The Eighth International Symposium on Operations Research and Its Applications (ISORA'09)
Zhangjiajie, China, September 20–22, 2009
Copyright © 2009 ORSC & APORC, pp. 75–82

Multi-criteria Student Project Allocation: A Case Study of Goal Programming Formulation with DSS Implementation

Li Pan[1] Sydney C. K. Chu[1] Guangyue Han[1]
Joshua Z. Huang[2]

[1]Department of Mathematics, University of Hong Kong, Hong Kong, China
[2]E-Business Technology Institute, University of Hong Kong, Hong Kong, China

Abstract This paper concerns a typical Student Project Allocation (SPA) problem involving distributing a set of projects to students of an undergraduate "Directed Studies in Mathematics" course at the Department of Mathematics of the University of Hong Kong. Among a set of projects, each student indicates a preference list over their eligible projects, while the Department wants to make the most allocations, with its preferences over the individual students according to their GPA's. We apply pre-emptive goal programming to a multi-criteria SPA model for allocating these projects to students with a DSS implementation. The numerical results illustrate the clear effectiveness and efficiency of this approach.

Keywords Student Project Allocation; Goal Programming; Decision Support System (DSS)

1 Introduction

In order to give a student an early experience on independent study, with the opportunity to do a small mathematics project close to research, the Department of Mathematics at the University of Hong Kong offers an undergraduate course "Directed Studies in Mathematics". For this course, the student is expected to do approximately 100 hours of independent work and attend meetings and seminars. And by the end of the course, the student would hand in a dissertation and may give an oral presentation as appropriate. For each academic year, the Mathematics teachers prepare a list of project briefs for the mathematics majors, who rank their eligible projects in order of their preferences. Each project is supervised by one supervisor and is allocated to one or two students within the Department, depending on the complexity of the project. The Department's preferences are to make the maximum possible number of acceptable allocations and assign projects to students with good academic performance, or high grade-points averages (GPA's).

Allocation of projects to students as part of a course is very common to universities with different criteria and specific conditions. It first appeared as the classical Hospital/Residents problem (HR) [6][7] in the 1980's, which was to distribute graduating medical students, or residents, to their first hospital posts. The algorithm in [7] found a stable matching of residents to hospitals, which was *resident-optimal*, in which each resident could obtain the hospital of his/her highest preference among all the stable matchings. The problem of allocating students to projects based on preference lists and capacity constraints —the so-called *Student Project Allocation* (SPA) *problem* — then followed as a generalization of HR. Later on and along the structured allocation and automated assessment approach, it was proposed in [10] a systematic approach to final-year (group) projects in an electrical engineering undergraduate course at Nanyang Technological University in Singapore. To allocate projects, a computer program (AssignProj, coded in C) tried to not only minimize the number of unassigned student groups but also take into account student preferences over projects. However, the model did not permit prospective supervisors' preferences. Besides, the algorithm obtained a feasible but not necessarily optimal solution; and the program could run for 24 hours for a large scale data to reach a solution. Optimization techniques, such as integer programming used in [2] and genetic algorithm adopted in [4], have been applied for the SPA problem. And most recently (in 2007 and 2008), the optimal and approximation algorithms for SPA problems have been studied with emphasis on stable matching and complexity issues [1][5]. The SPA problem of the type described in this paper takes into account student preferences over projects and supervisor/departmental preferences over students. Hence, it is by nature being multi-objective, and we solve this case of multi-criteria projects allocation problem instance by our Goal Programming (GP) model for the SPA problem developed in [3].

The remainder of this paper is structured as follows. In section 2, the original data of projects selection and its initial processing are discussed. Then, section 3 briefly reviews the GP formulation for our multi-criteria SPA (MC-SPA) model. The comparison and analysis of the GP solution with two other heuristic solutions, one derived manually by the Department and the other obtained from a greedy algorithm, are discussed in section 4. Section 5 presents a DSS implementation approach for this MC-SPA solution automation on spreadsheet. Concluding remarks are given in section 6.

2 Input data

2.1 Original Data

For the academic year 2008-2009, there are thirteen (13) projects in total for the course: Directed Studies in Mathematics; and twenty-five (25) students have submitted their preference lists for subsets of those projects online before the deadline. The projects are labelled from 1 to 13. And for all thirteen, except for the 5th and 8th which can each take two students, a project can only be allocated to one student. The original data including the 25 students' choices and their average GPA's are given in Fig. 1, in which each row represents one student's ranked choices and his/her GPA. Hence for each column of choices, the number j under the ith choice represents that the student's ith preference is the jth project. From the table, we can notice that students' selections tend to vary. One student selects only a single project but others may specify all. Hence, properly allocating those projects to students, trying to meet both students' and the Department's preferences is a critical issue.

