UNICA

UNIVERSITÀ
DEGLI STUDI
DI CAGLIARI

**Ph.D. DEGREE IN**

Mathematics and Computer Science

Cycle XXXV

**TITLE OF THE Ph.D. THESIS**

# New models and algorithms for the timetables of Italian high schools

Scientific Disciplinary Sector(s)

MAT/09


Ph.D. Student: Claudio Crobu

Supervisor Prof. Massimo Di Francesco

Co-Supervisor Prof. Enrico Gorgone


Final exam. Academic Year 2021/2022
Thesis defence: February 2024 Session

*To my parents, Erminia and Pietro, and my brother Roberto*

# Abstract

The thesis investigates new timetabling problems, which are motivated by the complex case of Italian schools. Generally speaking, the *High School Timetabling Problem* is aimed at giving the right order to the meetings between groups of students and teachers. However, it requires to pre-assign teachers to the classes, i.e. establishing which teacher will teach each subject in each class. This problem is called *Class Teacher Assignment Problem* and is faced by a mixed integer programming formulation, which also aims at easing the solution of the following *High School Timetabling Problem*. Moreover, the thesis shows that *Class Teacher Assignment Problem* can also be adapted to face the case where teachers must give lectures in more than one school (*Multi-school Timetabling Problem*). In particular, this thesis investigates a complex variant of the High School Timetabling problem w.r.t. a given assignment of teachers to classes. This problem presents requirements, which enforce to (i) provide teachers with the same idle times, (ii) avoid consecutive days with heavy workload, (iii) limit multiple daily lessons for each class, (iv) introduce shorter time units to differentiate entry and exit times. An integer programming model is presented for this problem, which is denoted by Italian High School Timetabling Problem [IHSTP]. However, requirements (i), (ii), (iii) and (iv) cannot be expressed according to the current XML High School Time-Table [XHSTT] standard. Since the [IHSTP] model is very hard to solve by an off-the-shelf solver, a two-step optimization method is presented: the first step optimally assigns teachers to lesson times and the second step assigns classes to teachers. An extensive experimentation is performed on the model by realistic and real instances from Italian schools, as well as benchmark instances from the literature. Finally, the experiments show that the method is effective in solving both this new problem and the simplified problem without the new requirements.

# Contents

.

# List of Figures

# List of Tables

.

# Glossary

*Articulated class*
A class made with the union of two or more *classes* with a small number of students.

*Block*
Two lessons for two pairs of *classes* and teachers who have to work together or separately in the same time slot.

*Class (or group)*
Group of students taking lessons from the same *curriculum* at the same time.

*Co-presence teacher (co-teacher)*
A teacher who always works together with another colleague.

*Curriculum*
It is the set of subjects in a class and the number of lessons for each subject.

*Daily period*
A time interval with a constant duration equal to the minimum lesson unit.

*Day off*
A day when the teacher does not teach.

*Double lesson*
A lesson with length of two periods which must be consecutive for the same class and teacher.

*Fractional time unit*
A period with a duration of a fraction of an hour.

*Free day*
A day without teaching commitments for a teacher.

*Full-time teacher*
A teacher with a weekly workload equal to a fixed number of hours.

*Idle time*
A pause between two non-consecutive lessons of a teacher.

*Lesson unit*
The minimum interval of time of a lesson (normally it is equal to one hour).

*Multiple lesson*
A lesson with length of some periods which must be consecutive for the same class and teacher.

*Part-time teacher*
A teacher with a reduced weekly workload compared to a *full-time teacher*.

*Split lesson*
A lesson which is not given in consecutive periods by a teacher in a class.

*Triple lesson*
A lesson with length of three periods which must be consecutive for the same class and teacher.

# Chapter 1

# Introduction

Educational timetabling is a scientific area motivated by the problems arising in over 18,000 universities ([UNI]), 65,000 schools ([TEC]) and a billion of students in hundreds of countries around the world.

The scientific research on educational timetabling is split into two main areas: curriculum-based timetables and class-based timetables [KS13],[N.14],[CDS23]). In the first case, individual students select which courses they want to attend (e.g. universities and Danish high-schools [DAN]). In the second case, groups of students are grouped into classes attending the same program of studies. For example, this arises in Italian high-schools, but similar settings also take place in Greece, Spain and France.

This thesis investigates the challenging problem of class-based timetabling, as motivated by Italian schools.

In September of each year, 5322 Italian high-schools face the problem of scheduling lessons for students and teachers [MIM]. In Italy schools are open from Monday to Saturday, sometimes up to Friday. Every group of student (from now on *class*) must have lessons for the same number of hours per day (usually 5).

In Italy, schools have to care about the surveillance of students, who are entrusted by family. As a result, the student timetable must be compact, i.e. without any break between lessons.

Instead, teachers' timetables may not be compact, that is, equipped with some breaks between a lessons (or idle times).

However, these breaks are not particularly appreciated by teachers, who label them as "empty" hours, thus underlining in some way an unwanted waste of time.

This work is carried out every year, over and over again, by some teachers who are part of the management staff. They are called upon to provide acceptable solutions to the requests made within a very short time period.

Several requirements must be taken into account in high-school timetabling. First of all, schools have to satisfy the plan envisaged for each year of a particular field of study, set by the Ministry of Education. This is quantified in a set of subjects foreseen for each class with the relative number of lesson hours per week.

Next, there are several requests, collected at the beginning of the year as wishes from the teachers (*"desiderata"*), usually concerning the choice of the free day, i.e. a day without teaching commitments for each teacher.

Additional strong and rigid constraints arise from the need for some teachers to be able to carry out important additional functions in the school environment.

Furthermore, there are other teachers who, against their will, also have to give lessons in other schools, sometimes far from the main school.

All of this, in short, would be called the high school timetabling problem. However, in reality some circumstances heavily influence the achievement of an effective solution within acceptable times for all the actors involved in the school: students, teachers, auxiliary and technical staff.

This problem cannot be faced as a whole and is decomposed in several phases. In the first, some important and truly strategic decisions are made by the vice principal of the school: the so-called assignment of classes to all teachers.

This activity takes into consideration many aspects from the most restricted to the merely desirable ones. It has the aim of producing a list of subject assignments to classes and teachers in such a way as to satisfy these requirements:

- each class of a certain year and a certain field of study must include a given standard set of subjects, each of which is taught by one or more teachers for a fixed number of hours per week;
- each full-time teacher must give lessons in his/her subject(s) of study for a total number of hours per week established by the Ministry, with rare exceptions;
- each part-time teacher must teach lessons for a smaller number of hours than their full-time colleagues;
- a class of the previous year attending lessons with a teacher should in principle keep the same teacher (rule of teaching continuity), obviously if the teacher is still working in the school;
- according to seniority rankings of the teachers, some desires may be satisfied in harmony with the previous constraints.

These considerations make it clear that the problem, as well as being strategic, is also rather complex and unfortunately, even in this case, the time horizon for the specific solution is always particularly short for determining optimal solutions.

In the following, those specific problems that arise from the all-encompassing school timetable problem will be defined, providing a correct context and a rigorous definition. In particular, the relevant temporal phases and the related problems to be addressed will be presented.

Figure 1.1 schematically illustrates the decision-making levels, as indicated in the context of Operations Research, the definition of the corresponding specific problem, the decision-makers and, above all, the objectives for each phase.

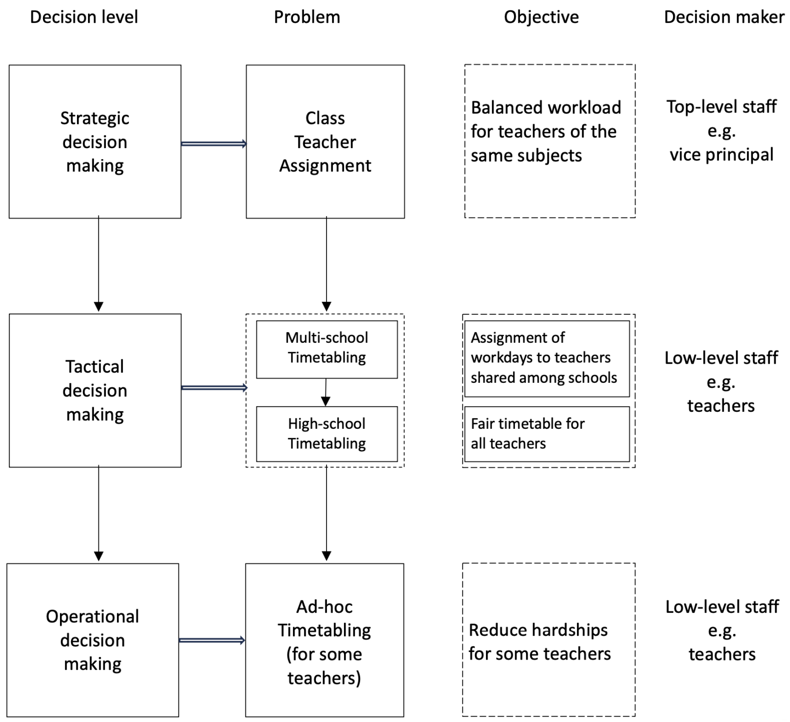



Figure 1.1: School timetabling process phases

As a result, four separate problems must be solved:

1. the problem of assigning each teacher a set of classes and subjects to teach (Class Teacher Assignment Problem);

2. the problem of splitting the working days of teachers operating in more than one school (Multi-school Timetabling Problem);

3. the classic problem of the school timetable (High-school Timetabling Problem).

4. the problem of adapting the school timetable for a few teachers due to particular needs (Ad-hoc Timetabling Problem)

This last problem is by its nature so simple and changeable that it is considered unnecessary to define it formally and propose a mathematical model.

In order to efficiently address the first problem in the former list, the thesis proposes a model to partition clusters of classes and teachers into (almost) independent subsets, to lead a separable timetabling problem, which is separable for each cluster. This thesis shows that same idea can be applied to multi-school timetabling problems.

This thesis is organized as follows. In chapter 2 the Italian High School Timetabling problem will be faced. In chapter 3 the Class Teacher Assignment problem will be investigated. For each of them, the mathematical models will be described, also including an efficient heuristic applicable in many cases.

Furthermore, extensive experimentation on both problems will be provided. In the conclusion of the thesis, promising research developments will be listed, also including integrations with research already done in the past.

# Chapter 2

# The Italian High-School Timetabling Problem

## 2.1 Overview of the Italian High-School Timetabling Problem

The High School Timetabling problem (HST) is a relevant research area, which aims to schedule lectures to time slots. Its characteristics are country-dependent [KS13] and several solution approaches were proposed [N.14]. The introduction of the XML High School Time-Table [XHSTT] format for the HST problem has provided a uniform way to support a variety of possible constraints. However, new requirements have emerged in the case of Italian high schools and some of them do not fit with this format. This research is motivated by such a case.

The recent reforms in 2008 and 2015 deeply changed the educational structure of Italian high schools to make a service more oriented to students and decrease system costs [MIU]. The reduction in the weekly extension of lessons has led to an irregular distribution of lectures. Moreover, when two classes of the same year (or level) have few students, schools are requested to merge them into the so-called *articulated class*, even if students have different curricula. Therefore, the students of articulated classes may have few subjects in common and must be split when the different characterizing subjects are taught. In addition, full-time teachers must give lessons for 18 hours a week and, if this workload is not complete, they must be enrolled in other schools. Yet, some teachers may have additional days-off to account for possible additional duties.

These reforms increased the number of idle times for teachers, who claim that this number must be the same for all of them for the sake of equity. The new rules may also result in the planning of timetables with heavy workloads in consecutive days and lead to the burn-out of teachers. Moreover, it is recommended to schedule school days of same duration for a class to plan the transportation of students smoothly, even if this situation leads to an increase in the number of lessons. In addition, it is important to diversify the entrance and exit times of classes to limit crowds, as emphasized by the recent pandemic event.

Although relevant research exists in the HST problem for Italian schools [Sch99, ADSV07], it dates back in time and the recent changes in requirements were not taken into account. In this thesis they are investigated and added to well-recognized requirements for the HST problem (e.g. assign teachers to classes, full-time and part-time teachers with one or more days off, surveillance in each class at any time slot.). The complete list of requirements is provided in Section 2.2. All in all, the new problem is denoted by the *Italian High-School Timetabling Problem* [IHSTP].

This thesis presents an Integer Programming formulation for the [IHSTP], which is denoted by IHSTT. Since large-scale instances cannot be solved efficiently by an off-the-shelf optimization solver, a two-stage decomposition is presented. In the first stage, teachers are assigned to time slots in the so-called *Teacher Profile Problem* [TPP] through a MIP formulation denoted by TP. In the second stage, a restricted version of IHSTT is solved from the solution of TP. The overall method is denoted by TP-IHSTT. Since some of the new requirements are not supported by the [XHSTT] format, all models are implemented by a general-purpose modeling language and solved by a MIP (Mixed-Integer Programming) solver.

The two-step method is extensively tested in several instances, in order to assess to what extent it can be adopted. More precisely, in the first part of the experimentation all requirements are considered and compare the solutions of the MIP solver for the IHSTT and those provided by the two-step method, in which each sub-problem is solved by the same MIP solver. In the second part a simpler problem without the new features of [IHSTP] is focused and the solutions of the MIP solver running IHSTT and the formulation proposed by [KSS15], which is denoted by KSS, are compared. The KHE heuristic [KHE] [Kin14] is also adopted to enrich the comparison. All variants are run without and with the TP step. In the second case, the methods are denoted by TP-KHE, TP-KSS and TP-IHSTT.

The experiments show the effectiveness of the two-step method, because it determines high-quality solutions for the problem at hand in terms of CPU times, costs and optimality gaps. Moreover, it is also effective for a simplified problem devoid of the new requirements: the method can be successfully applied both to KSS and IHSTT, but the results are far better in the latter case. Finally, IHSTT can effectively be used to solve some well-known benchmarks in the literature.

This chapter is organized as follows. In section 2.2 the specific requirements for Italian high schools are presented. In section 2.3 the related work is critically discussed, to compare Italian problem with the case of other countries. In section 2.4 a complete Integer Programming formulation for Italian high schools is defined. In section 2.5 the two-step method is presented and the [TPP] is described and formulated. In section 2.6 experimental results are presented.

## 2.2 Italian High School Timetabling Problem

In Italian schools each student belongs to a class (or group of students) sharing the same lessons according to a *curriculum*. All students in a class must follow the same set of subjects for a fixed number of weekly hours. Lessons are daily organized in time slots (e.g. 1 hour, but fractional lesson units are also possible) and must be placed in a time horizon, which normally spans over a week and is repeated periodically for the entire school year. Lessons may span over multiple consecutive time slots, to accommodate special needs as in-class works or lab activities. These lessons are called multiple lessons (e.g. double and triple lessons).

Each subject is taught by a teacher or, more rarely, by a teacher and a co-presence teacher (or co-teacher), who has to teach always together with another colleague supervising the activity. From now on, for the sake of simplicity, teachers and co-teachers will be denoted as teachers, unless one refers explicitly to co-teachers and non-co-teachers.

Teachers may give lessons on more than one subject in one or more classes. Schools open from Monday to Saturday (very seldom until Friday) and teachers must have a day off for rest. They are classified as full-time and part-time teachers. Full-time teachers have to teach for a fixed number of hours a week (typically 18 hours, but some reductions are possible to do some management tasks) and must work in others schools to complete their workload. Part-time teachers have a shorter workload according to their annual contract and may work for several schools. As a result, some teachers must receive more than one day-off from each school. For the same reasons, some teachers may not be available to teach in some specific times.

Clearly, a teacher should not be employed for very few time slots a day. Conversely, a workload spanning all time slots in a day is not recommended, to prevent burn-out. Time-slot breaks are possible between lessons, even if they are not always required or appreciated.

The objective is to build a timetable, i.e. assign each lesson to a specific time slot of each day, such that a number of requirements are satisfied. They are divided into mandatory (or hard) and desirable (or soft). For the sake of clarity, in the following all requirements are enumerated and denoted if they are hard or soft.

- R1 (hard) - Each class has to attend lessons for a given set of weekly days and a consecutive set of hours a day, as established by the school. For example, in a school all classes of the fifth year have to attend 32 hours a week and 5 hours a day, except on Tuesday and on Thursday, in which lessons are given for six hours. Every class of the second year has to attend lessons for 33 hours a week, in which the additional hour w.r.t. fifth year classes is given on Saturday.

- R2 (hard) - Every teacher has to teach for a fixed number of hours as established by national laws or school rules.

- R3 (hard) - Every teacher must have at least one day off a week. It can be determined according to two school-dependent policies: the day off can be *a priori* selected by the school or its decision is left *a posteriori* during timetable planning. Therefore, any methodological proposal must be able to deal with both policies.

- R4 (hard/soft) - A subset of teachers must/may receive additional days off according to specific conditions (e.g. employment in several schools, special contracts, additional administrative tasks, etc.). Unlike R3, these conditions affect whole days instead of specific daily parts.

- R5 (hard) - Since classes spend different time periods at school (on a daily and weekly basis), a lesson must be scheduled for a class only when the class is at school. For example, a fifth-year class cannot attend any lesson in the sixth hour on Saturday, if only five hours of lessons are scheduled for that class.

- R6 (hard) - A lesson must not be scheduled for a teacher in the case of specific commitments in specific periods of a day (e.g. employment in another school, special contracts, additional administrative tasks, etc.).

- R7 (hard) - Each class has to be taught by a given teacher for a fixed number of weekly hours. This number is called *week requirement* and is established by laws or school rules.

- R8 (hard) - A teacher-clash must be avoided: a teacher cannot teach simultaneously in two classes, unless they form an articulated class.

- R9 (hard) - A class-clash must be avoided: two teachers cannot teach the same class at the same time; the only exception is represented by the so-called co-teaching lessons (e.g. in some lab lessons).

- R10 (hard/soft) - The multiple lessons of a teacher in a class should be consecutive. It is important for multiple lessons of the same teacher in a class to be consecutive in a day. Clearly, a hard requirement for not splitting lessons could prevent the determination of a feasible timetable. As a result, both hard and soft options are possible. Moreover, consecutive lessons are welcome to have in-class works or written exams.

- R11 (hard) - For a limited number of hours, an articulated class must be divided into two or more groups attending different lessons with dedicated teachers. For example, a class could attend the lessons on the second foreign language with two different teachers at the same time: one for French and one for Spanish. The problem doubles in the case of co-teachers in articulated class: for example, if this class has two groups of students and the split groups must attend a lab lesson in co-teaching, four teachers must be involved with the class at the same time.

- R12 (hard) - Block lessons must be scheduled. These lessons take place at the same time for two or more classes, in order to share possible resources (e.g., gym or specialized language teachers). Blocks could also support the ordering of lessons by an optional offset, to enforce one lesson to precede another one in a class by a given number of periods.

- R13 (hard) - Preassigned lessons must be scheduled. In these lessons a teacher is already assigned to a class in a given period of a given day. They are often adopted when a teacher gives lessons for a short number of hours in a school.

- R14 (soft) - This requirement enforces a balanced distribution of the lessons among the workdays for a teacher in a class. This requirement can be denoted by horizontal distribution. For example, it holds for a teacher working for one hour on Monday, Tuesday, Thursday and Friday and for two hours on Wednesday (on Saturday no lessons are possible because the teacher must have a day-off).

- R15 (soft) - This requirement guarantees a balanced distribution among daily periods for a teacher in a class. This requirement can be referred to as vertical distribution. For example, it holds when a teacher gives lessons in a class no more than once in the first daily period, no more than once in the second daily period and so on.

- R16 (hard/soft) - Every teacher must/may give lessons in between a minimum and maximum number of *additional* days off. These numbers can be conveniently set to zero, if appropriate.

- R17 (soft) - Every teacher is willing to have a weekly timetable with no idle times between consecutive lessons. However, this requirement is often difficult to achieve in practice for every teacher.

- R18 (hard/soft) - The duration of each multiple lesson in a week must/may be limited between a minimum and a maximum number of periods.

- R19 (hard/soft) - Each teacher must/should not reach the maximum workload in two consecutive days.

- R20 (hard/soft) - A minimum and maximum number of double lessons must/should be scheduled in the week for some pairs of classes and teachers.

- R21 (hard/soft) - A minimum and maximum number of triple lessons must/should be scheduled in the week for some pairs of classes and teachers.

- R22 (hard/soft) - The timetable must/should avoid the occurrence of too many multiple lessons for a class in a particular day.
- R23 (hard/soft) - The daily number of periods of a teacher in a class must/should be in between a minimum and a maximum value.
- R24 (hard) - Fractional periods must be introduced to differentiate the times to start and end lessons for groups of classes. As a result, the duration of all lessons must be multiple quantities of this fractional time unit. This requirement could be hard and enforced for all classes, but it could also be ignored for all of them for the sake of equity.
- R25 (hard) - All teachers must have the same number of idle times. This requirement is set to be hard, because it must be enforced for all teachers or ignored for all of them for the sake of equity.

## 2.3 Related work

Several studies investigated the HST problem by Integer Programming. The problem characteristics are country-dependent and depend on the organizational model, which could be class-teacher (e.g. Australia, Bosnia, Brazil, Greece, Italy and South-Africa), course-based (e.g. USA) or a mix of them (e.g. Denmark, England, Finland and Netherlands). In the class-teacher model lessons are given to all students of a class, whereas in the course-based variant students attend lessons according to their individual plan, as in university course timetabling [BSS18]. In the first case compact timetables are built from classes, which do not have idle times, whereas teachers typically have. In the second case, the timetable of teachers has no idle times, which can take place for students. This work is in the area of class-teacher models and, to my knowledge, this is the first study investigating requirements R19, R22, R24 and R25.

Several constraints were defined in [RTR15] on Bosnia and Herzegovina, but computational experiments are not provided. Moreover, it also neglects requirements on days off, obligation to take lessons (R5), balance of lessons spread in the week (R14, R15), teachers' idle times (R17) and limits on multiple daily lessons for classes (R18).

A lot of research was carried in Brazil on the so-called Class-Teacher Timetabling Problem with Compactness Constraints (CTTPC) ([SUOM12], [DdAB12], [DdAB14], [DdAB16], [SC17], [SSC18], [SSCdS20]).

Owing to the specific characteristics of Brazilian schools, these papers do not consider requirements on irregular weekly class layout (R5), articulated class (R11), block lessons (R12), balance of lessons spread in the week (R14, R15), multiple daily lessons limit for classes (R18) and restrictions on triple lessons (R21).

The Danish HST problem was described in [SS12] and [SD14]. However, they do not take into account requirements on weekly workload of teachers (R2), additional days off (R4), split lessons (R10), articulated classes (R11), balance of lessons spread in the week (R14, R15), multiple daily lessons limit for classes (R18), limits for number of double (R20) and triple lessons (R21), and restrictions on daily class-teacher workload (R23).

