Simultaneous Hierarchical Clustering for Cell Formation Problems

Houman Mehrabadi

A Thesis in

The Concordia Institute for Information System Engineering

Presented in Partial Fulfillment of the Requirements of the Degree of

Master of Applied Science (Quality System Engineering) at

Concordia University, Montreal, Quebec, Canada

December 2011

Concordia University

School of graduate studies

This is to certify that the thesis prepared

By: Houman Mehrabadi

Entitled: Simultaneous Hierarchical Clustering for Cell Formation Problems

and submitted in partial fulfillment of requirements for the degree of

Master of Applied Science (Quality System Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by final examining committee:

Dr. C. Wang Chairman

Dr. A. Schiffauerova Examiner

Dr. O. Kuzgunkaya Examiner

Dr. S. Li Supervisor

Approved by ______________________

Chair of the Department or Graduate Program Director

December 2011 ______________________

Dean of Faculty

Abstract

Simultaneous Hierarchical Clustering for Cell Formation Problems

Houman Mehrabadi

A cellular manufacturing system consists of several work cells in which parts are processed under machines. Identification of parts and machines in each work cell, known as cell formation, is a major step in design of a cellular manufacturing system. This thesis presents a method for cell formation. The method uses clustering approach to identify work cells in three steps. First, production information represented in an incidence matrix is evaluated for finding the coupling relationships between parts and machines. In second step the gathered coupling information in the first step is used to reorder the incidence matrix rows and columns, and create a tree diagram. Using the tree diagram, work cells and parts and machines in each work cell are identified in the third step. Performance of the method is evaluated by solving two types of cell formation problem. The results indicate that the method can produce solutions as good as other methodologies. In comparison to clustering cell formation methodologies, the method has a flexible solution procedure that simultaneously groups parts and machines, and it does not need predetermined production information for executing cell formation.

Acknowledgments

I would like to thank my supervisor, Dr. Simon Li, for his guidance and lessons he taught me during my research. Also, I would like to thank my family and friends for their support and encouragement.

Contents

List of figures

List of tables

Chapter 1

1. Introduction

This chapter is an introduction to this thesis providing the background of cell formation problem, cell formation methodologies, the motivation for the research, and the outline of the thesis.

1.1 Background

Products or parts are made with different manufacturing systems depending on product variety and production order quantity (Hitomi, 1996). Examples for manufacturing systems are job shop, batch production, continuous production, cellular manufacturing, and virtual cellular manufacturing. Each of these manufacturing systems has its own advantage. For instance, job shop is suitable when a relatively high variety of products with few order quantities has to be made. With increase in order quantity and variety of the products, better productivity can be achieved with batch or continuous production systems. Cellular manufacturing is another option that provides higher productivity and flexibility when relatively high variety of products with considerable order quantity have to be made (Black and Hunter, 2003; Groover, 2008; Sule, 2009).

1.2. Cellular manufacturing

Cellular manufacturing is a manufacturing system based on group technology – a manufacturing approach in which similar parts are grouped together in order to achieve higher productivity (Seifoddini and Wolf, 1986). In cellular manufacturing a group of parts and machines is called a work cell. Inside each work cell, operations are performed on parts. A manufacturing plant may have a number of work cells.

1.3. Cell formation problem

Cell formation is a major step in the cellular manufacturing planning and implementation (Wemmerlöv and Hyer, 2002). Cell formation is a process that provides information about work cells. Given production information about parts and machines as input, the output of cell formation are the number of work cells and the parts and machines inside each work cell.

In order to achieve higher productivity from cellular manufacturing, certain design goals such as process completeness and resource usage should be established in cell formation problem. In addition, the cell formation problem is controlled by managerial and technical constraints. Examples are the number of possible new machines, separation of hazardous processes from other activities, work cell size, and number of operators / equipment assignable to a work cell (Wemmerlöv and Hyer, 2002). Cell formation design goals establish the criteria for the quality of cell formation output.

A design goal, process completeness, is explained as an example. Process completeness is the degree that a part is processed within its assigned work cell. A perfect process completeness for a part means that it is processed and completed in a single work cell. In practice, this may not be possible. If a part needs a process that cannot be performed by any of machines in the work cell, it has to visit another work cell where capable machine is located. This causes a movement out of work cell and called inter-cell movement, and the part is called exceptional part. Inter-cell movements may also cause a number of voids in the visiting cell. If the exceptional part is not completed in the second work cell, it will come back to its primary work cell (another movement). In figure 1.1 two solutions are represented by incidence matrix format, a perfect solution (a) and a solution with exceptional parts and voids (b).

Incidence matrix is an important visual presentation of parts and machines in cell formation. It is a two-dimensional matrix, and its rows and columns represent machines and parts. The elements of an incidence matrix equal either zero or one. If a part visits a machine, the corresponding element of incidence matrix equals one, otherwise it equals zero (zeros are not shown in figure 1.1). In other words, incidence matrix shows the required machines for each part. For example in figure 1.1 (b) machines 3, 5, and 6 required for producing part 6.

For fulfilling the process completeness, one of objectives in cell formation problem is to get a final solution with minimum number of inter-cell movements (exceptional parts) and the voids. In literature, this is established as a major objective in cell formation problem (Papaioannou and Wilson, 2010).

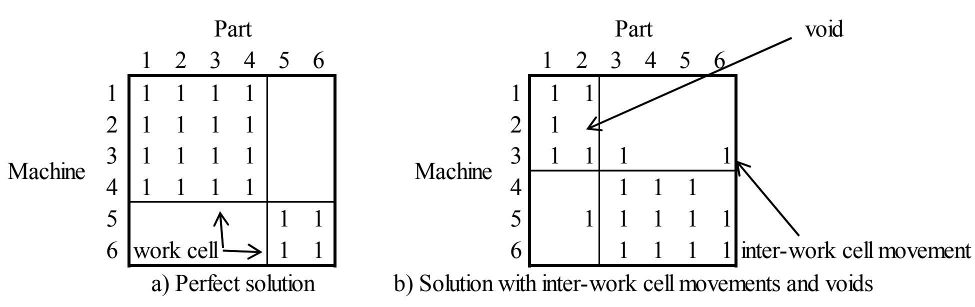

Figure 1.1 Cell formation problem solutions

1.4 Cell formation methodologies

During last three decades many research has been done on cell formation and many solution procedures have been developed based on different methodologies for cell formation. During this period, cell formation problem has been evolved from the basic machine-part problem to more complicated problems considering multiple objectives. At first, the focus of researchers was on basic machine-part cell formation problem which only considers the machines a part visits regardless of other production information. With progress in cell formation methodologies, the cell formation problem has been getting more practical with inclusion of more production information. Examples are considering the information about the order of machines a part visits (production sequence) or different ways of making a part using existing equipment (alternative routing). Based on production information, Different types of cell formation problem can be identified from the literature. Figure 1.2 shows four types of the cell formation problem. The first type of problem is machine-part using information of incidence matrix. In other three types, in addition to incidence matrix, extra production information such as production sequence, alternative routing, or both are considered in each type of cell formation problem.

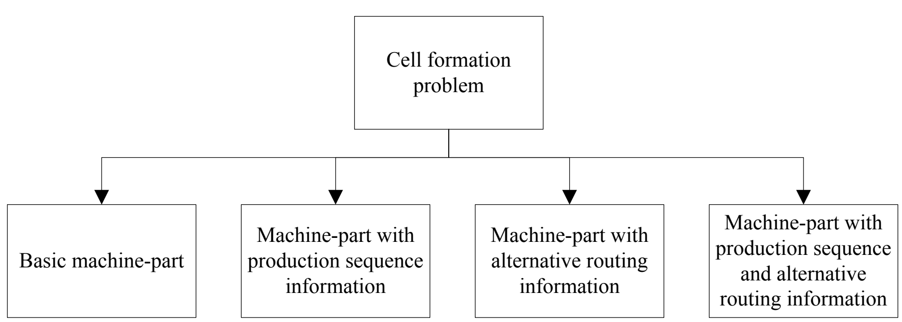

Figure 1.2 cell formation problems

1.5 Motivation

Selim et al. (1998) criticized the mathematical programming and similarity coefficient based methods for lack of "all objective/constraints" and not being useful for real world application, and proposed a comprehensive model to solve the cell formation with several objectives. The cell formation problem is a NP-hard problem (King and Nakornchai, 1982), and the mathematical programming computation time for getting a global optimum solution is not short. In addition, inclusion of many production factors and constraints adds to the complexity of model. Researchers have tried to reduce the number of variable and the constraints (Papaioannou and Wilson, 2011). For these reasons researchers started to apply optimization approaches for solving cell formation problem. Optimization methods can provide a sub-optimal solution in relatively reasonable amount of time while considering multiple objectives.

Through the development of similarity coefficient based methods during last two decades, newer coupling coefficients are considering more production factors and become closer to real manufacturing environment. Even though, optimization approaches have get more attention from academics in recent years, similarity based methods have couple of advantages over optimization methods. Comparing to mathematical programming / optimization methodologies, similarity based clustering methods are less complex but more flexible (Cheng et al. 1998; Yin and Yasuda, 2006). In addition to complexity, the solution procedure of mathematical programming/ optimization methods needs some input information such as number of work cells, size of work cells, and machines and parts in each work cells in order to perform the cell formation. Consequently, in final solution, the work cell configuration does not come from natural grouping of parts and machines and it is influenced by input information.

Despite of above advantages, similarity based clustering methodologies are criticized for some drawbacks. First, the grouping of machines and parts cannot be done simultaneously by clustering methods. Second, although similarity based clustering methods group the non-zero elements of incidence matrix, there is no guarantee to have a block diagonal form in final solution. A solution with block diagonal form has most of non-zero elements agglomerated adjacent to the diagonal line of final solution matrix. Such a format facilitates the identification of work cells as well as possible re-configuration of work cells if the solution is not satisfactory. Third, there is a need for pre-defined input such as number and size of the work cell into solution procedure. Fourth, hierarchical clustering methods has irreversibility problem. In other words, the steps toward grouping of any two machines (or parts) are unknown and non-traceable (Yin and Yasuda, 2006; Nair and Narendran 1998).

The objective of this research is improving the similarity based clustering approach for cell formation. A method based on clustering is proposed for cell formation in this research. The method for grouping machines and parts is based on a hierarchical clustering. The proposed method overcomes the above limitations of similarity based clustering methods by following improvements:

- Grouping of parts and machines are done simultaneously,
- There is no need for pre-defined information as input into solution procedure,
- Improve the flexibility of the solution procedure for incorporating new production factors.

1.6 Thesis structure

This thesis has six chapters. Chapter 1 provides the introduction and motivation for this research. A brief review of selected cell formation methodologies are presented in chapter 2. In chapter 3 the structure and details of proposed methodology is explained. The solution procedure of new method is demonstrated in chapters 4 and 5 by solving two different types of cell formation problem. Finally, thesis ends by a discussion on strengths of new methodology, contributions of this research, and the path work for future research.

Chapter 2

2. Literature review

The machine part cell formation problem is not a new research topic and many methods and solution procedures have been developed for it during the last three decades. In this section, after an introduction, three cell formation problem methodologies and their advantages/shortcomings are reviewed briefly. In end the three methods are compared to each other by three criteria.

2.1 Introduction

Group technology was the starting approach for solving the cell formation problem. Group technology is an approach in which similar parts are grouped together in order to increase manufacturing productivity by reducing waste and improve product quality (Singh, 1993). The origin of using group technology in manufacturing goes back to beginning of twentieth century, however it was not called group technology, and the research findings were not well-presented and organized until Mitrofanov defined the group technology concept (Gallagher and knight, 1973). The first confined application of group technology in manufacturing consisted of two methods, part classification and coding and production flow analysis. Part classification was developed by Mitrofanov, and production flow analysis by Burbidge (Irani et al., 1999). Burbidge (1989) outlined seven disadvantages of the classification and coding system in identification of group technology groups: the classification cannot capture the similarities (or differences) in Materials, size, tolerances, and order size. Also, Classification only groups the parts (not machines), and it is time consuming. Seifoddini (1989) redefined the two different methods that apply group technology to manufacturing, production flow analysis and clustering.

With more research on cell formation problem, the number of methodologies has increased and there is a need for a classification of cell formation methodologies

Several literature reviews and taxonomies have explored the different methodologies for cell formation problem. Due to extent of methodology development at the time of creation of the taxonomies and the purpose/ focus of each researcher, classifications differ in one or other way. In the first attempt for classification, three categories were identified, “informal methods, part coding analysis methods, and production based methods” (Papaioannou and Wilson, 2011). Wemmerlöv and Hyer (1986) developed a framework which classified the existing method into four categories based on the way part family/ machine group is identified. Singh (1993) reviewed seven classes of cell formation methodologies. Offodile et al. (1994) classification with focus on “assumptions, characteristics, properties and results” was only concentrated on production based methods. Singh and Rajamani (1996) described six methods of cell formation in detail. In Selim et al. (1998) classification, there are five general cell formation methodologies. Balakrishnan and Cheng (2007) reviewed the cell formation methodologies that have considered uncertainties in demand and resources. Papaioannou and Wilson (2011) reviewed the methodologies that use mathematical formulation in solving cell formation problem. While the number of sub-categories is different, they all follow similar classification system, based on solution procedure.

In this research, three cell formation methodologies are reviewed, machine component group analysis, similarity based clustering, mathematical optimization methods. In the following sections each method is reviewed and the advantages and drawbacks of each method are outlined.

2.2 Machine component group analysis

Machine component group analysis, a production flow analysis based solution methodology, reorders the rows and columns of a binary incidence matrix in order to keep most of the 1's on its diagonal. A highly dense diagonal can be breaks down into groups of the work cells. Many solution procedures had been developed based on this approach. Singh and Rajamani (1996) thoroughly discussed various machine component group analysis solution procedures. Singh (1993) pointed out two major draw backs of this method: not incorporating the real production information and not automatically addressing the bottle neck machines problem. Simplicity of procedure and simultaneous grouping of parts and machine are among the main features of this method (Singh and Rajamani 1996; Seifoddini, 1989).

2.3 Similarity coefficient based clustering methodology

Similarity coefficient based clustering method is another methodology that applies the group technology approach to solve the cell formation problem. Similarity based clustering methods are based on similarity coefficients and clustering algorithms, and follow a set of steps (Romesburg 1984; Vakharia, 1986; Yin and Yasuada, 2006). (1) Form the machine-part incidence matrix. (2) Measure similarity between machines (parts) pairs and construct a similarity matrix. (3) Use a clustering algorithm to process the values of similarity which results in constructing a tree. The tree shows the hierarchy of similarity between machines (parts). Groups of machines (part families) can be identified from the tree.

Similarity between machines (parts) is measured by coupling coefficients. Yin and Yasuda (2006) classified the similarity coefficients into two categories, general-purpose similarity coefficients and problem-oriented similarity coefficients. General purpose similarity coefficient value increases if the two objects are exactly similar to each other. The most common used example for this category is Jaccard similarity coefficient. McAuley (1972) used the Jaccard similarity coefficient for the cell formation problem for the first time.

