Available online at www.sciencedirect.com
ScienceDirect
Electronic Notes in DISCRETE MATHEMATICS
Electronic Notes in Discrete Mathematics 58 (2017) 31–38
www.elsevier.com/locate/endm

Variable neighborhood search accelerated column generation for the nurse rostering problem

Rafael A.M. Gomes[a] [1] Túlio A.M. Toffolo[a,b]
Haroldo Gambini Santos[a]

[a] *Department of Computing, Federal University of Ouro Preto, Brazil*
[b] *Computer Science Department, KU Leuven, Belgium*

Abstract

The Nurse Rostering Problem (NRP) is an optimization problem where nurses with specific skills must be assigned shifts in a schedule. The objective is to obtain a feasible solution while minimizing the number of soft constraint violations. This work presents a Variable Neighborhood Search accelerated Column Generation procedure for the NRP in addition to a Relax-and-fix Heuristic for obtaining feasible solutions. The algorithm improved the best known solutions by at least 10% for all 29 hidden instances from the Second International Nurse Rostering Competition (2014) with 4 weeks. The improved solutions have an optimality gap of at most 8%.

Keywords: nurse rostering · column generation · integer programming · heuristics

[1] Email: `rafael.amgomes@gmail.com`

http://dx.doi.org/10.1016/j.endm.2017.03.005
1571-0653/

1 Introduction

There has been much research regarding solution methods for the Nurse Rostering Problem (NRP) [1]. In 2010, the First Nurse Rostering International Competition (INRC-I) was proposed [5] and a significant number of different algorithms were presented for its large set of proposed instances. Most instances were exactly solved by [9] in 2012.

The Second International Nurse Rostering Competition (INRC-II) occurred during 2014 [10]. In contrast to the first competition, a smaller set of constraints was introduced. However, multiple planning horizons were considered, simulating real-world situations [3].

This work presents an Integer Programming formulation for the INRC-II problem with an exponential number of variables. Various Column Generation techniques are evaluated for handling this formulation. Besides solving the pricing problem using Integer Programming, we show how combining the method with heuristics significantly accelerates the method's convergence. Our computational experiments showed that methods for the production of good primal and dual bounds can be obtained using the proposed techniques.

The paper is organized as follows: Section 2 the NRP problem is formally stated by the proposed Integer Programming formulation. Section 3 introduces the Variable Neighborhood Search (VNS) heuristic applied to accelerate column generation. Section 4 presents Relax-and-fix Heuristics approaches for obtaining feasible solutions from linear-relaxed fractional solutions. Section 5 describes the computational experiments conducted and evaluates the performance of our proposal. Finally, conclusions and future works are detailed in Section 6.

2 The Nurse Rostering Problem Formulation for INRC-II

The present paper proposes a formulation for the INRC-II problem containing a large number of variables. The relaxation of this formulation is solved by Column Generation [2] and generally provides strong lower bounds.

In the proposed formulation, N, D, S and K_n represent the set of nurses, days, shifts and skills of nurses n, respectively. Ω_n represents the set of all possible allocation patterns for nurse n, λ_{np} indicates whether allocation $p \in \Omega_n$ is selected for nurse n (λ_{np}=1) or not (λ_{np}=0) and v_{dsk} measures the difference between the required and allocated number of nurses on day d and shift s for skill k. Equations and inequalities (1)–(5) present the proposed

formulation.
Minimize:

$$\sum_{n\in N}\sum_{p\in\Omega_n} c_{np}\lambda_{np} + \sum_{d\in D}\sum_{s\in S}\sum_{k\in K}\omega v_{dsk} \tag{1}$$

Subject to:

$$\sum_{p\in\Omega_n} \lambda_{np} = 1 \ \forall n \in N \tag{2}$$

$$\sum_{p\in\Omega_n} \alpha_{npdsk}\lambda_{np} - x_{ndsk} = 0 \ \forall n \in N, d \in D, s \in S, k \in K_n \tag{3}$$

$$v_{dsk} - (r^{\star}_{dsk} - r^{-}_{dsk}) \le 0 \ \forall d \in D, s \in S, k \in K_n \tag{4}$$

$$\sum_{n\in N} x_{ndsk} + v_{dsk} - r^{\star}_{dsk} = 0 \ \forall d \in D, s \in S, k \in K_n \tag{5}$$

The Objective function (1) minimizes all violations of soft constraints by way of variables λ_{np} and costs c_{np}. The variable v_{dsk} is penalized by weight ω for each unit below the optimal coverage. Constraints (2) require the selection of exactly one pattern for each nurse. Constraints (3) ensure that variable x is only active if a pattern in which the respective allocation exists is active. Finally, Constraints (4) and (5) measure how well-satisfied the minimum and optimal demands are, where r^{-}_{dsk} and $r^{\star}_{dsk}$ represent the minimum and optimal number of required nurses for day d, shift s and skill k, respectively.

