SIMULATION AND OPTIMIZATION MODELS FOR SCHEDULING MULTI-STEP SEQUENTIAL PROCEDURES IN NUCLEAR MEDICINE

A Dissertation

by

EDUARDO PEREZ ROMAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2010

Major Subject: Industrial Engineering

SIMULATION AND OPTIMIZATION MODELS FOR SCHEDULING MULTI-STEP SEQUENTIAL PROCEDURES IN NUCLEAR MEDICINE

A Dissertation

by

EDUARDO PEREZ ROMAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Lewis Ntaimo
	César Malavé
Committee Members,	Donald R. Smith
	Sheng-Jen Hsieh
Head of Department,	Brett A. Peters

May 2010

Major Subject: Industrial Engineering

ABSTRACT

Simulation and Optimization Models for Scheduling Multi-step Sequential Procedures in Nuclear Medicine. (May 2010)

Eduardo Pérez Román, B.Sc. Industrial Engineering, University of Puerto Rico

Co–Chairs of Advisory Committee: Dr. Lewis Ntaimo
Dr. César Malavé

The rise in demand for specialized medical services in the U.S has been recognized as one of the contributors to increased health care costs. Nuclear medicine is a specialized service that uses relatively new technologies and radiopharmaceuticals with a short half-life for diagnosis and treatment of patients. Nuclear medicine procedures are multi-step and have to be performed under restrictive time constraints. Consequently, managing patients in nuclear medicine clinics is a challenging problem with little research attention. In this work we present simulation and optimization models for improving patient and resource scheduling in health care specialty clinics such as nuclear medicine departments. We first derive a discrete event system specification (DEVS) simulation model for nuclear medicine patient service management that considers both patient and management perspectives. DEVS is a formal modeling and simulation framework based on dynamical systems theory and provides well defined concepts for coupling components, hierarchical and modular model construction, and an object-oriented substrate supporting repository reuse. Secondly, we derive algorithms for scheduling nuclear medicine patients and resources and validate our algorithms using the simulation model. We obtain computational results that provide useful insights into patient service management in nuclear medicine. For example, the number of patients seen at the clinic during a year increases when a group of stations are reserved to serve procedures with higher demand. Finally, we derive a stochastic online scheduling (SOS) algorithm for patient and resource management in nuclear medicine clinics. The algorithm performs scheduling decisions by taking into account stochastic information about patient future arrivals. We compare the results obtained using the SOS algorithm with the algorithms that do not take into consideration stochastic information. The SOS algorithm provides a balanced utilization of resources and a 10% improvement in the number of patients served.

To my family: their support has allowed me to complete this great task

ACKNOWLEDGMENTS

First and foremost, I would like to thank God. You have given me the tools to pursue my dreams. I could never have done this without the faith I have in You, the Almighty. I would like to thank my fiancée, Cinthya Alvarez. Her love and patience has given me the strength to complete this journey. I would also like to thank my family for the support they have given me over the years. My mother, Carmen A. Román, deserves special thanks for the endless encouragement. I can barely find the words to express all the wisdom, love and support she has given me. The support of my father, Rafael Pérez, is also greatly appreciated.

I am extremely grateful to my committee members. My advisors, Dr.Lewis Ntaimo and Dr.César Malavé, deserve special acknowledgments for the support and guidance they have provided me throughout the years. I thank Judy Meeks for the great job she does in the industrial and systems engineering department and for providing great advice.

I would like to thank my colleague graduate students, especially to my friends Mayra Méndez, Roberto Seijo, Joselyn Torres, Julian A. Gallego, and Cesar Rincón. Your friendship played a major role during this journey. Finally, I am extremely grateful of the people in charge of the St. Marys Student Catholic Center for providing a home away from home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

CHAPTER I

INTRODUCTION

The goal of this research is to develop models and algorithms for improving patient service and resources management in highly constrained health care environments such as nuclear medicine clinics (a sub-specialty of radiology). This research proposes a methodology towards achieving this objective, including a simulation model that integrates with the scheduling methods to manage patient service levels and resource productivity, scheduling algorithms, and optimizing methods for scheduling when dealing with random disruptions.

This research is motivated by the fact that health care facilities dedicated to the diagnosis and treatment of patients are becoming more critical in comprehensive health care systems. Diagnostic medicine procedures have increased 5-to-6 fold whereas the U.S. population increased by approximately 50% in the last few years [1]. This increase in demand has been identified as one of the contributors to the rise of health care costs in the U.S. [2, 3].

Physicians are becoming more prone in asking patients to undergo specialized procedures to make more accurate diagnoses. However, managing patients and resources in specialized clinics such as nuclear medicine remains a challenging problem. This may be attributed to the rise in demand for services and the nature of nuclear medicine procedures/tests.

These procedures require the use of radiopharmaceuticals with limited half-life, involve several steps that are constrained by strict time windows, and require multiple resources for completion. Moreover, schedules in nuclear medicine have to account

---

The journal model is *IEEE Transactions on Automatic Control.*

for other factors such as patient behavior, staff experience, service time variability, equipment failures, and radiopharmaceuticals delivery times. Consequently, patient management in nuclear medicine specialty clinics is a very challenging problem with very little research attention. Few commercial packages are available to assist nuclear medicine managers and they provide very rudimentary capabilities.

A. Problem Statement

Nuclear medicine uses painless and cost effective techniques to diagnose and treat diseases. Procedures in this field are essential in many medical specialties such as cardiology, pediatrics, and psychiatry. Nuclear medicine looks at both the physiology (functioning) and the anatomy of the body in establishing diagnosis and treatment. This health care discipline differs from X-ray, ultrasound and other diagnostic techniques because determines the presence of a disease based on biological changes rather than changes in anatomy.

Most nuclear medicine procedures require small amounts of radioactive materials (radiopharmaceuticals) to be performed. Nuclear pharmacies are dedicated to the compounding and dispensing of these radioactive materials. A centralized nuclear pharmacy usually serve as a "drugstore" to several nuclear medicine departments. When a particular radiopharmaceutical is required by a nuclear medicine department, a request is performed and a trained nuclear pharmacist prepares the product and dispenses it to the requesting clinic. Dosages, lead times, and appointment schedules depend upon the lead time required to supply the radiopharmaceuticals to the clinic.

Radiopharmaceuticals are introduced into the body of the patient by injection, swallowing, or inhalation. Radiopharmaceuticals require a time period that may range from hours to days to diffuse. The time required depends on the body tissue being examined and the radiopharmaceutical being used. These substances decay quickly and their effectiveness is only for a short period of time.

Nuclear medicine procedures are multi-step, involve multiple resources, and have to be performed under restrictive time constraints. A test may involve a single camera scan requiring either a short or long time (minutes to hours), or multiple scans during one day, or on multiple days. Most of these procedures can be divided in three major steps: radiopharmaceutical administration, imaging, and image interpretation. Radiopharmaceuticals are detected by gamma cameras that work with computers to provide images of the body part being examined. Several camera configurations are available, including single and multiple cameras (parallel machines). A nuclear medicine clinic typically has a certain number of cameras of each type. A sophisticated camera may cost $1 million, although some can be purchased for $200,000. These advance technology equipment may be subject to breakdowns if proper maintenance is not performed on them. The time required to obtain the images may also vary from minutes to hours [4]. When imaging is completed too soon after radiopharmaceutical administration a poor quality image could be obtained because the radiopharmaceutical was not diffused properly. On the other hand, if imaging is delayed too long the image obtained will be degraded because too much decay of the radiopharmaceutical occurred. Under these circumstances the procedure has to be suspended and re-scheduled for another day causing unnecessary patient exposure to radiation, unnecessary resource utilization, and unnecessary expenses in a new radiopharmaceutical ($100-$2000) to repeat the procedure. Suthummanon et al. [4] demonstrated in their study that the operational time (machine time and direct labor time) and the cost of radiopharmaceuticals have the most influence in cost per procedure in nuclear medicine.

Nuclear medicine clinics normally operate one shift every weekday and may not be open on weekends. Nuclear medicine procedures require the participation of several health professionals. Nuclear physicians are in charge of examining the patients before procedures, interpreting test results, and they can also administer radiopharmaceuticals to the patients. Technologists perform several important activities, including calibrating equipment, counseling patients, administering radiopharmaceuticals, patient imaging, and results interpretation. Registered nurses assist in the procedures and also administer the radiopharmaceuticals to the patients. Nurses also have an important role taking care of patients during unexpected emergencies like seizures and cardiac arrest. Resources availability have an important role in performing the procedures as planned. Nuclear medicine facilities have a certain number of resources on hand and they must plan accordingly. When a resource is not available guides in how to handled patients must be follow.

Patient scheduling in nuclear medicine department must take into account several practical needs. A schedule must satisfy the goals of both the departments and the patients. Patients should be scheduled as soon as possible, resources and radiopharmaceutical constraints must be observed and interference with other patients must be minimized. Patients are concerned with the level of service they receive which can be measured in terms of cost. Cost associated with nuclear medicine procedures include money, time, discomfort, possibly drugs reactions, radiation dose and the hypothetical risk of radiation-associated cancer [5]. Nuclear medicine departments are concerned with providing acceptable levels of patient service, the good utilization of expensive resources, and with the minimization of costs such as overtime.

A schedule could be disrupted by random unexpected events. Some examples include, patient not showing up to their appointments, emergency patients requiring priority service, camera breakdowns, radiopharmaceuticals not delivered on time, and imaging repetition by physician request. If a patient moves during the scan, the procedure has to be repeated. Remaining immobile during a procedure may be difficult for some patients, especially for elderly patients that suffer from arthritis. Patients being served, pending (i.e. scheduled later in the day) and scheduled for future days may be rescheduled. Under unexpected events rescheduling is used to minimize patient inconvenience and to achieve higher service efficiency when a disruption occurs. Depending upon the nuclear medicine department size, either a method to prescribe an optimal schedule or a heuristic, may be preferred for rescheduling. Opportunities for improving efficiency and techniques to hedge random disruptions are major concerns.

The limitations imposed by the behavior and life cycle of the radiopharmaceuticals, procedure protocols, random disruptions, and resources availability make the management of patients in nuclear medicine facilities a complex problem. This leads to unique challenging scheduling issues that must be solved in order to provide a higher level of patient service. There exist a need for identifying those factors patients consider relevant when they evaluate the service received at the clinic. For example, waiting too long to get an appointment could increase patient discomfort which can be considered as an indicative of a low quality of service, although this may vary depending on the patient. In addition, performance measurements used to evaluate the nuclear medicine department profitability have to be identified. This research is aimed to identify a list of factors considered important when evaluating service effectiveness in nuclear medicine clinics. This will be a combination of both patients and management perspectives. Our goal is to elevate the level of service offered to the patients and to improve the way resources are managed by properly scheduling patients through the system.

B. Research Objective and Tasks

The objective of this research is to develop models and algorithms for improving patient service and resources management in highly constrained health care environments such as nuclear medicine clinics (a sub-specialty of radiology). We propose a methodology towards achieving this objective, including a simulation model that integrates with the scheduling methods to manage patient service levels and resource productivity, algorithms for scheduling multi-steps medical procedures that are time constrained, and optimizing methods for scheduling when dealing with random disruptions. We provide a description for each task in the following subsections.

1. Task 1: Modeling and Simulation of Nuclear Medicine Patient Service Management

We first derive and implement a discrete event system specification (DEVS) simulation model for nuclear medicine patient service and resource management. The model incorporates both patient and resource scheduling algorithms within the simulation framework and is capable of representing any nuclear medicine facility since it is automatically reconfigurable. This model provides a novel decision support system for assisting managers not only in patient and resource scheduling but also in assessing their daily scheduling decisions and system performance. A real nuclear medicine specialty clinic was modeled using this simulation. We provide computational results that quantify important trade-offs among the scheduling strategy used by a real nuclear medicine clinic.

2. Task 2: Algorithms for Nuclear Medicine Appointment Scheduling

The second task is to derive algorithms to improve the way patients and resources are scheduled in specialty clinics such nuclear medicine clinics. These algorithms provide patient appointments for multi-step procedures by considering both patient's and manager's interests. The algorithms are validated using a simulation model of an actual nuclear medicine setting and historical data. Computational results that provide useful insights into managing patients and resources in nuclear medicine specialty clinics are discussed.

3. Task 3: Stochastic Online Algorithms for Nuclear Medicine Appointment Scheduling

The last task is to derive a stochastic online scheduling algorithm for patient scheduling in nuclear medicine clinics. The algorithm take advantage of existing historical information on patient and radiopharmaceutical arrivals to make more informed decisions. A preliminary computational study is presented that demonstrate the benefits of considering stochastic information when performing appointment scheduling in specialty clinics.

C. Organization of the Dissertation

The dissertation has six chapters. Chapter II contains a concise review of literature on modeling and simulation in health care systems as well as patient scheduling in health care. Chapter III discusses the first task (Section 1) where a discrete event system specification (DEVS) simulation model for nuclear medicine clinics is derived. Computational results are provided for a scheduling algorithm currently used in a real nuclear medicine clinic. Chapter IV presents the second task proposed in Section 2.

Here algorithms for patients and resources scheduling are derived and implemented within a simulation model. Computational experiments are provided and trade-offs among the proposed algorithms are discussed. In Chapter V a stochastic online algorithm (Section 3) for patient and resource scheduling in nuclear medicine clinics is derived and a preliminary computational study is presented. Finally, Chapter VI contains the concluding remarks and the future research directions.

CHAPTER II

LITERATURE REVIEW

Over the past years health care organizations have faced pressures to deliver quality services to their patients while facing rising of costs [6]. This fact has motivated researchers and health care professionals to identify new approaches to improve the efficiency of health care operations and to reduce delivery costs. In this chapter, we provide a literature review of some of the methodologies that have been used to study health care systems. We focus on those topics that are aligned with the methodology proposed in this work.

We first provide a background of the use of discrete event simulation in health care systems analysis. Discrete event simulation is an operations research modeling and analysis methodology that permits system evaluation [7]. In health care, discrete event simulation can be used as a forecasting tool to assess the potential impact on changes on patient flow, to examine resources allocation, and/or to investigate the complex relationships among different system variables. The use of discrete event simulation has become increasingly more accepted by health care decisions makers. This is due in part to the large number of successful studies published in literature, as well as the development of software packages.

In addition to discrete event simulation, we provide a description of other methodologies that have been used to analyze health care systems. Some of them include queueing theory, dynamic programming, and mathematical programming. Most of the research found in literature for these methodologies focuses in the patient scheduling problem. The problem is addressed by deriving models that optimize a performance measure of an appointment system.

The rest of the literature review discusses research work in online scheduling. We describe several problem settings in which this methodology has been implemented and discuss how it can be adapted to address the problem of managing patient and resources in health care systems. This chapter is organized as follows. In Section A we provide a literature review on the use of discrete event simulation to analyze health care system. In addition, we provide a subsection that focuses on the research work that studies health care specialty clinics. Section B provides a description of the existing analytical research work that addresses patient scheduling in health care systems. We also provide a subsection for the work that is specialized in health care specialty clinics. Section C discusses the research work published in online scheduling and provides a description of how this framework can be adapted to the problem studied in this work. Lastly, Section D describes the performance measures frequently used in literature to assess the performance of health care systems.

A. Discrete Event Simulation in Health Care

Discrete event simulation [7, 8] has been used frequently to study patient appointment systems in health care over the past years. Karnon [9] compare discrete event simulation modeling to several techniques, such as Markov chain analysis, and conclude that simulation is particularly well suited for modeling health care clinics due to the complexity of such systems. This technique has been used to forecast the impact of changes in the system, to examine resource needs or to investigate the relationships between variables in a system [10]. Most simulation studies in health care focus on outpatient scheduling.

Ho and Lau [11] considers various rules for scheduling patient in outpatient clinics with the goal of minimize the weighted sum of personnel's and patients' idle time. They show that the idle times incurred by any of the rules are caused by three factors:

probability of no-show, the coefficient of variations of service times, and the number of patients per clinical session.

Klassen and Rohleder [12] study the best time to schedule patients with large service time differences and variances using simulation. They try several scheduling rules and take into consideration patient no-shows. They conclude that the best time to schedule those patients is toward the end of the day. In addition, they analyze the best time slots to leave open for potentially emergency patients and found no conclusive policy.

Edwards et al. [13] develop simulation models to compare two clinics that used two different queueing systems; serial processing and quasi-parallel processing. In serial processing patients wait in a single queue whereas in quasi-parallel processing patients are directed to the shortest queue to maintain flow. The authors showed that in quasi-parallel processing waiting times can be reduced up to a 30%.

Hashimoto and Bell [14] conducted a simulation study in an outpatient clinic that involve multiple sequential providers including a nurse, a physician, and a discharger. Their primary issue was to quantify the trade-off of patient waiting times versus idle time for doctors and staff. They concluded that increasing the number of physicians without increasing the supporting staff significantly increased the length of stay for the patients.

Swisher et al. [15] develop a simulation model that encompass both the operations and information center of a health care outpatient clinic. They consider scheduling more patients with significatively larger mean service times in the morning sessions rather than the afternoon. A reduction in the physicians' overtime was found by implementing a reduction in the physicians lunch time periods.

Groothuis et al. [16] investigate two patient scheduling policies in a hospital cardiac catheterization lab using simulation. The first policy does not schedule patients after 4:00 PM whereas the second policy schedule a fixed number of patients each day. They apply these scheduling policies to three different clinic configurations using patient thruput and working day duration as the performance measures. They found that one of the configurations was able to accommodate in average two additional patients per day while having less days exceeding eight hours.

Swisher and Jacobson [17] use an object oriented discrete event simulation to evaluate different staffing options and facility sizes for a two physician family practice health care clinic. They describe a measure for clinic effectiveness that takes into account both patient and physician satisfaction. They use this measure to evaluate the clinic configuration. The authors conclude that no specific configuration gives the perfect balance between profit and satisfaction and that is up to the decision maker to evaluate the trade-offs.

Rohleder and Klassen [18] consider discrete event simulation to study a rolling horizon appointment scheduling. In this study two management policies are considered; overload rule and rule delay. In overload rules overtime and double booking are used when patient demand is high, while the rule delay policy determines when the overload rule is implemented. The authors conclude that the best scheduling policy will depend on the performance measures that are considered more important by the decision makers.

Guo et al. [19] present a simulation framework for a doctor's appointment clinic that server for the evaluation and optimization of scheduling rules. Their simulation model captures four components of outpatient scheduling systems: external demand for appointments, supply of provider time-slots, patient flow logic and the scheduling algorithm. They demonstrate the framework for a pediatric clinic case and discuss the challenges for adapting the framework to other settings.

Surgical (operating room) center scheduling has also been studied using simulation modeling. Fitzpatrick et al. [20] study several scheduling techniques for hospital operating rooms that include: first-come-first-serve, fixed, variable, and mixed block. Fixed policy schedules the same block of time in the same time-slot each day of the week. The variable policy schedules based on seasonal demand fluctuations and mixed block scheduling is a combination of fixed and variable. The authors conclude that variable block scheduling is superior to the other alternatives in terms of resource utilization, patient throughput, and patient waiting time.

Akkerman and Knip [21] use discrete event simulation to study the facility dimensioning problem in surgery center. They showed that the number of beds can be reduced if recovering patients that no longer require specialized services are transferred. Butler et al. [22] study the sensitivity of patient misplacement with respect to a variety of modifications in their bed allocation policy. They conclude that reducing the length of stay and reallocating rooms among the different services within a hospital could decrease patient misplacement.

Simulation has been also applied to study hospital emergency departments. In emergency rooms patients arrive without an appointment most of the time and require treatment over a large and varied set of conditions. Although patients arrivals are highly unpredictable, the treatment sequence can be controlled by the staff. Lopez-Valcarcel and Perez [23] use discrete event simulation to evaluate the staffing levels, the arrival rates, and the service times in an emergency department. They recommend that the arrival rate should not exceed twelve patients per hour. Altinel and Ulas [24] use discrete event simulation to study the emergency department bed sizing problem. El-Darzi et al. [25] analyze the patient flow to reduce emergency department stay and increase patient throughput in emergency departments.

Garcia et al. [26] study the impact of having a fast-track queue in an emergency department. The fast-track queue is used to serve no urgent patients in an emergency room. They conclude that assigning a limited amount of resources to a fast-track lane significatively patient waiting times. Mahapatra et al. [27] presents a similar study that analyzes the effects of having a fast lane for treating non critical patients in an emergency department. The authors show that waiting times can be reduced by 10% by adding a fast track unit to the emergency department.

Kirtland et al. [28] study several alternatives to improve patient flow in an emergency department. Their goal is to reduce time patient spent at the facility and determine the appropriate staffing levels using simulation. They found that by using a fast track lane in minor care, placing patients in the treatment area instead of sending them back to the waiting room, and using a point-of-care lab testing patient waiting times can be reduced by an average of thirty eight minutes.

McGuire [29] uses simulation to determine how to reduce the length of stay for patients in an emergency department. The author conclude that adding an additional clerk during peak hours, extending the operating hours of the fast-track lane, adding a holding area for waiting patients, and using physicians instead of residents in the fast track area significatively reduced the amount of time patients spent at the clinic.

Badri and Hollingsworth [30] analyze the effects of several scenarios on scheduling, limited staffing, and changing the patient demand patterns in an emergency room. They conclude that serving only those patients that needed urgent care and eliminating one or more doctors on each shift provides the better outcomes for the clinic under study.

Samaha et al. [31] use discrete event simulation to reduce patient length of stays in a hospital emergency department. The authors determined that the length of stay was a process related problem rather that resource dependent. They show that adding additional beds did not shorten the length of stay of the patients. For more information about simulation studies in health care systems we refer the reader to some survey papers that have been written addressing this topic [32, 6, 33, 34, 35].

1. Discrete Event Simulation of Specialty Clinics

Specialty clinics bring their own set of problem characteristics when scheduling patients and allocating space within health care facilities. However, there is limited research reported about the use of discrete event simulation for the problem of managing patient in specialty clinics such as nuclear medicine departments. Most of the research reported focuses in radiology departments.

Walter [36] develop a simulation model of a hospital radiology department to predict the effects of scheduling policies on the efficiency of the appointment system, as measured by the average patient queueing time and doctor idle time during the day. He conclude that staff time savings were possible by segregating patients into inpatient and outpatient sessions with a similar examination time distribution. Additionally, he found that overbooking yield a small increase in staff utilization while substantially increased patient waiting time at the clinic.

Kho and Johnson [37] developed a computer model that simulates patient flow through a radiology facility. This model was used to identify causes of congestion and low productivity and to predict effects of changes in the system. They conclude that performance can be improved by distributing patient demand for outpatient services evenly.

Johannes and Wyskida [38] develop a model for scheduling patients and clinical instruments in a nuclear medicine department that minimizes the equipment idle time. The authors used simulation to test the shortest-processing-time-first rule to schedule several patient classes in a nuclear medicine department. Only a limited number of procedures were studied and their heuristic assumes that the group of patients requiring service are known at the beginning of the day.

Sepúlveda et al. [39] use simulation to study a full service cancer treatment center with the objective of analyzing patient flow through the clinic, evaluate the impact of alternative floor layouts, using different scheduling options. The simulation of three key scenarios were used to identify patient bottlenecks and improved patient management at the center. Simulating different scheduling policies showed a 20% increase in the number of patients seen per day.

Centeno et al. [40] perform a simulation study in a radiology department which objective was to determine the adequate number of technologists assisting per operation while maximizing the utilization of the staff. The idle time of resources and number of daily procedures were consider as performance measures for the system. Ramakrishnan et al. [41] use simulation to study different scenarios of a radiology service department. Their goal was to identify patient flow changes in the computerized tomography (CT) scan area that would maximize patient throughput. An increase of 20% in patient throughput was found by making changes to the CT scan area. Other research work on the use of simulation to analyze staff allocations to improve patient flow in radiology departments includes O'Kane [42] and Klafehn [43].

B. Scheduling in Health Care

In this section we present a summary of analytical methodologies used for patient scheduling in health care facilities. Liao et al. [44] consider the problem of scheduling the optimal arrival times of K customers. Each request is to be assigned to one of K equal time durations. The authors studied a static and a dynamic versions of the planning arrival problem. For the dynamic case planning decisions are made at the beginning of each time slot, whereas for the static problem decisions are made at the beginning of the first time slot. Dynamic programming was applied to determine the optimal block sizes. The solution obtained is used as lower bound to solve the static problem by a branch-and-bound algorithm, which is restricted to small size problems.

Wang [45] studies both the static and dynamic case for a single server system with exponential service times and try to minimize the weighted sum of customer flow time and system completion time. The author shows that customer flow times can be represented by a phase type distribution and calculates the optimal appointment times using a recursive procedure. The author concludes that the optimal appointment intervals are not constants but dome shaped. He later extended the study to any service time distribution that can be approximated with a phase-type distribution [46].

Liu and Liu [47] consider the case of a queueing system with multiple doctors. They study the effects of patient no-shows and doctors lateness in an operation session divided into a number of identical periods and they schedule a group of jobs to serve at the beginning of each period. They develop a dynamic programming formulation that allow them to find optimal block sizes and use the results to solve the static problem. The authors compare their results with those obtained using exhaustive simulation.

Penneys [48] studies the effects of block versus sequential scheduling on patient waiting times, length of patient encounters, and physician patient free time in two busy dermatologist clinics. He concludes that under block scheduling physicians enter the exam room earlier, increase patient free time during the day, and the clinic finishes 35 minutes before on average whereas patient waiting times remain about the same under both policies. Chung [49] proposes a modified block scheduling method in which a double book of patients is performed at the beginning of each hour and the end of the hour is left open to catch up in case physician runs behind schedule. The author claims an improvement of 15% on patients waiting time.

Lau and Lau [50] address two problems relevant to outpatient and surgical scheduling: a cost estimation given a particular scheduling rule and finding the optimal schedule based on a particular sequence of arrivals. They present a procedure to solve the first problem when service times are nonidentical distributed. They evaluate the accuracy of their methods using simulation.

Cayirli et al. [51] use a classification scheme of "new/return" to analyze the effects of sequencing in appointment scheduling. They consider the effects of accepting patient walk-ins and the issue of patient punctuality. They claim that sequencing decisions have more impact on clinic performance than the choice of an appointment rule. They conclude that unexpected events such as walk-ins, no-shows, punctuality and session volume have a great influence in the effectiveness of appointment systems.

Robinson and Chen [52] consider the problem of finding optimal appointment times when a sequence of N patients has been specified. They formulate the problem as a stochastic linear programming and solve it using Monte-Carlo integration. They develop a heuristic that adjust appointment intervals using as a basis the structure of the optimal policy.

Denton and Gupta [53] study the problem of determining the optimal appointment times for a sequences of jobs with uncertain duration. They present a two-stage stochastic linear programming model to determine the optimal appointment intervals by considering the expected cost of patient waiting, server idling and cost of tardiness with respect to a chosen session length. They apply a decomposition approach to solve it for a general i.i.d service times.

Brahimi and Worthington [54] present a queueing model to outpatient clinics. They study the finite capacity multi-server queuing model with nonhomogeneous arrivals and general discrete service times distributions. They claim that by using analytical techniques, waiting times can be substantially reduced without increasing server idleness.

Gerchak et al. [55] study the reservation of surgical capacity for emergency cases on a daily basis when these rooms are also used for elective surgeries. They formulate the problem as a stochastic dynamic program assuming that surgery durations are independent and identically distributed. They claim that the optimal policy is a function of the patients waiting for deferrable surgeries. They develop an algorithm that finds the optimal number of deferrable surgeries to schedule on any given day.

Kaandorp and Koole [56] consider a single server case with exponential service and a single no-show probability for all patients. They propose a model that minimizes the weighted average of the expected waiting time, server idle time, and overtime. The authors claim that a local search can be used to obtain a globally optimal schedule because the model is multi-modular.

Gupta and Wang [57] use a Markov Decision Process model to study the problem of which appointment request to accept to maximize the revenue of a primary care clinic. They consider patient choices and claim that for a single server case an optimal solution can be provided. For more information about scheduling studies in health care systems we refer the reader to some survey papers that have been written addressing this topic [58, 59].

1. Specialty Clinics Appointment Scheduling

Several papers study variations of the problem of scheduling patient in specialty clinics using optimization techniques. For example, Conforti et al. [60] study optimization models for outpatient scheduling within a radiotherapy department, whereby patients have to visit the treatment center several times during the week.

Green et al. [61] address the problem of scheduling randomly arriving patients of different types in an MRI facility. They formulate the problem as a finite-horizon dynamic program for an appointment schedule that allows at most one patient per period and a single server, where only one patient can be served at a time. The authors derive properties of the optimal scheduling policies and identify a service sequence that minimize the expected total cost of serving patients in a diagnostic facility. Patrick et al. [62] study a similar problem but they characterize patients with different priorities. Kolisch and Sickinger [63] consider a similar problem but with two CT scanners. The authors compare three decision rules under three different appointment schedules.

Patrick and Puterman [64] consider the problem of scheduling patients in a CT scan department. They formulate an optimization problem that returns a reservation policy that minimizes the non-utilization of resources subject to an overtime constraint. Their approach assumes the use of a pool of patients that can be called to occupy unused time slots. The authors use simulation to demonstrate a reduction in outpatient waiting time.

Sickinger and Kolisch [65] propose a generalization of the Bailey-Welch rule as well as a neighborhood search heuristic for a medical service facility with two resources. The Bailey-Welch rule claims that for one server clinic the best performance in terms of patient waiting and server idle time is to schedule two patients for the first appointment space and one patient on the ones that follow. The authors analyze the impact of different problem parameters on the total reward.

Standridge and Steward [66] propose a simulation model that includes a control logic for patient scheduling. The system presented by the authors schedules patients within a simulation framework. Vermeulen et al. [67] devise an adaptive approach to automatic optimization of resource calendars in a CT scan facility. They implement a simulation model for a case analysis to demonstrate that their approach makes efficient use of resources' capacity.