Problem-oriented similarity coefficients measure the appropriateness of two objects in specific problems such as cell formation problem. In cell formation problem appropriateness of two objects to be in the same work cell is measured. Depending on cell formation problem objectives, value of appropriateness (coupling coefficient) of two objects varies. Gupta (1993) developed a coupling coefficient that considers alternative routing, production sequence, production volume, and operation time in cell formation process. Won and Kim (1997) and Won (2000) modified Jaccard similarity coefficient for incorporating alternative routing information in cell formation problem. Operation sequence information is also considered in many similarity coefficients. (Selvam and Balasubramanian, 1985; Seifoddini, 1988; Gupta and Seifoddini, 1990; Balasubramanian and Panneerselvam, 1993; Nair and Narendran, 1998; Sarker and Xu, 2000; Alhourani and Seifoddini, 2008). The other production factors that are considered in coupling coefficients are the production volume of the parts, tooling requirements, operation times, materials handling cost, and batch sizes. Consideration of more production factors into solution procedure makes cell formation closer to real manufacturing environment.

Comparing to other cell formation methodologies, similarity based clustering has several advantages and shortcomings. Yin and Yasuda (2006) explained two reasons that similarity coefficient clustering methods are more flexible than mathematical programming and optimization methods. First, Similarity coefficient based methods can be extended easily. As a result, more production factors easily can be considered in the method and several alternative solutions are available. Second, Similarity coefficient based methods have three independent stages. In other words, the cell formation problem is decomposed into three simpler steps. This feature of similarity based methods provides a more flexible solution procedure for cell formation problem. Another advantage of this method is final solution features. In practice, use and modification of binary matrices are easier comparing to other methods (Cheng et al. 1998).

The similarity based clustering method has some disadvantages comparing to the other methods. First, the machine groups and parts families can not be identified simultaneously by this method (Liu et al. 2010). Second, some similarity coefficients clustering methods (clustering and data reorganization, array sorting) cannot create a matrix with blocks on it diagonal, and work cell identification is complex (Boe and Cheng, 1991). In addition, some similarity coefficient methods suffer from chaining problem (Singh and Rajamani 1996; Miltunburg and Zhang, 1991). This problem happens when two machines with very high similarity coefficient values are assigned to same work cell (Kattan, 2007). Selim et al. (1998) found that similarity coefficients have often very limited criteria and cannot solve "multi-objective" cell formation problems. Also, similarity coefficient based clustering methods can only use binary incidence matrix (Gupta and Seifoddini, 1990).

2.4 Mathematical optimization methods

The two major methods of mathematical optimization for cell formation are mathematical programming and metaheuristics. In both methods, there is a mathematical model. The mathematical model consists of objective(s), constraints and variables to seek for the optimal cell formation solutions. The advantage of mathematical optimization methods is incorporating numerous numbers of production factors and constraints. In following sections, mathematical programming and metaheuristics are reviewed.

2.4.1 Mathematical programming methods

Kusiak (1987) used p-median model to form part families. Using p-median model for cell formation has three limits. First, this model does not group the machines. As a result extra steps needed to group the machines. Second, quality of solution depends on input information (number of part families) into model. Third, p-median is unable to solve medium size cell formation problem in a reasonable amount of time. The later limit of p-median model is due to cell formation problem. Cell formation problem is NP-hard (King and Nakornchai, 1982). P-median model was modified by several researchers for further improvements such as quality of solution (Deutsch et al., 1998) and computation time (Won, 2000; Won, 2004). Srinivasan et al. (1990) developed an assignment model to improve the quality of the solution and computation time. Rajmani et al. (1996) used a mixed integer programming method to simultaneously group machines and parts with consideration of alternative process plans, processing time, the capacities of machines, and cell size restrictions. Adil et al. (1996) used a non-linear integer programming for simultaneous grouping of machines and parts for better diagonalizations of incidence matrix considering alternative routing.

Selim et al. (1998) criticized the mathematical programming for lacks “all objective/constraints” and not useful for real world application, and proposed a comprehensive model to solve the cell formation with several objectives. The proposed model is hard to solve due to combinatorial complexity. For this reason, researchers have tried to reduce the number of variable and the constraints (Papaioannou and Wilson, 2011). Another drawback of this methodology is that numbers of the work cells, parts, and machines are inputs to the solution procedure. Consequently, no natural identification of the work cells is provided by these methods (Lee and Garcia-diaz, 1993).

2.4.2. Metaheuristics

Due to limitation of mathematical programming methods in solving medium and large cell formation problems, researchers started to use metaheuristics approaches to solve cell formation problems. Metaheuristics have been used to solve NP-hard problems in past two decades (Papaioannou and Wilson, 2010). Metaheuristic methods can reach to sub-optimal solutions in shorter computation time. The need of faster solution procedure has become so important that the computation timing has become a quality measure for cell formation problem methodologies and many researchers have compared their solution procedure with others by the computation time for getting the final solution. Solution procedure of metaheuristics starts with a feasible solution, and solution procedure produces, evaluates and improves possible solutions; this continues till solution procedure finds the best solution (Narendran and Srinivasan, 1999). Heuristics, a similar approach, can get a near to optimum solutions in a short period of time (Papaioannou and Wilson, 2011). Heuristics have been used in other methods such as clustering. However, heuristics methods have some limits. The limits are sensitivity to initial solution and incidence matrix, need for the predefined work cell size, and dependability to user judgment (Singh and

Rajamani, 1996). In metaheuristics the solution procedure in not only based on initial solution, and the user knowledge is managed and controlled to get better solution (Singh and Rajamani, 1996). The major metaheuristics methods for cell formation are genetic algorithms, simulated annealing, and tabu search.

Venugopal and Narendran (1992) used genetic algorithm method for the first time for solving a multi-objective cell formation problem. Gupta et al. (1996) solved cell formation problem considering production sequence information. Gravel et al. (1998) considered multiple routing for parts in solving cell formation problem with genetic algorithm. Onwubolu and Mutingi (2001) used genetic algorithms to solve cell formation problem with consideration of the machine loading. Gonçalves et al. (2004) used a hybrid algorithm approach, combined a local search heuristic with a genetic algorithm. James et al. (2007) also applied a hybrid approach by local search with a standard grouping genetic algorithm.

Souliah (1995) used a general simulated annealing algorithm to design cellular manufacturing layout. Chen et al. (1995) used a simulated annealing based heuristic to solve cell formation problems that is flexible to use non-binary matrix and easy to modify to consider more production information in cell formation problem. Su and Hsu (1998) developed a new method by using simulated annealing with the merits of genetic algorithm (Parallel Simulated Annealing). Transportation cost and loading balance inside and outside of work cells were the main objectives in their model. Arkat et al. (2004) used simulated annealing with consideration of alternative process routing and the production volume of the parts. They compared performance of simulated annealing with genetic algorithm. Wu et al. (2009) proposed a hybrid simulated annealing algorithms that consider multiple process routing for the parts.

Logendran and Ramakrishna (1995) used tabu search to solve a quadratic binary programming models considering machine duplication, multiple visit of a part on a machine, and lot splitting. Aljaber et al. (1997) developed a tabu algorithm to find work cells with the objective of minimum inter-cell movement. In their solution procedure, Machines and parts grouped simultaneously under a shortest spanning path model. Lozano et al. (1999) model feasibility of the work cells were examined by a tabu search algorithm without considering production sequence and alternative process routing information. Lei and Wu (2005) developed a hybrid algorithm based on similarity coefficient and tabu search considering alternative routing. Lei and Wu (2006) proposed a pareto-optimality-based tabu search algorithm that considered several objectives.

One drawback of metaheuristics is their sub-optimized solution that may not be the best solution. The other drawbacks of these methods are the need for input to solution procedure, and the complexity of solution procedure.

2.5 Comparison of cell formation methodologies

The four reviewed methodologies in the previous sections are compared by three criteria in this section: the method of part families and machine grouping identification, the need for production information input, and flexibility. The importance of each criterion in cell formation is discussed in the following paragraphs.

Identification of part families and machine grouping can be done in two different ways, sequential and simultaneous. In sequential way part families and machine groups identification is executed in two steps in one of following options (Wemmerlöv and Hyer, 1986):

- First part families are identified. Then, machines are assigned to part families,
- First machine groups are identified. Then, parts are assigned to machine groups.

In simultaneous way part families and machine groups are identified at the same time. The sequential way has two weaknesses comparing to simultaneous way. First, since there are two possible options to execute the sequential way, two different solutions are available from each option. This may cause confusion for decision making and planning. Second weakness happen when the first step is not done perfect and this may cause the degradation of final quality due to negative influence of first step on second step. Simultaneous identification of part families and machine groups has not these problems.

The solution procedure of some cell formation methodologies does not start to process information if they lack some input production information such as number of the work cells or size of the work cells. This input information prevent natural grouping in the solution procedure. For example, if from the beginning, the number of work cells is set to be three; the solution procedure of the methodology does not provide solutions with four work cells, even if the solution with four work cells has the higher quality than solution with three work cells.

According to Yin and Yasuda (2006), Flexibility of a cell formation can help to easily incorporate additional production factors. Higher flexibility of the cell formation method enables the method to easily solve different types of cell formation problems.

The comparison of cell formation methods by above three criteria is shown in table 3.1. None of the four methods can satisfy all three criteria. Similarity coefficient based clustering, specifically, is weak in two areas. Similarity based clustering methods cannot simultaneously group machines and parts, and there is a need for production input information into its solution procedure. In next chapters it will be shown that how the proposed methods, matrix-based clustering, can overcome these weaknesses of similarity-based clustering methods.

Method	Identify part families and machine groups simultaneously	No need for input information	Flexibility
Machine component analysis	✓	✓	-
Similarity coefficient based clustering	-	-	✓
Mathematical programming	✓	-	-
Metaheuristics	✓	-	-

Table 2.1 comparison of cell formation methodologies

Chapter 3

3. Matrix-based clustering method

In cellular manufacturing parts are completed in work cells. A work cell is a collection of machines and similar parts. Similar parts have some common features in design or manufacturing method. Creation of work cells is a major step of a cellular manufacturing system design and implementation (Wemmerlöv and Hyer, 2002).

In this chapter, it is shown that how matrix-based clustering identifies work cells in cell formation. First, the structure of the solution procedure, the relationship between the steps, and the input/output of the solution procedure are introduced. Next, each step of the solution procedure is explained in detail.

3.1 Work flow

The solution procedure of matrix-based clustering has three steps, coupling analysis, sorting analysis, and partitioning analysis. The order of execution of these steps is shown in figure 3.1. The first step in solution procedure is coupling analysis. It follows by sorting analysis. The third and final step of solution procedure is partitioning analysis. The input into solution procedure is production information. It includes information on required machines for each part and order of machines each part visits. In addition, it may include some more production information such as part order quantity, number of parts per batch.

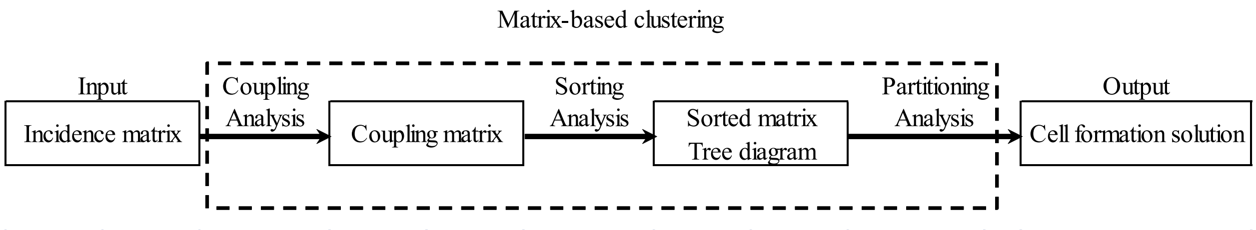

Figure 3.1 Matrix-based clustering solution procedure workflow

One efficient method for organizing input production information is using matrix. Matrix has a visual format and is easy to understand. Production information can be easily collected and represented by a matrix. It also can be used in cell formation computation. In cell formation, the matrix that is used for representing production information is known as incidence matrix.

Incidence matrix is a two-dimension matrix in which rows and columns represent machines and parts. Incidence matrix has two types, binary and non-binary. In a binary incidence matrix, the value of each entry is either one or zero. One means that corresponding part of that entry visits corresponding machine of that entry. If value is equal to zero, it means that corresponding part of that entry does not visit corresponding machine of that entry. In figure 3.2 a binary incidence matrix is shown. In figure 3.2, part 1 visits machine 1, and the corresponding element to part 1 and machine 1 in the incidence matrix is equal to one. Part 1 is not being processed under machine 2 and the corresponding element to machine 2 and part 1 has the value of zero. In other words:

$M = [a_{ij}], \qquad (i = 1, 2, 3, \ldots, m; j = 1, 2, 3, \ldots, n)$

$a_{ij} = 1$, if part i visit machine j

$a_{ij} = 0$, otherwise

		Part				
		1	2	3	4	5
Machine	1	1	0	0	1	0
	2	0	1	0	0	1
	3	0	1	1	0	1
	4	1	0	0	1	1

Figure 3.2 Incidence matrix

In a non-binary matrix, the value of each entry can be zero or non-zero (between 0 and 1). Zero value indicates that the corresponding part to the entry does not visit the corresponding machines of the entry. A non-zero value means that corresponding part of that entry visits corresponding machine of that entry. In addition, the value also carries a weight factor which stand for production information such as the degree to which a part may be processed under a machine (Lee and Wang, 1991) or ratio of time which a part is being processed under a machine (George et al., 2003).

The incidence matrix is the input for matrix-based clustering solution procedure. In the first step-coupling analysis, it is analyzed along other provided production information in order to measure the coupling between parts and machines. The result of coupling analysis is used in the sorting analysis for reordering the incidence matrix rows and columns, and creation of a block diagonal form from incidence matrix. In final step, partitioning analysis, the work cells are identified on block diagonal and output of solution procedure is the solution for cell formation problem. The solution of cell formation includes information on number of work cells and parts and machines in each work cell. In the following sections each step of solution procedure is explained in detail with examples.

3.2 Coupling Analysis

In this thesis, the notion of coupling is defined as the appropriateness of arranging two objects in the same work cell. There are two types of objects: machines and parts. Specifically, two machines are identified as coupled if they need to process some common parts. Similarly, two parts are identified as coupled if they require some common machines. In the context of similarity coefficient methods (Yin and Yasuda, 2006), the degree of coupling can be measured by similarity coefficients (such as Jaccard coefficient). In coupling analysis, the key step is the reasoning process why it is appropriate to put two objects in the same work cell according to available production information. Then, the coupling concept can help us to aggregate different considerations for solving cell formation problems.

Three types of coupling are classified in this thesis: part-part coupling, machine-machine coupling, and machine-part coupling. While the part-part and machine-machine couplings have been briefly discussed in the previous paragraph, the coupling of machine-part is relatively new in this research. In the context of cell formation, a machine and a part are identified as coupled if they are not associated with many other machines and parts, respectively. In the extreme situation, if Part P is only related to Machine M and vice versa, it is said that Part P and Machine M are strongly coupled. This notion of coupling between pairs of different sets is new in addressing cell formation problems, and it is one key to solve cell formation problems in this thesis.

