

# Automatic scheduling of nurses: What does it take in practice?

Elina Rönnberg, Torbjörn Larsson and Ann Bertilsson

**Linköping University Post Print**

N.B.: When citing this work, cite the original article.

Original Publication:

Elina Rönnberg, Torbjörn Larsson and Ann Bertilsson, Automatic scheduling of nurses: What does it take in practice?, 2012, in Systems Analysis Tools for Better Healthcare Delivery, eds P. Pardalos, P. Georgiev and P. Papajorgji, pp 151-178, Springer, ISBN 978-1-4614-5093-1

<http://dx.doi.org/>

Copyright: Springer

<http://www.springerlink.com/?MUD=MP>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-76079>

Chapter in the book:

*System Analysis Tools for Better Health Care Delivery*

Title:

Automatic scheduling of nurses: What does it take in practice?

Authors:

- Elina Rönnerberg, Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden (corresponding author, email: elron@mai.liu.se)
- Torbjörn Larsson, Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden
- Ann Bertilsson, Department of Mathematics, Linköping University, SE-581 83 Linköping, Sweden, and SCHEMAGI AB, Mjärdevi Science Park, Teknikringen 7, SE-583 30 Linköping, Sweden

Abstract:

Many hospital wards need to be staffed by nurses around the clock every day of the year, and because of that, many nurses have to work irregular hours and according to schedules that have a great impact on their personal lives. Today most schedules are made by hand or with limited computer support, and this is both difficult and very time consuming. A common consequence of making the schedules by hand is that the outcome is neither favourable for the nurses nor satisfactory for the running of the ward. The nurse scheduling problem is well suited for addressing with operations research methods, but a challenge for automated scheduling of nurses is the ability to adapt the scheduling to the specific conditions on each ward. The intention of this paper is to provide a piece of practical experience that can help bridge the gap between advanced method development and the use of automatic nurse scheduling in practice. Our approach is to take on the real life scheduling problem with all its details, and to use a straightforward meta-heuristic in order to deliver automatically generated schedules. The contribution of the paper is based on the result of two case studies, which will provide insights into examples of real world nurse scheduling, including evaluation and feedback from the wards.

# 1 Introduction

Many hospital wards need to be staffed by nurses around the clock every day of the week, and because of that, many nurses have to work irregular hours and according to schedules that have a great impact on their personal lives. Today there is a shortage of nurses in many countries, and in order to make the nursing profession more popular and to ensure high quality health care delivery, it is urgent to try to improve the working conditions for nurses. One possible and already ongoing improvement is that more flexibility and adaptation to personal requests is introduced in the scheduling.

Many nursing wards still use a traditional kind of scheduling, where the head nurse is responsible for creating a schedule by hand. Using this scheduling strategy leaves little room for flexibility and adaptation to personal requests from the nurses. Creating schedules by hand, or with limited computer support, is manageable, but both difficult and time consuming. The consequence of using this kind of scheduling strategy is that a lot of time and effort is spent on schedules that are neither favourable for the nurses nor satisfactory for the running of the ward. To instead use of a more modern approach for scheduling, of which self-scheduling is a common example, is an appealing alternative, but because of the increase in complexity that this entails, this kind of scheduling has proved to be difficult to use in practice.

The nature of the staff scheduling problems, with their typically huge numbers of possibilities for how to construct schedules, generally make them suitable for addressing with operations research methods. Furthermore, if staff is needed round the clock, as is often the case for nurses, the staff scheduling problem provides a true challenge from an operations research point of view. Ever since the fifties, operations research techniques have been applied to nurse scheduling problems, but as Burke et al. [7] state in their article *The state of the art of nurse rostering*, Section 4, even though a lot of research has been devoted to applying operations research techniques to nurse scheduling, there is still a gap between the mathematical models and solution methods found in articles and the flexibility that is needed to tackle real life nurse scheduling problems.

The intention of this paper is to provide a piece of practical experience that can help bridge the gap between advanced method development and the use of automatic nurse scheduling in practice. Our approach is to take on the real life problem with all its details, and to use a straightforward meta-heuristic in order to deliver automatically generated schedules. The contribution of this paper is based on the result of two case studies, which will provide insights into real world examples, including evaluation and feedback from the wards.

## 1.1 Background

This section will provide a brief overview of the nurse scheduling problem all the way from the strategic planning, which is done well in advance of the scheduling, to the last minutes updates made on a daily basis. Some references will be given, but as a general starting point for further reading, the surveys by Burke et al. [7] and Cheang et al. [8] are recommended.

### **The nurse scheduling process**

The process of ensuring that there are enough nurses present at all times comprises of numerous decisions based on different time horizons and different levels of details. These decisions can be divided into three planning phases, as illustrated in Figure 1.

*The long term planning* is a part of the overall strategic planning process of each ward. First the ward managers must estimate how many nurses with each of the necessary skills are needed during all possible time periods of the day. Given these needs, the managers can determine how the days should be divided into working shifts and the staffing demand for each of these shifts.

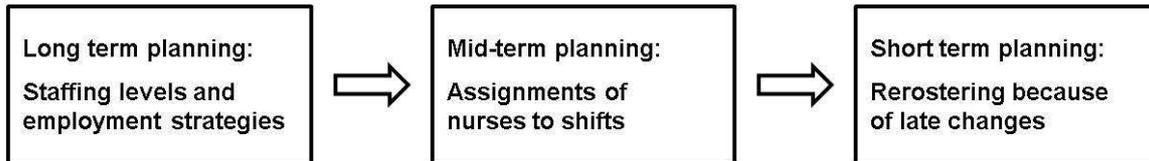


Figure 1: The three phases of nurse scheduling.

Based on this staffing demand, the kind of nurses necessary and their terms of employment can be decided. An example of such a condition is whether or not a nurse should work night shifts. The long term planning is done for the first time when a new ward is opened, and the plan is then updated regularly, for example annually, and as well as whenever major changes occur. More information about this phase of the planning can be found in Arthur and James [1], Ghosh and Cruz [14], Venkataraman and Brusco [22], and Fagerström et al. [13].

When the staffing demand is known and there is a given workforce of nurses, each nurse is assigned to a *schedule* specifying which shifts she should work, usually for a scheduling period of four to ten weeks. This phase in the planning process can be referred to as *the mid-term planning*, or *nurse rostering*. From an operations research point of view, this is the most well studied of the three phases of planning. Examples of surveys of staff rostering in general are Ernst et al. [11] and Ernst et al. [12], and surveys specialising in nurse rostering are Burke et al. [7] and Cheang et al. [8]. Mid-term planning is further described in the next section. The work presented in this paper mainly concerns mid term planning, and this is henceforth what we refer to when the term *scheduling* is used.

The schedule produced during the mid term planning should be seen as a viable plan, but the conditions for the staffing can change over time and necessitate some rescheduling. For example, a nurse can call in sick just before the shift starts, or the number of nurses needed can be different from the original estimate. Whenever there is a shortage of nurses for a shift, *the short-term planning* consists of deciding whether to use overtime, to call in a nurse on her day off, to call in a substitute nurse, or to try to manage despite the shortage. How to incorporate an optimisation tool in short term planning is dealt with in Bard and Purnomo [4], and in Moz and Pato [19] for example.

## Mid-term planning

There are two main approaches to creating a schedule for nurses; either the schedule can be unique for each scheduling period, or it can be cyclic, that is, the same schedule can be used repeatedly, period after period. In recent years, non-cyclic scheduling has gained a lot of popularity due to the flexibility that follows from creating a unique schedule each period. Because both these approaches are common in practice, in our work we have chosen to present case studies in which one of the wards uses non-cyclic schedules and the other uses cyclic ones. The details regarding the differences between cyclic and non-cyclic scheduling will be presented in more detail in Section 2.5.

Whether the schedules are cyclic or not, there are some fundamental considerations that need to be taken into account when creating them. On the ward, there is a given staff of nurses that should be assigned to shifts. The nurses have different skills and individual contracts stating which shifts they can work and for how many hours per week they should work. For each shift, for example a day, an evening, or a night shift, there are given demands for nurses with certain skills, and these demands should be fulfilled.

In addition to assigning the nurses to shifts in such a way that the staffing demands are fulfilled, there are *scheduling rules* that prescribe the properties of a schedule. Firstly, the

schedules must be consistent with the prevailing laws and regulations regarding staff scheduling. Secondly, there are always some quality aspects to consider for the schedule to be acceptable to the nurses, and thereby possible to use. A typical example of such a quality aspect is the even distribution of unpopular shifts.

