

# A Genetic-algorithm Approach for Balancing Learning Styles and Academic Attributes in Heterogeneous Grouping of Students

<https://doi.org/10.3991/ijet.v12i03.5803>

Anon Sukstrienwong  
Bangkok University, Bangkok, Thailand  
anon.su@bu.ac.th

**Abstract**—Cooperative learning is an instructional approach in which students work together in small groups in order to achieve a common academic goal. In the context of cooperative learning, students in classrooms tend to learn more by sharing their experiences and knowledge. In addition, a diversity of educational backgrounds and student learning styles can be used to build heterogeneous groups of students. In this paper, we propose an approach for the group composition, regarding the index of learning styles (ILS) questionnaire and prior educational knowledge in order to achieve the mechanism for equity among groups and ensure that heterogeneous students are distributed optimally within the group formation. This causes the search for an optimized group composition of all students to become a more complex and becomes a time-consuming task. Therefore, the proposed algorithm mimics the natural process of a genetic algorithm in order to achieve optimal solutions. In addition, we have implemented our algorithm to construct student groups. A case study shows that the algorithm enhances the quality of the group formation of heterogeneous students leading to better solutions.

**Keywords**—Heterogeneous grouping, genetic algorithms, group formation, index of learning styles

## 1 Introduction

Presently, the cooperative learning among students in groups is supported by several researchers because it is an educational technique used to manage learning activities in classrooms. The activities depend on the learning goals of the class in question. Cooperative learning involves groups of students working together to achieve assigned tasks. Felder and Brent suggested that cooperative learning can be used for any type of assignment within lecture classes, laboratories, or project-based courses [1]. In collaborative learning, members within a group must work together in order to collectively learn something, which is different from cooperative learning. This is because peers in cooperative learning divide the assignment into small tasks and are likely to solve these individually before combining their results in order to form final result. On the other hand, in collaborative learning members tend to interact and work to-

gether to complete their assignments. Previous studies related to collaborative learning have also emphasized the importance of heterogeneity for performance. Despite the differences between cooperative learning and collaborate learning, Razmerita argued that some educators use the two terms interchangeably [2].

Some studies have focused on improving cooperative learning for students, including the work by Balmeceida, Schiaffino, and Pace [3]. In this article, they claimed that different characteristics of group members might influence the group performance. Therefore, they have considered three characteristics of students to form groups: psychological styles, team roles, and social networks. Paredes, Ortigosa and Rodriguez [4] proposed a technique for integrating grouping mechanisms through each student's learning style. In this research, all students were required to answer a questionnaire in order to determine their learning style before generating the group formation. Martin and Paredes have argued that homogeneous groups are better at achieving specific objectives. However, when we constructed groups according to the students' educational backgrounds, experiences, and attitudes, the groups formed became heterogeneous [5]. Aller [6] discussed the ILS questionnaire formulated by Richard M. Felder and Linda K. Silverman. Presently, the detail of this questionnaire is available online [7]. Aller also argued the factors contributing to the success in an introductory computer science course. Currently, there are many researchers investigating the relationship between learning style and student performance in the field of computer science [8]. Also, Chen and Lin conducted an experiment to explore the relationship between learning styles based on ILS. Felder and Silverman [9] reported that students prefer to learn in various ways, for instance, by memorizing, hearing, seeking, reasoning logically, building a mathematical model and drawing analysis. Additionally, JuHou and Hu [10] proposed that the diversity of learner characteristics plays a critical role within the process of achieving learning goals.

Group formation is considered as an effective approach in promoting the academic success of cooperative learning [11-13]. However, in the real world, there are various factors that must be considered to construct heterogeneous groups of students. Factors based on class objectives can include: student attributes such as attitude, grade point average: current classes, and previous academic background. Therefore, the search for an optimized group composition of all students is a time-consuming task. The problem becomes even more complex. Moreover, it seems that consideration of the learning styles are usually ignored in forming heterogeneous groups of students. Besides, some researchers documented that learning style-based group formation helps promoting heterogeneity groups and has a beneficial impact on learning in cooperative learning environments, as seen in [14,15]. These reasons have given way to the proposal of a genetic-algorithm approach to form student groups by using both student learning styles and student educational background as the main criteria to allocate students in group formation.

The paper is divided into seven parts including this introductory section. The remainder is organized as follows: A general idea of heterogeneous grouping is presented in section two. The index of learning styles classified by the Felder and Silverman model, which is used in this paper, is briefly described in section three. Section four briefly demonstrates genetic algorithms (GAs). Section five provides details of the

proposed algorithm based on GAs and the mechanism for generating student groups. Section six demonstrates the design of our developed software named Genetic Algorithm for Forming Student Groups (GAFSG), including the user interface. Additionally, the case study is described. The results obtained are reported. The discussion related to the experiment is presented in this section as well. Finally, section seven includes conclusions and future work.

## **2 Heterogeneous Grouping**

Essentially, cooperative learning is heterogeneous in membership especially in terms of achievement motivation and task orientation [16]. At present, much literature exists studying on the advantage of building heterogeneous groupings in classrooms. In most works, researchers support that heterogeneous grouping is the practice of teaching students of different ages and ability levels together in the same classroom where between two and five students of varying abilities learn together. Some researchers confirm that the classroom employing heterogeneous grouping has a positive effect on students by improving their attitudes towards each other and their school work. Thousand, Villa and Nevin [17] stated that one of the principles of cooperative learning is the heterogeneous grouping of students. The key assumption is that a group works better when peers are balanced in terms of diversity, functional roles, and personality differences. Moreover, one important aspect of constructing cooperative learning groups within small groups is the maximization of the heterogeneity of the students. Therefore, students should be placed in groups that are mixed in terms of academic abilities and aptitudes, personality, social class, religion, language proficiency, race, and sex. As a consequence, heterogeneous grouping is also termed 'mixed ability' or 'achievement' grouping. Several advantages of heterogeneous grouping have been demonstrated in the surrounding literatures, such as the fact that it improves the academic achievement of students [18]. Additionally, it is argued by Jacobs [18] that heterogeneous grouping in elementary grade levels has produced academic and social benefits for high-ability, average-ability, and low-ability learners. According to the research bulletin of North Carolina Middle School, heterogeneous classrooms maximize greater learning opportunities for low-ability students, without being detrimental to high-ability students [19].

## **3 The Index of Learning Styles**

The learning styles of each student were classified by the Felder and Silverman model [8,20] and collected using the Index of Learning Styles (ILS) questionnaire before the group tasks began. Felder and Silverman proposed four dimensions of learning styles based on the preferred manner in which a learner constructs new knowledge for themselves. Some researchers have proven the reliability validity of the questionnaire such as Felder and Spurlin[21] and Litzinger et al. [22]. Additionally, Tanner and Allen [23] summarized the definition as follows: active learners are adept at administration and team-oriented project work; reflective learners do well at

individual research and design. Sequential learners always have good analytical skills to solve convergence (single answer) problem; global learners are always good synthesizers, able to draw material from multiple disciplines to solve problems that cannot be resolved in a conventional single-discipline approach.

The ILS questionnaire relates to students' preferred models for receiving information which is composed of 44 questions with two possible answers: 'a' or 'b'. The questionnaire is used to assess preferences on four dimensions of a learning style model, namely active/reflective (the process by which they receive it); sensing/intuitive (the type of information they receive); visual/verbal (the modality in which they receive it) and sequential/global (the order in which they receive it). As For each dimension, there are exactly two categories, 'a' and 'b'. If the total score on the first dimension was 3 for active, the ILS score for this dimension is 3a. Also, if total score was 8 for reflective, the score will be written as 8b. Since b is letter of larger ( $8 > 3$ ), the result on the active/reflective dimension is 5b. If one student has 1a to 3a for the active/reflective, it means the student has a mild preference for active rather than reflective learning. If the score is higher –for example, which is between 5a to 7a– the student has a moderate preference for active learning. Finally, if the score is between 9a and 11a, one can conclude that the student has a strong preference for active learning.

