Formation of machine groups and part families in cellular manufacturing systems using a correlation analysis approach

Wafik Hachicha and Faouzi Masmoudi and Mohamed Haddar

Ecole Nationale d'ingénieurs de Sfax

19. October 2006

Formation of machine groups and part families in cellular manufacturing systems using a correlation analysis approach

Wafik Hachicha[a] , Faouzi Masmoudi [a,b] and Mohamed Haddar [a,b]

a *Unité de recherche de Mécanique, Modélisation et Production,(U2MP).*
b *Département de génie mécanique, Ecole Nationale d'ingénieurs de Sfax, B.P. W, 3038 Sfax, Tunisia*

Abstract

The important step in the design of a cellular manufacturing (CM) system is to identify the part families and machine groups and consequently to form manufacturing cells. The scope of this article is to formulate a multivariate approach based on a correlation analysis for solving cell formation problem. The proposed approach is carried out in three phases. In the first phase, the correlation matrix is used as similarity coefficient matrix. In the second phase, 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 simultaneously machine groups and part families while maximizing correlation between elements. In the third stage, an algorithm is improved to assign exceptional machines and exceptional parts using respectively angle measure and Euclidian distance.

The proposed approach is also applied to the general Group Technology (GT) problem in which exceptional machines and part are considered. Furthermore, the proposed approach has the flexibility to consider the number of cells as a dependent or independent variable.

Two numerical examples for the design of cell structures are provided in order to illustrate the three phases of proposed approach. The results of a comparative study based on multiple performance criteria show that the present approach is very effective, efficient and practical.

Keywords: cellular manufacturing, cell formation, correlation matrix, Principal Component Analysis, exceptional machines and parts

---

1. Introduction

Increasing global competition has made many business leaders and policy makers turn their attention to such critical issues as productivity, quality and reducing manufacturing costs. Consequently, there have been major shifts in the design of manufacturing systems using innovative concepts. The adoption of cellular manufacturing (CM) has consistently formed a central element of many of these efforts and has received considerable interest from both practitioners and academicians.

The cellular manufacturing system which is based on the concept of Group Technology (GT) philosophy aims at increasing productivity and production efficiency by reducing throughput times (Burgess, Morgan & Vollmann, 1993). The two fundamentals problems associated with CM are: part-family formation and machine-cell formation. Part-family formation is to group parts with similar geometric characteristics or processing requirements to take advantage of their similarities for the design or manufacturing purpose. Machine-cell formation is to bring dissimilar machines together and dedicate them to the manufacture of one or more part families.

The problem of cell formation (CF) is to identify the part families and machine groups by rearranging the initial incidence matrix in a block diagonal form, with a minimum number of parts traveling between cells. Extensive work has been performed in the area of CF problem and numerous methods have been developed. A number of researches have published review studies for existing CF literature (refer to Sinha & Hollier, 1984; Joines, King & Culbreth, 1996; Shambu, 1996; Yin & Kazuhiko, 2005). The main used techniques are classification and coding systems, machine-component group analysis, mathematical and heuristic approaches, similarity coefficient based on clustering methods, graph-theoretic methods, knowledge-based and pattern recognition methods, fuzzy clustering methods, evolutionary approaches and neural network approaches.

Furthermore, the CF research in the literature can be divided into three categories, according to the formation logic used (indicated by J. Geonwook, 1998; Wang, 2003 and others):

(a) Grouping part families (e. g. in Kusiak, 1987) or machine cells only (e. g. in Rajamani et al, 1990)

(b) Forming part families and then machine cells (e. g. in Choobineh, 1988; Adenso-Diaz et al, 2005),

(c) Forming part families and machine cells simultaneously (e. g. in Adil et al, 1993).

Part family grouping procedures are used for identifying groups of parts that are similar to one another. Some approaches focus attention on grouping machine cells only but these procedures often assume that part families already have been formed. Part-machine grouping procedures are for identifying part families and machine groups sequentially or simultaneously.

The proposed methodology falls into the third category (i.e., forming part families and machines groups simultaneously). This approach consists in solving machine-part grouping problem, identifying exceptional machines and parts and solving CF problem mode by assigning theses exceptional elements. To this effect, an original technique is used proposing to use correlation as a new definition of similarity coefficient and to use the Principal Component Analysis (PCA) as a cluster method. These techniques allow the identification of part families and machine groups simultaneously and the identification of exceptional machines and exceptional parts.

Most of the existing CF methods suffer from one or more drawbacks. Their major common drawbacks include: the inflexibility in determining the number of cells (i.e. in some methods, the number of cells is a dependent variable, while in others it has to be identified in advance), and the limited industrial application due to the unavailability of software programs supporting them. New cell formation approaches that overcome these limitations are clearly needed. In response to this need, this paper proposes a new approach based on new similarity coefficient methods for the manufacturing cell formation. The proposed approach attempts to: prove its feasibility and validity to solve a cell formation problem, perform very well in terms of a number of well known criteria, have the flexibility in allowing the user either to identify the required number of cells in advance, or consider it as a dependent variable and be supported by available commercial software programs in order to facilitate industrial applications.

The outline of the paper is as follows: Section 2 describes methods manufacturing CF problem which use similarity coefficients approach, CF problem with exceptional machines and parts and performance criteria. Then the proposed approach is presented in Section 3. Afterwards, two numerical examples give presentation for each proposed methodology phase is shown in Section 4. Section 5 uses well-know CF problem from the literature and dedicated for illustrations results. Lastly, conclusion is made in Section 6.

2. Preliminaries

2.1. Similarity coefficients methods (SCM)

Different similarity coefficients have been proposed by researchers in different fields. A similarity coefficient represents the degree of commonality between two parts or two machines. SCM-based methods rely on similarity measures in conjunction with clustering algorithms. It usually follows a prescribed set of steps:

Step 1:

The initial machine-part incidence matrix is a binary matrix whose rows are machines and columns stand for parts. Where $\boldsymbol{a_{ij}} = 1$, means that machine $\boldsymbol{i}$ (1….m) is necessary to process part $\boldsymbol{j}$ (1….p) and $\boldsymbol{a_{ij}} = 0$, otherwise.