Given detailed information regarding the requirements described above, the task of creating a schedule for a nursing ward is complex, but doable. On many hospital wards, this work is still carried out manually, which is time-consuming and often results in schedules with undesired properties. Operations research methodology is well suited to addressing the mid-term planning problem, and a lot of research has been carried out within this field during the last five decades. Some examples can be found in Dowsland [9], Dowsland and Thompson [10], Bard and Purnomo [3], Berrada et al. [5], and Jaumard et al. [17].

As mentioned earlier, working a mixture of both weekdays and weekends, and both day- and night-time, has serious consequences on the personal lives of the nurses. The so called *preference scheduling* is a popular way of trying to improve the working conditions for the nurses as it takes their requests and opinions into consideration when creating the schedule. However, the ambition to also take the nurses' desires into account further complicates the scheduling process.

A rather extreme form of preference scheduling is *self-scheduling*, which is a process in which nurses themselves are responsible for creating a schedule for each scheduling period. There are several ways of carrying out this process, for examples see Bailyn et al. [2] and Karlsson [18]. Note however that even if the intentions behind using self-scheduling are good, it is often a real challenge to make it work well in practice when the schedule is created by hand.

## 1.2 Approach and objective

Based on discussions with health care representatives who act on different levels of their organisations, we have concluded in earlier studies that a key to success for the automated scheduling of nurses is the ability to adapt the scheduling to the specific conditions on each ward. Based on this observation, we have developed a model designed to be easily adapted to different work places. To present all the details of this model is beyond the scope of this paper, so instead, we describe the kind of issues the model can handle, since this helps to answer the question of what it takes to automatically create a schedule for nurses in practice.

In order to carry out the case studies, we have designed a tabu search strategy focused on modularisation and flexibility rather than on short solution times. From a research point of view, this search strategy is of little interest, because it merely includes the standard components of a regular tabu search, see for example Burke and Kendall [6] and Reeves [20]. For this reason, the tabu search strategy will be only briefly described in this paper, thereby allowing the focus to be on the results of two case studies carried out in order to demonstrate the wide variety of aspects to consider when scheduling nurses.

### A comment on the choice of solution method

The reason for choosing a tabu search approach is twofold. Firstly, solving integer linear programs, such as scheduling problems, all the way to optimality is typically very difficult. Secondly, doing so in our setting is not practically meaningful, since when it comes to a schedule for nurses, the difference between an optimal solution and a good solution can be rather insignificant, because it might be just a few shifts that differ.

Why not use a branch and bound approach in a commercial solver and settle for a good solution there? Branch and bound is a viable approach, but a tabu search dedicated to the nurse scheduling problem is more likely to be more easily adaptable to different work places than a traditional integer programming model. Furthermore, when comparing these two solution processes and how the relaxation of the problem is made, there is a significant difference. In

a traditional branch and bound approach, the integrality requirements are relaxed during the solution process, which means that the intermediate solutions do not constitute meaningful schedules. In our tabu search, the integrality is kept and instead some constraints are relaxed, which means the intermediate solutions during the search will always be complete schedules, though not necessarily fully feasible ones. In a nurse scheduling setting, the distinction between feasibility and infeasibility is not always clear, and rules are often expressed in terms of what is preferable, thus making it an advantage to easily be able to change which constraints should be relaxed and which should not, and to be able to produce good and near-feasible solutions during the search.

## The case studies

The two case studies to be presented have been carried out in three steps. First, we met with representatives of the ward in order to collect information about their ward and their scheduling. In the second step we ran the tabu search to find schedules to deliver to the ward, and during this step we had to ask the head nurse further questions about their scheduling conditions and suggest schedules in order to get feedback. The last step involved analysing and evaluating the final schedules.

As we will describe, the results of both the case studies show that we were able to automatically generate schedules that are at least as good as the manually constructed ones, which can be seen as a first step towards verifying that our approach is flexible enough for real life nurse scheduling problems.

Due to practical reasons, both case studies have been carried out in Sweden, which is clearly a limitation. From what we know from the literature (see the references in Section 1.1) the aspects to consider do however seem quite similar in other countries, although some of them take slightly different forms.

### 1.3 Outline of the paper

The two sections below will give a brief description of our mathematical model and tabu search approach. The two sections to follow then present the two case studies, and these are then followed by some concluding remarks.

## 2 Modelling framework

As described in the introduction to this paper, in order to produce schedules applicable to a real world nursing ward, it is crucial to be able to tailor the model to fit the rules and conditions on the ward in question. We do this in two steps, starting from a platform which we present in this section as our modelling framework. As can be seen in Figure 2, the first step is to specify the characteristics of the ward in order to be able to create a model tailored to the ward, and then in a second step, for each scheduling period, include period specific information, such as staffing demand and preferences from the nurses.

The overall design of the modelling framework is divided into the following blocks, which will be presented in the remainder of this section.

<b>Minimise</b>	The deviation from the staffing demand
<b>Maximise</b>	Fulfilment of requests from the nurses
	Fairness among the nurses
<b>Subject to</b>	Staffing demand
	Scheduling rules
	Quality aspects

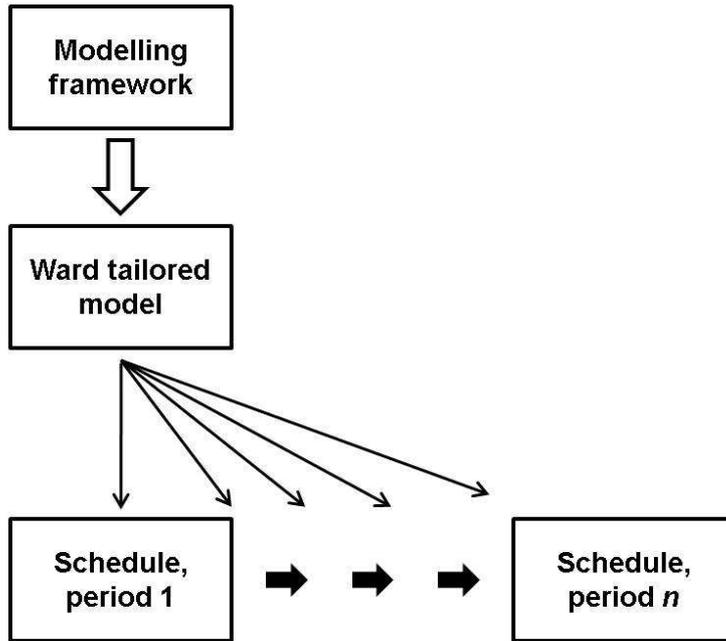


Figure 2: Within the modelling framework, the model is tailored for the needs of each ward and then supplied with data for each scheduling period.

In this section, the staffing demand, the scheduling rules, and the quality aspects will all be introduced as normal constraints. In reality some of these constraints are soft, which means that they only need to be fulfilled if it is possible. Which constraints are of this kind, and how important it actually is to fulfil them can vary from ward to ward. We ignore this distinction for the time being and will return to it when describing the tabu search in Section 3.

The aspects that need to be considered when automatically generating a schedule is heavily dependent upon the strategy the ward uses for their scheduling, and also on how their scheduling process is carried out. The modelling framework to be described could be used for cyclical as well as non-cyclical scheduling, and should preferably be used together with self-scheduling. The distinction between these different scheduling strategies and how they affect the automatic scheduling, will be addressed in Section 2.5.

## 2.1 Staffing demand and substitute nurses

In the course of one day, a ward normally needs a varying number of nurses who possess different kinds of skills; something which should be matched to the members of a nursing staff, who typically have different education, training and qualifications. In our model, the need for nurses is expressed in terms of a staffing demand, which is an estimate of how many nurses are needed during a shift. This staffing demand is accompanied by a profile describing the skills needed to fulfil the demand in question. For each shift, the staffing demand is given as a lower and an upper bound on the number of nurses needed.

Below follow some examples of differences between nurses on a ward:

- The nurses have different levels of education which place them in different categories, such as trained nurses and assistant nurses.
- There can be certain tasks that only some of the nurses have the proper training for.