## 4 Genetic Algorithms

Genetic algorithms (GAs) are inspired by the principles of Darwin's theory of evolution and are often used as evolutionary computational models for various fields of study [24]. Presently, GAs are recognized as a very powerful tool used in optimization and have been applied in various fields, including: computer science, engineering, education, and Stock market data mining optimization [24]. They encode a potential solution to a problem in a chromosome-like data structure, generally an array. Therefore, each individual is understood to be a feasible solution. A general design of a GA is presented in Figure 1.

For the initial population, GAs begin with generation 0 (Gen=0) with a random population. All individuals are evaluated by a fitness function. Three common operators of GAs – reproduction, crossover, and mutation –perform during one generation in order to generate new offspring. The objective of such operators is to preserve the chromosomes, or a part of them, which represent better solutions following natural selection principles [25]. Typically, the reproduction operator is more likely to select individuals to form the next generation of the algorithm by a fitness function. The crossover operator combines two chromosomes (parents) to generate to produce new chromosomes (offspring). There are many ways to implement it in GAs, such as single point crossover, two-point crossover, and heuristic crossover [26].

A common mutation operator involves a small change in a bit of a chromosome in order to ensure that the successive individual is not exactly the same. Mutation operator is a divergence operation, since it generates new offspring that helps to find new solutions. It is intended to break more members of a population out of a local minimum/maximum space. However, two basic parameters involved in GAs are crossover probability ( $p_c$ ) and mutation probability ( $p_m$ ). Basically, these probabilities determine the degree of solution accuracy and the convergence speed that GAs can obtain.

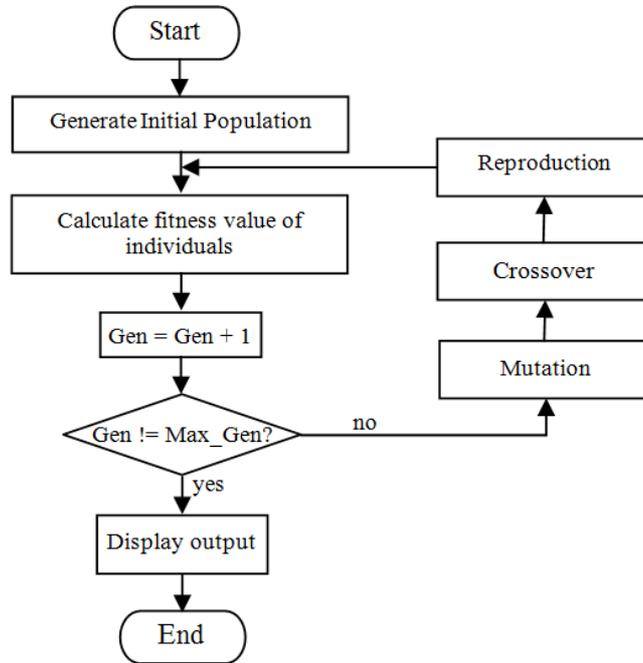


Fig. 1. General design of GAs.

Crossover probability indicates a ratio of how many couples of a population will be picked for mating. If  $p_c$  is 1.0 then all offspring is made by crossover. If it  $p_c$  is 0.0, a whole new generation is made from exact copies of chromosomes from the previous population. The value of  $p_m$  is usually between 0.005 - 0.1, since large values of  $p_m$  might transform GAs into a random search algorithm [27]. If a termination criterion is not met ( $Gen \neq Max\_Gen$ ), the common operators will be applied and move to the next generation ( $Gen = Gen + 1$ ). As GAs are a kind of heuristic search algorithm that are shown to be structurally similar, the best individual based on the fitness value in the last generation may simply approximate the solution to the problem. However, we can modify many different variations to improve efficiency such as two-point crossover, sizing initial population, and parallel genetic algorithms.

## 5 The Proposed Algorithm For Heterogeneous Grouping

In this section, we present an approach for balancing heterogeneous groups of students based on a diversity of educational backgrounds and student learning styles in the context of cooperative learning styles. Therefore, the main objective of the proposed algorithm aims to improve the quality of group composition. Euclidean distance between generated groups is deployed in the optimization process, therefore the group formation problem becomes even more complex. To achieve this aim, the objectives of the research are described below:

- To propose a new approach for a coalition of heterogeneous grouping related to the Index of Learning Style (ILS);
- To model a system that implements the proposed approach for group formation to achieve an equality of group attributes among the established groups;
- To facilitate teachers for generating the heterogeneous groups of students and provide an opportunity for students to learn from dissimilar peers.

It is worth emphasizing that the main goal of this paper is to help teachers form heterogeneous groups of students. Nevertheless, information has been gathered about the learning styles of students in order to balance heterogeneous students within the group formation in terms of learning styles and educational backgrounds. Therefore, a web-based application has been developed in order to meet this purpose.

### 5.1 The ILS Score

The ILS score is set as one of the major attributes in the proposed algorithm. Each individual can be allocated to one group related on the ILS score that the student made before the group process began. As mentioned previously, there are four dimensions of the ILS. In addition, there are two opposite categories of each dimension. Hence, we need to convert the value of ‘a’ to be a negative value and ‘b’ to be a positive value. For example, 2a is equal to -2 and 2b is equal to +2.

### 5.2 Problem Encapsulation

A class contains  $n$  students, denoted by  $S = (s_1, s_2, \dots, s_n)$ . Each student has its attribute represented in a  $k$ -dimensional vector, in the form of a vector field  $(v_1, v_2, \dots, v_k)$ . And, student  $i$  in  $k$ -dimensional space is represented as a vector field  $(v_1^i, v_2^i, \dots, v_k^i)$ . In this paper, the student attributes refer to course backgrounds, prerequisite courses, experiences and learning styles.

To apply a genetic algorithm, we encode the problem as a chromosome. Since there are  $n$  students to form student groups, the chromosome is composed of  $n$  fixed characters. Each element of the chromosome represents the group that the students belong to. If  $n$  students are divided into  $p$  groups, the characters of each chromosome can be represented in the form  $G_i$  where  $0 \leq i \leq p$ . For instance, if there are 16 students to form three groups ( $p=3$ ), the chromosome length will be equal to 16. The possible characters for each chromosome are  $G_1, G_2, \text{ or } G_3$ . Supposing a chromosome represents the solution of group formation where  $G_1 = \{s_2, s_8, s_{11}, s_{14}, s_{15}\}$ ;  $G_2 = \{s_4, s_6, s_9, s_{12}, s_{13}, s_{16}\}$  and  $G_3 = \{s_1, s_3, s_5, s_7, s_{10}\}$ , the chromosome named chromosome<sub>1</sub> can encode this formation as represented in Figure 2:

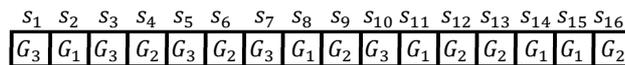


Fig. 2. An example chromosome ( $n=16$ ), where  $G_i$  representing the group  $i$  and  $0 \leq i \leq 3$ .