C. Stochastic Online Scheduling

Online scheduling occurs naturally in several applications areas. Contrary to offline optimization problems, data is not available in advance in online optimization. Online stochastic optimization assumes the distribution of future requests, or an approximation thereof, is available for sampling [68]. The typical case is the existence of either historical data or predictive models. Online stochastic optimization problems are also constrained by time constraints, meaning that they have a limited time to find a solution and make a decision for the problem.

Literature shows the benefits of taking into account future events when optimizing decision processes online. It has been shown that using additional stochastic information can improve the quality of solutions in scheduling applications such as: dynamic vehicle routing [69, 70, 71] , packet scheduling [72, 69, 73, 68], reservation systems [74], inventory management [70, 75], organ transplants [76], and elevator dispatching [77]. In these sample applications, stochastic information is exploited in widely different ways; however, the unifying theme seen throughout this research is that there are considerable advantages to taking account of stochastic information.

The common strategy is to predict the future requests using a statistical model by sampling the observations on the history. Chang et al. [72] study the multiclass packet-network scheduling problem. The authors use a Hidden Markov Model (HMM) to generate the tasks arrivals for each class with a particular weight and develop an algorithm named expectation that adapts an optimal offline algorithm into an online algorithm by sampling possible future tasks sequences from the HMM. Unfortunately, the expectation algorithm does not perform well under time constraints, since it must distribute its available optimizations across all requests.

This issue was recognized and addressed in [68] where a consensus algorithm was proposed. The consensus algorithm solve as many samples as possible and to select the request which is chosen most often in the sample solutions at time t. The consensus algorithm was shown to outperform the expectation algorithm on online packet scheduling under time constraints. However, as decision time increases, the quality of the consensus algorithm levels off and is eventually outperformed by the expectation algorithm. The regret algorithm proposed in [69] combines the features of both expectation and consensus algorithms. It evaluates every decision on all samples like the expectation algorithm and has the ability to avoid distributing the samples among decisions like the consensus algorithm.

Awasthi and Sandholm [76] consider the scheduling of human kidney transplants using a stochastic online framework. They propose an adaptation of the regrets algorithm proposed by Bent and Van Hentenryck [69]. Van Hentenryck et al. [74] consider the online stochastic reservation problem where the goal is to allocate requests that are received online to limited group of resources in order to maximize the profit (Multi-knapsack problem). The authors adapted the consensus and regret algorithms to their problem. A modification of the regret algorithm is presented that is based on a constant sub-optimality approximation of multi-knapsack problem [78]. The authors used two black-boxes to handle the stochastic arrivals of reservation requests for hotel rooms. One black-box is the sub-optimal approximation module and the other is the sampling module which relies on the observations of the past arrivals. This problem in general differs from the scheduling problems considered in [72, 68] and [79]; mostly because the approach followed is not about selecting the best request but rather about how best to serve a request.

The expectation algorithm has some resemblances to the sample average approximation method for non-dynamic stochastic programming [80, 81] where the solution depends of a deterministic part and a stochastic part. The deterministic part gives the immediate plan and the stochastic part gives a penalty for changing the plan to accommodate the best as possible the scenario that has become reality. One must average over many scenarios to find the best expected solution. The sample average approximation has been applied to the stochastic vehicle routing problem [82].

In this work we consider using sampling of possible future procedure requests to obtain an appointment date (day and time) for a patient. In other words, finding a schedule by considering the possible requests that may occur after the procedure request in hand. Conceptually the historical information provided by the clinic can be incorporated in the classical way into a multistage integer stochastic program [83] which would then be solved to come up with an optimal scheduling plan. However such stochastic optimization techniques are not capable of efficiently solving problems of the size of the required domain. In contrast, online stochastic algorithms are suboptimal but scalable ways of solving stochastic integer programs [84, 85]. The idea is to sample a subset of the future scenarios (trajectories), solve the offline problem on each of them using a two stage stochastic integer programming (SIP) model [86], assign a score to each possible action, and select the action that is the best overall.

D. Performance Measurements

Patient satisfaction in outpatient clinics may be difficult to quantify since it depends on the way patients perceive the service received. Several performance measures have been identified in the literature as the most commonly used for evaluating patient service satisfaction in health care clinics. Waiting time *Type 1*, is the time a patient waits from the time he/she calls for an appointment until the date of the appointment [18, 87, 88]. Waiting time *Type 2* is the time a patient waits from the time he/she arrives at the clinic to the time when service is started [18, 89, 90, 40]. The percentage of time a patient requests for an appointment and is satisfied [61, 57] and the time the patient spends in the system [91, 92] were the last ones identified.

Besides patient satisfaction, health care managers are concerned with the profitability of the business. In particular, nuclear medicine department managers understand that providing a high level of service to their patients is important for the business. But this requires improving other areas such as human resource overtime [89, 90], resource utilization [18, 90, 40], and patient throughput [41]. Those performance measures have been used commonly in literature to represent the management's perspective.

CHAPTER III

MODELING AND SIMULATION OF NUCLEAR MEDICINE PATIENT SERVICE MANAGEMENT IN DEVS *

A. Introduction

Health care costs in the U.S. have increased in recent years and now exceed those in other nations that provide similar, or better care for their citizens. Increased demand for specialized services has been identified as one of the causes of this trend in U.S. health care costs [3]. Speciality clinics such as nuclear medicine, a sub-specialty of radiology, use appointment scheduling systems to manage patients. These clinics are affected by many factors such as patient behavior, staff experience, service time variability, equipment failures, and radiopharmaceuticals management, which have an impact on the way the appointment systems perform. This chapter focuses on patient service management in nuclear medicine, which uses new technology to treat and diagnose patients. Nuclear medicine procedures/tests include positron emission tomography(PET) scan, imaging test, heart stress and radiotherapy for lymphoma. Most of the tests require administering a radioactive isotope or radiopharmaceutical in order to take high quality images deep within the body and involve multiple scans during one day, or on multiple days. To successfully perform a nuclear medicine test, all the resources needed for each step of the test must be available at specific times. If the test is not completed successfully, the patient must be re-scheduled for another day. Therefore, scheduling patients, radiopharmaceuticals and resources to avoid

*Reprinted with permission from "Modeling and Simulation of Nuclear Medicine Patient Service Management in DEVS" by E. Pérez, L. Ntaimo, C. Bailey, and P. McCormack, 2010. *Simulation*, Online First, Copyright [2010] by The Society for Modeling and Simulation International.

re-scheduling is a very challenging problem for nuclear medicine departments. Furthermore, the characteristics of patient and resource management in nuclear medicine make it a unique problem with little research work reported in the literature. The limitations imposed by the behavior and short life cycle of the radiopharmaceuticals, combined with the different types of patient arrivals, random disruptions and resource availability, make the management of patient service in nuclear medicine a complex problem.

The contributions of this research include the first (to the best of our knowledge) DEVS simulation model for nuclear medicine patient service management. The model represents an advance toward improving patient service in health care with innovations in the way the model is represented and implemented. The model incorporates both patient and resource scheduling algorithms within the simulation framework. This in essence provides a novel decision support system for assisting managers not only in patient and resource scheduling, but also in assessing their daily scheduling decisions on system performance. The simulation model enables system-level performance assessment, identification of potential bottlenecks, and integration of scheduling and patient flow analysis. A computational study to quantify important trade-offs between a patient and resource scheduling strategy currently used in a real nuclear medicine clinic and variations of this strategy is presented. While this work focuses on nuclear medicine, the results can be applied to many other systems that are not as complex as nuclear medicine. These include diagnostic imaging areas such as magnetic resonance imaging (MRI) and computed axial tomography (CT scan).

In nuclear medicine, a typical test requires at least three resources: a radiopharmaceutical; gamma camera; technologist; and, sometimes, a nurse or EKG (electrocardiogram) technician. Nuclear medicine equipment may cost up to a million dollars and therefore must be managed efficiently. The schedule for radiopharmaceutical delivery, injecting the patient, and scanning must adhere to a specified protocol since radioactivity decays over time. For example, a scan made too early before the radiopharmaceutical has diffused adequately, or too late after excessive decay has occurred, results in a poor image. If there is too much delay, the procedure must be terminated and repeated on another day, causing unnecessary exposure to radiation, poor utilization of resources, and increased cost. Furthermore, since relatively few pharmacies supply radiopharmaceuticals, the dosages and delivery schedules depend on the lead time (hours to days) required to supply the radiopharmaceuticals. A well-designed system for patient service management in nuclear medicine has to consider the goals of both managers and patients.

A viable approach to address the challenging problem of managing patient service in nuclear medicine is modeling and simulation (M&S). In this chapter, a discrete event M&S approach for managing patient service in nuclear medicine is considered. In particular, the *discrete event system specification* (DEVS) formalism [8] is used to derive a generic simulation model for a nuclear medicine patient service management that can be tailored to any real nuclear medicine clinic. DEVS is a formal M&S framework based on dynamical systems theory and provides well defined concepts for coupling components, hierarchical and modular model construction, and an object-oriented substrate supporting repository reuse. Modular construction is one of the most important characteristics of DEVS because it allows the modeler to design and construct each model independently for optimal efficiency. As long as models adhere to certain protocols, they can interact with each other. In this work, the patient and management perspectives are considered. Both points of view are very important for developing patient and resources scheduling policies, and for evaluating the performance of patient service and resource utilization. Patients are concerned with the level of service offered by the department while managers are also concerned with using their limited resources effectively.

The rest of the chapter is organized as follows. In Section B preliminaries on DEVS are provided. The simulation model derivation, hierarchical structure, operation and implementation is presented in Section C. Section D presents a computational study and some concluding remarks and directions for future research are presented in Section E.

B. Preliminaries

To provide a mathematical foundation for the proposed simulation models, we first review some preliminaries on DEVS. The reader familiar with DEVS may skip this section. In this work we use the parallel DEVS formalism [8] to construct the simulation model for a nuclear medicine department. Parallel DEVS is a revision of the classical DEVS formalism [8]. It uses a hierarchical approach to build complex models starting with the basic or *atomic* model, and then coupling the atomic models to create *coupled* (composite) models. An atomic model has to be in a defined state at any time and has input and output ports through which all interaction with the environment is mediated. External events arising outside the model are received through the input ports, and the model description determines how the model responds to them. All internal events arising within the model change its state and manifest themselves as events on the output ports to be transmitted to other models. Communication between models is enabled via the couplings.

Unlike classical DEVS, parallel DEVS allows all imminent components to be activated and send their outputs to other components of the system. DEVS has a well defined concept of system modularity and component coupling to form coupled models. This leads to the property of closure under coupling which justifies treating coupled models as components and enables hierarchical model composition construct. Since several DEVS simulators have already been developed and validated, the models developed in this work will be executed by the existing DEVS simulators. Consequently, the focus of this chapter is on simulation modeling and not simulation algorithms.

Let M denote an atomic model with a set of input ports $IPorts$, a set of input values (events) X_p, a set of output ports $OPorts$, and a set of output values (events) Y_p. We denote by (p, v) the port-value pair. Then a basic parallel DEVS is a structure defined as follows [8]:

$$DEVS = (X_M, Y_M, S, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, ta) \tag{3.1}$$

where,

$X_M = \{(p, v)|p \in IPorts, v \in X_p\}$ is the set of input ports and values;

$Y_M = \{(p, v)|p \in OPorts, v \in Y_p\}$ is the set of output ports and values;

S is the set of sequential states;

$\delta_{ext} : Q \times X_M^b \rightarrow S$ is the *external transition* function, where X_M^b is a set of *bags* over elements in X_M and Q is the set of total states;

$\delta_{int} : S \rightarrow S$ is the *internal state transition* function;

$\delta_{con} : Q \times X_M^b \rightarrow S$ is the *confluent transition* function;

$\lambda : S \rightarrow Y_M^b$ is the output function;

$ta : S \rightarrow R_{0,\infty}^+$ is the *time advance* function; and

$Q := \{(s, e)|s \in S, 0 \leq e \leq ta(s)\}$ is the set of total states, where s is the state and e is the elapsed time.

Note that a *bag* is a set with possible multiple occurrences of its elements. This allows parallel DEVS to handle multiple inputs. Equation (3.1) can be interpreted as follows: At any time the system is in some state s and if no external events occur, the system will not change state for a time $ta(s) \in [0, \infty]$. When this time expires the system outputs the value, $\lambda(s)$, and changes to state $s' = \delta_{int}(s)$. An output is only possible after an internal transition. If an external event $x \in X_M$ occurs when the system is in total state (s, e) with $e \leq ta(s)$, i.e., before expiration time, the system changes to state $s' = \delta_{ext}(s, e, x)$. The external transition function dictates the system's new state when an external event occurs, while the internal transition function dictates the system's new state when no events occurred since the last transition. The confluent function decides the next state in cases of collision between external and internal events.

The DEVS formalism includes the means to construct models from components. The specification includes the external interface, the components (DEVS models), and the coupling relations. Let EIC, EOC and IC denote the external input coupling, external output coupling and internal coupling, respectively. Then a coupled model N can be defined mathematically as follows:

$$N = (X, Y, D, \{M_d \mid d \in D\}, EIC, EOC, IC) \tag{3.2}$$

where,

$$X = \{(p, v) | p \in IPorts, v \in X_p\}$$

is the set of input ports and values and

$$Y = \{(p, v) | p \in OPorts, v \in Y_p\}$$

is the set of output ports and values. D is the set of component names, and for each

$d \in D$,

$$M_d = (X_d, Y_d, S, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, ta)$$

is a DEVS model with

$$X_d = \{(p, v) | p \in IPorts_d, v \in X_p\}$$

and

$$Y_d = \{(p, v) | p \in OPorts_d, v \in Y_p\}.$$

The external input coupling, EIC, connect external inputs to component inputs:

$$EIC \subseteq \{((N, ip_N), (d, ip_d)) \mid ip_N \in IPorts, d \in D, ip_d \in IPorts_d\}. \tag{3.3}$$

The external output coupling, EOC, connect external outputs to component outputs:

$$EOC \subseteq \{((N, op_d), (N, op_N)) \mid op_N \in OPorts, d \in D, op_d \in OPorts_d\}. \tag{3.4}$$

Lastly, the internal coupling, IC, connect component outputs to component inputs:

$$IC \subseteq \{((a, op_a), (b, ip_b)) \mid a, b \in D, op_a \in OPorts_a, ip_b \in IPorts_b\}. \tag{3.5}$$

We should point out that in DEVS no output port of a component may be connected to an input port of the same component, i.e., $((a, op_a), (b, ip_b) \in IC$ implies $a \neq b$. In other words, no direct feedback loops are allowed for each component. Armed with the above characterizations, we are now in a position to derive several atomic and coupled DEVS models for nuclear medicine patient service management.

C. Simulation Model

The practical setting of a nuclear medicine department involves several resources, which include humans and equipment, procedures/tests and performance measures.

We start by describing these entities in the context of model abstraction and then derive the corresponding atomic and coupled models that constitute the nuclear medicine simulation model.

1. Model Abstraction

We conceptualize a nuclear medicine department model involving human and equipment resources, stations, and patients. We classify these entities by considering their roles and the interactions they have within the model.

a. Human Resources

We distinguish between four types of human resources: technologists, nurses, physicians, and managers. We capture the behavior of each human resource by taking into account the expertise and experience. Human resources that have been executing their tasks for several years are expected to complete their tasks relatively faster than those who have less experience. The set of activities each type of human resources can perform depends on the expertise. Table I lists some of the activities that can be performed by each type of human resource.

In our simulation model we represent each human resource type as a separate atomic model, capable of receiving messages containing their schedules. A schedule includes times (and stations) when the human resource will serve patients. When it is time to serve a patient the human resource travels to the appropriate station according to the schedule. Travel time from the human resource's office to each station is known.

Table I. Human resources responsibilities in nuclear medicine

Human resource	Responsibilities
Technologist	Hydrate patient; Radiopharmaceutical preparation; Imaging
Nurses	Hydrate patient; Radiopharmaceutical administration; Draw doses
Physicians	Hydrate patient; Radiopharmaceutical administration; Draw doses
Managers	Hydrate patient; Radiopharmaceutical preparation; Radiopharmaceutical administration

b. Procedures/Tests

Nuclear medicine procedures/tests are essential in medical specialties such as cardiology, pediatrics and psychiatry. The procedures are usually requested by physicians by calling the nuclear medicine clinic to ask for an appointment for a patient. The procedures provide physicians with information about structure and function of the human body (diagnosis) but are also used for disease treatment. Table II lists several nuclear medicine procedures and their current procedural terminology (CPT) codes.

Each procedure images a specific organ and requires the administration of at least one radiopharmaceutical. The number of steps for each procedure may range from 3 to 11. The duration of each step may vary depending on the experience of the human resource in charge, however it must be completed within the time window stipulated by protocol for the procedure. As an example, the CVE/SP-M procedure is described in Table III. This procedure involves four steps and requires the use of two different

Table II. Examples of nuclear medicine procedures

CPT Code	Name
78465	Cardiovascular Event (CVE) Myocardial Imaging (SP-M)
78815	PET CT skull to thigh
78306	MSB-bone imaging (whole body)
78315	MSC-bone imaging (three phase)
78223	GIC-Hepatobiliary imaging
78472	CVJ-cardiac blood pool
78585	REB-Pulm perfusion / ventilation
78006	ENC-Thyroid imaging
78195	HEE-Lymphatic imaging
78464	CVD-Myocardial imaging

radiopharmaceuticals. This procedure takes a minimum of 95 minutes to complete and requires the involvement of at least two human resources. Table IV shows the steps for the PET/CT procedure. In this procedure only one radiopharmaceutical is needed.

Table III. Procedure 78465: cardiovascular event (CVE) myocardial imaging (SP-M)

Step	Activity	Time (min.)	Station	Human Resource
1	Hydrate patient and dispense dose	10	TRT 1, 2 & 3	Technologists; Nurse; Manager
2	Stress EKG	30	Treadmill 1 & 2;	EKG Tech
3	Patient Wait	30-60	Waiting room	
4	Imaging	30	Axis 1, 2, & 3; P2000A & B, P3000	Technologists; Nurse; Manager

Table IV. Procedure 78815: computed tomography (CT) skull to thigh

Step	Activity	Time (min.)	Station	Human Resource
1	Hydrate patient	10	TRT 1, 2 & 3	Technologists; Nurse; Manager
2	Dispense dose	5	Meridian; TRT 1, 2 & 3; Axis 1, 2, & 3; P2000A & B	Technologists; Nurse; Manager
3	Patient Wait	60	Waiting room	
4	Imaging	45	Axis 1, 2, & 3; Meridian; P3000, P2000A & B	Technologists; Manager

c. Stations and Equipment

A nuclear medicine department has separate stations for performing specific procedures/tests. We conceptualize stations as spaces where patients are served by human resources. Every station has at most one type of nuclear medicine equipment. We classify stations based on the type of equipment in the station. Nuclear medicine equipment includes different types of gamma cameras and treadmills. Our simulation model incorporates an atomic model for each equipment, an atomic model that represents a room, and a coupled model for a station. The station coupled model represents both the room space and the equipment in the station. Before starting any activity at any station, the model has to verify that all the entities needed to perform the procedure have arrived. For example, to administer a radiopharmaceutical the human resource, patient and the radiopharmaceutical needed to perform this activity must be in the station. The time spent by these entities in the station will depend on the protocol for performing the procedure/test, experience of the human resource, and the type of equipment involved.

d. Patients

Patient service requests are usually managed by a call center. A receptionist is in charge of taking care of these requests by finding an appointment for the patient. In the simulation model we represent the call center by a scheduler atomic model. This atomic model is in charge of patients' schedules. We also have a call generator atomic model that generates patient service requests during the simulation. Patients will always ask for one procedure and, in some cases, they will also provide a day of preference for the appointment. The scheduling of patients depends on the algorithm or rules available in a given nuclear medicine department. Schedule information is passed to all the models involved in the scheduling. First, human resources whose schedules were affected by the inclusion of a new patient are notified. Secondly, notifications are sent to the models that are in charge of generating patients and radiopharmaceuticals at the time of the appointment.

e. Performance Measures

Since our nuclear medicine patient service management approach involves both patient and management perspectives, models are designed to capture important information pertaining to both perspectives. We derive a transducer atomic model, which is responsible for collecting this information and for computing statistics of interest to the modeler. In particular, we use performance measures that have been seen in the literature. Table V gives the selected performance measures used to evaluate patient service satisfaction in health care clinics.

Besides being concerned about the quality of service they provide to their patients, nuclear medicine managers have to watch for the profitability and operation of the business. Table VI gives the performance measures that are commonly used in the

Table V. Performance measures for patient satisfaction in health care

Name	Description	References
Waiting time Type 1	Time patient wait from the time of calling for an appointment until the date of the appointment	[18]
Preference satisfaction ratio	Number of times patient request for an appointment is satisfied	[61]
Cycle time	Time patient spent on the system	[91], [92]

Table VI. Managers performance measures

Name	Description	References
Equipment utilization	Maximize utilization	[18],[90]
Human resource utilization	Most adequate number of human resources while maximizing utilization	[90],[40]
Patient throughput	Number of patients served per day	[41]

literature from a management's perspective in health care clinics.These measures are used in our simulation model to assess the system performance of a nuclear medicine clinic based on the patient and resource scheduling algorithm used by the clinic. We discussed the algorithm in detail in Section D.

2. Atomic and Coupled Models

We now discuss each of the atomic and coupled models that were developed as part of the overall simulation model. We derive an atomic model for each of the following human resources: Manager (MANGR), Technologist (TECH), Nurse (NURSE), Receptionist (RCPST) and Physician (PHYS). Similarly, we derive an atomic model for all Equipment (EQUIP) resources including Treadmill, Axis Gamma Camera, TRT Gamma Camera and P2000 Gamma Camera. We also derive an atomic model for nuclear medicine department rooms (ROOM) such as the Waiting Room, Treadmill Room, Axis Camera Room, TRT Camera Room and P2000 Camera Room. Note that the names of the rooms are based on the equipment inside the room. So by coupling the ROOM and EQUIP atomic models, we build a nuclear medicine department station (STATION) coupled model. We should point out that atomic models are generic representations of the different types of entities used in nuclear medicine. So in the simulation the models are instantiated with different attributes to represent the unique or different entities in nuclear medicine.

Next, we couple the above models to build a nuclear medicine department coupled model called NMD. To model call requests, patient arrivals, and radiopharmaceutical arrivals, we derive atomic models for Call Generator (CGEN), Scheduler (SCHED), Scheduled Patient Generator (PGENR), Radiopharmaceutical Generator (RPGENR) and Transducer (TRANSD). These models are coupled together to build an Experimental Frame (EF) coupled model. The overall simulation model results from coupling NMD and EF.

We are now in a position to provide mathematical descriptions of the models. However, due to space limitation we focus only on three atomic models, TECH, EQUIP, and SCHED, which are critical to the operation of a nuclear medicine department. In what follows, $\times$ denotes the cartesian product of sets and $\wedge$ denotes the logic AND operation.

a. TECH Atomic Model

We consider a TECH atomic model having input and output ports as shown in Figure 1. The model has three basic input ports, namely; “in”, “set”, and “update”, and two types of output ports, namely; "out" and "room$x$". The number of output ports of type "room$x$", $x = 1, \cdots, n$, depends on the number of rooms n in the nuclear medicine facility. The input port "in" is for receiving a message to activate the model at the beginning of the simulation. A new schedule for the technologist is sent through the "set" port, while any updates to the schedule are sent through the "update" port. The "out" output port allows for transmitting information to the TRANSD atomic model. The rest of the output ports transmit information to the rooms associated with the technologist. The TECH (technologist) atomic model

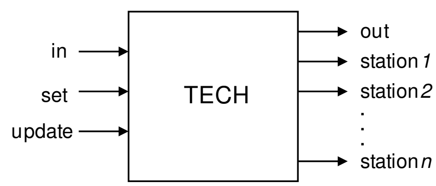

Fig. 1. A basic TECH (technologist) atomic model

has eight basic states, namely; *idle, get_schedule, waiting, update_schedule, travel_to, travel_from, serve_patient* and *wait_here*. We depict the behavior of the atomic model using the state (transition) diagram shown in Figure 2. Mathematically, the TECH atomic model can be defined in parallel DEVS as follows:

$$DEVS_{TECH} = (X_M, Y_M, S, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, ta) \tag{3.6}$$

where,

$$X_M = \{(p, v) | p \in IPorts, v \in X_p\}$$

is the set of input ports and values, where

$$IPorts = \{\text{"in"}, \text{"set"}, \text{"update"}\},$$

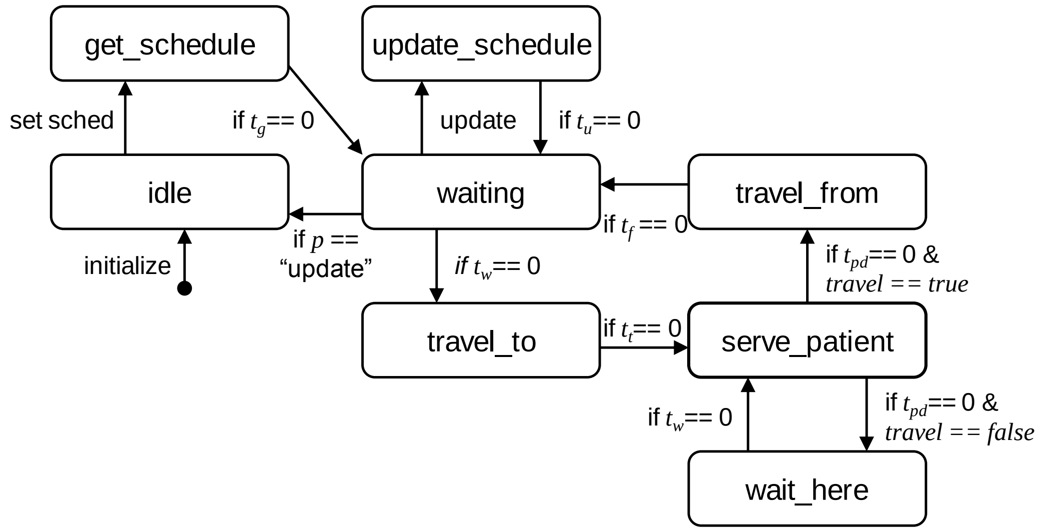

Fig. 2. State transition diagram for TECH atomic model

and for $p =$"in"; $X_{\text{in}} = V_1$; for $p =$"set", $X_{\text{set}} = V_2$; and for $p =$"update", $X_{\text{update}} = V_3$. The sets V_1, V_2 and V_3 are arbitrary sets. The set

$$Y_M = \{(p, v)|p \in OPorts, v \in Y_p\}$$

is the set of output ports and values, where

$$OPorts = \{\text{"out"}, \text{"station1"}, \text{"station2"}, \cdots, \text{"station}n\text{"}\}.$$

The sets Y_{out}, Y_{station1}, Y_{station2}, $\cdots$, $Y_{\text{station}n}$ are arbitrary sets. The set

$$S = \{\text{"idle"}, \text{"get\_schedule"}, \text{"waiting"}, \text{"travel\_to"}, \text{"serve\_patient"},$$

$$\text{"wait\_here"}, \text{"travel\_from"}\} \times \Re_{+,0} \times V_1 \times V_2 \times V_3$$ is the set of sequential states.

External Transition Function:

$$\delta_{ext}((phase, \sigma, sched), e, (p, v))$$

$$= (\text{"idle"}, \infty, sched), \quad \text{if } p == \text{"in"}$$

$= (\text{“get\_schedule”},\, t_g, sched),\quad \text{if } phase == \text{“idle”} \wedge p == \text{“set”}$

$sched = computeEventDelays();$

$= (\text{“waiting”},\, t_w, sched),\quad \text{if } phase == \text{“get\_schedule”}$

$t_w = \text{getWaitDelay}();$

$= (\text{“update\_schedule”},\, t_u, sched),\quad \text{if } phase == \text{“waiting”} \wedge p == \text{“update”}$

$sched = updateEventDelays();$

$= (\text{“idle”},\, \infty, sched),\quad \text{if } phase == \text{“waiting”} \wedge sched == \emptyset$

$= (phase, \sigma - e, sched)$, otherwise.

Internal Transition Function:

$\delta_{int}((phase, \sigma, sched), e, (p, v))$

$= (\text{“travel\_to”},\, t_t,\, sched),\quad \text{if } phase == \text{“waiting”} \wedge travel == true$

$t_t = getTravelDelay();$

$stationID = getStationID();$

$= (\text{“serve\_patient”},\, t_{pd},\, sched),\quad \text{if } phase == \text{“travel\_to”}$

$t_{pd} = getProcedureDelay();$

$= (\text{“wait\_here”},\, t_w,\, sched),\quad \text{if } phase == \text{“serve\_patient”} \wedge travel == false$

$t_w = getWaitDelay();$

$= (\text{“serve\_patient”},\, t_{pd},\, sched),\quad \text{if } phase == \text{“wait\_here”}$

$t_{pd} = getProcedureDelay();$

$= (\text{“travel\_from”},\, t_f,\, sched),\quad \text{if } phase == \text{“serve\_patient”} \wedge travel == true$

$t_t = getTravelDelay();$

$= (\text{“wait”},\ t_w,\ sched), \qquad \text{if } phase == \text{“travel\_from”}$

$t_w = getWaitDelay().$

Confluence Function:

$\delta_{con}(s, ta(s), x) = \delta_{ext}(\delta_{int}(s), 0, x).$

Output Function:

$\lambda(phase, \sigma, sched)$

$= (\text{station}i,\ msg_i) \quad \text{if } phase == \text{“travel\_to”} \wedge stationID == i$, where msg_i is the message to send to the Station coupled model for station $i = 1, ..., n$.

