

Noname manuscript No. (will be inserted by the editor)
--

Optimal student sectioning on mandatory courses with various sections numbers

David Schindl

Received: date / Accepted: date

Abstract In sufficiently large schools, courses are given to classes in sections of various sizes. Consequently, classes have to be split into various given numbers of sections. We focus on how to dispatch the students into sections of equal size, so as to minimize the number of edges in the resulting conflict graph. As a main result, we show that subdividing the students set in a regular way is optimal. We then discuss our solution uniqueness and feasibility, as well as practical issues concerning teacher assignments to sections and the case of an additional course with unequal section sizes requirements.

Keywords student sectioning · timetabling · conflict graph

1 Introduction

The task of timetabling in its usual form consists in assigning timeslots to courses such that two courses are not assigned a common timeslot if they have a common teacher or student. This formulation implicitly needs the assumption that the courses are already formed, i.e. the set of students associated to each course is already known. This student to course assignment is indeed not to discuss when the whole class does not need to be split, or has to be split in some unique, specific way (boys vs girls, beginner vs advanced level, etc.).

However, there are a number of cases where this association is not known in advance, and has therefore to be defined. This task is known as *student sectioning* and has to be achieved in the two following most frequent situations. First, when there are optional (elective) courses, and the students have expressed preferences over the set of proposed courses, but will not be assigned to all chosen courses. In that case, the goal of student sectioning is usually to maximize overall student satisfaction, while keeping some balance between different courses and avoiding conflicts in the timetable. The second case is when the same course is given several

D. Schindl
Haute Ecole de Gestion de Genève
Tel.: +41-22-3886572
E-mail: david.schindl@hesge.ch

times in a week due to capacity restrictions. These occurrences of a same course are called *sections*. Notice that both the above cases can happen simultaneously: an optional course may be split into sections.

If student sectioning is needed, it may be achieved before, during or after timetabling¹. In the first case, the goal will be to find a sectioning which will facilitate the subsequent timetable creation, typically by avoiding as much as possible potential conflicts, i.e. pairs of sections with common students. In the two latter cases, special care should be taken when moving a student into a section, to avoid conflicts in the timetable of that student. As will be detailed in the next section, most papers on student sectioning deal with approaches where the timetable is known when it comes to sectioning.

In this paper, we concentrate on the task of dispatching students into balanced sections of mandatory courses (i.e. each student must be assigned to exactly one section of each course), prior to timetabling, the major difficulty being the fact that the number of sections may vary significantly between different courses. Numerous issues can imply a variety of section numbers. These can for instance be a limited number of large lecture rooms or computer rooms, pedagogical constraints restricting the number of students to some course, or simply budget. In this context, given a class with 120 students, one may need to create 2 large sections of 60 students for Economics ex cathedra courses, 3 sections of 40 students for Applied Mathematics, 4 sections of 30 students for Informatics and 5 sections of 24 students for English. These kind of issues are particularly likely to arise in Universities of Applied Sciences, where many courses are supposed to be practical and need to be taught only in groups of small to moderate sizes, while the total number of students in a class can reach several hundreds. This is the case at the Geneva School of Business Administration, and the solution we present here is applied each semester in several departments of the school. Our main contribution is to show that a natural way of sectioning, which we will call "regular", permits to obtain an *a priori* optimal sectioning, in the sense that the number of pairs of potentially conflicting courses is minimized. We point out that our result solves the stated problem exactly.

The paper is organized as follows. Section 2 consists of a review of the literature related to student sectioning. Some terminology and a precise formulation of our problem are given in Section 3 and its solution and proof are presented in Section 4. We follow with a discussion in Section 5 and conclude with some recommendations and open questions.

2 Related literature

We first review contributions that are most related to the present work, with papers dealing with theoretical properties of student sectioning. The second part reviews contributions with heuristic approaches. We refer the interested reader to (Dostert et al., 2016) for a larger list of references on student sectioning.

¹ Although the word "timetabling" usually refers to the whole process of creating a timetable for each student and each teacher, in this paper we call timetabling the task of assigning timeslots to sections, as a problem separate from student sectioning, consisting in assigning students to sections.