3.2.1 Coupling coefficient

To quantify the degree of coupling for clustering analysis, coupling coefficients need to be formulated for specific applications. In this thesis, two specific applications will be addressed: simple machine-part cell formation (Chapter 4) and cell formation of production sequence information (Chapter 5). The detailed formulations of corresponding coupling coefficients will be provided in these respective chapters. Coupling coefficient value, regardless of type, has two properties:

- The value of coupling coefficient varies between 0 and 1. For any pair of machine and part, the coupling coefficient value of one is indicator of highest appropriateness for the pair to be in the same work cell. A non-zero value less than one shows a less degree of coupling between the pairs and coupling decreases when coupling coefficient value gets closer to zero. Zero value is indicator of no coupling.
- The value of the coupling coefficient is independent from the order of pair. For instance the value of coupling coefficient of pair (part 1, part 2) is equal to the pair (part 2, part 1) (Li, 2010).

3.2.2 Example

For better illustration, a small incidence matrix is used to demonstrate the coupling analysis. In this example the incidence matrix of the figure 3.2 is used as input into solution procedure. From this information, in coupling analysis, firstly, three coupling coefficients part-part, machine-machine , and machine- part are calculated, afterwards all values compiled into one single matrix.

Based on provided production information, incidence matrix, part-part coupling coefficient can be calculated for all pairs of parts. As mentioned before, the only reason two parts can be highly coupled if both visit the same machines. For example, if both parts need turning and both visit lathe, those two parts are highly coupled. Similarly, machine-machine coupling coefficient can be calculated from information from incidence matrix. Two machines are highly coupled if both process common parts. If two machines process completely different parts and have no parts in common, the coupling between two machines would be at least possible value, zero.

In figure 3.2, machines 1 and 4 process common and uncommon parts. Parts 1 and 4 are processed under both machines. In addition machine 4 process part 5. Machines 1 and 4 have two common parts and one uncommon part. Accordingly the machine-machine coupling coefficient for the pair is less than one. Machines 1 and 3 can be compared in same way. There are 5 uncommon parts and no common parts. Machines 1 and 3 have no coupling, and value of coupling coefficient for the pair is equal to zero.

The calculated value for part-part, machine-machine, and machine-part coupling coefficients are shown in figure 3.3. The coupling coefficient equations for three coupling coefficients will be explained in detail in chapter four. There are three matrices in Figure 3.3. Each matrix is for one type of coupling coefficient. Matrix (a) shows the values of part-part coupling coefficients. The values indicate the coupling value for each two parts for being at same work cell. For example, the coupling coefficients for part pairs (1, 4) and (4, 5) are 1 and 0.25 respectively. These values indicate that pair 1 and 4 has higher coupling than pair 4 and 5 for being in same work cell. The machine-machine coupling coefficients in figure (b) shows that machine 1 has the least level of appropriateness for being with machines 2 or 3 in a same work cell. However, it can be grouped with machine 4 since the machine-machine coupling coefficient of pair (1, 4) is higher than machine-machine coupling coefficients of pair (1, 2) or pair (1, 3). In figure (c) appropriateness of pair of a part and a machine is displayed by the values of machine-part coupling coefficient. Machine 1 and part 1 are more appropriate to be together than machine 3 and part 1 as the values of corresponding machine coupling coefficients differ significantly (0.5 versus 0).

(a)

Part	1	2	3	4	5
1	0	0	0	1	0.25
2	0	0	0.5	0	0.67
3	0	0.5	0	0	0.33
4	1	0	0	0	0.25
5	0.25	0.67	0.33	0.25	0

(b)

Machine	1	2	3	4
1	0	0	0	0.67
2	0	0	0.67	0.25
3	0	0.67	0	0.2
4	0.67	0.25	0.2	0

(c)

Machine \ Part	1	2	3	4	5
1	0.5	0	0	0.5	0
2	0	0.5	0	0	0.4
3	0	0.4	0.5	0	0.33
4	0.4	0	0	0.4	0.33

Figure 3.3 coupling coefficients: (a) part-part coupling coefficient, (b) machine-machine coupling coefficient, (c) machine-part coupling coefficient.

3.2.3 Coupling matrix

After quantifying the degree of coupling between machines and parts, the coupling values between any two objects can be recorded on a matrix, namely, coupling matrix (denoted as *CM*). The formulation of a coupling matrix is given below.

$$\text{Coupling matrix= CM} = \begin{bmatrix} w_r.CM_c & w_{rc}.CM_{rc}^{\ T} \\ w_{rc}.CM_{rc} & w_c.CM_r \end{bmatrix} \qquad (1)$$

Where CM_r is matrix of machine-machine coupling coefficients; CM_c is the matrix of part-part coupling coefficients; CM_{rc} is the matrix of machine-part coupling coefficients, and w_r, w_c, w_{rc} are the weights for each matrix. The choice of weights depends on priority of the user. In this research w_r and w_c are equal to one, and w_{rc} is set according to following criterion equation. Effect of w_{rc} value appears in sorting analysis; by increase in w_{rc}, the amount of non-zero entries along the diagonal of sorted matrix will increase (Li, 2009).

$$w_{rc} \geq \frac{\sum CM_r /(o^2 - o) + \sum CM_c (n^2 - n)}{2(\sum CM_{rc} /(o \times n))} \tag{2}$$

For the example in section 3.2.2, coupling matrix can be built using values from matrixes in figure 3.3. The result is shown in figure 3.4. The structure of coupling matrix is divided by lines to show its components, part-part coupling coefficients, machine-machine coupling coefficients, machine-part coping coefficients and its transpose. The value of w_{rc} equal 1.39974.

	Part 1	Part 2	Part 3	Part 4	Part 5	Machine 1	Machine 2	Machine 3	Machine 4
Part 1	0	0	0	1	0.25	0.70	0	0	0.56
Part 2	0	0	0.5	0	0.67	0	0.70	0.56	0
Part 3	0	0.5	0	0	0.33	0	0	0.70	0
Part 4	1	0	0	0	0.25	0.70	0	0	0.56
Part 5	0.25	0.67	0.33	0.25	0	0	0.56	0.47	0.47
Machine 1	0.70	0	0	0.70	0	0	0	0	0.67
Machine 2	0	0.70	0	0	0.56	0	0	0.67	0.25
Machine 3	0	0.56	0.70	0	0.47	0	0.67	0	0.2
Machine 4	0.56	0	0	0.56	0.47	0.67	0.25	0.2	0

Figure 3.4 Coupling Matrix

3.3 Sorting analysis

Coupling matrix is the input to sorting analysis. In sorting analysis, the coupling matrix is analyzed to find high coupling coefficient values and corresponding elements. After finding these elements, the order of rows (machines) and columns (parts) are changed in a way that highly coupled machines and parts are placed adjacent to each other. The change of order in incidence matrix is a NP-hard problem (Chen and Irani, 1993), and there are many ways to execute it. In sorting analysis the change of order of rows and columns is done through two steps. First a hierarchical clustering algorithm is used to construct a tree diagram (step 1). Afterwards, the tree diagram is used to construct the sorted matrix (step 2). Sorted matrix has the same dimension as incidence matrix and the only difference is the order of parts (columns) and machines (rows). In addition, in sorted matrix most of the non-zero entries are along the diagonal line. The above two steps is described in more detail by continuing the example from coupling analysis section.

3.3.1 Tree diagram

The input to sorting analysis is coupling matrix that is built in coupling analysis. In sorting, a hierarchical clustering approach is used to construct the tree diagram. Hierarchical clustering is a numerical classification method that creates a hierarchical structure of elements. Using the hierarchical structure, elements can be grouped to sub-groups in different ways: from a single group containing all elements to more groups with different number of elements or even numerous small groups that contain one element only. This hierarchical structure is represented by a tree matrix (Everitt et al., 2001). For tree diagram construction, the coupling matrix analyzed by the following algorithm (Li, 2011):

Step 1: The corresponding entities to highest value in coupling matrix used for the labeling the tree leaves. Branches can be made with combination of the leaves. The vertical axis is equal to the coupling coefficient value, *Step 2*: Up to dating coupling matrix by combining coupling values of picked entities in previous step. Combining is performed by using average distance formulation,

Step 3: Repeating above two steps till the coupling matrix cannot further reduced.

In the coupling matrix in the figure 3.3, parts 1 and 4 have the highest coupling coefficient value (1). The tree construction started with branching with these two elements. Afterwards coupling matrix revised by substituting columns corresponding part 1 and 4 with new column named P1, 4. The corresponding entities of column P1, 4 in the updated coupling matrix have average value of part 1 and 4 entities in the original coupling matrix. The updated coupling matrix is shown in figure 3.4. The updated coupling matrix is searched again for finding the highest coupling coefficient value. In the new matrix, the coupling coefficient value of 0.7 is highest and related to three pairs: (P1, P4 and M1), (P2 and M2), and (P3 and M3). These are new branches of the three. Coupling matrix again revised. These steps repeated until coupling matrix cannot be reduced, figure 3.6. All repetitions for getting tree diagram in figure 3.6 are explained in the appendix of this thesis, section A.1. The vertical axis in the tree diagram shows the values of the coupling coefficients and in horizontal axis the parts and machines are shown (P1 stands for part 1, and M2 stands for Machine 2).

		Part 1,4	Part 2	Part 3	Part 5	Machine 1	Machine 2	Machine 3	Machine 4
Part	1,4	0	0	0	0.25	0.70	0	0	0.56
Part	2	0	0	0.5	0.67	0	0.70	0.56	0
Part	3	0	0.5	0	0.33	0	0	0.70	0
Part	5	0.25	0.67	0.33	0	0	0.56	0.47	0.47
Machine	1	0.70	0	0	0	0	0	0	0.67
Machine	2	0	0.70	0	0.56	0	0	0.67	0.25
Machine	3	0	0.56	0.70	0.47	0	0.67	0	0.2
Machine	4	0.56	0	0	0.47	0.67	0.25	0.2	0

Figure 3.5 up dated coupling matrix

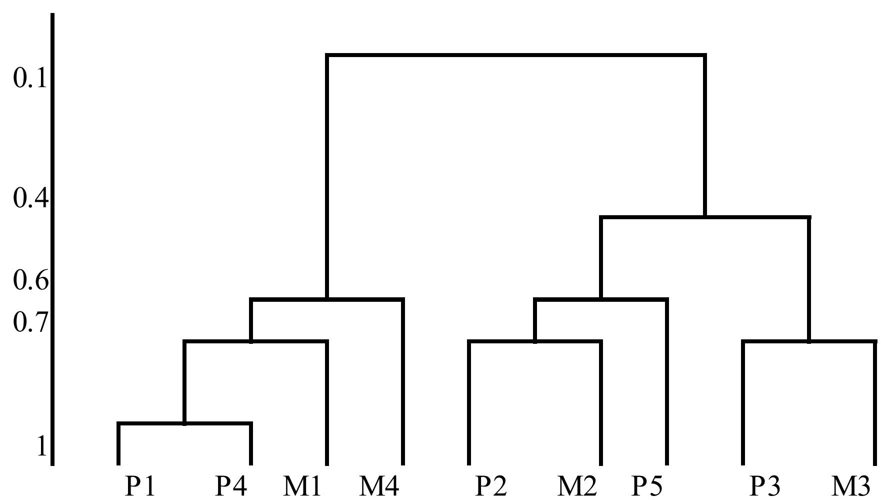

Figure 3.6 Tree diagram

3.3.2 Sorted matrix

In the second step of sorting analysis, tree diagram is used to create the sorted matrix. Sorted matrix can be constructed by reordering the rows and columns of incidence matrix with help of the sequence of nodes in the tree diagram. In figure 3.7 it is shown that how information in tree diagram can help to construct the sorted matrix. The parts and machines on each branch on tree diagram are base for reordering the row and columns of incidence matrix. The notable observation about sorted matrix is that it is a diagonal matrix in which non-zero elements are located around diagonal line of the matrix. In contrast, in incidence matrix, figure 3.2, the non-zero elements are dispersed in the matrix and do not follow a structure.

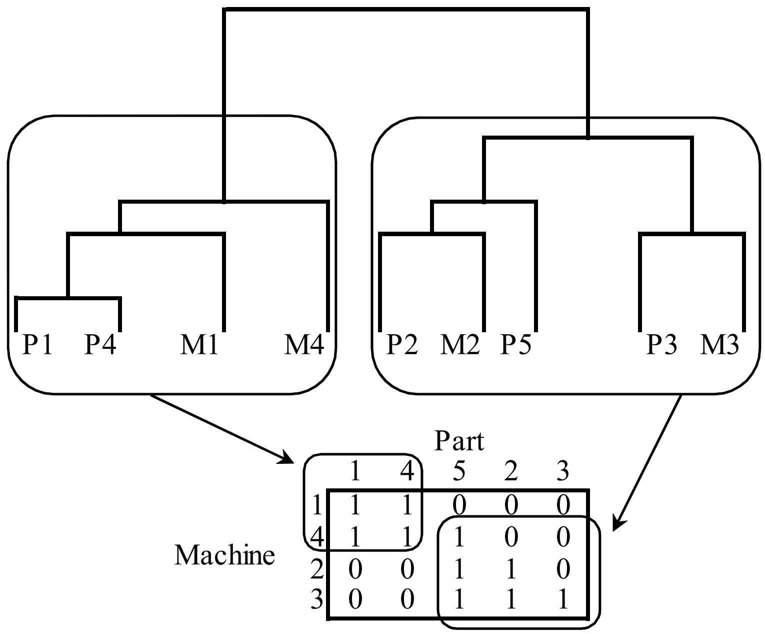

Figure 3.7 Sorted matrix construction

3.4. Partitioning analysis

In final step of matrix-based clustering, partitioning analysis, the work cells are identified on the sorted matrix. In other words output of the partitioning analysis is solution for cell formation problem. Due to structure of sorted matrix, the work cell identification is simple when the size of matrix is small or medium. However, with increasing the number of parts, machines, and off-diagonal non-zero elements in larger sorted matrixes, work cell identification becomes difficult. As a result an algorithm is needed for identifying work cells in sorted matrix.

In partitioning analysis the sorted matrix is divided into partitions by partition points. Each partition is a work cell. Partition points are imaginary points that separate blocks of non-zero entries in a sorted matrix. In Figure 3.8, it is shown how a partition point divides the sorted matrix into two work cells.

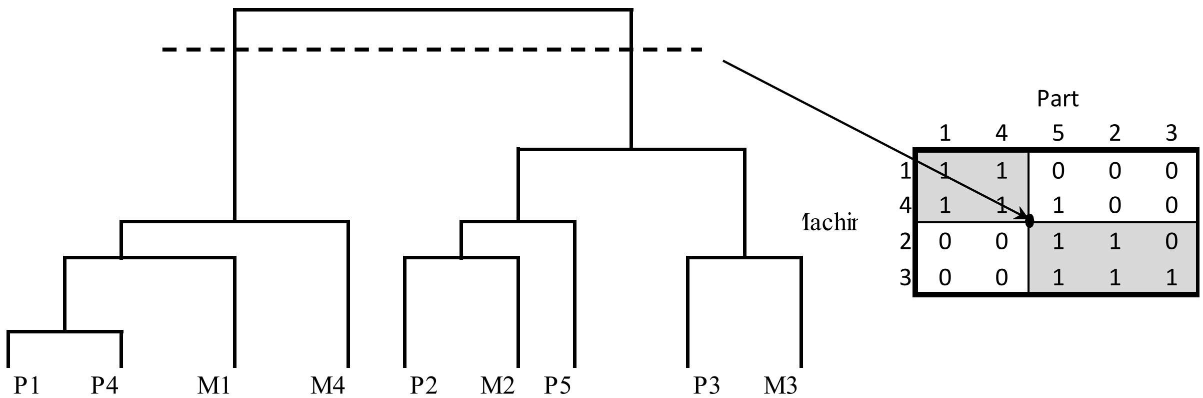

Figure 3.8 Partitioning analysis

There are several algorithms for generating partition points (Li, 2010). In matrix-based clustering a tree-based approach is used for finding the partition points in partitioning analysis. Tree-based approach uses the tree diagram from sorting analysis. In partition analysis the branches of the tree are cut and separated from each other. Each cut branch is a work cell which contains a number of parts and machines. Each cut point on the tree diagram is a partition point.

