Modeling and Solving Academic Load Distribution Problem

Lidia A. Zaozerskaya[1], Valentina A. Plankova[1], and Marina V. Devyaterikova[2]

[1] Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Omsk, Russia
zaozer@ofim.oscsbras.ru
plankova@ofim.oscsbras.ru
[2] Omsk Tank Automotive Engineering Institute, Omsk, Russia
devy_m@mail.ru

Abstract. We consider an academic load distribution problem (ALD) in which training courses consist of separate units. Each of the units can be assigned just to one teacher. The minimum and maximum possible amounts of an academic load are set for each teacher. Integer linear programming (ILP) models for various formulations of this problem are proposed. One model takes into account interpersonal relationships when appointing teachers to one discipline. It is shown that finding a feasible solution of the ALD problem is NP-complete. In this paper, the research of the bicriteria ILP model of the ALD problem is continued. It is established that the cardinality of a complete set of alternatives for this problem is polynomial. An auxiliary problem is proposed to find a feasible solution to the problem under study. A computational experiment with ILP models on random instances and on a real data instance is carried out.

Keywords: Teacher assignment problem · Integer linear programming · Bicriteria problem · NP-completeness · Pareto-optimal solution · Complete set of alternatives

1 Introduction

Many problems taking place in the sphere of education are successfully solved by discrete optimization models and methods. They are the problems of scheduling [1], tests generation of academic performance rating [11], etc. Academic load distribution (ALD) for teachers is an important stage of the qualitative training process organization. It is difficult to formalize this problem. It becomes actual in a big and various department staff, quite a number of the courses, including the presence of the load in several groups or at several faculties, in the case of

In: S. Belim et al. (eds.): OPTA-SCL 2018, Omsk, Russia, published at http://ceur-ws.org

the changeable loading, when there are special requirements for the academic load distribution at a higher school. In particular, the works [3, 8] concentrate on the construction of mathematical models for the ALD problem.

The ALD problem, in which the average number of the training courses assigned to each teacher is minimized, is considered in [3]. The mathematical model offered in the given work represents a special case of the transportation problem with the fixed surcharges. The work shows that the mentioned problem is NP-hard, represents the equivalent combinatory formulation of this problem and offers the branch-and-bound algorithm for its solving. However, in this case, it is not considered that each training course splits into the indivisible units corresponding to various kinds of the academic load, for example, lectures, seminars, laboratory works, tests, examinations, etc. As a result of the application of the model from [3], there may be obtained a solution wherein a unit of a training course is distributed among several teachers. It contradicts the existing practice at Russian higher schools.

The bicriteria formulation of the ALD problem in which the training courses consist of particular units each of which can be distributed just to one teacher is considered in [9]. The minimum and maximum possible amounts of academic load are preassigned for every teacher. In [9], ILP models for some formulations of this problem were proposed. The L-class enumeration algorithm [5, 10] which enables to solve problems of small dimension was suggested for one model.

In this article, we continue the study of the bicriteria model. According to it, one optimization criterion is to minimize the maximum number of courses assigned to teachers, and the other criterion is to maximize the total preferences of the "teacher–course" relations. In addition, a variant of the problem, considering interpersonal relations in the assignment of teachers to the same training course, is regarded. We show that the cardinality of the complete set of alternatives for this problem is polynomial.

We should note that in contrast to the problem from [3] finding a feasible solution of the investigated ALD problem is NP-complete. We construct an auxiliary ILP problem which can be used to organize the process of finding a solution to the investigated problem. Computational results obtained for random generated instances and for the real data of one department of Omsk State Technical University is described.

2 Formulations of Academic Load Distribution Problem and Mathematical Models

Let $I = \{1, ..., m\}$ be a set of teachers. The values a_i, c_i denote the maximum and minimum possible amounts of load specified for each teacher i, $i \in I$. Normally c_i differs from a_i by no more than 5%. Let $J = \{1, ..., n\}$ be a set of training courses and let t_j be a number of individual units of the course j. Denote a number of hours for the k-th unit of the course j by b_j^k, where $k \in K_j = \{1, ..., t_j\}$.

