# OPTIMIZATION MODELS FOR SCHEDULING AND RESCHEDULING ELECTIVE SURGERY PATIENTS UNDER THE CONSTRAINT OF DOWNSTREAM UNITS

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Ergin Erdem

In Partial Fulfillment
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Industrial and Manufacturing Engineering

May 2013

Fargo, North Dakota

# North Dakota State University
## Graduate School

**Title**

Optimization Models for Scheduling and Rescheduling Elective Surgery Patients under the Constraint of Downstream Units

**By**

Ergin Erdem

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

**DOCTOR OF PHILOSOPHY**

SUPERVISORY COMMITTEE:

Dr. Jing Shi
Co-Chair

Dr. Kambiz Farahmand/Dr. David L. Wells
Co-Chairs

Dr. Joseph Szmerekovsky

Dr. Xiuli Qu

Rachael Boyer

Approved:

| 07/02/2013 | Dr. Canan Bilen-Green |
|---|---|
| Date | Department Chair |

## ABSTRACT

Healthcare is a unique industry in terms of the associated requirements and services provided to patients. Currently, healthcare industry is facing challenges of reducing the cost and improving the quality and accessibility of service.

Operating room is one of the biggest major cost and revenue centers in any healthcare facility. In this study, we develop optimization models and the corresponding solution strategies for addressing the problem of scheduling and rescheduling of the elective patients for surgical operations in the operating room.

In the first stage, scheduling of the elective patients based on the availability of the resources is optimized. The resources considered in the study are the availability of the operating rooms, surgical teams, and the beds/equipment in the downstream post anesthesia care units (PACUs). Discrete distributions governing surgical durations for selected surgical specialties are developed for representing variability for duration of surgery. Based on the distributions, a stochastic mathematical programming model is developed.

It is indicated that with the increase of problem sizes, the model may not be solved by using a leading commercial solver for optimization problems. As a result, a heuristic solution approach based on genetic algorithm is also developed. It is found out that the genetic algorithm provides close results as compared to the commercial solver in terms of solution quality. For large problem sizes, where the commercial solver is unable to solve the problem due to the memory restrictions, the genetic algorithm based approach is able to find a solution within a reasonable amount of computation time.

In the second stage, the rescheduling of the elective patients due to the sudden arrival of the emergency patients is considered. A mathematical programming model for

minimizing the costs related with expanding the current capacity and disruption caused by the inclusion of the emergency patient is developed. Also, two different solution approaches are brought forward, one with using the commercial solver, and the other based on genetic algorithm. Genetic algorithm based approach can always make efficient decision regarding whether to accept the emergency patients and how to minimize the reshuffling effort of the original elective surgery schedule.

# ACKNOWLEDGMENTS

I would like to express my gratitude for the people who genuinely support me for during the dissertation.

Foremost, I would like to thank my advisor, Dr. Jing Shi. Without his support and help, I would not be able to complete my dissertation. His creative ideas and suggestions help me conduct the research and increase my expertise and knowledge. He helped me realize my potential by providing very valuable suggestions.

I would also want to thank Dr. Wells and Dr. Farahmand for serving as co-chairs in my dissertation committee. Their guidance and valuable suggestions helped me expand my perspective on the research topic.

Meanwhile, I would like to extend my gratitude to other committee members, Dr. Xiuli Qu, Dr. Joseph Szmerekovsky, and Rachael Boyer.

I sincerely acknowledge the continuous funding support from Department of Veterans Affairs, as well as National Science Foundation (Award # 1126570).

Lastly but not leastly, I would like to express my gratitude to my family, who is always by my side. They provided me the moral support and encourage me in my endeavors.

**TABLE OF CONTENTS**

## LIST OF TABLES

Table | Page

## LIST OF FIGURES

Figure Page

# 1. INTRODUCTION

Healthcare industry is inherently complex and depends on the interdisciplinary teams of trained professionals and paraprofessionals to meet the need of the individual patients and general population (United States Department of Labor, 2007). Increased need of the aging population along with the changing healthcare related technologies contributes to the ever increasing complexity of the healthcare industry. This industry is one of the fastest growing industries. In the U.S., since 1970’s, healthcare spending has grown at an average annual rate of 9.9% which is 2.5 percentage points higher than the growth of GDP (Kaiser Family Foundation, 2006).

Healthcare industry is facing challenges in terms of cost reduction and improvement of service. Substantial amounts of resources have been spent on healthcare industry. Current analysis indicates that 17.9% of GDP in the U.S. is spent on healthcare in 2012, reaching 2.8 trillion USD, and the trend is expected to continue in coming years (Plunkett Research, 2012). At the current rate of growth, the figure is expected to reach to 4.5 trillion USD in 2019, which will constitute 19.3% of the projected GDP in that particular year (Terry, 2010). In order to curb the increasing costs on healthcare, managerial aspects of clinic and hospital operations are being focused on more closely recently.

## 1.1. Operating Room Management and Scheduling

Among the most important cost and revenue centers, operating rooms carry important significance. It is one of the largest cost and revenue centers in the healthcare facility (Health Care Financial Management Association, 2005; Macario et al., 1995). It is estimated that in general 60-70% of all hospital admissions are generated by surgical interventions and the total expenses related with operating rooms constitute more than 40% of the expenses in a healthcare facility (Denton et al., 2007). Inefficient and inaccurate planning might cause delays and cancellations which might further lead to wastes and hence the increase of the total operation cost. The wastes should be mitigated or avoided (Gordon et al., 1998).

A closer attention to the current operating procedures for operating rooms will be beneficial for overall cost implications of healthcare facilities. The overall impact of operating rooms on the entire healthcare facility cannot be overlooked. This, in part, will be addressed by the thesis research.

Planning and scheduling operating rooms carry a specific importance. With the existence of the conflicting objectives, priorities of the stakeholders, and scarcity of the costly resources, managing operating rooms is a challenging task (Cardoen et al., 2010; Glouberman and Mintzberg, 2001). Usually, the scheduling of operating rooms involves many facets such as the patient safety and improved clinic outcomes, increasing the access of the corresponding resources utilized by the surgeons and corresponding clinical staff member, decreasing the related patient delays, improving overall satisfaction levels (e.g., patient, surgeons, clinic staff member, etc.), and increasing the efficiency related with the utilization of the corresponding resources (Alon and Schüpfer, 1999). Additionally, the increase of demand for the related surgical services due to aging population brings additional challenges for managing and scheduling operating rooms (Etzioni et al. 2003).

There are several factors that influence the optimal operating room planning. Considerations for the accompanying resources are one of the major factors in the planning of operating rooms and downstream clinic units. The clinic resources that are vital for the proper functioning of operating rooms and downstream units can be classified in different categories. One category is the human resources of clinical personnel comprising of surgical and support teams. Surgical teams usually serve during the peri-operative stage, whereas the support teams provide services during the pre and post-operative stages. The other category is facilities and equipment such as the specialized equipment used for performing specific surgeries, pre-surgical holding units for preparing the patient for surgical operation, and the post-operative holding units such as PACU (Guerriero and Gido, 2011).

### 1.2. Hierarchical Structure in Operating Room Planning

Three different approaches for classifying the management strategies for scheduling elective patients can be summarized as follows (Patterson, 1996; Guerriero and Guido, 2011),

- **Open scheduling:** This case is related with scheduling the elective cases in the medium and short term on the "first come, first served" basis. The schedule in the sense is defined by allocating the surgeries prior to the day of surgery with the aim of accommodating as many surgical cases as possible. Under this system, the surgeons might submit the cases until the day of surgery for being included in the schedule.
- **Block-scheduling:** In block scheduling practice, specific surgeons and surgical groups (SG) are assigned to available time blocks in operating rooms. Usually the time blocks are fixed with respect to the specialty and the time/date of the

week and month. To cite an instance, General Surgery operations might be scheduled on Mondays between 8:00am-12:00pm in operating rooms #1, #2, and #4. As such, the block scheduling involves two phases of decision planning. The first phase involves the construction of a cyclic timetable. The cyclic timetable can be defined as "a timetable that defines the number and available operating rooms, the hours that operating rooms will be open, and the SG and surgeons available for each operating room block" (Blake et al. 2002). The second problem phase involves filling up the time blocks with surgical cases such that the surgical operations can be performed within the scheduled time period. This practice involves creating a master surgical schedule. An example of master surgical schedule obtained from Mt. Sinai Hospital in Toronto, Canada is provided in Table 1 (Blake and Donald, 2002).

- **Modified block scheduling:** This practice involves the modification of block scheduling in such a way that some time blocks might be left open and some of them might be booked. Unused time blocks might be released before the time of surgery. Therefore, the open-scheduling practices might be applied. A master surgical schedule is constructed but there is no necessity that time blocks are assigned to any surgeon or surgical group (Dexter, 2000).

Block scheduling entails several advantages as listed below (Unibased System Architecture, 2011),

- The surgical team has a clear idea about the surgery schedule in advance and can adjust clinic appointments based on this schedule,

- The patients can be scheduled and the surgical teams can be dispatched based on the allocated OR times with respect to block scheduling practices,
- The workload for the surgeons, nurses, and operating room staff members can be evenly distributed among the days of the week,
- The admission to PACU and the corresponding intensive care unit (ICU) can be better planned through the distribution of load on different days of the week based on the generated block schedule of the surgeries.

Table 1. An example of master surgical schedule used in Mt. Sinai hospital of Toronto, Canada

| | **Main 1** | **Main 2** | **Main 3** | **Main 4** | **Main 5** | **Main 6** | **Main 7** | **Main 8** | **Main 9** | **Main 10** | **OPS 1** | **OPS 2** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mon | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-15:30 | Surg, Otol[1] 08:00-15:30 | Gyne 08:00-15:30 | Opht 08:00-15:30 | Not Staffed | Not Staffed | Oral 08:00-16:00 | Gyne 08:00-15:30 |
| Tue | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-15:30 | Otol 08:00-15:30 | Gyne 08:00-15:30 | Oral, Opht[2] 08:00-15:30 | Not Staffed | Not Staffed | Gyne 08:00-15:30 | Gyne 08:00-16:00 |
| Wed | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-17:00 | Surg 08:00-15:30 | Otol 08:00-15:30 | Gyne 08:00-15:30 | Gyne 08:00-15:30 | Not Staffed | Not Staffed | Gyne 08:00-16:00 | Opht 08:00-15:30 |
| Thu | Surg 08:00-17:00 | Surg 08:00-17:00 | Gyne 08:00-17:00 | Gyne 08:00-17:00 | Surg 08:00-15:30 | Gyne 08:00-15:30 | Open 08:00-15:30 | Opht 08:00-15:30 | Not Staffed | Not Staffed | Gyne 08:00-16:00 | Opht 08:00-15:30 |
| Fri | Surg 09:00-17:00 | Surg 09:00-17:00 | Surg 09:00-17:00 | Surg 09:00-17:00 | Surg 09:00-15:30 | Otol 09:00-15:30 | Gyne 09:00-15:30 | Opht 09:00-15:30 | Not Staffed | Not Staffed | Oral 09:00-15:30 | Gyne 09:00-16:00 |

[1]: Surg weeks 1 & 2; Otol weeks 3, 4, &5
[2]: Oral weeks 1, 2, & 3; Opht weeks 4 &5

The master surgical schedule have five surgical types, namely, General Surgery (Surg), Gynecology (Gyne), Ophthalmology (Opht), Otolaryngology (Otol), and Oral Surgery (Oral). The type of operation and the specified duration are indicated in the corresponding cell. For example, the General Surgery procedures are conducted in the main operating room #1 on Mondays between 08:00 and 17:00 am. Thursday time slots for the main operating room #7 are left open to be filled mainly on first-come, first-served basis.

This practice promotes flexibility in the system. The master surgical schedule presented in Table 1 presents the case of the modified block scheduling approach, where open time slots as well as reserved blocks are presented in the system.

In terms of the how actual scheduling practices are conducted, two different approaches are usually taken. The first one is advance scheduling, which involves scheduling surgical operation for a particular day. According to this approach, all surgical operations are scheduled at once (Magerlein and Martin, 1978). After surgical operation is assigned for a particular day, the assignment of the surgical operation to a specific operating room is conducted and the starting time of the surgical operation is determined. Advance scheduling depends on the various constraints such as the operating room time, beds, nursing and operating room staff, and the corresponding equipment used in operating rooms. Advance scheduling practices are in parallel with block scheduling approach where surgical cases are assigned to the surgical time blocks that have been determined based on master surgical schedule. The second approach is based on the first come-first served basis which is generally taken with the open scheduling practices (Magerlein and Martin, 1978). In this approach, usually, surgical cases are scheduled based on the order of the arrival (i.e., when the need for the surgical operation arises for a particular patient) based on the availability of resources one by one. During 1970’s, the first- come, first-served scheduling is usually the preferred approach (Rinde and Blakely, 1974). However, with the changing impetus on the cost and revenue considerations, advanced scheduling approaches gain momentum starting with 1980’s.

In general, open scheduling promotes more flexibility as compared to block-scheduling (Fei et al., 2009a). However, since early 1980’s, the paradigm in operating room planning is shifted to the block scheduling practices, and open scheduling is seldom implemented (Gabel et al., 1999). As previously stated, depending on the time block and variety of the surgical sequences, the block scheduling practices might decrease the total set-up times required for surgical operations. To cite an instance, suppose operating room #6 in Table 1 is reserved for General Surgery operations for the first and second week, and Otolaryngology operations for third, fourth and the fifth weeks as opposed to open scheduling where the General surgeons and Otolaryngology surgeons fill the time slot based on first come first served basis. Instead of wasting the precious amount of operating room time with set-up involved in switching between General Surgery and Otolaryngology, as per block scheduling practices, operating room might be prepped for Otolaryngology operations during third, fourth and fifth weeks with ease, because the conversion from General Surgery to Otolaryngology operations is only required once a month. Detailed operating room scheduling practices for the short term might be exercised for reducing set-up times in the open scheduling practices. On the other hand, block scheduling generally addresses those issues in better manner based on the grounds that master surgical schedule can be formed with the consideration of decreasing the total amount of set-up times as much as possible. Usually, the merits associated with block scheduling practices decrease, if the variety and type of the operations for incorporating seldom occurring cases increase. Increase in the arrival rate of emergency patients also presents a difficulty for the creation and implementing master surgical schedule (Guerriero and Guido, 2011).

In terms of hierarchical decisions of operating room planning, three decision levels can be identified. This can be summarized as follows (Kennedy, 1992; Wachtel and Dexter 2008; Vissers et al., 2001; Testi et al., 2007),

- **Strategic session planning:** The main objective in this phase is to distribute the operating room times among different operating rooms. It is considered as a case-mix planning problem.
- **Tactical level planning:** This stage involves developing the master surgical schedule (MSS) based on the decision given in the strategic level. The master surgical schedule is formed based on the operating room time allocated for each surgical group.
- **Operational level planning:** This level of planning involves scheduling elective patients on the daily basis. Note that prior to giving the decision, the master surgical schedule is already formed and the cases are assigned on a daily basis based for this schedule (Gabel et al., 1999). It might also involve the reservation of specialized equipment and last-minute changes to the elective surgery schedule (Guerriero and Guido, 2011).

Note that the distinction among the decision levels in the open scheduling practice is usually less strict than that in the block scheduling approach (Guerriero and Guido, 2011).

### 1.3. Surgical Patient Characteristics

Usually, the distinction of the patient characteristics is based on the nature of the surgical operation conducted on the patient. In that regard, the emergency patients are usually considered to be the urgent cases where the surgery should be conducted as soon as possible (Lamiri et al., 2008). It can be said that the emergency surgery is almost all the time unexpected. Guerriero and Guido (2011) specify the length of the time window for operating emergency patient as two hours. Typical cases that require urgent surgical intervention include but are not limited to the following (Encyclopedia of Surgery, 2012),

- Invasive resuscitation due to acute respiratory failure, pulmonary embolism, etc.,
- Blunt object penetrating chest, abdomen due to various traumas (i.e., car accidents, gun-shot wounds),
- Burns,
- Cardiac events such as heart attacks,
- Aneurysms,
- Brain injuries or similar urgent neurological conditions,
- Perforated appendix, ulcer, or peritonitis

The nature of the emergency surgeries necessitates prompt action where the surgeon and surgical team might have limited opportunity for collecting additional information on the patient’s medical history and current clinic condition as opposed to elective surgeries (Encyclopedia of Surgery, 2012).

On the other hand, elective surgeries are usually planned in advance that usually do not involve medical emergency. Since most of the surgical operations are elective in nature and can be planned in advance, researchers have done more work in developing models for scheduling of elective patients.

Apart from the two categories of elective surgeries and emergency surgeries, there is a third category called urgent cases. These cases refer to the non-elective cases in which the patient is sufficiently stable that surgeries can be postponed for a short time period, from several hours to 48 hours (Cardoen et al., 2010; Guerriero and Guido, 2011).

It should be noted that there are alternative ways of classifying the surgical patients based on the frequency of the occurrence. According to van Oostrum et al. (2008a), one such alternative way can be represented as,

- Frequent elective cases,
- Dummy elective cases which occur rather seldom,
- Emergency cases

In addition, add-elective cases are considered to be the cases that are scheduled to fill the remaining time capacities of operating rooms. The add-elective, emergency, and urgent cases are collectively termed the add-on cases (Guerriero and Guido, 2011).

## 1.4. Problem Definition

Considering the importance of the operating room planning and scheduling and the inherent complexity in the decision making process, in this research, our focus is on scheduling the elective patients and rescheduling the elective patients due to the admission of emergency patients by considering various resource constraints. These resources are related with operating rooms or downstream clinical units.

As such, the thesis research problem encompasses two stages. The first stage involves scheduling of elective patients with resource constraints in operating rooms and downstream clinic units. In addition to the current level of resources, the possibilities for expanding the current capacity to accommodate the schedule of elective patients are also considered. In essence, the successful operating room planning and scheduling involve careful consideration of the resources to put them in the best use. In that regard, scheduling of the elective patients dictates the use of the resources and is closely linked with the resource allocation. Considering the fact that operating rooms are the major revenue and cost center in a healthcare facility, it will be worthwhile to manage the resources in the most efficient manner by optimally planning and scheduling operating rooms. Moreover, operating rooms are tightly linked with the downstream clinical units such as PACU and other clinical units along the patient and material flow. Providing a sound approach for operating room planning and scheduling will not only help locally increase the efficiency of operating rooms, but also help increase the overall efficiency in the healthcare facility.

The second stage of the thesis research problem involves examining the effect of admitting the emergency patients to operating rooms. For healthcare facilities, there might not be any reserved operating room allocated exclusively for the emergency arrivals and elective and emergency patients might compete for the same set of resources. In those cases, current elective surgery might be disrupted and surgeries for the elective patients might need to be postponed or preponed to accommodate the emergency patients. Also, given the elective patient schedule, the problem of whether admitting or turning down the emergency patient is examined.

Other than PACU units, there are other pathways that patients might follow after surgical operation. For some of the cases, if the surgery is majorly invasive, and the patient is at high risk of complications, the patient might be transferred to intensive care unit (ICU) immediately after the surgical operations (Sutter Health, 2012; Iyer, 2001). On the other hand, especially after minor surgical operations that are conducted under local anesthesia, the patient might spend some time in an operating room, and might be discharged directly without being transferred to PACU or ICU units.

The thesis research focuses on the typical case where patients are transferred to PACU units after the surgical operation. After most of the surgical operation, as previously stated, unless the surgery is majorly invasive or involve potential complications, patient is transferred to PACU units. In that regard, in this thesis research, the typical pathway where patients having the surgical operation recover in the corresponding PACU units is considered for the analysis. Figure 1 describes the typical flow of patients on hospital floor whereas Figure 2 provides an overall view related with the problem structure.

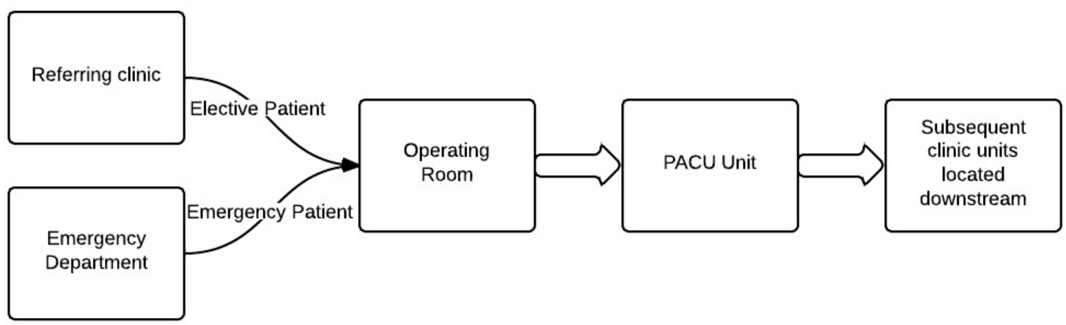


Figure 1. Typical flow of patients on hospital floor

Based on Figure 2, elective patients are referred for the surgical operations by the corresponding sub-specialty clinics. The first stage of the problem aims at creating the optimal surgery schedule based on the availability of the corresponding resources while considering the expansion of the current resource levels.

The output of the first stage problem constitutes the input for the second stage problem where the two types of decisions are given. The first decision is whether to admit emergency patient(s). If emergency patient(s) is/are admitted, the new schedule based on minimizing the disruption of the current surgery is formed. While forming the new schedule, availability of the resources and possibilities for expanding the current capacity are also considered. If the emergency patient(s) is/are not admitted, the original elective schedule is retained.

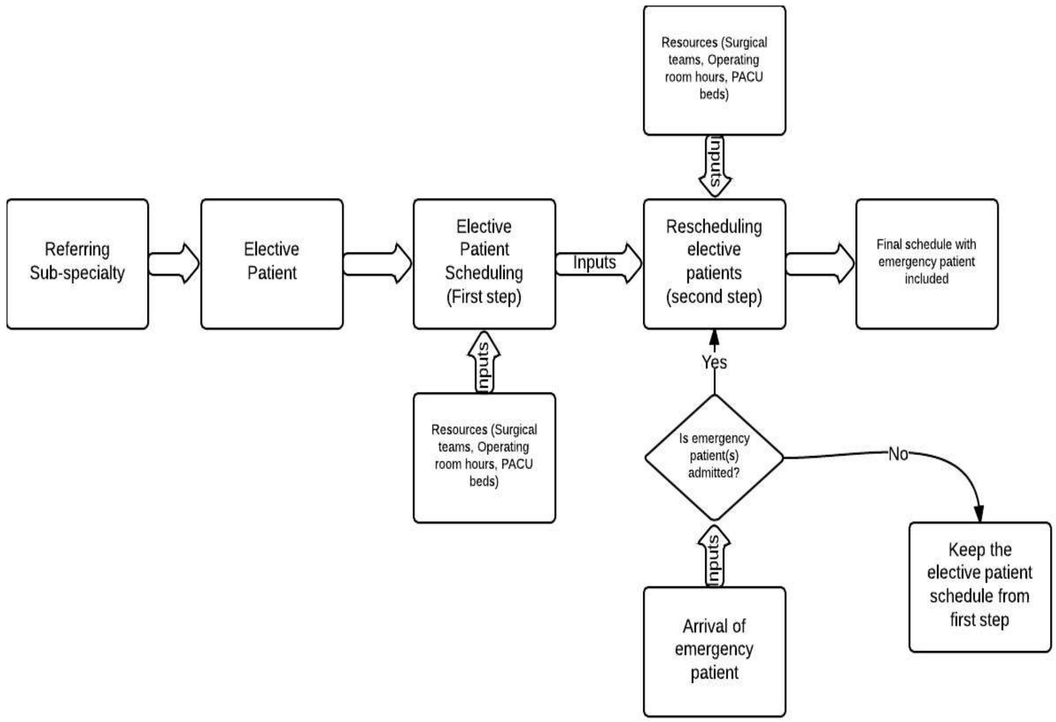


Figure 2. Schematic of the research problem structure

### 1.4.1. Additional considerations on elective patient scheduling

A sound scheduling practice should consider the efficient use of the resources. The PACU unit poses an important concern for the flow of operations in the healthcare facility and significantly affects the flow in operating rooms. The transfer of patients to the PACU units might be delayed for various reasons such as the non-availability of the PACU beds. Therefore, the integrated approach that simultaneously considers the availability of the PACU beds as well as the schedule of operating rooms is necessary.

Duration of surgical operations is another important consideration. Significant variations might exist in terms of the surgical durations due to various reasons. Some variations might be related to the surgeons, while others might be attributed to the requirements associated with the individual surgical operations. To cite an instance, significant variations might be observed for tumor removal operation due to the size of the tumor, and potential complications associated with removing the tumor. Therefore, a robust scheduling approach that accounts for the variation of surgery time duration might be necessary to utilize the resources associated with operating rooms and downstream clinic units in the best manner possible.

In order to improve planning and scheduling of operating rooms, expanding the current capacity of the resources utilized in operating rooms and downstream clinic units should also be considered. For this purpose, various expansion strategies might be considered. For operating rooms, the overtime practices might help expand the current capacity by increasing the number of working hours. Additionally, the possibility of hiring additional surgical teams is also an important consideration to provide a more flexible approach for governing the use of the available resources in the best manner possible. Also, expanding the capacity of the PACU units by incorporating additional equipment/bed will help level the utilization of resources in operating rooms and PACU units. While considering the expansion of the capacity of the available resources to facilitate the patient flow, the associated costs should not be ignored. In addition, expanding the current capacity may not serve the best interest of a healthcare facility if some efficiency issues are incurred.

### 1.4.2. Additional considerations on rescheduling elective patients

Making unprepared changes to the elective patient schedule is likely to cause inconveniences from various perspectives such as staff members, patients, and use of the equipment. This is elaborated in the following. To minimize the disruptions and reduce the costs for expanding the current capacity, in the second stage of the thesis research problem, the decision of including the emergency patients with respect to available resources and the consideration for expanding the current capacity are provided.

Disrupting the current elective schedule might create inconvenience for the surgical team members. For example, a cardiac surgeon might be assigned a time slot in an operating room, and he/she performs cardio-vascular surgical operation in the same day of every week, and the supporting staff takes care of pre-operative and peri-operative stages on that particular day. If the existing schedule is disrupted, it might be difficult to re-assign the required teams for performing those types of operation in a different date/time of the week.

Inconvenience might also arise for the elective patients. Given that the elective patients are already scheduled, they might not be receptive to the idea of preponing or postponing their surgical operations given the stressful nature of surgical operations. Postponing a surgical operation for a couple of hours for a particular patient might not cause a great deal of inconvenience. However, preponing the surgical operation or postponing the surgical operation to the next day or a couple of days in advance might lead to patient dissatisfaction and should be avoided as much as possible. The problem is aggravated if a short notice is given to the patients regarding the rescheduling of their surgical operations. Moreover, the disruption of the existing elective patient schedule might also lead to the readjustment of the medical equipment to different operating rooms, which causes more inefficiency. Certainly, the disruption of the existing schedule will most likely propagate to the downstream clinic units. The PACU might be overcrowded and bed blocking cases where the patients might not be transferred to the PACU units. Therefore, in order to provide smooth operation in the PACU units and ensure the leveling the corresponding resources in operating rooms and PACU units for increasing the efficiency, disruptions to the existing schedule should be minimized.

Clearly, it is worthwhile to consider the disruptions in the current elective schedule from different perspectives. One way to overcome this problem is to expand the current capacity to accommodate those changes. However, expanding the capacity also has some drawbacks since usually high cost is involved for increasing the level of the available capacity. This tradeoff should be carefully considered in practice. In particular, for facilities that have limited resources and unpredictable demands of emergency admissions, expanding the capacity is not a wise decision. In this case, the existing resources should be utilized in the best manner possible to accommodate the potential changes in the elective surgery schedule due to inclusion of the emergency patients. It is the exact situation that this thesis research is targeted at.

In many occasions, turning down the emergency patients might also be an option. The tradeoff between turning down and admitting the emergency patients is an important consideration. If the emergency patients are turned down, no changes in the elective patient schedule is required. However, the opportunity cost of turning down the emergency patients should not be ignored. If an emergency patient is admitted, then a sound approach for minimizing the disruption to the existing schedule should be developed which might involve expansion of the current capacity. Another facet of whether admitting or rejecting an emergency patient is that the decision should be given in a limited time window considering the nature of the emergency cases. For these types of emergency, the situation warrants that the medical intervention should be performed within an hour or so to decrease the mortality rates due to the effects of trauma and increase the chance of survival (Baez et al., 2006; Wilde, 2011). This concept is called the “Golden Hour” of trauma (Sacra and Martinez, 2009).

## 1.5. Outline of the Dissertation

In Chapter 2, we provide extensive review on the current literature on the planning and scheduling of operating rooms. The current management practices along with the hierarchical decision making process for short term and long term are discussed. In addition, the current work on scheduling and rescheduling of the elective patients are discussed in detail. Moreover, additional information on the stochastic nature of the surgical operations in operating rooms such as the variation of surgical durations and the uncertainties related with the arrival of emergency patients are discussed.

In Chapter 3, our main focus is on developing the corresponding solution methodologies for the scheduling of elective patients. Detailed information on the mathematical model and the accompanying genetic algorithm based approach is provided. A comparison between the genetic algorithm and the commercial solver based on GAMS platform in terms of the computation time and the solution quality is provided as well. Various other considerations, such as excluding the constraint of downstream units and adopting the deterministic values for the surgical durations are also explored. and the results are compared with the base stochastic problem.

In Chapter 4, the solution approaches developed for the rescheduling elective patients are examined. The link between the scheduling of the elective patients and the rescheduling is established. The problem is again formulated as a mathematical programming model, and solved by using the commercial solver and a genetic algorithm approach. A comparison between the two solution approaches is made in terms of the solution quality and computation time.

In Chapter 5, we draw the general conclusions, and point out future research directions.

# 2. LITERATURE REVIEW

As previously discussed, operating rooms and the downstream units are major cost and reveue center and the healthcare industry is directing a great deal of attention to this field to reduce the costs and improve the return on the financial assets. In parallel, there is a vast body of research literature in the operating room planning and scheduling field. It is indicated that nearly half of literature published after 1950 is from the 2000-2010 period. This reflects the increasing interest of researchers on the management and scheduling of operating rooms (Cardoen et al., 2010).

## 2.1. General Review on Operating Room and Scheduling

Magerlein and Martin (1978) provide a review on estimating the surgery times, as well as advance and day-to-day scheduling of the patients in surgical suites. In addition, a general outline for improving the overall performance of operating rooms is also provided with discussion on underlying reasons for failure to implement proposed scheduling schemes.

On the other hand, Blake and Carter (1997) provide a structured review on the surgical process and provide a unifying view on the current terminology on the surgery scheduling and operating room planning. Based on this approach, the framework for identification and reserving all resources that are external to operating room setting but vital for providing the required care for the patients undergoing surgical operations are proposed. In addition, they define the boundaries for strategic, administrative, and operational level decision making with respect to the operating room management. Przasnyski (1986) provide a literature review on cost containment and scheduling of specific resources with respect to operating rooms. Guerreiro and Guido (2011) also provide an extensive review on operating room scheduling with the main emphasis on the classification based on the decision levels. They mainly focus on the mathematical models for representing the set of the relevant problems in operating room management and describe the development of corresponding approaches to solve the formulated problems.

In the literature, in a broader context, some reviews are provided on the application of operations research and mathematical programming models in the general healthcare settings. Boldy (1976) provides a literature review on the application of mathematical programming for healthcare industry. The review primarily focuses on tactical and strategic health and social service problems. The author further discusses patient mix models where the primary consideration is to maximize the number of patients in a given time period subject to clinical constraints on departmental capabilities and minimum patient requirements. Similar approaches for reviewing the existing literature are taken by the Pierskalla and Brailer (1994), Smith-Daniels et al. (1988), and Yang et al. (2000). Although the literature reviews conducted by those authors do not specifically target at operating room scheduling and management, they carry importance with regard to application of operations research and optimization models in health care service delivery context.

### 2.2. Research on Strategic and Tactical Level Decisions

Operating room scheduling and planning at strategic level is usually seen as a resource allocation problem (Blake and Carter 1997). In that regard, the primary goal is to determine the number and type of surgeries to be performed based on the availability of resources (Guerriero and Guido, 2011). Dimensioning of the other resources that are critical in the management of operating rooms should also be specified at this level so that the system will function in a relatively smoother manner (VanBerkel and Blake, 2007). Usually, strategic planning is based on historical data and forecasts with planning horizon longer than one year generally, but some authors conclude that forecasts might not be accurate especially for time periods longer than a year (Masursky et al., 2008). Using the persistence based methods (i.e., using the last period realizations for the current period forecasts) might be a viable approach for getting accurate figures (Dexter et al., 1999a). Agnetis et al. (2012) address the long term planning in the operating room environment. They examine the tradeoff between the organizational simplicity which involves implementing the nearly same master surgical schedule that does not change completely every week versus dynamically adapting the master surgical schedule with regard to the waiting lists with the lean lists. The authors conclude that even introducing a very limited degree of variability might pay off in terms of resource efficiency and due date performance. Due date performance is measured by the amount of time elapsed by the actual surgery schedule and the date where the surgery is performed. They also investigate the scalability of the approach in a medium sized hospital in Italy.

On tactical level, the creation of cyclic master surgical schedules usually deals with satisfying the demand for surgical procedures, while considering the corresponding availability of surgical teams and dedicated equipment. In addition to financial considerations such as giving priority and providing additional time blocks for the surgical procedures with higher profit margin, minimizing operating room costs and maximizing utilization of critical resources and professional receipts are also considered (Kuo et al., 2003). In other studies, the equity assignment among surgical specialties (Blake et al., 2002; Blake and Donald, 2002) and the availability of beds in the downstream units are also considered (Beliën and Demeulemeester, 2007; Beliën et al., 2009). In addition, efficient allocation of operating room time among various surgical groups to reduce the waiting time of elective cases is also tackled (Zhang et al., 2009). Allocating operating room time is an important consideration for creating the master surgical schedule thus forming time blocks allocated for different surgical groups. Utilization of operating rooms as well as the minimization of the overtime practices for elective cases are also tackled in the literature (Fei et al., 2009b).

Tactical level decision usually involves the development of surgery schedule that usually last from 1-3 months to one year. Usually, the main objective is to create homogeneous groups in terms of surgery duration and length of stay (LOS), as well as the diagnosis related group and procedure codes (Guerriero and Guido, 2011). Creating homogeneous groups in terms of surgery duration and LOS generally leads to a balanced use of the clinical resources including surgical staff members and beds in PACU and hospital wards. Creating homogeneous groups in terms of the diagnosis related groups and procedure codes are usually helpful for decreasing set-up times in operating rooms. There are several clustering approaches in the literature that can help group the similar surgeries. In that regard, van Oostrum et al. (2008b) apply the hierarchical clustering approach to minimize the number and volume of the so-called dummy surgeries that cannot be grouped with other surgical procedures. The study, based on real data obtained from Beatrix Hospital located in Netherlands, aims at leveling of the hospital ward occupancy and the optimization of the operating room utilization.

