

FORMING STUDENT GROUPS USING BEES ALGORITHM



By

Adrian C. Ablazo

Eric Jones R. Bonga

Vladimir Christian B. Viscayno

**ATENEOD E DAVAO UNIVERSITY
SCHOOL OF ARTS AND SCIENCES
DAVAO CITY**

September, 2012

FORMING STUDENT GROUPS USING BEES ALGORITHM

A Mini-Thesis

Presented to the

Undergraduate Faculty of the

Computer Studies Division

Ateneo de Davao University

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science in Computer Science

By

Adrian C. Ablazo

Eric Jones R. Bonga

Vladimir Christian B. Viscayno

ATENEO DE DAVAO UNIVERSITY

SCHOOL OF ARTS AND SCIENCES

DAVAO CITY

September, 2012

TABLE OF CONTENTS

RECOMMENDATION FOR ORAL DEFENSE.....	III
RECOMMENDATION FOR ACCEPTANCE.....	IV
ATTESTATION OF PRIMARY AUTHORSHIP	V
ACKNOWLEDGEMENT	VI
TABLE OF CONTENTS	VII
LIST OF FIGURES.....	IX
LIST OF TABLES.....	X
ABSTRACT	XI
Chapter 1 INTRODUCTION.....	1
1.1. Background of the Study.....	1
1.2. Problem Statement	2
1.3. Research Objectives	2
1.4. Significance of the Study	2
1.5. Scope and Limitations.....	3
Chapter 2 REVIEW OF RELATED WORKS	4
2.1. Group Formation in Education.....	4
2.2. Different approaches to Group Formation	4
2.2.1. <i>Genetic Algorithm</i>	5
2.2.2. <i>Fuzzy Logic</i>	5
2.2.3. <i>Ant Colony Optimization</i>	6
2.3. Bees Algorithm.....	6
2.3.1. <i>Tuning a Fuzzy Logic Controller for a Robot Gymnast</i>	6
2.3.2. <i>Optimizing a Support Vector Machine for Wood Defect Classification</i>	7
2.3.3. <i>Manufacturing Cell Formation</i>	7
2.4. Comparison between Bees' Algorithm, Genetic Algorithm and Ant Colony Optimization	7
2.5. Theoretical Framework.....	9
2.5.1. <i>Genetic Algorithm</i>	9
2.5.2. <i>Bees Algorithm</i>	9
Chapter 3 PROJECT DESIGN AND METHODOLOGY	12
3.1. Conceptual Framework.....	12
3.2. Methodology	13
3.2.1. <i>Gather student data</i>	14
3.2.2. <i>Interview teachers</i>	14
3.2.3. <i>Develop the Criteria</i>	14
3.2.4. <i>Objective Function</i>	14
3.2.5. <i>Fitness Function</i>	15

	3.2.6. <i>Implement the BA using an automated group formation program</i>	15
	3.2.7. <i>Testing and Evaluation</i>	15
Chapter 4	THEORETICAL BACKGROUND	16
4.1.	Bees Algorithm	16
	4.1.1. <i>Bees in Nature</i>	16
	4.1.2. <i>BA in Group Formation</i>	17
Chapter 5	RESULTS AND DISCUSSION	19
5.1.	Gathering Student Data	19
	5.1.1. <i>Results of the 3rd year respondents</i>	20
	5.1.2. <i>Results from 4th year respondents</i>	23
5.2.	Interviewing the teachers	27
5.3.	Developing the Criteria	28
	5.3.1. <i>Objective Function</i>	32
	5.3.2. <i>Balanced Grouping</i>	34
5.4.	Implementing the Bees Algorithm	35
5.5.	Testing and Evaluation	38
	5.5.1. <i>CS3 grouping results vs. original grouping results</i>	38
	5.5.2. <i>CS4 grouping results vs. original grouping results</i>	42
Chapter 6	CONCLUSIONS AND RECOMMENDATIONS	45
6.1.	Conclusions	45
6.2.	Recommendations	46
	BIBLIOGRAPHY	47
	APPENDICES	48
	Appendix A: User Manual	48
	Appendix B: Relevant Codes	52

ABSTRACT

In the academe, project based subjects require a lot of energy, time, patience and man power to accomplish project objectives. That is why, student group formation in these subjects are important and critical, because an individual cannot work well if his or her group mates aren't capable of working together as a unit - because of deficiencies, such as lack of skills, training, or interest on the subject area or motivation. Heterogeneity is also important in group formation. This study adapts a mathematical approach which increases goodness heterogeneity of groupings and the Bees Algorithm, to form three-manned teams per class. The results of this study have reaped more heterogeneous groupings as automated further. It is concluded that the study would be used as a decision making tool.