If the whole tree diagram assumed to be a one branch, there would be one work cell in cell formation solution which is same as job shop manufacturing system. Coming down from top of the tree diagram, more branches can be identified which leads to higher number of work cells in the solution. Better solutions can be achieved with partition points from top branches. However, quality of solution depends on branches of the tree.

3.5 Grouping efficacy

In large cell formation problems, the quality of the final solution can be evaluated with grouping quality measures such as grouping efficacy. Grouping efficacy is one of popular measures for comparison of quality of final solution and it has been used in pervious researches (Kumar and Chandarasekharan, 1990; Sarker and Mondal, 1999; Yin and Yasuda, 2005). The formula of grouping efficacy is

$$\mu = \frac{N_1 - N_1^{out}}{N_1 + N_0^{in}} \tag{3}$$

Where:

N_1 : Total numbers of 1's in the work cells (total number of operations),

N_1^{out}: Total number of 1's outside the work cells (total number of inter-cell movements),

N_0^{in} : Total number of 0's inside the work cells (total number of voids),

μ: grouping efficacy.

The value of the coupling efficacy ranges from 0 to 1. The value of one is an indicator of highest quality, no inter-cell movement and no voids. Values less than one are indicator of existence of voids and inter-cell movement in the solution.

The grouping efficacy for the solution in figure 3.8 is calculated.

$$\mu = \frac{N_1 - N_1^{out}}{N_1 + N_0^{in}} = \frac{10-1}{10+1} = 0.8181$$

3.6 Comparison with traditional similarity coefficient methods

The solution procedure of matrix-based clustering was described in this chapter step by step. While using almost same approach as similarity based clustering methods, there are some differences between matrix-based clustering and other traditional clustering methods. The differences are highlighted in this section.

One feature of coupling analysis is the measurement of coupling coefficient of two different set of elements, machine and part. In similarity based clustering methods only similarity of one set of elements, part-part or machine-machine, is measured to form the initial work cells, and then a clustering algorithm is used to assign the remained set of element to the initial work cells. In other words, grouping of machines is done separately without consideration of both elements. On the other hand in matrix-based clustering, three types of coupling coefficient appropriateness are measured in pairs of part-part, machine-machine, and part-machine simultaneously.

Natural grouping of machines and parts is another different feature of matrix based clustering. In clustering methods there is a need for specifying the number of and size of the work cells as input into hierarchical clustering algorithms. This predefined input information limits the natural grouping of the machines and parts. For instance, if three work cells is defined as input to cell formation problem, solution procedure cannot provide solution with 2 or 4 work cells, even with higher quality.

The solution procedure of matrix-based clustering was described in this chapter. In the following chapters, the application of matrix-based clustering in solving two types of cell formation problem is presented in detail.

Chapter 4

4. Basic machine-part cell formation

In this chapter the matrix based clustering solution procedure for basic machine-part cell formation problem is discussed. First, the basic machine-part cell formation is introduced in problem statement. Next, coupling coefficients of coupling analysis are explained in detail. Afterwards, the performance of matrix-based clustering method is assessed with solving of two cell formation problems. The chapter ends with remarks on matrix-based clustering solution procedure.

4.1. Problem statement

Cell formation problem can be classified into different types based on input information. The machine-part cell formation problem is the first basic cell formation problem. In machine-part cell formation problem, the only provided information is the incidence matrix of parts and machines. In the incidence matrix, machines and parts are represented in rows and columns. Figure 4.1 shows an example of an incidence matrix. The output of cell formation problem is cell formation solution in which provides information on number of work cells and the machines and parts in each work cells.

		Part 1	2	3	4	5	6
Machine	1	1	1	1	0	1	0
	2	0	0	1	1	0	1
	3	0	1	0	0	1	0
	4	1	0	1	1	0	1
	5	0	0	1	1	0	1

Figure 4.1 Incidence matrix

4.2. Coupling Analysis

In coupling, the production information is studied and analyzed in order to evaluate the coupling between machines and parts. Coupling is a measure for appropriateness of machines and parts to be in same work cell. Three types of coupling are measured in matrix-based clustering: part-part, machine-machine, and machine-part. In order to quantify these couplings there is a need for a mathematical equation. Depending on input information, different equations may be used to calculate the coupling. In matrix-based clustering, the calculated value for coupling is known as coupling coefficient. In following sections, each coupling coefficient is explained in the basic machine-part cell formation problem context.

4.2.1. Part-part coupling coefficient

For measuring part-part coupling coefficient parts are compared to each other. In basic machine-part cell formation problem, part-part coupling coefficient between two parts equal to one if they are processed under same machines. For example, if two parts have same design features (i.e. physical shape, adds on, or surface finish), both parts would be processed under the same machine. For this reason, these two parts are more similar to each other rather than other parts that have different design features. As a result, the coupling coefficient of this pair equal one.

Coupling coefficient has a value between 0 and 1. One example on the incidence matrix can better illustrate the concept of part-part coupling coefficient. The following figure shows the same incidence matrix as the figure 4.1.

Machine \ Part	1	2	3	4	5	6
1	1	1	1	0	1	0
2	0	0	1	1	0	1
3	0	1	0	0	1	0
4	1	0	1	1	0	1
5	0	0	1	1	0	1

Figure 4.2 Incidence matrix

In this example, two pairs are discussed: (2, 3) and (3, 4). For comparing part 2 and part 3, the corresponding columns in the incidence matrix should be studied. Part 2 is processed under machines 1 and 3 (corresponding entry in incidence matrix is 1). Part 3 is processed under machines 1, 2, 4, and 5. The difference is the number of used machines and the type of the machines. Part 3 and 4, on the other hand, have more similarity to each other. Part 3 visits machines 1, 2, 4, and 5. Part 4 processed with machines 2, 4, and 5. The possibility of the grouping of part 3 with part 4 is more than being grouped with part 2. If we measure the similarity of the parts with coupling coefficient, logically, the coupling coefficient of pair (3, 4) should be higher than the coupling coefficient of pair (2, 3).

In mathematical term, the coupling coefficient of pair of parts depends on number of 1's and 0's in the two respective columns. Under this coupling logic all parts (columns) are compared to each other with the following coupling coefficient:

$$P_{\min/\max}(e_i, e_j) = \frac{\sum_{k=1}^{m} \min(a_{ki}, a_{kj})}{\sum_{k=1}^{m} \max(a_{ki}, a_{kj})} \qquad i, j \in [1, n] \tag{4}$$

Where two parts i and j are being compared, a_{ki} represents the entry of incidence matrix of part i and machine k, and m is the number of machines and n is number of parts. In this formula, the numerator counts the number of machines that process both parts and the denominator count the number of machines that are being visited by parts i and j. In comparison of any two parts, the numerator of the formula affects the coupling coefficient significantly. With increase in numerator value (both parts visit same machines), the similarity coefficient gets closer to 1. On the other hand if numerator decreases, the coupling coefficient will decrease toward zero. If there is no machine that processes the two parts, the numerator would be zero which results in zero value for coupling coefficient. For above example:

Pair (2, 3), $P_{\min/\max}(e_2, e_3) = \dfrac{1+0+0+0+0}{1+1+1+1+1} = 0.2$

Pair (3, 4), $P_{\min/\max}(e_3, e_4) = \dfrac{0+1+0+1+1}{1+1+0+1+1} = 0.75$

The values of coupling coefficient are supporting the discussion of comparing (2, 3) with (3, 4) pairs in the previous section. The part-part coupling coefficients for all the pairs in the incidence matrix are presented in figure 4.3.

Part	Part	1	2	3	4	5	6
	1	0	0.3333	0.5	0.25	0.3333	0.25
	2	0.3333	0	0.2	0	1	0
	3	0.5	0.2	0	0.75	0.2	0.75
	4	0.25	0	0.75	0	0	1
	5	0.3333	1	0.2	0	0	0
	6	0.25	0	0.75	1	0	0

Figure 4.3 part-part coupling matrix

4.2.2. Machine -machine coupling coefficient

The machine-machine coupling coefficient measures coupling of two machines. In machine-part cell formation problem, machine-machine coupling coefficient value starts from zero where two machines processed completely different parts and increases up to a maximum value of one, in a situation that two machines process same parts.

One example better describes the machine-machine coupling coefficient. The completed part A should have 3 studs and one bracket and a hole. The studs and bracket are assembled on part A by resistance welding, and the hole is made by a drilling. In this example the available machines are a drill machine, a welding robot, welder machine 1, welder machine 2, and a MIG mobile welding station. Among these machines only the two welders can weld the stud and bracket on the part A, welder 1 and welder 2. Drill machine can drill the hole feature on part A. Therefore the coupling coefficient of the pair consisting of welder 1, welder 2, and drill machine is higher than any other pairs when part A is being considered. As a result, these machines have more chance to group together in one work cell for making part A among the five machines.

Min/max coupling coefficient can be a good indicative of the machine-machine coupling coefficient. So the machine – machine coupling coefficient formula is:

$$M_{\min/\max}(e_i, e_j) = \frac{\sum_{k=1}^{n} \min(a_{ik}, a_{jk})}{\sum_{k=1}^{n} \max(a_{ik}, a_{jk})} \qquad i, j \in [1, m] \tag{5}$$

Where two machines i and j are being compared; a_{ik} is the entry in incidence matrix, corresponding machine i and part k ; n is number of the parts and m is number of the machines.

In this formula the numerator counts the parts that are processed under two machines that are being compared, and the denominator counts the number of parts that are being processed under the two machines. Similarly to part-part coupling coefficient, if numerator increases, the coupling coefficient of two machines increases toward one. If the number of parts that are processed by both machines decreases down to zero, the similarity coefficient reduced down to zero too. In the following section the machine-machine coupling coefficient of all machines is shown.

Production sequence for part A: bracket welding → stud welding → drilling→painting

		Part A
Machine	Welder 1	1
	Welder 2	1
	Welder robot	0
	Drill machine	1
	MIG mobile station	0

Figure 4.4 Production operation information for part A

$$M_{\min/\max}(e_{welder1}, e_{welder2}) = \frac{\sum_{k=1}^{n} \min(a_{ik}, a_{jk})}{\sum_{k=1}^{n} \max(a_{ik}, a_{jk})} = \frac{1}{3} = 0.33$$

$$M_{\min/\max}(e_{welder1}, e_{welder\_robot}) = \frac{\sum_{k=1}^{n} \min(a_{ik}, a_{jk})}{\sum_{k=1}^{n} \max(a_{ik}, a_{jk})} = \frac{0}{3} = 0$$

$$M_{\min/\max}(e_{welder1}, e_{drill\_machine}) = \frac{\sum_{k=1}^{n} \min(a_{ik}, a_{jk})}{\sum_{k=1}^{n} \max(a_{ik}, a_{jk})} = \frac{1}{3} = 0.33$$

		Machine				
		Welder 1	Welder 2	Welder robot	Drill machine	MIG mobile station
Machine	Welder 1	0	0.33	0	0.33	0
	Welder 2	0.33	0	0	0.33	0
	Welder robot	0	0	0	0	0
	Drill machine	0.33	0.33	0	0	0
	MIG mobile station	0	0	0	0	0

Figure 4.5 Machine –machine coupling coefficients

The above coupling coefficient values supports the previous discussion in which welder 1, welder 2, and drill machine have the highest potential to group together in a work cell to process part A.

4.2.3. Machine-part coupling coefficient

In this section the coupling of a machine for a part to be assigned to a same work cell will be discussed. The sole criterion in the basic machine-part cell formation problem is that a machine is highly coupled to a part when it can fulfill one of the operations on the part operation sequence list. In other words, the machine can process the part. As a result there is coupling between part and the machine. In the incidence matrix, a machine and a part are coupled when the corresponding element of incidence matrix is a non-zero element. However, it is possible that the part has more corresponding non-zeros with other machines (when the part has more than one operation on its operation sequence list). Therefore, there is possibility that the part groups with other machines in a different work cells. An example will better illustrate machine-part coupling. In the following example, two elements of figure 4.6 incidence matrix are being discussed, (2, 3) and (3, 5).

The corresponding element of machine 2 and part 3 is one; so both may be assigned to same work cell. However, part 3 is processed under other machines (1, 4, and 5) and part 3 may be grouped with them in another work cell as well. With the same logic, machine 2 may be grouped with other parts (4 and 6).

$$
\begin{array}{cc|c|c|c|c|c|c|}
 & & \multicolumn{6}{c}{\text{Part}} \\
 & & 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\
\hline
 & 2 & 0 & 0 & 1 & 1 & 0 & 1 \\
\hline
\text{Machine} & 3 & 0 & 1 & 0 & 0 & 1 & 0 \\
\hline
 & 4 & 1 & 0 & 1 & 1 & 0 & 1 \\
\hline
 & 5 & 0 & 0 & 1 & 1 & 0 & 1 \\
\hline
\end{array}
$$

Figure 4.6 Incidence matrix

For the second pair, (3, 5), the corresponding element in incidence matrix is equal to one, an indicative for appropriateness of machine 3 and part 5 to be assigned into the same work cell. Incidence matrix shows that part 5 has only another operation with machine 1. Machine 3, besides the part 5, only operates on part 2. Comparing the situation for two above pairs, the chance of grouping of pair (3, 5) into one work cell is more than the pair (2, 3).

With same logic every two pairs of machine-part can be compared to each other. First, each pair analyzed individually. The possibility of grouping of each pair is measured by counting and summation of the number of non-zeros in corresponding row (machine) and column (part). In second step, each two pairs (of machine-part) can be compared. A higher value in one pair indicates that components of that pair have more options for grouping to other machines/parts.

On the other hand, the components in the pair with the lower values have less chance of grouping to other machines/parts. As a result there is higher probability that later pair group together. In other words, the coupling coefficient of the pair with lower values is larger than the pair with higher values. Based on above reasoning, the coupling coefficient for machine – part formula is introduced:

$$R_{two-\mathrm{mode}}(m_i, p_j) = \frac{2a_{ij}}{\sum_{k=1}^{o} a_{ik} + \sum_{k=1}^{n} a_{kj}} \qquad (6)$$

In this formula, a_{ij} is the corresponding element of incidence matrix of machine i and part j, and the dominator is total of non-zeros along the corresponding row and column. If the corresponding element of machine I and part j is zero (part j does not visits machine j), numerator will be zero and consequently the machine-part coupling coefficient will be zero. If a_{ij} is equal to one, the value of machine-part coupling coefficient depends on denominator, the total number of 1's on the corresponding row of machine i and corresponding column of part j. The more 1's, the lower (machine *i*, part *j*) coupling coefficient and vice versa. The calculated machine-part coupling coefficients are shown in figure 4.7.

