January 2020

# Dynamic Resource Allocation For Coordination Of Inpatient Operations In Hospitals

Najibesadat Sadatijafarkalaei
*Wayne State University*

**DYNAMIC RESOURCE ALLOCATION FOR COORDINATION OF INPATIENT OPERATIONS IN HOSPITALS**

by

**NAJIBESADAT SADATIJAFARKALAEI**

**DISSERTATION**

Submitted to the Graduate School,

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

**DOCTOR OF PHILOSOPHY**

2020

MAJOR: INDUSTRIAL ENGINEERING

Approved By:

__________________________________________

Advisor Date

## DEDICATION

I would like to dedicate my work to my great family for their endless love, encouragement and support.

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my immense gratitude to my two co-supervisors. Dr. Ratna Babu Chinnam for his continuous support, motivation and academic advice throughout this dissertation project. He has been very supportive since the day I began working on my PhD program at Wayne State University. Dr. Evrim Dalkiran for her support and scientific advice in different steps of this research. Their guidance helped me to conduct all stages of this research. I really appreciate all of their contributions of time, support and ideas. I could not have imagined having better mentors for my doctoral research.

My special gratitude also goes to Dr. Ekrem Alper Murat and Dr. Yanchao Liu for their participation in the weekly meetings and providing invaluable comments and instructions. Their contributions toward the completion part of this dissertation led to significant improvements with optimization and mathematical modeling.

I am grateful to Dr. Dongxiao Zhu and Dr. Murat Yildirim, the members of the examining committee, who have improved this dissertation with their constructive feedback. I am grateful for their time and effort spent in reading the manuscript.

I would like to thank Dr. Mike Nauss and Dr. Seth Krupp, Directors for the main Emergency Department at the Henry Ford Health System, in providing me a great opportunity to access the required data, visiting the hospital operations for observation, and their feedback on my research. They gave me very helpful comments and provided access for close observation of the operations across the ED to Inpatient Unit network and studying the current process of hospital resource allocation and identifying potential opportunities for improvement.

I thank my fellow students in the ISE department, Tannaz Khaleghi, Elham Taghizadeh, Seungyup Lee and Maduka Balasooria for their stimulating discussions and advice.

Last but definitely not least, thanks to my family for their endless love and support during my PhD program. For my parents for giving birth to me and supporting me in all stages of my life. I must express my gratitude to Milad, my husband, for his continued love and encouragement. His support was in the end what made this dissertation possible.

TABLE OF CONTENTS


Dedication . . . . . ii

Acknowledgements . . . . . iii

List of Tables . . . . . viii

List of Figures . . . . . ix

Chapter 1 INTRODUCTION . . . . . 1

1.1 Research Framework . . . . . 4

1.2 Contributions . . . . . 7

Chapter 2 A Review on Reactive and Proactive Modeling of Resource Allocation in Health Systems . . . . . 10

2.1 Reactive Modeling of Resource Allocation in Health Systems . . . . . 11

2.1.1 Mathematical Modeling . . . . . 12

2.1.2 Simulation Modeling . . . . . 15

2.1.3 Queuing Modeling . . . . . 16

2.2 Proactive Modeling of Recourse Allocation in Health Systems . . . . . 17

2.3 Summary . . . . . 22

Chapter 3 INTEGRATED COORDINATION APPROACH USING MIXED-INTEGER PROGRAMMING . . . . . 24

3.1 Problem Description . . . . . 24

3.1.1 Patient to IU bed assignment . . . . . 26

3.1.2 EVS to IU dirty-bed assignment . . . . . 26

3.1.3 Patient to transport staff assignment . . . . . 27

3.2 Coordination Model Formulation . . . . . 28

3.2.1 Maintaining decision continuity between runs . . . . . 33

3.2.2 Special considerations . . . 34

3.3 Solution Approach . . . 35

3.3.1 Pre-processing . . . 36

3.3.2 Iterative optimization . . . 37

3.4 Computational Experiments . . . 39

3.4.1 Random Problem Instance Generation . . . 39

3.4.2 First-Come First-Served (FCFS) Approach . . . 42

3.4.3 Results . . . 43

Chapter 4 INTEGRATED PROACTIVE COORDINATION APPROACH USING STOCHASTIC OPTIMIZATION . . . 51

4.1 Problem Description . . . 51

4.2 Proactive Coordination Model Formulation . . . 52

4.2.1 First Stage . . . 53

4.2.2 Second Stage . . . 59

4.3 Solution Approach . . . 61

4.3.1 Scenario Generation . . . 61

4.3.2 Pre-processing . . . 62

4.3.3 Two stage optimization models . . . 63

4.4 Computational Experiments . . . 65

4.4.1 Random Problem Instance Generation . . . 65

4.4.2 Results . . . 69

Chapter 5 CONCLUSION AND FUTURE research . . . 74

5.1 Summary and Conclusion . . . 74

5.2 Future Research . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 75
5.2.1 Enterprise-wide resources allocation and coordination . . . . . . . . 75
5.2.2 Computationally effective solution approach using Machine Learning 76
References . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 78
Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 90
Autobiographical Statement . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94

## LIST OF TABLES


Table 1 Sets, indices, and parameters utilized in the optimization model . . . . 25

Table 2 Variables in the optimization model . . . . . . . . . . . . . . . . . . . . 28

Table 3 Parameter values utilized in the experimental study . . . . . . . . . . . 42

Table 4 Patient Waiting Time Breakdown (in minutes) . . . . . . . . . . . . . . 46

Table 5 Sets, indices, parameters, and variables utilized in the first step model . 54

Table 6 Sets, indices, parameters, and variables utilized in the second step model 58

Table 7 Average Waiting Time of Patient (minutes) . . . . . . . . . . . . . . . . 70

Table 8 Maximum Waiting Time of Patient (minutes) . . . . . . . . . . . . . . . 70

Table 9 Number of Boarded Patients . . . . . . . . . . . . . . . . . . . . . . . . 70

Table 10 Average Waiting Time of Boarded Patient (minutes) . . . . . . . . . . . 71

Table 11 Median Waiting Time of Patient (minutes) . . . . . . . . . . . . . . . . 71

Table 12 Average Percentage of Time for [Cleaning, Traveling, Idle] for EVS Staff 72

## LIST OF FIGURES

Figure 1 Hospital task coordination and resource allocation: Conceptualization of hospital services requiring multiple sub-services, tasks and resources (left) and ED patient admission to an IU as a special case. . . 4

Figure 2 Framework for coordination of care and resource allocation within the ED-to-IU network . . . . . . . . . . . . . . . . . . . . . . . . . 6

Figure 3 Comparison of reactive and proactive resource allocation approaches across the ED-to-IU network . . . . . . . . . . . . . . . . . . . . . 7

Figure 4 Flowchart for proposed solution approach . . . . . . . . . . . . . . . . 38

Figure 5 Distributions of Patient Admission and Discharge Times . . . . . . . . 40

Figure 6 Patient waiting time performance under the proposed coordination approach versus FCFS strategy . . . . . . . . . . . . . . . . . . . . . . 43

Figure 7 Performance of proposed approach in comparison with FCFS . . . . . 46

Figure 8 Resource utilization under proposed versus FCFS approaches . . . . . 48

Figure 9 EVS staff utilization under the proposed coordination approach through day 8 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49

Figure 10 EVS staff utilization under the FCFS strategy through day 8 . . . . . . 50

Figure 11 Reactive resource allocation approach across the ED-to-IU network . . 52

Figure 12 Flowchart for proposed solution approach . . . . . . . . . . . . . . . . 64

Figure 13 Distributions of Patient Arrival and Discharge Times . . . . . . . . . . 66

Figure 14 Patient waiting time performance under the proposed proactive (stochastic and deterministic) approach versus reactive (optimization, FCFS) approach . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 73

# CHAPTER 1 INTRODUCTION

Healthcare has become a large industry, with a high number of human and medical resources in healthcare systems and organizations with several complicated processes [78]. Healthcare is a particularly significant service industry because not only the quality and safety in delivering of care is critical [68], but also the associated expenses are very high [35]. According to a recent JAMA study, roughly 20–25% of U.S. healthcare spending is wasteful [83]. Spending grew 3.9% in 2017, reaching $3.5 trillion or $10,739 per person, accounting for 17.9% of the GDP [27].

Healthcare systems face several challenges such as increasing process complexity, inefficient utilization of resources, high pressure to enhance the quality of care and services, and the need to balance and coordinate the workload of health systems staff [44, 3, 29, 9]. Therefore, the need for effective and efficient processes for delivering healthcare services is imperative. Data-driven approaches including operations research and predictive modeling can overcome these challenges and improve the health systems performance in terms of quality, cost and patient satisfaction. These challenges lead to increased research interest in several domains of healthcare for many scientists. The key data-driven healthcare problems in different studies can be summarized as resource allocation and scheduling, logistics planning, medical treatment, preventive care and disease diagnosis with the main focus on hospitals processes and services. [19, 18, 70, 34, 5].

Hospitals are a key component of healthcare systems with many scarce resources such as caregivers (nurses, physicians) and expensive facilities/equipment. Unfortunately, when it comes to healthcare operations and flows (patient, resource/equipment, information), they are complex, manual, and reactive, resulting in delays, under-utilization of critical resources, and most importantly, compromised health outcomes. Process and resource management is regarded a high priority for healthcare systems in order to control costs while achieving high quality of care [46]. It is widely reported and understood that resource allocation and coordination is of utmost importance in managing the efficiency and effectiveness of healthcare systems and hospitals [60, 90].

Poor bed management and ineffective patient transfers are two important factors associated with hospital crowding, cost inefficiency, and patient dissatisfaction [2]. Patient transfer is a critical aspect of workflow within healthcare systems and the daily rate of patient transfer in inpatient departments of U.S. hospitals is 40−70% on average [49]. This high rate of patient transfer has far-reaching impact on resource utilization of the hospital. Therefore, providing efficient patient transfer and coordination is crucial to achieving cost efficiency and delivery of timely and appropriate care [54]. Moreover, patient transfers by definition involve different departments and efficient transfers cannot be achieved without inter-departmental coordination solutions and technologies [2]. Optimized bed assignment is the other critical aspect in care management and it is highly dependent on efficient coordination of different tasks including bed identification, cleaning, and assignment. Collectively, bed management and patient transfer are two important sub-services that significantly affect hospital performance and efficiency [54, 81]. Given the complexity of healthcare operations, bed management needs an integrated system-wide approach to provide resource and activity orchestration based on real-time information (about patients and resources) and optimized decision making [90]. Fortunately, most hospital systems in the developed world have employed some form of an Electronic Health Record (EHR) system in recent years to form a critical data backbone to support the realization of such orchestration platforms. Real-time information available in EHR systems can play a significant role in providing better operational coordination between different departments/services in the hospital through optimized task/resource allocation.

In this research, we particularly focus on the problem of resource and task coordination within the care network spanning the patient flow from Emergency Department to Inpatient Units (ED-to-IU network) to reduce ED patient admission waiting times. EDs are an important gateway to hospitals around the world and account for more than 50% of admissions in the U.S. hospitals [1]. Poor coordination in bed management (tracking, turnaround operations, and allocation) and delays in transferring admitted ED patients to inpatient beds leads to patient "boarding". This is a condition where a patient being admitted into the hospital at the end of ED treatment is "held up" within the ED due to delays attributable to factors such as admission approval, lack of clean inpatient unit beds, and patient transport resource shortage. Boarded patients not only occupy critical resources within the ED, limiting access to other patients seeking ED services, but also have a significant affect on healthcare cost, outcomes, and patient/staff satisfaction [2]. ED patient boarding is currently regarded an international crisis, and in the U.S., the Center of Medicare and Medicaid Services (CMS) has been requiring hospitals to report ED patient boarding statistics since 2014. While the median ED treatment time for admitted patients is 5.5 hours in 2016, the median boarding time is another 2 hours and 16 minutes and the problem is prevalent across all states and regions of the U.S. [26]. To address this pressing problem, we propose an integrated system-wide approach for real-time orchestration of tasks and resources across different departments within the ED-to-IU network in order to minimize ED patient boarding time.

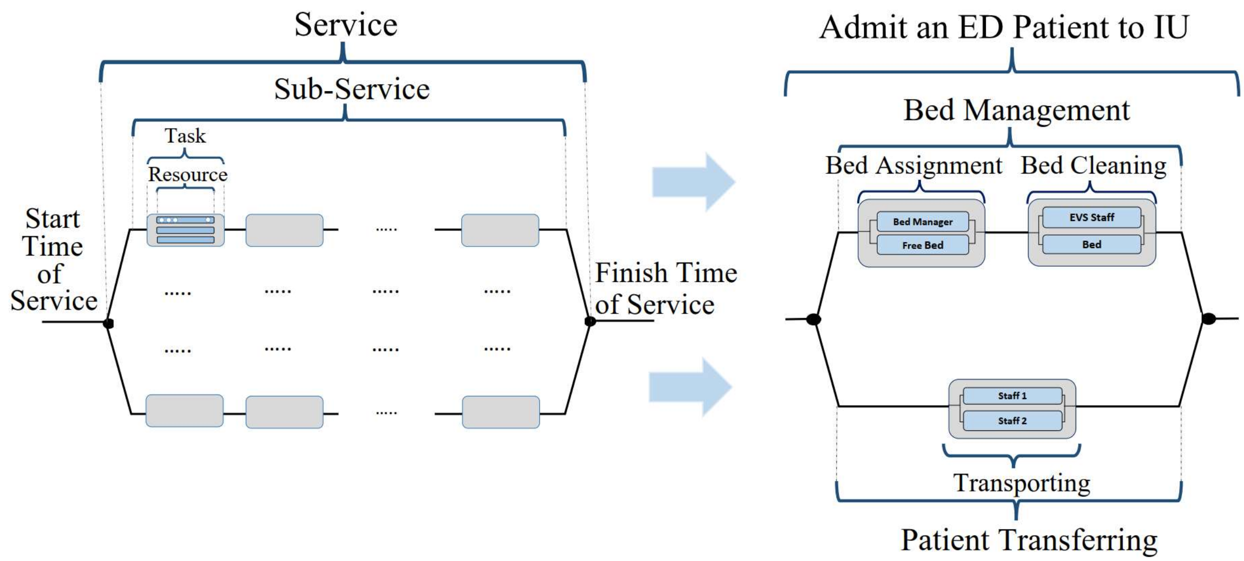



Figure 1: Hospital task coordination and resource allocation: Conceptualization of hospital services requiring multiple sub-services, tasks and resources (left) and ED patient admission to an IU as a special case.

Figure 1 demonstrates our research scope. The left side of Figure 1 illustrates the hierarchy embedded in a healthcare system service, where a service can consist of multiple sub-services and each sub-service includes different tasks and resources. While healthcare system involves several services as a complex system, in this study, we focus on the service of admitting an ED patient to inpatient units within the hospital. The right side of Figure 1 displays admitting an ED patient to an IU as a special case of healthcare system services. We consider bed management and patient transfer as two important sub-services that significantly affect hospital performance and efficiency. [54, 81].

### 1.1 Research Framework

Several studies have been conducted in the domain of healthcare resource allocation and scheduling. Since the costs of healthcare around the world are still rising, specifically in the U.S., proposing a novel and effective optimized approach for resource allocation is necessary [31]. There exist several different factors that affect health resource utilization and efficiency. An important factor is coordination of care across departments in order to achieve better resource allocation and quality of care. Coordinated resource allocation focuses on how well the right quantities of healthcare resources are managed and allocated among health services from an operations research perspective [31]. In our coordination framework, we consider different resources (inpatient beds, cleaning staff, transporters) and tasks (patient transportation, bed cleaning, and bed assignment) to provide an event-based dynamic approach where the event is defined according to the availability of a new patient, task or resource. We develop a mixed-integer programming model [8, 93] for solving the resource allocation problem, which consists of several real-world constraints/requirements such as patient gender matching during bed assignment to double rooms, appropriate and equitable resource assignment, resource availability time, patient/room isolation constraints, patient over-flow policies, staff shifts and changeovers.

We define a multi-agent system framework including a transportation team, an environmental services (EVS) team responsible for cleaning beds/rooms, emergency department, and inpatient units. We use real-time EHR information from hospital systems for coordination of different agents and develop an integrated optimization model for resource allocation. Our main decisions include: 1) IU bed assignment to patients, 2) EVS staff assignment for cleaning dirty beds, and 3) Transportation staff assignment to transport patients. Our framework is illustrated in Figure 2.

We focus on two scenarios in our proposed framework. In the first scenario, we consider the reactive approach of resource allocation in ED-to-IU network in hospitals. We assume that the resources are assigned after the admission decision to IU and disposition decision from the IU. We develop a deterministic dynamic real-time coordination model

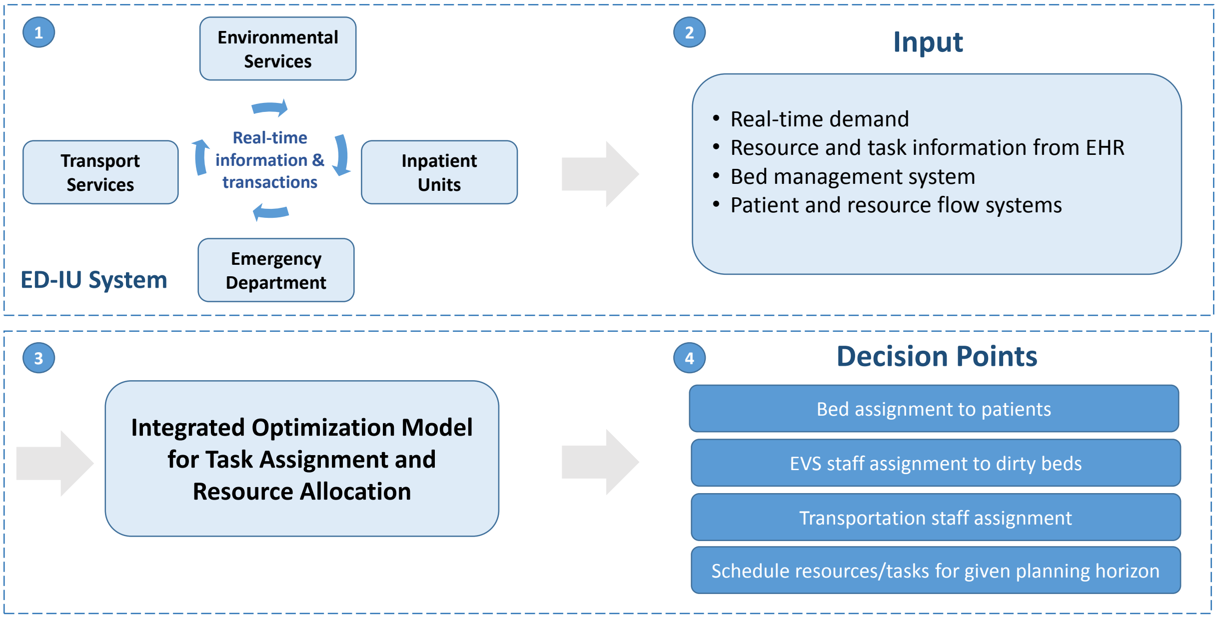



Figure 2: Framework for coordination of care and resource allocation within the ED-to-IU network

for resource and task assignment using mixed-integer programming. Next, we introduce a proactive approach to demonstrate how early task initiation [13] using available EHR information in upstream tasks like as triage and initial patient assessment can improve the reactive approach for resource coordination. We propose a proactive stochastic MIP model that significantly extends the reactive deterministic MIP model via incorporating the uncertainties in patient admissions. Our proposed proactive approach demonstrates that ED patient waiting times are further reduced when a reliable prediction of ED admission decision ahead of the actual admission decisions are available. We assume that when a new patient arrives at emergency department and is started the testing and treatment process, the information about the patient, such as patient's health history, provides reliable estimation of disposition decisions and admission times before the actual admission. This

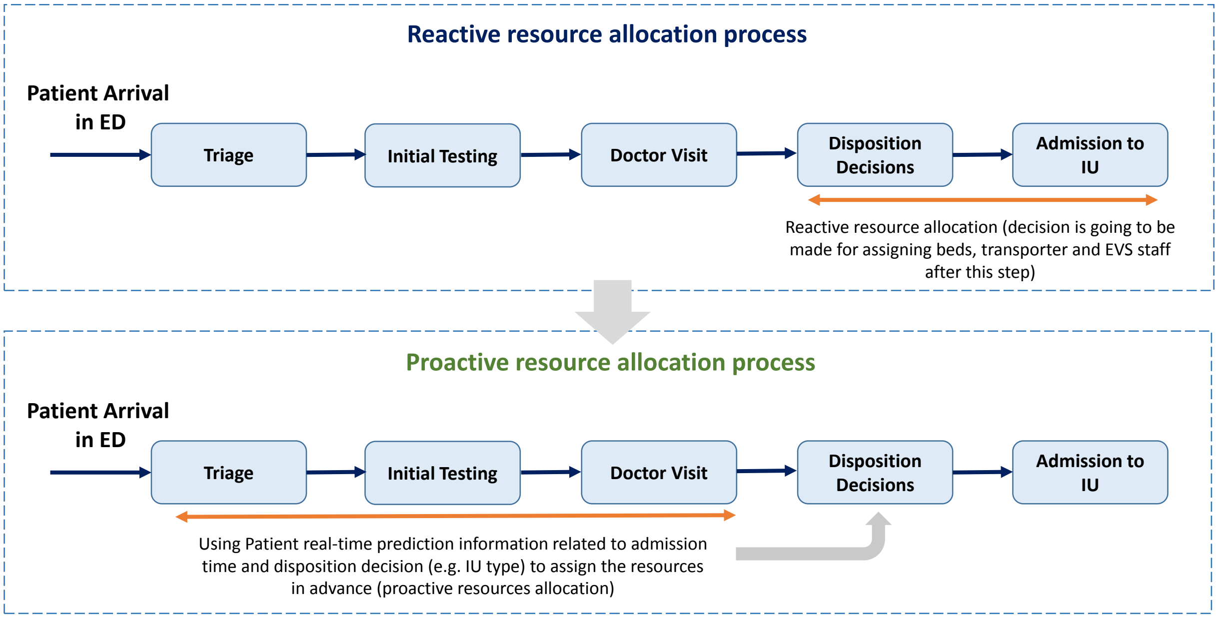



Figure 3: Comparison of reactive and proactive resource allocation approaches across the ED-to-IU network

estimation helps to improve the resource allocation to relevant tasks in a proactive manner regarding an impending demand. Figure3 demonstrates reactive and proactive resource allocation processes across the ED-to-IU network.

