Heterogeneous Group Forming Using Mixed-Integer Programming for an Integrated Engineering Design Course

Sinan Onal, Assistant Professor, sonal@siue.edu
Xin Chen, Assistant Professor, xchen@siue.edu
Industrial Engineering, Southern Illinois University Edwardsville

ABSTRACT

Cooperative learning is an effective technique in engineering education as it enhances student learning and ensures quality of outcome from each group. Integrated engineering design (Capstone) course is a good example of cooperative learning. Teams of students work on projects and perform tasks that range from assessment of user needs, design to production, and evaluation of economic/financial feasibility. Forming student groups may be challenging when the class size is large and multiple factors must be taken into consideration. In this study, we propose a mixed-integer programming (MIP) model to form heterogeneous groups based on students' gender (male or female), nationality (international or U.S.), grade point average (GPA), research interests (marketing, product design, project management, manufacturing, and finance), and their conflicts with other students. Each group is expected to have an average GPA similar to that of other groups. Each group should have students with different research interests while avoiding conflicts between students. Total 40 students in a Capstone course are considered in this study. The MIP model is implemented in the General Algebraic Modeling System (GAMS) and solved using the CPLEX solver. Results show that the MIP model efficiently identifies the optimal solution that forms groups of students.

I. INTRODUCTION

Several studies have confirmed that cooperative learning is an effective technique to enhance student performance in engineering education [1-3]. In cooperative learning, a group of students work together to complete tasks collectively toward project goals. Integrated engineering design (Capstone) courses are excellent venues for cooperative learning. A Capstone course concludes students' undergraduate study in industrial engineering. The course encompasses the integration of industrial engineering concepts and expands to enterprise applications. Teams of students work on projects and perform tasks that range from assessment of user needs, design to production, and evaluation of economic/financial feasibility.

However, cooperative learning may provide unsuccessful outcome if student groups are formed poorly [4]. Therefore, group forming is a critical task in cooperative learning since there are several dependent and independent variables in student's vector. Some instructors mix the students in groups based on their profile, learning style, academic performance so that students of all levels are represented in each group in a heterogeneous way, while others organize students by ability levels in a homogeneous way. The decision is made based on the purpose of the learning activity. In many Capstone courses, heterogeneous groups are formed because one of the objectives of the course is to balance students with complementary skills and different academic levels to obtain appropriate learning results for all groups.

Group formation can be based on students' profile such as ability, prior knowledge, learning style, and browsing behavior. However, group forming is especially difficult when the class size is large and independent variables are unbalanced (e.g., the number of male students almost doubles that of female students and international students are four times more than domestic students). Therefore, several computer based approaches have been proposed in the academic literature to provide effective solutions for assigning students to different groups.

Bekele [5] proposed a group forming algorithm based on Ant Colony Optimization for maximizing group heterogeneity by considering a large number of student characteristics: gender, group work attitude, interest in mathematics, achievement motivation, self-confidence, and shyness. Graf et al. [6] also used Ant Colony Optimization algorithm to build students groups in a heterogeneous way based on the students' personality traits such as interest for the subject, achievement motivation, self-confidence, and shyness and their performance level in the subject. Wang et al. [7] introduced a model, DIANA using genetic algorithm for forming heterogeneous groups based on the students' thinking styles. Cavanaugh et al. [8] presented a model called Team-Maker based on 'hill climbing' optimization algorithm with weighted criteria for building heterogeneous groups. The authors collected students' information such as gender, skills, and students' schedules through questionnaires. Moreno et al. [9] presented a multi-objective optimization method based on genetic algorithm for group forming by incorporating an arbitrary number of student characteristics. However, currently available computer based support systems do not consider students' research interests and conflicts between students in forming groups. In addition, most tools applied earlier are heuristic optimization tools which do not guarantee the optimal assignment of students to groups.