Keywords:

Student Group Formation, Bees Algorithm, Genetic Algorithm, Motivation, Balanced Groupings

Chapter 1

INTRODUCTION

1.1. Background of the Study

The continuous enlargement of industries and institutions and also the increasing complexity of projects have led to the spread of the workload from individuals to several teams. However, there are aspects that need to be addressed in forming groups. First, there is the consideration of factors of an individual: his/her skills, motivations and work ethic. There is also the fact that the personality of the individual, which is arbitrary and imprecise, is ruled out. Second, there is the problem of scalability and complexity. While forming groups out of ten or twenty people might seem feasible at first glance, the complexity of forming it might take more than a few minutes to more than an hour. Increasing the population to fifty or more people would increase the complexity exponentially, thus affecting the task's scalability.

There have been studies regarding group formation which involves the implementation of a certain algorithm or an automated method. One of these is the work of Ani, et. al (2010). They presented a method for group formation using Genetic Algorithm, where members of each group will be generated on a student's programming skill. Another is the work of Graf and Bekele (2002), which talked about a mathematical approach to form heterogeneous groups based on personality traits and the performance of students, using an Ant Colony Optimization algorithm. However, there were no studies using Bees Algorithm in group formation. Thus, the proponents wished to pursue this study in order to open up possibilities for more studies in optimal group formations.

1.2. Problem Statement

This study was conducted to investigate the Bees Algorithm and its implementation in the area of group formation. In addition, we sought to answer these following questions:

- ▲ What are the factors and criteria by which a student is categorized in a group?
- ▲ What are the properties or characteristics of a balanced grouping?
- ▲ How student groupings are formed using the bees algorithm?
- ▲ How can this algorithm be implemented in a program?

1.3. Research Objectives

The main objective in this study was to implement the Bees Algorithm in forming balanced groupings. The study also accomplished the following objectives:

- ▲ To identify the factors and criteria by which a student is categorized in a group.
- ▲ To identify the properties or characteristics of a balanced grouping.
- ▲ To explain how the student groupings are formed using the bees algorithm.
- ▲ To demonstrate the implementation of the Bees Algorithm in an automated group formation program.

1.4. Significance of the Study

The output of this study can initially be used by the undergraduate junior classes of Ateneo de Davao University's Computer Studies division. These classes have major subjects (like Systems Analysis and Design and Thesis) that have projects which require the flexibility of a well-balanced

student group to get the job done. Afterwards, other divisions and perhaps other schools can use the expected output to further aid their students in their major projects.

This study may serve as one of the first researches on the area of Bees Algorithm. This study will make way for more researchers in the study of group formation using other, more effective approaches.

1.5. Scope and Limitations

The study is conducted to figure out how group formation can be done using Bees Algorithm, in the Computer Studies (CS) cluster of the Ateneo de Davao University (AdDU).

The study will not be geared towards a development of an implementation that would cater to other clusters of AdDU, because the data that the system would be gathering would be the past grades of students in their pre-requisite subjects.

This study covers the implementation of the Bees Algorithm in group formation. It will cover the implementation in computer code, testing for optimal solutions, and testing for scalability.

It will only include factors which are precise, such as skill – through previous grades) and work ethic. Personality traits are not included in this study.

Chapter 2

REVIEW OF RELATED WORKS

2.1. Group Formation in Education

Collaboration has long been considered an effective approach to learning. However, forming optimal groups can be a time-consuming and complex task. Different approaches have been developed to assist teachers allocate students to groups based on a set of constraints. (Ounnas, et. al. 2009)

Forming effective student groups in the field of education has become quite a challenge for teachers nowadays. That is why there had been different methods and guidelines on group formations. But despite these methods, the results are not too promising. “Orphan students” among formed groups are inevitable.

According to Graf and Bekele (2002), students should work in small, mixed-ability groups of four members. A balanced group should contain: one high achiever, two average achievers, and one low achiever.

Group formation involves the shaping of groups containing a number of individuals in such a way that a group is considered “balanced”. Individuals with heterogeneous levels of motivation or skill on a subject area constitute a “balanced” group. A balanced grouping of individuals is important in huge projects, because these groupings are able to get the job done faster – unlike an imbalanced group or a single individual, working in a project.

2.2. Different approaches to Group Formation

Although there are no other papers specifically about Group formation using Bees Algorithm, there are related works with regards to group formation using *a certain algorithm or automated method*.