		Part					
		1	2	3	4	5	6
Machine	1	0.3333	0.3333	0.25	0	0.3333	0
	2	0	0	0.2857	0.3333	0	0.3333
	3	0	0.5	0	0	0.5	0
	4	0.3333	0	0.25	0.2857	0	0.2857
	5	0	0	0.2857	0.3333	0	0.3333

Figure 4.7 Machine-part coupling coefficient

4.3. Application: cell formation with binary incidence matrix

For assessment of matrix-based clustering performance in solving of cell formation problems, two previously solved cell formation problem were selected from the literature. The workflow of the matrix-based clustering explained step by step in each problem. The solutions of matrix-based clustering were compared with previous solutions from other methodologies.

The first example is a cell formation problem from research work of Boe and Cheng (Boe and Cheng, 1991). The incidence matrix is shown in figure 4.8. There are 35 parts and 20 machines in this example. Boe and Cheng solved this example with a close neighbor algorithms. Their proposed solution procedure solves the cell formation problem in two steps: machine grouping and part grouping.

Part

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
	1	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	1	0	1	0	0	1	1	1	0	0	0	0	1	0	1	0	1	1
	2	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
	3	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
	4	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
	5	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0
	7	1	0	1	0	1	0	1	0	0	0	0	1	0	0	1	0	1	0	1	1	0	1	1	1	0	1	0	0	1	1	1	1	0	0	0
	8	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
	9	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1
Machine	10	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
	11	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	1
	12	0	0	0	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	13	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	14	0	1	0	0	0	0	1	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0
	15	0	0	0	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0
	16	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0
	17	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	18	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0
	19	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0
	20	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0

Figure 4.8 binary incidence matrix (Boe and Chow, 1991)

4.3.1. Coupling analysis

The input to coupling analysis is the machine-part incidence matrix. In coupling analysis, three types of coupling are measured: part-part, machine-machine, and machine-part. The coupling coefficients for each type are described in previous sections. After calculation of all coupling coefficients, the concatenated coupling matrix can be built by the following structure.

$$\text{Concatenated coupling matrix= CM} = \begin{bmatrix} w_r.CM_c & w_{rc}.CM_{rc}^{\ T} \\ w_{rc}.CM_{rc} & w_c.CM_r \end{bmatrix} \tag{7}$$

In this example w_r and w_c are equal to 1, and w_{rc} is equal to 4.044 according to equation 1 in chapter 3. The coupling matrix is the output of the coupling analysis.

4.3.2. Sorting analysis

In sorting analysis, the coupling matrix from the previous step is analyzed, tree diagram created, and by information from the tree diagram a sorted matrix is made. The sorted matrix is a binary matrix of machine and parts with same dimension as incidence matrix. The only difference between the two is the order of parts and the machines in a way that the most of entries with value of 1 are located on the diagonal area of the sorted matrix. Sorted matrix and tree diagram are the outputs of the sorting analysis phase. Figure 4.9 shows the sorted matrix for the cell formation problem.

Machine \ Part	33	21	11	28	4	9	6	30	32	18	10	27	2	13	12	24	7	31	34	35	20	1	15	25	23	5	17	3	29	19	14	8	22	26	16
12	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
16	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0
8	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	1	0
7	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	1	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

Figure 4.9 sorted matrix

4.3.3. Partitioning analysis

In the output of the sorting analysis, sorted matrix, most of the non-zero elements are concentrated around diagonal of the matrix. With a primary visual inspection of the sorted matrix the agglomerations of non-zero element around diagonal can be noticed, figure 4.10.

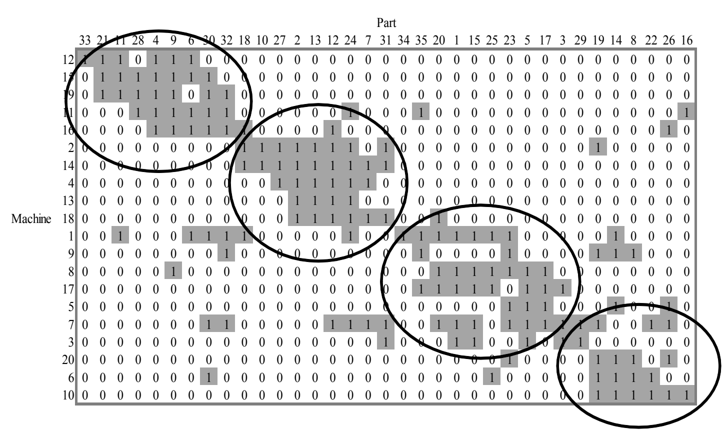

Figure 4.10 An approximate estimate of a four-work cell solution

Work cell identification is performed by partition points by placing the partition points on the matrix grid. Sorted matrix is the coordination plane. The origin (0, 0) is on top left corner of sorted matrix. There are two axes along matrix grid. The figure 4.11 illustrates the coordination system and a partition point (5, 10) on the sorted matrix in figure 4.9. Starting from first partition point, each partition point is put on the corresponding point on the sorted matrix.

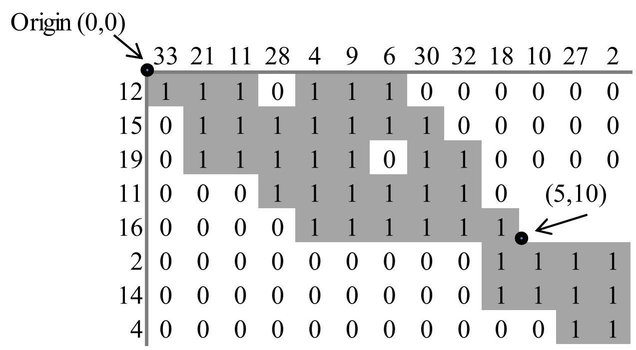

Figure 4.11 Partition points coordination system

4.3.4 Comments on quality of solutions

In addition to Boe and Cheng, this cell formation problem has been solved by different methodologies in the past, as indicated in table 4.1 (Onwubolu and Mutingi, 2001; James et al., 2007). There are solutions with four and five work cells for this problem. With matrix-based clustering, two solutions were produced, one with four work cells and the other with five work cells (Figure 4.12). The quality of solutions can be compared in two ways: comparing the grouping efficacy of solutions and comparing the content of the work cells.

Solution	Source	Methodology	Number of work cells	Grouping efficacy	Size of smallest work cell (machine x part)
1	Current research	Matrix-based clustering	5	0.5357	3 x 6
2	James et al.	Genetic algorithm	5	0.5797	1 x 6
3	Current research	Matrix-based clustering	4	0.5363	3 x 4
4	Boe and Cheng	Close neighbor algorithm	4	0.5158	3 x 8
5	Onwubolu and Mutingi	ZODIAK	4	0.5113	-
6		TSP-GA		0.5514	
7		Genetic algorithm		0.4444	

Table 4.1 Comparison of quality of solution

Grouping efficacy is a measure of quality based on number of inter-work cell movements and voids inside the work cells. The equation for grouping efficacy was explained in section 3.5. The quality of the solution was compared with solutions from previous researches, table 4.1. The values of grouping efficacy indicate that quality of matrix-based clustering solutions is very close to other methodologies. While it has better quality than the original solution, it lacks a bit comparing to others. However, higher grouping efficacy in other methods is not indicator of total superiority of their solutions. For instance, in the solution of James et al. (2007) there is a work cell with only one machine and four parts, which has many visiting parts from other work cells. This reduces the work cell balance and defies the purpose of cell formation in cellular manufacturing.

Solutions can be compared by content of work cells. In general when solutions are compared to each other some work cells are very similar. In solutions with four work cells, solution 3 and solution 4, the content of work cells are very similar in solutions, Table 4.2. Only three parts, 34, 35, and 23 are in different work cells in two solutions. In work cells with five work cells, work cells are similar in solution 1 and 2. However, more parts have different positions. One Important note is that in solution 2, from James et al. research, there is one work cell that has only one machine. This design has influenced the contents of other work cells and cause difference between solution 1 and 2.

	Solution 1											Solution 2										
	Part											Part										
Work cell 1	8	14	16	19	22	26						8	14	19	22	23	26					
Work cell 2	4	6	9	11	21	28	30	32	33			4	6	9	11	21	28	30	32			
Work cell 3	2	7	10	12	13	18	24	27	31			2	7	10	12	13	24	27	31			
Work cell 4	1	15	20	23	25	34	35					1	3	5	15	17	20	29				
Work cell 5	3	5	17	29								16	18	25	33	34	35					

	Solution 3											Solution 4										
	Part											Part										
Work cell 1	4	6	9	11	21	28	30	32	33			4	6	9	11	21	28	30	32	33	34	35
Work cell 2	1	3	5	15	17	20	23	25	29	34	35	1	3	5	15	17	20	25	29			
Work cell 3	2	7	10	12	13	18	24	27	31			2	7	10	12	13	18	24	27	31		
Work cell 4	8	14	16	19	22	26						8	14	16	19	22	23	26				

Table 4.2 Work cell content comparison in four solutions

	33	21	11	28	4	9	6	30	32	18	10	27	2	13	12	24	7	31	34	35	20	1	15	25	23	5	17	3	29	19	14	8	22	26	16
12	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
16	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Machine 18	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0
8	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	1	0
7	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	1	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

a) Solution 1 (with four work cells)

	33	21	11	28	4	9	6	30	32	18	10	27	2	13	12	24	7	31	34	35	20	1	15	25	23	5	17	3	29	19	14	8	22	26	16
12	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
16	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Machine 18	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0
8	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0	1	0
7	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	1	0	1	1	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	1	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1

b) Solution 3 (with five work cells)

Figure 4.12 Matrix-based clustering solutions

Part

Machine	30	32	6	34	35	4	9	11	21	28	33	1	5	15	17	20	25	3	29	31	2	10	12	13	18	24	27	7	8	14	19	23	26	16	22
7	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	0	0	1	0	0	1	0	1	0	0	1	1	1	0	1
1	1	1	1	1	1	0	0	1	0	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1	0	1	0	0	0
11	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0
16	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0
19	1	1	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	1	0	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	1	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
17	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	1	0	1	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0
9	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1
6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1
5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0

Figure 4.13 Original solution (Boe and Chow, 1991)

Part

Machine	16	18	25	33	34	35	2	7	10	12	13	24	27	31	1	3	5	15	17	20	29	8	14	19	22	23	26	4	6	9	11	21	28	30	32
1	0	1	1	0	1	1	0	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	1	0	1	0	0	1	1
2	0	1	0	0	0	0	1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	1	1	0	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	0	0	0	0	1	1
8	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
17	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0
9	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	1
10	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0
11	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1
12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0
16	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	1

Figure 4.14 James et al. solution (James et al., 2007)

4.4. Application: cell formation with non-binary incidence matrix

Binary machine-part matrix only contains one type of information which is necessary but not enough for a more realistic cell formation solution. By using a non-binary incidence matrix, more input information can be considered in basic machine part cell formation problem. For instance, each entry can represent the degree to which a part may be completed under a machine (Lee and Wang, 1991) or ratio of time which a part is being processed under a machine (George et al., 2003). From production viewpoint, the corresponding part and machine to higher value should be grouped together in one work cell.

4.4.1 Solution procedure

The matrix-based clustering solution procedure for a non-binary matrix is same as for a binary matrix. There are three steps: coupling analysis, sorting analysis and partitioning analysis. In coupling the same coupling coefficients as previous application are being used. Sorting analysis and partitioning analysis are performed likewise.

The cell formation problem for this section is from a previous research by Venugopal and Narendran (1992). A simulated annealing methodology was used to solve the problem. In this case, there are 19 parts and 12 machines. The non-binary incidence matrix in the figure 4.15 is the input into the coupling analysis. Every non-zero element in the matrix is representative of the actual workload of the corresponding part on the corresponding machine (Venugopal and Narendran, 1992).

Machine \ Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0.38	0	0	0	0	0	0.32	0	0	0	0	0	0	0	0	0	0	0.1	0	0.12
2	0	0.11	0	0	0	0	0	0	0	0	0	0	0.52	0	0	0	0	0	0	0
3	0	0.09	0.34	0	0.11	0	0	0	0	0	0	0	0	0	0.27	0	0	0	0	0.23
4	0	0	0.07	0	0	0	0	0.21	0	0	0.05	0.36	0	0	0	0	0.18	0	0	0.09
5	0	0.31	0	0	0	0	0	0	0	0	0.34	0	0	0	0.11	0.14	0.03	0	0	0
6	0.27	0	0	0.07	0	0.04	0.21	0.06	0.34	0	0	0	0	0	0	0	0	0	0	0
7	0	0.1	0	0	0	0.09	0.49	0	0	0	0	0	0	0	0	0	0.12	0.18	0	0
8	0	0.08	0.17	0	0	0	0	0	0	0	0	0	0	0	0.24	0	0	0.32	0	0.21
9	0.36	0	0	0.04	0	0	0.24	0.02	0.26	0	0	0	0	0	0.06	0	0	0	0	0
10	0	0.02	0	0.23	0	0	0	0.17	0	0.04	0	0	0.38	0	0	0	0	0	0.14	0
11	0.03	0	0.01	0	0	0	0	0	0	0.02	0	0.32	0.37	0	0	0.26	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0.58	0	0	0	0	0	0	0	0
13	0.01	0.21	0.02	0	0.29	0	0	0	0	0.18	0.26	0	0.06	0.03	0	0	0	0	0	0
14	0	0	0.1	0	0	0.17	0.01	0.13	0.03	0	0.02	0.01	0	0.19	0	0.11	0.2	0	0.14	0
15	0	0.27	0	0	0.22	0	0	0.09	0	0.27	0.18	0	0	0	0	0	0.06	0	0	0
16	0.21	0	0.01	0	0	0	0.32	0	0.31	0	0	0	0.07	0	0.04	0.01	0.03	0.02	0	0
17	0	0.12	0	0	0	0	0	0	0	0	0	0.39	0.31	0.18	0	0	0	0	0	0
18	0	0	0.02	0	0	0	0.04	0.14	0	0	0.01	0.33	0.35	0.12	0	0	0	0	0	0
19	0	0	0	0	0	0	0.33	0	0.36	0.05	0.1	0	0	0.09	0	0	0	0	0	0.08
20	0	0.14	0	0	0	0	0.52	0.16	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4.15 Incidence matrix, non-binary example

4.4.2 Comments on the quality of solutions

After performing the three stages, the matrix-based clustering provides a solution very similar to the original solution. Three work cells are same in both solutions. However, the numbers of the work cells differ. While in original solution there is large work cell, matrix-based clustering breaks that work cell into two smaller work cells. As a result solutions are different. In this example, Matrix-based clustering provides a solution with higher quality comparing to the original solution. The values for grouping efficacy as well as number of voids and exceptional part are shown in table 3.2.