2.2 Initial Data Processing

Since the 5th and 8th projects can each handle two students while all other projects only have capacity for one student, we treat the 5th project as two new different projects

student	Choices													Year GPA
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	
1	1	6	8	9										2.95
2	2	9	1											2.53
3	1	4												2.54
4	10	5	11											3.56
5	2	4	11	3	6	5								3.40
6	9	6	10											3.04
7	7	11	2	5	6	13	3							2.91
8	2	8	7	5	10	11	9	6						2.75
9	8	3	13											2.18
10	9	11	13	4										3.19
11	10	7	11	8	2	5								3.23
12	6	11	10	5	7	9	13	3	1	4	12	2	8	3.48
13	7	12												3.83
14	13													3.77
15	11	3	8	13										2.90
16	8													2.56
17	4	8	13	5	3	9								3.14
18	8	3	4											1.49
19	8	9	6	1	3	7	10	4	11	5	2	12	13	1.52
20	13	3	8											2.24
21	5	3	6	8	4	13								1.97
22	13	3	8	1	4	2								1.87
23	8	5	3											1.71
24	13	4	8	1	3	6	10	5						1.25
25	12	9	13											3.32

Figure 1: The original data of the SPA problem. (Shaded entries: Department's solution)

and rename them project 5 and project 14, with capacity for one student each. If a student selects the original project 5, he/she is treated as opting for both project 5 and project 14. Similarly, the original project 8 is doubled and labelled as the 8th and the 15th. After this pre-processing, there are now fifteen (15) projects in all, labelled from 1 to 15, in which projects 5 and 14, and projects 8 and 15 are respectively identical. Then these "new" 15 projects are assigned to students on a one-to-one basis.

Students' choices in Fig. 1 represent their preference orderings on the projects. In order to quantify and compare ranks overall, the so-called Analytical Hierarchy Process (AHP) Fundamental Scale [11] of $1, 3, 5, \ldots$ is adopted. That is, if a student takes a project as his/her first choice, then the priority cost of this choice is 1; as second choice, the priority cost is 3; as third choice, the priority cost is 5, and so on for a maximum of 15 projects. If a student does not select a project, we define the priority cost to be 0. Hence, the hierarchy of the students' choice list has been converted to numerical values, which can be used and compared in the next decision making procedure. As for the Department's preference, it bases on the students' GPA's, with the usual range from 1 to 4 points. Let G_i represent student i's GPA; and $S_i = 5 - G_i$ is therefore defined to be a positive parameter representing student i's GPA priority cost.

After the above data processing procedure, the input data for the SPA problem is as shown in Fig. 2, in which each row represents one student's choice and his/her GPA. The 25 students' choices are represented by a 25×15 matrix with entry $P_{i,j}$ representing the priority cost of student i over project j using the AHP fundamental scale.

student	Projects															Year GPA
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th	
1	1	0	0	0	0	3	0	5	7	0	0	0	0	0	5	2.95
2	5	1	0	0	0	0	0	0	3	0	0	0	0	0	0	2.53
3	1	0	0	3	0	0	0	0	0	0	0	0	0	0	0	2.54
4	0	0	0	0	3	0	0	0	0	1	5	0	0	3	0	3.56
5	0	1	7	3	11	9	0	0	0	0	5	0	0	11	0	3.40
6	0	0	0	0	0	3	0	0	1	5	0	0	0	0	0	3.04
7	0	5	13	0	7	9	1	0	0	0	3	0	11	7	0	2.91
8	0	1	0	0	7	15	5	3	13	9	11	0	0	7	3	2.75
9	0	0	3	0	0	0	0	1	0	0	0	0	5	0	1	2.18
10	0	0	0	7	0	0	0	0	1	0	3	0	5	0	0	3.19
11	0	9	0	0	11	0	3	7	0	1	5	0	0	11	7	3.23
12	17	23	15	19	7	1	9	25	11	5	3	21	13	7	25	3.48
13	0	0	0	0	0	0	1	0	0	0	0	3	0	0	0	3.83
14	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3.77
15	0	0	3	0	0	0	0	5	0	0	1	0	7	0	5	2.90
16	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	2.56
17	0	0	9	1	7	0	0	3	11	0	0	0	5	7	3	3.14
18	0	0	3	5	0	0	0	1	0	0	0	0	0	0	1	1.49
19	7	21	9	15	19	5	11	1	3	13	17	23	25	19	1	1.52
20	0	0	3	0	0	0	0	5	0	0	0	0	1	0	5	2.24
21	0	0	3	9	1	5	0	7	0	0	0	0	11	1	7	1.97
22	7	11	3	9	0	0	0	5	0	0	0	0	1	0	5	1.87
23	0	0	5	0	3	0	0	1	0	0	0	0	0	3	1	1.71
24	7	0	9	3	15	11	0	5	0	13	0	0	1	15	5	1.25
25	0	0	0	0	0	0	0	0	3	0	0	1	5	0	0	3.32

