100)$-$OP_7(130)$

Table 2 Operations capability of RMT in terms of processing time

	M_1	M_2			M_3			M_4	M_5		M_6	
	M11	M21	M22	M23	M31	M32	M33	M41	M51	M52	M61	M62
OP_1	140	–	85	–	90	–	–	–	125	–	–	–
OP_2	–	115	–	–	–	150	–	120	–	–	–	–
OP_3	–	–	–	110	–	–	–	–	135	–	160	–
OP_4	–	–	–	–	–	–	70	–	–	75	–	–
OP_5	100	–	65	–	–	85	–	–	–	–	–	95
OP_6	75	65	–	–	55	–	–	95	–	110	145	–
OP_7	–	–	–	80	–	–	105	–	–	–	–	60

From the above SSPOM and on the basis of proposed similarity coefficient (as given in the Eq. 7), the correlation matrix (S) is obtained as shown in Table 6.

Now cluster analysis is performed based on PCA method and by using Eq. 8. The computed eigenvalues for the correlation matrix (S) and their associated variance, and cumulative variance are listed, sorted in a descending order as shown in Table 7.

The graphical analysis is performed on a two dimensional scatter plot where each part is represented by a dot. The scatter plot indicates the relationship between two parts as shown in Fig. 3. There is high correlation between parts which are closely placed and thereby are strongly associated with each other such as (P_2 and P_6) and (P_4 and P_5). The correlation values are shown in Table 6. On the basis of correlation results, the following principle situations are recovered from the scatter plot:

i. Two neighboring parts having low distance measure belong to the same group such as P_2 and P_6 and P_4 and P_5.

Table 3 RMT Resources required for Manufacturing

Machine	Machine Module	Basic Module (BM)	Auxiliary Module (AM)
M_1	M11	{01, 02}	{01,03,05}
M_2	M21	{03,05}	{08,10,11}
	M22	{03,05}	{08,05,07}
	M23	{03,05}	{06,07,10}
M_3	M31	{01,04}	{02,09}
	M32	{01,04}	{02, 09,11}
	M33	{01,04}	{02,09,10}
M_4	M41	{02,05}	{03,08,}
M_5	M51	{03,06}	{04,09,10}
	M52	{03,06}	{04,07,09,}
M_6	M61	{04,07}	{01,03,07}
	M62	{04,07}	{01,03}

Table 4 Part Operation Incidence Matrix: POIM A

		P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}
POIM A =	OP_1	1	0	1	0	0	0	1	0	0	0	1
	OP_2	1	1	0	0	0	1	0	0	0	0	0
	OP_3	0	1	0	0	0	1	0	0	1	0	0
	OP_4	0	0	0	1	1	0	0	0	0	0	0
	OP_5	0	0	1	1	1	0	0	0	0	1	1
	OP_6	0	0	1	1	1	0	1	1	0	1	0
	OP_7	0	0	0	1	0	0	0	1	0	1	0

ii. Part group (P_4 and P_5) which is almost 180° to the other part group (P_6 and P_9) are negatively correlated and thus cannot belong to the same group.
iii. Two parts which are placed almost 90° to each other such as P_9 and P_8 are independent and thus cannot belong to the same group. It is clearly seen in the Fig. 3.

The co-ordinate (score) of each part on first two principle components is obtained from scatter plot as shown in Table 8.

The third phase of the methodology is to cluster parts, obtained from the preliminary solution space in second phase, into family by using AHKCA mentioned in the methodology. The algorithm is initialized by taking score values (coordinates) of parts from scatter plot as a starting point. The algorithm starts with the initial solution i.e. the number of parts and then computes and stores the euclidean distance between each part using Eq. 10. The parts having minimum distance are grouped first and these grouped parts are removed from the subsequent iteration. This process continues until all the parts are grouped together. The algorithm finally provides output in the form of dendrogram (Fig. 4).

