Unfinished or Inactive

Automating and Optimizing Weekly Course Offerings at an Educational Summer Camp

Zachary Siegel

February 27, 2019

Abstract

Parameters are defined to describe a pool of teachers and students, each with course preferences, abilities, availabilities, and capacities. Variables are defined to describe whether teachers and students are scheduled to teach or participate in a course offering in a given week. A linear objective function is defined to quantify the success of a given assignment, and both linear and integer constraints are defined to ensure the solution satisfies the system's logic. The complexity of the computation is analyzed in terms of the number of integer decision variables and the ideal performance of a MIP solver. The tractability is discussed for a scenario with roughly 100 weekly students and roughly 10 teachers.

1 Background

SCHEDULING PROBLEMS, etc.

2 Method: The Linear Model

The model solved to assign each student classes to take and each teacher classes to teach involves variables describing the system parameters, as well as variables describing the decisions. There is an objective function evaluating the quality of a solution, and there are linear and integer constraints to ensure the model corresponds to the scenario.

2.1 Possible Extensions

Parameters describing student/teacher pairwise distances from locations and willingness to work at or attend a given location would allow the natural extension of teachers and students to sites. A multi-stage process for assigning teachers and students to locations (for example, in job applications and signup forms) is easier to understand for everyone involved, but may miss out on some value.

Even further, the hiring process could be reduced to a “vetting” and/or “rating” procedure, and staff selection could be incorporated into a model to fill all necessary roles with a minimum cost in terms of commutes for all teachers. Depending on the scale, different rewards/penalties could be associated with ratings on separate dimensions (distance, experience, etc.), or one cumulative rating could be computed separately to reduce complexity of the optimization problem.

Ideally, rather than replacing the “human” elements of hiring, automation could allow an employer to focus exclusively on qualitative and personal impressions, assigning weighted values to each aspect of an employee’s fit in a given role. For example, different employers would assign different weights to experience, education level, commute distance, age, personal factors, etc., and would also give employees different scores on subjective personal factors. As a result, different employers using the exact same hiring procedure would have very different results.

Employees’ relative preferences for roles, shorter commutes, preferred locations, etc., could also be factored into the model with no additional computational cost, although general reported “preferences” could already account for much of this.

Flexibility is important, but a motivating goal is to allow for highly-specialized and individualized services across a network of businesses. There is no requirement that team members be generalists as long as roles are sufficiently filled.

2.2 System Scale

The instructors, students, courses, and weeks are enumerated as follows:

- The $n_{teachers}$ instructors are enumerated $1, \ldots, i, \ldots, n_{teachers}$
- The $n_{students}$ students are enumerated $1, \ldots, j, \ldots, n_{students}$
- The $n_{courses}$ courses are enumerated $1, \ldots, k, \ldots, n_{courses}$

- The n_{weeks} weeks are enumerated $1, \dots, l, \dots, n_{weeks}$

2.3 System Parameters

Let the following parameters describe the modeled scenario[1]:

- $p_{i,k}$ = the "preference" of teacher i for teaching course k on a scale of 0 (will not teach) to M_p (highest preference)
  - $p^0_{i,k} = \begin{cases} 1 & \text{teacher } i \text{ is able to teach course } k \\ 0 & \text{otherwise} \end{cases}$
- $q_{j,k}$ = the "preference" of student j for taking course k on a scale of 0 (will not take) to M_q (highest preference)
  - $q^0_{j,k} = \begin{cases} 1 & \text{student } j \text{ wants to take course } k \\ 0 & \text{otherwise} \end{cases}$
- $a^{teacher}_{i,l} = \begin{cases} 1 & \text{teacher } i \text{ available in week } l \\ 0 & \text{otherwise} \end{cases}$
- $a^{student}_{j,l} = \begin{cases} 1 & \text{student } j \text{ available in week } l \\ 0 & \text{otherwise} \end{cases}$
- $c^{student}_j$ = the maximum number of courses student j would like to take during the summer (it may not be possible for each student to take this many)
- $c^{location}_l$ = the maximum number of students that can be at the camp during week l. In this work, $c^{location}_l = c^{location} \ \forall \ l$, as the camps' capacities are constant from week to week
- $c^{classcap}_k$ = the maximum number of students that can be taught in class k at once
- $c^{sections}_{k,l}$ = the maximum number of sections of course k that can be held in week l
- $c^{curriculum} = 3$ is the maximum number of different courses a teacher can teach in one summer

[1]The variables p and q denote preferences; a represents availability; c represents capacities; and d represents costs.

