An ant-based clustering algorithm for manufacturing cell design

Y. Kao · S.C. Fu

Received: 17 May 2004 / Accepted: 18 November 2004 / Published online: 21 December 2005
© Springer-Verlag London Limited 2005

Abstract This paper proposes a new part clustering algorithm that uses the concept of ant-based clustering in order to resolve machine cell formation problems. The three-phase algorithm mainly utilizes distributed agents which mimic the way real ants collect similar objects to form meaningful piles. In the first phase, an ant-based clustering model is adopted to form the initial part families. For the purpose of part clustering, a part similarity coefficient is modified and used in the similarity density function of the model. In the second phase, the K-means method is employed in order to achieve a better grouping result. In the third phase, artificial ants are used again to merge the small, refined part families into larger part families in a hierarchical manner. This would increase the flexibility of determining the number of final part families for the factory layout designer. The proposed algorithm has been developed into a software system called the ant-based part clustering system (APCS). In addition to part family formation, APCS performs the tasks of machine assignment and performance evaluation. Finally, performance evaluation of the proposed algorithm was conducted by testing some well-known problems from literature. The evaluation results show that the algorithm is able to solve the cell formation problems effectively.

Keywords Ant algorithms · Cell formation problems · Cellular manufacturing

1 Introduction

Group technology (GT) mainly utilizes part similarities in design or manufacturing characteristics to form part families. Using the principle of GT, a batch production system achieves economies of scale and meets the objective of flexible production. Cell formation is one of the most important applications of GT, as well as being a very crucial component of the cellular manufacturing systems (CMS) design. For each part family, required machines are arranged together to form a machine cell or a manufacturing cell, and most family members can be manufactured within the cell. This kind of layout design problem is referred to as the cell formation problem (CFP).

A CFP requires an effective approach to form part families so that the association can be minimized between part families, while the similarity within a part family can be maximized. There had been a number of solutions proposed for this particular problem. Selim et al. [1] conducted systematic categorizations specifically focused on methods for part or machine grouping; here, the cluster analysis method was most frequently used. In conducting cluster analysis, two aspects should be taken into account: the definition of a similarity coefficient, and the selection of clustering methods.

Various similarity coefficients have been proposed by many researchers to measure the similarities (differences) between parts or machines [2]. Generally, these can be divided into three types: part similarity (difference) coefficients, machine similarity (difference) coefficients, and the association measure of machine and parts [3, 4].

For conventional clustering methods, there are two major categories: hierarchical and non-hierarchical methods [5]. The hierarchical clustering method (e.g., the single linkage method) often accompanies the similarity coefficient method in application. However, it is not easy for the method to obtain a good clustering result when the problem of interference of uneven part data distribution occurs. As for the non-hierarchical clustering method (such as the renowned K-means method), although good clustering quality can be obtained, the number of clusters must be given prior to part clustering. This imposes a challenge to the cellular layout designer; it is difficult to know beforehand what the optimum cell number should be.

To help resolve the above problems, a new part clustering algorithm that adopts the ant colony algorithm has been developed. In this paper, ant-based clustering models are briefly described...
in Sect. 2. In Sect. 3 and Sect. 4, an ant-based part clustering algorithm and its software system are developed, respectively. A numerical illustration is given in Sect. 5. A comparative study and a test of a large-sized problem are conducted in Sect. 6. Finally, conclusions regarding this proposed algorithm are offered in Sect. 7.

## 2 Ant-based clustering models

Observing an ant colony in nature, we can observe that all the food, dead ants, and larvae inside an ant nest will be quickly divided into different piles by the ants. The first ant colony clustering model was proposed by Deneubourg [6]. His model possesses the swarm intelligence of real ants, and was inserted into a robot for the object collecting task. Deneubourg also endowed artificial ants with a memorizing function, which keeps track of the types and quantities of objects that an ant has recently encountered. Hence, the ratio of the number of a given object in the memory buffer to the size of the memory buffer determines the similarity of the object currently in the surroundings. Each artificial ant uses the similarity ratio to determine the probability value of an encountered object being picked up or thrown down. Though it displays the basic functions of ant colony clustering, the Deneubourg model can be applied to gathering only two types of objects.

