



Efficiency and fairness criteria in the assignment of students to projects

Nahid Rezaeinia¹ · Julio César Góez¹ · Mario Guajardo¹

Accepted: 16 February 2021 / Published online: 24 April 2021
© The Author(s) 2021

Abstract

Teamwork has increasingly become more popular in educational environments. With the also increasing mobility trends in the educational sector, internationalization and other diversity features have gained importance in the structure of teams. In this paper, we discuss an assignment problem arising in the allocation of students to business projects in a master program in Norway. Among other problem features, the students state their preferences on the projects they most want to conduct. There are also requirements from the companies that propose the projects and from the program administration. We develop a bi-objective approach to consider efficiency and fairness criteria in this assignment problem. We test the model using real data of 2017 and 2018, in joint collaboration with the administrative staff of the program. In light of the good results, our proposed solutions have been implemented in practice in 2019 and 2020. The implementation of these solutions have been beneficial for the administration, the students, and the companies.

Keywords Project assignment · Multiobjective optimization · Efficiency · Fairness · OR in Practice · OR in Education

1 Introduction

The educational sector is experiencing a substantial growth in the last years. The UNESCO Institute for Statistics has forecast an increase from 5.3 million students to 8 million between 2017 and 2025. Likewise, teamwork has increasingly become more popular in educational environments (Hansen 2006). With the also increasing mobility trends in the educational sector, internationalization and other diversity features have gained importance in the structure of teams and teamwork (Kelly 2008). International mobility and increasing demand for

✉ Nahid Rezaeinia
nahid.rezaeinia@nhh.no

Julio César Góez
Julio.Goez@nhh.no

Mario Guajardo
Mario.Guajardo@nhh.no

¹ Department of Business and Management Science, NHH Norwegian School of Economics, 5045 Bergen, Norway

teamwork is particularly important in business environments, reflected in that many business schools have incorporated group projects and multicultural group working experiences into the curriculum, often in collaboration with companies and other organizations (Hansen 2006; Huxham and Land 2000). In general, collaboration in group projects is a good opportunity for students to learn teamwork, problem-solving, communication, leadership, and also to become more creative and more social in their future working places (Hansen 2006; Kelly 2008). In fact, it has been documented that positive group experiences help them to be more productive in industry settings and help them to succeed in their careers (Henry 2013). In this context, schools, universities and educational institutions play an important role in enabling teamworking experiences (Huxham and Land 2000).

The decision process of allocating students to groups and matching these groups with projects is often in charge of administrative staff at schools, who do not necessarily have the background and tools to generate a solution with all the desirable features. Moreover, as some criteria in the problem might be in conflict with each other, reconciling all of them into a single solution poses a challenge that can barely be addressed by manual techniques. A poor solution can have undesirable consequences not only in the specific features of the groups (such as gender balance and international composition), but also in the students' perception about the fairness of the decision process.

This paper reports the real-world implementation of a bi-objective modelling approach, to address one of such a problems arising in an international business master program in Norway. As a core activity of the program, the students have to complete a business project in collaboration with companies. These projects are posted in advance to the students, who then elicit a ranking of their preferred projects. The administration must form the groups and match them to the projects, taking into account the students' preferences, and specific requirements of the companies, among other aspects. The problem is a one-side preferences assignment while considering the other sides' requirements and conditions. Our bi-objective approach is based on integer programming, and incorporates concepts of efficiency and fairness regarding the preferences of the students. This approach has been implemented in practice during the last 2 years, replacing the traditional manual approach used before.

The remainder of this paper is organized as follows. Section 2 reviews relevant literature. Section 3 provides background on the applied setting that originates this work. Section 4 develops a series of mathematical modelling approaches for the problem, while Sect. 5 describes how these are used to compute a solution. Section 6 reports on the implementation in practice and numerical results. Concluding remarks are provided in Sect. 7.

2 Literature review

Random selection, self-selection and manually creating groups by a coordinator or instructor are common approaches to partition students into working teams. While these are easy to implement, they might fall into a number of shortcoming, such as lack of fit, isolation of some students, poor balance, and low level of satisfaction among students. These shortcomings have been addressed by a body of literature which finds its roots in classic assignment problems, including the celebrated work by Gale and Shapley (1962) on the admission of students to universities with two-sided preferences. In this section, we limit our review to papers in the education sector with the most similarities to our work, and to papers discussing efficiency and fairness criteria.

In the first stream, Krass and Ovchinnikov (2006) study the problem of assigning students to multiple non-overlapping groups considering diversity in skills, nationalities, genders, culture and academic backgrounds. They address the problem by an integer programming model, whose objective is to minimize the number of overlaps. The approach is applied to a case study arising at an MBA program in Canada. In another real-world case, Cutshall et al. (2007) introduces principles of equity and cohesion to form student teams in the core courses of a school of business in Indiana. Among other features, they attempt to match students with similar academic performance and also to avoid groups with lone female or lone international students (i.e., the number of female or international students in each group should be zero or at least two). They address the problem by an integer programming model, whose objective function is to minimize the maximum deviation of a team's academic performance.

Although both previous papers have successfully replaced manual methodologies easing the task of administrative staff and reporting positive results in practice, none of them considered the preferences of the students when forming the groups. In this respect, Lopes et al. (2008) develop a mixed integer programming model for the allocation of students to design projects, considering students preferences' and project sponsors' requirements, among other conditions. The model uses a single objective function that maximizes a weighted sum of the number of projects staffed, minus penalties for not satisfying students' preferences and other desirable conditions. The model is successfully applied to an engineering program at the University of Arizona. In a different problem, but of similar nature, Krauss et al. (2013) assign students to classes at an elementary school in New York City. In the assignment process, they considered the students designated friends list and also recommendations from parents, teachers and school therapists. As solution approach, they first use an integer programming formulation to construct a feasible solution. Then, they use a genetic algorithm to improve along criteria on the recommendations of parents and teachers, while penalizing the violations of the constraints originally considered in the integer model.

