A Solution Technique for Rostering in Hospital Departments

T. Adams, C. Walker, M. O'Sullivan
Department of Engineering Science
University of Auckland
New Zealand
tada955@auckland.ac.nz

Abstract

Rosters in hospital departments directly affect the number of patients that physicians are caring for each day, because the roster dictates who admits new patients. Therefore, balance of workload across physicians can be controlled to a degree by how the physicians are rostered. To build a roster exhibiting a good workload balance across physicians a mixed integer programme is solved, that utilises patient admission and discharge numbers.

Traditional rostering models do not include patient admissions and discharges. Including this data results in a model that is larger and more difficult to solve. In particular, although the 'vanilla' formulation can be solved for small departments (including admissions and discharge data over moderate time windows), it does not scale well, and so is less applicable for larger problem instances.

To reduce the time it takes to solve the model a new solution technique is developed, within the context of creating a roster for the General Medicine Department at Waitemata District Health Board.

Key words: Mixed integer programming, Rostering, Healthcare modelling.

1 Introduction

Keeping the number of patients each physician is caring for balanced is of concern to a hospital for two reasons. First, balanced workloads are fairer for the physicians. Second, unbalanced workloads lead to patients being transferred between physicians because one physician has too many patients and another too few. The transfer of patients between physicians is undesirable as it takes time for the new physician and the patient to become acquainted, both personally and medically.

Maintaining a continuous relationship between physician and patient is one of the most important concerns for hospitals and has been shown to reduce length of stay for some patients (Epstein et al. 2010). This continuity of care was the most important factor for Waitemata District Health Board (WDHB) when creating a new roster for the General Medicine (GM) department at Waitakere Hospital. By carefully planning which physician admits each day the workloads of the physicians can be balanced, which in turn improves the continuity of care. More details of the rostering problem at Waitakere Hospital are available in (Adams et al. 2016).

To create a roster that balances the workloads of the physicians data on the number of patients admitted and discharged each day is required. This data is used to estimate the physicians' workloads for a given roster. Traditional rostering approaches have not used this information and instead focused on minimising the cost to the hospital of staffing the roster, while meeting staffing requirements for physicians (Bailey 1985; Wolfe 1979).

More recently several attempts have been made to create rosters that improve continuity of care for patients. The number of patients handed over to another physician at the end of a shift is minimised, either by estimating the number of handoffs that occur at different times of day and structuring shifts to minimise this (Kazemian et al. 2014); or by developing a handoff metric based on how recently each physician worked (Smalley, Keskinocak, and Vats 2015).

Neither of these methods link patient admissions to physicians and therefore cannot calculate the number of patients each physician is caring for. Calculating the workloads of the physicians increases the size and complexity of the mixed integer programme used to create a roster. How much of an effect this has on the time taken to create a roster depends on both the number of physicians being rostered and the length of the time period over which the workloads of the physicians are to be balanced, called the planning horizon. Increasing either the number of physicians or the planning horizon increases the time to create a roster; therefore techniques that reduce this time are necessary to create rosters for departments with many physicians for a long planning horizon.

2 Methods

A mixed integer programme is used to create a roster which balances workloads across the physicians. To achieve this the model's objective function minimises the largest difference in workloads between any two physicians at any point in the planning horizon. Within the model, decision variables determine the roster for the physicians and the workloads of the physicians. The constraints of the model can be categorised as three distinct types: roster rules, workload calculation, and workload difference.

The first type of constraints make sure that the roster obeys all of the rules for a physicians' roster. These rules include things such as: a minimum number of physicians admitting patients each day; sequences of shifts such that if the first is worked the others must also be worked, for example a run of night shifts; combinations and sequences of shifts that cannot be worked.

The second type of constraints are used to calculate the workloads of the physicians. The workloads are determined using current assignment of shifts to physicians in the model and patient admission and discharge data. The data for each day in the planning horizon is divided into separate time periods, or segments. The workloads can then be calculated for each physician at the end of each segment.