Bed occupancy in the corresponding hospital wards is an important consideration during the preparation of surgical schedules and several studies have tackled this research topic. Beliën and Demeulemeester (2007) adopt performance measures such as the daily expected bed occupancy and its variance. The researchers calculate the expected bed shortages along with the probability of shortages each day, and they further develop the model based on the modification of objective function value so that the simple and repetitive surgical schedule might be formed. This practice helps level the average of the variance of bed occupancy in different hospital wards and the schedule might help decrease the number of instances where particular operating rooms need to be utilized among different surgical groups in the same day (Beliën et al., 2009).

Other than bed occupancy issues, some studies consider the maximization of operating room capacity and leveling of the patient outflow to the PACU and intensive care units during the construction of master surgical schedule. For instance, Vissers et al. (2005) formulate a mixed integer linear programming (MILP) model that minimizes the under- and over-utilization of the resources. Among the resources considered in this study are the over-time hours, intensive care beds, intensive care nursing staff, and medium care beds. The model is based on the deterministic LOS. Other factors such as the preference of surgeons for the time blocks, sequence activities, and the specialty capacity restrictions are also taken into account.

The distinction between stochastic and deterministic LOS is investigated in Adan et al. (2009). In that regard, they present an approach where the length of stay is stochastic in the intensive and medium care units. In line with this approach, they compare the results with those of Vissers et al. (2005). It is concluded that the new approach can lead to more robust planning of the intensive and medium care units in the presence of small variability in LOS. Meanwhile, focusing only operating rooms and disregarding the pre and post-operative units might create serious inefficiencies and bottle necks in the system especially in the presence of high LOS variability.

### 2.3. Elective Patient Scheduling

Dexter et al. (1999b) compare the performance of different scheduling algorithms to schedule add-on elective cases in operating rooms in terms of resource utilization. They conclude that the best fit descending algorithm with fuzzy constraints provides the best results in terms of operating room utilization. The best fit descending algorithm sorts add-on elective cases based on scheduled duration and assign the longest cases first to the operating rooms. During this assignment procedure, the fuzzy constraints generated based on the operating room time are also taken into consideration. In a similar token, Dexter et al. (1999c) employ the computer simulation approach for modeling operating room scheduling. They investigate the relationship between operating room utilization and various other factors such as the average length of time the patients wait for the surgeries. It is concluded that the practice of allocating block time for the elective cases based on the expected total hours for the elective cases provides best results for maximizing operating room utilization. It is also concluded that scheduling patients for the first available time block if the available time blocks exist within four weeks also help improve operating room utilization. If a case cannot be scheduled in those time blocks, it might be shifted into buffer time slots outside the block time. Hans et al. (2008) develop a constructive heuristics and local search methods to increase the utilization of operating rooms and reduce overtime practices. The approach can help free up some capacity in operating rooms. By clustering the types of surgeries with similar variability of duration, the “portfolio effect” is employed to reduce the overall variation and free up the capacity. Moreover, the solutions put the similar type of surgical operations in the same OR day so that the set-up times for operating rooms can be decreased and convenient schedules for surgeons can be created. However, the availabilities of surgical teams and OR staff are not taken into the consideration in the study.

Usually, scheduling decisions are provided at the operational level. The main decision given at the operational level is related with assigning the patients and then sequencing them in the operating rooms (Cardoen et al., 2009). The former is known as the advanced scheduling, and the latter is known as the allocation scheduling.

Many researchers employ the two-phase approach where the advanced scheduling and the allocating scheduling problems are solved concurrently (Guinet and Chaabane, 2003; Fei et al., 2006; Jebali et al., 2006; Fei et al., 2010; Wullink et al., 2008). While Guinet and Chaabane (2003) and Fei et al. (2006) address the problem focusing only in operating room settings, Jebali et al., (2006), Wullink et al. (2008), and Fei et al. (2010) develop approaches in an open-scheduling system where the corresponding problems are formulated as a two-stage hybrid flow shop problem involving other clinical units as well.

Another approach is to bring these two problems under the same umbrella by developing a unified approach (Roland et al., 2006; Roland et al., 2010). Roland et al., (2006) propose an approach where the opening costs of operating rooms as well as overtime hours is included in the objective function in a minimization problem setting. One unique contribution is that nonrenewable resources (i.e., pharmaceuticals and sterile materials) are also included in the mathematical model as the associated resource constraints, along with the renewable resources such as surgery teams. Roland et al. (2010) develop a slightly modified version of the model developed by Roland et al. (2006). The modified model incorporates additional flexibility in terms of human resources. Both approaches make use of heuristics based on genetic algorithm for solving the mathematical programming model.

Mulholland et al. (2005) implement a mathematical programming model for deciding the case-mix for the elective patients. They conclude that the optimal solution favors the surgical procedures that require inpatient care with improvements of 16.1% in hospital total profit margin and 3.6% in professional payments. In addition, the researchers conclude that with the changing case mix, under the optimal solution schedule, substantial changes in terms of general care and ICU resource utilization are required.

Some of the mathematical models feature multi-objective objective functions rather than the single criterion (e.g., monetary figures used by the Mulholland et al. (2005) in the objective function). To cite an instance, Ogulata and Erol (2003) employ a hierarchical multiple criteria mathematical programming approach for scheduling patients in operating rooms. The model considers three stages. In the first stage, the weekly patient acceptance planning is conducted, in the second stage, assignment of the patients to the surgeon groups are conducted. In the third, and the last stage, day-to-day scheduling of the patients is performed. They also provide case study from the large hospital in Turkey and conduct the sensitivity analysis based on the simulation study to verify the results provided by the mathematical model.

Zhang et al. (2009) develop a mathematical model for the weekly operating room allocation template with the objective being the inpatients' cost measured as the LOS. The clinical constraints and the case urgency priority are included in the formulation as well. In order to incorporate the effects of the randomness of the process, such as surgery time, demand, arrival time, and no-show rates, a simulation model is utilized and a case study is implemented.

Scheduling might be conducted in an inpatient and outpatient settings. Usually the distinction between inpatient and outpatient is based on the duration of the stay of the patient after the surgery. Usually when patients who undergo surgery stay overnight in the healthcare facility, they are considered as the inpatients, whereas outpatients are considered to be the ones who usually leave the hospital in the same day of admission. Although, most of the researchers do not indicate the type of patients, some authors make that distinction (Adan and Vissers, 2002). The researchers employ a mixed integer programming model to identify the number and mix of patients to be admitted to the hospital. Their model also takes the utilization of the key resources such as operating rooms and intensive care units with respect to the stated target levels into account.

Jebali et al. (2006) implement two-stage model in which in the first stage comprised of assigning the surgical cases to the corresponding operating room. In the second stage, they develop the model for sequencing of the surgeries in the given operating room. They compare the two strategies in which in first strategy, the decision of allocation of the surgical cases to operating rooms is not reconsidered in the second stage. Second strategy involves reconsidering the decisions given in the first stage in a less constrained setting.

The authors indicate that the first strategy provides fairly compatible results in terms of the solution quality as compared to the second strategy at decreased computation time.

In a similar fashion, Roland et al. (2010) develop a heuristic based approach for involving primarily with the human resource constraints and compare the solution with the mixed integer programming model. Planning and scheduling phases are conducted simultaneously. While preparing the schedule, the preferences of the staff members for operating rooms are also taken into consideration.

### 2.4. Emergency Patient Considerations

Most research efforts focus on scheduling and planning of operating rooms for the elective patients by assuming that dedicated resources (i.e., operating rooms, surgery teams) are allocated for the emergency patients. Nevertheless, some researchers study this problem for the patients from both of the streams (i.e., both emergent and elective) competing for the same resources. Wullink et al. (2007) employ a simulation study to make a comparison between the two policies. The first policy allocates a dedicated operating room for the emergency patients and the second policy allocates the reserved capacity for the operating rooms in which the elective patients also receive surgery. Three performance measures, namely the waiting time for the patients, staff overtime utilization, and operating room utilization are used. The results show that there is a drastic reduction in terms of waiting time for the emergency patients and slight increase in operating room utilization when some reserve capacity is allocated in the operating rooms for emergency surgeries. The researchers conclude that promoting the flexibility in terms of staff availability and promoting the practice of allocating reserve capacity in the operating rooms will streamline the flow in the overall system thereby reducing the waiting time for emergency patients in the clinic flow.

Similarly, Bowers and Mould (2004) employ a simulation study to investigate the effects of incorporating the elective surgeries in the operating rooms allocated for the trauma cases for the orthopedic cases. The researchers point out that incorporating elective patient cases in the emergency rooms promotes flexibility and increases the system-wide utilization. On the flip side, the associated risk of cancellation or deferral of the elective operation due to the arrival of emergency cases may increase. This is a viable option based on the premise that most elective patients prefer an earlier treatment at the expense of the associated risk of cancellation. It is pointed out that 13% of the elective patient demand for the orthopedic cases can be supplied by sharing operating rooms between the two types of patients. Although the results are dependent on the mix of patients admitted for the surgical operations, the practice of sharing operating rooms between the elective and emergency patients does not have significant effect on other performance measures such as the number of times the scheduled surgical operations cannot be completed in predefined time interval and overtime practices that are needed for a given time period.

However, not all studies support the practice of reserving allocated capacity for emergency cases in operating rooms that are primarily scheduled for the elective cases. For instance, Bhattacharya et al. (2006) suggest that utilizing the dedicated operating rooms for the elective cases might actually have beneficial effects in terms of the overall flow in the system by reducing the overtime utilizations. The authors, through retrospective study, demonstrate that overtime utilization by shifting add-on cases to the regular time hours in the dedicated operating rooms for elective patients actually reduces the overtime utilization of operating rooms by 72%. This improvement is highly unlikely, if same operating rooms might be utilized both for elective and emergency cases. This is due to the increased load and additional uncertainty introduced by the emergency cases. These findings, combined with previous studies, demonstrate that the trade-off between allocating the dedicated operating rooms for the emergency patients versus mixing these two streams of patients in the same operating rooms should be investigated carefully. The decision of allocating the operating rooms for multiple streams of patients is highly dependent on the actual clinic settings.

Everett (2002) develops a decision model for prioritizing the elective surgeries. By feeding the type of surgery and the urgency of the case, based on various performance measures, the schedules of the elective patients are created. This approach might also be used as a comparison tool for the different approaches in which the alternate policies can be formed in a multi-dimensional setting.

Cardoen and Demeulemeester (2008) evaluate the impact of the clinic pathways through discrete-event simulation. Based on multitude of performance measures, such as work-in-process inventory and percentage of timely starts, efficient frontiers are provided based on the sequencing rules of the different cases. The researchers consider not only operating room assignments where patients undergo surgical operations, but they also investigate different processes utilizing various clinical resources. To cite an instance, they consider the consultations where healthcare resources are allocated for diagnosing the health condition of the patient and deciding on the further action that needs to be taken. The operating room is utilized if the patient's healthcare condition necessitates surgery. The effects of including emergency cases as well as the late starts for elective cases are also considered. It is found that prioritizing the consultations (i.e., patient's appointments) for the returning patients, and the consultations of patients whose surgeries involve pre-surgical and post-surgical operations over new patients can improve system efficiency. Moreover, the researchers also consider the schemes of decreasing the number of breaks in the operating room schedule to respond to the emergency admissions in a timely manner.

Some researchers also consider LOS in the hospital ward where the patient is transferred after the surgical operation in their analysis. In that regard, Harper (2002) develops a framework for modeling the admission process for the hospital. The researcher considers LOS for the patient in the hospital ward as well as the operating room availability and workforce related constraints for running the surgery schedule. In the study, the operating room utilizations and corresponding bed requirements are regarded as the performance measures. It is indicated that the longest time first LTF policy (i.e., conducting first major surgeries, then intermediate surgeries, and finally minor surgeries) leads to reduced occupancies with the constant throughput, which is largely due to reduced variation in terms of total operating time. What differentiates Harper's study (2002) from Cardoen and Demeulemeester (2008) is that, Harper incorporates LOS in the corresponding ward as an input parameter for the model, where Cardoen and Demeulemeester's (2008) work mainly focus on the number of maximum bed requirements under different pooling strategies by assuming constant LOS.

Some also combine heuristics and simulation study to incorporate the emergency surgeries in the elective ones. For this purpose, van Der Lans et al. (2006) develop an approach for allocating the elective surgery schedules such that the length of the break-in moments can evenly be distributed during the current schedule of the elective surgeries.

The break-in moments are the moments that an elective surgery schedule is expected to finish and a new emergency surgery might be started. Evenly distributing these moments throughout the day potentially decreases the waiting time for the emergency cases that can be infused to the current operating room schedule without preempting the elective surgery schedules. Based on a simulation study, they demonstrate that allocating the slack capacity over all operating rooms with the break-in-moment optimization reduces the overall waiting time. The researchers also propose different construction and improvement heuristics for evenly spacing the break-in moment as much as possible during the course of the day.

Niu et al. (2007) develop a simulation based approach for determining the performance of the operating rooms. In their model, there are basically two different time slates that the patient might undergo surgical operation based on the nature of the surgical operation. During weekdays, between 7:30 am to 15:30 pm, surgical operations for emergency patients as well as elective patients are conducted. After 15:30 pm, and during weekends, two operating rooms are allocated for conducting surgical operations only for the emergency patient. The researchers consider various elements in operating rooms such as the transfer time of patients in and out of the operating rooms, operating room cleaning time, setup-time, holding time of the patient in the pre-operation area, etc. The authors compute the total LOS with respect to different levels of clinical resources using the discrete event simulation.

### 2.5. Incorporating PACU into the Elective Patient Scheduling

In order to capture the relationship between operating rooms and other clinical units, and present approaches for improving the performance in the overall clinic settings, researchers conduct studies that combine multiple clinical entities during the planning and scheduling of operating rooms. In that regard, ICU units (Dexter et al., 2002), hospital wards (Santibanez et al, 2007), holding area belonging to the pre-operative facility for preparing the patient for surgery (Mulholland et al., 2005; Denton et al., 2006), and the PACU units (Perdomo et al., 2006; Pham and Klinkert, 2008) are usually considered with the operating rooms.

Marcon and Dexter (2006) employ a simulation study to investigate the effects of the different sequencing rules for operating rooms on PACU availability. Contrary to the Harper's (2002) findings, the authors caution that scheduling the longest cases first leads to the increases in overtime utilization of operating rooms, PACU nurses, and the risk of delayed PACU admissions, and blocking of the operating rooms. It is suggested that scheduling the cases in a mixed order (i.e., shortest case first, longest case second, second shortest case third, etc.) and <u>h</u>alf <u>i</u>ncrease in OR time and <u>h</u>alf <u>d</u>ecrease in OR time (HIHD) rule (i.e., scheduling the shortest case first, third shortest case second, the longest case in the middle, and scheduling the other cases subsequently in terms of decreasing order of scheduled operating room time later on) provide better results as compared to other approaches. One of the differences is that the approach taken by Harper (2002) considers PACU units as a downstream clinical unit as opposed to the Harper (2002) where they consider the hospital wards as the immediate next downstream clinical unit where the patients are transferred after surgical operation.

VanBerkel and Blake (2008) employ a simulation study and demonstrate that the bed availability in the corresponding hospital ward might also be a bottleneck as well as operating room availability. In that manner, the bed availability in the hospital wards and the PACU units are important factors.

Interestingly, research is also directed at the modeling the scheduling of operating rooms as job-shop scheduling while incorporating PACU as a separate work station in the system. In this vein, Pham and Klinkert (2008) implement a job-shop scheduling approach for representing surgical operations in operating rooms. The system is modeled as multi-mode blocking job scheduling, where the node represents the available resources tied to the corresponding operation. The model takes into consideration the associated set-up times (e.g., preparation of operating rooms for the next surgical operation), as well as the cleaning times. The model also features the blocked sources where the nodes to the corresponding operation remain attached and cannot be utilized elsewhere, until the subsequent operation starts. For example, even though the surgical operation is completed, the bed in operating room will still be occupied unless the patient is transferred to the corresponding PACU unit.

Guinet and Chaabane (2003) consider the material planning for the operating room planning for the medium term (e.g., around a week). It incorporates some problem characteristics such as the hospitalization date and intervention deadline. In the objective function, the patient satisfaction and the resource utilization are also incorporated. A primal-dual heuristic approach is proposed based on the assignment problem that incorporates time window additive constraints and resource capacity.

Process redesign activities might also play a role in successful operating room management. Harders et al. (2009) develop an interdisciplinary approach for the process redesign in operating rooms. The results indicate that significant reductions in the non-operative time and non-anesthesia time might be realized. Better communication among the stakeholders paves the way for improving efficiency of operating rooms.

Another study on this field is the work of Hsu et al. (2003). It models the operating room activities as two stage no-wait flow shop scheduling problem, which take the PACU capacity into account. A Tabu-search based heuristic by iteratively solving those stages near optimality is proposed. In the first stage, the objective is to determine the minimum number of nurses in a single PACU such that the completion time in the PACU is not greater than the given threshold value regarding the time of the day. In the second stage, based on the optimization of the first-stage of the problem, considering the output of the first stage as input, the objective is to minimize the total make-span in the system. The proposed approach has been implemented in healthcare facility with satisfactory results.

An interesting practice that can be associated with the recovery of the patients after the surgical operation is letting the patients regain consciousness in operating rooms rather than the PACU unit in the absence of vacant beds in that particular unit. This approach allows the flexibility but might present a hindrance for the flow of the operations in the overall system. The corresponding problem is formulated as a four-stage hybrid flow shop problem with blockings (Augusto et al., 2010). These four stages are patient transports from hospital wards to operating rooms, surgery and recovery process, cleaning task, and patient transports from operating rooms to hospital wards respectively. The MILP model and a heuristic based approach based on Lagrangean relaxation are developed. Both approaches (i.e., allowing and not allowing recovering in operating rooms) are compared. It is concluded that when the ratio of PACU beds to the number of operating rooms drops below 1.5, allowing the recovery in operating rooms might be considered as a viable option.

Cowie and Corcoran (2012) also investigate the non-clinical factors delaying discharge from the post-anesthesia care units. The study indicates that not adequate number of beds in the post-operative ward for admittance and hectic work schedule of the ward nurses are the common reasons for delaying discharges from PACU units. In a larger context, delays in discharging the patients from PACU units have a detrimental effect on the operating room efficiency. The patients due to the blocking in the PACU units cannot be transferred to those downstream units in a timely manner, and this might cause the delay of surgeries in operating rooms. Therefore, improved discharge planning, restructured staffing, and alterations in operating room scheduling are beneficial for alleviating the problem.

Improvements in the utilization of operating room rooms and downstream ICU units might be realized by forming weekly master surgical schedules. In a study conducted by van Houdenhoven et al. (2008), implementing cyclic master surgical schedules leads to the increased utilization of OR rooms, and the decrease of unused operating room capacity to 6.3% in 4 weeks. The schedule, at the same time, yields simultaneous optimal loading for ICU workload. Table 2 summarizes the research on the inclusion of PACU unit for operating room planning.

Table 2. Summary of existing research conducted in incorporating PACU for OR planning and scheduling

| Source | Method | Summary | PACU units Specifics | Strengths of the study | Simplifying assumptions and critique of the study |
|---|---|---|---|---|---|
| Ballard and Kuhl (2006) | Discrete event simulation | Analysis on determining maximum capacity for operating rooms and downstream clinic units | The patients are transferred to the PACU units after the surgical operation. The duration of stay in PACU units is deterministic. | • Both inpatients and outpatients are considered in the PACU unit,<br>• Some patients can also be rerouted to hospital wards or discharged directly. | • Only a small subset of surgical operations is considered,<br>• Assignments of different surgical operations in the same operating room during the day are not allowed, |
| Cardoen et al. (2009) | Mixed integer programming and branch and bound based heuristics | Multi objective function involving equipment usage, infection related requirements, and presurgical tests | • Maximum bed occupancy in PACU units,<br>• Two stage PACU units are considered. | • Specified algorithms are based on tight branch and bound strategies are developed,<br>• Patient priorities are also included. | • Deterministic operating times,<br>• Assignment of patients to the particular day is not considered. |
| Cardoen et al. (2010) | Exact branch and price method based on column generation is used. | Multi objective function that incorporates equipment, infection requirements, and presurgical tests | • Maximum bed allocation in PACU unit,<br>• Two stage PACU units are considered as in the case of Cardoen et al. (2009) | The equipment requirements are also incorporated in the problem solution. | • The patient assignment to different operating rooms/days is not considered,<br>• Deterministic surgical duration is considered. |
| Denton et al. (2006) | Monte Carlo simulation combined with the heuristic approach | Multi objective function that considers patient intake, patient waiting time, and overtime requirements | The specific duration at the PACU unit is stochastic. | Combination of simulation approach and meta-heuristic method | • No expansion possibilities in operating rooms,<br>• Inclusion of emergency patient is not considered. |
| Fei et al. (2010) | • Hybrid genetic algorithm<br>• Column generation based heuristic | • The first stage for daily assignment is based on the set partitioning,<br>• The hybrid genetic algorithm with Tabu search for daily sequencing is used. | • Recovery time that can be shared between PACU units and operating rooms,<br>• The predetermined stay in PACU units. | • Scheduling and sequencing problems can be performed together,<br>• Hybrid job-shop scheduling problem | The daily assignment of patient decision might affect the sequencing stage. |
| Jebali et al. (2006) | A direct MIP formulation | • First step consists of assigning surgical operations,<br>• The second step consists of sequencing the assigned operations with the objective of improving operating room utilization. | The patients are either transferred to PACU units or ICU units. | • The availability of the PACU beds as bottleneck<br>• In addition, the cleaning time of operating room is also incorporated. | • The expansion of the existing resources is not included in the problem formulation. |

Table 2. Summary of existing research conducted in incorporating PACU for OR planning and scheduling (Cont'd)

| Source | Method | Summary | PACU units Specifics | Strengths of the study | Simplifying assumptions and critique of the study |
|---|---|---|---|---|---|
| Marcon and Dexter (2006) | Discrete event simulation is used. | The impact of the sequencing rules for the surgical cases on the workload of PACU nurses is studied. | Primary emphasis is on the workload of the PACU nurses. | Different sequencing rules are investigated. The effect of sequencing rules on the workload of PACU nurses is studied. | • No PACU beds are considered in the analysis,<br>• PACU units are analyzed in isolation. |
| Niu et al. (2006) | Discrete event simulation is used. | Various performance measures such as the OR efficiency and resource utilization is calculated. | PACU and post operating hospital wards are considered in tandem. | Length of stay is also considered in the analysis. | • No emergency patient arrival is considered,<br>• The expansion opportunities are not considered. |
| Pham and Klinkert (2008) | MILP model formulation | MILP formulation to minimize disruption due to considering add-on and emergency cases | Provides the pathways for patients to be transferred to ICU units after certain surgical operation. | • Flexibility to incorporate the emergency and add-on case,<br>• Minimize the disruption for including emergency cases using multi-mode blocking job-shop framework. | The model does not suggest any improvement heuristics. |
| Cowie and Corcoran (2012) | Retrospective study | Investigate effects of the non-clinical reasons for delaying admission to PACU units | Linking discharge with the PACU units. | A large sample of the patients is studied. | • The paper does not establish the analytical approach. |

## 2.6. Dealing with Stochasticity in Operating Room Planning

The surgical operations in operating rooms are subject to significant uncertainty due to various factors such as the arrival of emergency patients, variability in duration of surgical operations, no-show cases for the surgical operations, and the variability arising from workforce related causes such as absenteeism and sick leave. In order to deal with the stochasticity introduced by the various factors, considerable amount of research has been conducted. Harper (2002) provides a generic framework for determining the hospital capacity by employing simulation analyses. Three different sources of variability (i.e., LOS in hospital ward, duration of surgery, and arrival profiles of patient groups) are considered. One distinguishing feature of this study is that the author considers the capacity of a group of hospitals. In a similar manner, Persson and Persson (2007) develop another approach in which the surgery duration and the length of stay are taken into the consideration as the factors subject to variation. They employ a discrete event simulation model to analyze the impact of resource allocation policies in the Orthopedics department on waiting time and the utilization of emergency resources.

It is a common practice that some planned slacks are incorporated in the operating room schedule at the expense of utilization levels. Such practices help not only mitigate the problems associated with the stochastic effects on the surgery duration, but also deal with the inclusion of the emergency cases (van Oostrum et al., 2008a)

In order to deal with stochastic effects, other approaches are also proposed in the literature. Most of those approaches are based on the stochastic linear programming. Probabilistic constraints are also included in the mathematical models to decrease the associated probabilities of utilizing overtime hours. To address the increasing problem sizes, column generation approaches are usually implemented to find optimal or near-optimal solutions in a reasonable amount of time (van Oostrum et al., 2008b).

Overtime is closely linked with the stochasticity of surgery duration. Marcon et al. (2003) explore this aspect by developing a procedure for operating room planning to mitigate the no-realization risks while stabilizing the utilization rate of operating rooms. The mathematical model is developed based on the multi-knapsack problem where the available time for operating rooms is regarded as the corresponding knapsack. The authors incorporate two different cost functions. The first function involves minimizing the difference between the utilization rates of operating rooms, whereas the second cost function involves minimizing the no-realization risks (i.e., the associated risk that total actual duration for the surgical operations exceeds the allocated time for operating rooms). It is concluded that the amount of overtime utilized based on these two cost functions converge to each other when the number of surgical operations increases. The authors also outline the strategies for minimizing non-realization risks and conduct a simulation analysis to evaluate the performance of operating room schedules optimized by using the mathematical modeling approaches.

The uncertainty in surgery duration is also evaluated by the Denton et al. (2010). They develop a two stage stochastic MIP model for considering the daily schedule of operating rooms. The uncertainty in surgery duration is represented by the set of scenarios. The model is applicable both in open scheduling and block scheduling practices. Several heuristics are developed for this purpose for improving the utilization of operating rooms.

Post-operative care resources are also taken into account by various authors for reducing patient waiting time for elective surgeries and determining the capacity of the surgical suite (Persson and Persson, 2009; Ballard and Kuhl, 2006). Similarly, Perdomo et al. (2006) develop an MIP based model that incorporates the recovery rooms such that the sum of the completion times for all scheduled patients is minimized. In addition to the recovery beds, the cleaning time of operating rooms between different surgical operations is also considered. Based on this model, Augusto et al. (2008) develop another version for incorporating the transporters in charge of transporting the patient between operating rooms and pre and post-operative facilities. A heuristic based approach is then developed based on Lagrangian relaxation.

In a multiple operating room environment, Denton et al. (2007) develop a block scheduling approach. Two-stage MILP stochastic programming model is employed to implement this approach. The problem domain is especially suited for the cases where the number of open operating rooms varies from day to day. Their model aims at determining the optimal number of operating rooms and assignment of surgical cases to those operating rooms. At the same time, the model aims to schedule the surgeries to the fixed number of operating rooms. Associated with this approach, exact and heuristic approaches (e.g., largest processing time first) are developed and evaluated based on the real data. The researchers conclude that heuristic approaches perform fairly well.

Lamiri et al. (2008) develop a stochastic mathematical model for considering the operating room planning under the elective and emergency patient demands. In that regard, the overtime utilization as well as the underutilization is incorporated in the mathematical model. The authors prove that the stated problem is NP-Hard. They implement Monte Carlo optimization method that combines Monte Carlo simulation and mixed integer programming. It is demonstrated that increasing the computation budget (i.e., increasing the number of individual Monte Carlo simulations) helps converging solution to the real optimum value.

In a similar token, Lamiri et al. (2007) develop a column generation approach and combine Monte Carlo simulation with the column generation approach. The researchers demonstrate that near optimal solutions can be obtained in reasonable computation times by employing column generation approach that involves consideration of subset of possible columns. This approach involves a pricing problem for determining the promising columns to be included in the solution (Lubbecke and Desrosiers 2005; Barnhart et al., 1998). In a similar sense, in order to address the uncertainty induced by the duration of surgical operations, statistical approaches are also proposed by various researchers (Dexter and Ledolter, 2005; Dexter et al., 2007; Stepaniak et al., 2010; Dexter et al., 2009).

Other than the arrival and duration uncertainties, researchers have addressed other types of uncertainty. For instance, Dexter and Ledolter (2003) study the stochastic effects of contribution margin of the surgeons in the allocation of operating room capacity under maximization of the hospital's expected financial return. On the other hand, Cardoen and Demeulemeester (2008) study the effects of the resource unavailability for operating room planning and related clinical processes by employing a simulation model. Lamiri et al. (2009) compare several optimization methods in the presence of the shared operating rooms both for emergency and elective schedule cases for minimizing expected overtime costs and patients' related costs. For this purpose, an "almost" exact method combining Monte Carlo simulation and mixed integer programming is presented, and convergence properties are investigated. It is indicated that the Monte Carlo optimization method provides the convergence to the optimal solution at an exponential rate using modest number of samples. Among the heuristic methods, Tabu search provides the best results. An interesting observation is that, when variability with respect to arrival rate of emergency patients decreases, solution quality of the heuristic approaches deteriorates, which indicates that in a sense, it is easier to solve the stochastic problems as compared to deterministic counterparts. In Table 3, various approaches for the incorporating stochasticity in operating room and downstream clinic units are summarized.

Table 3. Summary of existing research incorporating stochasticity in operating room environment

| Source | Method used | Summary | How the stochasticity is incorporated | Strength | Simplifying assumptions and critique of the study |
|---|---|---|---|---|---|
| Ballard and Kuhl (2006) | Discrete event simulation | Analysis on the determining maximum capacity for operating rooms and downstream clinic units | Duration for the surgical operation is stochastic and handled with the simulation study. | Simulation study might handle different surgical durations. | • A comprehensive model for handling stochasticity in the surgical duration is missing,<br>• Overtime utilization of operating rooms is not included in the analysis. |
| Belien and Demeuleemeester (2009) | Mixed integer programming, MIP based heuristics, and simulated annealing | Various algorithms for building master cyclic schedule are implemented. | • Number of patients who undergo surgical operations in each time block is stochastic.<br>• Duration of stay for patient in downstream unit is stochastic. | Large set of the problem instances with the primary emphasis on the bed occupancy in the hospital wards | • Only surgery demand and OR capacity constraints are considered,<br>• Model does not consider the workforce requirements in the OR room and downstream units. |
| Bhattacharyya et al. (2010) | A retrospective study with the corresponding statistical analysis | Effect of the inclusion of the dedicated OR room for the emergency cases with corresponding performance measures | Stochasticity for duration of surgical operation is implicitly considered. | A fairly long observation period is adopted (1 year) and comparison is provided between before and after the implementation. | • Only the results of empirical observation are reported,<br>• Not supported by the mathematical models |
| Bowers and Mould (2004) | A discrete event simulation study | Effect of the allocating separate unit for emergency cases is considered. | • Stochasticity for duration of surgical operation,<br>• Demand for emergency trauma patient is stochastic | Performing surgical operation for emergency and elective patients is permitted in the same operating room. | The model does not consider PACU units and analysis is conducted only for operating rooms. |
| Cardoen and Demeulemeester (2008) | A discrete event simulation study | The effect of various scheduling rules on the current workforce level is investigated. | Duration of surgical operation and non-surgical clinic appointments are stochastic. | Effect of non-surgical clinic appointment on the existing workload of the surgical staff members is investigated. | The model does not consider the expansion possibilities of the current resource level. |
| Min and Yih (2010) | A sample average approximation method for two-stage stochastic model | Generate optimal schedule of elective surgery with uncertain surgery durations. | • Stochasticity in terms of surgical duration,<br>• The length of stay in the downstream SICU unit is stochastic. | Two stage stochastic model with the recourse variables for delaying surgery of elective patient | • Expansion of workforce levels is not considered. |

Table 3. Summary of existing research incorporating stochasticity in operating room environment (Cont'd)

| Source | Method used | Summary | How the stochasticity is incorporated | Strength | Simplifying assumptions and critique of the study |
|---|---|---|---|---|---|
| Dexter (2000) | A mixed integer programming model combined with retrospective study | Effect of moving last surgical case of elective schedule to new empty operating room on overtime utilization is studied. | Duration for surgical operation is stochastic. | Model takes unconventional approach for improving utilization by moving start time of surgical operation. | • Downstream clinical units is not considered,<br>• Based on single quantitative measure<br>• Assumption on the availability of an operating room for moving surgical operation might not be valid. |
| Hans et al. (2008) | Local and constructive heuristics based on the mean and variance of duration of surgery | Heuristic methods are used to minimize slack and overtime and number of cancelled surgical operations. | Duration for surgical operations is considered to be stochastic. | • Constructive and improvement heuristics for clustering similar-variance surgical operations in same operating room | • Availabilities of workforce and surgical teams are not considered,<br>• Additionally, expanding capacity of downstream PACU units is not permitted. |
| Lamiri and Xie (2007) | Sample average approximation and metaheuristic approaches | Comparison of several different methods for elective surgery planning based on the reduction of the overtime and undertime | • Total demand for the surgical operations is considered to be stochastic variable,<br>• Time for arrival of emergency patient is considered to be a stochastic variable. | • Sharing resources between elective and emergency surgery,<br>• Costs of scheduling the elective patient to the next planning period are considered. | • Current capacity is assumed to be constant.<br>• Duration for surgical operation is assumed to be constant |
| Lamiri and Xie (2008) | Monte Carlo simulation | An objective function that consists of the overtime costs and moving the patient for the next scheduling period is formed | • Demand for the surgical operations is considered to be stochastic. | • Both overtime and undertime utilization are taken into account, | • Approach does not take work force requirements into account.<br>• Downstream clinical units are not considered. |
| Niu et al. (2007) | Discrete event simulation | OR efficiency and resource utilization are calculated based on the scenarios and tested using simulation study. | • Duration for surgical operation and time spent in PACU units are stochastic. | • Preparation time of patients for the surgical operation | Overtime considerations for operating rooms and downstream clinic units are not taken into consideration. |
| Persson and Persson (2009) | Discrete event simulation combined with mathematical programming model | | Duration for surgical operation is stochastic. | • Expanding current capacity by adding beds<br>• Patients are prioritized based on waiting costs. | Inclusion of the emergency surgery in current elective schedule is not considered. |

## 2.7. Research Gaps and Contribution of Dissertation Research

Examining the literature, we conclude that significant research gap exists in terms of effectively establishing models for scheduling and rescheduling the elective patients. First of all, there are a few papers that examine the rescheduling of elective patients, but a comprehensive view on the scheduling and rescheduling of elective patients in the same problem setting is missing. Moreover, although the approaches developed in the literature identify the downstream units such as PACU units, the possibility of resource expansion for PACU units in terms of adding bed/equipment to facilitate flow of patients is not explored. Additionally, although there are models that deal with stochastic duration times and arrival of the emergency patients, those models do not establish the link between operating rooms and downstream clinic units. Although some models in the literature are used for building master surgical schedules, these models rely on additional constraints for limiting number of changeovers in operating room settings.