- The nurses can be divided into groups, within each category of nurses or across the categories. The staffing demand may, for example, include that at least one nurse from each group covers a certain shift.
- Two wards partly share a nursing staff, meaning that some of the nurses, but perhaps not all, can cover shifts on either of the two wards.

When tailoring the model to a new ward, one specifies the kinds of categories, groups and skills that are of relevance and what profiles are needed to cover a certain staffing demand. When it comes to the shifts, these are labelled to be either a day, an evening or a night shift, and it can be determined individually for each nurse when the shifts starts and ends.

If the staffing demand cannot be met by the nurses available, substitute nurses might need to fill the shortage, something which in practice costs money. For this reason, the use of substitute nurses should be kept at a minimum, hence making it an objective to minimise the number of under-staffed shifts. In addition, the number of over-staffed shift is minimised, because it is better to fill other not yet fully staffed shifts instead.

When creating the schedule, most of the staffing demand can be fulfilled in numerous ways by different combinations of nurses, but there is also a kind of staffing demand which usually concerns a predetermined nurse and shift, and because of that, the fulfilment of this demand is predetermined. Examples of such demands are holidays, administrative tasks, and meetings or courses.

## 2.2 Scheduling rules

An automatically generated schedule must of course be consistent with rules such as prevailing laws and regulations regarding staff scheduling, and our means for accomplishing this will be presented here. These types of rules are mainly of the following two kinds.

- Rules of a combinatorial nature, stating which shifts can be combined and not.
- Rules concerning the number of hours worked.

The combinatorial constraints are easily handled by stating which combinations of shifts are possible and not, either for all nurses on a ward simultaneously or, if the rules differ among the nurse, for each nurse individually. Rules concerning how many days in succession that a nurse is allowed to work are handled similarly, and these rules can also be made dependent on the kind of shifts worked.

When it comes to keeping track of the number of hours worked, every nurse is assigned a nominal number of hours to be worked per week, and this is the number of hours she should work on average according to her employment contract. This nominal number of hours typically varies among nurses. In practice, it is not possible to allow the nurses to work their nominal number of hours every week, and a limited deviation must be allowed. This deviation is called a *time status*, and it is updated every week. Its bounds control the accumulated maximum deviation from the nominal number of hours.

A ward may wish to further control the deviation from the nominal number of hours to be worked. For this purpose, the model offers the possibility to control the number of hours each nurse works both during the whole scheduling period and during a single week, which can be set to start on a Monday or on any other day of the week. These hours are then allowed to vary within certain intervals which are set for each workplace. Combining these three measures correctly is a good means for evening out the workload for each nurse during a scheduling period.

## 2.3 Quality aspects

In order to create a schedule that complies with the laws and regulations, it is sufficient to take into account the aspects of the scheduling described previously in this section, but considering these aspects only will not be sufficient to produce a schedule that can actually be used by a nursing ward. Below, we present the remaining considerations to be taken into account when creating a schedule; these will be referred to as quality aspects. At first glance, one might believe that quality aspects are less important than the rules described earlier, but this is actually not the case in practise. The quality aspects have shown to be just as important as the scheduling rules.

Legislation on working hours puts some limitations on how many days in succession a nurse is allowed to be at work, but on many wards there are additional rules whose purpose is to improve the working conditions for the nurses. A typical example thereof is a rule for how many nights in a row a nurse is allowed to work. These kinds of rules can be individual for each nurse, depending for example on her preferences and ability to work nights.

All the previously described rules for how to combine shifts have considered which shifts not to combine. It is also quite common that wards have policies for shifts that should be combined. A typical such example is that at least one of the nurses working the evening shift should also work the day shift the morning after, because then at least one of the nurses then is up-to-date about what happened the day before. In addition to the regular scheduling rules, a ward can also have a policy with additional rules on which shifts not to combine.

For natural reasons, some kinds of shifts, for example evening, night, and weekend shifts, are less popular among the nurses. These unpopular shifts should therefore, in some sense, be fairly distributed among the nurses. What is considered to be fair differs among wards, but it is crucial to distribute these kinds of shifts according to what is accepted on the particular ward. Some of the examples we have come across for night shifts are:

- All nurses shall work the same number of night shifts.
- For each nurse, the number of night shifts worked shall be proportional to the percentage of full-time she works.
- Each nurse is assigned to a category, depending on the percentage of full-time she works, and for each such category it is determined how many night shifts shall be worked.
- The head nurse determines for each nurse individually how many night shifts she shall work.

Fairness can be desired within a group of nurses only, or within the whole nursing staff. What also differs between wards is how they distribute the weekend shifts among the nurses. Some wards focus on the number of weekend shifts worked, while others only let the nurses work complete weekends, including at least shifts on both Saturday and Sunday, and therefore count whole weekends.

It is generally not possible to achieve complete fairness among the nurses, and therefore we keep track of the deviation between the number of an unpopular kind of shift that the nurse should work and the number that she has actually been assigned to. For some wards, it is crucial to achieve the best possible level of fairness, even at the expense of other preferences (this is typical for cyclical scheduling), while for other wards it is less important, as long as deviations are kept born in mind for the next scheduling periods and made up for in the long run.

## 2.4 Requests and fairness

One important benefit of using operations research methods for scheduling nurses is how easily the nurses' preferences can be taken into account when creating the schedule. Some preferences

come in the form of individually adjusted scheduling rules or quality aspects as described above, but most of them come as requests for when to work and not, and this kind of requests is in turn divided into two types, hard and soft. When creating our model, we introduced a system for making and handling these requests, and this system is presented here.

A hard request is directly translated into a staffing demand for the specific nurse and shift, and such a request must be approved beforehand by the head nurse. Examples of such requests are holidays, courses, and meetings. Some wards also allow for hard requests based purely on personal preferences, as for example to be free on a Wednesday evening because you want to go to the theatre. How and when the nurses can make hard requests is regulated by a policy created by the head nurse.

A soft request gives the nurse the possibility to have preferences about when to work and not, and this kind of request will be fulfilled if possible. The nurse can grade the request to be standard or important. When introducing the possibility to make requests, one challenge is to measure the fulfilment of requests in a fair way, including taking into account that:

- There are different types of soft requests.
- Nurses working different percentages of full-time work different numbers of shifts, which affects the number of requests that can be fulfilled.
- Nurses with many hard requests, for example those having a weeks holiday, have fewer days for which to make soft requests, and this affects the number of requests that can be fulfilled.
- Each nurse has her own strategy for making requests, using many or few, or perhaps even trying to cheat the system.

The level of request fulfilment must be measured in a fair way, without leaving the nurses a possibility to take unfair advantage of the system. The main principles for how this is accomplished is described in the article Rönnerberg and Larsson [21]. The calculations presented there lead to a normalisation of the fulfilment of requests, and the final outcome of these calculations is referred to as a score for each nurse.

One objective when creating the schedule is to maximise the total sum of scores, but if this objective is used alone, the scores of the resulting schedule are likely to vary significantly among the nurses, and therefore the schedule will probably not be acceptable to them. Fairness is the key to achieve acceptance, and therefore an additional objective is to maximise the score for the least favoured nurses. Furthermore, a high fulfilment of requests and a high level of fairness in every scheduling period is difficult to achieve without violating scheduling rules or allowing the schedule to have other unwanted properties. It is therefore vital to find the right balance between the objectives, and also to bear the score in mind between scheduling periods and strive for complete fairness in the long run.

## 2.5 Scheduling strategies

Depending on the context in which the schedule is to be used and on how the scheduling process is carried out, what is considered to be a good schedule and not can differ. The two main categories of scheduling strategies are cyclical and non-cyclical, which will both be described below. A third strategy, which will also be commented on, is self-scheduling, a kind of non-cyclical scheduling that has quickly grown in popularity during the past two decades. Our modelling framework is designed such that all of these strategies can be applied.

## **Cyclical schedules**

Using a traditional cyclical schedule means repeating the same schedule over and over again until the ward decides to change the schedule. Each schedule is typically of a period of between 4 and 10 weeks and is used for 6-12 months. Since the same schedule is used repeatedly it is very important that it is almost perfectly fair with respect to the score, the distribution of unpopular shifts, the number of hours worked, and quality aspects. This choice of scheduling strategy imposes the boundary restriction that the first week of the schedule can follow on from the last week. Further, because the nurses are bound to use the same schedule over and over again, they typically have a low level of influence on the scheduling when this kind of strategy is used.

