Computers & Operations Research 52 (2014) 29–38

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

A fix-and-optimize heuristic for the high school timetabling problem

Árton P. Dorneles [a], Olinto C.B. de Araújo [b], Luciana S. Buriol [a,*]

[a] Instituto de Informática, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, RS, Brazil
[b] Colégio Técnico Industrial de Santa Maria, Universidade Federal de Santa Maria, 97105-900 Santa Maria, RS, Brazil

ARTICLE INFO

Available online 5 July 2014

Keywords:
High school timetabling
Mixed integer linear programming
Fix-and-optimize
Variable neighborhood descent
Matheuristics
ITC-2011

ABSTRACT

The high school timetabling is a classical combinatorial optimization problem that takes a large number of variables and constraints into account. Due to its combinatorial nature, solving medium and large instances to optimality is a challenging task. When resources are tight, it is often difficult to find even a feasible solution. Among the different requirements that are considered in Brazilian schools, two compactness requirements must be met on a teacher's schedule: the minimization of working days and the avoidance of idle timeslots. In this paper, we present a mixed integer linear programming model and a fix-and-optimize heuristic combined with a variable neighborhood descent method. Our method uses three different types of decompositions - class, teacher and day – in order to solve the high school timetabling problem. The method is able to find new best known solutions for seven instances, including three optimal ones. A comparison with results reported in the literature shows that the proposed fix-and-optimize heuristic outperforms state-of-the-art techniques for the resolution of the problem at hand.

1. Introduction

A common task to all educational institutions is the provision of a class assignment that combines teachers, students, rooms and timeslots (periods) that, in practice, constitute a solution to the high school timetabling problem. High quality solutions are relevant for financial and pedagogical reasons. In addition to feasibility, such a solution must be improved as much as possible, satisfying requirements and constraints of different nature. The quality of a solution to the high school timetabling problem leads to better use of resources and economic gains, thus directly influencing the quality of teaching and learning, as well as working conditions for teachers.

In spite of it's relevance, the problem is often manually solved in small institutions. However, the automation of this task is becoming common, and nowadays it is widely adopted (or even mandatory) in large institutions.

Timetabling requirements are separated into *hard* and *soft* ones. By hard requirements we mean those that must be satisfied, while soft requirements are those that may be violated, but should be satisfied whenever possible. Soft requirements have different levels of importance and are oftentimes conflicting with each other. Thus, it may be impossible to satisfy all of them. Typically, the quality of a solution is associated directly to the satisfaction of soft requirements. The more soft requirements are satisfied, the better a solution is considered.

The timetabling problem first appeared in the scientific literature in the 1960's [1] and since then it has been subject of intense research. The most basic problem is to schedule a set of class-teacher events (or meetings) such that no teacher (nor class) is assigned to more than one lesson at any timepoint. This basic problem can be solved in polynomial time by a min-cost network flow algorithm [2]. However, in a real-world application, teachers can be unavailable in some periods. Therefore, when this constraint takes place, the resulting timetabling problem is NP-complete [3].

In fact, this problem has several NP-complete variants proposed in the literature and the set of objectives and requirements depends mostly on the context of the application, the school and the place where it is located [4–6]. For instance, in certain countries, including Brazil, it is common that a teacher works in more than one school, holding several jobs. In order to allow for this possibility, it is important to compact the lessons in each school to the minimal number of days. Furthermore, the solution also requires the reduction or elimination of *idle periods* between lessons in a teacher's schedule in compliance with pedagogical demands or personal preferences. Teachers can also request that his/her lessons be taught in two consecutive periods (*double lessons*). This set of requirements

* Corresponding author. Tel.: +55 51 3308 6827; fax: +55 51 3308 7308.
E-mail addresses: arton.dorneles@inf.ufrgs.br (Á.P. Dorneles), olinto@ctism.ufsm.br (O.C.B. de Araújo), buriol@inf.ufrgs.br (L.S. Buriol).

http://dx.doi.org/10.1016/j.cor.2014.06.023

define a problem called Class-Teacher Timetabling Problem with Compactness Requirements (CTTPCR).

This work extends previous studies about the CTTPCR. We propose a fix-and-optimize heuristic combined with a variable neighborhood descent (VND) method [7] using three different types of decompositions (class, teacher and day). We carry out extensive experiments to support the conclusions about the performance of our method in finding high quality solutions.

The remainder of the paper is organized as follows. Section 2 presents a review on the timetabling problem resolution. Section 3 presents the problem investigated in this study, the notation used to represent it, and a mixed integer linear programming formulation. Section 4 presents the proposed method for solving the problem. Section 5 presents experimental results considering both synthetic and real-world instances. Finally, Section 6 concludes the paper.

2. Related work

The timetabling problem has been intensively investigated since the 1960s [1]. We refer the reader to the excellent survey by Schaerf [6], which presents the main variants of the timetabling problem, its formulations and solution approaches.