We denote by l_{ij}^k preference coefficients of the unit k of the course j by the teacher i for all i, j, k. Also, these values can be interpreted as the coefficients of efficiency under this assignment.

Let s_j be a maximum number of units of one training course that are assigned to one teacher ($s_j \leq t_j$). We should note that if the values s_j is not limited, it may lead to distributions in which the teachers have a narrow specialization.

It is necessary to distribute the training courses among the teachers "uniformly" in the number of courses taking into account their preferences. The condition of "uniformity" can be accounted for in different ways. Now we describe a mathematical model of the problem at issue when a_i take the same value for all i.

We introduce Boolean variables x_{ij} and z_{ij}^k, where $i \in I, \quad j \in J, \quad k \in K_j$. Here $x_{ij} = 1$ if the teacher i gives the course j and $x_{ij} = 0$, otherwise; $z_{ij}^k = 1$, if the teacher i gives the k-th unit of the course j and $z_{ij}^k = 0$, otherwise. Let y be a non-negative integer variable.

Denote the vector of variables z_{ij}^k by $\bar{z}$. The ALD problem can be formulated as a bicriteria ILP problem in this way:

$$\text{minimize } y \tag{1}$$

$$\text{maximize } L(\bar{z}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{t_j} l_{ij}^k z_{ij}^k \tag{2}$$

subject to

$$c_i \leq \sum_{j=1}^{n} \sum_{k=1}^{t_j} b_j^k z_{ij}^k \leq a_i, \quad i \in I, \tag{3}$$

$$\sum_{i=1}^{m} z_{ij}^k = 1, \quad j \in J, \ k \in K_j, \tag{4}$$

$$x_{ij} \leq \sum_{k=1}^{t_j} z_{ij}^k \leq s_j x_{ij}, \quad i \in I, \ j \in J, \tag{5}$$

$$\sum_{j=1}^{n} x_{ij} \leq y, \quad i \in I, \tag{6}$$

$$y \geq 0, \ y \in Z, \ x_{ij}, z_{ij}^k \in \{0, 1\}, \ i \in I, \ j \in J, \ k \in K_j. \tag{7}$$

Further we will call the problem (1) – (7) as the base model and denote it by $M1$. The restrictions (3) provide for the assignment to the i teacher the units of the courses, the total amount of which satisfies the lower and upper bounds of the acceptable load. The equalities (4) show that each unit of any course should be assigned to just one teacher. The inequalities (5) describe the relationship of the variables x_{ij} and z_{ij}^k. The variable $x_{ij} = 1$ if and only if there exists the index k such that $z_{ij}^k = 1$, i.e., the course j is assigned to the teacher i if and only if when at least one unit of this course is assigned to him.

In the problem $M1$ the optimization criterion (1) with the conditions (6) means minimization of the maximum number of courses assigned to the teachers and the criterion (2) maximizes the total preference $L(\bar{z})$ in the load distribution. We should note that the function (1) optimizes the extracurricular load of the teachers, i.e., the number of preparations for classes.

It is easy to see that finding a feasible solution to the model $M1$ is NP-complete. This follows from the fact that a well-known NP-complete bin packing problem [2] can be reduced to the mentioned problem. Here m is the number of containers with the volumes a_i and $\sum_{j\in J} t_j$ is the number of items with the volumes b_{ij}^k which must be packed.

Consider the case when a_i take different values. Let us denote $a_{max} = \max_{i\in I} a_i$ and let y be the number of courses assigned to the teacher with the amount of the load a_{max}. Then we offer to model the condition of "uniformity" on the number of courses in the following way. The number of courses assigned to the teacher i should not exceed the value which is proportional to y with the factor $p_i = a_i/a_{max}$. Then the conditions (6) in the model $M1$ should be replaced by the following restrictions:

$$\sum_{j=1}^{n} x_{ij} \le p_i y + q, \quad i \in I, \tag{8}$$

where $q \in [0,\ 1]$ is a constant that controls the rounding of the values on the right side, for example, $q = 0.5$.

