Optimization-based Logistics Planning and Performance Measurement for Hospital Evacuation and Emergency Management

Esra Agca

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Industrial and Systems Engineering

Douglas R. Bish, Chair
Kimberly P. Ellis
Roger E. Glick
Raghu Pasupathy
Subhash C. Sarin

August 12, 2013
Blacksburg, Virginia

Keywords: Hospital Evacuation, Emergency Response, Transportation Planning, Resource Management, Equity Modeling, Integer Programming

Optimization-based Logistics Planning and Performance Measurement for Hospital Evacuation and Emergency Management

Esra Agca

(ABSTRACT)

This dissertation addresses the development of optimization models for hospital evacuation logistics, as well as the analyses of various resource management strategies in terms of the equity of evacuation plans generated. We first formulate the evacuation transportation problem of a hospital as an integer programming model that minimizes the total evacuation risk consisting of the threat risk necessitating evacuation and the transportation risk experienced en route. Patients, categorized based on medical conditions and care requirements, are allocated to a limited fleet of vehicles with various medical capabilities and capacities to be transported to receiving beds, categorized much like patients, at the alternative facilities. We demonstrate structural properties of the underlying transportation network that enables the model to be used for both strategic planning and operational decision making.

Next, we examine the resource management and equity issues that arise when multiple hospitals in a region are evacuated. The efficiency and equity of the allocation of resources, including a fleet of vehicles, receiving beds, and each hospital's loading capacity, determine the performance of the optimal evacuation plan. We develop an equity modeling framework, where we consider equity among evacuating hospitals and among patients. The range of equity of optimal solutions is investigated and properties of optimal and equitable solutions based on risk-based utility functions are analyzed.

Finally, we study the integration of the transportation problem with the preceding hospital building evacuation. Since, in practice, the transportation plan depends on the pace of building evacuation, we develop a model that would generate the transportation plan subject to the output of hospital building evacuation. The optimal evacuation plans are analyzed with respect to resource utilization and patient prioritization schemes. Parametric analysis of the resource constraints is provided along with managerial insights into the assessment of evacuation requirements and resource allocation.

In order to demonstrate the performance of the proposed models, computational results are provided using case studies with real data obtained from the second largest hospital in Virginia.

This work has been partially supported by the *National Science Foundation* under Grant # 0825611 and 1055360 and a Research Acceleration Program grant from *Carilion Clinic*.

To my dear parents, Zehra and İsmail,
and my dear siblings, Ebru and Erkan.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor Dr. Douglas R. Bish, for his invaluable guidance, encouragement, and patience. It has been a profoundly constructive experience to work with him in both research and teaching. I would like to thank my committee members, Dr. Subhash C. Sarin, Dr. Kimberly P. Ellis, Dr. Raghu Pasupathy, and Mr. Roger E. Glick, for their meticulous reviews of my work, valuable feedback, and support.

I have to thank the College of Engineering and the Department of Industrial and Systems Engineering for giving me the opportunity to teach engineering courses. I also thank the faculty of the Department of Engineering Education, especially Dr. Lisa McNair, for the training they provided that I greatly appreciate. I feel especially fortunate to have worked with Dr. Ebru Bish in teaching and I would like to extend my sincerest thanks to her for her support and advice.

I am thankful for all the friends I have had at Virginia Tech and for my friends who have always supported me from afar. Having such inspirational and encouraging friends have made this experience more nourishing and enjoyable.

I cannot thank my beloved family enough for their unconditional love and the great sacrifices they made throughout my journey. They have always believed in me and with everlasting gratitude, I dedicate my work to them.

Most importantly, I am grateful for having found my soul mate Emrah Aktunç who brings happiness and peace to my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Scope of Research

Emergency evacuations have been considered rare events, but their frequency has increased over the past few decades due to the rising number and magnitude of natural and man-made disasters that call for evacuations. The number of declared disasters has risen both in the U.S. and worldwide as reported by the Federal Emergency Management Agency (FEMA, 2013) and in the databases such as EM-DAT, maintained by the World Health Organization (WHO) Collaborating Centre for Research on the Epidemiology of Disasters (CRED, 2013). The destruction caused by disasters such as the terrorist attacks of September 11, 2001, Hurricanes Katrina and Rita in 2005, Haiti earthquake in 2010, Tōhuku earthquake in Japan in 2011, and Hurricane Sandy in 2012 raised awareness about the importance of emergency preparedness and response. The catastrophic effects of disasters have increased concern for improving the logistical efficiency and equity of emergency management operations (Van Wassenhove, 2006). The disaster management problems including, but not limited to, evacuation planning for affected populations and supply chain management for relief distribution are substantially relevant to operations research. Therefore, the need and value of disaster management studies in operations research and management sciences have been acknowledged (Altay and Green III, 2006).

Disaster management, which encompasses preparation, response, recovery, and mitigation phases, is highly complicated. One of the most challenging evacuation settings is the evacuation of hospital facilities and hospital evacuation incidents have recently been observed more often as a result of the increasing number of threats and the safety requirements. Sternberg et al. (2004) review 275 reported hospital evacuation incidents in the U.S. from 1970 to 1999 and show a significant increase in the number of such incidents per year. It must also be noted that this study excludes incidents after the year 2000 and that nine out of the ten costliest storms in the U.S. occurred after the year 2001 which suggests that the frequency of hospital evacuations have increased over the last decade. The growing need for emergency evacuation planning for hospital facilities is the primary motivation of the research presented in this dissertation.

Evacuating a hospital is generally much more complex than evacuating most other types of facilities (e.g., an office building) because of the complications inherent in the patient as evacuee. Patients require assistance and medical care throughout the evacuation process, and the evacuation does not end when the building is clear, but must also include transportation of patients to alternative hospital facilities. There are numerous factors that affect the evacuation plan including the nature of the threat, availability of resources and staff, the characteristics of the evacuee population, and the amount of risk that patients and staff are exposed to over time. The safety and health of patients is of fundamental importance, but safely moving patients to alternative hospital facilities, while under threat, is a very challenging task. Because of the complexities involved, emergency preparedness planning is indispensable for hospitals and hospitals must plan for evacuations. This is, in fact, required at the federal level by the Department of Health and Human Services' (HHS) Centers for Medicare and Medicaid Services (CMS). Hospital accrediting organizations, such as the Joint Commission (JC) (formerly the JC on Accreditation of Healthcare Organizations), assess the compliance of annual emergency management and evacuation plans of participating hospital facilities with the HHS requirements (see Joint Commission Accreditation Standard EM.02.02.03 and EM.02.02.11). However, these requirements do not specifically address the transfer of patients from hospitals to safer facilities. The U.S. Government Accountability Office (GAO) identified the lack of operational guidance for practitioners on patient transportation during disasters as a limitation of the National Disaster Medical System (NDMS) (GAO, 2006).

Every hospital must have evacuation plans to provide for the safety of patients when faced with an emergency, but there is not a standard plan that can respond to all possible emergencies successfully. The variability of evacuation plans is caused by the issues and complexities faced in constructing evacuation models. These issues include, but are not limited to, the characteristics of the threat and the evacuee population, risk to the occupants of the hospital, limited healthcare equipment to continue the treatment of patients during evacuation, and limited resources such as ambulances, buses, staff to transport patients, medical professionals and paramedics to treat patients. Although the lack of evacuation plans is the actual problem for many hospital facilities, practice indicates that the existence of evacuation plans do not necessarily guarantee that these are operational plans. A striking example of this was observed at the Memorial Medical Center in New Orleans during Hurricane Katrina in 2005 when the 246-page emergency plan was found to be lacking guidance for dealing with a complete power failure or how to evacuate the hospital if city streets were flooded (Fink, 2009, August 27). Another issue with many existing evacuation plans is that they are designed under optimistic assumptions that sufficient supply is guaranteed if hospitals have contracts with service providers. However, when a disaster necessitates evacuation for multiple hospitals in a region, there can be significant shortages in the supply received by each hospital. Hospital administrators interviewed by the GAO after Hurricane Katrina reported that they could not receive sufficient and timely vehicle supply even though they had contracted with transportation companies; because the transportation companies had contracts with multiple hospitals, required advance notice for dispatching vehicles, or suffered from the damage of the hurricane as well (GAO, 2006). Such examples clearly show that more thorough studies of the capability of hospitals during disasters need to be performed and more operational emergency plans need to be developed.

Given the catastrophic risks involved in hospital evacuations, the fundamental concern is the lack of appropriate evacuation models to aid strategic and operational emergency management decisions. We believe that operations research tools have a great potential to be successfully implemented for modeling and solving emergency hospital evacuation problems. First, mathematical programming formulations for hospital evacuations need to be developed and computational tractability needs to be ensured for practical use. Second, special structures of the hospital evacuation planning models need to be utilized to potentially improve solution performance. Finally, implications of optimal solutions resulting from the proposed models need to be analyzed with respect to resource management and equity issues as hospital evacuations are utilitarian by nature. The motivation for this proposed work relies on the aforementioned objectives.

1.2 Organization of the Dissertation

The remainder of this dissertation is organized as follows. Chapter 2 presents an overview of the existing literature on hospital evacuation incidents, building evacuation models, evacuation transportation planning, equity modeling for resource allocation problems, and patient prioritization during emergencies. Chapter 3 introduces the optimization-based framework for hospital evacuation transportation planning and discusses the structural properties and performance of the proposed integer programming formulation based on a large-scale case study with real data. In Chapter 4, resource management and equity issues that arise when multiple hospitals are evacuated are discussed and an equity modeling framework is developed based on social welfare economics. Chapter 5 introduces the integrated hospital building evacuation and transportation model and investigates the properties of the optimal evacuation plans in terms of resource management and patient prioritization schemes. Finally, in Chapter 6, concluding remarks on the dissertation are provided and directions for future research in hospital evacuation and emergency management are discussed.

Chapter 2

Literature Review

In this chapter, we present a review of the published literature on hospital evacuation planning and emergency response along with a review of equity modeling literature relevant to resource allocation problems. There is a rich literature on evacuations in general in social sciences that mostly study organizational preparedness practices, coordination efforts and decision-making processes amongst evacuation agencies, sociological impacts of disasters, and evacuee behavior. For examples of evacuation studies in social sciences, see the seminal report on evacuation behavior by Quarantelli (1980), the descriptive organizational decision-making model for emergency warning system by Sorensen and Mileti (1987), observation of the issues in nursing home evacuations by Vogt (1991), investigation of behavioral factors that affect mobilization time in evacuations by Sorensen (1987), and a review of decision-making process models for the evacuation of health care facilities by McGlown (2001). The literature relevant to our research on hospital evacuation planning in Chapters 3 - 5 is categorized as incident-based studies and mathematical modeling studies.[1] The mathematical modeling studies are further categorized based on their focus on building evacuations and hospital evacuation transportation problems. Finally, a review of the equity modeling literature based on welfare economics that is discussed further in Chapter 4 is provided.

[1]Parts of the discussion on the incident-based hospital evacuation literature will appear in the Journal of Emergency Management.

2.1 Incident-based Literature

Hospitals have to be evacuated for various reasons; examples include hurricanes, fires, floods, chemical leaks, earthquakes, bioterrorism, bomb threats, and loss of functionality. Due to the inherent complexity of hospital evacuation operations, healthcare facilities must learn from each other by sharing their emergency management experience. Emergency management is commonly described in four-phases: mitigation, preparedness, response, and recovery (Green, 2002). There is a vast literature on the response and recovery phases that mostly report hospital evacuation incidents and evaluate the applied strategies in terms of their efficiency and difficulty of implementation.

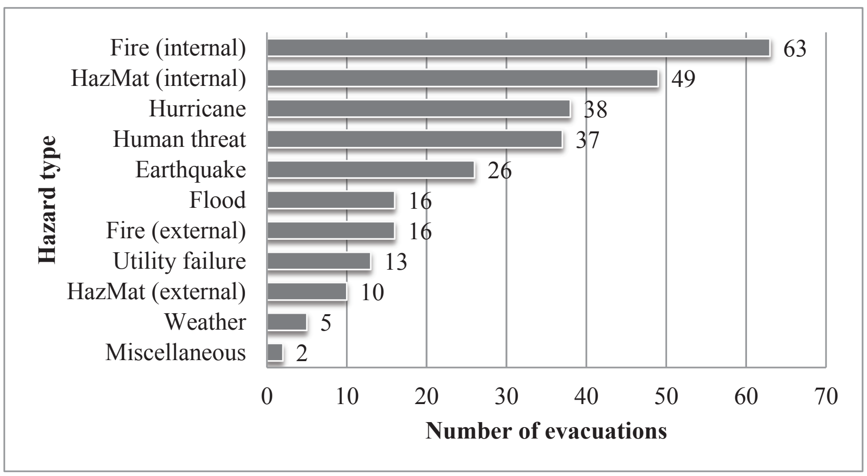
Reported incidents are caused by various types of threats depending on the characteristics of the hospital such as its location, size, and medical treatment capabilities. For example, the southeastern coastal areas of the United States is vulnerable to hurricanes, the Midwest is prone to tornadoes, and California is threatened by earthquakes. Therefore, studies on hospital evacuation incidents should be reviewed and evaluated based on the type of threat and the vulnerability of the hospitals.

Chavez and Binder (1995) describe the evacuation of the Sepulveda VA Medical Center in Los Angeles County after the Northridge, California, Earthquake in 1994. They discuss issues pertaining to the timing of evacuation, prioritization of patients, providing shelter for patients outside the hospital building while waiting for transport, and making arrangements with other facilities to receive the patients. They highlight the importance of establishing reliable communication systems, planning, education, and continuous training based on the hospital's disaster management life cycle.

Schultz et al. (2003) survey 91 acute care hospitals in Los Angeles County, 8 of which were evacuated, after the Northridge Earthquake. Their findings suggest that, although faced with the same disaster, hospitals adopt different strategies in patient triage (moving the most or least seriously ill patients) and arrangements with receiving facilities (independent or through the operations center). Based on the observations and implications of hospital evacuations after the Northridge Earthquake, Schultz et al. (2005) propose a standardized data collection tool to be used for recording pre- and post-event decision-making processes at the hospitals which would result in the accumulation of comparable data for evacuation research.

Maxwell (1982) reports the hospital evacuations following the 1979 nuclear accident at Three Mile Island near Harrisburg, Pennsylvania and describes the population reduction strategies such as discharging ambulatory and stable patients to reduce the number of evacuees for whom the resources will be used, cancellation of all surgeries except emergencies, and restriction of admissions to only life-threatening cases. This study also points out the lack of area-wide evacuation plans for hospitals that consider a wide range of threats.

Figure 2.1: Number of reported US hospital evacuations by hazard type, 1971-1999. (Adapted from Sternberg et al. (2004))



There are very few peer-reviewed publications that report the frequency of hospital evacuations. In a survey of 101 hospitals in Washington State, Burgess (1999) reports that 10 hospitals had to evacuate due to 11 hazardous material incidents between 1991–1996. The emergency department was evacuated in seven of these 11 incidents which is a useful information for evacuation planning and drill purposes.

As one of the few resources on hospital evacuation incident statistics, Sternberg et al. (2004) provide a review of the range of hazards causing 275 reported hospital evacuation incidents in the U.S. from 1971 to 1999. Based on the data set used in this study, more than half of the incidents were caused by internal or man-made hazards, but natural disasters that caused 30% of these incidents resulted in the most severe problems especially when multiple hospitals in a region were inoperable (see Figure 2.1). This study points out the necessity of mitigation investments and mutual aid across regions for hospitals threatened by hurricanes and earthquakes. It is also noted that hospitals should realize the severity of internal threats causing 56% of the reported incidents and consider the increasing complexity of their facilities in their emergency management efforts.

In order to identify the prevalence, causes, and challenges of hospital evacuations, Bagaria et al. (2009) describe the results of an online literature review commissioned by the World Health Organization (WHO) on hospital evacuations. Twenty one articles, which gave descriptions of 69 hospital evacuations, were selected for study. Of the hospitals studied, only four reported a specific evacuation plan. For these evacuations, the two most common challenges were communication and logistics, including both the movement of patients within the hospital and the movement of patients to alternative care facilities (and presumably the coordination of these movements). Communication problems are highlighted in many other articles such as Henry (1980), Augustine and Schoettmer (2005), Palmer et al. (2003), Cyganik (2001), and Maese (2009). Logistical challenges were also reported by Gray and Hebert (2007) for the hospital evacuations in New Orleans in response to Hurricane Katrina and its aftermath.

Gray and Hebert (2007), Hamilton et al. (2009), and Broz et al. (2009) provide a summary of the disaster response of hospitals in New Orleans, Houston, and Chicago, respectively, and all suggest the implementation of region-wide patient and evacuee tracking systems. Cocanour et al. (2002) describes the emergency evacuation of the Texas Medical Center due to the loss of electrical power, water, and telephone service after flooding during tropical storm Allison. During this incident, nursing staff and medical equipment such as ventilators were sent along with some intensive care unit (ICU) patients to other hospitals with available beds; however, these transfers were not tracked which made the recovery of equipment very difficult. This suggests that medical equipment tracking systems are necessary as well as patient and evacuee tracking systems for achieving efficiency in the recovery phase of hospital emergency plans.

Although patient transportation strategies implemented are reported in the incident-based literature, there are few studies that compare various allocation strategies and investigate the implementation issues in detail. Sternberg and Lee (2009) explain the challenges for healthcare transportation under catastrophic disasters and identify the institutional models that can improve the disaster healthcare transportation assignment. As a city that had to learn from its emergency medical services (EMS) performance during the events of September 11, 2001, New York City is used as a case study considering that it is prone to hazards such as terrorism, pandemic influenza, hurricane, winter storm, and technological hazards. The institutional models proposed include both centralized and decentralized assignment of vehicles a combination of which can be used during a disaster. Centralized vehicle assignment should be performed by an emergency operations center that coordinates the organizations involved, but this assignment model can be overwhelmed by the excessive amount of information. Decentralized vehicle assignment can be performed if centralized services are disrupted and if mutual adjustment among medical responders and transportation officials is advantageous. However, decentralized resource management models would lack of information about the overall resource limitations which may pose issues regarding equitable resource allocation. Equitable vehicle assignment is especially important when multiple hospitals in a region are forced to evacuate and we study the equity issues under centralized and decentralized resource management strategies in Chapter 4 where hospitals compete for the limited resources.

Sternberg and Lee (2009) present the special scenario in which transportation assignment is required for a hospital evacuation as an especially difficult problem. It is noted that staggered evacuation can occur as a result of staff or building limitations and can actually reduce the loading area congestion. We discuss the interaction between building evacuation and loading capabilities of a hospital in Chapter 5 where we study the integrated building evacuation and transportation problem.

We note that the above selection of incident-based publications is a comprehensive subset of existing literature in that it encompasses a wide range of threats that hospitals are exposed to and the most prevalent issues in hospital evacuations that we discuss below.

2.1.1 Issues and Challenges of Hospital Evacuations

There are numerous issues that complicate the evacuation of a hospital, and an evacuation plan should address the most relevant and important of these issues. For instance, some of these issues include the following:

1. there is a diverse evacuee population, consisting of staff, visitors, and patients that require various levels of care;
2. most patients must be transferred to alternative care facilities using a limited fleet of vehicles consisting of different vehicle types that offer various degrees of patient care;
3. patients must be moved to a staging area (for loading onto transport) using a limited number of staff having diverse skill sets;
4. the nature of the threat or circumstance requiring the evacuation is unknown, and the various threats can greatly change the nature of an evacuation;
5. in-hospital transportation tools and processes change (e.g., the beds used for evacuation may be different than typical hospital beds and evacuation may require the use of stairs rather than elevators).

We will next discuss these issues in more detail through examples from the literature.

Challenges presented by the evacuee population

Hospitals have a diverse evacuee population including staff, visitors, and patients. The patient population itself can be quite varied, especially at larger hospitals, and can range from patients that are only in the hospital for observation, and can thus potentially "self-rescue", to patients requiring special care, for instance, mechanical ventilation, patients that must remain in isolation due to a contagious disease, and patients that require help due to mental maturity/age, to name just a few of the possibilities. Patients present a unique evacuation challenge as they have specific care requirements and many are non-ambulatory, thus they require care and assistance during the evacuation. Moving patients, especially those in critical condition, almost certainly reduces the level of care they receive, and thus puts them at risk. Because of the assistance and care required by patients, other evacuees, especially staff, can often have a dual role of evacuee and rescuer. Furthermore, as mentioned in the introduction, because patients require medical care, the evacuation process does not end with the evacuation of the building, but includes another step, that of transferring the patients to alternative care facilities. In some cases family members would want to be transported with the patient (e.g., pediatric patients) or can transport a patient who can be safely discharged. Also, medical staff might need to be transferred to help care for patients sent to other locations.

For evacuation planning purposes, we would like to categorize patients based on their condition and care requirements, and any other factors that impact the movement of patients in an evacuation. AHRQ defines some standardized hospital bed/patient types, which might be useful for evacuation planning purposes (see `http://archive.ahrq.gov/research/havbed/definitions.htm`). Their categories are Adult Intensive Care, Medical/Surgical, Burn Intensive Care, Pediatric Intensive Care, Pediatric, Psychiatric, Negative Pressure/Isolation, and Operating Room. These bed/patient types are not rigorously defined and potentially do not consider other factors that might also play a role in evacuation requirements, for example, "Is the evacuee a bariatric patient?," "Does the patient require a ventilator?," and "Can the patient walk?" To understand the usefulness of this categorization scheme we must consider if they represent the "best" level of detail. For instance, the AHRQ definition for pediatric patients use 17 years of age as the cut-off, but a 17 year old is much different from a three year old; how does this impact the evacuation? Conversely, the patient classification scheme influences the logistics of the evacuation. When there are many bed/patient types, the chance that the patients must be sent further afield to a correct type of bed (and thus care level) is higher. This, among other things, can increase transportation risk and reduce the utilization of the transportation resources. What is the best number of categories to be used? Considering such consequences that are usually unforeseen, the patient categorization scheme used for evacuation planning needs to provide the essential information about the patients without imposing restrictions that do not improve the efficiency of the evacuation plans.

Hospital visitors and staff are not specifically included in the model below because as populations they are generally considered to be able to perform "self-rescue"; in other words, they are able to evacuate themselves without the help of outside resources.

Numerous potential threats and scenarios

Hospital evacuation planning is difficult because of the variety of possible threats that can require that an evacuation take place, and the large number of possible scenarios these threats can produce. Often threats are characterized based on causes, for instance, the threat can either be natural, such as tornado, hurricane, flood, wildfire, or earthquake, or man-made, such as hazardous material releases, fire, loss of utilities, terrorist attacks, or bomb threats. Threats can also be categorized based on whether there is any advance notice of them; for instance, a hurricane is detected, and its path forecasted, well before it makes landfall, but there is no notice of an earthquake at all. The amount of warning inherent with any particular threat is of course an important characteristic to consider when planning, but other aspects of a threat are also important. Threats can have different impacts on the hospital's ability to treat patients. For instance, a loss of electric power (see TJC's Sentinel Event Alert Issue 37, September 6, 2006 "Preventing adverse events caused by emergency electrical power system failures") can affect patients of different types in a much different manner (e.g., any patients requiring mechanical ventilation will be much more severely impacted than patients in for observation). Likewise, threats can have an impact on the hospital's ability to evacuate. For example, the loss of electric power prohibits the use of elevators, which, depending on the structure of the hospital, can greatly reduce the ability of the staff and responders to move patients out of the hospital. As another example, a hurricane threat can greatly reduce the ability to move patients to other facilities if there is massive traffic congestion caused by the regional evacuation, and a hurricane might require that multiple hospitals evacuate, potentially causing a shortage of vehicles (and requiring coordination among the evacuating hospitals). The threat can also be complex; an earthquake that necessitates an evacuation might also cause a loss of power. Conditions not related to the threat might also impose severe restriction; a hospital evacuation that occurs during cold winter weather might have to be performed in a different manner than a similar evacuation on a mild day. Also, based on the location, magnitude, and time-variant characteristics of the threat, the hospital may be evacuated partially or fully, immediately or over a time period. How all these factors impact the evacuation plan is of interest.

Hospitals can develop strategies for population reduction such as discharging ambulatory and stable patients to reduce the number of evacuees for whom the resources will be used (Maxwell, 1982). When there is an immediate threat such as structural damage due to earthquake or flood, it is reasonable to evacuate the ambulatory and stable patients first and the critically ill patients last from the building, but to reverse the order for transferring patients to suitable receiving facilities. For example during the evacuation of Sepulveda VA Medical Center due to the Northridge earthquake, bedbound patients who mostly required ICUs were the last to be evacuated from the building, but first to be transported to other facilities as soon as vehicles were available (Chavez and Binder, 1995). Again during the Northridge earthquake, in Los Angeles County, California, five hospitals evacuated the most critical patients first since critical patients require more resources and the resources are rapidly consumed after an earthquake, but one hospital concerned about the possibility of collapse evacuated the healthiest patients first since it allows the evacuation of a large number of patients in a short time (Schultz et al., 2003). Both strategies worked well for the corresponding hospitals as different conditions call for different actions even in the same disaster area.

Challenges presented by the building

The plan for evacuating the building itself must consider the constraints inherent in the design of the facility. For instance, because patients as evacuees require assistance, the building evacuation requires evacuation routes and access routes, where access routes are used by the staff and first responders to access the patients and assist them to leave the building using the evacuation routes. Staff and first responders might be required to use these routes multiple times in order to move all the patients from the building, and may need to carry equipment when using the access routes. The evacuation and access routes should be selected to minimize potential interference and congestion. Infrastructure elements such as hallways, stairwells, elevators, and doors can all restrict movement and cause congestion that can potentially slow the evacuation, especially if a particular element is used both for evacuation and access.

As mentioned earlier, the threat can effectively "change" the design of the facility, requiring different evacuation and access routes. If the threat renders the elevators unusable (because of loss of power, for instance), then the best evacuation plan for this scenario would be much different than a plan that can utilize the elevators. The design of the facility also helps dictate the vulnerability to this type of "change". For instance, the evacuation plan for a high-rise hospital might be much more affected by a power loss than a low-rise hospital. The design of the facility (including the parking lots and surface streets) can also determine where the viable staging areas, (i.e., areas where patients are loaded into vehicles to be transported to an alternative care facility), can be located. Finally, the placement of the patients from the various patient categories in the hospital (i.e., in the various wards) can impact the evacuation plan (i.e., the patient type that is the most difficult to move may or may not be in the location that is the most difficult to evacuate).

Challenges presented by available responders and their skill-sets

To assist patients out of the building and to provide medical care during the evacuation process, responders (e.g., medical staff, other hospital staff, and first responders) are required. As mentioned above, in addition to the patient population, the evacuation of hospital staff and visitors must also be addressed in evacuation plans, but considering the dual role of some participants. Regardless of their position, some or all of the staff members can be involved in the evacuation operation, but this requires that the personnel be used in a proper way given their skill set (including in some cases, their ability to navigate the hospital without getting lost). One challenge to planning for this aspect of the evacuation is the variability in the staffing based on the time of day, and also on threat characteristics (e.g., in response to a hurricane, staff might have to leave the hospital to help their families leave the evacuation zone).

Another challenge about the first responders is the difficulty of coordination. It is observed in both natural and man-made disasters that large numbers of medical volunteers unexpectedly arrive at the threatened or affected areas to provide aid. However, such additional personnel expected to improve and expedite the emergency response usually intensify the disaster. Many responders may not be trained or equipped for the tasks they undertake even if they are physicians and may not be aware of the incident command protocols. The arrival of unannounced responders at an emergency incident scene is called *convergent volunteerism* and can be counterproductive despite the good intentions of volunteers (Cone et al., 2003). Increasing use of social media for mass communication makes convergent volunteerism more relevant to disaster response planning, so, it is unrealistic to eliminate the unsolicited responders. Therefore, the potential benefit of volunteers must be utilized by providing volunteers with education and guidance (Auf der Heide, 2003). To avoid the problems caused by the convergent volunteerism of physicians, hospital evacuation plans should identify how the volunteers can participate in evacuation efforts such that they can use their specialized skills without disrupting the operations.

Constraints based on available means of transportation

Vehicles are required to transfer patients from the evacuating hospital or hospitals to alternative care facilities. When planning for this process, the vehicle options include advanced life support ambulances (ALS), basic life support ambulances (BLS), wheelchair vans, buses, helicopters, airplanes, and boats. The emergency manager must determine the best fleet (i.e., the best number of each type of vehicle) and how to best use this fleet. This is not an easy problem; the best solution is quite dependent on the other aspects of the evacuation and the state of the system such as the number and type of patients to be moved, the number and type of beds available at other hospitals, and the type of disaster.

However, dispatching these vehicles is not an easy task during disasters. Initially, there might not be many vehicles available at the receiving hospital, but more could be gathered in time and the fleet size would be dynamic. When vehicles are available, the first difficulty is determining the patients to assign to a limited number of vehicles. Then, the receiving hospitals should be determined considering the available beds and equipment capable of treating those patients. Staff availability (in terms of number and skill-set) at the receiving facilities should be considered when transferring patients. While the Texas Medical Center was being evacuated due to the loss of electrical power, water, and telephone service after flooding during storm Allison, ventilators and nursing staff were sent along with some ICU patients to another hospital with available beds (Cocanour et al., 2002). In such cases, records of the transferred resources, including both equipment and staff, should be kept as well as the record of patients to keep track of the amount of available resources.

Constraints based on equipment

In order to move the non-ambulatory patients within and outside the hospital, equipment like stretchers, beds, or wheelchairs are required as well as the equipment necessary to treat patients. The mobility of this equipment (or the availability of mobile versions) along with characteristics such as battery life, can greatly impact the evacuation options. Coordination of the use of this equipment can complicate the evacuation plan.

Constraints based on available beds at alternative care facilities

A plan for transferring patients to receiving hospitals must consider the number of available (licensed and staffed) beds in each facility, along with the number of beds that can be added in response to an emergency, i.e., surge capacity. Surge capacity for hospitals is the ability to expand rapidly to meet an increased demand for medical care and hospitals mostly have a significantly limited surge capacity (Kaji and Lewis, 2006). Also, the implications of using surge capacity should be considered as there are finite limits to a hospitals medical surge capacity and capability (Barbera and Mcintyre, 2007). Often, an evacuating hospital must make use of space in multiple alternative facilities, greatly increasing the logistical complexity of the evacuation. Furthermore, when considering bed/patient types, mismatches can reduce the available capacity of usable beds. Finding a proper bed for a patient is especially difficult in time of a disaster since many hospitals are often near their capacity under normal conditions, and by definition an evacuation is reducing the regional stock of hospital beds. The number and type of patients at the evacuating hospital, and the number and type of available beds at the potential receiving facilities, are always changing, thus, a hospital evacuation plan must be able to be dynamically updated. Hospital evacuations are also complicated when more than one hospital must be evacuated due to regional threats, like hurricanes or earthquakes. In these cases, the use of the available beds must be coordinated between the evacuating hospitals and/or a regional coordinating center.

Hospital evacuations are subject to more difficult issues than other building or regional evacuations due to their distinct evacuee population. The variety of patients with different conditions and medical care requirements results in particular assignment and scheduling constraints. All types of resources that can be used during the building evacuation or the transfer of patients to alternative health care facilities depend on the patient type. Having a mix of critical and ambulatory patients is especially a restricting factor when there are scarce transportation resources to be distributed among patients, and emergency evacuation plans need to determine the most efficient transportation schedule accordingly. There is a wide variety of evacuation planning models with different assumptions and measures of effectiveness that have been developed for hospital evacuations. However, the literature on evacuation practices based on the current plans indicates that these plans are not fully capable of dealing with the complexities of the hospital evacuation problems. Therefore, identifying the issues involved in the process and studying methods that would reduce the complexity of hospital evacuations is indispensable for the improvement of current emergency evacuation plans and optimization models.

2.2 Mathematical Modeling Literature

2.2.1 Building evacuations

General building evacuations have been extensively studied in the operations research literature for over fifty years. However, the repository of building evacuation models relies on the mobility of the evacuee population without any additional support for initiating movement. For a hospital, the evacuee population consists of patients who need assistance to move within or out of the building as we introduce modeling assumptions in Chapter 5. Therefore, hospital evacuations are much different than regular evacuations in this context. Although the existing building evacuation models provide a basis for hospital evacuation models, they do not meet the healthcare system requirements sufficiently.

The prevalent objective of general building evacuation problems is minimizing the total evacuation time (or building clearance time) and these problems are generally formulated as dynamic (time-expanded) network flow problems and solved by applying the corresponding network flow algorithms, such as maximal flow and minimum cost flow algorithms, see Ford and Fulkerson (1962) for these algorithms.

Chalmet et al. (1982) develop three deterministic network models for building evacuation assuming constant capacity and travel time for each arc: dynamic transshipment model, graphical model, and intermediate model. The dynamic model has the capability of triply optimizing the building evacuation by (i) minimizing the average time each person needs to exit the building, (ii) maximizing the number of people evacuating the building at each time period (equivalently, maximizing the flow in the first p periods for every p), and (iii) minimizing the time period in which the last evacuee leaves the building. Equivalencies for these objectives (also called the triple optimization result), for certain dynamic network flow problems with constant arc travel times, were derived in Jarvis and Ratliff (1982). The other two models are easier to use than the dynamic model and can provide insights on the bottlenecks, but they do not evolve over time, i.e., they are time-independent.

Building on the same type of models, Hamacher and Tufecki (1987) show that building evacuation models with multiple objectives (such as minimizing the total evacuation time and avoiding cycling of evacuees or evacuation with multiple priority levels for different parts of the building) can be solved as lexicographic minimum cost flow problems. Choi et al. (1988) incorporate flow dependent capacities into building evacuation network which adds side constraints to the problem. Chen and Miller-Hooks (2008) formulate the building evacuation problem with shared information on the changes in evacuation routes as a mixed integer linear program that minimizes the total evacuation time. This problem is solved using an exact algorithm based on Benders decomposition. For a more detailed review of the literature on evacuation models for buildings see Hamacher and Tjandra (2001).

The majority of evacuation models minimize some function of the evacuation time. Questioning the effectiveness of such objectives for different scenarios, Han et al. (2007) discuss various evacuation objectives minimizing measures of effectiveness including individual travel (or exposure) time, time-based risk exposure, and time- and space-based risk exposure. A combination of these measures of effectiveness can be implemented in multiobjective evacuation optimization problems. Correa et al. (2007) study the minimization of maximum latency of network flows with congestion and analyze the results for three different objective functions: (1) minimizing the maximum latency (min-max flow), (2) minimizing the average latency (system-optimal flow), and (3) maximizing fairness of latency of flow-carrying paths (Nash flow). It is shown that an optimal flow that has the minimum average latency is near-optimal for the objective of minimizing the maximum latency and close to being fair. Løvås (1995) assumes a building evacuation network with stochastic variables and discusses performance measures related to accident effects, evacuation time, queueing and waiting, network distances, and network redundancy. Probability distributions of the number of untrapped people and the number of safe evacuees at a time interval are considered as the most important measures in terms of safety. The expected time spent in a node, the average time spent in a node per visit, and the average number of people visiting a node are suggested to identify bottlenecks in the system.

The building evacuation studies reviewed above do not specifically answer the evacuation problems of healthcare facilities, but they certainly provide insights about evacuation modeling that are utilized in this dissertation. We next discuss the hospital evacuation modeling literature.

2.2.2 Hospital evacuations

The building evacuation literature reviewed above assume able-bodied evacuee populations. Hospital evacuations, unlike general building evacuations, involve evacuees that require extensive assistance; therefore, these problems need to be handled in a distinctive manner. Furthermore, a hospital evacuation involves more than the safe and efficient clearance of the building. An equally important aspect of the evacuation is to transport patients to appropriate alternative care facilities. The literature is quite scarce in hospital evacuation modeling pertaining to both the hospital building evacuation and the subsequent transportation of patients. There are only a few heuristic tools available for evacuation planning that do not necessarily generate plans at an operational level. One of these tools is the Hospital Evacuation Decision Guide (see `http://archive.ahrq.gov/prep/hospevacguide`), prepared for the Agency for Healthcare Research and Quality (AHRQ), that discusses the importance of estimating evacuation time, among other things, to support the decision to evacuate and the timing for this decision. The evacuation time (i.e., how long it takes to evacuate the hospital) is dependent on the resources available and how efficiently they are used. This metric is difficult to estimate. Because of this, a model to estimate the evacuation time, called the Mass Evacuation Transportation Model (METM) (see `http://archive.ahrq.gov/prep/massevac`), was developed for AHRQ. The objective of the METM is to produce an estimate of the time required to transport patients from the evacuating hospital(s) to alternative care facilities based on the vehicle loading, transit, and unloading times. METM uses a heuristic that sends the patients with the highest criticality to the closest hospitals with available capacity. This model makes a simplifying assumption that all beds are equivalently capable of serving patients with any level of criticality which, in practice, is inaccurate due to the wide range of medical requirements of patients. It also assumes that patient types determined by criticality are strictly assigned to a vehicle type which precludes flexible vehicle assignment. Moreover, the number of patients that can be loaded onto vehicles is only limited by the number of vehicles available and does not consider other possible limiting factors, such as staffing or physical space limitations. We note that the METM is currently not available due to problems found by the research team involved in the project of which this work is a substantial part.

Another study that estimates the evacuation time by Duanmu et al. (2010) focuses on the routing of hospital vehicles during a hurricane evacuation where the ambulances and general traffic compete for space in the regional traffic flow network. The ambulance trip times are estimated using a simulation model based on various hospital evacuation start times and multiple strategies that minimize the transportation time for patients are produced. This study does not consider any patient-specific attributes or requirements. Golmohammadi and Shimshak (2011) estimate the evacuation time for the hospital building evacuation using a predictive model that takes patient population and available resources as input and calculates the total evacuation time. Three patient types are defined based on mobility and the patients who are the fastest to evacuate are given the first priority. This patient prioritization rule is analogous to the shortest processing time rule in scheduling theory and can significantly increase the waiting time of the most critical patients.

Hospital evacuations can be approached from various modeling perspectives including project management, mathematical modeling, simulation, and hybrid models Taaffe et al. (2005). Although hospital evacuations are not specifically addressed, simulation is the most common approach in general evacuation studies among these options. For examples of simulation-based evacuation studies, see Sheffi et al. (1982), Tufekci (1995), de Silva et al. (1996), Hobeika and Kim (1998), and Taaffe et al. (2006). These simulation models are aimed at mass evacuations and do not clarify the transportation requirements of special evacuee populations such as hospital patients.