$= (\text{out},\ msg_i);$

Time Advance Function:

$ta(phase, \sigma, sched) = \sigma.$

Observe that the confluent function performs an internal transition before the external transition. In other words, no preemptions are allowed. The operation of the TECH atomic model can be described as follows. When an input is received on the “set” input port, the model transitions to the “get_schedule” state, where a *computeEventDelays*() method is called to retrieve the technologist's schedule for the day. The schedule is saved in “*sched*”. The model transitions to “waiting” state after the time required for obtaining the schedule (t_g) is elapsed. If an input is received on the “update” input port while the model is in the “waiting” state, a transition to the “update_schedule” state is performed and the method *updateEventsDelay*() is called to update the technologist's current schedule. After the time (t_u) required to complete the schedule update elapses, the model transitions back to the “waiting” state. The model goes into the "travel_to" state when its time to serve the next patient in the schedule ($t_w == 0$), and remains in that state until the travel time (t_t) to the patient location ($stationID$) has elapsed. It then goes into the "serve_patient" state. The state "serve_patient" involves performing a procedure based on the nuclear medicine test prescribed for the patient.

As an example, suppose that a technologist has to perform the nuclear medicine test CPT Code 78465 (see Table III) on a patient. This test involves the following activities: *hydrate patient*(10 minutes), *stress EKG* (30 minutes), *patient wait* (30-60 minutes), *imaging* (30 minutes), and *computer process* (15 minutes). Let us assume that the technologist has to perform step 1 and step 4 on the patient, and that the duration of each activity is deterministic. Then the TECH atomic model would transition from the "waiting" state to "travel_to" state at the scheduled time and remain in that state for a delay of t_t minutes, which is the amount of time it takes to travel to the station. Next, the model would transition to "serve_patient" for step 1 for a delay of $t_{pd} = 10$ minutes. After the delay elapses, the model would transition from "serve_patient" to "travel_from' for a delay of t_f minutes, that is, the amount if time it takes to travel from station 1 to the technologist's office. Finally, the model transitions from "travel_from" to "waiting".

Depending on the technologist's schedule, the TECH model either remains in the "waiting" for some delay t_w, or repeats the process if another patient has to be served. However, when its the scheduled time to travel to the imaging room, the model has to be in the "waiting" state, from where it would transition to "travel_to" for a t_t minutes delay, before moving to "serve_patient". This time to model goes into the "serve_patient" state to perform *imaging* for a delay of $t_{pd} = 30$ minutes. Upon completion of the delay, the model transitions to "travel_from" for a delay of t_f for the imaging room, and finally to the "waiting" state to wait for the next activity in the schedule.

b. EQUIP Atomic Model

The EQUIP atomic model has one input port "in" and one output port "out" as shown in Figure 3. The EQUIP atomic model has two states, "idle" and "busy". The behavior of this atomic model is depicted in Figure 4.

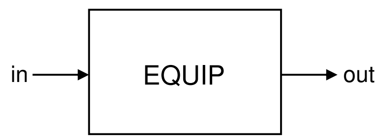

Fig. 3. A basic EQUIP (equipment) atomic model

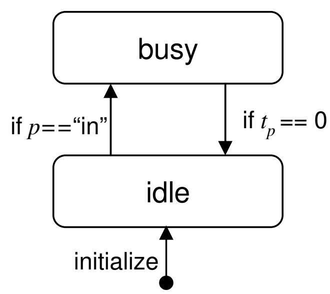

Fig. 4. State transition diagram for EQUIP atomic model

The model is initialized in "idle" state and transitions to the "busy" state if an input is received via the "in" port. A method is called to compute the amount of time (t_p) the model will stay "busy" just before transition. This time t_p depends on the activity performed on the equipment ($task$). Let us call this method *getTaskDuration*() and we will use it in expressing the model in parallel DEVS. When in "busy" state the model does not respond to any inputs, implying that the equipment is busy. Once the amount of time has elapsed, the model returns to the "idle" state. As we mentioned earlier, the EQUIP atomic model is coupled to a Room (ROOM) atomic model. The coupling between these two models will be discussed later in the chapter. The "out" port of the EQUIP atomic model is used to transmit information to the ROOM atomic model. We can now define EQUIP in parallel DEVS as follows:

$$DEVS_{EQUIP} = (X_M, Y_M, S, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, ta) \tag{3.7}$$

where,

$$X_M = \{(p, v)|p \in IPorts, v \in X_p\}$$

is the set of input ports and values, $IPorts = \{\text{"in"}\}$, and $X_{in} = V_1$ is an arbitrary set. The set

$$Y_M = \{(p, v)|p \in OPorts, v \in Y_p\}$$

is the set of output ports and values, $OPorts = \{\text{"out"}\}$, and Y_{out} is an arbitrary set. The set

$$S = \{\text{"idle"},\ \text{"busy"}\} \times \Re_0^+ \times V_1 \ \text{ is the set of sequential states.}$$

External Transition Function:

$$\delta_{ext}((phase, \sigma, task), e, (p, v))$$

$$= (\text{"busy"},\ t_p, task), \quad \text{if } phase == \text{"idle"} \wedge p == \text{"in"},$$

$$t_p = \text{getTaskDuration(task)};$$

$$= (phase, \sigma - e, task), \text{ otherwise.}$$

Internal Transition Function:

$$\delta_{int}((phase, \sigma, task), e, (p, v))$$

$$= (\text{"idle"},\ \infty,\ \text{task}), \quad \text{if } phase == \text{"busy"} \wedge t_p == 0$$

Confluence Function:

$\delta_{con}(s, ta(s), x) = \delta_{ext}(\delta_{int}(s), 0, x).$

Output Function:

$\lambda(phase, \sigma, task)$

$=$ (out, msg) if $phase ==$ "busy", where msg is the message to send to the ROOM atomic model.

Time Advance Function:

$ta(phase, \sigma, task) = \sigma.$

c. SCHED Atomic Model

The SCHED atomic model is in charge of accommodating patients and resources into the system schedule. We allow the modeler to use or implement a scheduling algorithm of their choice. The SCHED atomic model in shown in Figure 5. It has one input port "call_in" and three types of output ports, namely; "patient_out", "radioph_out" and "hres_x_out". The number of outport ports of type "hres_x_out" depends on the number of human resources available in the nuclear medicine facility. The information transmitted by these ports is used to update the human resources' schedules. The 'patient_out" and "radioph_out" output ports are used to send information to the Patient Generator (PGENR) atomic model and the Radiopharmaceutical Generator (RPGENR) atomic model, respectively.

The operation of the SCHED atomic model is depicted in Figure 6. The model has three basic states: "idle", "update_schedule", and "scheduling". The model is initialized in the "idle" state. A transition to the "scheduling" state occurs when the model is in the "idle" state and a message ($call_i$) is received at the "call_in" input port. A method, *getPatientSchedule*(); takes the information provided by the patient and

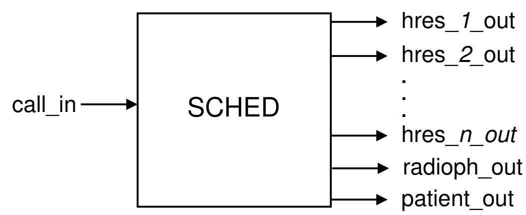

Fig. 5. A basic SCHED (scheduler) atomic model

performs the scheduling using the algorithm chosen by the user. If the scheduling is successful, the model transitions to the "update_schedule" state, where the schedules for the resources selected in serving the patient are updated. After completing the schedule updates, the model transitions to the "idle" state. Otherwise, if scheduling is unsuccessful, the model transitions from "scheduling" state back to the "idle" state.

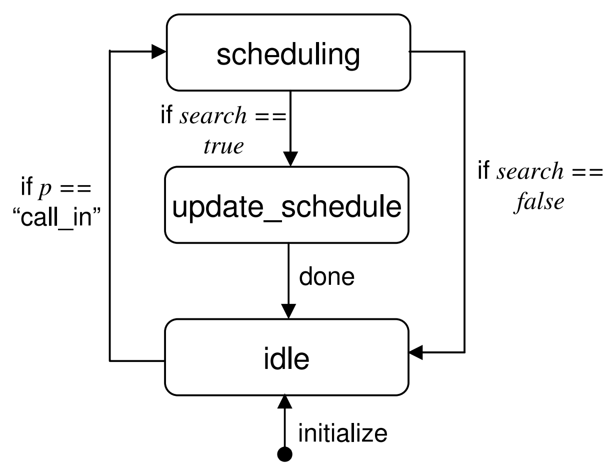

Fig. 6. State transition diagram for SCHED atomic model

Mathematically, the SCHED atomic model can be expressed in parallel DEVS as follows:

$$DEVS_{SCHED} = (X_M, Y_M, S, \delta_{ext}, \delta_{int}, \delta_{con}, \lambda, ta) \tag{3.8}$$

where,

$$X_M = \{(p, v)|p \in IPorts, v \in X_p\}$$

is the set of input ports and values, $IPorts = \{``call\_in"\}$, and $X_{call\_in} = V_1$ is an arbitrary set. The set

$$Y_M = \{(p, v)|p \in OPorts, v \in Y_p\}$$

is the set of output ports and values, and $OPorts = \{$"patient_out", "radioph_out", "hres_1_out", "hres_2_out", $\cdots$, "hres_n_out"$\}$, where $Y_{patient\_out}$, $Y_{radioph\_out}$, $Y_{hres\_1\_out}$, $Y_{hres\_2\_out}$, $\cdots$, $Y_{hres\_\mathrm{n}\_out}$ are arbitrary sets. The

$S = \{$"idle", "update_schedule", "busy"$\} \times \Re_{+,0} \times V_1$ is the set of sequential states.

External Transition Function:

$\delta_{ext}((phase, \sigma, call_i), e, (p, v))$

= ("scheduling", $t_s, call_i$), if $phase ==$ "idle" $\wedge\ p ==$ "call_in",

$appointment =$ getPatientSchedule($call_i$);

= $(phase, \sigma - e, call_i)$, otherwise.

Internal Transition Function:

$\delta_{int}((phase, \sigma, call_i), e, (p, v))$

= ("update_schedule", t_u, $call_i$), if $phase ==$ "scheduling" $\wedge\ search = true$;

= ("idle", ∞, $call\_i$), if $phase ==$ "update_schedule";

= ("idle", ∞, $call\_i$), if $phase ==$ "scheduling" $\wedge\ search = false$;

Confluence Function:

$\delta_{con}(s, ta(s), x) = \delta_{ext}(\delta_{int}(s), 0, x)$.

Output Function:

$\lambda(phase, \sigma, call_i)$

$=$ (patient_out, $patient_i$), if $phase ==$ "update_schedule", where $patient_i$ is the message to send to PGENR;

$=$ (radioph_out, $radioph_i$), if $phase ==$ "update_schedule", where $radiop_i$ is the message to send to the RPGENR;

$=$ (hres_i_out, msg_i), if $phase ==$ "update_schedule" $\wedge$ $hresID == i$, where msg_i is the message to send to the atomic model for human resource $i = 1, ..., n$.

Time Advance Function:

$ta(phase, \sigma, call_i) = \sigma$.

We omit the mathematical definitions of the rest of the atomic models (CGENR, PGENR, RPGENR and TRANSD) and instead devote the rest of this subsection to explain some of the coupled models used to create the simulation model. All the coupled models are coupled according to the three types of connections (*EIC, EOC*, and *IC*) defined in Equations 3.3, 3.4 and 3.5, respectively. We start with the STATION coupled model. As shown in Figure 7, the model is created by coupling EQUIP and ROOM. STATION has three input ports, namely; "patient_in", "radioph_in", and "hres_in". *EICs* exist between the input ports and the ROOM atomic model. Two *ICs* connect EQUIP with ROOM. Information is passed to EQUIP via ROOM when an input has been received on STATION's input ports. The STATION coupled model has two types of output ports, "patient_out" and "hres_n_out". The number of output ports of type "hres_n_out" depends on the number, n, of human resources in the nuclear medicine facility. The information transmitted by the outport ports is used to notify when a patient or human resource has been released from the room. This only happens when the ROOM atomic model receives information from the EQUIP atomic model notifying the service performed on the patient has been completed.

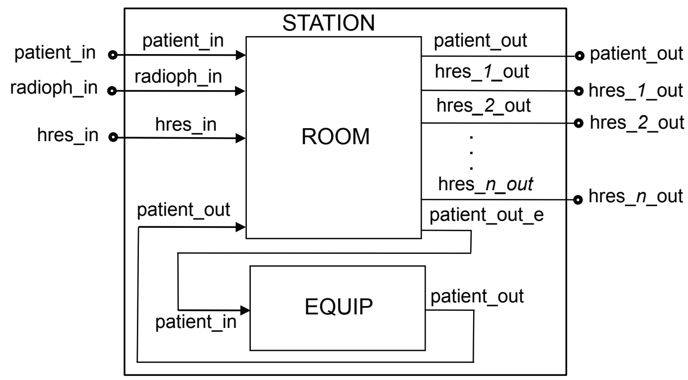

Fig. 7. The STATION coupled model

The next model is the NMD coupled model shown in Figure 8. This model is a representation of the nuclear medicine department (NMD) and is created by coupling the human resource atomic models (TECH, NURSE, RCPST, PHYSN, MANGR) to STATION. In the figure we only show human resource models TECH, NURSE and MANGR due to limitation in figure size.

The last coupled model is the Experimental Frame (EF) shown in Figure 9. The EF allows the modeler to specify the type of experiments that should be performed on NMD to enable answering questions of interest. Therefore, the EF is coupled to NMD (as depicted by the arrows) to create the overall simulation model for a nuclear medicine facility. The figure shows the atomic models that are part of EF and the way they are connected. CGENR is an atomic model of a telephone call center and is in charge of generating telephone call messages for patient appointment requests. This model allows the user to specify the telephone call arrival rate and the

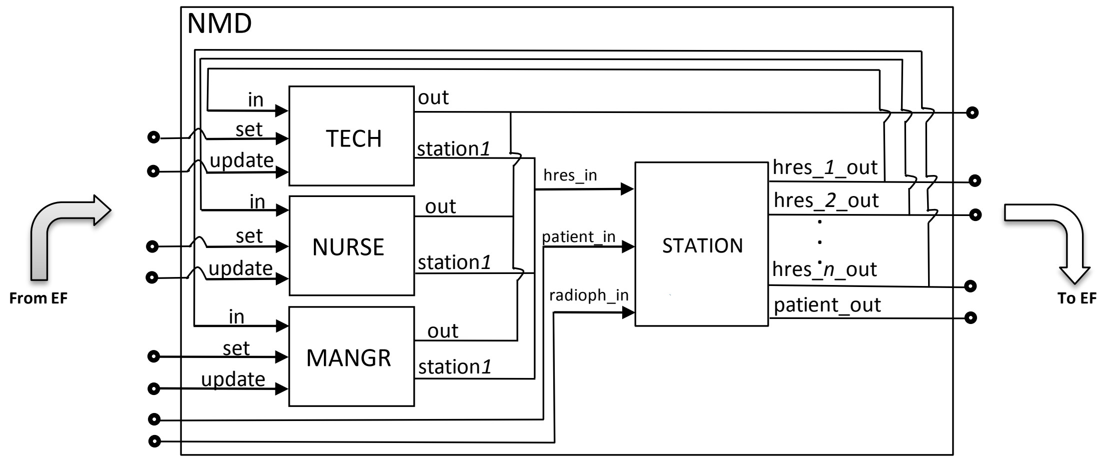

Fig. 8. The NMD coupled model

associated probability distribution. The generated appointment requests are received and processed by the SCHED atomic model. SCHED allows the user to select an algorithm for scheduling patients (and the needed resources) into the system. The schedule information is passed from SCHED to the RPGENR and PGENR atomic models. RPGENR models the ordering and arrival of radiopharmaceuticals at the facility at the scheduled time. PGENR models the actual arrival of patients to the nuclear medicine facility at their appointment times. To compute the performance measures of interest (described in Section 1), we created the transducer (TRANSD) atomic model. The TRANSD atomic model collects information from NMD and computes performance measures of interest.

3. NMD System Entity Structure

A system entity structure (SES) is used to plan, generate, and evaluate design of simulation-based systems. This is a scheme that organizes a set of possible structures of a system. A library of models is generated when all the components abstracted from the real system are implemented. The SES is used to classify these components

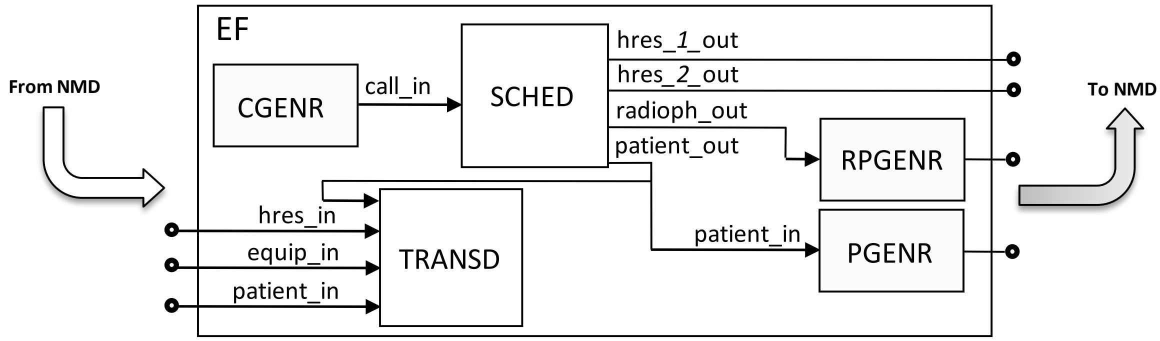

Fig. 9. The EF coupled model

by their characteristics and to organize them in a hierarchical composition. This representation allows the modeler to visualize the system as a whole. The goal of the SES is to synthesize a simulation model by traversing a model hierarchical structure. A SES represents not a single model structure, but a family of structures from which a candidate entity structure can be selected.

The SES for the NMD simulation model is shown in Figure 10. At the top level, the scheme shows the two major coupled models that define the system structure. The *Experimental Frame* (EF) branch is decomposed into three branches that are assigned to the *Transducer*(TRANSD), *Generator*(GENR) and *Scheduler*(SCHED) atomic models. The double line under the GENR branch means specialization. The Generator model is categorized into specialized entities such as the *Patient Generator*(PGENR), *Call Generator*(CGENR) and the *Radiopharmaceutical Generator*(RPGENR). The NMD branch is decomposed into two branches: Human Resource (HR) and Station (STATION). The HR branch is decomposed into four branches, each define a different type of human resource existing in nuclear medicine. The *Technologist* (TECH) can be specialized into Nuclear Radiology Technologist and EKG Technologist. The STATION branch is decomposed into two branches. A selection constraint, depicted as dotted arrow from Gamma Camera (GAMMC) and Treadmill

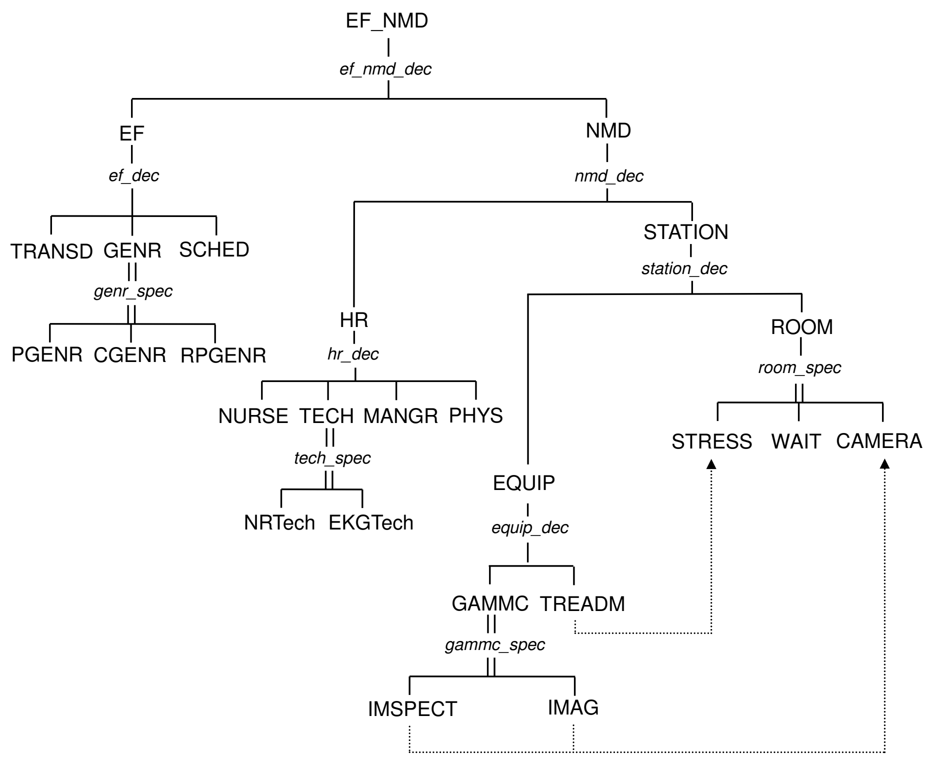

Fig. 10. System entity structure

(TREADM) specializes entities to ROOM. Specialization entities mean that those entities cannot be selected independently. Finally, GAMMC is specialized into image with SPEC capability (IMSPEC) and image (IMAG).

4. Model Implementation, Verification and Testing

We implemented the NMD simulation model in DEVSJAVA [93], a Java-based modeling and simulation software implementation of DEVS formalisms such as parallel DEVS. We tested and verified each atomic and coupled model using DEVSJAVA Simulation Viewer Version (SimView) 1.0.4. SimView allows the modeler to visually inspect the behavior of each model created in DEVSJAVA. Atomic models were verified first because they serve as building blocks for coupled models. Every component is represented with their input and output ports. Couplings among the various models are also represented for coupled models. Figure 11 shows a print screen of a SimView window for the EF coupled model.

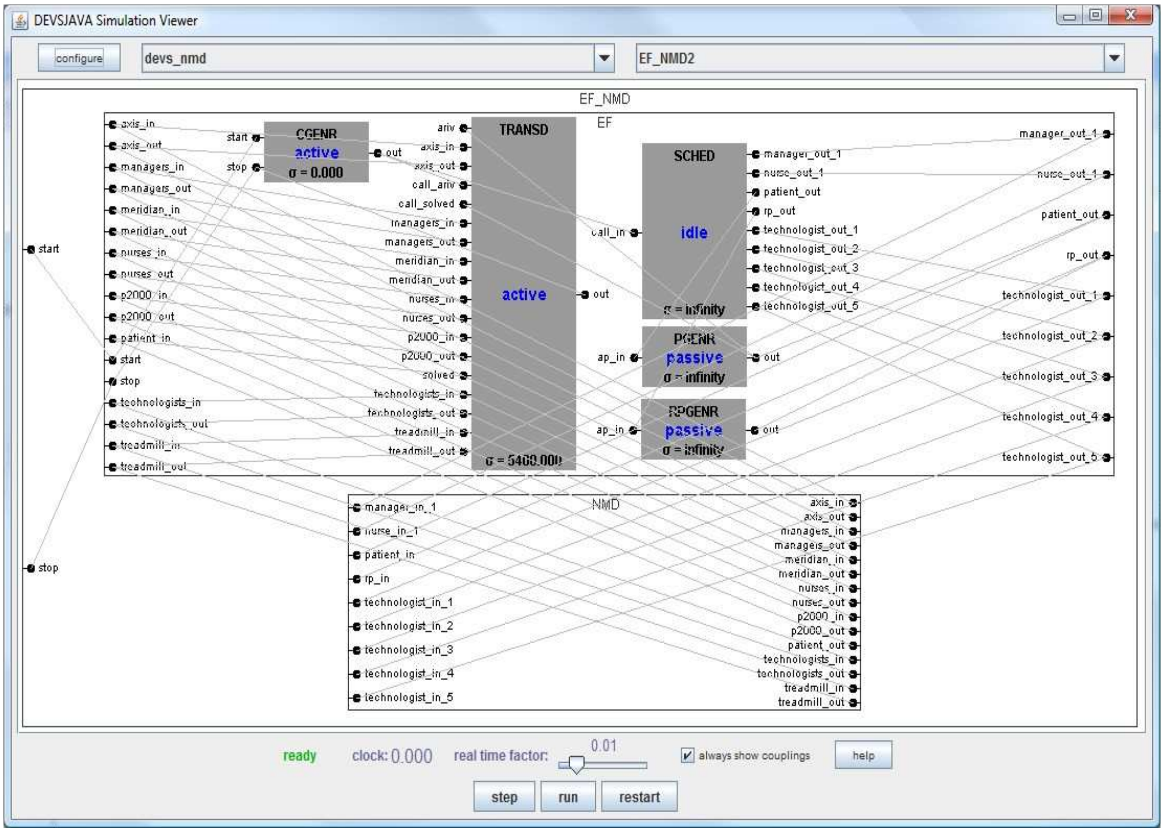

Fig. 11. SimView window

SimView has the advantage of having several convenient functionalities that include allowing the user to start and stop the simulation at any time during the simulation run, fast-forwarding or slowing down the simulation, and being able to input user defined parameters created for model verification and testing by simply clicking on a model's input port and selecting the desired option from the pop-up menu. To run a simulation, the user selects the appropriate model from the top menu on the SimView window and click the run button. During the simulation run the simulation clock is displayed on the window. Parameters and statistics of are displayed as well by positioning the mouse cursor on top of the model block.

D. Application

To validate the NMD simulation model, we applied it to a real setting and used historical data for a particular year. We implemented a patient and resource scheduling algorithm that was used in practice that year to schedule patients and resources in the simulation. Our simulation model validation is based on a real nuclear medicine setting, historical data and expert opinion. After validating the simulation model, we implemented and studied alternative scheduling algorithms to gain insights into patient service management and system performance in nuclear medicine. Next we describe the real nuclear medicine setting for our computational study. We then describe the experimental setup, report the simulation results and discuss our findings.

1. Real Nuclear Medicine Setting

We applied the NMD simulation model to the nuclear medicine department of the Scott and White Health System in Temple, Texas, U.S.A. This nuclear medicine clinic is one of the largest fully-accredited nuclear laboratories for general nuclear imaging and non-imaging, nuclear cardiology and positron emission tomography (PET) scan in the country. This facility operates five days a week from 8:00 am to 5:00 pm, and is not open on weekends. There are sixteen fulltime physician support staff budgeted in this department. Every member of the group performs specific tasks that depend on the staff specialty. The department has eight technologists and two EKG technologists. This staff group has several responsibilities that include the preparation and administration of the radiopharmaceuticals, drawing doses and imaging acquisition. Electrocardiogram (EKG) technicians perform stress exams for cardiac tests. A nurse assists with the radiopharmaceutical administration and drawing doses. The division manager can also assist with those activities in the absence of one of the regular staff. The department also has two fulltime nuclear medicine physicians, two radiology residents, and a staff cardiologist.

There are seven gamma cameras (one Philips PRISM 3000, two Philips PRISM 2000, three Philips AXIS and one Philips Meridian). Five of these cameras are planar, and are capable of doing 2D whole-body imaging and 3D Single Photon Emission Computed Tomography (SPECT). The other two cameras are also planar, one is SPECT capable of a small field of view only and the other is for imaging only. The stress cardiac area comprises a nurse station and three stress rooms. Two of the stress rooms have treadmills. The third room is for chemical stress testing for patients who cannot walk on a treadmill. All three are equipped with EKG capability. In the PET facility, there is one PET imaging camera, three patient preparation rooms for patient hydration and waiting time, and a radiopharmaceutical receiving room. Around 60 different procedures are performed in this department. Table II in Section 1 shows the procedures/tests that were performed more frequently at the clinic during the year of our study and we only use these procedures in our simulation.

Patient calls are answered by three receptionists. Patients may provide a preferred day of the week for their appointment. A search for an appointment is first done by trying to satisfy that preference. However, if an appointment is found where the patient waits more than a month to be served, the preference provided is disregarded and an alternative earlier appointment is provided. Resource scheduling is performed using a load balancing routine where each resource is scheduled in a round-robin manner. Nevertheless, some human resources (technologists) from the staff are fixed to specific stations. Human resources assigned to stations take care of the stations where equipment utilized the most is located. The clinic manager schedules patients and resources mainly based on experience.

2. Experimental Setup

We used the following configuration for the NMD simulation model based on the historical data: 7 gamma cameras, 2 stress rooms, 1 PET positron camera, 10 technologists, 1 nurse and 1 manager. The specific station names are TRT(1), TRT(2), TRT(3), Treadmill(1),Treadmill(2), Axis(1), Axis(2), Axis(3), P2000(1), P2000(2), P2000(3) and Meridian(1). The specific human resource names are Technologist(1), Technologist(2), Technologist(3), Technologist(4), Technologist(5), Technologist(6), Technologist(7), Technologist(8), Technologist(9) and Technologist(10). We assumed Poisson arrivals for patient appointment calls based on historical data. There were about 90 calls per day on average during that year. We computed average monthly arrival rates to use in the simulation and set the appointment call interarrival process to follow an exponential distribution with the following monthly mean interarrival times in minutes: January, 6.00; February, 6.25; March, 6.58; April, 6.67; May, 6.75; June, 6.88; July, 6.96; August, 7.04; September, 7.10; October, 7.29; November, 7.34; and December, 7.44.

To generate a nuclear medicine procedure/test for each patient, we used an empirical distribution for the procedures that were performed during that year. According to historical data about 70% of the patient calls were for outpatients, who made appointments in advance. The patients who required to be served immediately comprised the other 30%, half of which were inpatients who required to be served on the same day. The other half were emergency patients who needed to be served as soon as possible. Also, on average 1% of the patients arrived late for their appointments, 1% canceled their appointments, and 1% were no shows. Therefore, we did not include late arrivals, cancelations, and no shows in our simulation.