## **Non-cyclical schedules**

Non-cyclical scheduling means that a new schedule, usually for a period of 4-10 weeks, is created for each scheduling period. The advantage of creating a new schedule for each period is that it offers greater flexibility due to the possibility to take into account both changes on the ward and period specific requests from the nurses. A boundary condition is that the first weeks schedule is affected by the last weeks schedule from the previous scheduling period.

Instead of making the schedule very fair in each period, it can be made reasonably fair and then information about for example the score, the distribution of unpopular shifts, and the number of hours worked can be passed on to next period, and in this way a high level of fairness can be achieved in the long run.

## **Self-scheduling**

Self-scheduling is a general term used for the kind of scheduling processes where the nursing staff is jointly responsible for creating the schedule. This kind of scheduling exists in different forms around the world, but here our description is restricted to a kind of self-scheduling used in Sweden. Described briefly, this Swedish self-scheduling involves the following steps.

1. Without taking into account the staffing demand and other nurses' preferences, each nurse individually proposes a schedule for herself.
2. An improved and more feasible schedule is created through informal negotiations between the nurses.
3. A scheduling group consisting of around four nurses makes further improvements to the schedule.
4. The head nurse makes some final adjustments and approves the schedule.

Through contacts with several Swedish nursing wards, we have learnt that this kind of self-scheduling is much appreciated by the nurses, but unfortunately also associated with considerable problems. Making adjustments through informal negotiations often turns into a long and cumbersome process, which makes this kind of self-scheduling not only very time-consuming but also a source of conflicts. These problems however are not unique for this kind of self-scheduling, but are instead quite common for self-scheduling in general, see Bailyn et al. [2].

One great advantage of self-scheduling is that the scheduling is individualised and that the nurses can influence on when they work. The value of this influence is the reason why many Swedish wards use self-scheduling, despite the drawbacks presented. For a complete description of this kind of self-scheduling, the reader is referred to Rönneberg and Larsson [21].

### 3 A sketch of the tabu search design

The intention of this section is to provide a brief overview of how the tabu search algorithm explores the set of solutions on its quest for finding a solution good enough to be chosen as a schedule to be delivered to a ward. The main considerations in a tabu search, such as constraint relaxation, choice of neighbourhood, the evaluation function, and the search strategy, will be briefly commented on below.

#### 3.1 Constraint relaxation and neighbourhoods

As described in Section 2, the model is very flexible with respect to the constraints it is possible to include. It is common that some of these constraints are more important than others. Some constraints must always be complied with, while others serve more as guidelines and preferences. Because of this distinction, the constraints are divided into two categories, hard and soft, where a hard constraint has to be complied with, and a soft constraint is to be respected if possible. The constraints of the latter kind can in turn be ranked according to their importance. A common example of how this categorisation could be applied is to forbid a nurse to work six days in a row, while working five days in a row can be set to be unwanted but still feasible.

During the search, each intermediate solution is a schedule where the nurses have been assigned to shifts. On the one hand, to allow the search to explore feasible solutions only can prevent the search from reaching areas in the solutions space that can be of interest, and therefore some of the hard constraints are relaxed and the violations of these constraints are instead penalised in the evaluation function. On the other hand, in order to preserve the structure of the problem and also to keep the neighbourhoods within a reasonable size, some of the hard constraints are always required to be fulfilled during the search. As a consequence, the only conditions for a schedule to be a possible intermediate solution during the search is that all non-relaxed hard constraints and all hard requests are satisfied.

In general, a move is made by changing the shifts to be worked by a nurse, and all possible such moves constitute the neighbourhood. The neighbourhood is changed during the search, and sometimes only a subset of the full neighbourhood is explored. Such a subset is chosen primarily as it is considered to be promising with respect to the current phase, but also with some randomness included. In order to avoid discrimination if a tie occurs, it is broken by making a random choice. The strategy of including randomness in some decisions also helps to diversify the search and to avoid cycles.

#### 3.2 Evaluation function and search strategy

The three original objective functions introduced in Section 2 were: to minimise the deviation from the staffing demand; to maximise the fulfilment of requests; to maximise fairness among the nurses. These three objectives, together with the penalties for violating soft constraints and relaxed hard constraints, will constitute the evaluation function.

The different contributions to the evaluation function are grouped into three components in order to enable the search to be guided to focus on different aspects during different phases of the search. These three components are the following.

**Staffing:** To reduce deviation from the staffing demand.

**Hard constraints:** To reduce violation of the relaxed hard constraints.

**Preferences:** To reduce violation of soft constraints, to improve the fulfilment of requests, and to achieve fairness.

Each component is represented by a function obtained by weighing together the contributions to be taken into account in that component. The components are then weighed together to obtain the evaluations function.

The search starts from a solution, that is a schedule that fulfils the non-relaxed hard constraints and all requests made by the nurses. From the point of view of each individual nurse, this initial solution is an optimal one, but it does not constitute a complete feasible schedule. The aim of the tabu search is therefore to find a good and feasible solution which is as close as possible to the initial one. The change of focuses among the three components is obtained by changing the weights between the components, yielding the behaviour known as strategic oscillation. The phases changes are illustrated in Figure 3, where a sharp turn in the search represents a change of focus.

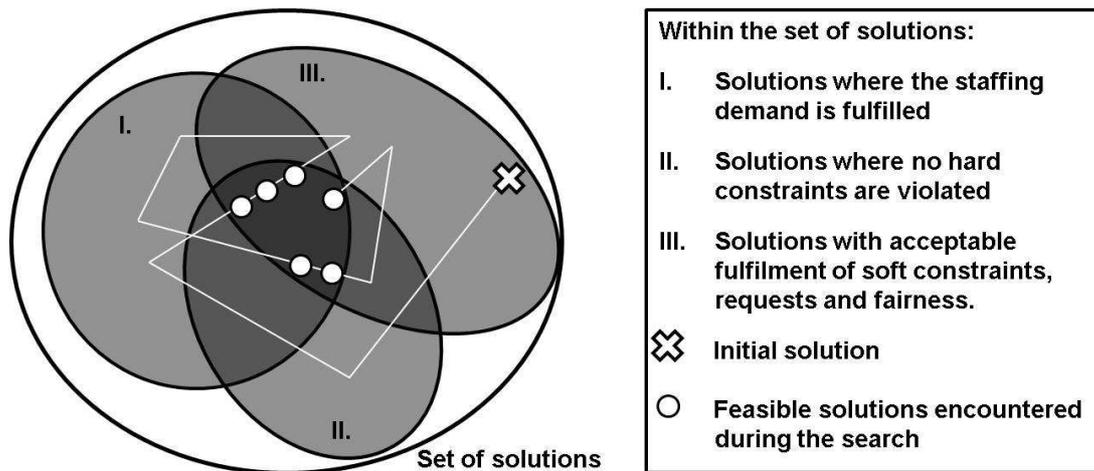


Figure 3: The white line illustrates how the set of solutions is explored and a sharp turn represents a change from one phase to another.

Changing focus during the search also serves as diversification, while keeping the same focus for a long time serves as intensification. When to change phases and what to focus on is controlled by factors such as the number of iterations within the same phase and the results of the moves during the phase, as well as randomly triggered events.

### 3.3 Tabu Lists, Aspiration and Termination

Static as well as dynamic tabu lists are used in order to avoid cycling. Some of these are effective within a phase while others are designed to control what happens between phases. The tenures of the lists are dependent on the size of the scheduling problem under consideration.

As an aspiration criterion, all tabu lists are ignored if a move leads to a solution that violates no hard constraints, has no deviation from the staffing demand, and has the best evaluation function value ever. The search terminates after a predetermined number of moves or when a predetermined number of feasible schedules has been found.

### 3.4 Outcome of the search

As will be demonstrated by the case studies in the next section, there can be significant differences between wards. One consequence of these differences is that different measures are of interest when evaluating outcomes of the automatic scheduling. What measures are of interest is considered both when tailoring the model to the ward and also when setting the parameters

for the tabu search. Examples of such parameters are the weights in the evaluation function and the tenures of the tabu lists. Making the right choices of parameter settings or not, can make the difference between success and failure. This lack of stability with respect to the parameter settings is of course a drawback of the solution method used, and also a reason for continued research to find a method less dependent on parameter settings.