In this paper, we present a mixed-integer programming (MIP) model to automatically form heterogeneous groups based on students' gender (male, female), nationality (international or U.S.), grade point average (GPA), research interests (marketing, product design, project management, manufacturing, and finance), and their conflicts with other students. Each group is expected to have an average GPA similar to that of other groups. Also, each group should have students with different research interests while avoiding conflicts between students. The rest of the paper is organized as follows: Section II presents brief description of the MIP model. Section III describes the proposed heterogeneous group forming formulation. Section IV introduces the results for a Capstone course and is followed by conclusions and future work in Section V.

II. MIXED-INTEGER PROGRAMMING

Forming groups may be completed manually if the class size is small. For example, to assign nine students to three groups according to their GPA, a trial-and-error method may be used to test all $\frac{C(9,3)\times C(6,3)}{P(3,3)} = 280$ possible group assignments and find the best assignment such that difference between average GPAs of three groups is the smallest. When the class size is large or additional constraints must be satisfied, e.g., each group must have at least one female student and one international student, to form groups manually will take long time and become intractable. Mathematical models such as an MIP model may be used to formulate the problem and a software package may be used to find the optimal solution.

Let I be a set of students in a class and J be a set of groups. $|I|$ is the total number of students and $|J|$ is the total number of groups. Total $|I|$ students must be assigned to total $|J|$ groups such that each group has the same number of students. $|I|$ must be divided exactly by $|J|$. If this is not the case, artificial students may be added to the set I to ensure that the remainder of $|I|$ divided by $|J|$ is zero. Let i represent a student in the class, $i = 1, 2, \ldots, |I|$ and $i \in I$. Let M represent a set of interests, i_1, i_2, …, $i_{|M|}$, where $|M|$ is the total number of interests. Each student has a prioritized list of interests his or her preference. The first one on the list, i.e., the most preferred interest, is assigned the highest score, a score of $|M|$. The second one on the list, i.e., the second most preferred interest, is assigned a score of $|M| - 1$. The last one on the list, i.e., the least preferred interest, is assigned a score of 1. In general, the m^{th} interest on the list is assigned a score of $|M| + 1 - m$.

Let $interest(i, m)$ represent a prioritized list of interests for student i. For instance, suppose there are five interests, i_1, i_2, i_3, i_4, and i_5. If student i's prioritized interests from the most preferred to least preferred are i_3, i_2, i_4, i_1, and i_5, then $interest(i, 1) = 3$, $interest(i, 2) = 2$, $interest(i, 3) = 4$, $interest(i, 4) = 1$, and $interest(i, 5) = 5$. Scores are assigned to student i's interests as follows: $i_3 = 5$, $i_2 = 4$, $i_4 = 3$, $i_1 = 2$, and $i_5 = 1$.

Multiple students are assigned to each group j, $j \in J$. Students in the same group are expected to have different interests, which are often aligned with students' skills. These different but complementary skills are necessary for the group to complete a high-quality project that covers various aspects of student's learning in four years. The ideal situation is that the most preferred interests of students in the same group are different. Each student contributes a set of skills that align with his or her most preferred interest; that set of skills is different than that of any other student in the same group and these complementary skill sets are the cornerstone of a comprehensive project. A minimum requirement for students in the same group to have different most preferred interests is that $|M| \geq \frac{|I|}{|J|}$, i.e., the total number of interests each student has must be greater than or equal to the number of students in a group.

If all students in the same group have different most preferred interests, each student is applying his or her most preferred interest to complete the project and each student's interest score is $|M|$. Because there are many constraints involved in assigning students into groups, it is possible that two or more students in a group have the same most preferred interests; to ensure different interests in a group, at least one student must apply his or her second most preferred interest or even less preferred interest to complete the project. The student's interest score will be $|M| - 1$ if his or her second most preferred interest is used by his or her group, and $|M| - 2$ if his or her third most preferred interest is used. In general, a student's interest score is $|M| + 1 - m$ when the student's m^{th} most preferred interest is used by his or her group.