To bridge the research gap, we propose a novel two stage framework for scheduling and rescheduling elective patients. Some unique contribution of the proposed models can be described below,

- In order to reduce the number of the changeovers for facilitating the master surgical schedule on the weekly basis where certain surgical operations can be grouped in particular day/time of the week, our scheduling model can implicitly handle that requirement by limiting availability of the desired surgical group to consecutive time periods of the desired day/time so that the block schedules can be formed accordingly. Since the proposed model provides this consideration implicitly rather than presenting explicit constraints to enforce it, resulting problem is smaller in size with respect to number of constraints.

- The proposed rescheduling model aims to minimize the disruption in the existing elective patient schedule. Although some models presented in the literature also aim to minimize the disruption of the elective surgery schedule, they do not explicitly consider the possibility of expanding the current capacity to minimize the amount of disruption. The proposed approach might be used for providing a comprehensive view on the tradeoff-minimizing the disruption to the existing schedule and expanding the capacity in monetary figures.
- Unlike the other approaches which prioritize patients explicitly by grouping them, the proposed models can implicitly take care of issue of patient prioritization. This provides additional flexibility on assessing the consequences of preponing or postponing surgical operations at individual patient level. Some patients might be more reluctant for their surgical operations being postponed or preponed. Therefore they might be assigned higher cost figures as compared to others.

Given the set of patients, and availability of resources, the developed framework allows scheduling elective patients and at the same time addresses forming the new elective schedule after the inclusion of the emergency patients. This approach is important in the sense that usually the existing approaches in the literature focus these two aspects separately. Our approach provides a comprehensive view on looking both scheduling and rescheduling aspects.

# 3. MODELING FOR DAY-TO-DAY SCHEDULING AND SEQUENCING OF ELECTIVE PATIENTS

As described in the first Chapter, our approach involves a two-stage approach. In the first stage, we develop a stochastic mathematical programming model for day-to-day scheduling and sequencing elective patients with regard to various constraints such as the availability of the surgical teams, operating room, and the bed/equipment in PACU units, where the patients recover from the effects of the anesthesia. PACU is located downstream of operating rooms in clinical settings. The elective patient list with the corresponding types of surgery is provided as the input for the mathematical model. Based on this, the model is used for assigning elective patients to the particular day and sequencing those patients on that particular day. Figure 1 presents a typical flow of patients on hospital floor. The constraints included in the first stage are the availabilities of operating rooms, surgical teams (i.e., surgeons, scrubbers, technicians, and anesthetists), PACU beds, and over-time constraints for operating rooms. Due to physical limitations and various other constraints, the number of beds/equipment that can be added to the existing capacity for PACU units is limited, and this perspective is incorporated in the model structure as well.

One critical factor for the first stage of the problem is the uncertainty related to the duration of surgical procedures. Our approach employs scenario generation to determine the joint probabilities that are used as model input. Based on the corresponding joint distributions and values of the decision variables associated with the outcome of the model, the time slots that are occupied both in operating rooms and PACU units can be identified. Based on this information, the corresponding make-span of daily operations in the operating room, as well as the need for various types of surgical teams and the number of occupied beds in PACU units can be determined. Finding the make-span of surgical operations in operating rooms helps determine the need for overtime requirements for operating rooms which is reflected in the objective function as a cost item as well.

In that regard, the corresponding costs of hiring additional surgical teams as well as the cost of adding additional bed/equipment in PACU are also reflected in the model by the corresponding coefficients in the objective function. As for the constraints, one of the main considerations is that simultaneously occupied beds in PACU cannot exceed an upper limit, and the number of simultaneous ongoing surgical operations cannot be more than the number of operating rooms that are available. In short, the constraints are considered with respect to the following resources:

- Availability of the PACU beds, which include the regular beds, and additional beds placed in terms of expanding the current capacity,
- Availability of the number and type of the surgical teams, which include the current number of surgical teams and additional surgical teams that can be hired on-need basis,
- Availability of operating rooms in which the number of ongoing operations cannot be more than the number of available operating rooms.

Based on the input of the list of patients for elective surgeries and corresponding surgery types, the objective is to,

- Assign the elective surgeries to a particular day,
- Provide the sequencing of surgical operations in that particular day

The mathematical programming model developed for this purpose addresses these two problems simultaneously. In the next section, we outline the mathematical programming model.

While surgical team member preferences and prioritization of patients can also be incorporated in the proposed scheduling model presented in this Chapter similar to the case of scheduling and sequencing rules implemented in a local hospital, differences arise in terms of the decision making process. In the local hospital, two-stage decision making process is conducted, where the scheduling and day-to-day sequencing decision making processes are carried out successively. In our approach, scheduling and sequencing decisions are carried in as single step.. Rather than assigning separate time blocks in terms of the master surgical schedule, by limiting the availability of the surgical teams to the specific time slots, we can explicitly identify time blocks in our model where specific surgeons/surgical teams are available. Additionally, equipment requirements might also be taken into consideration implicitly, by limiting the number of one type of surgical operation currently ongoing by limiting the number of surgical teams available for that particular operation. In short, the surgeon preferences and the master surgical schedule structure are taken into consideration. Our model can also take care of the patient priority issue. In the local hospital, the inpatients are operated later in the day, and the outpatients are operated early in day. The children and the patients with specific travel requirements are also given higher priorities. By introducing constraints in the model that prohibits some patients to be operated on certain specific time periods in the model, the patient priority issue can be taken into consideration.

### 3.1. Mathematical Programming Model

The mathematical programming model employs two-stage stochastic linear programming model with recourse variables. Table 4 presents the notation for index, set, and decision variables.

Table 4. Notation for index, sets, and decision variables

| | |
|---|---|
| **Indices** | |
| $d$ : | Day index; $d \in \{1, \dots, D\}$ |
| $i$: | Elective patient index; $i \in \{1, \dots, I\}$. |
| $j$: | Surgical operation type; $j \in \{1, \dots, J\}$ |
| $t$: | Time period index, $\mathrm{t} \in \mathrm{T}+1$ |
| $t'$: | Auxiliary time index, $\mathrm{t'} \in \mathrm{T}$ |
| $\omega$: | Scenario index $\omega \in \{1, \dots, \Psi\}$ |
| **Sets** | |
| $T_d^A$ : | {Set for overtime hours at day d} |
| $T_d^B$ : | {Set for regular working hours at day d} |
| **Decision variables** | |
| $O_{d\omega}^{OR}$ : | Amount of overtime utilized for the operating rooms at day $d$ under scenario ω |
| $O^{PACU}$ : | Amount of additional beds/equipment placed in PACU |
| $s_{it\omega}$ : | $\begin{cases} 1 \text{ if elective patient i occupies a bed in PACU unit under scenario } \omega \\ 0 \text{ otherwise} \end{cases}$ |
| $x_{it}$ : | $\begin{cases} 1 \text{ if surgery is scheduled to start at period t for elective patient i} \\ 0 \text{ otherwise} \end{cases}$ |
| $y_{it\omega}$ : | $\begin{cases} 1 \text{ if elective patient i has surgery at time period t under scenario } \omega \\ 0 \text{ otherwise} \end{cases}$ |
| $z_{jt\omega}$ : | Number of additional teams hired for surgical operations under scenario ω |

The model features various constraints such as the availability of the surgical teams, operating room, and PACU bed/equipment. The objective function consists of the following items:

- Cost of hiring additional surgical teams (with respect to number and type),
- Cost of overtime utilization of operating rooms,

- Cost of adding additional beds for PACU unit.

Elective patient list with the corresponding surgery types is provided as the input for the mathematical programming model. Based on this, the model is developed for scheduling and assigning the elective patients for operating rooms and sequencing elective patients in a particular day.

Based on this approach, a stochastic mathematical programming model with the corresponding parameters is developed and presented in Table 5.

Table 5. Notation for model parameters

| **Parameters** | |
|---|---|
| $B^{PACU}$: | Current capacity of PACU unit for the scheduling period (in terms of beds/equipment) |
| $B_d^{OR}$ : | Current capacity of operating rooms at day *d* (in terms of hours) |
| $C^{PACU}$ : | Unit expansion cost of PACU unit during planning period ($/bed & equipment) |
| $C_j^{ST}$ : | Hourly cost of hiring additional surgical team capable of performing surgical operation of type *j* |
| $C^{OR}$ : | Overtime utilization cost of one unit of operating room for the planning period ($/hour) |
| $m_{ij}$: | $\begin{cases} 1 \text{ if elective patient i required a surgery of type j} \\ 0 \text{ otherwise} \end{cases}$ |
| $N$: | Number of operating rooms in the system |
| $K_{j\omega}$ : | Operation time for surgery type *j* under scenario $\omega$ (stochastic variable) |
| $s_j$ : | Length of stay at PACU unit for surgery type *j* (hours) |
| $T$: | Number of time periods in the planning period |
| $U^{PACU}$ : | Upper limit on the over-utilization of PACU unit (bed/equipment) |
| $U^{OR}$ : | Upper limit on the over-time utilization cost of operating rooms (hours) |
| $\mu_{jt}$ : | Number of surgical teams available for performing surgery type *j* at time period *t* |
| $p(\omega)$: | Probability of scenario $\omega$ with respect to probability space $(\Omega, P)$ |

The stochastic mathematical programming model consists of two stages. The first stage involves determining the sequence and scheduling of elective patients. After the first stage, decisions are made, denoted by the corresponding $x_{it}$ variables, which indicate whether the surgical operation for patient *i* is scheduled to start at time period *t*. The corresponding pairs for $s_{it\omega}$ and $y_{it\omega}$ variables are represented and determined with respect to the corresponding outcome based on ($x$, $\omega$) pairs. Based on this initiation, $s_{it\omega}$ and $y_{it\omega}$ variables can be classified as the second stage decision variables, which are dependent on the corresponding $x_{it}$ variables based on scenario $\omega$. Note that all second stage decision variables have an effect on the corresponding objective function value whereas first stage variables, (i.e., $x_{it}$) have an indirect effect on the objective function value through determining the second stage variables. Note that time period $T+1$ indicates the outside planning period in which the patients are deferred to the next planning period. This means that not all the patients have to be operated in a particular planning period and some patients can be deferred to the next planning period indicated by $T+1$. This brings additional flexibility to the model based on the fact that in a high patient load environment, operating all the patients in the same planning period might be difficult, if not entirely impossible, due to the scarcity of the available resources such as operating room time, surgical teams, and PACU beds. By allowing patients to be operated in the next planning period, the model provides a more flexible approach. Moreover, in some cases, rather than over-utilizing or expanding the current bed/equipment capacity in PACU units, it might be preferable to defer some of the surgeries to the next planning period as well by considering the cost factor. Deferring patients to next planning period might be also realized based on prioritization of patients. Based on prioritization schemes, different cost figures might be assigned for deferring specific patients to the next planning period. In the following, we describe the mathematical model.

$$min\ E_{\omega}\left[Q(x,\omega)\right] \tag{1}$$

s.t.

$$\sum_{t=1}^{T+1} x_{it} = 1 \qquad \text{for } i = 1,\ldots,I \tag{2}$$

$$\sum_{i=1}^{I} x_{it} \le N \qquad \text{for } t = 1,\ldots,T \tag{3}$$

$$\sum_{i=1}^{I} y_{it\omega} \le N \qquad \text{for } t = 1,\ldots,T \tag{4}$$

$$\sum_{i=1}^{I} \sum_{t \in T_d{}^A} \sum_{j=1}^{J} y_{it\omega} = O_{d\omega}^{OR} \qquad \forall d \tag{5}$$

$$s_{it'\omega} \ge m_{ij} x_{it} \qquad \forall i,\ j,\ t = 1\ldots T, t' = t + K_{j\omega} - 1, \ldots, t + K_{j\omega} + S_j - 2 \tag{6}$$

$$y_{it'\varpi} \ge m_{ij} x_{it} \qquad \forall i,\ j, t = 1\ldots T, t' = t-1, \ldots, t + K_{j\omega} - 2 \tag{7}$$

$$\sum_{i=1}^{I} m_{ij} y_{it\omega} \le z_{jt\omega} + \mu_{jt} \qquad \forall j, \omega\ t = 1\ldots T \tag{8}$$

$$\sum_{i=1}^{I} s_{it\omega} \le B^{\text{PACU}} + O^{\text{PACU}} \qquad \text{for } t = 1,\ldots,T \tag{9}$$

$$O^{PACU} \le U^{PACU} \tag{10}$$

$$O_{d\omega}^{OR} \le U^{OR}, \qquad \forall \text{d} \tag{11}$$

$s_{it\omega}$, , $x_{it}$, , $y_{it\omega}$ are binary variables

$$Q(x,\omega) = C^{PACU} O^{PACU} + \sum_{\omega=1}^{\Psi} \sum_{d=1}^{D} C^{OR} O_{d\omega}^{OR} + \sum_{\omega=1}^{\Psi} \sum_{t=1}^{T} \sum_{j=1}^{J} C_j^{ST} z_{jt\omega} \tag{12}$$

$$E_\omega\left[Q(x,\omega)\right] = \sum_{w \in \Omega} p(\omega) Q(x,\omega) \tag{13}$$

Eq. 1 is the expectation of the overall objective function, represented by $Q(x,\omega)$, which is only dependent on the second stage variables. Eq. 2 ensures that the elective patients will be scheduled during the planning period. Eq. 3 ensures that the number of simultaneous starts for the elective patients cannot be more than the number of operating rooms. Eq. 4

ensures that the number of simultaneous ongoing operations cannot be more than the number of operating rooms. Eq. 5 is related with the overtime utilization of operating rooms. Eq. 6 ensures that the PACU beds are occupied for a certain time period, after the surgical operations are performed. Eq. 7 provides the link between the start and continuation of the surgical operation. Eq. 8 stipulates that the existing number of surgical teams (readily available surgical teams and the additional number of hired surgical teams) will be equal or more than the number of ongoing surgical operations of type *j*. Eq. 9 ensures that the total number of available PACU beds can satisfy the patient demand after the surgical operations are concluded. Eq. 10 determines the upper limits for the capacity expansion of PACU units. Eq. 11 satisfies overtime utilization of operating rooms. Eq. 12 is the overall cost function with respect to $O^{PACU}$, ${O_{d\omega}}^{OR}$, and $z_{jt\omega}$ variables. Those variables represent additional cost for expanding the current capacity of the available resources, namely expanding the current capacity of PACU units by adding additional bed/equipment, overtime utilization of operating rooms, and hiring additional surgical team(s). Eq. 13 is used for calculating the expected objective function value with respect to scenarios with the probabilities $\omega$.

### 3.2. Solution Approaches

A commercial optimization solver CPLEX in GAMS (version 23.2.1) is used for obtaining the results associated with the stochastic mathematical programming model. The first step in the GAMS based approach is defining the parameters associated with the scenarios based on the probabilities and surgical durations. After those parameters are defined, a loop is constructed to calculate the joint probabilities associated with the surgical durations. The joint probabilities of the scenarios are constructed to calculate the corresponding $p(\omega)$ values in Eq. 13. These values are used for calculating the second stage costs.

By using the expected values provided in Eq. 13, basically the problem is converted to a deterministic mathematical programming model. The parameter $p(\omega)$ denotes the joint probability for each scenario $\omega$ and corresponding $p(\omega)$ values are calculated based on the independent realization of each possibility associated with the surgical duration. By using the associated $p(\omega)$ values, one can calculate the expected value for the second stage costs associated with the model.

Figure 3 depicts steps that are taken in the GAMS implementation for converting the stochastic modeling problem to a deterministic one and solving it by the commercial solver.

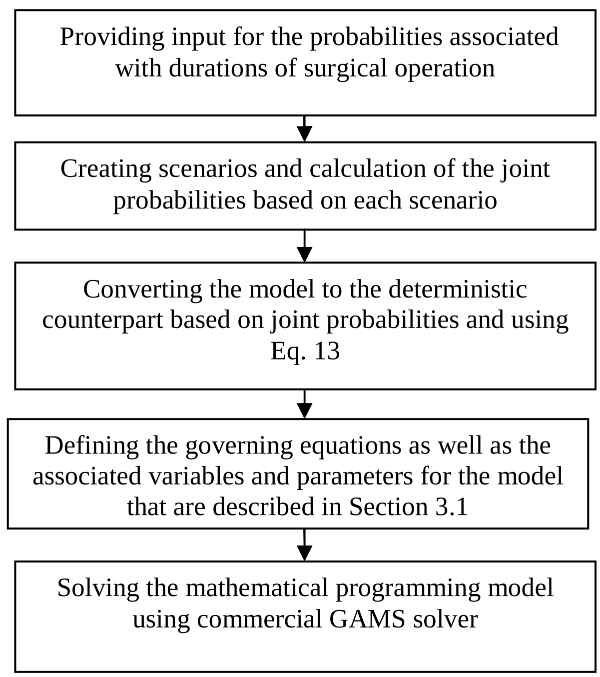


Figure 3. Flowchart for solving the stochastic mathematical programming model in GAMS

On the other hand, another approach that might be taken for solving the stochastic mathematical programming model is Monte Carlo simulation. Rather than calculating the joint probabilities based on each scenario, an approximation method is used for Monte Carlo simulation. There might be cases where an expectation function similar to the one provided in Eq. 13 cannot be computed exactly, and it might be approximated using Monte Carlo sampling methods. In this case, random samples belonging to solution space are drawn, and the second stage costs can be calculated using the following formula,

$$\hat{q}_N(x) = \frac{1}{N}\sum_{j=1}^{N} Q(x, \omega^j) \tag{14}$$

where $\hat{q}_N(x)$ is the approximation of second stage costs, $N$ is the total number of random samples that are drawn, and $\omega^j$ are the realizations of existing scenarios based on sampling. The Monte Carlo simulation approach provides an approximation, and the convergence properties of the Monte Carlo approach to actual values are dependent on the associated cost function and sample size.

However, this research does not employ the Monte Carlo simulation approach because it may be computationally intensive. Special tricks are often needed to improve the efficiency, but this makes the solution approach less straightforward. As such, we adopt genetic algorithm as a solution approach to compare with the solutions obtained from the commercial solver, GAMS.

### 3.3. Genetic Algorithm

In this section, we will provide information on the genetic algorithm approach in detail. Genetic algorithm is an evolutionary strategy consisting of different structures to represent the solution. It is basically considered to be a search heuristic that can mimic the process of evolution. With other heuristic algorithms such as the Tabu search, simulated annealing, etc., they are collectively named as meta-heuristic algorithms.

The work on genetic algorithm which is computer simulation of evolution started around 1950's by Barricelli (Barricelli, 1957). Through the years, researchers lay different components of evolutionary strategy. The methods of evolutionary strategy are in depth described by Fraser and Burnell (1970) and Crosby (1973). It came into the spotlight among research community by work of John Holland in the early and mid 1970's (Holland, 1975). Research in genetic algorithm remained largely theoretical, until the end of 1970's, however in 1980's commercial products based on genetic algorithm began to appear.

The first step of genetic algorithm is to create a random population of the individuals. These individuals (or chromosomes) are randomly generated and a collection of those chromosomes are included in the pool. This pool represents initial generation. To develop better solutions, the chromosomes are evolved throughout the generation. The evolution is conducted by various selection and recombination operators that will be discussed in the following sections.

Depending on the fitness value of the chromosomes, the chances of the chromosomes for passing its genotype to next generations are evaluated. This is done either,

- Stochastically (chromosomes have a better chance of selection based on their fitness function value);
- Deterministically (a chromosome with a better fitness function value will be selected for recombination operator for the next generation instead of chromosome that has lower fitness function value).

After selection based on the fitness function value, the individual genome is modified using different operators. The most commonly used ones are,

- Recombination (e.g., crossover),
- Mutation operators

A schematic representation of crossover operator is given in Figure 4 (Hackett, 1995).

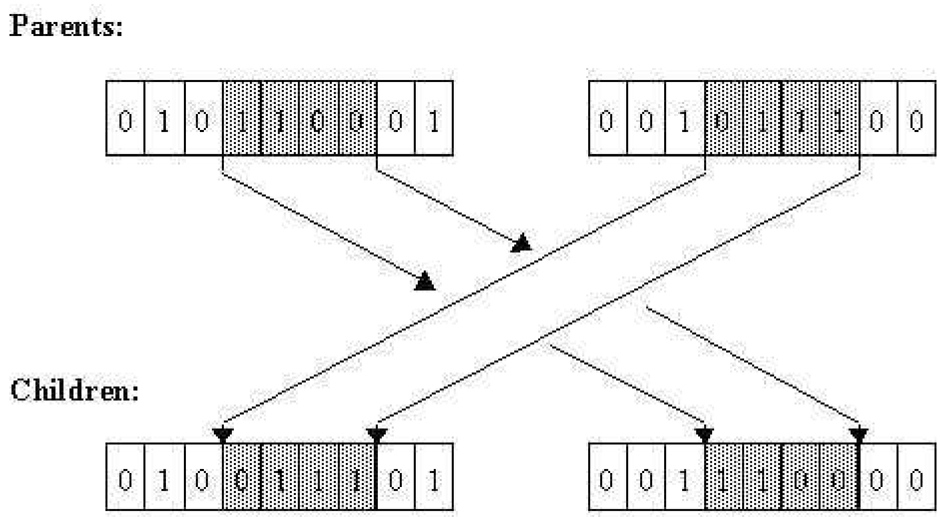


Figure 4. Example of recombination operator-traditional two-point crossover

Based on recombination and mutation operators, offsprings are created. These offsprings are evaluated based on the fitness function values to create a new population. In order to create a new population, offsprings as well as parents are evaluated and the chromosomes pertaining to next generations are selected. This might be either done by stochastically or deterministically similar to the schemes for selecting the list of the chromosomes for producing children. Another approach for selecting parents for producing offsprings and selecting chromosomes for next generation is combining various stochastic and deterministic approaches. In that regard, combining elitist and stochastic selection might be a viable approach. In this approach, a certain number of chromosomes are selected using the elitist (i.e., deterministic) selection and the remaining chromosomes are selected using some stochastic selection principles such as the roulette wheel selection. This decision depends on many factors such as problem domain, representation of solution, scaling of the fitness function, etc.

After a new population is selected, the previous steps described above are repeated. Commonly, the algorithm stops when either the number of generations exceeds the limit or a satisfactory fitness level has been reached. The satisfactory fitness function value might be determined based on the bounds established for problem domain.

A schematic representation of genetic algorithm steps is presented in Figure 5 as follows;

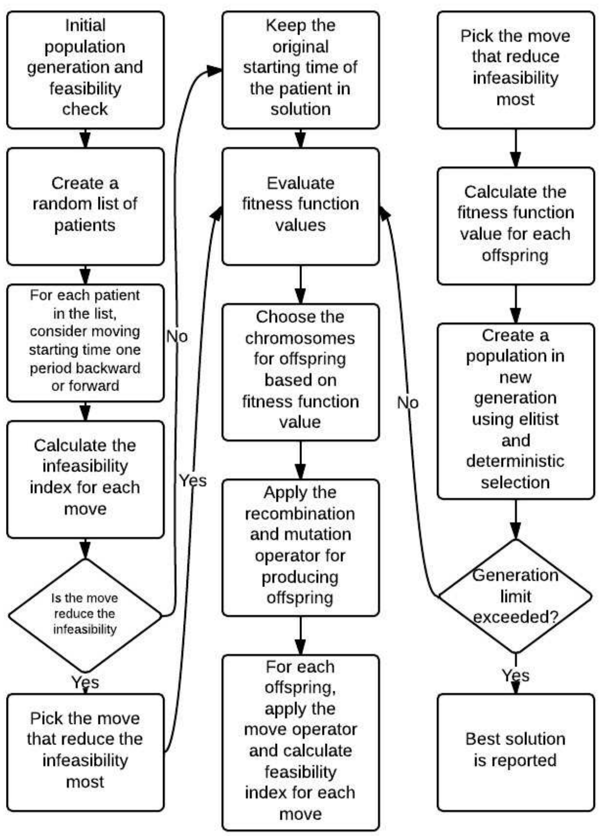


Figure 5. The genetic algorithm flowchart

As previously discussed, a typical genetic algorithm requires genetic representation of the solution. Usually, a standard representation of each candidate solution is the array of bits. In our research, we take another approach and present each unique solution by combination of three different parts. The chromosome structure will be described in the following section.

After defining the fitness function and the chromosome structure, the genetic algorithm proceeds through repetitive application of the three different operators, which are mainly the mutation, crossover, and selection for next generation operators. Those operators will be discussed in more detail in the following.

### 3.3.1. Representation of the solution

The chromosome representation for any given solution candidate consists of three parts. We demonstrate the chromosome representation using three different structures. These chromosome structures can be listed as follows,

- Starting time of the surgical operations,
- Sequence of patients that matches with the starting times of the surgical operations
- The number of patients undergoing surgery in each operating room, and the number of patients rolled to the next planning period.

These three parts of chromosomes are combined to represent the solution. All the bits in the chromosome consist of integer numbers. The first two parts are coupled for presenting the overall sequence and starting times for the operations. The third part actually complements the solution. It specifies the number of patients assigned to each operating room. It also specifies the number of patients rolled to the next planning period. An example of the chromosome representation is given in Figure 6.

$$\underbrace{3\ 5\ 9\ 4\ 43\ 13...\quad 37\ 5\ 34}_{\text{Starting time for surgical operations (First part of chromosome)}}$$

$$\underbrace{4\ 2\ 6\ 7\ ....\ 20\ 13}_{\text{Sequence of the elective patients (Second part of chromosome)}}$$

$$\underbrace{11\ 12\ 9\ 10\ 1}_{\text{The number of patients assigned to each operating room and rolled to next planning period (Third part of chromosome)}}$$

Figure 6. Sample chromosome representation

The first part of the chromosome structure provides us the information on the starting times of the patients. Based on Figure 6 as an example, we can see that the first patient in the sequence is scheduled to be operated at time period 3; the second patient is scheduled to be operated at time period 5, and so on.

The second part of the chromosome is the sequence of the patients. For determining the starting time of a specific patient, the interpretation should be conducted based the first and second part of the chromosome. Based on Figure 6 again, we see that the start time of the surgical operation for the fourth patient is scheduled for the time period 3. The second patient is scheduled to start at time period 5, the sixth patient is scheduled to start at the time period 9, and the twentieth patient is scheduled to start at the time period 5, and the thirteenth patient is scheduled to start at time period 34.

Note that it is possible that same time slots might be repeated in the first part of the chromosome representation in a multi-operating room environment. To cite an instance, in the sample chromosome representation, time slot 5 is repeated twice which means that the time slot 5 is selected as the operating time for two different patients (i.e., patient number 2 and the penultimate patient), but since they are operated in two different operating rooms, their schedules do not have time conflict and both patients can have the same starting time for their surgical operation.

The third part of the chromosome governs the number of surgeries performed during a given day. For example, based on Figure 4 again, in a 4-operating room scenario, we conclude that 11 patients are scheduled to be operated in the first operating room, 12 patients are scheduled to be operated in the second operating room, 9 patients are scheduled to be operated in the third operating room, and 10 patients are scheduled to be operated in the fourth operating room, and regarding the chromosome representation, one patient is scheduled to be operated in the next planning period which means that the patient will not be operated in that particular planning period.

### 3.3.2. Initial population generation

The first step for the initial population generation involves randomly sequencing elective patients. The first part of the chromosome which represents the starting times of the patients is generated by randomly assigning numbers between 1 and the last available time slot of an operating room. Using this scheme might lead to infeasibilities. To cite an instance, if the surgical operation is scheduled for the last time slot, and the duration of surgical operation is more than one hour, then time slots that cannot be assigned to any surgical operation is occupied with that particular surgical case. The same is true if the surgical operation is scheduled to start one period before the last available time slot and the duration for surgery is more than 2 hours.

When infeasibility is detected, new random number between 1 and number of available time slots is generated and feasibility is checked. However, this initial population generation scheme might also lead to other infeasibilities due to the scarcity of the resources. To cite an instance, two different surgical operations might be scheduled in the same operating room; or due to the lack of resources, it might not be possible to perform two or more surgical operations for the same time period in different operating rooms.

Note that usually PACU units operate for longer time periods as compared to operating rooms for accommodating patients undergoing surgical procedure. This convention is valid based on the fact that PACU units are considered to be a downstream clinic unit. For example, if the time spent in a downstream clinic unit (i.e., PACU unit) is one hour and the operating rooms run 12 hours per day, the PACU units should be operated at least 13 hours to accommodate all the surgical operations performed in the operating rooms. If the feasibility check regarding the time limitations is adopted for operating rooms, provided that PACU units continue to operate for an additional hour, infeasibility in terms of the time limitations does not exist for PACU units. However, other types of infeasibilities related with PACU units might arise such that the number of patients who simultaneously occupy PACU beds are greater than the number of available bed/equipment in PACU units. However, with respect to working hour constraints, the initial population creation scheme does allow infeasibilities. The second step is creating the sequence of patients for the second part of the chromosome by using random permutation of the patients. The infeasibility check that is performed in the previous step is not performed in this step.

The third step is assigning patients to different operating rooms. In this step, a random number between 1 and the total number of patients is generated for each operating room. In addition to that, a random number generated by the same scheme is assigned for the patients who are rolled to the next planning horizon, $N+1$ bits of the chromosome. After assigning the corresponding number of patients for each operating room, the total sum is calculated. If the total sum is equal to the number of patients, this step is concluded. However, it might be the case that the sum of total number of patients scheduled for the operating rooms and rolled to the next planning horizon might not match the total number of patients in this representation scheme. If the total sum is more than the number of patients, then it is apparent that the total sum should be reduced to match the total number of patients. In this case, a random number is uniformly selected between 1 and $N+1$ where $N$ represents the total number of operating rooms. To cite an instance, 1 indicates the first operating room, and $N+1$ means the group of patients rolled to the next planning horizon. After selecting the corresponding group of patients, this number is reduced by 1. Then in the next step, this process is repeated, until the total number of patients scheduled for the operating rooms and rolled to the next planning horizon is equal to the total number of patients. If the selected group is empty (e.g., the number of patients scheduled in that particular operating room is 0), then another group is selected for reducing number of patients in that specific bin.

In a similar manner, if the total number of patients operated in different operating rooms and the number of patients rolled to the next planning horizon is less than the total number of patients, then similar steps are taken. This time, the sum is incremented by 1 by selecting the specific bin and increasing it. Again, the feasibility checks are performed to ensure that number of patients in a specific bin cannot exceed the total number of patients. If such a case is encountered, another random number is generated.

In order to decrease the infeasibility of the initial population, we selectively include the generated solution in the initial population. For this purpose, a large number of randomly generated solutions (i.e., $\psi$) are considered for the analysis. If the generated solution is feasible, it is automatically included in the initial population. If the generated solution is infeasible in terms of the availability of operating rooms, surgical teams, or PACU beds, then for each infeasible solution, the randomly generated patient list is traversed, and for each patient, the start time of the operation is considered to be moved one period back or forth in an attempt to decrease the overall infeasibility. The move that decreases the level of infeasibility most is chosen, and a new schedule is formed based on this move. If the infeasibility is not reduced by preponing or postponing starting time of surgical operation, then the current schedule is kept.

This move procedure is repeated for every patient in the sequence until all the patients in the randomly generated list are traversed. In order to promote the variability of the pool of the initial population, the patient list is generated by using random permutation of the patients. At the end, the overall infeasibility in terms of the violation of the resource constraints for each solution is calculated.

After each solution in the initial population is considered, chromosomes (i.e., solutions) are sorted according to the ascending order of their infeasibility value. The top chromosomes in terms of their infeasibility values (i.e., the least infeasible ones) are selected for the initial population. In short, we generate $\psi$ random solutions, attempt to move the starting time for each patient in the randomly created list of the patients one time period forward or backward for decreasing the overall infeasibility and then calculate the feasibility index for each move and select the move that the infeasibility is reduced most. If the infeasibility is increased due to the move for each direction, then the original schedule is kept in the population. After all patients in the list are covered, we proceed with the new randomly generated solution and place it in the list of potential candidates for the consideration for the initial population. As previously suggested, after this step, all the solutions are ranked according to the ascending order of their infeasibility value, and the top chromosomes in the list are included in the initial population.

### 3.3.3. Crossover operator

For each generation, a certain number of chromosomes/solution pairs are selected for crossover operation. From those pairs, the first part of the chromosome for the offsprings is created using the two-point traditional crossover operator for the starting time for each operating room. In a two-point crossover, two points are selected for the parents, and then all the alleles between these two points are swapped between the parent chromosomes, rendering two offspring chromosomes. An example of the two point crossover is provided below;

Parent 1: 3 5 | 4 6 | 2 1

Parent 2: 6 3 | 5 2 | 1 4

In this representation, the crossover sites are located after the second allele and the fourth allele. Performing the swap, we obtain;

Offspring 1: 3 5 5 2 2 1

Offspring 2: 6 3 4 6 2 1

Note that under this scheme, the duplication of alleles is permitted and it might be possible to have two patients to have the same starting time for their surgical operations. Also another point worth noting is that, before performing the crossover operator, combining information from the first and second parts of chromosome, the starting times of the surgical operations for patients are sorted in the ascending order of their index values. The index values are represented by the second part of the chromosome. After this step, crossover operator is performed accordingly. An example is given in Figure 7.

$$\underbrace{5\ 2\ 9\ 8\ 4\ 3\ 6}$$

Starting time of surgical operations (First part of chromosome)

$$\underbrace{2\ 7\ 3\ 4\ 6\ 1\ 5}$$

Sequence of elective patients (Second part of chromosome)

$$\underbrace{3\ 5\ 9\ 8\ 6\ 4\ 2}$$

Sorted starting time according to the index values

Figure 7. Schema for sorting the first part of chromosome according to index values

For the second part of the chromosome, the partial mapped crossover (PMX) operator is utilized. The reason for using the PMX operator is that the sequence of patients constitutes the ordered chromosomes, in which applying 2-point or any traditional crossover operator might yield to infeasible sequences of patients. The infeasibility might be due to the representation of patient in the sequence for more than once, or not being represented at all. To prevent the representation problem, various crossover operators for the ordered chromosomes are developed. The PMX operator is one of them. It offers higher performance in some problem domains (Al-Dulaimi and Ali, 2008). For this type of crossover operator, the first parent donates a group of alleles and the corresponding group from the other parent is sprinkled about that particular group in the offspring. Once that is done, the remaining alleles are copied directly from the second parent. The steps are described below (Rubicite Interactive, 2012),

1. Randomly select a group of alleles from the first parent and copy them to the child. During this process, the index of the segments is recorded.
2. In the second step, by identifying the same segment positions in the second parent, pick the same value that has not been copied to the child. For each of the values;
   a. Record the index of this value in the second parent. Locate and record the value from first parent in the same position.
   b. Find the index of the same value in the second parent.
   c. If the index value in the second parent is part of the original alleles that has been selected, then go to step *a* and use this value.
   d. If the position is not the part of the original alleles that has been selected, then insert the value in step 1 into the offspring in that position.
   e. Copy any remaining positions from second parent to the offspring.

