Integrating Data Mining and Optimization Techniques on Surgery Scheduling

Carlos Gomes[1,2], Bernardo Almada-Lobo[1,3], José Borges[1,3], and Carlos Soares[1,2]

[1] INESC-TEC
[2] Faculdade de Economia/Universidade do Porto, Portugal
[3] Faculdade de Engenharia/Universidade do Porto, Portugal

Abstract. This paper presents a combination of optimization and data mining techniques to address the surgery scheduling problem. In this approach, we first develop a model to predict the duration of the surgeries using a data mining algorithm. The prediction model outcomes are then used by a mathematical optimization model to schedule surgeries in an optimal way. In this paper, we present the results of using three different data mining algorithms to predict the duration of surgeries and compare them with the estimates made by surgeons. The results obtained by the data mining models show an improvement in estimation accuracy of 36%. We also compare the schedules generated by the optimization model based on the estimates made by the prediction models against reality. Our approach enables an increase in the number of surgeries performed in the operating theater, thus allowing a reduction on the average waiting time for surgery and a reduction in the overtime and undertime per surgery performed. These results indicate that the proposed approach can help the hospital improve significantly the efficiency of resource usage and increase the service levels.

Keywords: Surgery Scheduling, Data Mining, Optimization.

1 Introduction

Technological advances, medical breakthroughs, better and more efficient services are constantly changing the world, improving the quality of life and, in the long run, life expectancy. Hospitals, which can now provide more and better care to patients who once were unable to receive treatment, face pressures due to the rise on demand and the elevated waiting time for treatments.

Looking at the operating theater, which is considered by many the largest budget consumer in hospitals [1], the waiting time problem is worsened due to the nature of the treatments involved (surgeries) and the immediate attention they naturally require. In fact, reducing waiting lists for surgery has always been a priority of sovereign governments and therefore of many researchers. The Portuguese government has successfully reduced the median waiting time for surgery from 8.6 to 3.3 months between 2004 and 2009, as a result of the introduction

S. Zhou, S. Zhang, and G. Karypis (Eds.): ADMA 2012, LNAI 7713, pp. 589–602, 2012.

of an incentive program to perform additional surgeries. Still, despite the efforts and success in tackling the long waiting lists for surgery, Portugal along with the United Kingdom ranked last among other 31 European countries regarding waiting times for treatment [2]. Moreover, incentive strategies imply great costs, and in the current economic context can easily be ceased. An OECD (Organisation for Economic Co-operation and Development) report also lists other countries with waiting time for surgery issues, such as Italy and Norway which have more than 25% of their patients waiting for more than 12 weeks [3].

In this work, we introduce a combination of data mining and optimization techniques applied to operating theater planning. In the process of scheduling surgeries, surgeons have to estimate empirically how long the combination of surgical procedures will take in order to book the operating room. The accuracy of these estimations will define the quality of the operating theater schedule, since every deviation from the estimates leads either to schedule disruptions (surgeries exceeding their allocated time) or to unoccupied time (surgeries finishing earlier than estimated). This wasted time is valuable, not only for the hospital but to the patients who see their health conditions and overall satisfaction quickly deteriorating throughout time. On the other hand, since scheduling is a complex combinatorial problem, subject to different rules and constraints, it represents an opportunity to use optimization techniques to improve its quality. Herein, we make use of data mining algorithms to estimate the duration of surgeries and of an optimization model to optimize the surgeries' schedule. We achieve good results with both techniques separately, but it is their combination that enhances the final scheduling solution and enables the automation of the entire surgery scheduling process. In summary, the contributions in this work are:

- Development of an automatic and effective mechanism to estimate surgery duration, based on historical surgery records, patient and surgeon information;
- An approach that optimizes operating theater resource utilization based on the predicted durations of surgeries.

Next we will provide further details about the problem and a review of related work in Sections 2 and 3, respectively. In Section 4 the estimation problem will be presented and the surgery scheduling optimization model will be described in Section 5. Experimental results are shown in Section 6 and final remarks are given in Section 7.

