Fair allocation of over- and understaffed shifts in nurse rostering

Lisa Garcia Tercero[1,2][0009-0005-9058-0508], Pieter Smet[1][0000-0002-3955-7725], Dries Goossens[2][0000-0003-0224-3412], and Greet Vanden Berghe[1][0000-0002-0275-5568]

[1] KU Leuven, Department of Computer Science, CODeS, Gebroeders de Smetstraat 1, 9000 Ghent, Belgium
{Lisa.GarciaTercero,Pieter.Smet}@kuleuven.be,
Greet.VandenBerghe@cs.kuleuven.be

[2] Ghent University, Faculty of Economics and Business Administration, Tweekerkenstraat 2, 9000 Ghent, Belgium
Dries.Goossens@ugent.be

1 Introduction

The increasing problem of nurse shortages worldwide poses a high threat to the quality of healthcare [1, 4, 5]. To prevent nurses from leaving the healthcare sector, it is crucial that their job satisfaction is taken into account during the rostering process. An important factor in this matter is to include fairness considerations. Fuchs, Schimmelpfeng and Brunner [3] provide an extensive review of the literature regarding fairness in hospital personnel scheduling, including a wide range of fairness measures to take into account such as the number of night shifts worked, the number of weekends worked and the total number of hours worked. While all of these measures are important to consider, some of them are already implicitly managed by laws and contracts. For example, the number of hours worked (on average) per week is usually determined in an employee's contract, and any overtime hours should be compensated. However, nurse shortages can also lead to understaffed shifts, increasing the workload for nurses on those shifts without any financial compensation. Besides job dissatisfaction, this can lead to reduced focus and quality of care on those shifts and an increased burnout risk for nurses in the long term. Despite these implications, fairness with regard to shift coverage has not been considered before in the literature, possibly because the impact of shift coverage violations is usually considered to be purely financial.
It is not obvious how to assess fairness with regard to understaffed shifts. As in many problems, improving fairness comes with a possible decrease in efficiency. An efficient roster avoids understaffed shifts altogether, but also needs to take into account that more shortages within a single shift create disproportionally more workload for the remaining nurses compared to distributing the same number of shortages over multiple shifts. In addition, it is not efficient to assign

more nurses to a shift than necessary, although it may be required in practice to fulfill all contractual obligations. In this extended abstract we discuss how to model efficiency and fairness for nurse rostering in terms of shift coverage, and we consider the trade-off between both.

2 Problem description and methodology

Table 1 denotes the key sets and parameters of a nurse rostering instance. Note that these can be extended to include any kind of hard constraints, but in this extended abstract we focus on the objectives related to shift coverage. An example of a formulation including additional constraints can be found in [2].

Table 1. Sets and parameters for a nurse rostering problem.

Sets	
D	Days in the decision horizon (index d)
S	Shift types (index s)
N	Nurses (index n)
Parameters	
δ_{ds}	Preferred number of nurses assigned to shift type s on day d
ω_{ds}^{over}	Cost per nurse exceeding the demand δ_{ds} for shift type s on day d
ω_{ds}^{under}	Cost per nurse falling short of the demand δ_{ds} for shift type s on day d
Decision variables	
x_{nds}	1 if nurse n is assigned to shift type s on day d, 0 otherwise
u_{dsc}	1 if there is a shortage of exactly c nurses on shift type s on day d, 0 otherwise $(0 < c \leq \delta_{ds})$
v_{ndsc}	1 if there is a shortage of exactly c nurses on shift type s on day d and nurse n is assigned to that shift, 0 otherwise $(0 < c \leq \delta_{ds})$
o_{ds}	number of nurses exceeding the demand δ_{ds} for shift type s on day d

Below we formulate the linking constraints between the decision variables. The first set of constraints links the overstaffing variables o and the understaffing variables u to the assignment variables x.

$$\sum_{n \in N} x_{nds} - o_{ds} + \sum_{c=1}^{\delta_{ds}} c \cdot u_{dsc} = \delta_{ds} \qquad \forall d \in D, s \in S \quad (1)$$

The other constraints link v to x and u, defining $v_{ndsc} = u_{dsc} \cdot x_{nds}$.

$$v_{ndsc} \leq u_{dsc} \qquad \forall n \in N, d \in D, s \in S, c \in [1, \ldots, \delta_{ds}] \quad (2)$$

$$v_{ndsc} \leq x_{nds} \qquad \forall n \in N, d \in D, s \in S, c \in [1, \ldots, \delta_{ds}] \quad (3)$$

$$v_{ndsc} \geq u_{dsc} + x_{nds} - 1 \qquad \forall n \in N, d \in D, s \in S, c \in [1, \ldots, \delta_{ds}] \quad (4)$$

