

# Improving Physician Schedules by Leveraging Equalization: Cases from Hospitals in U.S.

Pelin Damcı-Kurt<sup>1</sup>, Minjiao Zhang<sup>2</sup>, and Brian Marentay<sup>3</sup>

<sup>1</sup>Lightning Bolt Solutions, South San Francisco, CA 94080, pelin@lightning-bolt.com

<sup>2</sup>Kennesaw State University, Kennesaw, GA 30144, mzhang16@kennesaw.edu

<sup>3</sup>Lightning Bolt Solutions, South San Francisco, CA 94080, bmarentay@lightning-bolt.com

October 22, 2017

## Abstract

In this paper, we consider physician scheduling problems originating from a medical staff scheduling service provider based in the United States. Creating a physician schedule is a complex task. An optimal schedule must balance a number of goals including adequately staffing required assignments for quality patient care, adhering to a unique set of rules that depend on hospital and medical specialties, and maintaining a work-life balance for physicians. We study various types of physician and hospital requirements with different priorities, including equalization constraints to ensure that each provider will receive approximately the same number of a specified shift over a given time period. A major challenge involves ensuring an equal distribution of workload among physicians, with the end goal of producing a schedule that will be perceived by physicians as fair while still meeting all other requirements for the group. As the number of such equalization constraints increases, the physician scheduling optimization problem becomes more complex and it requires more time to find an optimal schedule. We begin by constructing mathematical models to formulate the problem requirements, and then demonstrate the benefits of a polyhedral study on a relaxation of the physician scheduling problem that includes equalization constraints. A branch-and-cut algorithm using valid inequalities derived from the relaxation problem shows that the quality of the schedules with respect to the soft constraints is notably better. An example problem from a hospitalist department is discussed in detail, and improvements for other schedules representing different specialties are also presented.

*Keywords: physician scheduling; optimization; soft constraints; workload distribution; valid inequalities; mixed-integer programming*

## 1 Introduction and background

Physician scheduling is an optimization problem that is complex and difficult to generalize. Although most constraints vary by hospital and medical specialty, there are certain requirements that should be met by any hospital, such as staffing required shifts for patient care. Some medical groups have a set list of requirements that rarely change, while others have rules that are constantly evolving. In either case, manually scheduling physicians becomes extremely difficult as input parameters such as schedule length, number of personnel and number of rules increase. Time spent on manual scheduling can add up to hundreds of hours per year, even with a team of people working to create a schedule. In many cases, hospitals and physicians settle for an unsatisfactory schedule due to

time constraints. Such inefficiencies can be particularly impactful, as physician schedules have a direct effect on the lives of physicians and the patients that they care for.

Studies show that physician burnout, a work-related syndrome involving emotional exhaustion, depersonalisation, and a sense of reduced personal accomplishment [16], has increased in practicing physicians in the United States from 2011 to 2014. The rates of burnout in primary care, family medicine, and emergency medicine were well above 50% in 2011 and they have increased in 2014 [18]. Research has focused on the consequences of physician burnout in a number of areas. For example, among physicians in a large health care organization, burnout and declining satisfaction were strongly associated with reductions in professional work effort over the following 24 months [19]. Furthermore, burnout is associated with lower patient satisfaction, reduced health outcomes, and possibly an increase in costs [4]. Addressing such a complex issue will require improvements in many tasks physicians are currently responsible for. The question, “How to create physician schedules that are both efficient and fair”, if answered correctly, can contribute to improving physician happiness and productivity. With the help of automated physician scheduling, we can find a balance between the needs of three stakeholders: physicians, hospitals and patients.

In this paper, we focus on general physician scheduling problems faced by a medical staff scheduling service provider based in the United States. The building blocks of a physician schedule are personnel, assignments (shifts), and a scheduling horizon (number of days scheduled). Input parameters for personnel include information such as name, assignments that they are eligible to work, and full time equivalent (FTE) percentage. The definition of assignments include a required number of minimum/maximum personnel and on which days they exist. A list of several classes of rules is available for each client to use in setting up their requirements. These rules can be defined as hard or soft requirements. Hard requirements must be met fully while soft requirements can be relaxed when conflicting with other rules. Rule classes offer flexibility and their usage depends on hospital and medical specialty requirements. The objective is to minimize the sum of penalties associated with each relaxed requirement that cannot be satisfied by the schedule.

In general, physician scheduling rules can be categorized in the following groups.

- Assignment demand rules: the option to choose a minimum and maximum number of personnel working a given assignment. For example, if a single person has to be on call during the daytime every day, then the minimum and maximum values for the Day Call assignment would each be set to one for every day of the week.
- Assignment compatibility rules: if more than one assignment is added into the system, then the scheduler has to set the compatibility of each assignment relative to all other assignments. For example, the Vacation and Day shift are incompatible assignments, meaning a single provider cannot work both assignments on the same day.
- Equalization rules: for a fair distribution of workload. They are specified as “Equalize allocation of *assignment(s)* across *personnel group*, allowing for at most a certain *difference* between personnel”, where *assignment(s)*, *personnel group*, and the certain *difference* value are user input parameters. These constraints generally span the given scheduling period.
- Conditional rules: requirements that can be stated as “if-then-else” statements. For example, “If Garfield is not on Vacation on Monday, then schedule him for either Day Call or Night Call”.
- Block scheduling rules: stated such as “John works Day Call in a block of at least 5 and at most 7 days in a row”.
- Numeric rules: constraints that help determine the minimum/maximum number of shifts

scheduled. For example, “Schedule Sarah for at most 15 shifts per month”.

The first and the most important step in building a physician schedule is to understand the exact needs of the physician group. The initial setup of requirements can take some effort because many scheduling administrators do not have a readymade list of requirements. Even when they do have predefined requirements, there are often constraints that they are using (but not aware of) when manually creating schedules. Thus, there is often a need to generate multiple schedules during implementation to add all the rules into the mathematical model. Moreover, some physician groups prefer to create several alternative schedules due to their requirement changes in every scheduling period. Because of the varying requirements of physician groups, different scheduling models are needed and will require different runtimes to obtain satisfactory schedules. The runtimes can range from a few seconds to as long as multiple days. Schedules are usually created using a mixed-integer programming (MIP) model, and its performance is of the utmost importance. The more time a schedule auto-generation takes, the longer administrators have to wait before they can review the schedule and suggest additional rules. Therefore, there is always a motivation to improve the performance of the MIP model and generate schedules of better quality (fewer soft rule violations) within the same time limit.