As we can see from both the incident-based literature and the mathematical models, a major issue in evacuation planning is patient prioritization. Childers et al. (2009) study the evacuation of a healthcare facility using dynamic programming and simulation to develop patient prioritization guidelines. The optimal patient prioritization that maximizes an expected reward in terms of the number of patients evacuated is determined based on the numbers of critical and non-critical patients remaining patients in the system. Jacobson et al. (2012) formulate the triage problem arising in the aftermath of mass-casualty events for distributing resources as a priority assignment problem in a clearing system with impatient jobs, similar to a job-scheduling problem, that maximizes the expected number of survivors. They suggest determining the priority of patients based not only on their medical condition, but also on resource limitations and use stochastic dynamic programming to identify optimal triage policies. See Argon et al. (2008) for a review of stochastic job scheduling literature in relation to the patient triage problem.

Hospital evacuation requires the allocation of limited resources to patients to mitigate the risk exposure and equity of this resource allocation becomes a measure of interest. Leclerc et al. (2012) review definitional issues in modeling equity for allocation of public resources and discuss the implications of different functional forms of equity in mathematical programming models. Bertsimas et al. (2012) study the trade-off between efficiency and fairness in resource allocation problems based on functions used in welfare economics and develop managerial prescriptions on the objective function selection for utility allocation. Equity and fairness issues have been studied in various contexts that involve resource allocation such as queuing theory (Avi-Itzhak et al., 2008), network flows (Correa et al., 2007), hazardous materials transportation (Gopalan et al., 1990), and organ transplantation allocations (Zenios, 2005).

Tayfur and Taaffe (2009) introduce a transportation planning model for hospital evacuations that minimizes the total costs (for vehicle leasing, staffing, and transportation) given a pre-specified evacuation duration. The model determines the staffing and fleet requirements that minimize the costs, so this model focuses on the financial burden of hospital evacuation rather than the patient safety. This model makes the restrictive assumption that all receiving hospitals are equidistant from the evacuating hospital, requiring one time interval of travel time, which is equivalent to the time required to load (and unload) the patients, in order to improve the solvability of the model.

2.3 Equity Modeling Literature

Equity, in its most general sense, is concerned with the fair distribution of resources, goods, income, rights, opportunities, or access to services among the entities in society. Being a universal concern, equity has been studied predominantly by philosophers, statisticians, sociologists, and economists. Distributional equity is important in a wide range of areas of public interest including taxation, the healthcare system, the education system, and the allocation of representation among political parties. In management sciences, equity has been studied in several contexts such as transportation systems design, risk analysis, facility location analysis, and allocation of helathcare services. As an essential part of public services, disaster management has been recently more concerned with equity after focusing mostly on the efficiency of service delivery. Similarly, in hospital evacuation planning, we will evaluate the performance of evacuation plans in terms of the equity of resource allocation. Due to its subjectivity, equity has no universally accepted definition and appropriate equity models should be developed for each specific problem. Equity models developed in economics are applicable to a wide range of management science problems and the relevant economics literature on equity measurement is reviewed below.

Definition and measures of inequality have been studied extensively in sociology and economics to compare the distribution of assets among social units (see Sen (1973) for a discussion of economic inequality measures). In economics, a basic theory for measuring inequality that is independent of welfare economics is developed. All valid measures of inequality are assumed to be positive-valued if assets are not equally distributed and zero otherwise. They are also assumed to be independent of which unit has what level of assets. According to this basic theory, the measures of inequality are consistent with three axioms: (1) principle of transfers, (2) population symmetry, and (3) scale invariance (Schwartz and Winship, 1980). *Principle of transfers* (also known as the Pigou-Dalton principle of transfers) requires the inequality to be reduced when assets are transferred from a unit to a worse-off unit. *Population symmetry* requires that two populations of equal size having the same mean assets have identical inequality. According to this axiom, two populations having the same mean assets but different numbers of units, m and n, can be compared by multiplying the first population by n and the second population by m. *Scale invariance* requires the inequality to remain unchanged when every unit's assets are multiplied by the same constant, in other words, increased or decreased by the same proportion. The most commonly used measures of inequality are the coefficient of variation, the mean relative deviation, the Gini coefficient, Theil's measure, and the standard deviation of the logarithms of each unit's assets. However, not all of these measures are consistent with the above three axioms as critiqued by Allison (1978) and Schwartz and Winship (1980).

An alternative approach to the basic theory depends greatly on welfare economics and utilizes the social welfare function for measuring inequality. This approach is closely related to the Lorenz curve that represents the cumulative proportion of assets possessed by the cumulative proportion of population ranked in ascending order of their income. If all assets are equally distributed, the Lorenz curve will be a straight diagonal line such that $x\%$ of assets are possessed by the poorest $x\%$ of the population, where $x \in [0, 100]$ (Atkinson, 1970). In case of any unequal levels of assets among the units in population, the Lorenz curve will be below this perfect equality line, which means that the poorest $x\%$ of the population possesses less than $x\%$ of the assets for some values of x. If a Lorenz curve for a given distribution A of assets is always below the Lorenz curve for another distribution B, then any scale-invariant measure of inequality that satisfies the principle of transfers will be larger for distribution A, representing a greater inequality.

Atkinson (1970) uses the relationship between the Lorenz curve and the social welfare function of utilities to define a measure of inequality that depends on a single inequity aversion parameter. The form of the utility function considered is based on one of the earliest discussions of the welfare function provided by Pratt (1964). In this study, utility $u(x)$ is defined as a function of assets or money (x). The risk premium $\pi(x, \tilde{z})$ is the value of risk $\tilde{z}$ such that the decision maker is indifferent between receiving a random risk $\tilde{z}$ and receiving a non-random amount $E(\tilde{z}) - \pi$. Utility function when $E(\tilde{z}) - \pi$ is received, $u(x + E(\tilde{z}) - \pi)$, is a strictly decreasing continuous function of the risk premium over all possible values of u. When the decision maker's local aversion to risk is measured, the risk premium for a small risk $\tilde{z}$ with $E(\tilde{z}) = 0$ and small variance σ_z^2 is considered. When the utility of receiving $E(\tilde{z}) - \pi$ and the expected utility of receiving $\tilde{z}$ are set equal by definition of the risk premium, the following relation is obtained.

$$\pi(x, \tilde{z}) = \frac{1}{2}\sigma_z^2 \left[-\frac{u''(x)}{u'(x)} \right] + o(\sigma_z^2)$$

The value $\alpha(x) = -\frac{u''(x)}{u'(x)} = -\frac{d}{dx}\log u'(x)$ is defined as the local measure of risk aversion (risk aversion in the small). This value is also a measure of concavity of u at point x. In order to define the relation between u and $\alpha(x)$, $-\alpha(x)$ is integrated (resulting in $\log u'(x) + c$, where c is a constant), exponentiated, and integrated again (resulting in $\exp^c u(x) + d$, where d is a constant). Since $\exp^c u(x) + d$ is equivalent to $u(x)$ as utility, u is equivalent to $\int \exp^{-\int \alpha}$.

If risk premium is viewed as a proportion of assets, $\pi^*(x, \tilde{z})$ is the proportional risk premium for a proportional risk $\tilde{z}$ such that the decision maker with assets x and utility function u is indifferent between receiving the random risk $x\tilde{z}$ and receiving the non-random amount $E(x\tilde{z}) - x\pi^*(x, \tilde{z})$. Then, $\alpha^*(x) = x\alpha(x)$ is the local proportional risk aversion at point x. The utility function takes the following forms given various constant values of $\alpha^*(x)$.

$$\begin{aligned} u(x) &\sim x^{1-c} && \text{if } \alpha^*(x) = c < 1 \\ u(x) &\sim \log x && \text{if } \alpha^*(x) = 1 \\ u(x) &\sim -x^{1-c} && \text{if } \alpha^*(x) = c > 1 \end{aligned}$$

These possibilities for the utility function provide the piecewise welfare function, $W_\alpha(u)$, defined as below.

$$W_\alpha(u) = \begin{cases} \sum_{k=1}^{n} \frac{u_k^{1-\alpha}}{1-\alpha} & \text{for } \alpha \geq 0,\ \alpha \neq 1, \\ \\ \sum_{k=1}^{n} \log(u_k) & \text{for } \alpha = 1 \end{cases}$$

The social welfare function, $W_\alpha(u)$, is implemented in the evacuation model proposed for multiple hospitals in Chapter 4 in order to evaluate the equity of optimal solutions. Utilities are defined as functions of evacuation risks that inherently represent the resources allocated to each entity.

After this review of the literature on emergency management, hospital evacuations, and equity modeling, we present the research within the scope of this dissertation including supplementary reviews of the relevant literature on specific topics discussed.

Chapter 3

Decision Support for Hospital Evacuation and Emergency Response

This chapter addresses the development of an optimization-based planning and decision support model for hospital evacuations. We propose an integer programming model for the evacuation transportation problem that minimizes the expected evacuation risk subject to limited resources and patient care requirements. The model allocates patients, categorized by criticality and care requirements, to a limited fleet of vehicles of various capacities and medical capabilities, to be transported to appropriate receiving hospitals considering the current available space in each hospital for each category of patient. We demonstrate the structural properties of the model that enable computational efficiency for solving large-scale problems for both planning and operations purposes and the robustness of the model under risk uncertainty.[1]

3.1 Introduction

Hospitals must take great care in emergency planning and management; the importance of hospital evacuation planning is highlighted by hospital accreditation standards (see Joint Commission Hospital Accreditation Standard EM.02.02.03 and others) requiring hospitals and other care facilities to develop evacuation plans and complete a significant amount of

[1]This chapter appeared in the Annals of Operations Research, DOI: 10.1007/s10479-011-0943-y.

disaster planning annually. For an emergency manager, an evacuation is probably the emergency response requiring the most extensive planning because of the complexity and risks involved. Hospitals have had to evacuate for many reasons; examples include hurricanes, fires, floods, chemical leaks, earthquakes, bomb threats, and loss of functionality, see, for example, Chavez and Binder (1995); Schultz et al. (2003); Sternberg et al. (2004). Evacuation is an important risk management tool, and should be considered whenever patients (and staff) are exposed to excessive risk. Although evacuation plans are subject to resource limitations and budget considerations, the main purpose of a hospital evacuation is to minimize risk to the patients (and staff).

Specifically, in this study we consider the problem of transporting patients from an evacuating hospital to a set of receiving hospitals. Transporting patients, many of them seriously ill, is inherently risky and one of the reasons hospital evacuations are so difficult. As expected, patients do not form a homogeneous population, and are categorized based on their condition and care requirements. Examples of these patient types include critical care, pediatric, pediatric critical care, and infectious disease patients (e.g., those having drug-resistant tuberculosis and thus requiring isolation, for instance). Hospital capacity is measured in terms of beds, and the beds are categorized much like the patients (the term "bed" denotes the equipment and staff that go along with the bed). Thus, when assigning patients to another hospital, the number of available beds for each patient type must be considered. For example, a critical care patient should not be placed in a pediatric bed or into a bed specified for less critical patients, as the equipment and staffing would not be appropriate. As we will see, patient type also plays an important role in risk assessment, and thus in the decision making process. The different transportation resources have varying capacities and provide different levels of care. For instance, Advanced Life Support (ALS) ambulances, staffed with an Emergency Medical Technician-Paramedic (EMT-P), Basic Life Support (BLS) ambulances, staffed with an Emergency Medical Technician (EMT) (who has less extensive training than an EMT-P), air ambulances, wheelchair vans, and buses can all be utilized, depending on availability and need.

The two main sources of evacuation risk are the threat risk (the reason for the evacuation) and the transportation risk. The threat risk characteristics that affect the evacuation plan are the impact of the threat on the various patient types and how threat evolves over time. For instance, a hurricane would pose no immediate threat risk to the patients (as it can be forecasted well before landfall), but eventually the threat risk might be considerable; a long term power outage would affect critical patients dependent on lifesaving equipment more than patients in the hospital for observation, while a fire could pose a serious risk for all patient types. In fact, in certain situations, a partial evacuation might be appropriate to reduce risk; certain patients might be safer if evacuated whereas others might not tolerate the transportation risk. The transportation risk is a function of the patient type, the selected vehicle's attributes, and the time required to transport the patient to the selected receiving hospital.

An evacuation plan should minimize the overall evacuation risk considering the limited resources available, i.e., the number of each type of vehicle and when they are available, and the bed availability at the receiving hospitals. These decisions are difficult to make in advance as the modeling parameters, such as the number of patients at the evacuating hospital, the available logistic resources, and the beds available at the receiving hospitals are dynamic. There are also many potential disaster scenarios that might require a hospital to evacuate, the characteristics of which can greatly impact the plan. Thus, evacuation decision support tools that can help produce an appropriate plan, given more detailed information on the particular situation, are essential. This is especially important as there is almost always a tight planning window and resources usually need to be organized quickly.

General building evacuations have been extensively studied in the Operations Research literature, for instance, see Chalmet et al. (1982); Hamacher and Tufecki (1987); Choi et al. (1988); Chen and Miller-Hooks (2008); for a more detailed review of the literature on evacuation models for buildings see Hamacher and Tjandra (2001). These evacuation studies generally use network flow models to minimize the overall duration of the evacuation or the average evacuation time. Equivalencies for these objectives, for certain problems, were derived in Jarvis and Ratliff (1982). Han et al. (2007) discuss potential evacuation objectives for regional evacuations, including the consideration of risk. Hospital evacuations, unlike a general building evacuation, have evacuees that require extensive assistance. Furthermore, a hospital evacuation involves more than the safe and efficient clearance of the building.

An equally important aspect of the evacuation is to transport patients to appropriate alternative care facilities. The literature that studies planning models for hospital evacuations is quite sparse. Childers et al. (2009) study the evacuation of a health-care facility using dynamic programming and simulation to develop a patient prioritization scheme. Tayfur and Taaffe (2009) introduce a transportation planning model for hospital evacuations that minimizes the total labor and equipment costs given a pre-specified evacuation duration. The web-based (see `http://archive.ahrq.gov/prep/massevac`) Mass Evacuation Transportation Planning Model (METPM), developed for the Agency for Healthcare Research and Quality (AHRQ), estimates the time required to evacuate one or more hospitals using a heuristic that sends the patients with the highest criticality to the closest hospitals with available capacity. This model assumes all beds are equivalent and patient types are strictly assigned to a vehicle type. Since duration is considered an important metric in the METPM, and is commonly used in other evacuation settings, in Section 3.4 we contrast this duration minimization approach with the risk minimization approach advocated in this work.

The remainder of this chapter is organized as follows. In Section 3.2 we present the modeling assumptions and formulation, a discussion of the structural properties of the model, and a discussion of risk and estimation issues. In Section 3.3 we present a case study based on the evacuation of a large hospital. In Section 3.4 we provide a numerical analysis using the case study, including a comparison of the risk minimizing and duration minimizing objectives and an analysis of the robustness of the risk minimizing solution. Finally, in Section 3.5 we conclude with a discussion on the practical implications of this work and indicate some future research directions.

3.2 Hospital Evacuation Transportation Problem

In this section we present the Hospital Evacuation Transportation Model (HETM), including a discussion of the assumptions and a model formulation. Then we examine an important aspect of the model's structure, followed by a general discussion of risk.

3.2.1 Formulation of the HETM

Consider an evacuating hospital H^0 having W_j patients of type $j \in P$, where P might include types such as critical, pediatric, and infectious. To evacuate H^0, a plan to transport the patients to a set of potential receiving hospitals (H), where hospital $i \in H$ has B_{ij} beds available for patients of type $j \in P$, must be developed. The study period is divided into T time intervals of equal length. The travel time from H^0 to receiving hospital $i \in H$ is τ_i time intervals. A set of vehicle types (V), having different capabilities is available; for example, vehicle types can include Advanced Life Support (ALS) ambulances, Basic Life Support (BLS) ambulances, and buses. The N_{kt}-parameter is the total number of vehicles of type $k \in V$ in use or waiting to be loaded in time interval t. This parameter has a temporal aspect because not all vehicles are immediately available, for instance, some must travel to reach H^0. The capacity of a vehicle of type k is C_k. The γ_k-parameter is the loading and unloading time for a vehicle of type $k \in V$. The number of vehicles that can be loaded at H^0 in any time interval is represented by the L-parameter, and is given in ambulance equivalencies; L_k represents a conversion factor for a vehicle of type k to ambulance equivalencies. For example, loading a bus requires more intervals than an ambulance and more resources (staff and space) per interval. The limitation on loading can be based on staffing limitations and/or the physical limitations of the loading area.

The objective is to evacuate the hospital in such a way as to minimize the total evacuation risk. Here we define risk as the probability that a generic adverse event E occurs, such as a major deterioration of patient health or death due to injury or lack of sufficient treatment. The evacuation risk is a combination of: 1) the threat risk, which is the risk patients are exposed to while waiting to be transported, and 2) the transportation risk, which is the risk incurred during travel. The threat risk parameter, α_{jt}, is the probability of event E for a patient of type $j \in P$ that remains in H^0 in time interval t. The *cumulative threat risk*, Λ_{jt}, is calculated in (3.1). Λ_{jt} is the probability of event E for a patient of type j that remains in the hospital through time interval t. A patient of type j transported in the first interval would not be exposed to any threat risk, thus $\Lambda_{j0} = 0$, and Λ_{jT} is the risk associated with not being evacuated from H^0. An important assumption used in calculating the threat risk is that the risks to which patients are exposed in different time intervals, i.e. α_{jt}, are independent, and that a patient does not change type during an evacuation, e.g., a non-critical patient does not become critical by being exposed to the threat risk

$$\Lambda_{jt} = 1 - \prod_{f=1}^{t}(1 - \alpha_{jf}), \quad \forall j \in P, \; t = 1, \cdots, T. \tag{3.1}$$

The transportation risk parameter, β_{jk}, is the probability of event E for a patient of type $j \in P$ being transferred by a vehicle of type $k \in V$ for one time interval. This risk is assumed to be constant through time since the same allocation made in different time intervals would not change the level of treatment supplied by a vehicle. The *cumulative transportation risk*, Θ_{ijk}, is calculated in (3.2). Θ_{ijk} is the risk for a patient of type $j \in P$, transported by a vehicle of type $k \in V$, to receiving hospital $i \in H$. When transporting patients to receiving facility i, a vehicle of type k is engaged for $2(\tau_i + \gamma_k)$ time units. However, only $(\tau_i + 2\gamma_k)$ time units contribute to the risk expression since patient transportation risk is incurred only as long as the patient is in the vehicle, including loading and unloading times

$$\Theta_{ijk} = 1 - (1 - \beta_{jk})^{(\tau_i + 2\gamma_k)}, \quad \forall i \in H, \; j \in P, \; k \in V. \tag{3.2}$$

The evacuation risk, R_{ijkt}, associated with the evacuation decision for a patient of type $j \in P$, is then calculated in (3.3) by combining the threat risk (Λ_{jt}) to which the patient is exposed before being transported in time interval t, and the transportation risk (Θ_{ijk}) based on the vehicle type k and the receiving hospital selected

$$R_{ijkt} = 1 - (1 - \Lambda_{jt})(1 - \Theta_{ijk}), \quad \forall i \in H, \; j \in P, \; k \in V, \; t = 1, \cdots, T. \tag{3.3}$$

As (3.1) - (3.3) indicate, the cumulative threat, transportation, and evacuation risk functions are non-linear.

Decision Variables:

x_{ijkt} : number of patients of type j transported to hospital i by vehicle type k starting in time interval t, $\forall i \in H,\ j \in P,\ k \in V,\ t = 1, \cdots, T$.

y_{ikt} : number of vehicles of type k that transport patients to hospital i starting at time interval t, $\forall i \in H,\ k \in V,\ t = 1, \cdots, T$.

Modeling Assumptions:

1. The hospital building can be evacuated such that patients of the appropriate type are available to satisfy the transportation plan within the limit established by the L-parameter.

2. Vehicles transport patients directly from the evacuating hospital H^0 to a receiving hospital $i \in H$, that is, a vehicle will not stop at multiple receiving hospitals to off-load patients at each.

3. Travel times between H^0 and the receiving hospitals have integral length, are known with certainty, and are independent of vehicle type. We note that it is easy to modify the HETM such that travel times are vehicle-dependent.

4. Both the time required to load a patient into a vehicle and the vehicle capacity are assumed to be independent of the patient type.

5. The loading time of a vehicle of type k, given by γ_k, is assumed to be equal to the unloading time. This assumption can be easily modified.

6. A patient of type $j \in P$ should be assigned to the corresponding type of bed, i.e., a type of bed that is fully equipped for treating this type of patient.

7. The transportation risk is assumed to be a function of the travel time, and the loading and unloading time.

The Integer Programming (IP) formulation of the model is as follows:

HETM:

$$\text{Minimize} \quad \sum_{i\in H}\sum_{j\in P}\sum_{k\in V}\sum_{t=1}^{T} R_{ijkt}x_{ijkt} + \sum_{j\in P}\Lambda_{jT}\left(W_j - \sum_{i\in H}\sum_{k\in V}\sum_{t=1}^{T} x_{ijkt}\right) \tag{3.4}$$

subject to

$$\sum_{i\in H}\sum_{k\in V}\sum_{t=1}^{T} x_{ijkt} \le W_j, \quad \forall j \in P \tag{3.5}$$

$$\sum_{j\in P} x_{ijkt} \le C_k y_{ikt}, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T \tag{3.6}$$

$$\sum_{i\in H} y_{ikt} + \sum_{i\in H}\sum_{f=1}^{\min\{2(\gamma_k+\tau_i)-1,t\}} y_{ik(t-f)} \le N_{kt}, \quad \forall k \in V,\ t = 1, \cdots, T \tag{3.7}$$

$$\sum_{k\in V}\sum_{t=1}^{T} x_{ijkt} \le B_{ij}, \quad \forall i \in H,\ j \in P \tag{3.8}$$

$$\sum_{i\in H}\sum_{k\in V}\sum_{f=t-\gamma_k+1}^{t} L_k y_{ikf} \le L, \quad \forall t = 1, \cdots, T \tag{3.9}$$

$$x_{ijkt} \ge 0 \text{ and integer}, \quad \forall i \in H,\ j \in P,\ k \in V,\ t = 1, \cdots, T \tag{3.10}$$

$$y_{ikt} \in \{0, 1, \cdots, \lceil L/L_k \rceil\}, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T. \tag{3.11}$$

The objective function (3.4) minimizes the sum of the evacuation risk (first term) for all the patients transferred from the evacuating hospital by time T and the sum of the cumulative threat risk (second term) for any patients not evacuated by time T. Constraint (3.5) defines the total number of patients, by patient type, in the evacuating hospital. Constraint (3.6) limits the number of patients that can be evacuated in interval t to the capacity of the vehicles selected to be loaded in t. Constraint (3.7) limits the number of vehicles, of each type, in use in interval t to those available in t, while (3.8) designates the number of beds of each type available in each potential receiving hospital. Constraint (3.9) limits the number of vehicles that can be loaded at any one time. The parameter L is the number of ambulances that can be loaded in a single time interval, and L_k is a conversion factor for other vehicle types. γ_k is the number of intervals required to load a vehicle of type k. Constraints (3.10)-(3.11) are the logical integrality and non-negativity constraints.

At this point it is appropriate to compare the HETM with other hospital evacuation models considered in the literature. Tayfur and Taaffe (2009) introduce a hospital evacuation resource requirement planning model that determines the staffing and fleet requirements to complete an evacuation within a certain time frame while minimizing costs (for vehicle leasing, staffing, and transportation). Thus, the Tayfur and Taaffe (2009) model has a much different strategic focus than the HETM model presented here. The Tayfur and Taaffe (2009) model also makes the simplifying assumption that all receiving hospitals are equidistant from the evacuating hospital, requiring one time interval of travel time, which is equivalent to the time required to load (and unload) the patients, thus each vehicle is engaged for four time intervals for each trip (this simplifying assumption is used to improve the solvability of the model). On the other hand, the METPM, developed for AHRQ, estimates the time required to complete an evacuation of one or more hospitals based on the vehicle loading, transit, and unloading times. The METPM requires that patients be classified by vehicle type (based on criticality) and strictly assigns patients to vehicles based on this classification. For instance, a patient with a BLS ambulance-classification will only be assigned to a BLS ambulance, not an ALS ambulance (an upgrade in care service) or a bus (a downgrade). Furthermore, in this model all beds are assumed to offer the same level of care, so the concept of a bed type is not used. The heuristic solution gives the most critical patients the highest priority and sends them to the closest hospital with remaining capacity. In the METPM the number of patients that can be loaded onto vehicles is only limited by the number of vehicles available, however, there are usually other limiting factors, including staffing limitations and limits on physical space, which can have a substantial impact.

3.2.2 Structural analysis of the HETM

While the HETM formulation does not closely resemble a vehicle routing problem (VRP) formulation, this transportation problem is a special case of the VRP (see Toth and Vigo (2002), for instance, for an overview of the VRP). In Figure 3.1 we provide a simple illustrative network including H^0, which functions somewhat like a depot in the VRP, along with two receiving hospitals and the connecting arcs with traversal times.

Unlike most VRP networks, the network in Figure 3.1 is not fully connected. This is based on our assumption that vehicles visit only one receiving hospital before returning to H^0. While using only such simple routes is a simplification of the VRP, this problem also includes

Figure 3.1: A VRP network representation.

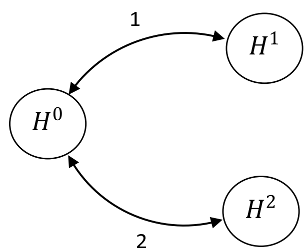
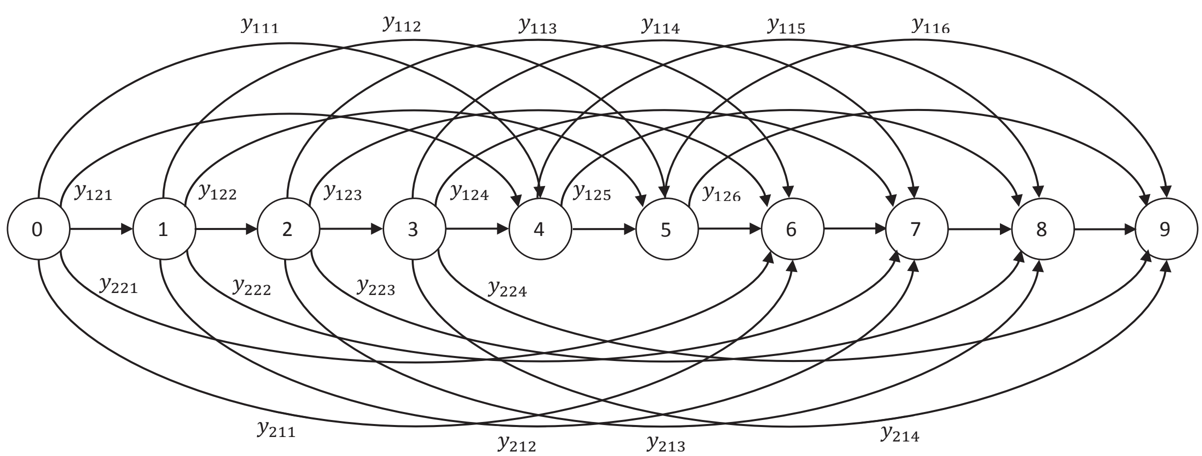


Figure 3.2: A network illustration.



many complications that preclude using a VRP-type modeling framework, including the non-linear risks. The HETM formulation is therefore based on a different network structure, as illustrated in Figure 3.2, which corresponds to the VRP network displayed in Figure 3.1 with two vehicle types. This is a time-expanded network, where node zero represents H^0 at the beginning of the evacuation and each of the following nodes represent H^0 at the end of the specified time interval. The y_{ikt}-label arcs are transportation arcs. For example, the arc labeled y_{121} represents a vehicle of type 2 going from H^0 to H^1 and back again, starting in time interval 1. The length of this arc represents the time a vehicle of type 2 is engaged when transporting a patient to H^1, including patient loading and unloading times (which are vehicle-dependent) and the round trip travel time. The unlabeled arcs directly connecting two consecutive nodes are vehicle-waiting arcs; flow over these arcs represents vehicles waiting to be utilized. The flow value into node t can be calculated as $N_{kt} - \sum_{i\in H} y_{ikt} - \sum_{i\in H}\sum_{f=1}^{\min\{2(\gamma_k+\tau_i)-1,t\}} y_{ik(t-f)}$ for vehicle type k, see 3.7. Each transportation arc has an associated x_{ijkt}-variable and evacuation risk parameter R_{ijkt} for each of the $|P|$ patient types. Thus, the network in Figure 3.2 represents the possible vehicle movements through time. A set of feasible y-variables essentially determines a viable sub-network, and given that sub-network, the HETM has a structure amenable to standard solution techniques. The structure is described in the following proposition.

Proposition 3.1. *For a feasible set of y-variables and a continuous relaxation of the x-variables, there exists an optimal solution in which the x-variables have integral values.*

Proof. A feasible set of y-variables must satisfy constraints (3.7) and (3.9), and be integer valued - thus satisfying (3.11). Given a set of feasible y-variables, which now can be considered as parameters, and a continuous relaxation of the x-variables, we transform the HETM constraint set (3.6)-(3.11) into the following:

$$\sum_{j\in P} x_{ijkt} \leq C_k y_{ikt}, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T \tag{3.12}$$

$$\sum_{k\in V}\sum_{t=1}^{T} -x_{ijkt} \geq -B_{ij}, \quad \forall i \in H,\ j \in P \tag{3.13}$$

$$x_{ijkt} \geq 0, \quad \forall i \in H,\ j \in P,\ k \in V,\ t = 1, \cdots, T, \tag{3.14}$$

where (3.12) corresponds to (3.6), (3.13) corresponds to (3.8) (we have modified (3.13) so that the left-hand-side has only coefficients of -1 or 0), and (3.14) corresponds to (3.11) with a continuous relaxation of the x-variables. The coefficient matrix $\mathbf{A}$ of this LP has entries of +1, -1, or 0. Furthermore, we can see from the constraints (3.12) and (3.13) that every column of $\mathbf{A}$ has a single +1 entry, a single -1 entry, and all remaining entries are 0. First we show by induction that $\mathbf{A}$ is *totally unimodular*, that is, every square submatrix of $\mathbf{A}$ has a determinant of +1, -1, or 0. Since all entries are +1, -1, or 0, every 1×1 submatrix has a determinant of +1, -1, or 0. Assume that this is true for every submatrix $(k-1) \times (k-1)$, where $k \geq 2$, then we know that the determinant of any $k \times k$ submatrix is +1, -1, or 0. This is valid because each column in the submatrix has: 1) all 0 entries, which would yield a determinant of 0; 2) a single non-zero entry of +1 or -1, which would yield a determinant of $\pm$ the determinate of the $(k-1) \times (k-1)$ submatrix, and thus by induction we get a determinant of +1, -1, or 0; or 3) exactly two non-zero entries of +1 and -1, which yields a determinant of 0. Hence, $\mathbf{A}$ is *totally unimodular*. As the right-hand-side of the constraint set is integer valued, this implies, by Cramer's Rule, that all extreme point solutions to the LP are all integer valued, see, for example, Bazaraa et al. (2005). ∎

Essentially, when the y-variables are set to integer values, the x-variables, even if still relaxed in the solution algorithm (for instance in a standard branch-and-bound algorithm), will have integer values. Without the property described in Proposition 3.1, the HETM requires much more effort to solve. We illustrate this in the numerical analysis (see Section 3.4) using the case study provided in Section 3.3.

3.2.3 Minimizing and estimating risk

We believe the risk minimization objective is the logical one to use in hospital evacuations, as evacuations are performed solely to move patients (and staff) away from a threat risk, while considering that moving patients is also risky. Risk minimization offers strategic flexibility, allowing the model to produce evacuation transportation plans that are appropriate to a wide range of potential threats and problem instances. While under normal circumstances it is preferable that critical patients are transported by ALS ambulances (if they must be transported, for instance, to a facility having more advanced treatment capabilities), under emergency circumstances this might not be the case. For instance, a threat might force the evacuation of the hospital building, moving patients to lawn and parking areas, before transport is readily available (as happened at the Sepulveda VA Medical Center after the Northridge earthquake, see Chavez and Binder (1995)). In this case, if the options for a critical patient are to wait for hours for an ALS ambulance, or to be immediately transported by a bus, the bus might be a better option, as it allows the patient to receive proper treatment much sooner, and is potentially better than the lawn, depending on the environmental conditions. In a risk minimization framework such trade-offs are considered, which is a strength of the HETM. Minimizing the duration or logistical costs of an evacuation does not offer a "currency" that allows such trade-offs to be made. Given a much different threat, such as a hurricane, which is usually forecasted well in advance of its landfall, a risk minimization model will place patients in the vehicle type that provides the best care, as the threat risk is minimal for an extended time.

While hospitals must consider costs when making decisions, evacuations are so rare that the logistical costs of an evacuation is not a primary concern, although with the proper cost data, a solution to the HETM can be evaluated for cost, allowing the emergency manager to determine whether more resources (and thus higher costs) reduce risk (which is not always the case) or are otherwise appropriate; this model finds the risk minimizing solution, given a certain set of resources, and thus considers cost indirectly as an input. Cost minimization also has a few potential pitfalls. For instance, to produce a realistic solution a performance target must be provided (in Tayfur and Taaffe (2009), a duration target, enforced by a penalty cost, is used), which begs the question, how is that target specified? This might be easy for a forecasted threat such as a hurricane, which have duration targets specified by the forecast, but most other threats do not have targets that are so easily specified. Furthermore, from a cost perspective, legal costs can potentially dwarf logistical costs, and thus a risk minimizing approach might be the most cost-effective approach.

Using the risk minimization objective to derive the evacuation plan requires the emergency manager to estimate the risks. While this is certainly challenging, and beyond the scope of this study, we note that hospitals make these decisions on a daily basis when transporting patients between hospitals; here it must be decided when the risk of inadequate facilities for a particular patient (something akin to the threat risk) overrides the risk of transporting the patient to a more appropriate facility. Also, every year hospitals are required to produce a detailed risk analysis for potential threats; this allows the emergency manager to focus on the more likely threats, and to help develop risk estimates *a-priori*. These risks are studied in the literature (see Kue et al. (2011), and references therein) and thus we assume that these risks can be based on data, where available, and experience-based knowledge and expert judgment as required. We note that in our numerical studies we found that this modeling-framework is not overly sensitive to estimation error, as we discuss below.

In the case study, Section 3.3, three threat risk scenarios are used, each having a threat risk parameter, α_{jt}, that changes through time based on a different functional form. These include a constant scenario where α_{jt} is constant in time ($\alpha_{jt} = \alpha_j$), a linear scenario where $\alpha_{jt} = mt$ (m represents the slope of the linear function), and an exponential scenario where $\alpha_{jt} = a\exp(t/b)$ (a and b are parameters of the exponential function). The cumulative threat risk, Λ_{jt}, is concave for a constant risk parameter, S-shaped (the slope first increases and then decreases) for a linear risk parameter, and has an increasing slope for an exponential risk parameter (see Figure 3.3). There is no reason that every patient type (we group patient types into risk groups in the case study) should have the same functional form of the threat risk parameter, but we do assume that, in general, the more critical patients are more suspectable to the threat risk. As we mentioned above, the estimation of risks is beyond the scope of this study, but we believe that the exponential threat risk with the given parameter (see Table 3.3) is representative of a prolonged loss of utilities (electricity and water). This would severely reduce the ability of the hospital to care for patients, but more so for critical patients who are dependent on lifesaving equipment, than less critical patients. For all patients the impact of the outages is initially minor (given that there are back-ups, for instance batteries, for all critical care equipment; without such back-up systems the risk function could take a much different form), but as time passes the risks increase, potentially in a drastic manner. We provide the other forms as a comparison, as not all disasters are likely to take the exponential form. But overall, the form of the cumulative threat risk, Λ_{jt} , can take is limited. While any proper risk form is acceptable, such a function must only take values between zero and one and must be non-decreasing as a function of time. We specifically use these different risk functions (constant, linear, and exponential) in Section 3.4 to test the robustness of the solution provided, that is, if we forecast the risk using the "wrong" functional form, how will the solution perform? We see that for the case study data provided below, the solution is quite robust.

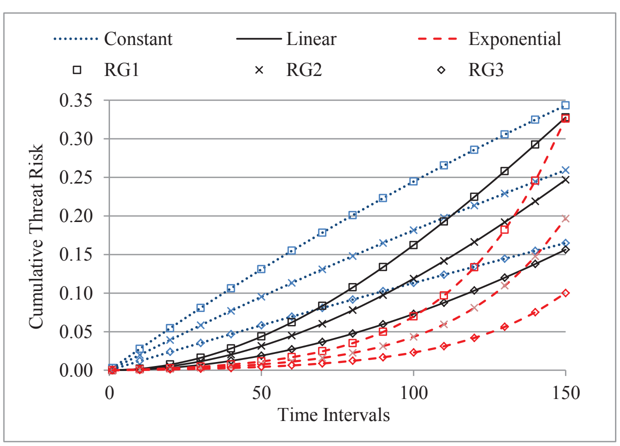
3.3 Case Study

In this section, we introduce a case study based on the evacuation of Carilion Medical Center (CMC). The CMC is a 765-bed primary care hospital combined with a 60-bed Level III Neonatal Intensive Care Unit (NICU) and is the second largest hospital in Virginia. It serves an immediate population of 250,000 and a regional referral population of over one million. The Emergency and Urgent Care Departments treat approximately 90,000 patients annually. The CRMH is the region's only Level I Trauma Center, and as such sees more than 1,200 of these high trauma patients annually. The CMC is a member of the Near Southwest Preparedness Alliance (NSPA), a mutual emergency management support and planning organization.

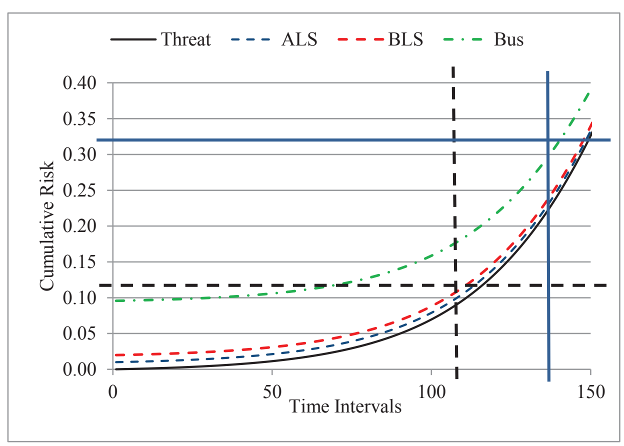
The CMC and the NSPA categorize the patient population into nine types for evacuation purposes, each type requiring a different level or type of care. These patient types are adult critical care patients (AdCC) requiring ICU, patients from the emergency department (ED), patients with infectious diseases in isolation beds (ISO), two types of medical surgery patients: the less critical MS-I and the more critical MS-II, patients in operating rooms (ORs), pediatric patients (Ped), pediatric critical care patients (PedCC), and psychiatric patients (Psy). The NSPA monitors the current number of patients, by type, in the member hospitals, and a snapshot data set was obtained. Table 3.1 displays how the 598 patients in the CMC, based on the snapshot data, are categorized by patient type.