## 4 Case study 1

The first case study was carried out on a hospital ward specialised in treating patients with infectious diseases. This case is a continuation of the pilot study presented in Rönnerberg and Larsson [21], which was carried out on this same ward. Within the scope of the previous pilot study, we used an AMPL (see [15]) implementation of a predecessor of our current model and CPLEX (see [16]) as solver for delivering schedules to the ward. During the long collaboration that we have had with this ward, we have delivered CLPEX-generated schedules which they have used. With regard to the three schedules presented in this section, these were the first ones created using our tabu search method, and they were therefore not used, but only evaluated. After these three periods, the collaboration continued and we have delivered a schedule for the ward to use, instead of them making one themselves.

Since all the schedules delivered so far have had the same characteristics, we find it sufficient to present the first three only. Moreover, during these three periods, the model remained almost unchanged, and because of that, some data will only be presented for one of the periods, to serve as an example.

The scheduling process used on this ward is self-scheduling, a brief description of which is found in Section 2.5 and a more thorough one in Rönnerberg and Larsson [21]. Because the nurses on this ward are used to self-scheduling, the fulfilment of requests for shifts is of great importance to them, and therefore the analysis of the result will focus on the fulfilment of requests. As will be presented, the ward uses no soft constraints, but only hard ones.

### 4.1 The ward

The ward has the capacity to care for 30 patients at a time, and it is staffed around the clock. The staff usually consists of about 15 trained nurses and 15 assistant nurses (in Swedish: *sjuksköterskor* and *undersköterskor*, respectively), with the difference between the two categories being that the trained nurses have a university degree in nursing, while the assistant nurses have only studied nursing at an upper secondary school. The trained nurses have more responsibilities and are permitted to perform more qualified tasks than the assistant nurses. The nursing ward is integrated with an open clinic, also specialised in infectious diseases, and some of the trained nurses work both on the ward and in the clinic.

The ward uses three types of shifts, day (approximately 6.45 a.m. – 3.15 p.m.), evening (approximately 1.30 p.m. – 9.45 p.m.), and night (approximately 8.45 p.m. – 7.00 a.m.). The times given are only approximate because there is a flexibility of  $\pm 30$  minutes in when the shifts start and end, depending on the nurses' contracts. The open clinic is only staffed during the daytime, 8.00 a.m. – 4.30 p.m.

The number of nurses employed on the ward varies somewhat, but in our example, taken from scheduling period 3, 16 trained nurses and 14 assistant nurses were working on the ward. Most of these nurses, namely 12 trained nurses and 13 assistant nurses, work all the three types of shifts, while the other nurses work day and evening shifts only. Not all of the nurses work full-time, and the percentage of full-time they work can be seen in Table 1. For trained nurses, full-time means 37 hours per week, if they work day and evening shifts only, and 36.2 hours if

they work all kinds of shifts. For the assistant nurses, the corresponding number of hours are 38.15 and 36.2, respectively.

Part of full-time	60%	75%	80%	85%	90%	100%
Trained nurses	0	4	2	1	1	8
Assistant nurses	1	6	1	4	0	2

Table 1: The number of nurses working a certain percentage of full-time.

For each shift during a week, the staffing demand specifies the minimum and the maximum need for trained nurses and assistant nurses, as well as the total number of nurses needed. Because the staffing demand was the same during each scheduling period and almost the same for all periods, we give only an example of a typical week in Table 2.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Day, min	3/1/7	3/1/7	3/1/8	3/1/7	3/1/7	2/1/5	2/1/5
Day, max	6/4/7	6/4/7	7/5/8	6/4/7	6/4/7	4/3/5	4/3/5
Evening, min	2/1/5	2/1/5	2/1/5	2/1/5	2/1/5	2/1/4	2/1/4
Evening, max	4/3/5	4/3/5	4/3/5	4/3/5	4/3/5	3/2/4	3/2/4
Night, min	1/1/3	1/1/3	1/1/3	1/1/3	1/1/3	1/1/3	1/1/3
Night, max	2/2/3	2/2/3	2/2/3	2/2/3	2/2/3	2/2/3	2/2/3

Table 2: The staffing demand for trained nurses/assistant nurses/total number of nurses.

The staffing demand for the open clinic usually requires need one trained nurse each day shift during weekdays. This demand can be fulfilled by a few of the trained nurses only, and the shifts at the clinic should be distributed evenly among these nurses.

Both night and weekend shifts should be evenly distributed among the nurses. However, if there would be a nurse working 50% of full-time or less, she should work only half as many shifts on both nights and weekends, as a nurse working full-time.

The combinatorial kinds of constraints used are:

- A nurse can work at most five days in succession.
- A nurse can be scheduled to work at most three consecutive night shifts.
- A nurse can work at most one shift per day.
- After working a night shift, a nurse may not be scheduled to work on the day or the evening shift the following day, or the day shift on the day after that.
- When working during a weekend, a nurse works either only days and evenings, or only nights.
- If working a weekend comprising of days and evenings, the nurse shall work both Saturday and Sunday, as well as either the evening shift on Friday or the day shift on Monday.
- There are no special rules for working nights over a weekend.

The number of hours worked is controlled by a time bank and intervals for the number of hours worked, both per week and in total during the period. The time bank is set to  $\pm 20$  hours and the number of hours worked is allowed to vary between 50% and 160% of the nominal values

within a week, from Monday to Monday and from Friday to Friday, and with  $\pm 20\%$  within the period.

## 4.2 The schedules

We will present the results from three scheduling periods, of respectively five, eight and eight weeks. As stressed earlier, the fulfilment of their requests is of great importance to the nurses on the ward. Table 3 shows the kinds of requests used for our scheduling on the ward, and how many requests of each kind that were made.

Type of request	Period 1	Period 2	Period 3
Hard request, holiday	7	54	37
Hard request, meetings etc.	54	17	51
Hard request, not to work	144	375	369
Hard request, to work	0	40	9
Soft requests, important not to work	0	48	0
Soft requests, important to work	0	25	0
Soft requests, to work	494	807	836

Table 3: The number of requests made by the nurses.

The solution times for the tabu search varied from a few hours to a few days. We find this acceptable since no major effort has been made to reduce these times, and also keeping in mind that a new schedule is needed only about five to ten times a year. The tabu search was able to find feasible schedules, both with respect to the staffing demand and the hard constraints, for all three scheduling periods.

An interesting measure for evaluating the outcome of the scheduling is to count the percentage of fulfilled requests each nurse receives. A histogram showing this result is presented in Figure 4, and as can be seen there, fulfilment is both quite high and quite fair, with a request fulfilment rate of at least 60% for each nurse.

It is worth mentioning that the results of period three are from a second attempt by the tabu search to find a schedule for this period. When running the tabu search the first time, it was discovered that the input data was not consistent with respect to the number of nights requested and the number of nights that should be worked by some of the nurses. After adjusting these numbers, the tabu search obtained the result presented here.

The remainder of this section is devoted to evaluating and analysing viability of automating the self-scheduling process by using our tabu search method. A cornerstone in the Swedish self-scheduling process presented is that nurses shall request as many shifts as they should work. The deviation between the requests made and the final schedules can therefore be represented by *swaps* of shift; that is, if a nurse does not work a requested shift, then she has to work some other shift. The swaps can be divided into two types, *primary swaps* and *secondary swaps*.