Solution	Source	Methodology	Number of inter-cell movements	Number of the voids	Grouping efficacy
1	Current research	Matrix-based clustering	59	24	0.395
2	Venugopal and Narendran (1992)	Simulated annealing	53	55	0.341

Table 4.3 Comparison of quality of solutions

Machine \ Part	16	13	12	4	19	14	6	8	17	11	2	10	5	7	9	1	20	3	15	18
11	0.26	0.37	0.32	0	0	0	0	0	0	0	0	0.02	0	0	0	0.03	0	0.01	0	0
2	0	0.52	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0	0	0	0
12	0	0	0.58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0.31	0.39	0	0	0.18	0	0	0	0	0.12	0	0	0	0	0	0	0	0	0
18	0	0.35	0.33	0	0	0.12	0	0.14	0	0.01	0	0	0	0.04	0	0	0	0.02	0	0
10	0	0.38	0	0.23	0.14	0	0	0.17	0	0	0.02	0.04	0	0	0	0	0	0	0	0
14	0.11	0	0.01	0	0.14	0.19	0.17	0.13	0.2	0.02	0	0	0	0.01	0.03	0	0	0.1	0	0
4	0	0	0.36	0	0	0	0	0.21	0.18	0.05	0	0	0	0	0	0	0.09	0.07	0	0
5	0.14	0	0	0	0	0	0	0	0.03	0.34	0.31	0	0	0	0	0	0	0	0.11	0
15	0	0	0	0	0	0	0	0.09	0.06	0.18	0.27	0.27	0.22	0	0	0	0	0	0	0
13	0	0.06	0	0	0	0.03	0	0	0	0.26	0.21	0.18	0.29	0	0	0.01	0	0.02	0	0
7	0	0	0	0	0	0	0.09	0	0.12	0	0.1	0	0	0.49	0	0	0	0	0	0.18
20	0	0	0	0	0	0	0	0.16	0	0	0.14	0	0	0.52	0	0	0	0	0	0
19	0	0	0	0	0	0.09	0	0	0	0.1	0	0.05	0	0.33	0.36	0	0.08	0	0	0
16	0.01	0.07	0	0	0	0	0	0	0.03	0	0	0	0	0.32	0.31	0.21	0	0.01	0.04	0.02
6	0	0	0	0.07	0	0	0.04	0.06	0	0	0	0	0	0.21	0.34	0.27	0	0	0	0
9	0	0	0	0.04	0	0	0	0.02	0	0	0	0	0	0.24	0.26	0.36	0	0	0.06	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.32	0	0.38	0.12	0	0	0.1
3	0	0	0	0	0	0	0	0	0	0	0.09	0	0.11	0	0	0	0.23	0.34	0.27	0
8	0	0	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0.21	0.17	0.24	0.32

Figure 4.16 Solution for non-binary incidence matrix by using Matrix-based clustering

Machine \ Part	1	7	9	5	2	10	11	12	13	4	6	8	14	16	17	19	15	3	18	20
1	0.38	0.32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.12
6	0.27	0.21	0.34	0	0	0	0	0	0	0.07	0.04	0.06	0	0	0	0	0	0	0	0
7	0	0.49	0	0	0.1	0	0	0	0	0	0.09	0	0	0	0.12	0	0	0	0.18	0
9	0.36	0.24	0.26	0	0	0	0	0	0	0.04	0	0.02	0	0	0	0	0.06	0	0	0
16	0.21	0.32	0.31	0	0	0	0	0	0.07	0	0	0	0	0.01	0.03	0	0.04	0.01	0.02	0
19	0	0.33	0.36	0	0	0.05	0.1	0	0	0	0	0	0.09	0	0	0	0	0	0	0.08
20	0	0.52	0	0	0.14	0	0	0	0	0	0	0.16	0	0	0	0	0	0	0	0
5	0	0	0	0	0.31	0	0.34	0	0	0	0	0	0	0.14	0.03	0	0.11	0	0	0
13	0.01	0	0	0.29	0.21	0.18	0.26	0	0.06	0	0	0	0.03	0	0	0	0	0.02	0	0
15	0	0	0	0.22	0.27	0.27	0.18	0	0	0	0	0.09	0	0	0.06	0	0	0	0	0
12	0	0	0	0	0	0	0	0.58	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0.12	0	0	0.39	0.31	0	0	0	0.18	0	0	0	0	0	0	0
10	0	0	0	0	0.02	0.04	0	0	0.38	0.23	0	0.17	0	0	0	0.14	0	0	0	0
11	0.03	0	0	0	0	0.02	0	0.32	0.37	0	0	0	0	0.26	0	0	0	0.01	0	0
2	0	0	0	0	0.11	0	0	0	0.52	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0.05	0.36	0	0	0	0.21	0	0	0.18	0	0	0.07	0	0.09
18	0	0.04	0	0	0	0	0.01	0.33	0.35	0	0	0.14	0.12	0	0	0	0	0.02	0	0
14	0	0.01	0.03	0	0	0	0.02	0.01	0	0	0.17	0.13	0.19	0.11	0.2	0.14	0	0.1	0	0
3	0	0	0	0.11	0.09	0	0	0	0	0	0	0	0	0	0	0	0.27	0.34	0	0.23
8	0	0	0	0	0.08	0	0	0	0	0	0	0	0	0	0	0	0.24	0.17	0.32	0.21

Figure 4.17 Original solution (Venugopal and Narendran, 1992)

4.5. Closing remark

In this chapter it was demonstrated that how matrix-based clustering can solve a basic machine-part cell formation problem. Two solved problems show that matrix-based clustering can provide equal or better solutions than other methods for cell formation problem. In addition, more information can be considered in matrix-based clustering solution procedure by using non-binary incidence matrix. By our knowledge, so far no similarity based clustering method can use non-binary incidence matrix as input into cell formation problem.

Chapter 5

5. Cell formation with production sequence information

In this chapter another type of cell formation problem is solved with matrix-based clustering method, cell formation problem with production sequence information. The structure of this chapter is similar to previous chapter. First, the problem is defined, then solution procedure is described according to type of the problem, and chapter finishes by solving a cell formation problem from previous literature.

5.1. Problem statement

Another type of cell formation problem is cell formation problem with production sequence information. In this type of problem, the input information into the problem is the part-machine incidence matrix and production sequence information. Part-machine incidence matrix shows the machines that each part visits. Production sequence is the information on the order of machines each part visits. The output of cell formation problem is cell formation solution which provides information on number of work cells and constitutes of each work cell.

In cellular manufacturing, production sequence can influence overall performance in different ways. The number of trip between machines affects machine handling cost, and number of intercellular moves (Alhourani and Seifoddini, 2007). It also can have an effect on layout of machines inside each work cell (Mahdavi and Mahadevan, 2007). Consideration of production sequence in cell design helps in implementation of just-in-time (Park and Suresh, 2003). Due to importance of production sequence information many efforts have been devoted to include it in cell formation problem (Yin and Yasuda, 2006).

In Similarity based clustering methods, several similarity coefficients have been developed to capture similarity between machines or parts with production sequence information. At the beginning, similarity coefficients were simple; however, in later researches more production factors such as part volume, batch size, and operation time were included in similarity coefficients (Yin and Yasuda, 2006; Alhourani and Seifoddini, 2007). In the following sections, the matrix-based clustering solution procedure for the cell formation problem with production sequence information is explained.

5.2. Coupling analysis

Due to structure of matrix-based clustering solution procedure, in which input data is used first in coupling analysis, coupling analysis should be capable to use all input information to measure coupling coefficients. In case of cell formation problem with production sequence information, coupling analysis use production sequence information in addition to incidence matrix to measure coupling coefficients. There are three different coupling coefficients in matrix-based clustering, machine-machine, part-part, and machine-part. These coupling coefficients use production sequence information to measure coupling.

5.2.1. Part-part coupling coefficient

When production sequence is considered, two parts are highly coupled if both parts visit common machines in similar order. Figure 5.1 presents the input information to a cell formation problem with production sequence information. Figure 5.1.a shows an incidence matrix, and figure 5.1.b shows the production sequence information for each part. In this example, parts 1 and 2 have higher coupling than parts 2 and 3. The reason is similarity of production sequence of parts 1 and 2. Parts 2 and 3 have very different production sequence.

		Part						
		1	2	3	4	5	6	7
Machine	1	1	1	0	1	0	0	0
	2	1	1	1	1	0	1	0
	3	0	1	0	0	1	1	1
	4	0	0	1	1	1	1	1
	5	1	0	1	0	0	0	0

a) Incidence matrix

	Part						
	1	2	3	4	5	6	7
Machine order	1	1	5	1	3	2	4
	2	2	4	2	4	3	3
	5	3	2	4		4	

b) Production sequence information

Figure 5.1 input for cell formation problem with production sequence information

Based on above reasoning coupling of parts can be calculated by coupling coefficient. In this chapter, two coupling coefficients are used in for measuring the coupling of parts: min/max and a similarity measure from research of Sarker and Xu (2000) which will be called Sarker-Xu coupling coefficient in this text. The complete description and examples on Sarker-Xu coupling coefficient are provided in the appendix.

The value of coupling coefficient from min/max and Sarker-Xu coupling are very close to each other. The small difference is due to measurement method of each coupling coefficient. In measuring coupling of any two parts, while Sarker-Xu coupling coefficient is calculated based on order of machines in production sequence, min/max coupling is measured by common machines for two parts. There is no reason that if two part visit common machines, they will visit them in same order. However, practically those parts that are made with common machines are put together in one work cell. Use of both coupling coefficient will be demonstrated and discussed in a cell formation example later in application section of this chapter.

5.2.2 Machine-machine coupling coefficient

Coupling relationship between machines depends on the production sequence of the parts. If the relative positions of the two machines in production sequence of parts be same for each part, the two machines are highly coupled. For example, coupling of machines 1 and 2 in figure 5.1 is higher than the coupling of machines 2 and 5. This can be observed by studying the production sequence of the parts. The positions of machines 1 and 2 in the production sequence of the parts that visit these two machines are very similar. Three parts visit both machines 1 and 2 and all parts first visit machine 1 and then machine 2. Two common parts visit machines 2 and 5. However the production sequence is different. Part 1 first visits machine 2 and then it visits machine 5(production sequence: $2\rightarrow5$). Part 3 production sequence is $5\rightarrow4\rightarrow2$. Not only the order of visit of common machines is not same, but also there is extra visit on machine 4 in middle between machines 5 and 2. In other words, there is distance between machine 5 and 2 in the production sequence of part 3. As a result, comparatively, considering production sequence, the coupling of machine pair 1 and 2 is higher than machine pair of 2 and 5. As discussed in above example, production sequence can influence on coupling relationship between any two machines in different ways. Coupling coefficient should be able to compare the machines based on machines position and distance.

Two coupling coefficients are used to measure machine-machine coupling in cell formation with production sequence information: min/max and a coupling coefficient based on a similarity measure from research of Nair and Narendran (1998). This coupling coefficient is called Nair-Narendran coupling coefficient in this text. The complete description and examples on this coupling coefficient are provided in the appendix.

The difference between Min/max and Nair-Narendran coupling coefficient is that in Nair-Narendran coupling measurement between the parts, the machines that are not first or last in production sequence of each part are distinguished from first and last machines. For example, in figure 5.1.b, the sequence operation of part 3 is 5→4→2. Machine 4 in this production sequence is not first or last. In final solution if part 3 grouped into a work cell that has machine 5 and 2 but not machine 4, there will be an inter-work cell movement in production of part 3. Part 3 should leave the work cell to another work cell where machine 4 is located, and come back to its initial work cell to visit machine 2. Adding extra importance to machine 4 in solution procedure may result in better solution with less inter-cell movements. Min/max coupling coefficient does not consider this criterion for measuring the coupling coefficients between machines. The use of both coupling coefficients will be demonstrated in an application later in this chapter.

5.2.3 Machine-part coupling coefficient

The machine-part coupling coefficient for cell formation problem with production sequence information is same as the one for machine-part cell formation problem, in previous chapter. If a part is processed under a machine, then machine and part have a degree of coupling to be grouped in same work cell. This coupling is influenced by the number of other parts which are processed under the machine; higher the number, lower the coupling and will result in lower coupling coefficient. The detail, numerical examples, and the coupling coefficient for machine-part (two mode coupling coefficient) are described in chapter 4.

5.2.4 Coupling matrix

After measurement of all coupling coefficients, coupling matrix can be constructed. The structure and method is explained in chapter three.

5.3 Sorting and partitioning Analysis

Sorting and partitioning analysis in cell formation problem with sequence information are executed exactly same as basic machine-part cell formation problem and are not repeated in this chapter. Detail information is available in chapters three and four.

5.4. Application

Solving a cell formation problem with different coupling coefficient is better indicator of matrix-based clustering performance in a cell formation problem with production sequence. In following cell formation problem, from a previous research, twenty parts are processed with eight machines. In their research, Nair and Narendran (1998) proposed a similarity coefficient for machines. Subsequently, clustering was performed using a non-hierarchical clustering algorithm. Figures 5.2 and 5.3 show the incidence matrix and the production sequence information.

		Part																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Machine	1	0	1	1	0	0	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0
	2	0	0	1	1	0	1	1	0	0	0	0	0	0	1	0	0	0	1	0	1
	3	0	1	0	0	0	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0
	4	0	0	1	1	0	1	1	0	0	1	0	0	0	0	1	0	0	1	0	1
	5	1	0	0	0	1	1	0	0	0	1	0	1	0	0	1	0	1	0	0	0
	6	1	0	0	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1
	7	0	0	1	1	0	1	1	0	0	0	1	1	0	0	0	0	0	1	0	1
	8	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1

Figure 5.2 incidence matrix

Part	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Machine order	6 5	1 3	2 1 7 8 4	2 4 7 8	6 5	2 4 7 8 5	8 4 7 2	1 3	1 6 3	6 4 5	7 3 1	5 7 6	1 3	1 2 3	4 5	1 3	3 5 1	4 2 8 7	1 3	4 2 6 7 8

Figure 5.3 Production sequence information

This cell formation problem is solved in three steps: coupling analysis, sorting analysis, and the partitioning analysis. In coupling analysis, three types of coupling are measured in order to build the coupling matrix: machine-machine, part-part, and machine-part. In this example, the cell formation is solved using two new coupling coefficients in addition to the three coupling coefficients used in the previous cell formation problem. Two similarity coefficients were chosen from literature and the cell formation problem solved with different combinations of coupling coefficients, table 5.1. The sorting and partitioning analysis was performed in same way as the machine-part cell formation problem in previous chapter.

5.4.1. Solution procedure

In coupling analysis, four different combinations of coupling coefficient used to solve the example, Table 5.1. In first solution, Sarker and Xu similarity coefficient is used to measure coupling coefficient of part pairs and two other coupling coefficients are same as previous chapter. In the second solution, Nair and Narendran similarity coefficient was used to measure machine-machine coupling coefficient and the two other are same as basic machine-part problem coupling coefficients. In the third solution, part-part and machine-machine dependencies were measured with the two coupling coefficients from other literature and two-mode was used for measuring machine-part coupling coefficients. In the last solution, the coupling coefficients of previous cell formation problem are used to capture sequence information of the example. Figure 5.4 shows the three solutions along the original solution. The result of third and fourth combination was same and displayed as solution three.