Figure 2: Input data for SPA problem after initial processing. (Shaded entries: GP solution)

3 Goal Programming Formulation

In this section, we briefly review a GP formulation for the multi-criteria student projects allocation model [3], or MC-SPA model for short. The most distinctive feature of this SPA problem is its being multi-objective. That is, we try to maximize the number of assigned projects, satisfying as much as possible both students' and Department's preferences. For the problem of allocating M projects and N students, the pre-emptive GP model is formulated as follows.

$$\begin{array}{lll}
\text{Goal 1:} & \text{Max} & Z_1 = \sum_{i=1}^{N} \sum_{j=1}^{M} x_{i,j} \\
\text{Goal 2:} & \text{Min} & Z_2 = \sum_{i=1}^{N} \sum_{j=1}^{M} P_{i,j} x_{i,j} \\
\text{Goal 3:} & \text{Min} & Z_3 = \sum_{i=1}^{N} \sum_{j=1}^{M} S_{i,j} x_{i,j}
\end{array}$$

$$\begin{array}{llll}
\text{Constraints:} & \sum_{i=1}^{N} x_{i,j} & \leq 1, & \forall j \\
& \sum_{j=1}^{M} x_{i,j} & \leq 1\ , & \forall i \\
& x_{i,j} & \leq P_{i,j}\ , & \forall i,j \\
& x_{i,j} & = 0 \text{ or } 1\ , & \forall i,j
\end{array}$$

The model has three hierarchical goals to achieve. The goal to allocate most projects to students is of primary importance and is maximized first. Next, using the target value obtained from the first goal, it tries to optimize the total satisfaction level of students' preferences, and then subsequently minimize the sum of GPA priorities of students assigned with projects. It introduces binary decision variables $x_{i,j}$, $i = 1,2,\ldots,N$; $j = 1,2,\ldots,M$, which is equal to 1 if student i is assigned project j and 0 otherwise. P in this formulation is an $N \times M$ coefficient matrix representing the priority costs between all pairs of students and projects, and S is modeled more generally as an $N \times M$ coefficient matrix representing the priority costs between the students and supervisors (beyond just GPA priorities of students).

4 Numerical Results

In this section, we apply GP to the MC-SPA model described in the previous section to the course selection data. The model is implemented in LINGO 10.0 for Windows [8], running on an Intel Celeron Duo processor based PC. It takes less than one second to solve each goal. The objective function value for goal 1 is 15, which means that all the 15 projects can be allocated to students. Using this target value for optimizing Goal 2 to minimize the sum of priority costs of allocated projects gives a value of 19 in this case, under the constraint that all the 15 projects should be distributed to students. This averages to 1.27 among the 15 students. Then, using the objective function value of 19 from the second goal and 15 from the first goal, optimizing Goal 3 produces the final SPA solution. The detailed assignments of projects to students are indicated by the shaded cell entries in Fig. 2. As also summarized in row 1 of Fig. 3 all the allocated projects are among the students' first two choices. This compares very well with the other two heuristic solutions with details to be given later. In fact, the result shows that nearly ninety percent (87%) of students are allocated their first choices, and only two students are allocated the projects of their second choices, while the average GPA of students of those allocated projects is 3.048, with their individual GPA's ranging from 1.97 to 3.83.

No. (percentage) of Projects Allocated	Student Preference						Avg. GPA	Min GPA
	1st	2nd	3rd	4th	5th	6th		
GP Model	13 (87%)	2 (13%)	0	0	0	0	3.048	1.97
Department's Manual Solution	7 (47%)	7 (47%)	0	0	0	1 (7%)	3.084	2.18
Greedy Algorithm	11 (73%)	2 (13%)	1 (7%)	1 (7%)	0	0	3.131	1.97

Figure 3: Numerical results of three methods.