By primary swaps is meant those swaps that are directly caused by the difference between the staffing demand and the requests made. The primary swaps are sufficient to redistribute nurses from over-staffed shifts to the ones with a shortage, but in order to create a schedule that is also consistent with the scheduling rules, further swaps might be required, and those are called secondary. The two types of swaps are illustrated in Figure 5. In this example, there are three shifts that need to be staffed, which can represent any one of the three shifts in a scheduling period, and they do not need to be consecutive or ordered. The staffing demand is three nurses for each shift. In their proposed schedules, three nurses request the first shift, two request the second, and four request the third one. In Case A, it is possible to let one of the

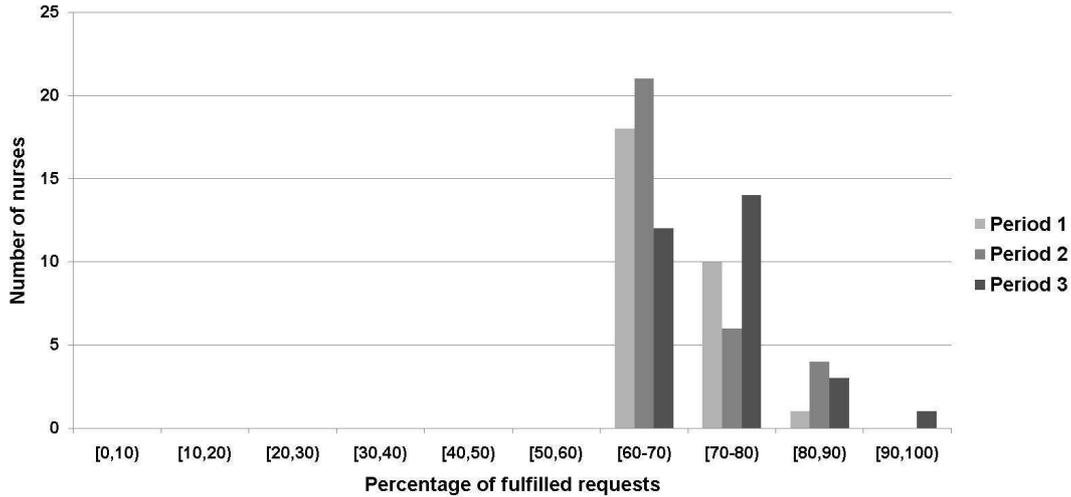


Figure 4: A histogram showing how many nurses that for each scheduling period got a certain number of requests fulfilled.

nurses who requested to work the third shift to work the second shift instead, which requires only a primary swap. In Case B, it is not possible for any of the nurses who requested the third shift to work the second one, for example because they do not have the right skills. Instead, one of these nurses is moved to shift one, which she is allowed to work, and then one of the nurses from shift one, who can work the second shift, is moved to the second shift. In this case, both a primary swap and a secondary swap are required.

In the analysis which follows, we only consider requests for working shifts. In Figure 6 there are three bars that represent the result for each scheduling period. The first bar shows the total staffing demand for the period. The second bar shows the number of requests that coincide with the staffing demand, that is to say the difference between the first and the second bar is the number of primary swaps required. The last bar represents the number of requests for working, of any strength, that were fulfilled by the tabu search. Hence, the difference between the second bar and the third bar is the number of secondary swaps, plus the number of requests not fulfilled because the schedule does not correspond to an optimal solution.

In our example, the total staffing demand is nine nurses, and for each of the shifts, the number of requests that coincide with the staffing demand are three, two, and three, respectively. In the kind of diagram described above, the first and the second bar would be nine and eight units high, respectively, with the difference corresponding to the single primary swap required. In Case A, the third bar would be of the same height as the second one, since only a primary swap is needed, while in Case B, the third bar would be one unit lower, because of the secondary swap.

The number of primary swaps required can clearly not be affected by any optimisation tool, or by any other approach for creating the schedule, because it depends solely on the difference between the staffing demand and the requests made by the nurses. If this difference should be large, there would be no reason to use self-scheduling at all, because it would then not be possible to meet many of the requests. However, as is illustrated in Figure 6, about 81% of the requests made in this case study coincide with the staffing demand; which we find very promising.

One welcome result is that there were as few secondary swaps as is shown in Figure 6. Although the secondary swaps are induced by the scheduling rules, the number of such swaps is highly dependent on how the ward is staffed. If all nurses work a high percentage of full-time,

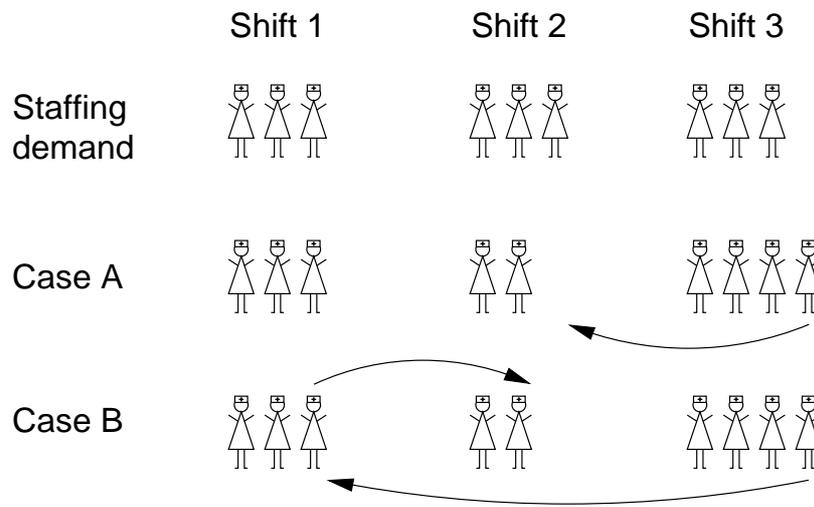


Figure 5: Illustration of the two kinds of swaps; in Case A only a primary swap is carried out, and in Case B both a primary and a secondary swap are carried out

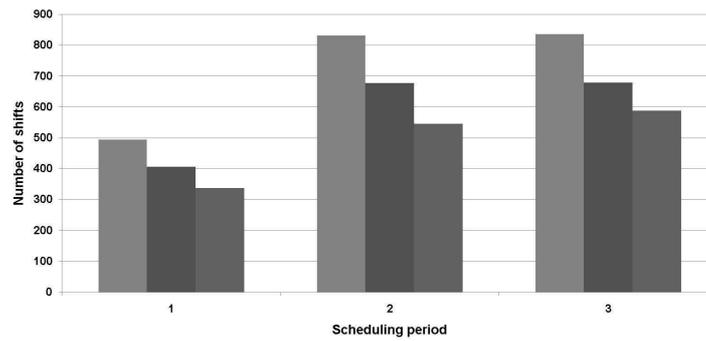


Figure 6: Fulfilment of requests for working a shift. For each scheduling period the first bar shows the number of requests made, the second one how many of these coincide with the staffing demand, and the third one shows the number of requests fulfilled.

then it is more difficult to create a schedule than if there are more nurses, who all work a lower percentage of full-time. The number of nurses who have the same skills and who therefore can fulfil the same demand is also important, since this affects the likelihood that a primary swap will be sufficient.

### 4.3 Feedback from the ward and discussion

The results presented show that the model and the method used were able to create feasible schedules for all three scheduling periods. The schedules were reasonably fair and all complied with at least 60 % of the requests made by each nurse.

The head nurse seems convinced that it will be possible to automatically generate schedules for the kind of self-scheduling that they use, but that how the automatic generation is implemented and presented to the nurses will be crucial if it is to receive acceptance. She also agrees that it is important to consider the schedule created by this tool as a suggestion that can then be adjusted by the nurses themselves, if such adjustment leads to an improvement. The schedules presented were in need of some minor adjustments, something that probably is, and always will be, inevitable; but compared to the work of creating the schedules manually from scratch, this workload is negligible.

The fact that our collaboration with this ward continued after the case study and that we have since delivered a schedule that is to be used, is a result that we believe speaks for itself.

## 5 Case study 2

The second case study was carried out at a home care unit in Norrköping, where they use a six week cyclical schedule, which is typically repeated for around six months. At this work place, the nurses are not allowed to make requests in a self-scheduling sense, due to the difficulties the head nurse has experienced in handling the requests when creating the schedule manually. Instead, the head nurse allows the nurses to make a very limited number of simple requests, such as not being scheduled on Tuesday evenings.

Instead of evaluating the schedule with respect to the fulfilment of requests, as we did in Case study 1, the measure of quality here is associated with the ability to comply with soft constraints, and to do it in a fair manner. The importance of complying with soft constraints is a consequence of using cyclical schedules; since the same schedule should be repeated over and over again, certain properties of the schedule are preferred to others, and fairness with respect to these is crucial.

The collaboration with the ward was initiated around six months before this paper was written and we have therefore only had the time to create one schedule for them, alongside their original scheduling procedure. At the same as this paper was written, the tabu search method presented was used to deliver a schedule for the ward.

### 5.1 The ward

The 29 nurses employed on the ward are all assistant nurses, and they are divided into groups depending on which geographical area they primarily serve. The groups, Group 1, Group 2, and Group 3, consist of 12, 10, and 7 assistant nurses respectively. Not all the nurses work full-time, and the percentage of full-time they work can be seen from Table 4.