Although the previous papers somehow incorporate multiple criteria, they do it by combining these criteria into a single objective function. In what is the closest article to ours, Magnanti and Natarajan (2018) address the problem of assigning students to projects using two optimization criteria sequentially: efficiency and fairness. In this context, efficiency is understood as the maximization of the total utility, that is, the sum of the utilities of the projects assigned to students. The utility associated to the assignment of a student to a specific project, is calculated according to the ranking of preferences elicited by the students before the optimization process. Among all the efficient solutions, Magnanti and Natarajan (2018) then proceed with a lexicographic max–min fairness criterion, which consists of minimizing the number of students assigned to their least preferred project and then repeat the process with the second-to-last preference and so on. They apply their approach in an undergraduate program at a university in Singapore and report positive impact in practice. While our problem and approach are similar, some important features differ. First, due to requirements of the coordinators and partner organizations, we incorporate some side constraints that differ from the side constraints in their application. In particular, we aim at incorporating balance on gender and nationality, and also specific requirements from the partners (such as wished skills and languages). Second, in addition to the first efficiency and then fairness approach to compute a solution, we also test the inverted sequence, that is, using the fairness criterion first and then efficiency. Previous literature in other contexts points to the importance of analysing the trade-off between fairness and efficiency. A solution based only on efficiency may become unacceptable for some agents (students in our case), while a solution based only on fairness may incur in a high efficiency loss or high price of fairness (Bertsimas et al. 2012; Nicosia et al. 2017). Moreover, motivated by a literature stream using quantitative fairness measures,

we are also interested in quantifying fairness. Besides avoiding the iterative process required to incorporate the lexicographic max–min fairness criterion, a fairness measure allows an easy and direct way to compare different solutions. For this purpose, we adopt Jain's index (Jain et al. 1984), a quantitative measure of fairness with advantageous features of being population size-independent, unaffected by scale, bounded and continuous. A broad range of works have used this index to measure fairness, including Sediq et al. (2012), Huaizhou et al. (2013), Guo et al. (2013) and Hoßfeld et al. (2016). Furthermore, our contribution reports results on a real-world case arising in a master program at the main business school in Norway. Our approach has been implemented in practice during the last 2 years, producing results that increase the utility of the students and improve fairness over the previous manual approach. The new approach does not only contribute to the satisfaction among the student community, but also eases the task of the administrative staff in terms of resource usage and also in terms of projecting transparency and objectiveness about the decision process.

3 Background

The Global Alliance in Management Education, also called CEMS (because of its former name Community of European Management School and International Companies), is a cooperation founded in 1988 involving business schools, universities, companies, and non-governmental organizations (NGOs). It currently consists of about 100 members from the five continents, with 33 of them in the educational sector. Normally, at most one school per country can have a partnership in the alliance. From Norway, NHH Norwegian School of Economics has been member of CEMS since 1992. The flagship program of the alliance is the CEMS MIM or CEMS Master's in International Management, a 1-year program for students who are pursuing a Master's Degree at a CEMS member institution. Students who complete the program successfully receive one master's degree from CEMS, in addition to the master's degree from their home institution.

A central piece of the CEMS MIM curriculum requires students to complete a semester-long business project, in collaboration with one of the companies or NGOs that are partners of CEMS. The business project usually consists of a real problem faced by one of these partners, and the students are expected to address the problem and complete a final report that is submitted to the organization and to academic censors. Participating in a business project may play an important role in the future career of students, because it provides them with the chance to earn early experience as consultant or potentially to find a job position in the partner organization. Likewise, for these organizations it is valuable to get assigned a suitable team of students, who can address the challenges involved in the project and provide useful input. Therefore, the assignment of students to business projects is an important decision and is usually the responsibility of the administrative staff.

At NHH, the decision process consists of several stages. First, the companies prepare project proposals supported by the administrative staff and academic supervisors. Then, a full-day workshop is organized where the partner organizations present their projects to the students. After the workshop, each student sends a ranking with her/his five most preferred projects to the administrative staff. The staff also gathers other information about the students, such as gender, nationality, language skills, etc. Then, the staff attempts to create groups and assign them to projects, with the main goal of addressing the students' preferences. In addition, the staff considers other aspects, such as gender diversity, nationalities and language requirements. Until 2018, the assignment of students to projects at NHH was assigned manu-

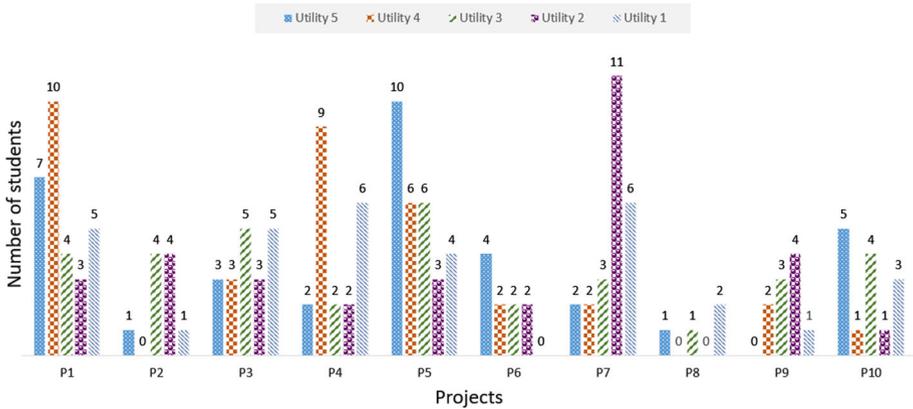


Fig. 1 Distribution of the students’ top-five preferences on 10 different projects