For a few decades variants of the timetabling problem were considered intractable by exact methods, since only small instances, or simplified variants of the problem, were shown to be solved in reasonable time. However, in the last years several improvements have been made in mixed integer linear programming (MIP) solvers [8], which motived new research studies using this approach [9].

The CTTPCR problem is originated from the Brazilian High School System and it was first defined by Souza and Maculan [10]. In Souza [11] a MIP formulation for CTTPCR was presented, as well as a set of instances, which became a basic testbed used until today. Recently, Santos et al. [12] proposed an extended formulation and applied a column generation algorithm, providing the best known lower bounds for the testbed CTTPCR instances.

In a previous work [13], we empirically showed that the minimization of the teachers' idle periods turns the resolution of the problem considerably more difficult through general purpose solvers. In addition, in that work we proposed a new MIP model by reformulating the idle periods requirement previously proposed by Santos et al. [12]. Even though the reformulation allows one to solve slightly larger instances of CTTPCR, it is still impracticable to solve large real-world instances from the testbed. In general, scheduling problems, which include the timetabling problem, are considered hard to solve directly via general purpose solvers [8].

One technique that is useful in this context is the use of matheuristics, which combine mathematical programming and heuristic methods [14]. Currently, matheuristics have been successfully applied to solve several optimization problems across a range of applications. Fix-and-optimize is a matheuristic that iteratively decomposes a problem into smaller subproblems. In each iteration of the algorithm, a decomposition process is applied aiming at fixing most of the decision variables at their value in the current solution. Since the resulting subproblem is composed only by a small group of "free" variables to be optimized, each subproblem can be solved fairly quickly by a MIP solver, when compared with the full model. The solution obtained in each iteration becomes the current solution when it improves the objective value. In further iterations of the algorithm, a different group of variables is selected to be optimized. This process is repeated until a termination condition is satisfied. The fix-and-optimize heuristic was proposed independently by Gintner et al. [15] and Pochet and Wolsey [16]. In the latter, the method was called *exchange*, designed to improve the relax-and-fix heuristic [17]. However, the name fix-and-optimize used by the former was adopted in the literature.

Regarding the educational timetabling problem, there are few studies in the literature exploring matheuristics. To the best of our knowledge, there is a limited number of publications related to course timetabling, including, e.g. [18,19]. Further, perhaps only one reference that makes use of matheuristics applied to high school timetabling has been published until now. In this work, Avella et al. [20] proposed a two-phase algorithm applied to a problem similar to CTTPCR. The first phase is a simulated annealing (SA) algorithm. The second phase consists in a large-scale neighborhood search that decomposes the problem into subproblems which are solved independently by a MIP solver. In each subproblem all teachers remain fixed, with the exception of a randomly chosen one.

3. Problem definition and modeling

The goal of the Class-Teacher Timetabling Problem with Compactness Requirements (CTTPCR) is to build a weekly timetable. The week is organized as a set of days D, and each day is split into a set of periods P. Let C be a set of classes and T a set of teachers. A class $c \in C$ is a group of students that follow the same course and have full availability. A *timeslot* is a pair, composed of a day and a class period (d,p), with $d \in D$ and $p \in P$, wherein all periods have the same duration. Teachers $t \in T$ may be unavailable in some timeslots.

The main input for the problem is a set of events E that should be scheduled. Typically, an event is a meeting between class and teacher to address a particular subject in a given number of lessons (*workload*) in a given room. Particularly in the Brazilian context, a class, a teacher and a room are pre-assigned to each event $e \in E$. In addition, each event defines how lessons are distributed over a week by requesting an amount of double lessons, restricting the daily limit of lessons, and defining whether lessons taught on the same day are consecutive or not.

A *feasible* timetable has a timeslot assigned to each lesson of events satisfying the hard requirements H1–H6 below:

- H1 The workload defined in each event must be satisfied.
- H2 A teacher cannot be scheduled to more than one lesson in a given period.
- H3 Lessons cannot be taught to the same class in the same period.
- H4 A teacher cannot be scheduled to a period in which she/he is unavailable.
- H5 The maximum number of daily lessons of each event must be respected.
- H6 Two lessons from the same event must be consecutive when scheduled for the same day, in case it is required by the event.

Besides feasibility regarding hard constraints, as many as possible of the soft requirements S1–S3 stated below should be satisfied:

- S1 Avoid teachers' idle periods.
- S2 Minimize the number of *working days* for teachers. In this context, working day means a day that the teacher has at least one lesson assigned to her/him.
- S3 Provide the number of double lessons requested by each event.

3.1. Problem formulation

In this subsection we present a MIP formulation for the CTTPCR problem considering all the hard and soft requirements mentioned above. The notation used in the problem formulation is presented in Table 1.