However, this crossover operator is not applied to every parent. For each pair of parents, the random number is generated. If the generated number is smaller than the probability of the crossover (i.e., $\varsigma$), then the crossover operator for the second part of the chromosome is applied. The crossover operator is applied within the same operating room, and since the number of patients in each operating room might be different, the crossover operator involving variable length is performed. Figure 8 represents the schematics of the variable crossover operator.

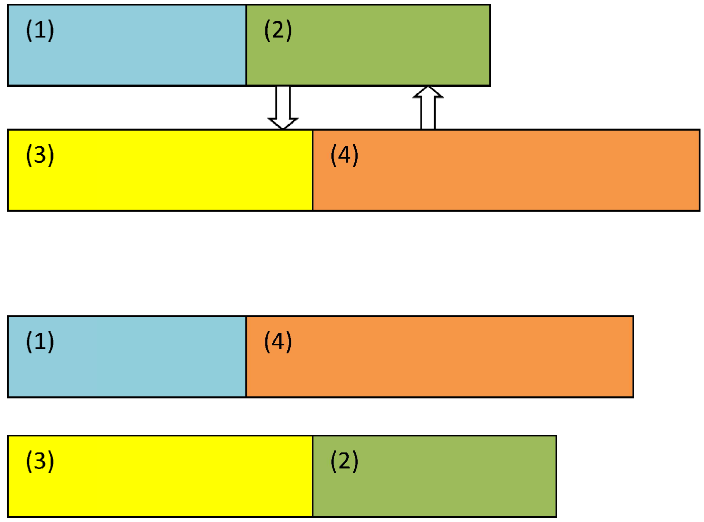


Figure 8. Schematic representation of variable crossover operator

The variable crossover operator with lower probability is adopted for preserving the order of the sequence of the patients and not disrupting the building blocks based on the sequence of the patients.

For the last part of the chromosome representation, a single point crossover is used. Note that the feasibility check is performed and steps are taken to equate the sum of the scheduled and rolled over patients to the number of total patients.

### 3.3.4. Mutation operator

The mutation operator for the chromosomes is performed on the bit by bit basis. For each bit, a random number uniformly distributed between 0 and 1 is generated. If the generated random number is below the mutation probability, then the bit is replaced with a random number between 1 and the total number of operating hours. To cite an instance, assuming that there are 5 operating days, and 12 hours available for operating room each day, a random number between 1 and 60 is generated, and the existing starting time for the particular patient is replaced with this number. For the second part of the chromosome, the exchange operator is performed. Again, a number that is uniformly distributed between 0 and 1 is generated for each bit in that particular chromosome part. If the generated number is below the mutation probability, then that particular patient is exchanged with another patient.

For the third part of the chromosome, mutation is performed by incrementing and decrementing the number of patients in a particular operating room. If the generated random number is below the mutation probability, then another random number is generated. If the number is below 0.5, then the value for that particular bit is decremented, otherwise it is incremented. After repeating this step for every bit in the chromosome structure, the total is checked to ensure that the number of patients in that particular representation is equal to the total number of patients. After the mutation operator is performed, the total infeasibility for the offspring is calculated. Similar to the case of initial population generation, for each patient in the randomly generated patient list, moving the start of the surgical operation one period backward or forward is considered. If the suggested move decreases the overall infeasibility, the move is performed. This is performed for every offspring.

### 3.3.5. Selection for the next generation

For selection for the next generation, a mixture of elitist and the roulette wheel selection is performed. A certain number of chromosomes are selected using the elitist selection with ranking of the chromosomes from the best to the worst with respect to the fitness function values. A fitness function value is calculated based on the objective function value and infeasibilities associated with that particular solution. The penalizing scheme for the infeasibilities for a particular solution is outlined in the next section. The remaining chromosomes are selected using the roulette wheel selection.

### 3.3.6. Penalizing scheme

Regarding penalizing scheme for evaluating the fitness function, we adopt the approach developed by Joines and Houck (2004) where the penalty coefficients can be calculated as follows,

$$g(\chi) = \rho^{\chi} \tag{15}$$

where $g(\chi)$ is the corresponding penalty coefficient for violating the constraint at generation $\chi$, and $\rho$ is a constant. After calculating the penalty coefficient, the total amount of infeasibilities determined by the violation of constraints depicted by Eqs.3-5 and Eqs. 8-11. The violation of constraint is calculated by taking the difference of left and right hand sides and multiplying it with a particular cost figure. This enables a wider search in scope at the initial generations with lower penalty coefficients, and towards the end of the execution of genetic algorithm with increased penalty coefficients.

### 3.3.7. Genetic algorithm parameters

The parameters of genetic algorithm are provided in Table 6. The values of the parameters are provided in Table 7. The genetic algorithm parameters are decided based on literature recommendations and pilot runs. Grefenstette (1986) shows that the bit-by-bit mutation rate around 0.01 provides better results, especially for the off-line performance. The off-line performance measures the average fitness function value of the best solution found so far throughout the generations and emphasizes improving the best solution found in every generation. De-Jong (1992) demonstrates that long term performance is improved by selecting a population size between 50 and 100. Grefenstette (1986) also indicates that the crossover rate of 0.45 provides better results in terms of the off-line performance as well. The population sizes varying between 30 and 80 are reported to provide better results in that regard. It has also been indicated that the performance of PMX improves with the increasing population size and is robust with respect to mutation rate (Buckles et al., 1990). In this study, we refer to the literature for an initial range of the parameter values and employ pilot runs for fine tuning of those parameters. The mutation probabilities and population size are selected based on the values taken from the literature. Other values are selected based on the pilot runs.

Table 6. Genetic algorithm parameters

| **Parameter** | **Notation** |
|---|---|
| Number of parent pairs selected for crossover | $\pi$ |
| Mutation probability for first part of chromosome structure | $\tau_1$ |
| Mutation probability for second part of chromosome structure | $\tau_2$ |
| Mutation probability for third part of chromosome structure | $\tau_3$ |
| Number of chromosomes/solutions selected for the next generation by elitist selection | $\upsilon$ |
| Number of chromosomes/solutions selected for the next generation by roulette wheel selection | $\varphi$ - $\upsilon$ |
| Population size | $\varphi$ |
| Number of randomly created solutions generated for the initial population | $\psi$ |
| Crossover probability for second and third part of chromosome depicting sequence of patients | $\varsigma$ |
| Limit on maximum generation number | $X$ |

Table 7. Genetic algorithm parameter values

| Parameter | Value |
|---|---|
| $\pi$ | 20 |
| $\tau_1$ | 1% |
| $\tau_2$ | 1% |
| $\tau_3$ | 1% |
| $\upsilon$ | 15 |
| $\varphi - \upsilon$ | 45 |
| $\varphi$ | 60 |
| $\psi$ | 1000 |
| $\varsigma$ | 0.05 |
| $\chi$ | 5000 |

### 3.4. Problem Parameters

The parameter selection for the problem is an important consideration. In order to find the values for the parameters used in the model, we resort to the literature and use similar values adopted in the literature previously.

Olejarz (2009) reports an average cost of $4000/day for an additional bed/equipment for PACU units. Park and Dickerson (2009) report the operating room cost is around 15-25 USD/minute during the overtime hours. According to Salary.com (2012), the hourly median rate for surgical team consisting of a surgeon, an anesthesiologist**,** and a registered nurse is $4,156/hour. These values are compiled separately from the salary.com site and calculated based on the individual values for the salaries of the surgeon, anesthetist, and the nurse assistant. Vogel et al. (2010) indicate that cost of deferring the surgery of a patient is $3,798/day. We assume a fixed value of $18,990 for deferring the surgery of a patient to next planning period. The corresponding values for input parameters that are included in the model are provided in Table 8.

Table 8. General problem parameters

| Description | Cost coefficients |
|---|---|
| Exceeding the regular capacity but not overtime capacity of the PACU unit | $4,000/bed-day (Olejarz, 2009) |
| Additional cost of hiring a surgical team | $4,156/hour |
| Cost of overtime for an operating room | $1500/hr (Park and Dickerson, 2009) |
| Rolling patient to next planning period | $18,990/patient (Vogel et al., 2010) |
| Length of planning period | 5 (days) |

Regarding the percentage of the patients with respect to different surgical types and duration of surgery, corresponding data is collected from a local medical center. Table 9 provides the characteristics for duration of the surgical operations. Based on the information obtained from the local medical center, the corresponding coefficients of variation are calculated for each surgical specialty. The surgical specialties having higher coefficient of variation are incorporated for forming the scenarios. For this purpose, the surgical specialties having coefficient of variation over 0.52 are selected. For determining the corresponding probabilities, the histograms of surgery duration pertaining to particular surgical specialty are used. For this purpose, corresponding duration for the surgical operations is grouped into the bins. Based on the relative frequencies of the corresponding bins, the corresponding probabilities are determined. Table 9 presents corresponding probabilities and coefficient of variation for each surgical specialty.

Coefficient of variation can be calculated with the following formula,

$$c_v = \frac{\sigma}{\mu} \tag{16}$$

where $c_v$ is the coefficient of variation, $\sigma$ is the sample standard deviation, and $\mu$ is the sample mean.

Table 9. Characteristics of the duration of surgical operation

| **Surgical Operation** | **Duration of operation (min)** | | | **Corresponding probability** | | | **Coefficient of variation** |
|---|---|---|---|---|---|---|---|
| Cardio-Vascular (CV) | 240 | | | 1 | | | 0.433205 |
| Ear-Nose-Throat (ENT) | 60 | 120 | | 0.6 | 0.4 | | 0.801798 |
| General Surgery | 60 | 120 | 180 | 0.2 | 0.5 | 0.3 | 0.535059 |
| Hand | 60 | | | 1 | | | 0.451006 |
| Neurology | 240 | | | 1 | | | 0.481006 |
| OB/GYN | 120 | | | 1 | | | 0.508653 |
| Ophthalmology | 60 | 120 | | 0.8 | 0.2 | | 0.523463 |
| Orthopedics | 120 | | | 1 | | | 0.483529 |
| Podiatry | 120 | | | 1 | | | 0.433205 |
| Urology | 60 | 120 | | 0.6 | 0.4 | | 0.688692 |

Based on the coefficient of variation analysis, we determine that the durations for Ophthalmology, Urology, ENT, and General Surgery should be represented by stochastic variables. In order to determine the corresponding probabilities, in the first step, the histograms based on corresponding surgical durations are drawn. These histograms are based on the bins having a width of 15 minutes. The histograms for surgical duration of General Surgery, Ophthalmology, ENT, and Urology are provided in Figures 9, 10, 11, and 12 respectively.

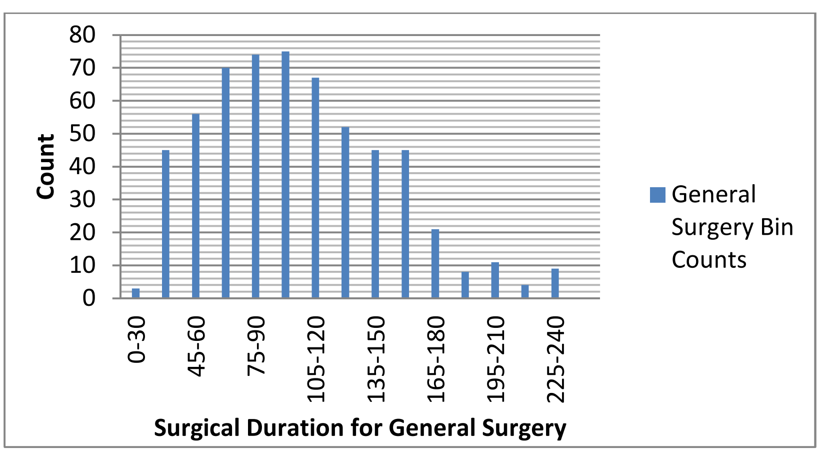


Figure 9. Histogram for surgical duration for General Surgery

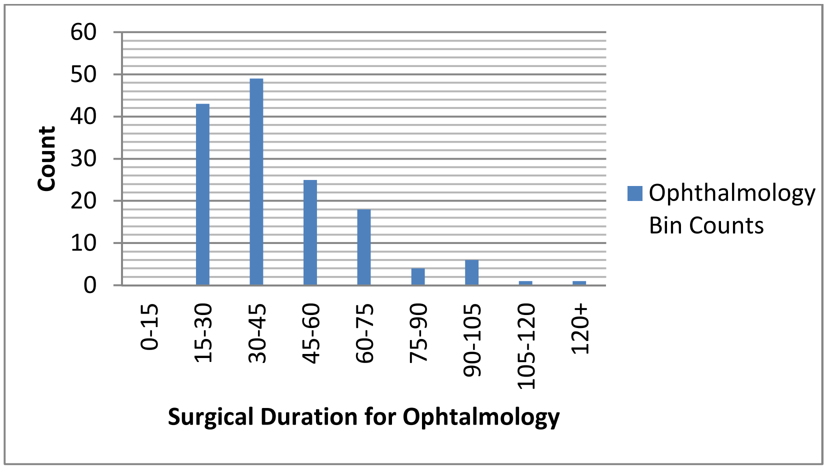


Figure 10. Histogram for surgical duration for Ophthalmology

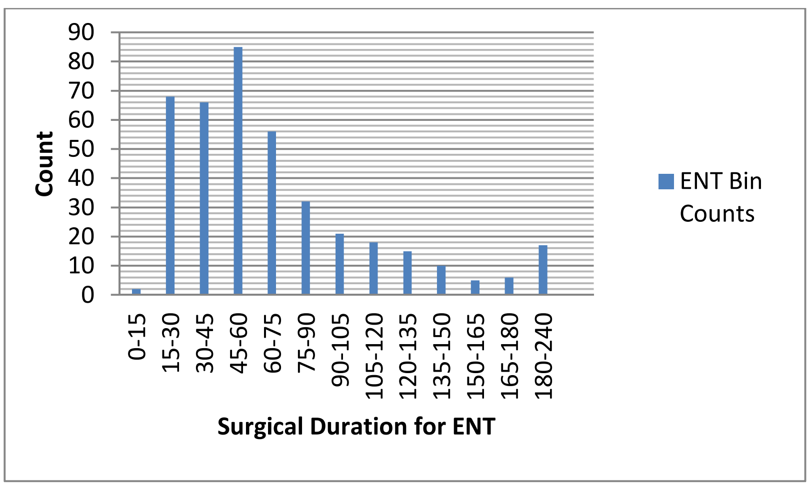


Figure 11. Histogram for surgical duration for ENT

In total, we included 24 scenarios based on the joint distribution of the duration of surgical operations. By analyzing the data collected from the local medical center for the period of July 1, 2011-June 30, 2012, the percentages and types of surgical operations are calculated. The surgery types with corresponding percentages are provided in Table 10.

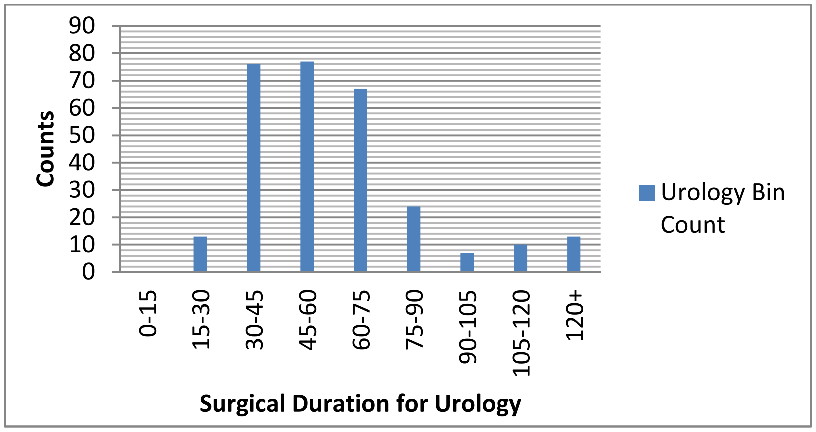


Figure 12. Histogram for surgical duration for Urology

Table 10. Corresponding percentages of patients belonging to each surgical specialty

| Surgery type | Percentage, % |
|---|---|
| Cardio-Vascular (CV) | 5 |
| Ear-Nose-Throat (ENT) | 15 |
| General Surgery | 30 |
| Hand | 5 |
| Neurology | 5 |
| Obstetrics and gynecology (OB/GYN) | 10 |
| Ophthalmology | 5 |
| Orthopedics | 15 |
| Podiatry | 5 |
| Urology | 5 |

In order to identify the effects of the number of operating room and elective patient load on the solution quality and computation times, an approach based on the experimental design has been utilized. For this purpose, 9 different scenarios are developed. Our approach involves running the genetic algorithm and GAMS based commercial solver for different sets of the input parameters of the problem, and providing the comparison based on solution quality and computation time between these two approaches.

As previously mentioned, the parameters that are being used in the experimental design are as follows,

- Elective patient load
- Number of operating rooms

Three different settings for the elective patient load (i.e., 50%, 85%, and 110%) and three for the number of operating rooms (i.e., 4, 6, and 8) are used for our analysis. Considering the complete experimental design, we have 9 scenarios in total. Based on these 9 scenarios, a comparison between genetic algorithm and the GAMS based commercial solver is performed.

### 3.4.1. Elective patient load

Elective patient load is an important parameter that determines the utilization of operating rooms and corresponding cost figures. A high elective patient load likely leads to postponing some surgical operations to the next planning period. The elective patient load for operating rooms can be expressed as,

$$L = \frac{\eta}{\kappa} \tag{17}$$

where $L$ is the elective patient load, $\eta$ is the total hours required for performing the surgeries in the elective patient list, and $\kappa$ is the number of regular hours for performing surgeries in operating rooms.

Note that total hours spent for conducting the surgical operations in operating rooms is determined based on the maximum duration time pertaining to the surgical operations. On the other hand, the calculation of the total working hours available for operating rooms excludes the overtime hours.

Based on those, the elective patient load that has three levels can be represented with respect to the generated scenarios. The elective patient loads are provided in Table 11. Regarding the PACU beds, for the 4 operating room scenarios, the current number of PACU beds is set to be 3. For the scenarios involving 6 operating rooms, 4 PACU beds are incorporated. Finally, for the scenarios with 8 operating rooms, the number of PACU beds is set to be 6.

Table 11. Elective patient loads utilized for the scenarios

| **Scenario** | **Elective patient load** |
|---|---|
| 1-3 | 50% |
| 4-6 | 85% |
| 7-9 | 110% |

### 3.4.2. Number of operating rooms

Number of operating rooms is determined based on the total operating rooms available for performing the surgical operation. Table 12 lists down the levels utilized for the scenario analysis based on the number of operating rooms. Based on these settings, final experimental design is formed. Based on information obtained from local medical center, we assume that, for operating rooms where overtime practices are allowed, the daily overtime limit is set to 2 hours for each operating room. These scenarios are provided in the Table 13.

Table 12. Number of operating rooms utilized for different scenarios

| **Scenarios** | **Number of operating rooms** |
|---|---|
| 1, 4, and 7 | 4 |
| 2, 5, and 8 | 6 |
| 3, 6, and 9 | 8 |

Table 13. Problem parameters with respect to scenarios

| Scenario | Number of operating rooms that overtime practice is employed | Number of PACU beds | Elective patient load | Number of operating rooms |
|---|---|---|---|---|
| 1 | 2 operating rooms | 3 | 50% | 4 |
| 2 | 4 operating rooms | 4 | 50% | 6 |
| 3 | 6 operating rooms | 6 | 50% | 8 |
| 4 | 2 operating rooms | 3 | 85% | 4 |
| 5 | 4 operating rooms | 4 | 85% | 6 |
| 6 | 6 operating rooms | 6 | 85% | 8 |
| 7 | 2 operating rooms | 3 | 110% | 4 |
| 8 | 4 operating rooms | 4 | 110% | 6 |
| 9 | 6 operating rooms | 6 | 110% | 8 |

### 3.5. Scenario Results

Based on the corresponding runs, the computation time and solution quality of both approaches (i.e., GAMS based approach and genetic algorithm) are reported. For a fair comparison of between the genetic algorithm and GAMS based commercial solver, equal amount of computation time is allotted for each approach and the results are compared. For this purpose, first genetic algorithm is run, and total computation time is recorded. The same amount of computation time is allocated for GAMS based commercial solver. In Table 14, a comparison between GAMS and genetic algorithm is provided.

Table 14. Comparison of genetic algorithm and GAMS based approach

| Scenario | GAMS solution ($) | Genetic algorithm solution ($) | Computation time (sec) | Relative difference (%) |
|---|---|---|---|---|
| 1 | 0 | 3823 | 3165 | N/A |
| 2 | 38476 | 40484 | 5578 | -4.96% |
| 3 | 43923 | 44817 | 7046 | -1.99% |
| 4 | 45815 | 48372 | 4562 | -5.29% |
| 5 | 59124 | 58245 | 7853 | 1.51% |
| 6 | 87839 | 91635 | 10131 | -4.14% |
| 7 | 147973 | 146785 | 6001 | 0.81% |
| 8 | N/A | 173427 | 10258 | N/A |
| 9 | N/A | 183472 | 13276 | N/A |

Examining the results, we see that usually with the increase of number of operating rooms, and elective patient load, the computation time increases considerably. To cite an instance for Scenario 2, the computation time is 5,578 seconds, whereas for the eighth scenario, the value rises to 10,258 seconds. For the scenarios involving heavy load and 8 operating rooms (i.e., Scenarios 8 and 9), the GAMS based approach is not able to find the optimal solution or near optimal solution due to the problem size. We see that for scenarios 1, 2, 3, 4, and 6, the GAMS based approach provides better results. To cite an instance, for scenario 2, the relative difference is 4.96% in favor of the GAMS based approach. For the scenarios 4 and 6, the difference is 5.29% and 4.14% respectively.

The progression of objective function value of the incumbent solution over the generations needs to be examined for the genetic algorithm. Figure 13 shows the progression of the best solution over the number of generations for the 5th Scenario.

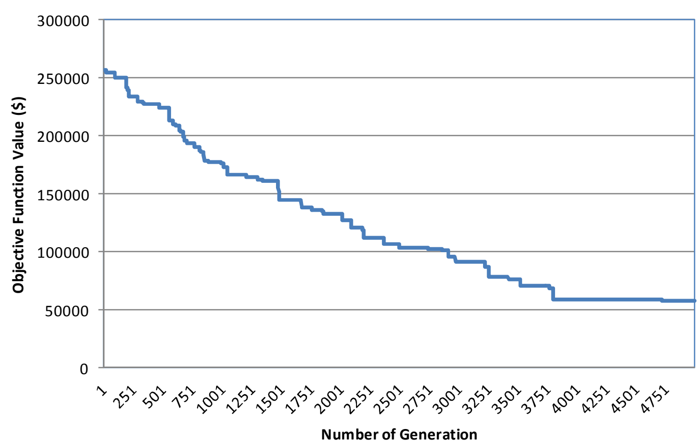


Figure 13. Progression of best objective function value for the genetic algorithm for the 5th scenario

Figure 13 indicates that the objective function value converges to the value of $58,245 approximately after 4,300 generations. The objective function value progresses in stepwise functions. Frequent stepwise decreases in objective function value are observed especially at the initial generations. This trend is replaced with less frequent decrements throughout the later generations. However, based on the amount of decrements, the trend is not obvious, the amount of reductions in the earlier generations usually equals to the amount of decrements in the later generations. Therefore, a trend does not exist in terms of the amount of reductions in the objective function value throughout the generations. Similarly, the progression of the objective function of best solution for Scenarios 2 and 8 are provided in Figures 14 and 15 correspondingly.

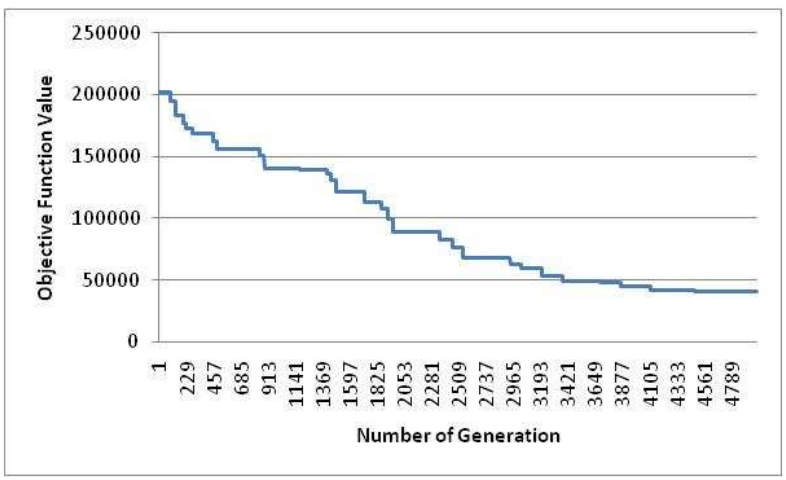


Figure 14. Progression of best objective function value for the genetic algorithm for the 2$^{nd}$ scenario

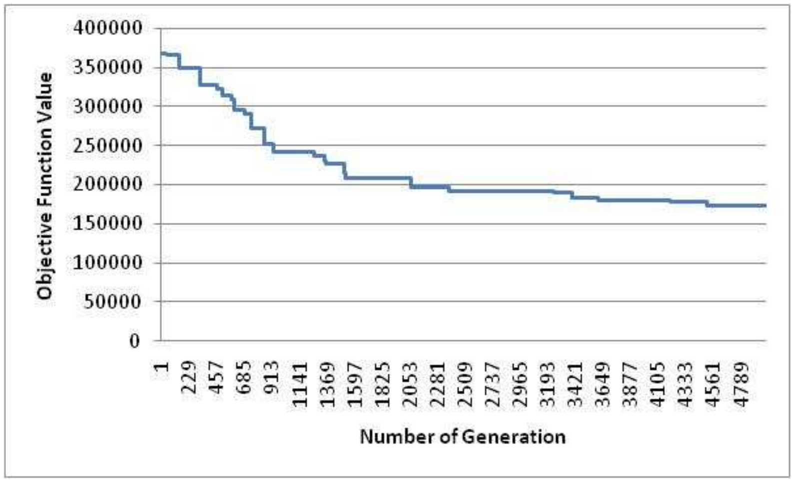


Figure 15. Progression of best objective function value for the genetic algorithm for the 8$^{th}$ scenario

Figure 16 depicts the progression of the solution quality provided by GAMS based approach for the 5$^{th}$ Scenario at predefined time intervals. These computation time intervals are set at 30, 60, 120, 300, 600, 1,800, 3,600, 7,200, 14,400, and 22,452 seconds. These time intervals indicate the computation times that the GAMS based approach allowed running for reaching the optimal solution. It can be seen that, there is a downward decrease indicating that with a higher computation time the solution quality improves. It is also worthwhile to indicate that up until the 14,400 seconds of computation time, the genetic algorithm provides better results.

For Scenario 2, the progression of objective function value of best solution provided by GAMS at the end of 30, 60, 120, 300, 600, 1,800, 3,600, and 6,294 seconds are provided in Figure 17.

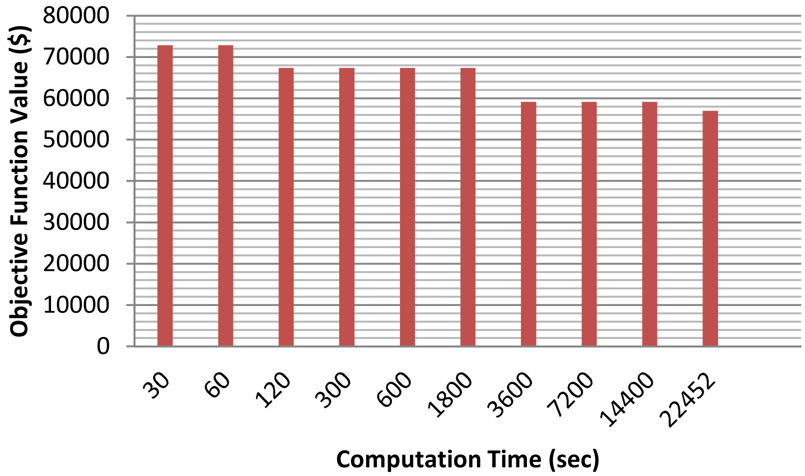


Figure 16. Progression of the GAMS based approach solution quality over computation time for the 5th scenario

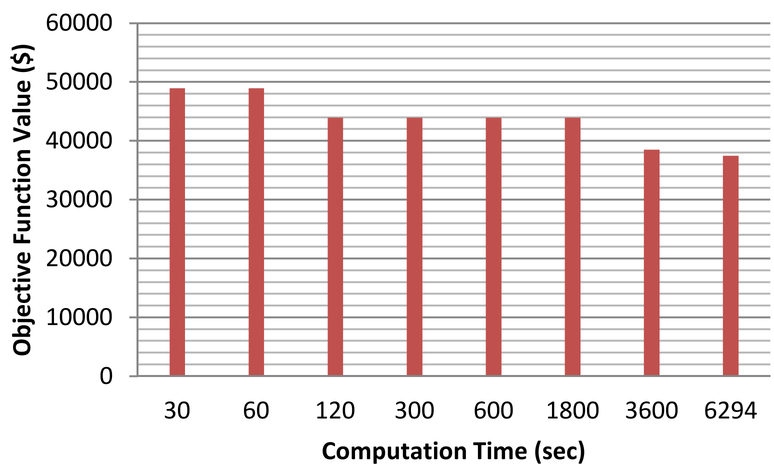


Figure 17. Progression of the GAMS based approach solution quality over computation time for the 2nd scenario

### 3.6. Exclusion of PACU Units

In order to investigate the effect of the downstream PACU units on the solution quality and computation time, an additional set of experiments is conducted. In those experiments, the downstream PACU units are excluded from the analysis by assuming that the number of

PACU beds is greater or equal to the number of operating rooms. Since the duration of stay in PACU bed is 1 hour and the duration of any surgical operation is equal to or longer than 1 hour, for the given problem structure, provided that the number of PACU beds are greater than or equal to the number of operating rooms, there will be no capacity restriction for downstream units. However, if the number of PACU beds is less than the total number of operating rooms, especially for the cases with high patient load, the bed blocking issues might occur in which the patients might not be transferred from an operating room to a PACU bed due to the lack of the adequate resources in a timely manner. Bed blocking might cause disruptions in the system, and further disrupts the schedule in operating rooms and flow of patients throughout the system. Since patients are not able to be transferred to the PACU beds, some might need to recover in operating rooms. This further wastes the dedicated resources in operating rooms since these resources are used for the different purposes than intended. The results are provided in Table 15. The third and fourth columns represent the best objective function values found by the GAMS and genetic algorithm based approaches without consideration of PACU units. The sixth and seventh columns represent the relative difference with respect to the objective function value obtained with PACU unit consideration.

Table 15. Comparison of the genetic algorithm and GAMS based approaches with and without PACU consideration

| Scenario | Number of PACU beds in original scenario | GAMS solution ($) | Genetic algorithm solution ($) | Computation time for genetic algorithm solution (sec) | Relative difference of GAMS solution with PACU unit consideration | Relative difference genetic algorithm solution with PACU unit consideration | Relative difference in terms of computation time |
|---|---|---|---|---|---|---|---|
| 2 | 3 | 36182 | 39475 | 5518.78 | -5.96% | -2.49% | -0.44% |
| 6 | 4 | 86385 | 88985 | 10098.93 | -2.89% | -1.66% | -0.42% |
| 8 | 6 | N/A | 170842 | 10923.65 | N/A | -1.49% | -0.10% |

It is indicated that the reductions in objective function value for both the GAMS and the genetic algorithm based approaches range from 1.49% to 5.96%. Additionally, there are slight reductions in terms of the computation time as compared to the problem instances with consideration of PACU units. The difference in computation time is less than 1% for all problem instances. This might be related with the fact that the capacity constraints for the fitness function calculations for the downstream units are excluded from the analysis for fitness value calculation in genetic algorithm. In terms of the GAMS based approach, the problem becomes more relaxed with less number of constraints, and reduced objective function values are obtained with lower computation times.

The decision of considering the downstream clinic units depends on the healthcare facility setting. If there are adequate number of PACU units (i.e., usually more than number of operating rooms), and the length of stay in PACU unit is shorter than the duration of surgery), in those cases, the mathematical programming model and GAMS based approach might be run without consideration of PACU units. However, for the cases where the duration of stay in PACU units is longer, or there are not adequate number of PACU beds, exclusion of downstream units in the analysis might yield to the bed blocking problem. This in turn might yield to higher system-wide costs because of the waste of dedicated resources in the operating room environment.

### 3.7. Comparison of Deterministic versus Stochastic Version

In order to investigate the effects of variability on duration of surgical operation, another set of runs is conducted using the deterministic surgical durations rather than the stochastic counterparts. For this purpose, two different levels for the duration of surgical operation (i.e., low and high settings) are selected. Basically, the low setting involves using the minimum surgical duration for each type of surgical operations as an input parameter. To cite an instance, for General Surgery, a surgical duration of 1 hour is specified for low setting. For other surgical specialties, the minimum duration for surgical operations is also specified. In the high setting, the maximum surgical duration for that particular surgical specialty is specified. To cite an instance, for the general surgery, high-setting involves a surgical duration of 3 hours. In total, we have three scenarios (i.e., scenarios 2,6, and 8), and 2 settings for each scenario, leading to 6 different cases. Table 16 provides information regarding these scenarios. The results are provided in Table 17.

Table 16. Scenario descriptions for deterministic surgical durations

| **Scenario** | **Setting** | **General Surgery (hour)** | **Ophthalmology (hour)** | **Urology (hour)** | **ENT (hour)** |
|---|---|---|---|---|---|
| 2 | Low | 1 | 1 | 1 | 2 |
| 6 | | 1 | 1 | 1 | 2 |
| 8 | | 1 | 1 | 1 | 2 |
| 2 | High | 3 | 2 | 2 | 3 |
| 6 | | 3 | 2 | 2 | 3 |
| 8 | | 3 | 2 | 2 | 3 |

Table 17. Comparison of genetic algorithm and GAMS based approaches for the deterministic version of the scheduling model

| **Scenario** | **Setting** | **Genetic algorithm solution ($)** | **GAMS solution ($)** | **Computation time for genetic algorithm (sec)** | **Relative difference (%)** |
|---|---|---|---|---|---|
| 2 | Low | $33,248 | $32,280 | 856.43 | -2.91% |
| 6 | Low | $84,228 | $83,126 | 1795.28 | -1.31% |
| 8 | Low | $164,456 | N/A | 1968.92 | N/A |
| 2 | High | $43,060 | $41,404 | 854.28 | -3.85% |
| 6 | High | $95,540 | $96,008 | 1791.92 | 0.49% |
| 8 | High | $192,446 | N/A | 1964.83 | N/A |

Examining the results, we see that as compared to the stochastic surgical duration time, the objective function values provided by genetic algorithm and the GAMS based approach are close to each other. To cite an instance, for the second scenario in the low setting, the relative difference is 2.91%. The stochastic counterpart has a relative difference of 4.96%. The same is also true for the other problem instances. It can be attributed to the fact that as the problem sizes get smaller, the solutions provided by the genetic algorithm and the GAMS based approaches converge to each other. Another observation is that compared to the genetic algorithm approach, the GAMS based approach provides better results except for one setting (i.e., Scenario 6, low-setting).