### 5.3 Quality of the Group formation

To achieve the mechanism for equity of academic attributes among heterogeneous groups, a particular fitness function is implemented. Due to the design of this algorithm, all students will be assigned into a certain group. Additionally, all attributes of a group are calculated by the average value of all students in the same group. Student  $i$  in  $k$ -dimensional space is represented as field  $(v_1^i, v_2^i, \dots, v_k^i)$ , where  $k$  is the number of attributes. Then, we can calculate the average value of all attributes from members. Therefore, a vector field of student group  $G_j$  is represented as  $G_j = (V_1^j, V_2^j, \dots, V_k^j)$ , where  $V_1^j = \frac{\sum_{i=1}^m v_1^i}{m}$ ,  $V_2^j = \frac{\sum_{i=1}^m v_2^i}{m}$ ,  $\dots$ ,  $V_k^j = \frac{\sum_{i=1}^m v_k^i}{m}$ ,  $m$  is the group size,  $p$  is the number of student groups, and  $1 \leq j \leq p$ . In the group formation, each student is allocated to only one group. Keep in mind that we try to generate student groups equally. If we divide  $n$  students into  $p$  groups, then each group will have  $m = \lfloor \frac{n}{p} \rfloor$  members. That is, in some cases, some groups may have fewer than  $\lfloor \frac{n}{p} \rfloor$  students. For instance, if a set of ten students is divided into three groups, the size of each group is  $\lfloor \frac{10}{3} \rfloor = 3$ . Certainly, one group of students must be  $\lfloor \frac{10}{3} \rfloor + 1 = 4$ .

The ultimate goal of the proposed algorithm is to minimize the distance among divided groups. If one chromosome forms student groups which are composed of  $p$  groups, namely  $G_1, G_2, \dots, G_p$ , the Euclidean Distance (ED) between any two groups can be calculated. As a result, the quality of the group formation becomes  $\sum_i^{p-1} \sum_{j=i+1}^p |G_i - G_j|$ . The big picture for the group formation is shown in Figure 3.

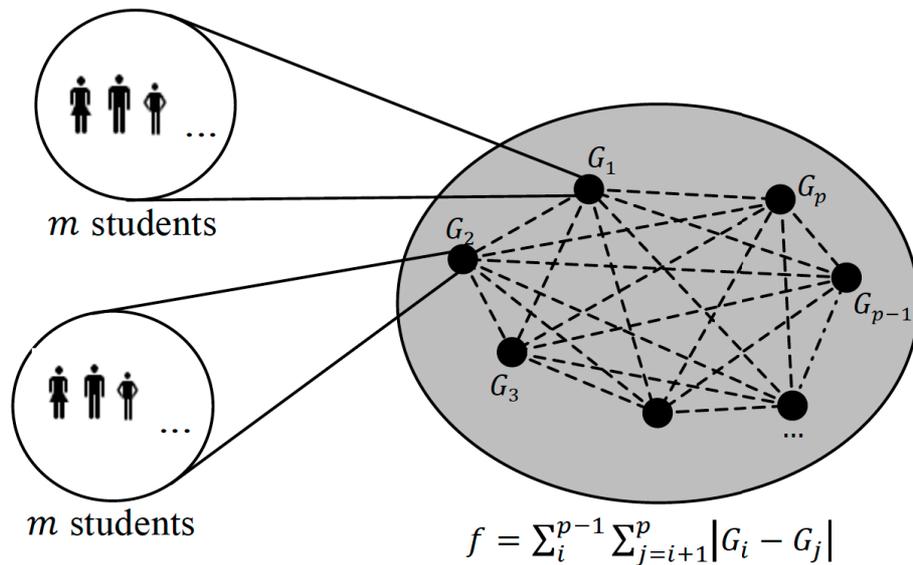


Fig. 3. Representation of the formation of  $p$  groups where group size =  $m$ .

The fitness function of the chromosome is expressed by equation (1).

$$\begin{aligned}
 f(\text{chromosome}) &= \sum_i^{p-1} \sum_{j=i+1}^p |G_i - G_j| \\
 &= \sum_i^{p-1} \sum_{j=i+1}^p \sqrt{\sum_{n=1}^k |V_n^i - V_n^j|^2}
 \end{aligned}
 \tag{1}$$

, where  $p$  denotes the number of student groups and  $k$  is the number of attributes.

Suppose 2-attribute vector of  $S_i$  is  $(v_1^i, v_2^i)$  and all vectors of 16 students, which are associated with the chromosome<sub>1</sub> in Figure 2, are presented in Table 1.

**Table 1.** 16 Student and 2-attribute vector

Student	2-attribute vector
$s_1$	(3.0,2.5)
$s_2$	(3.5,2.0)
$s_3$	(0.5,1.0)
$s_4$	(2.0,2.0)
$s_5$	(1.5,3.0)
$s_6$	(0.5,4.0)
$s_7$	(1.1,2.0)
$s_8$	(3.4,3.7)
$s_9$	(1.5,2.6)
$s_{10}$	(4.0,3.9)
$s_{11}$	(3.1,1.8)
$s_{12}$	(1.8,1.0)
$s_{13}$	(0.5,3.9)
$s_{14}$	(2.7,1.2)
$s_{15}$	(2.1,2.4)
$s_{16}$	(3.6,3.2)

Each attribute of  $G_i$ , where  $0 \leq i \leq 3$ , is calculated by the average value of all members in the group as demonstrated below:

$$\begin{aligned}
 G_1 &= \left( \frac{3.5+3.4+3.1+2.7+2.1}{5}, \frac{2.0+3.7+1.8+1.2+2.4}{5} \right) \\
 &= (2.96, 2.22) \\
 G_2 &= \left( \frac{2.0+0.5+1.5+1.8+0.5+3.6}{6}, \frac{2.0+4.0+2.6+1.0+3.9+3.2}{6} \right) \\
 &= (1.65, 2.78) \\
 G_3 &= \left( \frac{3.0+0.5+1.5+1.1+4.0}{5}, \frac{2.5+1.0+3.0+2.0+3.9}{5} \right) \\
 &= (2.02, 2.48).
 \end{aligned}$$

When all groups' vectors are identified, the fitness value for chromosome<sub>1</sub> can be evaluated by equation (1), which is equal to 2.87631 as presented below:

$$f(\text{chromosome}_1) = \sum_i^2 \sum_{j=i+1}^3 \sqrt{\sum_{n=1}^2 |V_n^i - V_n^j|^2}$$

$$\begin{aligned}
 &= \sqrt{(2.96 - 1.65)^2 + (2.22 - 2.78)^2} + \\
 &\sqrt{(2.96 - 2.02)^2 + (2.22 - 2.48)^2} + \\
 &\sqrt{(1.65 - 2.02)^2 + (2.78 - 2.48)^2} \\
 &= 2.88.
 \end{aligned}$$

#### 5.4 Single-Point Crossover Operator

For the initial population of Gen = 0, the algorithm begins by randomly initializing the characters of each chromosome. One important operator used in generating offspring is a single-point crossover. Two parents are randomly picked to generate two new offspring. A crossover point in the chromosome is randomly selected. This is the point where the parents can exchange part of the chromosomes. The first offspring has a portion of string from the beginning of the first chromosomes to the crossover point. In addition, the rest is copied from the second parent from the crossover point to the end. The second offspring operates in the opposite way. Therefore, the new offspring are alike to the parents.

In this paper, single-point crossover is used to generate offspring. Two chromosomes are randomly selected from the current generation as parents, namely parent1 and parent2. One crossover point between 1 and chromosome length n is randomly selected. Additionally, at each time of crossover, two offspring will be generated.