To enhance the performance of the MIP model used by the medical staff scheduling service provider, we conduct a polyhedral analysis on a *relaxation* of the MIP model to propose a branch-and-cut framework with new valid inequalities. We refer to it as the relaxation of the MIP model because it is not realistic to exhaustively include every possible rule of physician scheduling in our model. Instead, we intend to focus on the commonly existing requirements of assignment distribution to improve the schedule quality for a broad set of physician scheduling problems. In general, it is more time-consuming to solve MIP models for large-sized hospital groups, whose number of physicians scheduled can range from 10 to 250, the number of assignments scheduled from 20 to 150 and the schedule length from one week to one year. After analyzing many such large hospital groups, we narrowed down the common set of constraints to include *maximum assignment demand constraints* and *equalization constraints*. The maximum assignment demand constraints are written to limit the number of personnel that can be assigned to a certain shift. Hospitals do not want to assign more than the required number of personnel on a shift, which would correspond to unnecessary costs to compensate the additional personnel, so the maximum assignment demand constraints are defined as hard rules and can never be relaxed in this paper. The equalization constraints are needed to generate a balanced schedule by having an even distribution of shifts among all physicians, and can be defined as soft or hard, depending on other scheduling rules. Some groups state these assignment distribution requirements as percentages of acceptable deviation from the mean, while others focus on the specific number of assigned shifts. To test the enhanced algorithm for comparison with the current algorithm, we form a list of hospital groups that have maximum assignment demand constraints as well as equalization constraints and satisfy the assumptions of our algorithm.

The remainder of the paper is organized as follows. In Section 2, we review the literature on relevant staff scheduling problems. Section 3 describes the model relaxation and our solution procedure. In Section 4, we use Flagstaff Medical Center Hospitals and other physician groups as examples to illustrate the common scheduling requirements for hospitalists, the associated challenges, and the improvement on the quality of physician schedules using the proposed enhanced algorithm. Finally, Section 5 concludes the paper by summarizing our findings and discussing future work.

## 2 Literature review

The scheduling of employees is a multifaceted and time-consuming task. The basic goal is to assign employees to working shifts, taking into account organizational and regulatory rules, employee skills and preferences, required staffing, and other problem specific requirements. Although the staff scheduling problem has been extensively studied in the literature, previous studies have mostly focused on solving very particular problems that derive from practical needs. We refer the reader to several review papers on staff scheduling, such as [9], [10] and [25]. In the health care domain, the literature has mostly focused on the scheduling of hospital nurses. The nurse scheduling problem (NSP) assigns nurses to work shifts subject to a variety of hard and/or soft constraints such as legal regulations, hospital policies and nurse preferences. This is one of the many problems in healthcare that if solved properly adds major value to hospitals by reducing costs, increasing retention of employees and improving quality of healthcare. Common methods for solving NSP's are mathematical programming, heuristics and artificial intelligence. Mathematical programming models include traditional solution methods from linear programming, integer programming, goal programming and networks to solve NSP's. Cheang et al. [7] provide a bibliographic survey of many models and methodologies available to solve the nurse rostering problems. Burke et al. [5] review nurse rostering within the global personnel scheduling problem in healthcare and describe the role that nurse rostering plays within the wider context of longer term hospital personnel planning. Venkataraman and Brusco [26] develop two mixed-integer linear programs. The first model is used to determine aggregate labor requirements for a six-month scheduling horizon. The second model is used to disaggregate the nurse staffing plan into two-week labor schedules. They show that there are important interactions between staffing and scheduling policies on labor costs. Maenhout and Vanhoucke [15] present an integrated methodology for allocating a given workforce over multiple departments based on the hospital's nurse staffing policies, each ward's shift scheduling policies, and each nurse's characteristics. Multiple objectives are considered in their model, such as cost, job satisfaction as a result of scheduling, and the quality of nursing care. Similarly, Wright and Mahar [29] consider organizing a nurse workforce as efficiently as possible in order to reduce the associated costs and to increase employee satisfaction.

Physician scheduling requirements exhibit a few similarities to the NSP, such as the requirement of adequate coverage for demand and preferences for certain shifts. The objective functions used for the NSP include minimizing staffing cost, minimizing number of nurses scheduled, minimizing staff dissatisfaction and many others. In physician scheduling, staffing cost is not as relevant but minimizing deviations from the soft scheduling requirements is especially important. Furthermore, physician schedules have to include many medical specialty-specific requirements and are in general very different from nurse schedules. This implies that, depending on the department and the hospital, physician scheduling requirements can vary considerably. For example, hospitalist doctors deliver comprehensive medical care to hospitalized patients. A fundamental rule for hospitalists is to work at least a certain number of days in a row in order to minimize the number of possible handoffs of patients from one doctor to another. Emergency medicine however, does not particularly require continuity of work by doctors but requires adequate staffing levels. Due to the complexity of generalizing physician schedules, there is much less work in the literature for this problem compared to the NSP. Erhard et al. [8] give a review on physician scheduling and show that it has gained more interest in the last decade. To solve a physician scheduling problem, many researchers have applied mathematical programming, including but not limited to Sherali et al. [20] for assigning residents to night shifts, Rousseau et al. [17] for a customizable rostering problem, and Smalley and Keskinocak [21] for an optimization-based decision support system that generates weekly and daily schedules. These papers attempt to capture different requirements that come up in similar

physician scheduling settings. Their assumptions, however, are based on a relatively small data set.

