An Integer Programming-Based Local Search Algorithm for the Nurse Scheduling Problem

Juan Carlos Rivera, Sebastián Mesa
Functional Analysis and Applications Research Group
Universidad EAFIT
Medellín, Colombia

June 14, 2015

1 Introduction

Rostering and scheduling problems have been approached formally by researchers for more than 50 years. The rostering and scheduling of hospital personnel,which is a subset of this problem, results particularly complex due to the different needs hospitals have for their different kinds of personnel on the different parts of each day. Since hospitals work 24 hours a day, the benefits of a good schedule can be seen in the performance and well being of the workers and thus in the well being of the patients (Burke *et al.* , 2004).

Up until very little time ago, this kind of problems was solved manually and took a lot of time to be solved and even now, in spite of having computational tools, the hospitals that have them available to them don't use them to their full potential.

Since the decade of the 60's, articles of various authors that use a computational approach to try to solve the problem in general as well as specific to hospital personnel began to be published. The first approaches were mainly mathematical and used lineal models focused on an optimal solution in terms of money and employees' preferences. However, finding the optimal solution to most of the real problems may not be viable and even being so, it may take a considerable amount of time which is not practical since the administration seeks to generate quickly a good schedule that fits all the hard constraints of the problem and as many soft ones as possible.

Later, heuristic and metaheuristic approaches to the problem that focused on solving more complex and real versions of the problem began to appear. These approaches cover a broad range going from purely heuristic approaches using well known techniques such as annealing and taboo search, going through metaheuristic approaches that combine classic optimization with a heuristic component to refine a solution, and coming all the way to techniques that use artificial intelligence and expert systems. Some of these methods have shown promising results and are still being studied.

In the first International Nurse Rostering Competition, the first four places used very different approaches to solve the problem. Valois *et al.* (2010) used a strictly mathematical approach but also partitioned the problem into two subproblems; one for the day and one for the shift of the nurses. Nonobe (2010) used a metaheuristic technique in a constrained optimization problem, using binary variables for the problem and using taboo search. Lu & Hao (2010) used a multistart adaptive local search that used different neighborhoods and different strategies to explore those neighborhoods. Finally, Burke & Curtois (2010) used a previously develop staff rostering model and a variable depth search method.

2 Problem Description and Formulation

Based on the description proposed in The Second International Nurse Rostering Competition (Ceschia *et al.* , 2015) the Nurse Scheduling Problem can be described as follows:

Given a set $N = \{1, ..., n\}$ of n nurses, a schedule consist in assigning a subset of nurses to each shift $s \in S$ during each day $d \in D$ for each required skill $k \in K$, where S is the set of all available shifts, D is the set of days in planning horizon and K is the set of nurse skills. Nurses can work on any shift or they can have a day-off. In this paper, we model days-off as an additional shift, for instance $S = \{1, ..., h, h+1\}$, where h is the number of shifts and $h+1$ represents the day-off. Each nurse can have one or more skills and different requirements can be done for each skill, but in each working shift the nurse can cover exactly one skill request.

Nurses are also differentiate by different kind of contracts: full time, part time, on call, or other. The contracts limit the distribution and the number of assignments within the planning horizon, for instance, the minimum and maximum total number of assignments in the planning horizon, the minimum and maximum number of consecutive working days, the minimum and maximum number of consecutive days-off, the maximum number of working week-ends in the planning horizon, and if the nurses must work complete week-end or they can work just one day in a week-end.

There exists also constraints related to shift types. For instance, for each shift type (early, late, night, etc.), it is given the minimum and maximum number of consecutive assignments of that specific shift. A set of forbidden shift type successions is also given, for instance, if may not be allowed to assign to a nurse an early shift the day after a night one.

Requirements are specified for each day and each shift. The optimal and minimum number of nurses necessary to fulfill the working duties are given for each tuple shift-day.

The nurse assignments also must consider the border data. That means the last assignment in previous periods has influence in some constraints. For instance, if a nurse worked the last day before the period to be scheduled, that day count to compute the number of consecutive working days. In order to consider these features, two types of parameters are defined: parameters BD_{nsk} represent the number of consecutive working days of nurse $n \in N$ in the shift $s \in S$ and skill $k \in K$ until the first day of the planning horizon, and parameters IBD_{nsk} indicates if BD_{nsk} has a non-zero value ($IBD_{nsk} = 1$) or not ($IBD_{nsk} = 0$).