Figure 1 depicts the calculation of workloads of the physicians for one segment of one day. The number of admissions in the segment, taken from the data, is split evenly between the physicians that are working the admission shift in that segment. The number of patients that each physician was caring for in the previous segment and have not been discharged are added to their workload. Finally a fraction of the total workload for each physician is discharged; this fraction is based on the proportion of patients that were discharged in that segment in the data.

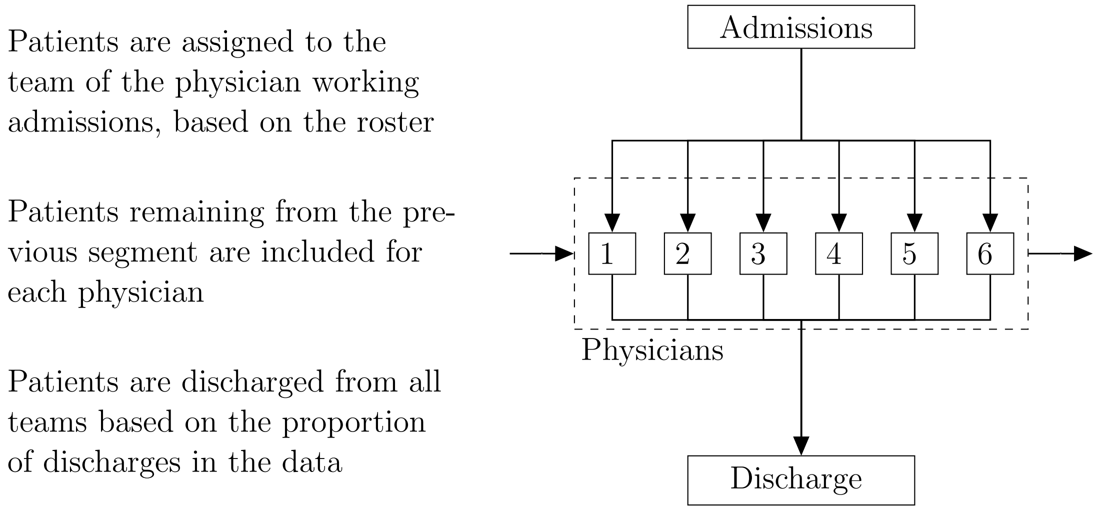

Figure 1: Calculating Physician Workloads

It is assumed that patients are split evenly by all admitting physicians in each segment. However this assumption creates an issue if the number of admitting physicians is different in the model than when the data was collected; for example when creating rosters for different scenarios of adding new physicians. The issue is that the workloads estimated within the model will not be exactly the same as the workloads actually experienced by the physicians when the data was collected. This difference means that the number of patients discharged cannot be subtracted from the physician workloads. Proportions of patients discharged are used instead, in each segment of each day. Using proportions for discharges means that the same number of patients are discharged as in the data, but non-integer numbers may be discharged from each physician.

The final type of constraints use the calculated workloads of the physicians to determine the largest difference in workloads between physicians in any segment. The mathematical model used to create the rosters is detailed in (Adams, O'Sullivan, and Walker 2016).

2.1 Solution Technique

For rosters that involve a small number of staff the model is able to find a solution in a reasonable amount of time. However as the number of staff in the model increases the time it takes to solve the model increases dramatically; for example when a larger hospital, or multiple departments in the same hospital, is modelled. The solution time also increases as more data is used since the workloads of the physicians have to be calculated over more days. This in turn leads to more constraints in the model.

Nearly all of the additional workload difference constraints are not binding when the optimal solution is found. This is because to determine the largest difference in workloads, the workload difference constraints are only strictly necessary in the segment that contains the largest difference in workloads. However if the workload difference constraints are not included on all of the other segments the model will only minimise the largest difference in that segment and ignore all of the others. Also there is no way to tell beforehand which segment will have the largest difference in workloads.