We implemented a scheduling algorithm that was used during the year of our study to schedule patients and resources. We refer to this algorithm as *fixed resource* (FR) algorithm and it can be summarized as follows:

- FR: Under this scheduling policy Technologist(1) and Technologist(2) are fixed to station Axis(1) and station Axis(2), respectively. The rest of the staff are available to be scheduled to the other stations as needed. The manager is available to perform procedures at any station.

To study the impact of patient and resource scheduling on system performance, we implemented the following four alternative variations of the FR scheduling algorithm:

- NFR: This is the *no fixed resource* scheduling policy where none of the human resources are fixed to any station and the manager is available to perform procedures at any station.

- FR_ALL: Under this scheduling policy, the human resources are fixed to specific stations as follows: Technologist(1) is fixed to station Axis(1); Technologist(2) is fixed to station Axis(2); Technologist(3) is fixed to station Axis(3); Technologist(4) is fixed to station P2000(1); Technologist(5) is fixed to station P2000(2); Technologist(6) is fixed to station P2000(3); and Technologist(7) is fixed to station Meridian(1). Technologist(8) is not fixed to any station while Technologist(9) and Technologist(10) (EKG technologists) are both fixed to stations Treadmill(1) and Treadmill(2). The manager is available to perform procedures at any station.

- NFR_NO_MNGR: This is the same as NFR but the manager is *no* longer available to perform any procedures unless the patient has to wait for the appointment for more than a month. This algorithm is aimed at studying the impact of not using the clinic manager to perform technologist duties unless necessary.

- FR_NO_MNGR: This scheduling policy is the same as FR but the manager is *no* longer available to perform any procedures unless the patient has to wait more than a month for the appointment. This algorithm is aimed at studying the impact of not using the clinic manager to perform technologist duties unless necessary.

We used the performance measures identified in the literature (Section 1) to quantify service levels based on both patient and management perspectives. Specifically, we used two performance measures for patient satisfaction described in Table V: patient waiting time *Type 1* and patient preference satisfaction ratio. For manager's perspective performance measures, we used the three measures listed in Table VI: equipment utilization, human resource utilization and patient throughput. We made 100 replications for each simulation run and used a scheduling time horizon of 12 months with a warm-up period of 3 months. To maintain independence among the replications, we used different seeds for the random number generators in the simulation. We computed the mean, standard deviation and confidence intervals for all the performance measures. All the simulations were conducted on a DELL Optiplex GX620 with a Pentium D processor running at 3.2GHz with 2.0GB RAM.

3. Simulation Results

We first report model validation results based on patient throughput using FR. Figure 12 shows a plot of the average number of patients served in a given month using FR and the actual historical values. As can be seen in the figure, the simulation results show a decreasing trend in the average number of patients served per month from January to December matching the historical trend. The FR values are within 10% of the actual values. They are all below the actual values because we only included the top ten procedures that were performed during that year (Table II) in the simulation. The average annual patient throughput for FR is 61.57 patients per day, compared to the actual annual patient throughput of 68.12 patients per day. For

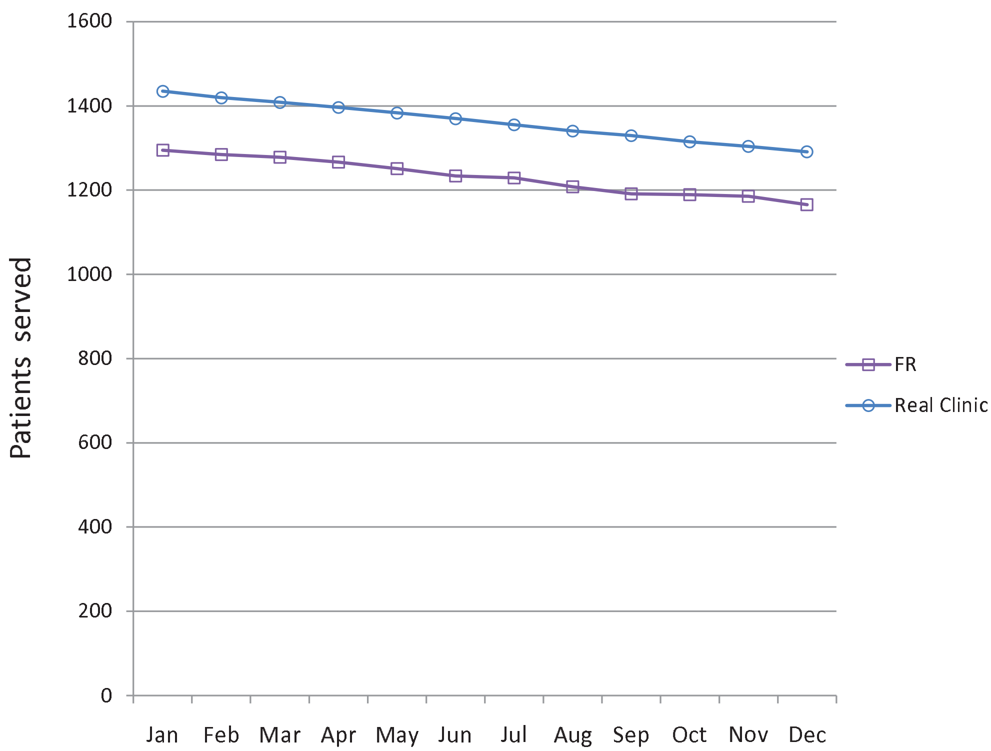

Fig. 12. Number of patients served per month: FR versus actual

patient perspective measures under FR, *Type 1* wait time (time a patient waits from call to appointment) is 5.48 days, while patient preference is 87.68%. The actual values for these two measures were not available for comparison. However, based on expert opinion, these results are consistent with the expected values for the year of our study.

Next we report simulation results for the rest of the alternative scheduling al-

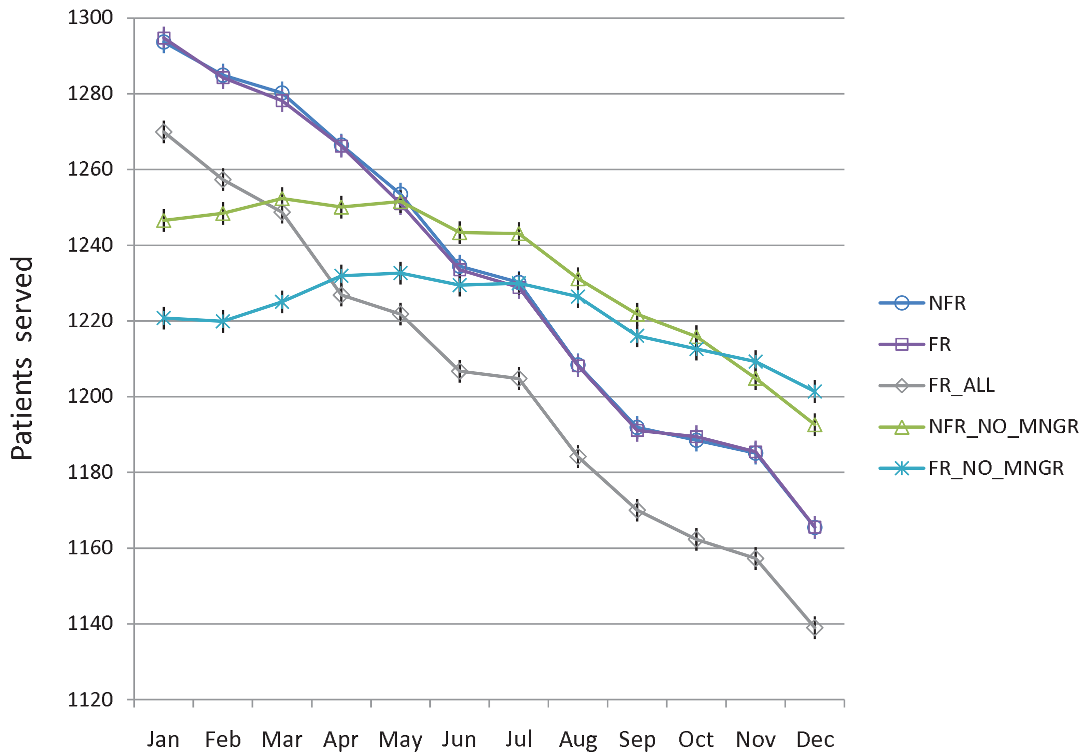

Fig. 13. Number of patients served per month under different algorithms

gorithms and compare them to FR. Patient throughput results are summarized in Figure 13. First, NFR provides results that are not significantly different from those obtained by FR for all the months. Recall that in NFR no resource is fixed to any station. Second, the results for FR_ALL in which all the technologists are fixed to specific stations (except Technologist 8), show a significant decrease in the average number of patients served in each month compared to FR and NFR. Third, the results for NFR_NO_MNGR (NFR without the manager) show a totally different trend. The number of patients served for January to April are significantly lower than those under FR and NFR, while those for June to December are significantly higher. The average number of patients served in May is not significantly different from that under FR and NFR. Finally, the results for FR_NO_MNGR (FR without the manager) shows a similar trend as NFR_NO_MNGR, but the lower values now range from January to June, while the higher values range from August to December. The average number of patients served in July is not significantly different from that under FR and NFR.

Table VII. Number of patients served, system throughput and simulation time

Algorithm	Statistic	Patients Served	Patients/hr	Patients/day	CPU Time (secs)
NFR	Mean	14782.84	6.84	61.6	614.44
	Std. Dev.	78.99	0.04	0.33	6.47
	CI Lower	14769.72	6.84	61.54	613.36
	CI Upper	14795.96	6.85	61.65	615.51
FR	Mean	14776.2	6.84	61.57	611.39
	Std. Dev.	78.65	0.04	0.33	9.09
	CI Lower	14763.14	6.83	61.51	609.88
	CI Upper	14789.26	6.85	61.62	612.9
FR_ALL	Mean	14448.88	6.69	60.2	554.14
	Std. Dev.	74.84	0.03	0.31	5.43
	CI Lower	14436.45	6.68	60.15	553.24
	CI Upper	14461.31	6.7	60.26	555.04
NFR_NO_MNGR	Mean	14801.24	6.85	61.67	651.12
	Std. Dev.	72.81	0.03	0.3	15.12
	CI Lower	14789.15	6.85	61.62	648.61
	CI Upper	14813.33	6.86	61.72	653.63
FR_NO_MNGR	Mean	14655.14	6.78	61.06	651.12
	Std. Dev.	67.19	0.03	0.28	14.86
	CI Lower	14643.98	6.78	61.02	648.65
	CI Upper	14666.3	6.79	61.11	653.59

The statistical results for the annual number of patients served, patient throughput and simulation CPU time for all the scheduling algorithms are given in Table VII. We report the mean, standard deviation (Std. Dev) and the 90% confidence interval (CI). The results for human resource and equipment (station) utilization are given in Figure 14 and Figure 15, respectively. Observe that utilization values for both human resource and equipment highly depend on the scheduling algorithm used. Finally, the results for patient perspective performance measures are given in Figure 16. The figure shows variation of patient waiting time *Type 1* and patient preference

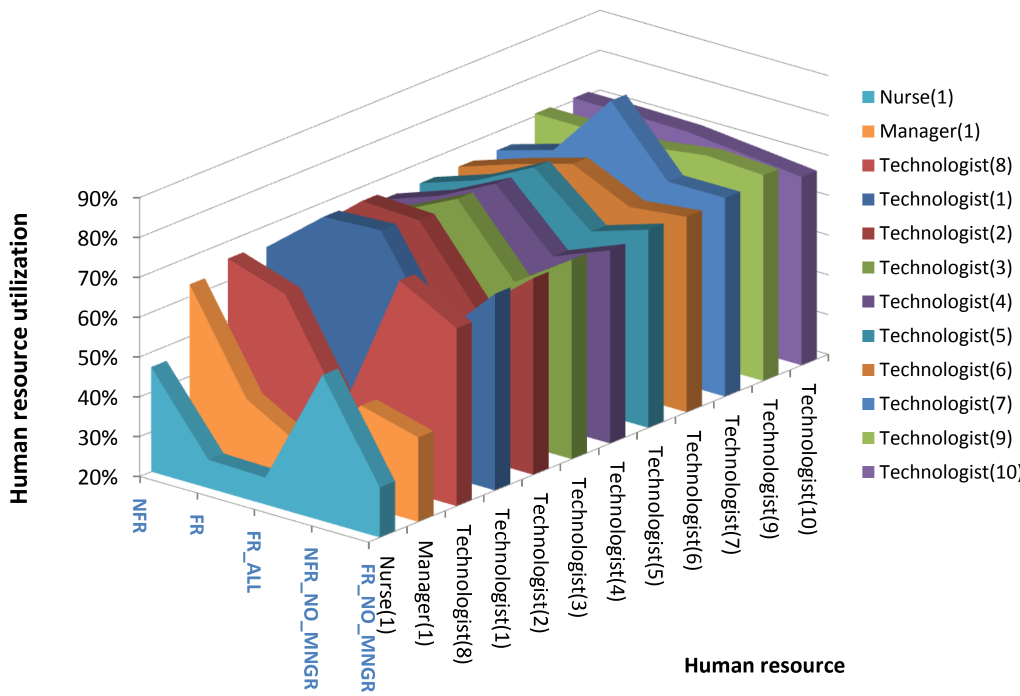

Fig. 14. Human resource utilization

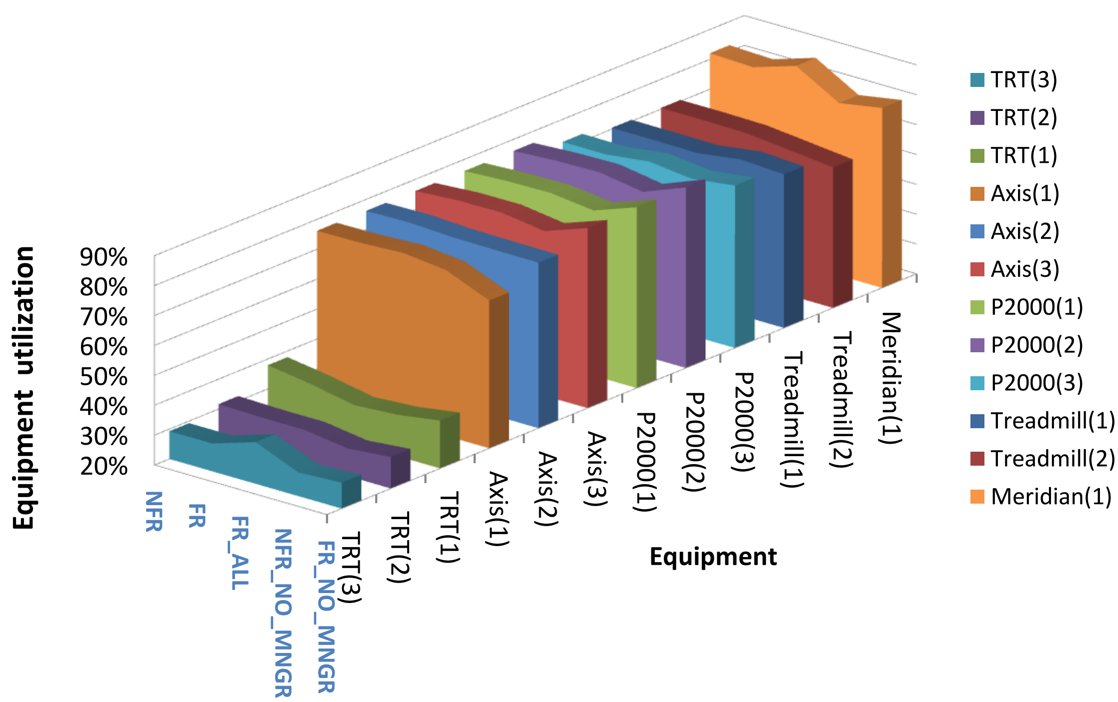

Fig. 15. Equipment (station) utilization

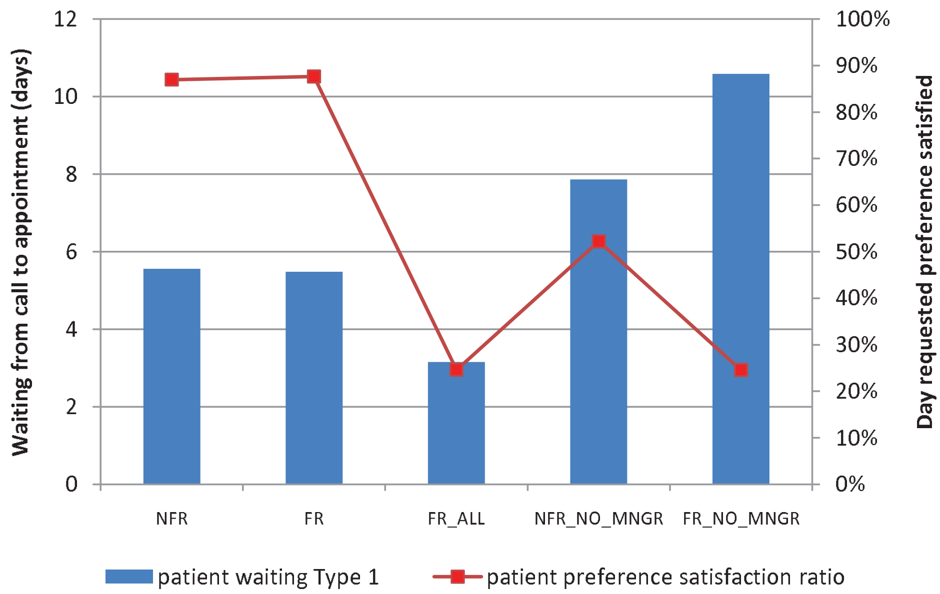

Fig. 16. Patient waiting *Type 1* and preference satisfaction ratio

satisfaction ratio (number of times patient request for an appointment is satisfied) with the scheduling algorithm used.

4. Discussion

Scheduling patients and limited resources in nuclear medicine to provide high levels of patient service while maximizing system performance is challenging. The type of algorithm used to schedule patients and resources has a high impact on system performance measures for both patients and managers. Among the five scheduling algorithms studied, only FR and NFR gives similar results in terms of patient throughput. However, even for these two algorithms, human resource and equipment utilization results are different. For example, FR has higher utilization for Technologist(1) and Technologist(2), who are fixed to stations Axis(1) and Axis(2), respectively. As a consequence, Axis(1), Axis(2), Axis(3), P2000(1) and P2000(2) have higher utilization.

Also, under FR utilization of TRT1, TRT2 and Meridian(1) is reduced. In fact, utilization of Nurse(1) and Manager(1) is reduced by almost half. Since Nurse(1) and Manager(1) perform some procedures on Axis(1) and Axis(2), their utilization decreases because they cannot use those stations anymore as they are dedicated to the two technologists. In terms of patient perspective measures, the results show that on average patients wait slightly less under FR than NFR. Also, there is slightly higher patient preference satisfaction. However, NFR not only provides similar patient throughput as FR, it also allows for equity in terms of evenly distributing the workload among human resources and equipment.

FR_ALL reveals that fixing all but one technologist to specific stations is not good for the system in terms of the average number of patients served per month. We observe that even though this patient throughput is significantly less than that under FR and NFR, utilization of resources under this policy is comparable to the other policies. We noticed that procedures with longer durations were scheduled more often under FR_ALL. However, FR_ALL has the lowest time for patient waiting time *Type 1*. This was caused by the fact that this policy was able to satisfy patient preference only 25% of the time. So most of the time patients were scheduled without taking into consideration their preferred appointment day, otherwise patients would end up waiting more than a month for an appointment.

When manpower is reduced as in NFR_NO_MNGR and FR_NO_MNGR where Manager(1) is no longer available to perform procedures, we observe that the system cannot satisfy demand from January to May when the demand is higher (Figure 13). This forces the system to schedule more patients further into the future, that is, longer patient waiting time *Type 1*. Consequently, there is a higher number of patients served from June (NFR_NO_MNGR)/August (FR_NO_MNGR) to December even though actual patient demand is lower during those months. In comparing

NFR_NO_MNGR and FR_NO_MNGR, we see that NFR_NOMNGR performs better than FR_NO_MNGR in terms of the number of patients served from January until October. FR_NO_MNGR serves more patients than NFR_NO_MNGR from November to December (see Figure 3). Overall, we see that with limited human resources, not fixing human resources to stations coupled with evenly distributing the workload among resources results in better system performance.

NFR and FR scheduling policies reported similar results for the patient perspective performance measures; however FR performs slightly better than NFR. An average waiting time *Type 1* of around 5.5 days and a patient preference satisfaction ratio that is close to 90% was obtained for both policies. The FR_ALL algorithm reported the lowest values for the average number of patients served during the year, average waiting time *Type 1*, and patient preference satisfaction ratio. Flexibility in terms of managing resources is limited when most of the technologists are fixed to stations. This fact forces the FR_ALL policy to disregard patient preference requests most of the time. If patient preferences are taken into account patients would end up waiting more than a month for their appointments. However, when patient preferences are not taken into account, patients tend to be scheduled earlier on the first appointment time slot available. NFR_NO_MNGR reported an average waiting time of 7.86 days and a patient preference satisfaction ratio of 52.23%. Under this policy the manager is not available most of the time and the staff capacity is less. This drop in capacity causes an increase in the average waiting time and a decrease in the patient preference satisfaction ratio when the figures are compared to those under NFR and FR. FR_NO_MNGR reported the highest waiting time and one of the lowest patient preference satisfaction ratios. Under FR_NO_MNGR flexibility for managing the resources available is limited because two technologists are fixed to stations and the manager is not available to perform procedures most of the time.

E. Summary

Managing patient service in nuclear medicine under limited resources is a very challenging problem with very little research attention. In this chapter, we use the discrete event system specification (DEVS) formalism to derive a generic simulation model for nuclear medicine patient service management that takes both patient and management perspectives. DEVS is a formal M&S framework based on dynamical systems theory and provides well defined concepts for coupling components, hierarchical and modular model construction, and an object-oriented substrate supporting repository reuse. We implement and validate the simulation model based on a real nuclear medicine setting and report computational results based on a scheduling algorithm and several patient and management performance measures. The results provide useful insights into patient service management in nuclear medicine. For example, fixing technologists to specific stations may result in reduced patient throughput unless the right number to fix to stations is carefully determined (e.g. through simulation). Also, reducing manpower even by a single technologist can result in scheduling patients further into the future during months with high patient demand. Thus it is up to each nuclear medicine clinic to select the 'best' scheduling policy based on which performance measures, for both patient and management perspectives, would provide a high level of service.

While this work focuses on nuclear medicine, we believe that results will find generality in patient service management in other health care settings. It also provides several future research directions. For example, the current simulation model can be extended to a stochastic model using stochastic DEVS (SDEVS), which allows for modeling atomic model state transitions as a stochastic process. One can also envision stochastic optimization algorithms for scheduling patients and resources, which can be based on mathematical programming or stochastic online optimization. Finally, the current work can be extended to a simulation-optimization setting, whereby the simulation model provides feedback to stochastic optimization scheduling algorithms with the objective of making optimal decisions based patient and the nuclear medicine management perspectives.

CHAPTER IV

PATIENT AND RESOURCE SCHEDULING OF MULTI-STEP MEDICAL PROCEDURES IN NUCLEAR MEDICINE

A. Introduction

Nuclear medicine is a sub-specialty of radiology that provides highly specialized services by means of new technology for diagnosis and treatment of patients. There has been a rise in demand for such specialized services in the U.S. and this has been attributed as a contributor to the increased health care costs, which surpassed those in other nations that provide similar services. Physicians are becoming more prone to asking patients to undergo specialized procedures in order to obtain more accurate diagnoses. However, scheduling patients and resources in specialized clinics such as nuclear medicine remains a challenging problem. This may be attributed to the increased demand in services and the nature of nuclear medicine procedures. In this chapter, we derive algorithms to assist nuclear medicine managers towards scheduling nuclear medicine patients and resources more efficiently. We consider both the patient's and the manager's perspectives.

Nuclear medicine procedures (tests or studies) are typically multi-step, involve multiple resources, and require the administration of a radiopharmaceutical (radioactive isotope, e.g., iodine-131) to the patient. This allows for images of specific body organs to be taken (scan) using gamma cameras that sense the radiation emitted by the radiopharmaceutical. Since radiopharmaceuticals have a short half-life (minutes), their decay imposes strict time constraints on scheduling patients and resources in order to get good quality scans. Thus scheduling patients in nuclear medicine requires very strict procedure protocols, which if not followed can result in poor scans. In this case, time, money and resources are wasted, and the patient has to be rescheduled for another day after having been exposed to radiation. Some nuclear medicine tests require only a single scan while others involve multiple scans in a day or multiple days. Each scan takes several minutes to hours to complete.

Nuclear medicine procedures require the utilization of the several resources such as a technologist, gamma camera, radiopharmaceutical, and sometimes, a nurse or EKG (electrocardiography) technician. The gamma cameras may cost up to a million dollars and thus have to be used and managed effectively. Since at many nuclear medicine clinics radiopharmaceuticals are prepared at remote radio-pharmacies from the clinic, scheduling of their delivery, patient injection and image acquisition requires lead time and must be carefully managed. Radiopharmaceuticals may cost up to several hundreds of dollars. The resources needed to perform each procedure step must be available at the scheduled times. A patient has to be rescheduled if the procedure is not completed successfully. Therefore, scheduling patients, resources, and radiopharmaceutical preparation and delivery is a challenging problem for nuclear medicine departments. Consequently, providing a high quality of service to the patient through the use of mathematical techniques is of great interest to nuclear medicine managers. However, the characteristics found in the management of patients and resources in nuclear medicine makes it a unique problem with limited research reported in the literature. Furthermore, very few commercial packages are available for patient service management and the few available do not have algorithms for scheduling patients and resources efficiently.

Several practical issues have to be considered to achieve a well designed system for patient service management in nuclear medicine. For example, scheduling decisions must satisfy the goals of both patients and managers. In this work, we consider both perspectives. Both points of view are important when designing scheduling policies to improve the service offered to patients and the way resources are utilized. The contributions of this chapter include mathematical algorithms for scheduling patients and resources in nuclear medicine that consider both patient and manager perspectives. These scheduling algorithms are implemented and tested using the discrete event simulation model proposed by [94]. We obtain computational results that provide useful insights into managing patient service and resources in nuclear medicine. While this work focuses on nuclear medicine, we believe that results can be applied to many other similar health care settings that may not be as complex as nuclear medicine. For example, our results can be applied to diagnostic imaging areas such as magnetic resonance imaging (MRI) and computed axial tomography (CT Scan).

The rest of the chapter is organized as follows: In Section 2 we review closely related work and provide preliminaries on nuclear medicine resources and procedures in Section 3. We derive algorithms for scheduling nuclear medicine patients and resources in Section 4. We report on a computational study to quantify important trade-offs among different patient and resource scheduling strategies in Section 5. We also discuss the results and highlight the insights into the complexity of nuclear medicine patient service management. We end the chapter with some concluding remarks and directions for further research in Section 6.

B. Nuclear Medicine Department Setting

A typical nuclear medicine department contains several entities that interact according to the requirements of the service requested. These include humans (staff), procedures/tests, stations, and patients. We describe these entities in the context of their interaction with the appointment scheduling system, and then summarize the performance measures used to evaluate the scheduling algorithms presented in this chapter.

Table VIII. Staff duties in nuclear medicine

Human resource	Responsibilities
Technologist	Hydrate patient; Radiopharmaceutical preparation; Imaging
Nurses	Hydrate patient; Radiopharmaceutical administration; Draw doses
Physicians	Hydrate patient; Radiopharmaceutical administration; Draw doses
Managers	Hydrate patient; Radiopharmaceutical preparation; Radiopharmaceutical administration

1. Nuclear Medicine Entities

a. Human Resources

The staff are comprised of four types of human resources: *technologists, nurses, physicians, and managers.* Each human resource possesses his/her own expertise and experience, which determine the set of activities they can perform and the amount of time required to complete each activity. Human resources that have more experience are expected to complete their tasks relatively quickly. Some of the activities that can be performed by the members of the staff are listed in Table VIII.

Table IX. Diagnostic procedures in nuclear medicine

CPT Code	Name
78465	Cardiovascular Event (CVE) Myocardial Imaging (SP-M)
78815	PET CT skull to thigh
78306	MSB-bone imaging (whole body)
78315	MSC-bone imaging (three phase)
78223	GIC-Hepatobiliary imaging
78472	CVJ-cardiac blood pool
78585	REB-Pulm perfusion / ventilation
78006	ENC-Thyroid imaging
78195	HEE-Lymphatic imaging
78464	CVD-Myocardial imaging (SP-R ORS)

b. Procedures (Tests)

Procedures/tests are requested by the patient's primary physician or attending physician. Unlike general outpatient clinics, in nuclear medicine patients attend to their appointments most of the time and *no shows* are not an issue. Nuclear medicine procedures provide physicians with information about the function of organs of the human body and are used for patient diagnosis and treatment. We present a list of nuclear medicine procedures with their current procedural terminology (CPT) codes in Table IX.