Since first-come first-served (FCFS) is extensively used for bed management and patient transfers in many healthcare systems, we compare the performance of our model with FCFS approach using data from a leading healthcare facility in SE-Michigan. The proposed reactive and proactive approaches are shown to significantly outperform FCFS practices prevalent in hospitals in terms of waiting times for admitted ED patients while also improving resource utilization, availability, and workload equity.

### 1.2 Contributions

The contributions of this research are summarized as following:

- First of all, we propose a novel and flexible real-time resource allocation and co-ordination model that can manage different types of resources. Our model is one of the the first models to propose dynamic optimized coordination within the ED-to-IU network (bed assignment, bed cleaning, and patient transfers) using mix-integer programming. The developed model provides assignment of resources to tasks, but also schedules the tasks for a rolling planning horizon while satisfying many real-world requirements and constraints. We consider several real-case constraints/requirements such as patient gender matching during bed assignment to double rooms, appropriate and equitable resource assignment, resource availability time, patient/room isolation constraints, patient over-flow policies, staff shifts and changeovers. Our model successfully coordinates bed and staff allocation and reduces the boarding time significantly.

- In addition to reactive approach which has been commonly employed in previous studies, we proposed the proactive approach (early task initiation) for both deterministic and stochastic versions of the developed MIP model. Our results show that resource utilizations are improved and patient boarding times are reduced by taking advantage of real-time EHR information.

- We propose tailored solution approaches for reactive and proactive models including several pre-processing methods to improve the expensive computations of optimization models. We also provide a specific and efficient sampling strategy for generating scenarios for the stochastic version of the MIP model.

The rest of this dissertation is organized as follows: In Chapter 2, we review the related works for reactive and proactive resource allocation models developed for health systems. Chapter 3 discusses the proposed MIP models and their formulations in detail based on the reactive approach. In Chapter 4, we analyze the proposed proactive approach based on deterministic and stochastic MIP models. Finally, in Chapter 5, we provide summary and conclusions of this dissertation and propose directions for future research.

## CHAPTER 2 A REVIEW ON REACTIVE AND PROACTIVE MODELING OF RESOURCE ALLOCATION IN HEALTH SYSTEMS

In recent years, many researchers have focused on the problem of improving the performance of health systems, particularly healthcare delivery services. As a main workstream, researchers addressed the resource allocation and coordination among different units of hospital [43]. Various studies considered different types of resources in their proposed models. Among the many scarce resources, the following are extensively studied: 1) staff doctors and nurses [30, 52], operating theaters for surgeries [57, 80, 89] and 3) inpatient unit beds (managed through emergency department to inpatient units network) [6, 28, 37]. Although these studies proposed approaches to improve health systems, only a few have addressed the complex interactions between patients and the various hospital units such as intensive care units (ICU) and emergency departments. The complex dynamics of the relationship between different units and resources has not been addressed in depth [43].

There exist many studies for resource allocation and patient flow management to improve the efficiency of healthcare systems focusing on the aforementioned scarce resources. While several studies have proposed reactive models, some studies have provided proactive resource allocation models by early task initiation in upstream processes using electronic health records (EHRs) including patient health history and previous admission records [13]. Proactive management and early task initiation enable emergency department (ED) to proactively contact to the relevant inpatient unit (IU) and request a bed for patient waiting to be admitted [10, 56]. In this chapter, we review the literature on reactive and proactive approaches for resources and tasks allocation in hospitals.

## 2.1 Reactive Modeling of Resource Allocation in Health Systems

In this section, we primarily focus on studies that have employed reactive modeling for healthcare task assignment and resource allocation. While majority of the approaches use mathematical programming [96, 85, 81, 92, 74, 23], several studies also report on the application of simulation modeling [77, 66, 58, 79, 86, 40] and queuing theory [69, 88, 82, 4, 67, 59] for resource allocation and bed capacity management [47]. Simulation models are often used for scenario analysis to simulate patient flows and process workflows, and can assist in staffing and bed capacity management decisions [47, 22]. There are also studies that promote discrete-event simulation to support bed allocation and manage patient flows [50, 32]. However, in general, simulation based decision support systems for operations management are not practical, and they are too demanding in terms of initial model calibration and continuous model refinement over time to keep pace with process/policy changes and data quality. Queuing theory is the mathematical study of waiting lines, or queues, and is able to consider time-dependent stochastic flows. In the context of hospital systems, it allows researchers to model the impact of patient arrival process, service duration and resource levels on patient flows and resource utilization. Other approaches such as scenario analysis can explore the impacts on the outcomes of the queuing system and provide resource utilization by systematic variations in the input parameters [14]. Bed assignment policy planning [67, 53], resource utilization management [14], and patient priority management [45] are applications of queuing theory that addressed in the healthcare domain. While queuing models can provide good insights into system dynamics for capacity planning, staffing, and policy development, they are not practical for real-time healthcare operations management.

### 2.1.1 Mathematical Modeling

Mathematical programming has been extensively employed for the scheduling and allocation of scarce resources to improve flow and utilization. We first focus on the studies with explicit bed assignment considerations. Luscombe and Kozan [64] proposed a new dynamic resource allocation and scheduling model and heuristic solution algorithms for bed assignment and task-resource sequencing in EDs. They reported on the performance of the proposed approach using historical data. More recently, Feng et al. [39] considered resource allocation in EDs and proposed a new stochastic multi-objective optimization model minimizing the average patient length of stay in hospital and medical resource wastage. Their proposed meta-heuristic solution approach integrates a non-dominated sorting genetic algorithm II with multi-objective computing budget allocation. Authors reported on the results of a computational study and a discrete event simulation model of the ED flow at a Taiwanese hospital. Both studies rely on heuristic solution approaches.

For dynamic patient admission scheduling problem, Ceschia and Schaerf [25] considered optimal multi-day assignment and scheduling of patients to inpatient beds subject to capacity and gender policy constraints. A meta-heuristic approach using simulated annealing and a complex neighborhood structure was proposed and evaluated through an experimental study. The results showed that their model is able to solve large problem instances within reasonable computational time. In another general bed assignment study, Thomas et al. [90] proposed an analytical decision support framework using mixed-integer goal programming to assign beds to patients. In addition to multiple goals, their formulation considers operating constraints such as staff and hospital requirements, unit utilization requirements, and gender mismatch requirements. Similar to our study, their approach collects input data as a snapshot of the hospital states from real-time data and workflow system. The proposed formulation however does not account for the coordination of bed assignment with other support services such as bed cleaning or patient transport. In addition, their model is not able to consider scheduling of tasks for resources across all supporting departments in order to achieve better system-wide resource allocation in the time horizon. More recently, Burdett and Kozan [21] proposed a deterministic integrated approach for resource allocation and task scheduling using flexible job shop scheduling where patients, beds, hospital inpatient units and health care activities are considered as jobs, single machines, parallel machines, and operations, respectively. They proposed a hybrid meta-heuristic algorithm for solving the problem. They applied numerical tests to evaluate the efficiency of their model. Since the flexible job shop scheduling problem is an NP-hard problem, the authors developed a hybrid Simulated Annealing (HSA) approach and constructive algorithms which achieve high quality near optimal solutions. Their model assigns patients to beds and other treatment locations (such as pre-operation, operating rooms and recovery units) and schedules the patient's activities in the assigned locations. This model takes a multi-stage approach and each job is a set of activities. The authors did not compare the proposed approach with current real-world practice to evaluate its performance. In addition, their model does not account for many real-world constraints.

Several other studies consider general resource-task assignment and scheduling without explicitly considering bed assignment. Punnakitikashem et al. [75] introduced a new integrated nurse staffing assignment model using stochastic integer programming to minimize the staffing cost and their workload. They applied their model on real data from a Northeast Texas hospital. As a solution framework, they developed three approaches including: 1) Bender's decomposition, 2) Lagrangian relaxation with Bender's decomposition, and 3) Nested Bender's decomposition. Hosseini et al. [51] developed a resource allocation and coordination model using multi-agent systems where there exist multiple agents and multiple tasks. They used Multi-agent Markov Decision Process (MMDP) in their approach. Zaerpour et al. [95] proposed a mixed integer programming (MIP)-based model for assigning time slots to the medical doctors in order to maximize the minimum service level across blocks of time. A branch and price heuristic algorithm is developed to solve practical problem instances. Authors evaluated the efficiency of their model based on numerical examples and two real-world case studies. Bastian et al. [12] developed a stochastic multi-objective auto-optimization model for resource allocation in fixed-input health systems such as Military Health System. The proposed approach is a strategic-level model with the goal of optimizing overall system performance. This model is proposed for better resource sharing across large healthcare units, however, it is not appropriate to apply for tactical or operational control.

Another branch of healthcare literature studies the ambulance allocation and coordination between hospitals. Although these studies do not focus on patient-bed allocation and coordination tasks, their proposed approach can be applied for general tasks/resources allocation problems including patient-bed allocation. Lopez et al. [61] proposed a multi-agent auction mechanism to coordinate the ambulances for emergency medical services. Their approach is a trust-based auction algorithm which only considers ambulance allocation. They evaluated the model performance through a simulation study. Billhardt et al. [17] proposed a novel coordination model for ambulance assignment that provides an integrated framework for dynamic redeployment of available ambulances along with a dynamic allocation of ambulances to patients. Their model is based on dynamic auction-based assignment of patients to ambulances where the goal is to optimize the total expected arrival times in each particular period. Authors developed three heuristic auction algorithms and tested them under different settings using real-world data. Lujak et al. [62] considered the problem of coordinating Emergency Medical Assistance (EMA) and hospitals for after-hours surgeries of urgent patients arriving by ambulance. They modeled this problem as a multi-agent system for task allocation and coordination for minimizing the average delay in assigning surgery teams for emergency patients. The authors developed an auction based solution approach for providing the best assignment solution for the whole system. They used a simulation study for evaluation and showed that their model outperforms first-come, first-served strategy. Lujak et al. [63] further extended this EMA coordination model for urgent out-of-hospital ST-segment elevation patients awaiting angioplasty. This extended model is a three-level optimization model where a globally efficient solution is proposed by using an auction algorithm in each level.

### 2.1.2 Simulation Modeling

Simulation models are used for scenario analysis to simulate patient flows and process workflows, and assist in bed management decisions. Many articles used simulation approaches to support bed allocation and manage patient flows [47]. Discrete event simulation (DES) is the most popular simulation approach applied to patient flow dynamics and provides analysis of different bed allocation scenarios between health systems units. Holm et al. [50] used a generic discrete event simulation for modeling of patient flow between the hospital wards. Using the simulation model, they generate utilization statistics given the numbers of beds for each ward and propose an allocation algorithm to optimally distribute available beds among the wards. Devapriya et al. [32] developed a decision support tool for bed capacity management based on DES. The model captures real-world patient flow data from various processes such as patient arrival and discharge and analyzes admission waiting time by arrival source and assigned bed, and occupancy rates. In another study, Mallor and Azcarate [65] introduced a simulation approach combined with an optimization model for Intensive Care Unit (ICU) bed capacity management. The objective function of optimization model is evaluated through simulation.

### 2.1.3 Queuing Modeling

Queuing models and techniques are widely applied in healthcare systems to improve resource management. Using queueing theory, Belciug and Florin [14] proposed an integrated framework for bed allocation and financial resource utilization. They model the patient flow using M/PH/c queuing system, where the patients arrivals is based on Poisson process, hospital beds are servers, and the patient length of stay is simulated using a phase-type distribution. The authors also provided an evolutionary optimization model to optimize both bed allocation and resource utilization, and presented what-if analysis to evaluate various options. In another study, Mathews and Long [67] provided a framework based on queuing and simulation for data-driven modeling of patient flow between ICU and step-down units (SDU) to analyze the impact of different bed allocation schemes. Patient type, patient arrival rate, time of patient transfer, service time, number of beds, patient priorities, and unit length of stay are estimated based on real data and inputted to the queuing model. Kilinc et al. [53] studied the dynamic assignment of ED admitted patients to hospital inpatient wards. They introduced a queuing framework and MDP model to provide effective mechanisms in order to minimize the risk of patient safety in ED while decreasing the number of secondary inpatient unit assignments for better quality of care.

## 2.2 Proactive Modeling of Recourse Allocation in Health Systems

To improve resource allocation and coordination in health systems especially in ED-IU network, several studies have been developed based on proactive strategy. The enriched electronic health records provide promising opportunity to predict real-time admission demand for different resources in health system while the patient is waiting to be admitted or get initial treatment. Developing proactive/progressive approach for improvement of patient flow and resource allocation (e.g. bed management) in emergency department have been increasingly developed in recent years and researchers have proposed different models including mathematical, simulation and queuing approaches as discussed in previous section.

Thompson et al. [91] addressed the problem of demand surges for inpatient beds in hospitals when patients face a long delays from admission to assign to the bed in floor. They proposed a decision support systems for bed management based on proactive transfers of patients between floor prior to the occurrence of a demand surge. In their approach in-house patients are transferred for the purpose of bed reallocation which is defined as proactive transfer, as opposed to as a last called and immediate decision to provide room for newly admitted patients which is defined as reactive transfer. Authors modeled the problem as a finite-horizon Markov decision process (MDP) and developed an approximation algorithm for solving the optimization problem. The authors implemented their approach on real-case problem and achieved significant cost saving by decreasing almost 50% of patient waiting time in the average to be admitted and being transferred to a floor. Peck et al. [71] applied predictive modeling to improve emergency department (ED) crowding problem by balancing the demand and supply of the resources. Their approach is to predict the number of emergency department (ED) patients who sequentially will be admitted to a hospital inpatient unit (IU) and proposed a new framework to utilize these predictions in order to improve the hospital resource allocation proactively. Their main contribution is to improve ED-to-IU patient flow by predicting admission demand when patients arrive to the ED in real time. Their framework defines aggregated individual patient admissions predictions as a measure of near-future IU bed demand which may be helpful for resource daily resources coordination in hospitals.

In another study, Peck et al. [72] applied discrete event simulation to investigate the patient flow effects on prioritization of inpatient units using prediction of admission and current state information in ED. Their results based on simulated hospital indicate that sharing prediction and crowding information in ED-IU network impacts on IU staff priorities which can lead in statistically significant improvement in patient boarding time. Gartner et al. [42] combined machine learning and mix-integer programming (MIP) to improve upstream planning for scares resource allocation decisions in hospitals, focusing on predicting of diagnosis related groups. The results of this study shows that early and accurate diagnosis group classification using machine learning, associated into an optimization resource allocation model, can increase the number of admitted patients and improve the utilization of resources such as operating rooms and beds. For machine learning part, authors selected subset of patients attributes as input of different classifiers and evaluated the performance of prediction model using appropriate metrics. In the other side, for MIP optimization model, authors considered several constrains in the model for maximizing the income margin of the patients that are admitted to or kept in the hospital minus resource over-utilization cost as optimization objective.

El-Rifai et al. [36] proposed a staff allocation model for the seasonal epidemic situation using on-calls to manage the uncertainties in demand and staff workload. An on-call scheduling policy is developed to make a balance between demand coverage and staff cost. The problem is formulated as two-stage stochastic integer linear model solved using a Sample Average Approximation (SAA) approach. The allocation model developed in this study is proactive where initial decision are made before the realization of the epidemic. The main focus of interest is emergency department in seasonal epidemic to reduce the overcrowding and balance staff workload. In the first stage, the allocation decisions are made based on estimation of demands using available incomplete data. The second stage handles decisions that are provided on a day-to-day basis. The first stage decision is provided at the beginning of the epidemic horizon and considers contracting staff to be on-call or on a regular duty in specific period. The Second stage decisions are provided consecutively at the beginning of each time horizon and specify the number of resources to call back to work among staff that are on-call. Several epidemic scenarios are formulated using real data for validation of the proposed approach.

Batt and Terwiesch [13] proposed early task initiation (ETI) approach (our focus in proactive approach) based on empirical study. They introduced a new load-dependent mechanism as a method of balancing workload by shifting conducting of some tasks to an upstream stage. For example, if in the triage step in ED, we can predict what tests will be ordered by the physician, we can order these tests at triage earlier and they will be started to be proceeded at triage stage. So this proactive strategy can decrease patient waiting to be seen by a physician and potentially eliminates one or more cycles in treatment process. This early task initiation (ETI) approach, which is a between-stages coordinator in which staff in an upstream stage proactively start tasks that are normally conducted by staff in a downstream stage. The results of ths study show that ETI achieves a reduction in treatment time by 20 minutes in average. In the other research, Barak-Corren et al. [11], proposed a progressive approach for improving patient flow in emergency departments. They applied logistic regression model to accurately predict patient's likelihood of hospitalisation at different stages of the treatment process. Their results indicated an accurate and early prediction of hospitalisation can speed up the bed coordination process and shorten the patient boarding time from ED to inpatient units.

Al-Refaie et al. [7] defined smart hospital as a health system that manage unexpected events and emergencies in real-time. In their study, they developed three optimization models for scheduling operating room during emergency events and proposed an hierarchical decision approach to integrate these three optimization models. First model consider newly opened rooms for ED patients, if there are more patients waiting in ED, the second model allocate emergency patients to untapped ranges and if still more beds is required and allocating all emergency patients to the untapped range is not feasible, then the operating room with the greatest free margin is rescheduled for both the emergency patients and the elective patients. The proposed framework act as proactive approach while first optimization model provides efficient resource utilization during ED events the other two optimization models improve the underutilized operating rooms time in a proactive way. In another study, Lee et al. [55] proposed predictive modeling for patient disposition decisions in emergency department. They applied a hierarchical multiclass classification approach to predict the appropriate inpatient unit for an admitted patient in ED in order to reduce boarding waiting times through the proactive initiation of admission process. Their work classifies the admitted ED patients into more granular classes so the results of the prediction is useful for unit-specific proactive coordination of tasks and resources across the ED-IU network. Their findings indicate there is a valuable predictive performance for the four admission classes with reasonable lead time for proactive resource coordination. For example, around 2.5 hours before the actual disposition decisions for the admitted patients.

Gonzalez et al. [43] focus specifically on delays in transferring patients from emergency department to inpatient units. They proposed a new Markov decision process (MDP) to enhance the efficiency of patient flow between ED and different patient units in hospital. The proposed approach is a proactive transfer method which estimates the demand of next period and determines how many patients in which period should be transfered to different hospital units. A dynamic programming approach is applied to provide an approximation of the optimal transfer policy, which specifies that a certain number of beds should be reserved in the different units according of the next period demand prediction. Authors considered three different decision-making levels based on hospital units: the intensive care (high complexity) unit, the intermediate care (medium complexity) unit, and the low complexity unit (WARD). In a recent research, Lee et al. [56] used the real-data from a major health system and showed emergency departments suffer from delays in patient boarding and reactive resource allocation is one of the main causes. To address this issue, they explored early downstream tasks initiation to reduce patient boarding in upstream. Specifically, they utilized the value of predicting ED patient disposition decisions to proactively model resource allocation mechanism as a fork-join queuing system. The proposed queuing network models complex operational interactions between tasks and resources and it is able to quantify the potential reduction in patient boarding delays based on function of bed request signal lead-time, accuracy of predicted decisions and two types patient arrival rate (ED admissions and non-ED admissions). The authors showed the proposed proactive inpatient bed allocation model can significantly reduce bed allocation delays for ED patients and does not increase waiting time for other admission sources.

### 2.3 Summary

Although the above reviewed studies provide various frameworks for resource assignment and coordination, none of them offer a comprehensive integrated framework for coordinating system-wide tasks across departments and resources simultaneously in real-time. Majority of these studies mainly focus on an isolated aspect of resource allocation in hospitals such as patient-to-bed assignment or staff allocation while there are remarkable opportunities to develop an integrated/system-wide resource allocation model by coordinating different services using real-time information.

In this research, we address this gap and develop an integrated resource allocation model for coordinating different tasks. Our approach proposes a real-time resource (bed, staff) and task (patient transportation, bed cleaning and bed assignment) allocation in parallel, which can improve management information systems in hospitals in order to achieve more efficient system level performance. Our model is an event-based approach where the event is defined according to the availability of patient, task or resource and can dynamically reassign the tasks/resources and update system information. It considers several real-world constraints as well as staffing shift changes affecting multiple resource types.

First, we provide the reactive approach which formulated as deterministic mixed integer optimization model to show how our resource/task coordination framework improve the baseline significantly and then we provide the proactive model to take the advantages of early tasks initiation and resources allocation as reviewed in this chapter to improve the proposed reactive model. In the proactive modeling, we consider two popular types of uncertainties in the literature including admission time and disposition decision.

# CHAPTER 3 INTEGRATED COORDINATION APPROACH USING MIXED-INTEGER PROGRAMMING

## 3.1 Problem Description

In this section, we propose an integrated model to optimize resource allocation within the ED-to-IU network. In our framework, we rely on real-time information available from the Electronic Health Record (EHR) system and sister IT platforms (e.g., Bed Management, Transport Planning, and Environmental Services systems) to provide coordination among three services within a fixed planning horizon: 1) Bed Management: Assignment of inpatient beds to admitted ED patients, 2) Bed Turnaround: Assignment of EVS staff to dirty beds for cleaning, and 3) Transportation: Assignment of transport staff to ED patients.

Our event driven approach treats the dynamic resource and task assignment and scheduling problem within a sequential optimization framework using a sliding planning horizon window. Within each planning horizon, a new problem instance is solved with deterministically known ED admitted patient pool and their attributes, resources' states and availability, and previously committed decisions of assignments and allocations. Shift from one planning horizon to the next is triggered either when a new epochal event (i.e., admission decision of a new patient or availability change of a resource such as bed, transport staff, EVS staff) occurs or a fixed duration is elapsed since the last optimization run. Between two consecutive optimizations, as new ED patients arrive and are treated and discharged or admitted into the hospital, we update resource availability and state information, arrival and discharge times of patients. Next, we discuss each service in detail, where the sets, indices, and parameters utilized in the formulation of this problem are given in Table 1.