Let $s(i, m)$ be a binary variable. $s(i, m) = 1$ indicates student i's m^{th} most preferred interest is used by his or her group and i's interest score is $|M| + 1 - m$; $s(i, m) = 0$ otherwise. The objective of the MIP model is to maximize the interest score among all students. Let f represent the minimum interest score among all students. Equation (1) calculates interest scores and ensure that f represents the minimum score. Since only one of each student's interests is used for his or her group, Equation (2) ensures that the summation of $s(i,m)$ is 1. Equation (3) is the objective function of the MIP model.

$$\prod_{m=1}^{|M|}(|M|+1-m)\,s(i,m) \geq f, \forall i \quad (1)$$

$$\sum_{m=1}^{|M|} s(i,m) = 1, \forall i \quad (2)$$

$$\max \quad f \quad (3)$$

In addition to the objective function (3) and two constraints (1) and (2), other constraints are needed to assign students into groups. Let $x(i,j)$ be a binary decision variable. $x(i,j) = 0$ indicates student i is not assigned to group j and $x(i,j) = 1$ indicates student i is assigned to group j. Let $g(i)$ be a binary variable that represents the gender of student i. $g(i) = 0$ indicates student i is male and $g(i) = 1$ indicates student i is female. In engineering disciplines, there are more male students than female students. Let $MinF$ and $MaxF$ be the minimum and maximum numbers of female students, respectively, required for each group. $MinF$ and $MaxF$ are parameters for the MIP model and may be determined based on the total number of female students and the number of groups. For instance, if 86 female students will be assigned to 9 groups, each group on average will have around 9 female students. $MinF$ and $MaxF$ may be set to 8 and 10, respectively. In general, $0 \leq MinF \leq MaxF \leq \frac{|I|}{|J|}$. When $MinF$ and $MaxF$ are close, the interval $(MinF, MaxF)$ is small and Equations (4) and (5) ensure that different groups have a similar mix of male and female students.

$$\sum_{i=1}^{|I|} g(i)x(i,j) \geq MinF, \forall j \quad (4)$$

$$\sum_{i=1}^{|I|} g(i)x(i,j) \leq MaxF, \forall j \quad (5)$$

Similar to gender distribution, another factor in forming groups is nationality. In engineering disciplines, there may be a significant number of international students. Let $n(i)$ be a binary variable that represents the nationality of student i. $n(i) = 0$ indicates student i is from the U.S. and $n(i) = 1$ indicates student i is an international student. Let $MinN$ and $MaxN$ be the minimum and maximum numbers of international students, respectively, required for each group. $MinN$ and $MaxN$ are parameters for the MIP model and may be determined based on the total number of international students and the number of groups. $0 \leq MinN \leq MaxN \leq \frac{|I|}{|J|}$. Equations (6) and (7) ensure that different groups have a similar mix of U.S. and international students.

$$\sum_{i=1}^{|I|} n(i)x(i,j) \geq MinN, \forall j \quad (6)$$

$$\sum_{i=1}^{|I|} n(i)x(i,j) \leq MaxN, \forall j \quad (7)$$

To ensure that the outcome of projects completed by different groups has similar quality, average academic performance of students in one group should be similar to that of other groups. GPA is a widely accepted indicator of student academic performance. Let $gpa(i)$ represent the GPA of student i. Let $MinGPA$ represent the minimum average GPA required for each group. $MinGPA$ is a parameter for the MIP model and may be determined based on the average GPA of all students. For instance, if the grand average GPA for all students is 3.098, $MinGPA$ may be set to 3.000. Equation (8) ensures that different groups have a similar average academic performance.

$$\frac{\sum_{i=1}^{|I|} gpa(i)x(i,j)}{|I|/|J|} \geq MinGPA, \forall j \tag{8}$$

An often overlooked factor in assigning students to groups is that there may be conflicts between students. Students who have a conflict or dislike each other do not want to be assigned to the same group. For example, two students who have worked on projects in the same group and did not get along would not want to be in the same group again. For best student learning, instructors should avoid assigning students with a conflict to the same group. Let $c(i,i')$ represent whether a conflict exists between students i and i'. $c(i,i') = 0$ indicates there is not a conflict and $c(i,i') = 1$ indicates a conflict exists. Equation (9) ensures that two students with a conflict are not assigned to the same group. It is also clear that each student may be assigned to only one group (Equation (10)) and each group must have exactly $\frac{|I|}{|J|}$ groups (Equation (11)).