In (Beyrouthy et al., 2008), an experimental study is reported about the impact on space utilization and timetabling feasibility of varying the maximum size and number of sections. They show that the latter parameter has the most impact on feasibility and observe that allowing schedule changes while assigning students has a significant impact on feasibility. A similar issue is handled with a graph theoretical approach in (Selim, 1988), where the author discusses on how many vertex splits permit to decrease the chromatic number, i.e. how many additional sections must be introduced in order to decrease the number of necessary timeslots to ensure feasibility. A lower bound is given in function of the total number of available courses and the number of courses to be chosen by each student. In (Dostert et al., 2016), the authors investigate the complexity statuses of fundamental versions of student sectioning. They show that the basic decision version of sectioning on a fixed timetable where each course uses exactly one timeslot is solvable in polynomial time with a network flow algorithm. On the other hand, they prove NP-completeness of three generalizations: when students attend only a subset of courses, when students have individual timeslots restrictions and when courses may use more than one timeslot. In a similar spirit, the authors of (ten Eikelder and Willemen, 2001) study both the complexities of basic versions of timetabling and student sectioning. They first show that the problem of assigning unique timeslots to sections (which is therefore a pure timetabling problem) with possibly some forbidden assignments changes from polynomial to NP-complete when some pairs of sections possibly require consecutive timeslots. They show then that the problem of student sectioning over a fixed timetable also becomes NP-complete by adding maximum section size constraints. The authors of (Cheng et al., 2003) consider the decision version of student sectioning where the timetable is fixed and part of the instance (no conflict is allowed). They show its NP-completeness even with a single student. They further propose a flow formulation, still for the single student case, which permits them in particular to show that the flow existence problem is NP-complete if there may be integral flow constraints. They finally discuss on how to model various aspects by modifying the objective function.

In (Carter, 2001), students have some mandatory courses and some optional ones, to choose from a list. Prior to timetabling, the author proposes to create an initial, called *homogeneous* sectioning by grouping students in clusters according to similarities between their optional course selections. The problem is then decomposed into smaller subproblems and greedy algorithms are run on each of them to assign timeslots to courses. After timetabling, some assignments are modified using an alternating tree approach, in order to decrease conflicts in students schedules. The authors of (Murray et al., 2007) present and discuss several practical issues in creating the timetable of a large university. The student sectioning part is managed in two phases: an initial student sectioning is created in the same homogeneous way as above and once the timetable is settled, the number of conflicts is lowered via a local search over the student to sections assignments. In (Müller and Murray, 2010), the same authors (except one) focus on student sectioning and propose a deeper approach. Students make choices over optional courses, which may be proposed in several sections. Sectioning is achieved in three phases. An initial homogeneous (as above) sectioning is first used only to guide the timetable construction towards timetables tending to admit satisfactory conflict-free assignments of students to courses. A second phase is run once the timetable is fixed, and consists of a variable neighborhood search on the set of not necessarily complete

student to sections assignments. The third phase allows student to remove themselves from courses or to subscribe on waiting lists for other courses in an online manner. If possible, assignments and waiting lists are automatically updated by possibly re-sectioning enrolled students. In (Hertz, 1991), the author studies the timetabling and student sectioning problem of a university, where there are some large courses which have to be repeated in several sections during the week. He approaches the problem with two subsequent tabu searches. The first one has as search space the set of timetables. In this algorithm, two sections of a same course are allowed to be scheduled at the same time, but with a penalty due to possible conflicts of common students. In the second search, the timetable is fixed and the search space is the set of student assignments to courses. Only students exchanges are allowed between different sections of a course, in order to try to diminish the number of conflicts in the student's schedules, while keeping the same sections sizes. In (Amintoosi and Haddadnia, 2005), student sectioning is achieved before timetabling. Students course selections must be respected, and as soft constraints, sections should be balanced, section sizes should belong to some range and the sectioning should be as homogeneous as possible in the sense of Carter (2001). To this aim, they represent students as vectors indicating their courses choices, and use a fuzzy data clustering algorithm. In (Alvarez-Valdes et al., 2000), the course timetable is fixed and revealed to students before they select courses. All selections must be respected, but there are several sections per course. A sectioning algorithm is then run, which consists of the two following phases. First, the sectioning problem for each student is modelled as a maximum independent set problem where vertices are sections and edges represent two sections the student cannot be assigned to (either because they belong to the same course, or because they have a timeslot in common). An enumerative procedure permits to store the best timetables for each student with respect to secondary criteria including balance of section enrollments, section maximum capacities and compact timetable. These timetables are then variables for another problem with the soft constraints expressed in a Lagrangian. This last problem is solved with a tabu search.

3 Terminology and problem formulation

We define here first precisely the terms we are going to use frequently in this paper. We follow the terminology used in Carter (2001). A *class* is a set of students needing all to follow the same set of courses. In our context, it is the set of all students. A *course* is a subject to be taught to all students, like Economics. Each course has to be taught a given number of times, called *sections*. A section may also refer to the set of students composing it and each student belongs to exactly one section of a given course. There are as many sections as the sum over all courses of the number of times it is taught. Students that belong to the same section of a given course do not necessarily for another course. We call a *division* a set of students, maximal under inclusion, belonging to the same section for each course. Each student belongs to exactly one division, and each section can be obtained by merging some divisions. We are concerned by the way of forming the sections, i.e. dispatching the students into sections, for each course.