Instead an approach is used where the workload difference constraints are only included for a small number of segments initially. Then when a solution (which corresponds to a roster) is found to this problem the largest difference in workloads of the currently constrained segments is compared to the largest difference in workloads on each of the unconstrained segments (for an example see Figure 2; constrained segments are yellow, unconstrained are green). If the largest difference in workloads in a currently unconstrained segment is larger than the largest difference in workloads in all of the currently constrained segments, then the solution is discarded and that unconstrained segment is added to the constrained segments. If no such unconstrained segments exist then the solution is kept.

One reason this approach works well is that after a solution has been discarded and workload difference constraints added, a solution to the problem with the new constraints can be found easily. This is because the workload difference constraints are only included in the model to provide a 'cost' for the roster, they do not affect the feasibility of a roster. Therefore when more workload difference constraints are added to the problem the roster (solution) that was found previously is still feasible, however its cost will have increased.

To demonstrate how this process works, information about the current solution and which segments were constrained was captured during the solution process. Figure 2 shows how the largest difference in workloads changes throughout the planning horizon of the roster in dark purple. The horizontal green line shows the maximum of these largest differences for the segments on which there are currently workload difference constraints and the yellow line segments show which segments are currently constrained.

It can be seen that the first solution found results in maximum differences in workload in unconstrained segments that exceed the maximum difference among the constrained segments. This is seen by the purple line, which represents the maximum difference in workloads in each segment, going above the green line, the maximum difference in workloads among the constrained segments. The constraints corresponding to the segments where the purple line is above the green line are added to the problem (see Figure 2 solution 2) and the solution process continues until another solution is found.

In the second and third solutions that are found the purple line only exceeds the green in one area and in both cases constraints are added that correspond to those segments. Finally a solution is found (number four) where the maximum difference in workloads in all segments, even the unconstrained ones, is less than or equal to the maximum difference in workloads among the constrained segments. This shows how we do not need to consider many of the workload difference constraints. The workload difference constraints that we did not consider are the green parts of the horizontal line, the ones we did consider are the yellow parts.