- $g =$ The minimum number of students required to enroll in order for a course to be held
- d_k = the cost of a teacher being assigned a course. Teachers teach fewer than $c^{curriculum}$ courses, but there may still be a cost associated with teaching multiple courses. Penalizing based on d_k (which may be the same $d_k = d \; \forall$ courses k) ensures that if teachers overall can learn fewer curricula, they will. If teachers are expected to teach roughly the maximum allowable number of courses, $d_k = 0 \; \forall \; k$ is acceptable.

2.4 Decision Variables

Let the following variables describe the decisions involved in constructing a schedule[2]:

- $x_{i,k,l} = \begin{cases} 1 & \text{teacher } i \text{ leads course } k \text{ in week } l \\ 0 & \text{otherwise} \end{cases}$
- $y_{j,k,l} = \begin{cases} 1 & \text{student } j \text{ takes course } k \text{ in week } l \\ 0 & \text{otherwise} \end{cases}$
- $x^0_{i,k} = \begin{cases} 1 & \text{teacher } i \text{ leads course } k \text{ in some week} \\ 0 & \text{otherwise} \end{cases}$
  - Note: $x^0_{i,k} = \max_l x_{i,k,l}$

3 The Mixed Integer Program

Maximize:

$$\sum_{i=1}^{n_{teachers}} \sum_{k=1}^{n_{courses}} \sum_{l=1}^{n_{weeks}} p_{i,k} x_{i,k,l} + \sum_{j=1}^{n_{students}} \sum_{k=1}^{n_{courses}} \sum_{l=1}^{n_{weeks}} q_{j,k} y_{i,k,l} - \sum_{i=1}^{n_{teachers}} \sum_{k=1}^{n_{courses}} d_k x^0_{i,k} \tag{1}$$

Such that:

- $x_{i,k,l}, y_{j,k,l}, x^0_{i,k} \in \{0,1\} \; \forall \; i,j,k,l$

[2]The variable x

- $x_{i,k,l} \leq p^0_{i,k} \ \forall \ i,k,l$
- $y_{j,k,l} \leq q^0_{j,k} \ \forall \ j,k,l$
- $x_{i,k,l} \leq a^{teacher}_{i,l} \ \forall \ i,k,l$
- $y_{j,k,l} \leq a^{student}_{j,l} \ \forall \ j,k,l$
- $\sum_l \sum_k y_{j,k,l} \leq c^{student}_j \ \forall \ j$
- $\sum_k y_{j,k,l} \leq \underbrace{1}_{\text{or } a^{student}_{j,k}} \ \forall \ j,l$
- $\sum_k x_{i,k,l} \leq \underbrace{1}_{\text{or } a^{teacher}_{i,k}}$
- $\sum_j \sum_k y_{j,k,l} \leq c^{location} \ \forall \ l$
- $\sum_j y_{j,k,l} \leq \sum_i c^{classcap}_k x_{i,k,l} \ \forall \ k,l$
- $\sum_i x_{i,k,l} \leq c^{sections}_{k,l} \ \forall \ k,l$
- $\sum_l y_{j,k,l} \leq \underbrace{1}_{\text{or } q^0_{j,k}} \ \forall \ j,k$
- $x_{i,k,l} \leq x^0_{i,k} \ \forall \ i,k,l$
- $\sum_k x^0_{i,k} \leq c^{curriculum} \ \forall \ i$
- $\sum_i g x_{i,k,l} \leq \sum_j y_{j,k,l} \ \forall \ k,l$