Consider a class with k mandatory courses to follow. Each course i , with $1 \leq i \leq k$, is taught to n_i sections of equal sizes. Consequently, for each i , our class

has to be split into n_i sections, each containing the fraction $\frac{1}{n_i}$ of the class. Notice that since we are working with fractions of classes, we are not concerned anymore with rounding issues. We thus now require sizes of sections to be exactly equal (not to within 1). This is a reasonable simplification that does not really impact the applicability of our result.

Section j of course i will be denoted by $s_{i,j}$, $1 \leq i \leq k$, $1 \leq j \leq n_i$. Duration of each course will be denoted d_i and may differ from one course to the other. Each section $s_{i,j}$ has to be assigned a fraction $\frac{1}{n_i}$ of the class and a set of d_i timeslots.

We define for each course $i \in \{1, \dots, k\}$ a function

$$s_i : [0, 1] \rightarrow \{1, \dots, n_i\}$$

Each function s_i represents the students assignment to one of the n_i sections of course i and will be called the *sectioning* i . For instance, if $s_3(x) = 4 \forall x \in [0.4, 0.45[$, it means that according to sectioning s_3 , each student ranked (for instance according to alphabetical order) between 40% and 45% belongs to the 4th section of the 3rd course. Using this notation, the 4th section of the 3rd course is given by $s_3^{-1}(4)$. The set of sectionings together will be simply called a *sectioning set*. Figure 1 is a schematic view of the following sectioning set $S = \{s_1, s_2, s_3\}$ with $k = 3$ courses, $n_1 = 2$, $n_2 = 3$ and $n_3 = 5$.

$$s_1(x) = \begin{cases} 1 & \text{if } x \in [0, 0.3[\cup [0.6, 0.8[\\ 2 & \text{if } x \in [0.3, 0.6[\cup [0.8, 1] \end{cases}$$

$$s_2(x) = \begin{cases} 1 & \text{if } x \in [0, 0.2[\cup [0.867, 1[\\ 2 & \text{if } x \in [0.2, 0.533[\\ 3 & \text{if } x \in [0.533, 0.867] \end{cases}$$

$$s_3(x) = \begin{cases} 1 & \text{if } x \in [0, 0.1[\cup [0.45, 0.55[\\ 2 & \text{if } x \in [0.1, 0.2[\cup [0.9, 1] \\ 3 & \text{if } x \in [0.2, 0.4[\\ 4 & \text{if } x \in [0.4, 0.45] \cup [0.55, 0.7[\\ 5 & \text{if } x \in [0.7, 0.9] \end{cases}$$

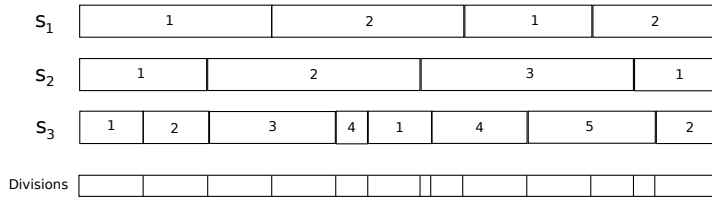


Fig. 1 Sectioning set example with $k = 3$ courses, $n_1 = 2$, $n_2 = 3$ and $n_3 = 5$.

For the timeslots assignment, we define the following function which we will call a *timetable* :

$$T : A \rightarrow \mathcal{P}(P)$$

such that $|T(s_{i,j})| = d_i, \forall i$, where A is the set of sections and P is the set of time periods or timeslots. For instance, if $T(s_{3,4}) = \{5, 8, 9\}$, the fourth section of course 3 is scheduled on timeslots 5, 8 and 9.

In a feasible timetable, two sections sharing a timeslot should not share a student. We consequently define potential conflict graphs of the two following kinds on the set of sections. Given a sectioning set $S = \{s_1, \dots, s_k\}$, two vertices $s_{i,j}$ and $s_{i',j'}$ are adjacent in the *students conflict graph* $G_S = (A, E_S)$, if and only if $s_i^{-1}(j) \cap s_{i'}^{-1}(j') \neq \emptyset$. Notice that for each course i , the vertex set $A_i := \{s_{i,j} : j = 1, \dots, n_i\}$ is a stable set in G_S . Minimizing the number of pairs of conflicting sections is then equivalent to minimizing the number of edges in the student conflict graph. Given a timetable T , two vertices $s_{i,j}$ and $s_{i',j'}$ are adjacent in the *timeslots conflict graph* $G_T = (A, E_T)$, if and only if $T(s_{i,j}) \cap T(s_{i',j'}) \neq \emptyset$.