2 Problem

Our work was conducted at a large Portuguese hospital, enabling the analysis and gathering of the necessary inputs to develop and validate the approach presented. The scope of the experiments is restricted to the outpatient department, also know as ambulatory surgery department, which deals with patients who are not hospitalized after the surgical intervention. On the other hand, inpatient surgeries require hospitalization for the patients' recovery, and this recovery is

the true bottleneck of the operating theater and hospital resources [4–6]. Therefore, we have decided to start by focusing on outpatients, for which the efficient time management of the operating theater becomes crucial to make the most of the existing resources. To illustrate how the predictions made by surgeons, which are currently used as basis for operating room scheduling, are deviated from reality , Figure 1 plots those estimates against the real surgery times recorded on this department between 2010 and 2011. It is possible to observe a ladder effect created by the coarse time granularity used by surgeons on their estimates (15, 30 or 60 min.). This prevents fine tuning of the surgery schedule and therefore reduces its quality. Furthermore, surgeons tend to make predictions according to their own interests and undervalue the objectives of the hospital. In fact, Macario states that surgeons intentionally overbook operating rooms to block them to other surgeons or underestimate the duration of surgeries in order to squeeze more surgeries in their available time [7]. Both scenarios lead to wasted time, but the latter actually produces schedule disruptions due to the delays caused, often leading to the postponing of surgeries.

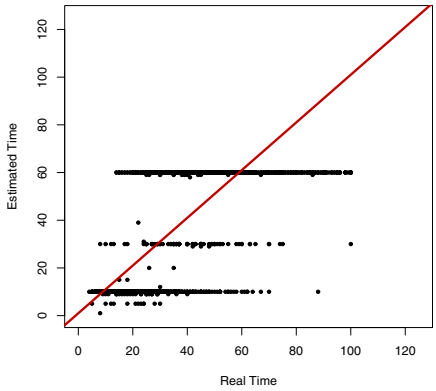

Fig. 1. Comparison between the estimated duration of surgeries made by surgeons and the real values. The diagonal line shows the desirable scenario where the estimated duration is always correct.

The economic and social effects of an improvement in efficiency through accurate estimates can be quantified in the increase of the number of surgeries carried out and the revenue obtained by hospitals. Concerning public health, this enables the possibility to reduce waiting lists for surgery with marginal costs. However, the challenge is not only methodological but also of change management. The introduction of decision support systems in this environment leads to a reduction in the decision making power of doctors and is prone to face resistance. Nonetheless, this work is only concerned with the former challenge and refrains from discussing the implementation issues.

3 Background

Operating theater planning is a major challenge throughout the scientific literature. The reviews by Cardoen et al., and Guerriero and Guido are good indicators of the vast amount of research that is being done in this field recently [1, 8]. Although there are many publications on modeling and predicting surgical durations, the most prevalent research field in the operating room is operational research, which typically addresses scheduling and planning problems. The operating theater planning is normally divided into three decision levels: (i) operational, (ii) tactical and (iii) strategic. Our work focuses on the first planning level. We combine methods from the operations research and data mining fields for the periodic (weekly) scheduling of patients to the available operating rooms. The tactical and strategic decision levels concern longer term decisions on capacity definition and allocation of resources to the different surgical specialties, also known as the master surgery schedule and case-mix planning problems.

The surgical process is naturally characterized by deep uncertainty [9], due to different sources of variability emerging since the patient arrival to his post-operative recovery. These factors affect the total duration of the patient's stay in the different areas of the operating theater. The most important, and the most studied in the literature, is the duration of the surgery, including anesthesia and the combination of surgical interventions [10]. Other significant sources of variability are the main surgeon performing the procedure and his team, the anesthesia type, and the patient's risk class, age and gender [11]. These factors increase the difficulty of the problem significantly, as there are many different possibilities (e.g., there is a wide variety of procedures and a large number of surgeons in hospitals) and they interact with each other (e.g., each surgeon is more effective on some surgeries than others and it may be easier to perform a given surgery on patients with some characteristics than others) [7]. Researchers have been modeling surgical times targeting different management decisions but most studies aim to predict surgery duration before it starts (off-line scheduling) [12, 13, 9, 10, 14]. Other tasks include predicting the time remaining during surgery execution (on-line scheduling) [15] and the duration of a series of surgeries, aiming to reduce the total overtime incurred into [16]. However, it is also recognized that, due to the uncertain nature of surgical procedures, it is often better to know the upper and lower bounds of the duration than a point estimate [14].