Table 1: Sets, indices, and parameters utilized in the optimization model

| | |
|---|---|
| **Sets** | |
| $I, J, K$ | resources (beds, transport staff, EVS staff, respectively) |
| $U$ | inpatient units (a unit contains one or more rooms) |
| $R$ | rooms (a room contains one or more beds) |
| $P$ | patients |
| $P_M/P_F, P_O$ | male/female patients, patients needing isolation; $P_M \subseteq P, P_F \subseteq P, P_O \subseteq P$ |
| $I_r$ | beds belonging to room $r$; $r \in R$, $I_r \subseteq$ I |
| $I_{clean}/I_{dirty}, I_O$ | beds with status clean/dirty, beds in isolation rooms; $I_{dirty} \subseteq I, I_{clean} \subseteq I, I_o \subseteq I$ |
| $\bar{I}, \bar{J}, \bar{K}$ | resources (beds, transporters, EVS staff) common to two consecutive planning cycles |
| $\bar{P}$ | patients common to two consecutive planning cycles |
| **Indices** | |
| $i, j, k$ | an individual resource (bed, transport staff, EVS staff, respectively) |
| $p$ | patient |
| $d$ | order of a task among a series of tasks |
| $u$ | inpatient unit |
| $r$ | room |
| clean/dirty | binary indices referring to status of a bed |
| **Parameters** | |
| $T_h$ | planning time horizon |
| $T^{now}$ | time at the beginning of a new planning horizon |
| $t_p$ | time that the admit decision is made for patient $p$ |
| $t_i^b$ | time that a bed $i$ becomes vacant |
| $t_{k1}^e$ | time that EVS staff $k$ becomes available for its initial service |
| $t_k^{ES}$ | time that the shift ends for EVS staff $k$ |
| $t_{j1}^t$ | time that transporter $j$ becomes available for its initial service |
| $n_r$ | existing number of patients in room $r$ at the start of model run |
| $G_i^M/G_i^F$ | 1 if gender of patient currently occupying bed $i$ is male/female, 0 otherwise |
| $g_p$ | 1 if gender of patient $p$ is male, 0 if female |
| $C_p$ | 1 if patient $p$ requires the overflow constraint to be a hard constraint, 0 otherwise |
| $s_e^b$ | service time for cleaning a bed |
| $R_{x,x'}$ | travel time between two locations ($R_{ip}$ is travel time for ED patient $p$ to IU bed $i$) |
| $h_{ip}$ | maximum score level allowed for patient $p$ to assign to bed $i$ |
| $H_p$ | average of IU preference concession for patient $p$ |
| $\bar{x}_{ip}$ | 1 if bed $i$ was assigned to patient $p$ in the previous run of the model, 0 otherwise |
| $\alpha$ | fraction of bed to patient assignments that cannot change from previous model run |
| $w_p$ | bed assignment priority weight for patient $p$ |
| $\beta_1$ | penalty coefficient for total sojourn time of beds in dirty state in the objective function |
| $\beta_2$ | penalty coefficient for violating overflow constraints in the objective function |
| $\beta_3$ | penalty coefficient for maximum patient boarding time in the objective function |
| $\gamma$ | least number of cleaning task assignments allowed in each iteration |
| $D^j, D^k$ | maximum number of assignments allowed for transporter $j$, EVS staff $k$ |

### 3.1.1 Patient to IU bed assignment

At each epochal event, a set of patients $(P)$ that are admitted from the ED to hospital's IUs need to be assigned to a hospital inpatient bed $(I)$. Each patient $(p \in P)$ is assumed to have several clinical and non-clinical attributes such as bed assignment priority weight $(w_p)$, preferred IU based on patient care requirements, isolation requirements, expected length of stay in IU, admission time $(t_p)$, and gender information $(g_p)$. The goal is to assign patients to beds so as to minimize the total patient waiting time across all patients for bed assignment and transfer while trying to limit individual patient waiting times from exceeding an acceptable threshold when feasible for fairness, while satisfying all patient care requirements. We seek to focus on total waiting time to promote system efficiency.

### 3.1.2 EVS to IU dirty-bed assignment

Upon the discharge of a patient from an inpatient unit, the vacant bed must be cleaned for succeeding patients by EVS staff. EVS is the department responsible for cleaning all sites inside the hospital, including inpatient rooms, emergency beds, hallways, etc. For the planning horizon, the attributes of each EVS staff $(k \in K)$ include shift schedule, availability time for task assignment (staff member could be in the midst of completing a task), service time needed for cleaning a bed, and the initial staff member location which affects the travel time to an assigned dirty bed. The goal of assigning EVS staff to dirty beds is to provide clean beds as early as possible according to their priority in order to optimize resource utilization. First priority is given to cleaning dirty-beds assigned to admitted patients. Dirty beds that are not assigned to any of the currently admitted patients are considered as second priority and are only cleaned if there is available EVS staff capacity. The aim of considering the cleaning of additional dirty beds, albeit not immediately needed, is to maximize the utilization of EVS staff and ready beds for future patients. We consider allocation of multiple tasks for each EVS resource through the time horizon. In particular, our model schedules bed cleaning tasks sequentially for each EVS staff member during the planning horizon (i.e., each staff member might be given a series of cleaning tasks) in order to improve resource utilization. The task assignments should account for workload equity among EVS staff and shift schedules.

### 3.1.3 Patient to transport staff assignment

Another task in patient hospitalization service is transporting patient from ED (initial location of patient) to the assigned clean bed in a particular IU. In order to minimize patient's waiting time, it is important to optimize transport start time and assignment of transport staff ($j \in J$) to patient ($p \in P$). Similar to EVS staff assignments, the future tasks of each transport staff are determined based on the number of patients needing transport within the planning horizon (i.e., each transport staff maybe assigned to more than one patient transfer during the planning horizon). The task assignments should also account for workload equity among transport staff and shift schedules.

These three sets of service and task assignments need to be coordinated using the real-time information available to improve the flow and utilization performance of the ED-to-IU network. Our approach aims to improve the system level performance by optimizing the collective wait times of the patients, lead-times for bed turnaround, and utilization of all the resources. The underlying premise of the proposed modeling and solution approach is that, by simultaneously accounting for the needs of multiple boarded ED patients and the different support services within the ED-to-IU network, integrated task assignment and resource allocation can significantly improve patient wait times, and in turn, patient satisfaction and health outcomes.

### 3.2 Coordination Model Formulation

The coordinated assignment problem is formulated as a mixed-integer linear program. The decision variables utilized in the optimization model are presented in Table 2.

Clearly, there are multiple objectives to be considered in carrying out the task assignments to streamline patient flow across the ED-to-IU network. In consideration for computational efficiency requirements (i.e., the need to execute the formulation within a reasonable and practical time frame) and fairness considerations (account for objectives of the different stakeholders), we take a weighted sum approach for optimization. In particular, the objective function in (3.1a) and (3.1b) minimizes a weighted sum of the total waiting time across all admitted patients considered for the planning horizon, total sojourn

Table 2: Variables in the optimization model

| Variables | |
|---|---|
| **Assignment variables** | |
| $x_{ip}$ | 1 if bed $i$ is assigned to patient $p$, 0 otherwise |
| $y_{jdp}$ | 1 if transporter $j$ for its $d^{th}$ service is assigned to patient $p$, 0 otherwise |
| $z_{kdi}$ | 1 if EVS staff $k$ for its $d^{th}$ service is assigned to dirty bed $i$, 0 otherwise |
| **Time variables** | |
| $t_i^c$ | time that clean bed $i$ becomes available (status clean) |
| $t_{kd}^e$ | time that EVS staff $k$ becomes available for its $d^{th}$ service |
| $t_{jd}^t$ | time that transporter $j$ becomes available for its $d^{th}$ service |
| $T_p$ | time that patient $p$ is served (transported to a clean bed) |
| $T^{pmax}$ | maximum patient waiting time $T_p$ for the run |
| $t_p^b$ | time that a bed is ready for patient $p$ |
| $s_p^t$ | service time for transporting patient $p$ |
| **Indicator variables** | |
| $\theta_{ip}^s$ | 1 if patient $p$ is assigned to bed $i$ in previous and current runs, 0 otherwise |
| $\delta_r$ | 1 if all patients in room $r$ are male, 0 if all females |
| **Penalty variables** | |
| $\delta_p$ | penalty for assigning patient p to a non-preferred IU |
| $\theta^S$ | total number of differing assignments from the previous model run |

times for beds in dirty status, and penalty terms for violating overflow constraints and maximum patient boarding time experienced. Since the magnitudes of the terms in the objective function are different, we use weights ($\beta_1$, $\beta_2$, $\beta_3$) in appropriate values to reflect the relative importance of the different cost components.

$$Min \quad \sum_{p \in P} w_p(T_p - t_p) \tag{3.1a}$$

$$Min \quad \beta_1 \sum_{i \in I_{dirty}} (t_i^c - T^{now}) + \beta_2 \sum_{p \in P} \delta_p + \beta_3 T^{pmax} \tag{3.1b}$$

The resources needed for a bed cleaning task are an EVS staff and an unoccupied dirty bed. Hence, the cleaning task of a dirty bed is completed upon the EVS staff $k$ available at $t_{kd}^e$ travels from his current position in emergency department or inpatient unit to the dirty bed $i$ and cleans the bed for $s_e^b$ time units, as formulated in constraints in (3.2) and (3.3). If it is the first cleaning task assignment for EVS staff $k$, he travels from a central location, which takes $R_{ki}$ time units as in constraints in (3.2). Otherwise, the travelling time is determined based on the location of the dirty bed cleaned in the previous assignment, as in constraints in (3.3). Additionally, EVS staff works in shifts and the constraints in (3.4) guarantee that the cleaning task of an EVS staff cannot go beyond the end of his shift. The parameter $t_i^b$ stores the input data for the time a bed becomes available. We define the availability of a dirty bed as the time the patient occupying the bed is discharged. Constraints in (3.5) imply that the cleaning task of dirty bed $i$ cannot start sooner than the patient occupying bed $i$ is discharged. For a clean bed, parameter $t_i^b$ represents the time the bed $i$ is available and clean, as utilized in constraints in (3.6).

$$t_i^c \geq t_{kd}^e + s_e^b + R_{ki} - (1 - z_{kdi})M, \quad \forall i \in I_{dirty}, k \in K, d = 1 \tag{3.2}$$

$$t_{i_2}^c \geq t_{kd}^e + s_e^b + \Big( \sum_{i_1 \in I_{dirty}} R_{i_1 i_2} z_{k(d-1)i_1} \Big) - (1 - z_{kdi_2})M, \quad \forall i_1, i_2 \in I_{dirty}, k \in K, d \geq 2 \tag{3.3}$$

$$t_i^c - (1 - z_{kdi})M \leq t_k^{ES}, \quad \forall i \in I_{dirty}, k \in K, d \leq D^k \tag{3.4}$$

$$t_i^c \geq t_i^b + s_e^b + \Big(1 - \sum_{k \in K} \sum_{d \leq D^k} z_{kdi}\Big)M, \quad \forall i \in I_{dirty} \tag{3.5}$$

$$t_i^c \geq t_i^b, \quad \forall i \in I_{clean} \tag{3.6}$$

Similar to the cleaning task resource requirements, the necessary resources, namely, a clean bed from the appropriate inpatient unit and a transport staff, should be available for transferring an admitted patient from emergency department to the assigned inpatient unit. In particular, the transfer task cannot be initiated until the assigned transport staff becomes available, as enforced in constraints in (3.7). The transport time from emergency department depends on the assigned inpatient unit and is defined via constraints in (3.8). Constraints in (3.9) enforce that the assigned bed should be clean and ready by the time the patient arrives, where the time that assigned bed becomes ready is formulated in (3.10). The constraints in (3.11) guarantee that the admitted patient is ready for transfer before the task starts. The last term in the objective function in (3.1b) is to minimize the time that the last patient is transferred to the assigned inpatient unit, $T^{pmax}$, which is defined via constraints in (3.12).

$$T_p \geq t^t_{jd} + s^t_p - (1 - y_{jdp})M, \quad \forall j \in J, \forall p \in P, d = 1, ..., D^j \tag{3.7}$$

$$s^t_p = \sum_{i \in I} R_{ip} x_{ip}, \quad \forall p \in P \tag{3.8}$$

$$T_p \geq t^b_p, \quad \forall p \in P \tag{3.9}$$

$$t^b_p \geq t^c_i - (1 - x_{ip})M, \quad \forall i \in I, \forall p \in P \tag{3.10}$$

$$T_p \geq t_p + s^t_p, \quad \forall p \in P \tag{3.11}$$

$$T^{pmax} \geq T_p, \quad \forall p \in P \tag{3.12}$$

The EVS and transport staff may be assigned to multiple non-overlapping tasks throughout the planning horizon. Since the order of assigned tasks are important, we introduce the index $d$ to represent the $d^{th}$ service assignment. The following constraints define the time the EVS staff $k$ / transporter $j$ becomes available for $d^{th}$ service, for $d \geq 2$. In particular, constraints in (3.13) enforce that the EVS staff becomes available for $d^{th}$ cleaning service after completing the $(d-1)^{th}$ bed cleaning assignment, if there is one. Similarly, constraints in (3.14) enforce earliest availability for transporter $j$ as $s^t_p$ time units after completing the transportation of $(d-1)^{th}$ patient, where $s^t_p$ is the travelling time from inpatient unit of patient $p$ to emergency department. Constraints in (3.15) and (3.16) guarantee simple time sequence relation between two consecutive task commencements respectively for EVS and transport staff, i.e., $d^{th}$ task is not started before $(d-1)^{th}$ task.

$$t^e_{kd} \geq t^c_i - (1 - z_{k(d-1)i})M, \quad \forall i \in I, \forall k \in K, d = 2, ..., D^k \tag{3.13}$$

$$t^t_{jd} \geq T_p + s^t_p - (1 - y_{j(d-1)p})M, \quad \forall p \in P, \forall j \in J, d = 2, ..., D^j \tag{3.14}$$

$$t^e_{kd} \geq t^e_{k(d-1)}, \quad \forall k \in K, d = 2, ..., D^k \tag{3.15}$$

$$t^t_{jd} \geq t^t_{j(d-1)}, \quad \forall j \in J, d = 2, ..., D^j \tag{3.16}$$

Next, we continue with the assignment restrictions. Constraints in (3.17) and (3.18) enforce that each patient is assigned a bed and a transport staff. We ensure that a bed is not assigned to more than one patient and one cleaning task via constraints in (3.19) and (3.20), respectively. Constraints in (3.21) enforce that a dirty bed must be cleaned if it is assigned to a patient. A transport staff and EVS staff can not be assigned more than one patient transfer and bed cleaning duty, respectively, at any service task, as formulated in (3.22) and (3.23). Constraints in (3.24) and (3.25) impose the restriction that a transport and an EVS staff will not be assigned to $d^{th}$ service task (transfer or bed cleaning), unless he is assigned to $(d-1)^{th}$ service task. Constraint in (3.26) implies that the clean beds are not assigned to EVS staff.

$$\sum_{i \in I} x_{ip} = 1, \quad \forall p \in P \tag{3.17}$$

$$\sum_{j \in J} \sum_{d \leq D^k} y_{jdp} = 1, \quad \forall p \in P \tag{3.18}$$

$$\sum_{p \in P} x_{ip} \leq 1, \quad \forall i \in I \tag{3.19}$$

$$\sum_{k \in K} \sum_{d \leq D^k} z_{kdi} \leq 1, \quad \forall i \in I \tag{3.20}$$

$$\sum_{p \in P} x_{ip} \leq \sum_{k \in K} \sum_{d \leq D^k} z_{kdi}, \quad \forall i \in I_{dirty} \tag{3.21}$$

$$\sum_{p \in P} y_{jdp} \leq 1, \quad \forall j \in J, d = 1, ..., D^j \tag{3.22}$$

$$\sum_{i \in I} z_{kdi} \leq 1, \quad \forall k \in K, d = 1, ..., D^k \tag{3.23}$$

$$\sum_{p \in P} y_{jdp} \leq \sum_{p \in P} y_{j(d-1)p}, \quad \forall j \in J, d = 1, ..., D^j \tag{3.24}$$

$$\sum_{i \in I} z_{kdi} \leq \sum_{i \in I} z_{k(d-1)i}, \quad \forall k \in K, d = 1, ..., D^k \tag{3.25}$$

$$\sum_{k \in K} \sum_{d \leq D^k} \sum_{i \in I_{clean}} z_{kdi} = 0 \tag{3.26}$$

Constraint in (3.27) enforces that at least $\gamma$ dirty beds are assigned to EVS staff. At the first iteration, $\gamma$ is set to zero and we increment it by one at each successive iteration. The iterative optimization approach is discussed in detail in Section 3.3.2.

$$\sum_{k \in K} \sum_{d \leq D^k} \sum_{i \in I} z_{kdi} \geq \gamma \tag{3.27}$$

### 3.2.1 Maintaining decision continuity between runs

The dynamic task assignment and scheduling problem is optimized by sequentially solving instances defined via a sliding planning horizon window. Task interruptions between two consecutive instances are not allowed, i.e., if a cleaning or transporting task is started in an instance and in progress in the consecutive one, the task will be completed without any interruptions. To further stabilize the solutions from one instance to the next, we require that a fraction of patients common to both instances are assigned to the same beds. The constraints in (3.28) indicate whether a patient is assigned to the same bed in two consecutive instances through the variables $\theta^s_{ip}$. Total number of such assignments are formulated in (3.29). Constraint in (3.30) enforces that a specified fraction of patient to bed assignments do not change between any two consecutive runs.

$$2\theta^s_{ip} \leq \bar{x}_{ip} + x_{ip} \leq \theta^s_{ip} + 1, \quad \forall i \in \bar{I}, \forall p \in \bar{P} \tag{3.28}$$

$$\theta^S = \sum_{i \in \bar{I}} \sum_{p \in \bar{P}} \theta^s_{ip} \tag{3.29}$$

$$\alpha(\sum_{i \in \bar{I}} \sum_{p \in \bar{P}} \bar{x}_{ip}) \leq \theta^S \tag{3.30}$$

### 3.2.2 Special considerations

We consider several real-world limitations in our coordination model, including patient and bed isolation, overflow between inpatient units, and gender matching. Constraints in (3.31) ensure that each patient who needs isolation is assigned a room in isolation. Furthermore, a patient requiring isolation is placed only in an empty room, other beds in that room are blocked and not assigned to any patient via constraints in (3.32).

$$\sum_{i \in I_O} x_{ip} = 1, \quad \forall p \in P_O \tag{3.31}$$

$$n_r + \sum_{i \in I_r} \sum_{\substack{p' \in P \\ p' \neq p}} x_{ip'} \leq M(1 - \sum_{i \in I_r} x_{ip}), \quad \forall p \in P_O, \forall r \in R \tag{3.32}$$

An admitted patient is assigned to a *preferred* medical specialty unit according to the needs of the patient. Nonetheless, a bed in other suitable inpatient units could also be assigned to the patient provided that all the beds in the preferred specialty unit are unavailable. This is referred to as “overflow" between the IUs. Constraints in (3.33) ensure that patients are placed in the most preferred units, as specified by the bed request. Otherwise, penalties are incurred for assigning patients to non-preferred IUs. For determining of this penalty, parameter $h_{ip}$ is defined for each bed-patient assignment type. Furthermore, these constraints ensure that patients are placed in a unit that provides a level of care that is more than the minimum level specified for the patient.

$$\sum_{i \in I} h_{ip} x_{ip} + \delta_p (1 - C_p) \geq H_p, \quad \forall p \in P \tag{3.33}$$

Gender matching requirements formulated via constraints in (3.34)-(3.37) ensure that all patients in a room have the same gender. Constraints in (3.34) and (3.35) imply that all the newly assigned patients to room $r$ must be male or female, respectively. Constraints in (3.36) and (3.37) enforce that a newly assigned, respectively, male and female patient to a bed in a room where there is occupancy by the opposite gender has to wait to be transported to the bed until the patients currently in the room are discharged.

$$\sum_{i \in I_r} \sum_{p \in P} x_{ip} g_p \geq \sum_{i \in I_r} \sum_{p \in P} x_{ip} - (1 - \delta_r) M, \quad \forall r \in R \tag{3.34}$$

$$\sum_{i \in I_r} \sum_{p \in P} x_{ip} (1 - g_p) \geq \sum_{i \in I_r} \sum_{p \in P} x_{ip} - \delta_r M, \quad \forall r \in R \tag{3.35}$$

$$G_i^F t_i^b \leq T_p + (1 - x_{jp} g_p) M, \quad \forall r \in R, \forall p \in P, \forall i, j \in I_r, i \neq j \tag{3.36}$$

$$G_i^M t_i^b \leq T_p + (1 - x_{jp} (1 - g_p)) M, \quad \forall r \in R, \forall p \in P, \forall i, j \in I_r, i \neq j, \tag{3.37}$$

### 3.3 Solution Approach

The solution approach is designed to address the need to dynamically coordinate resources within an evolving ED-to-IU network environment. We assume that ED patients are either discharged or admitted to the hospital without any anticipation (unplanned patient admissions). Our coordination model is dynamic, in the sense that resource to task assignments are updated in real-time whenever changes in the system, *epochal events*, indicate the existence of a better assignment solution. Herein we delineate the steps of the proposed solution approach by introducing the flowchart in Figure 4. The algorithm illustrates the procedure of our approach following an epochal event. When an epochal event occurs, the real-time information related to model parameters (given in Table 1) are updated. There are two types of epochal events handled by this method: 1) new patient arrival, 2) new available resource (an occupied bed becoming available upon discharge of a patient, EVS staff, transport staff). We first rely on pre-processing, and then the optimization model is executed to make new assignments and revise any prior assignments for enhanced ED-to-IU network flow.

### 3.3.1 Pre-processing

In the pre-processing step, we apply several techniques to reduce the size of the MIP problem to improve computational tractability. Specifically, we reduce the number of resources entertained during the model execution based on time horizon and resource availability time such that solution quality is not compromised. For example, if a bed is not suitable for any of the requests, we remove that bed from the optimization model. Additionally, pre-processing step limits the maximum number of tasks (cleaning, transporting) assigned to each resource (EVS staff, transporter) during the time horizon. For instance, if planning time horizon is 300 minutes (5 hours) and average bed cleaning time is 50 minutes, the maximum possible number of cleaning tasks for each EVS staff member that can be completed during the planning horizon is limited to 6. Furthermore, we specify symmetry cases (e.g. patients with similar attributes) and reduce the feasible solution region by defining constraints that ensure patients with the same attributes including gender and IU type are served in order of their admission times. Moreover, we utilize the solution generated by FCFS method as an initial solution for the optimization model.