Select a similarity coefficient and compute similarity values between machine (part) pairs and build a similarity matrix. An element in the matrix represents the sameness between two machines (parts). In this step a large number of research papers have used different types of similarity (Jaccard, Kulezynki, etc.) and dissimilarity coefficients (Hamming, Euclidean, Average Euclidean etc.) for determining part families or machines groups.
McAuley, 1972 suggested the use of Jaccard's similarity coefficient in the formation of GT cells. He defined a similarity coefficient between any two machines as the ratio of the number of parts that visit both machines to the number or parts that visit either or both machines.
Kusiak, 1987 sought to maximize the sum of similarity coefficients defined between pairs of parts using a linear integer programming model and defined the similarity coefficient between two parts as:

$S_{ij}^{p} = \sum_{k=1}^{m} \delta(a_{ki}, a_{kj})$, i≠j, j = 1, 2, …., p and $S_{ii}^{p} = 0$, where δ is the Kronecker product function. (1)

A commonality score was proposed by Wei & Kern (1989). It has been introduced to overcome the shortcomings of the Jaccard Similarity coefficients. The definition is presented as follows:
The commonality score between machine $\boldsymbol{i}$ and $\boldsymbol{j}$ is

$$S_{ij}^{p} = \sum_{k=1}^{m} \Gamma(a_{ki}, a_{kj}) \qquad (2)$$

where

$$\Gamma(a_{ik}, a_{jk}) = \begin{cases} p-1 \ \text{si} \ a_{ik} = a_{jk} = 1 \\ 1 \quad \text{si} \ a_{ik} = a_{jk} = 0 \\ 0 \qquad \text{si} \ a_{ik} \neq a_{jk} \end{cases} \qquad (3)$$

p is the part number and ***k*** is the ***k***th part in the initial incidence matrix. The commonality score not only recognizes the parts that require two machines for processing, but also the parts that do not require both machines. This is the advantage of commonality score (Yasuda & Yin, 2001). Many other definitions of similarity coefficient have been proposed for GT. For example in Gupta & Saifoddini, 1990; Genwook et al, 1998; Kitaoka et al, 1999.

Step 2:

Select a similarity coefficient and compute similarity values between machine (part) pairs and construct a similarity matrix. An element in the matrix represents the sameness between two machines (parts).

Step 3:

A clustering algorithm must transform the initial machine-part incidence matrix into the final matrix with structured form (blocks in diagonal). Use a clustering algorithm to process the values in the similarity matrix, which results in a diagram called a tree, or dendrogram, that shows the hierarchy of similarities among all pairs of machines (parts). Find the machine groups (part families) from the tree or dendrogram, check all predefined constraints such as the number of cells, cell size, etc.

Cell formation can be considered as a dimension reduction problem. In fact, a large number of interrelated machines need to be grouped into smaller set of independent cells. Few research have used multivariate analysis tool in cell formation problem. Kitaoka et al, 1999 proposed a double centering machine matrix for similarity of machines and parts as a similarity coefficient matrix. A quantification method is applied to find the eigenvalues and eigenvectors. Albadawi et al, 2005 suggested the use of Jaccard's similarity coefficient and proposed multivariate analysis by applying Principal Component Analysis for machine cells only.

2.2. The CF Problem with Exceptional Machines and Parts (PEMP)

Cell formation solutions often contain Exceptional Elements (EE). EE create interactions between two manufacturing cells. In most formation, there are usually exceptional parts and exceptional machines. An exceptional part can be viewed as parts that require processing on machines in two or more cells. An exceptional machine processes parts from two or more part families. These movements cause lack of segregation among the cells. This is in conflict with the main objective of GT which aims at independently operating cells. To effectively implement GT, a good clustering algorithm is needed such that the number of exceptional parts and exceptional machines are minimized (Cheng et al, 2001).

2.3. Performance criteria

The purpose of this section is to present objective criteria to evaluate the quality of clustering method. There are three principal criteria widely used in the literature:
The first is called the Percentage of Exceptional elements (PE) and defined as the ratio of the number of exceptional elements to the number of unity elements in the incidence matrix:

$$PE = \frac{EE}{UE} \times 100 \qquad (4)$$

Where UE denotes the number of unity elements in the incidence (i.e. total number of operations in the data matrix).

The second criterion is called Grouping Efficiency (GE) and defined by Chandrasekharan & Rajagopalan, 1986 as follows:

$$GE = \alpha \cdot \frac{UE - EE}{\sum_{k=1}^{Q} m_k p_k} + (1-\alpha)\left(1 - \frac{EE}{m \cdot p - \sum_{k=1}^{Q} m_k p_k}\right) \qquad (5)$$

where $\alpha \in [0,1]$ is a weighting parameter. A value of $\alpha = 0.5$ is commonly used. m_k and p_k denote, respectively, the number of machines in cell k and number of parts in family k. Q is the number of cells. m is the total number of machines and p is the total number of parts.
Note that when $\alpha = 1$, the Grouping Efficiency coincides with the third criterion: Machine Utilization (MU) which is defined as the frequency of visits to machines within cells.

3. Description of the proposed approach

The proposed approach consists in three phases as mentioned in figure 1.

3.1. Similarity coefficient matrix

The first phase consists in building a similarity matrix. The initial machine-part incidence matrix shown in (Eq. 6) is a binary matrix which rows are parts and columns stand for machines. Note that this proposed definition looks like the transpose of classical incidence matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & . & . & . & a_{1m} \\ a_{21} & a_{22} & & & & a_{2m} \\ . & & & & & \\ . & & & & & \\ . & & & & & \\ a_{p1} & a_{p2} & & & & a_{pm} \end{pmatrix} \quad (6)$$

Where $a_{ij} = 1$ if machine j is required to process part i and $a_{ij} = 0$ otherwise.

M_j is a binary row vector from the matrix A: $M_j^A [a_{1j}, a_{2j}, ..., a_{pj}]$

To make the initial matrix (A) more sufficiently meaningful and significant, its standardization is needed. Several methods of standardization are found in the literature (Schaffer & Green, 1996; Chaea & Wardeb, 2005 and others). In this article, the general standardization of the initial data set is used. It is expressed by:

$$M_j^B = \frac{M_j^A - E_j}{\sigma_j} \quad (7)$$

Where E_j is the average of the row vector M_j

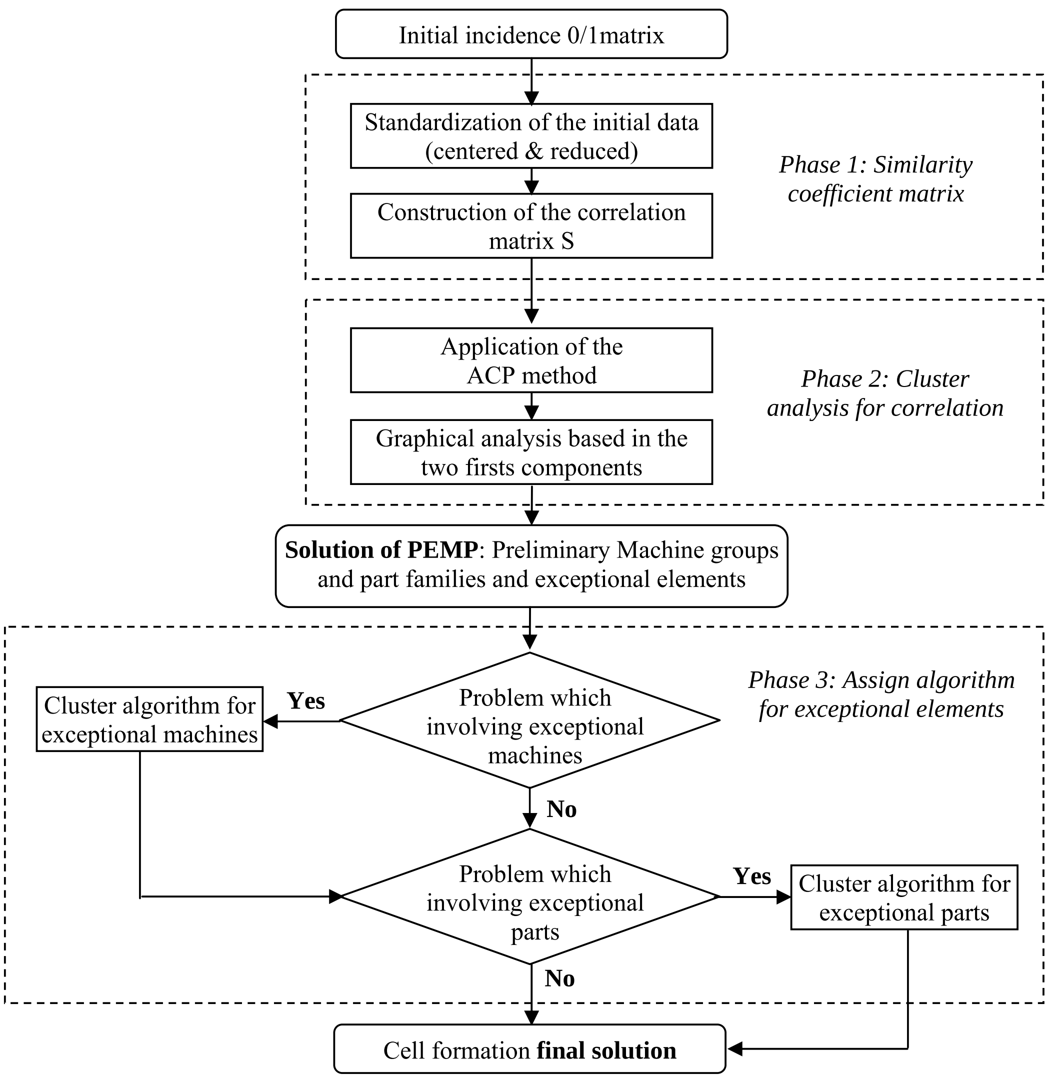


Fig. 1. Architecture of the proposed approach

$$E_j = \frac{\sum_{k=1}^{p} a_{kj}}{p} \tag{8}$$

And $$\sigma_j^2 = \frac{1}{p}\sum_{k=1}^{p}(a_{kj}\text{-}E_j)^2 \tag{9}$$

Applying Huyghens- Köning Theorem yields to $\sigma_j^2 = E_j - E_j^2$ (10)

The proposed similarity coefficient is based on the simple correlation matrix the incidence matrix. The correlation matrix S is defined as follows: (Gnanadesikan, 1997)

$$S = \frac{1}{p}B^{'}B \tag{11}$$

S_{ij} is m x m matrix which elements are given by: $S_{ii} = 1$ and $S_{ij} = \frac{1}{p}\sum_{k=1}^{p} b_{ik} b_{jk}$ (12)

3.2 Cluster analysis for correlation

In the second phase of the proposed approach, the machine groups and part families are identified by factor and graphical analysis. The objective is to find machine groups, part families and parts common machines using some classification scheme given by using Factor analysis representation of the data.

Factor analysis is a powerful multivariate analysis tool used to analyze the interrelationships among a large number of variables to reduce them into a smaller set of independent variables called factors. Factor analysis was developed in 1904 by Spearman in a study of human ability using mathematical models (Rummel, 1988). Since then, most of the applications of factor analysis have been used in the psychological field. Recently, its applications have expanded to other fields such as mineralogy, economics, agriculture and engineering. Factor analysis requires having data in form of correlations, and uses different methods for extracting a small number of factors from a sample correlation matrix. These methods include: common factor analysis, principal component analysis, image factor analysis, and canonical factor analysis. Detailed description of PCA method can be found in the relevant literature such as in (Labordere, 1977); (Rummel, 1988); (Harkat, 2003) (Gnanadesikan, 1997). and others.

PCA is the most widely used method. It is an investigation of the data that is largely widespread among users in many areas of science and industry. It is one of the most common methods used by data analysts to provide a condensed description. PCA 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 in 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. Thus, the study of principal components can be considered as putting into statistical terms the usual developments of eigenvalues and eigenvectors for positive semi-definite matrices.

The eigenvector equation where the terms $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_m$ are the real, nonnegative roots of the determinant polynomial of degree P given as:

$$\det(S - \lambda_i I) = 0 \;\; ; \;\; i \in <1, m> \tag{13}$$

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 explained PC of the variance in the data.

Where $$PC = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^{m}\lambda_k} = \frac{\lambda_1 + \lambda_2}{m} \tag{14}$$

Really, the number of principal components must be determined by using a specific technique as cross validation, Kaiser's criterion, reconstruction method etc. (Harkat, 2003 ; Ledauphin et al, 2004). In this application of ACP, the objective is clustering machines in group and parts in families. A binary decision is applied at each machine and part. Two principal components are enough to analyse correlation between elements (machines and parts).

As example, let us consider the graphical contain two machines and five parts. The data can be represented by a two dimensional scatter plot (figure 2) where each machine is represented by a line from the origin and each part is represented by a dot located at its weight in each line (machine).
Graphical clustering analysis is based on an angle distance measure. An angle distance measure θ, or normalized scalar product is used. It defined as:

$$\theta = \arccos\left(\frac{\sum_i x_i y_i}{\sum_i x_i^2 \sum_i y_i^2}\right) \qquad (15)$$

Where, x_i and y_i are the coordinates of P_i in the scatter plot.

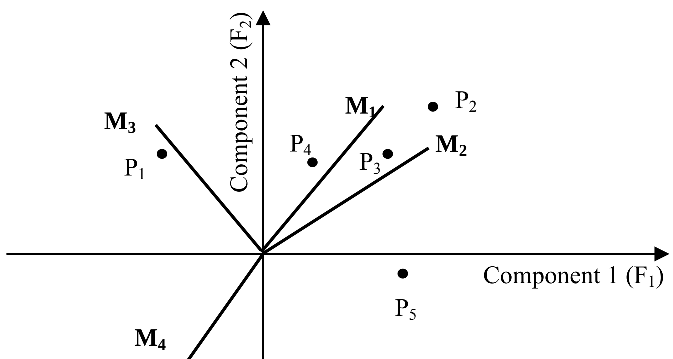


Fig. 2. Illustration of the scatter plot

Four principal situations for the classification of machines can be recovered:

- Two neighbor machines which have a low angle distance measure, consequently they belong to the same cell. Examples can be illustrated in the figure 2 by (M_1 and M_2).
- Two machines whose angle distance measure between them is almost 180°, this means that they are negatively correlated and mustn't belong to the same cell. Examples can be illustrated in figure 2 by (M_1 and M_4) and (M_2 and M_4).
- Two machines whose angle distance measure between them is almost 90°. This means that they are independent similarly and then they don't belong to the same cell. Examples can be illustrated in figure 2 by (M_1 and M_3) and (M_3 and M_4).
- If none of these three cases above is verified, the machine which called an exceptional machine is affected in the next phase of the proposed approach.