2.2.1. Genetic Algorithm

One of such is the work of Ani, et. al (2010) where they presented a method for group formation using genetic algorithm. In their study, members of each group were generated based on the students' programming skill. Genetic algorithm begins with an initial population represented by *chromosomes*. It can be taken and applied to a new population. The expectation is that the new population will be better than the old one. The offspring was selected according to the degree of fitness – in their case – according to the students' programming skill.

2.2.2. Fuzzy Logic

Another work related to group formation is by Strnad, et. al (2000), where they presented a new fuzzy-general analytical model for the problem of the project team formation. It builds on previous quantitative approaches, but adds several modeling enhancements like derivation of personnel attributes from dynamic quantitative data, complex attribute modelling, and handling of necessary over-competency. The flexibility of the requirements specification is improved using a special format that expresses the required team capabilities using fuzzy descriptors. A single compound objective function is then defined, which incorporates multiple opposing criteria that the solution should maximize. A special adaptation of the island genetic algorithm with mixed crossover was made in order to optimize the selection of multiple project teams with possible conflicting requirements.

2.2.3. Ant Colony Optimization

Graf, et. al (2002) presented their paper related to group formation, which talked about a mathematical approach to form heterogeneous groups based on personality traits and the performance of students. Using an Ant Colony Optimization (ACO) algorithm the heterogeneity of formed groups was maximized. ACO is a multi-agent meta-heuristic for solving NP-hard combinatorial optimization problems. The algorithm delivered stable solutions which are close to optimum for different datasets of 100 students. An experiment with 512 students was also performed demonstrating the scalability of the said algorithm.

2.3. Bees Algorithm

The objective of the Bees Algorithm is to search and explore good sites within a given problem space. Scout bees are sent out to randomly sample the problem space and locate good sites. A local search is continually applied through these good sites, although many scout bees are sent out each iteration, to search for additional good sites. (Brownlee, J., 2011)

Although there are no other papers specifically about group formation using Bees Algorithm, there are applications using or implementing the Bees Algorithm.

2.3.1. Tuning a Fuzzy Logic Controller for a Robot Gymnast

Eldukhri, et. al. (2007) used the Bees Algorithm to tune the parameters of a Fuzzy Logic Controller (FLC) developed to stabilize and balance an under-actuated two-link acrobatic robot in the upright position. A linear quadratic regulator was developed to obtain the scaling

gains needed to design the fuzzy logic controller. Simulation results confirmed that using BA to optimize the membership functions and the scaling gains of the fuzzy system improved the controller performance.

2.3.2. Optimizing a Support Vector Machine for Wood Defect Classification

Muhamad, et. al (2007), presented their paper, which describes the application of the Bees Algorithm to the optimization of a Support Vector Machine (SVM) for the problem of classifying defects in plywood. The algorithm, which is a swarm-based algorithm inspired by the food foraging behaviour of honey bees, was also employed to select the components making up the feature vectors to be presented to the SVM. The objective of the work was to find the best combination of SVM parameters and data features to maximize defect classification accuracy.

2.3.3. Manufacturing Cell Formation

Cellular manufacturing has received increasing attention in years. The key problem in designing manufacturing systems is cell formation, which as presented by Afify, et. al. (2007) in their paper, is concerned with grouping parts with similar processing requirements into part families and associated machines into machine cells. Their paper solves the cell formation problem using the Bees Algorithm.

2.4. Comparison between Bees' Algorithm, Genetic Algorithm and Ant Colony Optimization

Swarm intelligence is based on nature-inspired behaviour and is applied to optimization problems in various fields. The goal of optimization is to find the optimum in the smallest amount of iterations, where optimum means the best

from all possibilities chosen from a particular criterion. Pham, et al. (2006), presented a comparison of the Bees Algorithm, Genetic Algorithm and the Ant Colony Optimization, using the following seven benchmark functions: De Jong, Goldstein & Price, Branin, Martin & Gaddy, Rosenblock, Hyper Sphere, and Griewangk.

Function	Genetic Algorithm		Ant Colony Optimization		Bees Algorithm	
	Number of Iterations	Success %	Number of Iterations	Success %	Number of Iterations	Success %
De Jong	10160	100	6000	100	868	100
Goldstein & Price	5662	100	5330	100	999	100
Branin	7325	100	1936	100	1657	100
Martin & Gaddy	2844	100	1688	100	526	100
Rosenblock	10212	100	6842	100	631	100
Hyper sphere	15468	100	22050	100	7113	100
Griewangk	200000	100	50000	100	1847	100

Table 1: A Comparison between GA, ACO and BA.