We found that most of the research performed in this field determines the most important factors of variation between surgeries but does not come up with a generalized estimation model that could be applied transversely in the operating theater [17, 18]. As stated by Combes et al., the statistical models represent an average phenomenon and not the variation in the subsets of observations that that average represents [12]. Nonetheless, prediction models have been developed to successfully improve the accuracy of predictions of surgery duration. Wright et al., in the mid 1990s reduced the mean absolute error relative to surgeon estimates in almost 20% using regression-based methods [19]. Marinus et al. also show significant reductions in average overtime and undertime per surgery

on a vast set of surgical procedures using a regression model [13]. Stepaniak et al., estimate the effect of several medical variables that affect a surgery in two different hospitals in the Netherlands by means of ANOVA models and use those models to estimate the duration of surgical cases, obtaining a 15 % improvement [20]. An alternative approach, in which surgeons are given tools to improve their estimates, rather then being replaced by prediction models, was proposed by Combes et al. Motivated by a real case scenario, they developed a data exploration framework to help surgeons estimate the duration of surgeries based on their past performance. However, this study limited itself by applying the methodology to a small subset of gastric procedures [12].

4 Surgery Duration Estimation

The main goal of our work is to reduce the uncertainty of the estimated duration of surgeries. This can be addressed as a regression problem, where the surgery duration is the dependent variable y and the known environment settings and patient characteristics constitute the vector of independent variables $\mathbf{x}$ that influence the duration of a surgery: $y_i = f(\mathbf{x}_i)$. The data used to carry out this work was extracted from different databases, allowing us to enrich the dataset with characteristics of the patient, surgeon and surgical procedures.

4.1 Dataset

The data used in this work describes the outpatient surgeries performed in one of the largest Portuguese public hospitals from 2006 to 2011, containing approximately 9.500 completed surgical cases.

The most relevant attributes, selected using a attribute subset evaluator algorithm, are: Gender, Priority, Week Day, Shift, ICD Disease, ICD Procedure 1 and 2, Number of Interventions to Date, the existence of circulatory problems and the average total duration of the procedure.

In order to give an overview of the surgery duration distribution, Figure 2 plots the density histogram of surgery duration on the test set. It fits a log-normal distribution ($\mu = 26.123$, $\sigma = 1.862$ found by maximum likelihood estimation) for a significance level of 0.05, verifying the distribution type fit supported by Strum et al. and Spangler et al. [18, 17].

4.2 Evaluation

The accuracy of the methods was evaluated using standard error measures. The Mean Absolute Error (MAE):

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{1}$$

where y_i and $\hat{y}_i$ are the true and predicted duration values and n is the number of predictions (i.e., the number of surgeries for which a prediction was made). MAE

Table 1. Independent variables used for predictive modeling

#	Type	Description
1	Nominal	Patient gender
2	Numeric	Patient age
3	Ordinal	Patient priority
4	Numeric	Patient waiting time for surgery
5	Nominal	Surgery month
6	Nominal	Surgery weekday
7	Nominal	Surgery shift
8	Nominal	Patient diagnosed disease
9	Numeric	Number of interventions to be performed
10	Nominal	Intervention code 1
11	Nominal	Intervention code 2
12	Nominal	Intervention code 3
13	Numeric	Number of surgeries performed on the patient to date
14	Numeric	Number of interventions performed on the patient to date
15	Binary	If the patient has undergone surgery on other specialties
16	Nominal	Surgeon identification
17	Nominal	Surgeon gender
18	Numeric	Number of times the surgeon has dealt with this disease
19	Numeric	Number of times the surgeon has performed the main intervention
20	Binary	If the patient has other diagnosis
21	Binary	If the patient has any circulatory problem
22	Binary	If the patient has diabetes or renal problems
23	Binary	If the current diagnosis is recidivist
24	Numeric	Average total surgery duration
25	Numeric	Surgery real duration