Based on the number of tasks involved and their mix, the level of improvement from pre-processing varies. On average, the pre-processing step significantly improved the computational time for solving problem instances in our numerical experiments. For instance, for a problem with 6 patients in the queue, 4 EVS staff and 2 transporters, we reached an optimality gap of 0.3% after 15 minutes, while implementing the pre-processing step improved the optimality gap to 0.04% within 9.36 seconds. In another example with 8 patients in the queue, 2 EVS staff and 2 transporters, the optimality gap improved from 13% to 0.3% when the pre-processing step was implemented.

### 3.3.2 Iterative optimization

Our primary objective is to minimize the patient boarding, while we also seek to increase resource (bed, EVS staff, transporter) utilization. We follow an iterative optimization approach to prioritize reducing patient boarding over increasing resource utilization and to reduce the computational complexity. In the initial iteration, we use an aggregate objective function formulated as the summation of (3.1a) and (3.1b), where priority is given to minimizing patient waiting time by relaxing constraints in (3.27) through setting $\gamma = 0$. Let the function value of (3.1a) at the optimal solution be represented with $F_0$. In each subsequent iteration, we restrict total patient waiting time by the solution obtained at the first iteration, i.e., $\sum_p w_p(T_p - t_p) \leq F_0$, set the objective function to (3.1b) and increment $\gamma$ by one to maximize resource utilization. The iterative approach is continued until one of the following termination conditions is satisfied: 1) we reach maximum number of cleaning tasks or 2) the problem becomes infeasible. Utilizing the final feasible iteration of the current instance, the system information (e.g. $T^{now}$, planning horizon, re-

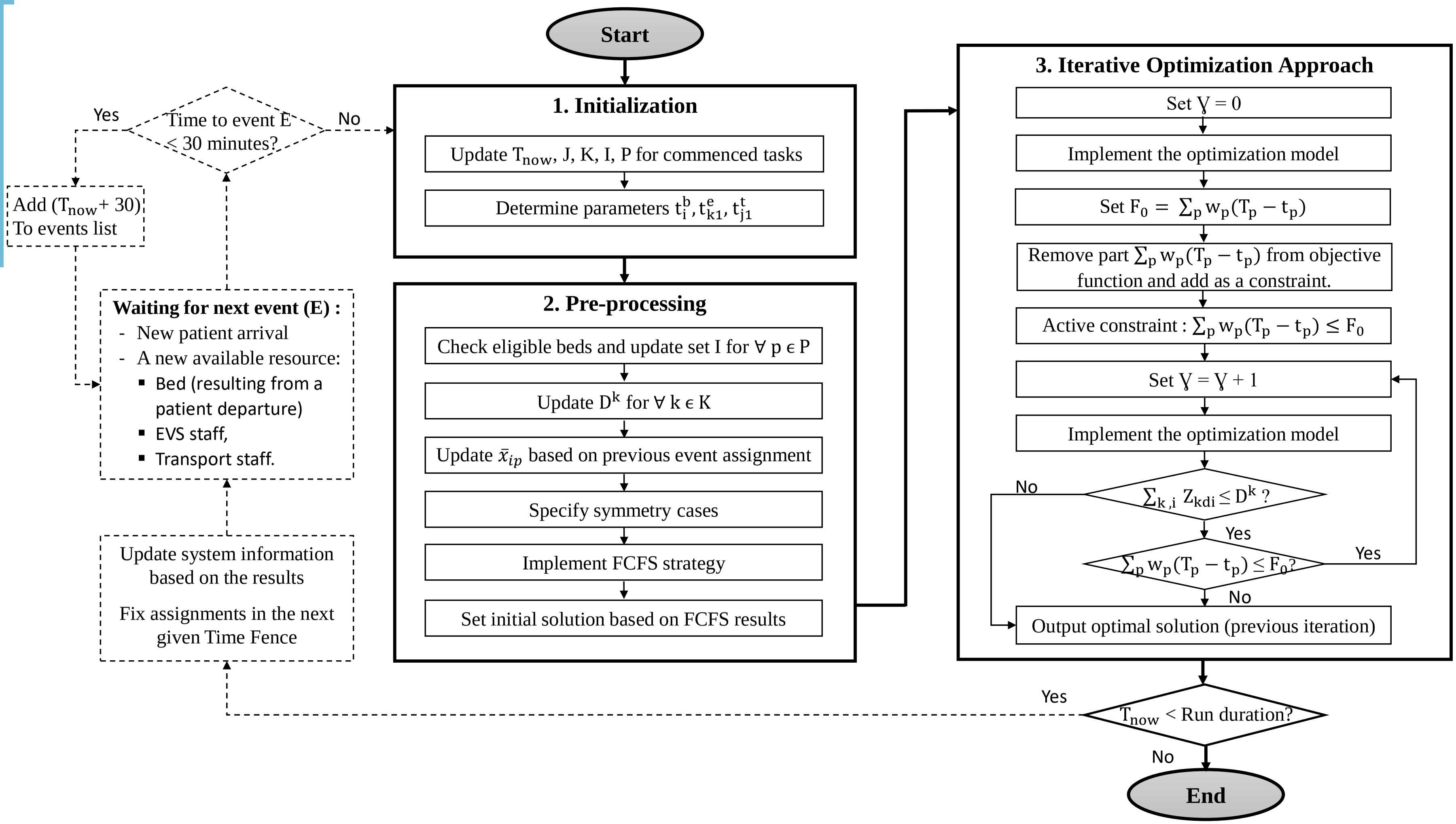



Figure 4: Flowchart for proposed solution approach

sources, patients) is updated. During the update, assignments corresponding to cleaning and transportation tasks starting within a defined time fence are fixed. Further details of the solution approach are given in Figure 2.

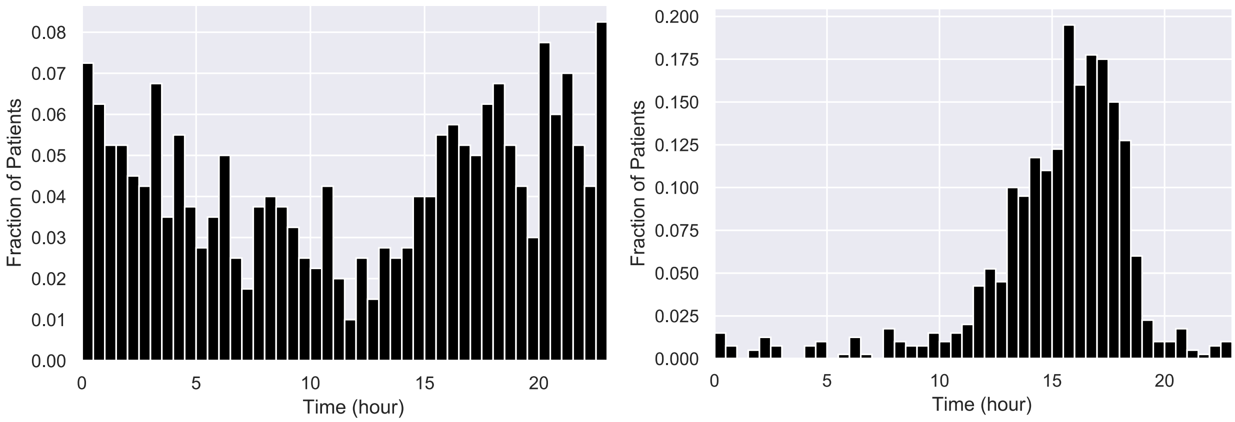
### 3.4 Computational Experiments

In this section, we conduct a computational study to demonstrate the effectiveness of the proposed approach for solving resource and task assignment problem using a test-bed of random instances generated based on data from Henry Ford Hospital (HFH) in Detroit, Michigan, a leading level-1 trauma center in Southeast Michigan. We compare the effectiveness of the proposed approach with the First-Come, First-Served (FCFS) method, prevalent in hospitals. The performance of each approach is evaluated in terms of patient waiting times and resource utilizations.

We tested our approach on a 3.10 GHz desktop with 16 GB of RAM under Windows OS with optimization solver Gurobi. We used maximum computational time of 15 CPU minutes and optimality gap of 0.1% as termination criteria. Given the time between any two consecutive optimization runs is at least 30 minutes, the 15-minute limit does not create any issue for resource planning. In our numerical experiments, 70% of the random instances were solved to the desired optimality within 3.35 CPU minutes, on average. The remaining instances were terminated prematurely with an average optimality gap of 23%.

#### 3.4.1 Random Problem Instance Generation

We generated a random problem instance based on admission and service data statistics from HFH to evaluate the proposed approach for deployment in real-world cases. The problem instance is run for 20 consecutive days of operations with the network initialized with all clean beds and all other resources being available. To eliminate the transient effect




(a) Patient Admission Times (b) Patient Discharge Times

Figure 5: Distributions of Patient Admission and Discharge Times

of initialization, we remove the results of the first 3 days and last 2 days and only consider results from the day 4 to 19 (namely, day 1 to day 15) for analysis and comparison.

The historical admission and discharge data from HFH covers the period from May 1, 2014 to December 15, 2016 and constitutes 243,745 ED visits and 41,942 IU admissions. Through this period, 10.6 patients arrive at the ED per hour and the admission rate is around 17%, on average. However, admissions and discharges are highly variable during the day, creating a challenging coordination problem. While the admission rates are higher in the evening and at night relative to mid-day, most of the discharges are observed in the afternoon. Utilizing historical data from HFH, we generated patient admission and discharge time distributions (by time of day) to improve the quality of our experiments, as displayed in Figures 5a and 5b. Consistent with HFH data, we assume patient admission likelihood to be the same for male and female patients. We consider the average length of stay of patients in IUs to be 2 days. Then, discharge times of patients from IUs are generated using Figure 5b.

HFH typically admits 40 patients per day with significant variability between days of the week. In our computational analysis, the number of admitted patients is varied from day to day using a Poisson distribution with a mean of 40. The number of admitted patients for the 15 days of interest are: 38, 42, 44, 43, 41, 44, 39, 45, 30, 28, 38, 40, 46, 37, and 46 respectively. This pattern has five consecutive days of excessive admissions from days 2 to 6 (varying from a minimum of 41 admissions to two days with 44 admissions) and we also see some sporadic peaks towards the end of the run.

In the study, we set the number of inpatient beds to 100, and consistent with HFH, half of the rooms have two beds and the other half have a single bed (34 beds are in single rooms and the rest in double rooms). The 100 beds are distributed across three typical types of IUs representing three levels of intensity of required care. At HFH, three main types of IUs are considered as general practice unit (GPU), telemetry unit (TU), and intensive care unit (ICU). According to HFH statistics, the GPU has more admissions than TU and ICU, and the admissions to TU and ICU are similar. By analysis of HFH historical data, we define the admission probability to the $IU_1$ (representing GPU) as 0.6 and consider the same probability for $IU_2$ and $IU_3$ (representing TU and ICU units, respectively) as 0.2 for an admitted ED patient. Also, we assume the same ratio for bed distribution as 60%, 20%, and 20% for $IU_1$, $IU_2$ and $IU_3$, respectively.

The average EVS and transport staff travel times from ED to $IU_1$, $IU_2$, $IU_3$ are set to be [5, 15, 25] minutes, respectively. Additionally, the average travel times between two IUs are [10, 20, 10] minutes for $IU_1$ and $IU_2$, $IU_1$ and $IU_3$, $IU_2$ and $IU_3$. Consistent with HFH, in all our experiments, the average duration for the bed cleaning task by EVS staff is assumed to be 50 minutes. We consider two EVS staff members and two transporters in each of the 8-hour shifts and the planning horizon for each coordination optimization is reasonably set to be 300 minutes (5 hours). The rest of the parameters are reported in

Table 3: Parameter values utilized in the experimental study

| Parameters | Values |
|---|---|
| $s_e^b$ | 50 minutes |
| $h_{ip}$ | $\begin{array}{cc} & p \text{ with preferred medical unit} \\ & \begin{array}{ccc} \text{IU}_1 & \text{IU}_2 & \text{IU}_3 \end{array} \\ \begin{array}{c} i \in \text{IU}_1 \\ i \in \text{IU}_2 \\ i \in \text{IU}_3 \end{array} & \begin{bmatrix} 1 & -M & -M \\ 0.3 & 1 & -M \\ 0.1 & 0.8 & 1 \end{bmatrix} \end{array}$ |
| $C_p$ | 0 for all patients |
| $H_p$ | 1 for all patients |
| $\alpha$ | 30% |
| $\beta_1$ | 0.01 |
| $\beta_2$ | 10 |
| $\beta_3$ | 0.01 |

Table 3.

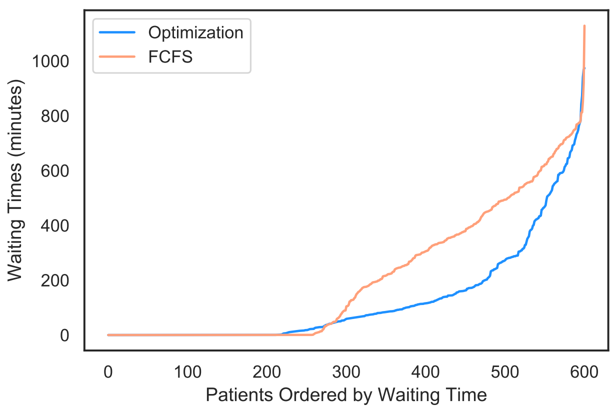
### 3.4.2 First-Come First-Served (FCFS) Approach

In the FCFS approach, the primary objective is to serve the currently admitted patients at the earliest, whereas the secondary objective is to increase resource utilization by cleaning extra dirty beds, as in the proposed optimization approach. First, patients are assigned to beds in ascending order of patient admission times, $t_p$, i.e., the patient admitted ahead of others is assigned, by order of preference, a clean bed, a dirty bed, or an occupied bed with earliest availability. Next, the dirty beds that are assigned to patients in the first step are cleaned in ascending order of bed availability times, $t_i^b$. Then, patients are assigned to transport staff in ascending order of the times that beds are clean and ready for patients, $t_p^b$. Lastly, dirty beds that are not currently assigned to any patient are cleaned in ascending order of bed availability times, if the cleaning task is not going to worsen any of the currently admitted patients' services. After all possible EVS staff - bed assignments are identified, the system information is updated.

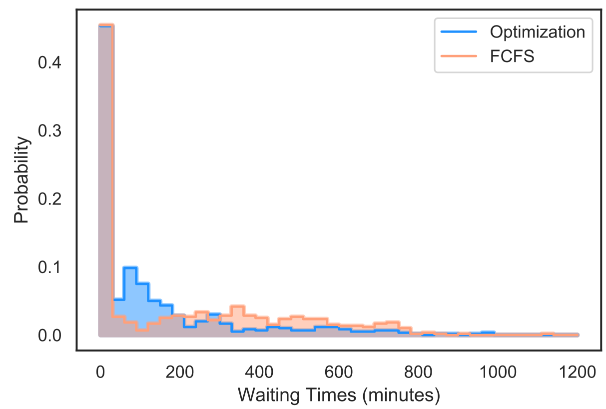
### 3.4.3 Results

We compare performance of the proposed coordination framework with FCFS benchmark method using a computational study. The comparison is based on patient waiting times (boarding), resource utilization, and equity based on performance statistics derived across each run spanning 15 consecutive days.

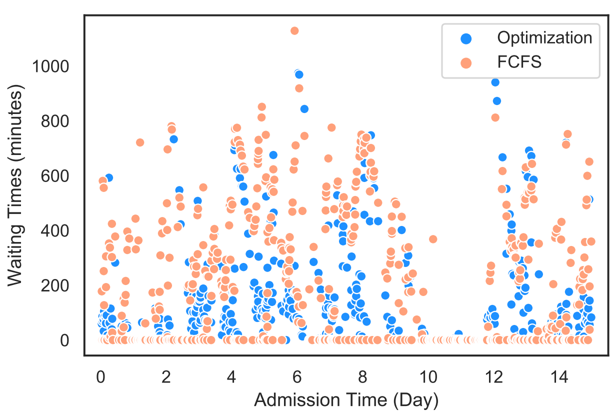
Figure 6 displays ED patient waiting times under the proposed coordination approach versus FCFS strategy considering various statistics. The results show that the proposed co-



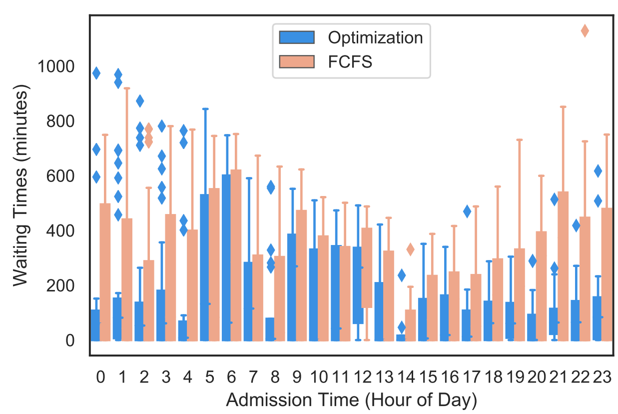

(a) Patient waiting times (sorted)




(b) Patient waiting time distribution




(c) Patient waiting times ordered by the admission time



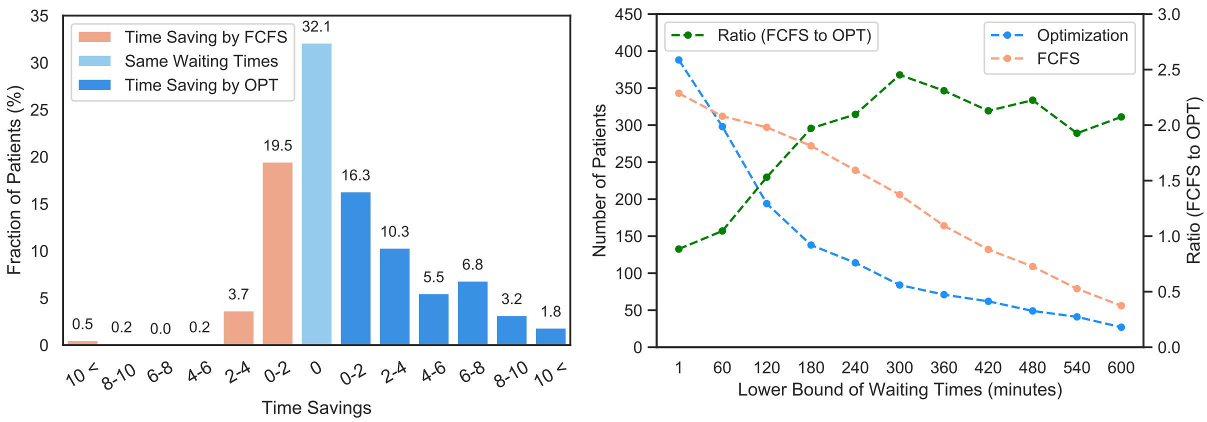

(d) Patient waiting time box plots through the day

Figure 6: Patient waiting time performance under the proposed coordination approach versus FCFS strategy

ordination approach led to significant improvement in patient waiting times and decreased frequency of patients with excessive waiting times. Although the range of patient waiting times are comparable for FCFS and coordination approach, there are remarkable differences between the distributions utilizing these two approaches. The performance profiles for FCFS and the proposed coordination approach are displayed in Figure 6a, where waiting times for all patients admitted during 15 days were sorted in non-decreasing order. As depicted by the first 211 data points in Figure 6a, both approaches served the patients without any boarding. For the next 69 data points, we observe that FCFS performed better than the proposed coordination approach and reduced the average waiting time by 11.8 minutes. For the following 315 patient data points, the coordination approach significantly outperformed FCFS, reducing patient boarding by 158.8 minutes, on average, where the maximum gap reached 255 minutes. Similar inference is deduced from Figure 6b, where the waiting times accumulated within the range [0, 200] using the coordination approach in comparison to a significantly dispersed distribution utilizing FCFS.

As displayed in Figure 5b, majority of the IU discharges occur in the afternoon, which generates a substantial demand for bed cleaning tasks. During this period, the bottleneck is primarily due to bed cleaning (EVS staff) rather than bed unavailability. On the other hand, we expect the bottleneck in the morning hours to be due to bed unavailability. This daily non-stationary behaviour of mismatch in supply and demand in IU beds creates a challenging resource management problem. Thus, we investigate performance of FCFS and coordination approach when waiting times are grouped by patient admission times. The scatter and box plots are displayed in Figures 6c and 6d, respectively. Both approaches demonstrated reduction in patient waiting in the afternoon in comparison to morning hours. However, the proposed coordination approach managed to keep the waiting times much lower and under control for a longer period, from 2 pm to 4 am, whereas FCFS approach quickly lost control and reached very long waiting times by 9 pm. Both approaches struggled from 5 am to 1 pm, where long waiting times are associated with bed unavailability which could be resolved by early morning discharges. While there are various studies focusing on improving daily discharge patterns to match supply and demand of IU beds, it is outside the scope of this research, and historical discharge patterns are utilized as input to our numerical study.

Figure 7a depicts the histogram of waiting time improvement for each patient utilizing the proposed coordination approach compared to FCFS. The computational results showed that 33.9% of patients benefited from the coordination approach, while 32.1% experienced the same waiting time under both approaches and 24.1% of patients had longer boarding under the coordination approach. The heavily right skewed distribution clearly shows that majority of patients were better served by the coordination approach, and over 81% of the remaining patients, who experienced longer waiting times with the proposed approach, had up to only two hours of additional boarding. Figure 7b reports the number of patients who experienced more than a specified waiting time utilizing the two approaches. The figure shows that more patients had less than 1-minute of boarding (practically, no boarding) using FCFS in relation to the proposed approach, whereas both approaches served comparable number of patients within one hour of their admissions. 50% more patients boarded longer than two hours when FCFS was utilized compared to the coordination approach. As the waiting time threshold was increased to three hours or more, the number of patients boarded longer than the threshold using FCFS became twice of those utilizing the




(a) Patient waiting time improvement (b) Frequency of patient waiting times

Figure 7: Performance of proposed approach in comparison with FCFS

proposed approach, demonstrating that the optimization approach decreased the number of patients experiencing relatively high waiting times.