The same method is used for the classification of parts: when a part is close to a line (machine), it is assigned to the cell which component this machine. Examples can be illustrated in the figure 2 by (P_1 and M_3) and (P_2, P_3, P_4 and M_1 M_2). Otherwise, it is an exceptional part which can be illustrated, for example, in the figure 2 by P_5. Exceptional parts are affected in the next phase of the proposed approach.

3.3 Assign algorithm for exceptional elements

The objective of the third phase is to assign exceptional parts and exceptional machines to preliminary cells (end of the second phase of the proposed approach in figure 1).
In most CF problem, there are usually exceptional parts and exceptional machines. For each type of exceptional element, an assign algorithm is proposed.

1. For the assignment of exceptional machines

This iteration continues until all exceptional machines are assigned to form machine groups.
Let e_m is the number of exceptional machines.

For k = 1 to e_m do

Step 1: Compute angle measure for each machine (different to M_k and not an exceptional machines)

$$\theta_{ik} = \min(|\theta_i - \theta_k|, 2\pi - |\theta_i - \theta_k|) \tag{16}$$

Where θ_i is the angle measure between M_i and the first principal component, $\theta_i \in]-\pi,\pi]$.

Step 2: Since the objective is to group machines with minimum angle distance, Machine M_i which have the smallest angle distance with M_k, is assigned to the machine groups M_i

End.

2. For the assignment of exceptional parts

An exceptional part can be viewed as a part that requires processing on machines in two or more cells. Let e_p is the number of exceptional part. The clustering algorithm for exceptional part is shown below:
This iteration continues until all exceptional parts are assigned to form part families.

For k = 1 to e_p do

Step 1: Compute Euclidean distance for each part (different to P_k and not an exceptional part). with the exceptional part P_k

$$d(P_k, P_i) = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{17}$$

Where x_i and y_i are the coordinates of P_i in the scatter plot (two principal components).

Step 2: Since the objective is to group parts with minimum distance, part P_i which have the smallest distance, is assigned to the part families P_i

End.

4. Numerical examples

4.1 Problem 1

In order to explain the methodology of the proposed approach, a manufacturing system is considered with seven machines (labeled M1-M7) and eleven parts (labeled P1-P11). This example is provided by Boctor (1991). The machine-part matrix A is shown in Eq. (18).

A = (18)

	M1	M2	M3	M4	M5	M6	M7
P1	1	1	0	0	0	0	0
P2	0	1	1	0	0	0	0
P3	1	0	0	0	1	1	0
P4	0	0	0	1	0	1	0
P5	0	0	0	1	0	0	1
P6	0	1	1	0	0	0	0
P7	1	0	0	0	1	0	0
P8	0	0	0	0	0	0	1
P9	0	0	1	0	0	0	0
P10	0	0	0	1	0	0	1
P11	1	0	0	0	0	1	0

Applying Eq. (8, 10 and 7) to the initial machine-part given in Eq. (18) yield the standardization matrix B given in Eq. (19).

For example, for machine M1

$$E_1 = \frac{4}{11} = 0.364$$

$$\sigma_1 = \sqrt{0.36-(0.36)^2} = 0.481$$

The member coefficient between part 1 and machine, b_{11} is calculated as follows

$$b_{11} = \frac{1-0.364}{0.481} = 1.32$$

$$b_{21} = \frac{0-0.364}{0.481} = -0.76$$

The same procedure be applied for the others elements of matrix B.

$$B = \begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \\ P6 \\ P7 \\ P8 \\ P9 \\ P10 \\ P11 \end{array} \begin{array}{ccccccc} M1 & M2 & M3 & M4 & M5 & M6 & M7 \\ \hline 1{,}32 & 1{,}63 & -0{,}61 & -0{,}61 & -0{,}47 & -0{,}61 & -0{,}61 \\ -0{,}76 & 1{,}63 & 1{,}63 & -0{,}61 & -0{,}47 & -0{,}61 & -0{,}61 \\ 1{,}32 & -0{,}61 & -0{,}61 & -0{,}61 & 2{,}12 & 1{,}63 & -0{,}61 \\ -0{,}76 & -0{,}61 & -0{,}61 & 1{,}63 & -0{,}47 & 1{,}63 & -0{,}61 \\ -0{,}76 & -0{,}61 & -0{,}61 & 1{,}63 & -0{,}47 & -0{,}61 & 1{,}63 \\ -0{,}76 & 1{,}63 & 1{,}63 & -0{,}61 & -0{,}47 & -0{,}61 & -0{,}61 \\ 1{,}32 & -0{,}61 & -0{,}61 & -0{,}61 & 2{,}12 & -0{,}61 & -0{,}61 \\ -0{,}76 & -0{,}61 & -0{,}61 & -0{,}61 & -0{,}47 & -0{,}61 & 1{,}63 \\ -0{,}76 & -0{,}61 & 1{,}63 & -0{,}61 & -0{,}47 & -0{,}61 & -0{,}61 \\ -0{,}76 & -0{,}61 & -0{,}61 & 1{,}63 & -0{,}47 & -0{,}61 & 1{,}63 \\ 1{,}32 & -0{,}61 & -0{,}61 & -0{,}61 & -0{,}47 & 1{,}63 & -0{,}61 \end{array} \quad (19)$$

The similarity matrix S is show in Eq. 20

$$S = \begin{array}{c} \\ M1 \\ M2 \\ M3 \\ M4 \\ M5 \\ M6 \\ M7 \end{array} \begin{array}{ccccccc} M1 & M2 & M3 & M4 & M5 & M6 & M7 \\ \hline 1{,}00 & & & & & & \\ -0{,}04 & 1{,}00 & & & & & \\ -0{,}46 & 0{,}54 & 1{,}00 & & & & \\ -0{,}46 & -0{,}38 & -0{,}38 & 1{,}00 & & & \\ 0{,}62 & -0{,}29 & -0{,}29 & -0{,}29 & 1{,}00 & & \\ 0{,}39 & -0{,}38 & -0{,}38 & 0{,}08 & 0{,}24 & 1{,}00 & \\ -0{,}46 & -0{,}38 & -0{,}38 & 0{,}54 & -0{,}29 & -0{,}38 & 1{,}00 \end{array} \quad (20)$$

The second phase consists in applying a cluster analysis for correlation which is based on PCA method. The computed eigenvalues for the matrix given in E.q 19 are listed and sorted in a descending order in table 1.

Table 1
Eigenvalues and associated percentage of variance

Components	**Eigenvalues**	**% of total variance**	**Cumulative %**
1	2.53	36.11	36.11
2	2.35	33.54	69.65
3	0.92	13.17	82.83
4	0.61	8.77	91.59
5	0.38	5.46	97.05
6	0.16	2.27	99.32
7	0.05	0.68	100.00
Sum total	7.00 (must be equal to m)	100	

The use of graphical analysis is based on a two dimensional scatter plot where each machine is represented by a line from the origin and each part is represented by a dot .The scatter plot indicates the relationship between machine and other machine, between machine and part and between part and other part. There should be high correlation among machines strongly associated with the same cell, and low correlation among machines that are associated with different cells.