ally. Although the staff tried to incorporate as much of these criteria as possible, the problem posed some challenges that could barely be addressed by hand. In particular, it turned practically impossible to not disappoint some of the students who ended up in projects far from their top preferences. In contrast, as some others were assigned to their most preferred project, it became hard to avoid perceptions of unfairness in the process. The difficulties of the problem are illustrated in Fig. 1, which shows a real data instance on the preferences of the students. The example involves 35 students who ranked their top-five choices among 10 projects. The choices are labeled as *Utility 1*, ..., *Utility 5*, where for a given student a value of 5 indicates that the project is the most preferred by the student, a value of 4 indicates the second most preferred project, and so on. Projects 1 and 5 are clearly the most popular, with seven and ten students ranking them as their top choice, respectively. With a maximum allowed of four students in each of these projects, it turned impossible to assign all these students to their top choice. On the other hand, only one student ranked project 8 as top choice and, moreover, none of the students ranked project 9 as top choice. As a fairness criterion, the administration will do their best to hopefully not assign students to their lowest ranked or non-ranked projects. However, depending on the number of students and available projects, it might be necessary to assign a group of students even to these less popular projects. Since it is within the interest of the program to keep the partnership with the companies that offer the projects, the administrative and academic staff will help them to make the project proposals attractive to the students. Nevertheless, avoiding situations like the one described in this example is not guaranteed, because the profile of preferences is realized according to the individual rankings of the students after the projects proposals have been presented to them.

To better illustrate the potential conflict between efficiency and fairness, we may consider alternative assignments of the 10 students who ranked project 5 as their first choice. In one possible assignment we can assign four students to their top choice, two students to their second choice, and four students to their third choice. The total utility for this assignment is 40. In another possible assignment, we can assign one student to her/his top choice, seven students to their second choice, and two students to their third choice. The total utility for this assignment is 39, which means a lower performance on the efficiency measure than the previous assignment. However, under a fairness criterion of assigning as few as possible students to a less preferred project, the second assignment is better because it assigns only two students to the third choice (in contrast to the four students assigned to the third choice by

the first solution). Now consider a third alternative, in which we assign two students to their top choice, six students to their second choice, and two students to their third choice. The total utility for this assignment is 40, which is as good as the first solution according to the efficiency criterion. Likewise, the third assignment is equivalent to the second solution according to the fairness criterion, since both solutions assign two students to their third choice. Even though these measures of efficiency and fairness do not drive the solutions in completely opposed directions, this example illustrates that, in general, a more efficient assignment is not necessarily fairer neither the other way around. Also, the example illustrates that considering both criteria may conduce to solutions that are better than solutions constructed using a single criterion.

If on top of these two criteria, we add the other aspects considered by the administrative staff, it is understandable that computing a solution became a daunting task for them and that students would sometimes feel disappointed. This motivated us to undertake a collaboration with the administration as to support their decision-making process and to provide the study program with better solutions in practice. Our work involved testing several approaches, which we present in the following section.

4 Mathematical modelling

The foundation of the approaches developed in this paper is mathematical optimization. This section provides the details of the different components of the optimization formulations that will be used later to develop our multi-objective approaches.

To build the mathematical formulation, we denote the set of students participating in the program as $S = \{s_1, \dots, s_n\}$, the set of projects proposed by participating companies as $\mathcal{P} = \{p_1, \dots, p_m\}$, and the set of required attributes as $\mathcal{C} = \{c_1, \dots, c_l\}$. Here, the attributes are for example expertise fields, skills, nationality and gender, and they will be used to model the criteria for the assignment of students to different groups. Using these sets, we define the parameters and decision variables needed to build the mathematical optimization model for the projects assignment problem. For the sake of simplicity, in the reminder of this section we use the indices s for an arbitrary student in S , p for an arbitrary project in \mathcal{P} , and c for an arbitrary attribute in \mathcal{C} .

First, we define the following parameters:

- u_{sp} : Defines the utility of a student s when assigned to a project p .
- a_{sc} : Defines the presence of an attribute c in a student s . This is a binary parameter, i.e., it is one if a student has an attribute c and zero otherwise.
- UB_p : Upper bound on the number of students that are needed for a project p .
- LB_p : Lower bound on the number of students that are needed for a project p .
- UB_{pc} : Upper bound on the number of students with attribute c that are needed for a project p .
- LB_{pc} : Lower bound on the number of students with attribute c that are needed for a project p .

The model has to decide which projects to assign and which students to match with those projects. We define two different sets of decision variables. First, in (1) we define a set of binary variables to decide the projects that will be assigned.

$$y_p = \begin{cases} 1, & \text{if project } p \text{ is selected,} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The variable y_p allows to disregard some projects. This situation may happen due to the constraints on the upper and lower number of students needed for each projects. It may occur that none of the students in S are able to satisfy some projects' requirements, or the ranking of some projects may be consistently lower than other projects, which may rule it out when balancing efficiency and fairness. Second, in (2) we define a set of binary variables used to decide the assignment of students to the different projects.

$$x_{sp} = \begin{cases} 1, & \text{if student } s \text{ is allocated to project } p, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

4.1 The assignment constraints

The initial problem of assigning the students to the different projects is defined by the assignment constraints (3)–(7).

$$\sum_p x_{sp} = 1 \qquad \forall s \in S \tag{3}$$

$$x_{sp} \leq y_p \qquad \forall s \in S, \forall p \in P \tag{4}$$

$$\sum_s x_{sp} \geq LB_p y_p \qquad \forall p \in P \tag{5}$$

$$\sum_s x_{sp} \leq UB_p y_p \qquad \forall p \in P \tag{6}$$

$$x_{sp}, y_p \in \{0, 1\} \qquad \forall s \in S, \forall p \in P \tag{7}$$

Constraints (3) ensure that each student is assigned to a single project. Constraints (4) ensure that a project p is selected if at least one student is assigned to it. Constraints (5) and (6) enforce the upper and lower bounds on the number of students that are needed for each project. Finally, constraints (7) limit the decision variables to be binary.