5.4.2 Comments on the quality of solutions

When solutions compared visually, none of them are perfect and all have inter-cell movements and voids in work cells. However, solution 1 and the original solution are more practical. They have less inter-cell movements and there is only on void in one of their work cells. These solutions are denser along diagonal line.

From machine and part grouping perspective, all solutions provide almost the same configuration with few exceptions. In solution 1, parts and machine are same as original solution except two parts, 11 and 14. The solution 2 provides a different combination of machine assignment and part families to the work cells, and this causes voids and exceptional part in the solution. In spite of these, two work cells in solution 2 are very similar to the other solutions. The original solution and solution 3 are identical. The result in this example indicates that matrix-based clustering creates a solution that has same work cells as other methods.

For further comparison of the solutions, grouping efficiency is measured by a grouping measure. First grouping efficacy of the solutions were calculated, table 5.1. The values show the same quality and confirm visual study observations. Solution 3 and the original solution have the same highest grouping efficacy, and solution 1 and 2 come after with less grouping efficacy.

Part

Machine	5	1	10	12	15	17	9	13	16	19	2	8	11	6	3	18	4	7	20	14
6	1	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
5	1	1	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
3	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1
1	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1
7	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0
4	0	0	1	0	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1

a) Solution1

Machine	15	10	5	1	12	20	6	18	4	7	3	14	11	17	13	16	19	2	8	9
5	1	1	1	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
4	1	1	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
6	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
8	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
7	0	0	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1

b) Solution 2

Figure 5.4 Matrix-based clustering and the original solution for the cell formation problem

Figure 5.4 (Continued)

		1	5	10	12	15	20	6	18	7	4	3	11	14	17	9	13	16	19	2	8
	6	1	1	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	5	1	1	1	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
	4	0	0	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
Machine	7	0	0	0	1	0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
	3	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1

c) Solution 3

		2	8	9	11	13	14	16	17	19	3	4	6	7	18	20	1	5	10	12	15
	3	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	1
	7	0	0	0	1	0	0	0	0	0	1	1	1	1	1	1	0	0	0	1	0
Machine	8	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0
	2	0	0	0	0	0	1	0	0	0	1	1	1	1	1	1	0	0	0	0	0
	5	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	1
	6	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0

d) Original solution (Nair and Narendran, 1998)

Figure 5.4 Matrix-based clustering and the original solution for the cell formation problem

Solution procedure	Machine-Machine coupling coefficient	Part-part coupling coefficient	Grouping efficacy
1	Min/Max	Sarker and Xu	0.7692
2	Nair and Narendran	Min/Max	0.6521
3	Nair and Narendran	Sarker and Xu	0.8225
4	Min/max	Min/Max	0.8225
Nair and Narendran	-	-	0.8225

Table 5.1 Comparison of quality of solutions

5.5 Closing remarks

In this chapter, it was shown that how matrix-based clustering can solve a cell formation problem with production sequence information. An application shows that matrix-based clustering can provide equal solution to other methodology for this type of cell formation problem. In addition, the following observations can be made from solution procedure:

- Min/max and two mode coupling coefficients can incorporate production sequence,
- Production information can be incorporated into cell formation problem by any of the coupling coefficients,
- The format of final solution from matrix-based clustering lets user to change the configuration of the work cells when needed. The agglomeration of non-zero element along the sorted matrix makes this possible. Since non-zero elements are close to each other. It is easier to revise the number and size of the work cells.

Chapter 6

6. Conclusions

In this chapter, the thesis ends with presenting a summary of research, a review of contributions and suggestions for future research.

6.1 Summary

Despite of many researches on cell formation in past decades, cell formation methods are yet far from being routine in manufacturing industry and research in this area is ongoing. A clustering method for cell formation is presented in this research. The matrix-based clustering identifies work cells by measuring coupling relationships between machines and parts and using coupling information to restructure the incidence matrix. Two different cell formation problems were solved by matrix-based clustering, simple machine-part and cell formation problem with production sequence information. The results indicate that matrix-based clustering is effective in providing quality solutions in both types of problem.

6.2 Contributions

Coupling analysis brings notable advantages for matrix-based clustering. First, there is no need to predefined work cell configuration input into cell formation problem. Therefore natural work cell formation is achievable in matrix-based clustering. Second, the measurement of coupling of two different set of elements of production system, machine and parts, to each other is a unique feature of matrix-based clustering. In other methodologies, by our knowledge, such a direct analysis is not available. In third feature of coupling analysis, decomposition of grouping into three sub-grouping makes the solution procedure simpler and flexible. In coupling each coupling coefficient can handle production factors separately in a simple way. In other methodologies, considering additional production factors make formulation very large and complex.

Another feature of matrix-based clustering is simultaneous grouping of machines and parts. In similarity based clustering methods, the machines or parts are grouped sequentially one after another. On the other hand, in matrix-based clustering coupling analysis, grouping of parts and grouping of machines are done separately and simultaneously.

The diagonal matrix structure of the final solution is another advantageous point of matrix-based clustering. The output of sorting/partitioning analysis has specific characteristic which can be used for practical production issues. For instance, if a work cell configuration does not provide a balanced solution for production, there is a need for changing work cell design. Sorted matrix in which its elements are agglomerated along diagonal helps to reconfigure the work cell easily. Whereas in other cell formation methods the final solution is not flexible enough to be reconfigured easily, matrix-based clustering has overcome this problem with its final solution format, the dense diagonal matrix. In short the main contributions of this research are:

- Simultaneous grouping of part and machines,
- No need for input production information in a clustering cell formation approach,
- A flexible method for cell formation.

The table 2.1 in chapter 2 can be updated by adding matrix-based clustering method, table 6.1.

Method	Identify part families and machine groups simultaneously	No need for input information	Flexibility
Machine component analysis	✓	✓	-
Similarity coefficient based clustering	-	-	✓
Mathematical programming	✓	-	-
Metaheuristics	✓	-	-
Matrix-based clustering	✓	✓	✓

Table 6.1 comparison of cell formation methodologies

6.3 Future work

Matrix-based clustering has the ability of solving different types of cell formation problems. In future studies it can be tailored to solve the cell formation problem with more production information such as alternative routing and simultaneous consideration of alternative routing and production sequence information.

The presented method only captures the dependencies of the part and machines. However, there are other resources that influence the work cell configuration as well. Examples are work force production equipment, and space. These two factors can have dependency relationship with both machines and parts. For instance, a part may be very heavy/big to move by one operator and there is a need for moving equipment inside a work cell. This leads to many considerations in work cell configuration. There is a need to capture more dependencies in matrix-based clustering.

Different methodologies have been used for clustering approach in cell formation problems; examples are integer programming or OAI methods such as genetic algorithm. A comparison between performance of matrix-based clustering and these methods can be another research area in this field.

References

Adil, G.K. et al. (1996). Cell formation considering alternative routings, *International Journal of Production Research,* 34(5), 1361-1380.

Alhourani, F., & Seifoddini, H. (2007). Machine cell formation for production management in cellular manufacturing systems. *International Journal of Production Research*, 45(4), 913-934.

Aljaber N. et al. (1997). A tabu search approach to the cell formation problem, *Computers and Industrial Engineering,* 32(1), 169-185.

Balakrishnan, J., & Cheng, C.C. (2007). Multi-period planning and uncertainty issues in cellular manufacturing: A review and future directions. *European Journal of Operation research,* 177(1), 281-309.

Balasubramanian, K.N. & Panneerselvam, R. (1993). Covering technique-based algorithm for machine grouping to form manufacturing cells, *International Journal of Production Research,* 31, 1479-1504.

Black, J.T., & Hunter S. L. (2003). *Lean manufacturing systems and cell design*. Dearborn, Michigan: Society of Manufacturing Engineers.

Boe W.J., & Cheng C.H. (1991). A close neighbour algorithm for designing cellular manufacturing systems. *International Journal of Production Research,* 29(10), 2097-2116.

Burbidge, J.L. (1989). *Production flow analysis for planning group technology*. Oxford: Clarendon Press.

Chen, C.Y., & Irani, S.A. (1993). Cluster first-sequence last heuristics for generating block diagonal forms for a machine-part matrix. *International Journal of Production Research*, 31(11), 2623-2647.

Chen, L. et al. (2007). Model-based rapid redesign using decomposition patterns. *Journal of Mechanical Design*, 129(3), 283-294.

Chen, C. L. et al. (1995). A Simulated annealing solution to the cell formation problem. *International Journal of Production Research,* 33, 2601-2614.

Cheng, C. H. et al. (1998). A TSP-based heuristic for forming machine groups and part families. *International Journal of Production Research*, 36(5), 1325-1337.

Deutsch, S.J. et al. (1998). Manufacturing cell formation using an improved p-median model, *Computers and Industrial Engineering,* 34, 135-146.

Everitt, B.S. et al. (2001). *Cluster Analysis* (Fourth edition). New York Oxford University Press.

Gallagaher, C.C., & Knight, W.A. (1973). *Group Technology*, London: Butterworth.

Gonçalves, J.F. et al. (2004). An evolutionary algorithm for manufacturing cell formation, *Computers and industrial Engineering,* 47, 247-273.

Groover M.P. (2008). *Automation, production systems and computer integrated manufacturing*(Third edition). Upper Saddle River: Pearson Education Inc.

Gupta, T. (1993). Design of manufacturing cells for flexible environment considering alternative routing, *International Journal of Production Research*, 31, 1259-1273.

Gupta, T., & Seifoddini 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(7), 1274-1269.

Hitomi, K. (1996). *Manufacturing systems engineering* (second edition). London: Taylor & Francis.

Irani, A.L. (1999). *Handbook of cellular manufacturing systems*. New York: John Wiley & Sons Inc.

Irani, A.L. et al. (1999). Introduction to cellular manufacturing systems. *Handbook of cellular manufacturing systems.* New York: John Wiley & Sons, Inc.

James, T.L. (2007). A hybrid grouping genetic algorithm for the cell formation problem, Computers & Operations Research, 34, 2059–2079.

Joines, J.A. et al. (1996). A comprehensive review of production oriented manufacturing cell formation techniques. *International Journal of flexible Automation and Integrated Manufacturing*, 3(3/4), 161-200.

Kattan, T. (2007). Cell formation problem - A Lagrangean relaxation to mathematical programming approach and a linear performance measure, Master thesis. University of Windsor Windsor, Canada.

Kim C. O. (2004). A two-phase heuristic algorithm for cell formation problems considering alternative part routes and machine sequences. *International Journal of Production Research*, 42 (18), 3911–3927.

King, J.R., & Nakornchai, V. (1982). Machine-component group formation in group technology. *International Journal of Production research,* 20(2), 117-133.

Kusiak, A. (1987). The generalized group technology concept. *International Journal of Production Research,* 25(4), 561-569.

Lee, H., & Garcia-Diaz, A. (1993). A network flow approach to solve clustering problems in group technology, *International Journal of Production Research*, 31(3), 603-612.

Lei, D. & Wu, Z. (2005). Tabu search approach based on a similarity coefficient for cell formation in generalized group technology, *International journal of Production Research,* 43(19), 4035-4047.

Lei, D. & Wu, Z. (2006). Tabu search for multiple-criteria manufacturing cell design, *International journal of Advance Manufacturing technology,* 28, 950-956.

Li, S. (2009). A revised two-phase method for decomposition of design problems, International Conference on Engineering Design, ICED'09. Stanford-California.

Li, S. (2010). Functional decomposition of the clustering approach for matrix-based structuring. *CD-ROM Proceedings of ASME Design Engineering Technical Conferences*, DETC2010-28768, Montreal-Quebec.

Li, S. (2011). A matrix-based clustering approach for the decomposition of design problems. *Research in Engineering Design*, Online First, (15 April 2011).

Liu, C. (2010) A heuristic algorithm for cell formation problem with consideration of multiple production factors, *International Journal of Advanced Manufacturing Technology,* 46, 1201-1213.

Logendran, R. & Ramakrishna, P. (1995). Manufacturing cell formation in the presence of lot splitting and multiple units of the same machine, *International Journal of Production Research,* 33(3), 675-693.

Lozano, S. et al. (1999). A one-step tabu search algorithm for manufacturing cell design, *Journal of the operational Research Society,* 50, 509-516.

Mahdavi, I., & Mahadevan, B. (2008). CLASS: An algorithm for cellular manufacturing system and layout design using sequence data, *Robotics and Computer-Integrated Manufacturing*, 24, 488-497.

McAuley, J. (1972). Machine grouping for efficient production, *Production Engineer*, (2), 53-57.

Miltenberg, J. & Zhang, W (1991). A comparative evaluation of nine well-known algorithms for solving the cell formation problem in group technology, *Journal of Operations management – special issue on Group Technology and Cellular manufacturing,*10(1), 44- 72.

Mosier, C.T. (1989). An experiment investigating the application of clustering procedures and similarity coefficients to the GT machine cell formation problem, *International Journal of Production Research*, 27(10), 1811-1835.

Nair G. J. & Narendran T.T. (1998). Case: A clustering algorithm for cell formation with sequence data, *International Journal of Production Research,* 36(1), 157-179.

Narendran, T.T., & Srinivasan, G. (1999). Machine-Component matrix clustering for cell formation, *Handbook of Cellular Manufacturing*, John Wiley & Sons, Inc.

Offodile, F.O. (1994). Cellular manufacturing: A taxonomic review framework. *Journal of Manufacturing Systems*, 13(3), 1994, 196-220.

Onwubolu, G.C. & Mutingi, M. (2001). A genetic algorithm approach to cellular manufacturing systems, *Computers and industrial Engineering,* 39(1-2), 125-144.

Park, S., & Suresh, N.C. (2003). Performance of Fuzzy ART neural network and hierarchical clustering for part-machine grouping based on operation sequences. *International Journal of Production Research*, 41(14), 3185-3216.

Papaioannou, G., & Wilson, J.M. (2011) The evolution of cell formation problem methodologies based on recent studies (1997-2008): review and directions for future research. *European Journal of Operational Research*, 206, 509-521.

Rajamani, D. et al. (1996). Design of cellular manufacturing systems, *International Journal of Production research,* 34(7), 1917-1928.

Ranson, G.M. (1972). *Group technology*. London, New York: McGraw Hill.

Sarker, B.R., & Mondal, S. (1999). Grouping efficiency measures in cellular manufacturing: a survey and critical review. *International Journal of Production Research*, 37(2), 285-314.

Sarker, B.R., & Xu, Y. (2000). Designing multi-product lines: job routing in cellular manufacturing systems, IIE Transactions, 32(3), 219-235.

Seifoddini, H. (1988). Incorporation of the production volume in machine cells formations in group technology applications, *Recent Developments in Production Research*, Elsevier Science Publishers B.V. Netherlands 562-570.

Seifoddini, H. (1989). A note on the similarity coefficient method and the problem of improper machine assignment in group applications, *International Journal of Production Research*, 72(7), 1161-1165.

Seiffodini, H., & Wolfe, P.M. (1986). Application of the Similarity Coefficient Method in Group Technology. *IIE Transactions*, 18 (3), 271-277.

Selim, H.M. et al. (1998). Cell formation in group technology: review, evaluation and directions for future research, *Computer and Industrial Engineering*, 34(1) 3-20.