The resulting dendrogram is based on the distances to K-means nearest group. In K-means algorithm, the measured distance is known as dissimilarity measure. Hence, the magnitude of the distance to nearest group of parts represents the dissimilarity

Table 5 Standard Sequence Part Operation Matrix: B

		P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}
B =	OP_1	1.580	−0.633	1.154	−0.633	−0.633	−0.633	1.580	−0.409	−0.409	−0.633	1.580
	OP_2	1.580	1.580	−0.867	−0.633	−0.633	1.580	−0.633	−0.409	−0.409	−0.633	−0.633
	OP_3	−0.633	1.580	−0.867	−0.633	−0.633	1.580	−0.633	−0.409	2.449	−0.633	−0.633
	OP_4	−0.633	−0.633	−0.867	1.580	1.580	−0.633	−0.633	−0.409	−0.409	1.580	−0.633
	OP_5	−0.633	−0.633	1.154	−0.633	−0.633	−0.633	1.580	−0.409	−0.409	−0.633	−0.633
	OP_6	−0.633	−0.633	1.154	1.580	1.580	−0.633	−0.633	−0.409	−0.409	−0.633	1.580
	OP_7	−0.633	−0.633	−0.867	−0.633	−0.633	−0.633	−0.633	2.449	−0.409	1.580	−0.633

Table 6 Correlation Matrix: S

		P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}
S =	P_1	1.000										
	P_2	0.300	1.000									
	P_3	0.091	−0.548	1.000								
	P_4	−0.200	−0.400	0.091	1.000							
	P_5	0.400	−0.400	0.091	0.949	1.000						
	P_6	0.300	0.999	−0.548	−0.400	−0.400	1.000					
	P_7	0.300	−0.400	0.730	−0.400	−0.400	−0.400	1.000				
	P_8	−0.258	−0.258	−0.354	−0.258	−0.258	−0.258	−0.258	1.000			
	P_9	−0.258	0.645	−0.354	−0.258	−0.258	0.645	−0.258	−0.167	1.000		
	P_{10}	−0.200	−0.400	−0.548	0.300	0.300	−0.400	−0.400	0.645	−0.258	1.000	
	P_{11}	0.300	−0.400	0.730	0.300	0.300	−0.400	0.300	−0.258	−0.258	−0.400	1.000

associated with the part. The ratio of the distance to the nearest group to the maximum distance measured in the dendrogram gives the percentage of dissimilarity of the particular parts group. In Fig. 4, parts P_2 and part P_6 are found at zero distance. So their percentage of dissimilarity is zero. It means that they are 100 % similar in operations. The distance between parts P_4 and P_5 is 0.1464 and their associated percentage of dissimilarity is calculated to 6.0 % and level of similarity is 94 %. The same procedure is repeated to calculate the other similarity level for remaining cluster of parts. The obtained results are summarized in Table 9.

The dendrogram shown in the Fig. 4 presents the different families that are formed and their associated similarity level is presented in the Table 9. At level 1, 10 families are formed named $P_{2,6,}P_1$,P_3, P_4,P_5,P_7,P_8,P_9,P_{10},P_{11}. At level 2, nine families are formed, two families composed of two parts P_2,P_6; P_4,P_5 and rest are composed of one part (i.e. P_1,P_3, P_7,P_8,P_9,P_{10},P_{11}).Similarly for different level, others set of part families are obtained. Finally, at level 10 only one family

Table 7 Principle components, Eigen value and Percentage Variance

No of principal components	Eigen value	% variance	Cumulative % variance
1	1.52e+000	50.04	50.04
2	8.53e-001	28.11	78.15
3	4.50e-001	14.83	92.97
4	1.50e-001	4.94	97.91
5	3.66e-002	1.21	99.12
6	1.58e-002	0.52	99.64
7	9.13e-003	0.30	99.94
8	1.88e-003	0.06	100.00
9	2.65e-007	0.00	100.00
10	5.60e-008	0.00	100.00
11	3.14e-008	0.00	100.00