After analyzing over 5500 department schedules across 57 medical specialties (which includes all clients), a recent report by a medical staff scheduling service provider [14] indicates that the most complex physician schedules correspond to groups which must provide 24/7 patient care coverage in an unpredictable setting, such as Emergency Medicine departments. Aligning with the report’s findings, we found more papers in the Emergency Room (ER) physician scheduling literature than any other specialty [2, 11, 24]. The same report also shows that scheduling complexity varies significantly by specialty and, in general, schedules involving hospital coverage are more complicated than those of office-based departments.

In addition to the individualized challenges of scheduling, almost all medical groups desire an equal workload. This is not only for work-life balance of physicians, but for patient safety as well. An unbalanced workload distribution implies excessive working hours for residents, which can in turn lead to medical errors and affect patient care [3, 12, 27, 28]. Any research on physician scheduling must consider a balance of workload among physicians as a hard or soft constraint, where a soft constraint can be violated with a certain penalty. Carter and Lapierre [6] extract characteristics of a generic ER physician scheduling problem from six hospitals in greater Montreal, Canada. Equalization of shifts in ER physician groups is done according to seniority levels. Similarly, Beaulieu et al. [2], Gendreau et al. [11] and Topaloglu [24] show results for their versions of ER physician scheduling that include fair distribution of responsibilities among physicians. Levin et al. [13] measure the distribution of workload balance among ER physician teams in terms of how some physicians are overworked while others remain idle. Stolletz and Brunner [23] introduce a set covering formulation that integrates physician preferences and fairness aspects into the scheduling model to solve the flexible shift scheduling problem of physicians in hospitals. Baum et al. [1] apply a mixed-integer program to schedule physicians in a radiology division under revenue and fairness aspects.

While this literature review shows many related problems and solution methods, none of them focus on solving a family of physician scheduling problems with equalization constraints. To the best of our knowledge, this is the first paper to consider improving the efficiency of solving an MIP model with equalization constraints, and as a result to improve the overall solutions for a class of physician scheduling problems.

### 3 Model description and solution procedure

In this section, we first present the formulation of the relaxation problem, which consists of maximum assignment demand and equalization constraints as described in Section 1. We then develop a class of valid inequalities for the relaxation and show that the separation problem can be solved in polynomial time.

We define  $P$  as the set of physicians,  $A$  as the set of assignments, and  $D$  as the set of dates that correspond to the individual days of the given scheduling period. Let  $x_{pda}$  be 1 if physician  $p \in P$  is assigned to assignment  $a \in A$  on date  $d \in D$ , and 0 otherwise. Variable  $x_{pda}$  exists only if physician  $p$  is eligible for assignment  $a$  and assignment  $a$  exists on date  $d$ . Only the physicians who are eligible for an assignment can work it. We let  $P_a$  be the set of physicians who are eligible for assignment  $a \in A$  and  $D_a$  be the set of dates when assignment  $a \in A$  exists. Then the maximum assignment demand constraints are formulated as

$$\sum_{p \in P_a} x_{pda} \leq \theta_{ad}^{\max} \quad \forall a \in A, d \in D_a, \tag{1}$$

where  $\theta_{ad}^{\max}$  represents the maximum demand for assignment  $a \in A$  on date  $d \in D_a$ . In theory, maximum demand constraints for assignments can be relaxed. However, they are not relaxed in practice because doing so can increase costs to compensate unnecessary personnel scheduled for an assignment. Demand constraints (1) exist and cannot be relaxed for all the medical groups studied in this paper.

Equalization rules are stated as “Equalize allocation of assignment group  $AG \subseteq A$  across personnel group  $PG \subseteq P$ , allowing for a difference of at most  $e$  between personnel”, where  $AG$ ,  $PG$  and  $e \in \mathbb{Z}_+$  are given parameters. In addition, parameter  $W_{ad}$  represents the weight of an assignment  $a \in AG$  on date  $d \in D_a$ , which can be different for different assignments depending on the hardness level. The following inequalities (2) and (3) formulate this requirement.

$$\sum_{a \in AG, d \in D_a} W_{ad} x_{p_1 da} \leq e + \sum_{a \in AG, d \in D_a} W_{ad} x_{p_2 da} \quad \forall p_1 \neq p_2 \in PG, \quad (2)$$

$$\sum_{a \in AG, d \in D_a} W_{ad} x_{p_2 da} \leq e + \sum_{a \in AG, d \in D_a} W_{ad} x_{p_1 da} \quad \forall p_1 \neq p_2 \in PG. \quad (3)$$

One can relax inequalities (2) and (3) by adding nonnegative variables  $\delta_{p_1 p_2}$  and  $\delta_{p_2 p_1}$ , respectively to the right-hand side (RHS). We assume all equalization rules can be relaxed in this paper.

Next, we describe a class of valid inequalities for the polytope defined by constraints (1)–(3). We introduce set  $AD$ , which consists of 2-tuple elements  $(a, d)$  with  $a \in AG$  and  $d \in D_a$ , to easily represent all of the assignment and date pairs considered in the equalization rule constraints. Among all physician groups that we have analyzed, the most common requirement is to equalize assignments with the same weight, particularly 1, and schedule at most one personnel on any assignment. Thus, we assume  $W_{ad} = 1$  and  $\theta_{ad}^{\max} = 1$  for all  $(a, d) \in AD$ . Intuitively, the following valid inequalities (4) imply that, given any two subsets that partition the complete set of (assignment, date) pairs, the maximum allowable difference between the total assignments on these two subsets is limited by an explicitly specified value.

**Proposition 1.** *We consider two nonempty sets,  $AD_1$  and  $AD_2$ , that partition set  $AD$  such that  $|AD_1| = k_1$ ,  $|AD_2| = k_2 = |AD| - k_1$ , and  $k_1 + e$  is an odd number. Let  $f = \frac{k_1 + e - 1}{2}$ . Then the following inequality is valid for the problem consisting of constraints (1)–(3) with integrality restrictions for  $x_{pda}$ .*

$$\sum_{(a,d) \in AD_1} x_{p_1 da} \leq f + \sum_{(a,d) \in AD_2} x_{p_2 da} + \delta_{p_1 p_2} \quad \forall p_1 \neq p_2 \in PG. \quad (4)$$

*Proof.* Note that  $f = \frac{k_1 + e - 1}{2} \geq 0$ , because by definition, set  $AD_1$  (i.e.  $k_1 \geq 1$ ) is nonempty and  $e \geq 0$ . Additionally,  $f$  is an integer since  $k_1 + e$  is an odd number. Let  $AD_r = \{(a, d) \in AD_1 : x_{p_1 da} = 1\}$  and  $|AD_r| = k_r$ . So  $\sum_{(a,d) \in AD_1} x_{p_1 da} = k_r$ . We consider the following two cases.