4.2 Side constraints

One of the strengths of using mathematical optimization as the foundation of our approach is the possibility of including side constraints, as it is also highlighted in Magnanti and Natarajan (2018). Side constraints appear when one has some additional conditions beyond ensuring a proper assignment and enforcing the limits on the number of students per project. Side constraints may for example aim at obtaining balance on gender and nationality, as well as ensuring that specific requirements from the partners are met. Those requirements usually cover wished skills and languages. The constraints (8) and (9) are used here to model such requirements. However, this approach does not limit the form of the side constraints that may be considered. In general, the approach of this paper allows for side constraints of any form.

$$\sum_s a_{sc} x_{sp} \geq LB_{pc} y_p \qquad \forall p \in P, \forall c \in C \tag{8}$$

$$\sum_s a_{sc} x_{sp} \leq UB_{pc} y_p \qquad \forall p \in P, \forall c \in C \tag{9}$$

Constraints (8) and (9) ensure the lower and upper bounds on the number of students with specific attributes needed for each project are satisfied.

4.3 Measuring the quality of an assignment

The main aim of the approach of this paper is to achieve a balance between efficiency and fairness. For that purpose, two different goals are defined, which are formulated using linear functions as follows.

To optimize efficiency, the linear function defined in Eq. (10) is used. Specifically, Eq. (10) maximizes the overall utility of an assignment measured as the summation of the students utility obtained when assigned to a project times the assignment variables.

$$U = \max \sum_s \sum_p u_{sp} x_{sp} \tag{10}$$

To work towards fairness, the linear function defined in Eq. (11) is used. Here, a utility level is fixed at \hat{u} and then it minimizes the number of students that will receive such utility. In the context of this work, the utilities are assigned based on the rank of a project. For example, if a project was ranked as k by a student in the preferences survey, then it provides that student with a utility of u_k .

$$F_k = \min \sum_P \sum_{s:u_{sp}=\hat{u}} x_{sp} \tag{11}$$

Equation (11) is used later in a sequential process to minimize the number of students who are assigned to projects that are not their top choice.

4.3.1 Measuring fairness

A limitation when using (11) is that it does not provide an overall measure of the achieved fairness of the reached assignment. To overcome that limitation, the Jain’s index (Jain et al. 1984) is used as an alternative to measure and optimize the assignment fairness. The Jain’s index formulation is provided in (12).

$$f(w) = \frac{(\sum_{i=1}^n w_i)^2}{n \sum_{i=1}^n (w_i)^2} \quad w_i \geq 0 \tag{12}$$

In (12) $w \in \mathbb{R}^n$, n is the number of participants, and w_i is the allocation given to the i -th user in a system. The Jain’s index (12) has the desired properties of population size independence, scale and metric independence, boundedness, and continuity. The index is broadly used to measure fairness in the assignment of resources in telecommunication networks, but it may also have applications in other areas. In particular, in this assignment problem the Jain’s index is computed using the utilities and assignment decisions, as shown in (13).

$$f(x) = \max \frac{(\sum_s \sum_p u_{sp} x_{sp})^2}{n \sum_s (\sum_p u_{sp} x_{sp})^2} \tag{13}$$

In (13), given the constraint (3), the term $\sum_p u_{sp} x_{sp}$ provides the utility of the assignment for a student s to a project. Hence, given a total utility for an assignment, the index (13) provides the assignment fairness measure for that level of utility. In other words, this index provides a fairness measure for the utility achieved when the assignment efficiency is optimized.

The Jain’s index is bounded by values that depend only on the number of participants. First, by definition the index is bounded above by 1, since $(\sum_s \sum_p u_{sp} x_{sp})^2 \leq n \sum_s (\sum_p u_{sp} x_{sp})^2$.

Second, the lower bound for the index is $1/n$, which happens when only one participant is assigned the total utility. As an illustration of this lower bound, consider a case where only one student is assigned to one of the projects in her/his preference list and all others are assigned to projects outside their preference lists. In this solution, the former student receives the total utility obtained with the assignment while the others obtain zero of that utility. The fairness index evaluated at this solution achieves its lower bound $1/n$. Notice that for the index to take a value below $1/n$, one would need that $(\sum_s \sum_p u_{sp} x_{sp})^2 \leq \sum_s (\sum_p u_{sp} x_{sp})^2$, which is mathematically not possible. The dependence only on the number of participants provides bounds that are known beforehand and can potentially be used by optimization solvers to speed up the solution process.

5 Lexicographic solution approaches

To optimize the projects assignment problem, a lexicographic approach is used to prioritize the different goals. For this, the required optimization models are build using the elements introduced in Sect. 4. The assignment constraints in Sect. 4.1 and the side constraints in Sect. 4.2 ensure a valid assignment. Then, the different possible objective functions are prioritized in different orders with the aim to research the effect choosing one measure over the other.

5.1 Prioritizing efficiency

When priority is given to efficiency over fairness the following orders is used. First, the problem (14) is solved to optimize the overall efficiency of the assignment.

$$\begin{aligned} \max \quad & \sum_s \sum_p u_{sp} x_{sp} \\ \text{s.t.} \quad & (3)–(9) \end{aligned} \tag{14}$$

Second, using the optimal efficiency of the first step, denoted by U^* , fairness is improved. Optimizing the utility function may result in multiple optimal solutions with the *Pareto Efficient* property. A Pareto Efficient solution cannot be more efficient regarding someone's assignment unless fairness is reduced in someone's else assignment (Magnanti and Natarajan 2018). Incorporating fairness aims to keep students' assignments as fair as possible, while keeping the same maximum utility. Two approaches are considered here to incorporate fairness.