For the low-settings, the objective function values provided both by the GAMS and genetic algorithm approaches are lower than the stochastic counter-parts. This is intuitive because the surgical operations take less time to perform, therefore less amount of resources is allocated for surgical operations. To cite an instance, for the second scenario, the genetic algorithm solution provides the objective function value of $40,484, while the deterministic counterpart generates a value of $33,248. The same is also true for the solutions provided by the GAMS based approach. For the same setting, the GAMS based approach provides the value of $32,280, whereas the stochastic counterpart generates a value of $38,476.

For the scenarios involving maximum setting, the same is true. Since more resources are used in terms of operating rooms and surgical teams, the results obtained by the deterministic models are higher in the maximum setting as compared to the stochastic counterparts. To cite an instance for genetic algorithm, for Scenario 2, high setting gives a value of $43,060, whereas the stochastic counterpart produces a value of $40,484.

Another finding is that it takes similar amount of time to run scenarios involving low and high settings. To cite an instance, the computation time for Scenario 2 under low setting is 856.43 seconds, whereas the computation time under the high setting is 854.28 seconds. It indicates that the duration length of surgical operations is not a significant factor for determining the computation times in a deterministic setting. The same is also true for other scenarios. Although there is a slight decrease of computation time with respect to the scenarios of high-setting, the decrease is not significant. An interesting observation is that the deterministic version of problem takes significantly less amount of time as compared to the stochastic version. To cite an instance, the computation time is decreased from 5,577.82 seconds to 856.43 seconds when the deterministic version is solved instead of the stochastic counterpart for Scenario 2 in low setting. The computation time is reduced by 85%. This is due to the fact that since a single scenario exists with respect to the duration of surgical operation in a deterministic version, the calculation for the fitness function as compared to the stochastic counterpart (i.e., where 24 scenarios exist based on joint distributions) takes much less time for genetic algorithm based implementation. This significantly affects the computation time. The same is also true for the scenarios involving computation time for the high setting. Rather than using the minimum and maximum durations of the surgeries, one can use the expected values of surgical duration for comparison purposes. To cite an instance, the expected value of surgical duration for General Surgery is 1.8 hours. However, in order to accommodate the fractional values for surgical durations, the resolution for time should be increased. Both mathematical programming model and genetic algorithm based approaches assume discrete values representing surgical durations. Rather than incorporating one hour time slots, it might be worthwhile to incorporate the half-hour time slots for the genetic algorithm and GAMS based approaches. However, increasing the time resolution increases the problem size. For GAMS implementation, the number of variables significantly increases, because of expanding the range for index of subscript *t*. This might significantly increase the problem size because most of the variables have *t* as their subscripts and dependent on time *t* as shown in Tables 4 and 5. This is also true for the genetic algorithm approach. This is because within the genetic algorithm many loops are performed based on time and thus it is expected that the computation time significantly increases if the time resolution is increased from one hour to half an hour.

### 3.8. Case Study

For illustration purposes, a sample problem involving four operating rooms with approximately 110% patient load and planning period of 5 days is solved. Tables 18-21 present the schedule of the elective patients. Note that the time slots occupied by patient surgeries in italic letters indicate the possibility of extending the surgery such that the particular time slot might be occupied. To cite an instance, for operating room #3, the 09:00 am-10:00 am and 10:00 am-11:00 am time slots of day 3 might be occupied for the General Surgery operation of patient 21 with some probability based on the scenario generated.

Table 18. Elective surgery schedule for operating room #1

| **Time** | **Day 1** | **Day 2** | **Day 3** | **Day 4** | **Day 5** |
|---|---|---|---|---|---|
| 08.00-09:00 | | | Ortho Pat 51 | ENT Pat 5 | ENT Pat 4 |
| 09:00-10:00 | Gn. Sur. Pat 14 | CV Patient 2 | Ortho Pat 51 | *ENT Pat 5* | *ENT Pat 4* |
| 10:00-11:00 | *Gn. Sur. Pat 14* | CV Patient 2 | Gn. Sur. Pat 23 | CV Patient 3 | Uro Pat 67 |
| 11:00-12:00 | *Gn. Sur. Pat 14* | CV Patient 2 | *Gn. Sur. Pat 23* | CV Patient 3 | *Uro Pat 67* |
| 12:00-13:00 | CV Patient 1 | CV Patient 2 | *Gn. Sur. Pat 23* | CV Patient 3 | Gn. Sur. Pat 22 |
| 13:00-14:00 | CV Patient 1 | Gn. Sur. Pat 26 | Ortho Pat 60 | CV Patient 3 | *Gn. Sur. Pat 22* |
| 14:00-15:00 | CV Patient 1 | *Gn. Sur. Pat 26* | Ortho Pat 60 | Ortho Pat 59 | *Gn. Sur. Pat 22* |
| 15:00-16:00 | CV Patient 1 | *Gn. Sur. Pat 26* | | Ortho Pat 59 | |
| 16:00-17:00 | | | | | |
| 17:00-18:00 | | | | | |

Table 19. Elective surgery schedule for operating room #2

| **Time** | **Day 1** | **Day 2** | **Day 3** | **Day 4** | **Day 5** |
|---|---|---|---|---|---|
| 08.00:09:00 | Gn. Sur. Pat 15 | Gn. Sur. Pat 19 | ENT Pat 7 | Opht Pat 49 | Gn. Sur. Pat 24 |
| 09:00-10:00 | *Gn. Sur. Pat 15* | *Gn. Sur. Pat 19* | *ENT Pat 7* | *Opht Pat 49* | *Gn. Sur. Pat 24* |
| 10:00-11:00 | *Gn. Sur. Pat 15* | *Gn. Sur. Pat 19* | OB/GYN Pat 46 | ENT Pat 6 | *Gn. Sur. Pat 24* |
| 11:00-12:00 | Ortho Pat 52 | ENT Pat 9 | OB/GYN Pat 46 | *ENT Pat 6* | ENT Pat 8 |
| 12:00-13:00 | Ortho Pat 52 | *ENT Pat 9* | Gn. Sur. Pat 20 | Neuro Pat 39 | *ENT Pat 8* |
| 13:00-14:00 | Ortho Pat 58 | OB/GYN Pat 42 | *Gn. Sur. Pat 20* | Neuro Pat 39 | OB/GYN Pat 43 |
| 14:00-15:00 | Ortho Pat 58 | OB/GYN Pat 42 | *Gn. Sur. Pat 20* | Neuro Pat 39 | OB/GYN Pat 43 |
| 15:00-16:00 | Gn. Sur. Pat 18 | Gn. Sur. Pat 17 | Ortho Pat 53 | Neuro Pat 39 | Gn. Sur. Pat 16 |
| 16:00-17:00 | *Gn. Sur. Pat 18* | *Gn. Sur. Pat 17* | Ortho Pat 53 | Ortho Pat 56 | *Gn. Sur. Pat 16* |
| 17:00-18:00 | *Gn. Sur. Pat 18* | *Gn. Sur. Pat 17* | | Ortho Pat 56 | *Gn. Sur. Pat 16* |

Table 20. Elective surgery schedule for operating room #3

| **Time** | **Day 1** | **Day 2** | **Day 3** | **Day 4** | **Day 5** |
|---|---|---|---|---|---|
| 08.00:09:00 | Ortho Pat 54 | OB/GYN Pat 44 | Gn. Sur. Pat 21 | Hand Pat 36 | Podi Pat 62 |
| 09:00-10:00 | Ortho Pat 54 | OB/GYN Pat 44 | *Gn. Sur. Pat 21* | *Uro Pat 69* | Podi Pat 62 |
| 10:00-11:00 | Uro Pat 68 | OB/GYN Pat 45 | *Gn. Sur. Pat 21* | Uro Pat 69 | ENT Pat 10 |
| 11:00-12:00 | *Uro Pat 68* | OB/GYN Pat 45 | | Opht Pat 48 | *ENT Pat 10* |
| 12:00-13:00 | Hand Pat 35 | ENT Pat 12 | ENT Pat 11 | *Opht Pat 48* | Gn. Sur. Pat 34 |
| 13:00-14:00 | Gn. Sur. Pat 33 | *ENT Pat 12* | *ENT Pat 11* | Gn. Sur. Pat 28 | *Gn. Sur. Pat 34* |
| 14:00-15:00 | *Gn. Sur. Pat 33* | Ortho Pat 57 | Opht Pat 50 | *Gn. Sur. Pat 28* | *Gn. Sur. Pat 34* |
| 15:00-16:00 | *Gn. Sur. Pat 33* | Ortho Pat 57 | *Opht Pat 50* | *Gn. Sur. Pat 28* | |
| 16:00-17:00 | | | | | |
| 17:00-18:00 | | | | | |

Table 21. Elective surgery schedule for operating room #4

| **Time** | **Day 1** | **Day 2** | **Day 3** | **Day 4** | **Day 5** |
|---|---|---|---|---|---|
| 08.00:09:00 | Neuro Pat 40 | Gn. Sur. Pat 30 | Podi Pat 63 | Gn. Sur. Pat 31 | OB/GYN Pat 41 |
| 09:00-10:00 | Neuro Pat 40 | *Gn. Sur. Pat 30* | Podi Pat 63 | *Gn. Sur. Pat 31* | OB/GYN Pat 41 |
| 10:00-11:00 | Neuro Pat 40 | *Gn. Sur. Pat 30* | Ortho Pat 55 | *Gn. Sur. Pat 31* | Ortho Pat 61 |
| 11:00-12:00 | Neuro Pat 40 | Gn. Sur. Pat 29 | Ortho Pat 55 | OB/GYN Pat 47 | Ortho Pat 61 |
| 12:00-13:00 | Gn. Sur. Pat 25 | *Gn. Sur. Pat 29* | Gn. Sur. Pat 27 | OB/GYN Pat 47 | Podi Pat 65 |
| 13:00-14:00 | *Gn. Sur. Pat 25* | *Gn. Sur. Pat 29* | *Gn. Sur. Pat 27* | ENT Pat 13 | Podi Pat 65 |
| 14:00-15:00 | *Gn. Sur. Pat 25* | Neuro Pat 38 | *Gn. Sur. Pat 27* | *ENT Pat 13* | Hand Pat 37 |
| 15:00-16:00 | Uro Pat 66 | Neuro Pat 38 | Gn. Sur. Pat 32 | Podi Pat 64 | |
| 16:00-17:00 | *Uro Pat 66* | Neuro Pat 38 | *Gn. Sur. Pat 32* | Podi Pat 64 | |
| 17:00-18:00 | | Neuro Pat 38 | *Gn. Sur. Pat 32* | | |

Examining the solution presented in Tables 18-21, it can be seen that overtime hours are scheduled for operating rooms #2 and #4 for the elective surgeries. For operating rooms #1 and #3, no overtime hours are scheduled. Note that in 4-operating room scenarios overtime hours are allowed for only 2 operating rooms. In our case, these are operating rooms #2 and #4. Since no overtime hours can be scheduled, the time slots between 16:00-18:00 are vacant for operating rooms 1 and 3. Additionally, the level of PACU beds are set to 3, however there is a need for expanding the capacity of the current PACU beds. To cite an instance, in the third day, at 10:00 am, for Orthopedics patient 52 and Podiatry patient 63, the surgical operations are concluded. For ENT patient 8, and General Surgery patient 22, with some probability, the surgical operations might be concluded as well. Therefore, 4 PACU beds are needed. According to Scenario, 3 PACU beds are already available; there is a need for expanding the current capacity.

Note that the empty time slots in elective surgery schedule are actually related with the constraints based on the availability of surgical teams. To cite an instance, if we prepone the start of the surgery for Patient 9 from 12:00 to 11:00 for reducing the idle time, then we need to increase the number of ENT team working simultaneously from 1 to 2 for that time period. The possibility of existence of alternative optimal solutions cannot be ruled out.

The model also considers the variation of surgical duration for Ophthalmology, Urology, General Surgery, and the Ear Nose and Throat (ENT) cases and adjusts the elective patient schedule accordingly. Note that the developed mathematical model can implicitly take the time preferences of surgical team into account, by limiting the availability of those teams during undesired timeslots in operating rooms. For instance, if the CV surgical team will not want to perform surgical operations during afternoons then the number of available teams for CV surgical operations for that that time period might be set to 0, and the overall schedule might be adjusted accordingly.

### 3.9. General Discussions - Elective Patient Scheduling

For the scheduling of the elective patients, we see that among the scenarios, the GAMS based approach overall performs better as compared to the genetic algorithm. However, for large problem sizes which involve 85% and 110% patient loads and 8 operating rooms, the GAMS based approach is unable to find the optimal or near optimal solution, while the genetic algorithm finds a near optimal solution within a reasonable amount of time. The failure for the GAMS approach to find an optimal solution can be attributable to the increased problem size. As the number of the patients and/or number of operating rooms increases, the GAMS based approach is not able to handle the increased number of variables and constraints associated with the larger problem sizes and thus fails to provide solutions.

Another observation is that when the downstream PACU units are disregarded, the objective function values reported by both the GAMS approach and the genetic algorithm decrease considerably. This is intuitive because in the scenarios selected for the analysis, the number of PACU beds is less than the total number of operating rooms. However, although the objective function decreases, it does not mean that not considering downstream units will decrease the overall costs. As previously mentioned, there might be a bed blocking instance, where the patients from operating rooms might not be transferred to the PACU units due to the non-availability of PACU beds. This definitely disrupts the flow of patients in the system and decreases the patient satisfaction, and for some cases, it might cause harm to the patients because the necessary surgical operation cannot be performed on the next patient scheduled time. For the cases where dedicated equipment and specialized care are not required after surgical operations, such as Podiatry and Hand surgeries, the patients might be transferred to hospital wards after surgery to alleviate the bed blocking problem. The patient might recover from the effects of anesthesia in the hospital ward. However, most of time, usually specialized equipment/care is required after surgery, especially for surgeries involving Neurology and Cardio-Vascular surgical specialties. In those cases, transferring patient directly to hospital wards might not be possible. For that reason, unless there is adequate number of beds/equipment in the downstream units, the downstream units should be considered in the analysis.

Another point that is worth mentioning is that the deterministic version of the problem takes less amount of time to solve. Additionally, for the deterministic version of problem, the objective function value is smaller than the stochastic counterpart under low setting. The opposite is true for the high-setting deterministic scenario where the objective function value is higher than that of the stochastic version. This is intuitive because fewer amounts of resources is allocated for the surgical operations and the additional cost of expanding the current capacity is reduced. On the other hand, for the high end of the duration for the surgical operations, the objective function values reported by GAMS and the genetic algorithm based approaches are substantially higher. This can also be attributable to the fact that additional resources need to be allocated for performing surgical operations. Since the duration for surgical operation is longer than in the case of stochastic version, the need for additional resources also arises, which in turn increases the objective function value. However, the stochastic version of the problem provides a more balanced view by incorporating the possibility of a surgical operation taking variable amount of time, thus provides a better allocation of resources. The stochastic modeling approach increases the flexibility of the system by alleviating the problem of allocating scarce resources in the best manner in a highly chaotic operating room environment.

# 4. RESCHEDULING OF ELECTIVE PATIENTS UPON THE ARRIVAL OF EMERGENCY PATIENTS

In this chapter, we develop the approach for rescheduling the elective patients upon arrival of emergency patients. The approach developed involves building a mathematical programming model, and a genetic algorithm based heuristic approach for rescheduling elective patients.

Although some of the healthcare centers reserve an operating room for performing emergency admissions, due to the scarce resources and less number of operating rooms, there might be a need for sharing operating rooms both for the emergency admissions and elective patients. In that case, if there is not any available resources for performing surgery on emergency patients, due to the nature of emergency admission, emergency patients should be given priority over the elective patients, and should be operated immediately. In the cases where the elective patient load is high, and no dedicated operating rooms are available for the emergency admissions, it is highly likely that the admission of emergency patients might lead to the disruption to the elective surgery schedule, because there are not enough open time slots for performing emergency surgery, and the elective patients need to be rescheduled. This will disrupt the current elective patient schedule.

The approach that will be outlined in this chapter basically addresses two different questions. The first question is whether to admit the emergency patient, and the second question is in case the emergency patients are admitted to healthcare facility, how the new schedule will be formed after the inclusion of the emergency patients. Rescheduling of the elective patients provides additional challenges most of the time in terms of the availability of the resources and prioritization of the patients. The objective is to minimize disruption to the existing schedule. Minimizing disruption is important for two reasons. First, it requires less planning for reallocating existing resources. To cite an instance, if the cyclic schedules are employed for the surgical staff, the change in the schedule might lead to employing more resources to accommodate the new surgical schedule, therefore increasing associated costs. Overtime practices for the surgical staff members might be utilized. Even the schedule is not disrupted based on time/date, then there might be a necessity for performing the surgical operation in a different operating room. This might also cause inconvenience because the dedicated equipment required for performing a specific operation should be moved to a new operating room. To make it worse, moving equipment might not be possible in certain instances. For example, to move the existing equipment for performing Cardio-Vascular surgical operation from the dedicated operating room to another one might not be possible due to the certain provisions required for operating that equipment such as additional cabling requirements. Therefore, the disruption to the existing schedule both in terms of time/date and operating room should be minimized with regard to the existing resources. Second, inconvenience might arise due to the time commitments associated with elective patients. Given that the elective patient is already scheduled, he/she might be hesitant to be rescheduled for another time period. Due to the stressful nature of surgical operations, providing a notice to the elective patient might cause inconvenience especially if it is a short one. Postponing the surgical operation for a couple of hours for a particular patient might not cause a great deal of inconvenience, however preponing the surgical operation, or postponing the surgical operation to the next day or a couple of days might lead to significant patient dissatisfaction and should be avoided as much as possible. For that reason, it will be worthwhile to minimize the amount of disruption in the existing elective schedule from the staff, equipment, and the patient perspective.

In this chapter, we develop the corresponding approach and build a model for minimizing the disruption in the elective patient schedule and the expansion of current resource levels upon the inclusion of emergency patients. Our objective in this stage is to develop a methodology for rescheduling the elective surgeries upon admission of the emergency surgeries. Due to the arrival of the emergency patients, the current elective surgery schedule might be disrupted, and the need might arise for rescheduling the elective surgeries scheduled. In order to model this problem, we develop a deterministic mixed integer linear programming model. We consider two distinct categories of patient admissions for surgical procedures. The first category is elective surgeries, which are already scheduled. The schedule of elective surgeries is treated as an input in our study. The second category is emergency patient arrivals. If an emergency patient is admitted, in the presence of shared operating rooms and surgical teams, the elective surgeries might need to be rescheduled. Upon the request for admission of emergency patients, the decision makers in a hospital must provide the timely decisions on (1) whether to admit or divert the emergency patient(s)? and (2) if any emergency patient is admitted, how to adjust the elective schedule to accommodate emergency admissions such that the disruption to the current schedule and the need for expanding the current resources are minimized?

If an emergency patient is admitted, he/she needs to be operated immediately and the changes need to be made in the elective surgery schedule accordingly. If the patient is not admitted, there will be no changes in the elective surgery schedule. These decisions are made under a variety of constraints so that the costs incurred due to the disruptions are minimized.

The constraints include surgical team availability (e.g., surgeon, scrubbers, technician, and anesthetists), operating room availability, over-time hour constraints for operating rooms, overutilization constraints at PACU, and PACU bed availability. Meanwhile, the cost items include the costs of delaying and preponing elective surgeries, the opportunity cost of diverting (or not admitting) the patients, the overtime and overutilization costs of operating rooms and the PACU units are also considered in the model respectively. The overtime hours are defined as the extended working hours beyond the regular working hours of operating rooms with the accompanying resources such as surgical teams. The overutilization in a PACU unit is defined in terms of the additional beds and supporting equipment/staff members needed in the PACU.

We develop the corresponding mathematical programming model to tackle the problem. Based on the initial results, it can be seen that depending on the problem size, it might not be possible to solve the given problem instance to optimality within the specified time period by employing commercial solvers. In order to overcome this problem, an evolutionary approach based on the genetic algorithm is developed and implemented.

### 4.1. Mathematical Programming Model

In order to make the optimal decisions on emergency surgery admission and elective surgery rescheduling, we develop a mixed integer linear programming (MILP) model to capture the patient flow between operating rooms and downstream clinic entities. The objective of the MILP model is to minimize the costs associated with the delaying and preponing elective surgery patients, and overtime/overutilization of operating rooms and PACU units.

### 4.1.1. Notation

The notation and descriptions of the indices and decision variables and parameters used in the MILP model are provided in Table 22 and Table 23.

Table 22. Notation for index, sets, and decision variables for rescheduling model

| | |
|---|---|
| **Indices** | |
| $d$ : | Day index; $d \in \{1, \ldots, D\}$ |
| $h$: | Emergency patient index; $h \in \{1, \ldots, H\}$. |
| $i$: | Elective patient index; $i \in \{1, \ldots, I\}$. |
| $j$: | Surgical operation type; $j \in \{1, \ldots, J\}$ |
| $t$: | Time period index, $\mathrm{t} \in \mathrm{T} \cup \{T+1\}, T+1$ indicates the time period outside the scheduling horizon |
| $t'$: | Auxiliary time index, $\mathrm{t'} \in \mathrm{T} \cup \{T+1\}, T+1$ indicates the time period outside the scheduling horizon |
| **Sets** | |
| $T^A$: | {Set for overtime hours} |
| $T^B$: | {Set for the time period for which it is not possible to perform the surgeries} |
| $T^C$: | {Set for regular working hours} |
| **Decision variables** | |
| $O_d^{OR}$ : | Amount of overtime utilized for the operating rooms at day $d$ |
| $O^{PACU}$ : | Amount of additional beds/equipment placed in PACU |
| $s_{it}$ : | $\begin{cases} 1 \text{ if elective patient i occupies a bed at PACU at time period t} \\ 0 \textit{ otherwise} \end{cases}$ |
| $s'_{ht}$ : | $\begin{cases} 1 \text{ if emergency patient h occupies a bed at PACU at time period t} \\ 0 \textit{ otherwise} \end{cases}$ |
| $x_{it}$ : | $\begin{cases} 1 \text{ if the surgery starts at the beginning of time period t for elective patient i} \\ 0 \textit{ otherwise} \end{cases}$ |
| $x'_{ht}$ : | $\begin{cases} 1 \text{ if the surgery starts at the beginning of time period t for emergency patient h} \\ 0 \textit{ otherwise} \end{cases}$ |
| $y_{it}$ : | $\begin{cases} 1 \text{ if elective patient i has surgery at time period t} \\ 0 \text{ otherwise} \end{cases}$ |
| $y'_{ht}$ : | $\begin{cases} 1 \text{ if emergency patient h has surgery at time period t} \\ 0 \text{ otherwise} \end{cases}$ |

Table 23. Notation for the parameters for mathematical programming model for rescheduling

| Parameters | |
|---|---|
| $B^{PACU}$: | Current capacity of the PACU unit for the scheduling horizon (in terms of beds/equipment) |
| $B_d^{OR}$ : | Current capacity of the operating rooms at day *d* (in terms of hours) |
| $C^{PACU}$ : | Unit expansion cost of the PACU during the planning period ($/bed & equipment) |
| $C^{OR}$ : | Overtime utilization cost of one unit of operating room for the planning period ($/hour) |
| $C_{jtt'}$ : | Cost of performing elective surgery scheduled at time period *t* and performed at time period *t'* |
| $g_{it}$ : | $\begin{cases} 1 \text{ if elective patient i is scheduled to have operation at time t} \\ 0 \text{ otherwise} \end{cases}$ |
| $m_{ij}$: | $\begin{cases} 1 \text{ if elective patient i needs surgery of type j} \\ 0 \text{ otherwise} \end{cases}$ |
| $m'_{hj}$: | $\begin{cases} 1 \text{ if emergency patient h needs surgery of type j} \\ 0 \text{ otherwise} \end{cases}$ |
| *N*: | Number of operating rooms in the system |
| $O_j$ : | Operation time for surgery type *j* (hours) |
| $r_j$ : | Cost of turning down the emergency patient requesting surgery type *j* |
| $S_j$ : | Length of stay at PACU for surgery type *j* (hours) |
| $t_s$: | Reference starting time (i.e., the time when the emergency patient arrives and the model is run) |
| *T:* | Number of time periods in the planning horizon |
| $U^{PACU}$ : | Upper limit on the overutilization of PACU (bed/equipment) |
| $U^{OR}$ : | Upper limit on the overtime utilization cost of the operating rooms (hours) |
| $\lambda_t$ : | Number of beds occupied in the PACU unit at time period *t* from the previous scheduling cycle |
| $\mu_{jt}$ | Number of surgical teams available for performing surgery type j at time period *t* |

### 4.1.2. MILP model formulation

$$min \quad \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=t_s}^{T+1}\sum_{t'=t_s}^{T+1} g_{it'}C_{jtt'}m_{ij}x_{it} + \sum_{h=1}^{H}\sum_{t=t_s}^{T+1}\sum_{j=1}^{J} r_j m'_{hj}(1-x'_{ht}) + C^{PACU}O^{PACU} + \sum_{d=1}^{D} C^{OR}O_d^{OR} \qquad (18)$$

s.t.

$$\sum_{i=1}^{I}\sum_{t\in T^C}\sum_{j=1}^{J} m_{ij}O_j x_{it} + \sum_{h=1}^{H}\sum_{t\in T^C}\sum_{j=1}^{J} m'_{hj}O_j x'_{ht} = B_d^{OR} + O_d^{OR} \qquad \forall d \qquad (19)$$

$$\sum_{i=1}^{I}\sum_{t\in T^A}\sum_{j=1}^{J} y_{it} + \sum_{h=1}^{H}\sum_{t\in T^A}\sum_{j=1}^{J} y'_{ht} = O_d^{OR} \qquad \forall d \tag{20}$$

$$\sum_{t\in T^B} y_{it} = 0 \qquad \forall i \tag{21}$$

$$\sum_{t\in T^B} y'_{ht} = 0 \qquad \forall h \tag{22}$$

$$\sum_{t=t_s}^{T+1} x_{it} = 1 \qquad \text{for } i = 1,\ldots,I \tag{23}$$

$$\sum_{i=1}^{I} x_{it} + \sum_{h=1}^{H} x'_{ht} \le N \qquad \text{for } t = t_s,\ldots,T \tag{24}$$

$$\sum_{i=1}^{I} y_{it} + \sum_{h=1}^{H} y'_{ht} \le N \qquad \text{for } t = t_s,\ldots,T \tag{25}$$

$$\sum_{i=1}^{I} m_{ij} y_{it} + \sum_{h=1}^{H} m'_{hj}\, y'_{ht} \le \mu_{jt} \qquad \forall j, t = t_s \ldots T \tag{26}$$

$$s_{it'} \ge m_{ij} x_{it} \qquad \forall i, j,\ t = t_s,\ldots,T, t' = t + t_s + O_j - 1,\ldots, t + t_s + O_j + S_j - 2 \tag{27}$$

$$s'_{ht'} \ge m'_{hj}\, x'_{ht} \qquad \forall i, j,\ t = t_s,\ldots,T, t' = t + t_s + O_j - 1,\ldots, t + t_s + O_j + S_j - 2 \tag{28}$$

$$y_{it'} \ge m_{ij} x_{it} \qquad \forall i, j, t = t_s,\ldots,T, t' = t + t_s - 1,\ldots, t + t_s + O_j - 2 \tag{29}$$

$$y'_{ht'} \ge m'_{hj}\, x'_{ht} \qquad \forall i, j, t = t_s,\ldots,T, t' = t + t_s - 1,\ldots, t + t_s + O_j - 2 \tag{30}$$

$$\sum_{t=t_s+1}^{T+1} x'_{ht} = 0 \qquad \forall \mathrm{h} \tag{31}$$

$$\sum_{i=1}^{I} s_{it} + \sum_{h=1}^{H} s'_{ht} + \lambda_t \le B^{\mathrm{PACU}} + O^{\mathrm{PACU}} \qquad \text{for } t = t_s,\ldots,T \tag{32}$$

$$O^{PACU} \le U^{PACU} \tag{33}$$

$$O_d^{OR} \le U^{OR}, \qquad \forall \mathrm{d} \tag{34}$$

$s_{it}$, $s'_{ht}$, $x_{it}$, $x'_{ht}$, $y_{it}$, and $y'_{ht}$ are binary variables

$O_d^{OR}$ and $O^{PACU}$ are non-negative integer variables.

Eq. 18 consists of four different terms. The first term indicates the cost of preponing or postponing the elective surgeries. The second term indicates the cost of turning down the emergency patients in terms of the lost revenue. The third term is the overtime utilization of operating rooms in terms of the available operating hours. The fourth term indicates the marginal cost of placing additional beds, equipment, and hiring additional staff for increasing the capacity of the PACU units.

Eq. 19 ensures that the total time for performing surgical operations for elective and emergency patients should be equal to the total regular and overtime utilization of the operating rooms. Eq. 20 stipulates that the surgical operations performed outside of the regular working hours are considered to be in the overtime hours. Eqs. 21 and 22 indicate that no operation is allowed other than the indicated regular and overtime hours for elective and emergency patients respectively. Eq. 23 ensures that all the elective patients need to be operated in the given operating cycle; if this is not possible, it will be assumed that they will be operated beyond the planning period indicated by time period $T+1$. Eqs. 24 and 25 ensure that the number of simultaneous new starts and ongoing operations cannot exceed the maximum number of the operating rooms. Eq. 26 indicates that the number of surgical operations that are performed simultaneously cannot exceed the number of surgical teams that are capable of performing those operations. Eqs. 27 and 28 indicate that the bed/equipment in the PACU units will be occupied following the surgical operation for a specified period of time during the recovery of the patients for elective and emergency patients respectively. Eqs. 29 and 30 ensure that the corresponding time slots for surgical operations are occupied during the specified duration following the start of surgical operations for elective and emergency patients respectively. Eq. 31 stipulates that no delay is permissible for the emergency patients arriving at the medical facility, and the opportunity cost of not operating the emergency patients in terms of lost revenue is reflected in Eq. 18. Eq. 32 is related with the bed/equipment/staff constraints that are available in the PACU units. Eq. 33 ensures that the number of operating hours cannot exceed the upper limit for the permissible operating hours for the operating rooms. Eq. 34 stipulates that the total number of additional beds/equipment added to the PACU units cannot exceed the permissible additional bed/equipment that can be placed in those units due to various considerations such as rules and regulations, physical place restrictions, etc.

The mathematical programming model involves the use of the binary variables as well as integer variables. The problem size increases exponentially with the increases in the number of patients and the type of surgical operations. In addition, the increase in the number of operating rooms also increases the problem size. The model has been implemented in the commercial optimization software package, GAMS solver. Unfortunately, it is found that the solver cannot provide efficient solutions for some scenarios using an Intel Quad Core PC – exact solutions cannot be obtained in a reasonable amount of time. The PC features an Intel® Core™ i5 processor at 2.8 GHz, and 8 GB memory. Clearly, the excessive computation time has negative impact on the usability of the approach and its potential merits. As mentioned earlier, the decisions should be made within 35-45 minutes – the shorter, the better. The urgency leaves little room for extended computation times required by improving the solution quality or possibly obtaining optimal solutions. Rather than searching for the exact optimal solutions, a heuristic approach is developed below to obtain the near optimal solutions in a reasonable time limit.

### 4.2. Genetic Algorithm

Figure 18 shows the steps of genetic algorithm to solve the MILP model. The notations in the genetic algorithm are summarized in Table 24. In the genetic algorithm, the first generation of chromosomes consists of the current elective surgery schedule and $\varphi$-1 chromosomes representing the random schedule of emergency patient and elective patients. Then, each new generation is obtained by applying the corresponding crossover and mutation operators on the chromosomes in the current generation, repairing new off-springs, and selecting the new generation from combined pool. The algorithm stops when the maximum generation limit is attained or the objective function value of the best feasible solution reaches 0.

Table 24. Notation for genetic algorithm parameters for rescheduling model

| **Parameter** | **Notation** |
|---|---|
| Number of parent pairs selected for crossover, or offspring pairs created | $\pi$ |
| Mutation probability for first part of chromosome structure | $\tau_1$ |
| Mutation probability for the second part of chromosome structure | $\tau_2$ |
| Mutation probability for the third part of chromosome structure | $\tau_3$ |
| Mutation probability for the third part of chromosome structure | $\tau_4$ |
| Number of chromosomes/solutions selected for the next generation by elitist selection | $\upsilon$ |
| Number of chromosomes/solutions selected for the next generation by roulette wheel selection | $\varphi$ - $\upsilon$ -*1* |
| Population size | $\varphi$ |
| Limit on maximum generation number | $X$ |
| Repair probability | $\psi$ |

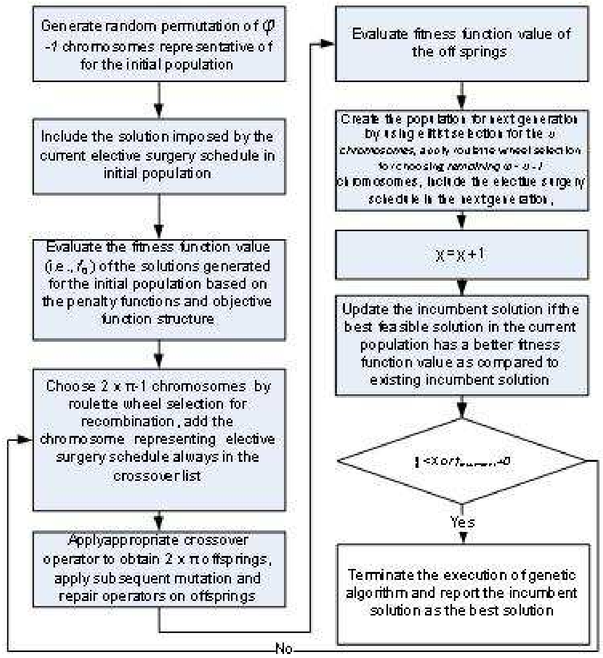


Figure 18. Proposed genetic algorithm flowchart for rescheduling model

#### 4.2.1. Representation of the solution

The chromosome representation for a solution consists of four parts. Figure 19 provides the chromosome representation;

$$\underbrace{3\ 5\ 9\ 13\ 4\ 45...\ \ 37\ 2\ 43\ 34}_{\text{(Sequence of patients)}}$$

$$\underbrace{3\ 4\ 4\ 6\ 7\ 18\ 20\ 32\ 34\ 38\ 42}_{\text{(The number of the empty time slots before each patient)}}$$

$$\underbrace{2\ 0\ 3\ 1\ 3\ 0\ldots\ 3\ 2\ 2\ 1\ 2\ 2\ 2\ 0}_{\text{(Number of patients that are operated in each operating room)}}$$

$$\underbrace{1}_{\text{The operating room that emergency patient is admitted}}$$

Figure 19. Proposed chromosome structure for rescheduling model

The first part represents the sequence of the elective patients scheduled for surgical operations. For example, the chromosome structure “3 5 9 13 4 45... 37 2 43 34” indicates that elective patient 3 is the first patient to be operated, whereas elective patient 34 is the last person who will undergo surgery in an operating room.