The constraints described below can be classified into hard constraints (H) and soft constraints (S) as follows.

Hard constraints: These constraints must be always satisfied by feasible solutions.

- **H1. Single assignment per day:** A nurse can be assigned to at most one shift per day.
- **H2. Under-staffing:** The number of nurses assigned to each shift $s \in S$ and each skill $k \in K$ during a day $d \in D$ must be at least equal to the minimum requirement MR_{sdk}. The parameter r_{nk} indicates if the nurse $n \in N$ has the skill $k \in K$ ($r_{nk} = 1$), or not ($r_{nk} = 0$). Here the minimum requirement for days-off is equal to zero ($MR_{h+1,d,k} = 0$).
- **H3. Shift type successions:** The shift type assignments of one nurse in two consecutive days must not belong to the set of prohibited successions P. Here, a prohibited shift succession $(s_1, s_2) \in P$ means that if the shift $s_1 \in S$ is performed by a nurse during a day, that nurse cannot perform shift $s_2 \in S$ on the next day.
- **H4. Missing required skill:** A skill requirement during a shift and a day must necessarily be fulfilled by a nurse having that skill.

Soft constraints: These constraints do not have to be satisfied by feasible solutions but it is desirable. They contribute (or they are penalised) on the objective function when they are not satisfied. Each one has a specific weight.

- **S1. Insufficient staffing for optimal coverage:** The number of nurses for each shift $s \in S$ and each skill $k \in K$ must be equal to the optimal requirement RO_{sdk} during the day $d \in D$. Each missing nurse is penalised according to the weight provided C_1. The decision variables M_{sdk} indicate the number of missing nurses with skill $k \in K$ in the shift $s \in S$ on each day $d \in D$. Extra nurses above the optimal value are not considered in the cost.

- **S2. Consecutive assignments per shift:** Minimum and maximum number of consecutive assignments per shift $s \in S$, $MINCAS_s$ and $MAXCAS_s$ respectively, should be respected. Their evaluation involves also the border data. Each extra or missing day is multiplied by the corresponding weight. The integer decision variables $NMCAS_{nsd}$ indicate the number of missing days while the binary decision variables $NECAS_{nsd}$ correspond to the number of exceeded days.

- **S3. Consecutive global assignments:** Minimum and maximum number of consecutive global assignments, $MINCAG_n$ and $MAXCAG_n$, should be respected. Their evaluation involves also the border data. Each extra or missing day is multiplied by the corresponding weight. The decision variables $NMCAG_{nd}$ (integer) and $NECAG_{nd}$ (binary) are, respectively, the number of missing and exceeded consecutive global days.

- **S4. Consecutive days off:** Minimum and maximum number of consecutive days off should be respected. Their evaluation involves also border data. Each extra or missing day is multiplied by the corresponding weight.

- **S5. Preferences:** Each assignment to an undesired shift is penalised by the corresponding weight. The parameter DS_{ns} indicates the desirable level of the shift $S \in S$ for the nurse $n \in N$.

- **S6. Complete week-end:** Some nurses must work complete weekends if her/his contracts indicate that. This feature is given by the parameter W_n which is equal to one if nurse $n \in N$ must work complete week-end or zero otherwise. Those nurses can work both week-end days or none. If they work only one of the two days, this is penalised by the corresponding weight. The decision variables MDW_{nd} count the missing days in a week-end for the nurse n.

- **S7. Total assignments:** For each nurse, the total number of assignments (working days) must be included within the limits (the minimum $MINWD_n$ and the maximum $MAXWD_n$) enforced by her/his contract. The difference (in either direction), multiplied by its weight, is added to the objective function. The decision variables $NMWD_n$ and $NEWD_n$ compute the number of missing working days and number of exceeded working days respectively.

- **S8. Total working week-ends:** For each nurse $n \in N$, the number of working week-ends must be less than or equal to the maximum $MAXWW_n$. The number of worked week-ends in excess is add to the objective function multiplied by the weight. A week-end is considered "working" if at least one of the two days is busy for the nurse, and in that case the binary decision variable WW_nd must be equal to 1. The decision variable $NEWW_n$ counts the number of exceeded working week-ends.