4 Optimal sectioning

Our main result tells how to choose a sectioning set S so as to minimize the number of edges of G_S , regardless of the timetable. Indeed, each edge in G_S represents a potential conflict, which will become real if it is also an edge in G_T . If sectioning is achieved before timetabling, the best we can do a priori is making $|E_S|$ as small as possible.

By $\text{GCD}(a, b)$ we denote the greatest common divisor of positive integers a and b and by $\text{LCM}(a, b)$ their least common multiple. These are linked by the formula $\text{GCD}(a, b)\text{LCM}(a, b) = ab$. For a graph $G = (V, E)$ and $V' \subseteq V$, we denote by $G[V']$ the subgraph of G induced by V' , that is to say the graph obtained from G by removing all vertices in $V \setminus V'$ and all edges with at least one endpoint in $V \setminus V'$.

Consider courses i and i' . Since the vertex sets A_i and $A_{i'}$ are stable in G_S for any sectioning set $S = \{s_1, \dots, s_k\}$, we have

$$\begin{aligned} \min_{s_1, \dots, s_k} |E(G_S)| &= \min_{s_1, \dots, s_k} \sum_{i, i' \in \{1, \dots, k\}, i < i'} |E(G_S[A_i \cup A_{i'}])| \\ &\geq \sum_{i, i' \in \{1, \dots, k\}, i < i'} \min_{s_i, s_{i'}} |E(G_S[A_i \cup A_{i'}])| \end{aligned}$$

Consequently, if we are able to find a sectioning set $S = \{s_1, \dots, s_k\}$ such that

$$|E(G_S[A_i \cup A_{i'}])| = \min_{s_i, s_{i'}} |E(G_S[A_i \cup A_{i'}])| \quad \forall \quad 1 \leq i < i' \leq k$$

we are done.

Next Proposition shows that it is actually the case, and the corresponding sectionings are the ones where all preimages $s_i^{-1}(j)$ are intervals. Figure 2 displays the optimal sectioning set corresponding to our example above.

Proposition 1 *If $s_i(x) = \lceil n_i x \rceil \forall i \in \{1, \dots, k\}$, then for each $i < i'$ we have*

$$\begin{aligned} |E(G_S[A_i \cup A_{i'}])| &= \min_{s_i, s_{i'}} |E(G_S[A_i \cup A_{i'}])| \\ &= n_i + n_{i'} - \text{GCD}(n_i, n_{i'}) \end{aligned}$$

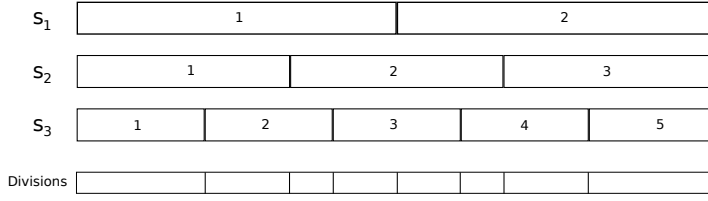


Fig. 2 Optimal sectioning with $k = 3$ courses, $n_1 = 2$, $n_2 = 3$ and $n_3 = 5$.

Proof We begin our proof with the following observations on properties we may assume on optimal sectionings s_i and $s_{i'}$.

Observation 1 There are optimal sectionings s_i and $s_{i'}$, with $s_i(x) = \lceil n_i x \rceil$.

Indeed, let \tilde{s}_i and $\tilde{s}_{i'}$ be optimal. Since $|\tilde{s}_i^{-1}(j)| = |s_i^{-1}(j)| = \frac{1}{n_i} \quad \forall j \in \{1, \dots, n_i\}$, there is a bijection σ from the $[0, 1]$ interval into itself, such that $\tilde{s}_i(\sigma(x)) = \lceil n_i x \rceil = s_i(x)$. Then choosing $s_{i'}(x) = \tilde{s}_{i'}(\sigma(x))$, s_i and $s_{i'}$ induce exactly the same conflict graph as \tilde{s}_i and $\tilde{s}_{i'}$, since $x \in s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ if and only if $\sigma(x) \in \tilde{s}_i^{-1}(j) \cap \tilde{s}_{i'}^{-1}(j')$. In particular, both conflict graphs have the same, minimal, number of edges. Since we restrict here only to two courses i and i' , the sets $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ are divisions.

From now on, we assume $s_i(x) = \lceil n_i x \rceil$.

Observation 2 There are optimal sectionings s_i and $s_{i'}$ such that the divisions $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ are intervals.

For each $1 \leq j \leq n_i$, an argument similar to the above, with σ being a bijection from the interval $[\frac{j-1}{n_i}, \frac{j}{n_i}[$ into itself, permits to assume that the divisions

$$s_i^{-1}(j) \cap s_{i'}^{-1}(j') = \left[\frac{j-1}{n_i}, \frac{j}{n_i} \left[\cap s_{i'}^{-1}(j'), \quad 1 \leq j' \leq n_{i'} \right.$$

are connected, and hence are intervals.