The French schools are investigated only in [Obe16]. However, it ignores the requirements on articulated classes (R11), block lessons (R12), balance of lessons spread in the week (R14, R15), limited idle times (R17), multiple daily lessons limit for classes (R18), limits on number of double (R20) and triple lessons (R21), and class-teacher workload (R23). Experimental results were provided only for one instance and presented very synthetically.

The Greek schools are investigated in [BDH97], [PVH03], [BDH09], [VGAH12] and [TIB20]. Unlike in the Italian case, there are no requirements on lessons spread in the week with respect to daily periods (R15), limits for number of double (R20) and triple lessons (R21).

The HST problem was investigated in Italy by [Sch99] and [ADSV07], who did not take into account the recent scholastic reforms. As a result, they could not consider the requirements on lessons spread in the week with respect to daily periods (R15), limits on the number of double (R20) and triple lessons (R21) and restrictions on the class-teacher workload (R23).

The HST problem was also generalized by [KSS15] and [FSCS17] to support the [XHSTT] format and adopt Integer Programming formulations. Although the set of requirements is wide, it is not exhaustive for the Italian case.

Table 2.1 reports which problem requirements are faced in the most recent literature on HST problem. Column R13 reports only methods dealing with this requirement explicitly. Nevertheless, every heuristic or MIP formulation can deal with R13 by fixing proper decision variables, even if it is not explicitly stated.

Therefore, one can notice that requirements R19, R22, R24 and R25 have not been investigated so far. This thesis covers this gap.

Moreover, the current version of [XHSTT] (XHSTT-2014 [XHS]) is described in [PKA+14] and does not support the new requirements R22, R24 and R25. The implementation of requirement R19 is possible, but it requires much effort and has some limitations. More precisely, at the moment this requirement must be implemented with a different value for any pair of consecutive days and any teacher. As a result, it would be simpler and more effective to *a priori* enforce the maximum workload between consecutive days, in order to simplify the implementation and decrease the memory issues.

The timetabling [XHSTT] logic is based on *a priori* enumeration of variable length sub-events covering an event (week requirement). However, another logic

| Year | Ref | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 | R10 | R11 | R12 | R13 | R14 | R15 | R16 | R17 | R18 | R19 | R20 | R21 | R22 | R23 | R24 | R25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1997 | [BDH97] | X | X | X | X | | X | X | | | | | | | X | | X | X | | | | | | | | |
| 1999 | [Sch99] | | X | X | | X | X | X | X | | | X | X | X | | | X | | X | | | | | | | |
| 2003 | [PVH03] | X | X | X | X | | | X | X | X | | X | | | X | | X | X | | | | | | | | |
| 2007 | [ADSV07] | X | X | X | X | X | X | X | X | X | X | | | X | X | | X | X | | | | | | | | |
| 2009 | [BDH09] | X | X | X | X | | X | X | X | X | X | | X | X | | | | X | X | | | | | | | |
| 2012 | [DdAB12] | | | X | X | | X | X | X | X | X | | | | | | X | X | | | X | | | X | | |
| 2012 | [SUOM12] | | | | X | | | | X | X | X | | | | | | | X | | | X | | | X | | |
| 2012 | [VGAH12] | X | | X | X | X | X | X | X | X | | | X | | X | | X | X | | | | | | | | |
| 2014 | [DdAB14] | | | X | X | | X | X | X | X | X | | | | | | X | X | | | X | | | X | | |
| 2014 | [SD14] | X | | X | | X | X | X | X | X | | | X | X | | | X | | | | | | | | | |
| 2015 | [KSS15] | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | | X | X | | X | | |
| 2015 | [RTR15] | X | X | | | | X | X | X | X | X | X | X | X | | | X | | | | X | X | | X | | |
| 2016 | [DdAB16] | | | X | X | | X | X | X | X | X | | | | | | X | X | | | X | X | | X | | |
| 2016 | [Obe16] | X | X | X | | X | X | X | X | X | X | X | | | | | X | | | | | | | | | |
| 2017 | [FSCS17] | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | | X | X | | X | | |
| 2017 | [SC17] | X | | | X | | X | X | X | X | X | | | | | | | X | | | X | | | X | | |
| 2018 | [SSC18] | X | | | X | | X | X | X | X | X | | | | | | | X | | | X | | | X | | |
| 2020 | [TIB20] | X | | X | X | X | X | | X | X | | | | | | | X | X | | | | | | | | |
| 2020 | [SSCdS20] | X | | | X | | X | X | X | X | X | | | | | | | X | | | X | | | X | | |

Table 2.1: Occurrence of the [IHSTP] requirements in the high school timetabling literature

is possible: to divide an event into sub-events with duration equal to 1 period. Since the second logic is expected to be less memory-consuming, it is of interest to develop a time-slot-based model and make a comparison to an event-based model. According to the example in [FSCS17], in an event with 4 periods, the KSS model can have 8 sub-events of different duration: 1, 1, 1, 1, 2, 2, 3 and 4. Conversely, in my model there are only 4 sub-events of 1 period each. Clearly this comparison is possible only in the problem without the new requirements. This work investigates this comparison, as well.

Finally, high school timetabling is a challenging research area and different communities of researchers and practitioners have worked on benchmark instances of this problem. This is also shown by the organization of the Third International Timetabling Competition (ITC2011) [PDGK+13].

IHSTT is compared to the Round 2 finalists, i.e. the evolutionary algorithm of team HFT [DH12] (Germany), the Simulated Annealing with Iterated Local Search of team GOAL [BFT+12], the hyper-heuristic [KOP12] (UK) and the Adaptive Large neighborhood Search of team LECTIO [SKS+12].

Their programs were run for 10 times on each instance with different random number seeds and the solutions were ranked from the highest to the lowest.

## 2.4 Mathematical model for the [IHSTP]

In this section the mathematical formulation for the [IHSTP] for a single school is presented. The formulation is based on the following sets: let $C$ be the set of classes (or groups of students), $T$ the set of teachers, $F \subseteq T$ the set of co-teachers, $D$ the set of days, $H$ the set of daily periods. $N$ is the set of possible periods in multiple lessons, for example $N = \{2, 3\}$ implies that only 2-period or 3-period multiple lessons are allowed. Note that the duration of a single period is shorter than the length of a lesson in the case of multiple lessons and/or fractional time units.

The following parameters are defined. Let $\chi_{ct}$ be the number of weekly lessons for class $c \in C$ with teacher $t \in T$ (this number is typically called *timetable requirement*). Preassigned lessons are denoted by parameter $\pi_{ctdh}$, which takes value 1 if a lesson has to be scheduled on day $d \in D$ at period $h \in H$ for class $c \in C$ with teacher $t \in T$, 0 otherwise.

In order to handle block lessons, consider any two classes $c', c'' \in C$ and any two teachers $t', t'' \in T$, and define the quantity $\phi_{c't'c''t''}$, which denotes the number of lessons that must be located in the same time slot for teacher $t' \in T$ in class $c' \in C$ and teacher $t'' \in T$ and class $c'' \in C$. Let $\mu_{ctf}$ be the number of weekly lessons of both teacher $t \in T \backslash F$ and co-teacher $f \in F$ in class $c \in C$.

Let $\underline{\epsilon}_{ct}$ and $\overline{\epsilon}_{ct}$ the minimum and the maximum duration of a multiple lesson for class $c \in C$ with teacher $t \in T$, whereas $\underline{\zeta}_{ct}$ and $\overline{\zeta}_{ct}$ are its minimum and maximum occurrence of multiple lessons of class $c \in C$ with teacher $t \in T$ in the week. It is denoted by $\theta_{ct}$ the penalty for the violation of multiple lessons of class $c \in C$ with teacher $t \in T$, denoting either the duration or the occurrence of the lesson on the basis of the instance requirement.

The following parameters are defined to link classes, days and periods. Let $\delta_{cdh}$ be a coefficient which takes value 1 if class $c \in C$ has to have a lesson on day $d \in D$ at period $h \in H$, 0 otherwise (note that, if $\delta_{cdh} = 0$, it is also possible for class $c \in C$ to have a lesson on day $d \in D$ at period $h \in H$). If a class $c \in C$ must not have a lesson on day $d \in D$ at period $h \in H$, the parameter $\beta_{cdh}$ has value 0, and 1 otherwise.

The following parameters are defined to link teachers, days and periods. Let $\gamma_{tdh}$ a boolean parameter with value 1 if teacher $t \in T$ is available to give a lesson on day $d \in D$ at period $h \in H$, 0 otherwise. According to the values of $\gamma_{tdh}$, one can easily detect the last assignable duty period $\nu_{td}$ for teacher $t \in T$ on day $d \in D$. In an assignable period a teacher must be available to teach even if his real activity depends on the timetabling. Moreover, teacher $t \in T$ must or could have in between $\underline{\alpha}_{td}$ and $\overline{\alpha}_{td}$ lessons on day $d \in D$.

The days-off of any teacher $t \in T$ are controlled by parameter $\tau_{ti}$, in which index $i$ takes integer values from 0 to 3. If $i = 0$, teacher $t \in T$ must have a day off on day $\tau_{t0} \in D \cup \{0\}$ (where 0 indicates a day off selected in the model solution); if $i = 1$, $\tau_{ti}$ represents the minimum number of additional days off of teacher $t \in T$ (since one day off must be guaranteed, the number of days off a week is at least $\tau_{t1} + 1$); if $i = 2$, $\tau_{ti}$ is the maximum number of additional days off for teacher $t \in T$ (hence, the number of days off a week is at most $\tau_{t2} + 1$); if $i = 3$, index $\tau_{ti}$ represents the (high) cost of violation of days off. Let $\tilde{D}_t$ be the singleton of the day off for teacher

$t \in T$: $\tilde{D}_t = \{\tau_{t0}\}$.

The following parameters are defined to link teachers and classes. Teacher $t \in T$ must or could have in between $\underline{\rho}_{ct}$ and $\overline{\rho}_{ct}$ lessons with class $c \in C$.

In order to introduce a possible fractional time duration for all classes of a school, consider an integer positive parameter $\eta$, which represents the number of daily periods in a single lesson. For example, if the lesson takes 1 hour and the daily periods of set $H$ represent 30 minutes, $\eta$ takes value 2. Since some lessons cannot have a duration multiple of $\eta$, they need to be removed from the planning of fractional time units. As a result, define the set of incompatible periods $\tilde{N}_\eta = \{n \in N | (n \bmod \eta) \neq 0\}$.

The first decision variable is denoted by $x_{ctdh}$. It takes value 1 if class $c \in C$ is assigned to teacher $t \in T$ on day $d \in D$ at period $h \in H$, 0 otherwise. Note that $x_{ctdh} = 0$ if $d \in D \cap \tilde{D}_t$, $t \in T$, $c \in C$, $h \in H$. Clearly, this is the main decision variable, because its entries with value 1 define the timetable. The following auxiliary variables are also defined:

- $a'_{td}$ is equal to 1 if at least one lesson of teacher $t \in T$ is scheduled on day $d \in D$, 0 otherwise;
- $a''_{ctd}$ is equal to 1 if at least one lesson of teacher $t \in T$ and class $c \in C$ is scheduled on day $d \in D$, 0 otherwise;
- $b_{c't'c''t''dh}$ takes value 1 if teacher $t' \in T$ has a lesson on class $c' \in C$ and teacher $t'' \in T$ has a lesson on class $c'' \in C$ in the same period $h \in H$ of the same day $d \in D$, 0 otherwise;
- $e_{ctfdh}$ is equal to 1 if teachers $t \in T$ and $f \in F$ have a lesson in class $c$ on day $d \in D$ at period $h \in H$, 0 otherwise;
- $m_{nctdh}$ is equal to 1 if a multiple lesson with duration $n \in N$ of teacher $t \in T$ starts at period $h \in H$ of day $d \in D$ in class $c \in C$ , 0 otherwise;
- $s^{min}$ are the minimum idle times for all teachers;
- $s^{max}$ are the maximum idle times for all teachers;
- $u'_{td}$ is the ordinal number of the first activity period of teacher $t \in T$ on day $d \in D$;
- $v'_{td}$ is the ordinal number of the last activity period of teacher $t \in T$ on day $d \in D$;
- $u''_{ctd}$ is the ordinal number of the first activity period of teacher $t \in T$ in class $c \in C$ on day $d \in D$;
- $v''_{ctd}$ is the ordinal number of the last activity period of teacher $t \in T$ in class $c \in C$ on day $d \in D$.

For the sake of clarity, constraints are clustered in types depending on the requirements presented in Section 2.2. The link between constraints and requirements is reported in Table 2.2, which also reports a brief description of the types of constraints. The model will adopt slack variables also for hard constraints in order to make a comparison to the model by [KSS15], where every constraint type could be hard or soft according to the specific instance at hand.

| **Requirements** | **Constraint** | **Description** |
|---|---|---|
| - | $C_0$ | Service constraints (required for the implementation of each requirement) |
| R1,R2,R7 | $C_1$ | Weekly requirement |
| R1 | $C_2$ | Class presence |
| R5,R9 | $C_3$ | Class unavailability |
| R6,R8 | $C_4$ | Teacher unavailability |
| R10 | $C_5$ | Split lessons |
| R3,R4 | $C_6$ | Days off |
| R9 | $C_7$ | Co-teaching |
| R11,R12 | $C_8$ | Block |
| R13 | $C_9$ | Pre-assigned lessons |
| R25 | $C_{10}$ | Equity in idle times |
| R17 | $C_{11}$ | Idle times |
| R18,R20,R21 | $C_{12}$ | Multiple lessons |
| R14 | $C_{13}$ | Horizontal distribution |
| R15 | $C_{14}$ | Vertical distribution |
| R16 | $C_{15}$ | Teacher workload restrictions |
| R23 | $C_{16}$ | Class/teacher workload restrictions |
| R22 | $C_{17}$ | Excessive multiple lessons |
| R19 | $C_{18}$ | Maximum workload |
| R24 | $C_{19}$ | Fractional time unit |

Table 2.2: Grouping of requirements in types of constraints

All constraints are described hereafter.

### $C_0$ - Service constraints.

$$a'_{td} \geq x_{ctdh} \qquad \forall c \in C, \forall t \in T, \forall d \in D, \forall h \in H \quad (2.1)$$

$$\nu_{td} a''_{ctd} \geq \sum_{h \in H} x_{ctdh} \qquad \forall c \in C, \forall t \in T, \forall d \in D \quad (2.2)$$

According to (2.1), any teacher $t \in T$ cannot be assigned to any class $c \in C$ in any period $h \in H$ of day $d \in D$ if he/she is not scheduled on this day. Constraint

(2.2) enforces that any teacher $t \in T$ cannot be assigned to any period in class $c \in C$ on day $d \in D$ if he/she is not scheduled in this class on this day.

$C_1$ **- Weekly requirement (R1, R2, R7, hard).** The sum of all lessons of teacher $t \in T$ in class $c \in C$ cannot differ from those required ($\chi_{ct}$). Since the satisfaction of this hard constraint could not be guaranteed, a non-negative integer variable $s_{ct}^{C_1}$ is introduced. More formally,

$$\sum_{d \in D} \sum_{h \in H} x_{ctdh} - s_{ct}^{C_1} \leq \chi_{ct} \qquad \forall c \in C, \forall t \in T \qquad (2.3)$$

$$\sum_{d \in D} \sum_{h \in H} x_{ctdh} + s_{ct}^{C_1} \geq \chi_{ct} \qquad \forall c \in C, \forall t \in T \qquad (2.4)$$

(2.3) and (2.4) is similar to the analogous constraint introduced in [Sch99].

Before introducing constraint types $C_2$ and $C_3$, it is worth noting that in each class there is at most a teacher $t \in T \setminus F$ and, if there is no teacher, the class cannot attend a lesson. These requirements can be directly enforced by the boolean parameters $\delta_{cdh}$ on class presence and $\beta_{cdh}$ on class availability, in fact

$\delta_{cdh} \leq \sum_{t \in T \setminus F} x_{ctdh} \leq \beta_{cdh} \qquad \forall c \in C, \forall d \in D, \forall h \in H$

However, these constraints are not implemented as reported above, because it is needed to penalize their violation. Therefore, in what follows, the constraints are considered separately and suitable auxiliary variables are introduced.

$C_2$ **- Class presence (R1, hard).** The following constraint enforces that each class must attend lessons in some periods and days of the weekly timetable:

$$\sum_{t \in T \setminus F} x_{ctdh} + s_{cdh}^{C_2} \geq \delta_{cdh} \qquad \forall c \in C, \forall d \in D, \forall h \in H \qquad (2.5)$$

Note that the constraint holds despite the non-negative integer variable, because a teacher could be assigned to a class in a daily period, even if the class does not have to attend a lesson in that period. Clearly, this situation must not be penalized unlike in the converse case.

$C_3$ - **Class unavailability (R5, R9, hard).** The following constraint enforces that a class could attend lessons in some periods and days only if it is available in these periods and days of the weekly timetable:

$$\sum_{t \in T \setminus F} x_{ctdh} - s_{cdh}^{C_3} \leq \beta_{cdh} \qquad \forall c \in C, \forall d \in D, \forall h \in H \qquad (2.6)$$

Note that the constraint holds despite the non-negative integer variable, because it is possible to have the availability of a class in a period of a day, but no teacher is assigned to the class. Clearly, this situation must not be penalized unlike in the converse case.

$C_4$ **- Teacher unavailability (R6, R8, hard).** Excluding the case of articulated classes, teacher $t \in T$ cannot be assigned to more than one class in each period of each day, i.e. $\sum_{c \in C} x_{ctdh} \leq 1$. The (un)availability of teachers is controlled by the boolean parameter $\gamma_{tdh}$ and the assignment of teachers when they are not available must be penalized. Therefore, a boolean variable $s^{C_4}_{tdh}$ is introduced, it takes value 1 of this critical situation occurs, 0 otherwise. Therefore, this constraint can be formulated as follows:

$$\sum_{c\in C} x_{ctdh} - s^{C_4}_{tdh} \leq \gamma_{tdh} a'_{td} \qquad \forall t \in T, \forall d \in D, \forall h \in H \tag{2.7}$$

$C_5$ **- Split lessons (R10, hard/soft).** Multiple lessons of any teacher $t \in T$ in class $c \in C$ must be consecutive on any day $d \in D$ (or without splits). This constraint can be enforced in period $h \in H$ by an upper bound of value $h$ on the period of the first lesson and a lower bound of value $h$ on the period of the last lesson for teacher $t \in T$ in class $c \in C$ on day $d \in D$, if this teacher is on duty in this class on this day. If $x_{ctdh} = 0$, these bounds must not be effective. More formally,

$$u''_{ctd} \leq (|H|+1) - (|H|+1-h) x_{ctdh} \qquad \forall c \in C, \forall t \in T, \forall d \in D, \forall h \in H \tag{2.8}$$

$$v''_{ctd} \geq h x_{ctdh} \qquad \forall c \in C, \forall t \in T, \forall d \in D, \forall h \in H \tag{2.9}$$

However, one must still link the time interval between the first and last teaching period to the number of lessons of a teacher in a day. The boolean variable $s^{C_5}_{ctd}$ is introduced to detect the split lessons of teacher $t \in T$ in class $c \in C$ on day $d \in D$ when it takes value 1, 0 otherwise. More formally,

$$a''_{ctd} + v''_{ctd} - u''_{ctd} \leq \sum_{h\in H} x_{ctdh} + s^{C_5}_{ctd}(|H|-2) \qquad \forall c \in C, \forall t \in T, \forall d \in D \tag{2.10}$$

A minor change in these constraints will be reported later to handle idle times.

$C_6$ **- Days off (R3, R4, hard/soft).** The overall number of days off must be in between the minimum and the maximum values, which are $1 + \tau_{t1}$ and $1 + \tau_{t2}$ for teacher $t \in T$, respectively. A non-negative integer variable $s^{C_6}_t$ is introduced to report how many times these constraints are not satisfied for teacher $t \in T$. Therefore,

$$1 + \tau_{t1} - s^{C_6}_t \leq |D| - \sum_{d\in D} a'_{td} \qquad \forall t \in T \tag{2.11}$$

$$|D| - \sum_{d\in D} a'_{td} \leq 1 + \tau_{t2} + s^{C_6}_t \qquad \forall t \in T \tag{2.12}$$

$C_7$ **- Co-teaching (R9, hard).** Co-teaching cannot be performed either when the class or the teacher or the co-teacher are not available in a daily period.

$$e_{ctfdh} \leq \beta_{cdh} \cdot \gamma_{tdh} \cdot \gamma_{fdh} \cdot x_{ctdh} \qquad \forall c \in C, \forall t \in T \setminus F, \forall f \in F, \forall d \in D, \forall h \in H \tag{2.13}$$

$$e_{ctfdh} \leq \beta_{cdh} \cdot \gamma_{tdh} \cdot \gamma_{fdh} \cdot x_{cfdh} \qquad \forall c \in C, \forall t \in T \setminus F, \forall f \in F, \forall d \in D, \forall h \in H \tag{2.14}$$

Moreover, one must guarantee exactly $\mu_{ctf}$ co-teaching lessons in a week and a possible violation must be taken into account. Therefore, a non-negative integer variable $s^{C_7}_{ctf}$ is introduced, it is an excess or lack of lessons for class $c \in C$ with teacher $t \in T$ and co-teacher $f \in F$:

$$\sum_{d\in D}\sum_{h\in H} e_{ctfdh}+s^{C_7}_{ctf} \geq \mu_{ctf} \qquad \forall c\in C, \forall t\in T\setminus F, \forall f\in F \quad (2.15)$$

$$\sum_{d\in D}\sum_{h\in H} e_{ctfdh}-s^{C_7}_{ctf} \leq \mu_{ctf} \qquad \forall c\in C, \forall t\in T\setminus F, \forall f\in F \quad (2.16)$$

$C_8$ **- Block lessons (R11, R12, hard).** Block lessons cannot be performed either when the first class or the second class or their teachers are not available in a daily period:

$$b_{c't'c''t''dh} \leq \beta_{c'dh}\cdot\beta_{c''dh}\cdot\gamma_{t'dh}\cdot\gamma_{t''dh}\cdot x_{c't'dh} \qquad \forall c', c''\in C, \forall t', t''\in T, \forall d\in D, \forall h\in H \quad (2.17)$$

$$b_{c't'c''t''dh} \leq \beta_{c'dh}\cdot\beta_{c''dh}\cdot\gamma_{t'dh}\cdot\gamma_{t''dh}\cdot x_{c''t''dh} \qquad \forall c', c''\in C, \forall t', t''\in T, \forall d\in D, \forall h\in H \quad (2.18)$$

Moreover, one must guarantee exactly $\phi_{c't'c''t''}$ block lessons in a week and a possible violation must be taken into account. Therefore, a non-negative integer variable $s^{C_8}_{c't'c''t''}$ is introduced, it is an excess or lack of block lessons for classes $c', c'' \in C$ with teachers $t', t'' \in T$:

$$\sum_{d\in D}\sum_{h\in H} b_{c't'c''t''dh}+s^{C_8}_{c't'c''t''} \geq \phi_{c't'c''t''} \qquad \forall c', c''\in C, \forall t', t''\in T \quad (2.19)$$

$$\sum_{d\in D}\sum_{h\in H} b_{c't'c''t''dh}-s^{C_8}_{c't'c''t''} \leq \phi_{c't'c''t''} \qquad \forall c', c''\in C, \forall t', t''\in T \quad (2.20)$$

In the case of articulated classes, one could represent a teacher by an alias (i.e. the pair of teachers $t'$ and $t''$ represent the same person).