To understand how the operator works, an example is illustrated in Figure 4. The first parent represents the group formation where  $G_1 = \{s_1, s_5, s_6, s_9, s_{11}\}$ ,  $G_2 = \{s_2, s_4, s_{10}, s_{12}, s_{14}, s_{16}\}$  and  $G_3 = \{s_3, s_7, s_8, s_{13}, s_{15}\}$ . And, the second parent represents the group formation where  $G_1 = \{s_2, s_3, s_8, s_{15}, s_{16}\}$ ,  $G_2 = \{s_4, s_6, s_9, s_{12}, s_{13}, s_{14}\}$  and  $G_3 = \{s_3, s_5, s_7, s_{10}, s_{11}\}$ . As shown in Figure 4, the crossover point is equal to 5. Characters from beginning of chromosome to the crossover point are copied from one parent. The rest is copied from the second parent. The result of the single-point crossover between the parents creates two new offspring. The first offspring called offspring1 represents the group formation where  $G_1 = \{s_1, s_5, s_8, s_{15}, s_{16}\}$ ,  $G_2 = \{s_2, s_4, s_6, s_9, s_{12}, s_{13}, s_{14}\}$  and  $G_3 = \{s_3, s_7, s_{10}, s_{11}\}$ . And, the second offspring named offspring2 represents the group formation where  $G_1 = \{s_2, s_6, s_9, s_{11}\}$ ,  $G_2 = \{s_4, s_{10}, s_{12}, s_{14}, s_{16}\}$  and  $G_3 = \{s_1, s_3, s_5, s_7, s_8, s_{13}, s_{15}\}$ .

For offspring1, we then calculate each attribute of  $G_i$  (where  $0 \leq i \leq 3$ ) as demonstrated below:

$$\begin{aligned}
 G_1 &= \left( \frac{3.0+1.5+3.4+2.1+3.6}{5}, \frac{2.5+3.0+3.7+2.4+3.2}{5} \right) \\
 &= (2.72, 2.96) \\
 G_2 &= \left( \frac{3.5+2+5+1.5+1.8+5+2.7}{7}, \frac{2+2+4+2.6+1+3.9+1.2}{7} \right) \\
 &= (1.65, 2.78)
 \end{aligned}$$

$$G_3 = \left( \frac{3.0+0.5+1.5+1.1+4.0}{5}, \frac{2.5+1.0+3.0+2.0+3.9}{5} \right)$$

$$= (2.02, 2.48).$$

The calculation of fitness value for offspring1 is presented below:

$$f(\text{offspring1}) = \sum_i^2 \sum_{j=i+1}^3 \sqrt{\sum_{n=1}^2 |V_n^i - V_n^j|^2}$$

$$= \sqrt{(2.72 - 1.65)^2 + (2.96 - 2.78)^2} +$$

$$\sqrt{(2.72 - 2.02)^2 + (2.72 - 2.48)^2} +$$

$$\sqrt{(1.65 - 2.02)^2 + (2.78 - 2.48)^2}$$

$$= 2.41.$$

For offspring2, each attribute of  $G_i$  (where  $0 \leq i \leq 3$ ) can be calculated as below:

$$G_1 = \left( \frac{3.5+0.5+1.5+3.1}{4}, \frac{2.0+4.0+2.6+1.8}{4} \right)$$

$$= (2.15, 2.60)$$

$$G_2 = \left( \frac{2+4+1.8+2.7+3.6}{5}, \frac{2+3.9+1+1.2+3.2}{5} \right)$$

$$= (1.65, 2.78)$$

$$G_3 = \left( \frac{3+5+1.5+1.1+3.4+5+2.1}{7}, \frac{2.5+1+3+2+3.7+3.9+2.4}{7} \right)$$

$$= (1.73, 2.64).$$

Finally, the calculation of fitness value for offspring1 is presented below.

$$f(\text{offspring2}) = \sum_i^2 \sum_{j=i+1}^3 \sqrt{\sum_{n=1}^2 |V_n^i - V_n^j|^2}$$

$$= \sqrt{(2.15 - 1.65)^2 + (2.60 - 2.78)^2} +$$

$$\sqrt{(2.15 - 1.73)^2 + (2.60 - 2.64)^2} +$$

$$\sqrt{(1.65 - 1.73)^2 + (2.78 - 2.64)^2}$$

$$= 1.11.$$

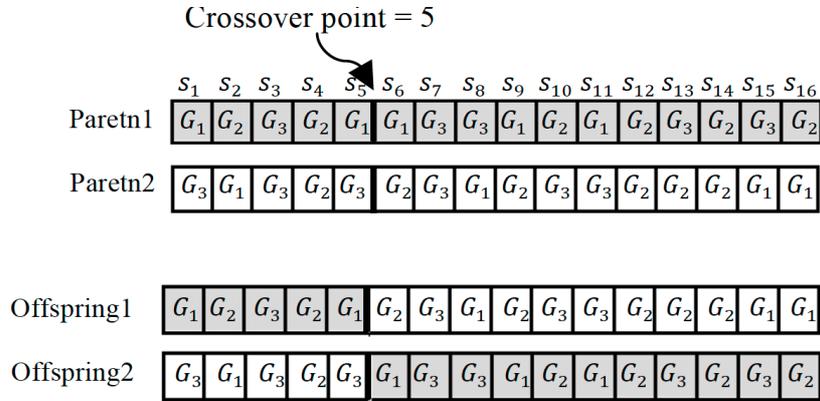


Fig. 4. An example of single point crossover,  $G_i$  representing the group  $i$  and  $0 \leq i \leq p$ .

For each generation, several chromosomes will be regenerated based on the single-point crossover operator.

The lowest fitness value of the chromosome implies that the chromosome generates equal heterogeneous groups of students in terms of student attributes. Therefore, the algorithm aims to search for the optimum solution by generating a chromosome of minimum fitness value. Therefore, two generated offspring will be selected into the next generation. As stated above, the best chromosome in the last generation may simply approximate the solution to the problem. Once the group formation has been formed, students cannot choose or remove group members by themselves without the teacher's permission; otherwise the fitness value of the algorithm would be incorrect.

## 6 System Design and Experimental Results

This section describes the primary components of the system including the user interfaces. Since the algorithm was proposed in order to construct student groups, the researcher has developed a web-based application named Genetic Algorithm for Forming Student Groups (GAFSG). The architecture for the system is based on the standard web paradigm presented in Figure 5. The software was implemented in the PHP language and was done on a notebook with an Intel Core i5-450M 2.4Ghz Nvidia GeForce GT330M, Ram: 2GB DDR3, and HDD: 500GB. GAFSG is widely accessible via the Internet using any Internet web browser.

In this design, each student is required to answer the online ILS questionnaire to identify their learning style. This part is the most important element as data must be obtained before the group formation begins, since the teacher grouped the students based on the learning styles. However, a student's learning style can be affected by different educational experiences, pre-test scores, and previous educational backgrounds, as suggested by Felkel and Gosky [28]. Therefore, the students' educational backgrounds were added as the main attribute for the algorithm.

## **6.1 Teacher Interface**

The teacher interface is designed for initializing group formation. It includes three major functions:

- Adding students by using the ‘Students’ menu. The screen shot for this menu is shown in Figure 6(a). Moreover, all student information can be entered into the system by either importing Microsoft Excel files by clicking ‘Import’ button or typing in student information directly.
- Generating student groups by using the ‘Groups’ menu. A teacher can specify the criteria for grouping students. Once the teacher has clicked this menu, the pop-up window appears in the center of the window. It allows the teacher to set the number of students per group and identify any attributes that need to be considered when grouping students. During the design phase, the teacher can modify the criteria at any time. The example of a pop-up window for this menu is presented in Figure 6(b).
- Adding questions for test and questionnaires related to the group formation by using the ‘Exams’ menu. This part is important since it gives teachers a way of measuring student learning. It distributes exams to students in order to help teachers identify the learning style of their students. The screen shot for this menu is shown in Figure 6(c).

The teacher of the course is able to activate an online ILS questionnaire for each student to complete (see Figure 6(d)). Moreover, the teacher needs to specify the attributes which are used to compose the group formation. The teacher can add the educational background or the pre-requisite courses related to the current course to the database of the system (see Figure 5). Once the responses to the online ILS questionnaire are collected, the teacher can generate student groups. During the design of the group formation, the teacher is able to adjust the criteria in group formation such as the group size, pre-requisite courses, and the dimension of ILS for grouping students learning styles by using “Groups menu”. In this way, the teacher can generate groups of students with different academic backgrounds and the dimension of ILS.