The second row of Fig. 3 represents a manual solution given by the Department. This distribution plan also successfully allocates all the 15 projects to students with nearly half (7) of the students obtaining their first choices, over ninety percent (93%) of them (7+7=14) receiving their fist and second choices and only one of the projects is assigned to a student as his/her 6th choice. The detailed assignments of projects to students are indicated by the shaded cell entries in Fig. 1. And for the third row, it is another distribution plan derived from a greedy algorithm, which treats the Department's GPA preference much more important than the students' preferences. The greedy algorithm is described as follows. First, it sorts all the 25 students according to their GPA's from the highest to the lowest. The first student is assigned his/her first choice. Consider next assigning the second student his/her first choice if the project is still available, otherwise his/her second choice. Then consider the third student, and assign him/her the fisrt unassigned project that he/she most prefers, and so on. During this processing, if a student being considered has all his/her selections already assigned, he/she is skipped over and we move on to consider the next one below him/her. Although it is a greedy approach, for this problem instance, it solves the SPA problem not too badly. Eleven out of fifteen students (73%) are assigned projects of their first choices, and the rest of 4 projects are allocated to students at least as their 4th choice. The average GPA has an expectedly high value of 3.131, but also accompanied by a low minimum GPA of 1.97.

From Fig. 3, we notice that the numerical results obtained from GP dominate the other two in satisfying students' preferences, since all projects are assigned to students their first or second choices, with nearly 90% of them first choices. Compared with the manual solution given by the Department, although the average GPA of the GP solution is slightly lower, the distribution plan itself is far better in meeting the overall students' preferences. Hence, the GP solution is deemed to be rather better than the Department's manual solution. As for the dominance between the results of the GP model and the greedy algorithm, the former totally dominates in meeting students' preferences, but its average GPA of students being assigned projects is lower than that derived from the greedy algorithm, since the greedy algorithm allocates projects sequentially from the student with the highest GPA to the one with the lowest.

Besides, if we consider re-allocating projects among only the 15 students already chosen by the Department, i.e., those 15 being assigned projects decided by the Department's manual solution, by the GP model to this restricted subset of students (being the first 15, cf. the shaded entries in Fig. 1), the allocation plan can indeed be improved as shown in Fig. 4 . This re-allocation has the number of students obtaining their first choices increased by one. No student gets his/her 6th choice, with a more equitable distribution.

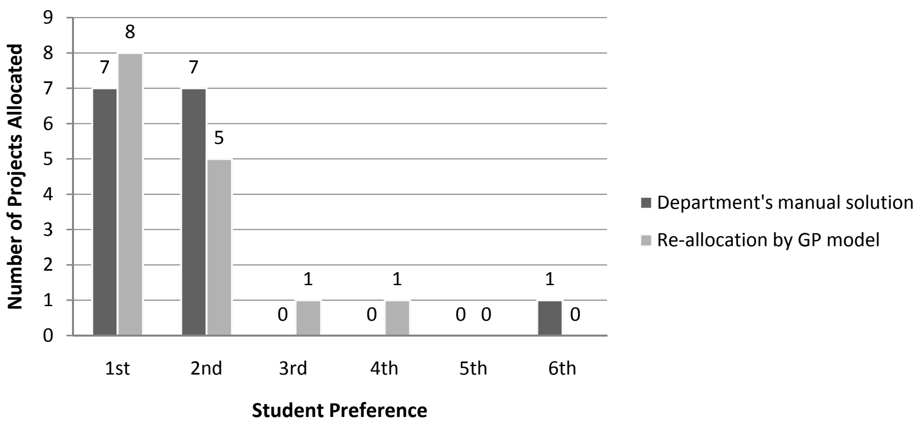

Figure 4: Distribution plans given by the Department and re-allocated by the GP model.

Furthermore, to analyze the trade-off effects between the last two goals of the GP model, i.e., between minimizing the students' preferences over projects and minimizing the Department's total priority costs over students, the non-dominated (or, efficient) frontier of the average GPA of students being assigned projects versus the total priority costs of allocated projects is plotted as shown in Fig. 5. As the target total priority cost is being relaxed, the average GPA of students allocated project is improved and reaches its

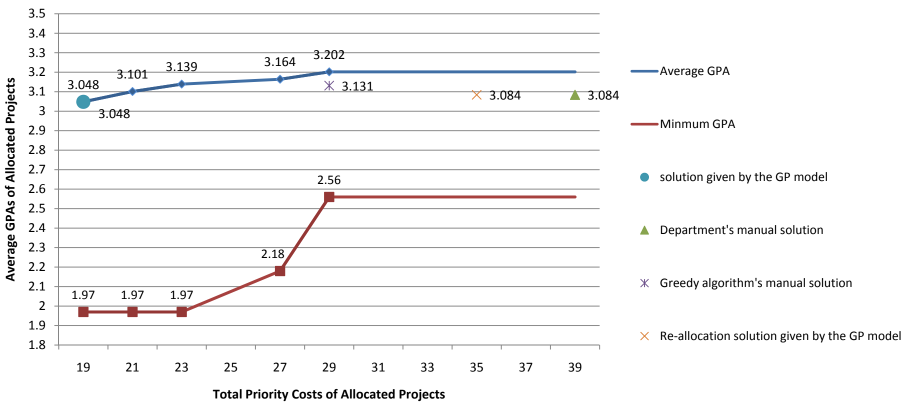

Figure 5: The efficient frontier between the total priority cost and the average GPA of allocated projects derived from the GP model.