2.1 The Pricing Problem

The pricing problem requires finding allocation patterns with negative reduced cost for each nurse, considering all planning horizon days. For each nurse, a subproblem is defined considering dual values μ and π_{dsk} for Constraints (2) and (3), respectively. Main decision variables are x_{dsk}, which are assigned value 1 if the nurse works on shift s of day d using skill k and zero otherwise.

The pricing problem may be solved exactly using an Integer Programming formulation very similar to that used in [9]. This formulation includes additional binary variables concerning the selection of contiguous work and resting days. Despite the formulation itself being rather weak, it may be strengthened with inequalities such as cliques and odd-holes derived from the Set Packing Polytope. The formulation proves extremely practical for solving medium-sized NRP instances, as shown in [9].

3 Variable Neighborhood Search Accelerated Column Generation

Heuristic procedures are employed to accelerate the production of feasible columns with negative reduced costs. In this phase, a solution is given by a vector S of size $|D|$ with each cell d_i representing the allocation $(^{s}/_{k})$ on a specified day where the index s represents the shift and k the skill. The value $(^{-}/_{-})$ corresponds to a resting day.

Initial feasible solutions begin from an empty solution where, for each non-satisfied demand, a qualified nurse is randomly selected and allocated. The procedure continues until every day/shift/skill has its demands satisfied. The local search phase begins once a feasible solution is obtained.

The local search phase is performed by the meta-heuristic Variable Neighborhood Search (VNS), proposed by [8]. VNS systematically perturbs the current solution (shake procedure) before executing a multi-neighborhood descent phase. The present paper implements the most common VNS variant, known as Sequential Variable Neighborhood Descent (SVND) [7], which employs the Variable Neighborhood Descent (VND) method as local search within the VNS.

Four neighborhood structures are considered within the VNS:
Change Allocation of one Day - $\mathcal{N}^{CA}(s)$: altering the allocation of a solution on a specific day by adjusting the work shift or skill.
Change Allocation Working Windows - $\mathcal{N}^{CW}(s)$: consists of changing allocations of contiguous working days (working windows) for the same allocation pattern.
Invert Working Windows - $\mathcal{N}^{IW}(s)$: working and resting windows are initially computed and then subsequently exchanged; uniform allocations are inserted into the new working windows.
Swap Allocation - $\mathcal{N}^{SA}(s)$: swaps the allocations of two different days.

Table 3 depicts an example with an initial solution s and a solution modified by each of the considered neighborhoods. The four neighborhoods are presented in the same order as used within the VND procedure.

Table 1
Illustration of local searches applied to solution s

	0	1	2	3	4	5	6	7	8	9	10
s	$^{-}/_{-}$	$^{D}/_{H}$	$^{L}/_{H}$	$^{L}/_{H}$	$^{N}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{E}/_{N}$	$^{E}/_{N}$
$s \oplus m^{CA}$	$^{-}/_{-}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{L}/_{H}$	$^{N}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{E}/_{N}$	$^{E}/_{N}$
$s \oplus m^{CW}$	$^{-}/_{-}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{D}/_{H}$	$^{D}/_{H}$
$s \oplus m^{IW}$	$^{D}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{D}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$
$s \oplus m^{SA}$	$^{-}/_{-}$	$^{D}/_{H}$	$^{L}/_{H}$	$^{L}/_{H}$	$^{N}/_{H}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{E}/_{N}$	$^{-}/_{-}$	$^{-}/_{-}$	$^{E}/_{N}$

The proposed VNS procedure maintains a pool (Elite Set) of the best solutions obtained during the search. Whenever a local optimum for all neighborhoods is generated, the shake procedure is activated whereby a percentage of allocations are randomly changed before continuing the search. The algorithm ends when the maximum number of iterations without improvement is reached or when the Elite Set size reaches a defined threshold.

4 Relax-and-fix Heuristics

Diving and Rounding heuristics [6] are employed to obtain integer feasible solutions. We use diving (also know as *Relax-and-fix*) to reach deep branching nodes while maintaining a good LP relaxation lower bound value [4]. The method dives on a branch-and-bound tree until a feasible solution is found. Once the master problem's linear relaxation is solved, a rounding procedure is applied that fixes all fractional variables above a threshold β to one before subsequently re-optimizing the model. Generating new columns when re-optimizing increases the changes of producing feasible solutions. For this reason, this cycle is repeated in the following iteration. Whenever no variable has a fractional value greater than β, the fractional variable with value closest to one is selected and then diving is performed again.

5 Computational Experiments

All algorithms were implemented in Java 8 and the experiments executed on an Intel®Xeon E5620 @ 2.40GHz 8-Core computer with 48 Gb of RAM running the CentOS Linux 7.2.1511 operating system. The column generation procedure is executed in parallel, such that the pricing problems of all nurses are solved concurrently.