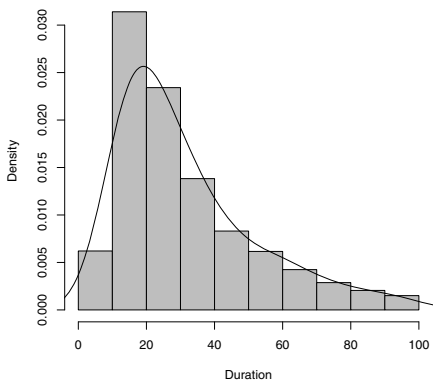

Fig. 2. Distribution of surgery durations (duration in minutes)

values are in the same scale as the dependent variable (i.e., surgery duration in this case). Alternatively, we can use Mean Absolute Percentage Error (MAPE), which rescales the values of MAE in the interval $[0, 1]$:

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \tag{2}$$

In many applications, such as this one, the impact of a small deviation is often irrelevant. So more importance should be given to larger ones. In such cases, the Mean Squared Error (MSE) measure may be more appropriate than MAE and MAPE:

$$MSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n} \tag{3}$$

An additional measure is the correlation between the predictions and the real values:

$$\tau = \frac{\sum_{i=1}^{n} (y_i \hat{y}_i) - n\bar{y}\bar{\hat{y}}}{(n-1)(s_y s_{\hat{y}})} \tag{4}$$

where $\bar{y}$ and s_y is the mean and variance of y.

These metrics do not take into account if the deviations are originated by over-or under-estimation. This is significant, due to the different effects they have on the operating theater schedule. Overestimation is when a surgery lasts less than predicted ($y_i < \hat{y}_i$), leading to idle time in the operating room time or under-utilization. Underestimation occurs when a surgery lasts longer than predicted ($y_i > \hat{y}_i$), having a greater impact on the operating room since it may disrupt and delay subsequent surgeries. Dexter proposes an asymmetric absolute error measure involving under-utilization and over-utilization costs [9]. As there are no estimates of those costs available in our application, we do not use this measure. However, in Section 6, we provide some analysis of the results in terms of over-and under-estimated time.

4.3 Experimental Setup

We compare three different data mining techniques: Linear Regression (LR), Random Forests (RF) and the M5 Rules (M5) algorithm [21]. However, rather than individual models, we have used ensembles of each of those models using the bootstrap aggregation algorithm (Bagging) (Bagging) [22], to reduce the variance and increase the estimation accuracy.

To develop and evaluate our models we separated the data into two sets. The first includes the surgeries from 2006 to 2009 and is used for training the models, while the second (surgeries in 2010 and 2011) is used for testing those models. The dataset consists of 25 variables, of which one is ordinal, five are binary, seven are numeric and the remaining 12 are nominal (Table 1). The identification of the surgical interventions and diseases are coded using the International Classification of Diseases norm (ICD-9) published by the World Health Organization.

5 Surgery Scheduling

The models obtained in the previous section are used to estimate the duration of a set of surgeries. The next step is to find the optimal schedule for those surgeries assuming that their duration is the predicted one. The capacity available (time) might be spread by different operating rooms, in different days and also different parts of the day (i.e., morning and afternoon shifts).

Our optimization model is also able to take into account constraints concerning the availability of surgeons. It is based on the 1-0 Multiple Knapsack Problem, proved to be NP-Hard [23]. In our case a knapsack is an operating room shift with a given time capacity. We want to fill the knapsacks with surgeries, which consist of pairs of patients and surgeons. Table 2 summarizes the notations used throughout this section.