At least one radiopharmaceutical is administered to the patient at the beginning of a procedure. Radiopharmaceuticals are requested in advance and they need to be at the clinic by the time of the patient appointment. Nuclear medicine procedures are multi-step and each step has a limited time duration. This time duration varies depending on the human resource performing the task; however, the procedure step has to be completed within the time window established by a protocol. Table X provides a description of a nuclear medicine procedure. The *MSC-bone imaging* procedure has

Table X. Procedure 78315: MSC-bone imaging (three phase)

Step	Activity	Time (min.)	Station	Human Resource
1	Hydrate patient	20	Meridian; TRT 1, 2 & 3; Axis 1, 2, & 3; P2000A & B	Technologists; Nurse; Manager
2	Imaging	15	Meridian; P2000A & B; Axis 1, 2, & 3	Technologists; Manager
3	Patient Wait	150-180	Waiting	
4	Imaging	45	Axis 1, 2, & 3; P3000; Meridian	Technologists; Manager

fours steps and a minimum completion time of 230 minutes. This procedure requires the utilization of one radiopharmaceutical, three stations, and three members of the staff.

c. Stations and Equipment

Nuclear medicine departments are subdivided into stations where procedures/tests are performed. Each station contains at least one type of equipment. Stations are classified depending on the equipment they contain. Nuclear medicine equipment include different types of gamma cameras and treadmills for cardiovascular tests. All the entities needed to perform a procedure step have to be present in the station before starting any activity. For example, in order to perform a scan the technologist and the patient have to be present at the station and the camera has to be configured to take the appropriate image. The time spent by these entities in the stations will depend on several factors such as the expertise of the human resource and the procedure protocol.

d. Patients

Requests for procedures are managed by a call center. A receptionist takes care of these requests by finding an appointment in the system for the patient. Requests are always for a single procedure and sometimes patients will provide a preference for the appointment date. The appointment provided to the patient is determined by the scheduling policies or algorithms used by the clinic.

2. Performance Measures

We evaluate the performance of our scheduling algorithms using measures that take into account the perspectives of both the clinic manager and the patients. The selected performance measures were identified as commonly used in literature. Table XI provide a description of performance measures used to quantify the level of patient satisfaction in health care clinics. Performance measures that consider the perspective of managers are provided in Table XII.

Table XI. Performance measures from patient's perspective

Name	Description	Reference
Waiting time type 1	Waiting time from the time of the procedure request until the time of the appointment	[18], [52]
Preference ratio	Number of times patients are scheduled on the date requested above all patient requests	[61]
Cycle time	Time patient spends in the system	[91], [92]

Table XII. Performance measures from manager's perspective

Name	Description	Refs.
Equipment utilization	The amount of time an equipment is used during operating hours	[18], [90]
Human resource utilization	The amount of time a human resource is used during operating hours	[90], [40]
Patient throughput	Number of patients served per day	[41]

To put all the pieces together, we now give a simple example (Figure 17) to illustrate patient/resource schedules in nuclear medicine. Figure 17 (a) shows two of the procedures performed in nuclear medicine. We list the step requirements for each procedure as follows: time duration ($time$), station (s), and human resource (r). For the purpose of this example, only one station and human resource is listed per procedure step but is important to keep in mind that several stations and human resources can be used to perform a procedure step in nuclear medicine (see Table X). Figure 17 (b) depicts a schedule for procedure 78815 where the patient is assigned to come at the beginning of the day. The schedule shows that to perform this procedure four resources are required at different times of the day. Figure 17 (c) shows the schedule for a second procedure (78465). Since some of the resources are unavailable at the beginning of the day the procedure has to be scheduled later in the day. The schedule shows that for this procedure five different resources are required. Also, the figure shows that no other procedure can be fitted into the schedule due to the unavailability of the resources at particular times. For instance, a second procedure 78465 cannot be schedule because the axis station and the technologist are unavailable during the time slots that will be required for the last step of the procedure.

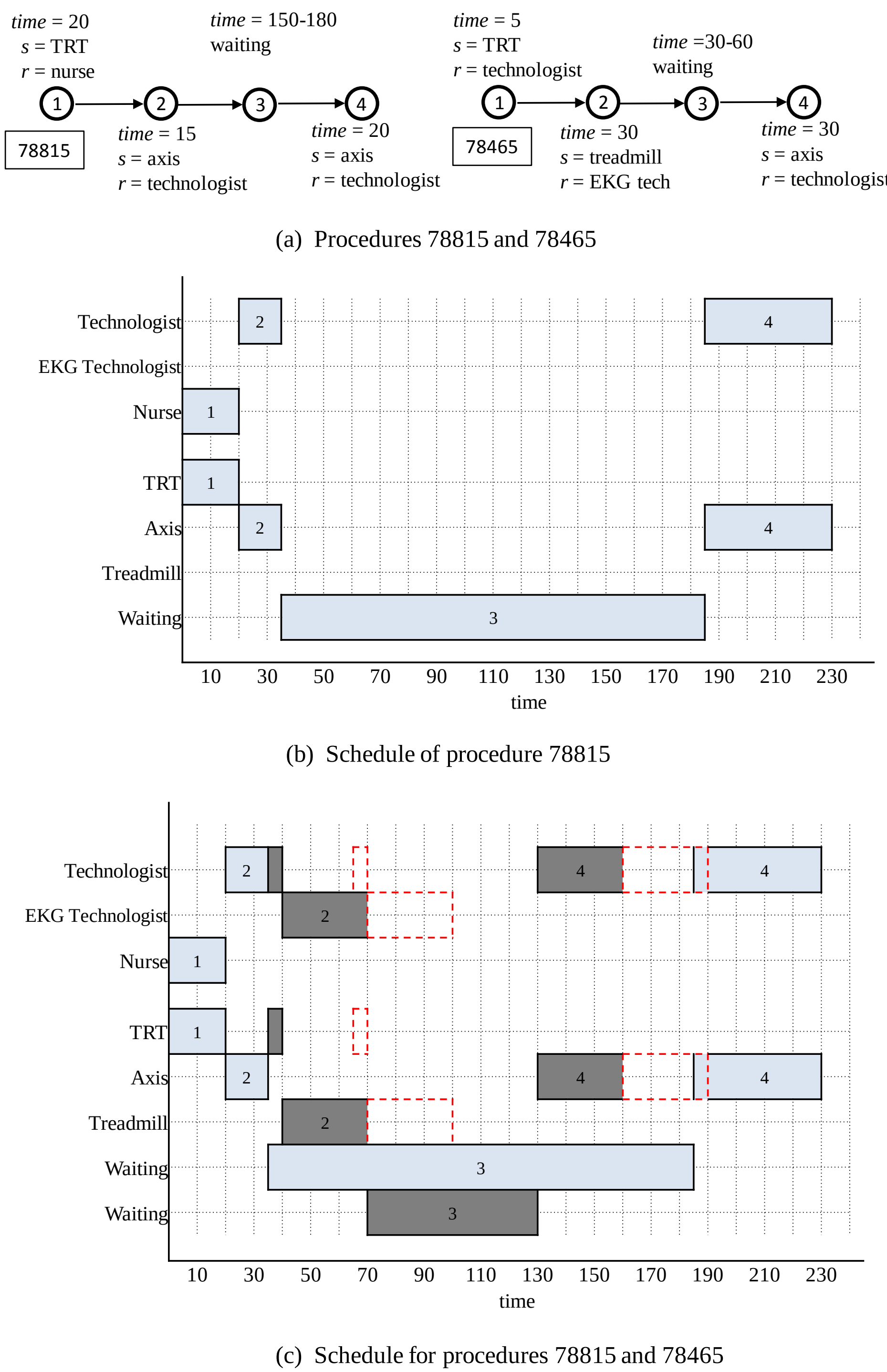

Fig. 17. Example showing one and two scheduled procedures

C. Patient and Resource Scheduling

We now turn to patient and resource scheduling in nuclear medicine and derive two algorithms named: *no fixed resource* (NFR) and *fixed resource* (FR). NFR algorithm schedules patients based on a patient's preferred day for the appointment. However, if the preference provided by the patient results in an appointment where the patient has to wait more than a month, an earlier appointment for an alternate day is considered. Like the NFR algorithm, the FR algorithm schedules patients by first considering their preferred day. However, unlike the NFR algorithm, some of the members of the staff are fixed to specific stations. In other words, some human resources are dedicated to specific stations. We derived the FR algorithm based on the real-life practical experience. We assume that no more than one patient can be scheduled to use the same resource at the same time and that the scheduling horizon is long enough so that no patient request is dropped.

We use the notation in Table XIII to mathematically describe the algorithms using pseudocode. We also use the following symbols: $\leftarrow$ denotes assignment; $==$ denotes (equality) comparison, and && denotes logic "and". We define the set of day and time slot pairs (d, t) for resource r as $X_r = \{(d, t)|\ 1 \leq d \leq h,\ 1 \leq t \leq \tau\}$. Similarly, we define the set of day and time slot pairs (d, t) for station s as $Y_s = \{(d, t)|\ 1 \leq d \leq h,\ 1 \leq t \leq \tau\}$. The sets X_r and Y_s include all the time slots that are already scheduled. The set of day and time slot pairs (d, t) for patient j is defined as $A_j = \{(d, t)|\ 1 \leq d \leq h,\ 1 \leq t \leq \tau\}$.

For easy of exposition, we first describe a *method* (function) we refer to as `CheckSchedule`() (Figure 18), which is implemented by both the NFR and FR algorithms. This method checks the availability of a human resource (when $\rho = r$) or a station (when $\rho = s$) during a given time interval t to $t + a_{kp}$, and returns a boolean

Table XIII. Scheduling algorithms sets and parameters

	Sets
J :	Set of patients, indexed j.
S :	Set of stations, indexed s.
R :	Set of human resources, indexed r.
P :	Set of nuclear medicine procedures, indexed p.
R_{kp} :	Set of human resources qualified to perform step k of procedure p.
S_{kp} :	Set of stations where step k of procedure p can be performed.
A_j :	Set of day and time slot pairs, (d, t), for patient j.
X_r :	Set of day and time slot pairs, (d, t), for human resource r schedule.
Y_s :	Set of day and time slot pairs, (d, t), for station s schedule.
	Parameters
d_j :	Call day for patient j.
t_j :	Call time for patient j.
p_j :	Procedure requested by patient j.
ω :	Number of days in a week.
μ :	Number of days in a month.
q :	Day of the week requested by the patient, indexed $q = 1, \cdots, 5$, where 1=Monday, 2=Tuesday, $\cdots$, 5=Friday .
n_p :	Total number of steps for procedure p, indexed $k = 1, \cdots, n_p$.
δ_p :	Number of time slots required for procedure p.
a_{kp} :	Number of time slots required for step k of procedure p
τ :	Total number of time slots in a day, indexed $t = 1, \cdots, \tau$.
m :	Number of days before arrival of radiopharmaceutical after placing order.
α :	First day that can be used to schedule an appointment for patient j.
ρ :	variable representing resource r or station s.
h :	Total number of days in the scheduling horizon.

indicating whether or not that time interval is available. If any one of the time slots within the interval t to $t+a_{kp}$ is occupied, that it is in the set X_ρ, the method returns $false$, otherwise it returns $true$. The method simply checks whether or not the time slots from time t to $t+a_{kp}$ are included in the current schedule.

The NFR and FR algorithms share the same overall structure described by the pseudocode in Figure 19. The set of patients J is initialized in line 1. Lines 2 and 3 define the time horizon (day and time) when patient requests will be received. Patient requests are received (line 4) and added to set J as they arrive (line 5). A method called `ServeRequest-Algorithm`(), where `Algorithm` denotes NFR or FR, takes patient information and finds an appointment (line 6). The two algorithms differ in the way they implement this method.

`CheckSchedule` (ρ, d, t, a_{kp})

$X \leftarrow X_\rho$
for $time = t$ **to** $t + a_{kp}$ **do**
    if $(d, time) \in X_\rho$
        return $false$
    else $time \leftarrow time + 1$
return $true$

Fig. 18. Pseudocode for `CheckSchedule`()

`Scheduling-Algorithm`

1 $J \leftarrow \{\emptyset\}, \;\; j = 0;$
2 **while** $d \leq \overline{h}$
3   **while** $t \leq \tau$
4     $(p_j, q) \leftarrow$`GetPatientRequest(j)`;
5   **do** $J \cup \{p_j\}, \;\; d_j \leftarrow d, \;\; t_j \leftarrow t, \;\; j \leftarrow j+1;$
6         $A_j \leftarrow$ `ServeRequest-Algorithm` $(j, p_j, d_j, t_j, q);$

Fig. 19. Pseudocode for `Scheduling-Algorithm`

1. NFR Algorithm

The NFR algorithm is invoked by the method `ServeRequest-NFR`() summarized in Figure 20. The algorithm allows the patient (j) to provide a preferred day of the week (q) for the appointment. Using this preferred day, a search for an appointment is performed based on the patient's requested procedure p_j. Once an appointment is found, the algorithm checks that the waiting time for the patient is not greater than a month. If it is, the algorithm checks for an alternative appointment (different day of the week) with lesser waiting time. We use a boolean variable θ to store the output returned by the `CheckSchedule`() method. Recall that θ takes a value of $true$ if a resource (human r or station s) is available, and $false$ otherwise.

The parameters of the procedure (p_j) requested by patient j and the set A_j are initialized in line 1. The starting day for searching for an appointment is determined in line 2 using the following information: day when the request was received (d_j), patient preferred day (q), and the number of days (m) needed to obtain the radiopharmaceutical for procedure (p_j). The scheduling horizon is defined in lines 3 and 4. Next a search for an available combination of human resource and station for each step (k) of the procedure is performed. A breath-first search to select a human resource from set R as well as to select a station from set S (lines 5-11). We decided to use breath-first search to ensure a balanced work assignment between the resources. If no human resource is available to serve one of the procedure steps (line 24), the current start time for the procedure is incremented (line 25). Using the new start time, the algorithm checks if the amount of time remaining on the day searched is enough to accommodate the procedure (line 26). If that is the case, the algorithm begins a new search for an appointment using the new time as a starting time. Otherwise, the algorithm moves to the following week to perform a new search on the

`ServeRequest-NFR` (j, p_j, d_j, t_j, q)

1 Initialize : $p \leftarrow p_j;\ \delta \leftarrow \delta_p;\ n \leftarrow n_p;\ A_j \leftarrow \{\emptyset\}$
2 Find : $\left\lceil \frac{d_j+m-q}{\omega} \right\rceil$ then $\alpha \leftarrow q + x\omega$
3 **for** $day = \alpha$ **to** $\overline{h}$ **do**
4   **for** $time = 1$ **to** τ **do**
5   **for** $k = 1$ **to** n **do**
6       $a \leftarrow a_{kp};\ \ R \leftarrow R_{kp};$ and $S \leftarrow S_{kp}$
7   **while** $R \neq \{\emptyset\}$
8         $\theta_r \leftarrow$ `CheckSchedule`$(r, day, time, a)$, where $r \in R$
9   **if** $\theta_r ==$ *true*
10   **while** $S \neq \{\emptyset\}$
11             $\theta_s \leftarrow$ `CheckSchedule`$(s, day, time, a)$, where $s \in S$
12   **if** $\theta_s ==$ *true*
13   **if** $k == 1$ `&&` $day - d_j > \mu$
14                 $A_j \leftarrow$ `NoPreference`(j, p, d_j, t_j)
15   **return** A_j
16   **else**
17                 $A_j \cup \{(day, time)\};\ X_r \cup \{(day, time)\};$
                   $Y_s \cup \{(day, time)\}$
18                 $time \leftarrow time + a;\ \ k \leftarrow k + 1;$ and **go to step 6**
19   **if** $S == \{\emptyset\}$
20             $time \leftarrow time + 1$
21   **if** $time + \delta > \tau$
22               $day \leftarrow day + \omega;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
                 and **go to step 4**
23   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
24   **if** $R == \{\emptyset\}$
25         $time \leftarrow time + 1$
26   **if** $time + \delta > \tau$
27           $day \leftarrow day + \omega;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
             and **go to step 4**
28   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
29   **return** A_j

Fig. 20. Pseudocode for NFR

`NoPreference` (j, p_j, d_j, t_j)

1 Initialize : $p \leftarrow p_j;\ \delta \leftarrow \delta_p;\ n \leftarrow n_p;\ A_j \leftarrow \{\emptyset\};\ \alpha \leftarrow m + d_j$
2 **for** $day = \alpha$ **to** $\overline{h}$ **do**
3   **for** $time = 0$ **to** τ **do**
4   **for** $k = 1$ **to** n **do**
5       $a \leftarrow a_{kp};\ R \leftarrow R_{kp};$ and $S \leftarrow S_{kp}$
6   **while** $R \neq \{\emptyset\}$
7         $\theta_r \leftarrow$ `CheckSchedule`$(r, day, time, a)$, where $r \in R$
8   **if** $\theta_r ==$ *true*
9   **while** $S \neq \{\emptyset\}$
10             $\theta_s \leftarrow$ `CheckSchedule`$(s, day, time, a)$, where $s \in S$
11   **if** $\theta_s ==$ *true*
12                 $A_j \cup \{(day, time)\};\ X_r \cup \{(day, time)\};$
                   $Y_s \cup \{(day, time)\}$
13                 $time \leftarrow time + a;\ k \leftarrow k + 1;$ and **go to step 6**
14   **if** $S == \{\emptyset\}$
15             $time \leftarrow time + 1$
16   **if** $time + \delta > \tau$
17               $day \leftarrow day + 1;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
                 and **go to step 4**
18   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
19   **if** $R == \{\emptyset\}$
20         $time \leftarrow time + 1$
21   **if** $time + \delta > \tau$
22           $day \leftarrow day + 1;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
             and **go to step 4**
23   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
24   **return** A_j

Fig. 21. Pseudocode for `NoPreference` method for NFR

day requested by the patient (line 27). The algorithm follows the same steps (lines 24-28) when no station is available (lines 19-23). A patient schedule is returned if a combination of human resource and station is found for each procedure step and if the waiting time for the appointment is less than a month (line 29). If the waiting time exceeds a month, the `NoPreference` () method (Figure 21) is invoked (lines 13-15).

The `NoPreference` () method schedules patients in the first space available in the scheduling horizon. The earliest the appointment can be scheduled is determined by the radiopharmaceutical arrival time (line 1). If no human resource is available to serve one of the procedure steps (line 19) the current start time for the procedure is incremented (line 20). Then the algorithm checks if the amount of time remaining on the day for the search is enough to accommodate the procedure (line 21). If that is the case, the method begins a new search for an appointment using the new time as a starting time. If the time is not enough, a new search begins on the next day (line 22). The same course of action is followed when no station is available (lines 14-18). The algorithm returns a patient schedule when a combination of human resource and station is found for each procedure step (line 24).

2. FR Algorithm

The FR algorithm is a variation of the NFR algorithm and has been used in a practical setting. This algorithm first tries to satisfy the preference provided by the patient; but when the waiting time is longer than a month, it performs a search for an alternative earlier appointment. FR and NFR algorithms differ in the way the human resources are assigned to patients. In the FR algorithm a group of human resources (e.g. two technologists) are assigned to always serve patients in specific stations. For example, technologist 1 and technologist 2 have to serve all patients whose procedures require

ServeRequest-FR (j, p_j, d_j, t_j, q)

1 Initialize : $p \leftarrow p_j;\ \delta \leftarrow \delta_p;\ n \leftarrow n_p;\ A_j \leftarrow \{\emptyset\}$
2 Find : $\left\lceil \frac{d_j+m-q}{\omega} \right\rceil$ then $\alpha \leftarrow q + x\omega$
3 **for** $day = \alpha$ **to** $\overline{h}$ **do**
4   **for** $time = 0$ **to** τ **do**
5   **for** $k = 1$ **to** n **do**
6       $a \leftarrow a_{kp};\ R \leftarrow R_{kp};$ and $S \leftarrow S_{kp}$
7   **while** $R \neq \{\emptyset\}$
8         $\theta_r \leftarrow$ `CheckSchedule`$(r, day, time, a)$, where $r \in R$
9   **if** $\theta_r == true$ `&&` $r == \hat{r}$
10   **if** $k == 1$ `&&` $day - d_j > \mu$
11             $A_j \leftarrow$ `NoPreference`(j, p, d_j, t_j)
12   **return** A_j
13   **else**
14             $A_j\ \cup\ \{(day, time)\};\ X_r\ \cup\ \{(day, time)\};\ Y_s\ \cup\ \{(day, time)\}$
15             $time \leftarrow time + a;\ \ k \leftarrow k + 1;$ and **go to step 6**
16   **else if** $\theta_r == true$
17   **while** $S \neq \{\emptyset\}$
18             $\theta_s \leftarrow$ `CheckSchedule`$(s, day, time, a)$, where $s \in S$
19   **if** $\theta_s == true$
20   **if** $k == 1$ `&&` $day - d_j > \mu$
21                 $A_j \leftarrow$ `NoPreference`(j, p, d_j, t_j)
22   **return** A_j
23   **else**
24                 $A_j\ \cup\ \{(day, time)\};\ X_r\ \cup\ \{(day, time)\};\ Y_s\ \cup\ \{(day, time)\}$
25                 $time \leftarrow time + a;\ \ k \leftarrow k + 1;$ and **go to step 6**
26   **if** $S == \{\emptyset\}$
27             $time \leftarrow time + 1$
28   **if** $time + \delta > \tau$
29               $day \leftarrow day + \omega;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
              and **go to step 4**
30   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
31   **if** $R == \{\emptyset\}$
32         $time \leftarrow time + 1$
33   **if** $time + \delta > \tau$
34           $day \leftarrow day + \omega;\ A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$
          and **go to step 4**
35   **else** $A_j \leftarrow \{\emptyset\};\ X_r \leftarrow \{\emptyset\};\ Y_s \leftarrow \{\emptyset\};$ and **go to step 5**
36   **return** A_j

Fig. 22. Pseudocode for FR

the use of stations s^1 and s^2. We use $\hat{r}$ to denote the human resources that may be fixed to stations. The FR algorithm is invoked by the method `ServeRequest-FR`() and the pseudocode is presented in Figure 22.

The pseudocode of the FR algorithm follows the major steps of the NFR algorithm. However, FR has an additional condition to identify those human resources that have been fixed to stations. A fixed human resource is always assigned to the same station. Consequently, when the availability of a fixed human resource is confirmed, the algorithm also confirms the availability of the station that share the same schedule (9-12). The algorithm returns a patient schedule if a combination of human resource and station is found for each procedure step and if the waiting time for the appointment is less than a month (line 36). The `NoPreference` () method is invoked (lines 20- 22) if the waiting time exceeds a month. A description of this method is provided in Figure 23. This method schedules patients in the first space available in the scheduling horizon and is similar to the one described in Section 1. However, this method includes the condition that identifies the human resources that are fixed to stations.

3. Algorithm Extensions

We now propose two extensions for the NFR algorithm. These two extensions are derived under the assumption that historical information about patient demand at the clinic studied is available. This information is used to reserve some of the clinic stations to be exclusively use by those procedures that are requested more often. We refer to these algorithms as NFR_FP (*No Fixed Resource Fixed Procedure*) and NFR_FP2, and they can be summarized as follows:

- NFR_FP: Determine the procedure that is requested the most at the clinic and

```
NoPreference (j, p_j, d_j, t_j)
1  Initialize : p ← p_j; δ ← δ_p; n ← n_p; A_j ← {∅}; α ← m + d_j
2  for day = α to \bar{h} do
3    for time = 0 to τ do
4      for k = 1 to n do
5        a ← a_{kp}; R ← R_{kp}; and S ← S_{kp}
6        while R ≠ {∅}
7          θ_r ← CheckSchedule(r, day, time, a), where r ∈ R
8          if θ_r == true && r == \hat{r}
9            A_j ∪ {(day, time)}; X_r ∪ {(day, time)}; Y_s ∪ {(day, time)}
10           time ← time + a; k ← k + 1; and go to step 5
11         else if θ_r == true
12           while S ≠ {∅}
13             θ_s ← CheckSchedule(s, day, time, a), where s ∈ S
14             if θ_s == true
15               A_j ∪ {(day, time)}; X_r ∪ {(day, time)}; Y_s ∪ {(day, time)}
16               time ← time + a; k ← k + 1; and go to step 5
17           if S == {∅}
18             time ← time + 1
19             if time + δ > τ
20               day ← day + 1; A_j ← {∅}; X_r ← {∅}; Y_s ← {∅};
                 and go to step 4
21             else A_j ← {∅}; X_r ← {∅}; Y_s ← {∅}; and go to step 5
22       if R == {∅}
23         time ← time + 1
24         if time + δ > τ
25           day ← day+1; A_j ← {∅}; X_r ← {∅}; Y_s ← {∅}; and go to step 4
26         else A_j ← {∅}; X_r ← {∅}; Y_s ← {∅}; and go to step 5
27     return A_j
```

Fig. 23. Pseudocode for `NoPreference` method for FR

the stations that can be used to perform this procedure. From the stations identified, reserve a subset to be exclusively used to serve this procedure. The size of the subset of stations depends on how often the procedure is requested. For instance, if a procedure is requested 40% of the time, the algorithm reserves 40% of the stations identified to serve this procedure.

- NFR_FP2: Reserve stations for the procedure requested the most as explained in NFR_FP. However, in this algorithm the stations reserved can also be used to schedule the second procedure requested the most. If the procedure does not have similar stations requirements; use the next procedure that is requested the most.

D. Application

To evaluate the performance of our algorithms we applied them to an actual nuclear medicine setting. The algorithms were implemented within the nuclear medicine department simulation presented by [94] using historical data for a year of operations. We first describe the configuration of the nuclear medicine setting we used and the experimental setup, and then report the computational results and findings.

1. Nuclear Medicine Setting

The scheduling algorithms presented in this chapter were used to study one of the largest fully accredited nuclear medicine laboratories in the country located at the Scott and White Health System in Temple, Texas. This nuclear medicine clinic operates nine hours days a week and is not open during weekends. The clinic has 12 stations which are named according to the type of equipment they contain. Table XIV provides the names of the stations and the type of equipment each station contains.

Table XIV. Stations of the Scott and White nuclear medicine clinic

Station name	Number	Equipment
Axis	3	Philips Axis Camera
P2000	2	Philips PRISM 2000 Camera
P3000	1	Philips PRISM 3000 Camera
Meridian	1	Philips Meridian Camera
Treadmill	2	Treadmill
TRT	3	Used for patient preparation

The staff is composed of technologists, nurses, and a manager. There are eight technologists and two EKG technologists in this clinic. The technologists have several responsibilities that include drawing doses and image acquisition. EKG technologists perform stress exams for cardiac tests and they are assigned to the Treadmill stations. A nurse is in charge of helping to draw doses. In the absence of one of the staff members, the division manager can perform that staff member's tasks (see Table VIII). Table IX of Section B shows the procedures that were performed more frequently at the clinic for a year of observation.

2. Experimental Setup

The configuration of the clinic used for test our algorithms was based on historical data. The clinic has twelve stations named: TRT(1), TRT(2), TRT(3), Treadmill(1), Treadmill(2), Axis(1), Axis(2), Axis(3), P2000(1), P2000(2), P2000(3), and Meridian(1). The staff has twelve members: Technologist(1), Technologist(2), Technologist(3), Technologist(4), Technologist(5), Technologist(6), Technologist(7), Technologist(8), Technologist(9), Technologist(10), Nurse(1), and Manager(1). A Poisson process was assumed for procedure request arrivals based on historical data. The monthly call interarrivals times in minutes followed an exponential distribution with the following means: January, 6.00; February, 6.25; March 6.58; April, 6.67; May, 6.75; June, 6.88; July, 6.96; August, 7.04; September, 7.10; October, 7.29; November, 7.34; and December, 7.44. Empirical distributions were used to generate a procedure request and an appointment preferred day for each patient

We conducted experiments to gain management insights into the impact of the scheduling approaches on patient service. In our computational study, we only considered the nuclear medicine procedures listed in Table IX. The performance measures listed in Section B were used to quantify service levels based on both patient and management perspectives. We performed computational experiments using the scheduling algorithms presented in Section C. Under the FR algorithm Technologist(1) and Technologist(2) are fixed to station Axis(1) and stations Axis(2). Under NFR_FP 40% of the stations are assigned to procedure 78465 and under NFR_FP2 these stations are shared with procedure 78815. A second set of experiments was performed to check the performance of the algorithms when capacity is added to the system.

The algorithms were implemented in Java within the simulation model for a nuclear medicine department [94]. The experiments involved 100 replications, using a one-year period of operations with a three-month warm-up period. To maintain independence of each replication, different seeds were used in the pseudo random number generators for each simulation run. All the experiments were executed on a DELL Optiplex GX620 with a Pentium D processor running at 3.2GHz with 3.0GB RAM.

3. Computational Results

Table XV shows the throughput and the computation times for the four algorithms. We report the mean, standard deviation (Std.Dev.), and a 95 % confidence interval for each performance measure. The results show that NFR and FR obtained similar figures for the average number of patients served during the year. However, these numbers were exceeded by the NFR_FP and NFR_FP2 algorithms, with NFR_FP obtaining the best performance with an average of 14,845 patients served during that year. The computer times (CPU secs.) were about the same for all the algorithms but it is evident that the computer times depend on the numbers of patients scheduled by the system.