3.1 Objective Function Explained

$$\underbrace{\sum_{i=1}^{n_{teachers}} \sum_{k=1}^{n_{courses}} \sum_{l=1}^{n_{weeks}} p_{i,k} x_{i,k,l}}_{\text{sum of teacher preferences for courses}} + \underbrace{\sum_{j=1}^{n_{students}} \sum_{k=1}^{n_{courses}} \sum_{l=1}^{n_{weeks}} q_{j,k} y_{i,k,l}}_{\text{sum of student preferences for courses}} - \underbrace{\sum_{i=1}^{n_{teachers}} \sum_{k=1}^{n_{courses}} d_k x^0_{i,k}}_{\text{sum of penalties for course}}$$

3.2 Constraints Explained

- $x_{i,k,l} \leq p^0_{i,k} \; \forall \; i, k, l$
  Teachers do not teach courses unless they are able to teach that course.
- $y_{j,k,l} \leq q^0_{j,k} \; \forall \; j, k, l$
  Students do not enroll in a course they do not wish to take.
- $x_{i,k,l} \leq a^{teacher}_{i,l} \; \forall \; i, k, l$
  Teachers do not teach in weeks in which they are not available.
- $y_{j,k,l} \leq a^{student}_{j,l} \; \forall \; j, k, l$
  Students do not enroll in courses in weeks in which they are unavaialable.
- $\sum_l \sum_k y_{j,k,l} \leq c^{student}_j \; \forall \; j$
  Students do not enroll in more courses than they are willing.
- $\sum_k y_{j,k,l} \leq \underbrace{1}_{\text{or } a^{student}_{j,k}} \; \forall \; j, l$
  Students do not enroll in more than one course per week. They may be able to enroll in no courses some weeks.
- $\sum_k x_{i,k,l} \leq \underbrace{1}_{\text{or } a^{teacher}_{i,k}}$
  Teachers do not teach more than one course per week. They may be able to teach no courses some weeks.
- $\sum_j \sum_k y_{j,k,l} \leq c^{location} \; \forall \; l$
  The total number of students in a week does not exceed the location's capacity.
- $\sum_j y_{j,k,l} \leq \sum_i c^{classcap}_k x_{i,k,l} \; \forall \; k, l$
  The number of students enrolled in a course in a given week cannot exceet the capacity of the sections of that course being taught that week.
- $\sum_i x_{i,k,l} \leq c^{sections}_{k,l} \; \forall \; k, l$
  The number of teachers teaching a given course in a given week does not exceed the number of sections of that course that can be taught in that week (with materials in mind).

- $\sum_{l} y_{j,k,l} \leq \underbrace{1}_{\text{or } q^0_{j,k}} \ \forall\ j,k$

  No student takes any course more than once. They may never be able to take some courses.

- $x_{i,k,l} \leq x^0_{i,k} \ \forall\ i,k,l$

  A teacher teaches a course in *some* week if and only if the teacher teaches that course in *at least one given* week.

- $\sum_{k} x^0_{i,k} \leq c^{curriculum} \ \forall\ i$

  No teacher teaches more than 3 different courses in a summer.

- $\sum_{i} g x_{i,k,l} \leq \sum_{j} y_{j,k,l} \ \forall\ k,l$

  The number of teachers teaching a course in a given week (the number of sections) should not exceed the minimum student population required to attain the minimum class size to offer each section of that course.

4 Computational Complexity

If every teacher can teach every course in every week and every student can take every course in every week, then the number of (binary) decision variables is exactly

$$\underbrace{n_{teachers} n_{courses} (n_{weeks} + 1)}_{x_{i,k,l} \ \forall\ i,k,l \text{ and } x^0_{i,k} \ \forall\ i,k} + \underbrace{n_{students} n_{courses} n_{weeks}}_{y_{j,k,l} \ \forall\ j,k,l}$$

However, many of the decision variables are constrained to zero, for example if teacher i cannot or will not teach course k, then $p_{i,k} = p^0_{i,k} = 0$, and so $x_{i,k,l} = x^0_{i,k} = 0 \ \forall\ l$. Similarly, if student j will not take course k, then $q_{j,k} = q^0_{j,k} = 0$, and so $y_{j,k,l} = 0 \ \forall\ l$. The computational cost of assigning zero-constrained variables is negligible.