Four principal situations for the classification of machines can be recovered:

- Two neighbor machines which have a low angle distance measure. Consequently they belong to the same cell. Examples can be illustrated in the figure 3 by (M_4 and M_7) and (M_1 and M_5).
- Two machines which the angle distance measure between them is almost 180°. This means that they are negatively correlated and not must be belong to the same cell.
- Two machines which the angle distance measure between them is almost 90°. This means that they independent, then they don't also belong to the same cell. Examples can be illustrated in figure 3 by (M_1 and M_3) and (M_4 and M_6).
- If no one of the three cases above is verified, the machine isn't affected to any cell. This means that it is an exceptional machine.

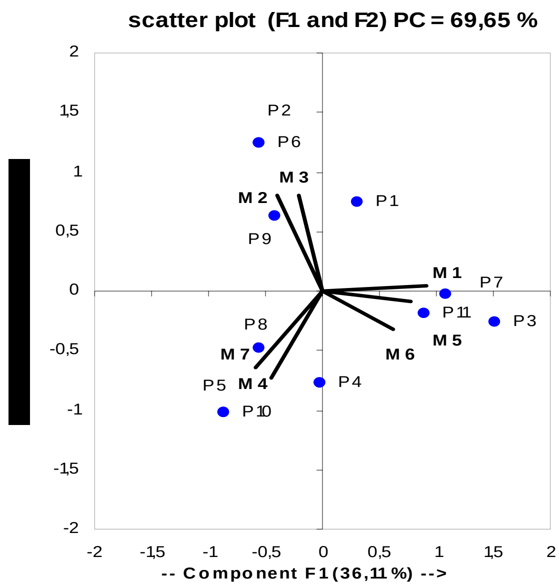


Fig. 3. Graphical illustration of the scatter plot

Applying the second phase of the proposed approach to these data sets yields the result shown in figure 3. We obtained the following results:

- The best grouping for the seven machines is to group them into tree cells: cell 1 consists of machines 2 and 3; cell 2 consists of machines 1, 5 and 6; while cell 3 consists of machines 4 and 7.

	M2	M3	M1	M5	M6	M4	M7
P2	1	1					
P6	1	1					
P9	0	1					
P3			1	1	1		
P7			1	1	0		
P11			1	0	1		
P5						1	1
P8						0	1
P10						1	1
P1	1		1				
P4					1	1	

(21)

- No exceptional Machines is identified
- Parts 1 and 4 were identified as exceptional parts.

The preliminary cell design is shown in Eq 21.

To complete the cell formation, the parts 1 and 4 need to be allocated to the machines cells by applying the cluster algorithm which it is the purpose of the third stage.

Table 2
Coordinates of each part in the scatter plot

Parts	Fist component : x_i	Second component : y_i	Assigned cell
P1	0.62	1.28	Exceptional part
P2	-1.11	2.34	Cell 1
P3	3.04	-0.51	Cell 2
P4	-0.03	-1.49	Exceptional part
P5	-1.72	-1.97	Cell 3
P6	-1.11	2.40	Cell 1
P7	2.17	-0.05	Cell 2
P8	-1.09	-0.91	Cell 3
P9	-0.82	1.21	Cell 1
P10	-1.72	-1.97	Cell 3
P11	1.78	-0.37	Cell 2

For each exceptional part, the problem consists to find a part k that minimizes the Euclidian distance between the exceptional parts.

$d(P_1, P_9) = \text{Min} \{ d(P_1, P_k); k = 1, 2, 3, \ldots..9; k \neq 1 \text{ and } 4\}$

$d(P_4, P_8) = \text{Min} \{ d(P_4, P_k); k = 1, 2, 3, \ldots..9; k \neq 1 \text{ and } 4\}$

Therefore P_1 and P_9 are in the same cell 1. P_4 and P_8 are in the same cell 3.
The results of final cell formation problem are shown in Eq. 22 which confirms that parts 1 and 4 are exceptional parts.

	M2	M3	M1	M5	M6	M4	M7
P1	1	0	1				
P2	1	1					
P6	1	1					
P9	0	1					
P3			1	1	1		
P7			1	1	0		
P11			1	0	1		
P4					1	1	0
P5						1	1
P8						0	1
P10						1	1

(22)

4.2 Numerical example involving exceptional machines and parts (11 x 22)

The problem is represented by a matrix which shown in Eq. 23. This problem is adopted to demonstrate the effectiveness of the proposed approach. It is a frequently cited in the literature GT problem which involving exceptional machines and exceptional parts (Chan & Milner, 1982; Cheng & all, 2001). Figure 4 shows the result of applying approach.
Applying the proposed approach to these data sets yields the result shown in figure 4. The following results are obtained:
-The best grouping for the seven machines is to group them into tree cells: cell 1 consists of machines 1, 4 and 5, cell 2 consists of machines 2, 3 and 6, while cell 3 consists of machines 7, 9 and 11.

- Machines 8 and 10 were identified as exceptional machines. Besides exceptional machine, the graphical analysis identifies easily exceptional parts 6 and 11. These elements were removed from the initial matrix. The removal of exceptional machines and exceptional parts allows the matrix presented in Eq. 24 to decompose.

Eq. 23

The part-machine incidence matrix for (11 x 22) problem

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
P1	1			1	1					1	
P2	1			1	1					1	
P3	1			1	1					1	
P4							1		1		1
P5		1	1			1		1		1	
P6			1				1	1			
P7					1					1	
P8		1				1				1	
P9							1	1	1		1
P10								1			1
P11	1	1		1						1	
P12		1	1			1		1		1	
P13			1					1			
P14							1	1	1		
P15	1			1	1					1	
P16	1			1	1					1	
P17							1		1		1
P18							1	1	1	1	1
P19		1	1			1		1			
P20	1			1	1			1			
P21	1			1	1						
P22	1			1	1						

Fig. 4

Graphical illustration of the scatter plot for (11x22) problem

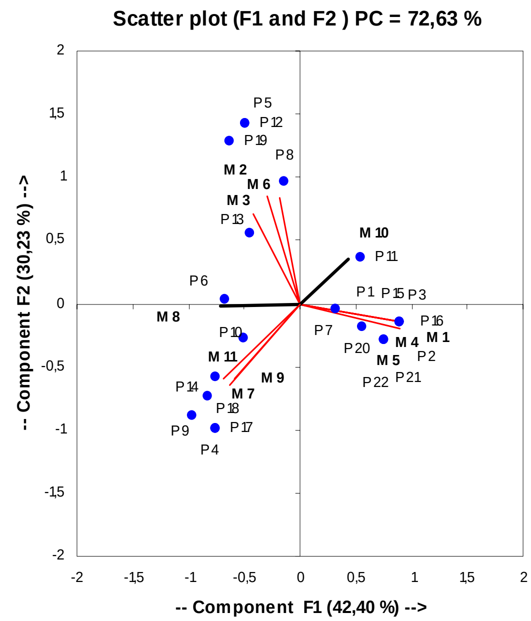


Table 3

Angle measure for each machine compared with the first principal component

Machines	$Cos^2\ \theta$	$Cos\ \theta$	$\theta\ ]-\pi, \pi]$	Assigned cell (phase 2)
M1	0.868	0.93	-0.37	Cell 1
M2	0.033	0.18	1.75	Cell 2
M3	0.175	0.42	2.00	Cell 2
M4	0.862	0.93	-0.38	Cell 1
M5	0.807	0.90	-0.45	Cell 1
M6	0.084	0.29	1.87	Cell 2
M7	0.580	0.76	-2.43	Cell 3
M8	0.995	1.00	-3.07	Exceptional Machine
M9	0.590	0.77	-2.45	Cell 3
M10	0.620	0.79	0.66	Exceptional Machine
M11	0.615	0.78	-2.47	Cell 3