The second part represents the number of open slots before a patient’s surgery by repeating the patient identification for the same amount of times. For example, the chromosome representation “2 2 3 4 4 6 7 18 20 32 34 38 42” indicates that there are two empty time slots just before the start of the surgical operation for patient 2. There is an open time slot before the surgery of patient 3. There are two open time slots before the surgery start for patient 4. There is one open time slot before the start of the surgery for patients 6, 7, 18, 20, 32, 34, 38, and 42. For the patient IDs not listed in the second part of a chromosome, there is no open time slot before their surgeries.

The third part of a chromosome governs the number and type of the surgeries performed during a given day. For example, the chromosome representation “2 0 3 1 3 0 … 3 2 2 1 2 2 2 0” demonstrates that the first 2 patients are operated in the first day, no patients are operated in the second day, 3 patients are operated in the third day, 1 patient is operated in fourth day, 3 patients are operated in the fifth day, and no patients are operated outside the planning cycle. In short, the first 6 numbers determine the number of patients operated in and out of the planning cycle for first operating room, and the second set of 6 numbers determine the number of the patients operated in second operating room, etc based on planning cycle of 5 days.

The length of the last part of a chromosome depends on the number of emergency patients arriving at that particular time period. Each bit in that chromosome part indicates whether an emergency patient is admitted or not, and the operating room assigned for surgery if the emergency patient is admitted. Based on Figure 19, we see that the emergency patient is admitted to the hospital, and assigned to operating room #1.

### 4.2.2. Evaluating the fitness of the population

This particular problem is a constrained minimization problem. The representation scheme of the solutions discussed above could not guarantee that a chromosome always represents a feasible solution. The availability of the surgical teams, the constraints satisfying that the operating room is not utilized during lunch hours, the availability of the corresponding beds in the downstream of the clinic flow (i.e., the PACU), the constraint satisfying that operating rooms do not operate after the overtime hours are not enforced by the representation scheme of the solutions proposed for the genetic algorithm. For example, it might be the case that two surgeries of the same type can start simultaneously according to the solution represented by the chromosome, although one surgery team is available to perform that particular surgery. The chromosome structure allows such a schedule although it is not feasible according to Eq. 26. In order to overcome this problem, the weights associated with the violation of the corresponding constraints are added to the objective function. In other words, the constraint violation is penalized through the inclusion of the penalty terms in the objective function.

### 4.2.3. Selecting the best-fit individuals for crossover operator for reproduction

After the fitness value of each solution is calculated, the next step is choosing the members of the population for crossover operation. For this purpose, we make use of the combination of roulette wheel selection with the elitist selection. A solution implied by the original schedule solution (i.e., rejecting the emergency admission, and performing the elective surgeries according to the original schedule) is always selected for the crossover operation. This scheme will allow the variations of the solutions derived from the original schedule solution be represented as offsprings.

### 4.2.4. Crossover operator

For each generation, a certain number of chromosomes/solution pairs are selected for cross-over operation. From those pairs, the first part of the chromosome for the offsprings is created using the partial mapped crossover (PMX) operator (Al-Dulaimi and Ali, 2008). This is because the traditional crossover operator might yield offsprings that have duplicate patients or some patients might not be represented in the chromosome solutions. Literature indicates that PMX operator provides consistent results in terms of the solution quality as compared to other approaches such as order and cycle crossover operators (Al-Dulaimi and Ali, 2008).

For the second and third parts of chromosomes, two-point traditional crossover operator is selected as opposed to single point crossover operator due to the fact that generally the former provides better results as compared to the latter (Sivanandam and Deepa, 2008). In two-point crossover operator, two crossover points are selected and the corresponding bits between these two points are copied to the offspring from the first parent. The bits outside those crossover sites are copied from the second parent. The second offspring is formed in the same manner by replacing the roles of the first and second parents. For the fourth part of chromosomes, the crossover operator is selected based on the length of that particular chromosome part. If only one emergency arrival is considered for admission, one of the parents is selected randomly and the corresponding value is copied to the offspring. If more than two emergency patients at a time are considered for admission, one point crossover operator is implemented.

### 4.2.5. Mutation operator

For the first two parts of the chromosomes, the mutation operator is applied on a bit-by-bit basis. For each bit, a random number uniformly distributed between 0 and 1 is generated. For the first part of the chromosomes, if the generated random number is smaller than $\tau_1$, the second bit is picked randomly, and these two bits are exchanged. In practice, it corresponds to switching the locations of two patients in the sequence of the patients represented in the solution. For the second part of the chromosomes, which indicates empty time slots in the schedule, if the generated random number is smaller than $\tau_2$, another random number will be generated. If the second random number generated is smaller than or equal to 0.5, the bit is deleted from the chromosome, indicating that the length of the vacant time for the operating room before the operation of the patient represented by this bit is reduced by one time unit. If the second random number generated is greater than 0.5, an additional copy of this bit is added to the chromosome. In other words, the length of the vacant time for that operating room before that patient is increased by one time unit.

For the third part of the chromosomes representing the number and type of the surgeries performed during a given day, mutation occurs in the form of generating a random number distributed uniformly between 1 and $(T+1) \times N$. The bit randomly selected is replaced with this new number. For the fourth part of the chromosomes representing whether the emergency patients are admitted or not, mutation occurs in the form of generating a random number distributed uniformly between 1 and $N+1$ for each bit in that chromosome part. The corresponding number in the bit is replaced with this new number.

### 4.2.6. Repair operator

After the crossover and mutation operators are applied to produce the offsprings, the repair operator is performed on the new generated chromosomes with certain probability. A uniformly distributed random number between 0 and 1 is generated for each chromosome, and if the random number is smaller than $\psi$, the repair operator is applied. The repair operator in general works for reducing the overall infeasibility by two different schemes. The first scheme is that if the start of the surgery for a particular patient is scheduled to start on an infeasible time period, the start of the surgery is delayed by some time periods to start in working hour to reduce the infeasibility associated with the solution. Another mechanism for the repair operator is to prepone the start of the surgical operation that is scheduled later to reduce the vacant time between those surgeries. Although the repair mechanism does not guarantee the feasibility because of various other constraints, they serve as an attempt to decrease the overall infeasibility.

### 4.2.7. Replacing least-fit population with new individuals

The elitist selection is used together with the roulette wheel selection to select the offsprings and parents that will create the next generation. After the fitness values of the offpsrings are calculated, the parents and offsprings are ranked in descending order of the fitness function value. The first $\upsilon$ chromosomes in the list are selected for the next generation. Additionally the remaining $\varphi$ - $\upsilon$ -1 number of solutions are selected based on the roulette wheel selection.

The last chromosome included in the new generation is the original schedule implied by the elective surgery schedule. This approach will help develop variations based on the elective surgery schedule which likely produce lower objective function values (i.e., better fitness function values). The solutions developed from the original elective schedule more likely yield solutions that honor the original schedule of elective surgeries, therefore reducing overall objective function value and increasing the corresponding fitness function.

The genetic algorithm parameters are decided based on literature recommendations and the pilot runs. In this study, we refer to the literature for the initial ranges of the parameter values and employ pilot runs for fine tuning of those parameters. As a result, the values presented in Table 25 are selected for our genetic algorithm implementation

Table 25. Genetic algorithm parameters for rescheduling model

| **Parameter** | **Value** |
|---|---|
| $\Pi$ | 15 |
| $\tau_1$ | 1% |
| $\tau_2$ | 1% |
| $\tau_3$ | 1% |
| $\tau_4$ | 1% |
| $Y$ | 10 |
| $\varphi$ - $\upsilon$ -1 | 49 |
| $\Phi$ | 60 |
| $X$ | 8000 |
| $\Psi$ | 50% |

Regarding penalizing scheme in evaluating the fitness function, we adopt the approach developed by Joines and Houck (1994) where the penalty coefficients can be calculated as follows,

$$g_w(\chi) = (\rho_w \chi)^\alpha \tag{35}$$

where $g_w(\chi)$ is the corresponding penalty coefficient for constraint $w$ at generation $\chi$, and $\rho_w$ and $\alpha$ are the constant values. This enables a wider search in scope at the initial generations with lower penalty coefficients, and toward the end of the execution of genetic algorithm with increased penalty coefficients. It is found that the values of 1.25 for $\rho_w$ and 0.5 for $\alpha$ give the best result in this study, and they are also within the suggested ranges (Joines and Houck, 1994).

Table 26 presents coefficients (i.e., cost and penalty figures) that are provided as input for the genetic algorithm as well as the mathematical programming model described in Section 4. Taheri et al. (2007), based on an actual applied study of a trauma center, estimate that admission of emergency patient will generate additional revenue of \$16,603 in the downstream clinic unit. Other cost figures are obtained from the first part of the study (Olejarz, 2009; Park and Dickerson, 2009; and Vogel et al., 2010).

Table 26. Corresponding cost parameters and penalty coefficients for rescheduling model

| | **Penalty calculation** | **Cost coefficients** |
|---|---|---|
| After hours operation | Eq. 35 with $\rho_w = 1.25$, $\alpha = 0.5$ | |
| Exceeding the regular+ overtime capacity of the PACU units | Eq. 35 with $\rho_w = 1.25$, $\alpha = 0.5$ | |
| Exceeding the upper limits of the availability of the surgical team | Eq. 35 with $\rho_w = 1.25$, $\alpha = 0.5$ | |
| Turning down the patient for the emergency admission | | \$16,603 (Taheri et al., 2007) |
| Operating the patient other than designated day (e.g., Wednesday instead of Monday) | | \$3,798/day (Vogel et al., 2010) |
| Operating the patient outside the planning horizon | | \$18,990 (Vogel et al., 2010) |
| Cost of overtime for an operating room | | \$1500/hr (Park and Dickerson, 2009) |

The fitness function of each chromosome is calculated as,

$$f_q = \frac{16,603}{obj_q},\ \mathrm{obj_q} \neq 0\ , \tag{36}$$

where $f_q$ is the fitness function of solution $q$, and $obj_q$ is the objective function value of same solution. Since the problem is a minimization type problem, a lower objective function value indicates a better solution which is reflected by a higher fitness function value. As such, the objective function value is inverted in Eq. 36. In this equation, the value of $16,603 is used as a numerator for calculating the fitness function in which the original elective surgery schedule will yield a fitness function value of 1 if no additional costs are involved in terms of exceeding the current capacity of operating rooms and the PACU units. This value is the cost of turning down an emergency patient according to literature (Taheri et al., 2007). In a sense, Eq. 36 provides a normalizing scheme where all the objective function values are scaled down with a reference value of 1 that belongs to the original elective surgery schedule with no overtime in operating rooms and overutilization in PACU units with the assumption that the emergency patient is not admitted.

### 4.3. Scenarios and Results

Runs are conducted to provide a comparison for GAMS and genetic algorithm based approaches in terms of the computation time and solution quality. The problem actually involves two different stages, where the first stage addresses scheduling of the elective patients, and in the second stage, the problem of rescheduling the elective patients is tackled. The output obtained from the first stage is fed as the input for the second stage of the problem. For this purpose, three scenarios are designated, each representing, low, medium and high patient load.

These scenarios reflect the general pattern where the first scenario features the 4 operating rooms with 50% patient load. In the second scenario, we have 6 operating rooms with 110% patient load, whereas in the third scenario, we have a total of 8 operating rooms with 85% patient load. These scenarios are reflecting the light, moderate and the high patient load cases respectively and obtained from the first stage of the problem. The scenarios that are being depicted are presented Table 27.

Table 27. Scenario descriptions for rescheduling elective patients

| **Scenario number** | **Elective patient load** | **Number of PACU beds** | **Number of operating rooms** | **Number of emergency arrivals** | **Overtime practices** |
|---|---|---|---|---|---|
| 1 | 50% | 3 | 4 | 1 | 2 operating rooms with 4 hours of overtime on daily basis |
| 2 | 110% | 4 | 6 | 1 | 4 operating rooms with 8 hours of overtime in total on daily basis |
| 3 | 85% | 6 | 8 | 1 | 6 operating rooms with 12 hours of overtime on daily basis |

The models are solved both using GAMS commercial solver and the genetic algorithm approach. The solution quality and computation time are compared. For determining the input for the rescheduling model, output from the scheduling model is used. The output of the first model discussed in Chapter 3 is basically the schedule of the elective patients. The schedule of the elective patients is fed into the rescheduling model. The solution quality and the computation time for the GAMS based approach and genetic algorithm are provided in the Table 28.

Table 28. Comparison of GAMS based approaches and genetic algorithm solutions

| **Scenario** | **Solution quality** | | **Computation time** | |
|---|---|---|---|---|
| | GAMS solution ($) | Genetic algorithm ($) | GAMS (sec) | Genetic algorithm (sec) |
| 1 | 0 | 0 | 0.28 | 503.81 |
| 2 | 30192 | 28692 | 2071 | 2071.03 |
| 3 | 32586 | 31788 | 3581 | 3581.23 |

Examining the results presented in Table 28, we see that for the Scenario 1, the objective function value of GAMS based approach is equal to the solution provided by the genetic algorithm albeit with different computation time. In less than a second, the GAMS based approach converges to an optimum solution which is $0, whereas for the genetic algorithm based approach it takes 503.81 seconds to reach to the same objective function value. For the second problem instance, the genetic algorithm provides better results as compared to GAMS based commercial solver for the same amount of computation time. For the third problem instance, the same is true where genetic algorithm providing better results as compared to GAMS based solver. Note that for all the problem instances, the computation time is below 1 hour. In general, it can be said that the computation time for the genetic algorithm increases linearly with the increasing number of patients/operating room however, for the GAMS based commercial solver, the number of variables increases in quadratic terms with the problem size. Since the same amount of computation time is allowed for both solution approaches, it can be seen that in general the genetic algorithm based approach takes the upper hand when the problem sizes gets bigger in terms of the higher number of operating rooms and heavy elective patient load.

### 4.4. Case Study

In order to demonstrate the changes in the elective patient schedule upon the arrival of the emergency patient, we develop a case where there are 4 operating rooms and approximately 110% elective patient load. This corresponds to Scenario 3 which is described in Chapter 3. One key criterion to evaluate the effectiveness of a solution approach is whether it is capable of providing near optimal or ideally optimal solutions in a limited time window. As mentioned earlier, the local medical center usually deals with emergency cases that require immediate attention such as trauma, and thus the decision makers have a limited time window of less than 1 hour before giving the corresponding decisions.

It is assumed that upon the arrival of the emergency patient, the rescheduling of elective patients up to 5 days is considered as in the case of the scheduling case. Table 29 provides the information on the general problem parameters.

Table 29. General problem parameter for the representative case for rescheduling model

| Parameter | Value |
|---|---|
| Number of PACU bed | 3 |
| Average stay in PACU | 60 minutes |
| Overtime options for operating room | 2 operating rooms can be allocated for 2 hour additional time slots each for overtime operations (4:00 pm-6:00 pm) |
| Emergency arrivals | A single trauma patient is brought to the hospital at the first day of the scheduling period (i.e., 08:00 am on Monday). |

For demonstration purposes, the evolution of objective function values of best feasible solutions reported by GAMS. It is indicated in Figure 20. The main purpose is to investigate if the solution obtained within 1 hour (or the decision time window) by the commercial solver is close to the final solutions provided by the GAMS based approach . If this is the case, the genetic algorithm approach will be less attractive and the need for using the genetic algorithm or other heuristic approaches for obtaining the near optimal solutions in the given time window of opportunity (35-45 minutes) will not be necessary. For this purpose, a comparison is made, where the solution quality after the certain computation time is reported. The solution quality (i.e., objective function value) after 1, 2, 3, 5, 15, 20, 30, 60, 90, 120, 240, and 360 minutes are reported.

Examining Figure 20, we see that the solution reported after 1 hour of computational time is approximately 6.5% worse as compared to the solution reported after 6 hours of computation time. This shows that there is a room for improvement, and the heuristic approach might be a viable option to efficiently obtain the solution for rescheduling of the elective patients. It is worthwhile to note that  the genetic algorithm provides an objective function value of $34,788 after approximately 21 minutes of computation time. This value is better than the objective function value reported by the GAMS based approach for the same time period.

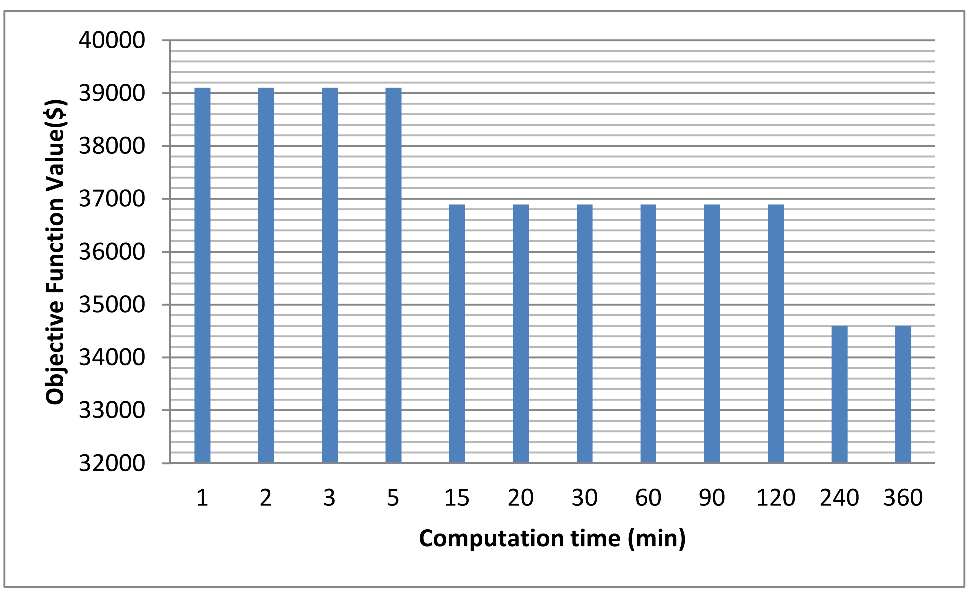


Figure 20. Progression of objective function value of the GAMS based approach

We also examine the evolution of solution quality with respect to the number of generation using the genetic algorithm approach for the representative case. The objective function values of the incumbent solutions as well as the average objective function values are provided in Figure 21.

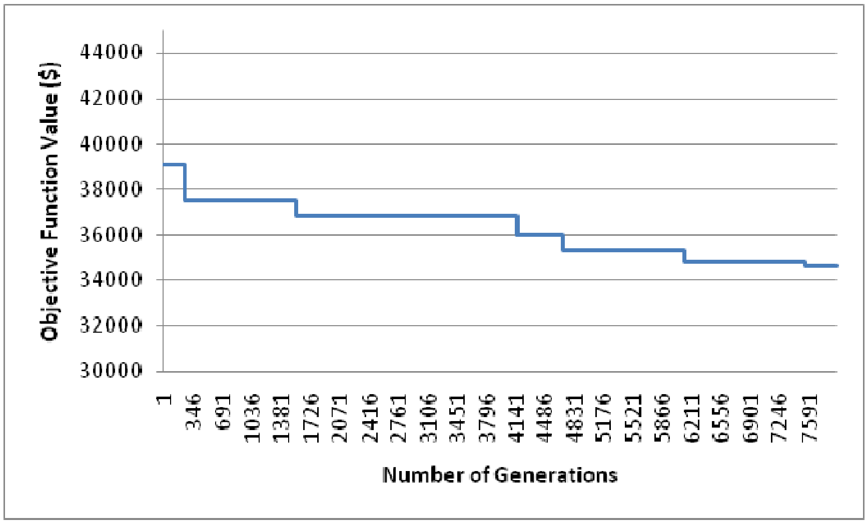


Figure 21. Progression of objective function values for genetic algorithm

As it can be seen in the Figure 21, the incumbent fitness function value has a stepwise character. When, better feasible solutions are found, the incumbent solution is updated with a lower objective function value.  Amount of decrease in the objective function value at initial generations are high, gradually decreasing over the generations.

The following example details the change of surgery schedule after the incorporation of the emergency patient for the representative case. The original schedules of elective surgery for operating rooms #1-#4 are provided in Tables 18-21, respectively. Inclusion of the emergency patient to the current schedule results in the changes in the current elective surgery schedule for all the operating rooms. The new schedules incorporating the admission of the emergency patient for those operating rooms are provided in Tables 30 - 33 respectively.

Table 30. Elective surgery schedule for operating room #1

| Time | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 |
|---|---|---|---|---|---|
| 08.00-09:00 | ***Trauma Pat 70*** | *Gn. Sur. Pat 29* | Ortho Pat 51 | ENT Pat 5 | ENT Pat 4 |
| 09:00-10:00 | ***Trauma Pat 70*** | *Gn. Sur. Pat 29* | Ortho Pat 51 | ENT Pat 5 | ENT Pat 4 |
| 10:00-11:00 | | ***Gn. Sur. Pat 14*** | Gn. Sur. Pat 23 | CV Patient 3 | Uro Pat 67 |
| 11:00-12:00 | | ***Gn. Sur. Pat 14*** | Gn. Sur. Pat 23 | CV Patient 3 | Uro Pat 67 |
| 12:00-13:00 | CV Patient 1 | | | CV Patient 3 | Gn. Sur. Pat 22 |
| 13:00-14:00 | CV Patient 1 | Gn. Sur. Pat 26 | Ortho Pat 60 | CV Patient 3 | Gn. Sur. Pat 22 |
| 14:00-15:00 | CV Patient 1 | Gn. Sur. Pat 26 | Ortho Pat 60 | Ortho Pat 59 | |
| 15:00-16:00 | CV Patient 1 | | | Ortho Pat 59 | |
| 16:00-17:00 | | | | | |
| 17:00-18:00 | | | | | |

Table 31. Elective surgery schedule for operating room #2

| Time | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 |
|---|---|---|---|---|---|
| 08.00:09:00 | Gn. Sur. Pat 15 | Gn. Sur. Pat 19 | ENT Pat 7 | Opht Pat 49 | Gn. Sur. Pat 24 |
| 09:00-10:00 | Gn. Sur. Pat 15 | Gn. Sur. Pat 19 | ENT Pat 7 | Opht Pat 49 | Gn. Sur. Pat 24 |
| 10:00-11:00 | Ortho Pat 52 | ENT Pat 9 | OB/GYN Pat 46 | ENT Pat 6 | ENT Pat 8 |
| 11:00-12:00 | Ortho Pat 52 | ENT Pat 9 | OB/GYN Pat 46 | ENT Pat 6 | ENT Pat 8 |
| 12:00-13:00 | Ortho Pat 58 | OB/GYN Pat 42 | Gn. Sur. Pat 20 | Neuro Pat 39 | OB/GYN Pat 43 |
| 13:00-14:00 | Ortho Pat 58 | OB/GYN Pat 42 | Gn. Sur. Pat 20 | Neuro Pat 39 | OB/GYN Pat 43 |
| 14:00-15:00 | Gn. Sur. Pat 18 | Gn. Sur. Pat 17 | | Neuro Pat 39 | Gn. Sur. Pat 16 |
| 15:00-16:00 | Gn. Sur. Pat 18 | Gn. Sur. Pat 17 | Ortho Pat 53 | Neuro Pat 39 | Gn. Sur. Pat 16 |
| 16:00-17:00 | | | Ortho Pat 53 | Ortho Pat 56 | |
| 17:00-18:00 | | | | Ortho Pat 56 | |

Table 32. Elective surgery schedule for operating room #3

| Time | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 |
|---|---|---|---|---|---|
| 08.00:09:00 | Ortho Pat 54 | OB/GYN Pat 44 | | Hand Pat 36 | Podi Pat 62 |
| 09:00-10:00 | Ortho Pat 54 | OB/GYN Pat 44 | Gn. Sur. Pat 21 | Uro Pat 69 | Podi Pat 62 |
| 10:00-11:00 | Uro Pat 68 | OB/GYN Pat 45 | Gn. Sur. Pat 21 | Uro Pat 69 | ENT Pat 10 |
| 11:00-12:00 | Uro Pat 68 | OB/GYN Pat 45 | | Opht Pat 48 | ENT Pat 10 |
| 12:00-13:00 | Hand Pat 35 | ENT Pat 12 | ENT Pat 11 | *Opht Pat 48* | Gn. Sur. Pat 34 |
| 13:00-14:00 | Gn. Sur. Pat 33 | *ENT Pat 12* | *ENT Pat 11* | Gn. Sur. Pat 28 | Gn. Sur. Pat 34 |
| 14:00-15:00 | Gn. Sur. Pat 33 | Ortho Pat 57 | Opht Pat 50 | *Gn. Sur. Pat 28* | |
| 15:00-16:00 | | Ortho Pat 57 | *Opht Pat 50* | | |
| 16:00-17:00 | | | | | |
| 17:00-18:00 | | | | | |

Table 33. Elective surgery schedule for operating room #4

| Time | Day 1 | Day 2 | Day 3 | Day 4 | Day 5 |
|---|---|---|---|---|---|
| 08.00:09:00 | Neuro Pat 40 | Gn. Sur. Pat 30 | Podi Pat 63 | Gn. Sur. Pat 31 | |
| 09:00-10:00 | Neuro Pat 40 | Gn. Sur. Pat 30 | Podi Pat 63 | Gn. Sur. Pat 31 | OB/GYN Pat 41 |
| 10:00-11:00 | Neuro Pat 40 | | Ortho Pat 55 | | OB/GYN Pat 41 |
| 11:00-12:00 | Neuro Pat 40 | Neuro Pat 38 | Ortho Pat 55 | OB/GYN Pat 47 | Ortho Pat 61 |
| 12:00-13:00 | Gn. Sur. Pat 25 | Neuro Pat 38 | Gn. Sur. Pat 27 | OB/GYN Pat 47 | Ortho Pat 61 |
| 13:00-14:00 | Gn. Sur. Pat 25 | Neuro Pat 38 | Gn. Sur. Pat 27 | ENT Pat 13 | Podi Pat 65 |
| 14:00-15:00 | | Neuro Pat 38 | Gn. Sur. Pat 32 | ENT Pat 13 | Podi Pat 65 |
| 15:00-16:00 | Uro Pat 66 | | Gn. Sur. Pat 32 | Podi Pat 64 | Hand Pat 37 |
| 16:00-17:00 | Uro Pat 66 | | | Podi Pat 64 | |
| 17:00-18:00 | | | | | |

Examining Tables 30 - 33, we can see that the new schedule incorporates the corresponding changes for 19 elective patients. Note that the inclusion of a trauma patient is indicated in bold, italic, and underlined letters in the corresponding tables, whereas schedule changes involving starting time in the same day are indicated by underlined letters. Schedule changes involving change of operating room is indicated by the italic letters. Schedule changes in starting times involving different days is indicated by the bold letters.

Emergency trauma patient 70 is operated in operating room #1 at 08:00 am immediately upon arrival. Patients 41, 61, 65, and 37 are postponed one time period ahead in operating room #4. Since the model does not penalize postponing or preponing in the same day, the objective function value is not affected by those moves. In a similar vein, the starting time for the surgery of patient 21 is postponed from 08:00 am to 09:00 am. The start time for the surgery of neurology patient 38 is preponed from 14:00 pm to 11:00 am. The cardio-vascular patient 21 is rolled to the next period. This move increases the objective function value by $18,990. The General Surgery patient 14 who is scheduled to receive surgical operation in operating room #1 on Monday 9:00 am is scheduled in the same operating room next day at 10:00 am. Starting time for General Surgery patient 32 is preponed from 15:00 pm to 14:00 pm in operating room #4. The starting time for surgeries for patients 9, 42, and 17 in operating room #2 on Tuesday are preponed by one time period. In a similar fashion, the starting times for the patients 52, 58, and 18 are preponed by one time period. Similar in the case with patients 8, 43, and 16, General Surgery patient 29 who is scheduled at operating room #4 at 11:00 am is moved to operating room #1 at 8:00 am.

It is worthwhile to mention that one patient (i.e., patient 21) is rolled to the next period, and surgical operation for one patient (i.e., patient 14) is postponed by one day. Note that under the current scheme, there is no need for expanding the PACU units, and 3 PACU beds can accommodate the current elective schedule. In the solution, it might be seen that adjustments are also made to reduce the number of PACU beds that are required from 4 to 3. To cite an instance, in operating room #3, the starting times of patients 41, 61, 65, and 37 are postponed by one period. If the original schedule had been kept, then on Friday at 10:00 am, there would be a need for 4 PACU beds at a result of the conclusion of surgical operations of patients 41, 62, 24, and 4 respectively. In order to avoid this, the starting time of surgical operations for those patients are postponed by one time period.

## 4.5. General Discussion

For rescheduling elective patients, the solution approach of using a commercial solver to solve the MILP model and that of adopting the genetic algorithm approach generate fairly consistent results in terms of the solution quality. In the situations where the patient load is low, usually no additional cost function is introduced. In these cases, using the commercial solver to solve the mathematical programming model might provide the optimal solution in the fraction of seconds. When the patient load is high, it is very likely that additional cost figures are introduced by employing overtime in operating rooms or expanding the capacity in the PACU unit, the computation time for both of the approaches increased considerably, however, both of the approaches provide compatible results in terms of objective function value. For the second and third Scenarios, that has been described in Table 27, the genetic algorithm performs better, whereas in the first Scenario, GAMS is able to find the optimal solution in much shorter computation time. Although the genetic algorithm can always provide comparable or better solutions within the time window for all scenarios, it is still important to use both solution approaches in practice, and they should complement each other. This is because the commercial solver can save a significant amount of computation time for some scenarios, and any time saving will be appreciated by the hospital staff and patients. Also, in practice, both approaches can be set up to run successively during the given day, and the process might be automated such that the results of the previous run are fed as the input for the subsequent run. In other words, whenever the request for emergency treatment arrives, both approaches can be run again to provide timely decisions whether to admit the emergency patient or not without the need for making corresponding changes in the genetic algorithm and GAMS code.

# 5. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

## 5.1. Conclusion

In this dissertation, the problem of scheduling and rescheduling elective patients is considered. The problem encompasses two stages. In the first stage, the primary objective is to schedule the elective patients under various constraints. As previously discussed, these constraints are surgical team constraints, the downstream clinical unit constraints (PACU bed/equipment availability), and operating room constraints. These constraints are important for determining the schedule of the elective patients. The existing schedule should consider the resource constraints so that the surgical operations can be performed. If the resource constraints are not considered, then the operations in the operating rooms and downstream units are disrupted.

The possibilities of the expansion of the existing capacity are incorporated in the problem both in rescheduling and scheduling models. The expansion might be carried out in different fashions in different departments. For operating rooms, the overtime hours might be employed to increase available operating time for operating rooms, or additional surgical teams might be hired for performing the surgical operations and expanding the existing capacity. Moreover, the number of PACU beds might be increased in order to increase the maximum number of patients in the PACU units from recovering the effects of the anesthesia. Providing venues for increasing the existing capacity might help to handle the high elective patient load cases.

The scheduling model provides flexible approaches. In case of high patient load and/or scarce resources along with limited means for expanding the current capacity, some of the patients might be scheduled to the next planning period. This increases the flexibility in the scheduling practices by providing more options. The patients who agree to be scheduled for the upcoming period might be identified, and those patients might be associated with low cost figures for deferring them to the next time period. This will be especially helpful in the tight constraint/high patient load environment.

In the second stage, rescheduling of elective patients upon inclusion of the emergency patients is considered. In that regard, the capacity constraints, such as the number of surgical teams present to perform the surgical operations, overtime hours and the current available PACU units are taken into consideration. The objective is to minimize the disruption to the existing schedule while minimizing the amount of additional resources to accommodate the inclusion of emergency patients and reshuffling the elective patients. Another aspect is that the suggested rescheduling model can also be used as a decision making tool for assessing and improving the original elective surgery schedule with regard to resource usage. Even without the admission of emergency patients, it can be adopted as a standalone approach for evaluating and improving the current elective surgery schedule with some modifications to the mathematical programming model and the genetic algorithm code. In that case, the purpose of shifting the elective patients is to better utilize the available resources. With proper modifications, the model can be also enhanced to assess the current elective surgery schedule with regard to block scheduling practices that might be applied in other healthcare settings.

In the scenarios that are incorporated in the rescheduling of the elective patient, the arrival of a single emergency patient at the beginning of the scheduling period is assumed. In fact, the proposed model and solution approaches can also handle multiple emergency patients arriving at the same time. This might be especially important for the cases where more than one patient requiring emergency treatment might be brought to one hospital due to various reasons, such as traffic accidents and terrorist attacks involving the injury of multiple people. Based on the available capacity, the model will help make decisions on whether to admit those patients. Depending upon the case settings, some emergency patients might be turned down while others might be admitted. If the solution is to be obtained by the genetic algorithm approach, the length of chromosome structure (i.e., the fourth component) should be adjusted and the appropriate crossover operator as discussed in previous sections should be implemented accordingly.

It can be seen that using the rescheduling and scheduling approaches successively might provide a viable approach for providing a comprehensive view on the problem. The first stage is scheduling the patients, whereas the second stage involves the rescheduling phase, where the elective patients are shuffled when an emergency patient arrives. The rescheduling approach also honors the existing schedule thus penalizing delaying and preponing the patients to the next/previous days while considering the existing resources.