## **6.2 Student Interface**

Since the researcher wished to form the heterogeneous groups of student based on their learning style and prior educational knowledge, all students were requested to answer an online questionnaire in order to identify the learning style. Therefore, the student interface is designed to collect educational background data and allow students to make the ILS questionnaire.

The example of the screenshot for students is illustrated in Figure 7. In order to access this part, students need to login. When they login to the system, they will see all exams or tests provided by the teachers. The incomplete status of any test indicates that the students have not done the exam.

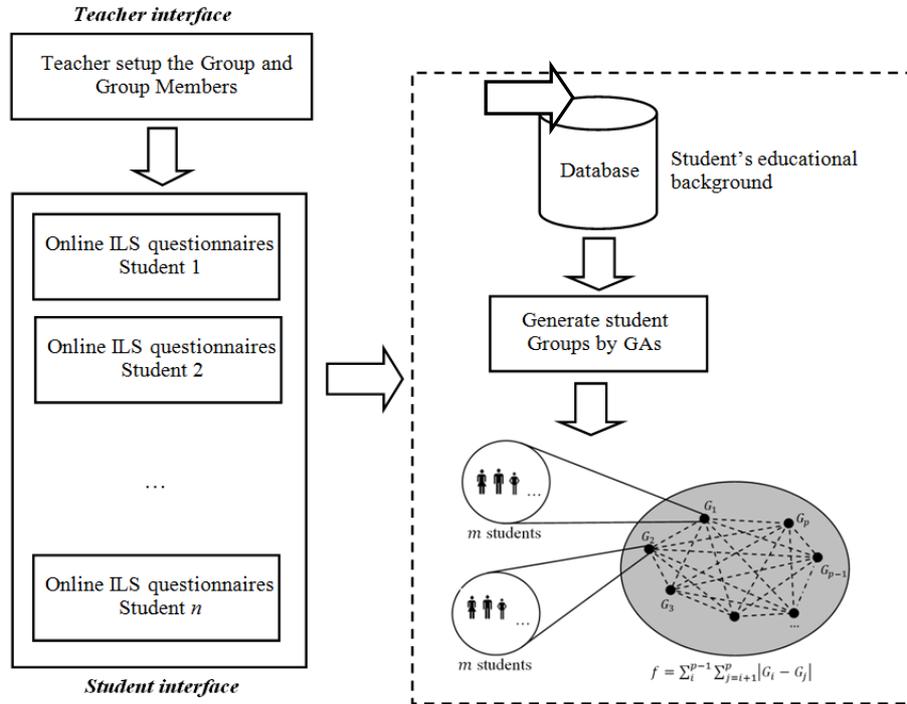
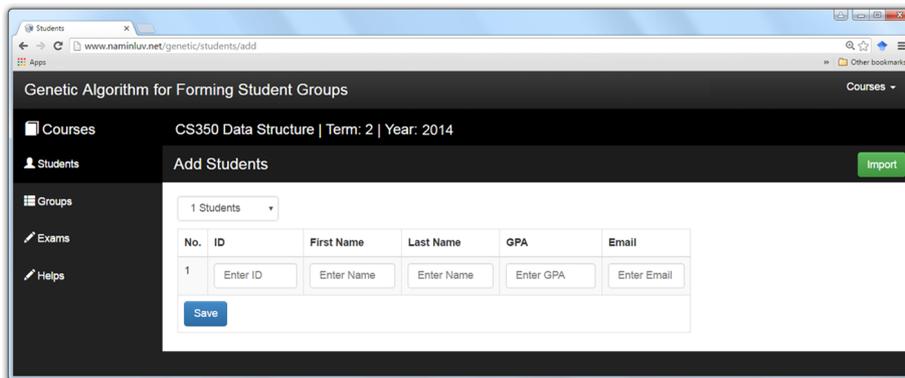
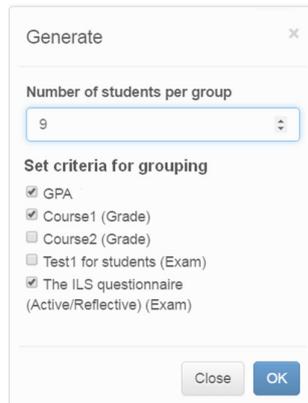


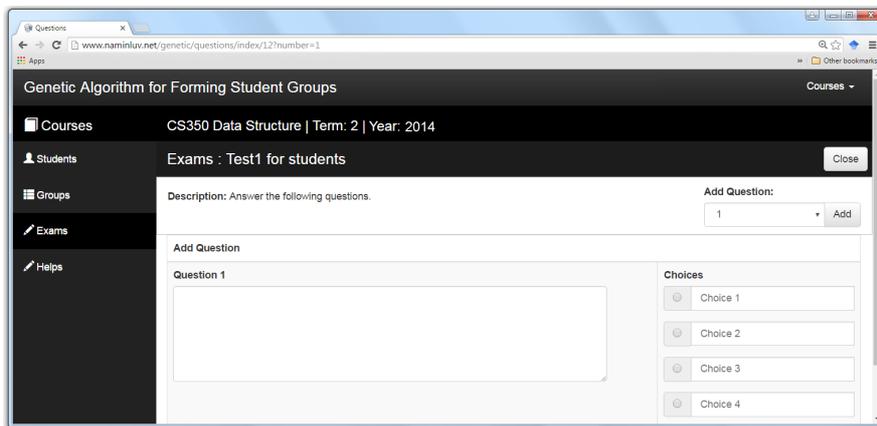
Fig. 5. The architecture of the GAFSG system for forming student groups.



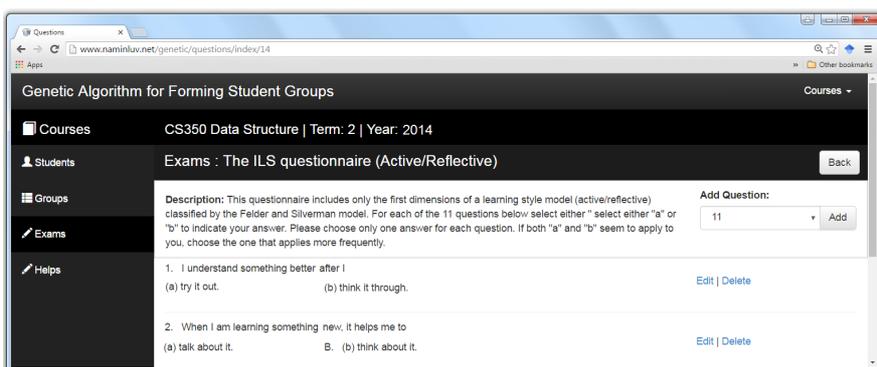
(a) Add Students.



(b) The example of the pop-up window for 'Group' menu.

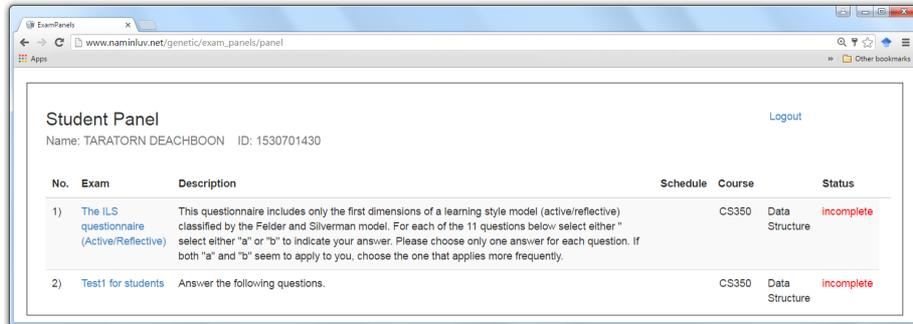


(c) Add questions.