Table 3.2 displays the 1,280 beds available at the 15 receiving hospitals considered in this

Figure 3.3: (a) The cumulative threat risk functions (Λ_{jt}) for the three functional forms and the three risk groups; (b) the cumulative threat and evacuation risks (R_{ijkt}) for the exponential case for three vehicle types for Risk Group 1 (RG1).



(a)



(b)

study, by patient type. For this study we use $T = 150$ ten-minute time intervals. Table 3.2 also displays the one-way travel time, in number of time intervals, between the evacuating hospital H^0 and the receiving hospitals. The number of patients at the evacuating hospital and the available beds at the receiving hospitals change on a daily basis.

Table 3.1: The number of patients in CMC (H^0) by type, $W_j, j \in P = \{1, \cdots, 9\}$.

Patient in CMC by type ($W_j, j \in P = \{1, \cdots, 9\}$)									Total
1	2	3	4	5	6	7	8	9	
AdCC	ED	PedCC	ISO	MS-II	ORs	MS-I	Ped	Psy	
72	51	2	1	286	34	96	19	37	598

Fleet scenarios

For this case study we examine two fleet scenarios. The first is the ambulance scenario, where there are 20 ALS ambulances and 20 BLS ambulances immediately available. After an hour an additional 15 ALS and 15 BLS ambulances are available, thus a fleet of 35 ALS ambulances, 35 BLS ambulances is available at the beginning of time interval seven (recall that an time interval is ten minutes in length). Each ambulance can transport a single patient. The second is the bus-ambulance scenario, which is identical to the first except that thirty minutes after the start of the evacuation five buses are available and each bus can transport 20 patients. We note that the local availability of ambulances is dynamic, and can vary depending on the time of day and the current work load (ambulances are used for emergency 911 support and for medical transport, and any number of them can be engaged at any time). The numbers used in this study are for illustration purposes. We believe part of the value of this tool is to allow emergency managers to perform "what if" analysis that can better enable them to make appropriate fleeting decisions and arrangements with transportation providers.

The physical loading capacity of the hospital, L, is measured in ambulance (ALS or BLS) equivalencies, and for this study is set to 10 ambulances every time interval. We assume that for each time interval a bus would require three times the resources of an ambulance to load,

Table 3.2: The number of available beds B_{ij} in each receiving hospital i, by patient type j, $i \in H = \{1, \cdots, 15\}, j \in P = \{1, \cdots, 9\}$, and the one-way travel time τ_i from the CMC to receiving hospital i.

Receiving Hospital	Available beds in H by patient type ($B_{ij}, i \in H, j \in P$)									Travel Time
	1	2	3	4	5	6	7	8	9	τ_i
	AdCC	ED	PedCC	ISO	MS-II	ORs	MS-I	Ped	Psy	(time intervals)
1	7	10	2	0	30	5	11	3	0	8
2	2	8	0	2	18	3	7	0	0	7
3	0	15	0	2	9	2	3	0	0	4
4	1	5	0	0	6	2	2	0	0	9
5	2	10	0	2	1	0	1	5	17	6
6	0	0	0	0	0	0	0	0	6	3
7	1	16	0	12	7	6	3	2	7	14
8	11	18	0	13	26	8	9	1	0	3
9	74	81	0	10	172	9	58	0	0	11
10	18	0	0	15	125	9	42	0	16	8
11	2	31	0	4	8	7	3	1	0	6
12	1	44	0	6	17	2	6	8	0	8
13	0	5	0	2	18	3	7	0	0	10
14	5	0	0	0	11	0	4	0	2	2
15	8	0	0	11	24	8	9	16	39	9
Total	132	243	2	79	472	64	165	36	87	

$L_{bus} = 3$; thus if three buses were being loaded in a particular interval, then there would only remain resources to load a single ambulance. Furthermore, an ambulance takes one interval to load ($\gamma_{ambulance} = 1$), while a bus requires two time intervals to load ($\gamma_{bus} = 2$). These parameters are rough estimates, and their values can be better estimated through the use of mock evacuations and other common training activities.

Risk scenarios

For this case study, we group the nine patient types into three risk groups based on criticality: patient types 1-3, the most critical patients, are in Risk Group 1 (RG1), patient types 4-6 are in Risk Group 2 (RG2), and patient types 7-9, the least critical patients, are in Risk Group 3 (RG3). The threat risk parameter, α_{jt}, is indexed on time and patient type, and depends on the nature of the threat. We study three threat risk scenarios, each using a different functional form for the threat risk parameter α_{jt}. These are: (i) a *constant* threat risk scenario, (ii) a *linearly increasing* threat risk scenario, and (iii) an *exponentially increasing* threat risk scenario. Table 3.3 gives the threat risk parameter values used in this study. As discussed in Section 3.2.3, we believe that the exponential threat risk with the given parameters is representative of a prolonged loss of utilities (electricity and water). Also provided in Table 3.3 are the transportation risk parameters, β_{jk}, for each risk group and transportation type combination.

Table 3.3: Risk group composition, threat risk functions (α_{jt}), and transportation risk parameters (β_{jk}).

Risk	Patient	Threat Risk			Transportation Risk		
Group	Types	Constant	Linear	Exponential	ALS	BLS	Bus
1	(1, 2, 3)	0.0028	0.000035t	$0.0000875\exp(t/30)$	0.001	0.002	0.010
2	(4, 5, 6)	0.0020	0.000025t	$0.0000625\exp(t/32)$	0.0001	0.0002	0.0005
3	(7, 8, 9)	0.0012	0.000015t	$0.0000375\exp(t/34)$	0.00005	0.00005	0.0001

Figure 3.3(a) shows the cumulative threat risk, Λ_{jt}, calculated using the data from Table 3.3, for the three functional forms and three risk groups, as a function of time. The cumulative threat risk function is concave for a constant risk parameter, S-shaped (the slope first increases and then decreases) for a linear risk parameter, and has an increasing slope for an exponential risk parameter. Figure 3.3(b) displays the cumulative exponential threat and evacuation risks for Risk Group 1, for the three vehicle types, for a hospital that has an eight interval travel time from H^0. The difference between the evacuation risk and the threat risk

is the effect of the transportation risk, which is non-linear. The influence of transportation risk, based on vehicle types, on the overall evacuation risk is considered by the model in assigning the best vehicle type to a patient. Figure 3.3(b) illustrates two concepts. *First*, if it is viable to transport Risk Group 1 evacuees using ALS and BLS by interval $T = 107$, see the dashed vertical line in Figure 3.3(b), then these patients will not be transported by bus, and if some are, they should be transported early so minimize the risk. The evacuation risk for patients that leave in the first interval is composed solely of cumulative transportation risk, this is the zero-intercept. *Second*, for the given time envelope ($T = 150$ time intervals) an evacuee will not be evacuated by bus after a certain time interval (see the solid vertical line) because this will entail more risk than remaining in the hospital. If the risk functions, which terminate after 150 intervals, represent the complete risk, this can help determine if partial evacuations is warranted.

In the next section, we examine the performance of the HETM and compare the risk-based optimization results with the duration-based results using the case study data.

3.4 Numerical Analysis

First we use the HETM to find the optimal evacuation transportation plan for the six scenarios (three risk scenarios, constant, linear, and exponential, combined with two fleet scenarios) outlined above for CMC. All computational runs were solved using IBM OPL IDE 6.3, ILOG CPLEX 12.1.0 software on a Dell Precision T5500 workstation with two 2.93 GHz CPU Xeon processors and 24 GB RAM. Table 3.4 shows the results.

For comparison purposes, if all patients remained in H^0 for all $T = 150$ time intervals, then they would be exposed to a total cumulative threat risk of 151.239, 143.964, and 119.092, respectively, for the constant, linear, and exponential risks. Thus we see that given these risk scenarios, evacuation is a prudent course.

Discussion of Proposition 3.1

We now examine the solution effort required to solve the HETM. The scenarios with three vehicle types, with 67,500 integer variables, have the largest number of such variables. Yet, despite the large size, the HETM takes at most a little over a minute to solve any of the

Table 3.4: Results for the HETM for the combination of the three risk and two fleet scenarios.

Fleet scenario	Risk scenario	Evacuation risk	Threat risk	Transportation risk	Evacuation duration (intervals)	Run time (s)
Ambulance	Constant	55.267	53.760	1.633	137	8.10
	Linear	28.268	26.761	1.556	135	8.64
	Exponential	10.410	8.860	1.562	133	10.26
Ambulance-Bus	Constant	26.249	23.688	2.664	61	28.38
	Linear	7.419	4.951	2.487	60	71.20
	Exponential	3.799	1.431	2.374	64	69.28

given scenarios. This performance is due to the special structure of the HETM described in Proposition 3.1. To illustrate the impact of the structure, consider the constant threat risk scenario with ambulances and buses. From Table 3.4 we see that it required 28.38 seconds to solve with a total evacuation risk of 26.268. A simple way to negate Proposition 3.1, without changing the optimal solution, is to set the bus capacity to 20.5 patients (the extra 0.5 patient capacity cannot be utilized by the integer x-variables, but with this change the right-hand-side is no longer integer, violating Proposition 3.1). In this case an optimal solution was not found in 24 hours, and the best integer solution found by CPLEX had an evacuation risk of 29.145, while the continuous solution was 26.003. Proposition 3.1 does not imply that the x-variables must be made continuous to gain this benefit. In fact, making the x-variables continuous has almost no performance impact; for the above scenario (with the bus capacity set back to 20) the solution time was 34.45 seconds.

Evacuation duration

Here we examine the evacuation duration from the HETM and compare it to some duration minimizing solutions. We observed that, using the given threat and transportation risk parameters, the generated plans have durations that range form 133 - 137 time intervals

(21.166 - 22.833 hours) using only ambulances, and 60 - 64 time intervals (10 - 10.666 hours) using ambulances and buses. Indeed, minimizing the overall evacuation duration is a commonly used evacuation objective; this is what the METPM heuristically minimizes. For comparison purposes, and to illustrate the advantages of risk minimization for the hospital evacuation problem, we provide the following duration minimizing objective and supporting constraints for the HETM:

$$\text{Minimize} \quad E \tag{3.15}$$

$$\text{subject to} \quad \frac{y_{ikt}}{L} \le b_{ikt}, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T \tag{3.16}$$

$$b_{ikt}(t + \tau_i + \gamma_k) \le E, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T \tag{3.17}$$

$$b_{ikt} \in \{0, 1\}, \quad \forall i \in H,\ k \in V,\ t = 1, \cdots, T \tag{3.18}$$

$$E \ge 0, \tag{3.19}$$

where (3.15) minimizes the non-negative variable E, which represents the overall evacuation time, and (3.16) forces the binary b_{ikt} variables to 1 if at least one vehicle of type k begins transporting patients to hospital i in time interval t. The evacuation completion time, E, is forced by (3.17) to be greater than or equal to the time that any vehicle reaches a receiving hospital. Constraints (3.18) and (3.19) are the binary and non-negativity restrictions for the variables. This definition of when an evacuation is complete, when the last patient arrives at a receiving hospital, is also used in the METPM. Without a risk minimization objective to drive the assignment of patients to vehicles, another strategy must be considered. For this purpose, we adopt the following three vehicle assignment strategies:

Restricted assignment: Each patient type is assigned to a vehicle type based on criticality, for instance critical patients are assigned to ALS ambulances, and non-critical patients are assigned to BLS ambulances.

Upgrade assignment: Much like restricted assignment, each patient type is assigned to a vehicle type, but patients can be upgraded to a vehicle that offers a higher level of care. Thus a patient type assigned to a BLS ambulance can also be served by an ALS ambulance.

Unrestricted assignment: Any patient type can be assigned to any vehicle type.

Replacing (3.4) with (3.15) and adding (3.16)–(3.19) to the HETM results in an *Unrestricted* assignment. By adding constraints that force the appropriate sets of x-variables to zero, we can obtain the *Restricted* and *Upgrade* assignment strategies. As mentioned earlier, the METPM assumes a restricted assignment of patient types to vehicle types.

Table 3.5 displays results for the bus-ambulance scenario for both the duration and risk minimization. Minimizing durations for the ambulance-only scenario presented some issues; for instance, the *Restricted* assignment produced a solution having a duration that was longer than 150 time intervals. This was a result of a mismatch between the number of patients in each risk group and the number of vehicles of each types; Risk Group 1 should be assigned to ALS ambulances, and Risk Group 3 to BLS ambulances, thus Risk Group 2, which has the largest number of patients, must share a vehicle type with another group, causing the other vehicle type to be under utilized. Even for the bus-ambulance scenario, we see from Table 3.5 that the duration for the *Restricted* assignment is much higher than the other assignment scenarios. Interestingly, we see that the transportation risk for the *Restricted* assignment is actually higher than for the *Upgrade*, this is because we have transportation risks that are always lower for vehicles having higher capabilities, thus upgrading can only

Table 3.5: Results for the Ambulance-Bus scenario for the duration minimizing solutions for the three specified assignment strategies.

		Duration minimization			Risk minimization		
		Restricted	Upgrade	Unrestricted	Constant	Linear	Exponential
Duration		144	96	55	61	60	64
Evac. risk	Constant	67.760	50.233	33.378	26.250		
	Linear	39.530	20.768	12.262		7.418	
	Exponential	17.544	7.170	7.641			3.799
Transportation risk		2.216	1.700	5.947	2.668	2.487	2.373
Run time (s)		633.880	3,123.050	19,114.120	28.380	71.200	69.280

decrease the transportation risk. Given this, the *Restricted* assignment will always be inferior to the *Upgrade* assignment. The *Unrestricted* assignment produces the shortest possible evacuation duration. While this tends to decrease threat risk, it also tends to increase the transportation risk (which, for the exponential case, allows the *Upgrade* assignment to produce a lower evacuation risk than the *Unrestricted* case). The *Unrestricted* case is the most computationally expensive assignment strategy, most likely due to its larger feasible region. As demonstrated in Table 3.5 the results from the risk minimization solutions have durations that are shorter than all assignment strategies save the *Unrestricted*. Conversely, we see that minimizing duration is not effective at minimizing risk; using the lowest risk results from each vehicle assignment strategy for each risk type we still have total evacuation risks that are 27.1%-88.7% greater than those from the risk-based model.

Table 3.6: The latest time of evacuation by patient type, $j \in P$.

Fleet scenario	Risk scenario	Patient Type 1	2	3	4	5	6	7	8	9
Ambulance	Constant	82	14	23	9	99	77	126	124	123
	Linear	82	14	28	3	99	84	124	124	121
	Exponential	46	17	17	1	98	67	123	122	118
Ambulance-Bus	Constant	36	11	18	3	45	33	51	50	45
	Linear	38	24	20	2	47	25	50	49	49
	Exponential	43	18	28	4	52	44	43	43	35

To illustrate another property of the evacuation strategy, Table 3.6 displays the interval in which the last patient of each type departs from the evacuating hospital. We observed that the last patient in Risk Group 2 departs before the last patient in Risk Group 3, despite there being many more Risk Group 2 patients. Table 3.7 shows the total number of patients from each risk group transported by each vehicle type. Here we see that many patients

Table 3.7: The number of the patients in each risk group transported by each type of vehicle in six scenarios.

Fleet scenario	Risk scenario	Risk Group 1			Risk Group 2			Risk Group 3		
		ALS	BLS	Bus	ALS	BLS	Bus	ALS	BLS	Bus
Ambulance	Constant	109	16	-	116	205	-	53	99	-
	Linear	125	0	-	110	211	-	46	106	-
	Exponential	125	0	-	105	216	-	57	95	-
Ambulance-Bus	Constant	102	23	-	18	85	218	7	43	102
	Linear	125	0	-	11	121	189	1	20	131
	Exponential	125	0	-	22	136	163	0	15	137

from Risk Group 2 are transported using buses, so the model determines that the added transportation risk is worthwhile to reduce the threat risk. Table 3.7 also shows that some patients in Risk Group 1 are transported using BLS ambulances for the same reason, but not by bus. Figure 3.3(b) shows that the evacuation risk using a bus is significantly higher than the risk using an ambulances, but a bus used in the first intervals is approximately as safe as an ALS ambulance used in interval 107 (for a receiving hospital eight time intervals from H^0). If the transportation risk were increased relative to the threat risk, a patient from Risk Group 1 could be transported by bus (we note that if a particular patient type should never go by bus, this can be ensured using a hard constraint). In the constant threat risk scenario, the model is more likely to place more critical patients into BLS ambulances or buses than the exponential form of the threat risk. As Figure 3.3(a) illustrates, the cumulative threat risk for a Risk Group 1 patient that remains in the hospital for 150 time units is about the same for all risk forms, the constant risk case more quickly accrues risk in the earlier time intervals, and thus it is of greater benefit to transport the critical patients early. Conversely, for the exponential risk the earlier periods have a relatively lower risk, and thus the transportation risk has a larger impact and patients are more often assigned to the safest transport.

Robustness

Robust evacuation solutions are very important. Of all the model inputs, the risk parameters are the most difficult to estimate. Thus, as we have presented three functional forms for the threat risk, it is appropriate to ask the question, what if the wrong functional form is used? To explore this, we solve the HETM using the three functional forms as "forecasted" risks (thus developing three evacuation plans) and evaluate each of these plans using the same three functional forms (given in Table 3.3) as if these were the "actual" risks. That is, if we use a plan that minimizes the exponential risk, but the actual risk is represented by the constant risk, how would the solution fare? The results from the nine possible combinations are displayed in Table 3.8, where the risks in bold-face type are those having the same forecasted and actual risks (thus they are the same as the results in Table 3.4. From these results, it seems that the evacuation plans produced are robust in regard to this "error". The forecasted risk is at most 5.2% greater than the actual risk (this occurs when the fore-

Table 3.8: Evacuation risks from plans that minimize the input risk, but are evaluated using the actual risk scenario.

Fleet scenario	Actual risk scenario	Forecasted risk scenario		
		Constant	Linear	Exponential
Ambulance	Constant	**55.268**	55.321	55.336
	Linear	28.335	**28.268**	28.304
	Exponential	10.530	10.444	**10.409**
Ambulance-Bus	Constant	**26.250**	26.525	26.929
	Linear	7.514	**7.418**	7.562
	Exponential	3.999	3.845	**3.799**

cast is for the constant risk function, and the actual risk is exponential). As a comparison, we see that these risks estimation errors still produce results that outperform the duration objectives in terms of risk (duration is not impacted by these errors).

Table 3.9: Evacuation risks from plans that minimize the input risk, but are evaluated using the actual risk scenario with varying transportation risk.

Fleet scenario	Actual risk scenario	Forecasted risk scenario ($1/2 \times \beta_{jk}$)			Forecasted risk scenario ($2 \times \beta_{jk}$)		
		Constant	Linear	Exponential	Constant	Linear	Exponential
Ambulance	Constant	55.275	55.327	55.336	55.289	55.332	55.335
	Linear	28.382	28.304	28.304	28.287	28.269	28.298
	Exponential	10.539	10.409	10.409	10.469	10.444	10.413
Ambulance-Bus	Constant	26.378	26.362	26.583	26.269	26.592	28.416
	Linear	7.843	7.459	7.435	7.49	7.436	8.345
	Exponential	4.333	3.939	3.829	3.966	3.829	3.872

As the evacuation plan produced by the HETM reduces risk by optimally balancing threat and transportation risk, we now examine how increasing and decreasing the forecasted transportation risk impacts the results. Table 3.9 shows the results from two cases, one where the forecasted transportation risk parameter β_{jk} is reduced by half, and another where it is doubled compared to the actual risks that are given in the case study. Thus these results include errors in the functional form of the threat risk and the magnitude of the transportation risk.

From Table 3.9 we find the largest increase in risk due to this error of 14.1% for the ambulance-bus case where the actual risk is exponential and the forecasted risk is constant and the transportation risk is under forecasted. In this case, the solution tends to suboptimally transport patients in vehicles with less care capabilities than is optimal, this is because the forecast is for a high threat risk early on, and a lower overall transportation risk, whereas the actual threat risk is lower early on, and the transportation risk is higher. For instance, in this case three patients from Risk Group 1 are transported by bus. This still outperforms the best of the duration solutions by a large margin. Overall, there is an average 1.7% increase in risk over all these scenarios due to the forecast error. Thus while forecasting risk will take some effort, these forecasts do not have to be extremely accurate to produce a very good plan.

3.5 Conclusions

Hospitals serve a particularly vulnerable segment of society, ranging from newborns to those requiring critical care due to illness or injury. Thus, while evacuating a hospital is never an easy choice, it is sometimes necessary, and special attention must be paid to planning and operational issues. Unlike a more traditional building evacuation, the evacuation process does not end when the building is cleared, in fact an important component is the transportation of patients to alternative care facilities. This is a difficult logistical problem, and the goal of this work is to produce a decision support tool that can assist in the planning and operational phases of an evacuation. The Hospital Evacuation Transportation Model (HETM) that we introduce here is an integer program that has a special structure that allows it to be solved in a reasonable amount of time for both planning and operational usage. The HETM develops hospital evacuation transportation plans that match patients, based on the patient's characteristics and care requirements, to vehicle types, based on the vehicles characteristics and current availability, and then determines the best receiving hospital for the patient. This is done in such a way as to minimize the total evacuation risk, both the threat risk (which is causing the evacuation) and the transportation risk.

While there are numerous modeling studies that focus on building evacuation in general, hospital evacuations, with their special characteristics, are not well studied. Significantly, hospital evacuations are an important candidate for study, as much more planning is required to evacuate a building when the evacuees are as dependent as patients are on the hospital for their care and many have, due to their condition, a much more limited ability for "self-preservation". This is also a very complex problem, and as with all complex problems the models that are used must include many assumptions, assumptions that were made so that the model could be solved. These assumptions must be explored in more detail, and more work must be done to understand their impact, and possible ways to relax these assumption, if that is deemed necessary. We firmly believe that the use of operations research models and methodologies can have a significant positive impact on the hospital evacuation planning process and the operational performance of the evacuation team.

Chapter 4

Equity Modeling and Resource Management for Hospital Evacuations

This chapter addresses the simultaneous evacuation of multiple hospitals, which can be in different hospital management groups, along with the equity issues that arise with the allocation of resources to the hospitals and patients in the system. An integer programming model is formulated to minimize the evacuation risk for the system given centralized resources that are shared among hospital groups. Decentralized resource management is introduced as an alternative strategy to determine the individual performance of each hospital group without resource sharing. Inequitable resource allocation decisions are observed in the risk-minimizing solutions. To address the distributional equity issues, an equity modeling framework is developed based on the social welfare function studied extensively in the economics literature. The social welfare function depends on utility functions for players, among whom the resources are allocated, and the inequity aversion parameter that represents the importance of equity for the decision maker. Utility functions for hospitals and patients are developed and their structural properties are discussed. The objective function for the proposed integer program is defined as the social welfare function to generate equitable solutions. Optimal solutions under various equity criteria are evaluated at the hospital- and patient-levels. The solution behavior of different utility functions are demonstrated through representative examples. The computational results of the proposed model are presented and managerial insights about resource management are provided.

4.1 Introduction

Hospitals are vulnerable to threats that can interrupt their operations. Emergencies that force hospitals to evacuate are rare, but the potential damage to the patients, staff, and visitors in hospitals is tremendous in many cases. Such disruptive events subject the hospital occupants to high risk and emergency evacuation is necessary to alleviate this risk. The Department of Health and Human Services' (HHS) Centers for Medicare and Medicaid Services (CMS) require emergency management and evacuation plans from hospitals and the hospital accrediting organizations such as the Joint Commission (JC) assess the compliance of annual plans of participating healthcare facilities with the HHS requirements (see, for example, JC Standard EM.02.02.03 EPs 4, 5, 9, 10, and EM.02.02.11 EP 3).

Disasters such as hurricane, flood, or wildfire may threaten hospitals and force them to evacuate and transfer patients to safer locations (for evacuation examples, see Maxwell (1982), Chavez and Binder (1995), Cocanour et al. (2002), Schultz et al. (2003), Augustine and Schoettmer (2005), Gray and Hebert (2007), and Bagaria et al. (2009)). It is likely to experience conditions under which a set of hospitals in the same region needs to be evacuated when a disaster, such as a hurricane or a wildfire, affects a wide area where multiple hospitals are located and calls for the transportation of patients from several hospitals to safer receiving facilities. When multiple hospitals in a region evacuate, the success of the evacuation transportation plan depends greatly on the efficient use of the available resources, including a fleet of vehicles, the capacity of evacuating hospitals for loading patients into vehicles, and beds at receiving facilities. However, the fairness of resource allocation to hospitals or patients is also a major concern for the decision maker responsible for evacuation planning in addition to, and mostly conflicting with, efficiency. The most efficient evacuation plans may not necessarily be equitable in resource allocation, and thus, risk distribution, among patients. The decision maker is faced with this trade-off between efficiency and equity of the resource allocation which does not have a straightforward explanation or solution due to both the subjectivity of fairness and the complexity of the hospital evacuation problem. First, equity must be defined based on the perceptions of players receiving the service such as evacuating hospitals or patients. Second, appropriate functional forms must be constructed to represent these perceptions of equity. We believe that making a fair assessment of each player and distributing resources equitably among players is possible by using optimization-based methods to generate transportation plans.

The problem considered in this paper is the transportation of patients to safer facilities when multiple hospitals are evacuated. Unlike the evacuation of a single hospital studied in Bish et al. (2011) where all available resources are used for the same hospital, the problem studied here involves a competition for the limited transportation resources and available beds at the receiving facilities. As the evacuee population consists of patients with various levels of care requirements, patients are categorized based on their health condition (e.g., pediatric, medical surgery, critical care) and the same categorization is used to define receiving beds that are equipped to serve these patients. Based on the patient requirements, vehicles with different capacity and level of care (e.g., advanced life support ambulances, air ambulances, wheelchair vans, buses) can be used to transport patients, if available. Transportation vehicles need to be scheduled efficiently as they are both limited in number, and thus, capacity, and costly to operate. The characteristics of the non-homogenous patient population impose restrictions on resource allocation as each patient of a certain category should be assigned to adequate receiving beds and vehicles that are capable of serving patients of this category. The decision maker not only needs to generate the best plan that can transfer patients with the minimum possible risk, but also must meet the equity criteria through the optimal assignment of vehicles to hospitals that satisfies the adequate allocation of beds to patients.

Securing sufficient transportation and medical care resources is essential for evacuation particularly when multiple hospitals compete for them. Hospitals can make arrangements for the fleet of vehicles to be used and for the beds available at the potential receiving facilities in advance or they can share the pool of resources during evacuation. Sternberg and Lee (2009) provide a list of institutional assignment models in healthcare transportation, a combination of which New York City (NYC) uses in emergencies. This list includes (i) assignments made by 911 dispatchers (which is the primary model used in NYC), (ii) assignments by a designated hospital that monitors and coordinates a group of hospitals, (iii) assignments to the nearest emergency department, (iv) specialized assignments (for patient types such as burn, ventilation, and isolation), (v) assignments made centrally by an Emergency Operations Center (EOC), and (vi) decentralized assignments made by mutual adjustments among medical responders and transportation officials. Pre-arrangement can be necessary especially for the specialized care patients who cannot afford waiting for an adequate transportation vehicle or for an appropriate bed to become available in time of emergency. However, in the absence of such pre-arrangements, hospitals should be able to share the evacuation resources efficiently in real-time. A designated coordinating hospital may not always be in operating condition or sending patients to the nearest emergency department may not prove to be efficient as the patient may have to be sent to an appropriate facility from there. The centralized assignment would seemingly be the ideal model to use, but it would not necessarily be equitable if certain players benefit from resource centralization significantly more than others. On the other hand, decentralized assignments can be efficient for a certain group of patients while another group can be jeopardized due to the lack of information about the overall resource limitations. It is possible that different resource management strategies would perform better in certain types of threats. Thus, emergency management would benefit greatly from the comparison of resource management strategies in alleviating the competition for limited resources and ensuring equitable resource allocations during a multi-hospital evacuation.

Evacuation studies in the literature utilize various objective functions most of which are time-based such as the building clearance time, evacuation time in which all evacuees leave the vulnerable area, or the latest time to reach the safer destination(s). Evacuation models in the literature mostly focus on regional or building evacuations and the models based on dynamic network flows can also maximize functions of flow reaching a single destination or multiple destinations. For a detailed review of the building evacuation models in the literature see Hamacher and Tjandra (2001). Equivalencies or comparison of the most common time- and flow-based evacuation objectives can be found in Jarvis and Ratliff (1982), Løvås (1995), and Correa et al. (2007). Han et al. (2007) examine the existing measures of effectiveness for evacuation in the literature and propose a framework that considers evacuation time, cumulative exposure, and time/space-based risk factors. The time-space-based risk and evacuation exposure objective function suggested by Han et al. (2007) is similar to the total evacuation risk function assumed in Bish et al. (2011), which additionally considers the risk incurred during transportation as well. These objectives essentially maximize the efficiency of evacuation and the risk-minimization approach is adopted in this study because it reflects the adverse effects of the threat causing the hospital evacuation and the transportation on patients adequately.

The fundamental objective of evacuation is to minimize the overall risk, which includes the threat risk and the transportation risk. The threat's location, impact, and magnitude over time greatly affect the evacuation plan. For example, disasters like earthquakes cannot be predicted with accuracy and this immediate threat would incur a high risk on all patients, whereas a hurricane can be predicted, there would be no threat risk initially, and the risk would increase in time. The other risk component, the transportation risk depends on the vehicle attributes, the category of the patient transported, and the travel time. Minimizing the overall evacuation risk provides transportation plans that schedule the transfer of patients as early as possible by the most adequate available vehicles such that the threat risk and transportation risk are balanced. However, an optimal evacuation plan that minimizes the overall risk for our problem with multiple hospitals does not necessarily minimize the risk for each evacuee or each hospital group. Hence, we need to consider certain measures of equity among patients or hospitals.

In terms of the consideration of risk, the hospital evacuation problem has similarities with transportation of hazardous materials. The selection of routes for transporting hazardous materials depends on the risk of using each link to the community and environment. Planning the shipments of hazardous materials in a region requires selecting multiple routes that have fair measures of risk so that no local community is jeopardized. Gopalan et al. (1990) developed an integer programming model to minimize the total risk of hazardous material shipments where the risk would be distributed equitably among the zones in the network. A risk and cost based framework for the routing of hazardous materials is proposed by Abkowitz and Cheng (1988) with a special emphasis on the estimation of risk on each link that includes the costs of fatalities, injuries, and property damages depending on the type of hazardous material shipped. Patients in the hospital evacuation problem are analogous to hazardous materials in that the costs are risk parameters attributed to the routes that patients follow in the evacuation network and they depend on the patient type. On the other hand, the decision maker in hospital evacuation seeks a fair distribution of risk among evacuating hospitals or patients while the a fair distribution of risk among the zones that are exposed to hazardous materials is desired.

Equity has been discussed, studied, and modeled as a concept in economics, decision making, and operations research as well as a philosophical, social, and political concept. We are especially interested in the welfare economics theory on equity that studies the equitable resource allocation to individual players considering each player's utilities. The social welfare function of utilities introduced by Atkinson (1970) that is based on risk aversion discussed by Pratt (1964) is of practical use since it depends on a single parameter that represents the importance of equity for the decision maker. We will use different utility definitions with the social welfare function and interpret the equity of solutions for various values of this inequity aversion parameter. Detailed reviews of economic theory on equity can be found in Young (1995) and Sen and Foster (1971).

Equity modeling is applicable to various public service and resource allocation problems in management science and operations research. Examples include job scheduling, queueing, call center design, communication and transportation network management, supply chain management, and healthcare planning. Savas (1978) discusses efficiency, effectiveness, and equity as the three performance measures for public services that depend on the spatial distribution of supply and demand and describes four general principles of equity (equal payment, equal outputs, equal inputs, and equal satisfaction of demand) that are inconsistent with each other yet fair for a specific perception of equity. Equity has been studied extensively in the context of facility location and the equity measures for facility location decisions are reviewed by Marsh and Schilling (1994) that also compare the behavior of optimal location solutions based on the functional form of these measures. The inevitable equity-efficiency trade-off is studied in several contexts including healthcare service allocation to patients such as cadaveric organs and kidney dialysis (Hooker and Williams, 2012; Bertsimas et al., 2013), humanitarian relief routing (Balcik et al., 2010), and air traffic control (Bertsimas et al., 2012). To investigate this trade-off in public service problems, Mandell (1991) employs the Gini coefficient, which is basically the average absolute difference between the service received by all pairs of players divided by twice the mean service, and develops a bicriteria mathematical programming model that first maximizes the level of output in each service area and then iteratively reduces the upper bound on the Gini coefficient while maximizing the output level. Bertsimas et al. (2012) consider the various objectives defined by the social welfare function along with the loss of efficiency (equity) as equity (efficiency) is regarded more important and provides near-tight and tight upper bounds on the price of fairness and the price of efficiency, respectively. Similar to the application areas involving resource allocation in the aforementioned studies, we provide an application of equity modeling in the context of hospital evacuation as an indispensable public service.

The remainder of this chapter is organized as follows. In Section 4.2, the model formulation for the evacuation of multiple hospitals is introduced and resource management strategies employed in this study are described. In Section 4.3, equity issues concerning the allocation of evacuation resources are discussed and the equity modeling framework based on welfare economics is explained. Additionally, the hospital- and patient-level approaches to equity are described and demonstrated through examples in Section 4.3. Section 4.4 presents a large-scale case study problem that is used for analyzing the performance of the proposed equity model with respect to risk distribution and resource management. Conclusions and directions for future research are discussed in Section 4.5.

4.2 Evacuation Transportation Model for Multiple Hospitals

In this section we present the formulation of the Hospital Evacuation Transportation Model for n evacuating hospitals (n-HETM) along with the modeling assumptions and the discussion of the model structure.

4.2.1 Formulation of the n-HETM

In this study, we consider the evacuation of multiple hospitals where there is a central decision maker for allocating a given set of resources among the hospitals. Resources including vehicles and receiving beds are centralized so that hospitals can share them throughout evacuation. Hospitals can belong to different management groups that can also perform evacuation in isolation from other groups. Receiving beds of a hospital group might be insufficient to accommodate all the patients that need to be transferred within that group. Furthermore, the hospital groups are not necessarily geographically clustered and sharing resources to transfer patients to a much closer hospital in another group can be preferred. The case where resources are not shared by hospital groups will be discussed further as the decentralized resource management strategy following the model formulation.

Let I be the set of evacuating hospitals, such that $|I| = n$, and P the set of patient types (e.g., critical care, pediatric, infectious, etc.), where W_p^i is the number of patients of type $p \in P$ in hospital $i \in I$. The patients in evacuating hospitals must be transported to a set of potential receiving hospitals (J), where hospital $j \in J$ has B_p^j beds available for patients of type $p \in P$. Each patient of type p should be assigned to a bed of type p, that is fully equipped for treating this type of patient. Let H be the set of all hospitals in the network, i.e., $H = I \cup J$. The study period is divided into T time intervals of equal length. The travel time from evacuating hospital i to receiving hospital j is τ^{ij} time intervals. Travel times between evacuating and receiving hospitals are known, have integral length, and are independent of vehicle type. The latter assumption is easy to modify such that travel times are vehicle dependent.

A set of vehicle types (V), having different capabilities, becomes available through time. Vehicle types can include Advanced Life Support (ALS) ambulances, Basic Life Support (BLS) ambulances, wheelchair vans, and/or buses. The vehicles are allowed to become available in time, so the fleet size is assumed to be dynamic through time. At time interval t, N_{vt} vehicles of type $v \in V$ become newly available for use at the evacuating hospitals. This parameter reflects the time required by vehicles that might have different origins (depots) to reach the evacuating hospitals. Hence, the changes in the vehicle supply can be easily represented by adjusting the values of the N_{vt}-parameters. The capacity of a vehicle of type $v \in V$ is C_v patients regardless of patient type, i.e., all patients require the same amount of space in each vehicle. The time required to load patients into a vehicle of type v is γ_v time intervals. The loading time is assumed to be independent of the patient type and equal to the time to unload patients to simplify the model. We note that the latter assumption is easy to modify such that the loading and unloading times are different. However, assuming that loading and/or unloading times depend on patient type complicates the problem. Vehicles transport patients directly from hospital i to hospital j without stopping at multiple hospitals. Assuming that evacuating hospitals share the available vehicles during evacuation, a vehicle can either visit an evacuating hospital in another group to take patients or return to the same hospital after transferring patients to an alternative care location.

In this study, partial evacuation is not considered as an option and it is assumed that patients of the appropriate type are available to satisfy the transportation plan that is limited by the loading capacity of the evacuating hospital. The number of vehicles that can be loaded at hospital $i \in I$ at any time interval is represented by the L^i-parameter. This parameter represents the loading capacity of the evacuating hospital subject to the aggregate limitation of available resources, such as personnel, staging area, and loading area. A vehicle can occupy this loading capacity for more than one time interval according to the loading time, γ_v. L^i-parameter is given in ambulance equivalencies; L_v represents a conversion factor for a vehicle of type $v \in V$ to ambulance equivalencies (i.e., a bus takes more space and effort to load than an ambulance). Due to this limitation, some vehicles may wait at an evacuating hospital to be loaded at a later time interval. Such vehicles will be represented as variables with the same evacuating hospital as their origin and destination in the formulation.