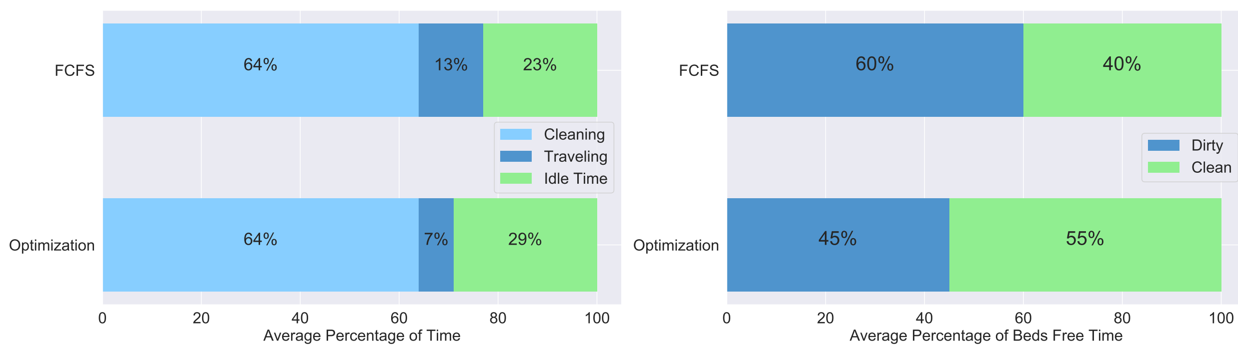
The boarding of a patient can be attributed to bed, EVS staff, and transporter unavailability. The time from patient admission decision until the transporter transfers the patient is defined as the total boarding. We further distribute the patient boarding to individual resources to compare the resource coordination under FCFS and the proposed optimization

Table 4: Patient Waiting Time Breakdown (in minutes)

| | | FCFS | Optimization | % Improvement |
|---|---|---|---|---|
| All patients ($n$=601) | Avg. waiting time | 211.1 | 129.3 | 38.7% |
| | Median waiting time | 89.7 | 59.1 | 34.1% |
| Boarded patients | # of patients | 345 | 390 | -13.0% |
| | Avg. waiting time | 367.7 | 199.3 | 45.8% |
| Delay due to transporter | # of patients experienced | 8 | 89 | -1012.5% |
| | Avg. waiting time | 5.8 | 15 | -158.6% |
| Delay due to bed preparation | # of patients experienced | 337 | 344 | -2.1% |
| | Avg. waiting time | 376.2 | 222 | 41.0% |
| Delay due to EVS staff | # of patients experienced | 281 | 256 | 8.9% |
| | Avg. waiting time | 239.9 | 71.9 | 70.0% |
| Delay due to bed unavailability | # of patients experienced | 143 | 140 | 2.1% |
| | Avg. waiting time | 367.6 | 320.8 | 12.7% |

approach. Keeping all the bed related decisions fixed, the maximum reduction in patient boarding by making the transporter available sooner is defined as transporter delay. The rest of the patient boarding is attributed to bed preparation delay. Similarly, the maximum reduction in bed preparation delay by making the EVS staff available sooner is defined as delay due to EVS staff. We define delay due to bed unavailability as the time between patient admission and the assigned bed becoming vacant. The rest of the bed preparation delay is due to physically cleaning the bed, and it is affected by unavailability of bed or EVS staff. Hence, it is not assigned to either resource.

The detailed statistics related to patient boarding is displayed in Table 4. The average boarding time with FCFS decreased by 38.7% from 211.1 minutes to 129.3 minutes when the proposed coordination approach was implemented. Furthermore, the median waiting time with FCFS decreased by 34.1% from 89.7 minutes to 59.1 minutes utilizing the coordination approach. When only boarded patients are considered, 345 patients experienced 367.7 minutes of average boarding using FCFS compared to 390 patients experiencing 199.3 minutes of average boarding with coordination approach, resulting a 45.8% reduction in average boarding. While FCFS chose to board fewer patients but for longer period of time, the proposed coordination approach chose to board only 13% more patients for significantly less period of time. A negligible portion of boarding can be attributed to transporter delay, where FCFS performed better than the coordination approach. The statistics show that the main reason for patient boarding is bed preparation delay. We observe comparable number of patients experiencing bed preparation delay with FCFS and the optimization approach, where the coordination approach reduced the average delay of 376.2 minutes with FCFS to 222 minutes, resulting a 41% decrease. The results demonstrate that the majority of improvement in patient boarding and bed preparation delay using optimization approach over the FCFS is due to EVS coordination. The optimization approach decreased the number of patients experiencing EVS delay from 281 to 256, corresponding to 8.9% reduction, and the average delay due to EVS from 239.9 to 71.9 minutes, resulting a 70% reduction, in comparison to FCFS. We also observe a decline in the number of patients waiting due to bed unavailability and average waiting time for a vacant bed, when optimization approach was utilized over the FCFS, while the improvement was less significant in comparison to the one with EVS.

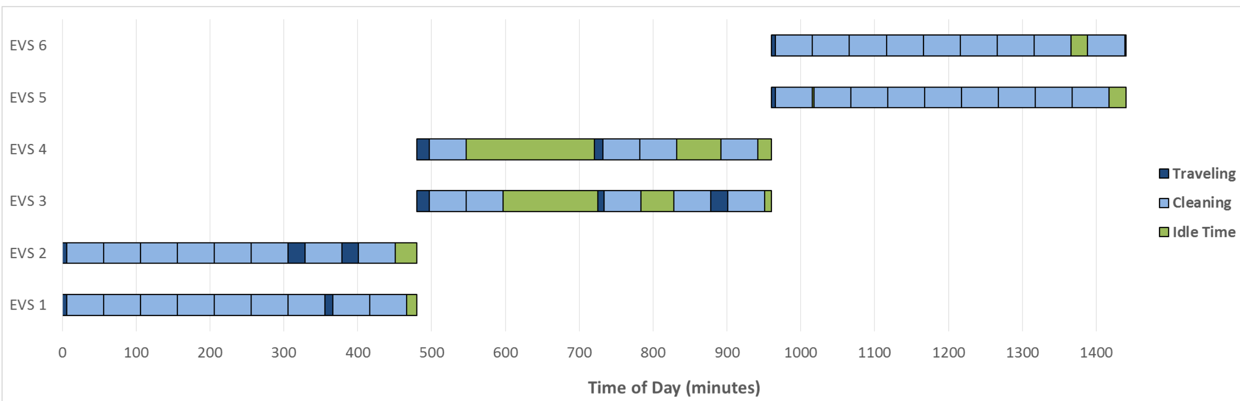



(a) EVS Staff - Average Utilization

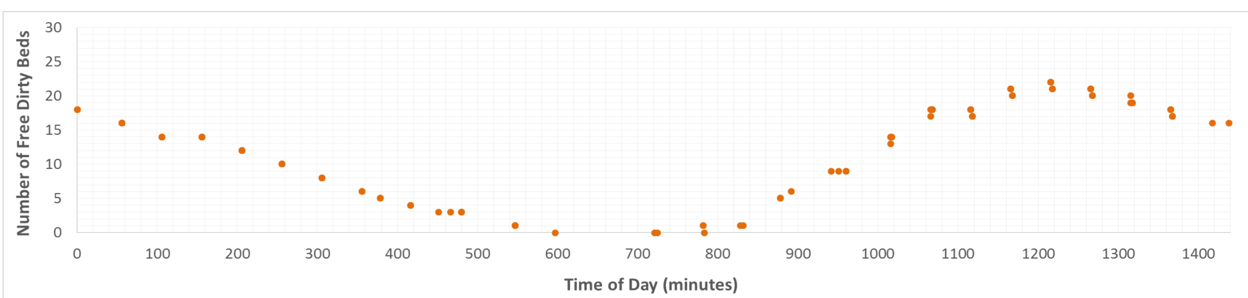
(b) Beds - Average Utilization

Figure 8: Resource utilization under proposed versus FCFS approaches

We also evaluated the utilization of EVS staff and IU beds under the two approaches. The results show that 39 dirty beds per day, on average, were cleaned utilizing either approach. However, the proposed coordination approach reduced the non-value added tasks (traveling between IUs) for EVS staff and increased the availability for other tasks such as area decontamination and hygiene management, as seen in Figure 8a. In order to fully understand the process, the cleaning, traveling, and idle times of all EVS staff through day 8 under the proposed approach and FCFS strategy are depicted in Figures 9 and 10. Furthermore, we analyze the proportion of time vacant IU beds are clean or dirty under FCFS and the coordination approach. As shown in Figure 8b, a vacant bed was dirty 60% of the time, on average, when FCFS approach was implemented, in comparison to 45% of the time, on average, with the proposed coordination approach. This confirms that the proposed approach increases effective utilization of the bed capacity by reducing the proportion of time that vacant beds are dirty.

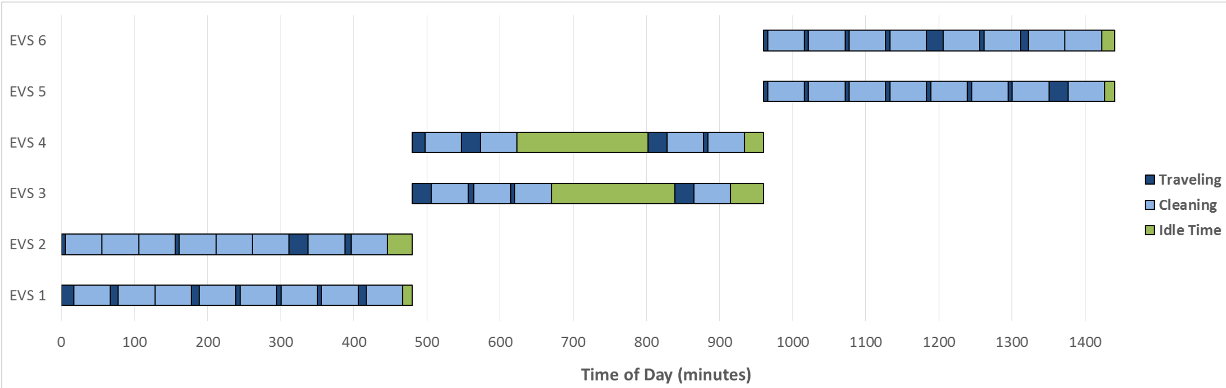



(a) EVS staff workload through the day

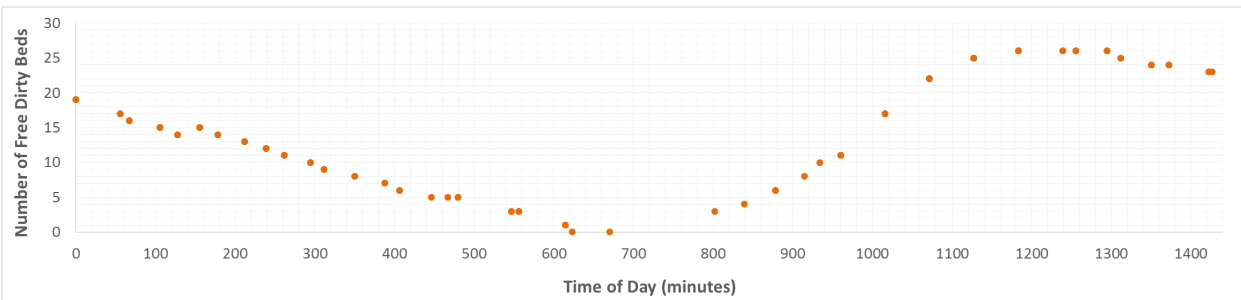



(b) Number of available dirty beds through the day

Figure 9: EVS staff utilization under the proposed coordination approach through day 8




(a) EVS staff workload through the day




(b) Number of available dirty beds through the day

Figure 10: EVS staff utilization under the FCFS strategy through day 8

# CHAPTER 4 INTEGRATED PROACTIVE COORDINATION APPROACH USING STOCHASTIC OPTIMIZATION

## 4.1 Problem Description

The common process of resource allocation within the ED-to-IU network for patients in hospitals starts after admission decision and bed request, as displayed in Figure 11. When the bed preparation is initiated only at the time of patient admission, it becomes a challenge to have the best use of limited resources such as EVS staff and transport staff. As discussed in previous chapter, the goal of assigning EVS staff to dirty beds is to provide clean beds as early as possible according to their priority in order to optimize resource utilization. First priority is given to cleaning dirty-beds assigned to admitted patients. Dirty beds that are not assigned to any of the currently admitted patients are considered as second priority. The motivation for cleaning additional dirty beds is to maximize the utilization of EVS staff and ready beds for future patients. In reactive approach, EVS staffs prepare inpatient beds according to the realized bed demand and there isn't any preference among the dirty beds (dirty beds that are not assigned to any admitted patients) for cleaning for future patients. Although reactive approach helps to reduce the misallocation of EVS staff to essential cleaning tasks, it is ineffective in reducing the delay related to the time needed to finish services associated with previous assignments for cleaning extra dirty beds. Our proposed proactive approach suggests utilizing reliable prediction of ED admission decisions ahead of the actual admission decisions to optimize the resource allocation in a proactive manner to reduce ED patients waiting times. We assume that when a new patient enters in emergency department and starts the testing and treatment process, the information about the patient, like as patient's health history, provides reliable estima-

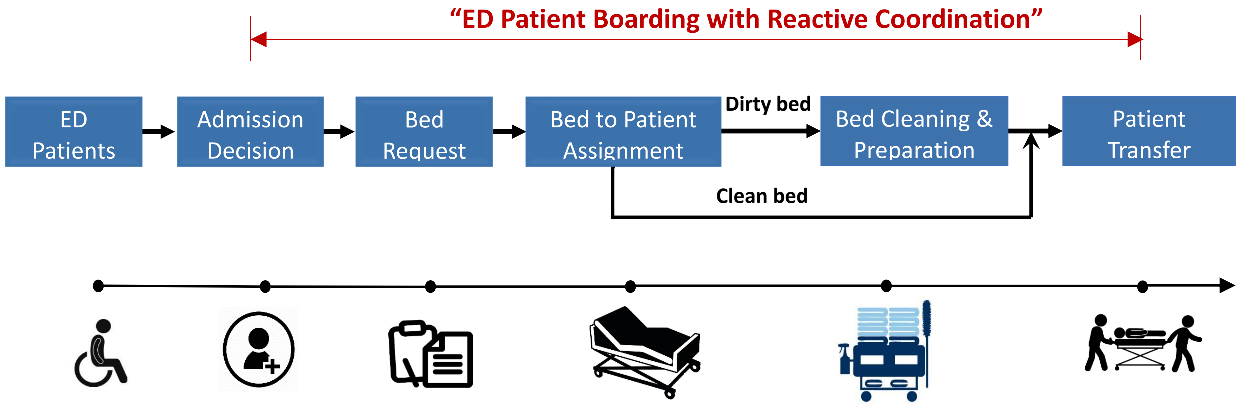



Figure 11: Reactive resource allocation approach across the ED-to-IU network

tion of disposition decision and admission time before the actual decision. This estimation helps to improve the allocation of resources to relevant tasks in a proactive manner regarding an impending demand. In this research, we propose a proactive coordination model to optimize resource allocation in ED-IU network in hospitals. In our framework, we capture real time information from ED and IUs to provide optimal coordination between three tasks: Bed Management, Bed Turnaround, and Transportation. The fundamental idea of the proposed modeling and solution approach is that, by considering impending admission and demands for resources, integrated task assignment and resource allocation can reduce ED boarding and, in turn, ED crowding.

### 4.2 Proactive Coordination Model Formulation

We have formulated the proactive resource coordination problem as a two-stage stochastic mathematical model [41, 38, 87]. We consider the reactive optimization model proposed in Chapter 4 as a base, and make necessary modifications to account for uncertainties. We developed two models for the problem, where the first model finds the best assignment of EVS staff to dirty beds by taking into account the ED patients that are waiting for admission decision as well as admitted patients. It uses the probability estimation for disposition decision of ED patients and the remaining length of stay estimation of patient in ED until admission decision. Therefore, we concern these two sources of uncertainty in the problem assumptions and formulation at stage one. These uncertainties decrease through the time since the growing information on the patient provides more accurate predictions for the patient. The optimal decision for this model gives the assignment of EVS staff to dirty beds. The second model builds on the result of the first model by considering only admitted patients. We use the assignment of EVS staff to dirty beds from stage one and availability time of clean beds as parameters for the stage two model. The optimal decision of this model shows the assignments of patients to beds and patients to transporters. Both models are formulated as mixed integer linear programs. The sets, parameters, and decision variables utilized in the first and second stage models are presented in Tables 5 and 6, respectively.

### 4.2.1 First Stage

The objective function of the first model considers the total waiting time of all patients across all scenarios and total sojourn times for beds in dirty status:

$$Min \quad \sum_{s\in S}\sum_{p\in P}(T_{ps}-t_{ps})+\beta_1\sum_{i\in I_{dirty}}(t_i^c-T^{now}) \tag{4.1}$$

The first set of constraints formulate time restrictions related to bed cleaning by EVS staff to find time that a dirty/clean bed becomes available. Constraints in (4.2) and (4.3) imply that cleaning service of bed i can start when EVS staff k is available after its previous service and the cleaning process takes $s_e^b$ units of time. If the EVS staff k is assigned to bed i for the first service, the traveling time of the staff for arriving to bed i is $R_{ki}$ as in

Table 5: Sets, indices, parameters, and variables utilized in the first step model

| | |
|---|---|
| **Sets** | |
| $I, J, K$ | resources (beds, transport staff, EVS staff, respectively) |
| $U$ | inpatient units (a unit contains one or more rooms) |
| $R$ | rooms (a room contains one or more beds) |
| $P$ | patients |
| $S$ | Scenarios |
| $P_M/P_F$ | male/female patients; $P_M \subseteq P, P_F \subseteq P$ |
| $I_r$ | beds belonging to room $r$; $r \in R$, $I_r \subseteq$ I |
| $I_{clean}/I_{dirty}$ | beds with status clean/dirty, beds in isolation rooms; $I_{dirty} \subseteq I, I_{clean} \subseteq I$ |
| **Indices** | |
| $i, j, k$ | an individual resource (bed, transport staff, EVS staff, respectively) |
| $p$ | patient |
| $s$ | scenario |
| $d$ | order of a task among a series of tasks |
| $u$ | inpatient unit |
| $r$ | room |
| clean/dirty | binary indices referring to status of a bed |
| **Parameters** | |
| $T_h$ | planning time horizon |
| $T^{now}$ | time at the beginning of a new planning horizon |
| $t_{ps}$ | time that the admit decision is made for patient $p$ in scenario $s$ |
| $t_i^b$ | time that a bed $i$ becomes vacant |
| $t_{k1}^e$ | time that EVS staff $k$ becomes available for its initial service |
| $t_k^{ES}$ | time that the shift ends for EVS staff $k$ |
| $t_{j1}^t$ | time that transporter $j$ becomes available for its initial service |
| $G_{is}^M/G_{is}^F$ | 1 if gender of patient currently occupying bed $i$ in scenario $s$ is male/female, 0 otherwise |
| $g_{sp}$ | 1 if gender of patient $p$ is male in scenario $s$, 0 if female |
| $s_e^b$ | service time for cleaning a bed |
| $R_{x,x'}$ | travel time between two locations ($R_{ip}$ is travel time for ED patient $p$ to IU bed $i$) |
| $\beta_1$ | penalty coefficient for total sojourn time of beds in dirty state in the objective function |
| $\gamma$ | least number of cleaning task assignments allowed in each run |
| $D^j, D^k$ | maximum number of assignments allowed for transporter $j$, EVS staff $k$ |
| **Variables** | |
| $x_{sip}$ | 1 if bed $i$ is assigned to patient $p$ in scenario $s$ , 0 otherwise |
| $y_{sjdp}$ | 1 if transporter $j$ for its $d^{th}$ service is assigned to patient $p$ in scenario $s$, 0 otherwise |
| $z_{kdi}$ | 1 if EVS staff $k$ for its $d^{th}$ service is assigned to dirty bed $i$, 0 otherwise |
| $t_i^c$ | time that clean bed $i$ becomes available (status clean) |
| $t_{kd}^e$ | time that EVS staff $k$ becomes available for its $d^{th}$ service |
| $t_{jds}^t$ | time that transporter $j$ becomes available for its $d^{th}$ service in scenario $s$ |
| $T_{ps}$ | time that patient $p$ is served (transported to a clean bed) in scenario $s$ |
| $t_{ps}^b$ | time that a bed is ready for patient $p$ in scenario $s$ |
| $s_{ps}^t$ | service time for transporting patient $p$ in scenario $s$ |
| $\delta_{sr}$ | 1 if all patients in room $r$ in scenario $s$ are male, 0 if all females |

constraint (4.2). Otherwise, the traveling time of EVS staff depends to the location of its $(d-1)^{th}$ service as in constraint (4.3). Also, we assume EVS staff works in shifts and the constraints in (4.4) enforce that the cleaning task of an EVS staff cannot go beyond the end of his shift. Constraints (4.5) restrict that cleaning of dirty bed i starts after bed gets empty and previous patient is discharged. Otherwise, if the bed is already clean, $t_i^c$ equals to the time that bed i is available and clean as in constraint (4.6).

$$t_i^c \geq t_{kd}^e + s_e^b + R_{ki} - (1 - z_{kdi})M, \quad \forall i \in I_{dirty}, k \in K, d = 1 \tag{4.2}$$

$$t_{i_2}^c \geq t_{kd}^e + s_e^b + \Big( \sum_{i_1 \in I_{dirty}} R_{i_1 i_2} z_{k(d-1)i_1} \Big) - (1 - z_{kdi_2})M, \quad \forall i_1, i_2 \in I_{dirty}, k \in K, d \geq 2 \tag{4.3}$$

$$t_i^c - (1 - z_{kdi})M \leq t_k^{ES}, \quad \forall i \in I_{dirty}, k \in K, d \leq D^k \tag{4.4}$$

$$t_i^c \geq t_i^b + s_e^b + \Big(1 - \sum_{k \in K} \sum_{d \leq D^k} z_{kdi}\Big)M, \quad \forall i \in I_{dirty} \tag{4.5}$$

$$t_i^c \geq t_i^b, \quad \forall i \in I_{clean} \tag{4.6}$$

Constraints in (4.7)-(4.11) show the time that patient p is transported to a clean bed in scenario s. These constraints enforce that patient arrives to the assigned bed after the bed is clean and ready. The transport time from emergency department depends on the assigned inpatient unit and is defined via constraints in (4.8).