Table XV. Scheduling algorithms results

Algorithm	Statistic	Patients Served	Patients/hr	Patients/day	CPU Time (secs)
NFR	Mean	14782.84	6.84	61.60	614.44
	Std. Dev.	78.99	0.04	0.33	6.47
	CI Lower	14767.17	6.84	61.53	613.15
	CI Upper	14798.51	6.85	61.66	615.72
FR	Mean	14776.20	6.84	61.57	611.39
	Std. Dev.	78.65	0.04	0.33	9.09
	CI Lower	14760.60	6.83	61.50	609.58
	CI Upper	14791.80	6.85	61.63	613.19
NFR_FP	Mean	14845.39	6.87	61.86	653.29
	Std. Dev.	80.16	0.04	0.33	48.76
	CI Lower	14829.48	6.87	61.79	643.61
	CI Upper	14861.29	6.88	61.92	662.96
NFR_FP2	Mean	14832.59	6.87	61.80	641.30
	Std. Dev.	81.78	0.04	0.34	18.86
	CI Lower	14819.01	6.86	61.75	638.17
	CI Upper	14846.16	6.87	61.86	644.43

Figure 24 shows the average number of patients served per month for each scheduling algorithm. All the algorithms exhibit a decreasing behavior due to the historical data used to define the interarrival times between patient requests in the simulation. However, two different trends are observed, one for NFR and FR and the other for NFR_FP and NFR_FP2. NFR and FR are able to schedule more patients from January to April. NFR_FP2 served more patients from May to June than

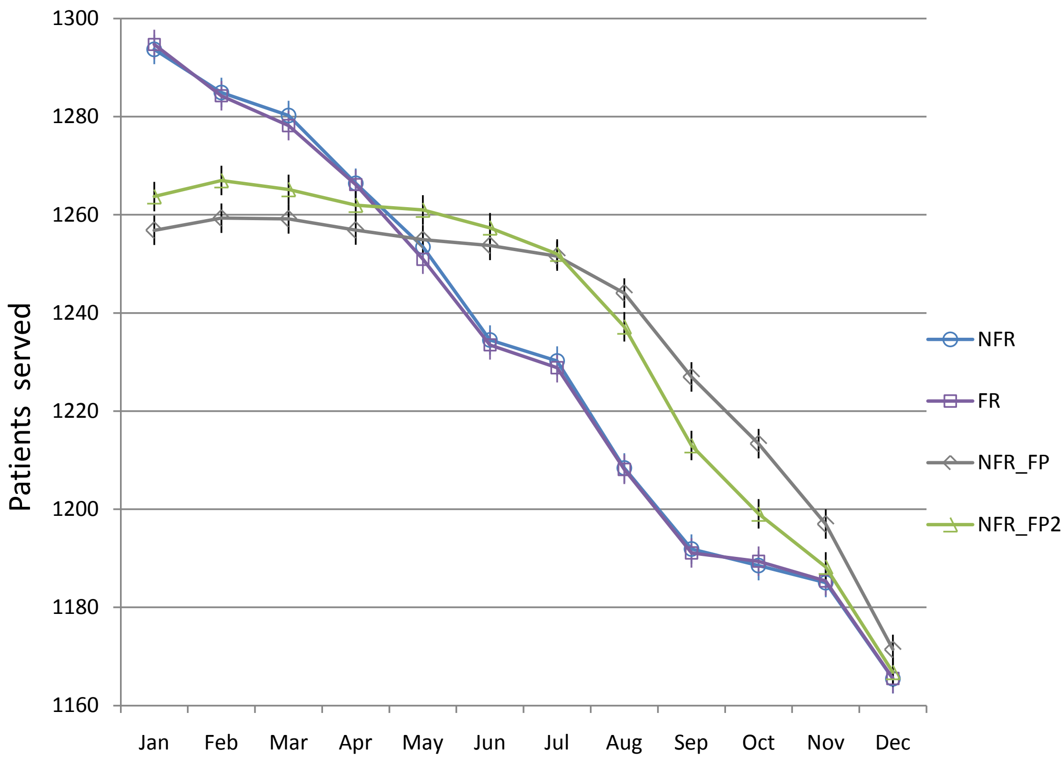

Fig. 24. Number of patients served per month using NFR, FR, NFR_FP, and NFR_FP2 algorithms

NFR_FP but both algorithms served the same number of patients in the month of July. From August to December, NFR_FP served more patients than NFR_FP2. The station reservation executed in the NFR_FP and NFR_P2 algorithms has two major consequences. First, patient schedules are arranged so that more patients are accommodated into the system. Second, the number of patients that are seen per month is limited by available system capacity. For instance, Figure 24 shows that during the first three months of the year when the patient requests were more frequent, the number of patients served under NFR_FP was never higher than 1270. Since NFR_FP has more stations reserved to one procedure (CPT 78845), the capacity for serving procedure requests is more limited compared to the other algorithms.

Figure 25 shows equipment utilization for each algorithm. Equipment utilization

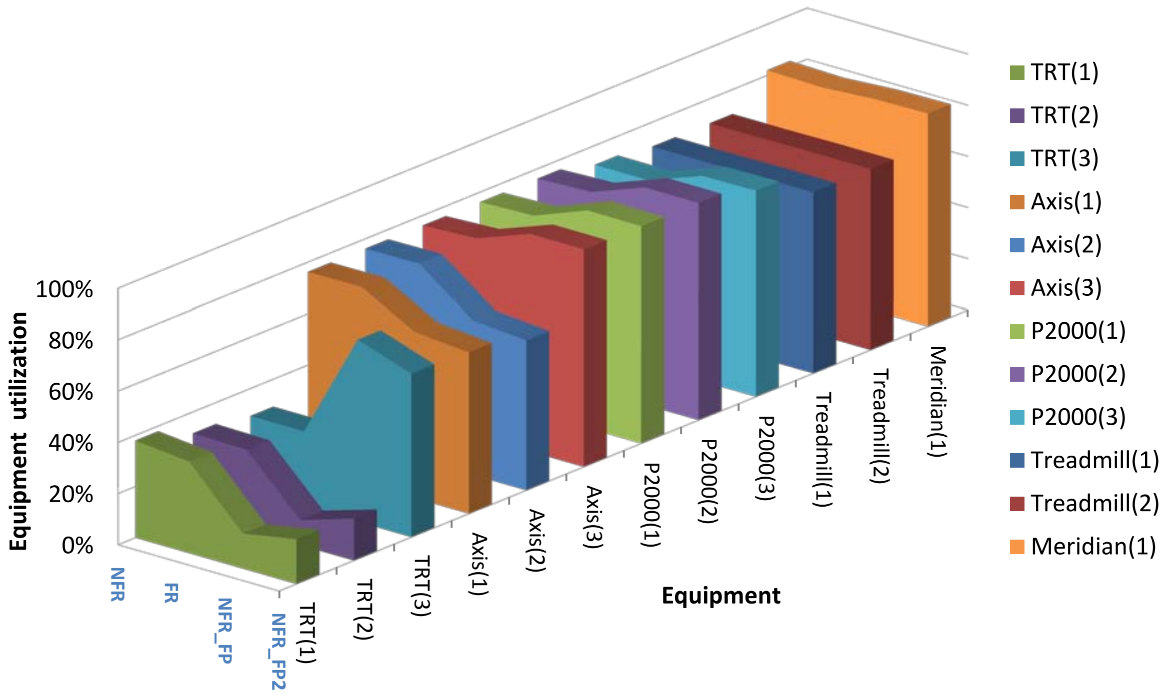

Fig. 25. Equipment utilization using NFR, FR, NFR_FP, and NFR_FP2 algorithms

under NFR and FR is about the same for all the equipment. NFR_FP algorithm shows a decrease in utilization for stations TRT(1) and TRT(2) when compared to NFR and FR. These two stations were only used to serve procedure CPT 78845. This procedure only uses TRT stations on the first step. Even though the procedure is scheduled very often, the utilization reflected is low because the procedure step has a small duration. Station TRT(3) shows a notable increase in utilization under NFR_FP when compared to NFR and FR. Since the other two TRT stations are assigned to a single procedure, TRT(3) is now used to serve all the other procedure requests that require the use of a TRT station. NFR_FP also shows a decrease in the utilizations of stations Axis(1) and Axis(2), which are stations now assigned to only serve procedure 1. This assignment also causes the increase in the utilizations of Axis(3), P2000(1), and P2000(2). Under the NFR_FP2 algorithm the equipment utilizations were similar to the NFR_FP, with small differences in the utilizations of the TRT stations. NFR_FP2 reserves TRT(1) and TRT(2) to serve procedures CPT 78845 and CPT 78815. This change allows more flexibility when scheduling procedures which causes an increase in the utilizations of stations TRT(1) and TRT(2) and a decrease in the utilizations of the station TRT(3).

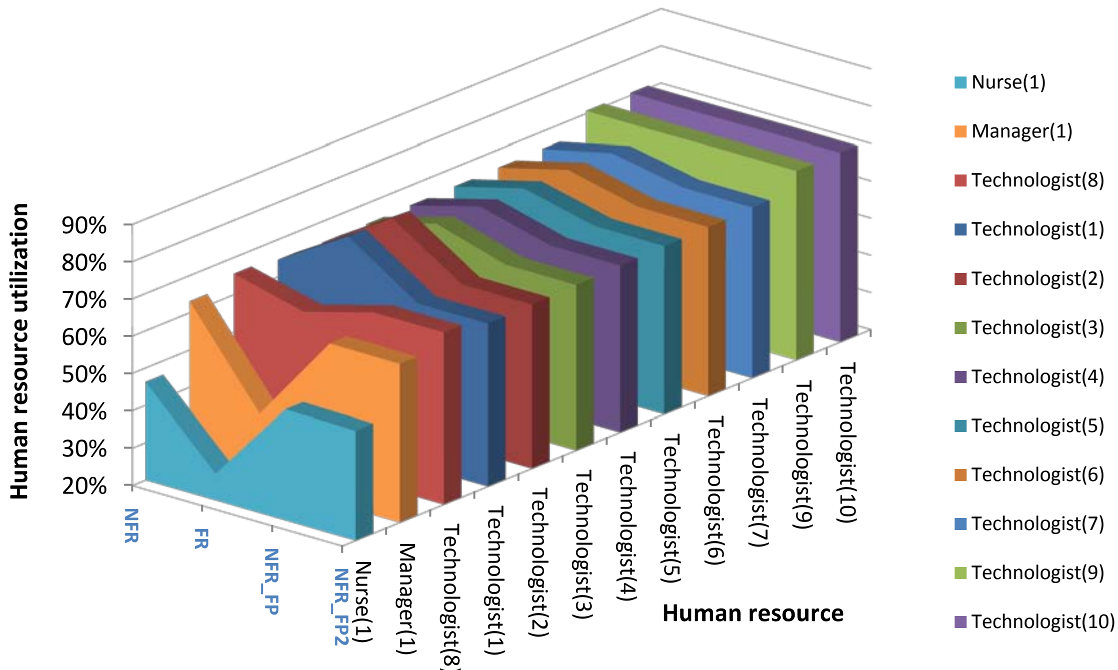

Fig. 26. Human resource utilization using NFR, FR, NFR_FP, and NFR_FP2 algorithms

Figure 26 depicts the human resource utilization at the clinic. Most of the algorithms show a balanced distribution of the work among the staff. A noticeable difference is under the FR algorithm, where the manager and the nurse have smaller utilization. Under FR, Technologist(1) and Technologist(2) are fixed to stations Axis(1) and Axis(2), respectively. These two human resources are assigned to camera stations that are used to perform procedures that are requested the most. Since the manager and the nurse have no access to these cameras under this algorithm, their utilization decreases and at the same time the utilization of both Technologist(1) and Technologist(2) increases.

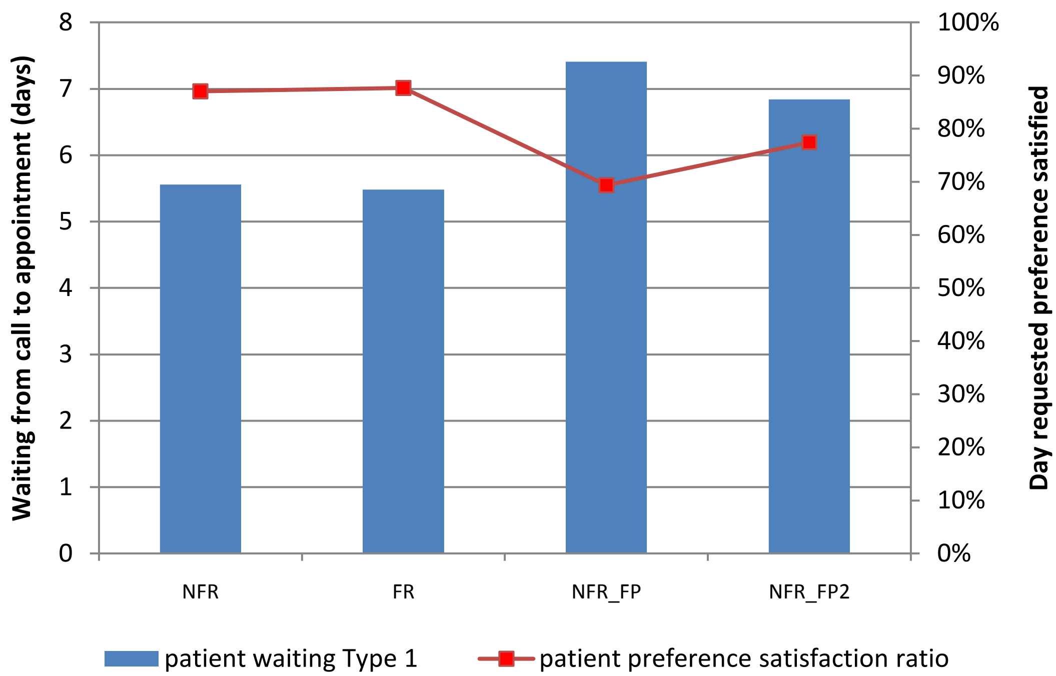

Fig. 27. Patient waiting *Type 1* and preference satisfaction ratio using NFR, FR, NFR_FP, and NFR_FP2 algorithms

Figure 27 shows the performance measures for patient service. NFR and FR algorithms performed better for both performance measures providing an average waiting time (Type 1) from call to appointment of about 5.5 days and day requested preference satisfaction ratio of about 87 %. NFR_FP algorithm provides the highest waiting from call to appointment with an average of 7.3 days and the lowest patient preference satisfaction ratio with 70 %. Even though this algorithm is able to accommodate more patients into the systems is evident that there is a price to pay in terms of the service provided to patients. By fixing several stations to one procedure, patient schedules are arranged so that more patients are accommodated into the system. However, this results in longer patient wait time and lesser patient preference satisfaction. NFR_FP2 demonstrates that by being less restrictive with one of the stations, patient service performance measures can be improved but with a tradeoff that results in a noticeable decrease in patient throughput towards the end of the year.

We now use the results obtained to study potential system capacity expansions that can lead to improved performance. The modifications are tested using the NFR_FP algorithm since it provided the highest throughput among the proposed algorithms Several modifications were considered but due to space limitation we only discuss the following two cases:

- NFR_FP_TRT: Under this policy station TRT(4) is added to the system.
- NFR_FP_MC : This policy adds more capacity (MC) to the system by including an additional EKG Technologist and the following stations: TRT(4), P2000(4), and Treadmill(3). A new EKG Technologist is required to operate the new Treadmill station.

Table XVI. Scheduling algorithms results for potential system capacity expansions

Algorithm	Statistic	Patients Served	Patients/hr	Patients/day	CPU Time (secs)
NFR_FP	Mean	14845.39	6.87	61.86	653.29
	Std. Dev.	80.16	0.04	0.33	48.76
	CI Lower	14829.48	6.87	61.79	643.61
	CI Upper	14861.29	6.88	61.92	662.96
NFR_FP_TRT	Mean	14838.80	6.87	61.83	653.28
	Std. Dev.	58.30	0.03	0.24	17.50
	CI Lower	14827.23	6.86	61.78	649.81
	CI Upper	14850.37	6.88	61.88	656.75
NFR_FP_MC	Mean	14875.53	6.89	61.98	743.72
	Std. Dev.	70.78	0.03	0.29	7.93
	CI Lower	14861.49	6.88	61.92	742.15
	CI Upper	14889.58	6.89	62.04	745.29

Table XVI shows the performance measures for the NFR_FP algorithm when additional stations are added to the system. NFR_FP_TRT shows an average throughput similar to NFR_FP. On the other hand the NFR_FP_MC obtained an improvement of around 30 more patients a year. This improvement in throughput was expected because several stations were added to the system.

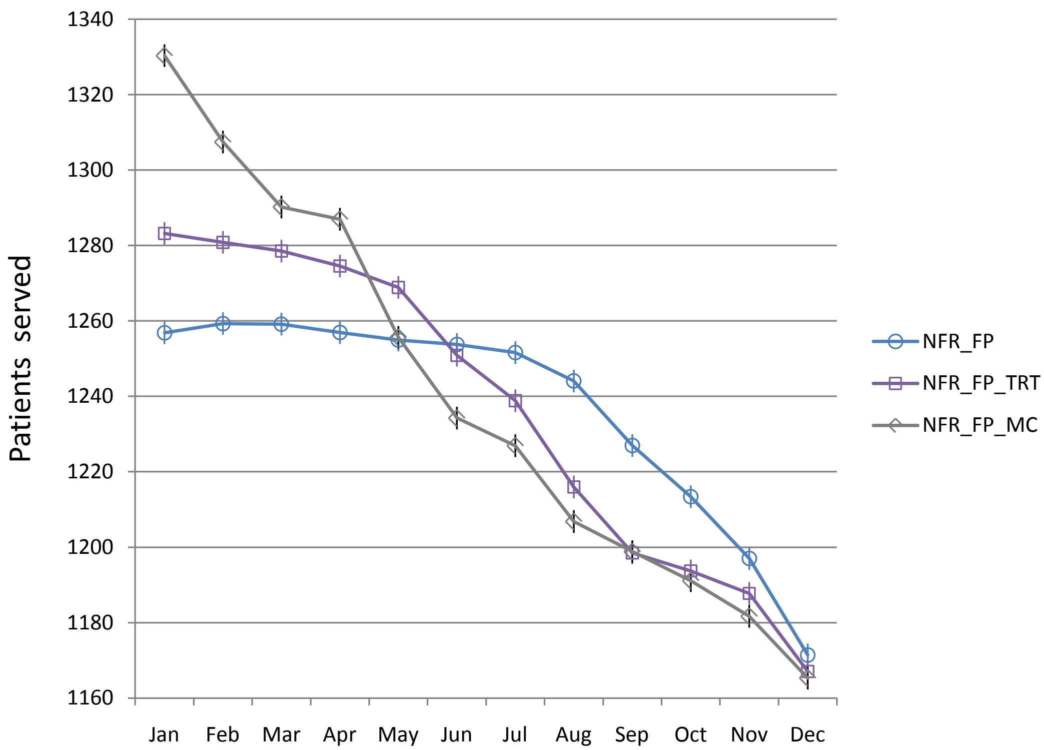

Fig. 28. Number of patients served per month using NFR_FP_TRT and NFR_FP_MC algorithms

Figure 28 shows the throughput per month for the algorithm under the new system configurations. The graph shows that adding station TRT(4) to the system helps by increasing the number of patients seen during the first months of the year where procedure requests are performed more frequently. Likewise, the figure shows that under the NFR_FP_MC the number of patients serve at the beginning of the year was improved even more.

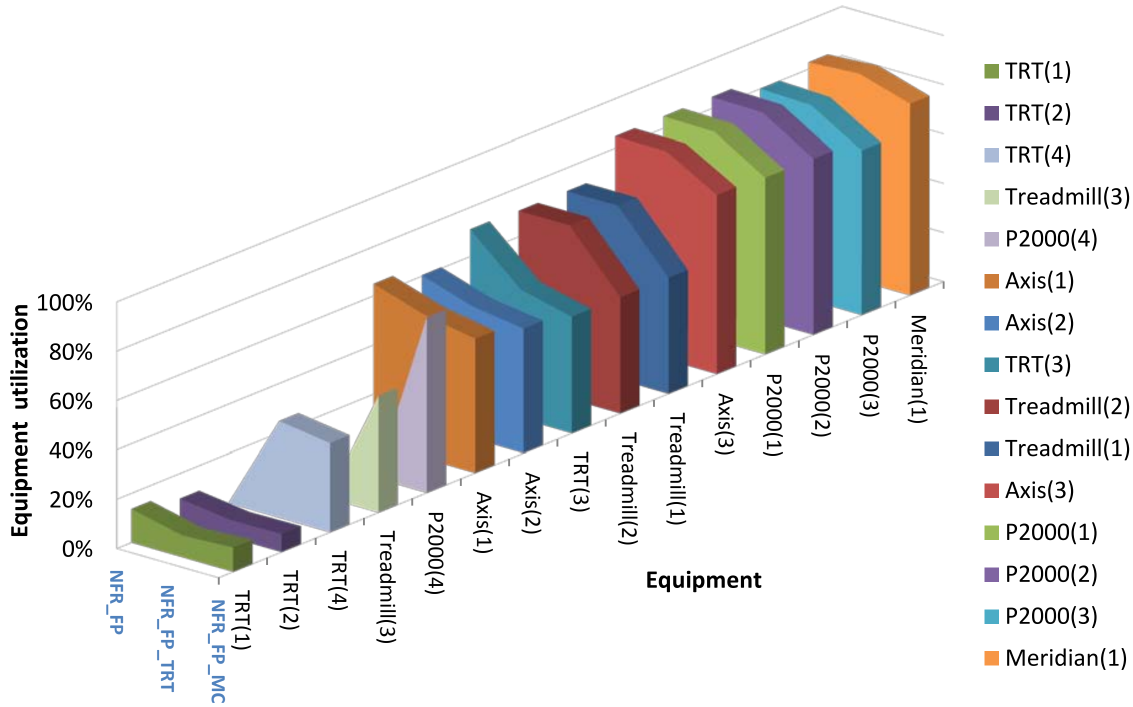

Fig. 29. Equipment utilization using NFR_FP_TRT and NFR_FP_MC algorithms

Figure 29 shows the effects of the system modifications on equipment utilization. The addition of a TRT(4) station results in reduced utilization of the other TRT stations under NFR_FP_TRT. The amount of work is now balanced between these stations, in particular between TRT(3) and TRT(4). Figure 30 shows the utilization for the staff when capacity is added to the system. The graph shows the utilization of the staff is balanced under the three alternatives.

Figure 31 displays the results obtained for the performance measures from patient's perspective. The results were improved significantly with the addition of the new stations. In terms of waiting times, a decrease of 1.3 days in average was achieved under NFR_FP_TRT and a decrease of 2.1 days average was achieved under NFR_FP_MC. On the other hand, an improvement in the patient preference satisfaction ration was achieved by both modifications with 85 % for both NFR_FP_TRT

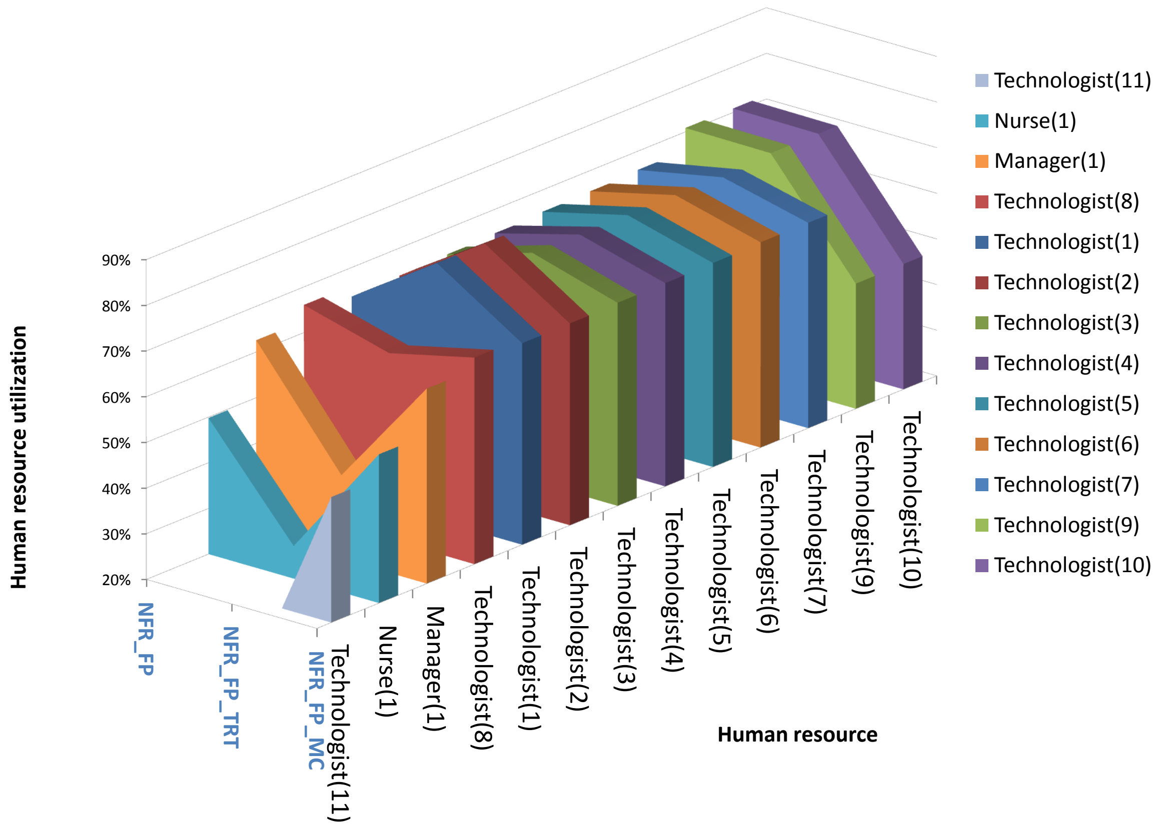

Fig. 30. Human resource utilization using NFR_FP_TRT and NFR_FP_MC algorithms

and 95% for NFR_FP_MC.

E. Summary

Managing patients in nuclear medicine departments is a challenging problem with limited research reported in the literature. The complexity involved in this health care setting makes this problem unique. In this chapter, we derive and implement algorithms for scheduling nuclear medicine patients and resources. The scheduling algorithms take into consideration the time constraints imposed by the decay of the

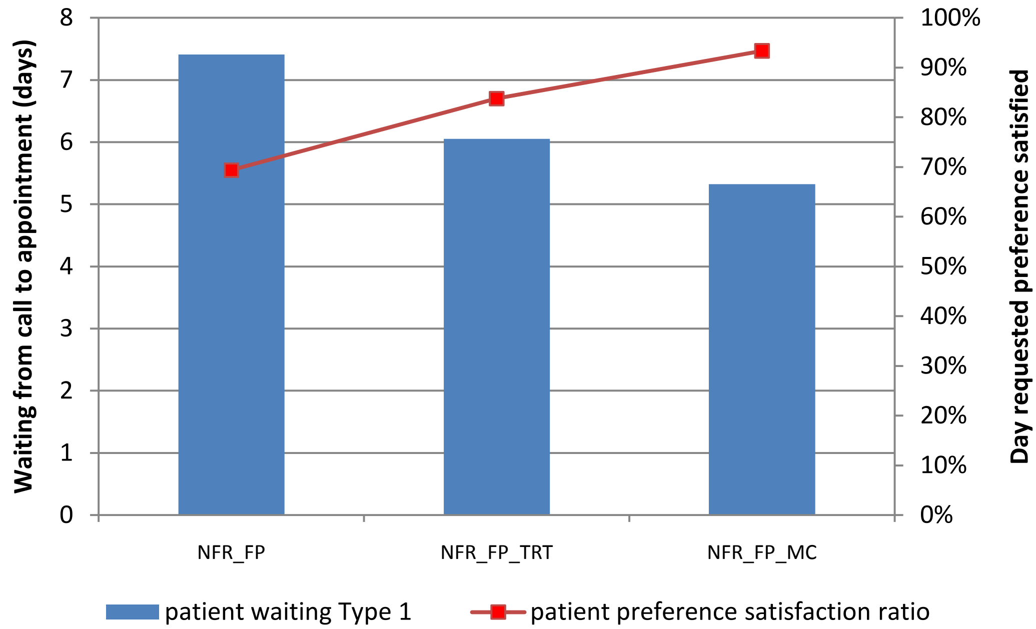

Fig. 31. Patient waiting *Type 1* and preference satisfaction ratio using NFR_FP_TRT and NFR_FP_MC algorithms

radiopharmaceuticals, which are required for most of the nuclear medicine procedures. The algorithms were implemented within a simulation framework and the experiments performed were based on historical data provided by an actual clinic. We obtain computational results that provide useful insights into patient service management in nuclear medicine. For example, no single patient and resource scheduling algorithm provides the best results relative to all performance measures. Thus, it is up to the nuclear medicine clinic to decide which algorithm to use under given demand and patient/management preferences. However, in terms of throughput the results show that reserving stations to be exclusively used by procedures that are requested frequently improves the number of patients served. The results also show that reserving stations for specific procedures affects patient service satisfaction by increasing the waiting times. In terms of capacity, the results showed that by adding only one station (TRT) patient service satisfaction can be improved significantly.While this work focuses on nuclear medicine, we believe it will also find generality in other health care settings. Further research include stochastic (online) optimization algorithms that take into account data uncertainties such as stochastic patient arrivals, patient no shows, equipment failures, and delayed radiopharmaceutical deliveries.

CHAPTER V

STOCHASTIC ONLINE PATIENT AND RESOURCE SCHEDULING OF MULTI-STEP MEDICAL PROCEDURES IN NUCLEAR MEDICINE

A. Introduction

Nuclear medicine is a branch of medical imaging that uses small amounts of radioactive materials to diagnose and treat a variety of diseases, including many types of cancers, heart disease and certain other abnormalities within the body. Medical imaging has become a major factor in the total cost of U.S. health care [95]. According to an analysis sponsored by the Blue Cross and Blue Shield Association, diagnostic imaging technologies cost between $65 billion and $75 billion in 2000, more than twice the cost of cardiovascular technologies or in vitro diagnostics [96]. In order to obtain more accurate diagnoses physicians are requesting patients to undergo medical imaging procedures more often. Suthummanon et al. [4] showed in their study that machine time, direct labor time, and radiopharmaceuticals have the most influential in the cost per procedure in nuclear medicine. However, scheduling patients and resources in medical imaging clinics such as nuclear medicine departments remains a challenge. This can be attributed to the increase in demand for this service and to the complexity in the nuclear medicine procedure protocols. In this chapter we derive a stochastic online scheduling algorithm for improving patient and resource scheduling in nuclear medicine clinics. This scheduling algorithm considers both the patient's and manager's perspectives.