$$x(i,j) + x(i',j) \leq 2 - c(i,i'), \forall i,i',j \tag{9}$$

$$\sum_{j=1}^{|J|} x(i,j) = 1, \forall i \tag{10}$$

$$\sum_{i=1}^{|I|} x(i,j) = \frac{|I|}{|J|}, \forall j \tag{11}$$

The most challenging task in formulating the MIP model is to ensure different interests in a group. Each student in a group applies one of his or her interests to complete the project, which is different than the interest of any other student in the same group. A set of constraints are written in two steps. First, difference between interests of any two students is calculated. The difference is between $1 - |M|$ and $|M| - 1$. Since $|M|$ is usually greater than 1, the difference has negative values, which cannot be used directly to indicate whether interests of two students are the same or different. The absolutely value of the difference is calculated. The relationship between the absolute value and difference between interests of two students may be described by piecewise linear functions shown in Figure 1.

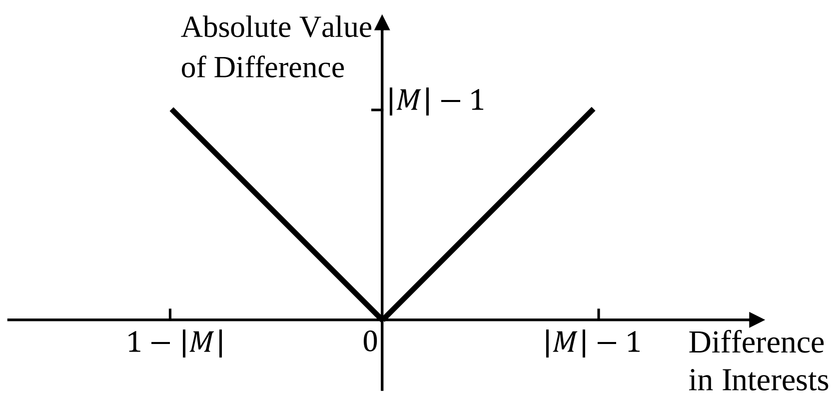

Figure 1. Difference in Student Interest Scores and Its Absolute Value

The difference between interests of two students i and i' may be expressed as a weighted sum of $1-|M|$, 0, and $|M|-1$. Let $a_1(i,i')$, $a_2(i,i')$, and $a_3(i,i')$ be the weights. Equation (12) expresses the difference using these three weights. Equation (13) ensures that the sum of weights is 1. At any time, the sum of two weights is 1 and the other weight is 0. Let $b_1(i,i')$ and $b_2(i,i')$ be two binary variables. Equations (14) ~ (17) ensure that either $a_1(i,i')+a_2(i,i')=1$ or $a_2(i,i')+a_3(i,i')=1$.

$$\prod_{m=1}^{|M|} interest(i,m)\, s(i,m) - \prod_{m=1}^{|M|} interest(i',m)\, s(i',m) = a_1(i,i')(1-|M|) + a_3(i,i')(|M|-1), \forall i,i' \tag{12}$$

$$a_1(i,i')+a_2(i,i')+a_3(i,i')=1, \forall i,i' \tag{13}$$

$$a_1(i,i') \le b_1(i,i'), \forall i,i' \tag{14}$$

$$a_2(i,i') \le b_1(i,i')+b_2(i,i'), \forall i,i' \tag{15}$$

$$a_3(i,i') \le b_2(i,i'), \forall i,i' \tag{16}$$

$$b_1(i,i')+b_2(i,i')=1, \forall i,i' \tag{17}$$

In the second step, the absolute value of difference between interests may be expressed as a weighted sum of $|M|-1$ using the same weights $a_1(i,i')$, $a_2(i,i')$, and $a_3(i,i')$. Let $absolute(i,i')$ be the absolute value of difference. Equation (18) calculate the absolute value of difference. Equation (19) ensures that any two students in the same group have different interests. When $absolute(i,i')=0$, students i and i' have the same interest; they cannot be assigned to the same group because $x(i,j)+x(i',j) \le 1$. When $absolute(i,i')>0$, students i and i' have different interests; they may be assigned to the same or different groups.