To model this optimization problem, we propose a Mixed Integer Linear Program (MILP). In that model, the binary decision variable x_{nsdk} is equal to 1 if the nurse $n \in N$ is assigned to the shift $s \in S$ during the day $d \in D$ to perform its skill $k \in K$.

The MILP can be described by the Equations (1) to (27):

$$\min Z = \Delta Z_1 + \Delta Z_2 + \Delta Z_3 + \Delta Z_4 + \Delta Z_5 + \Delta Z_6 + \Delta Z_7 \tag{1}$$

$$s.t. \sum_{s \in S} \sum_{k \in K} x_{nsdk} = 1, \quad \forall\ n \in N,\ d \in D \tag{2}$$

$$\sum_{n \in N} x_{nsdk} \cdot r_{nk} \geq RM_{sdk}, \quad \forall\ s \in S,\ d \in D,\ k \in K \tag{3}$$

$$\begin{aligned} &\sum_{k \in K} (x_{n,s_1,d-1,k} + x_{n,s_2,d,k}) \leq 1, \\ &\forall\ n \in N,\ d \in D \backslash \{1\},\ (s_1, s_2) \in P \end{aligned} \tag{4}$$

$$\sum_{n \in N} x_{nsdk} \cdot r_{nk} + M_{sdk} \geq RO_{sdk}, \quad \forall\ s \in S,\ d \in D,\ k \in K \tag{5}$$

$$\Delta Z_1 = C_1 \cdot \sum_{s \in S} \sum_{k \in K} M_{sdk} \tag{6}$$

$$\begin{aligned} &\sum_{d=d_0}^{d_f} \sum_{k \in K} x_{nsdk} + NMCAS_{nsd_0} \geq \\ &MINCAS_s \cdot \sum_{k \in K} (x_{nsd_0k} - x_{n,s,d_0-1,k}), \\ &\forall\ n \in N,\ s \in S,\ d_0 \in D \setminus \{1\},\ d_0 \leq |D| - MINCAS_s + 1, \\ &d_f = d_0 + MINCAS_s - 1 \end{aligned} \tag{7}$$

$$\begin{aligned} &\sum_{k \in K} BD_{nsk} + \sum_{k \in K} \sum_{d=1}^{d_f} x_{nsdk} + NMCAS_{n,s,1} \geq \\ &MINCAS_s \cdot \sum_{k \in K} (x_{n,s,1,k} - IBD_{nsk}), \\ &\forall\ n \in N,\ s \in S,\ d_f = MINCAS_s - BD_{nsk} \end{aligned} \tag{8}$$

$$\sum_{d=d_0}^{d_f}\sum_{k\in K} x_{nsdk} - NECAS_{nsd_0} \leq MAXCAS_s, \qquad (9)$$
$$\forall\ n\in N,\ s\in S,\ d_0\in D\setminus\{1\},\ d_f=\min\{d_0+MAXCAS_s, |D|\}$$

$$\sum_{k\in K} BD_{nsk} + \sum_{k\in K}\sum_{d=1}^{d_f} x_{nsdk} - NECAS_{n,s,1} \leq MAXCAS_s, \qquad (10)$$
$$\forall\ n\in N,\ s\in S,\ d_f = MAXCAS_s - BD_{nsk}+1$$

$$\Delta Z_2 = C_2\cdot\sum_{n\in N}\sum_{s\in S}\sum_{d\in D}(NMCAS_{nsd}+NECAS_{nsd}) \qquad (11)$$

$$\sum_{d=d_0}^{d_f}\sum_{s=1}^{h}\sum_{k\in K} x_{nsdk} + NMCAG_{nd_0} \geq$$
$$MINCAG_n\cdot\sum_{s\in S}\sum_{k\in K}(x_{nsd_0k} - x_{n,s,d_0-1,k}), \qquad (12)$$
$$\forall\ n\in N,\ d_0\in D\setminus\{1\},\ d_0\leq |D| - MINCAG_n+1,$$
$$d_f = d_0 + MINCAG_n - 1$$