$C_9$ **- Preassigned lessons (R13, hard).** The lessons of teacher $t \in T$ in class $c \in C$ have to be scheduled in period $h \in H$ of day $d \in D$ when the boolean parameter $\pi_{ctdh}$ takes value 1. Since lessons could also be scheduled when $\pi_{ctdh}$ is 0, the satisfaction of preassigned lessons can be enforced by

$$x_{ctdh} \geq \pi_{ctdh} \qquad \forall c\in C, \forall t\in T, \forall d\in D, \forall h\in H \quad (2.21)$$

However, the sake of consistency with the other constraints, the former expression is presented by a boolean variable $s^{C_9}_{ctdh}$, which takes value 1 only if the compulsory lesson of class $c \in C$ with teacher $t \in T$ in period $h \in H$ of day $d \in D$ is not scheduled:

$$x_{ctdh}+s^{C_9}_{ctdh} \geq \pi_{ctdh} \qquad \forall c\in C, \forall t\in T, \forall d\in D, \forall h\in H \quad (2.22)$$

$C_{10}$ **- Equity in idle times (R25, hard).** The same number of idle times among all teachers can be pursued by the minimization of the difference between the maximum and the minimum idle times among all teachers. Therefore, one needs to introduce a new non-negative integer variable $s^{C_{10}} = s^{max} - s^{min}$ and minimize its value. Clearly, the number of idle times of each teacher must be in between $s^{min}$ and $s^{max}$ in the weekly planning horizon. More formally,

$$s^{min} + s^{C_{10}} = s^{max} \tag{2.23}$$

$$\sum_{d \in D}(v'_{td} + a'_{td} - u'_{td} - \sum_{c \in C}\sum_{h \in H} x_{ctdh}) \leq s^{max} \qquad \forall t \in T \tag{2.24}$$

$$\sum_{d \in D}(v'_{td} + a'_{td} - u'_{td} - \sum_{c \in C}\sum_{h \in H} x_{ctdh}) \geq s^{min} \qquad \forall t \in T \tag{2.25}$$

$C_{11}$ **- Idle times (R17, soft).** The idle times of teacher $t \in T$ on day $d \in D$ can be derived from the first and the last activity daily period in a way similar to constraints $C_5$ on split lessons. More precisely, for each period daily and teacher a lower bound on the last activity period and an upper bound on the first activity period are computed. Their difference must be limited above by the number of lessons given by teacher $t \in T$ over all classes in a day. In order to guarantee the satisfaction of this constraint, an additional non-negative integer variable $s_{td}^{C_{11}}$ is introduced to report how many times a idle time occurs for teacher $t \in T$ over all periods $h \in H$ of day $d \in D$. Clearly, this variable will be minimized in this formulation. Therefore:

$$u'_{td} \leq (|H|+1) - (|H|+1-h)\sum_{c \in C} x_{ctdh} \qquad \forall t \in T, \forall d \in D, \forall h \in H \tag{2.26}$$

$$v'_{td} \geq h \sum_{c \in C} x_{ctdh} \qquad \forall t \in T, \forall d \in D, \forall h \in H \tag{2.27}$$

$$a'_{td} + v'_{td} - u'_{td} \leq \sum_{c \in C}\sum_{h \in H} x_{ctdh} + s_{td}^{C_{11}} \qquad \forall t \in T, \forall d \in D \tag{2.28}$$

$C_{12}$ **- Multiple lessons (R18, R20, R21, hard/soft).** Consider a multiple lesson of length $n$ starting in period 1 for teacher $t \in T$ in class $c \in C$ on day $d \in D$. As a consequence, in period $n+1$, $x_{ctd(n+1)}$ must take value 0. More formally:

$$\sum_{i=1}^{n} x_{ctdi} + 1 - x_{ctd(n+1)} \leq n + m_{nctd1} \qquad \forall n \in N \setminus \{|H|\}, \forall c \in C, \forall t \in T, \forall d \in D \tag{2.29}$$

Hence, $m_{nctd1}$ must take value 1 when $n$ consecutive lessons are followed by a period with no lesson between teacher $t$ and class $c$.

If the multiple lesson of length $n$ is scheduled in the last periods of a day, the former constraint is modified as follows:

$$1 - x_{ctd(\nu_{td}-n)} + \sum_{i=1}^{n} x_{ctd(\nu_{td}-n+i)} \leq n + m_{nctd(\nu_{td}-n+1)} \quad \forall n \in N \setminus \{|H|\}, \forall c \in C, \forall t \in T, \forall d \in D \tag{2.30}$$

The following constraint introduces for the special case in which multiple lessons span over all periods in a day:

$$\sum_{h\in H} x_{ctdh} \leq n-1+m_{nctd1} \qquad \forall n \in \{|H|\}, \forall c\in C, \forall t\in T, \forall d\in D \quad (2.31)$$

The case of multiple lessons starting after the first period and ending before the last one is still needed to introduce:

$$1-x_{ctd(h-1)}+\sum_{i=1}^{n} x_{ctd(h+i-1)}+1-x_{ctd(h+n)} \leq n+1+m_{nctdh} \; \forall n\in N\setminus\{|H|\}, \forall c\in C, \forall t\in T, \forall d\in D, \forall h\in\{2..(\nu_{td}-n)\} \quad (2.32)$$

Sometimes the minimum ($\underline{\zeta}_{ct}$) and the maximum ($\overline{\zeta}_{ct}$) number of multiple lessons of predefined length (ranging between $\underline{\epsilon}_{ct}$ and $\overline{\epsilon}_{ct}$) must be considered for some teachers in some classes. The weekly number of multiple lessons can be computed by variable $m_{nctdh}$, but an additional non-negative integer variable $s_{ct}^{C_{12}}$ must be adopted to compute the number of violations for teacher $t \in T$ in class $c \in C$:

$$\sum_{n=\underline{\epsilon}_{ct}}^{\overline{\epsilon}_{ct}} \sum_{d\in D} \sum_{h=1}^{|H|+1-n} m_{nctdh} + s_{ct}^{C_{12}} \geq \underline{\zeta}_{ct} \qquad \forall c\in C, \forall t\in T \quad (2.33)$$

$$\sum_{n=\underline{\epsilon}_{ct}}^{\overline{\epsilon}_{ct}} \sum_{d\in D} \sum_{h=1}^{|H|+1-n} m_{nctdh} - s_{ct}^{C_{12}} \leq \overline{\zeta}_{ct} \qquad \forall c\in C, \forall t\in T \quad (2.34)$$

A congruence check of $m_{nctdh}$ is needed: the sum of all lessons (multiple or single) must be equal to the week requirement

$$\sum_{n\in N} \sum_{d\in D} \sum_{h=1}^{|H|+1-n} (n \cdot m_{nctdh}) = \chi_{ct} \qquad \forall c\in C, \forall t\in T \quad (2.35)$$

$C_{13}$ **- Horizontal distribution (R14, soft).** The lessons of a teacher in a class should not be clustered either in the first part or in the second part of a week. If the weekly number of lessons $\chi_{ct}$ of teacher $t \in T$ in class $c \in C$ is even, it is enforced to have the same number of lessons in the two parts of the week; if $\chi_{ct}$ is odd, their difference should be one. Since this ideal balance could not be guaranteed in both previous cases, a non-negative integer variable $s_{ct}^{C_{13}}$ is introduced to report the difference between the periods of teacher $t \in T$ in class $c \in C$ in the two parts of the week. More formally:

$$\sum_{d=1}^{\lfloor |D|/2 \rfloor} \sum_{h\in H} x_{ctdh} - \sum_{d=\lceil |D|/2 \rceil+1}^{|D|} \sum_{h\in H} x_{ctdh} - s_{ct}^{C_{13}} \leq \lceil \frac{\chi_{ct}}{2} \rceil - \lfloor \frac{\chi_{ct}}{2} \rfloor \qquad \forall c \in C, \forall t \in T \quad (2.36)$$

$$\sum_{d=\lceil |D|/2 \rceil+1}^{|D|} \sum_{h\in H} x_{ctdh} - \sum_{d=1}^{\lfloor |D|/2 \rfloor} \sum_{h\in H} x_{ctdh} - s_{ct}^{C_{13}} \leq \lceil \frac{\chi_{ct}}{2} \rceil - \lfloor \frac{\chi_{ct}}{2} \rfloor \qquad \forall c \in C, \forall t \in T \quad (2.37)$$

Note that both of these constraints hold when $|D|$ is even or odd.

$C_{14}$ **- Vertical distribution (R15, soft).** The number of lessons of teacher $t \in T$ in class $c \in C$ should not be clustered in a specific period $h \in H$ over all days of the weekly planning horizon. Therefore, it is enforced an upper bound $\lceil \frac{\chi_{ct}}{|H|} \rceil$ and a lower bound $\lfloor \frac{\chi_{ct}}{|H|} \rfloor$ on the number of lessons scheduled for any teacher in any class in a given period. Since these bounds could be violated, a non-negative integer variable $s_{cth}^{C_{14}}$ is defined to report how many times they are not met for teacher $t \in T$ in class $c \in C$ in period $h \in H$. Therefore,

$$\sum_{d\in D} x_{ctdh} - s_{cth}^{C_{14}} \leq \lceil \frac{\chi_{ct}}{|H|} \rceil \qquad \forall c \in C, \forall t \in T, \forall h \in H \quad (2.38)$$

$$\sum_{d\in D} x_{ctdh} + s_{cth}^{C_{14}} \geq \lfloor \frac{\chi_{ct}}{|H|} \rfloor \qquad \forall c \in C, \forall t \in T, \forall h \in H \quad (2.39)$$

Due to (2.38) and (2.39) the lessons must have a balanced distribution over all daily periods.

$C_{15}$ **- Teacher workload restrictions (R16, hard/soft).** The number of activity periods of each teacher $t \in T$ in any day $d \in D$ must be in between the lower bound $\underline{\alpha}_{td}$ and the upper bound $\overline{\alpha}_{td}$, if at least a lesson is scheduled for teacher $t \in T$ on day $d \in D$ (this is checked by the values of variable $a'_{td}$). Since this situation should not occur, the non-negative integer variable $s_{td}^{C_{15}}$ is defined to report how many times the violation occurs.

$$\sum_{c\in C} \sum_{h\in H} x_{ctdh} - \eta s_{td}^{C_{15}} \leq a'_{td} \overline{\alpha}_{td} \qquad \forall t \in T, d \in D \quad (2.40)$$

$$\sum_{c\in C} \sum_{h\in H} x_{ctdh} + \eta s_{td}^{C_{15}} \geq a'_{td} \underline{\alpha}_{td} \qquad \forall t \in T, \forall d \in D \quad (2.41)$$

Note that the parameter $\eta$ accounts for the possible use of fractional periods.

$C_{16}$ **- Class/teacher workload restrictions (R23, hard/soft).** The number of daily activity periods of each teacher $t \in T$ with class $c \in C$ must be in between the lower bound $\underline{\rho}_{ct}$ and the upper bound $\overline{\rho}_{ct}$, if at least a lesson is scheduled for teacher $t \in T$ with class $c \in C$ on day $d \in D$ (this is checked by the values of variable

$a''_{ctd}$). Since this situation should not occur, the non-negative integer variable $s^{C_{16}}_{ctd}$ is defined to report how many times the violation occurs.

$$\sum_{h\in H} x_{ctdh} - \eta s^{C_{16}}_{ctd} \leq a''_{ctd}\overline{\rho}_{ct} \qquad \forall c\in C, \forall t\in T, \forall d\in D \quad (2.42)$$

$$\sum_{h\in H} x_{ctdh} + \eta s^{C_{16}}_{ctd} \geq a''_{ctd}\underline{\rho}_{ct} \qquad \forall c\in C, \forall t\in T, \forall d\in D \quad (2.43)$$

Note that the parameter $\eta$ accounts for the possible use of fractional periods.
$C_{17}$ **- Excessive multiple lessons (R22, hard/soft).** The number of periods with multiple daily lessons for class $c \in C$ on day $d \in D$ cannot be larger that a threshold value, which can be reasonably set to $\lceil\frac{|H|-1}{2}\rceil$ (e.g. half of the periods in a day, when $|H|$ is even). Since this situation should not occur, the non-negative integer variable $s^{C_{17}}_{cd}$ is defined to report how often it occurs.

$$\sum_{n\in N}\sum_{t\in T\setminus F}\sum_{h=1}^{|H|-(n-1)} n\cdot m_{nctdh} - s^{C_{17}}_{cd} \leq \lceil\frac{|H|-1}{2}\rceil \qquad \forall c \in C, \forall d \in D \quad (2.44)$$

Note that $n$ allows to consider duration of multiple lessons, which cannot take the trivial length of one period. Note that co-teachers are not involved in (2.44) because they always work with other teachers.

Yet, multiple lessons with duration $n$ cannot start after period $|H| - (n-1)$.
$C_{18}$ **- Teacher maximum workload (R19, hard/soft).** The workload of teacher $t \in T$ in any two consecutive days $d \in D$ and $(d+1) \in D$ must take value one period less than the sum of maximum workload in these days (i.e. $\overline{\alpha}_{td} + \overline{\alpha}_{t(d+1)}$). The non-negative integer variable $s^{C_{18}}_{td}$ is introduced to quantify the violation of this requirement.

$$\sum_{c\in C}\sum_{h\in H}(x_{ctdh} + x_{ct(d+1)h}) - s^{C_{18}}_{td} \leq \overline{\alpha}_{td} + \overline{\alpha}_{t(d+1)} - 1 \qquad \forall t\in T, \forall d\in D\setminus\{|D|\} \quad (2.45)$$

Although the implementation of this constraint is possible in the standard [XHSTT] format, in common real cases it should be replaced by several hundreds of constraints *Limit Busy Times*, when teachers and days are 100 and 6 respectively.
$C_{19}$ **- Fractional time unit (R24, hard).** When fractional time units are possible, the duration of lessons must be multiple of parameter $\eta$. Since this condition may not always be met, a non-negative integer variable $s^{C_{19}}_{nct}$ is defined to report how often it is not satisfied.

$$\sum_{d\in D}\sum_{h=1}^{|H|+1-n} m_{nctdh} - s^{C_{19}}_{nct} = 0 \qquad \forall n\in \tilde{N}_{\eta}, \forall c\in C, \forall t\in T \quad (2.46)$$

The objective function is a linear combination of the violation of all types of constraints. Let $o_i$ the overall violation of $i$-th constraint type and $\omega_i$ its weight. More formally, the violation of each constraint type is reported below:

$$o_1 = \sum_{c\in C}\sum_{t\in T} s^{C_1}_{ct} \quad o_2 = \sum_{c\in C}\sum_{d\in D}\sum_{h\in H} s^{C_2}_{cdh} \quad o_3 = \sum_{c\in C}\sum_{d\in D}\sum_{h\in H} s^{C_3}_{cdh} \quad o_4 = \sum_{t\in T}\sum_{d\in D}\sum_{h\in H} s^{C_4}_{tdh}$$

$$o_5 = \sum_{c\in C}\sum_{t\in T}\sum_{d\in D} s^{C_5}_{ctd} \quad o_6 = \sum_{t\in T} \tau_{t3} s^{C_6}_{t} \quad o_7 = \sum_{c\in C}\sum_{d\in D}\sum_{h\in H} s^{C_7}_{cdh}$$

$$o_8 = \sum_{c\in C}\sum_{d\in D}\sum_{h\in H} s^{C_8}_{cdh} \quad o_9 = \sum_{c\in C}\sum_{t\in T}\sum_{d\in D}\sum_{h\in H} s^{C_9}_{ctdh} \quad o_{10} = s^{C_{10}} \quad o_{11} = \sum_{t\in T}\sum_{d\in D} s^{C_{11}}_{td}$$

$$o_{12} = \sum_{c\in C}\sum_{t\in T} \theta_{ct} s^{C_{12}}_{ct} \quad o_{13} = \sum_{c\in C}\sum_{t\in T} s^{C_{13}}_{ct} \quad o_{14} = \sum_{c\in C}\sum_{t\in T}\sum_{h\in H} s^{C_{14}}_{cth} \quad o_{15} = \sum_{t\in T}\sum_{d\in D} s^{C_{15}}_{td}$$

$$o_{16} = \sum_{c\in C}\sum_{t\in T}\sum_{d\in D} s^{C_{16}}_{ctd} \quad o_{17} = \sum_{c\in C}\sum_{d\in D} s^{C_{17}}_{cd} \quad o_{18} = \sum_{t\in T}\sum_{d\in D\setminus\{|D|\}} s^{C_{18}}_{td} \quad o_{19} = \sum_{n\in N}\sum_{c\in C}\sum_{t\in T} s^{C_{19}}_{nct}$$

Hence, the objective function of IHSTT is:

$$f = \sum_{i=1}^{19} \omega_i o_i \tag{2.47}$$

The complete MIP model consists in minimizing $f$, subject to constraints (2.1)-(2.46). This model is very complex to solve as it is and, according to preliminary experiments, no significant gain is obtained by the removal of slack variables for hard constraints. Moreover, these variables help guarantee a full compatibility with [KSS15], where every constraint type could be hard or soft according to the specific instance at hand. A two-step method is proposed in the following section to solve the problem.

## 2.5 A two-step method for the [IHSTP]

Since the model for [IHSTP] is expected to be very hard to solve, a two-step method to determine high-quality solutions within a reasonable time interval is presented. The method is motivated by many possibilities for selecting the activity periods of each teacher, who gives lessons in a class for a limited number of periods w.r.t the the overall number of periods spent by students in the same class (e.g. a teachers must stay in a class for 4 hours a week and the same class attends lessons for 32 hours a week). Therefore, the [IHSTP] would be simplified if one *a priori* knows the schedule of teachers without details on the classes taught in each period.

Therefore, the proposed method decomposes [IHSTP] into two problems:

- The first problem assigns teaching periods to teachers to determine the so-called teacher profile. This problem is called Teacher Profile Problem [TPP].
- The second problem assign classes to teachers according to the solution of the [TPP] and results in a simplified version of the [IHSTP], which is called restricted [IHSTP] and denoted by [RIHSTP].

The details about the mathematical formulations of these problems are provided in Subsection 2.5.1 and Subsection 2.5.2. Figure 2.1 shows the connection between [TPP] and [RIHSTP].

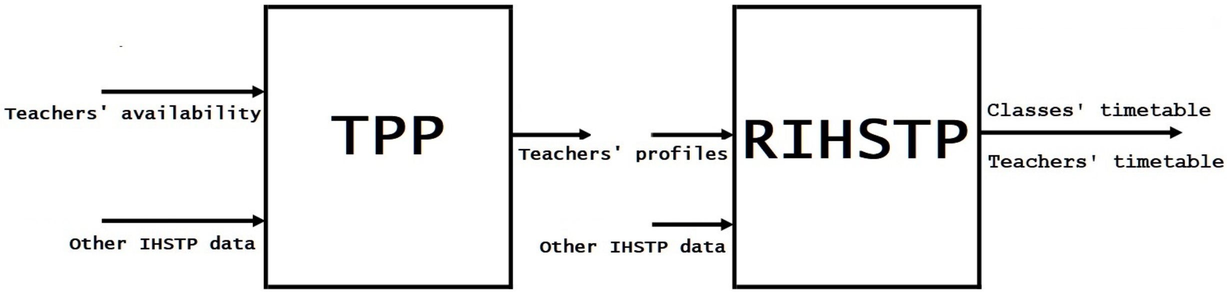



Figure 2.1: The connection between [TPP] and [RIHSTP]

### 2.5.1 The teacher profile problem [TPP]

#### Problem description

Relevant data for the [TPP] are the periods in which teachers are available and lessons are given for each class in each day. Classes may not spend the same number of periods at school, because usually the number of school days in a week is not an exact divider of the overall week requirements (e.g. 33 hours over 6 days from Monday to Saturday). Schools have two choices to face this situation: fixing in advance which days have extra periods or letting this decision to the optimization phase. In the first case, parameters $\beta_{cdh}$ and $\delta_{cdh}$ take the same value for all the weekly periods; in the second case they differ when the extra daily periods occur. Generally speaking, Italian schools prefer the first choice, because it results in a greater management control. Moreover, it would not be possible to determine the work shifts of the teachers if the attendance periods of classes at school are not known. Since the teacher profile is determined before the final timetable, in what follows the values of $\beta_{cdh}$ and $\delta_{cdh}$ are supposed to be identical.

[TPP] is aimed to obtain a subset of the profiles for each teacher who is not a co-teacher (or teacher profile), while taking into account some requirements of the [IHSTP], but their determination must be computationally viable. Clearly, the periods in a (non-co-)teacher profile must be consecutive in a day, in order to a priori minimize idle times. In the [TPP] daily profiles in which all teachers either start in the first period or end in the last period are considered. This assumption decreases the number of possible profiles and is also motivated by equity issues. In fact, teachers starting in the second period have an edge over those starting in the first one, because they wake up later and come across less congested roads in their trips. Similarly, teachers ending in the last hour are more tired than those ending before and can go home later. Therefore, two possible shifts are considered: the first shift starts in the first period, the last shift ends in the last period. Note that the profiles of co-teachers are not determined in the [TPP], because they may end up working with teachers with different profiles and it may be impossible to satisfy all the requirements at the same time.

Figure 2.2 shows an example on the construction of a profile. The teacher has a day off on Wednesday and is available to teach from period 1 to period 6 in the other days. Assume to select in the first shift 3 periods on Monday, 4 periods on Tuesday and 3 periods on Saturday (a). In the last shift assume to select 2 periods on Monday, 3 periods on Thursday and 3 periods on Friday (b). The shifts can be merged and result in the final teacher profile (c). Although the profile of Monday has one idle period, it is acceptable owing to the relevant workload in this day. Note that this profile also satisfies the horizontal and vertical distribution, as defined in requirements R15 and R16.