$$absolute(i,i') = \left(a_1(i,i')+a_3(i,i')\right)(|M|-1), \forall i,i' \tag{18}$$

$$x(i,j)+x(i',j) \le absolute(i,i')+1, \forall i,i',j \tag{19}$$

Equation (20) is the complete MIP model that combines Equations (1) ~ (19).

max $\quad f$

s.t.
$$\prod_{m=1}^{|M|}(|M|+1-m)\, s(i,m) \ge f, \forall i$$
$$\sum_{m=1}^{|M|} s(i,m) = 1, \forall i$$
$$\sum_{i=1}^{|I|} g(i)x(i,j) \ge MinF, \forall j$$
$$\sum_{i=1}^{|I|} g(i)x(i,j) \le MaxF, \forall j$$
$$\sum_{i=1}^{|I|} n(i)x(i,j) \ge MinN, \forall j$$
$$\sum_{i=1}^{|I|} n(i)x(i,j) \le MaxN, \forall j$$
$$\frac{\sum_{i=1}^{|I|} gpa(i)x(i,j)}{|I|/|J|} \ge MinGPA, \forall j$$

$$x(i,j) + x(i',j) \le 2 - c(i,i'), \forall i, i', j$$
$$\sum_{j=1}^{|J|} x(i,j) = 1, \forall i$$
$$\sum_{i=1}^{|I|} x(i,j) = \frac{|I|}{|J|}, \forall j$$
$$\prod_{m=1}^{|M|} interest(i,m)\, s(i,m) - \prod_{m=1}^{|M|} interest(i',m)\, s(i',m) = a_1(i,i')(1 - |M|) + a_3(i,i')(|M| - 1), \forall i, i'$$
$$a_1(i,i') + a_2(i,i') + a_3(i,i') = 1, \forall i, i'$$
$$a_1(i,i') \le b_1(i,i'), \forall i, i'$$
$$a_2(i,i') \le b_1(i,i') + b_2(i,i'), \forall i, i'$$
$$a_3(i,i') \le b_2(i,i'), \forall i, i'$$
$$b_1(i,i') + b_2(i,i') = 1, \forall i, i'$$
$$absolute(i,i') = \left(a_1(i,i') + a_3(i,i')\right)(|M| - 1), \forall i, i'$$
$$x(i,j) + x(i',j) \le absolute(i,i') + 1, \forall i, i', j$$

$$s(i,m), x(i,j), x(i',j), b_1(i,i'), b_2(i,i') = 0 \text{ or } 1$$
$$a_1(i,i'), a_2(i,i'), a_3(i,i') \ge 0$$

$$i = 1, 2, \dots, |I|, i' = 1, 2, \dots, |I|, j = 1, 2, \dots, |J|, m = 1, 2, \dots, |M| \qquad (20)$$

III. FRAMEWORK OF the PROPOSED HETEROGENEOUS GROUPS FORMING ALGORITM

a. Dataset Description

Student's demographic information has been collected through a survey and face to face meetings. In the survey and meetings, students' interests and conflicts within other students have been asked and collected. Total 40 students are in the Capstone course and Table 1 shows demographic information for eight students. There are 26 males and 14 females, and 8 U.S. and 32 international students. Average GPA is 3.033. Each student has a list of five prioritized research interests, and conflicts with other students. Each student's prioritized research interests are assigned a score of 1 to 5; 5 being the highest score assigned to the most preferred interest.