Now we will describe another variant of ALD problem. Let the maximum number of the courses r_i assigned to the teacher i are given, $i \in I$. The requirement of the diversity "uniformity" to the number of assigned courses can be also set explicitly

$$\sum_{j=1}^{n} x_{ij} \le r_i, \quad i \in I. \tag{9}$$

Consider the problem $M2$ which is obtained from the base model by adding constraints taking into account interpersonal relationships among the teachers. Let $T = \{(i_1, i_2) \mid i_1, i_2 \in I\}$ be the set that contains index pairs of teachers who are undesirable to be assigned to one course for any reason because these teachers have inconsistent relations [6]. Let us add appropriate restrictions

$$x_{i_1,j} + x_{i_2,j} \le 1,\ (i_1, i_2) \in T,\ j \in J. \tag{10}$$

They mean the possibility of assigning to any j course no more than one teacher from the pair (i_1, i_2).

3 Solving ALD Problem and Computational Experiment

Consider the bicriteria discrete optimization problem

$$\text{minimize } F(x) = (f_1(x), f_2(x))$$

subject to

$$x \in X,$$

where X is some finite set.

There are several approaches to solve this problem. One approach is finding all Pareto-optimal solutions or a part of this solutions. We denote the set of Pareto-optimal solutions by $\tilde{X}$, and a complete set of alternatives (CSA) by X^0 (see, for example, [4, 7]). Let X be the set of feasible solutions of the problem. CSA is any set X^0 ($X^0 \subseteq \tilde{X}$) which has the minimal cardinality and $F(X^0) = F(\tilde{X})$, where $F(X') = \{F(x) \mid x \in X'\}, \quad X' \subseteq X$. It is obvious that

$$X^0 \subseteq \tilde{X} \subseteq X.$$

We note that the cardinality of a complete set of alternatives for the constructed problems $M1$ and $M2$ is polynomial because the value of the second criterion does not exceed n.

To search for CSA, you can solve a series of the single-criterion problem (2)–(5), (7) with the corresponding condition (6), (8) or (9) and the restriction $y \leq \bar{Y}$, where $\bar{Y} = 1, 2, ..., n$. It is clear that the small values of $\bar{Y}$ are more interesting. If there is a change of the optimal value of the function L in the transition from the current value of $\bar{Y}$ to $\bar{Y} - 1$ or the optimal value L^* does not exist then a Pareto-optimal solution was obtained. This solution is also an element from CSA.

We introduce a non-negative variable v which can be interpreted as the maximum excess of values of a_i, $i \in I$, for some distribution of the units of the courses. Let us consider an auxiliary problem

$$\text{minimize } v \tag{11}$$

subject to

$$\sum_{j=1}^{n} \sum_{k=1}^{t_j} b_j^k z_{ij}^k \geq c_i, \quad i \in I, \tag{12}$$

$$\sum_{j=1}^{n} \sum_{k=1}^{t_j} b_j^k z_{ij}^k - a_i \leq v, \quad i \in I, \tag{13}$$

and conditions (4)–(6).

If a feasible solution of the ALD problem exists, then the optimal value of the objective function of the auxiliary problem is equal to zero. We introduce additional restrictions $y \leq \bar{Y}$ and $L \geq \bar{L}$ and solve the auxiliary problem, where $\bar{Y}$ and $\bar{L}$ are some given values. Changing the values of $\bar{Y}$ and $\bar{L}$, one can obtain a series of feasible solutions of the ALD problem or unfeasible solutions in which there is a small excess of some or all a_i. Such solutions can also be used in practice.

We conducted a computational experiment to investigate the possibility of applying the proposed models.

We present the results of using model $M1$ with restrictions (8) for solving a practical ALD problem. Data on the load for the 2010/2011 academic year provided by one Department of the Omsk State Technical University was used.