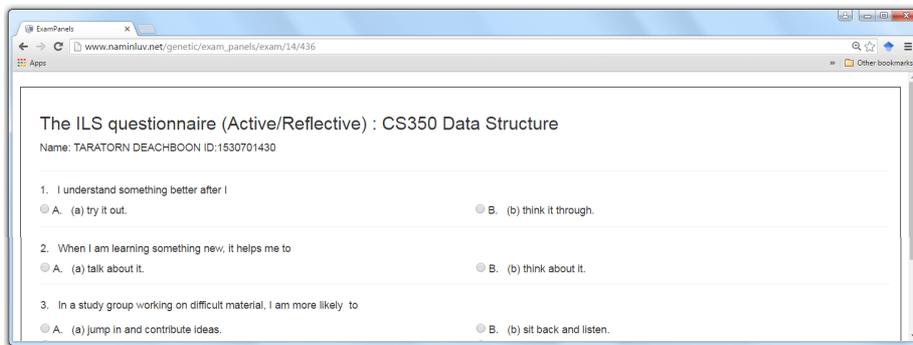


(d) The ILS questionnaire used in this paper classified by the Felder and Silverman model.

**Fig. 6.** A snapshot of the web-based application named GAFSG.



(a) Student Panel and the list of exams.



(b) The ILS questionnaire (Active/Reflective).

Fig. 7. The example of the screenshot for students.

### 6.3 Initial Parameters for the Algorithm

In our genetic algorithm, we have parameters as follows:

- Population size ( $M$ )
- Maximum number of generations ( $Max\_Gen$ )
- Crossover percentage on population ( $p_c$ ).

The program was run several times to see which initial values of parameters would rapidly direct the algorithm's search to the optimized formation. Execution time took a few seconds. Parameters for controlling the algorithm are illustrated in Table II.

Table 2. Initial parameters used for our algorithm

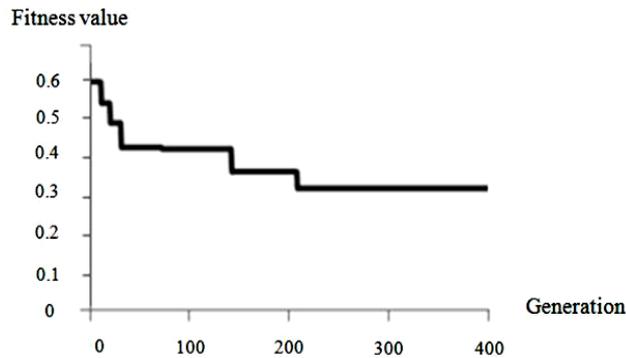
Constants	Value
$M$	300
$Max\_Gen$	400
$p_c$	0.30

#### 6.4 A Case Study and Discussion

In this section, the results of the case study are presented. The case study for generating student groups in the CS350 course (Data Structures) was conducted during the second semester of 2014 at Bangkok University. It was performed using 25 undergraduate students majoring in Computer Science. As mentioned previously, the teacher has to specify the type of academic attributes and the dimension of ILS for grouping students. In this paper, three academic attributes were used to generate heterogeneous groups of students.

As defined in the Felder and Silverman model [21], if a student is classified as active, the student tends to learn by trying things out and gains pleasure from working in the group. This type of students is more likely to understand information best by doing something active with it. Furthermore, according to Martín and Paredes [5], the active/reflective dimension is related to computer-supported collaborative learning (CSCL) and groupware which is designed to help students involved in a common task to achieve their goal. Hence, one attribute used for building the heterogeneity group is the first dimension of ILS (active/reflective). Therefore, at the beginning of the CS350 course, all students were asked to make the ILS online questionnaire. In addition, students were allowed to fill in their grade point average (GPA) and previous grades, since two other attributes for forming student groups are GPA and grade of pre-requisite courses in the Bachelor of Science Program in Computer Science at Bangkok University. Therefore, all groups have to be balanced in terms of the first dimension of ILS and the grade of two previous courses to ensure that all generated groups have an equal opportunity in performing well in the class.

Figure 8 shows the result generated by the algorithm for the first case study where the fitness value is 0.31.



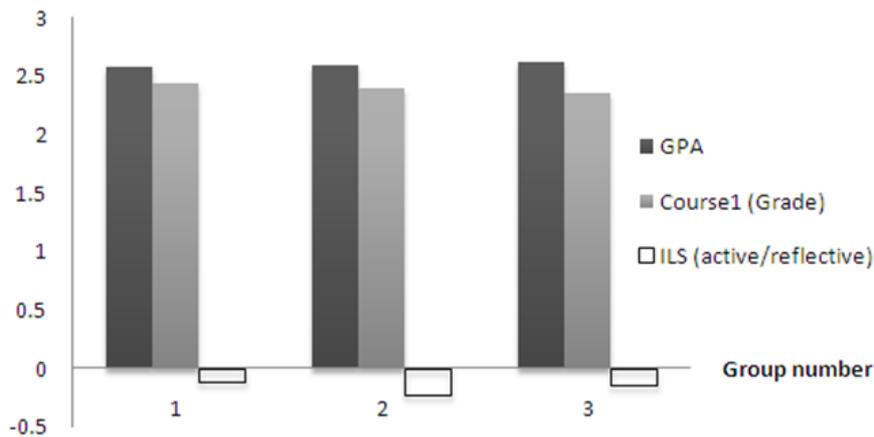
**Fig. 8.** Experimental result made by the proposed algorithm for generating student groups ( $n = 25$ , group size = 9, and number of groups = 3), where  $M = 300$ ,  $\text{Max\_Gen} = 400$  and  $p_c = 0.30$ .

Groups of students constructed by the algorithm are demonstrated in Figure 9. As stated earlier, the group size can be enlarged or reduced. In this case study, the researcher set  $p$  equal to 9. Therefore, there were three groups of students. The first two

established groups have nine students, but the third group has only seven members. As one can see from Figure 9(a), each generated group contains dissimilar students, but established groups are similar in all attribute values. Due to the design, the algorithm tries to make the equity among the generated groups. The average value of ILS, spans a small margin which is -0.22 to -0.11. Average Course1(Grade) ranges from 2.36-2.44. Furthermore, GPA is between 2.58-2.62. Consequently, the optimized group formation is established when the fitness value is as low as 0.31. It indicates that the researcher’s algorithm is able to distribute heterogeneous students optimally. Members of each group associated with Figure 9(a) are illustrated in Figure 9(b).

To guarantee the performance of the algorithm we compared the results to the self-selecting method made by the students themselves. The results of the self-selecting method are presented in Figure 10 (a) and Figure 10 (b). As one can see from these figures, most groups are greatly unbalanced. The groups differ greatly in most attributes, including the first dimension of ILS (active/reflective) score. Moreover, the fitness value is 9.92, which is higher than the fitness value of the proposed algorithm. It is confirmed by the value of ILS with a wide range from -2.25 to 2.5. Especially for group 3, all members are on the same category of this dimension (active). Hence group number 3 has a preference for active rather than reflective learning. Besides, the value of Course1(Grade) spans a wide gap from 1.94-3.0. Additionally, GPA spans from 2.22-3.06. Students with high GPA are likely in the same group, as presented in group 2 in Figure 10(b).

From these results, it implies that the self-selecting method generates the groups with no direction, making some groups better in some attributes, but lower in the others. Therefore, for the self-selecting method, fairness among the student groups is difficult to provide in terms of the learning styles and prior educational knowledge.



(a) Groups of students made by the algorithm, where fitness value = 0.31.