In what follows, all requirements of the profiles (or shifts) are enumerated.

1. **R26** (Shift selection). For non-co-teachers, the first and/or the second shift could be selected in a day.

2. **R27** (Duration of shifts). The length of shifts cannot be larger than the daily availability and teachers cannot be on duty in days which are not selected.

3. **R28** (Allocation of periods to shifts): Non-co-teachers must be on duty for all periods in a shift, if it is selected.

4. **R29** (Teacher profile definition). A period is part of a teacher profile if and only if the first shift or the second shift are selected.

5. **R30** (Profile consistency). Teachers cannot be assigned to profiles with periods in which they are not available. Moreover, the daily profiles cannot be selected in days off.

6. **R31** (Class surveillance). The profile of teachers must guarantee that each class is monitored by one of its non-co-teachers in each daily period.

7. **R32** (Alternated shifts). The profiles of teachers should encourage the alternation between the first and the last shift between any pair of consecutive days to incentivize a good vertical distribution possibly.

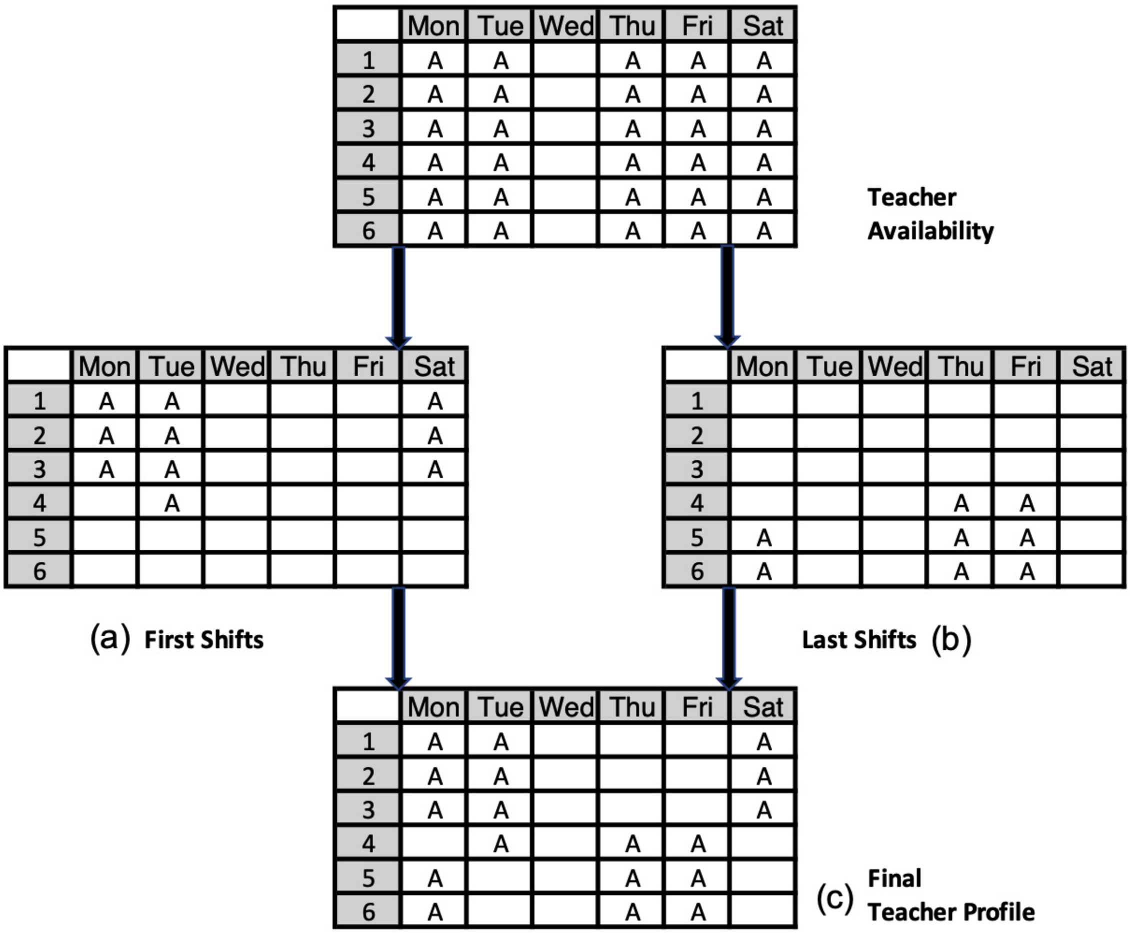



Figure 2.2: An example of optimized Teacher Profile

Finally, it is needed to restate a number of requirements of the [IHSTP] in terms of teacher profiles. More precisely, these requirements concern day off (**R33**, soft), additional days off (**R34**, soft), pre-assigned lessons (**R35**, hard), horizontal distribution (**R36**, hard), vertical distribution (**R37**, hard), block (**R38**, hard), fractional time unit (**R39**, hard), teacher workload restrictions (**R40**, soft), idle times (**R41**, soft), equity in idle times (**R42**, soft).

### Optimization model

The [TPP] formulation is based on the notation already presented for the [IHSTP]. However, some additional notation needs to be introduced. Let $\psi^c$ be a $|T\backslash F|$-column-vector, in which $\psi^c_t$ takes value 1 if teacher $t \in T\backslash F$ teaches in class $c \in C$, and 0 otherwise. Let $\psi^{c\top}$ be the transpose of $\psi^c$. Moreover, let $L_c$ be the set of classes with some teachers in common with class $c \in C$, including class $c$ itself.

It is possible to compute $L_c$ from $\psi^c$ as follows:
$L_c = \{c' \in C | \psi^{c\top}\psi^{c'} > 0\}$

Let $y_{tdh}$ be a decision variable, which takes value 1 if the teacher $t \in T$ is on duty in day $d \in D$ at period $h \in H$, 0 otherwise. Clearly, $y_{tdh}$ is the main decision variable of the [TPP], because all entries with value 1 represent the profile of teacher $t \in T$. The following variables of the IHSTT are also used with the same meaning in [TPP] model: $a'_{td}, b_{c't'c''t''dh}, u'_{td}, v'_{td}, s^{min}, s^{max}$. In addition, the following auxiliary variables are defined:

$f_{tdh}$ 1 if $h \in H$ is the last period of the first shift of teacher $t \in T$ on day $d \in D$, 0 otherwise;
$l_{tdh}$ 1 if $h \in H$ is the first period in last shift of teacher $t \in T$ on day $d \in D$, 0 otherwise;
$n'_{td}$ length of the first shift of teacher $t \in T$ on day $d \in D$;
$n''_{td}$ length of the last shift of teacher $t \in T$ on day $d \in D$;
$\tilde{m}_{ntdh}$ is equal to 1 if a block of duration $n \in N$ of teacher $t \in T$ starts at period $h \in H$ of day $d \in D$ in one of the shifts, 0 otherwise.

$HC_1$ - **Shift selection (R26, hard)**. The following constraint states that the first shift could be selected for any teacher in each day:

$$\sum_{h=1}^{\nu_{td}} \gamma_{tdh} f_{tdh} \leq 1 \qquad \forall t \in T \setminus F, \forall d \in D \quad (2.48)$$

Note that the first shift includes all periods between the first one and time slot such that $f_{tdh}$ has value 1. A similar constraint is formulated for the last shift:

$$\sum_{h=2}^{\nu_{td}} \gamma_{tdh} l_{tdh} \leq 1 \qquad \forall t \in T \setminus F, \forall d \in D \quad (2.49)$$

Clearly, the last shift includes all periods between the time slot for which $l_{tdh}$ has value 1 and the last one.

$HC_2$ - **Duration of shifts (R27, hard)**. The following constraints determine the duration of shifts for each teacher in each day from the values of variables $f_{tdh}$ and $l_{tdh}$ :

$$n'_{td} = \sum_{h \in H} h \gamma_{tdh} f_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D \quad (2.50)$$

$$n''_{td} = \sum_{h \in H} (\nu_{td} + 1 - h) \gamma_{tdh} l_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D \quad (2.51)$$

Moreover, (2.52) ensures that $a'_{td}$ takes value 1 when teacher $t \in T$ works on day $d \in D$ and, in this case, the duration of duty shifts are bounded by suitable values.

$$n'_{td} + n''_{td} \leq a'_{td} \sum_{h \in H} \gamma_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D \quad (2.52)$$

In a workday at least one lesson has to be given by a teacher:

$$n'_{td} + n''_{td} \geq a'_{td} \qquad \forall t \in T \setminus F, \forall d \in D \setminus \tilde{D}_t \quad (2.53)$$

$HC_3$ - **Allocation of periods to shifts (R28, hard)**. The following constraints link variable $y_{tdj}$ to $f_{tdh}$ and $l_{tdh}$:

$$y_{tdj} \geq f_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D, \forall h \in H, j \in \{1, .., h\} \quad (2.54)$$

$$y_{tdj} \geq l_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D, \forall h \in H, j \in \{h, .., \nu_{td}\} \quad (2.55)$$

$HC_4$ - **Teacher profile definition (R29, hard)**.
The values of $y_{tdh}$ are computed in the following constraint:

$$y_{tdh} = \sum_{i=h}^{\nu_{td}} f_{tdi} + \sum_{i=2}^{h} l_{tdi} \qquad \forall t \in T \setminus F, \forall d \in D, \forall h \in H \quad (2.56)$$

Note that three cases may occur: period $h \in H$ does not belong to any shift, or it is part of the first shift or part of the second shift.

$HC_5$ - **Profile consistency (R30, hard)**. Teachers cannot be assigned to profiles with periods in which they are not available. Moreover, profiles cannot be assigned to days which are not selected to give lessons:

$$y_{tdh} \leq a'_{td} \gamma_{tdh} \qquad \forall t \in T \setminus F, \forall d \in D, \forall h \in H \quad (2.57)$$

$HC_6$ - **Class surveillance (R31, hard)**. Since no class should be left unattended, for each daily period the number of classes must be equal to the number of teachers:

$$\sum_{t \in T \setminus F} y_{tdh} = \sum_{c \in C} \delta_{cdh} \qquad \forall d \in D, \forall h \in H \quad (2.58)$$

The previous constraint does not guarantee that each class is attended by one of its teachers (for the sake of simplicity, co-teachers are not considered). This is possible only if the sets of classes and teachers represent one partition or can be decomposed in several partitions (i.e when the subset of teachers gives lessons only in a subset of classes in a partition and vice versa). Therefore, it is needed to recall the definition of $L_c$ from the values of $\psi_t^c$, to report the partition associated with class $c \in C$. If $L_c \equiv C$, there is one partition, else there are at least two partitions. Therefore, for each period of any day and class, a balance must be guaranteed between the number of teachers of the class and the number of classes in the same partition, provided that the class is available:

$$\sum_{t \in T \setminus F : \psi_t^c = 1} y_{tdh} = \delta_{cdh} |L_c|$$

However, some classes may not be available in the same daily periods. As a result, the former formula is modified as follows:

$$\sum_{t \in T \setminus F : \psi_t^c = 1} y_{tdh} = \sum_{c' \in L_c} \delta_{c'dh} \qquad \forall c \in C, \forall d \in D, \forall h \in H \quad (2.59)$$

$HC_7$ - **Day off selection (R33, hard)**. Day off must be guaranteed for each teacher:

$$a'_{td} = 0 \qquad \forall t \in T \backslash F, \forall d \in D \cap \tilde{D}_t \quad (2.60)$$

$HC_8$ - **Preassigned lessons (R35, hard)**. Preassigned lessons must be scheduled

$$y_{tdh} \geq \pi_{ctdh} \qquad \forall c \in C, \forall t \in T \backslash F, \forall d \in D, \forall h \in H \quad (2.61)$$

(2.61) is very similar to (2.21) in constraint $C_9$ of the IHSTT.

$HC_9$ - **Horizontal distribution (R36, hard)**. Unlike in the [IHSTP], in the [TPP] the horizontal distribution of lessons is enforced on the overall activity of each teacher without paying attention to classes:

$$\sum_{d=1}^{|D|/2} \sum_{h \in H} y_{tdh} + \lceil \frac{\sum_{c \in C} \chi_{ct}}{2} \rceil - \lfloor \frac{\sum_{c \in C} \chi_{ct}}{2} \rfloor \geq \sum_{d=|D|/2+1}^{|D|} \sum_{h \in H} y_{tdh} \qquad \forall t \in T \backslash F \quad (2.62)$$

$$\sum_{d=1}^{|D|/2} \sum_{h \in H} y_{tdh} - \lceil \frac{\sum_{c \in C} \chi_{ct}}{2} \rceil + \lfloor \frac{\sum_{c \in C} \chi_{ct}}{2} \rfloor \leq \sum_{d=|D|/2+1}^{|D|} \sum_{h \in H} y_{tdh} \qquad \forall t \in T \backslash F \quad (2.63)$$

Note that (2.62)-(2.63) are similar to (2.36)-(2.37) in $C_{13}$ of the IHSTT.

$HC_{10}$ - **Vertical distribution (R37, hard)**. The same logic holds for the vertical distribution:

$$\sum_{d \in D} y_{tdh} \leq \lceil \frac{\sum_{c \in C} \chi_{ct}}{|H|} \rceil \qquad \forall t \in T \backslash F, \forall h \in H \quad (2.64)$$

$$\sum_{d \in D} y_{tdh} \geq \lfloor \frac{\sum_{c \in C} \chi_{ct}}{|H|} \rfloor \qquad \forall t \in T \backslash F, \forall h \in H \quad (2.65)$$

Clearly, (2.64)-(2.65) are similar to (2.38)-(2.39) in $C_{13}$ of the IHSTT.
$HC_{11}$ - **Block (R38, hard)**. Constrains on block lessons are enforced.

$$b_{c't'c''t''dh} \leq \beta_{c'dh} \cdot \beta_{c''dh} \cdot \gamma_{t'dh} \cdot \gamma_{t''dh} \cdot y_{t'dh} \quad \forall c', c'' \in C, \forall t', t'' \in T \backslash F, \forall d \in D, \forall h \in H \quad (2.66)$$

$$b_{c't'c''t''dh} \leq \beta_{c'dh} \cdot \beta_{c''dh} \cdot \gamma_{t'dh} \cdot \gamma_{t''dh} \cdot y_{t''dh} \quad \forall c', c'' \in C, \forall t', t'' \in T \backslash F, \forall d \in D, \forall h \in H \quad (2.67)$$

$$\sum_{d \in D} \sum_{h \in H} b_{c't'c''t''dh} = \phi_{c't'c''t''dh} \qquad \forall c', c'' \in C, \forall t', t'' \in T \backslash F \quad (2.68)$$

Note that (2.66)-(2.68) exhibit minor changes w.r.t. (2.17)-(2.20) in constraint $C_8$ of the IHSTT.

$HC_{12}$ - **Fractional time unit (R39, hard)**. The duration of both shifts of every teacher must be a multiple quantity of the fractional time unit $\eta$.

$$\sum_{h \in H} y_{tdh} \leq n - 1 + \tilde{m}_{ntd1} \qquad \forall n \in \{|H|\}, \forall t \in T \backslash F, \forall d \in D \quad (2.69)$$

$$\sum_{i=1}^{n} y_{tdi}+(1-y_{td(n+1)}) \leq n+\tilde{m}_{ntd1} \qquad \forall n\in N, \forall t\in T\backslash F, \forall d\in D \quad (2.70)$$

$$1-y_{td(h-1)}+\sum_{i=1}^{n} y_{td(h+i-1)}+1-y_{td(h+n)} \leq n+1+\tilde{m}_{ntdh} \quad \forall n\in N, \forall t\in T\backslash F, \forall d\in D, \forall h\in\{2,..,(\nu_{td}-n)\} \quad (2.71)$$

$$1-y_{td(\nu_{td}-n)}+\sum_{i=1}^{n} y_{td(\nu_{td}-n+i)} \leq n+\tilde{m}_{ntd(\nu_{td}-n+1)} \qquad \forall n\in N, \forall t\in T\backslash F, \forall d\in D \quad (2.72)$$

$$\sum_{d\in D}\sum_{h=1}^{\nu_{td}+1-n} \tilde{m}_{ntdh} = 0 \qquad \forall n\in \tilde{N}_{\eta}, \forall t\in T\backslash F \quad (2.73)$$

Moreover, (2.69)-(2.72) compute the length of every shift, while (2.73) guarantees that each shift must have a length multiple of $\eta$. Clearly, these constraints can be skipped if $\eta = 1$.

$SC_1$ - **Alternated shifts (R32, soft)**. The first shift and last shift are recommended to be alternate in consecutive days.

$$1-s_{tdh}^{SC_1} \leq y_{tdh}+y_{t(d+1)h} \qquad \forall t\in T\backslash F, \forall d\in D\backslash\{|D|\}, \forall h\in\{1,\nu_{td}\} \quad (2.74)$$

$$1-s_{tdh}^{SC_1} \geq y_{tdh}-y_{t(d+1)h} \qquad \forall t\in T\backslash F, \forall d\in D\backslash\{|D|\}, \forall h\in\{1,\nu_{td}\} \quad (2.75)$$

$$1-s_{tdh}^{SC_1} \geq y_{t(d+1)h}-y_{tdh} \qquad \forall t\in T\backslash F, \forall d\in D\backslash\{|D|\}, \forall h\in\{1,\nu_{td}\} \quad (2.76)$$

$$1-s_{tdh}^{SC_1} \leq 2-y_{tdh}-y_{t(d+1)h} \qquad \forall t\in T\backslash F, \forall d\in D\backslash\{|D|\}, \forall h\in\{1,\nu_{td}\} \quad (2.77)$$

$SC_2$ - **Days off placement (R34, soft)**. It is recommended to provide additional days off to each teacher:

$$\sum_{d\in D} a'_{td}+1+\tau_{t1}-s_t^{SC_2} \leq |D| \qquad \forall t\in T\backslash F \quad (2.78)$$

$$\sum_{d\in D} a'_{td}+1+\tau_{t2}+s_t^{SC_2} \geq |D| \qquad \forall t\in T\backslash F \quad (2.79)$$

Note that these constraints enforce the assignment of days off, when they were not indicated by teachers.

$SC_3$ - **Teacher workload restrictions (R40, soft)**. The following constraints play the same role of those in $C_{15}$, where $\sum_{c\in C} x_{ctdh}$ is replaced by $y_{tdh}$:

$$\sum_{h\in H} y_{tdh}+\eta s_{td}^{SC_3} \geq a'_{td}\underline{\alpha}_{td} \qquad \forall t\in T\backslash F, \forall d\in D \quad (2.80)$$

$$\sum_{h\in H} y_{tdh} \le a'_{td}\overline{\alpha}_{td} + \eta s_{td}^{SC_3} \qquad \forall t \in T\setminus F, d \in D \quad (2.81)$$

$SC_4$ - **Idle times (R41, soft)**. The following constraints play the same role of those in $C_{11}$, where $\sum_{c\in C} x_{ctdh}$ is replaced by $y_{tdh}$:

$$u'_{td} \le (\nu_{td}+1) - (\nu_{td}+1-h)y_{tdh} \qquad \forall t \in T\setminus F, \forall d \in D, \forall h \in H \quad (2.82)$$

$$v'_{td} \ge h \cdot y_{tdh} \qquad \forall t \in T\setminus F, \forall d \in D, \forall h \in H \quad (2.83)$$

$$a'_{td} + v'_{td} - u'_{td} \le \sum_{h\in H} y_{tdh} + s_{td}^{SC_4} \qquad \forall t \in T\setminus F, \forall d \in D \quad (2.84)$$

$SC_5$ - **Equity in idle times (R42, soft)**. It is recommended for the teachers to have the same minimum idle times:

$$\sum_{d\in D} s_{td}^{SC_4} \le s^{max} \qquad \forall t \in T\setminus F \quad (2.85)$$

$$\sum_{d\in D} s_{td}^{SC_4} \ge s^{min} \qquad \forall t \in T\setminus F \quad (2.86)$$

$$s^{min} + s^{SC_5} \ge s^{max} \quad (2.87)$$

**Teacher Profile Problem objective function**

The objective function of the [TPP] is the sum of all constraint deviation multiplied by a proper weight:

$$f = \omega_1 \sum_{t\in T\setminus F} \sum_{d\in D} \sum_{h\in\{1,\nu_{td}\}} s_{tdh}^{SC_1} + \omega_2 \sum_{t\in T\setminus F} s_t^{SC_2} + \omega_3 \sum_{t\in T\setminus F} \sum_{d\in D} s_{td}^{SC_3} + \omega_4 \sum_{t\in T\setminus F} \sum_{d\in D} s_{td}^{SC_4} + \omega_5 s^{SC_5} \quad (2.88)$$

The complete formulation of the [TPP] consists in minimizing $f$, subject to constraints (2.48)-(2.87)

## 2.5.2 The restricted [IHSTP]

This problem is obtained by replacing $\gamma_{tdh}$ with values of $y_{tdh}$ in IHSTT, as determined in the solution of the [TPP]. This substitution occurs in constraints (2.7), (2.13)-(2.14), (2.17)-(2.18). All in all, the two-step method is supposed to be effective owing to the larger number of null entries of $y_{tdh}$ as opposed to $\gamma_{tdh}$. The real effectiveness of the method will be evaluated in the following experimentation.

Figure 2.3 shows how the solution of the [TPP] can be adopted to obtain a possible solution of the restricted [IHSTP] for teacher A, who must give lessons in classes denoted by 3C, 4C and 5C. For example, according to the TPP, teacher A must give lessons on Monday from period 1 to period 3 and from period 5 to period 6. The restricted [IHSTP] assigns the selected work periods of teacher A to each class. In Figure 2.3, teacher A is assigned to class 3C from period 1 to period 2, class 5C in period 3, class 4C from period 5 to period 6.

| | Mon | Tue | Wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|
| 1 | A | A | | | | A |
| 2 | A | A | | | | A |
| 3 | A | A | | | | A |
| 4 | | A | | A | A | |
| 5 | A | | | A | A | |
| 6 | A | | | A | A | |

| | Mon | Tue | Wed | Thu | Fri | Sat |
|---|---|---|---|---|---|---|
| 1 | 3C | 5C | | | | 4C |
| 2 | 3C | 5C | | | | 4C |
| 3 | 5C | 4C | | | | 3C |
| 4 | | 3C | | 5C | 4C | |
| 5 | 4C | | | 5C | 3C | |
| 6 | 4C | | | 3C | 5C | |

Figure 2.3: A teacher timetable obtained as the solution of [RIHSTP] program from the Teacher Profile

## 2.6 Experimentation

### 2.6.1 Experimental settings and results

This experimentation has several objectives. First, it is aimed to show to what extent the IHSTT can be solved to tackle specific and realistic instances of Italian high schools, where requirements R19, R22, R24 and R25 are taken into account. Second, it is wanted to assess how much the two-step method is effective both in terms of running time and objective function, and how long the computation takes in each step. Third, it is faced a simplified setting without R19, R22, R24 and R25, and compare the results of a MIP solver to solve the IHSTT, the KSS model ([KSS15], [FSCS17]), as well as the outcomes from the KHE heuristic[1]. Fourth, in the simplified setting, it is extended and evaluated the teacher profile phase to the KSS model and the KHE heuristic, to have a deeper understanding on the effectiveness of the two-step method. Fifth, it is performed a "stress test" on the viability of the IHSTT model by complex instances coming from 4 real Italian schools. Moreover, IHSTT is also tested on some well-known benchmark instances taken from the ITC2011 competition. Last but not least, IHSTT is run on some well-known benchmark instances taken from [XHSTT] data-sets with short and long running times to evaluate solution quality and speed, respectively.