The objective function includes both an efficiency and a fairness objective. The efficiency objective aims to minimize the total staffing shortage over all shifts, taking into account that the greater the shortage on a given shift, the heavier the workload becomes for the remaining nurses covering that shift. This should be reflected in the objective function, which is why the penalty increases exponentially as more nurses fall short on a given shift. Moreover, an efficient solution also avoids overstaffed shifts. The priority balance between over- and understaffed shifts is determined by their respective weights, where ω_{ds}^{over} is usually much smaller than ω_{ds}^{under}.

$$\min \sum_{d\in D}\sum_{s\in S}\omega_{ds}^{over} o_{ds} + \sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\left(\sum_{c=1}^{\delta_{ds}} u_{dsc}\cdot 3^c\right) \tag{5}$$

The fairness metric considers the number of understaffed shifts each nurse has to cover. From a nurse's point of view, covering an overstaffed shift does not bring any major disadvantages like covering an understaffed shift does. Therefore overstaffing is included in the efficiency objective but not in the fairness objective. The formulation of a fairness objective function is less straightforward and may impact the final solution, which is why we designed three different fairness objectives.
(1) The first fairness objective aims to minimize the exponential cost of understaffed shifts for the worst-off nurse (min-max).

$$\min \max_{n\in N} \sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\sum_{c=1}^{\delta_{ds}} v_{ndsc}\cdot 3^c \tag{6}$$

(2) The second objective aims to minimize the difference between the best-off and the worst-off nurse (min-max spread).

$$\min \max_{n,n'\in N, n'<n} \left|\sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\sum_{c=1}^{\delta_{ds}} v_{ndsc}\cdot 3^c - \sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\sum_{c=1}^{\delta_{ds}} v_{n'dsc}\cdot 3^c\right| \tag{7}$$

(3) The third objective aims to minimize the aggregated difference between all pairs of nurses (min-total spread).

$$\min \sum_{n\in N}\sum_{n'=1}^{n-1} \left|\sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\sum_{c=1}^{\delta_{ds}} v_{ndsc}\cdot 3^c - \sum_{d\in D}\sum_{s\in S}\omega_{ds}^{under}\sum_{c=1}^{\delta_{ds}} v_{n'dsc}\cdot 3^c\right| \tag{8}$$

Additionally, the instances we consider to report computational results include the following hard constraints:

- A lower and upper limit on the number of hours/days worked per nurse;
- Per shift type, an upper limit on the number of these shifts worked per nurse;
- An upper limit on the number of weekends worked per nurse;

– A lower and upper limit on the number of consecutive days worked;
– A lower limit on the number of consecutive days off;
– A set of forbidden shift sequences (e.g., a night shift must not be immediately followed by an early shift);
– Each nurse has a fixed set of days off, on which they must not be assigned.

The size of the model is almost completely determined by the sizes of D, S and N. While the planning horizon can range from two weeks up to a whole year, in practice planning horizon of one or three months are common. The number of shift types per day is often equal to four (early, day, late, night), and a hospital department usually consists of 10 to 30 nurses.

Solving such an IP-model including the efficiency and one of the fairness objectives with different priorities allows us to infer the Pareto front of these two objectives. This provides insight in the "cost" of efficiency when improving fairness and vice versa. While hospitals may initially want to focus on efficiency to optimize the day-to-day workflow, an unfair distribution of these understaffed shifts carries implicit costs in the longer term, such as the risk of nurses having a burnout due to excessive workload. Moreover, the shape of the Pareto front is impacted by the modeling choices, such as the fairness objective function, the formulation of exponential penalties, and the assumption of which constraints are hard and which are soft. Investigating the impact of these specific choices on the solution outcome provides further valuable insights for the planning department of hospitals.

References

1. Chan, Z.C., Tam, W.S., Lung, M.K., Wong, W.Y., Chau, C.W.: A systematic literature review of nurse shortage and the intention to leave. Journal of Nursing Management **21**(4), 605–613 (2013). https://doi.org/https://doi.org/10.1111/j.1365-2834.2012.01437.x
2. Curtois, T., Qu, R.: Computational results on new staff scheduling benchmark instances. ASAP Res. Group, School Comput. Sci., Univ. Nottingham, Nottingham, UK, Tech. Rep (2014)
3. Fuchs, G., Schimmelpfeng, K., Brunner, J.O.: A roadmap for integrating fairness in personnel planning and scheduling in hospitals. Operations Research, Data Analytics and Logistics **45**, 200479 (2025). https://doi.org/https://doi.org/10.1016/j.ordal.2025.200479
4. Goodare, P.: Literature review: Why do we continue to lose our nurses? The Australian Journal of Advanced Nursing **34**(4), 50–56 (2017)
5. World Health Organization: Nursing and midwifery (July 2025), https://www.who.int/news-room/fact-sheets/detail/nursing-and-midwifery, accessed: 2026-03-24