$$T_{ps} \geq t_{jds}^t + s_{ps}^t - (1 - y_{sjdp})M, \quad \forall j \in J, \forall p \in P, \forall s \in S, d = 1, ..., D^j \tag{4.7}$$

$$s_{ps}^t = \sum_{i \in I} R_{ip} x_{sip}, \quad \forall p \in P, \forall s \in S \tag{4.8}$$

$$T_{ps} \geq t_{ps}^b, \quad \forall p \in P, \forall s \in S \tag{4.9}$$

$$t_{ps}^b \geq t_i^c - (1 - x_{sip})M, \quad \forall i \in I, \forall p \in P, \forall s \in S \tag{4.10}$$

$$T_{ps} \geq t_{ps} + s_{ps}^t, \quad \forall p \in P, \forall s \in S \tag{4.11}$$

Since the EVS staff can be assigned to multiple distinct cleaning services during the planning horizon, constraints in (4.12) determine the time that the EVS staff becomes available for $d^{th}$ service after finishing previous assigned task. Likewise, the transporters can be assigned to multiple separate tasks throughout the time horizon in each scenario. Constraints in (4.13) define the time that the transport staff becomes available for $d^{th}$ service after finishing previous task in each scenario. Constraints in (4.14) and (4.15) show that $d^{th}$ task is not started before $(d-1)^{th}$ task. These constraints ensure simple time sequence relation between two consecutive tasks for EVS and also for transport staff in each scenario.

$$t^e_{kd} \geq t^c_i - (1 - z_{k(d-1)i})M, \quad \forall i \in I, \forall k \in K, d = 2, ..., D^k \tag{4.12}$$

$$t^t_{jds} \geq T_{ps} + s^t_{ps} - (1 - y_{sj(d-1)p})M, \quad \forall p \in P, \forall s \in S, \forall j \in J, d = 2, ..., D^j \tag{4.13}$$

$$t^e_{kd} \geq t^e_{k(d-1)}, \quad \forall k \in K, d = 2, ..., D^k \tag{4.14}$$

$$t^t_{jds} \geq t^t_{j(d-1)s}, \quad \forall j \in J, d = 2, ..., D^j, \forall s \in S \tag{4.15}$$

Constraints in (4.16)-(4.25) are the assignment restrictions. These constrains enforce that one resource member (bed and transporter) can be assigned to at most one patient and each patient must be assigned to a bed and transporter in each scenario. Constraints in (4.20) and (4.21) state that at most one EVS staff is assigned to each dirty bed, and at most one dirty bed is assigned to an EVS member for their $d^{th}$ service. Constraints in (4.22) guarantees the dirty bed is assigned to one EVS staff for cleaning if the bed is assigned to a patient. Constraints in (4.23) and (4.24) ensure that a transporter and an EVS member are not assigned for their $d^{th}$ service if they are not assigned for $(d-1)^{th}$ service. Constraint

(4.25) implies that clean beds are not assigned to EVS staff.

$$\sum_{i \in I} x_{sip} = 1, \quad \forall p \in P, \forall s \in S \tag{4.16}$$

$$\sum_{j \in J} \sum_{d \leq D^j} y_{sjdp} = 1, \quad \forall p \in P, \forall s \in S \tag{4.17}$$

$$\sum_{p \in P} x_{sip} \leq 1, \quad \forall i \in I, \forall s \in S \tag{4.18}$$

$$\sum_{p \in P} y_{sjdp} \leq 1, \quad \forall j \in J, \forall s \in S, d = 1, ..., D^j \tag{4.19}$$

$$\sum_{k \in K} \sum_{d \leq D^k} z_{kdi} \leq 1, \quad \forall i \in I \tag{4.20}$$

$$\sum_{i \in I} z_{kdi} \leq 1, \quad \forall k \in K, d = 1, ..., D^k \tag{4.21}$$

$$\sum_{p \in P} x_{sip} \leq \sum_{k \in K} \sum_{d \leq D^k} z_{kdi}, \quad \forall i \in I_{dirty}, \forall s \in S \tag{4.22}$$

$$\sum_{p \in P} y_{sjdp} \leq \sum_{p \in P} y_{sj(d-1)p}, \quad \forall j \in J, \forall s \in S, d = 1, ..., D^j \tag{4.23}$$

$$\sum_{i \in I} z_{kdi} \leq \sum_{i \in I} z_{k(d-1)i}, \quad \forall k \in K, d = 1, ..., D^k \tag{4.24}$$

$$\sum_{k \in K} \sum_{d \leq D^k} \sum_{i \in I_{clean}} z_{kdi} = 0 \tag{4.25}$$

Constraints in (4.26)-(4.29) formulate gender restrictions that ensure all patients in a room have the same gender for each scenario. A newly assigned patient to a bed in a room where there is occupancy by the opposite gender has to wait to be transported to the bed until the current patients in the room are discharged.

$$\sum_{i \in I_r} \sum_{p \in P} x_{sip} g_{sp} \geq \sum_{i \in I_r} \sum_{p \in P} x_{sip} - (1 - \delta_{sr})M, \quad \forall r \in R, \forall s \in S \tag{4.26}$$

$$\sum_{i \in I_r} \sum_{p \in P} x_{sip}(1 - g_{sp}) \geq \sum_{i \in I_r} \sum_{p \in P} x_{sip} - \delta_{sr} M, \quad \forall r \in R, \forall s \in S \tag{4.27}$$

$$G_{is}^{F} t_i^b \leq T_{ps} + (1 - x_{sjp} g_{sp})M, \quad \forall r \in R, \forall p \in P, \forall s \in S, \forall i, j \in I_r, i \neq j \tag{4.28}$$

$$G_{is}^{M} t_i^b \leq T_{ps} + (1 - x_{sjp}(1 - g_{sp}))M, \quad \forall r \in R, \forall p \in P, \forall s \in S, \forall i, j \in I_r, i \neq j, \tag{4.29}$$

Table 6: Sets, indices, parameters, and variables utilized in the second step model

| | |
|---|---|
| **Sets** | |
| $I, J$ | resources (beds, transport staff, respectively) |
| $U$ | inpatient units (a unit contains one or more rooms) |
| $R$ | rooms (a room contains one or more beds) |
| $P$ | patients |
| $P_M/P_F$ | male/female patients; $P_M \subseteq P, P_F \subseteq P$ |
| $I_r$ | beds belonging to room $r$; $r \in R$, $I_r \subseteq$ I |
| **Indices** | |
| $i, j$ | an individual resource (bed, transport staff, respectively) |
| $p$ | patient |
| $d$ | order of a task among a series of tasks |
| $u$ | inpatient unit |
| $r$ | room |
| **Parameters** | |
| $t_p$ | time that the admit decision is made for patient $p$ |
| $t_{j1}^t$ | time that transporter $j$ becomes available for its initial service |
| $G_i^M/G_i^F$ | 1 if gender of patient currently occupying bed $i$ is male/female, 0 otherwise |
| $g_p$ | 1 if gender of patient $p$ is male, 0 if female |
| $R_{x,x'}$ | travel time between two locations ($R_{ip}$ is travel time for ED patient $p$ to IU bed $i$) |
| $\beta_2$ | penalty coefficient for maximum patient boarding time in the objective function |
| $D^j$ | maximum number of assignments allowed for transporter $j$ |
| **Variables** | |
| $x_{ip}$ | 1 if bed $i$ is assigned to patient $p$, 0 otherwise |
| $y_{jdp}$ | 1 if transporter $j$ for its $d^{th}$ service is assigned to patient $p$, 0 otherwise |
| $t_i^c$ | time that clean bed $i$ becomes available (status clean) |
| $t_{jd}^t$ | time that transporter $j$ becomes available for its $d^{th}$ service |
| $T_p$ | time that patient $p$ is served (transported to a clean bed) |
| $W_p$ | waiting time for patient $p$ |
| $W^{pmax}$ | maximum patient waiting time $W_p$ for the run |
| $t_p^b$ | time that a bed is ready for patient $p$ |
| $s_p^t$ | service time for transporting patient $p$ |
| $\delta_r$ | 1 if all patients in room $r$ are male, 0 if all females |

### 4.2.2 Second Stage

In this stage, the objective of the proposed mathematical model is to minimize the weighted sum of total waiting times of patients that are already admitted and maximum patient waiting time experienced:

$$Min \quad \sum_{p \in P} (T_p - t_p) + \beta_2 W_{pmax} \tag{4.30}$$

Similar to the previous model, the two important resources, a clean bed and a transport staff, must be available for transferring admitted patient to assigned bed, as enforced in constraints in (4.31)-(4.33). Constraints in (4.34) defines the time that assigned bed becomes ready. We utilize the results of previous step related to the availability time of clean beds for this model. In other words, the parameter $t_i^c$ stores the input data for the time a bed becomes clean and ready. We define the availability of clean bed based on their assignment to EVS staff in previous step. The constraints in (4.35) guarantee that the admitted patient is ready for transfer before the task starts. The constraints in (4.36) define the earliest availability time of a transport staff that equals to $s_p^t$ time unit after completing his previous task. The variable $s_p^t$ is the traveling time from emergency department to assigned inpatient unit. In the objective function, the term $W_pmax$ shows the maximum waiting time of patients admitted to inpatient units, which is defined in constrains (4.37) and (4.38). Then, we continue with the assignment restrictions. We guarantee that each patient is assigned to one bed and one transport staff via constraints in (4.39) and (4.40). Constraints (4.41) and (4.42) enforce that a bed and a transport staff is not assigned to more than on patient. Constraints in (4.43) impose the restriction that a transport staff should not be assigned to $d^{th}$ service, except the staff is assigned to $(d-1)^{th}$

service. Next, We formulated gender matching restrictions in constraints (4.44)-(4.47). These constraints guarantee that all patients in a room have the same gender.

$$T_p \geq t_{jd}^t + s_p^t - (1 - y_{jdp})M, \quad \forall j \in J, \forall p \in P, d = 1, ..., D^j \tag{4.31}$$

$$s_p^t = \sum_{i \in I} R_{ip} x_{ip}, \quad \forall p \in P \tag{4.32}$$

$$T_p \geq t_p^b, \quad \forall p \in P \tag{4.33}$$

$$t_p^b \geq t_i^c - (1 - x_{ip})M, \quad \forall i \in I, \forall p \in P \tag{4.34}$$

$$T_p \geq t_p + s_p^t, \quad \forall p \in P \tag{4.35}$$

$$t_{jd}^t \geq T_p + s_p^t - (1 - y_{j(d-1)p})M, \quad \forall p \in P, \forall j \in J, d = 2, ..., D^j \tag{4.36}$$

$$W_p \geq T_p - t_p, \quad \forall p \in P \tag{4.37}$$

$$W_{pmax} \geq W_p, \quad \forall p \in P \tag{4.38}$$

$$\sum_{i \in I} x_{ip} = 1, \quad \forall p \in P \tag{4.39}$$

$$\sum_{j \in J} \sum_{d \leq D^j} y_{jdp} = 1, \quad \forall p \in P \tag{4.40}$$

$$\sum_{p \in P} x_{ip} \leq 1, \quad \forall i \in I \tag{4.41}$$

$$\sum_{p \in P} y_{jdp} \leq 1, \quad \forall j \in J, d = 1, ..., D^j \tag{4.42}$$

$$\sum_{p \in P} y_{jdp} \leq \sum_{p \in P} y_{j(d-1)p}, \quad \forall j \in J, d = 1, ..., D^j \tag{4.43}$$

$$\sum_{i \in I_r} \sum_{p \in P} x_{ip} g_p \geq \sum_{i \in I_r} \sum_{p \in P} x_{ip} - (1 - \delta_r)M, \quad \forall r \in R \tag{4.44}$$

$$\sum_{i \in I_r} \sum_{p \in P} x_{ip}(1 - g_p) \geq \sum_{i \in I_r} \sum_{p \in P} x_{ip} - \delta_r M, \quad \forall r \in R \tag{4.45}$$

$$G_i^F t_i^b \leq T_p + (1 - x_{jp} g_p)M, \quad \forall r \in R, \forall p \in P, \forall i, j \in I_r, i \neq j \tag{4.46}$$

$$G_i^M t_i^b \leq T_p + (1 - x_{jp}(1 - g_p))M, \quad \forall r \in R, \forall p \in P, \forall i, j \in I_r, i \neq j, \tag{4.47}$$

## 4.3 Solution Approach

In this section, we provide a solution approach according to the proposed proactive coordination model in ED-to-IU network in hospitals. Figure 12 illustrates the steps of the proposed solution approach in detail. In this approach, the real time information are updated periodically after each $\Delta t$ units of time. The real time information are including model parameters (defined in previous section such as set of patients and resources, and resources availability times), the distribution of remaining length of stay of patients in ED before admission decisions, and patient's disposition. In this model, we consider all patients that entered in the emergency department before current time and waiting for next proceeding of the hospital. These patients are categorized in two types: 1) patients that entered to emergency department and still waiting for admission/discharge decision and 2) patients that entered to ED and admitted to inpatient unit and now they are waiting to be assigned and transferred to IU bed. After updating the real time information, we generate $N_s$ scenarios according to the uncertainties parameters in the model. Afterwards, the proposed preprocessing step and optimization model are executed in order to revise resource allocation in the system.

### 4.3.1 Scenario Generation

There are two types of uncertainties corresponding to patients entered to emergency department and waiting for admission decision, including 1) Disposition decision of the patients and 2) Remaining length of stay(RLOS) of patient in ED before admission decision. The statistics of each uncertainty are updated during the time until the patient is admitted. For each event, we generate $N_s$ scenarios based on the latest updated statistics for the proposed stochastic optimization model (first step). The samples of admission decision times for patient p are generated according to the updated distribution of RLOS of patient p in ED using Latin Hypercube Sampling (LHS) approach [48]. In LHS approach, the cumulative distribution is divided into segments, one for each scenario. A probability is randomly selected in each segment using a uniform distribution, and then mapped to the actual representative value of the variable's actual distribution. This method guarantees that the distribution function is sampled evenly. Moreover, we generate $N_s$ samples based on updated probabilities of disposition decision of patient p and assign to scenarios randomly. For example, if the probabilities of disposition decision ([Home, IU1, IU2, IU3]) for patient p equals to [0.4, 0.3, 0.2, 0.1] and we assume to generate 20 scenarios, the number of cases for each decision are [8, 6, 4, 2] cases for each patient and then we assign to scenarios randomly. Next, we combine all samples to finalize the $N_s$ scenarios.

#### 4.3.2 Pre-processing

In the preprocessing step, we apply several techniques to reduce the size of stochastic problem which leads to less complexity time. In this step, we reduce the number of resources based on time horizon considering to keep possible solutions. For example, if a bed is not suitable for any of the requests, we remove that bed from the optimization model. Also, preprocessing step limits the task duplication based on the time horizon. For instance, if time horizon is 180 minutes and average cleaning time is 50 minutes, the maximum duplication of cleaning tasks for each EVS staff member is 4. As another technique in this step, we specify symmetry cases (e.g. patients with similar attributes including gender and IU type) and reduce the feasible solution area by defining specific constraints that ensure patients with the same attributes including are served in order of their admission times. In our computational experiments, this step significantly reduced the computational time for implementing almost all large problem instances.

### 4.3.3 Two stage optimization models

We employ the optimization approach in two steps to consider uncertainties in the system, to prioritize ED patients that are admitted to hospital and to reduce the computational complexity. Since there are two uncertainties related to patients entered to emergency department, we consider a stochastic optimization model for the first step. Our primary goal of the first step is to find the best assignment of EVS staff to dirty beds by considering all scenarios. We follow an iterative approach for this step to balance the total number of clean beds between inpatient units. After each iteration, we find the demand for clean beds and also available clean beds (considering new assignments) in each IU. If the difference between demand and supply for IUs is imbalance, one new constraint will be added in order to limit the new assignments for cleaning tasks in each IU. For example, if the results of step 1 shows that the difference between number of requests for clean bed and available clean beds in IU1 and IU2 are 0 and 6 respectively, the new constraint will be added to the model to restrict the number of assignments of EVS staff to dirty bed for IU1 and IU2 to 3 assignments for each one. In the second step, we utilize the availability times of clean beds based on the result of first step and fix as parameters. Then we implement the second optimization model by considering just admitted patients. The results of this step reports the best assignments of patient to transport staff and patient to clean bed based on the current situation to minimize patients waiting times. After achieving the final

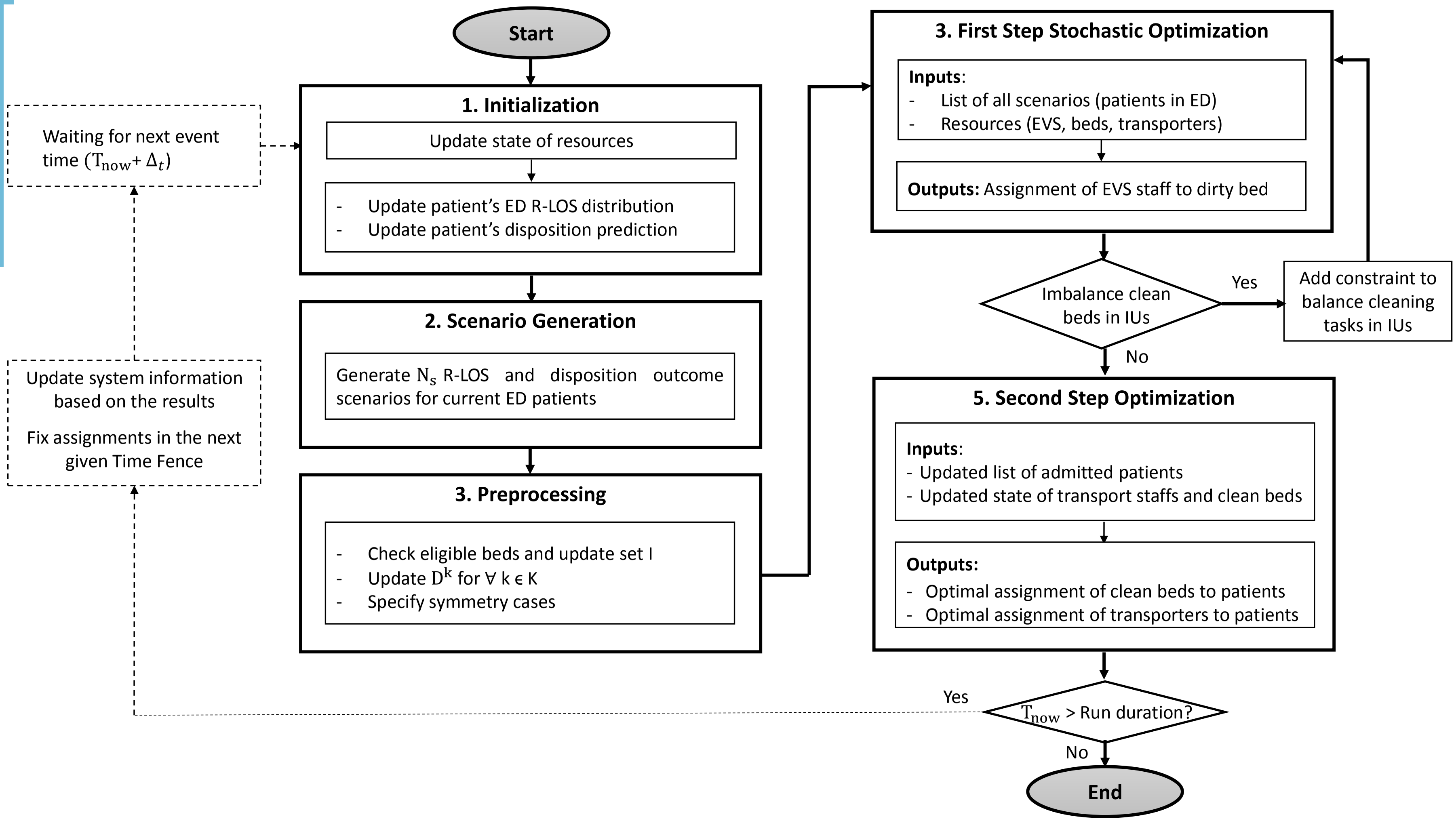



Figure 12: Flowchart for proposed solution approach

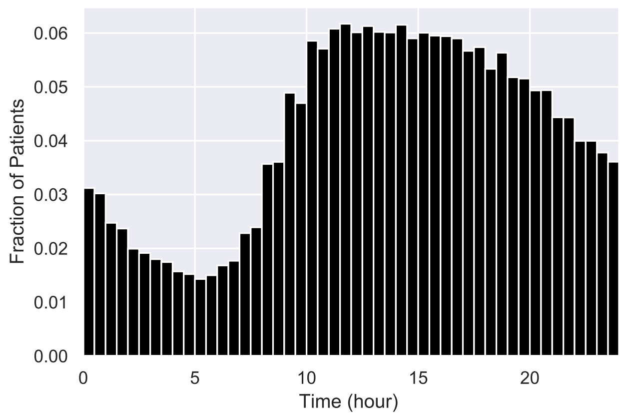
result of the current instance, the system information is updated. For the next time event, we fix the resource assignments that are starting during a defined time fence. Figure 12 demonstrates more detail of the proposed solution approach.