$\theta\ (M_8, M_7) = \text{Min}\ \{\theta(M_8, M_k);\ k = 1, 2, 3\ldots, 11;\ k \neq 8 \text{ and } 10\}$

$\theta\ (M_{10}, M_1) = \text{Min}\ \{\theta(M_{10}, M_k);\ k = 1, 2, 3\ldots, 11;\ k \neq 8 \text{ and } 10\}$

Therefore M_8 and M_7 are in the same cell 3.

And M_{10} and M_1 are in the same cell 1.

Chan & Milner (1982) and Cheng C. H. & all (2001) give the same result of general GT problem which involving exceptional parts and exceptional machines as presented in Eq. 24. However, the proposed approach beyond this result.

The third phase allow final cell formation problem to do the best assign of those exceptional machines and exceptional parts. Applying this additional phase to the solution under consideration yields the final cell design in Eq. 25 with 3 cells.

Eq. 24 Solution of PEPM for (11 x 22) problem

	M1	M4	M5	M2	M3	M6	M7	M9	M11	M8	M10
P1	1	1	1								1
P2	1	1	1								1
P3	1	1	1								1
P15	1	1	1								1
P16	1	1	1								1
P20	1	1	1							1	
P21	1	1	1								
P22	1	1	1								
P7	0	0	1								1
P5				1	1	1				1	1
P8				1	0	1					1
P12				1	1	1				1	1
P19				1	1	1				1	
P13				0	1	0				1	
P4							1	1	1		
P9							1	1	1	1	
P10							0	0	1	1	
P14							1	1	0	1	
P17							1	1	1		
P18							1	1	1	1	1
P11	1	1	0	1							1
P6					1		1	0	0	1	

Eq. 25 Final Solution with 3 cells for (11 x 22) problem

	M1	M4	M5	M10	M2	M3	M6	M7	M9	M11	M8
P1	1	1	1	1							
P2	1	1	1	1							
P3	1	1	1	1							
P15	1	1	1	1							
P16	1	1	1	1							
P20	1	1	1	0							1
P21	1	1	1	0							
P22	1	1	1	0							
P7	0	0	1	1							
P11	1	1	0	1	1						
P5				1	1	1	1				1
P8				1	1	0	1				
P12				1	1	1	1				1
P19					1	1	1				1
P13					0	1	0				1
P4								1	1	1	0
P9								1	1	1	1
P10								0	0	1	1
P14								1	1	0	1
P17								1	1	1	0
P18				1				1	1	1	1
P6						1		1	0	0	1

5. Computational results

Based on six well-know incidence matrices published in the literature, the performance of the proposed approach has been evaluated using the multiple performance criteria discussed in Section 2.3. Table 4 summarizes the special features and the sources of these data sets.

Table 4
Test cell formation problems

No.	Size (machines x parts)	N cells	References
1	5 x 7	2	Waghodekar & Sahu, 1984
2	7 x 11 (example 1)	3	Boctor,1991
3	8 x 20	3	Chandrasekharan & Rajagopalan, 1986
4	11 x 22 (example 2)	3	*Chan & Milner, 1982
5	14 x 24	4	King, 1980
6	16 x 43	5	*King & Nakornchai,1982 (introduced by Burbidge en 1973)

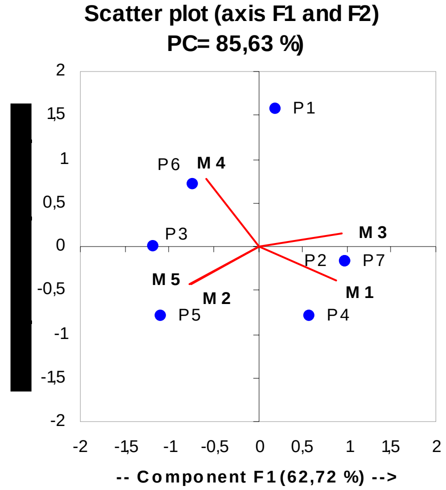
*Literature problem involving exceptional machines and exceptional parts (PEMP)

Applying the proposed approach to these data sets yields to solutions shown in Eq 26 – 29

Problem 1 (5 x 7)

The result of applying the proposed approach to the problem 1 shown in figure 5, Eq. 26 and Eq. 27

Fig. 5 Graphical illustration of the scatter plot for (5x7) problem	Eq. 26 Solution of PEPM for (5 x 7) problem



P$_1$ and P$_4$ are exceptional parts

	M2	M4	M5	M3	M1
P3	1	1	1		
P5	1	0	1		
P6	0	1	1		
P7				1	1
P2				1	1
P4			1	1	1
P1		1		1	0

Eq. 27 Solution with 2 cells for (5 x 7) problem

	M2	M4	M5	M3	M1
P3	1	1	1		
P5	1	0	1		
P6	0	1	1		
P1		1		1	0
P2				1	1
P4			1	1	1
P7				1	1

Problem 6 (16 x 43)

This example consist is a practical engineering situation. A 16-machines and 43-part problem (King & Nakornchai, 1982) is considered.

Fig 7.a Graphical illustration in scatter plot without exceptional machines for (16 x 43) problem

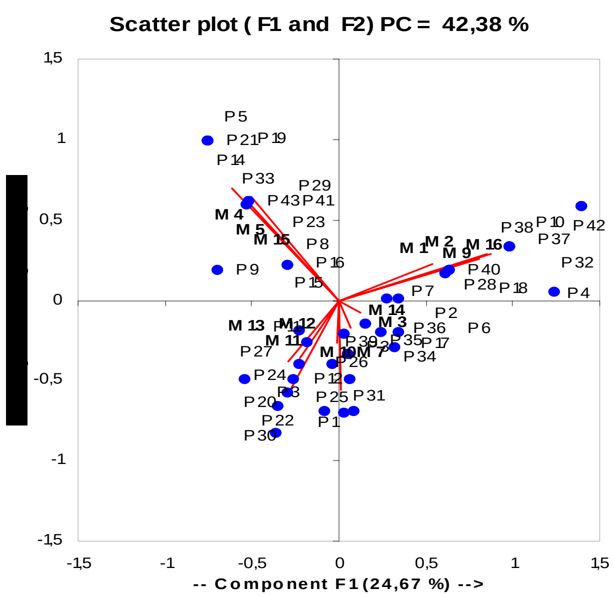
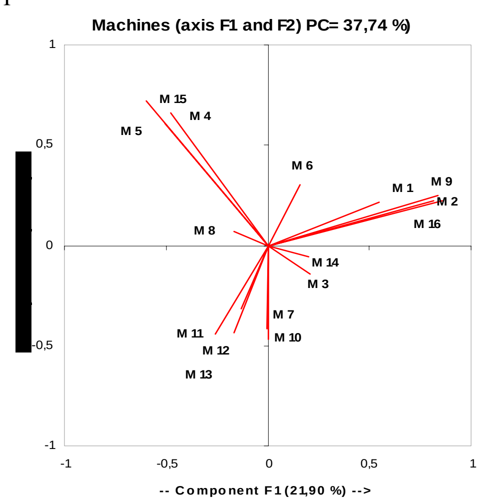