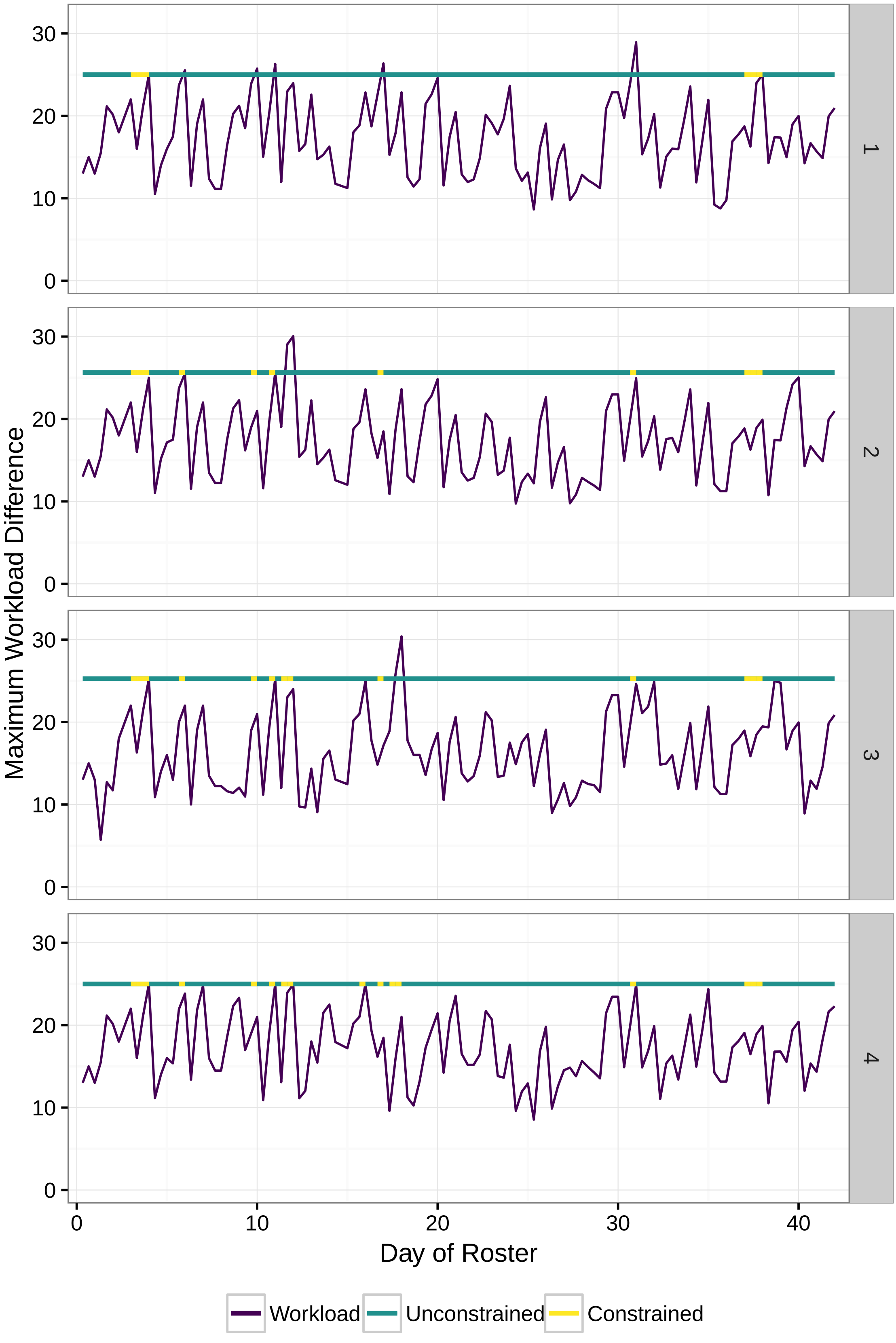

Figure 2: Time Series of Differences in Patient Workloads

3 Results

The effect of the solution technique is investigated using the rostering problem faced by the GM department at WDHB. The solution technique affects both the size of the problem and the solution time. The effects of this technique are explored by varying both the planning horizon of the roster and the number of physicians in the problem.

The planning horizon is varied from two to twelve weeks in steps of two weeks. Note that twelve weeks was the longest possible planning horizon due to the number of days of data available. Problems with five, six, eight and ten physicians were considered, because they were the configurations being considered by WDHB. All problems were solved using Gurobi on a Windows 64-bit computer with an Intel Core i7-4790 CPU @ 3.60 GHz and 8 GB of RAM.

Figure 3 depicts how long the model took to solve for varying lengths of planning horizon and number of physicians in the problem. Solution time increases with both length of planning horizon and number of physicians. The scale of the y-axis varies between subplots to allow

4 Discussion

The solution technique developed reduces the solution time for nearly all of the problem instances considered. The technique was not as effective for problems with fewer physicians (5 and 6) as it was for the problems with more physicians (8 and 10). One possible explanation for this is that some presolve techniques cannot be used, to enable the solver to add constraints into the problem as it is being solved. The presolve techniques are usually applied to the problem before solving to simplify the problem.

There is, therefore, a trade off when using this solution technique; removing constraints from the problem potentially reduces solve time, however removing the presolve potentially increases solve time. For the problems with fewer physicians the lack of a presolve is only just compensated by our improved solution technique. For the problems with more physicians the improved solution technique more than compensates for the lack of a presolve.

The solution technique is more effective for problems with a larger number of physicians because it decouples the problem from the length of the planing horizon. The objective of the problem no longer depends on the length of the planning horizon because workload definition constraints are only included for a fixed number of segments regardless of the length of schedule.