1.  $k_r \leq f$ .

This case is trivial since the RHS of inequality (4) is  $f$  plus non-negative variables.

2.  $k_r \geq f + 1$ .

Due to maximum demand constraint (1), the definition of set  $AD_r$ , and the fact that  $\theta_{ad}^{\max} = 1$  for all  $(a, d) \in AD$ , we have  $\sum_{(a,d) \in AD_r} x_{p_2 da} = 0$ . Then the equalization rule (2) implies that,

$$\sum_{(a,d) \in AD_1 \setminus AD_r} x_{p_1 da} + \sum_{(a,d) \in AD_r} x_{p_1 da} \leq e + \sum_{(a,d) \in AD_1 \setminus AD_r} x_{p_2 da} + \sum_{(a,d) \in AD_2} x_{p_2 da} + \delta_{p_1 p_2}. \quad (5)$$

Furthermore,  $\sum_{(a,d) \in AD_1 \setminus AD_r} x_{p_2 da} \leq k_1 - k_r \leq k_1 - (f + 1) = 2f - e + 1 - (f + 1) = f - e$ . After replacing  $\sum_{(a,d) \in AD_1 \setminus AD_r} x_{p_2 da}$  on the RHS of inequality (5) with  $f - e$ , we obtain inequality (4), which completes the proof.  $\square$

*Example.* To illustrate inequality (4), we consider a physician set  $P = \{p_1, p_2, p_3\}$ , a date set  $D = \{M, T, W, Th, F\}$ , an assignment group  $AG = A = \{\text{Off}\}$ , and a personnel group  $PG = \{p_1, p_2\}$ , with the maximum deviation  $e = 1$ . Then the maximum assignment demand constraints are

$$x_{p_1, d, \text{Off}} + x_{p_2, d, \text{Off}} + x_{p_3, d, \text{Off}} \leq 1 \quad \forall d \in D,$$

and the equalization constraints are

$$\begin{aligned} \sum_{d \in D} x_{p_1, d, \text{Off}} &\leq 1 + \sum_{d \in D} x_{p_2, d, \text{Off}} + \delta_{p_1 p_2}, \\ \sum_{d \in D} x_{p_2, d, \text{Off}} &\leq 1 + \sum_{d \in D} x_{p_1, d, \text{Off}} + \delta_{p_2 p_1}, \\ x_{p, d, \text{Off}} &\in \{0, 1\} \quad \forall p \in P, d \in D, \delta_{p_1 p_2}, \delta_{p_2 p_1} \geq 0. \end{aligned}$$

Therefore, the valid inequality (4) with  $AD_1 = \{(\text{Off}, \text{Th}), (\text{Off}, \text{F})\}$ ,  $AD_2 = \{(\text{Off}, \text{M}), (\text{Off}, \text{T}), (\text{Off}, \text{W})\}$ ,  $k_1 = 2$ , and  $f = \frac{k_1 + e - 1}{2} = 1$  is

$$x_{p_2, \text{Th}, \text{Off}} + x_{p_2, \text{F}, \text{Off}} \leq 1 + x_{p_1, \text{M}, \text{Off}} + x_{p_1, \text{T}, \text{Off}} + x_{p_1, \text{W}, \text{Off}} + \delta_{p_2 p_1}. \quad (6)$$

For example, if  $x_{p_2, \text{Th}, \text{Off}} = x_{p_2, \text{F}, \text{Off}} = 1$ , then by maximum assignment demand constraints we have  $x_{p_1, \text{Th}, \text{Off}} = x_{p_1, \text{F}, \text{Off}} = 0$ . To satisfy the equalization constraints with the maximum deviation  $e = 1$ , we have to have  $x_{p_1, \text{M}, \text{Off}} + x_{p_1, \text{T}, \text{Off}} + x_{p_1, \text{W}, \text{Off}} + \delta_{p_2 p_1} \geq 1$ . Thus, inequality (6) is valid.  $\square$

Next, we show that although there are exponentially many inequalities (4), their separation can be conducted in polynomial time.

**Proposition 2.** *Given a point  $(\bar{x}, \bar{\delta}) \in \mathbb{R}_+^{|A||D||P| + \frac{|P|(|P|-1)}{2}}$ , there is an  $O(|A||D| \log(|A||D|))$  algorithm to find the most violated inequality (4), if any.*

*Proof.* Because  $k_1 = |AD_1|$ ,  $f = \frac{k_1 + e - 1}{2}$  and  $\sum_{(a,d) \in AD} x_{p_2 da} = \sum_{(a,d) \in AD_1 \cup AD_2} x_{p_2 da}$ , inequality (4) is equivalent to

$$\sum_{(a,d) \in AD_1} x_{p_1 da} + \sum_{(a,d) \in AD_1} x_{p_2 da} - \frac{k_1}{2} \leq \frac{e - 1}{2} + \sum_{(a,d) \in AD} x_{p_2 da} + \delta_{p_1 p_2} \quad \forall p_1 \neq p_2 \in PG. \quad (7)$$

Given a provider pair  $p_1, p_2$  and a point  $(\bar{x}, \bar{\delta})$ , the RHS value of inequality (7) is constant. Thus, to find the most violated inequality (7), we maximize the LHS value of inequality (7) at point  $(\bar{x}, \bar{\delta})$ . Initially, we let  $AD_1 = \emptyset$ .