Fig 7.b
Graphical illustration of machines in scatter plot



The result of applying the proposed approach to the problem 6 is shown in figure 7.a, figure 7.b, Eq. 28 and Eq. 29. Parts 2, 7 and 9 are exceptional parts and machine 6 and 8 are exceptional machines. The solution is shown in table 13. This matrix is identical to that given by Chan & Milner (1982) and by Cheng C. H. & all (2001). Chan & Milner (1982) require the user to visually examine the immediate matrix and to manually identify exceptional machines and exceptional parts. Cheng

C. H. & all (2001) proposes heuristic branching rules which mathematical algorithm must be programmed to identity exceptional machines and exceptional parts. The proposed procedure approach is completely based on graphical analysis that requires only a factor analysis available on most statistical packages including XLSTAT, S-PLUS, SPSS and authors.

Eq. 28 Solution PEMP

	M1	M2	M9	M16	M3	M14	M11	M12	M13	M10	M7	M4	M5	M15	M6	M8
P4	0	0	1	0											0	0
P10	0	1	1	1											0	0
P18	0	0	1	1											0	0
P28	0	1	1	0											0	1
P32	0	1	1	1											1	0
P37	1	1	1	1											1	1
P38	0	1	1	1											0	1
P40	0	1	1	0											1	0
P42	1	1	1	1											1	0
P34					1	0									1	0
P35					1	1									0	0
P36					1	0									0	0
P6					0	1									1	0
P17					1	1									1	0
P3							1	0	1						0	1
P11							0	1	0						0	1
P20							1	0	0						0	1
P22							0	1	0						0	0
P24							1	1	1						0	1
P27							1	1	0						0	1
P30							1	1	0						0	0
P13										1	1				1	0
P25										1	1				0	0
P1										1	1				1	1
P12										1	0				1	1
P26										1	0				0	0
P39										1	0				1	0
P31										1	0				0	1
P5												1	1	1	0	0
P14												1	1	1	1	0
P19												1	1	1	1	1
P21												1	1	1	0	1
P23												1	1	0	1	1
P29												1	1	0	0	0
P43												0	1	1	1	1
P41												0	1	1	0	1
P8												0	1	0	1	1
P15												0	0	0	1	1
P16												0	1	0	0	0
P33												0	1	1	1	0
P2	0	1	1	1		1									1	1
P7	0	0	0	1	1										1	0
P9							1					1	1	0	0	1

Eq .29 Solution with 5 cells

	M1	M2	M9	M16	M4	M5	M15	M6	M8	M11	M12	M13	M3	M14	M10	M7
P2	0	1	1	1				1	1					1		
P4	0	0	1	0												
P10	0	1	1	1												
P18	0	0	1	1												
P28	0	1	1	0					1							
P32	0	1	1	1				1								
P37	1	1	1	1				1	1							
P38	0	1	1	1					1							
P40	0	1	1	0				1								
P42	1	1	1	1				1								
P5					1	1	1	0	0							
P14					1	1	1	1	0							
P19					1	1	1	1	1							
P21					1	1	1	0	1							
P23					1	1	0	1	1							
P9					1	1	0	0	1	1						
P29					1	1	0	0	0							
P43					0	1	1	1	1							
P41					0	1	1	0	1							
P8					0	1	0	1	1							
P15					0	0	0	1	1							
P16					0	1	0	0	0							
P31					0	0	0	0	1						1	
P33					0	1	1	1	0							
P3									1	1	0	1				
P11									1	0	1	0				
P20									1	1	0	0				
P22										0	1	0				
P24									1	1	1	1				
P27									1	1	1	0				
P30										1	1	0				
P34								1					1	0		
P35													1	1		
P36													1	0		
P6								1					0	1		
P7				1				1					1	0		
P17								1					1	1		
P13								1							1	1
P25															1	1
P1								1	1						1	1
P12								1	1						1	0
P26															1	0

Table 5
Summary of proposed approach results and the best-known results

No.	Size	N cell	Proposed approach results			Best-known results			
			PE	MU	GE	PE	MU	GE	Recent Reference
1	5 x 7	2	12.50	82.35	85.61	12.5	82.35	85.61	Wang, 2003
2	7 x 11	3	09.52	76.00	86.10	09.52	76.00	86.10	Stawowy, 2006
3	8 x 20	3	14.80	100.00	95.80	14.80	100.00	95.80	Albadawi & all, 2005
4	11 x 22	3	14.10	80.72	86.90				Cheng C. H. & all, 2001
5	14 x 24	4	03.28	68.60	83.90	03.30	68.60	83.60	Wang, 2003
6	16 x 43	5	22.22	63.22	79.00	23.01	64.23	79.40	Chan, F.T.S & all, 2005

Chan & Milner (1982) and Cheng C. H. & all (2001) give the same result of general GT problem which involving exceptional parts and exceptional machines which presented in Eq. 28. However, the proposed approach beyond this result. The third phase allow final cell formation problem to do the best assign of those exceptional machines and exceptional parts. Applying this additional phase to the solution under consideration yields the final cell design in Eq. 29 with 5 cells.

Table 5 summarizes the results of comparative study and the best-known results in the recent literature. Basically, the results of CF problem are the same as those found in recent literature such as Cheng C. H. & all, 2001; Wang, 2003; Albadawi & all, 2005. These recent researches were compared with former methods like rank order clustering (ROC, King 1980), direct clustering algorithm (DCA, Chan & Milner 1982), bond energy algorithm (BEA, McCormick & all 1972), single linkage clustering (SLC, McAuley 1972), GRAFICS method (Srinivasan & all 1990) and others. These recent researches demonstrated to be better in comparative studies. Therefore, it could be said that the proposed approach is valid. It is more flexible and able to get correlation information between each machine and part.

6. Final conclusion

In this paper, a new approach is presented for part-family and machine-cell formation. The main aim of this article is to formulate a correlation analysis model to generate optimal machine cells and part families in GT problems. The correlation matrix for similarity machine and part is used as similarity coefficient matrix. Principal Components Analysis (PCA) method is applied to find the optimal machine and part elements. Exceptional machines and parts are easily assigned to cells using cluster algorithm.

The proposed method is a logical and systematic approach to the design of cellular manufacturing systems which makes it easily portable into practice, is that it uses PCA, which are available in many commercial software packages and it can be performed on most statistical packages including SAS (1985), SPAD (1995), SPSS (1999), S-PLUS, XLSTAT, and others.