We may thus assume that each division $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ is an interval. Recalling that $|E(G[A_i \cup A_{i'}])|$ is equal to the number of sets $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ that are non empty, $|E(G[A_i \cup A_{i'}])|$ is simply equal to the number of divisions $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$.

We now show that $|E(G[A_i \cup A_{i'}])| \geq n_i + n_{i'} - \text{GCD}(n_i, n_{i'})$. Assume by contradiction that the number of divisions $s_i^{-1}(j) \cap s_{i'}^{-1}(j')$ is at most $n_i + n_{i'} - \text{GCD}(n_i, n_{i'}) - 1$. Of course, if there are, say p divisions, there are $p-1$ *separations*, i.e. pairs of adjacent (in $[0, 1]$) divisions (see Figures 1 and 2). This means that there are at most $n_i + n_{i'} - \text{GCD}(n_i, n_{i'}) - 2$ such separations. Further, exactly $n_i - 1$ of them are induced by the sections $s_i^{-1}(j)$. We call them *original*. Hence there are at most $n_{i'} - \text{GCD}(n_i, n_{i'}) - 1$ separations that are non original and exclusively induced by the sections $s_{i'}^{-1}(j')$. Let us call them *additional*.

Construct an auxiliary graph $H = (V_h, E_h)$ with $V_h = \{h_1, \dots, h_{n_{i'}}\} = \{s_{i'}^{-1}(j') : 1 \leq j' \leq n_{i'}\}$ and two vertices $h_{j'_1}$ and $h_{j'_2}$ being adjacent in H if and only if there is at least one pair of divisions $s_i^{-1}(j_1) \cap s_{i'}^{-1}(j'_1)$ and $s_i^{-1}(j_1) \cap s_{i'}^{-1}(j'_2)$ sharing an additional separation. The number of connected components of H is at least $|V_h| - |E_h| = n_{i'} - (n_{i'} - \text{GCD}(n_i, n_{i'}) - 1) = \text{GCD}(n_i, n_{i'}) + 1$. Consider the smallest one, and denote its number of vertices by q . Clearly, $q \leq \left\lfloor \frac{n_{i'}}{\text{GCD}(n_i, n_{i'}) + 1} \right\rfloor$. Observe

that for each connected component in G_S , the corresponding union of divisions in $[0, 1]$ is only delimited by original separations. As a consequence, the sum of its division lengths must be a multiple of $\frac{1}{n_i}$, say $\frac{a}{n_i}$, $a \in \mathbb{N}^*$:

$$\frac{a}{n_i} = q \cdot \frac{1}{n_{i'}} \Leftrightarrow an_{i'} = qn_i$$

So $an_{i'}$ is a multiple of both n_i and $n_{i'}$, thus $an_{i'} \geq \text{LCM}(n_i, n_{i'})$. On the other hand,

$$\begin{aligned} \frac{a}{n_i} &= q \cdot \frac{1}{n_{i'}} \leq \left\lfloor \frac{n_{i'}}{\text{GCD}(n_i, n_{i'}) + 1} \right\rfloor \cdot \frac{1}{n_{i'}} \\ &< \frac{n_{i'}}{\text{GCD}(n_i, n_{i'})} \cdot \frac{1}{n_{i'}} = \frac{1}{\text{GCD}(n_i, n_{i'})} \\ &= \frac{\text{LCM}(n_i, n_{i'})}{n_i n_{i'}} \end{aligned}$$

Hence $an_{i'} < \text{LCM}(n_i, n_{i'})$, a contradiction.

To show that for $s_i(x) = \lceil n_i x \rceil$ and $s_{i'}(x) = \lceil n_{i'} x \rceil$ we have $|E(G[A_i \cup A_{i'}])| = n_i + n_{i'} - \text{GCD}(n_i, n_{i'})$, we only need to observe that the points in $[0, 1]$ that are both original and additional separations are

$$a \frac{\text{LCM}(n_i, n_{i'})}{n_i n_{i'}}, \quad a \in \{1, \dots, \text{GCD}(n_i, n_{i'}) - 1\}$$

Indeed, separations that s_i and $s_{i'}$ have in common are the multiples of $\frac{1}{n_{i'}} = \frac{n_i}{n_i n_{i'}}$ and $\frac{1}{n_i} = \frac{n_{i'}}{n_i n_{i'}}$. So the total number of separations is

$$(n_i - 1) + (n_{i'} - 1) - (\text{GCD}(n_i, n_{i'}) - 1) = n_i + n_{i'} - \text{GCD}(n_i, n_{i'}) - 1$$

and the number of edges of the conflict graph, which is equal to the number of divisions, is $n_i + n_{i'} - \text{GCD}(n_i, n_{i'})$.