Rather than considering multiple weeks separately, as in the competition, different weeks are concatenated and the resultant single horizon is solved. The goal is thus to solve the complete schedule and enable better evaluation of strategies considering weeks separately. Furthermore, analyzing global optimal solutions may provide insights for strategies considering separate weeks.

The pricing procedure begins with many threads (one per nurse) running the VNS algorithm. If these threads succeed in finding new columns, they are inserted and the next iteration begins. Otherwise, the MIP-based column generation procedure is executed. Whenever the execution of the VNS heuristic or the MIP based pricing is concluded, all new columns found with negative reduced cost are inserted in the master problem.

Computational experiments demonstrate that the addition of a VNS heuristic within the column generation algorithm is a determining factor to accelerate the convergence of the optimal solution from the linear programming relaxation of the master problem, as demonstrated in Figure 1. Note that convergence is significantly delayed without the VNS procedure. Experiments shown in Figure 1 include executions with different values for the most influential parameters of the algorithm which according to our preliminary experiments were the size of the elite set (es) and the shake percentage (sp). The maximum number of iterations without improvement was set to 20.

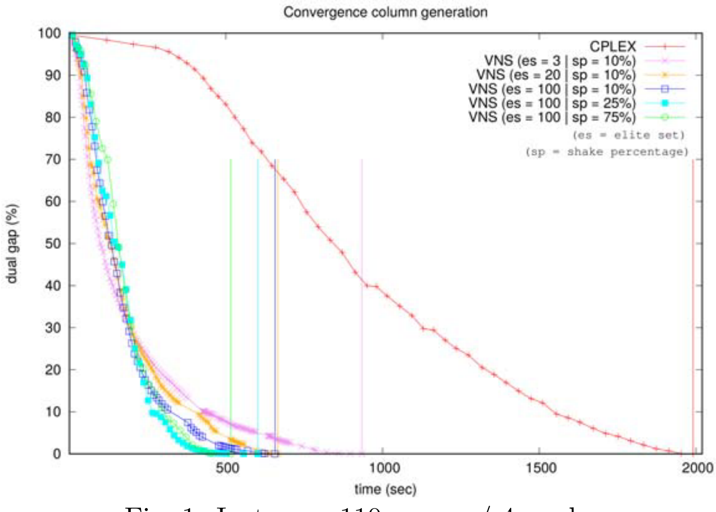

Fig. 1. Instance: 110 nurses / 4 weeks.

5.1 Best Results

The algorithm was executed on all 29 INRC-II Challenge instances containing between 35 and 110 nurses and the results are provided by Table 2. Instance names are abbreviated: n is the number of nurses, h refers to the history from previous week and w indicates the execution order of weeks forming the planning horizon. Table 2 presents the best known solution cost (BKS), the linear relaxation value (opt) and the linear relaxation rounded value ($\overline{opt}$)[2]. The solution cost obtained by the Relax-and-fix heuristic for each instance is presented. For all evaluated instances, the proposed algorithm obtained better solutions (obj) than the previous BKS, as can be seen the column gb which indicates the gap over of BKS. Note that the duality gap (dg) values demonstrate how the costs obtained are close to the linear relaxation optimal ($\overline{opt}$), indicating both solution and lower bound quality.

6 Conclusions and Future Research

This paper presented techniques capable of handling a large integer programming formulation for the Nurse Rostering Problem. By combining MIP and heuristic procedures, the proposed algorithm proved capable of not only providing strong lower bounds, but also of improving the best known solutions for all 29 hidden INRC-II instances with 4-week scheduling horizons. The final computed duality gap was less than 8% for all instances. A fundamental component of the method is a heuristic based on Variable Neighborhood Search procedure which significantly speeds up the convergence of the column generation procedure.

[2] objective function coefficients are multiple of 5.