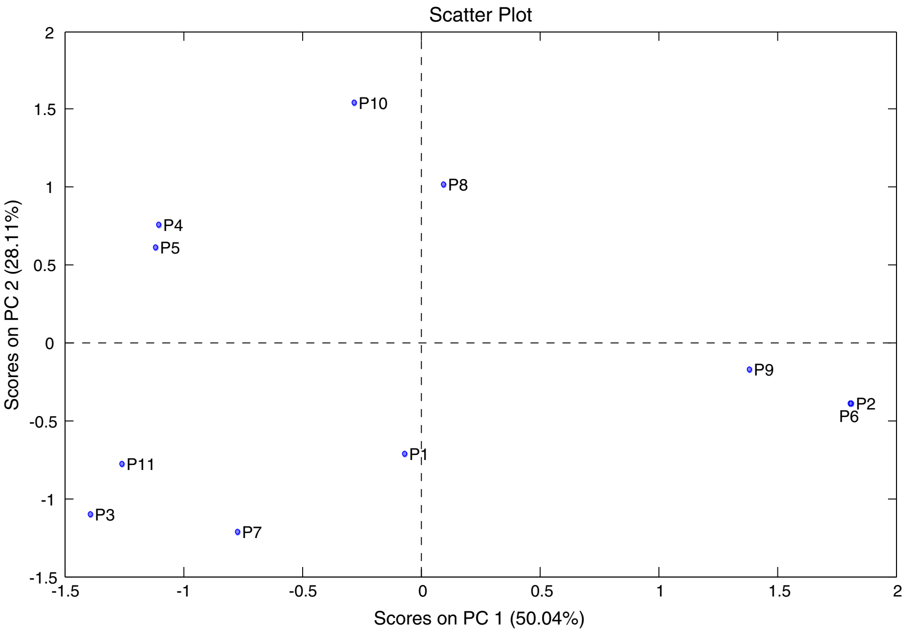

Fig. 3 Graphical representation of Scatter Plot

is formed composed of all parts at the lowest similarity. On the basis of the formed part families the system planner first configures the manufacturing system to produce the first part family. Once it is finished, the system is reconfigured to produce the second part family and so forth. Each system reconfiguration adds cost to the production of the parts. Hence there is a need to arrive at a suitable number of part families by selecting the dendogram level and the sequencing of the corresponding families to achieve minimum cost solution of the problem.

As defined in Section 3.2, cost of parameters regarding reconfigurability must be estimated on the basis of experience. Let us take cost of adding or removing from the

Table 8 Score (Co-ordinate) of parts

Part Number	First Principle Component (50.04 %)	Second Principle Component (28.11 %)
P1	−0.0699	−0.7105
P2	1.8086	−0.3856
P3	−1.3957	−1.1034
P4	−1.1100	0.7616
P5	−1.1208	0.6156
P6	1.8086	−0.3856
P7	−0.7736	−1.2148
P8	0.0917	1.0202
P9	1.3827	−0.1736
P10	−0.2835	1.5465
P11	−1.2622	−0.7806