In the sequel, we call *regular* the sectioning set of Proposition 1, i.e. $S = \{s_1, \dots, s_k\}$, with $s_i(x) = \lceil n_i x \rceil$, $i = 1, \dots, k$. Notice that unsurprisingly, regular sectioning follows the same principle as homogeneous sectioning proposed in (Carter, 2001), in the sense that if two students are involved in the same sections for courses 1 to $k-1$, it means that they are ranked closely and are therefore likely to be also involved in the same section for course k .

5 Discussion

5.1 Section numbers

Proposition 1 tells us that the minimum number of edges in the conflict graph is equal to

$$\begin{aligned} \min_{s_1, \dots, s_k} |E(G_S)| &= \sum_{i, i' \in \{1, \dots, k\}, i < i'} n_i + n_{i'} - \text{GCD}(n_i, n_{i'}) \\ &= (k-1) \sum_{i=1}^k n_i - \sum_{i, i' \in \{1, \dots, k\}, i < i'} \text{GCD}(n_i, n_{i'}) \end{aligned}$$

Hence, apart from sectioning regularly, revising section numbers (if possible) may also permit to diminish the number of potential conflicts. The first part $((k-1) \sum_{i=1}^k n_i)$ is simply the total number of sections. Of course, by creating less sections, we will tend to have less potential conflicts, but this is not helping. The second part tells us we should give preference to section numbers which have large common divisors, which is intuitively sound. For instance, if $n_1 = 4$, choosing $n_2 = 7$ would produce $4 + 7 - \text{GCD}(4, 7) = 11 - 1 = 10$ potential conflicts, whereas choosing $n_2 = 6$, or $n_2 = 8$ (depending on practical constraints) would both produce 8 potential conflicts.

5.2 Uniqueness

Observe that if there are only two different section numbers n_1 and n_2 , other sectioning sets may also minimize the edges in the conflict graph. For instance, if $n_1 = 3$ and $n_2 = 4$, the sectioning set \tilde{S} depicted in Figure 3 also produces $n_1 + n_2 - \text{GCD}(n_1, n_2)$ edges in $G_{\tilde{S}}$. However, this sectioning set is specific to the

s_1	1	2	3			
s_2	1	4	2	4	3	4

Fig. 3 Sectioning set \tilde{S} .

section numbers $n_1 = 3$ and $n_2 = 4$ and finding an optimal sectioning for a third course with a larger number of sections may become tedious, if not impossible. In contrast, our regular sectioning function $s_i = \lceil n_i x \rceil$ does not depend on other values n_j , $j \neq i$, permitting to add any number of sections with guaranteed optimality as long as regular sectioning is used.

5.3 Timetabling feasibility

Given any sectioning set, it is not clear whether a conflict free timetable exists or not, since there may be several additional constraints preventing it, like restrictive teachers availabilities. However, permuting sections (i.e. sets of students) of the same course may be allowed as long as teacher assignments remain unchanged. Given a timetable T , we call a sectioning set S *feasible* with respect to T if there exists a set of section permutations $\sigma_1, \dots, \sigma_k$, with σ_i over the set $\{1, \dots, n_i\}$, such that the resulting timetable is conflict free, i.e. such that $T(s_{i,j}) \cap T(s_{i,j'}) = \emptyset$

whenever $s_i^{-1}(\sigma_i(j)) \cap s_i^{-1}(\sigma_i(j')) \neq \emptyset$. Obviously, regular sectioning may not be feasible for any timetable, the simplest example being where all sections get the same timeslots. A less trivial and more useful question is the following: if there exists a feasible sectioning set for some timetable T , is the regular sectioning set also feasible? In the general case the answer is negative, and we illustrate it with a small example of 12 students, 2 courses of 3 and 4 sections, durations $d_1 = 1$ and $d_2 = 2$, and the timetable depicted in Table 1. Indeed, Figure 4 displays

Section $s_{i,j}$	$s_{1,1}$	$s_{1,2}$	$s_{1,3}$	$s_{2,1}$	$s_{2,2}$	$s_{2,3}$	$s_{2,4}$
$T(s_{i,j})$	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	{4, 5}

Table 1 Timetable incompatible with regular sectioning.

the corresponding timeslots conflict graph². If we apply the regular sectioning

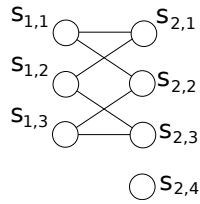


Fig. 4 Timeslots conflict graph G_T .

approach, the resulting sectioning conflict graph is, up to section permutations within the same course, depicted in Figure 5. Clearly, S is feasible with respect to

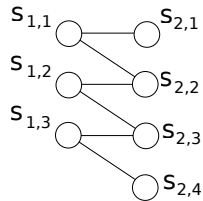
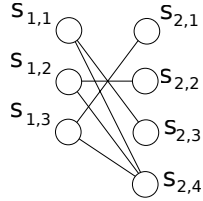


Fig. 5 Sectioning conflict graph G_S obtained with regular sectioning.