Nuclear medicine is divided in two major areas: diagnostic and therapeutic. Procedures in nuclear medicine require the administration of a radiopharmaceutical to the patient, involve several resources and are multi-step following a specific sequence.

Radiopharmaceuticals allow for the imaging (scans) or treatment of a specific organ of the human body. The short half-life of the radiopharmaceuticals imposes strict time constraints on scheduling patients and resources. To successfully complete a procedure every step has to be initiated and completed within a specific time window. If the procedure protocol is not followed a poor scan will result causing poor utilization of expensive resources and patient rescheduling. A scan could last from minutes to hours and a procedure may require multiple scans in a day or multiple days.

To perform a nuclear medicine procedure several resources are needed such as a technologist, a radiopharmaceutical, gamma camera, and sometimes a treadmill, a nurse or EKG technician. Radiopharmaceuticals are prepared in radio-pharmacies outside the clinics, therefore scheduling of their delivery, patient administration and image acquisition requires lead time and must be carefully managed. Gamma cameras are expensive, some of them may cost up to $1 million and thus have to be managed effectively. Resources required to serve a patient must be available at the time they are scheduled. Patients are re-scheduled if the procedure is not completed successfully. Therefore, scheduling of patients, resources, and radiopharmaceuticals is a very challenging problem for nuclear medicine departments. Consequently, finding ways to provide a high quality of service to the patient by using mathematical techniques is of great interest for nuclear medicine managers. The characteristics of this problem make it unique with very limited research reported in the literature.

Patient requests in nuclear medicine arrive during the day as the scheduling proceeds. The challenge is to schedule a sequence of appointment requests when not all of them are known to the scheduler in advance. In other words, when a patient request is received at the clinic, the receptionist has to provide an appointment to the patient without taking into consideration the requests that will be received in the future. This usually causes inefficiencies to the system such as lower utilization for some of the resources and longer waiting times for patients to get an appointment. This problem can be considered as an online scheduling problem. In online scheduling requests are not known in advance, rather they are revealed online during the day. Consequently, every time a request is received a decision is made without knowing the requests that will be arriving in the future.

Stochastic planning techniques are an alternative to address this problem. In this chapter an online stochastic scheduling algorithm for patient and resource management in nuclear medicine clinics is proposed. The idea is to select the best appointment date for each request received that will allow the clinic to maximize the cumulative value of the patients served over a long time horizon (maximize throughput). An online stochastic framework for scheduling patients and resources in nuclear medicine is developed. An algorithm is presented that uses a stochastic programming model to solve the offline problem and select the most appropriate appointment for a patient by taking into consideration future arrival requests.

The rest of the chapter is organized as follows: In Section B a description of the scheduling problem is provided and the stochastic algorithm is presented. In Section C the setting in which the algorithm was tested is described. A preliminary computational study is reported in Section D and a discussion of the results is given. The chapter ends with some concluding remarks and directions of future research in Section E.

B. Scheduling Problem

In this section we provide a description of the problem of scheduling patients and resources in nuclear medicine clinics from three different perspectives. First, a description is provided for the offline version of the problem in which it is assumed that all the patients requests that will be received at the clinic are known in advance. Secondly, we provide a description of the online scheduling problem. In this problem requests are revealed one at a time during the day and the appointment decision is made without taking into account future requests. Lastly, we discuss the stochastic online version of the problem where requests are also revealed during the day but possible future requests are taken into account when deciding on a patient appointment. The possible future requests are based on historical data.

1. The Offline Problem

The offline problem is defined for a day in which each resource r has a schedule that contains τ number of time-slots. Each schedule contains patients already scheduled and open spaces to serve new patient requests. In the offline version of the problem all patient requests are known in advance. A set J of patients asking for a procedure request J is used as input. Each request is characterized by the type of procedure p requested and a preferred day for the appointment (q). The goal is to find an assignment of a subset $B \subseteq J$ of patients asking for procedures satisfying the problem-specific constraints and maximizing the objective function. Therefore, we derive an integer programming (IP) model for multi-step medical procedure scheduling. The IP model allocates a subset B of the requests to resources schedules so that their capacities are not exceeded (only one patient can be served on each resource schedule time-slot) and the objective function is maximized. Two binary variables are associated to this problem, $x_{jt\ell}^{ik}$ and $w_{j\ell}^{ik}$. Variable $x_{jt\ell}^{ik}$=1 if patient j is scheduled to use resource i at time-slot t when procedure is started at time ℓ for the k step of the procedure, otherwise $x_{jt\ell}^{ik}$=0. Variable $w_{j\ell}^{ik}$=1 if resource i is selected to serve patient j in step k when procedure is started at time ℓ, otherwise $w_{j\ell}^{ik}$=0. Table XVII gives the notation and Figure 32 shows the IP model for the offline scheduling problem.

Table XVII. Scheduling problem sets and parameters

	Sets
J :	set of patients, indexed j.
I :	set of resources, indexed i.
S :	set of stations, indexed s.
R :	set of human resources, indexed r.
P :	set of nuclear medicine procedures, indexed p.
A :	set of radiopharmaceuticals, indexed a.
S_{kp} :	set of stations where step k of procedure p can be performed.
R_{kp} :	set of human resources qualified to perform step k of procedure p.
I_{kp} :	set of resources that can be used to perform step k of procedure p, I_{kp}=$\{R_{kp} \cup S_{kp}\}$.
I_{tj} :	set of resources that could be used in time-slot t to serve patient j.
A_{tj} :	set of radiopharmaceuticals that are required at time-slot t to serve patient j.
L_{itj} :	set of appointment star times that require the use of resource i at time-slot t for patient j.
K_{itj} :	set of procedure steps that require the use of resource i at time-slot t for an appointment for patient j.
T_{ij} :	set of time-slots where resource i could be used to serve patient j.
T_{aj} :	set of time-slots where radiopharmaceutical a could be used to serve patient j.
L_j :	set of possible start-times for patient j.
	Parameters
i :	subscript, for the i resource;
j :	subscript, for the j patient;
a :	subscript, for the a radiopharmaceutical;
p :	subscript, for the p procedure;
k :	subscript, for the k step of a procedure;
ℓ :	subscript, for the ℓ starting time-slot for a patient appointment;
t :	subscript, for the t time-slot, incremental time;
τ :	total number of time-slots in a day, indexed $t, \dots, \tau$;
s_{it} :	$s_{it} = 1$, if resource i is available at time-slot t; otherwise s_{it}=0;
s_{at} :	$s_{at} = 1$, if radiopharmaceutical a is available at time-slot t; otherwise s_{ia}=0;
d_{kp} :	number of time-slots required to complete step k of procedure p;
n_p :	total number of steps for procedure p, indexed $k, \dots, n_p$;
q :	day of the week requested by patient, indexed q=1,. . . ,5, where 1=Monday, 2=Tuesday, 3=Wednesday, 4= Thursday, 5=Friday.

$$Max : \sum_{j\in J} \sum_{i\in I_{1p}} \sum_{\ell\in L_j} w_{j\ell}^{i1} \tag{5.1a}$$

$$s.t. \sum_{k=1}^{n_p} \sum_{i\in I_{kp}} \sum_{t=\ell+\sum_{\hat{k}=0}^{k-1} d_{\hat{k}}-1}^{t=\ell+\sum_{\hat{k}=0}^{k} d_{\hat{k}}-1} x_{jt\ell}^{ik} \leq \sum_{k=1}^{n_p} 2\ d_{kp}, \quad j \in J,\ \ell \in L_j \tag{5.1b}$$

$$\sum_{j\in J} \sum_{k\in K_{itj}} \sum_{\ell\in L_{itj}} x_{jt\ell}^{ik} \leq s_{it}, \quad \forall i \in I_{tj},\ \forall t \in T_i \tag{5.1c}$$

$$\sum_{j\in J} \sum_{\ell\in L_{itj}} x_{jt\ell}^{ik} \leq s_{at}, \quad \forall a \in A_{tj},\ \forall i \in I_{tj},\ \forall t \in T_i,\ k = 1 \tag{5.1d}$$

$$\sum_{i\in R_{kp}} \sum_{\ell\in L_j} w_{j\ell}^{ik} \leq 1, \quad j \in J,\ k = 1, \ldots, n_p \tag{5.1e}$$

$$\sum_{i\in S_{kp}} \sum_{\ell\in L_j} w_{j\ell}^{ik} \leq 1, \quad j \in J,\ k = 1, \ldots, n_p \tag{5.1f}$$

$$x_{jt\ell}^{ik} - w_{j\ell}^{ik} = 0, \quad j \in J,\ i \in I_{kp}, \quad k = 1, \ldots, n_p, \quad \ell \in L_j, \quad t \in T_i \tag{5.1g}$$

$$\sum_{i\in R_{kp}} w_{j\ell}^{ik} - \sum_{i\in R_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad j \in J,\ \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.1h}$$

$$\sum_{i\in S_{kp}} w_{j\ell}^{ik} - \sum_{i\in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad j \in J,\ \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.1i}$$

$$\sum_{i\in R_{kp}} w_{j\ell}^{ik} - \sum_{i\in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad j \in J,\ \ell \in L_j, \quad k = 1, \tag{5.1j}$$

$$x_{jt\ell}^{ik},\ w_{j\ell}^{ik} \in \{0, 1\} \tag{5.1k}$$

Fig. 32. IP model for the offline problem

The objective function (Equation 5.1a) looks to maximize the number of patients that can be scheduled to be served at the clinic in a day. Equation 5.1b is the constraint used to verify that the number of resources assigned to the patient do not exceed those needed to perform the procedure requested. Equation 5.1c gives for each resource the time-slot by time-slot resource requirements that must be less than or equal to the resource availability at each time period. Similar constraints are used for radiopharmaceuticals (Equation 5.1d). Equation 5.1e and Equation 5.1f are used to select the staff and station per procedure step respectively, and also to decide the start-time of the appointment for each patient. Equation 5.1g is used to assure that the same resource is scheduled for the duration of a particular procedure step. Equation 5.1h and Equation 5.1i are used to verify that the staff and stations, respectively, selected to serve a patient follow the procedure sequence protocol. Equation 5.1j is used to match a station to a staff member for each step of the procedure requested by the patient. Finally, Equation 5.1k set values of the variables as binaries.

2. The Online Problem

In the online problem all the requests are not known in advance, rather they are revealed online (one at the time) during the day and they are scheduled when they are received. For example, if a sequence of requests $\xi = \langle \xi_1, \ldots, \xi_{t-1}, \xi_t \rangle$ is revealed at different times of the day, the requests $\xi_1, \ldots, \xi_{t-1}$ are already scheduled at time t when request ξ_t is received. At time t the problem is to decide how to schedule request ξ_t by keeping all the other patients already scheduled fixed.

To solve this problem an online framework is proposed. This framework has at its disposal a mathematical model to solve the problem. In this case, an adaptation of the IP model in Figure 32 presented in Figure 33 is used to solve the problem. The new mathematical model only schedules one patient at a time by keeping the patient

$$Max : \sum_{i \in I_{1p}} \sum_{\ell \in L_j} w_{j\ell}^{i1} \tag{5.2a}$$

$$s.t. \sum_{k=1}^{n_p} \sum_{i \in I_{kp}} \sum_{t=\ell+\sum_{\hat{k}=0}^{k-1} d_{\hat{k}}-1}^{t=\ell+\sum_{\hat{k}=0}^{k} d_{\hat{k}}-1} x_{jt\ell}^{ik} \leq \sum_{k=1}^{n_p} 2\ d_{kp}, \quad \ell \in L_j \tag{5.2b}$$

$$\sum_{k \in K_{itj}} \sum_{\ell \in L_{itj}} x_{jt\ell}^{ik} \leq s_{it}, \quad \forall i \in I_{tj},\ \forall t \in T_i \tag{5.2c}$$

$$\sum_{\ell \in L_{itj}} x_{jt\ell}^{ik} \leq s_{at}, \quad \forall a \in A_{tj},\ \forall i \in I_{tj},\ \forall t \in T_i,\ k = 1 \tag{5.2d}$$

$$\sum_{i \in R_{kp}} \sum_{\ell \in L_j} w_{j\ell}^{ik} \leq 1, \quad k = 1, \ldots, n_p \tag{5.2e}$$

$$\sum_{i \in S_{kp}} \sum_{\ell \in L_j} w_{j\ell}^{ik} \leq 1, \quad k = 1, \ldots, n_p \tag{5.2f}$$

$$x_{jt\ell}^{ik} - w_{j\ell}^{ik} = 0, \quad i \in I_{kp}, \quad k = 1, \ldots, n_p, \quad \ell \in L_j, \quad t \in T_i \tag{5.2g}$$

$$\sum_{i \in R_{kp}} w_{j\ell}^{ik} - \sum_{i \in R_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.2h}$$

$$\sum_{i \in S_{kp}} w_{j\ell}^{ik} - \sum_{i \in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.2i}$$

$$\sum_{i \in R_{kp}} w_{j\ell}^{ik} - \sum_{i \in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \tag{5.2j}$$

$$x_{jt\ell}^{ik},\ w_{j\ell}^{ik} \in \{0, 1\} \tag{5.2k}$$

Fig. 33. IP model for the online problem

schedules of previous requests fixed.

A general framework for the online problem named `Online_Schema()` is depicted in Figure 34. The algorithms presented in this chapter share the same online schema but differ in the way they implement the function `ServeRequest()`. `Online_Schema()` has a time horizon defined in days. The parameter $\overline{h}$ is used to define the total number of days in the scheduling horizon and the parameter b denotes a particular day of the horizon. The set G_j is used to save the appointment schedule found using the `ServeRequest()` function. An additional set Γ is used to save all the appointments scheduled in the system. Also, in what follows $\leftarrow$ is used to denote an assignment.

`Online_Schema`

1 $J \leftarrow \{\emptyset\},\ \ \Gamma \leftarrow \{\emptyset\},\ \ j = 0;$

2 **while** $b \leq \overline{h}$

3   **while** $t \leq \tau$

4         $(p_j, q) \leftarrow$ `GetPatientRequest(j)`;

5   **do** $J \cup \{p_j\},\ \ d_j \leftarrow d,\ \ t_j \leftarrow t,\ \ j \leftarrow j+1;$

6             $G_j \leftarrow$ `ServeRequest` $(j, p_j, d_j, t_j, q, \Gamma);$

7             $\Gamma \cup G_j;$

Fig. 34. The generic online algorithm

The first step of `Online_Schema()` (line 1) initializes the patient set J and the value of j. Lines 2 and 3 define the time horizon in which patient requests will be accepted. The function `GetPatientRequest()` gets the required information from the patient when the request is received (line 4). The patient is added to the set J in line 5 and the `ServeRequest()` function is invoked in line 6. The `ServeRequest()`

function use the information provided to find an appointment for the patient. We define a `ServeRequest-OS`() function for the online problem in Figure 35. Line 1 is used to call function `OptSolution`(). This function uses the information provided by the `Online_Schema`() framework to construct the IP mathematical model (Figure 33) and obtain a schedule for patient j. The appointment found is returned in line 2.

`ServeRequest-OS`$(j, p_j, d_j, t_j, q, \Gamma)$

1 $G_j \leftarrow$ `OptSolution`$(j, p_j, d_j, t_j, q, \Gamma)$;

2 **return** G_j

Fig. 35. The `ServeRequest-OS`() function

3. The Stochastic Online Problem

The online stochastic problem is an extension of the online problem. Again, the problem is to decide how to serve a request ξ_t by keeping all previous requests already scheduled fixed. However, this problem also accounts for possible future requests based on historical data to make more informed decisions (i.e. how to serve current request).

For this problem, we propose a stochastic online framework that has at its disposal a two-stage stochastic programming model to solve the problem. The first stage of the model is presented in Figure 36 and decides when to schedule the current patient request and which resources to use. The first stage of the model is similar to the IP formulation presented in Section 2 but differs in the objective function. An additional coefficient that accounts for the expected value of the objective function in the second stage is included that will allow for recourse/corrective actions for the

$$Max: \sum_{i \in I_{1p}} \sum_{\ell \in L_j} w_{j\ell}^{i1} + E[Q(x, \tilde{\omega}] \tag{5.3a}$$

$$s.t. \sum_{k=1}^{n_p} \sum_{i \in I_{kp}} \sum_{t=\ell+\sum_{\hat{k}=0}^{k-1} d_{\hat{k}}-1}^{t=\ell+\sum_{\hat{k}=0}^{k} d_{\hat{k}}-1} x_{jt\ell}^{ik} \leq \sum_{k=1}^{n_p} 2\ d_{kp}, \quad \ell \in L_j \tag{5.3b}$$

$$\sum_{k \in K_{itj}} \sum_{\ell \in L_{itj}} x_{jt\ell}^{ik} \leq s_{it}, \quad \forall i \in I_{tj},\ \forall t \in T_i \tag{5.3c}$$

$$\sum_{\ell \in L_{itj}} x_{jt\ell}^{ik} \leq s_{at}, \quad \forall a \in A_{tj},\ \forall i \in I_{tj},\ \forall t \in T_i,\ k = 1 \tag{5.3d}$$

$$\sum_{i \in R_{kp}} \sum_{\ell \in L_j} w_{j\ell}^{ik} \leq 1, \quad k = 1, \ldots, n_p \tag{5.3e}$$

$$\sum_{i \in S_{kp}} \sum_{\ell \in L_j} w_{j\ell}^{ik} \leq 1, \quad k = 1, \ldots, n_p \tag{5.3f}$$

$$x_{jt\ell}^{ik} - w_{j\ell}^{ik} = 0, \quad i \in I_{kp}, \quad k = 1, \ldots, n_p, \quad \ell \in L_j, \quad t \in T_i \tag{5.3g}$$

$$\sum_{i \in R_{kp}} w_{j\ell}^{ik} - \sum_{i \in R_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.3h}$$

$$\sum_{i \in S_{kp}} w_{j\ell}^{ik} - \sum_{i \in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \ldots, n_p, \tag{5.3i}$$

$$\sum_{i \in R_{kp}} w_{j\ell}^{ik} - \sum_{i \in S_{(k-1)p}} w_{j\ell}^{i(k-1)} = 0, \quad \ell \in L_j, \quad k = 1, \tag{5.3j}$$

$$x_{jt\ell}^{ik},\ w_{j\ell}^{ik} \in \{0, 1\} \tag{5.3k}$$

Fig. 36. SIP model for the online stochastic problem, First Stage

patient request the model is trying to schedule.

The second stage of the problem is presented in Figure 37 and represents a possible "scenario". A scenario in this problem is defined as sequence of future patient requests arriving at the clinic after the current request that needs to be scheduled. To generate a scenario ω for the second stage of the problem, a function `GetSample()` is defined. This function returns a set of requests over a time interval that starts at time t (time in which current request was received) and it is stored in set J^{ω}. Two new binary variables are defined for the second stage, $y^{ik\omega}_{j't\ell}$ and $z^{ik\omega}_{j'\ell}$. Variable $y^{ik\omega}_{j't\ell}$=1 if patient j' is scheduled to use resource i at time-slot t and when the procedure is started at time ℓ for the k step of the procedure. Otherwise $y^{ik\omega}_{j't\ell}$=0. Variable $z^{ik\omega}_{j\ell}$=1 if resource i is selected to serve patient j in step k and when the procedure is started at time ℓ. Otherwise $z^{ik\omega}_{j'\ell}$=0. The second-stage problem is similar to the offline problem presented in Section 1. This problem has a modification in Equation 5.4c. The solution (patient schedule) obtained on the first-stage of the problem is now considered a parameter in this constraint.

4. Stochastic Online Scheduling Algorithm

The online stochastic algorithm for this problem follows the same framework presented in Figure 34) but differ in the way the `ServeRequest()` function is implemented. A description for the `ServeRequest-SOS()` for the online stochastic problem is presented in Figure 38. Line 2 is used to establish the number of scenarios that are going to be used to define the problem. Again, a scenario is a set of requests over a time interval that starts at time t. The `GetSample()` function is used to generate the requests from a probability distribution (line 3). The requests generated for scenario η are stored in the set U_η and each scenario set is added to the set U. Once all the scenarios are generated, the information is passed to the `OptSolution()` which generates the for each outcome $\omega \in \Omega$ of $\tilde{\omega}$, $Q(x,\omega) =$

$$Max : \sum_{j' \in J^\omega} \sum_{i \in I_{1p}} \sum_{\ell \in L_{j'}} z_{j'\ell}^{i1\omega} \tag{5.4a}$$

$$s.t. \sum_{k=1}^{n_p} \sum_{i \in I_{j'k}} \sum_{t=\ell+\sum_{\hat{k}=0}^{k-1} d_{\hat{k}}-1}^{t=\ell+\sum_{\hat{k}=0}^{k} d_{\hat{k}}-1} y_{j't\ell}^{ik\omega} \leq \sum_{k=1}^{n_p} 2\ d_{kp}, \quad \ell \in L_{j'},\ \forall j' \in J^\omega \tag{5.4b}$$

$$\sum_{j' \in J^\omega} \sum_{k \in K_{itj'}} \sum_{\ell \in L_{itj'}} y_{j't\ell}^{ik\omega} + x_{jt\ell}^{ik\omega} \leq s_{it}, \quad \forall i \in I_{tj'},\ \forall t \in T_i \tag{5.4c}$$

$$\sum_{j' \in J^\omega} \sum_{\ell \in L_{itj'}} y_{j't\ell}^{ik\omega} \leq s_{at}, \quad \forall a \in A_{tj'},\ \forall i \in I_{tj'},\ \forall t \in T_i,\ k = 1 \tag{5.4d}$$

$$\sum_{i \in R_{kp}} \sum_{\ell \in L_{j'}} z_{j'\ell\omega}^{ik} \leq 1, \quad j' \in J^\omega,\ k = 1, \ldots, n_p \tag{5.4e}$$

$$\sum_{i \in S_{kp}} \sum_{\ell \in L_j} z_{j'\ell}^{ik} \leq 1, \quad j' \in J^\omega,\ k = 1, \ldots, n_p \tag{5.4f}$$

$$y_{j't\ell}^{ik\omega} - z_{j'\ell}^{ik\omega} = 0, \quad j' \in J^\omega,\ i \in I_{kp},\ \ell \in L_{j'},\ t \in T_i,\ k = 1, \ldots, n_p, \tag{5.4g}$$

$$\sum_{i \in R_{kp}} z_{j'\ell}^{ik\omega} - \sum_{i \in R_{(k-1)p}} z_{j'\ell}^{i(k-1)\omega} = 0, \quad j' \in J^\omega,\ \ell \in L_{j'},\ k = 1, \ldots, n_p, \tag{5.4h}$$

$$\sum_{i \in S_{kp}} z_{j'\ell}^{ik\omega} - \sum_{i \in S_{(k-1)p}} z_{j'\ell}^{i(k-1)\omega} = 0, \quad j' \in J^\omega,\ \ell \in L_{j'},\quad k = 1, \ldots, n_p, \tag{5.4i}$$

$$\sum_{i \in R_{kp}} z_{j'\ell}^{ik\omega} - \sum_{i \in S_{(k-1)p}} z_{j'\ell}^{i(k-1)\omega} = 0, \quad j' \in J^\omega,\ \ell \in L_{j'},\quad k = 1, \tag{5.4j}$$

$$y_{j't\ell}^{ik\omega},\ z_{j'\ell}^{ik\omega} \in \{0, 1\} \tag{5.4k}$$

Fig. 37. SIP model for the online stochastic problem, Second Stage

two-stage stochastic model and finds a solution or a schedule for the patient. This solution is stored in set G_j and returned to the overall framework.

`ServeRequest-SOS`$(j, p_j, d_j, t_j, q, \Gamma)$

$U \leftarrow \{\emptyset\}$
for η= 1 to numberScenario **do**
  $U_\eta \leftarrow$ `GetSample`(η)
  $U \cup U_\eta$
$G_j \leftarrow$ `OptSolution`$(j, p_j, d_j, t_j, q, \Gamma, U)$;
  return G_j

Fig. 38. The `ServeRequest`() function

C. Application

In this section we consider a small instance of the problem of managing patients and resources in nuclear medicine clinics. The goal of this example is provide a proof of concept of the benefits of using stochastic online scheduling when managing patients and resources in nuclear medicine.

Consider a nuclear medicine department with four human resources and four stations. The members of the human resource group are named as follows: Technologist, one EKG technologist, Nurse, and Manager. The stations are classified according to the equipment they contain. The clinic stations are named as follows: TRT, Treadmill, Axis, and P2000. For the purpose of the example we limit the number of procedures that can be requested to three. These procedures have been identified as the most requested by the Scott and White Health System in Texas and

Table XVIII. Highly requested nuclear medicine procedures

CPT Code	Name
78465	Cardiovascular Event (CVE) Myocardial Imaging (SP-M)
78815	PET CT skull to thigh
78306	MSB-bone imaging (whole body)

they are listed in Table XVIII. A Poisson process is assumed for procedure request arrivals with mean call interarrivals times of 6 minutes based on historical information provided by a real clinic. Empirical distributions were used to generate a procedure request and an appointment preferred day for the patient.

We conducted preliminary experiments to get insights into the impact of the stochastic online scheduling in patient and resource scheduling in nuclear medicine clinics. The results are compared with those obtained when solving the offline problem (all requests are known in advance) and with an online implementation of the *No Fixed Resource* (NFR) scheduling algorithm proposed in Chapter IV. The NFR algorithm schedules patients and resources based on the patient's preferred day of the week. However, if the patients wait for the appointment is more than one month, an earlier appointment is considered. The experiments were conducted using a time horizon of a month.

The stochastic online algorithm was implemented in C++ and solved using CPLEX. Two scenarios of future arrivals were generated every time a request was received. Each scenario considered the requests that may be received for a time period of a week. This information was used to construct the two-stage stochastic programming model and to find an appointment for the patient request. All the experiments were conducted on a DELL Optiplex GX 620 with a Pentium D processor running at 3.2GHz with 3.0GB RAM.

Table XIX. Average number of patients served using SOS algorithm

	NFR	SOS	Offline
Patients served	463.00	521.00	579.00
Patients served / day	23.15	26.05	28.95

D. Preliminary Results

The results for the number of patients served during the month are reported in Table XIX. We also report the number of patients served per day. The results show that the stochastic online scheduling algorithm (SOS) performs better than the NFR algorithm in terms of throughput. There is an increase of 12.53% in the number of patients served. The maximum number of patients that can be served for this problem is 579 which is obtained by solving the offline problem.

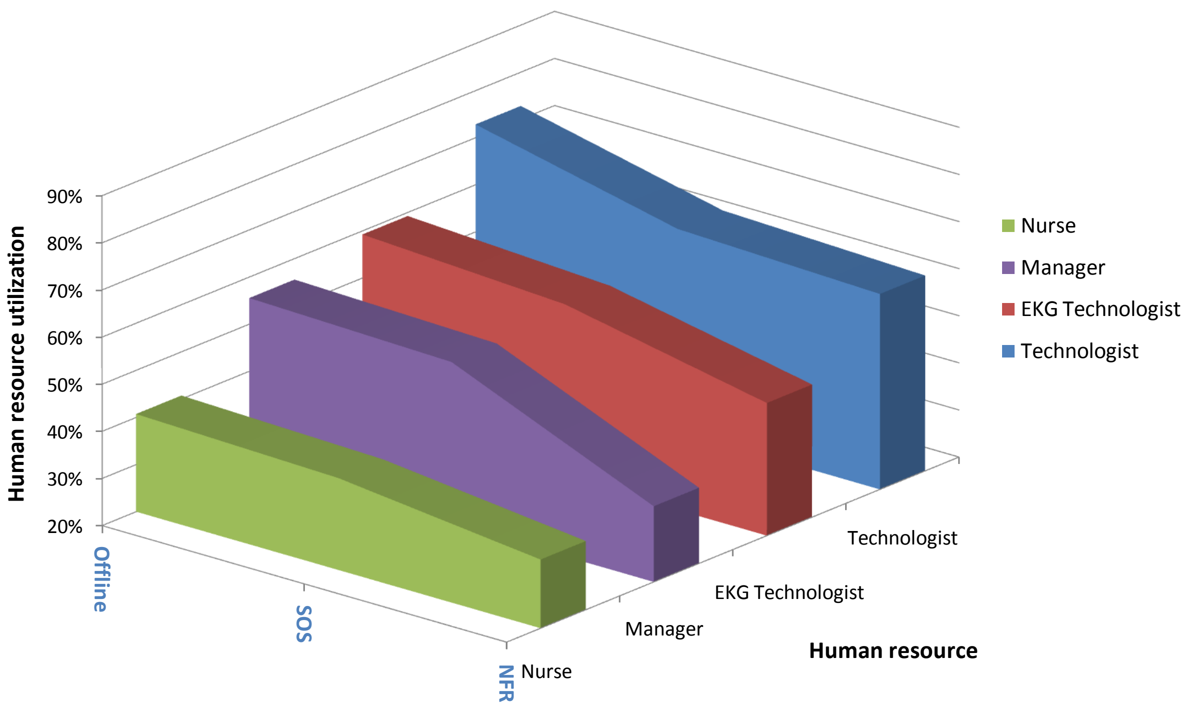

Fig. 39. Human resource utilization using SOS algorithm

We also report the utilization for the human resources. The results are plotted in Figure 39. The technologist is the human resource with highest utilization under both NFR and SOS algorithms. The graph also shows that under the SOS algorithm a higher utilization of all the human resources is obtained when compared to the NFR algorithm. This can be attributed to the higher number patients served under the SOS scheduling algorithm.