5.1.1 Lexicographic fairness

First, a lexicographic approach is used to improve fairness while enforcing the same efficiency level obtained when solving (14). In this approach, the first step is to solve Problem (15) to minimize the number of students assigned to projects ranked 1. To ensure the utility is not worsening, the utility is constrained to be at least as good as the one obtained solving (14).

$$\begin{aligned}
 & \min \sum_p \sum_{s:u_{sp}=1} x_{sp} \\
 & \text{s.t. (3)–(9)} \\
 & \sum_s \sum_p u_{sp}x_{sp} \geq U^*
 \end{aligned} \tag{15}$$

Using the solution of (15), the sequence of optimization problems (15) is solved. The sequence is obtained by changing the value of $k \in \{2, \dots, K - 1\}$. Here, K is the highest ranking a student can assign to a project. In other words, in each step the number of students assigned to a project with ranking k is minimized, starting with the lowest ranking and moving up one ranking at a time until the level before the maximum. That minimization is constrained to ensuring that at most F_ℓ^* students are assigned to a project with a rank $\ell \in \{1, \dots, k - 1\}$. Here, F_ℓ^* is the optimal value of the objective function at iteration ℓ in the sequence.

$$\begin{aligned}
 & \min \sum_p \sum_{s:u_{sp}=k} x_{sp} \\
 & \text{s.t. (3)–(9)} \\
 & \sum_s \sum_p u_{sp}x_{sp} \geq U^* \\
 & \sum_p \sum_{s:u_{sp}=\ell} x_{sp} \leq F_\ell^* \quad \forall \ell \in \{1, \dots, k - 1\}
 \end{aligned} \tag{16}$$

5.1.2 Jain’s index

The second approach uses the Jain index to improve fairness. Here, the Problem (17) is solved to maximize the Jain’s index while enforcing the same efficiency level found solving (14).

$$\begin{aligned}
 & \max \frac{(\sum_s \sum_p u_{sp}x_{sp})^2}{n \sum_s (\sum_p u_{sp}x_{sp})^2} \\
 & \text{s.t. (3)–(9)} \\
 & \sum_s \sum_p u_{sp}x_{sp} \geq U^*
 \end{aligned} \tag{17}$$

5.2 Prioritizing fairness

When priority is given to fairness over efficiency the approaches discussed in Sect. 5.1 are inverted. This results in two different approaches.

5.2.1 Lexicographic fairness

The first approach uses a lexicographic approach to sequentially improve fairness. The aim is to minimize the number of students that are assigned to low ranked projects. First, the Problem (18) is solved to minimize the number of students assigned to projects ranked 1.

$$\begin{aligned} \min \sum_p \sum_{s:u_{sp}=1} x_{sp} \\ \text{s.t. (3)–(9)} \end{aligned} \tag{18}$$

Using the solution of (18), the sequence of optimization problems (19) is solved.

$$\begin{aligned} \min \sum_p \sum_{s:u_{sp}=k} x_{sp} \\ \text{s.t. (3)–(9)} \\ \sum_p \sum_{s:u_{sp}=\ell} x_{sp} \leq F_\ell^* \quad \forall \ell \in \{1, \dots, K - 1\} \end{aligned} \tag{19}$$

Using the sequentially optimized fair assignment, the fairness level obtained is used as a lower bound to optimize efficiency in (20).

$$\begin{aligned} \max \sum_s \sum_p u_{sp}x_{sp} \\ \text{s.t. (3)–(9)} \\ \sum_p \sum_{s:u_{sp}=\ell} x_{sp} \leq F_\ell^* \quad \forall \ell \in \{1, \dots, K - 1\} \end{aligned} \tag{20}$$

5.2.2 Jain’s index

The second approach uses the Jain’s index to optimize fairness first. The Jain’s index is maximized solving Problem (21).

$$\begin{aligned} \max \frac{(\sum_s \sum_p u_{sp}x_{sp})^2}{n \sum_s (\sum_p u_{sp}x_{sp})^2} \\ \text{s.t. (3)–(9)} \end{aligned} \tag{21}$$

The optimal value of the Jain’s index found, denoted as J^* , is used as a lower bound for fairness to optimize efficiency in (22).

$$\begin{aligned} \max \sum_s \sum_p u_{sp}x_{sp} \\ \text{s.t. (3)–(9)} \\ \frac{(\sum_s \sum_p u_{sp}x_{sp})^2}{n \sum_s (\sum_p u_{sp}x_{sp})^2} \geq J^* \end{aligned} \tag{22}$$

6 Implementation and numerical results

The approaches proposed in Sect. 5 were implemented and tested with the data instances of 2017 and 2018, after the assignment through the old manual methodology had taken place. In light of the good results for those experimental instances, we supported the administration

Table 1 Overview of number of students, projects and required attributes

	2017	2018	2019	2020
Number of students	35	39	34	33
Number of projects	10	11	9	9
Number of requested attributes	10	21	3	9

to conduct the assignment in 2019 and 2020 by putting this new methodology into practice. In this section we summarize our numerical results.

6.1 Overview

The input for the problem are the projects' descriptions provided by the companies, the administration's requests, and the students rankings. With the information in place, we identify the different attributes and categories. These may vary from one to another year, but in general, we identify five main categories. First we have *profile*, which refers to a student's interests or field. Second, we consider the *language skills*, because for some projects it is important to have at least one student able to speak a specific language. Third, we consider *nationality*. Some projects need students with specific nationalities to facilitate the collaboration with a partner company taking part in the project. Fourth, we consider *gender* balance, responding to the requirement of the administration. What balance means precisely, depends on the gender distribution of the class, but for purpose of our formulation this translates into lower and upper bounds on the number of students of a specific gender assigned to each of the allocated projects. Fifth, we include a category named *home requirement*, which requires assigning at least one student from NHH to each group.

Table 1 provides an overview of the number of students and projects that participated in the business project program and the number of requested attributes during the years 2017 to 2020.