T if and only if there is a permutation of $\{s_{1,1}, s_{1,2}, s_{1,3}\}$ (or $\{s_{2,1}, s_{2,2}, s_{2,3}, s_{2,4}\}$) such that no pair of vertices is an edge in both conflict graphs. It is not the case, since there are 3 vertices of degree 2 among the $s_{2,j}$ vertices in G_T , whereas there are only 2 vertices of degree one among them in G_S . Now consider the same sectioning set \tilde{S} as in Figure 3, and described in Table 2 for our 12 students. The corresponding sectioning conflict graph is displayed in Figure 6. This graph is the

² For the sake of simplicity, we do not display edges between vertices of the same course, since these potential conflicts can never be realized.

Student	1	2	3	4	5	6	7	8	9	10	11	12
Section for course 1	1	1	1	1	2	2	2	2	3	3	3	3
Section for course 2	3	3	3	4	2	2	2	4	1	1	1	4

Table 2 Sectioning set \tilde{S} .**Fig. 6** Sectioning conflict graph $G_{\tilde{S}}$ obtained with sectioning set \tilde{S} .

bipartite complement of the timeslots conflict graph G_T , thus \tilde{S} is feasible with respect to T .

This example shows that if regular sectioning permits to minimize the number of edges, it may not be the only optimal solution in that sense, and there may exist timetables that are not feasible for S , but feasible for another sectioning set. However, in this example, the course $s_{2,2}$ is assigned timeslots 1 and 3, and we did not find such an example where all course are assigned only consecutive timeslots, a situation we may expect in practice. The existence or not of such an example remains an open question.

5.4 Teacher assignments

In most cases, teachers are indifferent to which sections they are assigned. The section assignments among teachers within a same course can therefore be freely permuted. We provide here some hints on how to exploit this flexibility prior to timetabling. We again base our discussion on the number of potential conflicts, and consider teachers like students, in the sense that they contribute to edges in the student conflict graph.

If we restrict attention to only one course, we observe that changing teacher assignments to sections will not change the conflict subgraph (up to isomorphism) induced by sections of this course, hence the number of edges.

If a teacher is assigned sections in different courses, it may be preferable to assign him or her similar sections (with many students in common). Indeed, these sections are likely to be already adjacent in the student conflict graph, so adding the teacher to them will create only few new edges, if any. Observe again that this similarity argument is consistent with homogeneous sectioning (Carter (2001)).

Finally, for the case where two teachers are assigned sections in different courses, a criterion could be based on their availabilities: it may be suitable to assign similar sections to teachers with complementary availabilities. This criterion suggests a preprocessing consisting in permuting teachers among sections in order to maximize the probability of existence of a feasible timetable. However, this task does not promise to be trivial: if two teachers have complementary availabilities

and are assigned similar sections, what about a third teacher having availabilities similar to the first one?

5.5 Courses with unbalanced sections sizes

This paper deals with the situation where all courses must have balanced sections. It happens in practice that some courses have some "subcourses" chosen by students, making perfect balancing impossible. We don't have a solution proven optimal for this more general case, but we give some hints based on experience.

Consider the following example. Assume that a class has mandatory courses with $n_1 = 3$ and $n_2 = 4$ sections. The resulting regular sectioning is displayed on Figure 7. Additionally, assume we have a course, say "Foreign Language", which

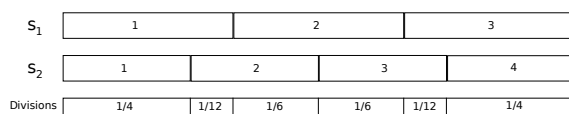


Fig. 7 Regular sectioning with 3 and 4 sections.

is made of one section of German, two sections of Chinese and three sections of Spanish, whereas 23 students subscribed for German, 40 for Chinese and 57 for Spanish. Requiring the sections of a same language to be balanced, we would need to create one section of 23 students (19.2% of the total), two sections of 20 students ($\frac{1}{6}$ each) and 3 sections of 19 students (15.8% each).

It is tempting to associate both Chinese sections to divisions 3 and 4 since they have exactly the right size. Doing this, we observe that there is no way to associate the remaining students in a connected manner, and the total number of new conflicts due to "Foreign Language" would be at least 16. Now if we look at the different connected ways of sectioning this course, we observe that the number of new conflicts of 16 is reachable too, and that this number depends on how we order the 6 sections. For instance, placing from left to right the German course, followed by both Chinese courses and the three Spanish courses, the number of such conflicts would be 17.