Table 2. Variables used in the optimization model

Symbol	Description
N	Set of patients
R	Set of operating rooms
S	Set of surgeons
D	Set of scheduling days
T	Set of shifts (morning/afternoon)
s_i	Surgeon assigned to patient i
d_i	Duration estimated for the surgery of patient i (minutes)
A_{rdt}	Availability of operating room r on day d and shift t (binary)
S_{jdt}	Availability of surgeon j on day d and shift t (binary)
u	Operating room clean up time (constant)
c	Shift capacity (constant)

5.1 Decision Variables and Objective Functions

As mentioned above, the mathematical model devised resembles the binary multiple knapsack model. Thus, the binary decision variables x_{irdt} define the assignment of a patient i to an operating room r on a given day d and shift t, and the binary decision variable y_{jrdt} assigns the surgeon j to an operating room r on day d and shift t. Note that the model presented assigns patients to a period of time and local (knapsack) but does not sequence patients inside a knapsack. The sequence can be obtained using a simple method such as random ordering, since we constrain the problem to fix a surgeon to one operating room per shift.

The main goal of our work is to increase the efficiency of the operating room, which can be translated to the maximization of surgeries performed in a given period:

$$Max_{f_1} = \sum_{i \in N} \sum_{r \in R} \sum_{d \in D} \sum_{t \in T} x_{irdt} \tag{5}$$

Yet, increasing surgeries performed decreases the operating room utilization, due to the setup (clean-up of rooms) times between surgeries. The following expression represents the maximization of the mean utilization of all operating rooms in R over the set of days in D.

$$Max_{f_2} = \frac{\sum_{i \in N} \sum_{r \in R} \sum_{d \in D} \sum_{t \in T} x_{irdt} d_i}{c \sum_{r \in R} \sum_{d \in D} \sum_{t \in T} A_{rdt}} \tag{6}$$

5.2 Constraints

First of all, given that there is a surgeon responsible for each patient, the binary variables y_{jrdt} are linked to x_{irdt} (Eq 7). This enables us to ensures that the surgeons cannot be assigned to multiple procedures simultaneously. Additionally, the following constraints are also included in our model:

$$x_{irdt} \leq y_{jrdt}, \forall i \in N, r \in R, d \in D, t \in T, j \in S : j = s_i \tag{7}$$

$$\sum_{i \in N} x_{irdt}(d_i + u) \leq A_{rdt} c, \forall d \in D, r \in R, t \in T \tag{8}$$

$$y_{jrdt} \leq S_{jdt}, \forall j \in S, d \in D, r \in R, t \in T \tag{9}$$

$$\sum_{r \in R} \sum_{t \in T} y_{jrdt} \leq 1, \forall j \in S, d \in D \tag{10}$$

$$x_{irdt}, y_{jrdt} \in \{0, 1\} \tag{11}$$

Equation 8 represents the constraint for the available capacity per operating room. It prevents overtime, i.e., planning surgeries on a shift that take longer than the available time, and restricts patients from being assigned to operating rooms if the room is unavailable on a given day and shift. Equation 9 limits the allocation of surgeons according to their availability. Equation 10 prevents the surgeons from being allocated to different operating rooms in the same shift and day. Finally, Equation 11 states that the decision variables are binary.

5.3 Experimental Conditions

The experiments are limited to a single week and we compared the results obtained with the optimization model based on the predicted duration of surgeries with the observed schedule. We used a patient waiting list composed by 394 real patients, the total duration of the surgery is known, as well as the estimates made by the surgeons. These figures will be used to compare the performance in terms of planned and real utilization rate, number of surgeries performed, overtime and undertime. We assume there is only one operating room with one 4 hour shift available per day, on a Monday to Saturday week (Friday off). It is also assumed a constant clean-up time of 17 minutes between surgeries, and that every surgeon (8 in total) and patient is immediately available for surgery at the scheduled time. The experiments were performed using CPLEX 12.2.

6 Results

Initially we will focus on the quality of the predictions (Section 6.1) and then we analyze the results of the combination of the predictions with the optimization model (Section 6.2).

6.1 Surgery Duration Estimation

Table 3 compares the accuracy of the three data mining algorithms tested: Linear Regression (LR), Random Forests (RF) and M5 Rules (M5) with two baselines. The first is the surgeon estimates and the second is the median duration, that is, the median of the surgeries of the same type.