$$\sum_{s=1}^{h}\sum_{k\in K} BD_{nsk} + \sum_{s=1}^{h}\sum_{k\in K}\sum_{d=1}^{d_f} x_{nsdk} + NMCAG_{n,1} \geq$$
$$MINCAG_n\cdot\sum_{s=1}^{h}\sum_{k\in K} x_{n,s,1,k}, \qquad (13)$$
$$\forall\ n\in N,\ d_f = MINCAG_n - BD_{nsk}$$

$$\sum_{d=d_0}^{d_f}\sum_{s=1}^{h}\sum_{k\in K} x_{nsdk} - NECAG_{nd_0}\leq MAXCAG_n, \qquad (14)$$
$$\forall\ n\in N,\ d_0\in D\setminus\{1\},\ d_f=\min\{d_0+MAXCAG_n, |D|\}$$

$$\sum_{s=1}^{h}\sum_{k\in K} BD_{nsk} + \sum_{s=1}^{h}\sum_{k\in K}\sum_{d=1}^{d_f} x_{nsdk} - NECAG_{n,1}\leq MAXCAG_n, \qquad (15)$$
$$\forall\ n\in N,\ d_f = MAXCAG_n - BD_{nsk}+1$$

$$\Delta Z_3 = C_3\cdot\sum_{n\in N}\sum_{d\in D}(NMCAG_{nd}+NECAG_{nd}) \qquad (16)$$

$$\Delta Z_4 = C_4\cdot\sum_{n\in N}\sum_{s\in S}\sum_{d\in D}\sum_{k\in K} DS_{ns}\cdot x_{nsdk} \qquad (17)$$

$$\sum_{k\in K}\sum_{s=1}^{h} x_{nsdk} - MDW_{nd} \leq \sum_{k\in K}\sum_{s=1}^{h} x_{n,s,d-1,k} + (1 - W_n),$$
$$\forall\ n \in N,\ d \in \{7, 14, 21, 28\} \tag{18}$$

$$MDW_{nd} + \sum_{k\in K}\sum_{s=1}^{h} x_{nsdk} \geq \sum_{k\in K}\sum_{s=1}^{h} x_{n,s,d-1,k} + (1 - W_n),$$
$$\forall\ n \in N,\ d \in \{7, 14, 21, 28\} \tag{19}$$

$$\Delta Z_5 = C_5 \cdot \sum_{n\in N}\sum_{d\in D'} MDW_{nd}, \quad \text{where } D' = \{7, 14, 21, 28\} \tag{20}$$

$$\sum_{d\in D}\sum_{s=1}^{h}\sum_{k\in K} x_{nsdk} + NMWD_n \geq MINWD_n, \quad \forall\ n \in N \tag{21}$$

$$\sum_{d\in D}\sum_{s=1}^{h}\sum_{k\in K} x_{nsdk} - NEWD_n \leq MAXWD_n, \quad \forall\ n \in N \tag{22}$$

$$\Delta Z_6 = C_6 \cdot \sum_{n\in N} (NMWD_n + NEWD_n) \tag{23}$$

$$\sum_{s=1}^{h}\sum_{k\in K} (x_{nsdk} + x_{n,s,d-1,k}) \leq 2 \cdot WW_{nd}$$
$$\forall\ n \in N,\ d \in \{7, 14, 21, 28\} \tag{24}$$

$$\sum_{s=1}^{h}\sum_{k\in K} (x_{nsdk} + x_{n,s,d-1,k}) \geq WW_{nd}$$
$$\forall\ n \in N,\ d \in \{7, 14, 21, 28\} \tag{25}$$

$$\sum_{d\in D'} WW_{nd} - NEWW_n \leq MAXWW_n,$$
$$\forall\ n \in N, \quad \text{where } D' = \{7, 14, 21, 28\} \tag{26}$$

$$\Delta Z_7 = C_7 \cdot \sum_{n\in N} NEWW_n \tag{27}$$

The objective function (1) is the sum of all soft constraint penalisations.

Equations (2) to (4) are referred to the hard constraints. Equation (2) means that each nurse must be assigned to only one shift on each day, as it is explained by constraint H1. Remind that here the set of shifts include days-off ($s = h+1$). In Equation (3) the minimum required number of nurses with each skill on each shift and day is limited as expressed in constraints H2. The presence of the parameter r_{nk} assures that the assigned nurses have the required skill as it is asked by constraints H4. Equation (4) guarantees that constraints H3 about consecutive shift assignments are satisfied.