1. For each assignment and date pair  $(a, d) \in AD$ , we let  $\sigma_{ad} = \bar{x}_{p_1 da} + \bar{x}_{p_2 da} - \frac{1}{2}$ . If  $\sigma_{ad} > 0$ , then we let  $AD_1 \leftarrow AD_1 \cup \{(a, d)\}$ . Therefore, the maximum value of the LHS of inequality (7) is  $\sum_{(a,d) \in AD_1} \sigma_{ad}$ .
2. If  $f < e$ , then it implies  $|AD_1| < e + 1$  because  $f = \frac{|AD_1| + e - 1}{2}$ . We sort  $\{\sigma_{ad}\}_{(a,d) \in AD \setminus AD_1}$  in non-increasing order and add the assignment and date pairs associated with the first  $(e + 1 - |AD_1|)$  biggest  $\sigma_{ad}$  values into set  $AD_1$ . Then the updated value of  $f$  satisfies  $e \leq f$ .
3. If  $e$  is even, then  $k_1$  has to be odd by definition (since  $k_1 + e$  is defined to be odd). Similarly, if  $e$  is odd, then  $k_1$  has to be even. Thus, if the choice of  $f$  does not satisfy this restriction, we add one more  $(a, d)$  pair to set  $AD_1$  from the non-increasing sorted set of  $\{\sigma_{ad}\}_{(a,d) \in AD \setminus AD_1}$  values.

If  $\sum_{(a,d) \in AD_1} \sigma_{ad} > \frac{e-1}{2} + \sum_{(a,d) \in AD} \bar{x}_{p_2 da} + \bar{\delta}_{p_1 p_2}$ , then we have found the most violated inequality (7) with the determined sets  $AD_1$  and  $AD_2 = AD \setminus AD_1$ . Otherwise, there is no violated inequality (7) with the given provider pair  $p_1, p_2$  at point  $(\bar{x}, \bar{\delta})$ .

Note that the complexity of Step 1 is  $O(|A||D|)$  and the complexity of Steps 2 and 3 is  $O(|A||D| \log |A||D|)$ . Therefore, for inequality (4) we have a separation algorithm of complexity  $O(|A||D| \log(|A||D|))$ .  $\square$

Next, we show computational results using inequalities (4) in a branch-and-cut-algorithm for several real physician scheduling problems.

## 4 Case studies

In this section, we use Flagstaff Medical Center Hospitals (FMCH) as an example to elaborate on requirements for hospitalist scheduling and analyze specific constraints in addition to showing results from other physician groups with different specialties. Common aspects of a hospitalist schedule include the need to assign physicians a certain number of shifts per month/schedule, equalizing night and weekend shifts, linking of shifts and not having providers scheduled the day after a night shift. There are block scheduling rules such as “day shifts must be worked five to seven days in a row” to specify the number of consecutive days a provider works given an assignment. These requirements are important for hospitalist groups because they work with patients that have been hospitalized. Their goal is to minimize hand-offs, as transferring of patient information from one physician to another presents a greater risk of miscommunication and mistakes. Smalley et al. [22] maximizes continuity in a pediatric intensive care unit by introducing a handoff continuity score (HCS) into their physician scheduling problem. The value of HCS helps in measuring the continuity of a schedule. Schedules may also need to accommodate provider contracts requiring them to work a given number of shifts for the fiscal year. Additionally, equalizing less desired shifts (nights and weekends) is crucial for every provider to feel like they are being treated equally. Next, we explore the details of specific hospitalist groups’ schedules.

### 4.1 Flagstaff Medical Center Hospitals

FMCH is a hospitalist group that schedules 36 providers on several assignments every three months (quarterly). Their main assignments are Day, Night, Admitting and Swing. All of the assignments require at least one provider to work it except for the Admitting and Swing assignments, which can be left unfilled. All of the shifts can be worked by at most one physician, except for a certain

number of days where two providers can be assigned. Note that these special days are not included in the equalization constraints. The only hard rules that cannot be relaxed for FMCH group are:

- A provider can work an assignment only if there are available slots, i.e. the “maximum number of personnel” component of assignment demand rules.
- None of the assignments can be worked together, so a provider can only work at most one assignment every day, i.e. the assignment compatibility rules.

FMCH has a number of equalization, conditional, block scheduling and numeric rules in their setup. It is important to note, however, that these requirements may change in the next scheduling quarter, depending on factors such as new hires and changes to hospital policies. Furthermore, any individual rule generally spans multiple personnel, assignments and dates, so the same rule can be relaxed many times in a given scheduling period if it is defined as a soft rule. The critical issue to address in these cases is what penalty value to set for each soft rule. It takes a few schedule runs to determine which rules are in fact the most important, so the penalties are adjusted until a balance is achieved. The penalty ranges from 1,000 to 1,000,000, with a larger penalty indicating a higher cost for violating the corresponding rule.

FMCH began implementation by defining an ideal set of rules. These rules were a starting point and over time they evolved considerably. The number of hard rules was limited because they can cause the scheduling problem to be infeasible. If there is no possible solution to a set of rules, it can take anywhere from a few minutes up to several hours to determine the conflicting requirements. With FMCH, initially all equalization requirements were defined as hard rules, but this was not practical due to previously entered provider requests for leave and other assignments. In the final setup all of the ideal rules were defined as soft with high priorities, with additional rules having a lower priority. These additional rules were included so that if an ideal rule was violated it would be in a more controlled way. Next, we give a select list of rules that can be relaxed under their respective categories.

Equalization rules consider the complete scheduling period and different sets of physicians:

- Equalize allocation of Admitting allowing for a difference of at most 4 between personnel.
- Equalize allocation of Day shifts allowing for a difference of at most 6 between personnel.
- Equalize allocation of Night shifts allowing for a difference of at most 3 between personnel.
- Equalize allocation of Swing allowing for a difference of at most 4 between personnel.
- Equalize allocation of shifts that can be worked on weekends allowing for a difference of at most 4 between personnel.