The KHE is a well-known freeware open-source C program, that implements an advanced heuristic described in [Kin14] and supports the [XHSTT] format. Although KHE does not have any time limit for optimization, the option of multiple separate threads can be introduced to obtain better solutions. Clearly, the KSS model supports [XHSTT] format and can be solved by any MIP solver, but additional implementations were performed to process the main decision variables in order to scale to larger problems instances.

The experimentation is organized according to two experimental settings, which differ for which constraints are hard, soft or disabled. These settings are called *Setting1* and *Setting2*. For the sake of clarity it is denoted by *Setting1* the experimentation with requirements R19, R22, R24 and R25, whereas in *Setting2* they are ignored. Therefore, *Setting1* represents the current case of Italian schools and *Setting2* is the simplified problem. Table 2.3 reports the types of constraints and the methods run in these settings.

The values of coefficients $\omega_i$ in the objective function of IHSTT indicate whether the $i$-th constraint is hard, soft or disabled. They can be set by schools according to their policies. When two options are reported (e.g. hard and disabled) in Table 2.3, some instances consider one option and other instances the other option. In this experimentation, hard constraints have values of $\omega_i$ equal to $100,000$. Soft-constraints have values of $\omega_i$ much lower than $100,000$ and typically range between 1 and 100. If $i$-th constraint is not used (or disabled), the corresponding value of $\omega_i$

[1] *http://jeffreykingston.id.au/khe/*

| Constraints | Requirements | *Setting1* IHSTT/TP-IHSTT | *Setting2* KHE/KSS/IHSTT/TP-KHE/TP-KSS/TP-IHSTT |
|---|---|---|---|
| C1 | R1,R2,R7 | Hard | Hard |
| C2 | R1 | Hard | Hard |
| C3 | R5,R9 | Hard | Hard |
| C4 | R6,R8 | Hard | Hard |
| C5 | R10 | Soft/Hard/Disabled | Soft/Hard/Disabled |
| C6 | R3,R4 | Hard | Hard |
| C7 | R9 | Hard/Disabled | Hard/Disabled |
| C8 | R11,R12 | Hard/Disabled | Hard/Disabled |
| C9 | R13 | Hard/Disabled | Hard/Disabled |
| **C10** | **R25** | **Hard** | **Disabled** |
| C11 | R17 | Soft | Soft |
| C12 | R18,R20,R21 | Soft/Hard | Soft/Hard |
| C13 | R14 | Soft | Disabled |
| C14 | R15 | Soft | Disabled |
| C15 | R16 | Hard | Hard |
| C16 | R23 | Hard | Hard |
| **C17** | **R22** | **Soft** | **Disabled** |
| **C18** | **R19** | **Soft** | **Disabled** |
| **C19** | **R24** | **Disabled/Hard** | **Disabled** |

Table 2.3: Types of constraints in the two experimental settings.

is 0. The details about the values are reported for each instance in tables 2.13 and 2.14 of Appendix C.

In *Setting1* the solutions provided by a MIP solver on IHSTT and TP-IHSTT are compared. The outcomes of *Setting1* are reported in Table 2.5; in addition, the value of slack variables are reported in Table 2.15 and Table 2.16. In *Setting2* the KHE heuristic, the same MIP solver running KSS and IHSTT, as well as TP-KHE, TP-KSS, TP-IHSTT are compared. These outcomes are reported in Table 2.6.

Although [XHSTT] benchmark instances exist in the literature, 24 new specific instances are built for this experimentation in order to capture all novelties arising in this problem setting. They describe real situations or realistic conditions according to expert-based opinions. The first 20 instances are grouped according to their size: in the first group, instances are denoted from 1 to 9 and their size ranges from small to medium; in the second group, instances are denoted from 10 to 20 and their size ranges from medium to huge. Finally the last 4 instances are grouped according to their complexity.

Instance 1 comes from the timetable data of a *real* middle school with 5 daily periods. Instance 2 is more realistic for high-schools, because it has 6 daily periods. Instance 3 adds to Instance 2 co-teaching with one full-time teacher and one part-time teacher. Instance 4 is more complex than the previous ones, because it features 3 articulated class (one teacher must teach two class in the same time) and blocks. Instance 5 represents a school in which classes have 5 hours and half every day, to have uniform entrance and exit times among all school classes. Instance 6 and

Instance 8 introduce fractional time units (of 30 and 15 minutes, respectively), to allow splitting the entrance of students (in two or four groups, respectively), reducing overcrowding. This is an emerging issue owing to the COVID-19 pandemic. Instance 7 is quite realistic, because every class has 12 teachers in a week. Instance 9 is more complex and larger than Instance 7: it features 18 classes and teachers with variable week requirements ranging from 2 to 4 time slots. The second group (instances between 10 and 20) is generated using a common block with 12 teachers and 6 classes with fixed week requirement (3 periods). Instances 21 and 22 are provided by mid-schools and have up to 24 classes. Instance 23 has 42 classes and come from a high-school with scientific, classic and linguistic curriculum. Instance 24 describes a technical school with 40 classes with several curricula, articulated classes and labs.

| Instance | $\|C\|$ | $\|T\|$ | *Requir. (1)* | $\|D\|$ | $\|H\|$ | $\|D\| \cdot \|H\|$ | $\|C\| \cdot \|D\| \cdot \|H\|$ | $\eta$ | $\Delta$ | $\rho$ | $\overline{\rho}$ | $\underline{\alpha}$ | $\overline{\alpha}$ | *CoTea (2)* | *Artic (3)* | *Blocks (4)* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 13 | 50 | 6 | 5 | 30 | 150 | 1 | 60’ | 1 | 2 | 2 | 5 | - | - | - |
| 2 | 3 | 6 | 18 | 6 | 6 | 36 | 108 | 1 | 60’ | 1 | 2 | 3 | 5 | - | - | - |
| 3 | 3 | 8 | 36 | 6 | 6 | 36 | 108 | 1 | 60’ | 1 | 2 | 3 | 5 | 21 | - | 3 |
| 4 | 6 | 6 | 36 | 6 | 6 | 36 | 216 | 1 | 60’ | 1 | 2 | 3 | 5 | - | 18 | 90 |
| 5 | 3 | 6 | 18 | 6 | 11 | 66 | 198 | 1 | 30’ | 2 | 7 | 3 | 10 | - | - | - |
| 6 | 3 | 6 | 18 | 6 | 12 | 72 | 216 | 2 | 30’ | 2 | 4 | 6 | 10 | - | - | - |
| 7 | 6 | 12 | 72 | 6 | 6 | 36 | 216 | 1 | 60’ | 1 | 2 | 3 | 5 | - | - | - |
| 8 | 3 | 6 | 18 | 6 | 24 | 144 | 432 | 4 | 15’ | 4 | 8 | 4 | 12 | - | - | - |
| 9 | 18 | 36 | 216 | 6 | 6 | 36 | 648 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | 648 |
| 10 | 18 | 36 | 216 | 6 | 6 | 36 | 648 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 11 | 24 | 48 | 288 | 6 | 6 | 36 | 864 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 12 | 30 | 60 | 360 | 6 | 6 | 36 | 1080 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 13 | 36 | 72 | 432 | 6 | 6 | 36 | 1296 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 14 | 42 | 84 | 504 | 6 | 6 | 36 | 1512 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 15 | 42 | 85 | 504 | 6 | 6 | 36 | 1512 | 1 | 60’ | 1 | 1 | 3 | 5 | 18 | - | - |
| 16 | 48 | 96 | 576 | 6 | 6 | 36 | 1728 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 17 | 54 | 108 | 648 | 6 | 6 | 36 | 1944 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 18 | 60 | 120 | 720 | 6 | 6 | 36 | 2160 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 19 | 78 | 156 | 936 | 6 | 6 | 36 | 2808 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 20 | 156 | 312 | 1872 | 6 | 6 | 36 | 5616 | 1 | 60’ | 1 | 1 | 3 | 5 | - | - | - |
| 21 | 23 | 44 | 230 | 6 | 5 | 30 | 690 | 1 | 60’ | 1 | 2 | 1 | 5 | - | - | - |
| 22 | 24 | 44 | 288 | 5 | 6 | 30 | 720 | 1 | 60’ | 1 | 2 | 1 | 5 | - | - | - |
| 23 | 42 | 79 | 491 | 6 | 6 | 36 | 1512 | 1 | 60’ | 1 | 3 | 1 | 5 | - | - | - |
| 24 | 40 | 103 | 697 | 6 | 6 | 36 | 1440 | 1 | 60 | 1 | 4 | 1 | 6 | 270 | 30 | 30 |

Table 2.4: Description of the Italian schools’ instances

The most important problem data of each instance are reported in Table 2.4, where *Requir. (1)* indicates the number of timetable requirements, $\Delta$ the duration of a single period (in minutes), *CoTea (2)* number of lessons in co-teaching, *Artic. (3)* number of lessons in articulated classes, *Blocks (4)* number of block lessons. The data files of instances in *Setting2* are also available in [XHSTT] format[2], whereas this is not possible for *Setting1*, because some constraints of [IHSTP] are not supported in [XHSTT] standard. In this experimentation, it is adopted the modeling language IBM OPL to call the MIP solver CPLEX 20.1 for implementing and solving all models. All the experiments are performed on a computer with an Intel I5-4460

[2] *https://github.com/ClaudioCrobu/IHSTP*

3.20 GHz 4-core CPU equipped with 32 GBytes of DDR3 RAM and 1 TBytes SSD drive running Ubuntu 20.04 LTS. The time limit is 3 hours.

| ***Setting1*** | *IHSTT* | | | | | *TP-IHSTT* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Instance* | *Time* | *Obj* | *Gap* | *Idle times* | *LBGap* | *Time* | *Obj* | *Idle times* | *LBGap* | *TP time* |
| 1 | TL | 2625 | 2917 | 0.9 | 100 | **40.5** | **87** | **0.0** | **0** | 19.5 |
| 2 | TL | 1970 | 228 | 2.3 | 100 | **4.1** | **600** | **1.0** | **0** | 1.8 |
| 3 | TL | 2364 | 224 | 2.1 | 100 | **160.9** | **730** | **0.9** | **4** | 1.4 |
| 4 | TL | 2436 | 121 | 1.8 | 100 | **60.0** | **1100** | **1.0** | **0** | 45.6 |
| 5 | TL | 2610 | 226 | 3.8 | 100 | **304.7** | **800** | **1.0** | **25** | 81.4 |
| 6 | TL | 3904 | 225 | 5.7 | 100 | **8.8** | **1200** | **2.0** | **0** | 7.0 |
| 7 | TL | 3615 | 198 | 2.5 | 100 | **42.1** | **1215** | **1.0** | **0** | 0.1 |
| 8 | TL | 8688 | 262 | 11.3 | 100 | **234.9** | **2400** | **4.0** | **0** | 145.8 |
| 9 | TL | 21600 | 454 | 4.7 | 100 | **270.2** | **3840** | **1.0** | **0** | 22.1 |
| 10 | TL | 21600 | 500 | 4.8 | 100 | **21.7** | **3600** | **1.0** | **0** | 6.4 |
| 11 | TL | 33600 | 600 | 5.0 | 100 | **27.0** | **4800** | **1.0** | **0** | 8.7 |
| 12 | TL | 42000 | 600 | 5.4 | 100 | **45.9** | **6000** | **1.0** | **0** | 11.3 |
| 13 | TL | 64800 | 800 | 5.6 | 100 | **38.3** | **7200** | **1.0** | **0** | 13.6 |
| 14 | TL | 75600 | 800 | 5.4 | 100 | **46.2** | **8400** | **1.0** | **0** | 16.3 |
| 15 | TL | 75600 | 800 | 5.3 | 100 | **86.7** | **8400** | **1.0** | **0** | 16.2 |
| 16 | TL | 86409 | 800 | 5.7 | 100 | **50.9** | **9600** | **1.0** | **0** | 21.9 |
| 17 | TL | 108012 | 900 | 5.3 | 100 | **57.7** | **10800** | **1.0** | **0** | 29.0 |
| 18 | TL | 146189 | 1118 | 5.5 | 100 | **272.8** | **12000** | **1.0** | **0** | 42.3 |
| 19 | TL | 175146 | 1023 | 5.4 | 100 | **75.2** | **15600** | **1.0** | **0** | 27.6 |
| 20 | TL | 379861 | 1118 | 5.1 | 100 | **2499.1** | **31203** | **1.0** | **0** | 183.0 |

Table 2.5: Results of *Setting1* (idle times are expressed in hours, all remaining times in seconds; TL = Time Limit = 10800 seconds)

Table 2.5 focuses on *Setting1* and is organized into three groups of columns. The first column lists the instances, the second group reports the outcomes of the IHSTT, the third group shows the results of TP-IHSTT. For example, in instance 5 the TP is solved in 81.4 seconds and the overall two-step method in 304.7 seconds. All the instances cannot be solved by the IHSTT within the time limit.

Two quality measures are reported in the results: the objective value (denoted by *Obj*) and the mean of the idle times (denoted by *Idle times*). For example, according to the solution of the IHSTT, in instance 16 this value is 5.7 hours, but it is obtained at the time limit, when constraint C10 (*Equity Idle Times*) is not satisfied. The same instance is effectively solved by the TP-IHSTT, which returns a much lower value of idle times for all teachers.

Two types of gaps are reported. The column *Gap* indicates the relative difference between *Obj* and the lower bound $L_B$ returned by the MIP solver. It is computed as:

$$Gap = \left\lceil 100 \cdot \frac{Obj - L_B}{L_B} \right\rceil \tag{2.89}$$

When $L_B = 0$, no value of *Gap* is reported. Moreover, the reported value of *Gap* is $\infty$ when hard constraints are not satisfied. This gap is not reported in the group of columns TP-IHSTT, because it takes always value zero.

The column *LBGap* indicates the relative difference between *Obj* and the lower bound $\overline{L}_B$ at the root node after CPLEX cuts:

$$LBGap = \left[100 \cdot \frac{Obj - \overline{L}_B}{Obj}\right] \tag{2.90}$$

The best outcomes are emphasized in bold.

Table 2.6 pertains to *Setting2* and is organized into eight groups of columns. The first group lists the instances, the following six groups report the outcomes of KHE, KSS, IHSTT, TP-KHE, TP-KSS and TP-IHSTT, the last group shows the common time of the first phase of the 2-phase methods.

In order to make a fairer comparison on KHE, it was used with the option of parallel threads. KHE is run with different values of threads and reported the best one in column *Threads*. For the first group of instances 1, 10, 100 and 1000 threads number were used; the option with 1000 threads was not used for the second group of instances because of a memory problem. It is remarked that the concept of threads in KHE is different from the one adopted in CPLEX which corresponds to the CPU cores. Since KHE does not compute a lower bound, in the computation of *Gap*, this is replaced by the best upper bound computed by the other methods. The string MEM means that the computer's available memory was insufficient for building the instance.

In the last row of Table 2.6 the average rank of each method is reported. It is computed by assigning value 1 to the minimum objective function in the group, value 2 to the second and so on, but in case of equality the method with the minimum time is considered. According to this logic, the best results are emphasized in bold.

| ***Setting2*** | *KHE* | | | *KSS* | | | *IHSTT* | | | *TP-KHE* | | | *TP-KSS* | | | *TP-IHSTT* | | *TP* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Instance* | *Time* | *Obj* | *Threads* | *Time* | *Obj* | *Gap* | *Time* | *Obj* | *Gap* | *Time* | *Obj* | *Threads* | *Time* | *Obj* | *Gap* | *Time* | *Obj* | *Time* |
| 1 | 350.2 | 4000 | 1000 | TL | 200 | - | 593.1 | **0** | 0 | 195.2 | 12000 | 1000 | 6.9 | 0 | 0 | **5.6** | **0** | 3.6 |
| 2 | 388.8 | (5,1200) | 1000 | TL | **600** | 0 | TL | **600** | 0 | 408.1 | (6,600) | 1000 | **2.7** | **600** | 0 | 2.8 | 600 | 1.6 |
| 3 | 339.2 | (3,900) | 1000 | 6550.1 | **600** | 0 | TL | 700 | 17 | 214.1 | (4,600) | 1000 | 2.8 | 600 | 0 | **2.5** | **600** | 1.4 |
| 4 | 784.4 | (11,2500) | 1000 | **2335.6** | **1100** | 0 | TL | 1100 | 0 | 447.7 | (15,1100) | 1000 | **67.2** | **1100** | 0 | 78.3 | 1100 | 64.7 |
| 5 | 476.1 | 3600 | 1000 | TL | (1,3600) | $\infty$ | TL | **800** | 33 | 569.1 | (20,1300) | 1000 | **75.0** | **600** | 0 | 82.6 | 600 | 64.8 |
| 6 | 1059.0 | (4,1800) | 1000 | TL | **1200** | 0 | TL | **1200** | 0 | 622.1 | (5,1200) | 1000 | 9.3 | 1200 | 0 | **7.6** | **1200** | 6.6 |
| 7 | 1195.3 | 1300 | 1000 | TL | 1400 | 17 | 7645.1 | **1200** | 0 | 406.6 | 1200 | 1000 | **5.5** | **1200** | 0 | 15.5 | 1200 | 0.1 |
| 8 | 2821.1 | (4,5600) | 1000 | | MEM | | TL | **2400** | 0 | 2164.3 | (6,2400) | 1000 | | MEM | | 170.6 | **2400** | 165.3 |
| 9 | 2951.6 | **3900** | 1000 | TL | 4800 | 33 | TL | 4200 | 17 | 104.1 | 3600 | 100 | 20.6 | 3600 | 0 | **16.5** | **3600** | 14.5 |
| 10 | 342.1 | 4000 | 100 | TL | 7000 | 94 | TL | **3900** | 8 | 107.0 | 3600 | 100 | 20.3 | 3600 | 0 | **16.9** | **3600** | 14.2 |
| 11 | 406.7 | 5400 | 100 | TL | (3,13500) | $\infty$ | TL | **5300** | 10 | 136.7 | 4800 | 100 | 21.4 | 4800 | 0 | **17.3** | **4800** | 12.7 |
| 12 | 490.1 | 6700 | 100 | TL | (5,17500) | $\infty$ | TL | **6500** | 8 | 181.5 | 6000 | 100 | 27.8 | 6000 | 0 | **21.2** | **6000** | 16.1 |
| 13 | 693.2 | 8300 | 100 | | MEM | | TL | **7900** | 10 | 190.3 | 7200 | 100 | | MEM | | **25.6** | **7200** | 19.5 |
| 14 | 871.9 | **9600** | 100 | | MEM | | TL | 9800 | 17 | 245.9 | 8400 | 100 | | MEM | | **30.0** | **8400** | 22.2 |
| 15 | 830.5 | **9400** | 100 | | MEM | | TL | 9700 | 15 | 247.0 | 8400 | 100 | | MEM | | **32.4** | **8400** | 22.4 |
| 16 | 1094.6 | **10900** | 100 | | MEM | | TL | 11100 | 16 | 480.2 | 9600 | 100 | | MEM | | **196.1** | **9600** | 183.4 |
| 17 | 1286.4 | **12400** | 100 | | MEM | | TL | 12400 | 15 | 456.5 | 10800 | 100 | | MEM | | **87.3** | **10800** | 73.3 |
| 18 | 1302.8 | **13400** | 100 | | MEM | | TL | 14100 | 18 | 602.4 | 12000 | 100 | | MEM | | 203.0 | **12000** | 183.5 |
| 19 | 2053.6 | 17900 | 100 | | MEM | | TL | **17800** | 14 | 855.9 | 15600 | 100 | | MEM | | **269.6** | **15600** | 232.9 |
| 20 | 6557.6 | **37200** | 100 | | MEM | | TL | 39500 | 27 | 624.7 | 31200 | 10 | | MEM | | 388.4 | 31200 | 303.2 |
| Average rank | | *1.90* | | | *2.55* | | | ***1.45*** | | | *2.20* | | | *2.15* | | | ***1.20*** | |