Table 3. Estimation errors

Metric	Surgeon	Median	LR	RF	M5
MAE	13.84	11.19	9.55	9.63	8.89
MSE	312	224	171	158	144
MAPE	47%	54%	39%	44%	37%
Correlation	0.69	0.75	0.78	0.79	0.81

The results show there is clear advantage of using the data mining algorithms for predicting surgery duration, when compared to the estimates made by surgeons, which is the method currently used in the hospital, and also to the median estimates. The M5 algorithm shows an average absolute improvement of 4.95 minutes per surgical case, which represents an improvement of almost 36% in estimation accuracy compared to the surgeon estimates. Additionally, the mean squared error decreased to half, meaning that the largest deviations were greatly reduced. The estimates are plotted against the real durations in Figure 3, showing that the ladder effect seen in Figure 1 disappeared.

Focusing just on the results obtained by the M5 algorithm and the surgeon predictions and separating them according to whether they over- or underestimate the true duration, it is noticeable that the greatest improvement comes from surgeries whose prediction was overestimated by the surgeons (Table 4). This observation shows that surgeons are mostly prone to over-estimate the duration of surgeries, thus apparently preferring to block their colleagues [7]. On the other hand, as expected, the data mining algorithms distribute their errors evenly.

Table 4. Total surgery overtime and undertime relative to the predictions (minutes)

	Surgeon	M5	Difference	Improvement
Overtime	11605	9894	1711	-15%
Undertime	18714	9588	9126	-49%

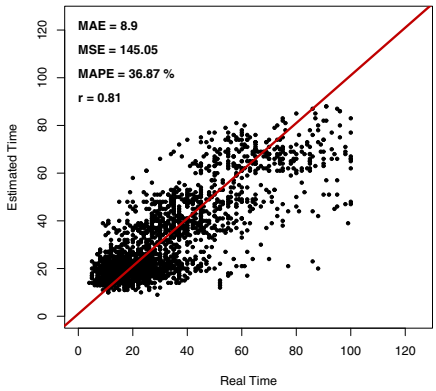

Fig. 3. Comparison between the bagged M5-Rules algorithm estimates and the real duration

Although the absolute gain per surgery may seem small (4.95 minutes) the relative gain of 36% is significant, where one could assume that the M5 algorithm provides 36% more time to perform surgeries. Considering that the average duration of surgeries is 31 minutes and that the approach proposed here could reduce the wasted time by 10837 minutes in 2010 and 2011, then this system would have enabled up to an extra 343 surgeries to be performed. Despite being a rough estimation, this still indicates that a significant improvement in efficiency could be achieved with the approach proposed here.

6.2 Schedule Optimization

Two experiments were executed to schedule an entire week of surgeries. In the first, we compare the schedule obtained by maximizing the number of surgeries performed (Eq. 5) based on the duration estimates made by the surgeons with the true schedule. This isolates the effect of the optimization component in the results. The second experiment compares the schedules generated by the optimization model using each objective functions, number of surgeries (Eq. 5) and operation room utilization (Eq. 6). In both cases the optimization is based on the durations estimated by the data mining models. This experiment aims to compare the two objective functions.

The results for the first experiment are presented in Table 5. As expected, the optimization model increased the number of surgeries scheduled but the magnitude is surprising: almost three-fold. In fact, the optimization model performed abnormally due to the selection of surgeries with very low estimated duration. This resulted in a schedule consisting solely on surgeries with underestimated duration, which naturally leads to overtime in the use of the operating theater. The surgeries scheduled would require 124% of the available time.