TABLE 1: STUDENT'S DEMOGRAPHIC INFORMATION

i	$g(i)$	$n(i)$	$gpa(i)$	$interest(i,m)$					i', where $c(i,i') = 1$			
				$m=1$	$m=2$	$m=3$	$m=4$	$m=5$				
1	0	0	2.429	4	3	5	2	1				
2	0	1	2.750	1	3	5	4	2				
3	1	1	2.804	1	2	4	5	3	28	34	2	16
4	0	1	2.750	1	3	2	5	4				
5	0	0	3.891	4	3	2	1	5				
6	0	1	2.372	4	3	1	2	5	7	11	3	38
7	0	1	3.100	3	2	5	4	1				
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮				
40	0	1	2.782	4	3	5	2	1				

b. Formulation

Based on the data collected for the senior design project, there are total 40 students, $|I| = 40$. Ten groups are to be formed, $|J| = 10$. Each group has $\frac{|I|}{|J|} = 4$ students. Each student has five interests, $|M| = 5$. $MinF = 1$ and $MaxF = 2$. Since the majority of this particular class is international students, there are not enough U.S. students such that each group has at least one U.S. student. $MinN = 3$ ensures that no group has more than one U.S. student and the limited number of U.S. students may be evenly distributed among groups. $MaxN$ and its corresponding constraint are not used because $MaxN = 4$ and its corresponding constraint is redundant. $MinGPA = 3.0$, which is slightly less than 3.033, the grand average GPA of all 40 students. Equation (21) is the complete MIP model based on data for the senior design class.

max $\quad f$

s.t.
$$\prod_{m=1}^{5}(6-m)\, s(i,m) \geq f,\ \forall i$$
$$\sum_{m=1}^{5} s(i,m) = 1,\ \forall i$$
$$\sum_{i=1}^{40} g(i)x(i,j) \geq 1,\ \forall j$$
$$\sum_{i=1}^{40} g(i)x(i,j) \leq 2,\ \forall j$$
$$\sum_{i=1}^{40} n(i)x(i,j) \geq 3,\ \forall j$$
$$\frac{\sum_{i=1}^{40} gpa(i)x(i,j)}{4} \geq 3.0,\ \forall j$$
$$x(i,j) + x(i',j) \leq 2 - c(i,i'),\ \forall i, i', j$$
$$\sum_{j=1}^{10} x(i,j) = 1,\ \forall i$$
$$\sum_{i=1}^{40} x(i,j) = 4,\ \forall j$$
$$\prod_{m=1}^{5} interest(i,m)\, s(i,m) - \prod_{m=1}^{5} interest(i',m)\, s(i',m) = -4a_1(i,i') + 4a_3(i,i'),\ \forall i, i'$$
$$a_1(i,i') + a_2(i,i') + a_3(i,i') = 1,\ \forall i, i'$$
$$a_1(i,i') \leq b_1(i,i'),\ \forall i, i'$$
$$a_2(i,i') \leq b_1(i,i') + b_2(i,i'),\ \forall i, i'$$
$$a_3(i,i') \leq b_2(i,i'),\ \forall i, i'$$
$$b_1(i,i') + b_2(i,i') = 1,\ \forall i, i'$$
$$absolute(i,i') = 4\big(a_1(i,i') + a_3(i,i')\big),\ \forall i, i'$$
$$x(i,j) + x(i',j) \leq absolute(i,i') + 1,\ \forall i, i', j$$

$$s(i,m),\ x(i,j),\ x(i',j),\ b_1(i,i'),\ b_2(i,i') = 0 \text{ or } 1$$
$$a_1(i,i'),\ a_2(i,i'),\ a_3(i,i') \geq 0$$

$$i = 1, 2, \ldots, 40,\ i' = 1, 2, \ldots, 40,\ j = 1, 2, \ldots, 10,\ m = 1, 2, \ldots, 5 \qquad (21)$$

IV. RESULTS

The MIP model (Equation (21)) is implemented in the General Algebraic Modeling System (GAMS) and solved using the CPLEX solver. Table 2 shows the optimal solution to the MIP model. All students apply either the most preferred or second most preferred interest to complete the senior design project. The optimal value, i.e., the maximum $f = 4$. The group average GPA is between 3.006 (group 5) and 3.105 (group 1). Each group has at least one female student and at most two female students. Eight groups have one U.S. student in the group and two groups do not have any U.S. student. Students with conflicts are not in the same group. One issue is that there are 14 students who apply their second most preferred interest.