Conditional rules:

- Day assignments are divided into seven categories. For each category, there is a conditional rule that says if a provider is assigned to a Day assignment of a certain category on a given day, then they have to work the same assignment the day before and the following day.
- A generalized rule says that if a provider has been working on any shift (for a number of days in a row), then their blocks of time off must be a minimum of three days. This is important to give providers time to recuperate.
- A more specific set of rules says that if a provider has worked for  $n$  days in a row on day shifts, then their blocks of time off must be a minimum of  $n - 2$  days, where  $n = 5, 6, 7$ .
- Do not schedule certain shifts one after the other (e.g. Day shifts after a Night shift). One specific rule is “If a provider is on Swing shift today, then do not schedule them on a Day

shift tomorrow.”

Block scheduling rules:

- Nocturnists (who specialize in working overnight shifts) work in a block of at least 3 and at most 6 days in a row.
- Providers other than nocturnists work in a block of at least 5 and at most 7 days in a row. Assignments include all Day shifts.
- Providers other than nocturnists work in a block of at least 2 and at most 6 days in a row on the Swing shift.
- Rules differ for providers with FTE values less than 100%.

Numeric rules:

- Staff assignments Admitting and Swing everyday.
- Nocturnists should work between 11 and 13 Night shifts a month.
- 100% FTE providers should work between 41 and 43 shifts (including Day shifts, Night shifts, Admitting and Swing) every quarter.
- 80% FTE providers should work between 32 and 33 shifts every quarter.
- 50% FTE providers work between 21 and 22 shifts every quarter.

The relaxation of soft constraints can be due to a number of reasons such as conflicting soft rules, shift/vacation requests of physicians or not finding an optimal solution within the given time limit. The schedule being studied for FMCH has 641 requests in total, which the model works around when building the schedule. It is common for requests entered by physicians on certain assignments and days to violate some soft rules, especially during holidays due to vacation requests. The main objective becomes staffing assignments adequately so other rules become less important.

## 4.2 Comparison of FMCH Schedules

We now compare the schedules created with the current system and the new algorithm that utilizes valid inequalities (4). The objective is to find a schedule that minimizes the sum of the penalty associated with the violated soft rules. For FMCH as well as other cases presented in Subsection 4.3, we use Mosel with FICO Xpress 8.0.5, on a machine with Intel(R) Xeon(R) CPU X5690 at 3.47GHz with 16GB RAM. Time limit and MIP tolerance parameters depend on the difficulty of each specific problem. We incorporate inequalities (4) to the branch-and-cut algorithm by using the exact separation algorithm given in Proposition 2. Cuts are added for any branch-and-bound node with depth of less than or equal to 20.

With the time limit of 6 hours, the current system produces a FMCH schedule with a penalty of 269,256,000. After computation with the enhanced algorithm, the new system yields a schedule with a total penalty of 262,712,000, a reduction of 6,544,000. Comparing the sum of violation values, we see that the new algorithm relaxed fewer rules. We can also use the percentage of how close we were able to get to the optimal solution to describe the improvement in schedules, with a lower percentage indicating a better schedule. The optimality gap of the current system’s schedule is 66.70%, while we obtain a gap of 47.25% with the new system. If we extend the time limit, the optimality gap will be closed further. The computation is terminated at the current time limit for comparison purpose and to balance with the client’s concern over time consumption.

Figure 1 shows the differences in rule violations. The relaxed set of rules are very similar in both schedules, but the number of times they were relaxed is higher with the current algorithm for

eight rule types (R20019, R307, R20034, R20048, R304, R20035, R318, R306) and with the new algorithm for three rule types (R20021, R345, R309). The rules, shown on the horizontal axis in both Figures 1a and 1b, are listed from left to right in non-increasing order of penalties. Their

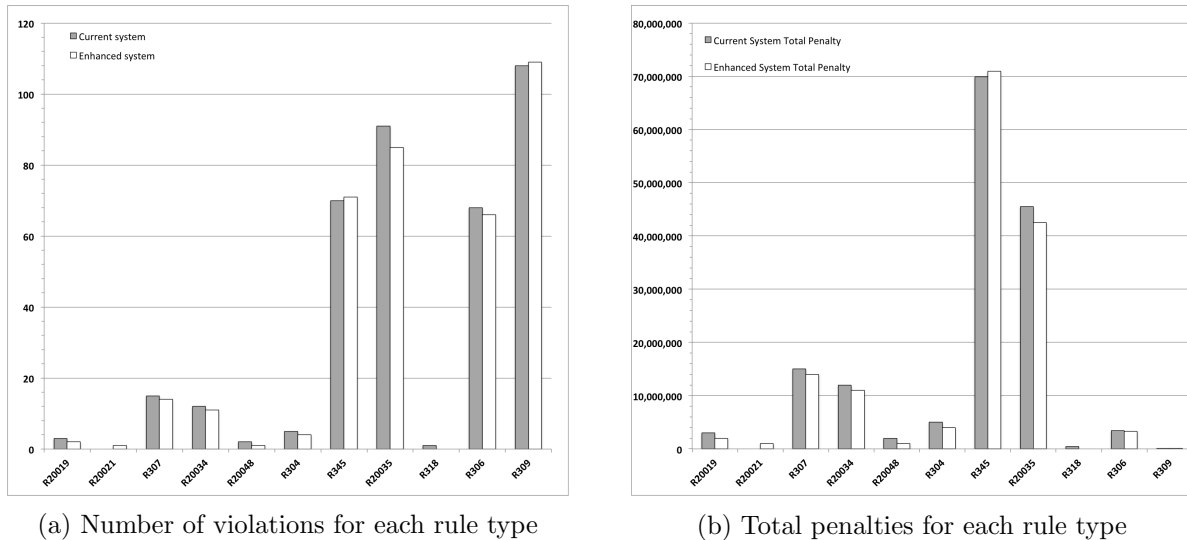


Figure 1: Comparison of rule violations for FMCH schedule generated by the current and enhanced algorithms.

descriptions and penalty values are presented in Table 1.

Table 1: Descriptions and penalties for rules shown in Figure 1.