Table 5. Comparison of the real schedule with the one optimizing the number of surgeries (Eq. 5) based on the estimates made by the surgeons

	True Schedule						Maximize Surgeries					
	Mon	Tue	Wed	Thu	Sat	Total	Mon	Tue	Wed	Thu	Sat	Total
Number of Surgeries	4	2	3	4	5	18	10	10	10	10	10	50
Planned Utilization	45%	37%	32%	58%	69%	48%	21%	23%	21%	21%	21%	21%
Real Utilization	35%	29%	10%	58%	63%	39%	126%	131%	135%	118%	111%	124%
Overtime	40	18	55	18	54	185	252	259	273	233	216	1233
Undertime	17	0	4	20	39	80	0	0	0	0	0	0

Table 6 compares the schedules obtained by the optimization of each of the two objective functions, using data mining estimations. As stated, the first one maximizes the number of surgeries performed while the second the utilization of the operating rooms. By maximizing the utilization, less surgeries are carried out than what happened in fact, due to the clean-up time (setup) between surgeries. Therefore, the results illustrate that we may obtain results that are better in terms of the use of resources but worse in terms of the reduction of surgery waiting lists.

Table 6. Comparison of the optimized schedules using the two objective functions (Eq. 5 and Eq. 6) based on the predictions by the data mining models

	Maximize Surgeries						Maximize Utilization					
	Mon	Tue	Wed	Thu	Sat	Total	Mon	Tue	Wed	Thu	Sat	Total
Number of Surgeries	7	7	7	7	8	36	3	3	3	3	3	15
Planned Utilization	29%	29%	29%	29%	42%	31%	82%	90%	69%	74%	88%	80%
Real Utilization	47%	60%	48%	48%	65%	53%	75%	75%	75%	54%	75%	71%
Overtime	43	73	49	44	56	265	40	54	15	47	36	192
Undertime	0	0	4	0	0	4	24	19	29	0	5	77

Additionally, when comparing the schedule obtained by maximizing the number of surgeries using each of the two methods to estimate duration, surgeon estimates (right-hand side of Table 5) and data mining models (left-hand side of Table 6), we observe that the latter not only provides more accurate predictions (as observed in Section 6.1) but also generates a more reasonable schedule: the real utilization of the operating rooms is under 100%. Finally, when compared to the real schedule (left-hand side of Table 5), the schedule obtained using the method to optimize the number of surgeries based on the predictions made by the data mining models (left-hand side of Table 6), we observe a very significant improvement in both evaluation measures: twice as many surgeries are performed, resulting in an increase in real utilization time from 39% to 53%. These results clearly show the advantage of using data mining to predict the duration of surgeries.

7 Conclusions

Although the stimulus initiatives introduced by the Portuguese government to reduce waiting lists for surgery have shown positive results, they imply significant costs and are now being reduced due to the current financial crisis. We address the problem of optimizing the schedules of operating rooms with a combination of data mining and optimization techniques. To the best of our knowledge this is a novel approach. It was tested on outpatient surgeries, a subset of all the surgeries for which the accurate prediction of the duration is particularly important because it is the most important factor for the optimization of operating room usage. Our results show that it is possible to significantly increase the utilization of the rooms and the number of surgeries performed.

The performance achieved by the data mining and optimization models is sufficient to encourage the development of this work and its experimentation on different surgical departments. However, the optimization model will have to be adjusted to take into consideration downstream resources following the operating room (e.g., recovery wards).

Acknowledgements. This work was partially supported by FCT projects "An integrated framework for operating room capacity planning and scheduling" (PTDC/EGE-GES/102681/2008) and "Evolutionary algorithms for Decision Problems in Management Science" (PTDC/EGE-GES/099741/2008).