Based on the problem setting described above, the goal is producing an evacuation transportation plan for multiple hospitals that minimizes the total evacuation risk where risk is defined as the probability that an undesired event occurs. This event can be in the form of a major deterioration of patient health or death due to injury or lack of sufficient treatment. The evacuation risk is defined as a combination of: 1) the threat risk that patients are exposed to while waiting to be transported, and 2) the transportation risk incurred during travel. The cumulative threat risk, Λ^i_{pt}, calculated in (4.1), is the probability of the undesired event for a patient of type p that remains in hospital i through time interval t, where λ^i_{pt} is the probability of the undesired event for a patient of type p that remains in hospital i in time interval t. λ^i_{pt} at different time intervals are assumed to be independent and a patient does not change type during an evacuation, e.g., a non-critical patient does not become critical by being exposed to the threat risk.

$$\Lambda_{pt}^{i} = 1 - \prod_{f=1}^{t}(1 - \lambda_{pf}^{i}), \quad \forall i \in I,\ p \in P,\ t = 1, \cdots, T. \tag{4.1}$$

The cumulative transportation risk, Θ_{pv}^{ij}, is calculated in (4.2), where θ_{pv} is the probability of the undesired event for a patient of type $p \in P$ transferred by a vehicle of type $v \in V$ for one time interval. θ_{pv} is assumed to be constant through time since the same allocation made in different time intervals would not change the level of treatment supplied by a vehicle. When transporting patients from hospital i to j, a vehicle of type v that arrives at hospital k is engaged for $(\tau^{ij} + 2\gamma_v + \tau^{jk})$ time units, however, only $(\tau^{ij} + 2\gamma_v)$ time units contribute to the risk expression since transportation risk is incurred only as long as the patient is in the vehicle, including loading and unloading times.

$$\Theta_{pv}^{ij} = 1 - (1 - \theta_{pv})^{(\tau^{ij} + 2\gamma_v)}, \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V. \tag{4.2}$$

The evacuation risk, R_{pvt}^{ij}, associated with the evacuation decision for a patient calculated in (4.3) by combining the cumulative threat risk (Λ_{pt}^{i}) to which the patient is exposed before being transported in time interval t and the cumulative transportation risk (Θ_{pv}^{ij}) based on the vehicle type k and the receiving hospital selected. As (4.1) - (4.3) indicate, the cumulative threat, transportation, and evacuation risk functions are non-linear.

$$R_{pvt}^{ij} = 1 - (1 - \Lambda_{p(t-1)}^{i})(1 - \Theta_{pv}^{ij}), \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T. \tag{4.3}$$

The decision variables and the model formulation are provided below.

Decision Variables:

x_{pvt}^{ij} : number of patients of type p transported from evacuating hospital i to receiving hospital j by a vehicle of type v starting in time interval t, $\forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T$.

ν_{vt}^{i} : number of vehicles of type v initially assigned to evacuating hospital i at time t, $\forall i \in I,\ v \in V,\ t = 1, \cdots, T$.

y_{vt}^{ij} : number of vehicles of type v that move from hospital i to hospital j starting in time interval t, $\forall i,\ j \in H,\ v \in V,\ t = 1, \cdots, T$.

***n*-HETM:**

$$\text{Minimize} \quad \sum_{i\in I}\sum_{j\in J}\sum_{p\in P}\sum_{v\in V}\sum_{t=1}^{T} R^{ij}_{pvt}x^{ij}_{pvt} \tag{4.4}$$

subject to

$$\sum_{j\in J}\sum_{v\in V}\sum_{t=1}^{T} x^{ij}_{pvt} = W^i_p, \quad \forall i\in I,\ p\in P \tag{4.5}$$

$$\sum_{i\in I}\sum_{v\in V}\sum_{t=1}^{T} x^{ij}_{pvt} \le B^j_p, \quad \forall j\in J,\ p\in P \tag{4.6}$$

$$\sum_{p\in P} x^{ij}_{pvt} \le C_v y^{ij}_{vt}, \quad \forall i\in I,\ j\in J,\ v\in V,\ t=1,\cdots,T \tag{4.7}$$

$$\sum_{i\in I} \nu^i_{vt} \le N_{vt}, \quad \forall v\in V,\ t=1,\cdots,T \tag{4.8}$$

$$\nu^i_{vt} + \sum_{\substack{j\in J:\\ t>\tau^{ij}}} y^{ji}_{v(t-\tau^{ij})} + y^{ii}_{v(t-1)} = y^{ii}_{vt} + \sum_{j\in J} y^{ij}_{vt}, \quad \forall i\in I,\ v\in V,\ t=1,\cdots,T \tag{4.9}$$

$$\sum_{\substack{i\in I:\\ t>\tau^{ij}+2\gamma_v}} y^{ij}_{v(t-\tau^{ij}-2\gamma_v)} + y^{jj}_{v(t-1)} = y^{jj}_{vt} + \sum_{i\in I} y^{ji}_{vt}, \quad \forall j\in J,\ v\in V,\ t=1,\cdots,T \tag{4.10}$$

$$\sum_{j\in J}\sum_{v\in V}\sum_{f=t-\gamma_v+1}^{t} L_v y^{ij}_{vf} \le L^i, \quad \forall i\in I,\ t=1,\cdots,T \tag{4.11}$$

$$x^{ij}_{pvt} \ge 0 \text{ and integer}, \quad \forall i\in I,\ j\in J,\ p\in P,\ v\in V,\ t=1,\cdots,T \tag{4.12}$$

$$\nu^i_{vt} \in \{0,1,\cdots,N_{vt}\}, \quad \forall i\in I,\ v\in V,\ t=1,\cdots,T \tag{4.13}$$

$$y^{ij}_{vt} \in \{0,1,\cdots,\lfloor L^i/L_v\rfloor\}, \quad \forall i\in H,\ j\in H,\ v\in V,\ t=1,\cdots,T. \tag{4.14}$$

The objective function (4.4) minimizes the total expected evacuation risk. Constraint (4.5) limits the total number of patients of each type transported to the initial population size while (4.6) defines the number of beds of each type available at each alternative care location. (4.7) represents the vehicle capacity restriction on the number of patients transferred and (4.8) restricts the total number of vehicles assigned to evacuating hospitals at each time interval. Constraints (4.9) and (4.10) represent the conservation of flow over the planning horizon at each evacuating hospital and at each receiving hospital, respectively. (4.11) restricts the number of vehicles that can be loaded at each time interval due to the loading capacity. (4.12) - (4.14) are the logical integrality and non-negativity constraints. Considering several evacuating hospitals expands the problem discussed in Bish et al. (2011) and increases the number of variables (arcs) in the evacuation transportation model (network) as well as requiring additional constraints on shared resources. The n-HETM is an integer programming (IP) model and the continuous relaxation of the x- and ν-variables provides integer-valued optimal solutions due to the structure of the constraint matrix as explained in Proposition 4.1.

Proposition 4.1 . *For a feasible set of y-variables and a continuous relaxation of the x- and ν-variables, there exists an optimal solution for the n-HETM in which the continuous variables have integral values.*

Proof. A feasible set of y-variables must satisfy constraints (4.9) - (4.11) and be integer valued, i.e., satisfy (4.14). Such a set of y-variables can be considered as parameters. The corresponding feasible values of all ν-variables are directly computed by the equality constraint (4.9) and, therefore, must be integer and must conform to constraint (4.8). Then, a continuous relaxation of the x-variables transforms the n-HETM constraint set as follows: (4.5) - (4.7) and (4.12) into the following:

$$\sum_{j\in J}\sum_{v\in V}\sum_{t=1}^{T} x_{pvt}^{ij} = W_p^i, \quad \forall i \in I,\ p \in P \tag{4.15}$$

$$\sum_{i\in I}\sum_{v\in V}\sum_{t=1}^{T} x_{pvt}^{ij} \leq B_p^j, \quad \forall j \in J,\ p \in P \tag{4.16}$$

$$\sum_{p\in P} x_{pvt}^{ij} \leq C_v y_{vt}^{ij}, \quad \forall i \in I,\ j \in J,\ v \in V,\ t = 1, \cdots, T \tag{4.17}$$

$$x_{pvt}^{ij} \geq 0, \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T \tag{4.18}$$

where (4.15) - (4.17) correspond to (4.5) - (4.7) and (4.18) corresponds to (4.12) with a continuous relaxation of the x-variables. If the constraint coefficient matrix of an LP is totally unimodular and the right-hand-side values are all integer valued, the LP has integer basic solutions by Cramer's Rule (Bazaraa et al., 2005). The right-hand-side values of the constraints above are integer-valued and we show that the constraint coefficient matrix satisfies the conditions for total unimodularity as follows.

Camion (1965) proved that a matrix is totally unimodular if and only if every square Eulerian submatrix formed from it is singular, where a submatrix is Eulerian if both the sum of each row and the sum of each column are even. It can be shown that the coefficient matrix of the above LP is totally unimodular by examining the characteristics of its submatrices.

The entries of the coefficient matrix $\mathbf{A}$ of the LP above are $+1$ or 0. Each x-variable appears once in each of the three constraint sets (4.15) - (4.17), therefore, every column of $\mathbf{A}$ has three $+1$ entries and all remaining entries are 0. Since the sum of each row and column of a Eulerian matrix must be even, all 1×1 Eulerian submatrices of $\mathbf{A}$ consist of a 0 entry and have a determinant of 0. The 2×2 Eulerian submatrices can have either four $+1$ entries or four 0 entries, and have a determinant of 0. Similarly, the $k \times k$ Eulerian submatrices, such that $k \geq 3$, can have either zero or two entries of $+1$ in each column, because each column of $\mathbf{A}$ has three $+1$ entries, and an even number of $+1$ entries in each row, resulting in a determinant of 0. We have shown that all square Eulerian submatrices of $\mathbf{A}$ have a determinant of 0, or, equivalently, are singular matrices. Thus, by Camion's theorem, $\mathbf{A}$ is totally uniomodular. Since the right-hand-side of the constraint set is integer valued, by Cramer's Rule, all extreme point solutions to the LP are integer valued (Bazaraa et al., 2005). ∎

The constraint matrix structure of the n-HETM described above significantly reduces the number of integer variables that need to be processed by the branch-and-bound algorithm and allows the integer program to be solved with much less effort than it would require without this structure.

Special case for vehicles with single patient capacity:

If all vehicles can only transport a single patient at a time, i.e., $C_v = 1, \ \forall v \in V$, the number of decision variables can be reduced because the total number of patients transported on an arc by a certain vehicle type at a given time is equal to the number of vehicles of that type used, i.e., $\sum_{p \in P} x_{pvt}^{ij} = y_{vt}^{ij}$, thus constraint (4.7) is no longer necessary. In order to track the vehicles staying at the hospitals, $y_{vt}^{ii}, \ \forall i \in I, \ v \in V, \ t = 1, \cdots, T$ and $y_{vt}^{jj}, \ \forall j \in J, \ v \in V, \ t = 1, \cdots, T$ variables are still required. The constraints (4.9) - (4.11) are replaced by the following constraints (4.9′) - (4.11′).

$$\nu^i_{vt} + \sum_{\substack{j\in J:\\ t>\tau^{ij}}} \sum_{p\in P} x^{ji}_{pv(t-\tau^{ij})} + y^{ii}_{v(t-1)} = y^{ii}_{vt} + \sum_{j\in J}\sum_{p\in P} x^{ij}_{pvt}, \quad \forall i\in I,\ v\in V,\ t=1,\cdots,T \tag{4.9′}$$

$$\sum_{\substack{i\in I:\\ t>\tau^{ij}+2\gamma_v}} \sum_{p\in P} x^{ij}_{pv(t-\tau^{ij}-2\gamma_v)} + y^{jj}_{v(t-1)} = y^{jj}_{vt} + \sum_{i\in I}\sum_{p\in P} x^{ji}_{vt}, \quad \forall j\in J,\ v\in V,\ t=1,\cdots,T \tag{4.10′}$$

$$\sum_{j\in J}\sum_{v\in V}\sum_{f=t-\gamma_v+1}^{t} L_v\big(\sum_{p\in P} x^{ij}_{pvf}\big) \le L^i, \quad \forall i\in I,\ t=1,\cdots,T \tag{4.11′}$$

As a result of the substitution of x-variables for the vehicles in transit, the number of variables becomes $(|I||J||P|+|I|^2+|J|^2+|I|)|V|T$. The number of variables in the original formulation is $(|I||J||P|+|I+J|^2+|I|)|V|T$. By Cauchy-Schwartz inequality,

$$|I+J|^2 \ge |I|^2+|J|^2, \ \text{ then, } (|I||J||P|+|I+J|^2+|I|)|V|T \ge (|I||J||P|+|I|^2+|J|^2+|I|)|V|T.$$

Therefore, a significant reduction in the number of variables can be achieved by the assumption of single patient capacity for the vehicles. In fact, this is consistent with the fact that ambulances serve one patient at a time as the most frequently used medical transport vehicles in practice. However, the continuous relaxation of variables would not provide integer-values optimal solution to the problem with the single patient capacity assumption for vehicles.

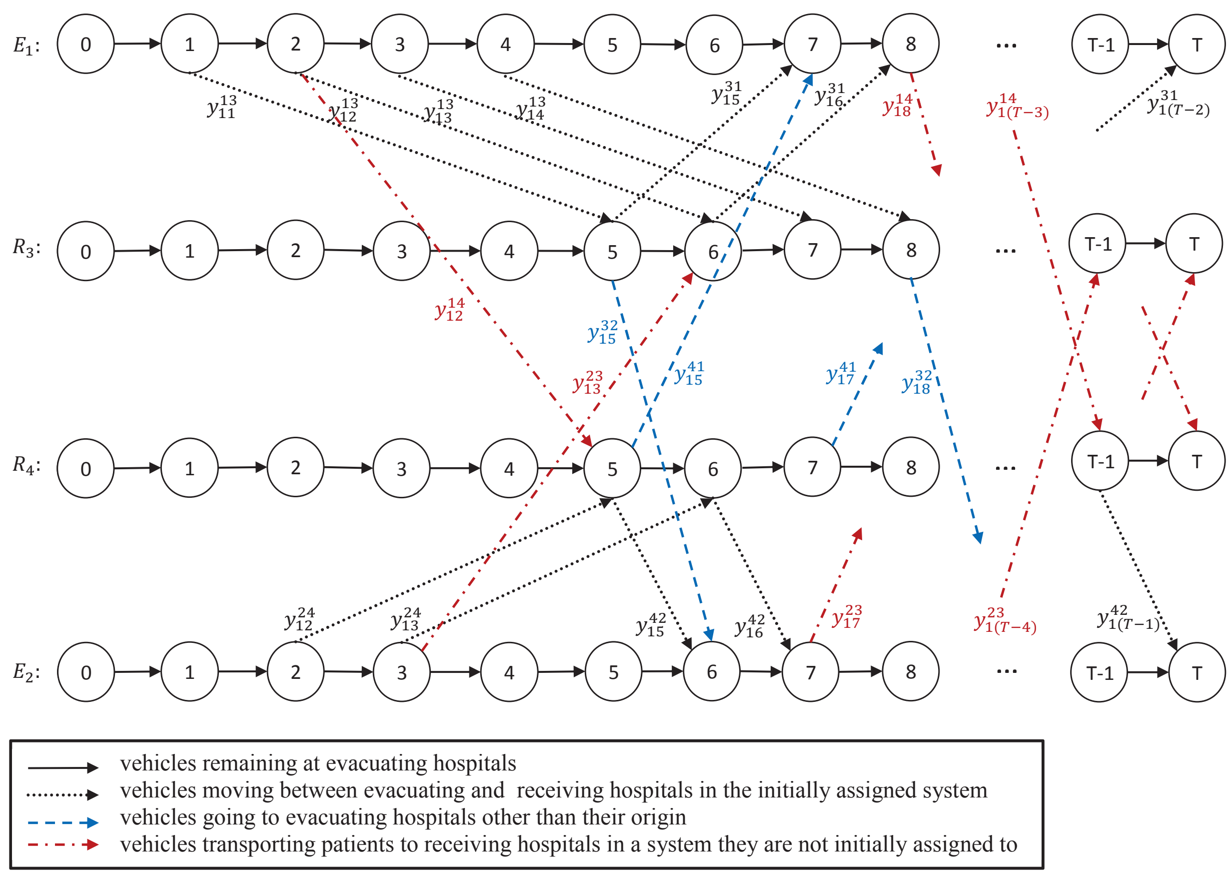
4.2.2 Resource management strategies

In the hospital evacuation problem studied, there are three main types of resources with different characteristics: (1) the hospital-specific loading capacity (L^i), (2) the fleet of vehicles (N_{vt}), and (3) the receiving beds (B^j_p). We consider centralized and decentralized resource management strategies in this study based on whether the latter two resources are shared by evacuating hospitals. The *loading capacity* of each hospital is controlled by the hospital and is not a resource that hospitals can share. It is a reusable resource and vehicles can use a portion of the loading capacity for more than one time interval, based on the loading time parameter. Each hospital can increase its loading capacity via improvements in emergency response training, staffing levels, staging capacity, or physical loading area. Unlike loading capacity, the *fleet of vehicles* can be shared by evacuating hospitals. The fleet size can potentially be increased by the central authority and the hospitals have less control over the fleet size than the loading capacity. Vehicles are also reusable resources and are occupied for multiple time intervals defined by the loading/unloading time and travel times. The *receiving beds* in different management groups can also be shared by evacuating hospitals. Unlike loading capacity and fleet of vehicles, receiving beds are not reusable and each bed is assigned to at most one patient during evacuation. It is difficult to increase the number of available beds especially during a disaster; however, if the bed supply is insufficient to accommodate patients, the study area can be expanded to include more receiving hospitals.

The performance of the evacuation transportation plan is determined by a set of limiting resources. If the solution can be improved by increasing the amount of a certain type of resource, then, *ceteris paribus*, this type of resource is currently a limiting resource. The physical loading capacity would be the limiting resource if there are enough number of vehicles available to transport patients at any time interval, but not all of these vehicles can be used due to the loading capacity. Similarly, the fleet size is the limiting resource when the loading capacity cannot be fully utilized due to the lack of enough vehicles. Also, the number and distribution of available beds affect the usage of vehicles through travel times. Given a large fleet size, the evacuation plan may be initially restricted by the loading capacity; but, at some point in evacuation, the restricting resource may become the fleet size due to long travel times. Since the loading capacity and fleet of vehicles are reusable resources in this time-expanded evacuation network, there can be multiple limiting resources depending on the time of observation.

The risk-minimizing n-HETM formulation is based on *centralized resource management strategy* where a central decision maker distributes all available vehicles and receiving beds among the evacuating hospitals. This provides the vehicles the flexibility of traveling to any evacuating hospital and transporting patients to any receiving hospital with the appropriate type of beds. The fleet management strategy used determines the network structure of the problem as well. The evacuation transportation network with two groups each having one evacuating hospital and one receiving hospital can be visualized by Figure 4.1. Assume that Group I consists of E_1 and R_3 and Group II consists of E_2 and R_4. The set of nodes in a row represents the corresponding hospital at the end of each time interval. Note that in this example, neither hospital can transport patients initially since the vehicles, that might be sent from a depot, have not arrived yet. Assuming the travel time from a central depot to the evacuating hospitals are $\tau^{d1} = 1$ and $\tau^{d2} = 2$, the first transportation arcs for evacuating hospital E_1 and E_2 are shown at time intervals 1 and 2, respectively.

Figure 4.1: A network illustration for two evacuating hospitals under centralized resource management.



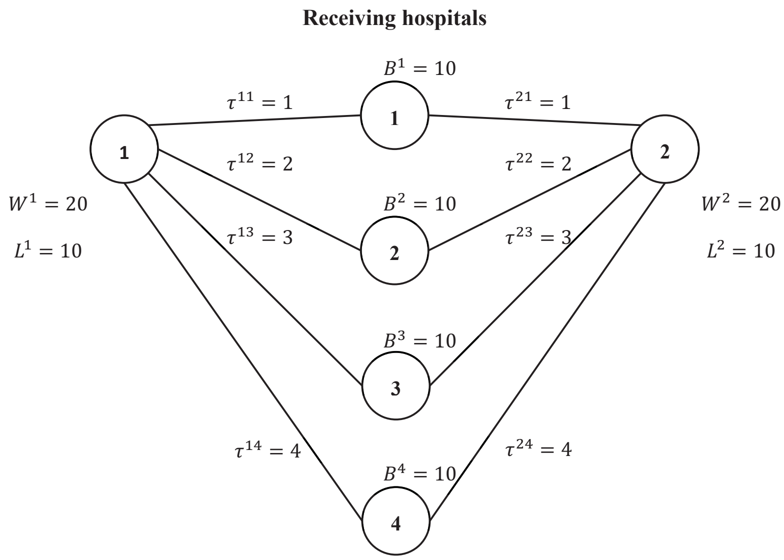
The dotted transportation arcs are associated with y_{vt}^{ij}- and y_{vt}^{ji}-variables where E_i and R_j are in the same group. The colored transportation arcs represent the vehicles that cross the system boundaries and travel to hospitals in a group that they are not initially in. The blue dashed arcs are associated with y_{vt}^{ji}-variables and the red dashed-and-dotted arcs are associated with y_{vt}^{ij}-variables, where E_i and R_j are in different groups. For example, in time interval 5, there are incoming vehicles to both R_3 and R_4 from E_1 and E_2 respectively, but some of these vehicles do not travel back to their origin. Instead, the vehicles from E_1 are sent to E_2 and vice versa via the blue dashed arcs. All the transportation arcs that originate from an evacuating hospital include the loading and unloading time (each assumed to be one time interval) plus the travel time. For example, the arc associated with the variable y_{12}^{14} has a length of four time intervals since the travel time between E_1 and R_4 is two time intervals and the loading and unloading times are one interval each. The transportation arcs originating from a receiving hospital have a length equal to only the travel time between the corresponding hospitals since the associated vehicles are not carrying any patients. The arcs connecting the consecutive nodes represent the vehicles that remain at an evacuating or receiving hospital because they could not be loaded at the corresponding time interval due to resource limitations and will be loaded and dispatched at a later time interval. These arcs are denoted by the y_{vt}^{ii} or y_{vt}^{jj}-variables in the n-HETM formulation.

In order to explore the equity of system-level risk minimization with centralized resources, consider a symmetric evacuation network where the evacuating hospitals are completely identical and receiving hospitals with enough number of beds to accommodate the patients are equidistant. Then the evacuation plan would depend mainly on the loading capacity of evacuating hospitals and the fleet size. It may be intuitive that the optimal evacuation plan would be equitable in such a symmetric network. However, it is observed that the optimal solution does not necessarily provide an equitable plan even in this case, as demonstrated with the example below.

Consider two evacuating hospitals with 20 patients of the same type as shown in Figure 4.2. Both evacuating hospitals have a vehicle loading capacity of 10 ambulances at a time. There are four receiving hospitals, each with 10 available beds for this single patient type which necessitates the utilization of all available beds. The receiving hospitals are equidistant from the evacuating hospitals such that $\tau^{ij} = j, \forall i \in \{1, 2\},\ j \in \{1, 2, 3, 4\}$. Note that the patient and vehicle indices are omitted in notation since only one patient type and one vehicle type are assumed.

The threat and transportation risk parameters used in computations are $\lambda = 0.008$ and $\theta = 0.004$. Assume that there are 10 ambulances available for transportation at the beginning

Figure 4.2: A symmetric network illustration with two evacuating hospitals.



of evacuation, all resources are shared, and, for simplicity, patients can be loaded into vehicles instantaneously, i.e., $\gamma = 0$. By the symmetry of the network, patients at the two evacuating hospitals should ideally receive the same level of service that is measured in terms of the risks determined by the time of the transportation assignment and the travel time to the assigned destination. Therefore, the average patient risk observed by the population in each evacuating hospital, denoted by $\overline{R}^i$, is of interest in order to evaluate the fairness of the generated plan from the perspective of evacuating hospitals. Equitable solutions would imply that the average risks for patients at evacuating hospitals are as close to each other as possible and small values of the range $Range_h = |\overline{R}^1 - \overline{R}^2|$ are desired for fairness.

First, we find the system-level minimum total risk, R^*, by solving the n-HETM. The least equitable optimal solution is found by minimizing the risk of one hospital at a time subject to the additional constraint (4.20) that defines R^* as the upper bound on the system-level risk, as shown below, in addition to the original constraints (4.5)-(4.14). The solutions of these optimization problems provide the range of optimal average patient risk for each hospital, denoted as $[\overline{R}^i_{LB}, \overline{R}^i_{UB}],\ \forall i \in I$.

$$\text{Minimize} \quad \sum_{j\in R}\sum_{p\in P}\sum_{v\in V}\sum_{t=1}^{T} R^{ij}_{pvt}x^{ij}_{pvt} = R^i \tag{4.19}$$

$$\text{subject to} \quad R \leq R^* \tag{4.20}$$

The initial problem of minimizing the system-level risk has multiple optimal solutions. The least equitable solution within the optimal solution pool is found using the objective (4.19) and constraint (4.20). The two extreme optimal solutions with the most and the least hospital-level equity for $L = 10, N = 10$ are shown in Figures 4.3 and 4.4, respectively, along with the optimal system-level average risk, $\overline{R}^*$, the average risk for each hospital, $\overline{R}^1$ and $\overline{R}^2$, and the time at which the last patient is loaded into an ambulance for each hospital, LT^1 and LT^2.

Despite the symmetry of the network, the optimal solution shown in Figure 4.4 results in a gap between the average risk experienced by patients in two evacuating hospitals. The system-level optimization provides the best total risk, but it does not necessarily provide an equitable plan in terms of the average risk experienced by patients in the two hospitals even for such a symmetric network. The equity of the evacuation plan can be improved by taking advantage of the network symmetry. The synchronous evacuation plan which is also another optimal solution produces the most equitable evacuation plan as shown in Figure 4.3. Note that the optimal solution consists of variables that satisfy the condition that $x^{1j}_t = x^{2j}_t, \forall j \in \{1, 2, 3, 4\},\ t = 1, \cdots, T$ in addition to the original n-HETM constraints (4.5)-(4.11).

The bottleneck or restricting resource of the evacuation plan can be either one of the two reusable resources, loading capacity (L) and fleet size (N), as discussed above. Table 4.1 below summarizes the most and the least equitable optimal solutions of various (L, N) scenarios including the average risk ($\overline{R}^*$), average threat risk ($\overline{R}_\Lambda$), and average transportation risk ($\overline{R}_\Theta$). Based on the reduction in $\overline{R}^*$ as the resource levels are varied, the restricting resource on the performance of the optimal evacuation plan are listed in the last column of Table 4.1. Recall that, for a specific problem instance, loading capacity and fleet size can become the bottleneck at different epochs of the planning horizon due to the dynamic changes in the level of available resources.

Threat risk depends on the time at which patients are loaded onto vehicles to be trans-

Figure 4.3: The most equitable optimal solution for the symmetric network example.

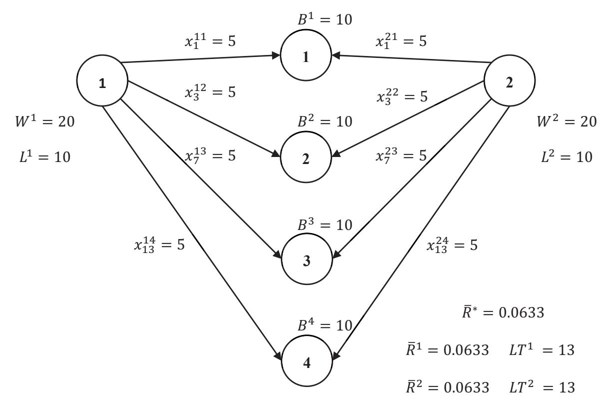
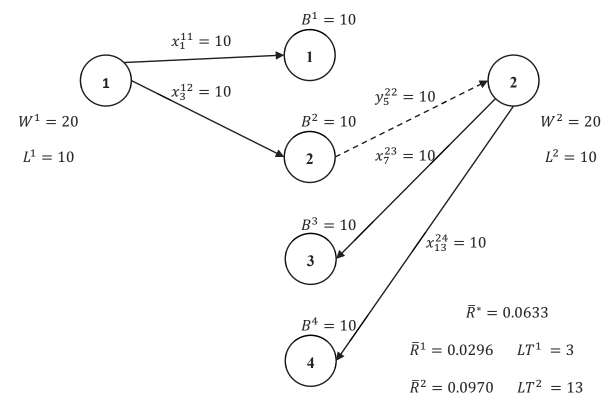


Figure 4.4: The least equitable optimal solution for the symmetric network example.



ferred and transportation risk depends on the length of time they travel. We observe that increasing the fleet size can only reduce the threat risk, mainly because patients can be transported earlier if there are more vehicles available unless the loading capacity is the single restricting resource, whereas the transportation risk may not necessarily be reduced because of the constant travel time parameters and available bed distribution of the underlying network. As a result of the more significant reduction in threat risk for the hospital that has a higher risk than that of the other hospital in the least equitable solution, for a fixed L-value, as N increases, the equity of the least equitable optimal solution is improved, that is, the difference of the average risks for the two hospitals diminishes as shown in Table 4.1 for the symmetric network example. For values of N above $2W$, the optimal solution remains the same since $N = 2W$ provides the best possible evacuation plan where the loading capacity is the only bottleneck.

If the loading capacity is reduced to 5 vehicles at a time for both hospitals, there is a single optimal solution for all N values that is the most equitable solution. This is a result of a combination of assumptions: First, the integer-valued travel times to receiving hospitals are distinct and an ordinal ranking can be used to assign patients to the hospitals. Second, since sending a patient to a closer hospital at a given time would always result in lower risk, it would be optimal to allocate half of the capacity of each receiving hospital to the patients from each evacuating hospital. In this symmetric example, the number of beds in

Table 4.1: The most and the least equitable risk-minimizing solutions for various (L, N) scenarios for the symmetric network example.

				Most equitable		Least equitable						Bottleneck
L	N	$\overline{R}^*$	R_{max}	$\overline{R}^1_\Lambda = \overline{R}^2_\Lambda$	$\overline{R}^1_\Theta = \overline{R}^2_\Theta$	$\overline{R}^1$	$\overline{R}^2$	$\overline{R}^1_\Lambda$	$\overline{R}^2_\Lambda$	$\overline{R}^1_\Theta$	$\overline{R}^2_\Theta$	resource
5	10	0.0633	0.1206	0.0540	0.0100	0.0633	0.0633	0.0540	0.0540	0.0100	0.0100	N
	20	0.0411	0.0697	0.0315	0.0100	0.0411	0.0411	0.0315	0.0315	0.0100	0.0100	L, N
	30	0.0373	0.0546	0.0277	0.0100	0.0373	0.0373	0.0277	0.0277	0.0100	0.0100	L
	40	0.0373	0.0546	0.0277	0.0100	0.0373	0.0373	0.0277	0.0277	0.0100	0.0100	L
10	10	0.0633	0.1206	0.0540	0.0100	0.0296	0.0970	0.0238	0.0843	0.0060	0.0139	N
	20	0.0373	0.0622	0.0276	0.0100	0.0315	0.0430	0.0238	0.0315	0.0080	0.0119	N
	30	0.0315	0.0470	0.0218	0.0100	0.0277	0.0354	0.0199	0.0238	0.0080	0.0119	L, N
	40	0.0296	0.0393	0.0199	0.0100	0.0277	0.0316	0.0199	0.0199	0.0080	0.0119	L
15	10	0.0633	0.1206	0.0540	0.0100	0.0296	0.0970	0.0238	0.0843	0.0060	0.0139	N
	20	0.0373	0.0622	0.0276	0.0100	0.0267	0.0478	0.0199	0.0354	0.0070	0.0129	N
	30	0.0296	0.0470	0.0199	0.0100	0.0276	0.0316	0.0199	0.0199	0.0080	0.0119	N
	40	0.0277	0.0393	0.0179	0.0100	0.0257	0.0296	0.0179	0.0179	0.0080	0.0119	L

each receiving hospital is divisible by two; therefore, we can send 5 patients, which is the maximum number of patients that can be sent at a time due to parameter L, from each evacuating hospital to a receiving hospital and share the beds equally. Any unequal capacity sharing among the patients of two evacuating hospitals would force delayed evacuation of some patients and increase the risk. However, if the travel time to all receiving hospitals from both evacuating hospitals were equal, i.e., $\tau^{ij} = \tau, \ \forall i \in E, j \in R$, the minimum total evacuation risk can be obtained by multiple optimal transportation plans as the basis of the optimal solution can be changed without affecting the time intervals at which patients reach their destination by simply swapping the destination hospitals for pairs of patients. Inequity can also be driven by a number of beds that is not divisible by two. For example, if $B^j = 11$ for any receiving hospital j, then, the problem has multiple optimal solutions at various levels of equity.

In addition to the hospital-level equity, one can evaluate the equity of optimal solutions from the patients' perspective. At the patient-level the even distribution of risk among patients is desirable. Therefore, the optimal solutions in which the maximum patient risk is minimized would be considered equitable. The most and least equitable optimal solutions for the min-max risk fairness criterion at the patient-level for the same (L, N) scenarios as above are provided in Table 4.2. Our first observation is that the maximum patient risk, R_{max}, is lower than the R_{max} of the risk-minimizing optimal solutions from Table 4.1 for the cases where $N \geq 20$. However, the average risk, $\overline{R}^i, \forall i \in \{1, 2\}$, of the most equitable optimal solutions at the hospital-level are increased at the expense of distributing the risk more equitably among patients.

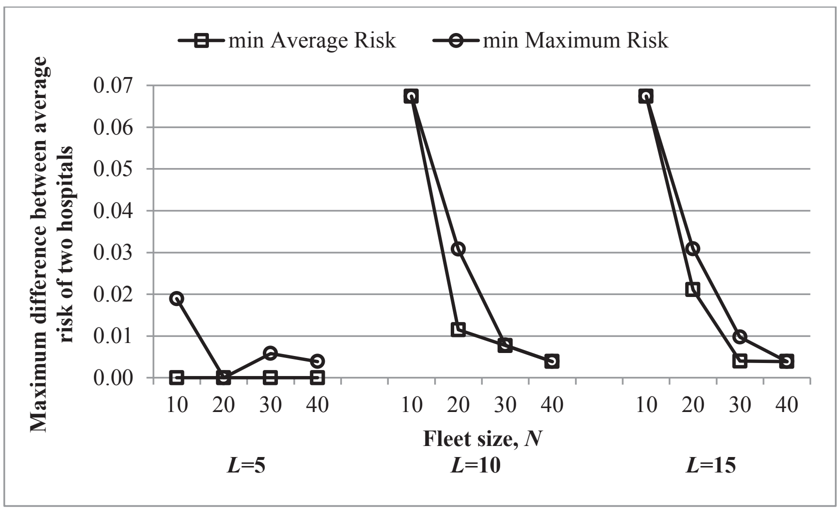
Table 4.2: The most and the least equitable min-max patient risk solutions for various (L, N) scenarios for the symmetric network example.

			Most equitable					Least equitable					
L	N	R^*_{max}	$\overline{R}^1 = \overline{R}^2$	$\overline{R}^1_\Lambda$	$\overline{R}^2_\Lambda$	$\overline{R}^1_\Theta$	$\overline{R}^2_\Theta$	$\overline{R}^1$	$\overline{R}^2$	$\overline{R}^1_\Lambda$	$\overline{R}^2_\Lambda$	$\overline{R}^1_\Theta$	$\overline{R}^2_\Theta$
5	10	0.1206	0.0670	0.0578	0.0578	0.0100	0.0100	0.0576	0.0766	0.0502	0.0654	0.0080	0.0119
	20	0.0584	0.0411	0.0315	0.0315	0.0100	0.0100	0.0411	0.0411	0.0315	0.0315	0.0100	0.0100
	30	0.0470	0.0374	0.0277	0.0277	0.0100	0.0100	0.0345	0.0403	0.0277	0.0277	0.0070	0.0129
	40	0.0432	0.0374	0.0277	0.0277	0.0100	0.0100	0.0354	0.0393	0.0277	0.0277	0.0080	0.0119
10	10	0.1206	0.0739	0.0660	0.0631	0.0084	0.0115	0.0296	0.0970	0.0238	0.0843	0.0060	0.0139
	20	0.0584	0.0400	0.0296	0.0311	0.0108	0.0092	0.0258	0.0565	0.0199	0.0432	0.0060	0.0139
	30	0.0393	0.0320	0.0218	0.0226	0.0104	0.0096	0.0277	0.0355	0.0199	0.0238	0.0080	0.0119
	40	0.0297	0.0297	0.0199	0.0199	0.0100	0.0100	0.0277	0.0316	0.0199	0.0199	0.0080	0.0119
15	10	0.1206	0.0705	0.0615	0.0610	0.0098	0.0102	0.0296	0.0970	0.0238	0.0843	0.0060	0.0139
	20	0.0584	0.0377	0.0280	0.0280	0.0100	0.0100	0.0238	0.0546	0.0179	0.0413	0.0060	0.0139
	30	0.0393	0.0306	0.0199	0.0218	0.0109	0.0090	0.0257	0.0355	0.0179	0.0238	0.0080	0.0119
	40	0.0316	0.0277	0.0179	0.0179	0.0100	0.0100	0.0257	0.0296	0.0179	0.0179	0.0080	0.0119

Considering the most equitable optimal solutions listed in Tables 4.1 and 4.2, we observe that given a symmetric network, there exists an optimal solution for both total risk minimization and maximum risk minimization problems where the average hospital risks are equal, i.e., $\overline{R}^1 = \overline{R}^2$. This average risk of the optimal solution for total risk minimization is a lower bound on that for maximum risk minimization.

The minimax risk objective results in greater upper bounds and smaller lower bounds on average patient risks for hospitals. This leads to a greater difference between the average patient risk of hospitals in the least equitable optimal solution than the total risk minimization objective. The greatest difference between average patient risks of the evacuating hospitals is shown in Figure 4.5. We also observe that as the fleet size is increased, the range for the average hospital risk gets tighter faster with total risk minimization than it does with minimax risk objective.

Figure 4.5: The maximum difference between average risk of two evacuating hospitals (from the least equitable optimal solution) for various (L, N) scenarios.



The symmetry of the above network can be broken in several ways by introducing: (1) uneven loading capacities, (2) uneven travel times, (3) uneven receiving bed distribution, and (4) uneven patient populations at the evacuating hospitals. Each of these changes can significantly change the optimal resource allocation, but we are especially interested in the impact of loading capacity as a hospital-specific resource that can be adjusted for improved evacuation performance. Therefore, we will begin with modifying the loading capacity to achieve an asymmetric network. When the loading capacity of one of the evacuating hospitals is increased, the intuitive expectation of the hospital with higher capacity would be reduced evacuation risks. However, the optimal solution minimizing the overall risk may have different outcomes.