Table 2.6: Results of *Setting2* (All times are expressed in seconds; TL = Time Limit = 10800 seconds; $\infty$ = Infeasible; MEM = memory exhausted)

| **Statistics** | *Setting1* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | *IHSTT* | | | *TP-IHSTT* | | | *TP* | | |
| *Instance* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* |
| 1 | 2668 | 6068 | 26816 | 1268 | 2472 | 8481 | 1312 | 1990 | 8090 |
| 2 | 1053 | 2834 | 12296 | 720 | 1758 | 5886 | 647 | 734 | 3600 |
| 3 | 1605 | 3828 | 15884 | 1083 | 2558 | 8849 | 647 | 734 | 3600 |
| 4 | 2595 | 6516 | 26352 | 1734 | 4059 | 13077 | 1817 | 1918 | 9608 |
| 5 | 1593 | 4904 | 22286 | 990 | 2664 | 9456 | 1127 | 1192 | 7920 |
| 6 | 1521 | 6002 | 26120 | 918 | 2670 | 9126 | 827 | 1114 | 5922 |
| 7 | 3136 | 6922 | 32086 | 2290 | 4345 | 15224 | 251 | 242 | 906 |
| 8 | 2457 | 18608 | 112868 | 1580 | 8401 | 37561 | 2248 | 9922 | 95587 |
| 9 | 9682 | 7152 | 79146 | 6402 | 4288 | 27596 | 3877 | 4356 | 22752 |
| 10 | 9769 | 7212 | 80616 | 6480 | 4302 | 27756 | 3877 | 4320 | 21672 |
| 11 | 13025 | 9616 | 107488 | 8640 | 5760 | 37296 | 5169 | 5760 | 28896 |
| 12 | 16281 | 12020 | 134360 | 10800 | 7170 | 46260 | 6461 | 7200 | 36120 |
| 13 | 19537 | 14424 | 161232 | 12960 | 8640 | 55872 | 7753 | 8640 | 43344 |
| 14 | 22793 | 16828 | 188104 | 15120 | 10080 | 65352 | 9045 | 10080 | 50568 |
| 15 | 22793 | 16832 | 188156 | 15120 | 10080 | 65388 | 9045 | 10080 | 50568 |
| 16 | 26049 | 19232 | 214976 | 17280 | 11520 | 74496 | 10337 | 11520 | 57792 |
| 17 | 29305 | 21636 | 241848 | 19440 | 12852 | 82728 | 11629 | 16200 | 71496 |
| 18 | 32561 | 24040 | 268729 | 21600 | 14328 | 92304 | 12921 | 18000 | 79440 |
| 19 | 42329 | 31252 | 349336 | 28080 | 18642 | 120276 | 16797 | 18720 | 93912 |
| 20 | 84657 | 62504 | 698672 | 56160 | 37290 | 240504 | 33593 | 46800 | 206544 |

| **Statistics** | *Setting2* | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *KSS* | | | *IHSTT* | | | *TP-KSS* | | | *TP-IHSTT* | | | *TP* | | |
| *Instance* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* | *#Var.* | *#Con.* | *#NZ* |
| 1 | 11119 | 14274 | 88779 | 2134 | 5425 | 21022 | 4444 | 2022 | 27104 | 1070 | 2194 | 7130 | 1144 | 1879 | 6906 |
| 2 | 7836 | 9768 | 65460 | 873 | 1841 | 8096 | 3474 | 1362 | 22242 | 594 | 1506 | 4662 | 640 | 718 | 4402 |
| 3 | 8097 | 10295 | 64721 | 1321 | 2750 | 11726 | 3424 | 1599 | 21341 | 932 | 2259 | 7209 | 640 | 718 | 3402 |
| 4 | 14424 | 18020 | 121436 | 2294 | 4817 | 16938 | 6391 | 2555 | 41447 | 1482 | 3555 | 10629 | 1385 | 1615 | 6711 |
| 5 | 16800 | 28146 | 269184 | 1350 | 4446 | 17700 | 6111 | 3278 | 79656 | 834 | 2364 | 7584 | 1120 | 1186 | 7878 |
| 6 | 18672 | 30720 | 273216 | 1260 | 5508 | 21510 | 4827 | 3090 | 44955 | 684 | 2202 | 6948 | 827 | 1114 | 5922 |
| 7 | 16460 | 20182 | 123918 | 2715 | 6457 | 26880 | 7228 | 3500 | 43350 | 1923 | 3932 | 13174 | 264 | 253 | 983 |
| 8 | | MEM | | 1980 | 17682 | 104130 | | MEM | | 1053 | 6637 | 29061 | 610 | 910 | 3894 |
| 9 | 39126 | 48450 | 217794 | 8070 | 7530 | 59955 | 16848 | 7956 | 58392 | 4938 | 2766 | 20274 | 4020 | 4164 | 20160 |
| 10 | 39348 | 48888 | 218880 | 8100 | 5544 | 61020 | 16956 | 8064 | 58608 | 4968 | 2772 | 20520 | 4020 | 4164 | 20160 |
| 11 | 53904 | 66624 | 293280 | 10800 | 7392 | 81360 | 22608 | 10752 | 78144 | 6624 | 3696 | 27360 | 5360 | 5552 | 26880 |
| 12 | 18978 | 10261 | 84576 | 13500 | 9240 | 101700 | 28260 | 13440 | 97680 | 8280 | 4620 | 34200 | 6700 | 6940 | 33600 |
| 13 | | MEM | | 16200 | 11088 | 122040 | | MEM | | 9936 | 5544 | 41040 | 8040 | 8328 | 40320 |
| 14 | | MEM | | 18900 | 12936 | 142380 | | MEM | | 11592 | 6468 | 47880 | 9380 | 9716 | 47040 |
| 15 | | MEM | | 18900 | 12936 | 142380 | | MEM | | 11592 | 6468 | 47880 | 9380 | 9716 | 47040 |
| 16 | | MEM | | 21600 | 14784 | 162720 | | MEM | | 13248 | 7392 | 54720 | 10720 | 13984 | 59520 |
| 17 | | MEM | | 24300 | 16632 | 183060 | | MEM | | 14904 | 8316 | 61560 | 12060 | 15732 | 66960 |
| 18 | | MEM | | 27000 | 18480 | 203400 | | MEM | | 16560 | 9240 | 68400 | 13400 | 17480 | 74400 |
| 19 | | MEM | | 35100 | 24024 | 264420 | | MEM | | 21528 | 12012 | 88920 | 17420 | 22724 | 96720 |
| 20 | | MEM | | 70200 | 48048 | 528840 | | MEM | | 43056 | 24024 | 177840 | 34840 | 45448 | 193440 |

Table 2.7: Statistics of the number of variables, constraints and non-zeros in IHSTT,TP-IHSTT programs for *Setting1* (Table 2.5) and in KSS,IHSTT,TP-KSS and TP-IHSTT programs for *Setting2* (Table 2.6)

Table 2.7 reports the number of variables (*#Var.*), constraints (*#Con.*) and non-zero coefficients (*#NZ*) in the constraints of both *Setting1* and *Setting2*. This table clearly shows the decrease in the size of instances when one switches from IHSTT to TP-IHSTT (*Setting1*), as well as from KSS to IHSTT with or without the TP step (*Setting2*).

| ***Setting2*** | *KHE TL1* | | *KHE TL2* | | *KSS model size* | | *KSS TL1* | | *KSS TL2* | | *IHSTT model size* | | *IHSTT TL1* | | *IHSTT TL2* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Instance* | *Obj* | *Threads* | *Obj* | *Threads* | *#var* | *#con* | *Obj* | *Gap* | *Obj* | *Gap* | *#var* | *#con* | *Obj* | *Gap* | *Obj* | *Gap* |
| 21 | 25900 | 10 | 25900 | 10 | 39330 | 17908 | **1600** | - | 1200 | - | 11766 | 26621 | 2500 | - | **400** | - |
| 22 | **4300** | 100 | **4300** | 100 | 42228 | 29261 | 7000 | - | 6100 | - | 12668 | 28679 | 6200 | - | **4300** | - |
| 23 | **1100** | 10 | **1100** | 10 | MEM | | MEM | | MEM | | 26628 | 59637 | 1410700 | - | 30100 | - |
| 24 | (22,1300) | 10 | (22,1300) | 10 | MEM | | MEM | | MEM | | 29290 | 52158 | **638700** | 2400 | **48500** | 79 |
| *Average rank* | *1.75 = (3+1+1+2)/4* | | *1.75 = (3+1+1+2)/4* | | | | *2.50 = =(1+3+3+3)/4* | | *2.75 = =(2+3+3+3)/4* | | | | *1.75 = (2+2+2+1)/4* | | *1.25 = (1+1+2+1)/4* | |

Table 2.8: Results of *Setting2* for instances 21-24 (All times are expressed in seconds; TL1 = Time Limit 1 = 10800 seconds; TL2 = Time Limit 2 = 43200 seconds; $\infty$ = Infeasible; MEM = memory exhausted)

Table 2.8 focuses on the last four instances. They are solved by KHE, KSS and IHSTT within two time limits TL1 and TL2, which amount to 10800 seconds and 43200 seconds, respectively. As for KHE, the objective function Obj and the optimal number of threads are reported. Next, data and outcomes of the KSS are shown. More precisely, the model size in terms of number of variables (*#Var.*) and constraints (*#Con.*). The outcomes are the objective (*Obj*) and the gap (*Gap*) computed by (2.89) are reported. The same organization of results is adopted for IHSTT. The average rank is computed as in Table 2.6, but it is separately computed for TL1 and TL2.

Although IHSTT was motivated by the case of Italian schools, it is aimed to show its compatibility with respect to some benchmark instances and its capability in

| ***Without time limits*** | | | |
|---|---|---|---|
| *Instance* | *Best known* | *KSS* | *IHSTT* |
| Brazil 1 | 41* | 41 | 41 |
| Brazil 2 | 5* | 5 | 5 |
| Brazil 3 | 24* | 26 | 26 |
| Brazil 4 | 51* | 61 | 59 |
| Brazil 5 | 19* | 30 | 41 |
| Brazil 6 | 35* | 60 | 98 |
| Brazil 7 | 53 | 122 | 113 |

Table 2.9: Brazilian instances without any time limits (* means *optimum*)

obtaining good solutions even in this case. The experimentation is divided into two parts. In the first part, some [XHSTT] instances are solved without any time limits, according to the policy of round 1 of the ITC2011 competition. This experimentation is reported in Table 2.9 on instances from *Brazil 1* to *Brazil 7*, for which the best known solutions are taken from [XHS]. The solutions of IHSTT are compared to those for KSS, as reported in [KSS15]. In the second part, four instances of the Round 2 of the ITC2011 competition are solved to evaluate the IHSTT formulation with respect to some well-known benchmarks. The rules of Round 2 are rigorously followed. The time limit is set by the ITC2011 benchmark utility and only one thread is used. The results of IHSTT are reported in Table 2.10 and compared to all participants of Round 2, as reported in [ITC]. Moreover, Table 2.11 compares the size of these instances for IHSTT and KSS according to [KSS15].

Finally, since this problem is naturally affected by symmetry, all tests are rerun while disabling the symmetry control options in CPLEX. However, these results are never better than those presented so far.

### 2.6.2 Analysis of results

#### ***Setting1***

Consider the columns denoted by IHSTT in Table 2.5. They show that all instances use the overall available time to determine low quality upper-bounds. Moreover, the lower bounds at the root node after CPLEX cuts are always zero. In addition, the average idle times are not acceptable.

Consider the groups columns denoted by TP-IHSTT in Table 2.5. In this case, all instances are optimally solved within the time limit. The column "*TP time*" shows that an acceptable time is spent for solving the TP. The time spent in the

| **ITC2011** | | *GOAL* | | *HySST* | | *Lectio* | | *HFT* | | *IHSTT* | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Instance* | *Seed* | *Obj* | *Rank* | *Obj* | *Rank* | *Obj* | *Rank* | *Obj* | *Rank* | *Obj* | *Rank* |
| BrazilInstance2 | 102545520 | 1.00063 | 3 | 1.00078 | 4 | 0.00046 | 1 | 7.00189 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 109328591 | 1.00054 | 3 | 1.00075 | 4 | 0.00057 | 1 | 5.00183 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 234546972 | 1.00087 | 4 | 1.00081 | 3 | 0.00028 | 1 | 7.00180 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 317604170 | 1.00051 | 3 | 1.00078 | 4 | 0.00019 | 1 | 6.00186 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 584363925 | 1.00054 | 3 | 1.00069 | 4 | 0.00047 | 1 | 5.00198 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 65843198 | 1.00063 | 3 | 1.00082 | 4 | 0.00038 | 1 | 6.00207 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 792992094 | 1.00063 | 3 | 1.00087 | 4 | 0.00025 | 1 | 7.00195 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 802033156 | 1.00066 | 3 | 1.00072 | 4 | 0.00034 | 1 | 6.00165 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 856676505 | 1.00066 | 3 | 1.00072 | 4 | 0.00034 | 1 | 7.00210 | 5 | 0.00103 | 2 |
| BrazilInstance2 | 96247109 | 1.00051 | 3 | 1.00078 | 4 | 0.00053 | 1 | 7.00189 | 5 | 0.00103 | 2 |
| BrazilInstance3 | 102545520 | 0.00132 | 3 | 0.00096 | 1 | 0.00159 | 4 | 29.00264 | 5 | 0.00126 | 2 |
| BrazilInstance3 | 109328591 | 0.00134 | 3 | 0.00126 | 1 | 0.00175 | 4 | 31.00288 | 5 | 0.00126 | 1 |
| BrazilInstance3 | 234546972 | 0.00138 | 3 | 0.00123 | 1 | 0.00153 | 4 | 28.00285 | 5 | 0.00126 | 2 |
| BrazilInstance3 | 317604170 | 0.00087 | 1 | 0.00111 | 2 | 0.00112 | 3 | 30.00306 | 5 | 0.00126 | 4 |
| BrazilInstance3 | 584363925 | 0.00117 | 2 | 0.00096 | 1 | 0.00150 | 4 | 26.00264 | 5 | 0.00126 | 3 |
| BrazilInstance3 | 65843198 | 0.00135 | 3 | 0.00123 | 1 | 0.00171 | 4 | 32.00276 | 5 | 0.00126 | 2 |
| BrazilInstance3 | 792992094 | 0.00129 | 2 | 0.00132 | 3 | 0.00136 | 4 | 29.00273 | 5 | 0.00126 | 1 |
| BrazilInstance3 | 802033156 | 0.00137 | 3 | 0.00135 | 2 | 0.00167 | 4 | 29.00303 | 5 | 0.00126 | 1 |
| BrazilInstance3 | 856676505 | 0.00120 | 1 | 0.00133 | 3 | 0.00149 | 4 | 29.00288 | 5 | 0.00126 | 2 |
| BrazilInstance3 | 96247109 | 0.00111 | 2 | 0.00102 | 1 | 0.00149 | 4 | 32.00288 | 5 | 0.00126 | 3 |
| BrazilInstance4 | 102545520 | 17.00099 | 4 | 5.00221 | 3 | 1.00188 | 2 | 64.00258 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 109328591 | 18.00090 | 4 | 3.00241 | 3 | 2.00202 | 2 | 67.00243 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 234546972 | 18.00093 | 4 | 2.00238 | 3 | 1.00172 | 2 | 66.00246 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 317604170 | 18.00093 | 4 | 4.00242 | 3 | 2.00185 | 2 | 66.00234 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 584363925 | 17.00111 | 4 | 3.00233 | 2 | 4.00265 | 3 | 68.00243 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 65843198 | 17.00102 | 4 | 3.00210 | 3 | 3.00201 | 2 | 63.00225 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 792992094 | 17.00102 | 4 | 4.00223 | 3 | 2.00215 | 2 | 67.00243 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 802033156 | 18.00083 | 4 | 5.00227 | 3 | 3.00212 | 2 | 68.00195 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 856676505 | 16.00107 | 4 | 3.00239 | 3 | 3.00200 | 2 | 68.00222 | 5 | 0.00165 | 1 |
| BrazilInstance4 | 96247109 | 16.00104 | 4 | 3.00235 | 3 | 2.00150 | 2 | 68.00258 | 5 | 0.00165 | 1 |
| BrazilInstance6 | 102545520 | 4.00234 | 4 | 3.00273 | 3 | 0.00250 | 1 | 22.00438 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 109328591 | 4.00225 | 4 | 2.00270 | 3 | 0.00192 | 1 | 23.00363 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 234546972 | 4.00236 | 4 | 3.00281 | 3 | 0.00204 | 1 | 24.00369 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 317604170 | 4.00222 | 4 | 3.00240 | 3 | 0.00218 | 1 | 22.00360 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 584363925 | 4.00230 | 4 | 3.00284 | 3 | 0.00323 | 1 | 23.00438 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 65843198 | 4.00228 | 4 | 2.00229 | 3 | 0.00183 | 1 | 25.00387 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 792992094 | 4.00246 | 4 | 3.00298 | 3 | 0.00241 | 1 | 21.00423 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 802033156 | 4.00210 | 4 | 3.00256 | 3 | 0.00191 | 1 | 24.00372 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 856676505 | 4.00207 | 3 | 4.00270 | 4 | 0.00261 | 1 | 22.00384 | 5 | 0.00703 | 2 |
| BrazilInstance6 | 96247109 | 4.00228 | 4 | 3.00291 | 3 | 0.00239 | 1 | 23.00369 | 5 | 0.00703 | 2 |
| ***Final ranking*** | | *GOAL* | ***3.33*** | *HySST* | ***2.88*** | *Lectio* | ***2.00*** | *HFT* | ***5.00*** | *IHSTT* | ***1.78*** |

Table 2.10: Results of simulated *ITC2011* - Round 2 (Time limit used for IHSTT = 556 seconds - CPLEX parameter threads was set to 1 - *Obj* corresponds to *Cost* in ITC2011 tables)

| | **IHSTT** | | | | **KSS** | | | |
|---|---|---|---|---|---|---|---|---|
| *Instance* | *#Var* | *#Con* | *#NZ* | *#Var×#Con* | *#Var* | *#Con* | *#NZ* | *#Var×#Con* |
| Brazil Instance 2 | 1E+04 | 5E+03 | 5E+04 | 5E+07 | 3E+04 | 1E+04 | 1E+05 | 3E+08 |
| Brazil Instance 3 | 1E+04 | 6E+03 | 6E+04 | 9E+07 | 3E+04 | 2E+04 | 2E+05 | 6E+08 |
| Brazil Instance 4 | 3E+04 | 1E+04 | 1E+05 | 4E+08 | 5E+04 | 2E+04 | 2E+05 | 1E+09 |
| Brazil Instance 6 | 4E+04 | 2E+04 | 2E+05 | 6E+08 | 6E+04 | 3E+04 | 3E+05 | 2E+09 |
| Mean relative model size | **100%** | | | | **446%** | | | |
| Mean relative non-zeros size | **100%** | | | | **232%** | | | |

Table 2.11: Model size comparison on *ITC2011*

TP step is on average 68% of the total running time for first group of instances and 80% for second group, if the default parameters are used for the configuration of CPLEX. The lower bounds at the root node after CPLEX cuts are often equal to the final integer solution. Generally speaking, the two-step method returns lower values of the average idle times, i.e. higher-quality timetables from the viewpoint of teachers in real applications. Moreover, the left side of Table 2.7 shows the decrease in the number of variables, constraints and non-zeroes entries if one switches from IHSTT to TP-IHSTT. Therefore, the two-step method looks a promising approach for solving [HST] problems and it is worth investigating its viability also in *Setting2*.

### ***Setting2*** **without TP step**

Consider the columns denoted by KHE, KSS and IHSTT in Table 2.6. In the first group of instances (1-9), IHSTT outperforms KSS: KSS performs better for 4 times out of 9, while IHSTT determines the optimal solutions for 6 instances out of 9. Furthermore KSS does not get the first feasible solution within the time limit for two instances; such a situation never occurs to IHSTT. Although KHE software does not give guarantees of optimality, the comparison to IHSTT shows that it performs better only for one time and obtains infeasible solutions only five instances out of nine.

In the second group of instances IHSTT works better in five instances out of eleven. The other solutions have an optimality gap ranging from 15% to 27%. KSS gets a feasible solution one time out of eleven, in two instances it does not return the first feasible solution within the time limit. Therefore, IHSTT is always superior to KSS in all instances. KHE always obtains feasible solutions. Hence, it looks better than KSS and slightly worse than IHSTT.

Moreover, Table 2.7 shows the decrease in the number of variables, constraints and non-zeroes entries if one switches from KSS to IHSTT.

### ***Setting2*** **with TP step**

Consider the columns denoted by prefix **TP-** in Table 2.6. In the first group of instances (1-9), TP-IHSTT proves to be superior to TP-KHE and TP-KSS in 5 instances out of 9. KHE does not take advantage of the TP step and in seven cases it worsens w.r.t. the case without the TP step. The benefits of the TP step are instead very clear for both KSS and IHSTT, as they show significant improvements in gaps and optimization times.

In the second group of instances (10-20) TP-KHE improves all solutions owing to the TP step. The comparison between TP-KSS and TP-IHSTT indicates a much better effectiveness of TP-IHSTT in terms of running times. Furthermore, TP-KSS is more demanding from the point of view of memory use, as the 8 largest instances cannot be solved and compared to TP-IHSTT.

Yet, Table 2.7 shows the decrease in the number of variables, constraints and non-zeroes entries if one switches from TP-KSS to TP-IHSTT.

According to the former results, it is of interest to run IHSTT for a larger time limit, to possibly obtain optimal solutions for all instances in Table 2.4. A final experimentation is carried out with a time limit of 24 hours and the optimal solutions are eventually obtained for all instances. These solutions are equal to those of the two-step method, i.e. the proposed method returns the optimal solutions for all instances in Table 2.4. Moreover, the method exhibits a considerable speed-up in running times.

### Real instances from 4 Italian schools

According to Table 2.8, KSS can determine feasible solutions for two instances out of four within the usual time limit of 3 hours = 10800 seconds. In both cases, the value of Gap cannot be computed by (formula), because $L_B = 0$. If the running time is increased to 12 hours = 43200 seconds, the improvement is marginal in the first two instances, whereas it is still not possible to build the model for the last two instances. On the other hand, IHSTT can provide feasible solutions for all instances after 3 hours and a significant improvement is obtained after 12 hours. This is an experimental confirmation on its better use of memory. Finally, KHE can provide either very good and very poor solutions and no improvement can be obtained by larger running times.

### Comparison to benchmarks

Table 2.9 shows that IHSTT can obtain two best known solutions. All in all, it seems to have equivalent performances as opposed to KSS. Table 2.10 shows that this formulation leads to good-quality results in the time available with respect to the meta-heuristics taking part in the round 2 of this competition. Moreover, Table 2.11 compares the model dimensions of these instances according to the data reported in [KSS15]. According to Table 2.11, in these instances the average increases in memory size and number of nonzeros from IHSTT to KSS are 4.46 and 2.32, respectively. The memory size is computed by the product between the number of constraints and the number of decision variables.