Based on such kind of observations, we also recommend to keep connectedness while sectioning such a course, but optimality remains on open question as well as characterizing the best sections order. If for some reason connectedness has to be dropped, the homogeneity criterion mentioned in Sections 2 and 4 tells us to avoid merging divisions that are too distant. This has been experienced in practice when it comes to timetabling: courses involving non adjacent and distant divisions have many potential conflicts and are therefore more difficult to schedule.

Finally, it is important to notice that in our example, we assume we can arbitrarily decide which language is associated to which division, since it amounts to permuting the students. However, if there is another such course, this freedom is partially lost, since it is then only allowed to permute divisions in the same subcourse.

6 Conclusion

In this paper, we address the problem of student sectioning on mandatory courses with various numbers of sections, before knowing the timetable. We prove that sectioning students in a regular way permits to minimize the total number of potential conflicts. We then discuss implications of this result, uniqueness of this optimal solution, the existence of a feasible timetable, some issues about teachers assignments and the case where a course has to be sectioned in an unbalanced way.

As an output, we give the following recommendations:

- For all course that are mandatory, section regularly;
- If you can choose or influence the numbers $\{n_1, \dots, n_k\}$ of sections by course, choose numbers that have a large common divisor;
- Assign a teacher to similar sections (having a large intersection);
- Assign teachers with complementary availabilities to similar sections;
- Section in a connected manner even on courses with unbalanced section sizes;
- If connected sectioning is impossible, avoid merging divisions that are too distant according to the chosen student ordering.

We leave the reader with the following open questions:

- Are there section numbers n_1, \dots, n_k on k courses, with durations d_1, \dots, d_k , a non-regular balanced sectioning \tilde{S} and a timetable T where all courses are assigned consecutive timeslots, such that \tilde{S} is feasible with respect to T , but regular sectioning is not?
- Is the problem "BSS + (E)" defined in (Dostert et al., 2016) still NP-complete, when restricted to the case where courses are assigned only consecutive timeslots? What about the variation of "BSS + (E)" where sections have to be balanced?
- In the case of an additional course with some imposed unbalanced section sizes, is there still a connected sectioning which is optimal? If yes, how to order the sections?

References

- Alvarez-Valdes, R., Crespo, E., and Tamarit, J. M. (2000). Assigning students to course sections using tabu search. *Annals of Operations Research*, 96(1):1–16.
- Amintoosi, M. and Haddadnia, J. (2005). Feature selection in a fuzzy student sectioning algorithm. In Burke, E. and Trick, M., editors, *Practice and Theory of Automated Timetabling V (PATAT 2004)*, pages 147–160, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Beyrouthy, C., Burke, E. K., Landa-Silva, D., McCollum, B., McMullan, P., and Parkes, A. J. (2008). Conflict inheritance in sectioning and space planning. In *Practice and Theory of Automated Timetabling VI (PATAT 2006)*, Montreal.
- Carter, M. W. (2001). A comprehensive course timetabling and student scheduling system at the university of waterloo. In *Practice and Theory of Automated Timetabling III (PATAT 2000)*, pages 64–84, London, UK. Springer-Verlag.

- Cheng, E., Kruk, S., and Lipman, M. (2003). Flow formulations for the student scheduling problem. In Burke, E. and De Causmaecker, P., editors, *Practice and Theory of Automated Timetabling IV (PATAT 2002)*, pages 299–309, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Dostert, M., Politz, A., and Schmitz, H. (2016). A complexity analysis and an algorithmic approach to student sectioning in existing timetables. *Journal of Scheduling*, 19(3):285–293. Copyright - Springer Science+Business Media New York 2016.
- Hertz, A. (1991). Tabu search for large scale timetabling problems. *European Journal of Operational Research*, 54(1):39 – 47.
- Laporte, G. and Desroches, S. (1986). The problem of assigning students to course sections in a large engineering school. *Computers & Operations Research*, 13(4):387 – 394.
- Müller, T. and Murray, K. S. (2010). Comprehensive approach to student sectioning. *Annals of Operations Research*, 181:249–269.
- Murray, K., Müller, T., and Rudová, H. (2007). Modeling and solution of a complex university course timetabling problem. In *Practice and Theory of Automated Timetabling VI (PATAT 2006)*, pages 189–209.
- Selim, S. M. (1988). Split vertices in vertex colouring and their application in developing a solution to the faculty timetable problem. *The Computer Journal*, 31(1).
- ten Eikelder, H. and Willemen, R. (2001). Some complexity aspects of secondary school timetabling problems. In Burke, E. K. and Erben, W., editors, *Practice and Theory of Automated Timetabling III*, volume 2079 of *Lecture Notes in Computer Science*, pages 18–27. Springer Berlin Heidelberg, Springer Berlin Heidelberg.