The three types of shifts they use are day shifts, evening shifts, and split shifts, which means that a nurse works half a shift in the morning and half a shift in the evening. The length of a shift is determined individually for each nurse, depending on the percentage of full-time she works and also on whether it is a weekend or a weekday. A day shift starts at 7.00 a.m. or

Part of full-time	50 %	55%	60%	70%	75%	77.5%	80%	85%	90%	100%
Assistant nurses	2	1	1	1	3	2	2	1	2	14

Table 4: The number of nurses working a certain percentage of full-time.

at 8.00 a.m. and the length of the shift is between 4 and 8 hours. The evening shifts are also between 4 and 8 hours long, and the split shifts are 9 or 10 hours long.

The staffing demands for the evening and the split shifts are only expressed in terms of the minimum number of nurses needed, and can be seen in Table 5. For the day shifts, there is a total minimum number of nurses needed, of whom some nurses must work a long day shift.

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Day, total	3/3/1/14	3/3/1/14	3/3/1/14	3/3/1/14	3/3/1/14	1/1/1/5	1/1/1/5
Day, long	1/1/1/5	1/1/1/5	1/1/1/5	1/1/1/5	1/1/1/5	1/1/1/5	1/1/1/5
Evening	1/1/1/6	1/1/1/6	1/1/1/6	1/1/1/6	1/1/1/6	0/0/0/1	0/0/0/1
Split	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/0	0/0/0/5	0/0/0/5

Table 5: The staffing demand for Group 1 / Group 2 / Group 3 / total number of nurses.

Three of the nurses have a fixed schedule created by the head nurse, and two other nurses each have a fixed schedule for half of the period because they have specific tasks to perform then. In each scheduling period, every nurse is required to attend two meetings, one with her group, and one with the whole nursing staff. These meetings are each scheduled as a two-hour afternoon shift.

When it comes to the scheduling rules, some of them apply to all the nurses while some may apply only to some nurses. The constraints will be presented such that regular constraints are presented first, and the exceptions are given directly thereafter. The hard constraints of a combinatorial nature are the following.

- Each nurse must be consecutively off duty from Friday to Sunday, or from Saturday to Monday, at least once during the scheduling period.
- Each nurse must be off duty at least 14 days during the scheduling period.
- A nurse should never work more than two consecutive evenings. Exception: There is a nurse who works evenings only.
- A nurse can work at most one shift per day.
- A nurse is not allowed to work a solitary day in between days on which she is off duty. Exception: The nurse working 50% of full-time is excluded from this rule.
- If a nurse works during a weekend, then she must work both on Saturday and Sunday. Exceptions: The nurses working 50% or 60% of full time, and some nurses with a fixed schedule.
- The evening shift on Fridays shall only be staffed with nurses working the weekend to follow.
- If a nurse works during a weekend, she has to be off duty on either Monday or Tuesday the following week.

- Around two thirds of the nurses are allowed to work at most 4 consecutive days, the remaining one third is allowed to work at most 5 consecutive days. Exception: One nurse is allowed to work at most 3 consecutive days.

The number of shifts scheduled per week can deviate from between 55% and 135% from a nurse’s nominal number of shifts to be worked. The time bank, which in this case is measured in numbers of shifts, was set to  $\pm 3$  shifts during the period. An objective for the whole scheduling period was to minimise the deviation between the nominal and the scheduled numbers of hours.

The head nurse wanted four soft constraints to be taken into consideration. Ideally, she would like all of these to be complied with, but from experience she knows that this is not possible. For future reference, when the result is to be analysed, these soft constraints are listed below.

- **Five days in a row:** The one third of the nurses allowed to work 5 days in a row prefer not to do so, and therefore they have a soft constraint that they shall work at most 4 days in a row. Exception: One nurse likes to work 5 days in a row.
- **Solitary day off:** A nurse should preferably not be off duty on a solitary day between two days of work.
- **Evening, then off duty:** Ideally, a nurse shall not be off duty the day after she works an evening shift. Exception: One nurse prefers this.
- **Two evenings in a row:** If possible, the nurses should not be scheduled to work two evenings in a row. Exception: The nurse who work evenings only.

Only a few requests were made by the nurses, and an example of such a request is a nurse who wanted to be off duty on Tuesday evenings. Because the number of requests was so very small, the reasonable ones were all complied with, and they will not be discussed further. When evaluating the fairness, the requests were not of interest; instead fairness was associated with the number of unpopular shifts each nurse had to work and with the fulfilment of soft constraints. The head nurse decided that the unpopular shifts should be distributed with respect to the percentage of full-time a nurse works, and with some support from us, she then concluded that a fair distribution of these shift would correspond to what is presented in Table 6.

Part of full-time	50%	55%	60%	70%	75%	77.5%	80%	85%	90%	100%
Weekend shifts	3	6	2	4	4	4	6	6	6	6
Evening shifts	5	24	7	5-6	5-6	5-6	5-6	5-6	5-6	7-8
Split shifts	0	0	0	2	2	2	3	3	3	3
Friday evening shifts	1	2	1	1	1	1	1	1	1	1-2

Table 6: The number of unpopular shifts of each kind a nurse working a certain percentage of full-time should be assigned to.

To achieve this distribution of shifts was a very important objective during the tabu search. As can be seen from the table, the head nurse had a special agreement with the nurse working 60% as well as with the nurse working 55%, who works evenings and weekends only. Nurses with a schedule fixed beforehand are not shown in this table.

## 5.2 The schedules

For the analysis of this case study, the following four main issues are of interest.

- Does the schedule comply with the hard constraints?
- Is there a deviation between the staffing demand and the staffing levels?
- Is the distribution of unpopular shifts fair?
- How many times, if any, are the soft constraints violated for each nurse?

The tabu search was set to run for a few hours, and the outcome of the search was highly satisfactory, making it easy to answer the first three questions. The schedule complied with all the hard constraints as well as with the staffing demand. The distribution of unpopular shifts was made according to Table 6, which was a very welcome result since we did not know beforehand if this was possible. The deviation between the nominal and the scheduled numbers of hours was at most 3 hours per nurse, something which is acceptable in practice. As was expected, some of the soft constraints were violated, and this will be further discussed.

The violations of soft constraints are presented in Figure 7, with the number of occasions specified for each nurse. Because the constraints are not exactly the same for all the nurses, it could be rather difficult to draw any conclusions from the figures presented without taking the exceptions from the rules into consideration. In Figure 7, the nurses are numbered from 1 to 29. The first 16 nurses are allowed to work at most four consecutive days, and they follow the regular constraints only. Nurses 17–29 are associated with some kinds of exceptions to be further discussed.

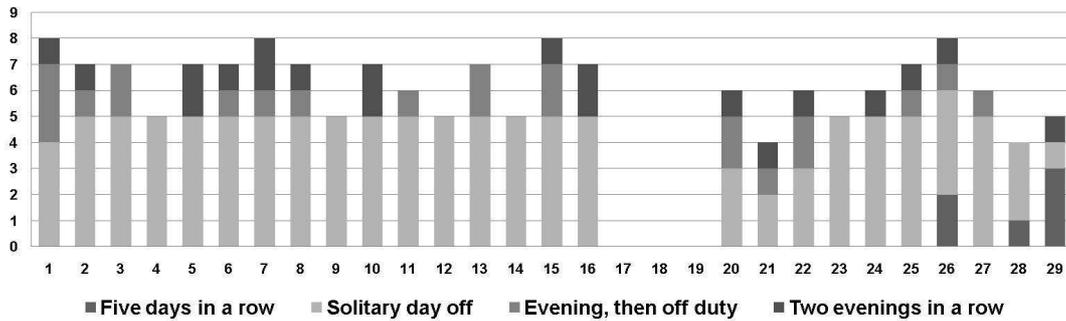


Figure 7: For each nurse, the bar represents the number of times a soft constraint of each type has been violated in the schedule.

Nurses 17–19 have fixed schedules for the whole scheduling period, and they are therefore not affected by the soft constraints. For nurses 20 and 21, the schedule is fixed for one half of the scheduling period each, which means that they are less affected by the soft constraints, something which is also true for nurse 22, who also has a partly fixed schedule. Nurse 23 works evening shifts only, which means that some of the soft constraints do not apply to her. Number 24 is the nurse who wants to be off duty the day after she has worked an evening shift, and therefore the occurrences of such events are not included in the figure. Nurses 25–29 do not work many weekends, which makes it more difficult to avoid having them working 5 days in a row. Out of these five, nurse 29 prefers to work 5 consecutive days.