### Other useful data about experimentation

Table 2.12 shows a summary of [IHSTP]-[TPP] slack variables.
Table 2.13 shows the values of weights of constraints used in *Setting1*.
Table 2.14 shows the values of weights of constraints used in *Setting2*.
Table 2.15 shows constraints slacks' values in IHSTT (*Setting1*).
Table 2.16 shows constraints slacks' values in TP-IHSTT (*Setting1*).

| Var. | Type | Description |
|---|---|---|
| $s_{ct}^{C_1}$ | Integer non-negative | Not assigned weekly lessons for class $c$ and teacher $t$ |
| $s_{cdh}^{C_2}$ | Boolean | Not assigned required lesson for class $c$ on day $d$ at period $h$ |
| $s_{cdh}^{C_3}$ | Boolean | Violated availability periods of class $c$ on day $d$ at period $h$ |
| $s_{tdh}^{C_4}$ | Boolean | Violated availability periods of teacher $t$ on day $d$ at period $h$ |
| $s_{ctd}^{C_5}$ | Boolean | Split lessons for class $c \in C$ and teacher $t$ on day $d$ |
| $s_{t}^{C_6}$ | Integer non-negative | Lack/excess of days off for teacher $t$ |
| $s_{ctf}^{C_7}$ | Integer non-negative | Lab lessons for class $c$, teacher $t$ and co-teacher $f$ in excess or in lack |
| $s_{c't'c''t''}^{C_8}$ | Integer non-negative | Block lessons for classes $c', c''$ with teachers $t', t''$ in excess or in lack |
| $s_{cdh}^{C_9}$ | Boolean | Not assigned preassigned lesson for class $c$ on day $d$ at period $h$ |
| $s^{C_{10}}$ | Integer non-negative | Difference between maximum and minimum idle times for teachers |
| $s_{td}^{C_{11}}$ | Integer non-negative | Idle times for teacher $t$ on day $d$ |
| $s_{l}^{C_{12}}$ | Integer non-negative | Violation for multiple lessons limit $l \in L$ |
| $s_{ct}^{C_{13}}$ | Integer non-negative | Violation of ideal weekly lessons' distribution for class $c$ and teacher $t$ |
| $s_{cth}^{C_{14}}$ | Integer non-negative | Violation of ideal daily lessons' distribution for class $c$ and teacher $t$ for period $h$ |
| $s_{td}^{C_{15}}$ | Integer non-negative | Violation of under-load/over-load limits for teacher $t$ on day $d$ |
| $s_{ctd}^{C_{16}}$ | Integer non-negative | Violation of under-load/over-load limits for class $c$/teacher $t$ on day $d$ |
| $s_{cd}^{C_{17}}$ | Integer non-negative | Presence of multiple lessons overload for class $c$ on day $d$ |
| $s_{td}^{C_{18}}$ | Integer non-negative | Presence of two consecutive heavy days $d, d+1$ for teacher $t$ |
| $s_{td}^{C_{19}}$ | Integer non-negative | Violation of fractional time units for teacher $t$ on day $d$ |

Table 2.12: IHSTT Slack variables summary

| Instance | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 | C15 | C16 | C17 | C18 | C19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100,000 | 100,000 | 100,000 | 100,000 | 1,000 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 0 | 1,000 | 0 |
| 2 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 3 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 4 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 5 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 6 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100 | 1 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 1 |
| 7 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 8 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100 | 1 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 1 |
| 9 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 10 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 11 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 12 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 13 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 14 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 15 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 16 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 17 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 18 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 19 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |
| 20 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 100,000 | 100 | 0 | 10 | 3 | 100,000 | 100,000 | 1 | 1,000 | 0 |

Table 2.13: Weights of constraints C1-C19 used in *Setting1* (Table 2.5)

| Instance | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 | C15 | C16 | C17 | C18 | C19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100,000 | 100,000 | 100,000 | 100,000 | 1,000 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 2 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 3 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 4 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 5 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 6 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100 | 1 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 7 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 8 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100 | 1 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 9 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 10 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 11 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 12 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 13 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 14 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 15 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 16 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 17 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 18 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 19 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |
| 20 | 100,000 | 100,000 | 100,000 | 100,000 | 0 | 100,000 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 100,000 | 100,000 | 0 | 0 | 0 |

Table 2.14: Weights of constraints C1-C19 used in *Setting2* (Table 2.6)

| Instance | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 | C15 | C16 | C17 | C18 | C19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0 | 8 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 18 | 0 | 8 | 30 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21 | 0 | 12 | 48 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 0 | 8 | 52 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 70 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | 0 | 16 | 48 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 0 | 96 | 176 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 216 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 216 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 336 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 420 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 648 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 756 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 756 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 864 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1,080 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1,440 | 0 | 128 | 303 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1,716 | 0 | 252 | 342 | 0 | 0 | 0 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3,744 | 0 | 370 | 587 | 0 | 0 | 0 | 0 | 0 |

Table 2.15: Constraints C1-C19 slacks' values in IHSTT (*Setting1* Table 2.5)

| Instance | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 | C15 | C16 | C17 | C18 | C19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 28 | 3 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 36 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 48 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 60 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 72 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 84 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 84 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 96 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 108 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 120 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 156 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 312 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

Table 2.16: Constraints C1-C19 slacks' values in TP-IHSTT (*Setting1* Table 2.5)

| Var. | Type | Description |
|---|---|---|
| $s_{tdh}^{SC1}$ | Boolean | 1 if teacher $t$ teaches in the same period $h$ in two consecutive days $d, d+1$, 0 otherwise |
| $s_t^{SC2}$ | Integer non-negative | Violation of minimum/maximum days off required |
| $s_{td}^{SC3}$ | Integer non-negative | Violation of under-load/over-load limits for teacher $t$ on day $d$ |
| $s_{td}^{SC4}$ | Integer non-negative | Idle times for teacher $t$ on day $d$ |
| $s^{SC5}$ | Integer non-negative | 0 if all teachers have the same minimum idle times, positive otherwise |

Table 2.17: TP Slack variables summary

# Chapter 3

# The Class Teacher Assignment Problem

## 3.1 Overview of the Class Teacher Assignment Problem

Hultberg and Cardoso [HC97] studied *Teacher Assignment Problem* and the conditions for partitioning sets of teachers, in order to simplify the Timetabling problem, but only in specific and unrealistic conditions. Moreover, this work addresses a simplified university course problem instead of a high-school problem.

No additional reference was detected on the *Class Teacher Assignment Problem*, despite it arises every year in any Italian high school.

In this problem, it is necessary to assign lessons to teachers according to the subjects in the curriculum of each class. The school has a set of teachers who have to guarantee these lessons, while taking into account several constraints depending on their contracts:

- full-time teachers must work for 18 hours a week, whereas part-time teachers must be on duty for less than 18 hours (usually 9 or 12) a week;
- some teachers must hold positions in more than one school and must guarantee the due service in all the schools;
- some teachers have management support roles (staff) and need to devote part of the hours of the standard teaching position (18 hours) to carry out their service tasks for the school community.

Furthermore, other constraints concern the number of hours available on various days due to the presence or absence of surveillance personnel.

Furthermore, some constraints concern the so-called teaching continuity, which almost obligatorily requires to assign a class to a teacher on a given subject if the

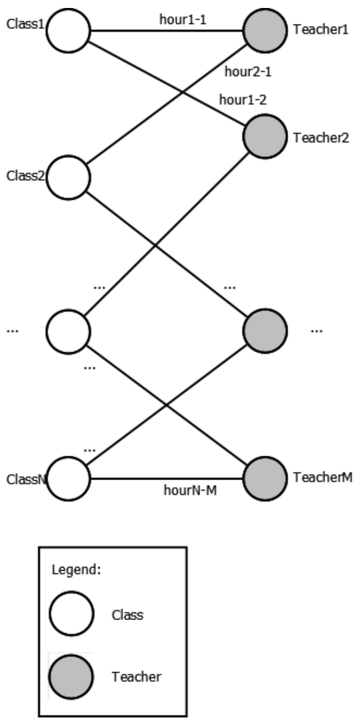



Figure 3.1: Class Teacher Assignment bipartite graph

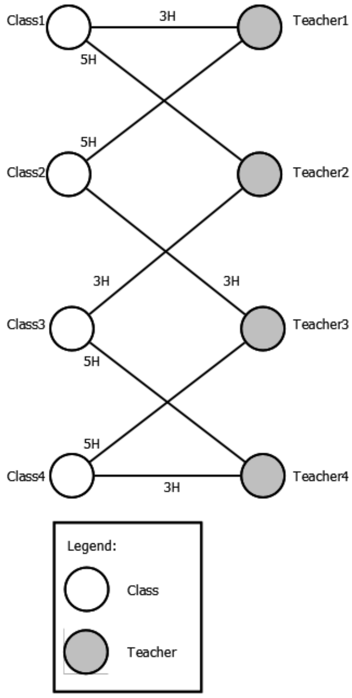
same teacher had taught that subject to the same students in the previous course year (for example: if a teacher taught the subject MATH in the previous year in class 3A, in the current the teacher should give lessons of MAT in class 4A).

It is worth noting that some solutions of the *Class Teacher Assignment Problem* can simplify the solution of the subsequent timetabling problem: grouping classes and teachers into clusters (or quasi-partitions) with the smallest possible number of connections leads to a timetabling problem (almost) separable for each cluster.

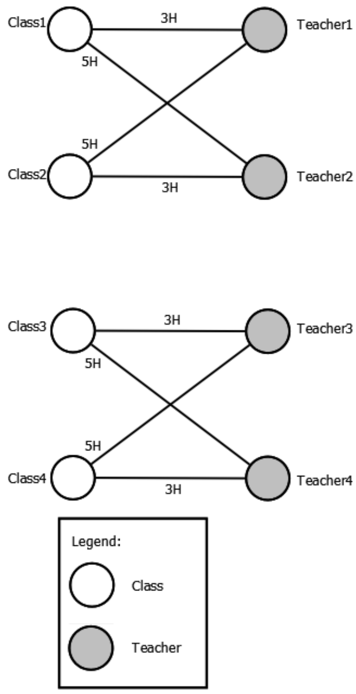
From a mathematical point of view, the problem corresponds to the partitioning of the bipartite graph having on one side the set of classes and on the other the set of teachers. Figure 3.1 shows an example with $|N|$ classes and $|M|$ teachers. The edges represent possible assignments of teachers to classes and report the numbers of hours of each teacher in each class.

Figure 3.2a a numerical example of the *Class Teacher Assignment Problem* and solution with one partition only, whereas Figure 3.2b reports a solution in which two partitions are obtained for an instance with same problem data.

In the latter case, the resolution of the timetabling of the original instance would lead to separable timetabling problems, which result in obvious benefits in terms of reduction of running times and improvements in the quality of solutions.




(a) One-partition bipartite graph




(b) Two-partition bipartite graph

Figure 3.2: Example of Class Teacher bipartite graphs

## 3.2 The Class Teacher Assignment Problem

### 3.2.1 Notation

Instances of the Italian High-School Timetabling Problem can be solved only if appropriate teachers were previously assigned to each class. The notation and the modeling of the Class Teacher Assignment Problem (CTAP) is presented hereafter.

In a school there are several classes (groups of students). Let $G$ be the set of classes.

Several subjects are taught ($S$) in classes by a group of teachers ($T$) and co-teachers ($C$). The co-teachers are teachers who always give lessons in the presence of another teacher. Each teacher $t \in T$ must be on duty for a total of $\theta_t$ hours per week.

Each class $g \in G$ must weekly attend a specific subset of subjects $\tilde{S}_g$ for a fixed number of teaching hours $\lambda_{gs}, s \in \tilde{S}_g$. Each class belongs to a group called class-group (or sections) because of the inner of schools. The parameter $\gamma_g$ specifies the class-group which the class $g \in G$ belongs to. Moreover, $\hat{C}$ is the set of class-groups and $\hat{C} = \bigcup_{g \in G} \gamma_g$.

Each subject may or may not require the presence of a co-teacher to assist the primary teacher: the parameter $\mu_s$ indicates the number of co-teachers required for the subject $s \in S$. Some subjects may require a linked lab subject: $L_s \subseteq S$ is the set of linked subjects for the subject $s \in S$ (it is empty when $s$ is not linked to any subject).

In order to obtain a partitioning in the CTAP, the number of desired partition is introduced as a data. Let $\nu$ the desired number of partitions, which may be different from the obtained number of partitions.

Each class $g \in G$ must belong to one and only one partition. The parameter $\pi_g^G$ is used to indicate a preferential partition for class $g \in G$: it has the value 0 when there is no preference, otherwise it has a non-negative integer value identifying the desired partition. Let $\overline{G}$ be the set of the classes pre-assigned to a partition: $\overline{G} = \{g \in G | \pi_g^G > 0\}$

The set of teachers of subjects $s \in S$ could be primary teachers of set ($\tilde{T}_s \subseteq T$) or co-teachers ($\tilde{C}_s \subseteq C$). The parameter $\sigma_{gst}$ takes value 1 if the teacher $t \in T$ must be assigned to class $g \in G$ for the subject $s \in S$, 0 otherwise.

Each teacher $t \in T$ must belong to one partition at least. The parameter $\pi_t^T$ is used to indicate a preferential partition for teacher $t \in T$: it has the value 0 when there is no preference, otherwise it has a non-negative integer value identifying the desired partition. Let $\overline{T}$ be the set of the teachers preassigned to a partition: $\overline{T} = \{t \in T | \pi_t^T > 0\}$

In the Class Teacher Assignment Problem one must assign a teacher $t \in \tilde{T}_s$ and possibly a co-teacher $c \in \tilde{C}_s$ to each class$g \in G$ for each $s \in \tilde{S}_g$ of the subjects that the students of the class must follow in order to minimize:

1. the deviation from the desired number of partitions $\nu$ (this criterion is called DESIRED NUMBER OF PARTITIONS)
2. the difference between the maximum and the minimum number of classes in each partition (this criterion is called PARTITION BALANCING)
3. the weighted number of remaining connections between partitions (this criterion is called ALMOST SEPARATE PARTITIONS)
4. the maximum number of classes assigned to each teacher who must teach a subject, for all subjects (this criterion is called NUMBER OF CLASSES ASSIGNED TO A TEACHER)
5. the maximum number of class-groups assigned to each teacher who must teach a subject, for all subjects (this criterion is called NUMBER OF CLASS-GROUPS ASSIGNED TO A TEACHER)
6. the maximum number of different subjects assigned to each teacher (this criterion is called NUMBER OF SUBJECTS ASSIGNED TO A TEACHER)
7. the number of different class-groups assigned to each partition (this criterion is called NUMBER OF CLASS-GROUPS ASSIGNED TO A PARTITION)

For each of the previous criteria a non-negative integer weight must be selected. The value 0 is taken if corresponding criterion is not to be used.

### 3.2.2 A Mathematical model for the CTAP

Decision variables:
$x_{gst}$ takes value 1 when teacher $t \in T$ is assigned to teach the subject $s \in S$ in class $g \in G$, 0 otherwise;
$y_{gp}$ takes value 1 when class $g \in G$ belongs to partition $p \in P$, 0 otherwise;
$z_{tp}$ takes value 1 when teacher $t \in T$ belongs to partition $p \in P$, 0 otherwise;
$w_p$ takes value 1 when partition $p \in P$ is not empty, 0 otherwise.

Secondary and slack variables:

$b^{S4}_{gt}$ takes value 1 when the class $g \in G$ is assigned to teacher $t \in T$; ;

$b^{S5}_{\hat{c}t}$ has the value 1 when at least one class of class-group $\hat{c} \in \hat{C}$ is assigned to teacher $t \in T$ ;
$b^{S6}_{st}$ takes value 1 when at least one class is assigned to teacher $t \in T$ for teaching a given subject $s \in S$;
$b^{S7}_{p\hat{c}}$ takes value 1 when at least one class of class-group $\hat{c} \in \hat{C}$ belongs to partition $p \in P$.

Additional decision variables:

$a^{S1}$ is the absolute value of the difference between $\nu$ and the optimal number of partitions;

$a^{S2max}$ is the number of classes in the largest partition;

$a^{S2min}$ is the number of classes in the smallest partition; ;

$a^{S3}_{pqgst}$ has the value 1 when the teacher $t \in T$ is assigned to class $g \in G$ for teaching subject $s \in S$ and $g, t$ belongs to two different partitions $p, q \in P$, 0 otherwise;

$a^{S4max}_s$ is the maximum number of classes assigned to teachers for a given subject $s \in S$;

$a^{S4min}_s$ is the minimum number of classes assigned to teachers for a given subject $s \in S$;

$a^{S5}_t$ is the number of class-groups assigned to teacher $t \in T$;

$a^{S6}_t$ is the number of different subjects assigned to teacher $t \in T$;

$a^{S7}_p$ is the number of different class-groups assigned to a partition $p \in P$.

**Objective function**

$$min\ \omega_1 a^{S1} + \omega_2 (a^{S2max} - a^{S2min}) + \omega_3 \sum_{p\in P} \sum_{q\in P\setminus\{p\}} \sum_{g\in G} \sum_{s\in \tilde{S}_g} \sum_{t\in \tilde{T}_s} \lambda_{gs} \cdot a^{S3}_{pqgst} +$$

$$+\omega_4 \sum_{s\in S} (a^{S4max}_s - a^{S4min}_s) + \omega_5 \sum_{t\in T} a^{S5}_t + \omega_6 \sum_{t\in T} a^{S6}_t + \omega_7 \sum_{p\in P} a^{S7}_p$$

where $\omega_i$ $(i = 1, ..., 7)$ are non-negative integer weights.

***Hard constraints***

$HC_1$ - Symmetry breaking constraint: if some partition is not used (not enabled) its identification index must follow those of enabled partitions. Thus, all equivalent partitions between them can be deleted. Example: if only two (of three) partitions are needed $(1, 1, 0)$ is equivalent to $(0, 1, 1)$ or $(1, 0, 1)$.

$$w_p \le w_{p-1} \qquad \forall p \in P \setminus \{1\} \tag{3.1}$$

$HC_2$ - Classes and teachers feasibility: all classes must be used and all teachers must be assigned to at least one class.

$$\sum_{p\in P} \sum_{g\in G} y_{gp} = |G| \tag{3.2}$$

$$\sum_{p\in P} \sum_{t\in T} z_{tp} \ge |T| \tag{3.3}$$

$HC_3$ - Classes' subjects weekly hours balance: all classes must follow their subjects for the number of hours per week established by their curriculum.

$$\sum_{t \in \tilde{T}_s \cup \tilde{C}_s} \lambda_{gs} x_{gst} = \mu_s \lambda_{gs} \qquad \forall g \in G, \forall s \in \tilde{S}_g \tag{3.4}$$

$HC_4$ - Teachers' weekly hours balance: all teachers must work the number of hours per week established by their contract.

$$\sum_{g \in G} \sum_{s \in \tilde{S}_g} \lambda_{gs} x_{gst} = \theta_t \qquad \forall t \in T \tag{3.5}$$

$HC_5$ - Pre-assigned partitions: the pre-assigned classes and teachers must belong to the given partition.

$$y_{g\pi_g^G} = 1 \qquad \forall g \in \overline{G} \tag{3.6}$$

$$z_{t\pi_t^T} = 1 \qquad \forall t \in \overline{T} \tag{3.7}$$

$HC_6$ - Partition activation for a class:

$$w_p \geq y_{gp} \qquad \forall p \in P, \forall g \in G \tag{3.8}$$

$HC_7$ - Partition activation for a teacher:

$$w_p \geq z_{tp} \qquad \forall p \in P, \forall t \in T \tag{3.9}$$

$HC_8$ - Upper bound for variable $w_p$:

$$w_p \leq \sum_{g \in G} y_{gp} \qquad \forall p \in P \tag{3.10}$$

$HC_9$ - Linking between variables:

$$y_{gp} + x_{gst} \leq 1 + z_{tp} \qquad \forall p \in P, \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \cup \tilde{C}_s \tag{3.11}$$

$HC_{10}$ - For a given class the same subject cannot be assigned to more teachers or co-teachers (i.e. no teachers' clash):

$$\sum_{t \in T} x_{gst} \leq 1 \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \tag{3.12}$$

$$\sum_{c \in C} x_{gsc} \leq 1 \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall c \in \tilde{C}_s \tag{3.13}$$

$HC_{11}$ - Class assignments to teachers:

$$g_{gt}^T \geq x_{gst} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \cup \tilde{C}_s \tag{3.14}$$

$HC_{12}$ - Pre-assigned class-teachers must be met:

$$x_{gst} \geq \sigma_{gst} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \qquad (3.15)$$

$HC_{13}$ - Linked subjects: some subjects may have lab-lessons accordingly the same teacher $t \in T$ who is assigned to a class $g \in G$ for a subject $s' \in S$ must be assigned for the linked lab subject $s'' \in S$.

$$x_{gs't} = x_{gs''t} \qquad \forall g \in G, \forall s' \in \tilde{S}_g, \forall s'' \in L_{s'}, \forall t \in \tilde{T}_{s'} \cap \tilde{T}_{s''} \quad (3.16)$$

***Soft constraints***

$SC_1$ - The non-negative integer variable $a^{S1}$ quantifies the difference between desired and actual number of partitions:

$$\nu - \sum_{p \in P} w_p + a^{S1} \geq 0 \qquad (3.17)$$

$$\nu - \sum_{p \in P} w_p - a^{S1} \leq 0 \qquad (3.18)$$

$SC_2$ - The non-negative integer variables $a^{S2min}$ and $a^{S2max}$ are the lower and the upper bound of the cardinality of the smallest and the largest partition of classes, respectively:

$$a^{S2min} \leq \sum_{g \in G} y_{gp} + |G|(1 - w_p) \qquad \forall p \in P \quad (3.19)$$

$$a^{S2max} \geq \sum_{g \in G} y_{gp} - |G|(1 - w_p) \qquad \forall p \in P \quad (3.20)$$

The term $|G|(1 - w_p)$ is helpful when the partition $p \in P$ is not activated to prevent the estimate of $a^{S2min}$ and $a^{S2max}$ from being compromised.

$SC_3$ - The connection between class $g \in G$ and teacher $t \in T$ in different partitions $p, q \in P$ through subject $s \in S$ are indicated by the value 1 of the boolean variable $a^{S3}_{pqgst}$:

$$y_{gp} + z_{tq} + x_{gst} \leq 2 + a^{S3}_{pqgst} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s, \forall p \in P, \forall q \in P \setminus \{p\} \quad (3.21)$$

$a^{S3}_{pqgst}$ takes value 1 only when teacher $t \in \tilde{T}_s$ is assigned to teach the subject $s \in \tilde{S}_g$ in class $g \in G$ and teacher and class are in different partitions.

$SC_4$ - $b^{S4}_{gt}$ is 1 if the class $g \in G$ is assigned to teacher $t \in T$, 0 otherwise. $a^{S4max}_s$ and $a^{S4min}_s$ are the maximum and minimum number of classes assigned to teachers for subject $s \in S$:

$$b^{S4}_{gt} \geq x_{gst} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \quad (3.22)$$

$$a_s^{S4max} \geq \sum_{g \in G} b_{gt}^{S4} \qquad \forall s \in S, \forall t \in \tilde{T}_s \quad (3.23)$$

$$a_s^{S4min} \leq \sum_{g \in G} b_{gt}^{S4} \qquad \forall s \in S, \forall t \in \tilde{T}_s \quad (3.24)$$

$SC_5$ - $b_{\hat{c}t}^{S5}$ is 1 if the teacher $t \in T$ is assigned to the class-group $\hat{c} \in \hat{C}$; $a_t^{S5}$ is the number of different class-groups assigned to teacher $t \in T$:

$$b_{\hat{c}t}^{S5} \geq x_{gst} \qquad \forall \hat{c} \in \hat{C}, \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \quad (3.25)$$

$$a_t^{S5} = \sum_{\hat{c} \in \hat{C}} b_{\hat{c}t}^{S5} \qquad \forall t \in T \quad (3.26)$$

$SC_6$ - $b_{st}^{S6}$ is 1 if the subject $s \in S$ is assigned to teacher $t \in T$. $a_t^{S6}$ is the number of different subjects assigned to teacher $t \in T$

$$b_{st}^{S6} \geq x_{gst} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \quad (3.27)$$

$$a_t^{S6} = \sum_{s \in S} b_{st}^{S6} \qquad \forall t \in T \quad (3.28)$$

$SC_7$ - $b_{p\hat{c}}^{S7}$ is 1 if at least one class of the class-group $\hat{c} \in \hat{C}$ is assigned to partition $p \in P$. $a_t^{S7}$ is the number of different class-groups assigned to teacher $t \in T$

$$1 + b_{p\hat{c}}^{S7} \geq x_{gst} + y_{gp} \qquad \forall g \in G, \forall s \in \tilde{S}_g, \forall t \in \tilde{T}_s \quad (3.29)$$

$$a_p^{S7} = \sum_{\hat{c} \in \hat{C}} b_{p\hat{c}}^{S7} \qquad \forall p \in P \quad (3.30)$$

## 3.3 Experimental settings and results

In the Class Teacher traditional assignment there is no need to create partitions of classes and teachers why the only objective is to satisfy the various requirements as much as possible. Instead in the Class Teacher partitioning assignment an additional objective is to partition the set of classes and reduce as much as possible the set of teachers assigned to multiple partitions.

