

Davidson College Classroom Matching Problem

Hien Tran, Miguel Donaldo

April 21, 2023

Abstract

Your abstract.

1 Introduction

One of the key challenges in managing a university's academic schedule is matching available classrooms to all the courses that need to be offered. This task becomes even more complex when professors' availabilities and preferences are taken into account. The problem is to find an optimal assignment of classrooms to courses that takes into consideration the professors' availability and preferences, in order to create a schedule that is both efficient and satisfactory for all parties involved.

To solve this problem, we can turn to Linear Optimization Problems (LOPs), which are mathematical models that can be used to optimize the allocation of resources. By formulating the problem as an LOP, we can use algorithms to find an optimal solution that minimizes conflicts and maximizes the utilization of resources. The goal is to create a schedule that satisfies all constraints while minimizing the cost, such as classroom usage or professor preferences.

This problem is especially important in higher education institutions where schedules must be made well in advance and are often complex due to the wide range of courses offered, the availability of professors, and the need for efficient classroom usage. An optimized scheduling system can save time and resources while providing a better learning experience for students and a more productive work environment for professors.

2 Methodology

2.1 Overview

In this section, we will discuss two linear programming models that we developed using binary linear programming. In both models, we utilized binary decision variables to represent whether a course is assigned to a particular classroom at a given time. We also imposed several constraints to ensure that the results align with logical expectations.

The two models differ primarily in their objective functions. In the first model, we set the objective function to 0, effectively treating all feasible solutions as equally optimal. This model is useful when we want to verify the existence a feasible distribution of courses across classrooms and time slots, without considering other factors such as the preferences of professors or students.

In contrast, in the second model, we set the objective function to maximize the satisfaction score. This model takes into account the preferences of professors and students, aiming to maximize their overall satisfaction with the assigned schedule. To achieve this objective, we added additional constraints that limit the number of courses assigned to each professor, prevent scheduling conflicts for students enrolled in multiple courses, and maximize the utilization of available classrooms.

By developing these two models, we can provide different scheduling options to cater to different priorities and preferences. We can use the first model to ensure a feasible distribution of courses across classrooms and time slots, while the second model can be used to maximize the satisfaction of professors and students. The choice of model will depend on the specific needs and objectives of the school and the stakeholders involved in the scheduling process.

2.2 Data Preparation

The data preparation for our mathematical model involved several steps to ensure that we had accurate and relevant information for optimization.

In this case, the first step was to gather the necessary data to optimize the matching of classrooms to courses, taking into account professors' availabilities and preferences. To begin, we collected real data from a semester at Davidson College, which allowed us to identify the courses and the corresponding professors. This involved extracting data from the school's course schedule website and carefully processing it to create a dataset containing course codes, course names, and professor names.

In addition to the course and professor data, we also required data on the classrooms and time periods available for scheduling. To obtain this information, we conducted a comprehensive survey of the school, and consulted the school course schedule website to manually gather information on the availability of classrooms and the time periods in which they could be used. We carefully compiled this data into a dataset containing classroom codes, classroom names, and available time periods for each classroom. Lastly

Once all the necessary data was collected, we preprocessed the data to make it usable for optimization. This involved meticulously cleaning the data to remove any missing information.

2.3 Assumptions

There are several assumptions that we made along the way in order to simplify the model:

- **Equal Class Size and Classroom Capacity:** Our scheduling model is based on the assumption that the maximum number of students allowed in each class must not exceed the capacity of any classroom available for scheduling. Therefore, any course may be assigned to any classroom without constraints, as long as the capacity of the classroom is sufficient to accommodate the number of enrolled students. This assumption simplifies the scheduling problem and makes it more tractable, while ensuring that the solution remains feasible and realistic.
- **Fixed Weekly Meeting Time for All Courses:** We assume that all courses have a fixed weekly meeting time of exactly 150 minutes. We have chosen to only consider time periods that satisfy this condition, which means that a small portion of courses whose weekly class time is longer or shorter than 150 minutes were excluded from our analysis. This assumption simplifies the scheduling problem by providing a consistent basis for course scheduling and ensuring that all courses are treated equally with respect to their weekly meeting time. However, it is worth noting that in reality, there may be variation in the weekly meeting time of courses, and this may impact the scheduling of classes.
- **No Preferences in Classroom Location:** Our scheduling model is based on the assumption that professors' classroom preferences are not considered in the course scheduling process. Specifically, we assume that professors do not have a preference for specific classrooms or locations within the school. In reality, however, it is common for professors to have preferred classrooms or teaching locations, based on factors such as comfort, familiarity, or accessibility.