Assume that the loading capacity of the second hospital is incremented such that $L^2 =$

$L^1 + 1$. The range of average patient risk for the hospitals in the risk-minimizing solutions for the above network example with unequal loading capacities are provided in Table 4.3. As the loading capacity of the second hospital is incremented, the optimal solution having the greatest possible risk for this hospital can only be improved, i.e., $\overline{R}^2_{UB}$ is non-increasing with L^2. For example, the optimal solutions with the greatest $\overline{R}^2$ for the cases where $L^1 = 5, L^2 = 6, N = 20$ and $L^1 = 5, L^2 = 6, N = 30$ are shown in Figures 4.6 and 4.7, respectively. The second hospital can send one more patient at a time, so, vehicles are transferred from the first hospital to the second one so that more patients are transferred earlier. This vehicle sharing strategy reduces the risk for both evacuating hospitals compared to the case where loading capacities are equal.

Table 4.3: The range of average risk for two hospitals among risk-minimizing solutions for various (L, N) scenarios when the symmetry is broken by changing L^2.

L^1,L^2	N	$\overline{R}^*$	$[\overline{R}^1_{LB}$,	$\overline{R}^1_{UB}]$	$[\overline{R}^2_{LB}$,	$\overline{R}^2_{UB}]$
5, 6	10	0.0633	0.0633	0.0633	0.0633	0.0633
	20	0.0407	0.0407	0.0428	0.0386	0.0407
	30	0.0362	0.0373	0.0385	0.0339	0.0350
	40	0.0362	0.0373	0.0385	0.0339	0.0350
10, 11	10	0.0633	0.0296	0.0970	0.0296	0.0970
	20	0.0373	0.0315	0.0440	0.0306	0.0430
	30	0.0313	0.0283	0.0356	0.0271	0.0344
	40	0.0294	0.0279	0.0318	0.0271	0.0310
15, 16	10	0.0633	0.0296	0.0970	0.0296	0.0970
	20	0.0373	0.0267	0.0488	0.0257	0.0478
	30	0.0296	0.0276	0.0327	0.0265	0.0316
	40	0.0275	0.0257	0.0300	0.0249	0.0292

In this example, the incremented loading capacity at the second hospital results in an increased initial fleet size for this hospital, i.e., 20 ambulances are assigned when $L^2 = 6$

Figure 4.6: The optimal solution with the greatest average risk for hospital 2 ($\overline{R}^2$) when $L^2 = L^1 + 1 = 6$ and $N = 20$.

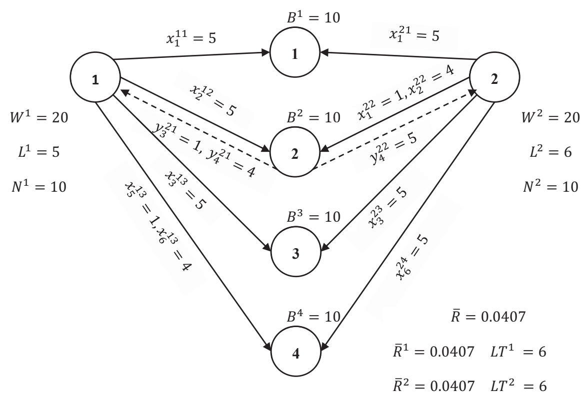
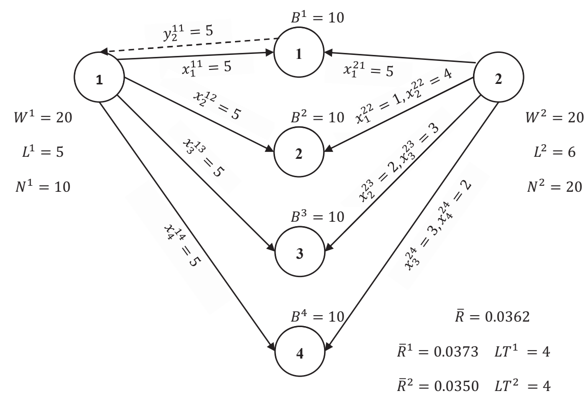


Figure 4.7: The optimal solution with the greatest average risk for hospital 2 ($\overline{R}^2$) when $L^2 = L^1 + 1 = 6$ and $N = 30$.



whereas 10 ambulances are assigned when $L^2 = 5$. However, the hospital that initially has more vehicles transfers the excess vehicles to the other hospital at the earliest time interval that the vehicles reach a receiving hospital. For example, in Figure 4.7, the vehicle transfer is from hospital 2 to hospital 1 since the initial vehicle distribution is 20 ambulances at hospital 2 and 10 at hospital 1. This vehicle transfer strategy is observed in many of the optimal solutions listed in these tables, including both the equal and unequal loading capacity cases, and the transfer always occurs at the earliest possible time interval.

As opposed to the centralized resource management strategy that is completely flexible, resource sharing can be impossible at times when the evacuating hospitals are isolated as a result of the threat disrupting the communication channels in a region. There can also be certain regulations or obligations such that a hospital group will deploy a fleet of vehicles for the evacuation of only their own hospitals or that the receiving facilities in each group will give higher priority to incoming patients from their own hospitals. We will define a second strategy named the *decentralized resource management strategy* under which each hospital group performs evacuation on their own without sharing resources with each other. These two strategies are employed in this study, because decentralized resources yield the most restrictive resource allocation and the most efficient (risk-minimizing) resource allocation is achieved by completely centralizing the resources; but we note that more moderate allocations of vehicles and prioritization rules for receiving beds can be developed aside from these two extreme strategies. These two strategies and the corresponding network structures are described below.

Decentralized resource management

The base case we will be using for benchmarking the solutions that consider different fairness criteria is the risk-minimizing model under decentralized resource management strategy. Let J^i be the set of receiving hospitals affiliated with evacuating hospital i and I^j be the set of evacuating hospitals affiliated with receiving hospital j. From here on, the set of affiliated hospitals will be called a *group* of hospitals. Each hospital group uses both their own fleet of vehicles and the beds at the receiving hospitals within their group. The vehicle flow conservation constraints (4.8) - (4.10) of the n-HETM are modified as in constraints (4.21) - (4.24) to reflect this strategy. The n-HETM-D does not utilize the ν-variables since the fleet size and mix is known for each hospital group.

***n*-HETM-D:**

$$\sum_{v\in V}\sum_{t=1}^{T}(y_{vt}^{ij}+y_{vt}^{ji})=0, \quad \forall i\in I,\ j\notin J^i \tag{4.21}$$

$$\sum_{j\in J}\sum_{v\in V} y_{v0}^{jj}=0 \tag{4.22}$$

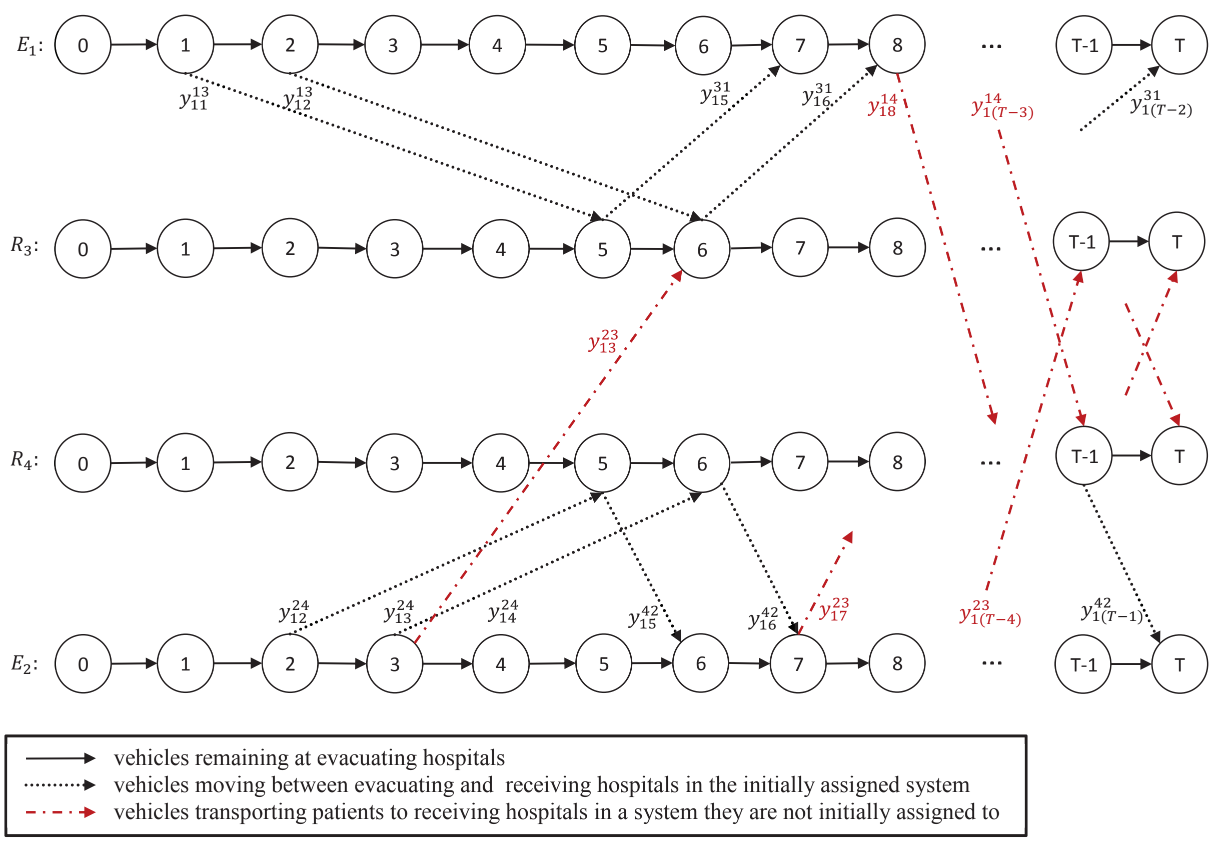
$$N_{vt}+\sum_{\substack{j\in J:\\ t>\tau^{ij}}} y_{v(t-\tau^{ij})}^{ji}+y_{v(t-1)}^{ii}=y_{vt}^{ii}+\sum_{j\in J} y_{vt}^{ij}, \quad \forall i\in I,\ v\in V,\ t=1,\cdots,T \tag{4.23}$$

$$\sum_{\substack{i\in I:\\ t>\tau^{ij}+2\gamma_v}} y_{v(t-\tau^{ij}-2\gamma_v)}^{ij}+y_{v(t-1)}^{jj}=y_{vt}^{jj}+\sum_{i\in I} y_{vt}^{ji}, \quad \forall j\in J,\ v\in V,\ t=1,\cdots,T \tag{4.24}$$

Decentralized management of resources can have its advantages if each hospital group has enough resources to transport and provide continuous care for its patients. However, this is generally not the case since demand for resources is higher than usual in emergencies. Then, the evacuation problem may be infeasible with only the regular resource levels and planning for surge capacity might be required which in practice encompasses several resources including receiving beds, staff, transportation vehicles, and medical equipment. Based on our modeling assumptions, capacity of the receiving hospitals can cause infeasibility of the problem under decentralized strategy and feasibility can be achieved by introducing a surge hospital with enough number of receiving beds. Even if the decentralized strategy is feasible to implement, the resulting evacuation plan would impose high risks on patients if lower risks are achievable by transportation of patients to closer hospitals in other hospital groups or by simply sharing the fleet of vehicles.

As a modified version of the decentralized fleet strategy, each hospital group can be forced to use its own resources throughout the evacuation while making excess receiving beds at each group available to the other one after all possible patient-to-bed assignments are made within each group. Figure 4.8 demonstrates the decentralized strategy on an evacuation network with two hospital groups each having one evacuating and one receiving hospital such that E_1 and R_3 are in the first group and E_2 and R_4 are in the second group.

Figure 4.8: A network illustration for two evacuating hospitals under decentralized resource management.



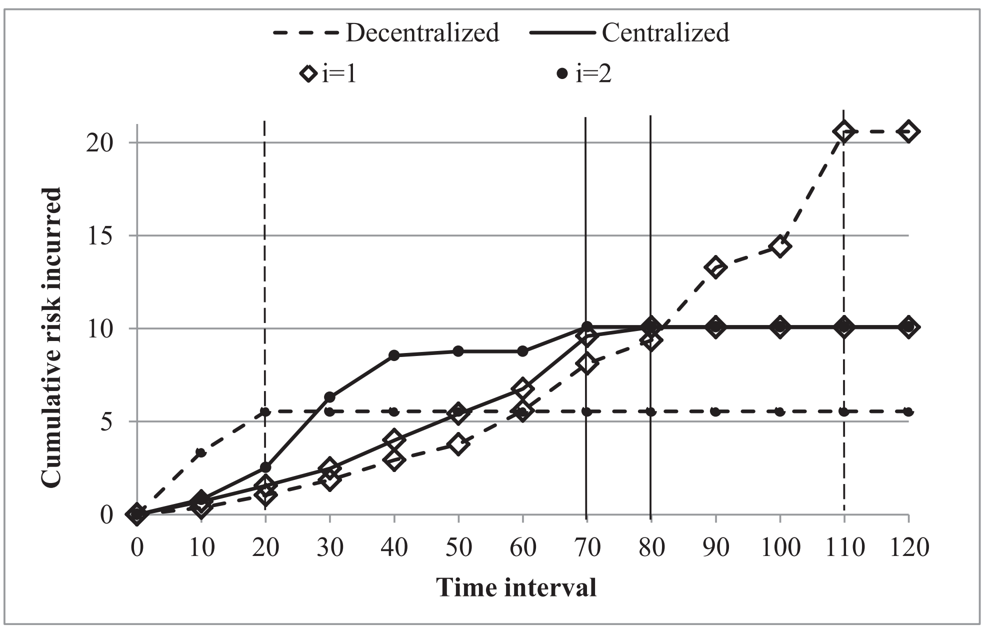
As in Figure 4.1, the vehicles arrive at the evacuating hospitals, possibly from a central depot, in 1 and 2 time intervals. The dotted arcs represent the y_{vt}^{ij} and y_{vt}^{ji}-variables where both E_i and R_j are in the same group as before. The second type of transportation arcs correspond to the y_{vt}^{ij} and y_{vt}^{ji}-variables where E_i and R_j are in different groups. Such inter-group transportation arcs would exist only if a group cannot accommodate its whole evacuee population using the beds at its receiving facilities.

Although the number of available beds in a group might be enough to serve its evacuee population, the level of treatment that can be provided along with those beds may not be a perfect match for every patient. As explained above, the bed types represent the care level that is provided to patients admitted to those beds. Since one of the main assumptions of the n-HETM is that every patient is assigned to an adequate bed of the same type, the evacuation plan would search for the beds that the remaining patients in each group can be assigned to according to their type. In this figure, patients at E_2 are assigned to beds of appropriate type at R_4 and the remaining patients for whom a matching bed is not found in R_4 are sent to R_3. Similarly, patients from E_1 who cannot be accommodated at R_3 are sent to R_4. In order to solve the n-HETM under this strict strategy, a two-phase solution algorithm can be used if the number of excess beds in other hospital groups is not known initially. First, the n-HETM is solved for each group individually and the number of remaining patients in the evacuating hospitals and the number of excess beds in the receiving hospitals, by type, are reported. Then, the data is adjusted by supplementing the necessary capacity for each group from one another and the n-HETM is solved for each group. However, if the hospital groups are already sharing information about the number and mix of excess beds, then, the excess beds in other groups can be included in the initial data set, which is similar to introducing surge capacity. The optimal evacuation plan can include vehicles that carry patients to the other group's receiving hospital before all matching beds in the group are occupied as the vehicles represented by variable y_{13}^{23} in Figure 4.8.

In both of the above networks in Figures 4.1 and 4.8, the incoming flow to node t for an evacuating hospital is calculated in terms of ν- and y-variables both of which are integer based on the conservation of flow constraints (4.7) and (4.8). Given a feasible set of y-variables, the ν-variables would be integer valued as well. The incoming flow to node t for a receiving hospital is calculated in terms of the integer y-variables based on the conservation of flow constraints (4.9). There is an x-variable and an evacuation risk parameter R^{ij}_{pvt} associated with each transportation arc in the network for each patient type. As shown in Proposition 4.1, if a set of feasible y-variables is known, the network can be reduced to a feasible sub-network, given which a solution with integer valued x-variables can be found.

For convenience, we introduce some notation for the risks resulting from different resource management strategies. Let $\overline{R}^i_C$ and $\overline{R}^i_D$ be the average patient risks for hospital $i \in I$ for the risk-minimizing solutions under centralized and decentralized strategies, respectively. We will use the notation $\overline{R}_{C,h}$ and $\overline{R}_{D,h}$ for the average hospital risks under the two strategies where $\overline{R}_h = \frac{\sum_{i \in I} \overline{R}^i}{|I|}$ in general.

Figure 4.9: The impact of resource centralization on the cumulative risk and evacuation time.



There are two possible outcomes of centralizing the fleet of vehicles and the set of receiving beds: (1) risk can be reduced for all evacuating hospitals, i.e., $\overline{R}^i_C \leq \overline{R}^i_D$, $\forall i \in I$, and (2) risk can be increased for at least one evacuating hospital while it is reduced for the others, i.e., $\exists i \in I : \overline{R}^i_C > \overline{R}^i_D$. The outcome depends on the network properties including the demand (number of patients to be transported) and supply (number of available receiving beds) distribution over the network as well as the transportation times between hospitals. These two cases will be demonstrated in Section 4.3 following the introduction of the equity modeling framework. Figure 4.9 illustrates the impact of centralizing resources on the accumulation of risk for a problem instance with two evacuating hospitals where one of the hospitals helps reduce the risk by increasing its own risk. The time at which the last patient is transported from the evacuating hospitals are marked by the vertical lines for both hospitals under the two resource management strategies. Decentralized resources causes a much longer evacuation for the first hospital while the second hospital completes evacuation in just about 20 time intervals. This can be due to a smaller population size and/or vast amount of resources at the second hospital compared to the first hospital. When all resources are shared, the two hospitals compromise for the system-optimal solution where the second hospital completes evacuation much later (in 70 time intervals) than it can by not sharing resources. We note that the system-optimal solution provided is the most equitable one, but since the problem can have multiple optima all optimal solutions do not necessarily result in equal average patient risks for hospitals as in this example. As a result of resource centralization, the evacuation time, and therefore, risk, for the first hospital is significantly improved (total risk is reduced by 50%), but at the expense of an even more significant 80% increase in the second hospital's total risk. So, centralization can improve the risk distribution to some extent, but additional criteria for fairness of this distribution should be set for more reliable results. For this reason, we discuss issues with measures of equity that must be considered and propose an equity modeling framework for hospital evacuation planning next.

4.3 Equity Modeling Framework for Hospital Evacuations

Equity is a major concern in the delivery of public resources in addition to effectiveness and efficiency. Distributional equity is relevant for any problem that involves resource allocation such as hospital evacuation planning that requires the fair allocation of resources to patients over time to alleviate risks. There is no universally accepted definition for equity or a consensus on equity measures that decision makers should use since it is a highly subjective notion. However, an equity framework can be developed for this specific problem using coherent definitions and criteria of equity.

Measures of inequality have been studied extensively in sociology and economics to compare the distribution of assets or income among social units or players (see Sen (1973) for a discussion of economic inequality measures). In economics, a basic theory developed independent of welfare economics introduced the most commonly used measures of inequality such as the coefficient of variation, the mean relative deviation, the Gini coefficient, Theil's measure, and the standard deviation of the logarithms of each unit's assets (Allison, 1978). However, it is shown that these measures may not evaluate inequality properly when the equity ranking of the compared income distributions change at different percentages of population (Schwartz and Winship, 1980). Atkinson (1970) proposed a measure of income inequality, called the Atkinson index, based on this approach. The Atkinson index is an equally distributed equivalent measure in that it is defined based on the level of per capita income that would give the same level of social welfare as a given income distribution, if equally distributed. Each social unit or player i has a non-negative utility, denoted by the utility function u^i, and the importance of equity is denoted by the parameter $\alpha \geq 0$. The welfare function of utilities depends on this single parameter, such that the greater the α, the more important equity is in decision making. Since the utility distribution is expected to become fairer as α is increased, this fairness scheme is also called *α-fairness*. The piecewise non-linear social welfare function, $W_\alpha(u)$, is defined as follows:

$$W_\alpha(u) = \begin{cases} \displaystyle\sum_{i=1}^{n} \frac{(u^i)^{1-\alpha}}{1-\alpha} & \text{for } \alpha \geq 0,\ \alpha \neq 1, \\ \\ \displaystyle\sum_{i=1}^{n} \log(u^i) & \text{for } \alpha = 1 \end{cases} \tag{4.25}$$

The corresponding inequity measure, namely the Atkinson index, is

$$A_\alpha = \begin{cases} 1 - \left[\dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \left(u^i / \overline{u} \right)^{(1-\alpha)} \right]^{\frac{1}{1-\alpha}} & \text{for } \alpha \geq 0,\ \alpha \neq 1, \\ \\ 1 - \left[\displaystyle\prod_{i=1}^{n} \left(u^i / \overline{u} \right) \right]^{\frac{1}{n}} & \text{for } \alpha = 1 \end{cases}$$

where $\overline{u} = \frac{1}{n} \sum_{i=1}^{n} u^i$. Note that the Atkinson index equals zero if all utilities are the same, i.e., $A_\alpha = 0$ if $u^i = \overline{u},\ \forall i = 1, \cdots, n$, and equals 1 if there is no distribution, i.e., $A_\alpha = 1$ if $u^i = 0,\ \forall i = 1, \cdots, n$. This index can be intuitively interpreted as follows: If $A_\alpha = 0.1$, this means that if the income is equally distributed, we can achieve the same level of social welfare with $1 - A_\alpha = 0.9 = 90\%$ of income.

Given the problem-specific utility function and the α value, this non-linear social welfare function must be maximized to produce the optimal utility distribution among players. This problem requires the maximization of a non-linear objective function of continuous decision variables subject to the linear set of constraints of the n-HETM. In this study, we use the welfare economics approach to measure the equity of evacuation plans based on the utility received by the players, i.e., either evacuating hospitals or patient groups of different types. There are three particular α values of interest corresponding to different equity criteria. The welfare function represents the utilitarian criterion when $\alpha = 0$, the proportional fairness criterion when $\alpha = 1$, and the max-min fairness criterion as $\alpha \to \infty$, as explained next. Table 4.4 summarizes the objective functions corresponding to these three fairness criteria.

When $\alpha = 0$, the objective is maximizing the sum of player utilities and argmax $\sum_{i=1}^{n} u^i$ is the *utilitarian* solution. The utilitarian criterion provides the optimal system-level solution and does not consider distributional equity among players. If the players are evacuating hospitals, the most efficient allocation of resources is achieved resulting in the greatest sum of utilities for the hospitals. However, the most efficient resource allocation may be quite inequitable from either the hospitals' or patients' perspective if some hospital or patient is exposed to much higher risks than others for the sake of the overall good. Since the utilitarian objective is neutral to the inequity of resource allocation or risk distribution among evacuating hospitals or patients, we need to consider other equity criteria as well.

Proportional fairness solution, achieved when $\alpha = 1$, is argmax $\sum_{i=1}^{n} \log(u^i)$. This criterion is based on the Nash solution to the bargaining problem for two players in which a change in the allocation of resources is favorable only if the percentage increase in one player's utility exceeds the percentage decrease in the other player's utility (Nash, 1950). When there a more than two players, a proportionally fair resource allocation would be one where the total proportional change in the utilities of players is non-negative. That is, the proportionally fair utility distribution, u^i_{Pr}, must satisfy the condition $\sum_{i=1}^{n} \frac{u^i_{Pr} - u^i}{u^i_{Pr}} \geq 0$, for any feasible utility distribution u^i. The objective function for this criterion is equivalent to maximizing the product of utilities, i.e., argmax $\sum_{i=1}^{n} \log(u^i)$ = argmax $\prod_{i=1}^{n} u^i$, since $\sum_{i=1}^{n} \log(u^i) = \log(\prod_{i=1}^{n} u^i)$ and the logarithm function is strictly increasing on $u^i > 0$.

Max-min fairness solution, achieved as $\alpha \to \infty$, is based on the Rawlsian justice introduced by Rawls (1971) and the Kalai and Smorodinsky (1975) solution to Nash's bargaining problem. This type of objectives either minimize the largest loss or, equivalently, maximize the smallest gain for resource allocation problems. In the hospital evacuation context, max-min fairness is achieved when the minimum player utility is maximized by argmax $\min_{i=1,\cdots,n} \{u^i\}$. This extreme fairness criterion results in the most even utility distribution possible among players.

Table 4.4: Summary of the equity modeling framework.

	Fairness criteria		
	Utilitarian ($\alpha = 0$)	Proportional fairness ($\alpha = 1$)	Max-min fairness ($\alpha \to \infty$)
Objective	Max $\sum_{i \in I} u^i$	Max $\sum_{i \in I} \log(u^i)$	Max $\min_{i \in I} \{u^i\}$

The social welfare function based on the Atkinson index is an appropriate choice for the hospital evacuation problem due to its characteristics that are aligned with the most commonly suggested criteria to take into account when selecting an (in)equity measure. These criteria are reviewed in detail by Marsh and Schilling (1994). We discuss the relevant criteria found in the literature in the context of hospital evacuations.

Selection criteria of inequity measures for hospital evacuations

1. **Appropriateness:** The inequity measure selected should be appropriate managerially such that it reflects the values and preferences of the decision maker or the players. It should be intuitive and comprehensible enough to interpret. The decision maker in a hospital evacuation planning problem considers the risk of each evacuation decision to choose the best plan and the social welfare function can be adapted to this framework in that the utilities of players can be represented in terms of risks. The importance of equity for the decision maker is represented by the inequity aversion parameter, α, which enables us to comprehend and explain the results. The Atkinson index is a flexible measure since it becomes more sensitive to inequalities at the bottom of the utility distribution as α increases unlike other measures such as the Gini coefficient that does not give varying weights to different parts of the utility distribution. In this respect, the social welfare function provides an appropriate measure of the equity of hospital evacuation plans.

2. **Impartiality:** An inequity measure should be impartial among the players or groups being evaluated by avoiding any association with characteristics such as social class, race, gender, or age. Measuring equity using the social welfare function is impartial since all the players are treated equally by this function with an arbitrary indexing.

3. **Analytic tractability:** Computational complexity of the inequity measure is an undeniably important aspect to consider. Some measures such as the linear Gini coefficient or the maximum difference between utilities of pairs of players are easy to solve compared to measures like coefficient of variation. It is also crucial to identify whether an optimal solution is required or an approximation is sufficient. Due to its additive form, the solution of the social welfare function does not require an excessive computational effort.

4. **Scale invariance:** An inequity measure is scale invariant if the degree of equity remains unchanged when the utilities of all players are multiplied by a constant. Scale

invariance holds for the social welfare function since the optimal utility distribution among hospitals (or patients) would be the same when each hospital's (or patient's) utility is multiplied by the same constant. The only thing that would change is the optimal objective function value, $W^*_\alpha(u)$.

5. **Principle of transfers:** The condition suggested by Pigou (1912)-Dalton (1920) is that the equity increases as income is transferred from a rich player to a poorer player as long as the rich player remains richer than the poor player. This principle holds for the social welfare function since the inequity measure, A_α, cannot increase if a transfer of $\delta > 0$ is made from a player with utility u^i to a player with utility u^j such that $u^i - \delta > u^j + \delta$. The social welfare function's sensitivity to utility transfers depends on the parameter α.

6. **Pareto optimality:** As the solution improves, none of the players should be worse off according to Pareto optimality. In the hospital evacuation setting, this would mean that as $W_\alpha(u)$ increases, there should not be any player that has a decreased utility. However, this principle does not hold for most of the inequity measures including the social welfare function since a utility transfer decreasing one player's utility may result in a better system-level solution, i.e., higher $W_\alpha(u)$.

7. **Decomposability:** The equity of the whole system should be defined as a function of equity within the components according to decomposability principle. This is especially important when the system is divided into several groups as is the case with hospital evacuations. The inequity of the whole evacuation system is the sum of the inequities for all players, therefore the social welfare function satisfies the decomposability principle.

8. **Normalization:** The inequity measures can be normalized such that their values range from 0 to 1, where 0 represents perfect equity. Normalized measures make it possible to compare distributions of utility with different units of measure. As mentioned above, the Atkinson index equals 0 when all player utilities are the same and 1 when the utilities are all zero, therefore, this is a normalized measure.

9. **Principle of population:** The inequity measure should be independent of the population size, n, i.e., the number of players, so that the equity levels of different groups can be compared. For finite n, the Atkinson index depends on the population size, but it can be divided by its upper bound to make it independent of n. Since a single group of n players is considered in our study, this principle can be avoided.

According to the above selection criteria, the social welfare function based on the Atkinson index is flexible in terms of the decision maker's aversion to inequity, impartial among players, computationally tractable, scale invariant, complying with the principle of transfers, decomposable, and normalized. Therefore, this function allows an adequate equity measurement of hospital evacuation plans.

The equity modeling framework explained above requires utility functions to be defined for the players involved. An evacuation plan can be evaluated in terms of its equity from the perspective of different players involved in the operations, such as patients, hospital management, medical transport providers, and first responders. Satisfying the expectations of all stakeholders simultaneously is usually not feasible. The definition of equity would vary among these players and different definitions can lead to extremely different resource management decisions. The major players concerned with fairness in hospital evacuation planning are evacuating hospitals and patients. Therefore, we focus on the equity of resource allocation, thus, risk distribution, at the hospital- and patient-levels.

One of the significant aspects of hospital evacuation planning is the non-homogenous evacuee population and providing the adequate resources timely to especially patients in critical condition is of the highest importance. Patients grouped according to their medical requirements could be considered as players; however, characterizing fairness of resource allocation among patient types is problematic. Normalizing and scaling the risks for these groups of patients is not coherent with the above criteria for equity measurement. When there is variability in the mix and location of available beds among receiving hospitals for each patient type, it is not reasonable to expect converging per patient risks for each group. The equity of evacuation plans is not discussed at the patient type-level in this study due to the issues regarding the normalization and scaling of risks. Therefore, we concentrate on the hospital-level and patient-level equity as described next.

4.3.1 Hospital-level equity

The evacuation plans that are regarded equitable by the evacuating hospitals should result in (1) the minimum possible system-level average patient risk, (2) the minimum gap between the average patient risks for different evacuating hospitals, and (3) an improvement over the average patient risks for hospitals based on the decentralized resource management strategy. To represent these equity criteria, we consider two utility function definitions for the evacuating hospitals: (1) utility that is complementary to risk and (2) benchmark-based utility defined as risk reduction. Note that utility based on the first definition is always non-negative, but it must be restricted to be non-negative when the second definition is being used.

Utility complementary to risk

In the n-HETM formulation, the penalty coefficients are the cumulative risk parameters corresponding to each evacuation decision. The utility of each decision can be defined as the complementary value to its risk, $1 - R_{pvt}^{ij}$. Therefore, the utility of an evacuating hospital would be an aggregate measure of these individual utilities. In order to make the hospital-level utilities comparable with each other, we are interested in the average patient utility complementing risk for hospital $i \in I$ considering the cases where evacuating hospitals may have different population sizes. Note that, at this level, different patient types are treated equally in the utility function as the equity of interest is among evacuating hospitals rather than patient types.

$$u_1^i = \frac{\sum_{j\in R}\sum_{p\in P}\sum_{v\in V}\sum_{t=1}^{T}(1 - R_{pvt}^{ij})x_{pvt}^{ij}}{\sum_{p\in P} W_p^i} = 1 - \frac{R^i}{W^i} = 1 - \overline{R}^i \in [0,1], \quad \forall i \in I \tag{4.26}$$

where $W^i = \sum_{p\in P} W_p^i$ and $R^i = \sum_{j\in R}\sum_{p\in P}\sum_{v\in V}\sum_{t=1}^{T} R_{pvt}^{ij} x_{pvt}^{ij}$.

Benchmark-based utility

Equity measures can be defined in terms of the deviation from a given level of utility for each player. Morton et al. (2003) considers different benchmark-based utility functions for a portfolio allocation problem where they define the benchmark as the *target* utility level and try to minimize the deviation from that level. The benchmark can also be interpreted as the worst-case utility level that the decision maker is trying to move away from. Bertsimas et al. (2012) uses an air traffic control example and defines the utility function as the delay reduction for airlines compared to a scheduling procedure that prioritizes flights based on their original schedule. This benchmark is selected such that the schedules generated by the maximization of the social welfare function of airline delay reductions always improve upon the current plan.

In the n-HETM, the two types of resources that can be shared by evacuating hospitals are the fleet of vehicles and the receiving beds. Under the decentralized resource management strategy, the evacuating hospitals are assumed to use their own fleet and the receiving beds at the hospitals in their own group independently. Such a restricted resource allocation strategy can only generate evacuation plans that are inferior to the plans under the centralized resource management strategy in terms of the overall risk. Therefore, we can evaluate the relative improvement in risk as we change our resource management strategy and fairness criterion.

In order to account for the inherent differences among evacuating hospitals and the asymmetry in the evacuation network, we will be using the utilitarian solution under the decentralized resource management strategy as the base case and compare the optimal solution under centralized resource management strategy to this base case solution. This utility function represents the average patient risk reduction for hospital $i \in I$.

$$u_2^i = \frac{R_D^i - \sum_{j \in R} \sum_{p \in P} \sum_{v \in V} \sum_{t=1}^{T} R_{pvt}^{ij} x_{pvt}^{ij}}{\sum_{p \in P} W_p^i} = \frac{R_D^i - R^i}{W^i} = \overline{R}_D^i - \overline{R}^i \in [0,1], \quad \forall i \in I \tag{4.27}$$

where $\overline{R}_D^i$ is the average patient risk for hospital i for the risk-minimizing solution under decentralized resource management strategy.

Discussion of the hospital-level utility functions

In order to understand the solution behavior of utility function u_2^i compared to u_1^i, consider the relationship between the utility levels based on the two definitions. The utility functions are defined as

$$u_1^i = 1 - \overline{R}^i \in [0,1] \text{ and } u_2^i = \overline{R}_D^i - \overline{R}^i \in [0,1].$$

By definition, u_2^i is non-negative, which implies that solutions that make at least one of the hospitals worse off compared to the decentralized strategy solution are not acceptable when using u_2^i in optimization. The definition of u_1^i does not imply such a requirement, but if we add the non-negative risk reduction constraint, $\overline{R}^i \leq \overline{R}_D^i$, when using u_1^i in optimization, the two problems would be defined on the same set of feasible solutions which we will interpret in terms of the average patient risk distribution among evacuating hospitals, $\left(\overline{R}^1, \cdots, \overline{R}^{|I|}\right)$. We will refer to u_1^i subject to non-negative risk reduction as $u_{1'}^i$, i.e., $u_{1'}^i = 1 - \overline{R}^i : \overline{R}^i \leq \overline{R}_D^i$.

The risk measures obtained from the optimal solutions when the inequity aversion parameter is α and the welfare function is defined by utility function $f = 1, 1', 2$ for u_1^i, $u_{1'}^i$, and u_2^i, respectively, are represented by the following notation.

Average patient risk for hospital $i \in I$:

$$\overline{R}_f^{i*}(\alpha) = \frac{R_f^{i*}(\alpha)}{\sum_{p \in P} W_p^i}$$

Average of average patient risks for evacuating hospitals:

$$\overline{R}_{fh}^{*}(\alpha) = \frac{\sum_{i \in I} \overline{R}_f^{i*}(\alpha)}{|I|}$$

Minimum and maximum average patient risks for evacuating hospitals:

$$R_{fh,min}^{*}(\alpha) = \min_{i \in I} \left\{ \overline{R}_{fh}^{*}(\alpha) \right\}$$
$$R_{fh,max}^{*}(\alpha) = \max_{i \in I} \left\{ \overline{R}_{fh}^{*}(\alpha) \right\}$$

Range of average patient risks for hospitals in the system:

$$Range_{fh}^{*}(\alpha) = |R_{fh,max}^{*}(\alpha) - R_{fh,min}^{*}(\alpha)|$$

Average patient risk of the system:

$$\overline{R}_f^*(\alpha) = \frac{\sum_{i \in I} R_f^{i*}(\alpha)}{\sum_{i \in I} \sum_{p \in P} W_p^i}$$

Minimum and maximum patient risks of the system

$$R_{f,min}^*(\alpha) = \min_{i \in I, j \in J, p \in P, v \in V, t=1,\cdots,T} \left\{ R_{pvt}^{ij} : x_{pvt}^{ij} > 0 \right\}$$
$$R_{f,max}^*(\alpha) = \max_{i \in I, j \in J, p \in P, v \in V, t=1,\cdots,T} \left\{ R_{pvt}^{ij} : x_{pvt}^{ij} > 0 \right\}$$

Range of patient risks of the system

$$Range_f^*(\alpha) = |R_{f,max}^*(\alpha) - R_{f,min}^*(\alpha)|$$

The above notation will be used in the following discussion of the two hospital-level utility functions and the related examples from hereon.

Utilitarian criterion:

At $\alpha = 0$, the objective functions

$$\text{Max} \quad \sum_{i \in I} \left(1 - \overline{R}^i\right) \quad \text{and} \quad \text{Max} \quad \sum_{i \in I} \left(\overline{R}_D^i - \overline{R}^i\right)$$

defined by utility functions u_1^i and u_2^i, respectively, are both equivalent to minimizing the sum of average patient risks for all hospitals, i.e., minimizing $\sum_{i \in I} \overline{R}^i$, when the two problems have the same feasible solution space as a result of adding the non-negative risk reduction constraint to the problem using u_1^i. Therefore, $\overline{R}_{1'}^{i*}(0) = \overline{R}_2^{i*}(0)$ and the objective based on $u_{1'}^i$ is interchangeable with the objective using u_2^i when $\alpha = 0$.