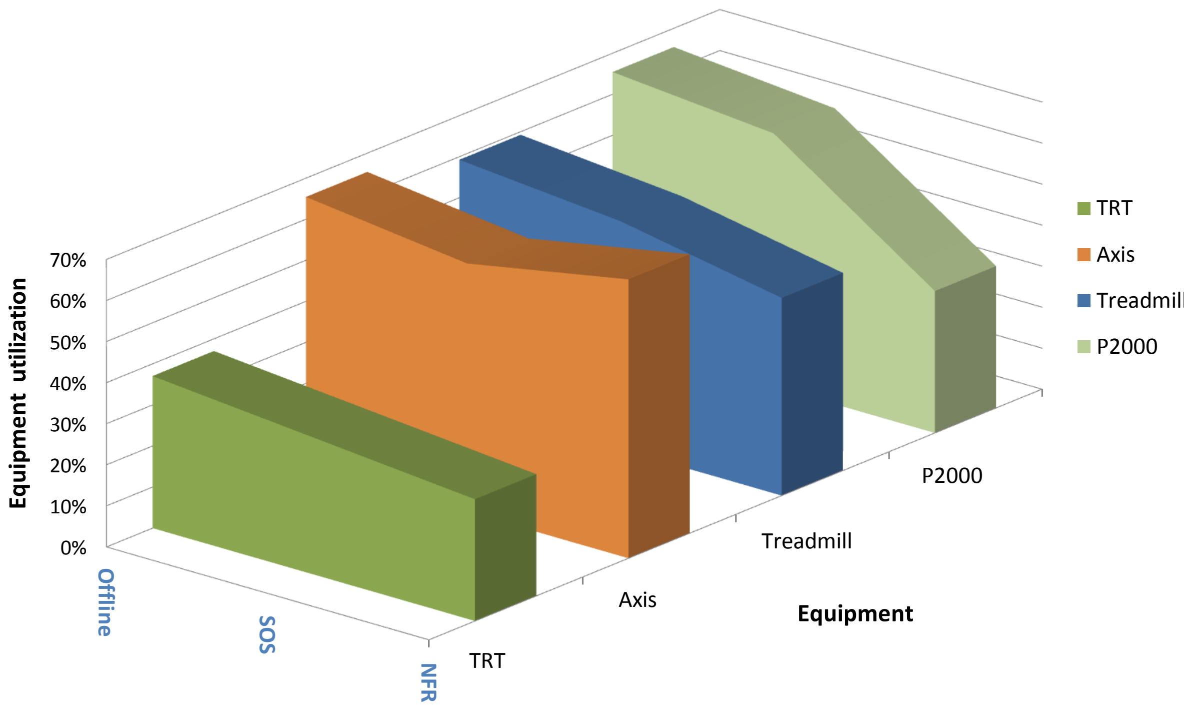

Fig. 40. Equipment (station) utilization using SOS algorithm

Figure 40 depicts the utilization of the stations (equipment). The plot shows that the station with highest utilization for both the NFR and SOS scheduling algorithm is the Axis station. Since the SOS algorithm was able to accommodate more patients into the system the utilization of the remaining stations; TRT, Treadmill, and P2000 was higher when compared to the NFR algorithm.

E. Summary

Appointment scheduling in specialized clinics such as nuclear medicine departments is a very challenging problem. Radiopharmaceutical properties require procedures to be performed following strict protocols that must be adhered to by the staff. In this chapter we derive a stochastic online scheduling algorithm for patient and resource scheduling in nuclear medicine departments. This algorithm has at its disposal a two-stage stochastic programming model which is used for solving the problem of scheduling patients and resources in nuclear medicine. We obtain preliminary computational results that provide proof of the concept for the potential of considering stochastic information when scheduling patients in health care clinics. For example, the number of patients served for a month was significatively larger under the SOS scheduling algorithm when compared to the NFR scheduling algorithm. However, the performance of the SOS algorithm can be improvement since the offline solution of the problem provides a significatively larger number of patients served. Therefore we believe that the SOS algorithm can be improved in such a way that the results obtained can be closer to the offline solution.

Even though the SOS algorithm show some potential in improving the solution of the problem there are some disadvantages attached to using the SOS algorithm framework versus other methods. For example, the number of scenarios required to obtain a "good" solution for a problem instance can be large. This can have some repercussion in terms of the time required to obtain a solution.

Even though this work focus in nuclear medicine we believe it can also benefit other health care settings or disciplines. Further research include the extension of the algorithm to account for other uncertainty sources such as resource unavailability. In addition, we will like to characterize the algorithms solutions with the objective of developing a heuristic that can find similar solutions with less computational effort.

CHAPTER VI

CONCLUSIONS AND FUTURE RESEARCH

This dissertation presents a study of patient and resource scheduling in specialty health care clinics such as nuclear medicine departments. This study responds to the fact that medical imaging clinics such as nuclear medicine clinics have become a factor in the total cost of U.S. health care. Equipment utilization, direct labor time, and radiopharmaceutical prices have been identified as the most influential factors in the cost per procedure in nuclear medicine and there is a need for improving the existing appointment systems to improve the way resources and patients are managed.

The strategies presented were designed to improve the management of patients and resources in nuclear medicine clinics by taking into consideration both patients and management perspectives. While this work focuses in nuclear medicine we believe the solution strategy presented will find generality in other health care settings or applications.

A. Conclusions

Given the rise in health care costs patient scheduling in medical clinics has been a subject of study during the past few years. This work proposes a methodology for improving patient service and resource management in highly constrained health care environments such as nuclear medicine. In our study we found that implementing some of the proposed ideas will result in a better patient service and in a better utilization of the resources in health care nuclear medicine clinics.

We follow a methodology toward achieving our objective. First, we present a simulation model that integrates with scheduling methods to manage patient service levels and resource productivity. We validated the simulation model based on a real nuclear medicine setting and report computational results based on a scheduling algorithm. The results show that fixing some of the human resources to specific stations may reduce the patient throughput of the system unless it is carefully determined (e.g. through simulation). In addition the results show that a balanced workload is achieved under those algorithms where no human resources are fixed to stations.

In addition, we derive scheduling algorithms for scheduling multi-steps medical procedures that are time constrained in health care specialty clinics. The scheduling algorithms take into consideration the time constraints imposed by the decay of the radiopharmaceuticals, which are required for most of the procedures. The algorithms were implemented within the simulation framework and the experiments were based on historical data. The results obtained provide good insights into patient and resource management in nuclear medicine clinics. No single algorithm provides the best results relative to all performance measures. Thus, the nuclear medicine clinic could decide which algorithm to implement based on priorities. The results show that reserving stations to be exclusively used to serve those procedures that are requested more often improves throughput. However, this may also increase the average waiting time for the patients.

Finally, we developed a stochastic online scheduling algorithm for managing patients and resources by tanking into consideration stochastic information about patient future requests. This algorithm has at its disposal a two-stage stochastic programming model that is used to solve the scheduling problem. The preliminary results obtained provide some insights to the benefits of considering stochastic information when scheduling patients in health care clinics. We applied the stochastic online scheduling (SOS) algorithm to a small problem instance derived from a real clinic. The *No Fixed Resource* (NFR) algorithm was also implemented to this problem and both performance were compared. The results show that the number of patients served under the SOS was significantly higher.

B. Future Research

Scheduling in highly constrained environments is an open challenge problem in health care. This type of problem not only affect specialty clinics such as nuclear medicine departments but also other health care settings like outpatient surgery centers. The work presented in this dissertation represents a step forward in addressing this problem but additional questions that can be easily extended into future research directions.

First, specialty clinics are usually part of complex integrated systems (hospitals). In such cases, a macroscopic analysis of multi-facility systems may be needed to improve the performance of specialty clinics rather than focus on individual units. Discrete event simulation can be used to capture the interaction of major service departments and support services in the hospitals. At the facility level the discrete event simulation model presented in this work can be extended to a stochastic discrete event simulation model where additional stochastic information can be included to improve the model representation of reality.

Further research also include extensions of the scheduling algorithms proposed in this work. These algorithms can be extended to account for different patient behaviors such as late cancelations, patient no-shows or emergencies. It will be interested to investigate if a correlation exist between the amount of time patients have to wait for an appointment and patient no-shows.

In the future one can consider characterizing the solutions provided by the stochastic online scheduling algorithm with the objective of developing a heuristic algorithm that can find similar solutions with less computational effort. Even though this work focus in managing patients and resources in nuclear medicine, we believe it can also be applied to other health care settings or disciplines.

In addition, the simulation model presented in this work can be extended to a simulation optimization setting. We can envision a framework in which the simulation model provides feedback to the scheduling algorithm with the objective of making optimal decisions based on patient and the nuclear medicine management perspectives. Finally, there is a need of quantifying the economic impact of enhancing alternatives. For example, the impact of purchasing additional equipment that is considered critical for the clinic operation and/or the cost of cross training staff.

REFERENCES

[1] F.A. Mettler, M. Bhargavan, B.R. Thomadsen, D.B. Gilley, J.A. Lipoti, M. Mahesh, J. McCrohan, and T.T. Yoshizumi, "Nuclear Medicine Exposure in the United States, 2005-2007: Preliminary Results," in *Seminars in Nuclear Medicine*, May 2008, pp. 384–391.

[2] W.A. Richland, "Future of Nuclear Medicine, Part 1: Marketing Research Forecasts," *Journal of Nuclear Medicine*, vol. 39, no. 2, pp. 27–30, 1998.

[3] J. Zhang, "Growth in U.S. Health-care Spending Slows Again," *The Wall Street Journal*, p. A2, January 9, 2007.

[4] S. Suthummanon, V.K. Omachonu, and M. Akcin, "Applying Activity-based Costing to the Nuclear Medicine Unit," *Health Services Management Research*, vol. 18, no. 3, pp. 141–150, 2005.

[5] D.D. Patton, "Cost-effectiveness in Nuclear Medicine," in *Seminars in Nuclear Medicine*, May 1993, pp. 9–30.

[6] S.H. Jacobson, S.N. Hall, and J.R. Swisher, "Discrete-event Simulation of Health Care Systems," *International Series in Operations Research and Management Science*, vol. 91, pp. 211–252, 2006.

[7] A.M. Law and W.D. Kelton, *Simulation Modeling and Analysis*, New York : McGraw-Hill Higher Education, 1997.

[8] B.P. Zeigler, T.G. Kim, and H. Praehofer, *Theory of Modeling and Simulation*, Orlando, FL : Academic Press, 2000.

[9] J. Karnon, “Alternative Decision Modelling Techniques for the Evaluation of Health Care Technologies: Markov Processes Versus Discrete Event Simulation,” *Health Economics*, vol. 12, no. 10, pp. 837–848, 2003.

[10] L. Moreno, R.M. Aguilar, C.A. Martin, J.D. Pineiro, J.I. Estevez, J.F. Sigut, and J.L. Sanchez, “Patient-centered Simulation to Aid Decision-making in Hospital Management,” *Simulation*, vol. 74, no. 5, pp. 290–304, 2000.

[11] C.J. Ho and H.S. Lau, “Minimizing Total Cost in Scheduling Outpatient Appointments,” *Management Science*, vol. 38, no. 12, pp. 1750–1764, 1992.

[12] K.J. Klassen and T.R. Rohleder, “Scheduling Outpatient Appointments in a Dynamic Environment,” *Journal of Operations Management*, vol. 14, no. 2, pp. 83–101, 1996.

[13] R.H. Edwards, J.E. Clague, J. Barlow, M. Clarke, P.G. Reed, and R. Rada, “Operations Research Survey and Computer Simulation of Waiting Times in Two Medical Outpatient Clinic Structures,” *Health Care Analysis*, vol. 2, pp. 164–169, 1994.

[14] F. Hashimoto and S. Bell, “Improving Outpatient Clinic Staffing and Scheduling with Computer Simulation,” *Journal of General Internal Medicine*, vol. 11, no. 3, pp. 182–184, 1996.

[15] J.R. Swisher, S.H. Jacobson, B. Jun, and O. Balci, “Simulation of the Queston Physician Network,” in *Proceedings of the 1997 Winter Simulation Conference*, Dec. 1997, pp. 1146–1154.

[16] S. Groothuis, G. Van Merode, and A. Hasman, “Simulation as Decision Tool for Capacity Planning,” *Computer Methods and Programs in Biomedicine*, vol. 66,

pp. 139–151, 2001.

[17] J.R. Swisher and S.H. Jacobson, "Evaluating the Design of a Family Practice Health Care Clinic Using Discrete-event Simulation," *Health Care Management Science*, vol. 5, no. 2, pp. 75–88, 2002.

[18] T.R. Rohleder and K.J. Klassen, "Rolling Horizon Appointment Scheduling: A Simulation Study," *Health Care Management Science*, vol. 5, no. 3, pp. 201–209, 2002.

[19] M. Guo, M. Wagner, and C. West, "Outpatient Clinic Scheduling: A Simulation Approach," in *Proceedings of the 2004 Winter Simulation Conference*, Dec. 2004, pp. 1981–1987.

[20] K.E. Fitzpatrick, J.R. Baker, and D.S. Dave, "An Application of Computer Simulation to Improve Scheduling of Hospital Operating Room Facilities in the United States," *International Journal of Computer Applications in Technology*, vol. 6, pp. 215–215, 1993.

[21] R. Akkerman and M. Knip, "Reallocation of Beds to Reduce Waiting Time for Cardiac Surgery," *Health Care Management Science*, vol. 7, no. 2, pp. 119–126, 2004.

[22] T.W. Butler, G.R. Reeves, K.R. Karwan, and J.R. Sweigart, "Assessing the Impact of Patient Care Policies Using Simulation Analysis," *Journal of the Society for Health Systems*, vol. 3, no. 3, pp. 38–53, 1992.

[23] B.G. Lopez-Valcarcel and B. Perez, "Evaluation of Alternative Functional Designs in an Emergency Department by Means of Simulation," *Simulation*, vol. 63, no. 1, pp. 20, 1994.

[24] I.K. Altinel and E. Ulas, "Simulation Modeling for Emergency Bed Requirement Planning," *Annals of Operations Research*, vol. 67, no. 1, pp. 183–210, 1996.

[25] E. El-Darzi, C. Vasilakis, T. Chaussalet, and P.H. Millard, "A Simulation Modelling Approach to Evaluating Length of Stay, Occupancy, Emptiness and Bed Blocking in a Hospital Geriatric Department," *Health Care Management Science*, vol. 1, no. 2, pp. 143–149, 1998.

[26] M.L. Garcia, M.A. Centeno, C. Rivera, and N. DeCario, "Reducing Time in an Emergency Room via a Fast-track ," in *Proceedings of the 1995 Winter Simulation Conference*, Dec. 1995, pp. 1048–1053.

[27] S. Mahapatra, C.P. Koelling, L. Patvivatsiri, B. Fraticelli, D. Eitel, and L. Grove, "Pairing Emergency Severity Index5-level Triage Data with Computer Aided System Design to Improve Emergency Department Access and Throughput," in *Proceedings of the 2003 Winter Simulation Conference*, Dec. 2003, pp. 1917–1925.

[28] A. Kirtland, J. Lockwood, K. Poisker, L. Stamp, and P. Wolfe, "Simulating an Emergency Department "Is As Much Fun As..."," in *Proceedings of the 1995 Winter Simulation Conference*, Dec. 1995, pp. 1039–1042.

[29] F. McGuire, "Using Simulation to Reduce Length of Stay in Emergency Departments," *Journal of the Society for Health Systems*, vol. 5, no. 3, pp. 81–90, 1997.

[30] M.A. Badri and J. Hollingsworth, "A Simulation Model for Scheduling in the Emergency Room," *International Journal of Operations and Production Management*, vol. 13, pp. 13–24, 1993.

[31] S. Samaha, W.S. Armel, and D.W. Starks, "The Use of Simulation to Reduce the Length of Stay in an Emergency Department," in *Proceedings of the 2003 Winter Simulation Conference*, Dec. 2003, pp. 1907–1911.

[32] T. Eldabi, R.J. Paul, and T. Young, "Simulation Modelling in Healthcare: Reviewing Legacies and Investigating Futures," *Journal of the Operational Research Society*, vol. 58, no. 2, pp. 262–270, 2007.

[33] J.B. Jun, S.H. Jacobson, and J.R. Swisher, "Application of Discrete-event Simulation in Health Care Clinics: A Survey," *Journal of the Operational Research Society*, vol. 50, no. 2, pp. 109–123, 1999.

[34] R.W. Klein, R.S. Dittus, S.D. Roberts, and J.R. Wilson, "Simulation Modeling and Health-care Decision Making," *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, vol. 13, no. 4, pp. 347–354, 1993.

[35] A.R. Mahachek, "An Introduction to Patient Flow Simulation for Health-care Managers," *Journal of the Society for Health Systems*, vol. 3, no. 3, pp. 73–81, 1992.

[36] S.D. Walter, "A Comparison of Appointment Schedules in a Hospital Radiology Department," *British Journal of Preventive and Social Medicine*, vol. 27, no. 3, pp. 160–167, 1973.

[37] J.W. Kho and G.M. Johnson, "Computer Simulation of Hospital Health Care Delivery System," in *Proceedings of the 1976 Winter Simulation Conference*, Dec. 1976, pp. 12–20.

[38] J.D. Johannes and R.M. Wyskida, "A Nuclear Medicine Patient/Instrument

Scheduling Model," *The International Journal of Management Science*, vol. 6, no. 6, pp. 523–530, 1978.

[39] J.A. Sepúlveda, W.J. Thompson, F.F. Baesler, M.I. Alvarez, and L.E. Cahoon III, "The Use of Simulation for Process Improvement in a Cancer Treatment Center," in *Proceedings of the 1999 Winter Simulation Conference*, Dec. 1999, pp. 1541–1548.

[40] M.A. Centeno, C. Albacete, D.O. Terzano, M. Carrillo, and T. Ogazon, "A Simulation Study of the Radiology Department at JMH," in *Proceedings of the 2000 Winter Simulation Conference*, Dec. 2000, pp. 1978–1984.

[41] S. Ramakrishnan, K. Nagarkar, M. DeGennaro, M. Srihari, A.K. Courtney, and F. Emick, "A Study of the CT Scan Area of a Healthcare Provider," in *Proceedings of the 2004 Winter Simulation Conference*, Dec. 2004, pp. 2025–2031.

[42] P.C. O'Kane, "A Simulation Model of a Diagnostic Radiology Department," *European Journal of Operational Research*, vol. 6, no. 1, pp. 38–45, 1981.

[43] K.A. Klafehn, "Impact Points in Patient Flows through a Radiology Department Provided through Simulation," in *Proceedings of the 1987 Winter Simulation Conference*, Dec. 1987, pp. 914–918.

[44] C.J. Liao, CD Pegden, and M. Rosenshine, "Planning Timely Arrivals to a Stochastic Production or Service System," *IIE Transactions*, vol. 25, no. 5, pp. 63–73, 1993.

[45] P.P. Wang, "Static and Dynamic Scheduling of Customer Arrivals to a Single-server System," *Naval Research Logistics*, vol. 40, no. 3, pp. 345–360, 1993.

[46] P.P. Wang, "Optimally Scheduling N Customer Arrival Times for a Single-server System," *Computers and Operations Research*, vol. 24, no. 8, pp. 703–716, 1997.

[47] L. Liu and X. Liu, "Dynamic and Static Job Allocation for Multi-server Systems," *IIE Transactions*, vol. 30, no. 9, pp. 845–854, 1998.

[48] N.S. Penneys, "A Comparison of Hourly Block Appointments with Sequential Patient Scheduling in a Dermatology Practice," *Journal of the American Academy of Dermatology*, vol. 43, no. 5, pp. 809–813, 2000.

[49] M.K. Chung, "Tuning Up Your Patient Schedule," *Family Practice Management*, vol. 9, no. 1, pp. 41–45, 2002.

[50] H.S. Lau and A.H.L. Lau, "A Fast Procedure for Computing the Total System Cost of an Appointment Schedule for Medical and Kindred Facilities," *IIE Transactions*, vol. 32, no. 9, pp. 833–839, 2000.

[51] T. Cayirli, E. Veral, and H. Rosen, "Designing Appointment Scheduling Systems for Ambulatory Care Services," *Health Care Management Science*, vol. 9, no. 1, pp. 47–58, 2006.

[52] L.W. Robinson and R.R. Chen, "Scheduling Doctors' Appointments: Optimal and Empirically-based Heuristic Policies," *IIE Transactions*, vol. 35, no. 1, pp. 295–307, 2003.

[53] B. Denton and D.B. Gupta, "A Sequential Bounding Approach for Optimal Appointment Scheduling," *IIE Transactions*, vol. 35, no. 11, pp. 1003–1016, 2003.

[54] M. Brahimi and D.J. Worthington, " Queueing Models for Out-Patient Appointment Systems–A Case Study," *Journal of the Operational Research Society*, vol.

42, no. 9, pp. 733–746, 1991.

[55] Y. Gerchak, D. Gupta, and M. Henig, "Reservation Planning for Elective Surgery Under Uncertain Demand for Emergency Surgery," *Management Science*, vol. 42, no. 3, pp. 321–334, 1996.

[56] G.C. Kaandorp and G. Koole, "Optimal Outpatient Appointment Scheduling ," *Health Care Management Science*, vol. 10, no. 3, pp. 217–229, 2007.

[57] D. Gupta and L. Wang, "Revenue Management for a Primary-Care Clinic in the Presence of Patient Choice," *Operations Research*, vol. 56, no. 3, pp. 576–592, 2008.

[58] T. Cayirli and E. Veral, "Outpatient Scheduling in Health Care: A Review of Literature," *Production and Operations Management*, vol. 12, no. 4, pp. 519–549, 2003.

[59] D. Gupta and B. Denton, "Appointment Scheduling in Health Care: Challenges and Opportunities," *IIE Transactions*, vol. 40, no. 9, pp. 800–819, 2008.

[60] D. Conforti, F. Guerriero, and R. Guido, "Optimization Models for Radiotherapy Patient Scheduling," *4OR: A Quarterly Journal of Operations Research*, vol. 6, no. 3, pp. 263–278, 2007.

[61] L.V. Green, S. Savin, and B. Wang, "Managing Patient Service in a Diagnostic Medical Facility," *Operations Research*, vol. 54, no. 1, pp. 11–25, 2006.

[62] J. Patrick, M.L. Puterman, and M. Queyranne, "Dynamic Multi-priority Patient Scheduling for a Diagnostic Resource," *Operations Research*, vol. 56, no. 6, pp. 1507–1525, 2008.

[63] R. Kolisch and S. Sickinger, "Providing Radiology Health Care Services to Stochastic Demand of Different Customer Classes," *OR Spectrum*, vol. 30, no. 2, pp. 375–395, 2008.

[64] J. Patrick and M.L. Puterman, "Improving Resource Utilization for Diagnostic Services through Flexible Inpatient Scheduling: A Method for Improving Resource Utilization," *Journal of the Operational Research Society*, vol. 58, no. 2, pp. 235–245, 2007.

[65] S. Sickinger and R. Kolisch, "The Performance of a Generalized Bailey-Welch Rule for Outpatient Appointment Scheduling under Inpatient and Emergency Demand," *Health Care Management Science*, vol. 12, pp. 408–419, 2009.

[66] C.R. Standridge and D. Steward, "Using Expert Systems for Simulation Modeling of Patient Scheduling," *Simulation*, vol. 75, pp. 148–156, 2000.

[67] I.B. Vermeulen, S.M. Bohte, S.G. Elkhuizen, H. Lameris, P.J.M. Bakker, and H. La Poutré, "Adaptive Resource Allocation for Efficient Patient Scheduling," *Arificial Intelligence in Medicine*, vol. 46, no. 1, pp. 67–80, 2009.

[68] R. Bent and P. Van Hentenryck, "The Value of Consensus in Online Stochastic Scheduling," in *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2004, pp. 219–226.

[69] R. Bent and P. Van Hentenryck, "Regrets Only! Online Stochastic Optimization under Time Constraints," in *International Conference on Automated Planning and Scheduling (ICAPS)*, June 2004, pp. 501–506.

[70] P.A.N. Bosman and H. La Poutre, "Online Transportation and Logistics Using Computationally Intelligent Anticipation," *Advances in Computational Intel-*

ligence in Transport, Logistics, and Supply Chain Management, vol. 144, pp. 185–208, 2008.

[71] R.W. Bent and P. Van Hentenryck, "Scenario-based Planning for Partially Dynamic Vehicle Routing with Stochastic Customers," *Operations Research*, vol. 52, no. 6, pp. 977–987, 2004.

[72] H. Chang, R. Givan, and E. Chong, "On-line Scheduling via Sampling," in *Artificial Intelligence Planning Systems (AIPS)*, June 2000, pp. 62–71.

[73] B. Hamidzadeh, L.Y. Kit, and D.J. Lilja, "Dynamic Task Scheduling using Online Optimization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 11, no. 11, pp. 1151–1163, 2000.

[74] P. Van Hentenryck, R. Bent, L. Mercier, and Y. Vergados, "Online Stochastic Reservation Systems," *Annals of Operations Research*, vol. 171, no. 1, pp. 101–126, 2009.

[75] C. Helmberg and S. Rohl, "A Case Study of Joint Online Truck Scheduling and Inventory Management for Multiple Warehouses," *Operations Research*, vol. 55, no. 4, pp. 733–752, 2007.

[76] P. Awasthi and T. Sandholm, "Online Stochastic Optimization in the Large: Application to Kidney Exchange," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, June 2009, pp. 405–411.

[77] D. Nikovski and M. Branch, "Marginalizing Out Future Passengers in Group Elevator Control," in *Uncertainity in Artificial Intelligence (UAI)*, June 2003, pp. 443–450.

[78] B. Dean, M. Goemans, and J. Vondrak, "Approximating the Stochastic Knapsack Problem: The Benefit of Adaptivity," *Mathematics of Operations Research*, vol. 33, no. 4, pp. 208–217, 2004.

[79] A.M. Campbell and M.W.P. Savelsbergh, "Decision Support for Consumer Direct Grocery Initiatives," *Transportation Science*, vol. 39, no. 3, pp. 313–327, 2005.

[80] A.J. Kleywegt, A. Shapiro, and T. Homem-de Mello, "The Sample Average Approximation Method for Stochastic Discrete Optimization," *SIAM Journal on Optimization*, vol. 12, no. 2, pp. 479–502, 2002.

[81] A. Shapiro, *Handbooks in Operations Research and Management Science*, New York : Elsevier, 2003.

[82] B. Verweij, S. Ahmed, A.J. Kleywegt, G. Nemhauser, and A. Shapiro, "The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study," *Computational Optimization and Applications*, vol. 24, no. 2, pp. 289–333, 2003.

[83] J. Dupacová, "Multistage Stochastic Programs: The State-of-the-art and Selected Bibliography," *Kybernetika*, vol. 31, no. 2, pp. 151–174, 1995.

[84] L. Mercier and P. Van Hentenryck, "Performance Analysis of Online Anticipatory Algorithms for Large Multistage Stochastic Integer Programs," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, June 2007, pp. 1979–1984.

[85] L. Mercier and P. Van Hentenryck, "*Amsaa*: A Multistep Anticipatory Algorithm for Online Stochastic Combinatorial Optimization," *Lecture Notes in*

Computer Science, vol. 5015, pp. 173–187, 2008.

[86] R. Schultz, L. Stougie, and M. H. Van Der Vlerk, "Two-Stage Stochastic Integer Programming: A Survey," *Statistica Neerlandica*, vol. 50, pp. 404–416, 1996.

[87] J. Atlason, M.A. Epelman, and S.G. Henderson, "Call Center Staffing with Simulation and Cutting Plane Methods," *Annals of Operations Research*, vol. 127, no. 1, pp. 333–358, 2004.

[88] J. Atlason, M.A. Epelman, and S.G. Henderson, "Optimizing Call Center Staffing Using Simulation and Analytic Center Cutting Plane Methods," *Management Science*, vol. 54, no. 2, pp. 295–309, 2008.

[89] P.M. Vanden Bosch and D.C. Dietz, "Minimizing Expected Waiting in a Medical Appointment System," *IIE Transactions*, vol. 32, no. 9, pp. 841–848, 2000.

[90] B.T. Denton, A.S. Rahman, H. Nelson, and A.C. Bailey, "Simulation of a Multiple Operating Room Surgical Suite," in *Proceedings of the 2006 Winter Simulation Conference*, Dec. 2006, pp. 414–424.

[91] V. Podgorelec and P. Kokol, "Genetic Algorithm Based System for Patient Scheduling in Highly Constrained Situations," *Journal of Medical Systems*, vol. 21, no. 6, pp. 417–427, 1997.

[92] T.R. Rohleder and K.J. Klassen, "Using Client-variance Information to Improve Dynamic Appointment Scheduling Performance," *Omega*, vol. 28, no. 3, pp. 293–302, 2000.

[93] B.P. Zeigler and H. Sarjoughian, *Introduction to DEVS Modeling & Simulation with JAVA^{TM}: Developing Component-based Simulation Models*, Tucson, Arizona, 2003, Accessed from http://www.acims.arizona.edu.

[94] E. Pérez, L. Ntaimo, C. Bailey, and P. McCormack, "Modeling and Simulation of Nuclear Medicine Patient Service Management in DEVS," *Simulation*, 2010, Online first.

[95] P. Wing, M.H. Langelier, and A.P. De, "Nuclear Medicine Scientists: Findings and Recommendations Based on a 2006 Survey," *Journal of Nuclear Medicine*, vol. 48, no. 4, pp. 11–24, 2007.

[96] J. Brice, "Utilization Review Takes Aim at Imaging Expenses ," *Diagnostic Imaging*, vol. 26, no. 12, pp. 26–31, 2004.

VITA

Name:

Eduardo Pérez Román

Address:

3131 TAMU

College Station,

TX 77843-3131

Email:

eduardo.inin@gmail.com

Education:

B.S. Industrial Engineering,

University of Puerto Rico at Mayaguez, 2004

Ph.D. Industrial Engineering,

Texas A&M University, 2010