### 4.4 Computational Experiments

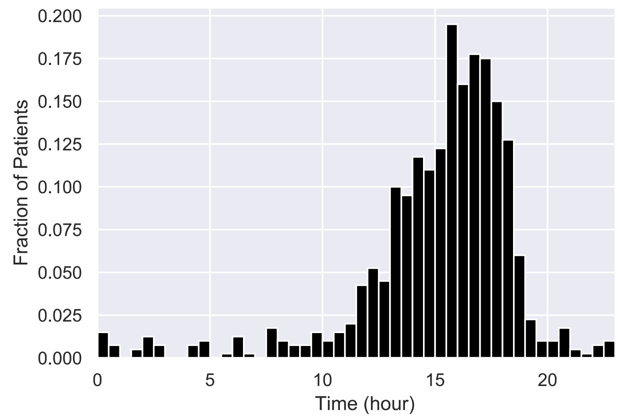
In this section, we provide a computational study to evaluate the effectiveness of proposed proactive resource and task assignment approach. Also, we investigated whether or not considering uncertainties in the model has an impact on the performance. Therefore, two versions of both the deterministic proactive model and the stochastic proactive model were tested. For the purpose of experimentation, we generate random instances based on data statistics from Henry Ford Hospital in Detroit, Michigan. We compare the performance of the proposed approach with the reactive optimization model and First-Come, First-Served method. The performance has been evaluated in terms of patients waiting times across all replications of the experiment setting. We tested our approach on a 3.10 GHz desktop with 16 GB of RAM under Windows OS with optimization solver Gurobi. We used maximum computational time of 15 CPU minutes and optimality gap of 0.1% as termination criteria. Given the time between any two consecutive optimization runs is at least 30 minutes, the 15-minute limit does not create any issue for resource planning. The following part outlines how the problem instances are generated.

#### 4.4.1 Random Problem Instance Generation

We generated five random instances based on data statistics from Henry Ford Hospital to evaluate the proposed approach. For the simulated experiments, we implemented these instances for 8 consecutive days simulated data in ED-IU processes. To eliminate the transient effect of initialization, we remove the results of the first 3 days and only con-




(a) Patient Arrival Times




(b) Patient Discharge Times

Figure 13: Distributions of Patient Arrival and Discharge Times

sider results of the last 5 days (namely, day 1 to day 5) for analysis and comparison. We utilized the historical data from HFH for generating distribution of patient arrival time to emergency department and distribution of patient discharge time from inpatient units, as shown in Figures 13a and 13b. We assume that the disposition decision of ED patients will be determined after a random variable (LOS in ED) with mean 4.5 hours and standard deviation 2 hours from when patients arrive in the emergency department. We generate the LOS of patients in ED from lognormal distribution with mean 5 and variance 1.2. Also, we suppose the real time information are updated periodically after each 30 minutes to when patients' emergency medical treatment is complete and they leave the emergency department to either go home (discharged patients) or to a hospital bed (admitted patients).

The procedure for updating the statistics related to remaining length of stay (RLOS) of patient in ED is as follows:

1. We generate the initial estimation of LOS in ED until disposition decision time for patients based on HFH, that equals to the initial value for expected remaining length of stay (RLOS) in ED. The standard deviation (SD) of RLOS for patient p is computed

by multiplying coefficient of variation (COV) and mean of RLOS. In this study we consider 0.5 as COV.

- $SD(RLOS_p(t)) = COV \times E(RLOS_p(t))$

2. The mean and standard deviation of RLOS will be updated after each $\Delta$t (30 minutes for this study) from patient arrival time until expected RLOS is less than 15 minutes. We use logistic function [76] to define the changes from initial stage to saturation.

   (a) $E(RLOS_p(t + \Delta t)) = E(RLOS_p(t)) + \delta_1 \times f(T_{now} - T_{arrival}(p))$

      - $RLOS_p(t) =$ remaining LOS of patient p in ED at time t
      - $\delta_1 = 30 \times Random\ number\ from\{-1, 1\}$
      - $f(x) = \frac{a}{(1+b\ c^{-x})};\ a, b = 1, c = 0.5$

   (b) $SD(RLOS_p(t + \Delta t)) = COV \times E(RLOS_p(t + \Delta t))$

3. We find disposition decision (admission/discharge) time of patients based on their final updated expected RLOS by getting sample from RLOS distribution:

   - Disposition decision time $= T_{now}$ + random value from RLOS distribution

The procedure for updating the probabilities of the disposition decision for patient $p$ in ED is as follows:

1. We generate disposition decision of patients based on HFH statistics :

   - 30% of ED patients are admitted to inpatients units with ratio $[0.6, 0.2, 0.2]$ to $[IU1, IU2, IU3]$ and 70% of ED patients are discharged.

- We consider initial probability for disposition decision [0.7, 0.18, 0.06, 0.06] for [Discharge, IU1, IU2, IU3] for all patients

2. After each Δt, the probability of disposition decision will be updated based on random values from Dirichlet distribution [94] function with parameter $\alpha_t$:

   - The total number of updates ($N_u$) for each patient equals to number of updates for his admission time.
   - $Pr_p(t) =$ probabilities of dispositions provided by random values from $Dir(\alpha_t)$
   - We suppose $\alpha_0$ equals to [11, 3, 1, 1] for [Discharge, IU1, IU2, IU3]
   - $\alpha_{t+\Delta t} = \alpha_t + \delta_2 \times f(T_{now} - T_{arrival}(p))$
     - We suppose the final update for $\alpha$ should be 100 for patient disposition and 1 for others. For example if patient p will be assigned to IU1 , the final alpha ($\alpha_T$) is [1, 100, 1, 1]
     - $\delta_2 = \frac{\alpha_T - \alpha_0}{N_u}$
     - $f(x) = \frac{a}{(1+b\, c^{-x})};\ a, b = 1, c = 0.5$

In the study, the average number of admitted patients in a day is considered 40 and the inter-arrival times of patients are distributed based on HFHS historical dataset. We assume 50%-50% ratio for patient's gender (male and female). In addition, we set the number of beds as 100 in total while those are distributed in the rooms with one or two capacity of patients, half of the rooms have two beds and the other half have a single bed. In the computational study, we consider three types of Inpatient Units (for three levels of intensity of care). According to historical data from HFHS, average admission probability of a patient to each of the Inpatient Units IU1, IU2, IU3 are [0.6, 0.2, 0.2], respectively. Hence, we assume the same ratio for bed distribution as 60%, 20%, and 20% for IU1, IU2, and IU3 respectively. Other parameters in this model are completion times of the tasks. The average traveling times from ED to IU1, IU2, IU3 are respectively [5, 15, 25] minutes. Also, the average traveling times between two IUs are [10, 20, 10] minutes for IU1 to IU2, IU1 to IU3, IU2 to IU3. In this experiment, the average duration of the cleaning task by EVS staff is 50 minutes. We consider two EVS staffs and two transporters in each of the 8-hour shifts and the time horizon for each coordination optimization is set as three hours.

### 4.4.2 Results

In the experimental study, we compare the performance of both the stochastic and the deterministic models of proactive coordination approach with the FCFS benchmark method and also the reactive coordination model proposed in previous chapter. The comparison is based on patients waiting times in 5 simulated days for 5 different instances. In general, the proactive approaches outperform reactive approach and FCFS strategy in particular the result of stochastic coordination model are impressive. Table 7 compares the average of ED patient waiting time for four approaches for all instances. The results show that the proposed proactive approaches decreases patients waiting times for different number of admitted patients. Also, the stochastic proactive model leads to higher improvement in patients waiting times especially when the demands are higher. For instance, in example 5 with average 45 admitted patients has higher improvement of average boarding times in comparison with example 4 where average number of admitted patient is 38. As presented in Table 8 , the maximum values of the waiting times for proactive methods are decreased with minimizing the average waiting times. When only boarded patients

Table 7: Average Waiting Time of Patient (minutes)

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic | Number of Patients per Day |
|---|---|---|---|---|---|
| 1 | 190.321 | 110.225 | 91.304 | 78.953 | [39, 38, 42, 44, 43] |
| 2 | 315.319 | 181.708 | 124.068 | 105.947 | [48, 38, 40, 39, 42] |
| 3 | 177.357 | 108.199 | 89.240 | 75.199 | [49, 37, 39, 43, 37] |
| 4 | 148.604 | 82.758 | 64.314 | 59.415 | [44, 42, 26, 41, 39] |
| 5 | 513.141 | 232.280 | 194.138 | 143.763 | [46, 52, 54, 40, 35] |

Table 8: Maximum Waiting Time of Patient (minutes)

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic | Number of Patients per Day |
|---|---|---|---|---|---|
| 1 | 956.825 | 976.300 | 937.875 | 872.057 | [39, 38, 42, 44, 43] |
| 2 | 1695.900 | 1279.802 | 1203.972 | 1120.287 | [48, 38, 40, 39, 42] |
| 3 | 1819.254 | 1613.543 | 1613.543 | 1060.408 | [49, 37, 39, 43, 37] |
| 4 | 678.467 | 665.647 | 681.879 | 665.039 | [44, 42, 26, 41, 39] |
| 5 | 1010.732 | 1219.985 | 997.572 | 852.584 | [46, 52, 54, 40, 35] |

are considered, lower number of patients experienced lower average waiting times using proactive approaches (stochastic and deterministic) compared with reactive optimization model and FCFS approach, see Tables 9 and 10. Furthermore, as displayed in Table 11, the median waiting time of boarded patients with proactive approach improves significantly rather than reactive approach .

Figure 14 displays sorted waiting times of all patients for all approaches for 5 instances during 5 days. The results show that the proactive approach consistently outperform the

Table 9: Number of Boarded Patients

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic |
|---|---|---|---|---|
| 1 | 119 | 128 | 109 | 106 |
| 2 | 124 | 158 | 118 | 108 |
| 3 | 100 | 112 | 98 | 84 |
| 4 | 92 | 107 | 92 | 86 |
| 5 | 191 | 190 | 191 | 174 |

Table 10: Average Waiting Time of Boarded Patient (minutes)

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic |
|---|---|---|---|---|
| 1 | 329.464 | 177.351 | 172.556 | 153.436 |
| 2 | 526.380 | 238.043 | 217.644 | 203.065 |
| 3 | 363.581 | 197.986 | 186.676 | 183.520 |
| 4 | 310.130 | 148.459 | 134.220 | 132.647 |
| 5 | 609.858 | 277.509 | 230.730 | 187.553 |

Table 11: Median Waiting Time of Patient (minutes)

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic |
|---|---|---|---|---|
| 1 | 338.742 | 91.748 | 107.302 | 89.428 |
| 2 | 485.901 | 138.324 | 57.882 | 53.317 |
| 3 | 303.311 | 102.914 | 48.167 | 66.777 |
| 4 | 287.707 | 95.653 | 45.347 | 38.558 |
| 5 | 635.323 | 208.801 | 139.893 | 96.176 |

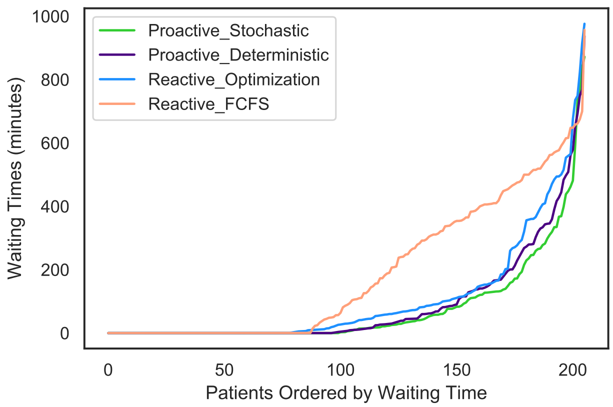
reactive models, for all problem instances. The proactive approach leads to significant improvement in patients waiting times and decreases the frequency of patients with excessive waiting times. We note that, although the range of patients waiting times can be the same for all methods, there is remarkable difference between the distributions of these solutions. As depicted, Proactive strategy increases the number of patients without any boarding in all examples. Also, we observe that stochastic optimization model performed better than the deterministic optimization model in proactive approach and reduced the number of boarded patients during 5 days. Instance 5 (Figure 14e), shows a clear higher improvement by stochastic optimization model as there is higher number of admitted patients in this example.

We also evaluated the utilization of EVS staffs for the proactive, reactive, and FCFS approaches. The results of this computational study show that the solution of all approaches have approximately the same average workload of cleaning of dirty beds in all approaches.

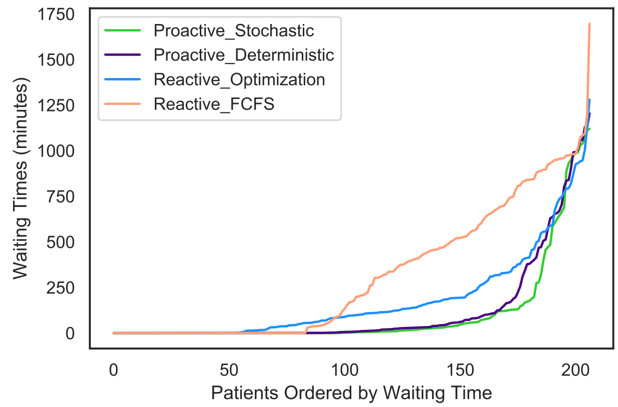
Table 12: Average Percentage of Time for [Cleaning, Traveling, Idle] for EVS Staff

| Example | FCFS | Reactive Optimization | Proactive Deterministic | Proactive Stochastic |
|---|---|---|---|---|
| 1 | [71%, 14%, 15%] | [71%, 2%, 27%] | [70%, 4%, 26%] | [70%, 5%, 25%] |
| 2 | [68%, 13%, 19%] | [64%, 3%, 33%] | [66%, 5%, 29%] | [67%, 5%, 28%] |
| 3 | [67%, 12%, 21%] | [64%, 2%, 34%] | [65%, 6%, 29%] | [66%, 3%, 31%] |
| 4 | [74%, 11%, 15%] | [73%, 2%, 25%] | [73%, 3%, 24%] | [73%, 3%, 24%] |
| 5 | [79%, 15%, 6%] | [81%, 8%, 11%] | [77%, 7%, 16%] | [80%, 8%, 12%] |

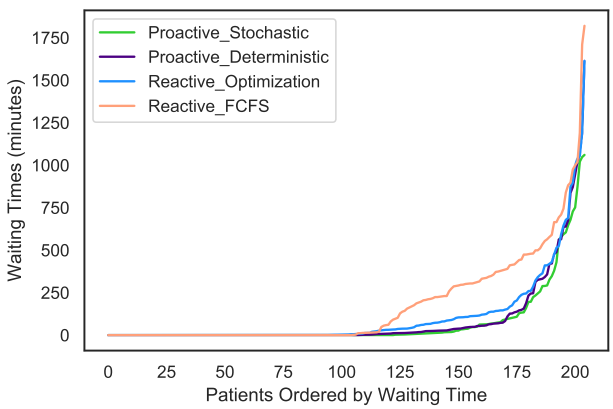
However, the proactive and reactive optimization models can reduce the workload of EVS staff and increase their idle times by assigning them optimally rather than FCFS strategy (see Table 12 ).



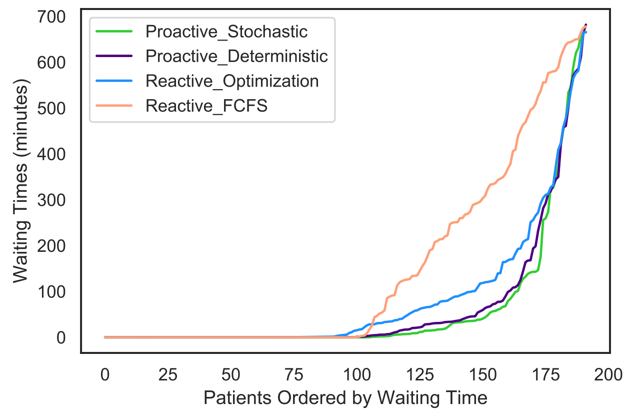

(a) Example 1- Patient waiting times (sorted)



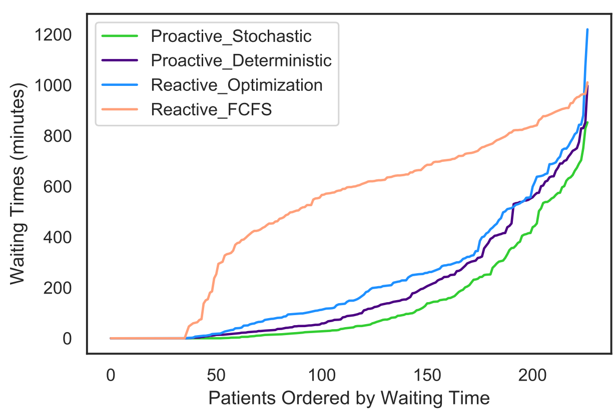

(b) Example 2- Patient waiting times (sorted)




(c) Example 3- Patient waiting times (sorted)




(d) Example 4- Patient waiting times (sorted)




(e) Example 5- Patient waiting times (sorted)

Figure 14: Patient waiting time performance under the proposed proactive (stochastic and deterministic) approach versus reactive (optimization, FCFS) approach

# CHAPTER 5 CONCLUSION AND FUTURE RESEARCH

## 5.1 Summary and Conclusion

In this research, we proposed a dynamic resource allocation and task assignment optimization model to improve system performance by minimizing waiting time experienced by emergency department (ED) patients considered for hospital admission (i.e., boarding). In particular, we developed an effective mixed-integer formulation to solve the proposed coordination problem. Our method coordinates different departments/teams including emergency department, inpatient units, environmental services, and transportation. The proposed coordination approach not only provides assignment of resources to tasks, but also schedules the tasks for a rolling planning horizon while satisfying many real-world requirements and constraints. The proposed coordination approach is shown to significantly outperform first-come first-served practices prevalent in hospitals in terms of ED patient waiting times while also improving resource utilization, availability, and workload equity.

We focus on two perspectives in our proposed framework. At the first one, we consider the reactive approach of resource allocation as a dominant process in ED-to-IU network in hospitals. The main assumption here is assigning the resources after actual admission time to IU and disposition decision. We develop a deterministic dynamic real-time coordination model for resource and task assignment using mixed-integer programming. Then we take the advantages of early task initiation and introduce a proactive approach to demonstrate how early task initiation using available EHR information in upstream tasks like as triage and initial patient assessment can improve the reactive approach for resource coordination. We propose the proactive optimization model based on reactive deterministic MIP model and involve uncertainties into the model and provide the proactive stochastic MIP model. We consider two sources of uncertainty with the problem assumptions and formulation in this step: 1) Disposition (admission/discharge) decision time and 2) The disposition decision (e.g. IU type that patient will be assigned).

Our proposed proactive approach shows if we consider the reliable predicted information of ED admission decisions ahead of the actual admission decisions, we would be able to optimize the resource allocation in a proactive manner to reduce ED patients waiting times. We assume that when a new patient enters in emergency department and starts the testing and treatment process, the information about the patient like as patient's health history provides reliable estimation of disposition decisions and admission times before the actual decisions. This estimation helps to improve allocating the resources to relevant tasks in a proactive manner regarding an impending demand.

We validate the effectiveness of the model using data from a leading healthcare facility in SE-Michigan, U.S. Our proposed coordination methodology is also applicable to other healthcare departments for resource management and improved patient satisfaction.

## 5.2 Future Research

There are several avenues for future research. We propose two major opportunities as future work which can extent and improve the current research. The first direction focuses on extending the problem modeling and the second one is providing ideas for improving the solution approach, specifically for large scale problems.

### 5.2.1 Enterprise-wide resources allocation and coordination

In this research, our focus limited to resources allocation and coordination in ED-to-IU network while the same framework can be extended to multiple departments as a multi-agent system to share their tasks and resources which optimize multiple objectives. For example, our framework can include elective patients and provides bed allocation as well as staff coordination for transportation and bed cleaning across different departments including ED. In addition, for proposing more comprehensive optimization model, other resources/tasks and more types of uncertainties can be added to the model.

### 5.2.2 Computationally effective solution approach using Machine Learning

As mentioned in chapter 3 and 4, we consider several decision variables and real-world constraints in the proposed MIP models which make them computationally expensive. Although our solution approaches improved execution time significantly, still there is a need to improve the solution approach specifically for supporting large-scale healthcare facilities. Development of more computationally efficient heuristics could be one direction for future direction. On the other side, recently, few studies utilized machine learning to solve the large scale online Mixed Integer Programming (MIP) problems [16, 33, 15, 24, 20] which could be an interesting future workstream of this research. For better understanding, an MIP model is solved on a common basis, maintaining notable similarities in formulation structures and solution outcomes but varies in model coefficients. For example, our proposed MIP model shares important similarities in terms of model structure and solution outcomes across different model coefficients (e.g number of patients and number of resources). This provides a great opportunity to integrate Machine Learning (ML) algorithms to explore relationship between an MIP model's formulation structure and its solution values to improve the computation performance. Accordingly, if we consider an standard formulation of optimization model as $z = min_{Ax \leq B, x \in X} C^T x$, different instances of model are varying only in formulation coefficients $A$, $B$ and $C$. Therefore, The main task here is to generate the training set based on comprehensive multiple runs and predict the probability that a decision variable (in our case, binary variable) to get value 1 (or zero) in the optimized solution. Since we consider many real-world constraints in our model, there are many formulation coefficients in our model, so we propose deep learning algorithms [73, 84] as appropriate choice of prediction model for this high dimensional problem.

**REFERENCES**

[1] R. Abelson. E.r.s account for half of hospital admissions, study says. https://www.nytimes.com/2013/05/21/business/half-of-hospital-admissions-from-emergency-rooms.html, 2013.

[2] J. Abraham and M. C. Reddy. Challenges to inter-departmental coordination of patient transfers: a workflow perspective. *International Journal of medical informatics*, 79(2):112–122, 2010.

[3] J. Adler-Milstein, C. M. DesRoches, P. Kralovec, G. Foster, C. Worzala, D. Charles, T. Searcy, and A. K. Jha. Electronic health record adoption in us hospitals: progress continues, but challenges persist. *Health affairs*, 34(12):2174–2180, 2015.

[4] R. R. Adudodla, G. Babu, S. Kumar, and S. Sharma. Bed allocation in hospitals-a case study. *International Journal of Innovative Technology and Research*, 3(5):2455–2458, 2015.

[5] M. Ahmadian, R. Blake, and E. Elahi. Suppliers learning: Aggregate and individual levels. In *2018 Annual Conference (47th)*, page 752, 2018.

[6] M. A. Ahmed and T. M. Alkhamis. Simulation optimization for an emergency department healthcare unit in kuwait. *European journal of operational research*, 198(3):936–942, 2009.

[7] A. Al-Refaie, T. Chen, and M. Judeh. Optimal operating room scheduling for normal and unexpected events in a smart hospital. *Operational Research*, 18(3):579–602, 2018.

[8] M. J. Alves and J. Clímaco. A review of interactive methods for multiobjective integer and mixed-integer programming. *European Journal of Operational Research*,

180(1):99–115, 2007.