The models developed in the previous chapters consider the resource constraints such as the availability of the surgical teams, availability of PACU beds, and the working hours of operating rooms. The approach presented in the scheduling and rescheduling phases is a flexible one in which the block scheduling practices can be implemented implicitly by limiting the availability of the surgical teams to the specified time periods. These time periods might constitute the blocks for the surgery teams and surgical groups and can be used towards the creation of the master surgical schedule in a cyclic pattern. It might be important for applying the cyclic master surgical schedule in terms of leveling workforce requirements for surgical and support teams. The surgeons and the support staff might prefer working specific day/time of the week, and by collecting the information of the preferences, and by inputting the surgical team preferences in terms of the availability of those teams, the model might be used for finding the optimum schedules. Additionally, the number of the changeovers from one type of surgical operation to another type might be minimized by adjusting the availability of the surgical teams. To cite an instance, rather than making a particular team available in disjoint time periods, the availability of a team might be arranged in such a way that the team is available in the consecutive time periods. This serves for two purposes. The first purpose is that it will increase the convenience of the workforce by allowing them to work for consecutive time periods rather than working on disjoint time intervals. The second purpose will be reducing the number of changeovers from one surgical operation to another one. This will reduce the time for preparing operating rooms for the next surgical operations. It will be logical to assume that the preparation time of an operating room for the same type of surgery is generally less than the preparation time of the same operating room for a different type of surgery. In that regard, the minimization of changeovers in a particular operating room increases the efficiency of the operating room by increasing the total time actually spent for performing surgical operations and reducing the time spent for preparing operating rooms for subsequent operations

In general, we see that both for rescheduling and scheduling phases, the genetic algorithm based approaches provide compatible solution in terms of the objective function values. Especially, in the rescheduling case, where the computation time is important in terms of the giving decision whether accepting and rejecting the emergency patient, rule of thumb might be developed for employing either genetic algorithm or mathematical modeling based approach separately or using them together. A composite index representing the problem size might be developed for this purpose to determine which approach to be used.

For the genetic algorithm used in scheduling and rescheduling elective patients, two different approaches based on the representation of the problem are developed. In scheduling the elective patients, the chromosome representation is based on the starting time of the surgical operation for the patients, whereas in rescheduling the elective patient, the solution is mainly represented by the sequence of the patients operated in operating rooms. Representing the sequence of the patient who will undergo surgical operation work better for the rescheduling based on the fact that elective patient schedule is always included in the solution pool and the original elective patient schedule is always selected for the crossover and subsequent operations. This practice provides a suitable starting point for forming the elective patient schedule where the patients can be reshuffled to present the best strategy for forming the new schedule. Additionally, most of the cases, the new optimal or near-optimal solution is the variation of the original elective solution based on the sequence of the patients, and therefore adopting the sequence based representation provides better results.

However, for scheduling the elective patients, since there is no initial feasible solution to start with, representing the solution in terms of the starting time provides better results by searching the solution space in a higher resolution. Therefore, for scheduling of the elective patients, the chromosome representation based on the starting time of the surgical operations is adopted.

## 5.2. Policy Implications and General Considerations

In this section, we discuss different policy implications and the general consideration regarding the proposed approaches as follows.

Some states necessitate that it is not possible to turn down the emergency patients. The mathematical programming model can handle the no turning down rule by either employing high cost figures associated with the turning down the emergency patients in the objective function or imposing the corresponding constraint that the emergency patients should be operated in that particular time period of arrival.

The prioritization of patients might be also incorporated in the mathematical model. To cite an instance, for the rescheduling of elective patients upon arrival of the emergency patient, the model tries to minimize the disruption in the current elective schedule by assigning the monetary cost figures for shifting the elective patient from one particular time slot to another one. Additionally by imposing additional constraints, as previously stated, some of the patients (i.e., children, outpatient, or patients with some travel restrictions) might be operated in the earlier hours during the day.

Both the scheduling and rescheduling models aim to improve the patient access to the surgical operations by increasing throughput. Throughput in this problem setting might be loosely defined as the number of patients who undergo surgical operations. In order to improve the throughput, cost figures are associated for the deferring the patient to the next planning period, therefore given the level of the resources, during the planning horizon, the models aim to improve the patient access by minimizing number of patients who do not undergo surgical operations for that planning period.

Both scheduling and rescheduling models use the duration of stay in the PACU units as the input parameter on patient basis. There might be the case that due to the complications and other considerations, the patient might need to stay longer in the PACU unit, or might be transferred to ICU instead of the PACU unit. Longer stays in the PACU unit might be handled with changing the input parameters for duration of the stay in the PACU unit. Since both models consider making corresponding decisions at the patient level and treat patients on individual basis, duration of stay in the corresponding PACU unit might be adjusted for each of the patient/surgical operations on case by case basis. For transfers to the ICU unit, new downstream clinical units should be defined, and incorporated in the model.

Additional constraints might be imposed to provide the suitable time slots where no elective patient is scheduled for that particular time slot. If an emergency patient arrives, without the need for the rescheduling of the elective patients, the emergency patient might be operated in that particular time period. This type of policy brings an improvement in operating room planning. The number of available time slots for operating emergency patients might be adjusted based on the historical data. If during the week, for some time periods, more emergency patients are likely to arrive, the time slots that are available for emergency surgical operations might be increased as compared to the other time periods. That will help to create the emergency patient friendly schedules that the possibility of rescheduling elective patients is minimized.

## 5.3. Future Research Directions

Based on the thesis research results, we point out the future research directions as follows,

1. Incorporating constraints in the rescheduling mathematical model for limiting the length of the notice provided to the patients regarding the change of the schedule.
2. Developing the mathematical models featuring continuous distributions rather than the discrete distributions for governing the duration of the surgical operations for the scheduling model. The same approach might also be followed for the rescheduling model as well. A mathematical programming model featuring stochasticity in duration of surgical operation for rescheduling elective patients might be developed.
3. Incorporating additional constraints for reducing the number of changes of the surgical operation performed in an operating room in a given day. In order to increase the effectiveness of the system, the constraints limiting the number of surgical operations might also be incorporated in the mathematical model both for scheduling and rescheduling mathematical programming models.
4. Another consideration is incorporating the prioritization of patients in the scheduling model. Some patients might prefer to be operated at specific time of the day. To cite an instance, children under certain age are usually given the priority for the earlier time slots within the given day. In an outpatient setting, patients who need to travel long distances back to their homes are also given priority during the day for the earlier time slots. These can be incorporated in the

mathematical model, since the model identifies admissions based on the individual patient level.

5. Developing other heuristic approaches. These include but are not limited to; ant colony optimization, Tabu search, simulated annealing, and other metaheuristic approaches.
6. Developing repair schemes for improving the overall solution feasibility of the solution throughout the generations for the genetic algorithm implementation. This might be especially a viable approach for the cases with the high elective patient load and large variations in terms of the surgical duration. The corresponding move algorithm developed in scheduling and rescheduling sections might be further improved to provide better solution in a limited amount of time.
7. Using Bender's decomposition approach. That approach is usually employed to solve the class of optimization problems that possess the specific structure. In that regard, the structure of the problem is exploited. Bender decomposition is also extensively used in stochastic programming models (Infanger, 1994; Nielsen and Zenios, 1997). Additional detail is provided in Appendix C.
8. Implementing Lagrangian relaxation methods to increase the performance of mathematical modeling based solutions might be a viable approach.

## REFERENCES


Adan IJBF, and Vissers JMH, 2002. Patient mix optimisation in hospital admission planning: A case study. *International Journal of Operations and Productions Management* 22(4): 445-61.

Adan J, Bekkers J, Dellaert N, Vissers J, and Yu X, 2009. Patient mix optimization and stochastic resource requirements: a case study in cardiothoracic surgery planning. *Health Care Management Science* 12:129-41.

Agnetis A, Coppi A, Crsini M, Dellino G, Meloni C, and Pranzo M, 2012. Long term evaluation of operating theater planning policies. *Operations Research for Healthcare* 1(4):95-104.

Al-Dulaimi BF, and Ali HA, 2008. Enhanced traveling salesman problem solving by genetic algorithm technique (TSPGA). *World Academy of Science Engineering and Technology* 38:296-302.

Alon E, and Schüpfer G, 1999. Operating room management. *Der Anaesthesist* 48(10): 689-97.

Augusto V, Xie X, and Perdomo V, 2008. Operating theatre scheduling using Lagrangian relaxation. *European Journal of Industrial Eng* 2(2):172–89.

Augusto V, Xie X, and Perdomo V, 2010. Operating theatre scheduling with patient recovery in both operating rooms and recovery beds. *Computers and Industrial Engineering* 58(2):231–38.

Baez AA, Lane PL, Sorondo B, and Giraldez EM, 2006. Predictive effect of out-of-hospital time in outcomes of severely injured young adult and elderly patients. *Prehospital and Disaster Medicine*. 21(6):427-30.

Ballard S, and Kuhl M. 2006. The use of simulation to determine maximum capacity in the surgical suite operating room. *Proc. of Winter Simulation Conference,* Monterey, California, USA, December 3-6, 2006.

Barnhart C, Johnson EL, Nemhauser GL, Savelsbergh MWP, and Vance PH, 1998. Branch and price: Column generation for solving huge integer programs. *Operations Research*, 46, 316-29.

Barricelli NA, 1957. Symbiogenetic evolution processes realized by artificial methods. *Methodos*: 143–82.

Beliën J, and Demeulemeester E, 2007. Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research* 176(2):1185–204.

Beliën J, Demeulemeester E, and Cardoen B, 2009. A Decision Support System for cyclic master surgery scheduling with multiple scheduling. *Journal of Scheduling* 12(2):147-161.

Bhattacharya T, Vrahas MS, Morrison SM, Kim E, Wiklund RA, Smith RM, and Rubash HE, 2006. The value of the dedicated orthopaedic trauma operating room. *The Journal of TRAUMA Injury, Infection, and Critical Care* 60 (6):1336-41.

Blake JT, Dexter F, and Donald J, 2002. Operating room manager's use of integer programming for assigning block time for surgical groups: A case study. *Anesthesia & Analgesia* 94:143-148.

Blake JT, and Donald J, 2002. Mount Sinai uses integer programming to allocate operating room time. *Interfaces* 32(2):63-73.

Blake JT, and Carter MW, 1997. Surgical process scheduling: A structured review. *Journal of Health Systems* 5(3):17-30.

Boldy D, 1976. A review of the application of mathematical programming to tactical and strategic health social services problem. *Operational Research Quarterly* 27(2):439-48.

Bowers J, and Mould G, 2004. Managing uncertainty in orthopaedic trauma theatres. *European Journal of Operational Research* 154:599-608.

Buckles, BP, Petry PE, Kuester, RI, 1990. Schema survival rates and heuristic search in genetic algorithms. *In: Tools for Artificial Intelligence Proceedings of the 2nd International IEEE Conference*, 86-91.

Cardoen B, Demeulemeester E, and Beliën J, 2009. Optimizing a multiple objective surgical case sequencing problems. *International Journal of Production Economics*

119(2):354-66.

Cardoen B and Demeulemeester E, 2008. Capacity of clinical pathways--a strategic multi-level evaluation tool. *Journal of Medical Systems*, 32(6):443-52.

Cardoen, B, Demeulemeester E., and Beliën J, 2010. Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3):921-32..

Cowie B, and Corcoran P, 2012. Postanesthesia care unit discharge delay for nonclinical reasons. *Journal of PeriAnesthesia Nursing*, 27(6):393-8.

Crosby JL, 1973. Computer simulation in genetics. London: John Wiley & Sons.

De Jong KA, 1992. Are genetic algorithms function optimizers? *Proceedings of the Second International Conference on Parallel Problem Solving from Nature*.

Denton B, Miller A, Balasubramanian H, and Huschka T, 2010. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operational Research*, 58(4):802–16.

Denton B, Viapiano J, and Vogl A, 2007. Optimization of surgery sequencing and scheduling decisions under uncertainty. *Healthcare Management Science*, 10(1): 13-24.

Denton B, Rahman, AS, Nelson, H, and Bailey AC, 2006. Simulation of a multiple operating room surgical suite. *In Proceedings of the 2006 Winter Simulation Conference.*

Dexter F, 2000. A strategy to decide whether to move the last case of the day in operating room to other operating room to decrease overtime labor costs. *Anesthesia Analgesics* 91:925-928.

Dexter F, Epstein R, Lee J, and Ledolter J, 2009. Automatic updating of times remaining in surgical cases using Bayesian analysis of historical case duration data and “instant messaging” updates from anesthesia providers. *Anesthesia and Analgesia* 108:929–40.

Dexter F, and Ledolter J, 2005. Bayesian prediction bounds and comparisons of operating room times even for procedures with few or no historic data. *Anesthesiology* 103:1259–67.

Dexter F, Macario A, and Ledolter J, 2007. Identification of systematic underestimation (bias) of case durations during case scheduling would not markedly reduce overutilized operating room time. *Journal of Clinical Anesthesiology* 19:198–203.

Dexter F, Macario A, Qian F, and Traub R, 1999a. Forecasting surgical groups' total hours of elective cases for allocation of block time. Application of time series analysis to operating room management. *Anesthesiology* 91:1501–8.

Dexter F., Macario, A., Traub RD, 1999b. Which algorithm for scheduling add-on elective cases to maximizes operating room utilization? Use of bin packing algorithms and fuzzy constraints in operating room management. *Anesthesiology* 91(5):1491–1500.

Dexter F, and Ledolter J, 2003. Managing risk and expected financial return from selective expansion of operating room capacity: Mean-variance analysis of a hospital's portfolio of surgeons. *Anesthesia and Analgesia* 97:190-5.

Dexter F, Blake JT, Penning, DH, and Lubarsky DA, 2002. Calculating a potential increase in hospital margin for elective surgery by changing operating room time allocations or increasing nursing staying to permit completion of more cases: A case study. *Anesthesia and Analgesia* 94:138-42.

Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA, 1999c. An operating room scheduling strategy to maximize the use of operating room block time: Computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. *Anesthesia and Analgesia* 89(1):7–20.

Encyclopedia of Surgery, 2012. Emergency Surgery. URL: http://www.surgeryencyclopedia.com/Ce-Fi/Emergency-Surgery.html. Accessed on January 20, 2012.

Etzioni DA, Liu JH, Maggard MA, and Ko CY, 2003. The aging population and its impact on the surgery workforce. *Annals of Surgery* 238(2):170-7.

Everett JE, 2002. A decision support simulation model for the management of an elective surgery waiting system. *Health Care Management Science* 5:89-95.

Fei H, Meskens N, and Chu C, 2006. An operating theatre planning and scheduling problem in the case of a block scheduling strategy. *In Proc. of Winter Simulation Conference* 1:422–8.

Fei H, Meskens N, and Chu C, 2010. A planning and scheduling problem for an operating theatre using an open scheduling strategy. *Computers and Industrial Engineering* 58:221–30.

Fei H, Chu C, Meskens N, 2009a. Solving tactical operating room planning problem by a column-generation based heuristic procedure with four criteria. *Annals of Operations Research* 166:91-108.

Fei H, Neskens N, Combes C, and Chu C, 2009b. The endoscopy scheduling problem: A Case study with two specialised operating rooms. *International Journal of Production Economics* 120(2):452-62.

Fraser A, and Burnell D, 1970. Computer Models in Genetics. New York: McGraw-Hill.

Gabel R, Kulli J, Stephen LB, Spratt D, and Ward D, 1999. Operating Room Management. London, Butterworth-Heinemann.

Glouberman S, and Mintzberg H, 2001. Managing the care of health and the cure of disease. Part 1: Differentiation. *Health Care Management Review*, 26:56-69.

Goldberg DE, 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Boston, MA, USA. Addison-Wesley Publishing Company Inc.

Gordon T, Paul S, Lyles A, and Fountain J, 1998. Surgical unit time utilization review: resource utilization and management implications. *Journal of Medical Systems* 12(3): 169-179.

Grefenstette JJ, 1986. Optimization of control parameters for genetic algorithms. *IEEE Transactions on Systems Man and Cybernetics* 16(1):122-8.

Guerriero F, and Guido R, 2011. Operational research in the management of the operating theatre: A survey. *Health Care Management Science*, 14:89-114.

Guinet A, and Chaabane S, 2003. Operating theatre planning. *International Journal of Production Economics* 85(1):69–81.

Hackett P, 1995. A comparison of selection methods based on the performance of a genetic program applied to the cart-pole problem. Bachelor of Science Thesis. Griffith University, Gold Coast Campus, Queensland.

Harders, M, Mark M, Weight S, Sidhu T, Krupka D, and Sandberg W, 2009. Increasing operating room throughput via parallel processing may not require extra resources. *Anesthiology* 110(2):435.

Harper PR, 2002. A framework for operational modeling of hospital resources. *Health Care Management Science* 5:165-73.

Health Care Financial Management Association, 2005. Achieving operating room integrity through process integration, Technical Report.

Holland J, 1975. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. Cambridge, London, United Kingdom: The MIT Press.

Hsu V, de Matta R, and Lee CY, 2003. Scheduling patients in an ambulatory surgical center. *Naval Research Logistics* 50:218–38.

Infanger G, 1994. Planning under uncertainty-solving large scale stochastic linear programs. Boyd & Fraser, 1994.

Jebali, A, Alouane, ABH, and Ladet P, 2006. Operating room scheduling. *International Journal of Production Economics*, 99(1-2): 52-62.

Joines J, and Houck CR, 1994. On the use of non-stationary penalty functions to solve nonlinear constrained optimization problems with GA's, *Proceedings of IEEE, International Conference on Evolutionary Computation* 579-584.

Kalvelagen, E. Benders decomposition with GAMS. URL: http://www.amsterdamoptimization.com/pdf/benders.pdf, accessed March 21, 2012.

Kennedy M, 1992. Bin-packing, knapsack, and chance constrained approaches to operating room scheduling. Ph.D. thesis, Rensselaer Polytechnic Institute School of Engineering Department of Decision Sciences Engineering.

Kuo P, Schroeder R, Mahaffey S, and Bollinger R, 2003. Optimization of operating room allocation using linear programming techniques. *Journal of the American College of Surgeons* 197(6):889–95.

Lamiri M, Dreo J, and Xie X, 2007. Operating room planning with random surgery times. *In: Proceedings of the 3rd Annual IEEE Conference on Automation Science and Engineering*, Sept 22–25. Scottsdale, AZ, USA

Lamiri M, Grimaud F, and Xie X, 2009. Optimization methods for a stochastic surgery planning problem. *International Journal of Production Economics* 120(2):400-10.

Lamiri M, Xie X, Dolgui A, and Grimaud F, 2008. A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research* 185(3):1026-37.

Lubbecke ME, and Desrosiers J, 2005. Selected topics in column generation. *Operations Research* 53:1007–23.

Macario A, Vitez TS, Dunn B, and McDonald T, 1995. Where are the costs in perioperative care? Analysis of hospital costs and charges for inpatient surgical care. *Anesthiology* 83(6):1138-44.

Magerlein, JM and Martin JB, 1978. Surgical demand scheduling: A review. *Health Services Research* 13:418-33.

Marcon E, Kharraja S, and Simonnet G, 2003. The operating theatre planning by the follow-up of the risk of no realization. *International Journal of Production Economics* 85:83-90.

Marcon E, and Dexter JE., 2006. Impact of surgical sequencing on post anesthesia care unit staffing. *Health Care Management Science,* 9: 87–98.

Martin RK, 1999. Large scale linear and integer optimization: A unified approach. Kluwer, 1999.

Masursky D, Dexter F, O’Leary C, Applegeet C, Nussmeier N, 2008. Long-term forecasting of anesthesia workload in operating rooms from changes in a hospital’s local population can be inaccurate. *Anesthesia and Analgesia* 106:1223–31.

McGowan JE, Truwit JD, Cipriano P, Howell RR, VanBree M, Garson AJR., and Hanks JB, 2007. Operating room efficiency and hospital capacity: Factors affecting operating room use during maximum hospital census. Journal of the American College of Surgeons 204(5):865-71.

Min D, and Yih, Y. 2010. An elective surgery scheduling problem considering patient priority. *Computers and Operation Research*, 37:1091–9.

Mitsuo M, and Cheng R, 1997. Genetic algorithm and engineering design. New York: John Wiley &Sons.

Mulholland W, Abrahamse P, and Bahl V, 2005. Linear programming to optimize performance in a department for surgery. *Journal of American College of Surgeons,* 200(6):861-868.

Nielsen SS, and Zenios SA, 1997. Scalable parallel benders decomposition for stochastic linear programming. *Parallel Computing* 23:1069-88.

Niu Q, Peng Q, ElMekkawy T, Tan YY, Bryant H, and Bernaerdt, 2007. Performance analysis of the operating room using simulation. *In: Proceedings of the 2007 CDEN and CCEE Conference*.

Ogulata SN, and Erol R, 2003. A hierarchical multiple criteria mathematical programming approach for scheduling general surgery operations in large hospitals. *Journal of Medical Systems* 27(3):259-70.

Olejarz, D, 2009. ICU capacity boosted to largest in Michigan. Henry Ford News, Henry Ford Health System, URL: http://www.henryfordhealth.org/body.cfm?id=46335&action=detail&ref=986, accessed December 21, 2011.

Park KW, and Dickerson C, 2009. Can efficient supply management in the operating room save millions? *Current Opinion in Anesthesiology* 22:242-8.

Patterson P, 1996. What makes a well-oiled scheduling system? *OR Manager* 12(9):19-23.

Perdomo V, Augusto V, and Xie X, 2006. Operating theatre scheduling using Lagrangian relaxation. *In: Service systems and service management international conference*, 2006.

Persson M, and Persson J, 2009. Health economic modeling to support surgery management at a Swedish hospital. *Omega* 37(4):853–86.

Persson M, and Persson, JA, 2007. Analysing management policies for operating room planning using simulation. Working paper, Blekinge Institute of Technology, Sweden, 2007.

Pham, D-N, and Klinkert A, 2008. Surgical case scheduling as a generalized job shop scheduling problem. *European Journal of Operational Research* 185(3):1011-25.

Pierskalla WP, and Brailer DJ, 1994. Applications of operations research in health care delivery, *In: Proc of Operations Research and Public Sector*, 469-505, North-Holland, 1994.

Plunkett Research, 2013. Introduction to the health care industry. URL: http://www.plunkettresearch.com/health-care-medical-market-research/industry-and-business-data, accessed March 23, 2013.

Przasnyski Z, 1986. Operating room scheduling: A literature review. *AORN Journal* 44(1): 67-79, 1986.

Rinde A, and Blakely T, 1974. OR resource utilization. Technical Report. Chicago Hospital Council, Chicago.

Roland B, Di Martinelli C, Riane F, and Pochet Y, 2010. Scheduling an operating theatre under human resource constraints. *Computers and Industrial Engineering* 58(2):212–220.

Roland B, Di Martinelly C, and Riane F, 2006. Operating theatre optimization: a resource-constrained based solving approach. *In: International conference on service systems* *and service management* 1:443–448.

Rubicite Interactive, 2012. Crossover technique-PMX, URL: http://www.rubicite.com/Tutorials/GeneticAlgorithms/CrossoverOperators/PMXCrossoverOperator.aspx, accessed March 12, 2013.

Sacra JC, and Martinez R, 2009. Trauma Systems, in: W.R. Roush (Ed.), *Principles of EMR systems*, second edn 41, American College of Emergency Physicians, Texas,

Santibanez P, Begen M, and Atkins D, 2007. Surgical block scheduling in a system of hospitals: An application to resource and wait list management in a British Columbia health authority. *Health Care Management Science* 10:269-82.

Sivanandam, SN, and Deepa SN, 2008. Introduction to Genetic Algorithms, first edn., Heidelberg: Springer-Verlag.

Smith-Daniels VL, Schweikhart SB, and Smith-Daniels DE, 1988. Capacity management in healthcare services: Review and future research directions. *Decision Sciences* 19:889-919.

Stepaniak P, Heij C, and de Vries G, 2010. Modeling and prediction of surgical procedure times. *Statistica Neerlandica* 64:1–18.

Taheri PA, Maggio PM, Dougherty J, Neil C, Fetyko S, Harkins, DR, and Butz DA, 2007. Trauma center downstream revenue: the impact of incremental patients within a health system. *Journal of Trauma* 62(3):615-9.

Testi A, Tanfani E, and Torre G, 2007. A three-phase approach for operating theatre schedules. *Health Care Management Science* 10(2):163-72.

The Henry J. Kaiser Family Foundation, 2006. Snapshots: Comparing projected growth in health care expenditures and the economy. URL: http://www.kff.org/insurance/snapshot/chcm050206oth2.cfm, accessed March 10, 2013.

Unibased System Architecture, 2010. Block scheduling in perioperative environment. URL: http://www.unibased.com/blockschedulingor.html, accessed January 17, 2012.

United States Department of Labor, 2007. Health care industry information. URL:

http://www.doleta.gov/BRG/Indprof/Health.cfm, accessed January 20,2013.

Van der Lans M, Hans E, Hurink JL, Wullink G, van Houdenhoven M, and Kazemier G, 2006. Anticipating urgent surgery in operating room departments. Working Paper, University of Twente, The Netherlands.

Van Oostrum J, Parlevliet T, Wagelmans A, and Kazemier G, 2008b. A method for clustering surgical cases to allow master surgical scheduling. In: Conference Proceedings, ORAHS, 2008b.

Van Oostrum J, van Houdenhoven M, Hurink J, Hans E., Wullink G, and Kazemier G, 2008a. A master surgical scheduling approach for cyclical scheduling in operating room departments. *OR Spectrum* 30(2):355-74.

VanBerkel P, and Blake J, 2007. A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. *Health Care Management Science* 10(4):373–385.

Van Houdenhoven M, van Oostrum J, Wullink G, Hans E, Hurink JL, Bakker J, and Kazemier G, 2008. Fewer intensive care unit refusals and a higher capacity utilization by using a cyclic surgical case schedule. *Journal of Critical Care* 23(2):222-6

Vissers J, Bertrand J, and De Vries G, 2001. A framework for production control in health care organizations. *Production Planning and Control* 12(6):591–604.

Vissers JMH, Adan IJBF, and Bekkers JA, 2005. Patient mix optimization in tactical cardiothoracic surgery planning: A case study. *IMA Journal of Management Mathematics* 16(3):281-304.

Vogel TR, Dombrovskiy, VY, and Lowry SF, 2010. In-hospital delay of elective surgery for high volume procedures: The impact on infectious complications. *Journal of the American College of Surgeons* 211(6):784-90.

Wachtel R, and Dexter F, 2008. Tactical increases in operating room block time for capacity planning should not be based on utilization. *Anesthesia & Analgesia* 106(1):215–22.

Wilde, ET, 2012. Do emergency medical system response times matter for health outcomes? Working Paper, Columbia University.

Wullink G., van Houdenhoven M, Hans, EW, van Oostrum JM, van der Lans M, and Kazemier G, 2007. Closing emergency operating rooms improves efficiency. *Journal of Medical Systems* 31.

Yang Y, Sullivan KM, Wang PP, and Naidu KD, 2000. Applications of computer simulation in medical scheduling. In: *Proceedings of the Joint Conference on Information Sciences*..

Zhang B, Murali P, Dessouky M, and Belson D, 2009. A mixed integer programming approach for allocating operating room capacity. *Journal of the Operational Research Society*, 60:663 –73.

## APPENDIX A. GAMS CODE FOR MATHEMATICAL PROGRAMMING MODEL FOR SCHEDULING

```
sets
t TIME SLOT INDEX 56 REFERS TO THE OUTSIDE THE PLANNING CYCLE 5*11=33 /1*55/
j   THE DIAGNOSIS TYPE INDEX 1 /1*10/
i   THE PATIENT INDEX  /1*70/
h   operating time states /1*3/
ww    scenarios


d DAY INDEX /1*5/ ;

alias (t,tprime);
alias (h,h2,h3,h7,h10);
set w /1*24/;
set current(w);
current('1')=yes;



parameter probxx;
parameter p(w);

Table  lambda(j,h)  probability of surgiucal operation of type  j  lasts period of  h

        1               2               3
1       1
2       0.6             0.4
3       0.2             0.5             0.3
4       1
5       1
6       1
7       0.8             0.2
8       1
9       1
10      0.6             0.4                     ;



table G(j,h) operating time of the surgery having diagnosis type j
          1   2    3
      1  4
      2  1   2
      3  1   2    3
      4  1
```

```
    5  4
    6  2
    7  1   2
    8  1
    9  2
    10 1   2 ;



loop((h,h2,h3,h7,h10),
   probxx =
          lambda('2',h2) *
           lambda('4',h3) *
          lambda('7',h7)  *
          lambda('10',h10)


          ;

      O('1',current) = G('1','1');
      O('2',current) = G('2',h2);
      O('3',current) = G('3',h3);
      O('4',current) = G('4','1');
      O('5',current) = G('5','1');
      O('6',current) = G('6','1');
      O('7',current) = G('7',h7);
      O('8',current) = G('8','1');
      O('9',current) = G('9','1');
      O('10',current) = G('10',h10);
      p(current) = probxx;
*     s(ss) = s(ss-1)          current(w+1)$current(w) = yes;
      current(w-1)$current(w) = no ) ;


parameter su(j) the time of stay in the intensive care unit related with diagnosis type j;
su(j) = 1;

parameter B(d) number of current operating hours for the operating room;
b(d) = 48;

parameter probxx;

SCALAR
bicu the regular capacity of the number of beds in the ICU /3/
cicu THE COST OF ADDING ADDITIONAL BED IN icu DURING THE PLANNING HORIZON
/20000/
cor THE HOURLY COST OF OPERATING THE OPERATING ROOM  /1500/
uicu THE UPPER LIMIT ON THE NUMBER OF ADDITIONAL BEDS THAT CAN BE PLACED
IN icu /1/
```

uor the upper limiot on the number of available overtime hours that can put in ICU 4*(16:00-18:00)/4/
N nUMBER OF OPERATING ROOM /4/
Over of overtime operating room /2/
surg /4156/;

table tau(t,j) NUMBER OF SURGERY TEAMS THAT ARE AVAILABLE FOR PERFORMING THE SURGERY TYPE J AT TIME PERIOD T

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 8 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 9 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 10 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 11 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 12 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 13 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 14 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 15 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 16 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 17 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 18 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 19 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 20 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 21 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 22 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 23 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 24 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 25 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 26 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 27 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 28 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 29 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 30 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 31 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 32 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 33 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 34 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 35 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 36 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 37 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 38 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 39 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 40 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| 41 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |

```
42      0   1   0   1   0   1   1   1   1   0
43      0   1   0   1   0   1   1   1   1   0
44      0   1   0   1   0   1   1   1   1   0
45      0   1   0   1   0   1   1   1   1   1
46      0   1   0   1   0   1   1   1   1   1
47      0   1   0   1   0   1   1   1   1   1
48      0   1   0   1   0   1   1   1   1   1
49      0   1   0   1   0   1   1   1   1   1
50      0   1   0   1   0   1   1   1   1   1
51      0   1   1   1   0   1   1   1   1   1
52      0   1   1   1   0   1   1   1   1   1
53      0   1   1   1   0   1   1   1   1   1
54      0   1   1   1   0   1   1   1   1   1
55      0   1   1   1   0   1   1   1   1   1
;
table m(i,j)    WHETHER THE PATIENT I HAS THE DIAGNOSIS TYPE OF J
        1   2   3   4   5   6   7   8   9   10
1       1   0   0   0   0   0   0   0   0   0
2       1   0   0   0   0   0   0   0   0   0
3       1   0   0   0   0   0   0   0   0   0
4       1   0   0   0   0   0   0   0   0   0
5       0   1   0   0   0   0   0   0   0   0
6       0   1   0   0   0   0   0   0   0   0
7       0   1   0   0   0   0   0   0   0   0
8       0   1   0   0   0   0   0   0   0   0
9       0   1   0   0   0   0   0   0   0   0
10      0   1   0   0   0   0   0   0   0   0
11      0   1   0   0   0   0   0   0   0   0
12      0   1   0   0   0   0   0   0   0   0
13      0   1   0   0   0   0   0   0   0   0
14      0   1   0   0   0   0   0   0   0   0
15      0   0   1   0   0   0   0   0   0   0
16      0   0   1   0   0   0   0   0   0   0
17      0   0   1   0   0   0   0   0   0   0
18      0   0   1   0   0   0   0   0   0   0
19      0   0   1   0   0   0   0   0   0   0
20      0   0   1   0   0   0   0   0   0   0
21      0   0   1   0   0   0   0   0   0   0
22      0   0   1   0   0   0   0   0   0   0
23      0   0   1   0   0   0   0   0   0   0
24      0   0   1   0   0   0   0   0   0   0
25      0   0   1   0   0   0   0   0   0   0
26      0   0   1   0   0   0   0   0   0   0
27      0   0   1   0   0   0   0   0   0   0
28      0   0   1   0   0   0   0   0   0   0
29      0   0   1   0   0   0   0   0   0   0
30      0   0   1   0   0   0   0   0   0   0
31      0   0   1   0   0   0   0   0   0   0
32      0   0   1   0   0   0   0   0   0   0
33      0   0   1   0   0   0   0   0   0   0
34      0   0   1   0   0   0   0   0   0   0
```

```
35      0   0   1   0   0   0   0   0   0   0
36      0   0   0   1   0   0   0   0   0   0
37      0   0   0   1   0   0   0   0   0   0
38      0   0   0   1   0   0   0   0   0   0
39      0   0   0   0   1   0   0   0   0   0
40      0   0   0   0   1   0   0   0   0   0
41      0   0   0   0   1   0   0   0   0   0
42      0   0   0   0   0   1   0   0   0   0
43      0   0   0   0   0   1   0   0   0   0
44      0   0   0   0   0   1   0   0   0   0
45      0   0   0   0   0   1   0   0   0   0
46      0   0   0   0   0   1   0   0   0   0
47      0   0   0   0   0   1   0   0   0   0
48      0   0   0   0   0   1   0   0   0   0
49      0   0   0   0   0   0   1   0   0   0
50      0   0   0   0   0   0   1   0   0   0
51      0   0   0   0   0   0   1   0   0   0
52      0   0   0   0   0   0   0   1   0   0
53      0   0   0   0   0   0   0   1   0   0
54      0   0   0   0   0   0   0   1   0   0
55      0   0   0   0   0   0   0   1   0   0
56      0   0   0   0   0   0   0   1   0   0
57      0   0   0   0   0   0   0   1   0   0
58      0   0   0   0   0   0   0   1   0   0
59      0   0   0   0   0   0   0   1   0   0
60      0   0   0   0   0   0   0   1   0   0
61      0   0   0   0   0   0   0   1   0   0
62      0   0   0   0   0   0   0   1   0   0
63      0   0   0   0   0   0   0   0   1   0
64      0   0   0   0   0   0   0   0   1   0
65      0   0   0   0   0   0   0   0   1   0
66      0   0   0   0   0   0   0   0   1   0
67      0   0   0   0   0   0   0   0   0   1
68      0   0   0   0   0   0   0   0   0   1
69      0   0   0   0   0   0   0   0   0   1
70      0   0   0   0   0   0   0   0   0   1     ;


binary variables x(i,t) whether the operation is scheduled at time period t for patient i
                 a(i,t,w) whether the operation is continued at time period t for patient i
                 s(i,t,w) whether patient i occupies the bed at ICU at time period t
                 u(i) if the patient is rolled to another horizon
                 z(t,j) number of additional surgical teams hired
integer variables
                  OOR(d) overtime utilization of operating room at day d)
                   OICU number of extra beds placed in ICU unit

variable mun


equations
```

```
*equation1(d,w)  this constraint determines the total utilization of the operating room
equation2(d,w) this constraint determines the overtime utilization of the operating oom equation3(w)
the constraints satisfying that it is is not possible to operate the operating room equation4(i)     the
elective patient should be either performed in planning horizon should be rooled to the next horizon

equation5(t,d,w)     the number of ongoing operations cannot be more than the number of operating
room
equation6(j,t,w) the constraint that determines the occupation of beds in ICU based on the x(i t)
values
equation7(i,j,t,tprime,w)
equation8(i,j,t,tprime,w) the constraint that is used for connecting a(i t) with x(i t)

equation9(t,w)   the number of occupied nbeds in ICU at an instant cannot be more than the number
of regular beds+number of extra beds for ICU
equation10       number of extra beds cannot be more than the upper limit on the number of beds for
ICU
equation11(d)    overtime utilization cannot be more than the upper limit for Operating room
equation12(t,d,w)

object  ;

*equation1(d,w).. sum((i,t)$((ord(t)ge (1+((ord(d)-1)*11)) and (ord(t)le 8+(ord(d)-1)*11))),a(i,t,w))
=l= B(d);
equation2(d,w).. sum((i,t)$((ord(t)ge (9+((ord(d)-1)*11)) and (ord(t)le 10+(ord(d)-1)*11))),a(i,t,w))
=e= OOR(d);
equation3(w).. sum((i,t)$((mod(ord(t)-1,11) ge 10)),a(i,t,w))=e= 0;
equation4(i).. sum(t,x(i,t))+u(i)=e= 1;

equation5(t,d,w)$((ord(t)ge (1+((ord(d)-1)*11)) and (ord(t)le 8+(ord(d)-
1)*11)))..sum(i,a(i,t,w))=l=N;
equation12(t,d,w)$((ord(t)ge (9+((ord(d)-1)*11)) and (ord(t)le 10+(ord(d)-1)*11)))..
sum(i,a(i,t,w))=l=over;

equation6(j,t,w)..sum(i, a(i,t,w)*m(i,j))=l=tau(t,j)+z(t,j);
equation7(i,j,t,tprime,w)$((ord(tprime) ge (ord(t)+O(j,w))) and (ord(tprime) le (ord(t)+O(j,w)+Su(j)-
1)) ).. s(i,tprime,w) =g= x(i,t)*m(i,j);
equation8(i,j,t,tprime,w)$((ord(tprime) ge (ord(t))) and(ord(tprime) le (ord(t)+O(j,w)-1)))..
a(i,tprime,w) =g= x(i,t)*m(i,j);
equation9 (t,w)..  sum(i,s(i,t,w))=l=bicu+OICU;
equation10.. oicu =l= uicu;
equation11(d).. OOR(d) =l= uor;

object..
sum(w,p(w)*OICU*CICU)+sum((d,w),p(w)*oor(d)*cor)+sum((j,t),z(t,j)*surg)+sum(i,3798*5*u(i))
=e= mun;
Model ergin /all/ ;

option limcol=0,limrow=0,solprint=off;
option reslim=3600;
Solve ergin using mip minimizing mun;
```

```
display x.l;
```