This experimentation has two objectives:

- compare the application on some schools of Class Teacher traditional assignment versus Class Teacher partitioning assignment using a variable number of partitions

- compare the total time (assignment plus timetabling) to solve the tested instances

In Table 3.1 the results of the experimentation on the mathematical formulation for the Class Teacher Assignment are summarized. Data from 5 different schools were used: 3 are realistic and were created taking into account the real constraints for the subjects in the study courses of the high-school in Italy; the other 2 come from real middle schools (their data have not been modified in any way).

For each of these 5 schools it proceeded in this way: in the first instance of the experiment, a class-teacher assignment was performed without requiring to partition the classes (number of partitions required = 1); next it was requested to divide the problem into several separable parts (partitions).

| #Instance | Source | #Partitions | #Classes | #Teachers | Shared teachers | CTAP time | IHSTT time1 | IHSTT time2 | IHSTT time3 | IHSTT time4 | Total time | Total time P | Function | Idle times | Time improvement | Obj improvement |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Realistic #1 | 1 | 6 | 26 | 0 | 0.5 | 393.7 | | | | 394.2 | 394.2 | 0 | 0.0 | - | - |
| 2 | Realistic #1 | 2 | 3+3 | 13+13 | 0 | 0.8 | 33.5 | 35.0 | | | **69.3** | **35.8** | 0 | 0.0 | 469% | 0% |
| 3 | Realistic #2 | 1 | 9 | 26 | 0 | 1.5 | 1,315.1 | | | | 1,316.6 | 1,316.6 | 0 | 0.0 | - | - |
| 4 | Realistic #2 | 3 | 3+3+3 | 13+13+13 | 0 | 24.0 | 33.9 | 36.7 | 39.6 | | **134.2** | **63.6** | 0 | 0.0 | 881% | 0% |
| 5 | Realistic #3 | 1 | 18 | 50 | 0 | 182.8 | 21,600.0 | | | | 21,782.8 | 21,782.8 | 10,000 | 2.0 | - | - |
| 6 | Realistic #3 | 2 | 9+9 | 25+25 | 0 | 185.4 | 10,800.0 | 10,376.0 | | | **21,361.4** | **10,985.4** | 4,800 | 1.0 | 2% | 108% |
| 7 | School #1 | 1 | 23 | 44 | 0 | 3,600.5 | 10,800.0 | | | | 14,400.5 | 14,400.5 | 16,000 | 1.8 | - | - |
| 8 | School #1 | 2 | 18+5 | 31+13 | 0 | 53.5 | 7,200.0 | 1,200.0 | | | 8,453.5 | 7,253.5 | 4,600 | 1.0 | 70% | 72% |
| 9 | School #1 | 3 | 9+5+9 | 16+13+17 | 2 | 3608.7 | 2,400.2 | 701.6 | 2,400.3 | | 9,110.8 | 6,009.0 | 1,500 | 0.3 | 58% | 428% |
| 10 | School #1 | 4 | 6+5+6+6 | 14+13+17+14 | 13 | 3642.8 | 552.9 | 635.8 | 456.1 | 943.8 | **6,231.4** | **4,586.6** | 0 | 0.0 | 131% | ∞% |
| 11 | School #2 | 1 | 24 | 44 | 0 | 2,132.5 | 10,800.0 | | | | 12,932.5 | 12,932.5 | 24,800 | 3.4 | - | - |
| 12 | School #2 | 2 | 6+18 | 13+32 | 1 | 3,690.2 | 1,200.0 | 7,200.0 | | | 12,090.2 | 10,890.2 | 8,100 | 1.8 | 19% | 85% |
| 13 | School #2 | 3 | 6+6+9 | 13+17+18 | 4 | 3,758.5 | 1,200.0 | 2,400.0 | 2,400.0 | | 9,758.5 | 6,558.5 | 5,900 | 1.3 | 48% | 154% |
| 14 | School #2 | 4 | 6+6+6+6 | 12+15+17+15 | 14 | 3,983.5 | 1,200.0 | 1,200.0 | 1,200.0 | 1,200.0 | **8,783.5** | **5,183.5** | 3,100 | 0.7 | 64% | 383% |

Table 3.1: Experimentation results (all time are expressed in seconds)

The outcomes are reported in Table 3.1. In the first column (#Instance) there is the unique identification number of the instance. In the second column (Source) there is a reference to the source of the data, i.e. the school from which the data necessary to assign the chairs were taken. In the third column (#Partitions) one can read the number of partitions requested (and obtained). In the fourth column (#Classes) the number of classes contained in each partition is reported (in order with respect to the partition number). In the fifth column (#Teachers) one reads the corresponding number of teachers per partition. In the sixth column (#Shared teachers) the number of teachers shared between different partitions is shown. In column seven (CTAP time) one can read the time (in seconds) relating to the run of the standard MIP Solver (CPLEX 12.10) to complete the CTAP phase. After the CTAP phase, timetabling was completed using the MIP Solver applied to the IHSTT model described in the chapter 2. In the next 4 columns 8, 9, 10, 11 (IHSTT time1 ... IHSTT time4) the completion times of the IHSTT timetabling phase for each of the previously generated partitions were reported. Column 12 (Total time) shows the cumulative time of the CTAP phase and the consequent IHSTT phase(s) for each instance. Column 13 (Total time P) indicates the sum of the CTAP times and the time of the longest IHSTT phase to simulate the execution of the entire process on a parallel machine with an adequate number of independent real computers available. Column 14 (Function) reports the final value obtained for the objective function for IHSTT (aggregating by sum the contributions relating to any partitions). Column 15 (Idle times) reports the average value of idle times for teachers. Finally, the Time improvement and Obj improvement utility columns show the percentage improvement value obtained respectively for the elapsed time and for the objective function of the partitioned instances compared to the corresponding non-partitioned instance.

As can be easily seen, in the groups of realistic instances there is a notable improvement in execution times and, when this does not occur (due to the time limit having been reached), however, a notable improvement in the objective function is highlighted.

In the group of real instances there is always a significant improvement in execution times and an even more marked improvement in the objective function.

Within the limits of the experimentation and the overall time used for it (about 3 months) it can be stated that the technique which uses an IHSTT timetabling preceded by a CTAP phase is very competitive in terms of execution times, particularly in terms of quality of optimization.

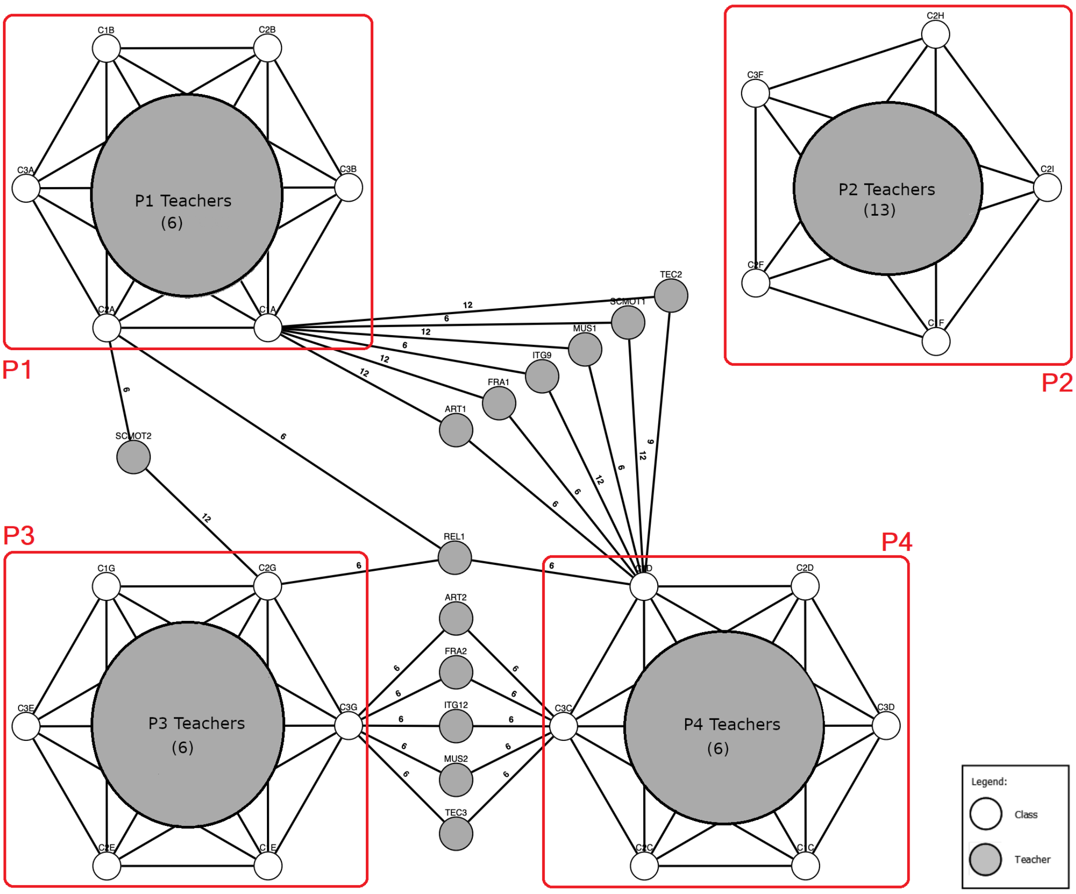



Figure 3.3: Instance #10 (4 partitions)

In figure 3.3 shown on instance #10 (4 partitions), it is also possible to note the set of 13 teachers who give lessons to classes located in at least 2 partitions at different times (teacher REL1 even has classes in 3 distinct partitions). In the internal area of the partitions, for greater simplicity, the teachers within each partition have not been reported, because they are not considered particularly significant. Furthermore, the classes belonging to the partitions were represented as cliques because they actually share the same conditions (same teachers).

From figure 3.3 one can see how the CTAP and the Multi-School Timetabling Problem are connected to each other. In particular, by replacing the partition entity with the school entity and the teacher shared between multiple partitions with the teacher shared between multiple schools, one infer that the Multi-School Timetabling Problem is modeled like the CTAP with all the consequences from a practical point of view.

# Chapter 4

# Conclusion and future developments

This thesis faced timetabling problems, which were motivated by the case of Italian high schools. More precisely, two problems were investigated: the Class Teacher Assignment Problem and the Italian High School Timetabling Problem.

The Class Teacher Assignment Problem has been modeled while accounting for all its useful features by an integer programming formulation. Its the most innovative feature is the possibility to create partitions of classes and teachers, to solve separated timetabling problems on a bipartite graph. The experimentation shows that this partitioning can lead to the resolution of the timetabling problems in a shorter overall time with a better quality in solutions. The Multi-School Timetabling Problem is derived from Class Teacher Assignment Problem by a suitable substitution of some entities.

This thesis also investigated the Italian High School Timetabling Problem, once the assignment of teachers to classes is made. It has well-established characteristics like co-teachers, articulated classes, multiple lessons, additional days-off, as well as quality indicators, such as the horizontal and vertical distributions of lessons. However, it exhibits new features which have not been investigated so far: fractional time units, equity in idle times, avoidance of consecutive heavy days and excessive workload for classes. All in all, this problem is more complex than those in the literature on the class-teacher paradigm. Moreover, the generalized HST problem based on KSS [KSS15] [FSCS17] does not incorporate all requirements in the [IHSTP].

A mixed integer programming model (denoted by IHSTT) has been proposed for [IHSTP], in order to pursue the maximum compatibility with KSS. Since KSS timetabling is based on the decomposition into sub-events and the IHSTT is built on equally-sized sub-events, the larger cardinality of the sets of sub-events makes IHSTT a more suitable approach for solving also realistic size instances of a simplified problem, in which the new requirements are omitted.

In order to obtain fast solutions for both the complete and the simplified [IHSTP], a two-step method is proposed. In the first step, the TP model is solved to cleverly decrease the initial solution space of IHSTT and determine the profiles of teachers. Next, the IHSTT model with restricted data is solved very effectively in the second step. The two-step method results in good-shaped timetables and suitable computing times even for the most complex problem instances. Although the method does not guarantee the optimality, it returns the optimal solutions for all instances motivating this research. Moreover, the two-step method is quite general and could be applied to other class-teacher problems for countries with a similar problem setting.

Some possible research developments are listed below:

- investigating the applicability of the partitioning CTAP and TPP also to the [XHSTT] standard instances;
- allowing the IHSTT model to have a better compatibility with [XHSTT] online database [XHS];
- carrying out extensive experimentation related to the multiple school timetabling problem with some shared teachers among two or more schools;
- planning temporary timetables in which some teachers may not be available or they have to be substituted;
- planning timetables according to teachers' preferences.

# Acknowledgments

First of all, my thanks go to Professors Massimo Di Francesco and Enrico Gorgone who supported me during my doctorate, encouraging me in all the moments I was tired, helping me to achieve my goals.

I am also deeply grateful to Professor Antonio Frangioni for his excellent advice and Professor Sebastiano Seatzu who first taught me to love Mathematics.

Last but not least I would like to thank my beloved wife Antonietta and my dear daughter Federica for their invaluable support and patience during these long four years.

# Bibliography


[ADSV07] P. Avella, B. D'Auria, S. Salerno, and I. Vasil'ev. A computational study of local search algorithms for italian high school timetabling. *Journal of Heuristic*, 13:543 – 556, 2007.

[BDH97] T. Birbas, S. Daskalaki, and E. Housos. Timetabling for greek high schools. *The Journal of the Operational Research Society*, 48(12):1191 – 1200, 1997.

[BDH09] T. Birbas, S. Daskalaki, and E. Housos. School timetabling for quality student and teacher schedules. *Journal of Scheduling*, 2009.

[BFT+12] Samuel S Brito, George HG Fonseca, Túlio AM Toffolo, Haroldo G Santos, and Marcone JF Souza. A SA-VNS approach for the high school timetabling problem. *Electronic Notes in Discrete Mathematics*, 39:169–176, 2012.

[BSS18] Niels-Christian F. Bagger, Matias Sørensen, and Thomas R. Stidsen. Benders' decomposition for curriculum-based course timetabling. *Computers & Operations Research*, 91:178–189, 2018.

[CDS23] Sara Ceschia, Luca Di Gaspero, and Andrea Schaerf. Educational timetabling: Problems, benchmarks, and state-of-the-art results. *European Journal of Operational Research*, 308(1):1–18, 2023.

[DAN] Danish School site. `https://take-the-leap.wep.org/it/come-funziona-il-sistema-scolastico-danese/`. Accessed: 2023-10-8.

[DdAB12] A.P. Dorneles, O.C.B. de Araujo, and L.S. Buriol. The impact of compactness requirements on the resolution of high school timetabling problem. In *Congreso Latino-Iberoamericano de Investigacion Operativa - Rio de Janeiro, Brazil 2012 September 24-28*, volume Conference Paper, 2012.

[DdAB14] A.P. Dorneles, O.C.B. de Araujo, and L.S. Buriol. A fix-and-optimize heuristic for the high school timetabling problem. *Computers & Operations Research*, 52:29 – 38, 2014.

[DdAB16] A.P. Dorneles, O.C.B. de Araujo, and L.S. Buriol. A column generation approach to high school timetabling modeled as a multicommodity flow problem. *European Journal of Operational Research*, 256:685 – 695, 2016.

[DH12] J. Domros and J. Homberger. An evolutionary algorithm for high school timetabling. *Proceedings of the Ninth International Conference on the Practice and Theory of Automated Timetabling (PATAT 2012)*, pages 485–488, 01 2012.

[FSCS17] G.H.G. Fonseca, H.G. Santos, E.G. Carrano, and T.J.R. Stidsen. Integer programming techniques for educational timetabling. *European Journal of Operational Research*, 262:28 – 39, 2017.

[HC97] Tim H Hultberg and Domingos M Cardoso. The teacher assignment problem: A special case of the fixed charge transportation problem. *European Journal of Operational Research*, 101(3):463–473, 1997.

[ITC] ITC2011 site. https://www.utwente.nl/en/eemcs/dmmp/hstt/itc2011/results/round2/. Accessed: 2022-01-05.

[KHE] KHE site (a software library for high school timetabling and nurse rostering). http://jeffreykingston.id.au/khe/. Accessed: 2021-04-07.

[Kin14] J. Kingston. KHE14: An algorithm for high school timetabling. In *Proceedings of the 10th International Conference of the Practice and Theory of Automated Timetabling - York, United Kingdom 2014 August 26-29*, volume Conference Paper, pages 269–291, 2014.

[KOP12] Ahmed Kheiri, Ender Ozcan, and Andrew Parkes. Hysst: hyper-heuristic search strategies and timetabling. In *Proceedings of the Ninth International Conference on the Practice and Theory of Automated Timetabling (PATAT 2012)*, 01 2012.

[KS13] S. Kristiansen and T.R. Stidsen. A comprehensive study of educational timetabling, a survey. *DTU Management Engineering*, Technical Report 8, 2013.

[KSS15] S. Kristiansen, M. Sorensen, and T.R. Stidsen. Integer programming for the generalized high school timetabling problem. *Journal of Scheduling*, 18(4):377 – 392, 2015.

[MIM] Ministero dell'Istruzione e del Merito site - principali dati della scuola - avvio anno scolastico 2023/2024. https://www.miur.gov.it/web/guest/-/

focus-principali-dati-della-scuola-avvio-anno-scolastico-2023-2024-. Accessed: 2023-10-10.

[MIU] Ministero Pubblica Istruzione site. https://archivio.pubblica.istruzione.it/riforma_superiori/nuovesuperiori/index.html#regolamenti. Accessed: 2021-04-14.

[N.14] Pillay N. An overview of school timetabling research. *Annals of Operational Research*, 218:261 – 293, 2014.

[Obe16] I. Ober. A variant of the high-school timetabling problem and a software solution for it based on integer linear programming. In *PATAT-2016 Udine, Italy 2016 August 23-26*, volume Conference Paper, pages 283 – 294, 2016.

[PDGK+13] Gerhard Post, Luca Di Gaspero, Jeffrey Kingston, Barry Mccollum, and Andrea Schaerf. The third international timetabling competition. *Annals of Operations Research*, 239, 02 2013.

[PKA+14] Gerhard Post, Jeffrey H Kingston, Samad Ahmadi, Sophia Daskalaki, Christos Gogos, Jari Kyngas, Cimmo Nurmi, Nysret Musliu, Nelishia Pillay, Haroldo Santos, et al. Xhstt: an xml archive for high school timetabling problems in different countries. *Annals of Operations Research*, 218(1):295–301, 2014.

[PVH03] K. Papoutsis, C. Valouxis, and E. Housos. A column generation approach for the timetabling problem of greek high schools. *Journal of Operational Research Society*, 54:230 – 238, 2003.

[RTR15] S. Ribic, R. Turcinhozic, and A.M. Ribic. Modelling constraints in school timetabling using integer linear programming. In *Proceedings of the 2015 XXV International Conference on Information, Communication and Automation Technologies (ICAT), 2015 October 29-31*, volume Conference Paper, pages 1–6, 2015.

[SC17] L. Saviniec and A.A. Costantino. Effective local search algorithms for high school timetabling problems. *Applied Soft Computing*, 60:363 – 373, 2017.

[Sch99] A. Schaerf. Local search techniques for large high school timetabling problems. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 29(4):368–377, 1999.

[SD14] M. Sorensen and F.H.W. Dahms. A two-stage decomposition of high school timetabling applied to cases in denmark. *Computers & Operations Research*, 43:36 – 49, 2014.

[SKS+12] Matias Sorensen, Simon Kristiansen, Thomas R Stidsen, et al. International timetabling competition 2011: An adaptive large neighborhood search algorithm. In *Proceedings of the ninth international conference on the practice and theory of automated timetabling (PATAT 2012)*, page 489, 2012.

[SS12] M. Sorensen and T.R. Stidsen. High school timetabling - modeling and solving a large number of cases in denmark. *PATAT-2012 Son, Norway 2012 August 28-31*, Conference Paper:359 – 364, 2012.

[SSC18] L. Saviniec, M.O. Santos, and A.M. Costa. Parallel local search algorithms for high school timetabling problems. *European Journal of Operational Research*, 265:81 – 98, 2018.

[SSCdS20] L. Saviniec, M.O. Santos, A.M. Costa, and L.M.R. dos Santos. Pattern-based models and a cooperative parallel metaheuristic for high school timetabling problems. *European Journal of Operational Research*, 280:1064 – 1081, 2020.

[SUOM12] H.G. Santos, E. Uchoa, L.S. Ochi, and N. Maculan. Strong bounds with cut and column generation for class-teacher timetabling. *Annals of Operational Research*, 194:399 – 412, 2012.

[TEC] Tecnica della Scuola site. https://www.tecnicadellascuola.it/montessori-nel-mondo-quante-sono-e-dove-sono-le-scuole. Accessed: 2023-10-8.

[TIB20] I.X. Tassopoulos, C.A. Ilioupoulou, and G.N. Beligiannis. Solving the greek school timetabling problem by a mixed integer programming model. *Journal of Operational Research Society*, 71:117 – 132, 2020.

[UNI] University of Turin site. https://politichediateneo.unito.it/ranking-internazionali/. Accessed: 2023-10-8.

[VGAH12] C. Valouxis, C. Gogos, P. Alefragis, and E. Housos. Decomposing the high school timetable problem. *PATAT-2012 Son, Norway 2012 August 28-31*, Conference Paper:209 – 221, 2012.

[XHS] XHSTT site. https://www.utwente.nl/en/eemcs/dmmp/hstt/. Accessed: 2021-04-07.