As for the preference survey, at the beginning of the semester, students taking part at the business project semester at NHH are asked to rank up K projects. To illustrate, Table 2 shows a partial overview of the results of the students' preferences survey for the year 2017, where $K = 5$. A student s that ranked a project p in its list will have an utility u_{sp} equal to the rank given to that project. If a project p was not ranked by the student s , it will be assigned a utility $u_{sp} = 0$.

6.2 Numerical results

The optimization models proposed in Sect. 5 were implemented in AMPL. The linear models were solved using CPLEX, version 12.8. For the non-linear models, we used BARON version 18.12.26. Additionally, to speed up BARON solution times, we provided the solver with the Jain's index bounds presented in Sect. 4.3.1. The computational runs are set to a time limit of one hour. Each of the instances from 2017 to 2020 were run using all the four proposed approaches. For the remaining of this section we use the following abbreviations to identify each of the approaches:

- PELF: first Prioritize Efficiency and then optimize using the Lexicographic Fairness;
- PEJI: first Prioritize Efficiency and then optimize fairness using the Jain's index;

Table 2 Partial overview of students' preferences in 2017

Students	Projects									
	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}
s_1	4				3	5	1		2	
s_2	3			4	5		2	1		
s_3		2	3	4	5					1
s_4	4			2	3	5	1			
s_5		2	5	3	4			1		
s_6	4					5	1	3	2	
s_7	5				2		1		4	3
s_8	1		2	4	5		3			
s_9	4	2		1	3					5
s_{10}	3		1	5	4		2			

Table 3 Summary of assigned projects

Year	Assigned projects
2017	$p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_{10}$
2018	$p_1, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{11}$
2019	$p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9$
2020	$p_1, p_2, p_3, p_4, p_5, p_7, p_8, p_9$

- PLFE: first Prioritize fairness using the Lexicographic Fairness and then optimize Efficiency;
- PJIE: first Prioritize fairness using the Jain's index and then optimize Efficiency.

6.2.1 Experimental results

In Table 3 we summarize the projects selected each of the years 2017, 2018, 2019 and 2020. For each year, all the approaches returned the same set of projects. For the sake of clarity, note that each year has a list of projects that is different and independent from any other year. Hence, the list of projects for each year is unique and valid only for that year. Moreover, all approaches disregard some projects due to the low preferences and the few number of students that ranked them. For example, from the distribution of students' preferences in the year 2017 shown in Fig. 1, the projects 8 and 9 were in the situation described. In particular, note that only one student ranked project 8 as first choice and no one did it for project 9, which resulted in those two projects being excluded from the final assignment.

Table 4 summarizes the experimental results for 2017. The last column shows the assignment that was done manually in 2017 by the CEMS administration staff at NHH. The results for 2017 reached different assignments with small difference in total utilities and fairness. In all these assignments, 100% of the students are assigned to their top three choices. All the approaches reached an optimal solution within few seconds. Note that the approaches PELF and PLFE are based on integer linear models, while the approaches PEJI and PJIE are based on mixed integer non-linear models. Compared to the manual assignment, none of the solutions obtained have students assigned to their two bottom choices, while the manual assignment that was used in 2017 had 9% of students assigned to their two bottom choices.

Table 4 Efficiency and fairness results for 2017 using all the proposed approaches

	PELF	PLFE	PEJI (Percentage of students assigned)	PJIE	Manual assignment
Utility 5	68%	63%	68%	63%	51%
Utility 4	26%	34%	26%	34%	17%
Utility 3	6%	3%	6%	3%	23%
Utility 2	0	0	0	0	3%
Utility 1	0	0	0	0	6%
Utility 0	0	0	0	0	0
Total utility	162	161	162	161	142
Fairness	98.4%	98.61%	98.4%	98.61%	92.3%

Table 5 Details of the PELF solutions for the 2017 instance

	Max efficiency	Min assignments with utility 3 (Number of students assigned)	Min assignments with utility 4
Utility 5	25	24	24
Utility 4	7	9	9
Utility 3	3	2	2
Utility 2	0	0	0
Utility 1	0	0	0
Utility 0	0	0	0
Total utility	162	162	162

The two best solutions for the 2017 instance were found by the approaches PELF and PLFE. First, recall that the approach PELF prioritize efficiency by maximizing the total sum of the student's utilities. The solution obtained maximizing utility assigns 25 students to their top choice, 7 to their second ranked choice, and 3 to their third ranked choice. Given that there are not students ranked in the three bottom choices, the lexicographic fairness process is initialized constraining the lower ranked projects assignments to zero. Then, it starts minimizing the number of students assigned to their third ranked choice. After that process is finished, the solution found has 24 students assigned to their top choice, 9 assigned to their second ranked choice, and 2 assigned to their third ranked choice. To summarize, the consideration of fairness reduced the number of students assigned to the top choice and third ranked choices by one and two correspondingly, and increased the number of students assigned to their second ranked choice by 2. Notice that the PELF approach ensures that the two solutions will have the same utility of 162. Hence, fairness is optimized over the set of optimal solutions to the problem optimizing efficiency. Hence, the solution found is Pareto efficient. The details of these results are summarized in Table 5.