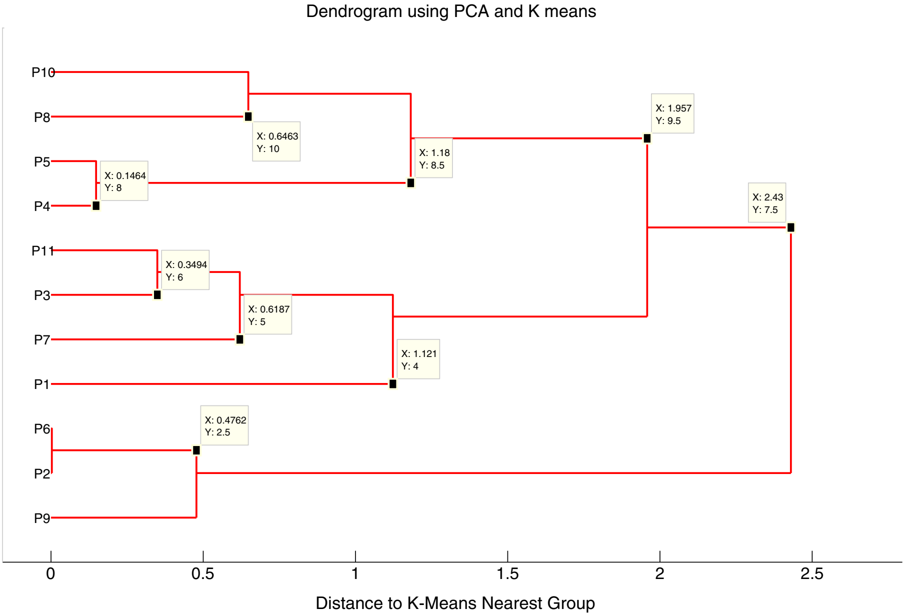

Fig. 4 Hierarchical Clustering of parts

system a module k from a machine m (α_{mk}) equal to 1; cost of removing from the system a machine m currently not required (β_m) equal to 10; Cost of adding a required machine m (γ_m) equal to 10 and cost per unit time of underutilization of module k in a machine m (ε_{mk}) equal to 7 for all machine m and module k.

An example of cost estimation for family's reconfiguration from P_4 to $P_5$is presented in Table 10.This table shows that machine M_1 is not used for producing P_5 in new system. Therefore machine M_1 has to be removed and its corresponding cost (β) is incurred. On the contrary machine M_6 is included in producing the part P_5 in system; therefore a corresponding cost (γ) is incurred. Machine M_2 is required for producing both parts family, so no cost is incurred because it remains in new system.

Table 9 Percentage of similarity and Formed part family

Dendrogram level	Distance between parts group	Similarity level (%)	Formed part families	Number of families
L1	0.000	100	$P_{2,6}$,P_1,P_3, P_4,P_5,P_7,P_8,P_9,P_{10},P_{11}	10
L2	0.1464	94.0	$P_{2,6}$,P_1,P_3, $P_{4,5}$,P_7,P_8,P_9,P_{10},P_{11}	09
L3	0.3494	85.6	$P_{2,6}$,P_1,$P_{3,11}$, $P_{4,5}$,P_7,P_8,P_9,P_{10}	08
L4	0.4762	80.4	$P_{2,6,9}$,P_1,$P_{3,11}$, $P_{4,5}$,P_7,P_8, P_{10}	07
L5	0.6187	74.5	$P_{2,6,9}$,P_1,$P_{3,11,7}$, $P_{4,5}$, P_8, P_{10}	06
L6	0.6463	73.4	$P_{2,6,9}$,P_1,$P_{3,11,7}$, $P_{4,5}$, $P_{8,10}$	05
L7	1.121	53.9	$P_{2,6,9}$,$P_{1,3,11,7}$, $P_{4,5}$, $P_{8,10}$	04
L8	1.180	51.4	$P_{2,6,9}$,$P_{1,3,11,7}$, $P_{4,5,8,10}$	03
L9	1.957	19.5	$P_{2,6,9}$,$P_{1,3,11,7,4,5,8,10}$	02
L10	2.430	2.8	$P_{2,6,9,1,3,11,7,4,5,8,10}$	01

Table 10 Cost of reconfiguration from part family P_4 to P_5

Module	Machine	P_4	P_5	Parameter	Cost
M11	M_1	1	0	β	10
M21	M_2	0	0	–	
M22		1	1	–	0
M23		0	0	–	
M31	M_3	1	0	α	1
M32		0	1	α	1
M33		0	0	–	
M41	M_4	0	0	–	
M51	M_5	1	0	α	1
M52		0	1	α	1
M61	M_6	0	0	–	
M62		0	1	γ	10
					24

Now module M31and M32 are required only for one family, therefore they must be removed and included, respectively. Similarly module M51 and M52 are required only for one family, therefore module adding and removing cost is incurred, respectively.