Table 2
Final results

Instance			BKS	Relaxation			Relax-and-fix Heuristic			
n	h	w		*time*	*opt*	$\overline{opt}$	*time*	*obj*	$gb(\%)$	$dg(\%)$
35	0	1-7-1-8	1630	1129	1337.1	1340	3269	**1425**	-12.58	6.34
	0	4-2-1-6	1800	896	1596.9	1600	5124	**1615**	-10.28	0.94
	0	5-9-5-6	1755	980	1492.5	1495	6872	**1540**	-12.25	3.01
	0	9-8-7-7	1540	986	1318.1	1320	4475	**1365**	-11.36	3.41
	1	0-6-9-2	1500	672	1285.4	1290	5359	**1385**	-7.67	7.36
	2	8-6-7-1	1490	1370	1250.7	1255	6453	**1335**	-10.40	6.37
	2	8-8-7-5	1255	1580	1075.6	1080	5586	**1085**	-13.55	0.46
	2	9-7-2-2	1705	1052	1493.8	1495	6204	**1525**	-10.56	2.01
	2	9-7-2-2	1650	984	1461.9	1465	12340	**1480**	-10.30	1.02
70	0	3-6-5-1	2700	832	2387.6	2390	3640	**2460**	-8.89	2.93
	0	4-9-6-7	2430	747	2197.5	2200	4943	**2330**	-4.12	5.91
	0	4-9-7-6	2475	854	2170.9	2175	9465	**2315**	-6.46	6.44
	0	8-6-0-8	2435	647	2285.0	2285	1795	**2400**	-1.44	5.03
	0	9-1-7-5	2320	847	2129.1	2130	3395	**2225**	-4.09	4.46
	1	1-3-8-8	2700	726	2510.3	2515	3457	**2615**	-3.15	3.98
	2	0-5-6-8	2520	822	2299.6	2300	2990	**2415**	-4.17	5.00
	2	3-5-8-2	2615	922	2340.8	2345	5032	**2405**	-8.03	2.56
	2	5-8-2-5	2540	897	2315.9	2320	7580	**2390**	-5.91	3.02
	2	9-5-6-5	2615	677	2387.0	2390	2495	**2480**	-5.16	3.77
110	0	1-4-2-8	2710	3344	2438.0	2440	13084	**2560**	-5.54	4.92
	0	1-9-3-5	2920	2455	2560.0	2560	9624	**2640**	-9.59	3.13
	1	0-1-6-4	2850	3735	2665.5	2670	24585	**2690**	-5.61	0.75
	1	0-5-8-8	2820	2052	2594.1	2595	12838	**2705**	-4.08	4.24
	1	2-9-2-0	3345	3919	3088.5	3090	11570	**3170**	-5.23	2.59
	1	4-8-7-2	2805	6758	2600.0	2600	8350	**2630**	-6.24	1.15
	2	0-2-7-0	3005	4926	2860.0	2860	10882	**2960**	-1.50	3.50
	2	5-1-3-0	2925	3535	2720.0	2720	9079	**2770**	-5.30	1.84
	2	8-9-9-2	3415	9578	3020.6	3025	15184	**3140**	-8.05	3.80
	2	9-8-4-9	3135	2482	2829.5	2830	11311	**3005**	-4.15	6.18

Future research aims to develop more advanced Branch-and-Cut-and-Price procedures based on the proposed formulation.

References

[1] Burke, E. K., P. De Causmaecker, G. V. Berghe and H. Van Landeghem, *The state of the art of nurse rostering*, Journal of Scheduling **7** (2004), pp. 441–499.

[2] Dantzig, G. B. and P. Wolfe, *Decomposition principle for linear programs*, Operations Research **8** (1960), pp. 101–111.

[3] De Causmaecker, P., P. Demeester, G. V. Berghe and B. Verbeke, *Analysis of real-world personnel scheduling problems*, in: *Proceedings of the 5th International Conference on Practice and Theory of Automated Timetabling, Pittsburgh*, 2004, pp. 183–197.

[4] Fischetti, M. and A. Lodi, *Heuristics in mixed integer programming*, in: J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh and J. C. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*, John Wiley & Sons, Inc., 2010 .

[5] Haspeslagh, S., P. DeCausmaecker, A. Schaerf and M. Stølevik, *The 1st international nurse rostering competition 2010*, Annals of Operations Research **218** (2012), pp. 221–236.

[6] Joncour, C., S. Michel, R. Sadykov, D. Sverdlov and F. Vanderbeck, *Column generation based primal heuristics*, Electronic Notes in Discrete Mathematics **36** (2010), pp. 695–702, International Symposium on Combinatorial Optimization (ISCO 2010).

[7] Mjirda, A., R. Todosijević, S. Hanafi, P. Hansen and N. Mladenović, *Sequential variable neighborhood descent variants: an empirical study on the traveling salesman problem*, International Transactions in Operational Research **24** (2017), pp. 615–633.

[8] Mladenović, N. and P. Hansen, *Variable neighborhood search*, Computers & Operations Research **24** (1997), pp. 1097–1100.

[9] Santos, H. G., T. A. M. Toffolo, R. A. M. Gomes and S. Ribas, *Integer programming techniques for the nurse rostering problem*, Annals of Operations Research **239** (2014), pp. 225–251.

[10] Sara, C., N. T. T. Dang, P. De Causmaecker, S. Haspeslagh and A. Schaerf, *The 2nd International Nurse Rostering Competition*, in: E. Özcan, E. K. Burke and B. McCollum, editors, *Proceedings of the 10th International Conference of the Practice and Theory of Automated Timetabling (PATAT 2014)*, York, UK, 2014, pp. 554–556.