maximum possible of 3.202. Nodes on the frontier correspond to the five non-dominated allocation plans, e.g. the MC-SPA GP solution marked by the round dot in Fig. 5. The decision maker can adopt a final distribution plan according to his/her own assessment. The manual solutions given by the Department and the greedy algorithm, as well as the re-allocation plan by GP, are indeed verified in Fig. 5 to be mathematically dominated solutions.

5 DSS Implementation of MC-SPA GP model

To facilitate successful applications, the GP code of the MC-SPA model can be improved into a Decision Support System (DSS), integrating the GP model as an optimizer engine with its front-end interface and back-end reporter integrated in the same Excel spreadsheet, into which the input data records can be placed and outputs of distribution plans and the efficient frontier plots displayed.

In our case, the DSS is being developed as an Excel-based system to generate allocation plans automatically by giving the input information of students' selection information and their GPA's. The input data will be processed and passed to a LINGO solver to find solutions based on the pre-emptive GP model. The DSS will capture the solution into the Excel spreadsheet as well as the efficient frontier as consisting of individual non-dominated solutions. The user-friendly spreadsheet based interface provides the ultimate flexibility for users to make changes to the original problem information (its data and parameters) and to perform what-if analysis.

6 Conclusion

In this paper, we apply a pre-emptive GP formulation to solve a multi-criteria SPA problem imposed by the Department of Mathematics of the University of Hong Kong. The MC-SPA GP model is implemented in LINGO 10.0 and computationally solved efficiently on a desktop PC, which technically stems from the benefit of its underlying assignment problem network structure with its totally unimodular coefficient matrix. Compared with the manual solutions given by the Department and the greedy algorithm, the GP model has produced an effective and efficient distribution plan of all projects being allocated to students with best priority costs and highest possible GPA's. With the implementation of its DSS, we highly expect that the approach can be used for future courses or project selections in a user-friendly and effective manner.

References

[1] D. J. Abraham, R. W. Irving and D. F. Manlove: "Two Algorithms for the Student-Project Allocation Problem", Journal of Discrete Algorithms, Vol.5, No.1, pp.73-90, 2007.

[2] A. A. Anwar and A. S. Bahaj: "Student Project Allocation Using Integer Programming", IEEE Transactions on Education, Vol.46, No.3, pp.359-367, 2003.

[3] S.C.K. Chu, L. Pan: "Student Project Allocation Using Goal Programming", in Proceedings of the INFORMS 2009 International Conference on Service Science, Hong Kong, Aug 2009, 6pp.

[4] P. R. Harper, V. de Senna, I. T. Vieira and A. K. Shahani: "A Genetic Algorithm for the Project Assignment Problem", Computers and Operations Research, Vol.32, pp.1255-1265, 2005.

[5] D. F. Manlove and G. O'Mally: "Student-Project Allocation with Preferences over Projects", Journal of Discrete Algorithms, Vol.6, No.4, pp.553-560, 2008.

[6] A. E. Roth: "The evolution of the labor market for medical interns and residents: a case study in game theory", Journal of Political Economy, Vol.92, pp.991-1016, 1984.

[7] A. E. Roth: "On the Allocation of Residents to Rural Hospitals: A General Property of Two-sided Matching Market", Econometrica, Vol.54, No.2, pp.425-427, 1986.

[8] L. Schrage, Optimization Modeling with LINGO, 5th ed., Lindo Systems Inc., Chicago, 2005.

[9] R. E. Steuer, Multiple Criteria Optimization: Theory, Computation, and Application, Krieger, 1989.

[10] C. Y. Teo and D. J. Ho: "A Systematic Approach to the Implementation of Final Year Project in an Electrical Engineering Undergraduate Course", IEEE Transactions on Education, Vol.41, No.01, pp.25-30, 1998.

[11] Wikipedia, the free encyclopedia: accessed on 26 March 2009 "Analytic Hierarchy Process (AHP)", http://en.wikipedia.org/wiki/Analytic_Hierarchy_Process , 2009.