# APPENDIX B. GENETIC ALGORITHM CODE FOR THE SCHEDULING MODEL IN MATLAB

```
tic
clear;
%   stream0 = RandStream('mt19937ar','Seed',0);
%   RandStream.setDefaultStream(stream0);
pop_size=60;
pat_size=132;
max_tries=20;
cross_size=20;
infeasibility=0;
inf_cost=0;
infeasibility_regular=0;
infeasibility_down=0;
infeasibility_up=0;
feasibility=0;
flag=0;
gen=1000;
global or_room;
or_room=8;

initial_pop_try=60;

pop_start1=zeros(pop_size,pat_size+5);
pat_list1=zeros(pop_size,pat_size);
u1=zeros(pop_size,or_room+1);

pop_start=zeros(initial_pop_try,pat_size+5);
pat_list=zeros(initial_pop_try,pat_size);
u2=zeros(initial_pop_try,or_room+1);

B=zeros(pop_size+cross_size,1);
IX=zeros(pop_size,1);

off_pop_start=zeros(pop_size,pat_size+5);
off_pat_list=zeros(pop_size,pat_size);
off_u1=zeros(pop_size,or_room+1);

IX1=zeros(pop_size,1);
B1=zeros(pop_size+cross_size,1);

best_u1=zeros(1,or_room+1);
best_pop_start=zeros(1,pat_size+5);
best_pat_list=zeros(1,pat_size);
```

```
roul_pop1=zeros(pop_size,1);
roul_pop2=zeros(pop_size+cross_size,1);

temp=0;

merge_pat_list=zeros(pop_size+cross_size,pat_size);
merge_u1=zeros(pop_size+cross_size,or_room+1);
merge_pop_start=zeros(pop_size+cross_size,pat_size+2);
crosslist_pop1=zeros(cross_size,1);
elit_selection=10;
mut_prob=0.02;
num_gen=1000;
ploy=zeros(1,num_gen);
ploy1=zeros(1,num_gen);

ploy(1)=0;
ploy1(1)=0;
global mult;
global prob;
global burak;
global oper_time;
global surgery_availability;
global diag_type;
global additional_hire;
global surg_specialty;
global num_scenarios;
global num_extensioni;
global num_regular;
global num_working;
global oicu;
global bicu;
global num_days;

mult=1.001;

burak=zeros(1,surg_specialty);
burak(surg_specialty+1)=1;
num_working=10;
num_extensioni=14;
num_regular=8;
surg_specialty=10;
num_days=5;
additional_hire=ones(num_extensioni*surg_specialty);

  surgery_availability=[
1    1    2    1    1    1    1    1    1    2    ;
1    1    2    1    1    1    1    1    1    2    ;
2    1    2    1    1    1    1    1    1    2    ;
1    1    2    1    1    1    1    1    1    2    ;
1    1    2    1    1    1    1    1    1    2    ;
1    1    2    1    1    1    1    1    1    2    ;
```

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | ; |

```
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
1	1	2	1	1	1	1	1	1	2	;
0	0	0	0	0	0	0	0	0	0	;
0	0	0	0	0	0	0	0	0	0	;
0	0	0	0	0	0	0	0	0	0	;
0	0	0	0	0	0	0	0	0	0	];

 diag_type=[
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
1	0	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	1	0	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
0	0	1	0	0	0	0	0	0	0	;
```

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | ; |

```
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	1	0	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	1	0	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	1	0	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	;
0	0	0	0	0	0	0	0	0	1	];

oper_time=[
	4	0 0;
	1	0 0;
	1	2 3;
	1	0 0;
```

```
        4          0  0;
        1          2  0;
        2   0  0;
        2          0  0;
        1          2  0;
        2          3  0];

  product=1;

     for k=1:surg_specialty
        product=product*nnz(oper_time(k,:));
        burak(surg_specialty-k+1)=burak(surg_specialty-k+2)*nnz(oper_time(surg_specialty-k+1,:));
     end;
        num_scenarios=product;

   inten_stay=ones(1,surg_specialty);
   bicu=5;
   oicu=1;
   prob=[0.03    0.075    0.045    0.03     0.075    0.045    0.02     0.05
      0.03          0.02     0.05     0.03     0.03     0.075    0.045    0.03
      0.075        0.045    0.02     0.05     0.03     0.02     0.05     0.03];
   traverse_list=zeros(1,pat_size);

 for i=1:initial_pop_try
    pat_list(i,:)=randperm(pat_size);
      traverse_list=zeros(1,pat_size);
    for t=1:or_room+1

       u2(i,t)=randsample((round(((pat_size/or_room))-
(0.1*round(pat_size/or_room))):(round((pat_size/or_room)+(0.1*round(pat_size/or_room))))),1);
    end;
       u2(i, or_room+1)=max(0,pat_size-sum(u2(i,1:or_room)));

if sum(u2(i,:))<pat_size

        while sum(u2(i,:))<pat_size
        k2=randsample(or_room,1);
        u2(i,k2)=min(pat_size,u2(i,k2)+1);
        end;
     end;
          if sum(u2(i,:))>pat_size

        while sum(u2(i,:))>pat_size
        k2=randsample(or_room,1);
        u2(i,k2)=max(0,u2(i,k2)-1);
        end;
     end;
     for y=1:or_room
        if y==1
           pop_start(i,1:u2(i,1))=randsample(num_extensioni*num_days,u2(i,1));
        else
```

```
           pop_start(i,sum(u2(i,1:y-
1))+1:sum(u2(i,1:y)))=randsample(num_extensioni*num_days,u2(i,y));
           end;
       end;
       for t2000=1:sum(u2(i,1:or_room))

           t=pop_start(i,t2000);

   while (mod(t,num_extensioni)==0) || (mod(t,num_extensioni)>num_working)

       t=randsample(num_extensioni*num_days,1);

   end;
    pop_start(i,t2000)=t;

         end;



        [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);
        if flag==0
            pop_start(i,pat_size+1)=infeasibility;
              pop_start(i,pat_size+2)=flag;
              pop_start(i,pat_size+3)=inf_cost;
              pop_start(i,pat_size+4)=feasibility;
              pop_start(i,pat_size+5)=1/(feasibility+inf_cost);
        else
            infeasibility_regular=infeasibility;

            traverse_list(1:sum(u2(i,1:or_room)))=randperm(sum(u2(i,1:or_room)));

            for koray=1:sum(u2(i,1:or_room))
                if (mod(pop_start(i,traverse_list(koray)),num_extensioni)>1) &&
(mod(pop_start(i,traverse_list(koray)),num_extensioni)<num_working)
           if mod(pop_start(i,traverse_list(koray)),num_extensioni)==1
               (pop_start(i,traverse_list(koray),num_extensioni))
           end;

                    pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))-1;

            [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);
            infeasibility_down=infeasibility;
               pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))+2;
               [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);
            infeasibility_up=infeasibility;
                if infeasibility_regular<infeasibility_down  && infeasibility_regular<infeasibility_up
                    pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))-1;
                else if infeasibility_down<infeasibility_regular  && infeasibility_down<infeasibility_up
                    pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))-2;
                    end;
                end;
```

```
            else if (mod((pop_start(i,traverse_list(koray))),num_extensioni)==num_working)
        pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))-1;
        [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);
        infeasibility_down=infeasibility;

            if infeasibility_regular<infeasibility_down
                pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))+1;
                end;

         else if (mod((pop_start(i,traverse_list(koray))),num_extensioni)==1)
        pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))+1;
        [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);
        infeasibility_up=infeasibility;
            end;
             if infeasibility_regular<infeasibility_up
                pop_start(i,traverse_list(koray))=pop_start(i,traverse_list(koray))-1;

                end;
                end;
        end;
        end;
         [inf_cost, infeasibility, feasibility, flag]=erman(pop_start,pat_list,u2,i);

        pop_start(i,pat_size+4)=feasibility;
          pop_start(i,pat_size+1)=infeasibility;
              pop_start(i,pat_size+3)=inf_cost;
              pop_start(i,pat_size+2)=flag;
              pop_start(i,pat_size+5)=feasibility+inf_cost*(mult^gen);
             end;
    end;

  [B1 IX1]=sort(pop_start(:,pat_size+1), 'ascend');

   for i1=1:pop_size
       u1(i1,:)=u2(IX1(i1),:);
       pop_start1(i1,:)=pop_start(IX1(i1),:);
         pat_list1(i1,:)=pat_list(IX1(i1),:);
   end;

  for gen=2:num_gen
      for i1=1:pop_size

      if pop_start1(i1,pat_size+2)==1
          pop_start1(i1,pat_size+5)=
1/(pop_start1(i1,pat_size+4)+pop_start1(i1,pat_size+3)*(mult^(gen)));
      end;
      end;

 total_fitness=sum(pop_start1(:,pat_size+5));
 [B IX]=sort(pop_start1(:,pat_size+5),'descend');
```

```
%   roul_pop1=zeros(60,1);
  for i=1:pop_size
        fitness=sum(B(1:i));
        roul_pop1(i)=fitness/total_fitness;
  end;
    for i3=1:cross_size
   y=rand;

  for j=1:pop_size-1

      if ((y>roul_pop1(j)) && (y<=roul_pop1(j+1)))

%           if ((flag==0) | (IX(j+1)~=60))
          crosslist_pop1(i3)=IX(j+1);
%           i3=i3+1;
          break;

      else if  y<=roul_pop1(1)

%             if ((flag==0) | (IX(1)~=60))
          crosslist_pop1(i3)=IX(1);
%           i3=i3+1;
          break;
      end;
          end;
      end;

  end;

  off_pop_start=zeros(cross_size,pat_size+5);
off_pat_list=zeros(cross_size,pat_size);
off_u1=zeros(cross_size,or_room+1);

  for j=1:2:cross_size

      yson=rand;
if yson<=0.01

    cross_locations=randsample(2:pat_size,2);
    cross_locations=sort(cross_locations);


off_pat_list(j,cross_locations(1):cross_locations(2))=pat_list1(crosslist_pop1(j),cross_locations(1):cr
oss_locations(2));

off_pat_list(j+1,cross_locations(1):cross_locations(2))=pat_list1(crosslist_pop1(j+1),cross_locations(
1):cross_locations(2));

    for i1=cross_locations(2)+1:pat_size
```

```
        if
nnz(pat_list1(crosslist_pop1(j+1),i1)==pat_list1(crosslist_pop1(j),cross_locations(1):cross_locations(
2)))==0

          off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),i1);

        else

            ind=find(pat_list1(crosslist_pop1(j),:)==pat_list1(crosslist_pop1(j+1),i1));
              while nnz((pat_list1(crosslist_pop1(j+1),ind))==off_pat_list(j,:))>0

                ind=find(pat_list1(crosslist_pop1(j),:)==pat_list1(crosslist_pop1(j+1),ind));

              end;
               off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),ind);

          off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),ind);

              end;
              end;

        for i1=cross_locations(2)+1:pat_size

        if
nnz(pat_list1(crosslist_pop1(j),i1)==pat_list1(crosslist_pop1(j+1),cross_locations(1):cross_locations(
2)))==0
          off_pat_list(j+1,i1)=pat_list1(crosslist_pop1(j),i1);

        else

            ind=find(pat_list1(crosslist_pop1(j+1),:)==pat_list1(crosslist_pop1(j),i1));

              while nnz((pat_list1(crosslist_pop1(j),ind))==off_pat_list(j+1,:))>0

                ind=find(pat_list1(crosslist_pop1(j+1),:)==pat_list1(crosslist_pop1(j),ind));

              end;
               off_pat_list(j+1,i1)=pat_list1(crosslist_pop1(j),ind);

              end;
        end;

    for i1=1:cross_locations(1)-1

       if
nnz(pat_list1(crosslist_pop1(j+1),i1)==pat_list1(crosslist_pop1(j),cross_locations(1):cross_locations(
2)))==0

          off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),i1);

        else
```

```
          ind=find(pat_list1(crosslist_pop1(j),:)==pat_list1(crosslist_pop1(j+1),i1));
            while nnz((pat_list1(crosslist_pop1(j+1),ind))==off_pat_list(j,:))>0

              ind=find(pat_list1(crosslist_pop1(j),:)==pat_list1(crosslist_pop1(j+1),ind));

            end;
             off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),ind);

        off_pat_list(j,i1)=pat_list1(crosslist_pop1(j+1),ind);

            end;
            end;

       for i1=1:cross_locations(1)-1

        if
nnz(pat_list1(crosslist_pop1(j),i1)==pat_list1(crosslist_pop1(j+1),cross_locations(1):cross_locations(
2)))==0
        off_pat_list(j+1,i1)=pat_list1(crosslist_pop1(j),i1);

      else

          ind=find(pat_list1(crosslist_pop1(j+1),:)==pat_list1(crosslist_pop1(j),i1));

            while nnz((pat_list1(crosslist_pop1(j),ind))==off_pat_list(j+1,:))>0

              ind=find(pat_list1(crosslist_pop1(j+1),:)==pat_list1(crosslist_pop1(j),ind));

            end;
             off_pat_list(j+1,i1)=pat_list1(crosslist_pop1(j),ind);

            end;
      end;

else

    off_pat_list(j,:)=pat_list1(crosslist_pop1(j),:);
    off_pat_list(j+1,:)=pat_list1(crosslist_pop1(j+1),:);


end;
         cross_location1=0;
for t1=1:or_room
   if t1==1 && min((u1(crosslist_pop1(j),t1)),(u1(crosslist_pop1(j+1),t1)))~=0
   cross_location1=randsample(min((u1(crosslist_pop1(j),t1)),(u1(crosslist_pop1(j+1),t1)))-1,1);

   off_pop_start(j,1:cross_location1)=pop_start1(crosslist_pop1(j+1),1:cross_location1);

off_pop_start(j,cross_location1+1:u1(crosslist_pop1(j),t1))=pop_start1(crosslist_pop1(j),cross_locati
on1+1:u1(crosslist_pop1(j),t1));
```

```
%off_empty(j+1,cross_locations(1):cross_locations(2))=pop1_empty(crosslist_pop1(j+1),cross_locat
ions(1):cross_locations(2));
    off_pop_start(j+1,1:cross_location1)=pop_start1(crosslist_pop1(j),1:cross_location1);

off_pop_start(j+1,cross_location1+1:u1(crosslist_pop1(j+1),t1))=pop_start1(crosslist_pop1(j+1),cros
s_location1+1:u1(crosslist_pop1(j+1),t1));

    elseif (u1(crosslist_pop1(j+1),t1))==0

        off_pop_start(j,:)=pop_start1(crosslist_pop1(j),:);
    off_pop_start(j+1,:)=pop_start1(crosslist_pop1(j+1),:);

    elseif t1~=1 && min((u1(crosslist_pop1(j),t1)),(u1(crosslist_pop1(j+1),t1)))~=0
        cross_location1=randsample((min(u1(crosslist_pop1(j),t1),u1(crosslist_pop1(j+1),t1))-1),1);

    off_pop_start(j,sum(u1(crosslist_pop1(j),1:t1-1))+1:sum(u1(crosslist_pop1(j),1:t1-
1))+cross_location1)=pop_start1(crosslist_pop1(j+1),sum(u1(crosslist_pop1(j+1),1:t1-
1))+1:sum(u1(crosslist_pop1(j+1),1:t1-1))+cross_location1);
    off_pop_start(j,sum(u1(crosslist_pop1(j),1:t1-
1))+cross_location1+1:sum(u1(crosslist_pop1(j),1:t1)))=pop_start1(crosslist_pop1(j),sum(u1(crosslis
t_pop1(j),1:t1-1))+cross_location1+1:sum(u1(crosslist_pop1(j),1:t1)));


%off_empty(j+1,cross_locations(1):cross_locations(2))=pop1_empty(crosslist_pop1(j+1),cross_locat
ions(1):cross_locations(2));
    off_pop_start(j+1,sum(u1(crosslist_pop1(j+1),1:t1-1))+1:sum(u1(crosslist_pop1(j+1),1:t1-
1))+cross_location1)=pop_start1(crosslist_pop1(j),sum(u1(crosslist_pop1(j),1:t1-
1))+1:sum(u1(crosslist_pop1(j),1:t1-1))+cross_location1);
    off_pop_start(j+1,sum(u1(crosslist_pop1(j+1),1:t1-
1))+cross_location1+1:sum(u1(crosslist_pop1(j+1),1:t1)))=pop_start1(crosslist_pop1(j+1),sum(u1(cr
osslist_pop1(j+1),1:t1-1))+cross_location1+1:sum(u1(crosslist_pop1(j+1),1:t1)));

    end;
end;
    t5=rand;
    if t5<=0.01

    y=randsample(or_room,1);

    off_u1(j,1:y)=(u1(crosslist_pop1(j),1:y));
    off_u1(j+1,1:y)=(u1(crosslist_pop1(j+1),1:y));

    off_u1(j+1,y+1:or_room+1)=(u1(crosslist_pop1(j),y+1:or_room+1));
    off_u1(j,y+1:or_room+1)=(u1(crosslist_pop1(j+1),y+1:or_room+1));

    else
     off_u1(j,:)=(u1(crosslist_pop1(j),:));
    off_u1(j+1,:)=(u1(crosslist_pop1(j+1),:));
```

```
  end;

  while sum(off_u1(j,:))>pat_size
    t=randsample(or_room+1,1);

    while off_u1(j,t)==0
        t=randsample(or_room+1,1);
    end;

    off_u1(j,t)=off_u1(j,t)-1;
  end;

  while sum(off_u1(j,:))<pat_size

    t=randsample(or_room+1,1);

     while off_u1(j,t)==pat_size
        t=randsample(or_room+1,1);
    end;

    off_u1(j,t)=off_u1(j,t)+1;
  end;

  while sum(off_u1(j+1,:))<pat_size
    t=randsample(or_room+1,1);

    while off_u1(j+1,t)==pat_size
        t=randsample(or_room+1,1);
    end;
    off_u1(j+1,t)=off_u1(j+1,t)+1;
  end;

  while sum(off_u1(j+1,:))>pat_size
    t=randsample(or_room+1,1);
   while off_u1(j+1,t)==0
        t=randsample(or_room+1,1);
    end;

    off_u1(j+1,t)=off_u1(j+1,t)-1;

  end;
 end;

 for j=1:2:cross_size

i10=sum(off_u1(j,1:or_room));
i11=sum(off_u1(j+1,1:or_room));
i12=nnz(off_pop_start(j+1,1:pat_size));
i13=nnz(off_pop_start(j,1:pat_size));

if i10-i13>0
```

```
    for t1000=1:(i10-i13)

          y=randsample(num_extensioni*num_days,1);
      while  (mod(y,num_extensioni)==0) || (mod(y,num_extensioni)>num_working)
      y=randsample(num_extensioni*num_days,1);
    end;

      off_pop_start(j,i13+t1000)=y;

    end;
end;
if i10-i13<0
    for t1000=0:(i13-i10-1)
        off_pop_start(j,i13-t1000)=0;
    end;
end;

if i11-i12>0
    for t1000=1:(i11-i12)
        y=randsample(num_extensioni*num_days,1);
      while  (mod(y,num_extensioni)==0) || (mod(y,num_extensioni)>num_working)
      y=randsample(num_extensioni*num_days,1);
    end;

      off_pop_start(j+1,i12+t1000)=y;
end;
end;
if i11-i12<0
    for t1000=0:(i12-i11-1)
        off_pop_start(j+1,i12-t1000)=0;
    end;
end;
  end;

 for j=1:cross_size
for i1=1:pat_size

          y1=rand;

          if y1<=mut_prob

               y2=randsample(pat_size,1);

                while y2==i1
             y2=randsample(pat_size,1);
                end;

                   temp=off_pat_list(j,y2);
                   off_pat_list(j,y2)=off_pat_list(j,i1);
                   off_pat_list(j,i1)=temp;
          end;
```

```
   end;
for i1=1:sum(off_u1(j,1:or_room))

           y=rand;
           if y<=mut_prob

              y1=rand;

              y2=randsample(0:num_extensioni*num_days-1,1);
              tries=0;
              flag=0;

              if y1<=0.5

                while ((off_pop_start(j,i1)-y2<=0) ||(mod(off_pop_start(j,i1)-y2,num_extensioni)==0) ||
(mod(off_pop_start(j,i1)-y2,num_extensioni)>num_working)) && (tries<max_tries)
                     y2=randsample(0:num_extensioni*num_days-1,1);
                     tries=tries+1;
                      if tries==max_tries
                         flag=1;
                     end;
                end;

                if flag==0
                off_pop_start(j,i1)=off_pop_start(j,i1)-y2;
                end;
              else
                  while ((off_pop_start(j,i1)+y2>num_extensioni*num_days) ||
(mod(off_pop_start(j,i1)+y2,num_extensioni)==0) ||
(mod(off_pop_start(j,i1)+y2,num_extensioni)>num_working)) && (tries<max_tries)
                     y2=randsample(0:num_extensioni*num_days-1,1);
                     tries=tries+1;
                     if tries==max_tries
                     flag=1;
                end;
                  end;
                if flag==0
                off_pop_start(j,i1)=off_pop_start(j,i1)+y2;
                 end;

              end;
           end;

end;

for i1=1:or_room+1

           y=rand;
           if y<=mut_prob

              y1=rand;
```

```
            if y1<=0.5

            y2=randsample(or_room+1,1);
            while y2==i1
               y2=randsample(or_room+1,1);
            end;

            if off_u1(j,i1)>0 && off_u1(j,y2)<pat_size

            off_u1(j,i1)=off_u1(j,i1)-1;
             off_u1(j,y2)=off_u1(j,y2)+1;

            end;

            end;

             if y1>0.5

            y2=randsample(or_room+1,1);
            while y2==i1
               y2=randsample(or_room+1,1);
            end;

            if  off_u1(j,y2)>0 && off_u1(j,i1)<pat_size

            off_u1(j,i1)=off_u1(j,i1)+1;
             off_u1(j,y2)=off_u1(j,y2)-1;

            end;

            end;
         end;
end;
  end;

  for j=1:2:cross_size

 i10=sum(off_u1(j,1:or_room));
 i11=sum(off_u1(j+1,1:or_room));
 i12=nnz(off_pop_start(j+1,1:pat_size));
 i13=nnz(off_pop_start(j,1:pat_size));

 if i10-i13>0
    for t1000=1:(i10-i13)

          y=randsample(num_extensioni*num_days,1);
      while  (mod(y,num_extensioni)==0) || (mod(y,num_extensioni)>num_working)
      y=randsample(num_extensioni*num_days,1);
    end;

      off_pop_start(j,i13+t1000)=y;
```

```
    end;
end;

if i10-i13<0
    for t1000=0:(i13-i10-1)
        off_pop_start(j,i13-t1000)=0;
    end;
end;

if i11-i12>0
    for t1000=1:(i11-i12)
        y=randsample(num_extensioni*num_days,1);
       while  (mod(y,num_extensioni)==0) || (mod(y,num_extensioni)>num_working)
      y=randsample(num_extensioni*num_days,1);
    end;

      off_pop_start(j+1,i12+t1000)=y;
end;
end;
if i11-i12<0
    for t1000=0:(i12-i11-1)
        off_pop_start(j+1,i12-t1000)=0;
    end;
end;

 end;
     for i=1:cross_size

         traverse_list=zeros(1,pat_size);
          [inf_cost, infeasibility, feasibility, flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size);
        if flag==0
            off_pop_start(i,pat_size+1)=infeasibility;
              off_pop_start(i,pat_size+2)=flag;
              off_pop_start(i,pat_size+3)=inf_cost;
              off_pop_start(i,pat_size+4)=feasibility;
              off_pop_start(i,pat_size+5)=feasibility+inf_cost*mult^(gen);
        else
            infeasibility_regular=infeasibility;
              traverse_list(1:sum(off_u1(1:i)))=randperm(1:sum(off_u1(i,1:i)));
            for koray=1:sum(off_u1(1:i))
                if (mod((off_pop_start(i,traverse_list(koray))),num_extensioni)>1) &&
(mod(off_pop_start(i,traverse_list(koray)),num_extensioni)<num_working)
            off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))-1;
            [inf_cost, infeasibility, feasibility, flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size);
            infeasibility_down=infeasibility;
               off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))+2;
               [inf_cost, infeasibility, feasibility,
flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size);
            infeasibility_up=infeasibility;
                if infeasibility_regular<infeasibility_down  && infeasibility_regular<infeasibility_up
```

```
                off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))-1;
            else if infeasibility_down<infeasibility_regular  && infeasibility_down<infeasibility_up
                off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))-2;
                end;
            end;

            else if (mod((off_pop_start(i,traverse_list(koray))),num_extensioni)==num_working)
        off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))-1;
        [inf_cost, infeasibility, feasibility,
flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size_size);
        infeasibility_down=infeasibility;

            if infeasibility_regular<infeasibility_down
                off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))+1;

                end;

         else if (mod((off_pop_start(i,traverse_list(koray))),num_extensioni)==num_working)
        off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))+1;
        [inf_cost, infeasibility, feasibility, flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size);
        infeasibility_up=infeasibility;
            end;
             if infeasibility_regular<infeasibility_up
                off_pop_start(i,traverse_list(koray))=off_pop_start(i,traverse_list(koray))-1;

                end;
                end;
        end;
        end;
         [inf_cost, infeasibility, feasibility, flag]=erman(off_pop_start,off_pat_list,off_u1,cross_size);

        off_pop_start(i,pat_size+4)=feasibility;
          off_pop_start(i,pat_size+1)=infeasibility;
              off_pop_start(i,pat_size+3)=inf_cost;
              off_pop_start(i,pat_size+2)=flag;
              off_pop_start(i,pat_size+5)=1/(feasibility+inf_cost*(mult^gen));

             end;
    end;

 merge_pat_list=[pat_list1; off_pat_list];
 merge_pop_start=[pop_start1 ;off_pop_start];
 merge_u1=[u1 ;off_u1];

 [B1 IX1]=sort(merge_pop_start(:,pat_size+5), 'descend');
   total_fitness=sum(merge_pop_start(:,pat_size+5));
   if (merge_pop_start(IX1(1),pat_size+5)>best_pop_start(pat_size+5)) &&
(merge_pop_start(IX1(1),pat_size+2)==0)
       ploy(gen)=merge_pop_start(IX1(1),pat_size+5);
       best_u1=merge_u1(IX1(1),:);
       best_pop_start=merge_pop_start(IX1(1),:);
```

```
        best_pat_list=merge_pat_list(IX1(1),:);
      else if (gen>=2);
          ploy(gen)=ploy(gen-1);
           end;
    end;

    for i=1:pop_size+cross_size_size
        fitness=sum(B1(1:i));
         roul_pop2(i)=fitness/total_fitness;
    end;

    for j=1:elit_selection;

%         pop1_final(j,:)=merge_population(IX1(j),:);
        pop_start1(j,:)=merge_pop_start(IX1(j),:);
        u1(j,:)=merge_u1(IX1(j),:);
        pat_list1(j,:)=merge_pat_list(IX1(j),:);
%          pop1_empty(j,:)=merge_empty(IX1(j),:);
    end;  for i=elit_selection+1:pop_size
   y=rand;
  for j=1:pop_size+cross_size_size-1      if ((y>roul_pop2(j)) && (y<=roul_pop2(j+1)))
%          pop1_final(i,:)=merge_population(IX1(j+1),:);
        pop_start1(i,:)=merge_pop_start(IX1(j+1),:);
        u1(i,:)=merge_u1(IX1(j+1),:);
        pat_list1(i,:)=merge_pat_list(IX1(j+1),:);
%          pop1_empty(i,:)=merge_empty(IX1(j+1),:);
        break;

      else if y<=roul_pop2(1)
%          pop1_final(i,:)=merge_population(IX1(1),:);
        pop_start1(i,:)=merge_pop_start(IX1(1),:);
        u1(i,:)=merge_u1(IX1(1),:);
        pat_list1(i,:)=merge_pat_list(IX1(1),:);
%          pop1_empty(i,:)=merge_empty(IX1(1),:);
        break;
      end;  end;
  end;
  end;
      end;
  ploy1(gen)=mean(pop_start1(:,pat_size+1));
    toc;
```

# APPENDIX C. BENDER'S DECOMPOSITION ALGORITHM CONSIDERATIONS

Bender's decomposition method is usually employed to solve the class of optimization problems that possess the specific structure. In that regard, the structure of the problem is exploited. Bender's decomposition is also extensively used for the stochastic programming models (Infanger, 1994; Nielsen and Zenios, 1997). The structure of the solution methodology can be provided as follows (Martin, 1999),

Consider the optimization problem having the following form;

$$\text{Minimize}\, \mathrm{c}^{\mathrm{T}} x + f^{T} y \tag{C.1}$$

s.t.

$$Ax + By \geq b \tag{C.2}$$

$$y \in Y \tag{C.3}$$

$$x \geq 0 \tag{C.4}$$

Assuming that y is fixed for some integer values, the model takes the following form,

$$\text{Minimize}\, \mathrm{c}^{\mathrm{T}} x \tag{C.5}$$

$$Ax \geq b - B\bar{y} \tag{C.6}$$

$$x \geq 0 \tag{C.7}$$

The model therefore can be represented as,

$$\underset{\mathrm{y}}{\text{Minimize}} \left[ f^{T} y + \min\, \mathrm{c}^{\mathrm{T}} x \middle| Ax \geq b - By \right] \tag{C.8}$$

The dual of the inner LP problem can be represented as;

$$\text{Maximize}\ (\mathrm{b} - \mathrm{B}\bar{\mathrm{y}})^{\mathrm{T}} u \tag{C.9}$$

$$A^{T} u \leq c \tag{C.10}$$

$$u \geq 0 \tag{C.11}$$

In Bender's decomposition algorithm, two different set of problems (i.e., the master problem and the sub-problems are solved sequentially, and based on the results obtained from the sub-problem, the cuts are generated and progressively added to the master problem to update the corresponding bounds. A restricted master problem has the following form;

$$\text{Minimize} z \tag{C.12}$$

s.t.

$$z \geq f^T y + (b - By)^T \bar{u}_{k,} \; k = 1 \ldots K \tag{C.13}$$

$$(b - By)^T \bar{u}_l \leq 0, l = 1 ... L \tag{C.14}$$

$$y \in Y \tag{C.15}$$

and sub-problem of the following form,

$$\text{Maximize } (\mathrm{b} - \mathrm{B}\bar{\mathrm{y}})^{\mathrm{T}} u \tag{C.16}$$

s.t.

$$A^T u \leq c \tag{C.17}$$

$$u \geq 0 \tag{C.18}$$

Based on this notation, the algorithm for Bender's Decomposition can be stated as follows; (Kalvelagen, 2005)

[Initialization]

Y:=Initial feasible integer solution

LB := -∞

UB := ∞

**while** UB-LB>ε

[solve subproblem]

$\max_u \{f^T \bar{y} + (b - B\bar{y})^T u \mid A^T u \le c, u \ge 0\}$

**If t**he sub-problem is unbounded then

Obtain unbounded ray $(b - By)^T \bar{u} \le 0$

Add cut $(b - By)^T \bar{u} \le 0$ to the master problem

**Else**

Get extreme point $\bar{u}$

Add cut $z \le f^T \bar{y} + (b - By)^T \bar{u} \le 0$ to the master problem

Update upper bound

**End if**

[solve master problem]

$\min_y \{z \mid cuts, y \in Y\}$

LB := $\bar{z}$

**end while**

Bender's decomposition method can be applied by our problem domain for the first part of the problem. The master problem will be formed by using the variant of Eqs. C.13-C.15. Bender's decomposition method will be adopted and implemented for the problem. For this purpose, the problem will be divided into portions, the master problem and the sub-problem. Eqs. C.2, C.3, C.5, C.10, and C.11 are used in the master problem. On the other hand, for each iteration, the Eqs. C.4, C.8, and C.9 are included in the sub-problem. The master and sub problems are linked with Eqs. C.6 and C.7.