This decoupling is more effective on problems with more physicians because as the number of physicians increases the effect of the length of the planning horizon on the solution time increases. This means that an additional day in the planning horizon is more expensive for a problem with 10 physicians than one with 5, in terms of the effect on solution time. With the new solution technique the length of the planning horizon can be extended without influencing the solution time as much, especially for problems with a large number of physicians.

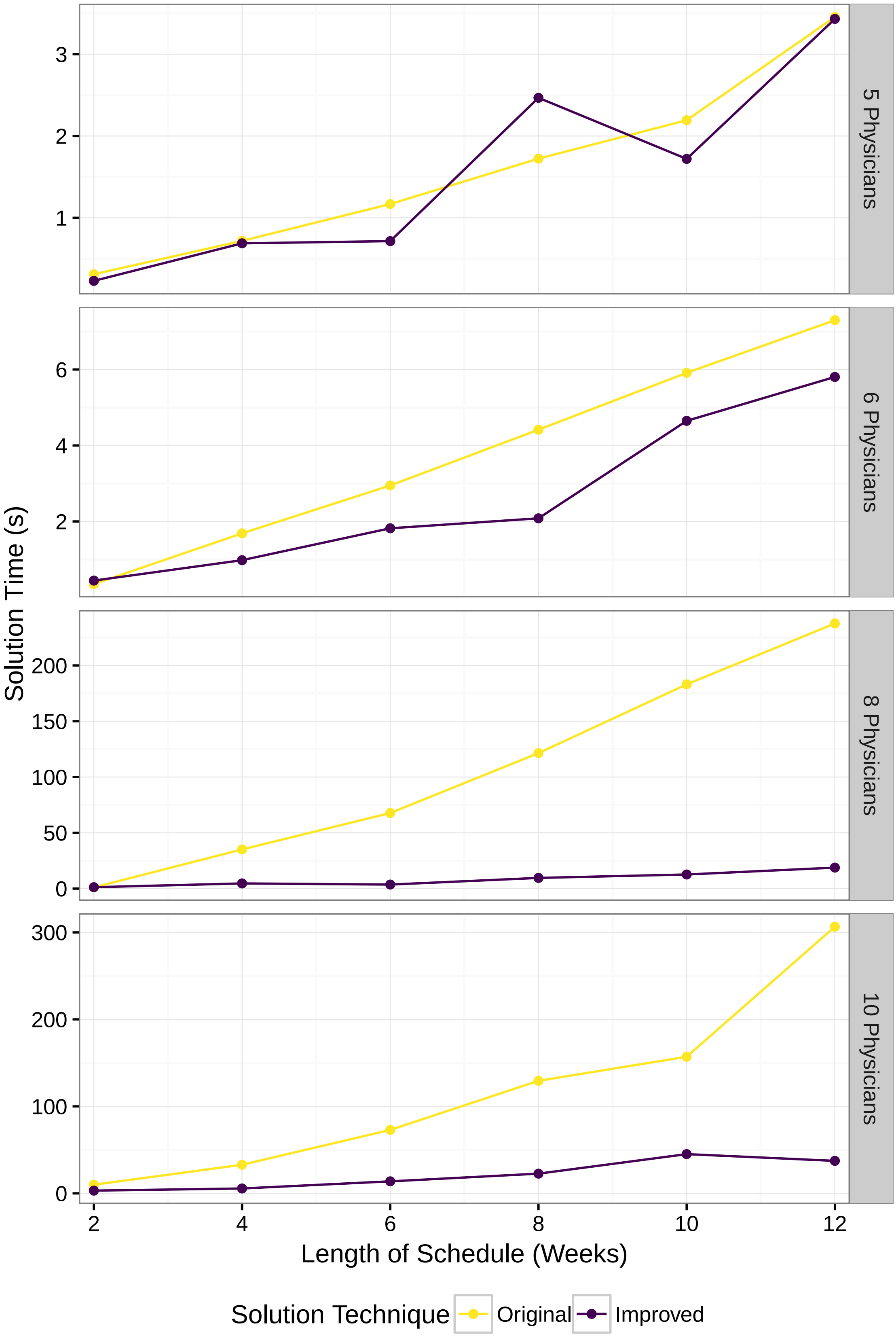

Figure 3: Time Taken to Solve Model

A further consideration is whether the roster could be improved by exploring additional objectives. Currently the largest difference in workloads in any segment of the planning horizon is minimised. This objective attempts to reduce how unfair the most unfair point in the planning horizon (the segment with the largest difference in workloads) is. Once the minimum value of the largest difference in workloads in the worst segment, and roster, is found, a lexicographic approach could be implemented that constrains this value and optimises a secondary objective.

A natural secondary objective would be to minimise the sum of the largest differences in workloads in each segment. One issue with this objective is that it is not compatible with the solution technique developed because it would mean that all of the workload definition constraints would become binding and therefore would all need to be included to find the optimal solution. This may be alleviated by the fact that an incumbent solution is readily available from the solution with the original objective.

If solution time is an issue for that secondary objective, another option is to maximise the minimum largest difference in workloads across the planning horizon. The effect of minimising the maximum largest difference and then maximising the minimum largest difference is to reduce the variability of the largest difference across the planning horizon. The cyclic nature of the roster means that all physicians should benefit from a similar reduction in workload if this variation is small. The advantage of this objective is that a solution technique similar to the one developed in this paper could be used.

5 Conclusions