The total amount of all types of the Department's load is 8571 hours. The number of different types of load including training courses is 103. The number of teachers equals 13. The total upper load is equal to 8590 hours.

In the first stage, the personal load is assigned. This is the leadership of the department, the scientific supervision of graduate students, undergraduates, etc. As a result in the second stage, it remains to distribute 5591 hours of 88 training courses among 11 teachers. The total number of the courses units equals 229. The number of units for every course is in the range from 1 to 6. The majority of courses (83%) consists of no more than three units. The upper bounds of the load for the teachers are given by the vector $a = (557, 560, 601, 741, 591, 560, 625, 658, 360, 360, 68)$. The preference coefficients l^k_{ij} belong to the set $\{1, ..., 10\}$.

For finding Pareto-optimal solutions we construct single-criterion ILP problem (2) – (5), (7), (6) with restriction $y \leq \bar{Y}$. This task has 6778 variables and 2200 constraints. This task was solved using the GAMS solver on PC with Intel Core i3 Processor (3.30 GHz). Computational experiments for various values $\bar{Y}$ were carried out. Pareto-optimal solutions are obtained for $\bar{Y} = 12, ..., 15$. The values L^* are equal to 1984, 2029, 2069 and 2076 respectively.

We compared our results with the actual load distribution in the department that was obtained without the use of mathematical models. In the actual load distribution, the total excess of the upper bounds of the teachers' load (a_i) was 225 hours, and the total underperformance of lower bounds (c_i) was 37 hours. Each teacher was assigned from 7 to 18 training courses.

Note that all our solutions satisfy the upper and lower bounds of the load of teachers. More "uniform" distributions on number of disciplines are obtained. For $\bar{Y} = 12$ the number of courses for each teacher is proportional to the values of a_i and varies from 6 to 12. In the actual distribution the number of courses varies from 7 to 18. In the mentioned solution, the total preference is decreased by 13% compared to the actual distribution. Note that it is not a valid solution for the $M1$ problem. Extracurricular preparation of the teachers, i.e., the number of preparations for classes is decreased by 14%.

Briefly, we present the results of a computational experiments for model $M1$ under restrictions (8) with random initial data. We used the GAMS solver and the CPLEX solver for corresponding single-criterion tasks. We have considered several series of tasks. For example, the S_1 series consists of tasks of small dimension with $m = 8$, $n = 16$, all t_j=4, $a_i \in [400,\ 800]$ and $l^k_{ij} \in [1,\ 10]$. The corresponding ILP models have 642 variables and 339 constraints. We received from 1 to 6 Pareto-optimal solutions for each $\bar{Y}$ from 2 to 11. The computing time was from 1 second to 30 minutes. With decreasing $\bar{Y}$, the solution time increased. Note that the restrictions (8) and the value of $\bar{Y}$ set the upper limit of the number of courses assigned to teachers. For small values $\bar{Y}$ we obtained optimal solutions, in which the mentioned bounds for all teachers are reached, i.e. "uniform" distributions are obtained. The most difficult are the tasks in which the total distributed load is close to the amount of the load of the lower bounds of the teachers.

The S_2 series consists of tasks of large dimension with $m = 22$, $n = 45$, all $t_j \in [2,\ 9]$, $a_i \in [211,\ 1268]$ and $l_{ij}^k \in [1,\ 10]$. These tasks are similar to practical tasks. The corresponding ILP model $M1$ has 5962 variables and 2139 constraints. The computing time depends on the given values of $\bar{Y}$ and s_j and varies from a few seconds to several hours. There are tasks for which no feasible solution has been found. In this case, we used the auxiliary problem AP. Changing the values of $\bar{Y}$ and $\bar{L}$, we managed to obtain a series of feasible solutions in acceptable time. If a task does not have feasible solutions, then it is possible to obtain a load distribution with an insignificant excess of the upper bounds a_i.