Rule Number	Description	Penalty
R20019	100% FTE works: 41-43 times per scheduling period	1000000
R20021	80% FTE works: 32-33 times per scheduling period	1000000
R307	Day: Schedule a min of 4 max of 7 in a row	1000000
R20034	Work in a block of at most 6 days in a row	1000000
R20048	Night: Schedule a min of 3 and max of 6 in a row	1000000
R304	4 days off after any working shifts of 6 days in a row	1000000
R345	Day: Schedule a min of 5 max of 7 in a row	999000
R20035	Work in a block of at most 5 days in a row	500000
R318	Equalize Swing assignment	444000
R306	4 days off after any work	50000
R309	5 days off after any work	1000

Furthermore, the new algorithm is able to satisfy one more equalization rule: “Equalize allocation of Swing, allowing for a difference of at most 4 between personnel”, which is violated by the current algorithm. Table 2 shows how the number of Swing shifts for each physician changes in order to satisfy this equalization rule. We refer to an equalization rule as relaxed when at least one pair of doctors does not satisfy it. For example, in Table 2, with the current algorithm, the difference of number of swing assignments between physician 1 and physician 3 is 6, which is greater than the maximum allowed difference, 4. This leads to the relaxation of the equalization rule. With the

Table 2: Number of Swing shift assignments for five physicians in FMCH with the current and enhanced system

Physician	Number of swing shift assignments	
	Current System	Enhanced System
1	8	6
2	8	8
3	2	4
4	6	7
5	9	7

enhanced system, it shows that the equalization requirement “allowing for a difference of at most 4 between personnel” is met exactly, which is an improvement over the current system brought by the enhanced algorithm.

### 4.3 Scheduling Improvements for Other Physician Groups

The new algorithm has also been used to improve the efficiency and quality of schedule generation for many other physician groups. In this subsection, we discuss the scheduling results of eight different physician groups, which were chosen to demonstrate the variability in setups among different specialties. For confidentiality reasons, we refer to them as Clients 1 through 8. Each group has their own equalization requirements, which are defined as relaxable and are elaborated below. In addition, Table 3 shows the specialty, number of equalization rules and current time limits for each client setup.

- Client 1 has equalization rules that allow for a difference of at most 1 shift between personnel every month. Only a certain set of assignments are equalized. They have providers with different FTE percentages, so shift numbers are equalized within each percentage group.
- Client 2 has rules to equalize the Monday through Thursday distribution of a particular assignment across all call physicians, with a deviation of 1 shift being allowed on any given day.
- Client 3 equalizes different assignments for four different provider groups. Three of the rules require a difference of at most 2 between personnel while the fourth rule allows for at most 7.
- Client 4 equalizes Night, Day, Weekend, Hospitalist and Teaching assignments across six different provider groups. The number of shifts assigned per provider can differ by at most 0, 1, 2 or 7 depending on the provider group and assignment.
- Client 5 equalizes all Call assignments (21 total), Day and Night shifts. The maximum allowed deviation between personnel is 1 shift in all rules.
- Client 6 equalizes two Day Call, two Screening and five “MSK” shifts. The maximum allowed deviation between personnel is 1 shift for all except MSK shifts, which are allowed to have a difference of at most 2.
- Client 7 has equalization rules based on provider seniority. They equalize assignments such as Sick Call, Triage, Overnight and Weekend. Depending on the group of doctors and the assignment, the maximum allowed deviation of assignments per month ranges from 1 to 4.

- Client 8 has equalization rules over a 1 year scheduling period, one of the longest scheduling horizons. Providers with 100%, 50% and 25% FTE are equalized with separate rules. Assignments include Eve, Night and Weekend.

Table 3: Number of equalization rules and schedule runtime limits for eight different clients.

Client	Specialty	Number of Equalization Rules	Time Limit(s)
1	Radiology	17	14400
2	Anesthesiology	8	10800
3	Radiology	4	10800
4	Pediatrics	44	7200
5	Anesthesiology	23	14400
6	Radiology	7	10800
7	Emergency Medicine	9	25200
8	Hospitalist	12	55000

The time limit for the algorithm and other constraints vary for each client schedule, as shown in Table 3. Time limits are set after running multiple schedules for the same client. The goal is to find a balance between model runtime and schedule quality. Therefore, being able to obtain a better schedule (which is measured with the weighted sum of penalties for each relaxable constraint) within the same time limit for a given client rule setup is a performance indicator of the enhanced algorithm. Again we use the percentage of how close the result is to the optimal solution to describe the improvement in schedules. Table 4 illustrates that the enhanced algorithm was able to improve the quality of schedules by at least 0.2%, up to 7.13%, with an average improvement of 2.49%.

Table 4: Comparison of solutions (within the same time limit) obtained by current system and system with enhancements to equalization rules for eight different client schedules.

Client	Number of physicians, assignments, days	Closeness of solution to optimal solution (%)	
		Current system	Enhanced system
1	32, 52, 72	9.11%	8.23%
2	30, 35, 403	10.73%	9.22%
3	38, 34, 219	15.57%	8.44%
4	17, 14, 158	46.51%	43.82%
5	36, 47, 73	95.75%	95.56%
6	39, 68, 93	10.15%	9.73%
7	64, 60, 68	3.37%	1.65%
8	104, 105, 65	60.29%	54.95%

## 5 Conclusions

Physician scheduling presents very challenging problems. Our goal in this research is to obtain better schedules and make work-life more enjoyable for physicians. Current state-of-the-art MIP solvers like Xpress, CPLEX and Gurobi can have difficulty in finding the best available schedules.

In this work, we highlight the use of advanced optimization tools for this problem and show how we can work with solvers to improve the solutions for a class of physician scheduling problems that include equalization constraints.

While the benefits of this project cannot be measured directly in dollars, any improvement in the quality of a schedule is valuable for physicians and hospitals. In most scheduling problems it is impossible to find the perfect schedule, so a balance is required between computing times and schedule quality. Having fewer violations of rules can directly impact a physician’s life. For example, in FMCH’s case, we managed to have one fewer equalization rule be violated which can mean freeing up time for a provider who had been covering more than their fair share. In the long term, these types of schedule improvements can reduce physician burnout and increase a hospital’s physician retention rate. There are considerable differences in the resulting schedules, and the number of rule violations has decreased overall with the new algorithm.

We have studied one set of requirements, namely equalization of shift assignments among providers, but there are many more rule types that can be considered for further analysis. As a next step, we are interested in examining other common rule types to find better quality schedules.