**Fitness: 0.31**

Group: 1

[Print](#) [View All Groups](#)

ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1560700336	PANUPON	SAMEKO	2.40	2	-3	<a href="#">Change Group</a>
1560700732	WICHA	VAINIK	2.42	2	-1	<a href="#">Change Group</a>
1560700534	SORAVI	PREAMSR	2.59	2.5	-1	<a href="#">Change Group</a>
1560701045	RUNGRUD	TAKHUL	2.65	3	5	<a href="#">Change Group</a>
1560700518	SORNTA	PEETAWATTANAKU	2.59	2	1	<a href="#">Change Group</a>
1560700625	Wachir	Pimsawan	3.57	4	-1	<a href="#">Change Group</a>
1560700013	SUTHINA	SONTHIKHU	2.75	2.5	-1	<a href="#">Change Group</a>
1530701430	TARATOR	DEACHBOO	2.37	2	1	<a href="#">Change Group</a>
1550700783	PATIPA	SAKDAMI	1.90	2	-1	<a href="#">Change Group</a>

Group: 2

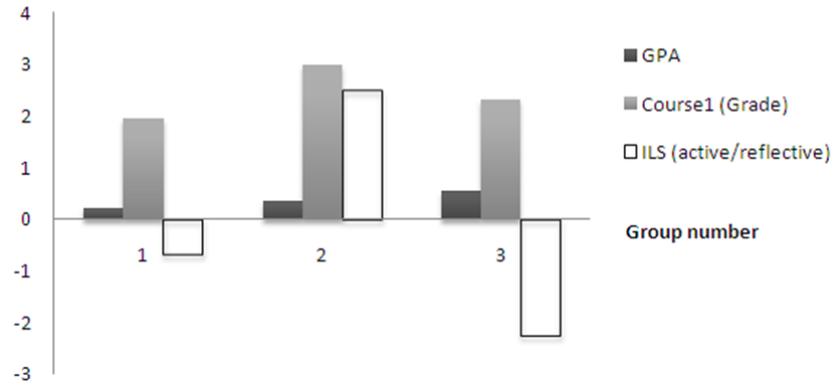
ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1560700302	KORRAPA	PHUMPHUNGPHU	3.10	3.5	-1	<a href="#">Change Group</a>
1560700120	THANAWA	KUMPOO	2.53	2	-1	<a href="#">Change Group</a>
1540701099	THANAPHO	JINDAAPIRAK	1.93	1.5	-1	<a href="#">Change Group</a>
1560700823	KAMPANO	MANOTULAKA	2.99	3	-1	<a href="#">Change Group</a>
1560700708	NATTAPON	MAUNGJUNTHUE	2.16	1	1	<a href="#">Change Group</a>
1560700674	PUWADO	SRITU	3.15	2.5	5	<a href="#">Change Group</a>
1560700021	SURAKI	KORNONGKLAN	3.22	3.5	-3	<a href="#">Change Group</a>
1560700476	ATIKO	ASSAWARITDAMRON	2.36	2	-3	<a href="#">Change Group</a>
1530703113	THITIPU	KUNASINSATHI	1.87	2.5	2	<a href="#">Change Group</a>
<b>Average:</b>			<b>2.59</b>	<b>2.39</b>	<b>-0.22</b>	<b>Actions</b>

Group: 3

ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1550700262	PONGPA	AUMPAIPANNASU	2.64	2	-1	<a href="#">Change Group</a>
1570700862	ary	Tippawa	2.96	2	5	<a href="#">Change Group</a>
1560700146	SUKI	PHUTTHICHO	2.71	2	3	<a href="#">Change Group</a>
1550701203	THIT	RUEANGRI	1.82	2	-5	<a href="#">Change Group</a>
1550700296	WITAWA	TUNSAKUNWWATTAN	2.68	2.5	-1	<a href="#">Change Group</a>
1560701185	Sittipor	Suksapa	3.32	4	5	<a href="#">Change Group</a>
1560700187	WISAROO	LANGKUNSA	2.19	2	-7	<a href="#">Change Group</a>
<b>Average:</b>			<b>2.62</b>	<b>2.36</b>	<b>-0.14</b>	<b>Actions</b>

(b) List of students in each group.

**Fig. 9.** Results of the proposed algorithm (25 students).



(a) Groups of student self-selecting method where fitness value = 9.92.

**Fitness: 9.92**

Group: 1

[Print](#) [View All Groups](#)

ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1560700732	WICHA	VAINIK	2.42	2	-1	<a href="#">Change Group</a>
1560700518	SORNTA	PEETAWATTANAKU	2.59	2	1	<a href="#">Change Group</a>
1530701430	TARATOR	DEACHBOO	2.37	2	1	<a href="#">Change Group</a>
1550700783	PATIPA	SAKDAMI	1.90	2	-1	<a href="#">Change Group</a>
1560700120	THANAWA	KUMPOO	2.53	2	-1	<a href="#">Change Group</a>
1560700708	NATTAPON	MAUNGJUNTHUE	2.16	1	1	<a href="#">Change Group</a>
1560700476	ATIKO	ASSAWARITDAMRON	2.36	2	-3	<a href="#">Change Group</a>
1530703113	THITIPU	KUNASINSATHI	1.87	2.5	2	<a href="#">Change Group</a>
1550701203	THIT	RUEANGRI	1.82	2	-5	<a href="#">Change Group</a>
<b>Average:</b>			<b>2.22</b>	<b>1.94</b>	<b>-0.67</b>	<b>Actions</b>

Group: 2

ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1560701045	RUNGRUD	TAKHUL	2.65	3	5	<a href="#">Change Group</a>
1560700625	Wachir	Pimsawan	3.57	4	-1	<a href="#">Change Group</a>
1560700302	KORRAPA	PHUMPHUNGPHU	3.10	3.5	-1	<a href="#">Change Group</a>
1560700823	KAMPANO	MANOTULAKA	2.99	3	-1	<a href="#">Change Group</a>
1560700674	PUWADO	SRITU	3.15	2.5	5	<a href="#">Change Group</a>
1570700862	ariy	Tippawa	2.96	2	5	<a href="#">Change Group</a>
1560700146	SUKI	PHUTTHICHO	2.71	2	3	<a href="#">Change Group</a>
1560701185	Sittipor	Suksapa	3.32	4	5	<a href="#">Change Group</a>
<b>Average:</b>			<b>3.06</b>	<b>3.00</b>	<b>2.50</b>	<b>Actions</b>

Group: 3

ID	First Name	Last Name	GPA	Course1: Grade	ILS	Actions
1560700336	PANUPON	SAMEKO	2.40	2	-3	<a href="#">Change Group</a>
1560700534	SORAVI	PREAMSR	2.59	2.5	-1	<a href="#">Change Group</a>
1560700013	SUTHINA	SONTHIKHU	2.75	2.5	-1	<a href="#">Change Group</a>
1540701099	THANAPHO	JINDAPIRAK	1.93	1.5	-1	<a href="#">Change Group</a>
1560700021	SURAKI	KORNONGKLAN	3.22	3.5	-3	<a href="#">Change Group</a>
1550700262	PONGPA	AUMPAIPANNASU	2.64	2	-1	<a href="#">Change Group</a>
1550700296	WITAWA	TUNSAKUNWIWATTAN	2.68	2.5	-1	<a href="#">Change Group</a>
1560700187	WISAROO	LANGKUNSA	2.19	2	-7	<a href="#">Change Group</a>
<b>Average:</b>			2.55	2.31	-2.25	<b>Actions</b>

(b) List of students in each group

Fig. 10. Self-selecting method (25 students).

## 7 Conclusion and Future Work

In this paper, an approach called GAFSG was presented in order to generate student groups in a heterogeneous way by using GAs. The proposed approach adapts some dimensions of the Felder-Silverman model to grouping students. The paper uses the concept of the heterogeneous grouping of students to achieve fairness among groups and ensure that heterogeneous students are distributed optimally. The quality of the groups generated by our proposed algorithm is compared to the self-selecting method. The algorithm was implemented as a web-based application that enables the teachers to form student groups.