The solution space of the problems based on $u_{1'}^i$ and u_2^i is more restricted than that of the problem based on u_1^i. Therefore, the sum of optimal average patient risks over all hospitals based on u_1^i is less than or equal to that based on $u_{1'}^i$ and u_2^i, i.e.,

$$\sum_{i \in I} \overline{R}_1^{i*}(0) \leq \sum_{i \in I} \overline{R}_{1'}^{i*}(0) = \sum_{i \in I} \overline{R}_2^{i*}(0).$$

In terms of the transfer of risks between hospitals compared to the solution for u_2^i, u_1^i can result in solutions where the average patient risk is increased for some $i \in I$ as long as the sum of average patient risks over all hospitals is reduced. This requires that the total increase in per patient risk for hospitals is less than or equal to the total decrease. That is,

$$\sum_{i\in I}\left(\overline{R}_1^{i*}(0)-\overline{R}_2^{i*}(0)\right)^+ \le \sum_{i\in I}\left(\overline{R}_2^{i*}(0)-\overline{R}_1^{i*}(0)\right)^+ .$$

Proportional fairness criterion:

At $\alpha = 1$, the objective functions based on $u^i_{1'}$ and u^i_2, respectively, are

$$\text{Max } \sum_{i\in I}\log\left(1-\overline{R}^i\right) \quad \text{and} \quad \text{Max } \sum_{i\in I}\log\left(\overline{R}^i_D-\overline{R}^i\right).$$

By the property that the logarithm of a product equals the sum of the logarithms of the factors, these objectives are equivalent to the following:

$$\text{Max } \prod_{i\in I}\left(1-\overline{R}^i\right) \quad \text{and} \quad \text{Max } \prod_{i\in I}\left(\overline{R}^i_D-\overline{R}^i\right).$$

The optimal solution of the two problems with these objective functions depends on the constants 1 and $\overline{R}^i_D$, $\forall i \in I$, respectively, due to the logarithm function. Consider the following two cases.

Case 1: If $\overline{R}^i_C \le \overline{R}^i_D$, $\forall i \in I$, risk reduction by resource centralization is possible for both hospitals.

Let $|I| = 2$ for simplicity. The objective functions based on $u^i_{1'}$ and u^i_2 are as follows.

$$\begin{aligned}\prod_{i\in I}\left(1-\overline{R}^i\right) &= \left(1-\overline{R}^1\right)\left(1-\overline{R}^2\right) = 1-\overline{R}^1-\overline{R}^2+\overline{R}^1\overline{R}^2\\ \prod_{i\in I}\left(\overline{R}^i_D-\overline{R}^i\right) &= \left(\overline{R}^1_D-\overline{R}^1\right)\left(\overline{R}^2_D-\overline{R}^2\right) = \overline{R}^1_D\overline{R}^2_D-\overline{R}^2_D\overline{R}^1-\overline{R}^1_D\overline{R}^2+\overline{R}^1\overline{R}^2\end{aligned}$$

In maximizing the first function, the values of $\overline{R}^i$ relative to 1 are of interest and the maximum product can be achieved by solutions in which only some hospital's average patient risk is not improved according to the decentralized strategy solution, i.e., $\exists i \in I : \overline{R}^{i*}_{1'}(1) = \overline{R}^i_D$. However, when the second function is maximized, the values of $\overline{R}^i$ relative to $\overline{R}^i_D$ are of interest and the optimal solution should imply a risk distribution such that all hospitals have improved average patient risks, i.e.,

$\overline{R}_2^{i*}(1) < \overline{R}_D^i, i \in I$. As long as an improvement is possible for both hospitals, a solution where $\exists i \in I : \overline{R}_2^{i*}(1) = \overline{R}_D^i$ cannot be optimal for the problem using utility function u_2^i under the proportional fairness criterion. Therefore, the problems using utility functions $u_{1'}^i$ and u_2^i are not equivalent, thus, their optimal risk distributions are not equal, due to the base values of $\overline{R}_D^i$ that are less than 1. This result holds regardless of the relationship between the $\overline{R}_D^i$ values of different hospitals, for example, if $\overline{R}_D^i = \overline{R}_D^k, \ \forall i, k \in I, i \neq k$.

Case 2: If $\exists i \in I : \overline{R}_D^i \leq \overline{R}_C^i$, risk reduction by resource centralization is not possible for at least one hospital.

If the decentralized strategy solution is the best possible solution for some hospital $i \in I$, that is, improvement in risks is not possible for this hospital by centralizing the resources, the problem under proportional fairness criterion cannot be solved since $\log\left(\overline{R}_D^i - \overline{R}^i\right) = \log(0)$ is not defined.

We have shown that the problems using $u_{1'}^i$ and u_2^i are not equivalent under the proportional fairness criterion and we will demonstrate that $u_{1'}^i$ does not necessarily dominate u_2^i in terms of the proximity of average patient risks for hospitals using Examples 1 and 2 below.

Max-min fairness criterion:

As $\alpha \to \infty$, the goal is maximizing the utility for the player with the minimum utility to achieve fairness. The objective functions using utility functions u_1^i and u_2^i are

$$\text{Max} \quad \min_{i \in I} \left\{1 - \overline{R}^i\right\} \quad \text{and} \quad \text{Max} \quad \min_{i \in I} \left\{\overline{R}_D^i - \overline{R}^i\right\}.$$

The optimal risk distribution for the first problem is such that all hospitals have the same average patient risk, i.e., $\overline{R}_1^{i*}(\infty) = \overline{R}_1^*, \ \forall i \in I$. In this case, risk is distributed equitably at the hospital-level by distributing the utilities u_1^i evenly among the evacuating hospitals.

The optimal solution for the second problem equalizes the deviation from the decentralized strategy risks, $\overline{R}_D^i - \overline{R}_2^i$, for each hospital. Therefore, if $\overline{R}_D^i = \overline{R}_D, \ \forall i \in I$, the utility, thus, the average patient risk, is distributed equally among the hospitals and

$\overline{R}_2^{i*}(\infty) = \overline{R}_2^{*}(\infty),\ \forall i \in I$. However, if the average decentralized patient risks for the hospitals are uneven, $\exists i, k \in I, i \neq k : \overline{R}_D^{i} \neq \overline{R}_D^{k}$, the resulting risk distribution is also uneven such that $\exists i, k \in I, i \neq k : \overline{R}_2^{i*}(\infty) \neq \overline{R}_2^{k*}(\infty)$.

When the max-min fairness criterion is adopted, the problem using utility function u_1^i is guaranteed to minimize the difference between the average patient risks for evacuating hospitals. The risk distribution may not be equal when the non-negative risk reduction requirement is added to the problem (as shown for Example 3 in Figure 4.16 below). On the other hand, the gap between average patient risks for evacuating hospitals can be greater when utility function u_2^i is used and it depends on the base values $\overline{R}_D^{i}$. This max-min fairness behavior of u_2^i does not provide a desirable distribution of risk among hospitals and u_1^i performs better than u_2^i according to the max-min fairness criterion.

The aforementioned aspects of the hospital-level utility functions will be demonstrated through examples after introducing the patient-level utility function.

4.3.2 Patient-level equity

The discussion of equity at the hospital-level indicates that a benchmark-based utility definition such as u_2^i is not as effective in ensuring equitable resource allocation decisions as a utility definition that is solely based on the risks according to the strategy in effect. Therefore, it is more appropriate to use a utility definition complementing the risk as in u_1^i at the patient-level as well. Then, we can define the utility of each individual patient as

$$u_1^{ijpvt} = \begin{cases} 1 - R_{pvt}^{ij} & \text{if } x_{pvt}^{ij} > 0,\ \forall i \in I, j \in J, p \in P, v \in V, t = 1, \cdots, T \\ \\ 0 & \text{otherwise} \end{cases} \tag{4.28}$$

Based on this definition, the objective functions to be maximized at the three specific values of the inequity aversion parameter α are as follows.

$$W_\alpha(u) = \begin{cases} \displaystyle\sum_{i \in I} \sum_{j \in J} \sum_{p \in P} \sum_{v \in V} \sum_{t=1}^{T} (1 - R_{pvt}^{ij}) x_{pvt}^{ij} & \text{for } \alpha = 0, \\ \displaystyle\sum_{i \in I} \sum_{j \in J} \sum_{p \in P} \sum_{v \in V} \sum_{t=1}^{T} \log\left(1 - R_{pvt}^{ij}\right) x_{pvt}^{ij} & \text{for } \alpha = 1, \\ \displaystyle\min_{i \in I, j \in J, p \in P, v \in V, t = 1, \cdots, T} \left\{1 - R_{pvt}^{ij} : x_{pvt}^{ij} > 0\right\} & \text{as } \alpha \to \infty \end{cases}$$

Note that when $\alpha = 0$, the objective is equivalent to the original objective function (4.4) of the n-HETM and the optimal solution is the system-level risk-minimizing solution. As $\alpha \rightarrow \infty$, the problem would be equivalent to minimizing the risk on the patient with the greatest evacuation risk. The objective function (4.29) and accompanying constraints (4.30)-(4.33) below can be used for this purpose, along with the original constraints (4.5) - (4.14), where K_{pvt}^{ij} are the indicator variables for the patient transportation events that are scheduled.

$$\text{Minimize} \quad R_{max} \tag{4.29}$$

$$\text{subject to} \quad x_{pvt}^{ij}/L^i \leq K_{pvt}^{ij}, \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T \tag{4.30}$$

$$R_{pvt}^{ij} K_{pvt}^{ij} \leq R_{max}, \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T \tag{4.31}$$

$$K_{pvt}^{ij} \in \{0, 1\}, \quad \forall i \in I,\ j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T \tag{4.32}$$

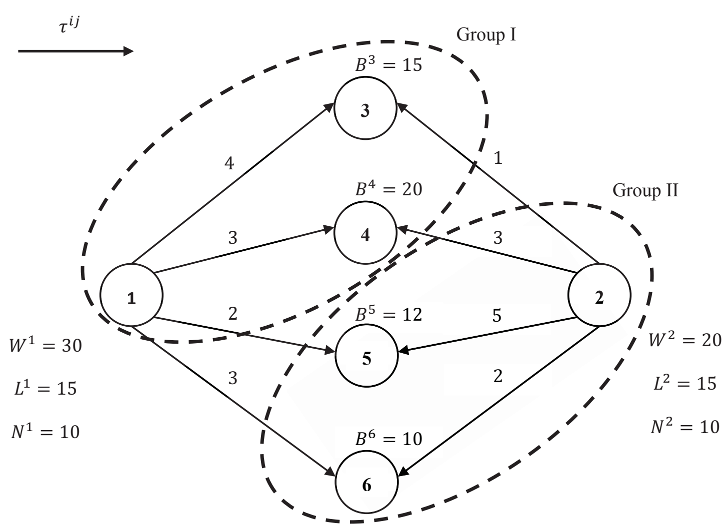
$$R_{max} \geq 0. \tag{4.33}$$

Based on the equity model described above at the hospital- and patient-levels, the equitable solution desired must (1) minimize the system-level risk, (2) minimize the range of player utilities, and (3) result in positive risk reduction compared to the decentralized resource management solution. Although optimal solutions maximizing the welfare function would not necessarily satisfy all three criteria, these desired properties guide the decision maker in choosing the most suitable equity criteria to employ in evacuation planning for a given scenario. We next illustrate the properties of the utility functions discussed above using examples solved for hospital-level equity and patient-level equity. The problems maximizing the non-linear welfare function are programmed in AMPL format and solved using the MINOS 5.5 solver that employs a reduced-gradient method to solve the linearly constrained problems with a non-linear objective function.

Example 1: ($\overline{R}_C^i \leq \overline{R}_D^i$, $\forall i \in I$) Consider an evacuation network of two hospital groups each having one evacuating hospital. Group I consists of hospitals labeled 1, 3, and 4 while Group II consists of hospitals labeled 2, 5, and 6. There are 30 patients of a single type at hospital 1 and 20 patients of the same type at hospital 2 to be evacuated, i.e., $W^1 = 30, W^2 = 20$. The travel times between the hospitals and the number of available beds at the receiving hospitals are given in Figure 4.10.

Let the only vehicle type available for transportation be ALS ambulances with a capacity of one patient at a time, a physical loading requirement of one unit, and loading time of one time interval, i.e., $C_1 = 1, L_1 = 1$, and $\gamma_1 = 1$. Each evacuating hospital has a capacity of loading 15 ambulances at a time, i.e., $L^1 = L^2 = 15$. There are 20 ambulances available in total and the fleet is equally split between the two hospital groups when the resources are decentralized, i.e., $N^1 = N^2 = 10$. For simplicity, assume that all vehicles are available at the beginning of the planning horizon and no additional vehicles become available afterwards, i.e., $N_{vt} = 0, t = 2, \cdots, T$. The threat risk parameter, λ^i_{pt}, is defined to be constant through time as $\lambda = 0.08$ for both evacuating hospitals. The transportation risk, θ_{pv}, is defined as $\theta = 0.04$.

Figure 4.10: (Example 1) An asymmetric network with two evacuating hospitals where both hospitals benefit from resource centralization.



The risk measure values evaluated at the optimal hospital- and patient-level equity solutions are shown in Table 4.5. The two hospital-level utility functions and the patient-level utility function were used in the welfare function maximization at various α values and the results for the three α values of interest are listed. To illustrate the behavior of the hospital-level utility functions as the fairness criterion changes, $\overline{R}^{i*}_f(\alpha)$ values at various α values are plotted in Figure 4.11 along with the decentralized strategy solution as a benchmark. In this example, both hospitals can benefit from centralization of resources, i.e., $\overline{R}^i_C \leq \overline{R}^i_D, i = 1, 2$. Therefore, u^i_1 is equivalent to $u^i_{1'}$. Both hospitals have lower average risks at all α levels compared to the decentralized strategy. As we have explained above, the optimal solutions of u^i_1 (which, in this example, is equal to $u^i_{1'}$) and u^i_2 are the same at $\alpha = 0$.

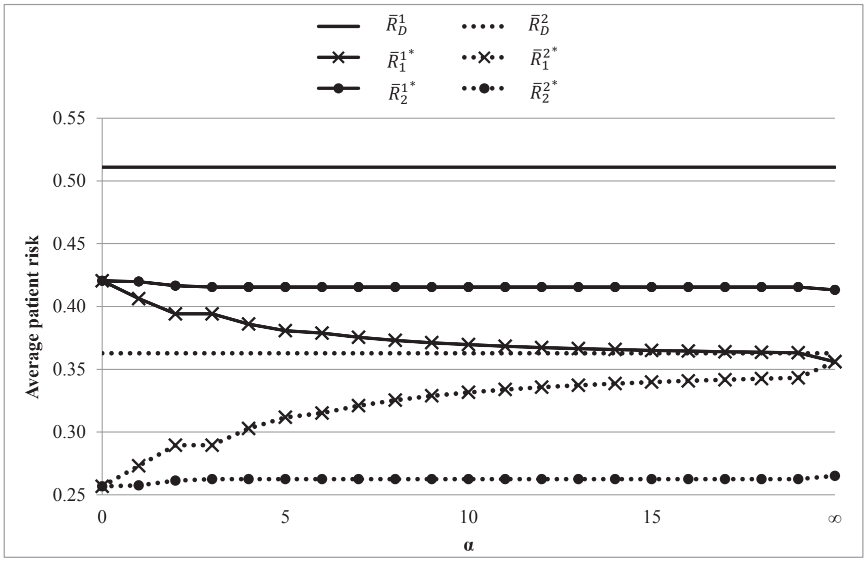
Table 4.5: (Example 1) Risk measures for optimal hospital- and patient-level equity solutions.

			Hospital-level Equity						Patient-level Equity		
			$u^i_1 \equiv u^i_{1'}$			u^i_2			$u^{ijpvt}_1 \equiv u^{ijpvt}_{1'}$		
	Decentr.	Centr.	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}^{1*}_f(\alpha)$	0.5109	0.3941	0.4204	0.4063	0.3560	0.4204	0.4198	0.4132	0.3941	0.3531	0.3910
$\overline{R}^{2*}_f(\alpha)$	0.3628	0.2895	0.2567	0.2731	0.3560	0.2567	0.2575	0.2651	0.2895	0.3612	0.3659
$\overline{R}^*_{fh}(\alpha)$	0.4369	0.3418	0.3386	0.3397	0.3560	0.3386	0.3387	0.3392	0.3418	0.3571	0.3784
$Range^*_{fh}(\alpha)$	0.1481	0.1046	0.1637	0.1332	0	0.1637	0.1623	0.1481	0.1046	0.0081	0.0251
$R^*_{f,min}(\alpha)$	0.2200	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875
$R^*_{f,max}(\alpha)$	0.7357	0.6464	0.5823	0.6157	0.6464	0.5823	0.5823	0.6157	0.6464	0.5459	0.5459
$\overline{R}^*_f(\alpha)$	0.4517	0.3523	0.3550	0.3530	0.3560	0.3550	0.3550	0.3540	0.3523	0.3563	0.3809
$Range^*_f(\alpha)$	0.5157	0.4590	0.3948	0.4282	0.4590	0.3948	0.3948	0.4282	0.4590	0.3585	0.3585

At $\alpha = 1$, $u^i_{1'}$ results in a smaller range of average patient risks for the two hospitals than u^i_2, thus, provides a more equitable solution at the hospital-level. By definition of u^i_2, the product of risk reductions for hospitals is maximized when $\overline{R}^{1*}(1)$ is increased and $\overline{R}^{2*}(1)$ is decreased compared to the $u^i_{1'}$ solution. This creates a wider gap between the average patient risks of the two hospitals, which is an undesired result in terms of equitable risk distribution among hospitals. In this case, $u^i_{1'}$ performs better than u^i_2 with a more equitable risk distribution among hospitals and a lower system-level average risk $\overline{R}^*$. However, this is not always the case and a different outcome can be observed as we demonstrate in Example 2.

As shown in Figure 4.11, when u^i_1 is used in optimization, as α is increased, there is a

Figure 4.11: Optimal average patient risks for hospitals in Example 1 according to decentralized strategy, u_1^i, and u_2^i.



utility transfer from hospital 2 to hospital 1, or a risk transfer from hospital 1 to hospital 2, that eventually results in the average patient risks for two hospitals to converge, where $Range_{1h}^*(\infty) = 0$. As the risk is distributed more equitably among hospitals, the average risk $\overline{R}_{1h}^*$ is increased and this increase can be interpreted as the *cost of equity*. The optimal objective function values using utility function u_1^i, $W_\alpha(u_1^i)$, are shown in Figure 4.12 along with the diminishing gap between the average patient risks for hospitals. However, when utility function u_2^i is used, the average patient risks for hospitals have a greater disparity as α approaches infinity and $\overline{R}_{2h}^*$ is decreased. As shown in Figure 4.13, the gap between average patient risks for hospitals can only be reduced to some extent using u_2^i. This result for the risk reduction-based utility function is not desirable as it does not decrease inequity among the hospitals when the objective is maximizing the minimum utility. Therefore, u_1^i is a more appropriate utility function than u_2^i for our risk-based hospital evacuation setting.

In addition to hospital-level equity, the decision maker can also be concerned about the patient-level equity. An interesting result of aiming for hospital-level equity is that the range of individual patient risks, $Range^*(\alpha)$, increases with α for both u_1^i, and u_2^i. This shows that while the decision maker is aiming for an equitable risk distribution among hospitals, the

Figure 4.12: (Example 1) Optimal welfare function values and gap between average patient risks for hospitals using u_1^i.

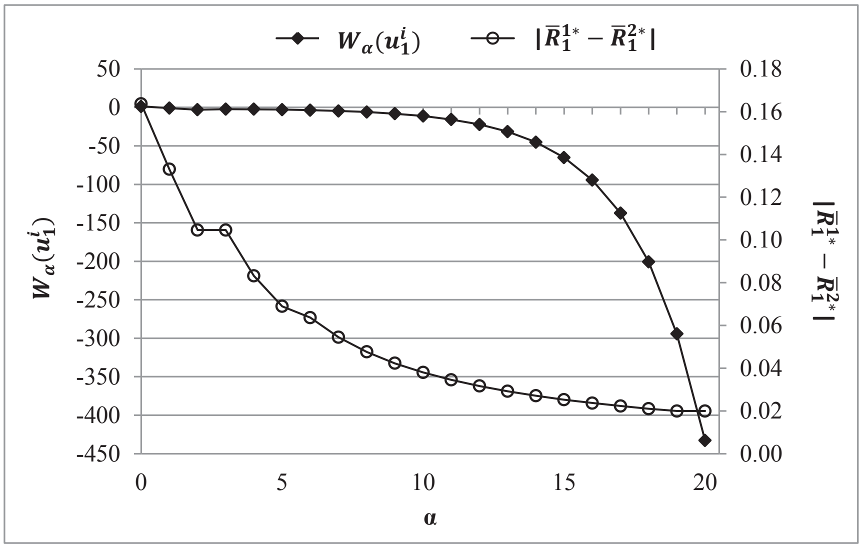
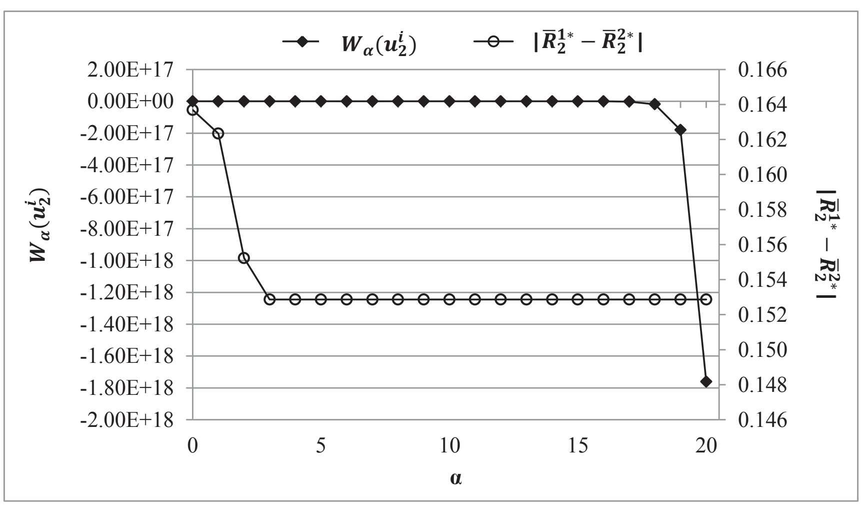


Figure 4.13: (Example 1) Optimal welfare function values and gap between average patient risks for hospitals using u_2^i.



patient-level equity is jeopardized. When the goal is achieving patient-level equity, the range of individual patient risks is decreased as α is increased at a cost of increased system-level risk. The cost of equity is the increase in system-level risk and it is higher for equity among individual patients than among hospitals. We can also observe that the optimal patient-level equity solutions can only have lower $\overline{R}_h^*$ values than the optimal hospital-level equity solutions. However, the corresponding range of average patient risks for hospitals, $Range_{fh}^*$, is reduced when patient-level equity is sought instead of hospital-level equity except for u_1^i as $\alpha \rightarrow \infty$.

Based on these optimal solutions, it is observed that there is no single utility function that dominates the others in terms of the system-level risk, the range of average patient risks for hospitals, and the range of patient risks at all α values. However, these results can guide the decision maker depending on the priorities in evacuation planning. For example, if the importance of equity among hospitals overrides that among patients, utility function u_1^i should be preferred as $\alpha \rightarrow \infty$. However, if a more moderate risk distribution that also provides the lowest range of risks at the patient-level in addition to the lowest range of risks at the hospital-level, patient-level utility function u_1^{ijpvt} should be preferred at $\alpha = 1$ at a cost in terms of system-level risk.

Example 2: $(\exists i \in I : \overline{R}_C^i > \overline{R}_D^i)$ Changing the travel times of the first example as shown in Figure 4.14, we can obtain a network in which one of the evacuating hospitals can be exposed to higher risk as a result of centralization of resources whereas the other hospital benefits from this, i.e., $\overline{R}_C^1 \geq \overline{R}_D^1, \overline{R}_C^2 \leq \overline{R}_D^2$. This can be observed when the utility function u_1^i is used in optimization where $\alpha = 0$ as shown in Table 4.6. Since non-negative risk reduction is not forced as in u_2^i, utility is transferred from hospital 1 to hospital 2, or risk is transferred from hospital 2 to hospital 1, resulting in a higher average patient risk for hospital 1 than its average patient risk under decentralized strategy.

The optimal average patient risks, $\overline{R}^{i*}(\alpha)$, for the utility functions $u_{1'}^i$ and u_2^i at various α values are shown in Figure 4.15 along with the decentralized strategy solution as a benchmark. When non-negative risk reduction is required, unlike the solutions of Example 1 discussed above, the proportional fairness solution for $u_{1'}^i$ creates a wider gap than the u_2^i solution in this example. Therefore, either one of the utility functions $u_{1'}^i$ and u_2^i can be dominant in terms of the equitable average patient risk distribution depending on the problem instance.

Figure 4.14: (Example 2) An asymmetric network with two evacuating hospitals where one hospital can be harmed by resource centralization.

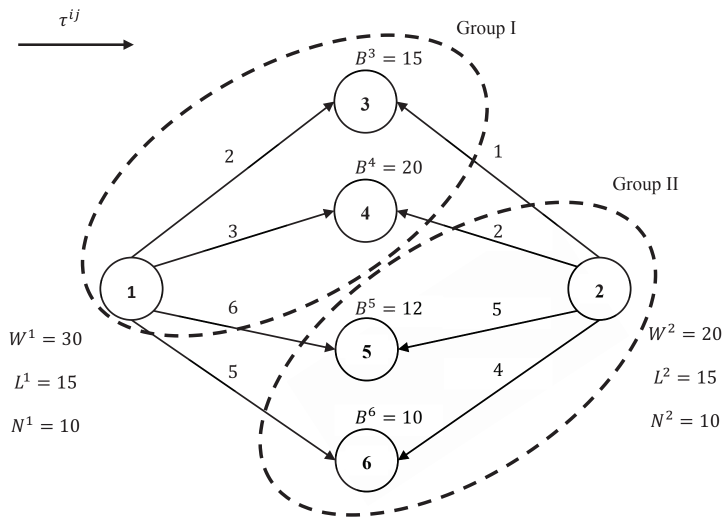


Table 4.6: (Example 2) Risk measures for optimal hospital-level equity solutions.

			Hospital-level Equity								
			u_1^i			$u_{1'}^i$			u_2^i		
	Decentr.	Centr.	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}_f^{1*}(\alpha)$	0.4392	0.4603	0.4948	0.4603	0.3893	0.4392	0.4392	0.3893	0.4392	0.3928	0.3817
$\overline{R}_f^{2*}(\alpha)$	0.4635	0.2636	0.2255	0.2636	0.3893	0.3007	0.3007	0.3893	0.3007	0.3821	0.4060
$\overline{R}_{fh}^{*}(\alpha)$	0.4513	0.3619	0.3602	0.3619	0.3893	0.3699	0.3699	0.3893	0.3699	0.3874	0.3939
$Range_{fh}^*(\alpha)$	0.0243	0.1967	0.2692	0.1967	0	0.1385	0.1385	0	0.1385	0.0107	0.0243
$R_{f,min}^*(\alpha)$	0.2200	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875	0.1875
$R_{f,max}^*(\alpha)$	0.6747	0.6304	0.6872	0.6304	0.6458	0.6304	0.6304	0.6458	0.6304	0.6150	0.6458
$\overline{R}_f^{*}(\alpha)$	0.4489	0.3816	0.3871	0.3816	0.3893	0.3838	0.3838	0.3893	0.3838	0.3885	0.3914
$Range_f^*(\alpha)$	0.4548	0.4430	0.4997	0.4430	0.4584	0.4430	0.4430	0.4584	0.4430	0.4276	0.4584

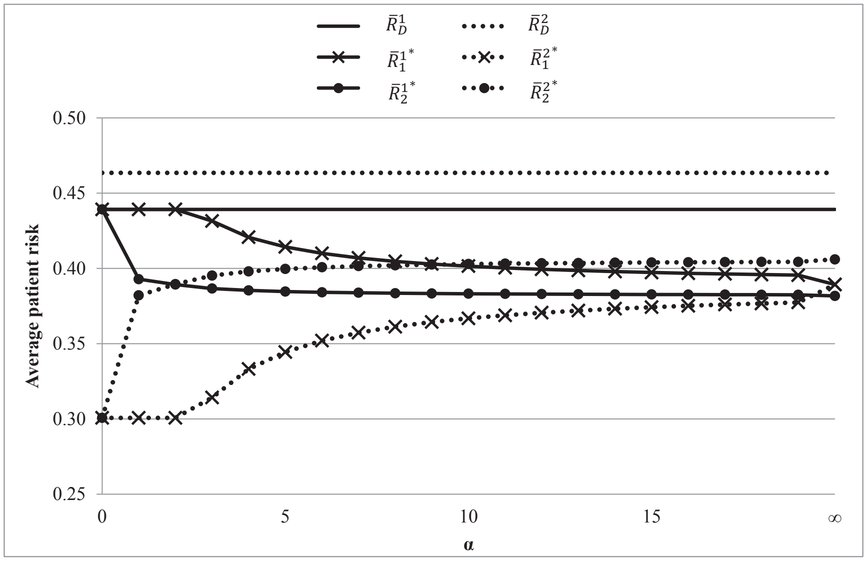
As α is increased when $u_{1'}^i$ is used, the risk distribution becomes more equitable and the average patient risks of two hospitals converge as α approaches infinity. However, when utility function u_2^i is used, the max-min fairness criterion does not necessarily provide an equal risk distribution; it rather ensures that the amounts of risk reduction for the hospitals are equal. The average patient risks of two hospitals diverge as $\alpha \to \infty$, which is, again, an undesired outcome of using u_2^i.

Table 4.7: (Example 2) Risk measures for optimal patient-level equity solutions.

	Patient-level Equity					
	u_1^{ijpvt}			$u_{1'}^{ijpvt}$		
	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}_f^{1*}(\alpha)$	0.4603	0.4670	0.3609	0.4392	0.4392	0.3878
$\overline{R}_f^{2*}(\alpha)$	0.2636	0.2567	0.4766	0.3007	0.3007	0.4184
$\overline{R}_{fh}^{*}(\alpha)$	0.3619	0.3619	0.4187	0.3699	0.3699	0.4031
$Range_{fh}^{*}(\alpha)$	0.1967	0.2103	0.1157	0.1385	0.1385	0.0306
$R_{f,min}^{*}(\alpha)$	0.1875	0.1875	0.2200	0.1875	0.1875	0.1875
$R_{f,max}^{*}(\alpha)$	0.6304	0.6304	0.5815	0.6304	0.6304	0.5815
$\overline{R}_f^{*}(\alpha)$	0.3816	0.3829	0.4071	0.3838	0.3838	0.4000
$Range_f^{*}(\alpha)$	0.4430	0.4430	0.3616	0.4430	0.4430	0.3941

When the patient-level risk measures are reviewed at various α values, it is observed that, the range of individual patient risks, $Range^*(\alpha)$, is non-increasing as α is increased for $u_{1'}^i$ as in Example 1. Maximizing the welfare function as the aversion to hospital-level inequity increases does not improve the patient-level equity, while the risk distribution among hospitals become more equitable. As observed based on Example 1 results, maximizing the welfare function for hospital-level equity results in increasing $Range^*$ whereas patient-level equity solutions shown in Table 4.7 result in decreasing $Range^*$ as α is increased. However, since the equitable utility distribution among patients results in higher average risk for the system, the cost of ensuring patient-level equity is higher than the cost of ensuring hospital-level equity. Also, ensuring equity among hospitals provides $\overline{R}_h^*$ values less than or equal to the values achieved when ensuring equity among individual patients.

Figure 4.15: Optimal average patient risks for hospitals in Example 2 according to decentralized strategy, $u_{1'}^i$, and u_2^i.



Example 3: $(\exists i \in I : \overline{R}_C^i \geq \overline{R}_D^i)$ Consider the evacuation network in Figure 4.16 where there are three hospital groups each with one evacuating and one receiving hospital. There are 10, 6, and 4 patients of the same type at evacuating hospitals 1, 2, and 3, respectively. The loading capacity of each hospital is 5 ambulances at a time interval and there are 15 ambulances in total distributed equally among three evacuating hospitals under the decentralized resource management strategy. The threat risk is $\lambda = 0.08$ and the transportation risk is $\theta = 0.04$ as in the previous two examples.

The average patient risk distribution of the optimal solutions using different utility functions and the three fairness criteria are summarized in Table 4.8. In this example, the minimum possible average patient risk for hospital 3 is $\overline{R}_D^3$ and this hospital's risk cannot be improved by resource centralization, it can only get worse as in the risk-minimizing centralized strategy solution and the three fairness criteria solutions using u_1^i. This is also the reason why the problem adopting the proportional fairness criteria using utility function u_2^i is infeasible. Since for any solution under the centralized strategy $\overline{R}^3 \geq \overline{R}_D^3$ and this criterion forces $\overline{R}^{3*}$ to equal $\overline{R}_D^3$, the undefined $\log\left(\overline{R}_D^3 - \overline{R}^{3*}\right) = \log(0)$ makes the problem infeasible.

Figure 4.16: (Example 3) An evacuation network with three evacuating hospitals.

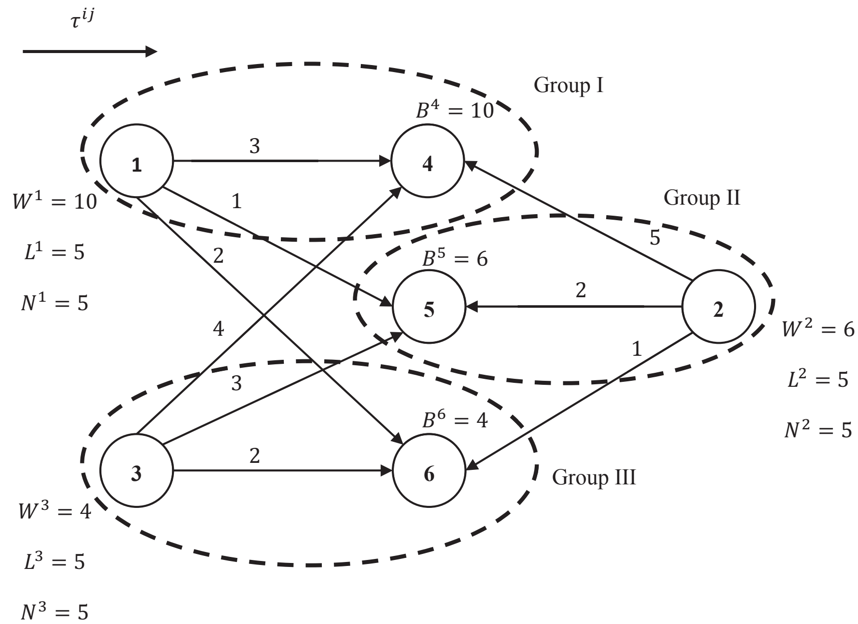
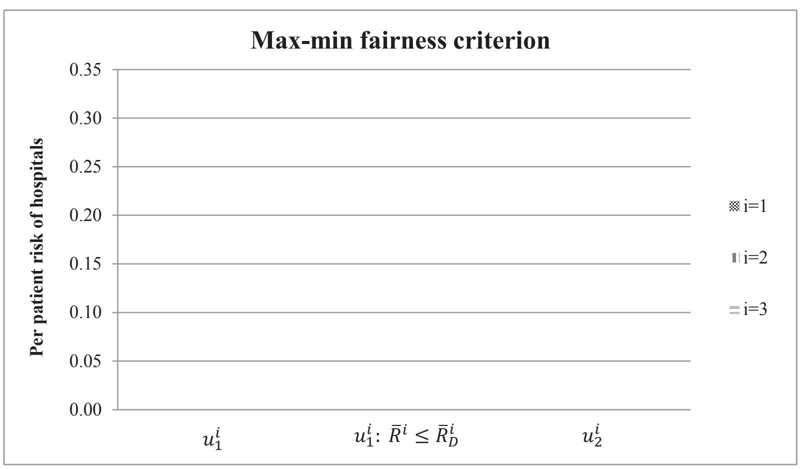


Table 4.8: (Example 3) Risk measures for optimal hospital-level equity solutions.

	Decentr.	Centr.	u_1^i			$u_{1'}^i$			u_2^i		
			$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}_f^{1*}(\alpha)$	0.3985	0.2777	0.2777	0.2777	0.2657	0.3340	0.3340	0.3129	0.3340	infeasible	0.3366
$\overline{R}_f^{2*}(\alpha)$	0.2568	0.2183	0.2183	0.2183	0.2657	0.2304	0.2304	0.2568	0.2304	-	0.2467
$\overline{R}_f^{3*}(\alpha)$	0.2200	0.2811	0.2811	0.2811	0.2657	0.2200	0.2200	0.2200	0.2200	-	0.2200
$\overline{R}_{fh}^{*}(\alpha)$	0.2918	0.2590	0.2590	0.2590	0.2657	0.2614	0.2614	0.2632	0.2614	-	0.2677
$Range_{fh}^{*}(\alpha)$	0.1786	0.0629	0.0629	0.0629	0	0.1141	0.1141	0.0929	0.1141	-	0.1166
$R_{f,min}^{*}(\alpha)$	0.2200	0.1875	0.1875	0.1875	0.1875	0.2200	0.2200	0.1875	0.2200	-	0.1875
$R_{f,max}^{*}(\alpha)$	0.5459	0.3662	0.3662	0.3662	0.4159	0.4169	0.4169	0.4169	0.4169	-	0.4169
$\overline{R}_{f}^{*}(\alpha)$	0.3203	0.2605	0.2605	0.2605	0.2657	0.2801	0.2801	0.2775	0.2801	-	0.2863
$Range_{f}^{*}(\alpha)$	0.3260	0.1787	0.1787	0.1787	0.2284	0.1969	0.1969	0.2294	0.1969	-	0.2294

In this example, the solutions for u_1^i according to the utilitarian ($\alpha = 0$) and proportional fairness ($\alpha = 1$) criteria are the same. This is due to the fact that any transfer of utility among the hospitals in this network results in a negative aggregate proportional change in hospital-level utilities. So, it is not worth redistributing the risk among hospitals in terms of the welfare function value. The utilitarian and proportional fairness solutions are the same for $u_{1'}^i$, but the non-negative risk reduction constraint changes the u_1^i solution such that $\overline{R}^{3^*} = \overline{R}_D^3$ and, in exchange for this risk reduction, $\overline{R}^{1^*}$ and $\overline{R}^{2^*}$ are increased. This risk distribution results in a higher range of hospital risks, $Range_h^*$ and higher system-level average risk, $\overline{R}^*$. In fact, at each α value, optimal u_1^i solution has the lowest system risk, the smallest range of hospital risks, and the smallest range of patient risks. Therefore, in this example, u_1^i dominates $u_{1'}^i$ and u_2^i in terms of these three equity requirements.

Figure 4.17: Average patient risk distribution among evacuating hospitals using different utility functions for Example 3.