Second, the PLFE approach is used to prioritize fairness. This approach starts minimizing the assignments to projects ranked below the top choice. In the first iteration, the assignments of students to projects with utility zero is minimized. After that iteration, the solution obtained has 8 students assigned to the their top choice, 4 to their second ranked choice, 13 to their third choice, and 10 to their fourth choice. There are no students assigned to their bottom choice or to a project that they did not rank. The process continues until it minimizes the number

of students assigned to their second choice. In every iteration, the optimization continues to satisfy the assignment levels in the lower ranked projects found in previous iterations. The final solution has 22 students assigned to their top choice, 12 to their second ranked choice, and one to their third choice. There are no students assigned to their two bottom choices or to a project that they did not rank. After the last fairness iteration, the approach optimizes efficiency constrained to ensure the fairness level achieved in each iteration of the lexicographic fairness process. Notice that, in general, given the utility structure of our instances at the end of the lexicographic fairness, the overall utility is fixed and optimizing efficiency will not change the solution. This is reflected in the results summary presented in Table 6. Comparing the solutions found with the two approaches, PEFL and PLFE, there is a trade off between efficiency and fairness. As expected, PEFL yields a higher overall utility, while PLFE delivers a solution with less students assigned to their third ranked project. Here, the difference between the utility achieved when efficiency is prioritized and the utility achieved when fairness is prioritized provides the price of fairness, which is equal to one utility unit in this instance.

The numerical results of the experiment using the 2018 instance are summarized in Table 7. All approaches found an optimal solution with the same level of utility and fairness. The optimal solutions found assigned all students to one of their two top choices. Also, the solutions found for the 2018 instance improves both utility and fairness when they are compared to the solution manually obtained by the staff. In addition, the solutions obtained with our approaches satisfied all the companies' requirements, while the manual assignment did not manage to satisfy all the requirements. That highlights the complexity of the problem, which was not trivial to handle manually and took many hours of work for the administration staff. However, with these optimization based approaches the solution obtained was compliant and it took only a few seconds to find it.

6.2.2 Implementation results

For the years 2019 and 2020, the administrative staff has used the optimization based approach proposed in this paper. The results obtained are summarized in Table 8, which includes the performance of the assignments proposed for 2019 and 2020. In both years 2019 and 2020 all approaches were able to reach optimality within the same time limit. In particular, the PJIE approach took about 6 min to obtain an optimal solution, which was the longest time. From Table 1 we can see that the conditions for the CEMS assignment problem may change significantly from year to year. Indeed, the assignment problems faced in 2017 and 2018 were more demanding on the number of requested attributes by the projects. Note that in 2020, the students ranked up four projects and a project with rank 4 is the topmost preference.

The results in Table 8 reveal that in the year 2019 all approaches except PJIE led to the same solution, while in 2020 all approaches led to the same solution. In both years all the students have been assigned to some of the projects stated in their top choices, and with the great majority assigned to one of their two most preferred choices. In practice, it has been positive for the administration being able to verify that the solution so obtained has such a performance along the different criteria and approaches. Being supported by optimization techniques has also improved the ability to conduct the process in a fast and objective way. In addition, both sides students and companies can rely now in that the decision process provides them with an assignment that take into account all the requirements and preferences.

As the numerical results of the implementation are affected by how the preferences of the students realized in practice, it is interesting to analyze other scenarios for comparison purposes. In particular, we may construct *best-case* and *worst-case* scenarios as a referential

Table 6 Details of the PLFE solutions for the 2017 instance

	Min assignments with utility 0	Min assignments with utility 2	Min assignments with utility 3 (Number of students assigned)	Min assignments with utility 4	Max efficiency
Utility 5	8	9	18	22	22
Utility 4	4	10	16	12	12
Utility 3	13	16	1	1	1
Utility 2	10	0	0	0	0
Utility 1	0	0	0	0	0
Utility 0	0	0	0	0	0
Total utility	115	133	157	161	161

Table 7 Efficiency and fairness results for 2018 using all the proposed approaches

	PELF	PLFE	PEJI (Percentage of students assigned)	PJIE	Manual assignment
Utility 5	71%	71%	71%	71%	38%
Utility 4	29%	29%	29%	29%	38%
Utility 3	0%	0%	0%	0%	21%
Utility 2	0	0	0	0	3%
Utility 1	0	0	0	0	0%
Utility 0	0	0	0	0	0
Total utility	184	184	184	184	161
Fairness	99.09%	99.09%	99.09%	99.09%	96%

Table 8 Implementation results for 2019 and 2020

	2019				2020			
	PELF	PLFE	PEJI	PJIE	PELF	PLFE	PEJI	PJIE
	(Percentage of students assigned)							
Utility 5	60%	60%	60%	3%	–	–	–	–
Utility 4	36%	36%	36%	76%	70%	70%	70%	70%
Utility 3	2%	2%	2%	18%	27%	27%	27%	27%
Utility 2	2%	2%	2%	3%	3%	3%	3%	3%
Utility 1	0	0	0	0	0	0	0	0
Utility 0	0	0	0	0	0	0	0	0
Total utility	153	153	153	129	121	121	121	121
Fairness	97.65%	97.65%	97.65%	98.08%	97.93%	97.93%	97.93%	97.93%

basis to compare the realized solution. In a best-case scenario, the preferences of the students would be such that everyone is allocated to her/his most preferred choice. It is easy to find a profile of preferences for this best-case scenario, by simply finding a feasible solution to the assignment problem with side-constraints. If a feasible solution exists, one can define the topmost preference of each student as the project to which she/he is assigned in this solution, to render an overall solution where the total utility is equal to the number of students multiplied by the highest utility. Note this solution also achieves the maximum of 100% in the fairness index (since everyone is assigned to a project ranked at the same level). In our case, we have verified that all approaches quickly reach such idealistic best solution, which in 2019 corresponds to a total utility equal to 170. Comparing to the results displayed in Table 8, we can see that the realized scenario of preferences allows to find a solution which reaches about 90% of the total utility of the best-case scenario. For the 2020 instance, the idealistic solution scores 132 in the total utility criterion, while the solution to the realized scenario reported in Table 8 scores 121, that is, about 92% in comparison to the best-case scenario. Note in the best-case scenario the preferences of the students are perfectly split among the assigned projects, which makes possible to provide all students with the highest utility. This involves some heterogeneity in the preferences of the students who are assigned to different projects (and homogeneity in the preferences of the students assigned to a same project). In contrast, we may think of a case where the preferences of all the students are