Lumer and Faiaeta [7] improved upon Deneubaug’s model by adding the Euclidean distance formula to the similarity density function and giving ants three kinds of abilities: speed, short-term memory, and behavior exchange. They named it the LF model. For data clustering, data objects and artificial ants are randomly distributed on a 2D chessboard, and then the artificial ants move randomly on the board to collect similar objects and move dissimilar objects away. Ants decide by themselves whether to pick up an encountered object or lay down a carried object. The decisions are made based on the calculated values of the similarity density function and the probability transfer functions. They proved that the LF model is better than the Deneubourg model, and thus, all subsequent research has been based on the LF model for improvements and applications.

Monmarche et al. [8] combined the LF model with the K-means algorithm for data clustering. Their approach can be divided into four major phases: (1) the LF model is used for initial data clustering; (2) the K-means algorithm is used to purify the initial clustering result; (3) the ant algorithm is used in data pile clustering; and (4) the K-means algorithm is used again to purify the result of pile clustering. They have also proved that the grouping results obtained by this approach are closer to the number of actual clusters and contain fewer errors than simply using K-means or ISODATA alone.

Because artificial ants are characterized by flexibility, randomness, non-centralized control, self-organization, and so forth, ant-based clustering models can be applied to solve complicated problems like part clustering. Therefore, this research applies the LF model to part family formation, and thus helps to resolve the cell formation problem.

### 3 An ant-based part clustering algorithm

This section provides a detailed description of the proposed part clustering algorithm. The proposed algorithm can be divided into three phases: (1) producing part families; (2) refining part families; and (3) combining part families.

The purpose of the first phase is to form initial part families. We first spread artificial ants and parts randomly on a 2D chessboard. The number of grids is described by the formula proposed by Monmarche et al. [8], e.g., \( c^2 = np \times 4 \), where \( c \) represents the length and width of the chessboard, while “np” describes the quantity of parts. It was discovered during the course of the experiment that similar parts will not congregate easily during a clustering process if there are too many artificial ants. Yet, on the other hand, too few artificial ants will lead to a less effective clustering process. Therefore, this research recommends that the acceptable number of artificial ants could be \(["np"/10]\).

Artificial ants perform part clustering by means of calculating the similarity density and probability transfer functions. Each artificial ant uses the similarity density function to judge the similarity of the surrounding parts. The similarity density function of the LF model adopts the Euclidean distance function to measure the similarity of objects. In order to make the LF model more suitable for the part clustering task, a part similarity coefficient replaces the Euclidean distance function, as shown in Eq. 1. In this redefined formula, the similarity coefficient proposed by Islam et al. [9] is adopted and modified, as shown in Eq. 2. The \( k \) value in Eq. 2 can be determined by trial and error, and will not be elaborated on in this paper.

\[
f(P_k) = \frac{\sum s(P_k, P_l)}{nP_l}, P_l \in n^2
\]

\[
s(P_k, P_l) = \frac{(a-k) + \sqrt{(a-k) \times d}}{(a-k) + b + c + d + \sqrt{(a-k) \times d}}
\]

where,

- \( P_k \): part held (or encountered) by an artificial ant
- \( P_l \): part located in the surrounding area (eight grids) of an artificial ant
- \( nP_l \): total number of parts located in eight grids around an artificial ant
- \( f(P_k) \): similarity density function to measure the similarity of a part \( P_k \) to the surroundings
- \( s(P_k, P_l) \): similarity between parts \( k \) and \( l \)
- \( a \): number of machines required by both parts \( k \) and \( l \)
- \( b \): number of machines required only by part \( k \)
- \( c \): number of machines required only by part \( l \)
- \( d \): number of machines not required by either part \( k \) or part \( l \)
- \( k \): adjusting value of an integer, whose range is \( 0 \leq k \leq nm/2 \)
- \( nm \): total number of machines
- \( n^2 \): surrounding area that is recognizable to an artificial ant (e.g., the area of \( n \times n \) grids), as shown in Fig. 1