The risk distribution among hospitals for the max-min fairness solutions using u_1^i, $u_{1'}^i$, and u_2^i are compared in Figure 4.17. An even distribution of the average patient risk among the three hospitals is observed for u_1^i. The additional constraint $\overline{R}^i \leq \overline{R}_D^i$ results in a different max-min fairness solution for utility function $u_{1'}^i$ in which hospitals 2 and 3 are forced to remain at their risk levels under the decentralized strategy. The risk distribution can become even more unequitable when u_2^i is used. Note that there are multiple optima for the problems using $u^i_{1'}$ and u^i_2 adopting max-min fairness criterion and the solutions provided in Table 4.8 are only one of the optima for these two problems. Various feasible combinations of $\overline{R}^1 \leq \overline{R}^1_D$ and $\overline{R}^2 \leq \overline{R}^2_D$ result in the same welfare function value as the optimal objective function value is $\min_{i\in I}\left\{\overline{R}^i_D - \overline{R}^i\right\} = \overline{R}^3_D - \overline{R}^{3*} = 0$ in both cases.

Table 4.9: (Example 3) Risk measures for optimal patient-level equity solutions.

	Patient-level Equity					
	u_1^{ijpvt}			$u_{1'}^{ijpvt}$		
	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}^{1*}_f(\alpha)$	0.2777	0.7890	0.2624	0.3226	0.3985	0.3277
$\overline{R}^{2*}_f(\alpha)$	0.2183	0.8051	0.2699	0.2441	0.2568	0.2526
$\overline{R}^{3*}_f(\alpha)$	0.2811	0.4797	0.2811	0.2200	0.2200	0.2200
$\overline{R}^{*}_{fh}(\alpha)$	0.2590	0.6913	0.2711	0.2622	0.2918	0.2667
$Range^*_{fh}(\alpha)$	0.0629	0.3254	0.0188	0.1026	0.1786	0.1077
$R^*_{f,min}(\alpha)$	0.1875	0.2811	0.1875	0.1875	0.1875	0.1875
$R^*_{f,max}(\alpha)$	0.3662	0.8182	0.3123	0.4169	0.7860	0.4169
$\overline{R}^{*}_{f}(\alpha)$	0.2605	0.7320	0.2684	0.2785	0.3203	0.2836
$Range^*_{f}(\alpha)$	0.1787	0.5371	0.1248	0.2294	0.5985	0.2294

The risk measures obtained when patient-level equity is sought are shown in Table 4.9. The average patient risk distribution among the three hospitals is more equitable and the system-level risk is lower when non-negative risk reduction is not required, i.e., u_1^{ijpvt} is employed, except for the proportional fairness criterion where the average patient risks are set to their values according to the decentralized strategy. The range of individual patient risks is also tighter without the non-negative risk reduction constraint. Comparing the patient-level equity results to the hospital-level equity results in Table 4.8, the lowest range of average patient risks for hospitals and the lowest range of individual patient risks are obtained for u_1^{ijpvt} as $\alpha \to \infty$. Moreover, the system risk is only slightly higher than the minimum system risk. Therefore, from both the hospitals' and the patients' perspectives, the solution for u_1^{ijpvt} satisfying the max-min fairness criterion would be preferred as the most equitable solution.

The above examples illustrate the performance of the utility functions in ensuring equity among evacuating hospitals and among patients. We next present a case study based on real data and provide an analysis of the optimal evacuation plans generated under various equity criteria.

4.4 Case Study

We introduce a case study based on real data in this section to demonstrate and analyze the performance of the n-HETM. To compare the resource management strategies discussed above, we use groups of hospitals that use a fleet of vehicles and a set of receiving beds isolated from other groups in case the evacuation operations are carried out with decentralized resources. A network that consists of three hospital groups, two of which each has one evacuating hospital, is considered in this study. All hospitals in this study are assumed to be members of an emergency preparedness alliance for mutual emergency management planning and support among hospitals in case of local disasters. Group I consists of a 765-bed evacuating hospital (H-I) and five receiving facilities, and Group II consists of a 146-bed evacuating hospital (H-II) and three receiving facilities. Group III has only one receiving hospital serving as a surge hospital that supplies extra receiving beds to be utilized when the first two groups use decentralized resources.

The patient population is categorized into three types, each type requiring a different level or type of care. Type 1 patients are the most critical ones, such as adult or pediatric critical care patients requiring intensive care, and type 3 patients are the least critical patients, such as outpatients or psychiatric patients. The instantaneous data used in this case study, which includes 360 patients in H-I, 90 patients in H-II, and 495 beds available at the 8 receiving hospitals by patient type, is displayed in Table 4.10. The length of a time interval is assumed to be 10 minutes and the one-way travel times between the evacuating hospitals and the potential receiving hospitals in number of 10-minute time intervals are provided in Table 4.10.

Table 4.10: The number of patients (W_p^i), the number of beds available (B_p^j), and the travel time between hospitals (τ^{ij}).

	Number of Beds B_p^j			Travel Time τ^{ij} (intervals)	
Hospital	1	2	3	H-I	H-II
Group I					
H-I (W_p^I)	135	160	65		
1	30	45	15	4	8
2	0	15	0	4	8
3	0	30	10	8	3
4	55	25	30	6	11
5	35	50	0	7	11
Group II					
H-II (W_p^{II})	45	25	20		
6	10	0	15	8	11
7	35	20	5	2	4
8	20	15	5	7	3
Group III					
9	20	0	10	15	15
TOTAL (beds)	205	200	90		

It is assumed that the vehicles can be immediately dispatched and sent to evacuating hospitals once the evacuation is initiated. In case the vehicles are dispatched from further locations such as depots, the time it takes for vehicles to reach evacuating hospitals can be easily changed by adjusting the N_{vt}-parameters. The vehicle types used in this study are ALS and BLS ambulances both capable of carrying one patient at a time. The time required to load a patient into an ALS or BLS ambulance is assumed to be one time interval that is also equal to the time to unload a patient from an ambulance. There are 35 ALS ambulances and 35 BLS ambulances in total that become available at the beginning of the planning horizon. When the resources are decentralized, 25 ALS and 25 BLS ambulances are assigned to H-I and 10 ALS and 10 BLS ambulances are assigned to H-II. The evacuating hospitals have a loading capacity of 10 ambulances at a time interval. These parameters can be better estimated by using observations of mock evacuations or training activities.

The threat risk parameter, λ_{pt}^{i}, is indexed on time, patient type, and evacuating hospital and depends on the nature of the threat. During a disaster, hospitals can be affected in different ways by the threat depending on their location or building conditions. For example, an earthquake might cause a prolonged loss of power at a hospital which would severely reduce the ability of the hospital to care for patients, but more so for critical patients who are dependent on lifesaving equipment than less critical patients. At the same time, another hospital affected by the earthquake might be dealing with fire that drastically increases the threat risk for all patients in that everyone in the hospital would be threatened by suffocation. For accurate comparison of resource management strategies, the same threat and transportation risk parameters are assumed for both hospital groups in this case study. Table 4.11 gives the values of the threat risk parameters, λ_{pt}^{i}, for each patient type and the transportation risk parameters, θ_{pv}, for each patient type and transportation type combination used in this study.

Table 4.11: Threat risk functions (λ_{pt}^{i}) and transportation risk parameters (θ_{pv}).

Patient Type	Threat Risk	Transportation Risk	
		ALS	BLS
1	$0.0012\exp(t/35)$	0.001	0.002
2	$0.0009\exp(t/35)$	0.0005	0.001
3	$0.0006\exp(t/35)$	0.0001	0.0002

The cumulative threat risk function, Λ_{pt}^{i}, used in this case study having an exponential risk parameter, λ_{pt}^{i}, has an increasing slope. For example, we consider the exponential threat risk with the given parameters, to represent the case of a utility loss (electricity or water) where the threat risk on patients with utility-dependent treatment increase drastically in a short amount of time. There are many other risk functions available, including more complicated combined functions and there is no reason that each risk group should have the same functional form for a certain scenario. As the n-HETM requires the risks as input parameters, any proper risk form is acceptable. These risk functions, which vary for different patient types, imply the relative criticality of patient types in a particular type of emergency. In this study, we are assuming that the threat that causes the evacuation affects all patient types, but the resistance to the effects of this threat varies among patient types. For example, pediatric patients are more resistant to the effects of the threat in question than the pediatric critical care patients.

The optimal hospital-level equity solutions for the case study are presented in Table 4.12. The problem under centralized strategy is a large-scale problem with 20,744 decision variables and 6,216 constraints. Evacuating hospital H-I benefits from resource centralization at the expense of increased risk for H-II, because patients from H-I are transported to closer receiving hospitals in Group II which forces some patients in H-II to travel further. When the

Table 4.12: Risk measures for optimal hospital-level equity solutions for the case study.

			u_1^i			u_2^i		
	Decentr.	Centr.	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}^{1*}$	0.1008	0.0585	0.0682	0.0682	0.0557	0.0682	0.0703	0.0848
$\overline{R}^{2*}$	0.0255	0.0435	0.0123	0.0123	0.0557	0.0123	0.0109	0.0095
$\overline{R}_h^*$	0.0632	0.0510	0.0402	0.0402	0.0557	0.0402	0.0406	0.0471
$Range_h^*$	0.0753	0.0150	0.0559	0.0559	0	0.0559	0.0594	0.0753
R_{min}^*	0.0072	0.0052	0.0053	0.0053	0.0052	0.0053	0.0043	0.0050
R_{max}^*	0.3510	0.1713	0.1439	0.1439	0.1820	0.1439	0.1484	0.2336
$\overline{R}^*$	0.0857	0.0555	0.0551	0.0570	0.0538	0.0550	0.0584	0.0680
$Range^*$	0.3438	0.1660	0.1385	0.1385	0.1768	0.1385	0.1441	0.2286
Solution time (s)	1.15	7.54	6.93	16.00	6.40	7.00	14.00	7.18

hospital-level equity is concerned, the desired solutions are the ones with the smallest range of the average patient risks for evacuating hospitals, $Range_h^*$. The range-minimizing behavior of utility function u_1^i as the decision maker becomes more inequity-averse is confirmed by the results for the case study. The undesired behavior of u_2^i that results in wider ranges for hospitals even though the decision maker becomes more inequity-averse is also observed. Therefore, the utility function defined as the complementary value to average patient risk for each evacuating hospital would be preferred to generate evacuation plans that distribute risk equitably among the evacuating hospitals.

The optimal solutions that yield hospital-level equity result in increased maximum risk among patients while the minimum patient risk does not deviate much as α is increased. Therefore, the range of individual patient risks increases with the inequity aversion parameter for both utility functions. While u_1^i and u_2^i are employed at increasing values of α, the average risk of the system, $\overline{R}^*$, is also increased. Although u_1^i ensures improved equity among hospitals, the cost of hospital-level equity is observed in terms of the wider gap among individual patient risks and the increased system-level risk. Using the benchmark-based utility function improves neither the hospital-level equity nor the patient-level equity in terms of the risk distribution, because the welfare function based on u_2^i results in an equitable distribution of the gain from resource centralization, that is the risk reduction, among hospitals.

The average distances traveled by patients of each type at the two evacuating hospitals based on the hospital-level equity solutions are shown in Table 4.13. The positive impact of resource centralization on the evacuation plan is reflected on the distance traveled by patients of all three types at H-I and H-II except for patients of type 1 at H-II who have to travel further as a result of sharing the receiving beds in Group II with patients of Group I. Note that both u_1^i and u_2^i result in the same distances under the utilitarian and proportional fairness criteria even though the optimal solutions are different. For patients of type 1 and 3 at H-II, the utilitarian and proportional fairness solutions for both u_1^i and u_2^i require even shorter distances to travel than the centralized strategy. It is observed that, for u_1^i under the max-min fairness criteria, the average distances traveled by patients of each type are balanced at H-I by reducing the distances for patient types 1 and 3, whereas they are increased for

Table 4.13: Average distance traveled (in number of time intervals) by patients of type p at each evacuating hospital based on the optimal hospital-level equity solutions.

Evacuating Hospital	p	Decentr.	Centr.	u_1^i $\alpha = 0, 1$	u_1^i $\alpha \to \infty$	u_2^i $\alpha = 0, 1$	u_2^i $\alpha \to \infty$
H-I	1	6.81	5.00	5.59	4.92	5.59	5.99
	2	5.88	5.09	5.09	5.10	5.09	5.22
	3	7.23	5.69	6.69	5.84	6.69	7.48
H-II	1	3.56	5.89	3.56	6.25	3.56	3.56
	2	3.40	3.00	3.00	3.00	3.00	3.00
	3	7.25	6.00	3.25	5.88	3.25	3.25

patients of type 1 and 3 at H-II. On the other hand, the solution for u_2^i under the max-min fairness criteria increases the average distance for patients of all types at H-I while keeping the distances same for patients at H-II. This is the impact of the model's effort to distribute the risk reduction based on the risks in decentralized strategy equitably between H-I and H-II.

The patient-level equity solutions for the case study are summarized in Table 4.14. The range of individual patient risks is minimized as the decision maker becomes more risk-averse while the range of average patient risks for hospitals is increased. When all the hospital- and patient-level equity solutions are considered, the utility function that provides the best combination of the system-level risk, the range of average patient risks for hospitals, and the range of individual patient risks is u_1^{ijpvt} as $\alpha \to \infty$. However, this solution causes a much higher risk for evacuating hospital H-II as the cost of minimizing the range of individual patient risks.

In order to demonstrate the impact of resource levels on the equity of the optimal solutions, we change the loading capacity of each evacuating hospital keeping that of the other hospital fixed at 10 ambulances at a time. The average patient risk for the evacuating hos-

Table 4.14: Risk measures for optimal patient-level equity solutions for the case study.

	u_1^{ijpvt}		
	$\alpha = 0$	$\alpha = 1$	$\alpha \to \infty$
$\overline{R}^{1*}$	0.0585	0.0579	0.0785
$\overline{R}^{2*}$	0.0435	0.0371	0.1650
$\overline{R}_h^*$	0.0510	0.0475	0.1217
$Range_h^*$	0.0150	0.0208	0.0865
R_{min}^*	0.0052	0.0019	0.0039
R_{max}^*	0.1713	0.1713	0.0847
$\overline{R}^*$	0.0537	0.0538	0.0618
$Range^*$	0.1660	0.1693	0.0807
Solution time (s)	7.54	12.00	180.00

Figure 4.18: The impact of loading capacity perturbations on the average patient risk for the system.

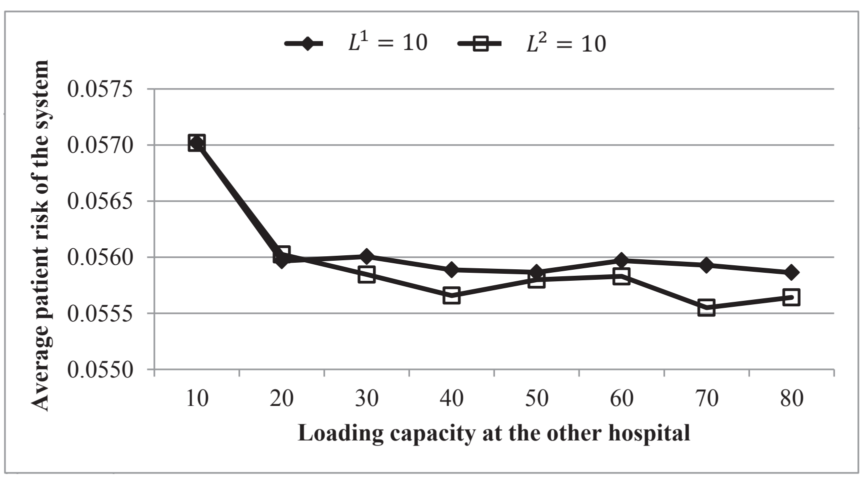
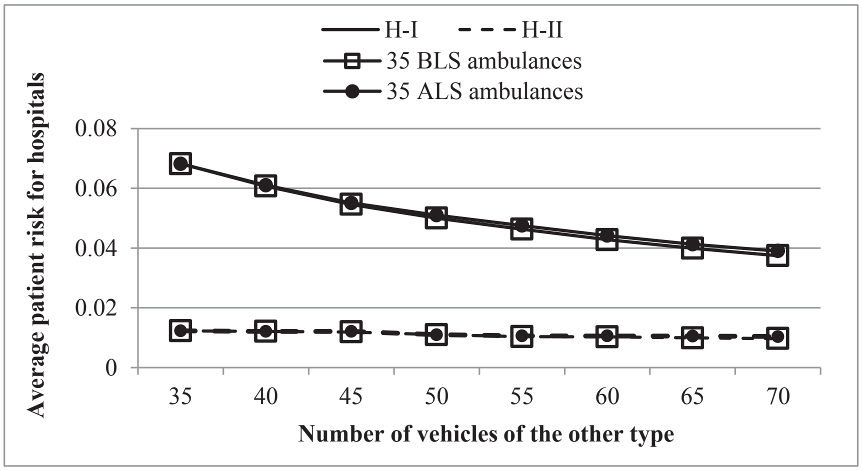


Figure 4.19: The impact of fleet size perturbations on the average patient risk for the hospitals.



pitals is observed to converge, as shown for earlier examples, when utility function u_1^i is used under the max-min fairness criteria. The system-level average patient risks for various loading capacity levels using u_1^i under utilitarian criterion ($\alpha = 0$) are plotted in Figure 4.18. Increasing the loading capacity of H-I has a slightly greater benefit in terms of the average patient risk of the system due to the larger evacuee population using this resource. The average patient risk of each hospital and of the system are only slightly reduced as the loading capacity of either hospital is increased since the fleet size is also a limiting resource in this case. Increasing the number of either ALS or BLS ambulances (keeping the number of vehicles of the other type at 35) is observed to decrease the risk significantly for patients at H-I, but only slightly for patients at H-II as shown in Figure 4.19. Thus, increased fleet size results in a tighter range of average patient risks for hospitals than increased loading capacity. We note that the implementation of the utility function u_2^i requires the separation of the fleet between the hospital groups as the fleet size changes, because this utility function is based on the performance of hospital groups under decentralized strategy. Therefore, specific conditions must be set for the fleet assignment to each hospital group in carrying out an analysis on the impact of fleet size on equity.

4.5 Conclusions

The simultaneous evacuation of multiple hospitals in a region is a challenging problem due to the fairness issues regarding resource allocation among hospitals and patients. In this chapter, we have provided an integer programming model formulation for the evacuation transportation planning for multiple hospitals where there is a central authority that distributes the set of available resources. We discuss the equity issues that arise when the available resources need to be allocated to multiple evacuating hospitals and define the decentralized resource management strategy that provides the best possible individual performance of each hospital group with dedicated resources. Due to the similarities of the hospital evacuation problem and public resource allocation problems, we introduce an equity modeling framework based on welfare economics. We especially focus on the utilitarian, proportional fairness, and min-max fairness criteria at the hospital- and patient-levels.

At the hospital-level, a utility function defined as the complementary value to risk and a benchmark-based utility function, that takes the decentralized resource management strategy solution as basis, are employed. The former utility function is desirable for minimizing the range of average patient risks of the evacuating hospitals, whereas the latter utility function is desirable for distributing the risk reduction among evacuating hospitals equitably. The first utility function is shown to outperform the benchmark-based utility function as it minimizes the range of average patient risks for hospitals whereas the second function increases this range. At the patient-level, a utility function defined as the complementary value to risk is employed. The optimal solutions for the three fairness criteria at both levels are evaluated using illustrative examples characterized by the possibility of improvement in average patient risk for evacuating hospitals. The case study is used for further analysis on the fairness of optimal solutions and the implications of resource levels on equity.

The issues of fairness are extremely difficult in the context of hospital evacuations. There are several interpretations of equity from the perspectives of different players and ethical concerns about resource allocation to patients. We have focused on equity according to the evacuating hospitals and according to individual patients. The proposed equity modeling framework allows for the implementation of different utility functions and for the evaluation of the optimal evacuation plans based on the assumed definition of equity. The performance of utility functions other than the functions defined here can be investigated and the model behavior based on the changes in resource levels can be further analyzed as directions for future research in this area.

Chapter 5

Integration of Building Evacuation and Transportation Planning

Hospital evacuations require moving patients out of the hospital building and transferring them to alternative facilities for continued care. The fundamental assumption to initiate an evacuation transportation model such as the HETM proposed by Bish et al. (2011) is that the building can be evacuated such that patients of the appropriate type are available to satisfy the transportation plan within the vehicle loading capacity, L, which is a simplification of all limitations observed during building evacuation. However, most of the patients in a hospital would require aid from staff to move and patients may not always be moved outside the building fast enough to utilize all available vehicles due to the constraints posed by building structure, staff availability, or moving equipment availability. Therefore, the number of patients that can be transported to receiving facilities depends on the number of patients at the staging area and building evacuation limitations can be better captured by incorporating this phase into evacuation models. In this chapter, we integrate the building evacuation with the subsequent transportation of patients and propose an integer programming model. The impact of building evacuation capabilities on the transportation plan is investigated through illustrative examples. Managerial insights on the optimal resource levels for staff and vehicles are provided.

5.1 Introduction

Hospitals are as vulnerable to natural or man-made disasters as any other facility; however, the evacuation of a hospital is significantly complicated mainly due to the special characteristics of patients and their medical requirements. The evacuation of a hospital requires the movement of patients from the building to the staging area and the transportation of patients to appropriate receiving facilities by available vehicles. The transportation of patients from evacuating hospitals to alternative receiving hospitals has been discussed in the preceding chapters with the fundamental assumption that the building can be evacuated such that patients of the appropriate type are available to satisfy the transportation plan within the physical vehicle loading capacity, L. However, the movement of patients out of the building to the staging area is likely to impose a bottleneck on the evacuation in practice. In this chapter, we integrate the building evacuation with the subsequent transportation of patients.

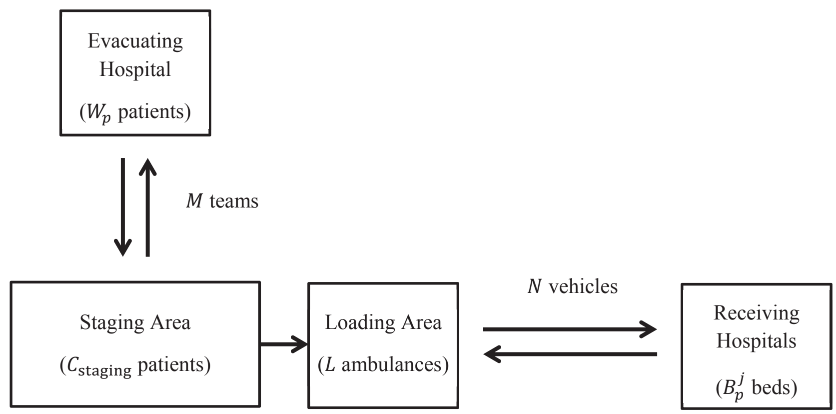
Building evacuations are extensively studied in civil engineering, operations research, and management science as discussed in Chapter 2 where the building evacuation literature is presented. Hamacher and Tjandra (2001) provide a detailed review of the evacuation models most of which are modeled for building evacuation problems. Chalmet et al. (1982) study the emergency evacuation of buildings with bottlenecks and propose network building evacuation models. Hamacher and Tufecki (1987) show that lexicographical optimization can be used for solving building evacuation problems with multiple objectives such as minimizing the total evacuation time along with (a) avoiding unnecessary movements in the building and (b) considering the priority of evacuating different parts of the building. Choi et al. (1988) study maximum flow, minimum cost, and minimax objectives for building evacuations by network flows with side constraints where the side constraints arise from variable arc capacities. Løvås (1998) presents nine models of wayfinding behavior in emergency evacuations. Chen and Miller-Hooks (2008) formulate a building evacuation problem with the objective of minimizing the total time until all evacuees reach the exits where online instructions are shared with the evacuees during evacuation. All of the above studies base their models on the assumption that evacuees have no mobility restriction and do not need assigned personnel for moving out of the building. However, hospital building evacuations are significantly more complicated because patients require medical staff to navigate the building and reach the staging area. Studies on hospital buildings which have such distinct evacuation issues are scarce and, to our knowledge, the integration of building evacuation and transportation problems for hospitals has not been approached from an optimization-based perspective before.

In this chapter, we introduce an integer programming model, in Section 5.2, that integrates the movement of patients from the hospital building to the staging area and their subsequent transportation to receiving hospitals with adequate capacity. This model has a complex structure as a result of the assumptions on capacity and the dependencies among variables. However, it is observed that some medium-scale problems (with around 12,000 variables and 5,000 constraints) can be solved in under five minutes. We discuss some of the assumptions that can be relaxed or modified to improve the tractability of the model as well as to better represent the real-life evacuation issues.

5.2 Model Formulation

In this study, we consider a single evacuating hospital in order to analyze the interaction between the building evacuation and the transportation phases in detail. The proposed model can be easily expanded to include multiple evacuating hospitals which may then lead to additional resource sharing and equity issues to be considered in decision making as discussed in Chapter 4. The complete evacuation process studied is depicted in Figure 5.1. The first phase of a hospital evacuation is the process of moving patients from their location in the hospital to the staging area(s). The evacuating hospital building has several floors, wards, hallways, stairs, elevators, and exits. There are two types of movement inside the hospital building: horizontal and vertical. Horizontal movement would be performed to transfer patients to another ward or section on the same floor whereas vertical movement is required for assisting patients from their original floor to either another floor that is safer or to the exits. In this study, we are not considering any shelter-in strategies and assuming that the threat in question requires the movement of patients out of the hospital building. Therefore, the horizontal movement inside the building is not considered and the constructional limitations for moving patients from their ward to the staging area using the vertical evacuation routes in the building are represented in terms of the time required for this movement.

Figure 5.1: Building evacuation and transportation process.



Let P the set of patient types (e.g., critical care, pediatric, infectious, etc.) where there are W_p patients of type $p \in P$ in the evacuating hospital. These patients must be transported to a set (J) of potential receiving hospitals, where hospital $j \in J$ has B_p^j beds available for patients of type $p \in P$. Let H be the set of all hospitals in the network, such that the evacuating hospital is labeled as hospital 1 and receiving hospitals are labeled as hospitals $2, 3, \cdots, |H| = |J|+1$, i.e., $H = \{1\} \cup J$. The study period is divided into T time intervals of equal length. The travel time from evacuating hospital to receiving hospital $j \in J$ is τ^j time intervals. Travel times between evacuating and receiving hospitals are known, have integral length, and are independent of vehicle type. We note that the latter assumption is easy to modify such that travel times are vehicle type-dependent to reflect their transportation capabilities.

The movement of patients from the hospital wards to the staging area with the aid of evacuation teams precedes the transportation of patients to alternative hospitals. Due to different levels of aid required by patients of different types, the evacuation teams are categorized similar to the vehicles to represent their capabilities over set K. All teams are assumed to be at the staging area at the beginning of the planning horizon and a team is assumed to aid one patient at a time. The total number of idle and busy teams of type $k \in K$ at the evacuating hospital at time t is represented by M_{kt}. The time it takes a team to reach the ward of a patient of type p starting from the staging area is τ_p time intervals and is assumed to be equal to the time it takes to move a patient from the ward to the staging area. This assumption can be easily modified such that the time to move a patient from the ward to the staging area is longer. Additionally, the staging area of the hospital has a capacity of C_{staging} patients at a time. Patients reaching the staging area can occupy space for multiple time intervals if they have to wait while other patients are loaded before them or if all vehicles are en route.

The assignment of teams to patients would be completely unrestricted without any additional rules defined in the model. However, more advanced teams must be preferred for assisting critical patients and any team can assist the non-critical patients. For example, neonatal intensive care unit (NICU) patients need constant respiratory support and an experienced nurse must be holding and assisting the baby patient along with other staff members who move necessary medical equipment for the patient. Therefore, an upgrade team assignment strategy is implemented such that more critical patients can only be assisted by advanced teams, but less critical patients can be assisted by any team as long as there are enough teams available. The set of team types that can assist patients of type p is denoted by K_p. The restricted assignments for a given set of patient types and team types are defined as an additional constraint set in the model formulation.

A set of vehicle types (V), e.g., Advanced Life Support (ALS) ambulances, Basic Life Support (BLS) ambulances, wheelchair vans, and buses is given. At time interval t, the total number of busy and idle vehicles of type $v \in V$ is N_{vt}. This parameter reflects the actual time required by vehicles that might have different origins (depots) to reach the evacuating hospitals. The capacity of a vehicle of type $v \in V$ is C_v patients. The time required to load a vehicle of type $v \in V$ is γ_v time intervals and it is assumed to be equal to the unloading time for that vehicle. The vehicle loading and unloading times are assumed to be independent of the patient type for simplicity and the impact of changing this assumption will be discussed after the model formulation is introduced. Vehicles transport patients directly from the evacuating hospital to hospital $j \in J$ without stopping at multiple hospitals and return to the evacuating hospital as needed.

The number of vehicles that can be loaded at the evacuating hospital in any time interval based on the limitations of the evacuating hospital's loading area is represented by the L-parameter, and is given in ambulance equivalencies; L_v represents a conversion factor for a vehicle of type $v \in V$ to ambulance equivalencies (i.e., a bus takes more effort to load than an ambulance). The loading area is considered as a separate area from the staging area and patients occupy the loading area as long as the vehicle loading time, γ_v. If the number of vehicles available to be loaded at a time interval exceeds L, the excess number of vehicles have to wait at the evacuating hospital before being loaded. Such vehicles will be represented as variables with the evacuating hospital as their origin and destination in the formulation.

Based on the given problem definition, the objective is generating a building evacuation and transportation plan that minimizes the total evacuation risk where risk is defined as the probability that an undesired event occurs. This event can be in the form of a major deterioration of patient health or death due to injury or lack of sufficient treatment. The evacuation risk is defined as a combination of: 1) the threat risk that patients are exposed to while waiting to be transported, and 2) the transportation risk incurred during travel. The cumulative threat risk, Λ_{pt}, calculated in (5.1), is the probability of the undesired event for a patient of type p that remains in the evacuating hospital through time interval t, where λ_{pt} is the probability of the undesired event for a patient of type p that remains in hospital i in time interval t. λ_{pt} at different time intervals are assumed to be independent and a patient does not change type during an evacuation, e.g., a non-critical patient does not become critical by being exposed to the threat risk.

$$\Lambda_{pt} = 1 - \prod_{f=1}^{t}(1 - \lambda_{pf}), \quad \forall p \in P,\ t = 1, \cdots, T. \tag{5.1}$$

The cumulative transportation risk, Θ^j_{pv}, is calculated in (5.2), where θ_{pv} is the probability of the undesired event for a patient of type $p \in P$ transferred by a vehicle of type $v \in V$ for one time interval. θ_{pv} is assumed to be constant through time since the same allocation made in different time intervals would not change the level of treatment supplied by a vehicle. When transporting patients to hospital j, a vehicle of type v that returns to the evacuating hospital is engaged for $2(\tau^j + \gamma_v)$ time units, however, only $(\tau^j + 2\gamma_v)$ time units contribute to the risk expression since transportation risk is incurred only as long as the patient is in the vehicle, including loading and unloading times.

$$\Theta^j_{pv} = 1 - (1 - \theta_{pv})^{(\tau^j + 2\gamma_v)}, \quad \forall j \in J,\ p \in P,\ v \in V. \tag{5.2}$$

The evacuation risk, R^j_{pvt}, associated with the evacuation decision for a patient is calculated in (5.3) by combining the cumulative threat risk (Λ_{pt}) to which the patient is exposed before being transported in time interval t and the cumulative transportation risk (Θ^j_{pv}) based on the vehicle type and the receiving hospital selected. As (5.1) - (5.3) indicate, the cumulative threat, transportation, and evacuation risk functions are non-linear.

$$R^j_{pvt} = 1 - (1 - \Lambda_{p(t-1)})(1 - \Theta^j_{pv}), \quad \forall j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T. \tag{5.3}$$

Parameters:

T	:	number of time intervals in the study period
W_p	:	number of patients of type $p \in P$ initially at the evacuating hospital
M_{kt}	:	total number of teams of type $k \in K$ at the evacuating hospital at time t
τ_p	:	number of time intervals required to reach a patient of type $p \in P$, which is equal to the time required to move a patient from the ward to the staging area
C_{staging}	:	maximum number of patients allowed at the staging area at a time
B^j_p	:	number of beds of type $p \in P$ available at hospital $j \in J$
τ^j	:	number of time intervals required to travel from the evacuating hospital to hospital $j \in J$
N_{vt}	:	total number of busy and idle vehicles of type $v \in V$ at time t
C_v	:	number of patients that can be loaded onto a vehicle of type $v \in V$
γ_v	:	number of time intervals required to load/unload a vehicle of type $v \in V$
L	:	number of vehicles in ambulance equivalencies that can be loaded at the staging area of the evacuating hospital at a time interval
L_v	:	conversion factor for a vehicle of type $v \in V$ to ambulance equivalencies

for loading capacity

Λ_{pt} : cumulative threat risk for a patient of type $p \in P$ that remains in the evacuating hospital through time interval t

Θ^j_{pv} : cumulative transportation risk for a patient of type $p \in P$ transferred by a vehicle of type $v \in V$ from the evacuating hospital to hospital $j \in J$

R^j_{pvt} : total evacuation risk for a patient of type $p \in P$ transferred by a vehicle of type $v \in V$ from the evacuating hospital to hospital $j \in J$ starting at time t

Decision Variables:

s_{pkt} : number of patients of type p moved from their ward in the evacuating hospital to the hospital's staging area by a team of type k starting in time interval t, $\forall p \in P,\ k \in K,\ t = 1, \cdots, T$.

x^j_{pvt} : number of patients of type p transported from the staging area of the evacuating hospital to the staging area of receiving hospital j by a vehicle of type v starting in time interval t, $\forall j \in J,\ p \in P,\ v \in V,\ t = 1, \cdots, T$.

ν_{vt} : number of vehicles of type v assigned to the evacuating hospital at time t, $\forall v \in V,\ t = 1, \cdots, T$.

y^{ij}_{vt} : number of vehicles of type v that move from hospital i to hospital j starting in time interval t, $\forall i,\ j \in H,\ v \in V,\ t = 1, \cdots, T$.

BETM:

$$\text{Minimize} \quad \sum_{j \in J} \sum_{p \in P} \sum_{v \in V} \sum_{t=1}^{T} R^j_{pvt} x^j_{pvt} + \sum_{p \in P} \Lambda_{pT} \left(W_p - \sum_{j \in J} \sum_{v \in V} \sum_{t=1}^{T} x^j_{pvt} \right) \tag{5.4}$$

subject to

$$\sum_{t=1}^{T} s_{pkt} = 0, \quad \forall p \in P,\ k \notin K_p \tag{5.5}$$

$$\sum_{p \in P} \sum_{f=0}^{\min\{2\tau_p - 1, t\}} s_{pk(t-f)} \leq M_{kt}, \quad \forall k \in K,\ t = 1, \cdots, T \tag{5.6}$$

$$\sum_{k \in K} \sum_{t=1}^{T} s_{pkt} \leq W_p, \quad \forall p \in P \tag{5.7}$$

$$\sum_{p\in P}\left(\sum_{k\in K}\sum_{\substack{f=1:\\ t>2\tau_p}}^{(t-2\tau_p)} s_{pkf} - \sum_{j\in J}\sum_{v\in V}\sum_{f=1}^{t} x^j_{pvf}\right) \leq C_{\text{staging}}, \quad \forall t = 1,\cdots,T \tag{5.8}$$

$$\sum_{k\in K}\sum_{\substack{f=1:\\ t>2\tau_p}}^{(t-2\tau_p)} s_{pkf} - \sum_{j\in J}\sum_{v\in V}\sum_{f=1}^{t} x^j_{pvf} \geq 0, \quad \forall p\in P,\ t = 1,\cdots,T \tag{5.9}$$

$$\sum_{v\in V}\sum_{t=1}^{T} x^j_{pvt} \leq B^j_p, \quad \forall j\in J,\ p\in P \tag{5.10}$$

$$\sum_{p\in P} x^j_{pvt} \leq C_v y^{1j}_{vt}, \quad \forall j\in J,\ v\in V,\ t=1,\cdots,T \tag{5.11}$$

$$\nu_{vt} \leq N_{vt} - N_{v(t-1)}, \quad \forall v\in V,\ t=1,\cdots,T \tag{5.12}$$

$$\nu_{vt} + \sum_{\substack{j\in J:\\ t>\tau^j}} y^{j1}_{v(t-\tau^j)} + y^{11}_{v(t-1)} = \sum_{j\in H} y^{1j}_{vt}, \quad \forall v\in V,\ t=1,\cdots,T \tag{5.13}$$

$$\sum_{j\in J}\sum_{i\in H}\sum_{v\in V} y^{ji}_{v1} = 0 \tag{5.14}$$

$$y^{1j}_{v(t-\tau^j-2\gamma_v)} \geq y^{j1}_{vt}, \quad \forall j\in J,\ v\in V,\ t=2,\cdots,T \tag{5.15}$$

$$\sum_{j\in J}\sum_{v\in V}\sum_{f=t-\gamma_v+1}^{t} L_v y^{1j}_{vf} \leq L, \quad \forall t=1,\cdots,T \tag{5.16}$$

$$s_{pkt} \geq 0 \text{ and integer}, \quad \forall p\in P,\ k\in K,\ t=1,\cdots,T \tag{5.17}$$

$$x^j_{pvt} \geq 0 \text{ and integer}, \quad \forall j\in J,\ p\in P,\ v\in V,\ t=1,\cdots,T \tag{5.18}$$

$$\nu_{vt} \in \{0,1,\cdots,N_{vt}-N_{v(t-1)}\}, \quad \forall v\in V,\ t=1,\cdots,T \tag{5.19}$$

$$y^{ij}_{vt} \in \{0,1,\cdots,\lfloor L/L_v\rfloor\}, \quad \forall i,\ j\in H,\ v\in V,\ t=1,\cdots,T. \tag{5.20}$$

The objective function (5.4) minimizes the total evacuation risk where the first summation represents the total risk for patients transported to alternative facilities and the second summation represents the total threat risk for patients moved out of the building, but are not transported to alternative facilities by the end of planning horizon. The restrictions of upgrade team assignment strategy are set by (5.5). Constraint (5.6) limits the number of busy teams to the total number of teams available. The total number of patients of each type moved out from the evacuating hospital is limited by the initial population size in constraint (5.7) and by the staging area capacity in constraint (5.8). Additionally, constraint (5.9) limits the number of patients transported to the number of patients available at the staging area. Constraint (5.10) defines the number of beds of each type available at each alternative care location. Constraint (5.11) represents the vehicle capacity restriction on the number of patients transferred at each time interval. Constraint (5.12) restricts the additional number of vehicles assigned to the evacuating hospital at each time interval by the number of vehicles that become newly available at that time interval. Constraint (5.13) represents the conservation of flow at the evacuating hospital. Similarly, constraints (5.14) and (5.15) stand for the conservation of flow at each receiving hospital. Constraint (5.16) restricts the number of vehicles that can be loaded at each time interval due to a physical loading capacity. Finally the logical integrality and non-negativity of decision variables are defined by constraints (5.17) - (5.20). We next present a structural property of the BETM that allows for tractability for even large-scale problem instances.