Selvam, R.P. & Balasubramanian, K.N. (1985). Algorithmic grouping of operations sequences. *Engineering Cost and Production Economics,* 9, 125-134.

Singh, N. (1993). Design of cellular manufacturing systems: An invited review, *European Journal of Operation Research*. 69(3), 284-291.

Singh, N., & Rajamani D. (1996*). Cellular manufacturing systems: design, planning, and control*. London: Chapman & Hall.

Souilah, A. (1995). Simulated annealing for manufacturing systems layout design, *European Journal of Operational Research,* 82, 592-614.

Srinivasan, G. et al. (1990). An assignment model for the part-families problem in group technology, *International Journal of production Research,* 28, 145-152.

Su, C.T. & Hsu, C.M. (1998). Multi-objective machine-part cell formation through parallel simulated annealing, *International Journal of Production Research,* 36, 2185-2207.

Sule, D.R. (2009). *Manufacturing facilities*(Third edition). Boca Raton, Florida CRC Press.

Kumar, C.S., & Chandraskharan, 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), 233-243.

Vakharia, A.J. (1986). Methods of cell formation in group technology, a framework for evaluation, *Journal of Operations Management – special combined issue*, 6(3), 257-271.

Venugopal, V. & Narendran, T.T. (1992). A genetic algorithm approach to the machine-component grouping problem with multiple objectives, *Computers and Industrial Engineering,* 22(4), 469-480.

Venugopal, V & Narendran, T.T. (1992). Cell formation in manufacturing systems through simulated annealing: An experimental evaluation, *European Journal of Operations research,* 63, 409-422.

Wemmerlöv, U., & Hyer, N. (1986). Procedures for the part family/machine group identification problem in cellular manufacturing. *Journal of Operation Management*, 6, 125-147.

Wemmerlöv, U., & Hyer, N. (2002). *Reorganizing the factory: Competing through cellular manufacturing*. Portland, Oregon, Productivity press.

Won, Y. K. (2000). New p-median approach to cell formation with alternative rout plans. *International Journal of Production Research,* 38, 229-240

Won, Y. K. & Kim, S. H. (1997). Multiple criteria clustering algorithms for solving the group technology problem with multiple process routings. *Computers and Industrial Engineering,* 32, 207-220.

Won, Y. & Lee, K. C. (2004). Modified p-median approach for efficient GT cell formation, *Computers & Industrial Engineering,* 46, 495-510.

Wu, T.H. et al. (2009). Hybrid simulated annealing algorithm with mutation operator to the cell formation problem with alternative process routings, *Expert Systems with Applications,* 36, 3652-3661.

Yin, Y., & Yasuda, K. (2006). Similarity coefficient methods applied to the cell formation problem: a taxonomy and review. *International Journal of Production Economics*, 101(2), 329-352.

Yin, Y., & and Yasuda, K. (2005). Similarity coefficient methods applied to the cell formation problem: a comparative investigation. *Computers & Industrial Engineering*, 48(3), 471-489.

Yu T. L. et al. (2009). Dependency Structure Matrix, Genetic Algorithms, and Effective Recombination, *Evolutionary Computation* 17(4), 595–626.

Appendix

A.1 Tree diagram construction in sorting analysis

In this appendix section the construction of tree diagram in sorting analysis step of matrix-based clustering solution procedure is explained step by step by executing on an example earlier discussed in chapter 3. The method for tree construction in sorting analysis is based on classic hierarchical clustering. The algorithm for tree construction has the following steps (Li, 2011):

Step 1: The corresponding entities to highest value in coupling matrix used for the labeling the tree leaves. Branches can be made with combination of the leaves. The vertical axis is equal to the coupling coefficient value,

Step 2: Updating coupling matrix by combining coupling values of picked entities in previous step. Combining is performed by using average distance formulation,

Step 3: Repeating above two steps until the coupling matrix cannot further reduced.

The algorithm is used to create tree diagram from coupling matrix in figure 1.A. In that coupling matrix, the highest coupling coefficient value is 1 corresponding to parts 1 and 4. These are the first branches of the tree diagram. Coupling matrix is updated by combining the columns of P1 and P4. The new columns known as “P1, P4” and its value equal to average of values in column P1 and P4 in coupling matrix in figure 1.A. The updated matrix is shown in figure A.2.

	Part 1	2	3	4	5	Machine 1	2	3	4
Part 1	0	0	0	1	0.25	0.70	0	0	0.56
2	0	0	0.5	0	0.67	0	0.70	0.56	0
3	0	0.5	0	0	0.33	0	0	0.70	0
4	1	0	0	0	0.25	0.70	0	0	0.56
5	0.25	0.67	0.33	0.25	0	0	0.56	0.47	0.47
Machine 1	0.70	0	0	0.70	0	0	0	0	0.67
2	0	0.70	0	0	0.56	0	0	0.67	0.25
3	0	0.56	0.70	0	0.47	0	0.67	0	0.2
4	0.56	0	0	0.56	0.47	0.67	0.25	0.2	0

Figure A.1 Iteration 1

	Part P1,P4	P2	P3	P5	Machine M1	M2	M3	M4
Part P1,P4	0	0	0	0.25	0.70	0	0	0.56
P2	0	0	0.5	0.67	0	0.70	0.56	0
P3	0	0.5	0	0.33	0	0	0.70	0
P5	0.25	0.67	0.33	0	0	0.56	0.47	0.47
M1	0.70	0	0	0	0	0	0	0.67
M2	0	0.70	0	0.56	0	0	0.67	0.25
M3	0	0.56	0.70	0.47	0	0.67	0	0.2
M4	0.56	0	0	0.47	0.67	0.25	0.2	0

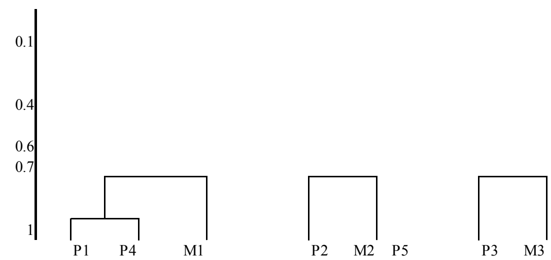

Figure A.2 Iteration 2

In updated coupling matrix in figure A.2, the highest coupling coefficient value is 0.70 which related to three pairs: (P3 and M3), (P2 and M2), and (P1, P4 and M1). These are another level of branches on the previous branch, figure A.2. Again columns of each pair are combined and coupling matrix updated. Result is in figure A.3.

	P1,P4,M1	P2,M2	P3,M3	P5	M4
P1,P4,M1	0	0.35	0	0.125	0.615
P2, M2	0.35	0	0.405	0.615	0.125
P3,M3	0	0.405	0	0.400	0.1
P5	0.25	0.615	0.405	0	0.47
M4	0.56	0.125	0.1	0.47	0

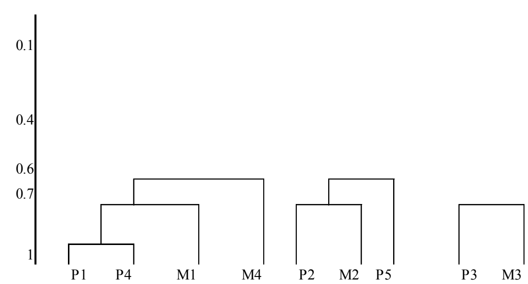

Figure A.3 Iteration 3

In updated coupling matrix in figure A.3, the highest coupling coefficient value is 0.615. This is related to two pairs: (P1, P4, M1 and M4) and (P2, M2 and P5). These are another level of branches on top of previous branches. The coupling matrix updated and showed in figure A.4.

	P1,P4,M1.M4	P2.M2,M5,P5	P3,M3
P1,P4,M1,M4	0	0.307	0.05
P2,P5 M2,M5	0.307	0	0.405
P3,M3	0.05	0.405	0

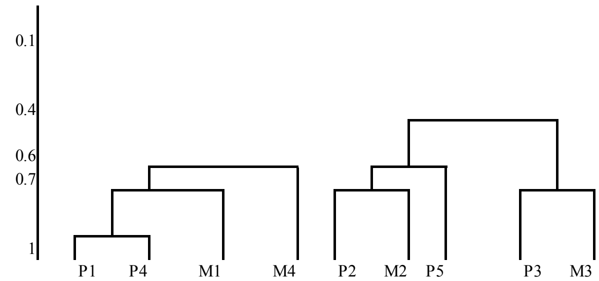

Figure A.4 Iteration 4

In updated coupling matrix in figure A.4, the highest coupling coefficient value is 0.405. This is related to the pair of (P2, M2, M5, P5 and P3, M3). These are another level of branches in the tree diagram. The coupling matrix updated and showed in figure A.5.

	P1,P4,M1.M4,	P2,P5 M2,M5,P3,M3
P1,P4,M1.M4,	0	0.178
P2,P5 M2,M5,P3,M3	0.178	0

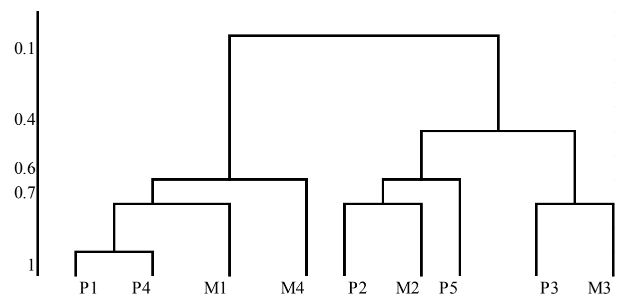

Figure A.5 Iteration 5

The matrix in figure A.5 is the last updated coupling matrix, and it is not possible to further update it. The only remaining pair contains all parts and machines of the incidence matrix. This pair is the highest branch level in the tree diagram. Now tree diagram is completed and sorted matrix can be constructed upon it. According to branches and parts and machines in each branch, rows and columns of incidence matrix can be reordered to get the sorted matrix.

A.2 Alternative coupling coefficients

In chapter 5, it was demonstrated that similarity measures in clustering methods can be used as coupling coefficients in matrix-based clustering. In the following two sections, the similarity measures that were used in chapter five are explained with examples.

A.2.1 Sarker-Xu similarity measure for parts considering production sequence information

Sarker and Xu (2000) used a similarity coefficient to divide part into groups. The only information needed for their similarity coefficient is production sequence. In their methodology, for any two part p and q the part-part similarity coefficient can be calculated by equation 8.

Coupling coefficient part-part = $\frac{N_{pq}^{s}}{N_p}$ (8)

Where

N_{pq}^{s} is the number of *in – sequence operations* of part p to part q

N_p is total number of operations of part p.

According to Sarker and Xu "*in-sequence operations*" are those operations that are in same relative order in the operation sequence of one part with respect to the operation sequence of another part" (Sarker and Xu, 2000). The incidence matrix and production sequence information of examples is shown in figure A.6

Machine \ Part	1	2	3	4	5	6	7
1	1	1	0	1	0	0	1
2	1	1	1	1	0	1	0
3	0	1	0	0	1	1	1
4	0	0	1	1	0	1	0

a) Incidence matrix

Part	1	2	3	4	5	6	7
Machine order	1 2	1 3 2	2 4	1 2 4	3	2 3 4	1 3

b) Production sequence information

Figure A.6 input information for cell formation examples

The similarity coefficients for all part pairs were calculated and results are in figure A.7 For parts 3 and 6 in figure A.6-a, the similarity coefficient is calculated. Part 3 goes through machines 2 and 4. Part 6 visits machines 2, 3 and 4. When the sequence information this two part is compared, it can be seen that these two parts go through same machine with the same sequence, machines 2 and 4. In result, $N_{3,6}^{s}$=$N_{6,3}^{s}$=2. However the number of operation is different:

For 3 and 6, N_{pq}^{s} is 2. N_p is 2. similarity coefficient is equal to 2/2=1

For 6 and 3, N_{pq}^{s} is 2. N_p is 3. similarity coefficient is equal to 2/3= 0.67

		Part						
		1	2	3	4	5	6	7
	1	0	1	0.5	1	0	0.5	0.5
	2	0.667	0	0.333	0.667	0.333	0.667	0.667
	3	0.5	0.5	0	1	0	1	0
Part	4	0.667	0.667	0.667	0	0	0.667	0.333
	5	0	1	0	0	0	1	1
	6	0.333	0.667	0.667	0.667	0.333	0	0.333
	7	0.5	0	0.5	0.5	0.5	0.5	0

Figure A.7 Part-part similarity coefficients

Based on matrix-based clustering rules in chapter three, section 3.2.1, the coupling coefficient of any two entities should be equal for a pair regardless of order. The matrix in figure A.7 is not symmetric. In other words the similarity coefficient of N_{pq} is not equal to N_{qp}. In order to equalize this two value for all pairs the average value of $N_{pq}$and $N_{qp}$is used to represent the part-part coupling coefficients, figure A.8.

		Part						
		1	2	3	4	5	6	7
	1	0	0.834	0.5	0.834	0	0.417	0.5
	2	0.834	0	0.417	0.667	0.667	0.667	0.334
	3	0.5	0.417	0	0.834	0	0.834	0.25
Part	4	0.834	0.667	0.834	0	0	0.667	0.417
	5	0	0.667	0	0	0	0.667	0.75
	6	0.417	0.667	0.834	0.667	0.667	0	0.417
	7	0.5	0.334	0.25	0.417	0.75	0.417	0

Figure A.8 Modified part-part coupling coefficients in a symmetric matrix

A.2.2 Nair and Narendran measure for similarity of machines considering production sequence information

Nair and Narendran (1998) developed a similarity coefficient for solving a cell formation problem with production sequence information. Coupling coefficient of machine i and l are calculated with equation 10.

$$s(i,l) = \frac{c_i + c_l}{t_i + t_l} \tag{9}$$

Numerator is the total common part visits to and from to machines i and l, and denominator are total part visits to machines i and l. The pair of i and l are compared based on visits from each part. Depending on a part if it visits both machines, or visits only one machine or neither of machines, the coupling coefficient varies from one to zero. In addition, the order of operations affects the coupling coefficient value: any common middle operations, if exist, increase the value of the coupling coefficients more than common first and last operations and uncommon operations. If Middle operations are not being carried out in a work cell, they can cause more than two inter-movements in production system. Nair and Narendran defined the components of the formula in following way:

$$c_i = \sum_{j=1}^{n} \sum_{p=1}^{n_{ji}} w_j c_{jip} \tag{10}$$

c_i = Total number of movements to and from machine i by parts which visits machines i and l.

w_j = Weight of part j,

c_{jip} = 0 if the part j does not visit machine i,

c_{jip}=1 if the part j visits both machines either in its first or last operation,

c_{jip}=2 if the part j visits both machines in one of its middle operations (not first or last).

The calculation for c_l, t_i, and t_l is very similar to above calculations. Figure A.9 shows the coupling coefficients that are calculated for the machines of incidence matrix in figure A.6 using above equation.

		Machine 1	Machine 2	Machine 3	Machine 4
Machine	1	0	0.7	0.5	0.28
	2	0.7	0	0.5	0.78
	3	0.5	0.5	0	0.34
	4	0.28	0.78	0.34	0

Figure A.9 machine-machine coupling coefficients