2.4 Models

2.4.1 Definition of Variables

To provide clarity, we will begin by defining the relevant variables. Let I , T , P , and C denote the sets of all courses, time periods, professors, and classrooms, respectively. We define $x_{itc} \in X = \{x_{itc} : i \in I, t \in T, c \in C\}$ as a binary decision variable representing whether course i is assigned to classroom c during time period t . For each course i , $T_i \subseteq T$ represents the set of available time periods, and α_{it} denotes the satisfaction score if course i is assigned to time period t . For each professor $p \in P$, $I_p \subseteq I$ denotes the courses taught by professor p . Finally, let $T_{overlap} = \{(t_1, t_2) : t_1 \neq t_2 \wedge overlap(t_1, t_2) = \text{TRUE}\}$ be the set containing a pairwise unordered tuple of time periods that overlap.

2.4.2 Model 1: Basic Model

Our basic model, we only want to verify the feasibility of our dataset where every course can be assigned to a classroom at a particular time. Our Basic Model is

$$\max 0$$

subject to

$$\sum_{i \in I} \sum_{t \in T} \sum_{c \in C} x_{itc} = |I| \quad (1)$$

$$\sum_{i \in I} x_{itc} \leq 1 \quad \forall t \in T, c \in C \quad (2)$$

$$\sum_{t' \in T_i} \sum_{c \in C} x_{it'c} = 1 \quad \forall i \in I \quad (3)$$

$$\sum_{c \in C} \sum_{i' \in I_p} x_{i'tc} \leq 1 \quad \forall t \in T, p \in P \quad (4)$$

$$\sum_{c \in C} \sum_{i' \in I_p} x_{i't_1c} + x_{i't_2c} \leq 1 \quad \forall (t_1, t_2) \in T_{overlap}, p \in P \quad (5)$$

$$\sum_{i \in I} x_{it_1c} + x_{it_2c} \leq 1 \quad \forall (t_1, t_2) \in T_{overlap}, c \in C \quad (6)$$

$$x_{ijk} \in \{0, 1\} \quad \forall i \in I, t \in T, c \in C \quad (7).$$

In the objective function, we set it to 0 to get a feasible solution. Constraint (1) ensures every course get assigned to one classroom at a specific time. Constraint (2), (6) ensure that every classroom is only occupied by only one course at a time and at overlapping time periods. Constraint (3) guarantees that every course is assigned only to one of the available time periods of the professor who teaches it. Constraint (4), (5) ensure that no classes taught by the same professor occur at the same time or the overlapping time periods. We included constraint (7) to specify that all decision variables are binary.

2.4.3 Model 2: Ranked availability

In this model, we simply modify the objective function to maximize the satisfaction score

$$\max \sum_{i \in I} \sum_{t \in T} \sum_{c \in C} \alpha_{it} x_{itc}$$

subject to

$$\sum_{i \in I} \sum_{t \in T} \sum_{c \in C} x_{itc} = |I| \quad (1)$$

$$\sum_{i \in I} x_{itc} \leq 1 \quad \forall t \in T, c \in C \quad (2)$$

$$\sum_{t' \in T_i} \sum_{c \in C} x_{it'c} = 1 \quad \forall i \in I \quad (3)$$

$$\sum_{c \in C} \sum_{i' \in I_p} x_{i'tc} \leq 1 \quad \forall t \in T, p \in P \quad (4)$$

$$\sum_{c \in C} \sum_{i' \in I_p} x_{i't_1c} + x_{i't_2c} \leq 1 \quad \forall (t_1, t_2) \in T_{overlap}, p \in P \quad (5)$$

$$\sum_{i \in I} x_{it_1c} + x_{it_2c} \leq 1 \quad \forall (t_1, t_2) \in T_{overlap}, c \in C \quad (6)$$

$$x_{ijk} \in \{0, 1\} \quad \forall i \in I, t \in T, c \in C \quad (7).$$

In the objective function, we set it to maximize the satisfaction score α_{it} corresponding to course i is assigned to time period t . All constraints are similar to the Basic model.