Cost of under utilization of resources deal with incurred cost due to not using machines or modules. These costs have to be calculated for each family formed in any level of dendrogram. As the RMS is configured for producing a specific family, when producing a family composed of one part the system is installed with the capacity and functionality needed allowing full utilization of the resources. Consequently, the cost of under-utilization for families composed of one part is zero. On the basis of cost estimation parameters, Table 11 shows the cost of under-utilization in family P_2,

Table 11 Cost of under-utilization in Family $P_2,P_6,P_9,P_1,P_3,P_{11},P_7,P_4,P_5,P_8,P_{10}$

Module	Machine	P_2	P_6	P_9	P_1	P_3	P_{11}	P_7	P_4	P_5	P_8	P_{10}	Parameter	Cost
M11	M_1	0	0	0	0	0	0	0	0	0	0	0	11ε	77
M21	M_2	1	0	0	1	1	0	0	0	0	0	1	15α+2ε	29
M22		0	0	0	1	1	1	1	1	1	0	0		
M23		1	0	1	0	0	0	0	0	0	0	0		
M31	M_3	0	0	0	0	1	0	0	1	1	1	1	12α+4ε	40
M32		0	1	0	0	0	1	0	0	0	0	0		
M33		0	0	0	0	0	0	0	1	1	0	0		
M41	M_4	0	0	0	0	0	0	1	0	0	0	0	10ε	70
M51	M_5	0	1	0	0	0	0	0	0	0	0	0	α+10ε	71
M52		0	0	0	0	0	0	0	0	0	0	0		
M61	M_6	0	0	0	0	0	0	0	0	0	0	0	3α+8ε	59
M62		0	0	0	0	0	0	0	1	0	1	1		
														346

P_6,P_9,P_1,P_3,P_{11},P_7,P_4,P_5,P_8,P_{10}. For example focusing on machine M_2, module M23 is not used in producing part P_1 and the cost (α) is incurred. Similary producing part P_2 module M22 is not used so another cost (α) is incurred. Now, Machine M_2 is not required for producing P_6 and P_8, therefore corresponding cost (ε) is incurred each time. Similarly, by adding all these cost, finally under-utilization cost for machine M_2 is $15\alpha+2\varepsilon$.

Once both cost for reconfiguration and under-utilization have been calculated, the mathematical model formulated in Section 3.2 is solved using optimization software LINGO 7.0 giving the optimal solution of 319. The binary variables selected are $K_7=1$, $T_{\{P2,6,9\},\{P1,3,11,7\},7}=1$, $T_{\{P1,3,11,7\},\{P4,5\},7}=1$, $T_{\{P4,5\},\{P8,10\},7}=1$ and $T_{\{P8,10\},\{P2,6,9\},7}=1$. This means that optimal solution corresponds to the dendrogram level 7 which, looking at Table 9, corresponds to the families, namely $\{P_{2,6,9}\}$, $\{P_{1,3,11,7}\}$, $\{P_{4,5}\}$ and $\{P_{8,10}\}$. The variable T_{ijl} indicate the minimum cost order in which the four family must be sequenced. In this case there are six possible sequences: namely $P_{2,6,9}$ -$P_{1,3,11,7}$ - $P_{8,10}$ - $P_{4,5}$; $P_{2,6,9}$ -$P_{1,3,11,7}$ -$P_{4,5}$ - $P_{8,10}$; $P_{2,6,9}$ -$P_{4,5}$-$P_{1,3,11,7}$ - $P_{8,10}$; $P_{2,6,9}$ -$P_{4,5}$- $P_{8,10}$ - $P_{1,3,11,7}$; $P_{2,6,9}$ - $P_{8,10}$ -$P_{4,5}$-$P_{1,3,11,7}$ and $P_{2,6,9}$ - $P_{8,10}$ -$P_{1,3,11,7}$ -$P_{4,5}$. The sequence obtained by Lingo 7 is the second one. The minimum cost solution is validated computing solution using Eqs. 11, 12 and 13. The minimum cost solution from all possible family sequences for all levels of the dendrogram is presented in the Table 12.