The first test function was De Jong's, for which BA could find the optimum 120 times faster than ACO and 207 times faster than GA. The second function was Goldstein and Price's, for which BA reached the optimum almost 5 times faster than ACO and GA. With Branin's function, BA had a 15% improvement compared with ACO and 77% improvement compared with GA. In Ronsenbrock's function in two dimensions, BA has good improvement over the two other methods. In the Hyper sphere model of six dimensions, BA needed half of the number of iterations compared with GA and one-third of that required for ACO. And lastly, in Griewangk's 10-dimensional function, BA could reach the optimum 10 times faster than GA and 25 times faster than ACO. All these functions had a success rate of 100%. Given these data, it can be concluded that Bees Algorithm is much more practical and faster in attaining an optimum compared to Genetic Algorithm and Ant Colony Optimization.

2.5. Theoretical Framework

2.5.1. Genetic Algorithm

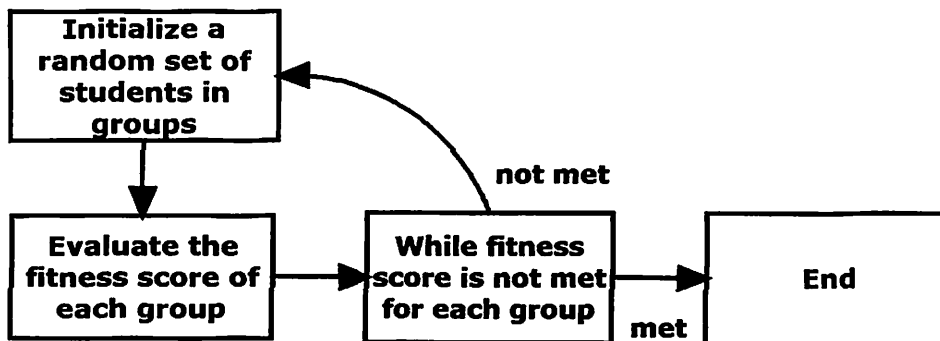


Figure 1: Genetic Algorithm Framework

Ani et. al (2010) came up with this framework which shows the processes of the Genetic Algorithm (GA) in the student group formation problem. It basically had three steps which are: *initialize a random set of students in groups*, *evaluate the fitness score of each group*, and if the fitness score of each group isn't high enough – *repeat the entire process – until an optimal solution is generated*.

In their study, they presented GA in such a way that it begins with an initial group represented by *chromosomes*. A *chromosome* is a set of solution from one group. It can be taken and applied to a new group. The expectation is that the new group or *offspring* will be better than the old one. The *offspring* will be selected according to the degree of fitness. This solution is repeated until the best possible solution is achieved.

2.5.2. Bees Algorithm

The Bees Algorithm (BA) is an optimization algorithm inspired by the natural foraging behaviour of honey bees to find the optimal solution. The algorithm requires a number of parameters to be set, namely: number of scout bees (n), number of sites selected out of n visited sites (m),

number of best sites out of m selected sites (e), number of bees recruited for best e sites (nep), number of bees recruited for the other ($m-e$) selected sites (nsp), initial size of patches (ng_h) which includes site and its neighbourhood and stopping criterion.

1. Initialise population with random solutions.
2. Evaluate fitness of the population.
3. While (stopping criterion not met)
//Forming new population.
4. Select sites for neighbourhood search.
5. Recruit bees for selected sites (more bees for best e sites) and evaluate fitnesses.
6. Select the fittest bee from each patch.
7. Assign remaining bees to search randomly and evaluate their fitnesses.
8. End While.

The algorithm starts with the n scout bees being placed randomly in the search space. The fitnesses of the sites visited by the scout bees are evaluated in step 2. In step 4, bees that have the highest fitnesses are chosen as "selected bees" and sites visited by them are chosen for neighbourhood search. Then, in steps 5 and 6, the algorithm conducts searches in the neighbourhood of the selected sites, assigning more bees to search near to the best e sites. The bees can be chosen directly according to the fitnesses associated with the sites they are visiting. Alternatively, the fitness values are used to determine the probability of the bees being

selected. Searches in the neighbourhood of the best e sites which represent more promising solutions are made more detailed by recruiting more bees to follow them than the other selected bees. Together with scouting, this differential recruitment is a key operation of the Bees Algorithm.

However, in step 6, for each patch only the bee with the highest fitness will be selected to form the next bee population. In nature, there is no such a restriction. This restriction is introduced here to reduce the number of points to be explored. In step 7, the remaining bees in the population are assigned randomly around the search space scouting for new potential solutions. These steps are repeated until a stopping criterion is met. At the end of each iteration, the colony will have two parts to its new population representatives from each selected patch and other scout bees assigned to conduct random searches (Pham et al., 2006).