References

1. Cardoen, B., Demeulemeester, E., Beliën, J.: Operating room planning and scheduling: A literature review. European Journal of Operational Research 201(3), 921–932 (2010)
2. Bjornberg, A., Garrofe, B.C., Lindblad, S.: Euro health consumer index. Technical report, Health Consumer Power House (2009)
3. Siciliani, L., Hurs, J.: Explaining waiting-time variations for elective surgery across oecd countries. Economic Studies, vol. (38). OECD (2004)
4. Chang, J.K., Calligaro, K.D., Lombardi, J.P., Dougherty, M.J.: Factors that predict prolonged length of stay after aortic surgery. Journal of Vascular Surgery 38(12), 335–339 (2003)
5. Lazar, H.L., Fitzgerald, C., Gross, S., Heeren, T., Aldea, G.S., Shemin, R.J.: Determinants of length of stay after coronary artery bypass graft surgery. Circulation 92(9), 20–24 (1995)
6. Tu, J., Mazer, C.: Can clinicians predict icu length of stay following cardiac surgery? Canadian Journal of Anesthesia 43(8), 789–794 (1996)
7. Macario, A.: Truth in scheduling: Is it possible to accurately predict how long a surgical case will last? Anesthesia & Analgesia 108(3), 681–685 (2009)
8. Guerriero, F., Guido, R.: Operational research in the management of the operating theatre: a survey. Health Care Management Science 14(1), 89–114 (2011)
9. Dexter, F., Traub, R.D., Qian, F.: Comparison of statistical methods to predict the time to complete a series of surgical cases. Journal of Clinical Monitoring and Computing 15(1), 45–51 (1999)

10. Li, Y., Zhang, S., Baugh, R.F., Huang, J.Z.: Predicting surgical case durations using ill-conditioned cpt code matrix. IIE Transactions 42(2), 121–135 (2009)
11. Strum, D.P., May, J.H., Vargas, L.G.: Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models. Anesthesiology 92(4), 1160–1167 (2000)
12. Combes, C., Meskens, N., Rivat, C., Vandamme, J.P.: Using a kdd process to forecast the duration of surgery. International Journal of Production Economics 112(1), 279–293 (2008)
13. Eijkemans, M.J.C., van Houdenhoven, M., Nguyen, T., Boersma, E., Steyerberg, E.W., Kazemier, G.: Predicting the unpredictable: A new prediction model for operating room times using individual characteristics and the surgeon's estimate. Anesthesiology 112(1), 41–49 (2010)
14. Stepaniak, P.S., Heij, C., Mannaerts, G.H.H., Quelerij, M.D., Vries, G.D.: Modeling procedure and surgical times for current procedural terminology-anesthesia-surgeon combinations and evaluation in terms of case-duration prediction and operating room efficiency: a multicenter study. Anesthesia & Analgesia 109(4), 1232–1245 (2009)
15. Dexter, F., Epstein, R.H., Lee, J.D., Ledolter, J.: Automatic updating of times remaining in surgical cases using bayesian analysis of historical case duration data and instant messaging updates from anesthesia providers. Anesthesia & Analgesia 108(3), 929–940 (2009)
16. Alvarez, R., Bowry, R., Carter, M.: Prediction of the time to complete a series of surgical cases to avoid cardiac operating room overutilization. Canadian Journal of Anesthesia 57(11), 973–979 (2010)
17. Spangler, W., Strum, D., Vargas, L., May, J.: Estimating procedure times for surgeries by determining location parameters for the lognormal model. Health Care Management Science 7(2), 97–104 (2004)
18. Strum, D.P., May, J.H., Vargas, L.G.: Modeling the uncertainty of surgical procedure times: comparison of log-normal and normal models. Anesthesiology 92(4), 1160–1167 (2000)
19. Wright, I.H., Kooperberg, C., Bonar, B.A., Bashein, G.: Statistical modeling to predict elective surgery time. comparison with a computer scheduling system and surgeon-provided estimates. Anesthesiology 85(6), 1235–1245 (1996)
20. Stepaniak, P.S., Heij, C., De Vries, G.: Modeling and prediction of surgical procedure times. Statistica Neerlandica 64(1), 1–18 (2010)
21. Holmes, G., Hall, M., Frank, E.: Generating Rule Sets from Model Trees. In: Foo, N.Y. (ed.) AI 1999. LNCS, vol. 1747, pp. 1–12. Springer, Heidelberg (1999)
22. Breiman, L.: Bagging predictors. Machine Learning 24(2), 123–140 (1996)
23. Khuri, S., Bäck, T., Heitkötter, J.: The zero/one multiple knapsack problem and genetic algorithms. In: Proceedings of the 1994 ACM Symposium on Applied Computing, SAC 1994, pp. 188–193. ACM (1994)