Table 12 shows the minimum cost solution of the problem at each level of the dendrogram. At level 1, 45 possible sequences are found; out of which minimum cost solution at level 1 is shown in the Table 12. The total cost obtained at this level is 377 which include only the reconfiguration cost, because at this level, families are formed by a single part. Therefore, for each family, RMS is configured specifically for one part which means that there is no under-utilization cost although there are frequent reconfigurations which incurred cost. Similarly number of possible solutions and their corresponding total cost are found at the each level of dendrogram (Table 12). In

Table 12 Total cost of optimal Sequence of part families at different level

Level	No. of possible sequence	Minimum cost sequence of families at level	Reconfiguration cost	Under-Utilization cost	Total cost
1	45	$P_{2,6}$ -P_1 -P_3 - P_4 -P_5 -P_7 -P_8 -P_9 -P_{10} -P_{11}	377	0	377
2	36	$P_{2,6}$ -P_1 -P_3 -$P_{4,5}$ -P_7 -P_8 -P_9 -P_{10} -P_{11}	345	24	369
3	28	$P_{2,6}$ -P_1 -$P_{3,11}$- $P_{4,5}$ -P_7 -P_8 -P_9 -P_{10}	278	79	357
4	21	$P_{2,6,9}$ -P_1 -$P_{3,11}$ -$P_{4,5}$ -P_7 -P_8 - P_{10}	215	138	353
5	15	$P_{2,6,9}$ -P_1 -$P_{3,11,7}$ -$P_{4,5}$ -P_8 -P_{10}	156	192	348
6	10	$P_{2,6,9}$ -P_1 -$P_{3,11,7}$ -$P_{4,5}$ - $P_{8,10}$	123	214	337
7	**6**	$\mathbf{P_{2,6,9}}$ **-**$\mathbf{P_{1,3,11,7}}$ **-**$\mathbf{P_{4,5}}$ **-** $\mathbf{P_{8,10}}$	**<u>62</u>**	**<u>257</u>**	**<u>319</u>**
8	2	$P_{2,6,9}$ -$P_{1,3,11,7}$ -$P_{4,5,8,10}$	41	282	323
9	1	$P_{2,6,9}$ -$P_{1,3,11,7,4,5,8,10}$	28	315	343
10	1	$P_{2,6,9,1,3,11,7,4,5,8,10}$	0	346	346

Entries in bold and underline show the minimum cost solution of the problem nut shell, it is concluded that low values of the dendrogam level gives high reconfigurations cost and lower under-utilization cost. And the contrary occurs for the large value of dendrogram level.

5 Conclusion

This work has presented a novel methodology for grouping parts into families on the basis of operation similarity which is a central issue in the design of reconfigurable manufacturing systems. In the first phase, a correlation analysis model is formulated to group the parts and Correlation matrix is used as the similarity coefficient matrix. Finally, Principle Component Analysis and Agglomerative Hierarchical K-means algorithm is applied to find the level of similarity in parts. The obtained part families are based on compactness of the family formation on the basis of operational similarity. Furthermore it uses PCA, which is available in many commercial software packages. However the goodness of the grouping of part is to be tested as the future scope. The second phase consists in a Mixed Integer Linear Programming (MILP) model that selects the families to consider and their processing sequence. This is done using a minimum cost criterion that consider both the costs of reconfiguration between two consecutive families (adding/removing needed/unneeded machines and modules) and cost due to the under-utilization of the assigned resources. The major difficulty is the value of both costs which are unknown, and accuracy estimation is required. This has been solved dividing the costs into four parameters, which are easier to estimate. Finally, the model has been validated through the results obtained with an example.

The selection of families among several possibilities is a highly complex problem (NP complete), and therefore calculations for resolution of the model grow exponentially together with the number of parts. Therefore, it is advisable to solve approximately with metaheuristic. The proposed work can be further extended to accommodate other factors such as, capacity of machines and intra family part sequence (i.e. minimum make span).