According to the case study, the algorithm performed better than the self-selecting method. The results have shown that this algorithm achieves the mechanism for equity among constructed groups in terms of academic abilities and optimally distributes heterogeneous students within the group formation. Additionally, the paper demonstrates a web-based application that implements the proposed mathematical approach by using Genetic Algorithms. The designed software system has also been presented.

Since heterogeneous groups of students enhance the learning progress in collaborative learning, for future work the researcher will try to identify the relationship between the performance of student groups and each dimension of learning style within the groups which are generated by this proposed method. Another issue for future work is to demonstrate the scalability of the algorithm and prove that the algorithm is able to search for stable solutions which are close to the optimum for different datasets.

## 8 References

- [1] R. M. Felder and R. Brent, "Cooperative learning," In Active learning: Models from the analytical sciences, ACS Symposium Series. 2007;970:34-53. <https://doi.org/10.1021/bk-2007-0970.ch004>
- [2] L. Razmerita and B. Armelle, "Collaborative Learning in Heterogeneous Classes: Towards a Group Formation Methodology", in Proceedings of the 3rd International Conference on Computer Supported Education (CSEDU 2011), Noorwijkhout, Netherlands, 2011, pp. 189-194.
- [3] J. M. Balmaceda, S. N. Schiaffino, and J. A. D. Pace, "Using constraint satisfaction to aid group formation in CSCL. Inteligencia Artificial," Revista Iberoamericana de Inteligencia Artificial, vol. 17, no. 53, 2014, pp. 35-45.
- [4] P. Paredes and A. Ortigosa, P. Rodriguez, "Method for Supporting Heterogeneous-Group Formation through Heuristics and Visualization," J. UCS, vol. 16, no.19, pp. 2882-2901, 2010.
- [5] E. Martín and P. Paredes, "Using Learning Styles for Dynamic Group Formation in Adaptive Collaborative Hypermedia Systems," In ICWE Workshops; 2004 July 28-30; Munich, Germany.
- [6] J. Allert, "Learning style and factors contributing to success in an introductory computer science course," In Advanced Learning Technologies. In Proceedings of the IEEE International Conference on Advanced Learning Technologies; 2004 30 Aug.-1 Sept; Joensuu, Finland. IEEE:2004. p.385-389. <https://doi.org/10.1109/icalt.2004.1357442>
- [7] R.M. Felder and B. A. Soloman, "Index of Learning Styles," 2015; <http://www.ncsu.edu/felder-public/ILSpage.html>.
- [8] L. Chen and J. M. C. Lin, "Learning styles and student performance in java programming courses," in Proceedings of the 2011 International Conference on Frontiers in Education: Computer Science and Computer Engineering; Las Vegas, USA., 2011, pp. 53-58.
- [9] R.M. Felder and L.K. Silverman, "Learning and teaching styles in engineering education". Engr. Educ.,78 (7), 674-681. 1988.
- [10] H. JuHou and H. Hui, "EBFO-based Personalized Learning Materials Recommendation in Cooperative Learning Group," Journal of Convergence Information Technology, vol. 8, no. 10, 2013, pp. 1057-1065. <https://doi.org/10.4156/jcit.vol8.issue10.128>
- [11] T. Y. Tang and K. C. Chan, "Feature construction for student group forming based on their browsing behaviors in an e-learning system", in Proceeding of 7<sup>th</sup> Pacific Rim International Conference on Artificial Intelligence, Trends in Artificial Intelligence, Tokyo, Japan. 2002, pp. 512-521. [https://doi.org/10.1007/3-540-45683-x\\_55](https://doi.org/10.1007/3-540-45683-x_55)
- [12] S. Graf and R. Bekele, "Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization", Intelligent Tutoring Systems. Springer Berlin Heidelberg, 2006, pp. 217-226. [https://doi.org/10.1007/11774303\\_22](https://doi.org/10.1007/11774303_22)
- [13] R. E. Slavin, "When Does Cooperative Learning Increase Achievement", Psychological Bulletin. Vol. 94, no. 3, 1983, pp. 429-445. <https://doi.org/10.1037/0033-2909.94.3.429>
- [14] M. Kyprianidou, S. Demetriadis, T. Tsiatsos, and A. Pombortsis, "Group formation based on learning styles: can it improve students' teamwork?", Educational Technology Research and Development, 60(1), 2012, pp.83-110. <https://doi.org/10.1007/s11423-011-9215-4>
- [15] N. Herrmann, "Creativity, learning, and the specialized brain in the context of education for gifted and talented children", In Adapted from an address to the 7th world conference on gifted and talented children, Salt Lake City, Utah, 1989.

- [16] D. W. Johnson, R. T. Johnson and K. A. Smith, Cooperative learning. Interaction Book Company, 1989.
- [17] J. S. Thousand, R. A. Villa, and A. I. Nevin, “Creativity and collaborative learning: The practical guide to empowering students, teachers, and families”, Baltimore: Paul H. Brookes Publishing Co., 2002.
- [18] A. J. Kruse, “The Benefits of Heterogeneous Groupings over Homogeneous Groupings”, 2011: <http://andrewkruse.efoliomn.com>
- [19] A. Strahan, K. Hartman, and S. Sikes, “North Carolina Middle School Association Research Bulletin, Heterogeneous Grouping,” 2015; [http://www.ncmle.org/ncmsa\\_bulletins/ncmsa\\_heterogeneous\\_grouping.html](http://www.ncmle.org/ncmsa_bulletins/ncmsa_heterogeneous_grouping.html).
- [20] R.M. Felder, “Reaching the second tier: learning and teaching styles in college science education”, J. Coll. Sci. Teach., 23(5) 1993, pp. 286 -290.
- [21] R. M. Felder and J. Spurlin. “Applications, reliability and validity of the index of learning styles,” International journal of engineering education, vol. 21, no. 1, pp. 103-112, 2005.
- [22] T. A. Litzinger, S. H. Lee, and J. C. Wise, “A study of the reliability and validity of the Felder-Soloman Index of Learning Styles,” Journal of Engineering Education, vol. 96 no. 4, 2007, pp. 310-319. <https://doi.org/10.1002/j.2168-9830.2007.tb00941.x>
- [23] K. Tanner and D. Allen, “Approaches to biology teaching and learning: learning styles and the problem of instructional selection—engaging all students in science courses”, Cell Biology Education, 3(4), 2004, pp. 197-201. <https://doi.org/10.1187/cbe.04-07-0050>
- [24] L. Lin, L. Cao, J. Wang, and C. Zhang, “The applications of genetic algorithms in stock market data mining optimization”, In Proceedings of Fifth International Conference on Data Mining, Text Mining and their Business Applications, 2004, pp. 273-280.
- [25] D. Whitley, “A genetic algorithm tutorial”, Statistics and computing, vol. 4, no. 2, 1994, pp. 65-85. <https://doi.org/10.1007/BF00175354>
- [26] Y. Kaya, and M. Uyar, A novel crossover operator for genetic algorithms: ring crossover. arXiv preprint arXiv:1105.0355. 2011.
- [27] N. Stark, G. F. Minetti, and C. Salto, A new strategy for adapting the mutation probability in genetic algorithms. In XVIII Congreso Argentino de Ciencias de la Computación. 2012.
- [28] B. H. Felkel and R. M. Gosky, “A study of Reliability and Validity of the Felder-Soloman Index of Learning Styles for Business Students”, 2015; <http://archives.math.utk.edu/ICTCM/VOL24/C004/paper.pdf>

## 9 Author

**Anon Sukstrienwong** is a lecture at Bangkok University, school of Information Technology and Innovation, Bangkok, Thailand (Email:anon.su@bu.ac.th).

Submitted, 28 April 2016. Published as resubmitted by the authors on 11 October 2016.