## 6 Acknowledgement

We are thankful to Flagstaff Medical Center Hospitals for allowing us to disclose their name and data in this research paper.

## References

- [1] Baum, R., Bertsimas, D., and Kallus, N. (2014). Scheduling, revenue management, and fairness in an academic-hospital radiology division. *Academic radiology*, 21(10):1322–1330.
- [2] Beaulieu, H., Ferland, J. A., Gendron, B., and Michelon, P. (2000). A mathematical programming approach for scheduling physicians in the emergency room. *Health Care Management Science*, 3(3):193–200.
- [3] Block, L., Wu, A. W., Feldman, L., Yeh, H.-C., and Desai, S. V. (2013). Residency schedule, burnout and patient care among first-year residents. *Postgrad. Med. J.*, 89(1055):495–500.
- [4] Bodenheimer, T. and Sinsky, C. (2014). From triple to quadruple aim: Care of the patient requires care of the provider. *Ann Fam Med*, 12(6):573–576.
- [5] Burke, E., Causmaecker, P., Vanden Berghe, G., and Van Landeghem, H. (2004). The state of the art nurse rostering. *Journal of Scheduling*, 7(6):441–499.
- [6] Carter, M. W. and Lapierre, S. D. (2001). Scheduling emergency room physicians. *Health Care Management Science*, 4(4):347–360.
- [7] Cheang, B., Li, H., Lim, A., and Rodrigues, B. (2003). Nurse rostering problems - a bibliographic survey. *European Journal of Operational Research*, 151:447–460.
- [8] Erhard, M., Schoenfelder, J., Fügner, A., and Brunner, J. O. (2016). State of the art in physician scheduling. Working paper.

- [9] Ernst, A. T., Jiang, H., Krishnamoorthy, M., Owens, B., and Sier, D. (2004a). An annotated bibliography of personnel scheduling and rostering. *Annals of Operations Research*, 127:21–144.
- [10] Ernst, A. T., Jiang, H., Krishnamoorthy, M., and Sier, D. (2004b). Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153:3–26.
- [11] Gendreau, M., Farland, J., Gendron, B., Hail, N., Jaumard, B., Lapierre, S., Pesant, G., and Soriano, P. (2006). Physician scheduling in emergency rooms. *Proc. 6th Internat. Conf. Practice Theory Automated Timetabling (PATAT'06), Brno, Czech Republic*, pages 2–14.
- [12] Halbach, M. M., Spann, C. O., and Egan, G. (2003). Effect of sleep deprivation on medical resident and student cognitive function: A prospective study. *Am. J. Obstet. Gynecol.*, 188(5):1198–1201.
- [13] Levin, S., Aronsky, D., Hemphill, R., Han, J., Slagle, J., and France, D. (2007). Shifting towards balance: measuring the distribution of workload among emergency physician teams. *Annals of Emergency Medicine*, 50(4):419–423.
- [14] Lightning Bolt Solutions (June 2016). Physician scheduling complexity by specialty. Accessed March 29, 2017.
- [15] Maenhout, B. and Vanhoucke, M. (2013). An integrated nurse staffing and scheduling analysis for longer-term nursing staff allocation problems. *Omega*, 41:485–499.
- [16] Maslach, C., Jackson, S. E., and Leiter, M. P. (1996). *Maslach burnout inventory manual, 3rd edn.* Palo Alto: Consulting Psychologists Press.
- [17] Rousseau, L., Pesant, G., and Gendreau, M. (2002). A general approach to the physician rostering problem. *Annals of Operations Research*, 115(1):193–205.
- [18] Shanafelt, T. D., Hasan, O., Dyrbye, L. N., Sinsky, C., Satele, D., Sloan, J., and West, C. P. (2015). Changes in burnout and satisfaction with work-life balance in physicians and the general US working population between 2011 and 2014. *Mayo Clin Proc*, 90(12):1600–1613.
- [19] Shanafelt, T. D., Mungo, M., Schmitgen, J., Storz, K. A., Reeves, D., Hayes, S. N., Sloan, J. A., Swensen, S. J., and Buskirk, S. J. (2016). Longitudinal study evaluating the association between physician burnout and changes in professional work effort. *Mayo Clin Proc*, 91(4):422–433.
- [20] Sherali, H. D., Ramahi, M. H., and Saifee, Q. J. (2002). Hospital resident scheduling problem. *Production Planning and Control*, 13(2):220–233.
- [21] Smalley, H. K. and Keskinocak, P. (2016). Automated medical resident rotation and shift scheduling to ensure quality resident education and patient care. *Health Care Management Science*, 19(1):66–88.
- [22] Smalley, H. K., Keskinocak, P., and Vats, A. (2015). Physician scheduling for continuity: An application in pediatric intensive care. *Interfaces*, 45(2):133–148.
- [23] Stolletz, R. and Brunner, J. O. (2012). Fair optimization of fortnightly physician schedules with flexible shifts. *European Journal of Operational Research*, 219(3):622–629.
- [24] Topaloglu, S. (2006). A multi-objective programming model for scheduling emergency medicine residents. *Computers and Industrial Engineering*, 51(3):375–388.

- [25] Van Den Bergh, J., Beliën, J., De Bruecker, P., De Boeck, L., and Demeulemeester, E. (2013). Personnel scheduling: a literature review. *European Journal of Operations Research*, 226(3):367–385.
- [26] Venkataraman, R. and Brusco, M. J. (1996). An integrated analysis of nurse staffing and scheduling policies. *Omega*, 24(1):57–71.
- [27] Volpp, K. G. M. and Grande, D. (2003). Residents’ suggestions for reducing errors in teaching hospitals. *N Engl J Med*, 348(9):851–855.
- [28] West, C. P., Tan, A. D., Habermann, T. M., Sloan, J. A., and Shanafelt, T. D. (2009). Association of resident fatigue and distress with perceived medical errors. *JAMA*, 302(12):1294–1300.
- [29] Wright, P. D. and Mahar, S. (2013). Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega*, 41:1042–1052.