fully homogeneous, that is, one of the projects is ranked as top-choice by all the students, another project is ranked as second choice by all the students, and so on. We could regard this as a worst-case scenario, in the sense that some few students will be assigned to their ranked projects, while all the others will be assigned to a non-ranked project. When running this scenario for our 2019 and 2020 data instances, the solution assigns between 12 and 15% of the students to a ranked project, and 40% of the students to a non-ranked project. The total utility is between 50 and 60, and the fairness between 48 and 50%. The large spread between these results and the results obtained for the best-case scenario reveals that the performance of the solutions may vary within a broad range. Moreover, note that an even worse scenario could be constructed by digging deeper into the constraints that enforce some students to be assigned to a project because of the required attributes (for example, when a project requires a German speaker and the only students who speak that language did not rank that project among their top choices). In this way, some data instances could eventually lead to solutions with total utility equal to zero. Fortunately, as shown by our results above, the scenarios realized in practice have allowed us to find solutions much closer to the best-case than the worst-case scenario.

7 Concluding remarks

This paper proposed a decision support tool for deciding the assignments of students to projects, taking into account the students' preferences and other problem requirements. Since, in general, it is not possible to assign all students to their most preferred project, we have developed a bi-objective integer optimization approach taking into account efficiency and fairness criteria. We have implemented the approaches to support the decision-making process of the administration at an international master program. Our solutions have been adopted in practice during the last 2 years and the plan is to continue with this application in the forthcoming years. The results are positive for the students and the companies, as well as the decision process turns easier to handle for the administration. Since the backgrounds of these stakeholders do not necessarily include optimization, our close collaboration with the administration has facilitated their understanding and the adoption of the new methodology.

Although our focus has been practice-oriented and in the specific case of a master program at NHH, our mathematical formulations allow flexibility to consider a diverse set of requirements and side constraints that may be needed in different setups, such as similar programs at other universities or institutions. Likewise, our work opens avenues to conduct future research involving more methodological aspects. For example, to the best of our knowledge, this is the first time that the Jain's index has been used to measure the fairness in an assignment problem involving people. While the index has appealing properties and provides a single measure of fairness (in contrast to the lexicographic order), it inherently involves a non-linear expression that affects the nature of the problem. A possible reformulation of the optimization models involving the Jain's index may lead to more efficient solution approaches. Also, studying different structures of preferences in a computational study remains of interest not only to investigate the computational performance of the approaches but also to analyze how they balance the trade-off between efficiency and fairness. Moreover, studying different utility functions and different measures of fairness are also topics for further research.

Acknowledgements The authors would like to express their gratitude to Marie Van Beijeren, Senior Consultant of the CEMS at NHH, for her assistance in data acquisition and support in writing this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bertsimas, D., Farias, V. F., & Trichakis, N. (2012). On the efficiency-fairness trade-off. *Management Science*, 58(12), 2234–2250.
- Cutshall, R., Gavrimeni, S., & Schultz, K. (2007). Indiana University's Kelley School of business uses integer programming to form equitable, cohesive student teams. *Interfaces*, 37(3), 265–276.
- Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1), 9–15.
- Guo, C., Sheng, M., Zhang, Y., & Wang, X. (2013). A Jain's index perspective on α -fairness resource allocation over slow fading channels. *IEEE Communications Letters*, 17(4), 705–708.
- Hansen, R. S. (2006). Benefits and problems with student teams: Suggestions for improving team projects. *Journal of Education for Business*, 82(1), 11–19.
- Henry, T. R. (2013). Forming productive student groups using a massively parallel brute-force algorithm. *Proceedings of the World Congress on Engineering and Computer Science*, 1, 23–25.
- Höbfeld, T., Skorin-Kapov, L., Heegaard, P. E., & Varela, M. (2016). Definition of QoE fairness in shared systems. *IEEE Communications Letters*, 21(1), 184–187.
- Huaizhou, S., Prasad, R. V., Onur, E., & Niemegeers, I. (2013). Fairness in wireless networks: Issues, measures and challenges. *IEEE Communications Surveys & Tutorials*, 16(1), 5–24.
- Huxham, M., & Land, R. (2000). Assigning students in group work projects: Can we do better than random? *Innovations in Education and Training International*, 37(1), 17–22.
- Jain, R. K., Chiu, D.-M. W., & Hawe, W. R. (1984). *A quantitative measure of fairness and discrimination*. Hudson: Eastern Research Laboratory, Digital Equipment Corporation.
- Kelly, P. (2008). Achieving desirable group-work outcomes through the group allocation process. *Team Performance Management*, 14(1/2), 22–38.
- Krass, D., & Ovchinnikov, A. (2006). The University of Toronto's Rotman School of Management uses management science to create MBA study groups. *Interfaces*, 36(2), 126–137.
- Krauss, B., Lee, J., & Newman, D. (2013). Optimizing the assignment of students to classes in an elementary school. *INFORMS Transactions on Education*, 14(1), 39–44.
- Lopes, L., Aronson, M., Carstensen, G., & Smith, C. (2008). Optimization support for senior design project assignments. *Interfaces*, 38(6), 448–464.
- Magnanti, T. L., & Natarajan, K. (2018). Allocating students to multidisciplinary capstone projects using discrete optimization. *Interfaces*, 48(3), 204–216.
- Nicosia, G., Pacifici, A., & Pferschy, U. (2017). Price of fairness for allocating a bounded resource. *European Journal of Operational Research*, 257(3), 933–943.
- Sedıq, A. B., Gohary, R. H., & Yanikomeroglu, H. (2012). Optimal tradeoff between efficiency and Jain's fairness index in resource allocation. In *2012 IEEE 23rd international symposium on personal, indoor and mobile radio communications-(PIMRC)* (pp. 577–583). IEEE.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.