[9] A. Ansaripour, A. Mata, S. Nourazari, H. Kumin, et al. Some explicit results for the distribution problem of stochastic linear programming. *Open Journal of Optimization*, 5(04):140, 2016.

[10] O. M. Araz, D. Olson, and A. Ramirez-Nafarrate. Predictive analytics for hospital admissions from the emergency department using triage information. *International Journal of Production Economics*, 208:199–207, 2019.

[11] Y. Barak-Corren, S. H. Israelit, and B. Y. Reis. Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. *Emergency Medicine Journal*, 34(5):308–314, 2017.

[12] N. D. Bastian, T. Ekin, H. Kang, P. M. Griffin, L. V. Fulton, and B. C. Grannan. Stochastic multi-objective auto-optimization for resource allocation decision-making in fixed-input health systems. *Health care management science*, 20(2):246–264, 2017.

[13] R. J. Batt and C. Terwiesch. Early task initiation and other load-adaptive mechanisms in the emergency department. *Management Science*, 63(11):3531–3551, 2017.

[14] S. Belciug and F. Gorunescu. Improving hospital bed occupancy and resource utilization through queuing modeling and evolutionary computation. *Journal of biomedical informatics*, 53:261–269, 2015.

[15] D. Bertsimas and B. Stellato. Online mixed-integer optimization in milliseconds. *arXiv preprint arXiv:1907.02206*, 2019.

[16] D. Bertsimas and B. Stellato. The voice of optimization. *Machine Learning*, pages 1–29, 2020.

[17] H. Billhardt, M. Lujak, V. Sánchez-Brunete, A. Fernández, and S. Ossowski. Dynamic coordination of ambulances for emergency medical assistance services. *Knowledge-Based Systems*, 70:268–280, 2014.

[18] S. Brailsford and J. Vissers. Or in healthcare: A european perspective. *European journal of operational research*, 212(2):223–234, 2011.

[19] S. C. Brailsford, P. R. Harper, B. Patel, and M. Pitt. An analysis of the academic literature on simulation and modeling in health care. In *Operational Research for Emergency Planning in Healthcare: Volume 2*, pages 231–251. Springer, 2016.

[20] F. Bravo and Y. Shaposhnik. Mining optimal policies: A pattern recognition approach to model analysis. *INFORMS Journal on Optimization*, pages ijoo–2019, 2020.

[21] R. L. Burdett and E. Kozan. An integrated approach for scheduling health care activities in a hospital. *European Journal of Operational Research*, 264(2):756–773, 2018.

[22] T. A. Cardona, I. G. Guardiola, and E. Cudney. Simulation of va hospital length of stay for analyzing additional inpatient bed capacity. In *IIE Annual Conference. Proceedings*, pages 1320–1325. Institute of Industrial and Systems Engineers (IISE), 2017.

[23] T. Cardoso, M. D. Oliveira, A. Barbosa-Póvoa, and S. Nickel. An integrated approach for planning a long-term care network with uncertainty, strategic policy and equity considerations. *European Journal of Operational Research*, 247(1):321–334, 2015.

[24] A. Cauligi, P. Culbertson, B. Stellato, D. Bertsimas, M. Schwager, and M. Pavone. Learning mixed-integer convex optimization strategies for robot planning and control. *arXiv preprint arXiv:2004.03736*, 2020.

[25] S. Ceschia and A. Schaerf. Modeling and solving the dynamic patient admission scheduling problem under uncertainty. *Artificial intelligence in medicine*, 56(3):199–205, 2012.

[26] CMS. Hospital compare datasets. https://data.medicare.gov/data/hospital-compare/, 2018. [Online; accessed 2018-06-11].

[27] CMS. Center for medicare and medicaid services, department of health and human services (hhs). https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html, 2019. [Online; accessed 2019-12-17].

[28] J. K. Cochran and K. T. Roche. A multi-class queuing network analysis methodology for improving hospital emergency department performance. *Computers & Operations Research*, 36(5):1497–1512, 2009.

[29] C. Davy, J. Bleasel, H. Liu, M. Tchan, S. Ponniah, and A. Brown. Effectiveness of chronic care models: opportunities for improving healthcare practice and health outcomes: a systematic review. *BMC health services research*, 15(1):1–11, 2015.

[30] S. De Vuyst, H. Bruneel, and D. Fiems. Computationally efficient evaluation of appointment schedules in health care. *European Journal of Operational Research*, 237(3):1142–1154, 2014.

[31] B. T. Denton. Handbook of healthcare operations management. *New York: Springer*, 10:978–1, 2013.

[32] P. Devapriya, C. T. Strömblad, M. D. Bailey, S. Frazier, J. Bulger, S. T. Kemberling, and K. E. Wood. Stratbam: a discrete-event simulation model to support strategic

hospital bed capacity decisions. *Journal of medical systems*, 39(10):130, 2015.

[33] J.-Y. Ding, C. Zhang, L. Shen, S. Li, B. Wang, Y. Xu, and L. Song. Optimal solution predictions for mixed integer programs. *arXiv preprint arXiv:1906.09575*, 2019.

[34] D. Dobrzykowski, V. S. Deilami, P. Hong, and S.-C. Kim. A structured analysis of operations and supply chain management research in healthcare (1982–2011). *International Journal of Production Economics*, 147:514–530, 2014.

[35] D. D. Dobrzykowski. Examining heterogeneous patterns of electronic health records use: a contingency perspective and assessment. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, 7(2):1–16, 2012.

[36] O. El-Rifai, T. Garaix, and X. Xie. Proactive on-call scheduling during a seasonal epidemic. *Operations Research for Health Care*, 8:53–61, 2016.

[37] A. Elalouf and G. Wachtel. An alternative scheduling approach for improving patient-flow in emergency departments. *Operations Research for Health Care*, 7:94–102, 2015.

[38] S. S. Fazeli, S. Venkatachalam, R. B. Chinnam, and A. Murat. Two-stage stochastic choice modeling approach for electric vehicle charging station network design in urban communities. *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[39] Y.-Y. Feng, I.-C. Wu, and T.-L. Chen. Stochastic resource allocation in emergency departments with a multi-objective simulation optimization algorithm. *Health care management science*, 20(1):55–75, 2017.

[40] E. Finch, V. Sharma, J. Abel, M. Al-Hussein, K. Lennerts, and U. Pfründer. Simulation application for resource allocation in facility management processes in hospitals. *Facilities*, 2007.

[41] D. Fouskakis and D. Draper. Stochastic optimization: a review. *International Statistical Review*, 70(3):315–349, 2002.

[42] D. Gartner, R. Kolisch, D. B. Neill, and R. Padman. Machine learning approaches for early drg classification and resource allocation. *INFORMS Journal on Computing*, 27(4):718–734, 2015.

[43] J. González, J.-C. Ferrer, A. Cataldo, and L. Rojas. A proactive transfer policy for critical patient flow management. *Health Care Management Science*, 22(2):287–303, 2019.

[44] S. Gurtner and K. Soyez. *Challenges and opportunities in health care management*. Springer, 2015.

[45] M. S. Hagen, J. K. Jopling, T. G. Buchman, and E. K. Lee. Priority queuing models for hospital intensive care units and impacts to severe case patients. In *AMIA Annual Symposium Proceedings*, volume 2013, page 841. American Medical Informatics Association, 2013.

[46] C. Harris, P. Cortvriend, and P. Hyde. Human resource management and performance in healthcare organisations. *Journal of Health Organization and Management*, 21(4/5):448–459, 2007.

[47] L. He, S. C. Madathil, A. Oberoi, G. Servis, and M. T. Khasawneh. A systematic review of research design and modeling techniques in inpatient bed management. *Computers & Industrial Engineering*, 127:451–466, 2019.

[48] J. C. Helton and F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, 81(1):23–69, 2003.

[49] A. L. Hendrich, J. Fay, and A. K. Sorrells. Effects of acuity-adaptable rooms on flow of patients and delivery of care. *American Journal of Critical Care*, 13(1):35–45, 2004.

[50] L. B. Holm, H. Lurås, and F. A. Dahl. Improving hospital bed utilisation through simulation and optimisation: with application to a 40% increase in patient volume in a norwegian general hospital. *International journal of medical informatics*, 82(2):80–89, 2013.

[51] H. Hosseini, J. Hoey, and R. Cohen. A coordinated mdp approach to multi-agent planning for resource allocation, with applications to healthcare. *arXiv preprint arXiv:1407.1584*, 2014.

[52] G. C. Kaandorp and G. Koole. Optimal outpatient appointment scheduling. *Health care management science*, 10(3):217–229, 2007.

[53] D. Kilinc, S. Saghafian, and S. Traub. Dynamic assignment of patients to primary and secondary inpatient units: Is patience a virtue. *HKS Working Paper No. RWP17-010*, 2018.

[54] A. Kulshrestha and J. Singh. Inter-hospital and intra-hospital patient transfer: Recent concepts. *Indian journal of anaesthesia*, 60(7):451, 2016.

[55] S.-Y. Lee, R. B. Chinnam, E. Dalkiran, S. Krupp, and M. Nauss. Prediction of emergency department patient disposition decision for proactive resource allocation for admission. *Health care management science*, pages 1–21, 2019.

[56] S.-Y. Lee, R. B. Chinnam, E. Dalkiran, S. Krupp, and M. Nauss. Proactive coordination of inpatient bed management to reduce emergency department patient boarding. *International Journal of Production Economics*, page 107842, 2020.

[57] J.-M. Lehtonen, J. Kujala, J. Kouri, and M. Hippeläinen. Cardiac surgery productivity and throughput improvements. *International Journal of Health Care Quality Assurance*, 20(1):40–52, 2007.

[58] S. Levin and M. Garifullin. Simulating wait time in healthcare: accounting for transition process variability using survival analyses. In *2015 Winter Simulation Conference (WSC)*, pages 1252–1260. IEEE, 2015.

[59] X. Li, P. Beullens, D. Jones, and M. Tamiz. An integrated queuing and multi-objective bed allocation model with application to a hospital in china. *Journal of the Operational Research Society*, 60(3):330–338, 2009.

[60] E. Litvak and M. Bisognano. More patients, less payment: increasing hospital efficiency in the aftermath of health reform. *Health Affairs*, 30(1):76–80, 2011.

[61] B. López, B. Innocenti, and D. Busquets. A multiagent system for coordinating ambulances for emergency medical services. *IEEE Intelligent Systems*, 23(5), 2008.

[62] M. Lujak and H. Billhardt. Optimizing emergency medical assistance coordination in after-hours urgent surgery patients. In *European Conference on Multi-Agent Systems*, pages 316–331. Springer, 2014.

[63] M. Lujak, H. Billhardt, and S. Ossowski. Distributed coordination of emergency medical service for angioplasty patients. *Annals of Mathematics and Artificial Intelligence*, 78(1):73–100, 2016.

[64] R. Luscombe and E. Kozan. Dynamic resource allocation to improve emergency department efficiency in real time. *European Journal of Operational Research*, 255(2):593–603, 2016.

[65] F. Mallor and C. Azcárate. Combining optimization with simulation to obtain credible models for intensive care units. *Annals of Operations Research*, 221(1):255–271, 2014.

[66] F. Mallor, C. Azcárate, and J. Barado. Control problems and management policies in health systems: application to intensive care units. *Flexible Services and Manufacturing Journal*, 28(1-2):62–89, 2016.

[67] K. S. Mathews and E. F. Long. A conceptual framework for improving critical care patient flow and bed use. *Annals of the American Thoracic Society*, 12(6):886–894, 2015.

[68] K. L. McFadden, S. C. Henagan, and C. R. Gowen III. The patient safety chain: Transformational leadership's effect on patient safety culture, initiatives, and outcomes. *Journal of Operations Management*, 27(5):390–404, 2009.

[69] I. Mohamed, I. El-Henawy, and R. Z. El-Din. Estimating bed requirements for a pediatric department in a university hospital in egypt. *Modern Management Science & Engineering*, 4:101, 2016.

[70] N. Mustafee, K. Katsaliaki, A. Gunasekaran, M. D. Williams, M. Fakhimi, and J. Probert. Operations research within uk healthcare: a review. *Journal of Enterprise Information Management*, 2013.

[71] J. S. Peck, J. C. Benneyan, D. J. Nightingale, and S. A. Gaehde. Predicting emergency department inpatient admissions to improve same-day patient flow. *Academic Emergency Medicine*, 19(9):E1045–E1054, 2012.

[72] J. S. Peck, J. C. Benneyan, D. J. Nightingale, and S. A. Gaehde. Characterizing the value of predictive analytics in facilitating hospital patient flow. *IIE Transactions on*

*Healthcare Systems Engineering*, 4(3):135–143, 2014.

[73] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, and S. Iyengar. A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5):1–36, 2018.

[74] N. A. Pujowidianto, L. H. Lee, G. Pedrielli, C.-H. Chen, and H. Li. Constrained optimizaton for hospital bed allocation via discrete event simulation with nested partitions. In *2016 Winter Simulation Conference (WSC)*, pages 1916–1925. IEEE, 2016.

[75] P. Punnakitikashem, J. M. Rosenberber, and D. F. Buckley-Behan. A stochastic programming approach for integrated nurse staffing and assignment. *IIE Transactions*, 45(10):1059–1076, 2013.

[76] L. J. Reed and J. Berkson. The application of the logistic function to experimental data. *The Journal of Physical Chemistry*, 33(5):760–779, 2002.

[77] F. F. Rodrigues, G. S. Zaric, and D. Stanford. Discrete event simulation model for planning level 2 âĂIJstep-downâĂİ bed needs using nems. *Operations research for health care*, 17:42–54, 2018.

[78] G. Royston. One hundred years of operational research in healthâĂŤuk 1948–2048. In *Operational Research for Emergency Planning in Healthcare: Volume 2*, pages 316–338. Springer, 2016.

[79] R. Saltzman, T. Roeder, J. Lambton, L. Param, B. Frost, and R. Fernandes. The impact of a discharge holding area on the throughput of a pediatric unit. *Service Science*, 9(2):121–135, 2017.

[80] P. Santibáñez, V. S. Chow, J. French, M. L. Puterman, and S. Tyldesley. Reducing patient wait times and improving resource utilization at british columbia cancer agen-

cyâĂŹs ambulatory care unit through simulation. *Health care management science*, 12(4):392, 2009.

[81] R. Schmidt, S. Geisler, and C. Spreckelsen. Decision support for hospital bed management using adaptable individual length of stay estimations and shared resources. *BMC medical informatics and decision making*, 13(1):3, 2013.

[82] X. Shen and X. Wang. Improving the health-care delivery process at hospital emergency services by a better use of inpatient bed information. *Electronic Commerce Research and Applications*, 14(1):14–22, 2015.

[83] W. H. Shrank, T. L. Rogstad, and N. Parekh. Waste in the us health care system: estimated costs and potential for savings. *Jama*, 322(15):1501–1509, 2019.

[84] A. Shrestha and A. Mahmood. Review of deep learning algorithms and architectures. *IEEE Access*, 7:53040–53065, 2019.

[85] S. Sitepu, H. Mawengkang, et al. Modeling an integrated hospital management planning problem using integer optimization approach. *JPhCS*, 890(1):012101, 2017.

[86] V. Sonmez, E. Dasdemir, O. M. Aydin, and M. C. Testik. A simulation based decision support tool for hospital bed capacity planning. In *Proceedings of Global Conference on Healthcare Systems Engineering*, page 2. Seyda Serdar Asan, PhD.–Istanbul–Turkey Copyright the Proceedings of the âĂę, 2014.

[87] E. Taghizadeh, M. Abedzadeh, and M. Setak. A multi objective reliable location-inventory capacitated disruption facility problem with penalty cost solve with efficient meta historic algorithms. *arXiv preprint arXiv:1711.09400*, 2017.

[88] H. Takagi, K. Misue, and Y. Kanai. Queuing network model and visualization for the patient flow in the obstetric unit of the university of tsukuba hospital. In *2014*

*Annual SRII Global Conference*, pages 147–156. IEEE, 2014.

[89] A. Testi, E. Tanfani, and G. Torre. A three-phase approach for operating theatre schedules. *Health Care Management Science*, 10(2):163–172, 2007.

[90] B. G. Thomas, S. Bollapragada, K. Akbay, D. Toledano, P. Katlic, O. Dulgeroglu, and D. Yang. Automated bed assignments in a complex and dynamic hospital environment. *Interfaces*, 43(5):435–448, 2013.

[91] S. THOMSON, M. NUNEZ, R. GARFINKEL, and M. D. DEAN. Efficient short term allocation and reallocation of patients t0 floor of a hospital during demand surges. *Operations Research*, 57(2):261–273, 2009.

[92] P.-F. Tsai and F.-M. Lin. An application of multi-attribute value theory to patient-bed assignment in hospital admission management: an empirical study. *Journal of healthcare engineering*, 5, 2014.

[93] J. P. Vielma. Mixed integer linear programming formulation techniques. *Siam Review*, 57(1):3–57, 2015.

[94] T.-T. Wong. Generalized dirichlet distribution in bayesian analysis. *Applied Mathematics and Computation*, 97(2-3):165–181, 1998.

[95] F. Zaerpour, D. P. Bischak, and M. B. Menezes. Coordinated lab-clinics: A tactical assignment problem in healthcare. *European Journal of Operational Research*, 263(1):283–294, 2017.

[96] L. Zhou, N. Geng, Z. Jiang, and X. Wang. Multi-objective capacity allocation of hospital wards combining revenue and equity. *Omega*, 81:220–233, 2018.

**ABSTRACT**

**DYNAMIC RESOURCE ALLOCATION FOR COORDINATION OF INPATIENT OPERATIONS IN HOSPITALS**

by

**NAJIBESADAT SADATIJAFARKALAEI**

**December 2020**

**Advisors:** Dr. Ratna Babu Chinnam & Dr. Evrim Dalkiran

**Major:** Industrial Engineering

**Degree:** Doctor of Philosophy

Healthcare systems face difficult challenges such as increasing complexity of processes, inefficient utilization of resources, high pressure to enhance the quality of care and services, and the need to balance and coordinate the staff workload. Therefore, the need for effective and efficient processes of delivering healthcare services increases. Data-driven approaches, including operations research and predictive modeling, can help overcome these challenges and improve the performance of health systems in terms of quality, cost, patient health outcomes and satisfaction.

Hospitals are a key component of healthcare systems with many scarce resources such as caregivers (nurses, physicians) and expensive facilities/equipment. Most hospital systems in the developed world have employed some form of an Electronic Health Record (EHR) system in recent years to improve information flow, health outcomes, and reduce costs. While EHR systems form a critical data backbone, there is a need for platforms that can allow coordinated orchestration of the relatively complex healthcare operations. Infor-

mation available in EHR systems can play a significant role in providing better operational coordination between different departments/services in the hospital through optimized task/resource allocation.

In this research, we propose a dynamic real-time coordination framework for resource and task assignment to improve patient flow and resource utilization across the emergency department (ED) and inpatient unit (IU) network within hospitals. The scope of patient flow coordination includes ED, IUs, environmental services responsible for room/bed cleaning/turnaround, and patient transport services. EDs across the U.S. routinely suffer from extended patient waiting times during admission from the ED to the hospital's inpatient units, also known as ED patient 'boarding'. This ED patient boarding not only compromises patient health outcomes but also blocks access to ED care for new patients from increased bed occupancy. There are also significant cost implications as well as increased stress and hazards to staff. We carry out this research with the goal of enabling two different modes of coordination implementation across the ED-to-IU network to reduce ED patient boarding: Reactive and Proactive. The proposed 'reactive' coordination approach is relatively easy to implement in the presence of modern EHR and hospital IT management systems for it relies only on real-time information readily available in most hospitals. This approach focuses on managing the flow of patients at the end of their ED care and being admitted to specific inpatient units. We developed a deterministic dynamic real-time coordination model for resource and task assignment across the ED-to-IU network using mixed-integer programming.

The proposed 'proactive' coordination approach relies on the power of predictive analytics that anticipate ED patient admissions into the hospital as they are still undergoing

ED care. The proactive approach potentially allows additional lead-time for coordinating downstream resources, however, it requires the ability to accurately predict ED patient admissions, target IU for admission, as well as the remaining length-of-stay (care) within the ED. Numerous other studies have demonstrated that modern EHR systems combined with advances in data mining and machine learning methods can indeed facilitate such predictions, with reasonable accuracy. The proposed proactive coordination optimization model extends the reactive deterministic MIP model to account for uncertainties associated with ED patient admission predictions, leading to an effective and efficient proactive stochastic MIP model.

Both the reactive and proactive coordination methods have been developed to account for numerous real-world operational requirements (e.g., rolling planning horizon, event-based optimization and task assignments, schedule stability management, patient overflow management, gender matching requirements for IU rooms with double occupancy, patient isolation requirements, equity in staff utilization and equity in reducing ED patient waiting times) and computational efficiency (e.g., through model decomposition and efficient construction of scenarios for proactive coordination). We demonstrate the effectiveness of the proposed models using data from a leading healthcare facility in SE-Michigan, U.S. Results suggest that even the highly practical optimization enabled reactive coordination can lead to dramatic reduction in ED patient boarding times. Results also suggest that signification additional reductions in patient boarding are possible through the proposed proactive approach in the presence of reliable analytics models for prediction ED patient admissions and remaining ED length-of-stay. Future research can focus on further extending the scope of coordination to include admissions management (including any necessary approvals from insurance), coordination needs for admissions that stem from outside the ED (e.g., elective surgeries), as well as ambulance diversions to manage patient flows across the region and hospital networks.

**AUTOBIOGRAPHICAL STATEMENT**

Najibesadat Sadatijafarkalaei is a Ph.D candidate in the Industrial and Systems Engineering Department at Wayne State University, Detroit, Michigan. She received her second Master's degree in Computer Science at Wayne State University in 2020 and achieved her first Master's degree in the Industrial and Systems Engineering from AmirKabir University of Technology in Tehran, Iran, 2013. Najibesadat got her bachelors' degree in Industrial Engineering from Sharif University of Technology in Tehran, Iran, 2011.

Her main research interests are in Machine Learning, Healthcare Data Analytics, Applied Operations Research and Operation Management. Najibesadat's papers have been published in top conferences and journals. Here is the link for her Google scholar page: https://scholar.google.com/citations?user=6FsrzF4AAAAJ&hl=en