TABLE 2: OPTIMAL SOLUTION TO THE MIP MODEL IN EQUATION (21)

i	$g(i)$	$n(i)$	$gpa(i)$	Group j	$interest(i,m)$		i', where $c(i,i')=1$				
					$m=1$	$m=2$					
10	1	0	3.563	1	4						
17	0	1	3.231	1		5					
29	0	1	3.038	1	3		22	30	13	38	3
32	0	1	2.587	1	1		22	39	15		
Group Average GPA			**3.105**								
2	0	1	2.750	2		3					
8	0	1	2.519	2		1					
33	1	0	3.325	2		4					
39	0	1	3.492	2		5					
Group Average GPA			**3.022**								
14	0	1	2.771	3	1		22	13	11	2	
19	0	1	3.038	3	3		13	4			
30	1	1	2.654	3		5	6	8	22		
37	0	0	3.643	3	4						
Group Average GPA			**3.027**								
4	0	1	2.750	4	1						
5	0	0	3.891	4	4						
7	0	1	3.100	4		2					
22	1	1	2.577	4	3						
Group Average GPA			**3.080**								
3	1	1	2.804	5	1		28	34	2	16	
12	0	1	3.163	5		3	9	2	11		
20	1	1	3.231	5	5						
27	0	0	2.824	5	4						
Group Average GPA			**3.006**								
9	0	1	3.231	6	1		12	26	2	31	
21	0	1	3.288	6	5						
23	1	1	2.722	6		3	22	15	39		
40	0	1	2.782	6	4						
Group Average GPA			**3.006**								
6	0	1	2.372	7	4		7	11	3	38	
15	1	1	3.269	7	1						
25	0	1	3.327	7	2						
28	1	1	3.118	7	3						
Group Average GPA			**3.022**								
13	0	1	2.692	8	1		6	4	8		

18	0	0	3.250	8	4										
36	1	1	3.308	8		2									
38	1	1	2.757	8	3										
Group Average GPA			**3.022**												
1	0	0	2.429	9	4										
16	1	1	3.000	9		5									
26	1	1	3.308	9	1		9	11							
34	0	1	3.346	9	3		4	13	30	22	8				
Group Average GPA			**3.021**												
11	0	1	2.735	10		4	12	26	8	6	4	13	7	2	3
24	0	1	3.154	10	1		22	13	8	4	38	16	30		
31	1	1	3.182	10		2									
35	0	0	3.093	10		3									
Group Average GPA			**3.041**												

To reduce the number of students who apply their interest that is not their most preferred interest, the objective function of the MIP model may be revised (Equation (22)). Table 3 shows the optimal solution to the MIP model with the revised objective function. There are only seven students that apply their second most preferred interest (bold and italic) to complete the project. The MIP model is solved within one hour. This indicates that the MIP model is computationally efficiently for forming groups for a large class with many constraints.

$$\max \quad f - 10^2 \sum_{i=1}^{|I|} s(i,2) - 10^4 \sum_{i=1}^{|I|} s(i,3) - 10^6 \sum_{i=1}^{|I|} s(i,4) - 10^8 \sum_{i=1}^{|I|} s(i,5) \qquad (22)$$

TABLE 3: OPTIMAL SOLUTION TO THE REVISED MIP MODEL

	Group 1				**Group 2**				**Group 3**				**Group 4**				**Group 5**			
Student	10	13	21	22	3	19	32	37	2	29	31	35	9	20	38	40	4	5	28	30
Interest	4	1	5	3	***2***	3	1	4	1	3	***2***	4	1	5	3	4	1	4	3	***5***
	Group 6				**Group 7**				**Group 8**				**Group 9**				**Group 10**			
Student	1	7	26	34	6	15	24	33	11	18	23	25	12	16	17	27	8	14	36	39
Interest	4	***2***	1	3	4	1	***5***	3	3	4	1	2	1	3	***5***	4	3	1	***2***	4

V. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a mixed-integer programming (MIP) model for forming heterogeneous groups of students based on the mapping of students' gender (male, female), nationality (international or U.S.), grade point average (GPA), research interests (marketing, product design, project management, manufacturing, and finance), and their conflicts with other students. Each group has an average GPA similar to that of other groups. Also, each group has students with different research interests while avoiding conflicts between students. An experiment with 40 students was performed to demonstrate the scalability of the algorithm. Experiment results show that the proposed model is computationally efficient and provides the optimal solution within an hour. Qualitative assessments of group learning and performance have been performed. Students were asked to rate herself/himself and her/his team members on attributes in three categories (Table 4) from 1 (low) to 5 (high): professionalism, initiative, and effectiveness.

TABLE 4: DESIGN REVIEWS PEER/SELF EVALUATION

<table>
<tr><th>PROFESSIONALISM</th><th>INITIATIVE</th></tr>
<tr><td>• Prompt in attendance at team meetings.
• Complete in delivering agreed-upon parts of project.
• Organized in seeking information from resources.
• Volunteers appropriately during team meetings when tasks need to be accomplished.
• Pulls fair share with regard to overall work load.</td><td>• Develops ideas constructively with others.
• Makes helpful suggestions on ways of accomplishing projects.
• Is a good listener.
• Seeks input from quieter team members.</td></tr>
<tr><td colspan="2">EFFECTIVENESS
• Meets deadlines
• Demonstrates knowledge in the subject area.
• Able to solve problems.
• Effectively incorporates information presented in the course and from other experiences into the project.</td></tr>
</table>

Average scores for professionalism, initiative, and effectiveness are 4.2, 4.4, and 4.1 respectively. The students also provided additional comments. Two U.S. students commented that they felt disconnected during team meetings since international students spoke in languages other than English. Future research will consider forming either heterogonous or homogeneous groups for even larger classes, e.g., first-year and second-year undergraduate classes with 100 ~ 200 students, in various engineering disciplines and beyond.

REFERENCE

1. Nielsen, T., Hvas, A. E., & Kjaergaard, A., *Student team formation based on learning styles at university start: does it make a difference to the student?* Reflection Education, 2009. **5**(2): p. 85-103.
2. Dewiyanti, S., Brand-Gruwel, S., Jochems, W., & Broers, N. , *Students experiences with collaborative learning in asynchronous Computer-Supported Collaborative Learning environments.* Computers in Human Behavior, 2007. **23**(1): p. 496-514.
3. Francescato, D., Mebane, M. E., Porcelli, R., Attanasio, C., & Pulino, M. , *Developing professional skills and social capital through computer supported collaborative learning in university contexts.* International Journal of Human–Computer Studies, 2007. **65**(2): p. 140-152.
4. Francescato, D., Porcelli, R., Mebane, M., Cuddetta, M., & Renzi, P., *Evaluation of the efficacy of collaborative learning in face-to-face and computer-supported university contexts.* Computers in Human Behavior, 2006. **22**(2): p. 163-176.
5. Bekele, R., *Computer-assisted learner group formation based on personality traits*, in *Informatics*2005, University of Hamburg.
6. Graf, S., & Bekele, R. . *Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization*. in *Intelligent tutoring systems 8th international conference*. 2006. Jhongli, Taiwan.
7. Wang, D.Y., Lin, S. S. J., Sun, C. T. , *DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups.* Computers in human Behaviors, 2007. **23**(4): p. 1997-2010.
8. Cavanaugh, R., Ellis. M., Layton, R., & Ardis, M. , *Automating the Process of Assigning Students to Cooperative-Learning Teams*, in *2004 American Society for Engineering Education Annual Conference & Exposition*2004: Salt Lake City, Utah.
9. Moreno, J., Ovalle, D. A., Vicari, R. M., *A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics.* Computers & Education, 2012. **58**(1): p. 560-569.