Computational experiences show that the proposed approach does not require long computing times and gives the same solution than which proposed in recent literature. Although the present approach focuses on the compactness of formation solution only, it can readily accommodate other manufacturing information such as production volume, sequence and alternative routings. Extending the proposed approach to this direction is our interesting research perspective.

References

Adenso-Diaz, B., Lozano, S., Eguira, I., (2005). Part-machine grouping using weighted similarity coefficients. *Computers & Industrial Engineering*, 48, 553-570.

Adil, G. K., Rajamani, D., Strong, D., (1993). A mathematical model for cell formation considering investment and operational cost. *European Journal of Operational Research*, 69, 330-341.

Albadawi, Z., Bashir, H. A., Chen, M., (2005). A mathematical approach for the formation of manufacturing cell, *Computers & Industrial Engineering*, 48, 3-21.

Arvindh, B., Irani, S., A., (1991). Principal components analysis for evaluating the feasibility of cellular manufacturing without initial-part matrix clustering. *International Journal of Production Research.* 32, 1909-1938.

Burgess, A. G., Morgan, I. & Vollmann, T. E., (1993). Cellular manufacturing: its impact on the total factory. *International Journal of Production Research*, 31, 2059-2077.

Boctor F. F., (1991). A linear formulation of the machine-part cell formation problem. *International Journal of Production Research*, 29, 343-356.

Chaea, S. S., Wardeb, W., D, (2005). Effect of using principal coordinates and principal components on retrieval of clusters. *Computational Statistics & Data Analysis.*

Chan, H., & Milner, D., (1982). Direct clustering algorithm for group formation in cellular manufacturing. *Journal of Manufacturing Systems*, 1(1), 65–67.

Chan, F.T.S, Lau, K.W., Chan, P.L.Y, Choy, K.L,(2005). Two-stage approach for machine-part grouping and cell layout problems, *Robotics and Computer-Integrated Manufacturing*.

Chandrasekharan, M. P., & Rajagopalan, R. (1986). An ideal seed non-hierarchical clustering algorithm for cellular manufacturing. *International Journal of Production Research*, 24, 451–464.

Cheng, C. H., Goh, C. H., Lee, A., (2001). Designing group technology manufacturing systems using heuristic branching rules. *Computers & Industrial engineering*, 40, 117-131.

Choobineh, F., (1988). A framework for the design of cellular manufacturing systems. *International Journal of Production Research*, 26, 1161-1172.

Delagarge, J., (2000). « Initiation à l'analyse des données » ; 3 ième édition, Ed Dunod.

Genwook, J., Leep, H. R., Parsaaei, H. R., (1998). A cellular manufacturing system based on new similarity coefficient which considers alternative routes during machine failure. *Computers & Industrial engineering*, 34, 21-36.

Gnanadesikan, R., (1997). Methods for Statistical Data Analysis of Multivariate Observations. Wiley, New York, pp. 7–15.

Gupta, T., Saifoddini, H., (1990). Production data based similarity coefficient for machine-component grouping decisions in the design of a cellular manufacturing system. *International Journal of Production Research*, 28, 1247-1269.

Harkat, M. F., (2003). Détection et Localisation de Défauts par Analyse en Composantes Principales. PhD Thesis Institut National Polytechnique de Lorraine.

Joines, J. A., King, R., E., Culbreth, C., T., (1996). A comprehensive Review of Production-Oriented Manufacturing Cell Formation Techniques. Research funded by the NCSU Furniture Manufacturing and Management Center, North Carolina State University, DDm-92-15432.

Kern, M., Wei, C., (1991). The cost of elimining exceptional elements in group technologie cell formation, *International Journal of Production Research*, 29, 1535-1547.

King, J. R., (1980). Machine-component grouping in production flow analysis: an approach using a rank order clustering algorithm. *International Journal of Production Research*, 18(2), 213–232.

King J. R., Nakornchai V., (1982). Machine-component group formation in group technology: review and extension, *International Journal of Production Research*, 20, 117–133.

Kitaoka, M., Nakamura, R., Seriza, S., Usuki, J. (1999). Multivariate analysis model for machine-part cell formation problem in group technology, *International Journal Production Economics*, 60-61, 433-438.

Kumar, K. R., Chandrasekharan, M. P., (1990). Grouping efficacy: a quantitative criterion for goodness of block diagonal forms of binary matrices in group technology. *International Journal of Production Research*, 28(2), 233–43.

Kusiak, A., (1987). The generalized group technology concept. *International Journal of Production Research*, 25, 561-569.

Labordere, A. H., (1977). Analyse de données : application et méthodes pratiques. Ed Masson.

Ledauphin, S., Hanafi, M., Qannari, E., (2004). Simplification and signification of principal components. Chemometrics and Intelligent Laboratory Systems, 74, 277– 281.

McAuley, D., (1972). Machine grouping for efficient production. *The Production Engineer*, 51, 53–57.

Panchalavarapua P. R. , Chankongb, V., (2005). Design of cellular manufacturing systems with assembly considerations. *Computers & Industrial Engineering*, 48, 449–469

Rajamani, R., Singh, R., Aneja, Y. P., (1990). Integrated design of cellular manufacturing systems in the presence of alternative process plans. *International Journal of Production Research*, 28, 1541-1554.

Rummel, R. J. (1988). Applied factor analysis. Evanston: Northwestern University Press.

SAS Institute, Inc. (1985). SAS user's guide-statistics 5. Cary, NC: SAS Institute.

Slomp, J., V.C, Boppana,. N. C, Suresh (2005). Design of virtual manufacturing cells: a mathematical programming approach. *Robotics and Computer-Integrated Manufacturing,* 21, 273–288

Seifoddine, H., Djassemi, M,. (1995). Merits of the production volume based similarity coefficient in machine cell formation. *Journal of Manufacturing Systems*, 14, 35-44.

Shambu, G., (1996). Performance evaluation of cellular manufacturing systems: a taxonomy and review of research. *International Journal of Operations & Production Management*, 8, 81-103.

Sinha, R. K., Hollier, R., H,. (1984). A review of production control problems in cellular manufacture, *International Journal of Production Research*, 22, 773-789.

Stawowy, A., (2006) Evolutionary strategy for manufacturing cell design. *The international Journal of Management Science: Omega*, 34, 1 – 18.

SPSS, Inc. (1999). SPSS professional statistics. Chicago, IL: SPSS, Inc.

Waghodekar, P. H., & Sahu, S. (1984). Machine-component cell formation in group technology: MACE. *International Journal of Production Research*, 22, 937–948.

Wang, J., (2003). Formation of machine cells and part families in cellular manufacturing systems using a linear assignment algorithm. *Automatica*, 39, 1607-1615.

Yasuda, K., Yin, Y., (2001). A dissimilarity measure for solving the cell formation problem in cellular manufacturing. *Computers & Industrial Engineering*, 39, 1–17.

Yin, Y., Kazuhiko, K., (2006). Similarity coefficient methods applied to the cell formation problem: A taxonomy and review. *International Journal of Production Economics*. 101, 329–352.