Equations (5) to (27) are referred to the soft constraints. Constraint (5) allows to compute the number of nurses required to reach optimal number of nurses with each skill on each shift and each day, according to constraint S1. The number of missing nurses is given by the decision variables M_{sdk}. Note that in these constraints only nurses with the specific skills are considered, so constraint H4 is also satisfied. Equation (6) define the total violation cost of constraints S1. In all tested instances the parameter C_1 is set to 30.

In constraints (7) to (11) the consecutive assignments per shift constraints are represented, according to constraints S2. Equation (7) is related to the minimum consecutive assignments per shift for days after the first day in the planning horizon, while equation (8) corresponds to the first day of the planning horizon. Similarly, constraints (9) and (10) describe the maximum consecutive assignments per shift for days after and the first day on the planning horizon. In equation (11) the auxiliar variable ΔZ_2 cumulates the total violation cost of constraints S2. Note that as day-off are treated as an additional or dummy working day, the constraints S4 are also represented by these equations. In all tested instances the parameter C_2 is set to 15 for real working days ($s \neq h+1$) and 10 for days-off ($s = h+1$).

Constraints S3 are modeled by Equations (12) to (16). As Constraints (7) to (11), Equation (12) represents the minimum number of global consecutive working days after first day, Equation (13) limits the minimum number of global consecutive working days in the first day of planning horizon, equations (14) and (15) have similar meanings but respect to the maximum number of global consecutive working days, and equation (16) compute the violation cost of constraints S3. The weights for consecutive days is $C_3 = 30$.

The Constraint (17) indicates the total cost of undesired assignment shifts explained by soft constraints S5 The value of parameter C_4 is set to 10. The Equations (18) and (19) allow to compute the missing days in a working week-end for nurse who must work complete week-ends. Equation (20) compute the violation cost associated to complete working week-end constraints. Here is supposed that the first day in the planning horizon is Monday. The violation weight of these constraints is 30.

Equations (21) and (22) are referred to the minimum and maximum number of working days of nurses in the planning horizon. Equation (23) cumulates the violation cost of these constraints. The parameter C_6 is set to 20.

Constraints (24) and (25) allow to compute variables WW_{nd} while equations (26) use that variables to calculate the number of exceeded working weekends. Finally, Equation (27) compute the total cost of exceeded working week-ends.

3 Conclusions and Future work

The scheduling problem in general is an increasingly difficult problem due to the growing number of specializations in every field and the requirements in time and money they have. Most, if not all, of the different approaches that have been used to create solutions to the problem should be taken into consideration when building a new model and the strategies to solve it because even now some forms of mathematical approaches, which could be thought to be less efficient than more complex kinds of approaches, have shown very good results. The implementation of the model with heuristic strategies and other ideas taken from the first places of the First Nurse Rostering competition are left for future work. Strategies taken into consideration for implementation are the partition of the problem into sub problems for day and shifts and heuristic techniques such as variable neighborhood search.

References

Burke, Edmund K., & Curtois, Tim. 2010. *An ejection chain method and a branch and price algorithm applied to the instances of the first international nurse rosteringcompetition.* INC2010.

Burke, Edmund K., Causmaecker, Patrick De, Berghe, Greet Vander, & Landeghem, Hendrik Van. 2004. The State of the Art of Nurse Rostering. *Journal of Scheduling*, 441–499.

Ceschia, Sara, Dang, Nguyen Thi Thanh, Causmaecker, Patrick De, Haspeslagh, Stefaan, & Scheerf, Andrea. 2015. *Second International Nurse Rostering Competition (INRC-II) -Problem Description and Rules-.* Tech. rept.

Lu, Zhipeng, & Hao, Jin-Kao. 2010. *Adaptive local search for the first international nurse rosteringconpetition.* INC2010.

Nonobe, Koji. 2010. *An approach using a general constraint optimization solver.* INC2010.

Valois, Christos, Gogos, Christos, Goulas, George, Alefragis, Panayiotis, & Housos, Efthymios. 2010. *A systematic two phase approach for the nurse rostering problem".* INC2010.