A computational experiments with model M2 which taking into account the interpersonal relations among the teachers showed a significant increase of computing time.

4 Conclusion

We propose some formulations of the ALD problem including the bicriteria model. It is shown that the cardinality of a complete set of alternatives is polynomial and finding a feasible solution of this problem is NP-hard. Computational experiments was carried out using the CPLEX solver and the GAMS solver for problems with random input data.

One of the constructed mathematical models was successfully applied to solve a practical problem. Distribution of the academic load of the department, depending on its size, usually takes from one to several weeks. The application of the constructed mathematical models significantly shortens this time.

The proposed models can be modified by additional requirements for load distribution, for example, when assigning no more than one new course to each teacher, uniform load distribution in semesters, etc. To solve problems of large dimension, it is advisable to develop heuristic algorithms.

The mathematical models of the mentioned problem can be used in the decision-making support systems for University management.

Acknowledgement. For the first author the research was supported by RFBR grant 16-01-00740. For the second author the work was supported by the program of fundamental scientific researches of the SB RAS – I.1.3., project – 0314-2016-0009.

References

1. Avella, P., Vasil'ev, I.: A computational study of a cutting plane algorithm for university course time-tabling. Journal of Scheduling 8(6), 497–514 (2005), (in Russian)
2. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W.H. Freeman & Co. New York, NY, USA (1979)

3. Hultberg, T.H., Cardoso, D.M.: The teacher assignment problem: a special case of the fixed charge transportation problem. European Journal of Operational Research 101, 463–473 (1997)
4. Emelechev, V.A., Perepelitsa, V.A.: The complexity of discrete multicriteria problems. Discrete Math. Appl. 4(2), 89–117 (1994)
5. Kolokolov, A.A.: Regular partitions and cuts in integer programming. In: Korshunov, A.D. (eds.): Discrete Analysis and Operations Research. Mathematics and Its Applications. vol. 355, pp. 59–79. Springer, Dordrecht (1996)
6. Kolokolov, A.A., Ziegler, I.A.: Resheniye nekotorykh zadach formirovaniya tselevykh grupp s uchetom logicheskikh ogranicheniy (Solving of some target groups formation problems with logical restrictions). Omskiy nauchnyy vestnik, seriya "Pribory, mashiny i tekhnologii" 154 (4), 103–107 (2017), (in Russian)
7. Kolokolov, A.A., Zaozerskaya, L.A.: Solving a bicriteria problem of optimal service centers locations. Journal of Mathematical Modelling and Algorithms 12 (2), 105–116 (2013)
8. Sultanova, S.N., Tarkhov, S.V.: Modeli i algoritmy podderzhki prinyatiya resheniy pri raspredelenii uchebnoy nagruzki prepodavateley (Models and algorithms of decision support in the distribution of academic load of teachers). Vestnik UGATU 7(3), 107–114 (2006), (in Russian)
9. Zaozerskaya, L.A.: On solving academic load distribution problem. In: Abstracts of the 17th Baikal international school-seminar "Methods of Optimization and Their Applications" (July 31 - August 6, 2017, Maksimikha, Buryatia). P. 129. ESI SB RAS, Irkutsk (2017)
10. Zaozerskaya, L.A.: Ob odnom algoritme perebora L-klassov dlya resheniya zadachi o pokrytii mnozhestva (On L-class enumeration algorithm for set covering problem). In: Trudy 11-y Baykal'skoy mezhdunar. shkoly-seminara "Metody optimizatsii i ikh prilozheniya". vol. 1, pp. 139–142. Irkutsk (1998), (in Russian)
11. Zaozerskaya, L.A., Plankova, V.A.: Primeneniye modeley diskretnoy optimizatsii dlya razrabotki avtomatizirovannoy sistemy kontrolya znaniy (The using discrete optimization models at development of an automatical knowledge checking system). Vestnik NGU, Seriya "Informatcionnyye tekhnologii" 6(1), 47–52 (2008), (in Russian)