A solution technique has been developed that reduces the time to create rosters for hospital departments, which balance the patient workloads of the physicians. This technique was implemented on a rostering problem for WDHB and was effective in reducing the time taken to generate a roster. With this new approach in hand it is possible to both extend the planning horizon considered by utilising more data, and create rosters for departments with more physicians.

This approach has been used to create a roster for the GM department at Waitakere Hospital. This roster was implemented on the 7th of December 2015 and has succeeded in balancing the workloads of the physicians working there. Both the number of patients being transferred to North Shore Hospital because of physician workload and the time in which acute patients are seen have been reduced.

There are several other avenues that could now be explored. First the choice of which segments to constrain initially could be examined, to further reduce the solution time. Second instead of using historical data for admissions and discharges, potential admission scenarios could be generated and used. This would allow hospitals to estimate the effect of increasing admissions on physician workloads and their rosters.

References

Adams, T., M. O'Sullivan, J. Christiansen, and P. Muir an C. Walker. 2016. Rostering general medicine physicians to balance workloads across inpatient wards: a case study. Under review by British Medical Journal: Innovations.

Adams, T., M. O'Sullivan, and C. Walker. 2016. Cyclic rostering for general medicine. Under preparation for submission to European Journal of Operations Research.

Bailey, J. 1985. "Integrated days off and shift personnel scheduling." *Computers & Industrial Engineering* 9 (4): 395 – 404.

Epstein, K., E. Juarez, A. Epstein, K. Loya, and A. Singer. 2010. "The impact of fragmentation of hospitalist care on length of stay." *Journal of Hospital Medicine* 5 (6): 335–338.

Kazemian, P., Y. Dong, T. R. Rohleder, J. E. Helm, and M. P. Van Oyen. 2014. "An IP-based healthcare provider shift design approach to minimize patient handoffs." *Health Care Management Science* 17 (1): 1–14.

Smalley, H. K., P. Keskinocak, and A. Vats. 2015. "Physician Scheduling for Continuity: An Application in Pediatric Intensive Care." *Interfaces* 45 (2): 133–148 (April).

Wolfe, Harry B. 1979. "Disaggregation planning, scheduling, and allocation of nursing staff." In *Disaggregation*, 563–575. Springer.