References

1. Koren, Y., Jovane, F., Heisel, U., Moriwaki, T., Pritschow, G., Ulsoy, A.G., VanBrussel, H.: Reconfigurable Manufacturing Systems. CIRP Ann. **48**(2), 527–540 (1999)
2. Xiaobo, Z., Jiancai, W., Zhenbi, L.: A stochastic model of a reconfigurable manufacturing system Part 1: a framework. Int. J. Prod. Res. **38**(10), 2273–2285 (2000)
3. Lokesh,.K., Jain, P.K.: A model and optimization approach for reconfigurable manufacturing system configuration design. Int. J. Prod. Res. **50**(12), (2011).
4. Abdi, M.R., Labib, A.W.: Grouping and selecting products: the design key of Reconfigurable Manufacturing Systems (RMSs). Int. J. Prod. Res. **42**(3), 521–546 (2004)
5. Galan, R., Racero, J., Eguia, I., Garcia, J.M.: A systematic approach for product families formation in Reconfigurable Manufacturing Systems. Robot. Comput. Integr. Manuf. **23**, 489–502 (2007)
6. Vakharia, A., Wemmerlov, U.: A comparative investigation of hierarchical clustering techniques and dissimilarity measures applied to the cell formation problem. J. Oper. Manag. **13**, 117–138 (1995)
7. Yin, Y., Yasuda, K.: Similarity coefficient methods applied to cell fprmation problem: a taxonomy and review. Int. J. Prod. Econ. **101**, 329–352 (2006)

8. Papaioannou, G., Wilson, J.M.: The evolution of cell formation problem methodologies based on recent studies (1997–2008): review and directions for future research. Eur. J. Oper. Res. **206**, 509–521 (2010)
9. Jolliffe, I.T.: Principal Component Analysis. Springer, New York (1986)
10. Preisendorfer, R.: Principal Component Analysis in Meteorology and Oceanography. Elsevier Science (1988).
11. Pearson, K.: On lines and planes of closest fit to systems of points in space. Philos. Mag. **2**, 559–572 (1901)
12. Tuncer, Y., Tanik, M.M., Alison, D.B.: An overview of statistical decomposition techniques applied to complex systems. Comput. Stat. Data. Anal. **52**(5), 2292–2310 (2008)
13. Horenko, I., Dittmer, E., Schütte, C.: Reduced Stochastic Models for Complex Molecular Systems. Comput. Vis. Sci. **9**(2), 89–102 (2006)
14. Rothenberger, M.A., Dooley, K.J., Kulkarni, U.R., Nada, N.: Strategies for software reuse: a principal component analysis of reuse practices. IEEE Trans. Softw. Eng. **29**(9), 825–837 (2003)
15. Barbieri, P., Adami, G., Piselli, S., Gemiti, F., Reisenhofe, E.: A three-way principal factor analysis for assessing the time variability of freshwaters related to a municipal water supply. Chemom. Intell. Lab. Syst. **62**(1), 89–100 (2002)
16. Lokesh, K., Jain, P.K.: Concurrent part machine group formation with important production data. Int. J. Simul. Model. **9**(1), 5–16 (2010)
17. Shehroz, S.K., Ahmad, A.: Cluster center initialization algorithm for K-means clustering. Pattern Recogn. Lett. **25**, 1293–1302 (2004)
18. Chaea, S.S., Wardeb, W.D.: Effect of using principal coordinates and principal components on retrieval of clusters. Comput. Stat. Data Anal. **50**(6), 1407–1417 (2005)
19. Gnanadesikan, R.: Methods for Statistical Data Analysis of Multivariate Observations. Wiley-Interscience, New York (1997)
20. Rummel, R.J.: Applied Factor Analysis. Northwestern University Press, Evanston, USA (1988)
21. Lawler, E., Lenstra, J., Kan, A., Shmoys, D.: The Traveling Salesman Problem. Wiley, New York (1992)