Studying the sixteen nurses following the regular soft constraints, one can conclude that it was difficult to avoid solitary days off, something which occurred 4–5 times for each of these nurses. The explanation to this is found when studying the other constraints and the staffing demand. The combination of having a schedule with both weekdays and weekends, along with the rules for how to work weekends (for example being free on Monday or Tuesday) makes it difficult to create many occurrences of consecutive days off, because the nurses have to alternate

between working on weekdays and weekends. That each nurse has 4–5 solitary days off should be compared to the total number of days off for each nurse, which is at least 14.

The constraint preventing a nurse from having a day off after working an evening shift, as well as the constraint that forbids working two evenings in a row, are violated only a few times each, something which we find to be an acceptable result for a schedule to be usable in practice. Furthermore, from Figure 7, it can be seen that the distribution of violated soft constraints is reasonably fair, and this is more important than the actual number of violations if the schedule is to be acceptable to the ward.

### 5.3 Feedback from the ward and discussion

Since the tabu search succeeded in complying with all the hard constraints, the staffing demand, and with the distribution of unpopular shifts, a lot of attention during this case study was given to the soft constraints. As described earlier, some of them were more easily complied with than others, and the ones to be prioritised was decided by the head nurse. During this study, we have focused on finding a best possible schedule for the ward in this particular scheduling period, and this has been done empirically, in close collaboration with the head nurse. If one were to further evaluate the behaviour of the method, it could be of interest to study the kinds of solutions that it is possible to create, and preferably perform tests on many instances from the same ward.

The nurses' responses varied between very satisfied with the result and a bit more negative because some soft constraints were violated. The head nurse's response to this criticism was that the soft constraints are also violated with manual scheduling, and that a great advantage of using our tabu search method is that she does not need to decide for which nurses and when this should happen. Furthermore, she also consider "the computer" to be more objective and fair in its decisions.

Usually the head nurse works for one to two weeks creating a schedule, and by using our automatically generated schedule, this time was reduced to 4 hours. These four hours were used for inspecting and evaluating the result of the scheduling, and also for making small adjustments, such as extending or shortening shifts for the nurses in order to even out the deviation of  $\pm 3$  hours between the nominal and the scheduled numbers of hours for the period.

The most telling result of this case study is the fact that the ward wants us to deliver a schedule the next time they need a new one, and that we will do so.

## 6 Concluding remarks

The long-term goal of our work is to develop an optimisation tool that functions as a practical tool for the automatic scheduling of nurses. That we managed to deliver usable schedules of high quality to both the wards in the case studies is very promising – and a huge leap in the right direction.

One objective of this paper was to illustrate the wide variety of issues to consider when scheduling nurses in practice. The two wards were chosen for the case studies because they differ significantly, and also because they are typical within their respective categories of Swedish nursing wards.

The tabu search method used does not play a central role in this paper, but instead merely serves as a means to be able to carry out the computations needed for the case studies. For future research within this area, it could be interesting to develop a more efficient method, preferably less dependent on parameter settings, for solving the nurse scheduling problems described in our modelling framework.

Something that is common for all the wards that we have come across is the great importance of fairness, even though there are significant differences between the definitions of fairness. In

Case study 1, fairness is associated with the fulfilment of requests only, while in Case study 2, it is associated with the soft constraints. It is also possible to combine the two aspects and let the fulfilment of requests and soft constraints interact. However, it can then be a challenge to obtain the right priorities between the two aspects.

One challenge when working with a new ward is to understand what is essential in their scheduling, and in order to be able to successfully deliver a schedule, it is of crucial importance to understand their values and traditions. During our work, the responses from the nurses have been both expectant and sceptical; expectant because of the time-consuming work and difficulties associated with the manual scheduling process, and sceptical mainly because they are afraid of losing control over the scheduling. Because of the nurses' scepticism, it is important to present the outcome of the automatic scheduling pedagogically and to emphasise that the optimisation tool only offers a qualified suggestion for a schedule. If it is considered beneficial, the nurses are allowed to make minor adjustments themselves.

The great benefit that our approach brings should be the time and effort saved if the head nurse is handed a schedule that is both feasible and fair. Other benefits are the objectivity of a computerised planning tool and the decrease in lead time for constructing a schedule.

## **7 Acknowledgements**

Thanks to all the health care representatives that we have been in contact with for sharing your knowledge about nurse scheduling. Special thanks to Elisabet Shimekaw and Ann Andersson, head nurses on the wards in Case studies 1 and 2, respectively, and also to all the nurses working on these wards.

This work was carried out in collaboration between the Department of Mathematics, Linköping University, and SCHEMAGI AB, with financial support from HNV/Vinnova, and with support from LEAD.

## References

- [1] T. Arthur and N. James. Determining nurse staffing levels: A critical review of the literature. *Journal of Advanced Nursing*, 19:558–565, 1994.
- [2] L. Baily, R. Collins, and Y. Song. Self-scheduling for hospital nurses: An attempt and its difficulties. *Journal of Nursing Management*, 15:72–77, 2007.
- [3] J. F. Bard and H. W. Purnomo. Preference scheduling for nurses using column generation. *European Journal of Operational Research*, 164:510–534, 2005.
- [4] J. F. Bard and H. W. Purnomo. Hospital-wide reactive scheduling of nurses with preference considerations. *IIE Transactions*, 37:589–608, 2005.
- [5] I. Berrada, J. A. Ferland, and P. Michelon. A multi-objective approach to nurse scheduling with both hard and soft constraints. *Socio-Economic Planning Sciences*, 30:183–193, 1996.
- [6] E. K. Burke and G. Kendall, editors. *Search methodologies: Introductory tutorials in optimization and decision support techniques*. Springer, New York, NY, 2005.
- [7] E. K. Burke, P. de Causmaecker, G. van den Berghe, and H. van Landeghem. The state of the art of nurse rostering. *Journal of Scheduling*, 7:441–499, 2004.
- [8] B. Cheang, H. Li, A. Lim, and B. Rodrigues. Nurse rostering problems — a bibliographic survey. *European Journal of Operational Research*, 151:447–460, 2003.
- [9] K. A. Dowsland. Nurse scheduling with tabu search and strategic oscillation. *European Journal of Operational Research*, 106:393–407, 1998.
- [10] K. A. Dowsland and J. M. Thompson. Solving a nurse scheduling problem with knapsacks, networks and tabu search. *Journal of the Operational Research Society*, 51:825–833, 2000.
- [11] A. T. Ernst, H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. An annotated bibliography of personnel scheduling and rostering. *Annals of Operations Research*, 127:21–144, 2004.
- [12] A. T. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier. Staff scheduling and rostering: A review of applications, methods and models. *European Journal of Operational Research*, 153:3–27, 2004.
- [13] L. Fagerström, A-K. Rainio, A. Rauhala, and K. Nojonen. Professional assessment of optimal nursing care intensity level: A new method for resource allocation as an alternative to classical time studies. *Scandinavian Journal of Caring Science*, 14:97–104, 2000.
- [14] B. Ghosh and G. Cruz. Nurse requirement planning: A computer-based model. *Journal of Nursing Management*, 13:363–371, 2005.
- [15] *ILOG AMPL CPLEX System, Version 10.0, User’s Guide*. ILOG, 2006.
- [16] *ILOG CPLEX 10.0, User’s manual*. ILOG, 2006.
- [17] B. Jaumard, F. Semet, and T. Vovor. Case study: A generalized linear programming model for nurse scheduling. *European Journal of Operational Research*, 107:1–18, 1998.
- [18] P. Karlsson. Rapport om kartläggning av arbetstidslösningar inom landstinget. Stockholms läns landsting, 2005.

- [19] M. Moz and M. V. Pato. Solving the problem of rostering nurse schedules with hard constraints: New multicommodity flow models. *Annals of Operations Research*, 128:179–197, 2004.
- [20] C. R. Reeves, editor. *Modern heuristic techniques for combinatorial problems*. McGraw-Hill, London, 1995.
- [21] E. Rönnberg and T. Larsson. Automating the self-scheduling process of nurses in Swedish healthcare: a pilot study. *Health Care Management Science*, 13:35–53, 2010.
- [22] R. Venkataraman and M. J. Brusco. An integrated analysis of nurse staffing and scheduling policies. *Omega: The International Journal of Management Science*, 24:57–71, 1996.