Lemma 5.1 . *If a matrix* $\mathbf{A}$ *is full rank, the total unimodularity of* $\mathbf{A}$ *is preserved under the following three elementary row (column) operations: (1) exchanging two rows (columns), (2) multiplying a row (column) by -1, and (3) adding a row (column) to another row (column).*

The proof of Lemma 5.1 is provided in Schrijver (1998).

Proposition 5.1 . *For a feasible set of* y*-variables and a continuous relaxation of the* s*-,* x*-, and* ν*-variables, there exists an optimal solution for the BETM in which the continuous variables have integral values.*

Proof. A feasible set of y-variables must satisfy constraints (5.13) - (5.16) and be integer valued, i.e., satisfy (5.20). Such a set of y-variables can be considered as parameters. The corresponding feasible values of all ν-variables are directly computed by the equality constraint (5.13) and are integer-valued. Therefore, constraints (5.12) and (5.19) are satisfied, too. Then, a continuous relaxation of the s- and x-variables transforms the BETM constraint set as follows: (5.6) - (5.11) and (5.17) - (5.18) into the following:

$$\sum_{p\in P}\sum_{f=0}^{\min\{2\tau_p-1,t\}} s_{pk(t-f)} \leq M_{kt}, \quad \forall k\in K,\ t=1,\cdots,T \tag{5.21}$$

$$\sum_{k\in K}\sum_{t=1}^{T} s_{pkt} \le W_p, \quad \forall p \in P \tag{5.22}$$

$$\sum_{p\in P}\left(\sum_{k\in K}\sum_{\substack{f=1:\\ t>2\tau_p}}^{(t-2\tau_p)} s_{pkf} - \sum_{j\in J}\sum_{v\in V}\sum_{f=1}^{t} x^j_{pvf}\right) \le C_{\text{staging}}, \quad \forall t = 1,\cdots,T \tag{5.23}$$

$$-\sum_{k\in K}\sum_{\substack{f=1:\\ t>2\tau_p}}^{(t-2\tau_p)} s_{pkf} + \sum_{j\in J}\sum_{v\in V}\sum_{f=1}^{t} x^j_{pvf} \le 0, \quad \forall p \in P,\ t = 1,\cdots,T \tag{5.24}$$

$$\sum_{v\in V}\sum_{t=1}^{T} x^j_{pvt} \le B^j_p, \quad \forall j \in J,\ p \in P \tag{5.25}$$

$$\sum_{p\in P} x^j_{pvt} \le C_v y^{1j}_{vt}, \quad \forall j \in J,\ v \in V,\ t = 1,\cdots,T \tag{5.26}$$

$$s_{pkt} \ge 0, \quad \forall p \in P,\ k \in K,\ t = 1,\cdots,T \tag{5.27}$$

$$x^j_{pvt} \ge 0, \quad \forall j \in J,\ p \in P,\ v \in V,\ t = 1,\cdots,T \tag{5.28}$$

where (5.21) - (5.26) correspond to (5.6) - (5.11) and (5.27) - (5.28) correspond to (5.17) - (5.18) with a continuous relaxation of the s- and x-variables, respectively. If the constraint coefficient matrix of an LP is totally unimodular and the right-hand-side values are all integer-valued, the LP has integer basic solutions by Cramer's Rule (Bazaraa et al., 2005). The right-hand-side values of the constraints above are integer-valued and we show that the constraint coefficient matrix satisfies the conditions for total unimodularity as follows.

We observe that for each t, the sum of rows for constraints (5.9) for all patient types is the negative of constraint (5.8). We apply elementary row operations on the constraint coefficient matrix to eliminate constraints (5.8) and (5.9). The reduced row echelon form of $\mathbf{A}$ consists of constraints (5.21) - (5.22) and (5.25) - (5.26). Therefore, $\mathbf{A}$ is an $m \times n$ matrix where $m = |K|T + |P| + |J||P| + |J||V|T$, $n = |P||K|T + |J||P||V|T$, and $\min\{m, n\} = m$. Since the rank of $\mathbf{A}$ is m, $\mathbf{A}$ is full rank and this reduction would preserve total unimodularity by Lemma 5.1.

Camion (1965) proved that a matrix is totally unimodular if and only if every square Eulerian submatrix formed from it is singular, where a submatrix is Eulerian if both the sum of each row and the sum of each column are even. It can be shown that the coefficient matrix of the above LP is totally unimodular by examining the characteristics of its submatrices.

The entries of the coefficient matrix $\mathbf{A}$ of the LP above are $+1$ or 0. Each s-variable appears in $2\tau_p$ consecutive constraints in constraint set (5.21) and once in constraint set (5.22). Each x-variable appears once in each of the two constraint sets (5.25) and (5.26), therefore, every s-variable column of $\mathbf{A}$ has $2\tau_p + 1$ entries of $+1$ and all remaining entries are 0 and every x-variable column of $\mathbf{A}$ has two entries of $+1$ and all remaining entries are 0. Since the sum of each row and column of a Eulerian matrix must be even, all 1×1 Eulerian submatrices of $\mathbf{A}$ consist of a 0 entry and have a determinant of 0. The 2×2 Eulerian submatrices can have either four $+1$ entries or four 0 entries, and have a determinant of 0. The $k \times k$ Eulerian submatrices, such that $k \geq 3$, can have either $c \leq k$ entries of $+1$ in each column, such that $c \leq 2\tau_p$ and $c \pmod 2 = 0$, because each column of $\mathbf{A}$ has $2\tau_p + 1$ entries of $+1$, and an even number of $+1$ entries in each row, resulting in a determinant of 0. We have shown that all square Eulerian submatrices of $\mathbf{A}$ have a determinant of 0, or, equivalently, are singular matrices. Thus, by Camion's theorem, $\mathbf{A}$ is totally uniomodular. Since the right-hand-side of the constraint set is integer valued, by Cramer's Rule, all extreme point solutions to the LP are integer valued (Bazaraa et al., 2005). ∎

The structure of the coefficient matrix of the BETM described in Proposition 5.1 allows for reducing the number of integer variables to only the number of y-variables. Therefore, the branch-and-bound algorithm is required to branch on only the y-variables and the integer program becomes substantially easier to solve.

The favorable structural property of the BETM is achieved by making a set of assumptions as explained above. These assumptions can be modified to obtain a more realistic representation of the actual evacuation process. However, these modifications would cause the tractability of the problem to deteriorate, rendering the model impractical for use in planning and operations. Moreover, the optimal solutions for the more detailed problem may not outperform the solutions for a simplification of the system. One such modification would be allowing patients of different types to be transported by a vehicle. In addition to the vehicle capacity parameter, C_v, a space requirement parameter for each patient type would be defined. Let C_{pv} be the units of space required by a patient of type $p \in P$ in a vehicle of type $v \in V$. Constraint 5.29 that is similar to a knapsack problem constraint would replace (5.11) to reflect this modification.

$$\sum_{p \in P} C_{pv} x^j_{pvt} \leq C_v y^{1j}_{vt}, \quad \forall j \in J,\ v \in V,\ t = 1, \cdots, T \tag{5.29}$$

Consider three patient types with space requirement parameters $C_{1v} = 3, C_{2v} = 2$ and $C_{3v} = 1$ where $C_v = 6$. The above constraint would be $3x^j_{1vt} + 2x^j_{2vt} + x^j_{1vt} \leq 6y^{1j}_{vt}$. The constraint coefficients of the x-variables different from 0 and 1 are allowed in this constraint which violates the special structure of the constraint coefficient matrix that allows for integer solutions for the LP relaxation given a set of feasible integer y-variables.

An alternative way to allow patients of different types to be transported by a vehicle would be defining the upper bound on the number of patients of each type that can be transported by a vehicle of type v, denoted by n_{pv}. The following constraint would be added to the BETM formulation to implement this modification along with constraint (5.11).

$$x^j_{pvt} \leq n_{pv} y^{1j}_{vt}, \quad \forall p \in P,\ j \in J,\ v \in V,\ t = 1, \cdots, T \tag{5.30}$$

For example, if $C_v = 6$, $n_{1v} = 2, n_{2v} = 3$, and $n_{3v} = 6$, the feasible solution must satisfy the constraints $\{x^j_{1vt} \leq 2y^{1j}_{vt}, x^j_{2vt} \leq 3y^{1j}_{vt}, x^j_{3vt} \leq 6y^{1j}_{vt}, x^j_{1vt} + x^j_{2vt} + x^j_{3vt} \leq 6y^{1j}_{vt}\}$. Therefore, if there is a single vehicle assigned for this trip, i.e., $y^{ij}_{vt} = 1$, the feasible solutions would be $(x^j_{1vt}, x^j_{2vt}, x^j_{3vt}) \in$ {(0,0,6), (0,1,5), (0,2,4), (0,3,3), (1,0,5), (1,1,4), (1,2,3), (1,3,2), (2,0,4),(2,1,3), (2,2,2), (2,3,1)}. Unlike the above modification, this additional constraint does not have an adverse impact on the tractability of the model because the constraint coefficient matrix structure is maintained.

Another aspect of the BETM that can be modified for a more realistic representation of the process would be defining the threat risk for patients based on which phase of the evacuation they are in. We define the threat risk to be in effect from the beginning of the evacuation until the time at which a patient is loaded onto a vehicle to capture the effects of waiting on a patient. However, when the total waiting time of the patient is not spent at the staging or loading area, but can also be spent waiting inside the hospital building, the threat risk can be quite different depending on the conditions at the location. Patients can be in continuous care inside the hospital until a team of staff reaches them if the hospital has not experienced a utility loss disrupting patient care. Then, patients may be exposed to lower, and probably a different form of, threat risk than they will be exposed to at the staging area. However, if there is an internal threat such as a fire or a chemical leak, the threat risk accrued while patients are waiting to be assisted out may be significantly higher than the risk accrued at the staging area. The distinction of threat risk inside and outside the hospital building would require the use of time epochs at which patients exit the building and leave the staging area. This modification can have an adverse effect on the tractability of the model and additional simplifying assumptions may be required. We consider this concept of location-based threat risks as a potential extension of this study.

A Simplified Version of the BETM:

The BETM has a complex structure mainly due to the interaction between the vehicle flow and patient flow. To improve tractability of the model, the following simplifying assumptions can be made: (1) there is a single team type, i.e., $|K| = 1$, (2) there is a single vehicle type, i.e., $|V| = 1$, and (3) each vehicle can transport one patient at a time, i.e., $C_v = 1$. These assumptions lead to the following model in which the only decision variables are the s- and x-variables.

BETM-Simplified:

$$\text{Minimize} \quad \sum_{j \in J} \sum_{p \in P} \sum_{t=1}^{T} R_{pt}^{j} x_{pt}^{j} + \sum_{p \in P} \Lambda_{pT} \left(W_p - \sum_{j \in J} \sum_{t=1}^{T} x_{pt}^{j} \right) \tag{5.31}$$

subject to

$$\sum_{p \in P} \sum_{f=1}^{\min(2\tau_p - 1, t)} s_{p(t-f)} + \sum_{p \in P} s_{pt} \leq M_t, \quad \forall t = 1, \cdots, T \tag{5.32}$$

$$\sum_{t=1}^{T} s_{pt} \leq W_p, \quad \forall p \in P \tag{5.33}$$

$$\sum_{p \in P} \sum_{\substack{f=1: \\ t > 2\tau_p}}^{(t-2\tau_p)} s_{pf} - \sum_{j \in J} \sum_{p \in P} \sum_{f=1}^{t} x_{pf}^{j} \leq C_{\text{staging}}, \quad \forall t = 1, \cdots, T \tag{5.34}$$

$$\sum_{\substack{f=1: \\ t > 2\tau_p}}^{(t-2\tau_p)} s_{pf} - \sum_{j \in J} \sum_{f=1}^{t} x_{pf}^{j} \geq 0, \quad \forall p \in P, \ t = 1, \cdots, T \tag{5.35}$$

$$\sum_{t=1}^{T} x_{pt}^{j} \leq B_p^{j}, \quad \forall j \in J, \ p \in P \tag{5.36}$$

$$\sum_{j \in J} \sum_{p \in P} \sum_{f=1}^{\min\left\{2(\tau^j+\gamma)-1, t\right\}} x^j_{p(t-f)} + \sum_{j \in J} \sum_{p \in P} x^j_{pt} \leq N_t, \quad \forall t = 1, \cdots, T \quad (5.37)$$

$$\sum_{j \in J} \sum_{p \in P} \sum_{f=t-\gamma+1}^{t} x^j_{pf} \leq L, \quad \forall t = 1, \cdots, T \quad (5.38)$$

$$s_{pt} \geq 0 \text{ and integer}, \quad \forall p \in P, \ t = 1, \cdots, T \quad (5.39)$$

$$x^j_{pt} \geq 0 \text{ and integer}, \quad \forall j \in J, \ p \in P, \ t = 1, \cdots, T \quad (5.40)$$

The major change in the formulation is in vehicle flow constraints that involve y-variables. The original constraints (5.11) - (5.16) are replaced by (5.37) and (5.38) to represent the additional assumptions. The simplified BETM is no longer a linear program as the constraint coefficient matrix is no longer totally unimodular. However, the optimal solution for continuous x-variables (s-variables) given a feasible set of s-variables (x-variables) is integer-valued.

The BETM-Simplified is a degenerate problem, that is, an extreme point of the feasible polyhedron is defined by more than one basis since the number of hyperplanes that pass through an extreme point is greater than the number of decision variables. Degeneracy results in zero-valued basic variables which show that a different basis can represent the same optimal solution. This degenerate problem has a unique optimal solution, therefore, the dual problem has multiple optimal solutions.

In order to investigate the impact of resource levels on the optimal evacuation plan and to determine the required resource levels to achieve a risk target, a parametric analysis on the resource constraints will be performed in Section 5.3. We next define the resource utilization measures that will be used in the following analysis.

Resource Utilization:

The teams, the staging area, the loading area, and the vehicles are re-usable resources and their utilization provides information about which resources are restricting the evacuation performance. In other words, resource utilization information guides the decision maker in determining the resources in the system to invest in for further improvement of the evacuation performance. We define utilization for the four re-usable resources as follows.

$$\text{Team utilization} = \frac{\sum_{p\in P}\sum_{k\in K}\sum_{t=1}^{T} 2\tau_p s_{pkt}}{\sum_{k\in K} M_{k1} ET}, \tag{5.41}$$

where ET is the evacuation duration defined as the time at which the last patient reaches a receiving facility.

$$\text{Fleet utilization} = \frac{\sum_{j\in J}\sum_{v\in V}\sum_{t=1}^{LLT} 2(\tau^j + \gamma_v) y_{vt}^{1j}}{\sum_{v\in V} N_{v1}(ET - FLT + 1)}, \tag{5.42}$$

where LLT is the last load time at which the last patient leaves the staging area and is loaded into a vehicle and FLT is the first load time at which the first patient is loaded into a vehicle. Then, $ET - FLT + 1$ is the total vehicle-time in which vehicles can be used and the numerator is the total vehicle-time that vehicles are en route or loading or unloading patients. Similarly, the staging area and loading area utilizations are calculated as below.

$$\text{Staging area utilization} = \frac{\sum_{p\in P}\sum_{t=1}^{T}\left(\sum_{k\in K}\sum_{f=1}^{t-2\tau_p} s_{pkf} - \sum_{j\in J}\sum_{v\in V}\sum_{f=1}^{t} x_{pvf}^{j}\right)}{C_{\text{staging}}(LLT - FLT + 1)} \tag{5.43}$$

$$\text{Loading area utilization} = \frac{\sum_{j\in J}\sum_{p\in P}\sum_{v\in V}\sum_{t=1}^{T} x_{pvt}^{j}}{L(LLT - FLT + 1)} \tag{5.44}$$

5.3 Model Analysis

In this section, we will present the computational results for a case study and analyze the results with respect to performance measures including risks, resource utilization, patient prioritization schemes, building clearance time, and evacuation completion time. A network that consists of one evacuating hospital and nine receiving hospitals is considered in this case study. All hospitals in this study are assumed to be members of an emergency preparedness alliance for mutual emergency management planning and support among hospitals in case of local disasters.

Patients are categorized into three types based on their medical condition and requirements and there are 135, 160, and 65 patients of types 1, 2, and 3, respectively, type 1 being the most critical, such as adult or pediatric critical care patients, and type 3 being the least critical, such as outpatients or psychiatric patients. The time it takes to move the patients out of the building depends on their type and given as $\tau_1 = 3, \tau_2 = 2, \tau_3 = 1$ to be consistent with the care required by each patient type. Patients can be transferred to adequate beds available at the nine receiving hospitals. The instantaneous data used in this case study, which includes 360 patients in the evacuating hospital and 535 beds available at the nine receiving hospitals by patient type, is displayed in Table 5.1. The length of a time interval is assumed to be 10 minutes and the one-way travel times between the evacuating hospital and the receiving hospitals in number of 10-minute time intervals are provided in Table 5.1.

It is assumed that the vehicles are available for immediate use once the evacuation is initiated and can be dispatched as soon as the first patients arrive at the staging area. In case the vehicles are dispatched from further locations such as depots, the time it takes for vehicles to reach evacuating hospitals can be easily changed by adjusting the N_{vt}-parameters. The vehicle types used in this study are ALS and BLS ambulances both capable of carrying one patient at a time. Transportation risk is assumed to be higher for BLS ambulances for all patient types due to the level of treatment that can be supplied during patient transfer. Also, the more critical the patient, the greater the difference between the transportation risk of an ALS ambulance and a BLS ambulance for that patient is. The time required to load a patient into an ALS or BLS ambulance is assumed to be one time interval that is also equal to the time to unload a patient from an ambulance. Initially, the evacuating hospital is assumed to have a physical loading capacity of 10 ambulances at a time interval.

The threat risk parameter, λ_{pt}, is indexed on time, patient type, and evacuating hospital and depends on the nature of the threat. The cumulative threat risk, Λ_{pt}, used in this case study is a function of an exponential risk parameter and has an increasing slope. The threat that causes the evacuation is assumed to affect all patients in the evacuating hospital, but the resistance to the effects of this threat varies among patient types. Table 5.2 gives the

Table 5.1: The number of patients (W_p), the number of beds available (B_p^j), and the travel time between hospitals (τ^j).

	Number of Beds B_p^j			Travel Time τ^j (intervals)
Hospital (j)	1	2	3	
1	30	45	15	4
2	0	15	0	4
3	0	30	10	8
4	55	25	30	6
5	35	50	0	7
6	20	0	50	15
7	10	0	15	8
8	35	20	5	2
9	20	15	5	7
TOTAL (beds)	205	200	130	
W_p	135	160	65	

values of the threat risk parameters, λ_{pt}, for each patient type and the transportation risk parameters, θ_{pv}, for each patient type and transportation type combination used in this study.

Table 5.2: Threat risk functions (λ_{pt}) and transportation risk parameters (θ_{pv}).

Patient Type	Threat Risk	Transportation Risk	
		ALS	BLS
1	$0.0012\exp(t/35)$	0.001	0.002
2	$0.0009\exp(t/35)$	0.0005	0.001
3	$0.0006\exp(t/35)$	0.0001	0.0002

We begin by assuming that the staging area capacity is 50 patients at a time, the loading capacity of the hospital is 10 ambulances at a time, and 70 ambulances are available in total for transportation at the beginning of planning horizon. Moreover, we assume that there are 30 teams available to assist patients out of the building. We will perform a sensitivity analysis with respect to the changes in the levels of these three resources to observe their impact on the optimal plan. We will provide managerial insights into optimal resource levels for risk-minimization. The results based on the optimal solution for the initial problem instance described above are shown in Table 5.6.

Table 5.3: Optimal results for the case study ($M = [15, 15], C_{\text{staging}} = 50, L = 10, N = [35, 35]$).

	System	By patient type		
		1	2	3
W	360	135	160	65
Total risk	25.8457	19.0563	6.2310	0.5584
Average risk	0.0718	0.1412	0.0389	0.0086
Threat risk	24.1627	18.2770	5.4205	0.4652
Transportation risk	1.8461	0.9042	0.8480	0.0939
Building clearance time	64	64	49	22
Evacuation duration	67	67	51	23
Team utilization	82.29%			
Fleet utilization	95.24%			
Staging area utilization	0.32%			
Loading area utilization	53.73%			
Solution time (s)	45.16			

5.3.1 Optimal resource levels given a risk-based target

Making the optimal allocation decisions given the set of limited resources is essential for hospital evacuation planning. However, as the complexity of resource management increases, it is even more valuable to have guidelines to determine the sufficient level of resources required to achieve a target evacuation performance. We define this target as the optimal evacuation risk when the level of a set of resources is fixed and the other resources are assumed to be unlimited. This target risk is then used to determine the optimal levels for the resources initially assumed to be unlimited.

In the hospital evacuation process, the resources are (1) teams, (2) staging area capacity, (3) loading capacity, (4) fleet, and (5) receiving beds. The staging area capacity, the loading capacity, and the number of receiving beds are constant through time. The staging area capacity impacts the utilization of the teams and the building clearance time, because teams can move more patients than the number of patients that can be loaded if there is a positive staging area capacity. The loading area can be designed as lanes assigned to different types of vehicles or as a common loading dock. Although the utilization of the loading area fluctuates over time, the loading capacity is fixed. Unlike the staging and loading capacities, the available number of beds at receiving hospitals only decrease over time as beds are not re-usable. However, the teams of staff and the vehicles are dynamic resources and the number of available teams and vehicles may change in time as a result of additional staff called in or vehicles requested from other hospitals or private service providers. As the dynamic resources in our evacuation model, additional teams of staff to assist patients out of the building and additional vehicles to transport patients to receiving facilities may be easier to acquire than expanding the staging and loading area or securing additional receiving beds; but investing in the static resources of the evacuating hospital may reduce the required number of teams or vehicles. In order to determine the optimal number of teams and vehicles required to achieve a risk target, we assume that these numbers are constant through time. Given the set of limited resources consisting of receiving beds, staging area capacity, and loading capacity, we determine the optimal levels for teams and fleet through optimization as demonstrated below.

The target risk may be achieved by multiple resource options each with a different combination of the number of teams and the fleet size. The optimal building evacuation decisions would fully utilize the teams and if there are a lot of teams, there may be too many patients waiting at the staging area due to the limited loading capacity and fleet size. On the other hand, if there are too many vehicles and not enough teams, vehicles may have to be idle as they wait for patients to arrive at the staging area. Therefore, the goal is balancing the vertical and horizontal evacuation operations by finding the best combination of resource levels.

Consider the case study introduced above with only one type of teams and one type of vehicles (ambulances), i.e., $|K| = 1, |V| = 1$. All teams and vehicles are assumed to be available at the beginning of the evacuation, therefore, the number of teams and the fleet size are constant through time, i.e. $M_{kt} = M_k, N_{vt} = N_v, \ \forall t = 1, \cdots, T$. The above set of beds at receiving hospitals shown in Table 5.1, the staging area capacity of 50 patients at a time, and the loading capacity of 10 ambulances at a time are given. As there are 360 patients, teams and vehicles would be unlimited if there is one team and one vehicle ready to be allocated to each patient at any time. Therefore, initially, we assume that there are 360 teams available to assist patients out of the building and 360 ambulances are available for transportation at the beginning of evacuation.

The risk measures based on the optimal solution of the BETM are shown in Table 5.4 and the system's total risk of 11.1717 is used as the target risk level to determine the minimum number of teams and vehicles required, minimizing $\sum_{k \in K} M_k$ and $\sum_{v \in V} N_v$, respectively, such that $R \leq R^*$, where R is the total evacuation risk and R^* is the optimal objective function value of the BETM with unlimited number of teams and vehicles. The minimum amount of resources required to execute the optimal evacuation plan with this target total risk are 70 teams and 140 vehicles. These resource levels may be higher than a hospital of this size can acquire even when regional resources are included. There may not be enough teams of staff especially if the evacuation is initiated at a time of day or day of week when staffing level is lower than usual, such as at night or on the weekend. The required fleet size may not be reasonable considering the cost of purchasing and operating the medical transportation vehicles along with the staff involved in transportation. We will evaluate the evacuation performance given more reasonable and realistic resource levels and compare it with these results later.

The building clearance time and evacuation duration when the minimum amount of resources are used are shown in Table 5.4 as the time measures. All patients inside the

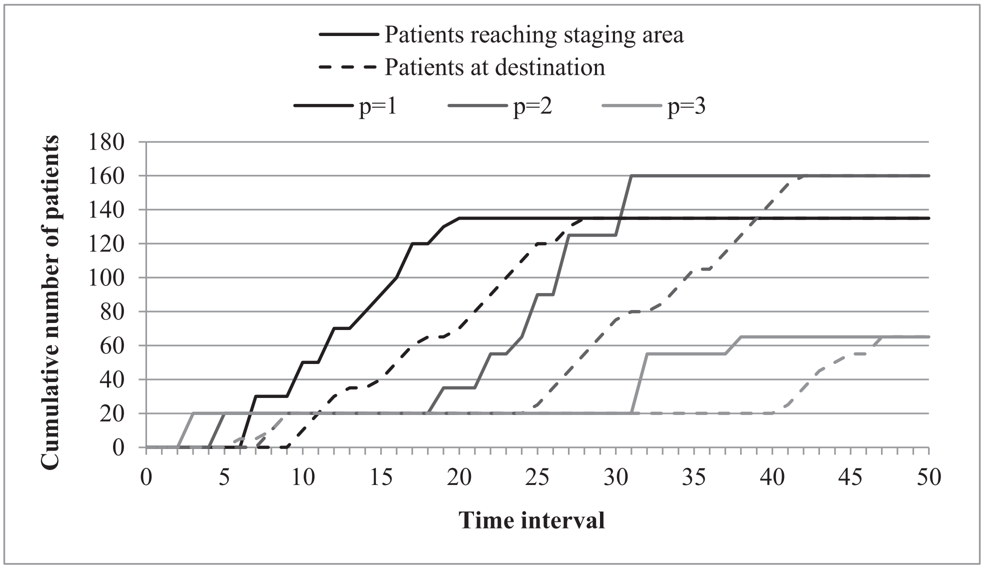
Table 5.4: Optimal results for the case study ($M = 70, C_{\text{staging}} = 50, L = 10, N = 140$).

	System	By patient type		
		1	2	3
W	360	135	160	65
Total risk	11.1717	3.5464	5.8234	1.8020
Average risk	0.0310	0.0263	0.0364	0.0277
Threat risk	9.7045	2.6733	5.2772	1.7540
Transportation risk	1.5083	0.8923	0.5666	0.0495
Building clearance time	37	17	29	37
Evacuation duration	47	28	42	47
Team utilization	61.00%			
Fleet utilization	67.93%			
Staging area utilization	21.76%			
Loading area utilization	100.00%			
Solution time (s)	1.03			

hospital building are moved to the staging area in 37 time intervals (6.16 hours) and their transportation to alternative facilities is completed in 47 time intervals (7.83 hours) as shown in Figure 5.2. The latest building clearance time and evacuation duration are observed for patient type 3, the least critical patients with the lowest threat and transportation risk parameters. However, patients of type 3 do not have the lowest average evacuation risk because of both being transported at later time intervals and the greater distance they have to travel compared to the other patients. These patient prioritization results are for this problem instance and can be different depending on the network parameters.

The utilization levels of resources for the example are calculated by (5.41) - (5.44) and shown in Table 5.4 for the minimum required number of teams and vehicles. If 360 teams and 360 vehicles are available, team utilization is only 13.71% and fleet utilization is only 28.35%. Therefore, both the team and fleet utilizations are increased by using the minimum number of resources required. However, the staging area utilization is 59.72% if there are

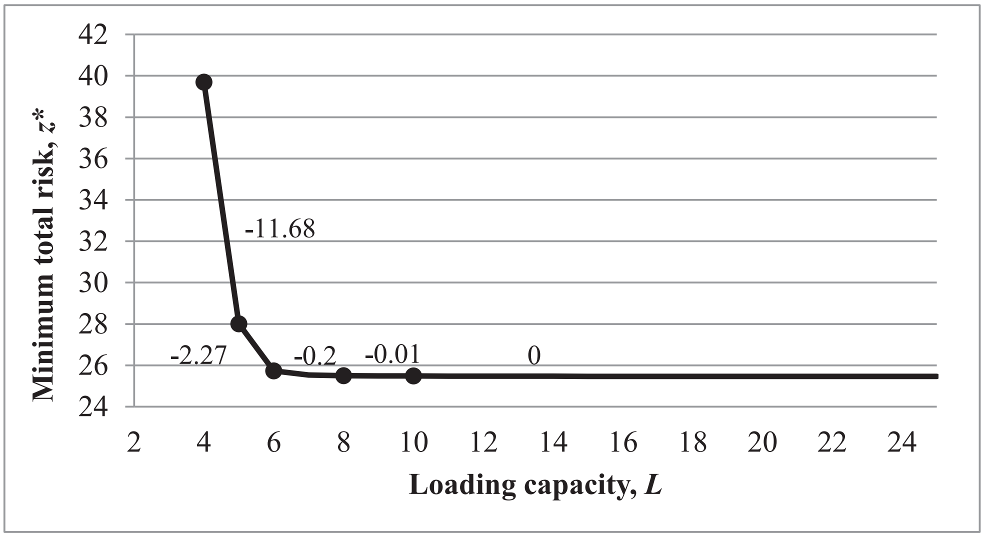
Figure 5.2: The cumulative number of patients who are moved to the staging area and who reach their destinations by time interval t by type.



unlimited teams and vehicles and is reduced to 21.76% when the minimum required number of teams and vehicles are used; because the number of patients moved to the staging area is limited by the reduced number of teams and this results in less number of patients waiting at the staging area to be loaded into vehicles. If the conditions of the staging area are worse than the conditions inside the hospital in terms of patient care, the reduced waiting time at the staging area is an indirect improvement in the evacuation performance. We also note that the optimal solution of this problem instance is achieved even when there is no staging area capacity available since staging area is never the restricting resource.

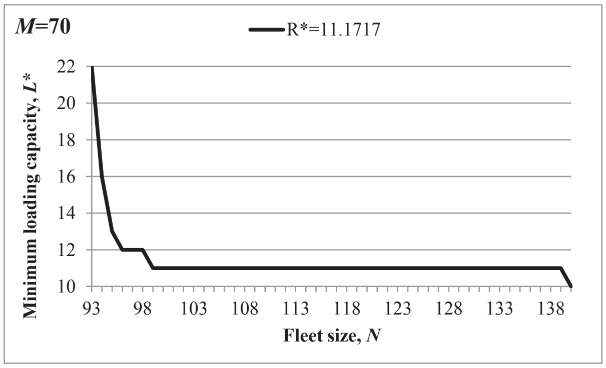
The perturbation function of the loading capacity constraint is shown in Figure 5.3. The left- and right-shadow prices for the loading capacity that refer to the reduction in the optimal total risk are labeled on the perturbation function. The problem becomes infeasible only if the loading capacity is below 4 ambulances. Although increasing the loading capacity does not result in a significant risk reduction above 10 ambulances, there may be other consequences of increased loading capacity regarding the other resource requirements and investing in loading capacity may prove to reduce the minimum required resource levels to obtain the same risk for the system.

Figure 5.3: The cumulative number of patients who are moved to the staging area and who reach their destinations by time interval t by type.

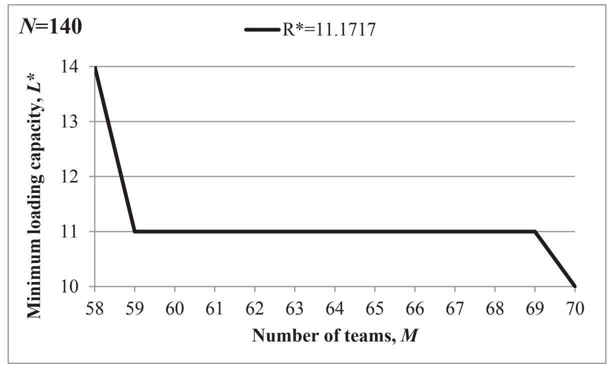


Defining the target risk as the minimum evacuation risk given a loading capacity of 10 ambulances at a time, i.e., $R^* = 11.1717$, we would like to identify the resources which the hospital management should invest in to achieve its target. By fixing the number of teams at the minimum requirement of 70 teams, the minimum loading capacity required, L^*, is obtained given the feasible number of vehicles ($93 \leq N \leq 140$) such that $R \leq R^*$. The resulting combinations of the optimal loading capacity for the feasible fleet size values are shown in Figure 5.4 (a). It is observed that if the loading capacity is increased by only one ambulance, the minimum required number of vehicles can be significantly reduced from 140 to 99. If the management's goal is reducing the fleet size, at least 93 vehicles are required, but the hospital has to have a loading capacity of 22 ambulances at a time instead of their current capacity of 10 ambulances. If the loading capacity expansion is a more expensive investment than acquiring more vehicles, the minimum loading capacity requirement can be dropped to 12 ambulances at a time by adding only three vehicles to the fleet. If, instead, the fleet size is fixed at 140 vehicles, the L^* values given the feasible number of teams ($58 \leq M \leq 70$) are found as shown in Figure 5.4 (b). Increasing the loading capacity to 11 ambulances reduces the minimum number of teams required from 70 to 59.

Figure 5.4: The minimum loading capacity required to achieve risk target R^* (a) for feasible fleet size values when $M = 70$ and (b) for feasible number of teams when $N = 140$.



(a)

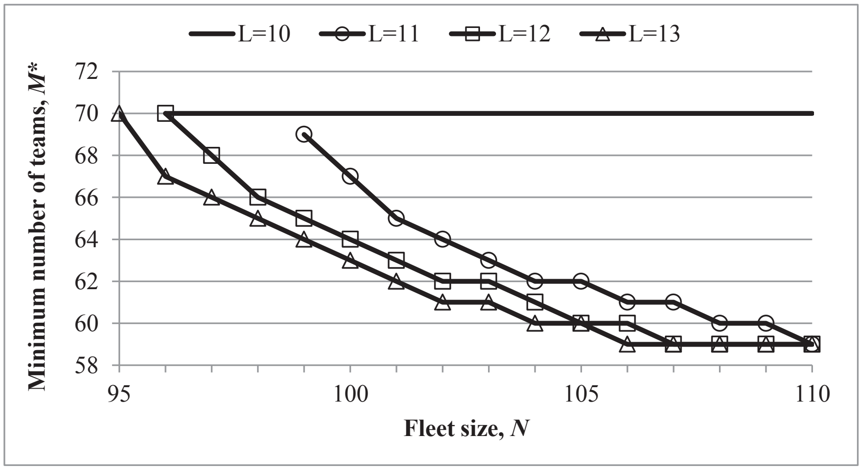


(b)

Figure 5.4 shows the possible improvement in minimum resource requirements if the management invests in incrementing the loading capacity when either one of the resource levels is fixed. Additionally, the set of feasible resource levels where neither the number of teams nor the fleet size is fixed would be useful in decision making. Maintaining the same risk target and increasing the loading capacity, the minimum number of teams for the feasible fleet size values are shown in Figure 5.5. The problem is feasible for $N \geq 99$ when $L = 11$, for $N \geq 96$ when $L = 12$, and $N \geq 95$ when $L = 13$. At a given fleet size, the minimum required number of teams can only decrease as the loading capacity increases. Given this set of optimal resource levels, the hospital management can make an informed decision by choosing the least costly option based on the cost of each type of resource. Let c_k be the operating cost of a team and c_v be the cost of operating a vehicle. For example, if $c_k = 2$ and $c_v = 3$ for this problem, the best option is $(M^* = 65, N^* = 101)$ at a total cost of 433.

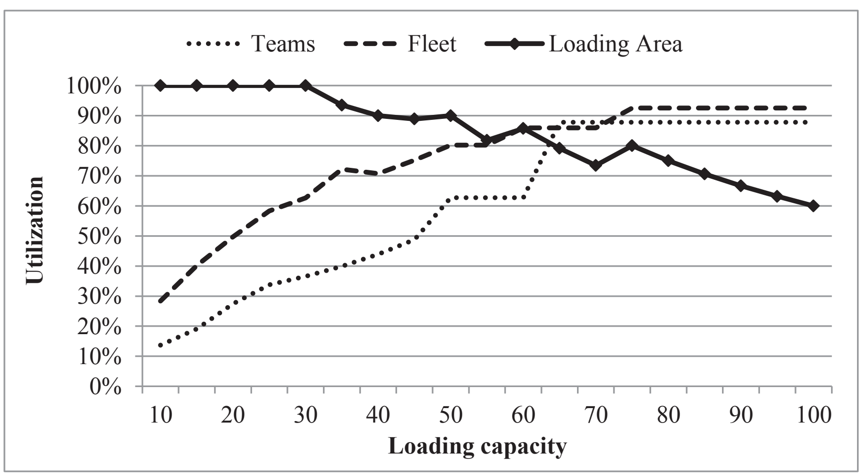
In this problem instance, the bottleneck resource is clearly the loading capacity of 10 ambulances at a time that is always fully utilized. Given the receiving hospital capacities and the staging area capacity as above and assuming 360 teams and 360 vehicles, as the loading capacity is increased, the utilization of teams and vehicles is shown in Figure 5.6. Both the team utilization and fleet utilization are monotonically increasing as L is increased.

Figure 5.5: The minimum required number of teams for feasible fleet size values to achieve the risk target $R^* = 11.1717$ when $L = \{10, 11, 12, 13\}$.



However, the loading area utilization does not necessarily decrease monotonically and may fluctuate instead, depending on the divisibility of the number of patients by the loading capacity.

Figure 5.6: Utilization of teams and fleet versus loading capacity.



As the bottleneck resource, the impact of increasing the loading capacity on the minimum resource requirements for the teams and vehicles is of interest. If $L = 11$ given that there are 360 teams and 360 vehicles, the risk-minimizing solution has a total risk of 10.1724 as shown in Table 5.5. The minimum number of teams and vehicles required to achieve this risk target are 77 and 154, respectively. Compared to the requirements when $L = 10$, besides the additional loading capacity of one ambulance, 7 more teams and 14 more vehicles are required to reduce the risk from 11.1717 to 10.1724 and the evacuation time from 47 time intervals to 44 intervals. The resource utilization levels for different combinations of number of teams and fleet size are given in Table 5.5.

Table 5.5: Optimal results for the case study ($M = 77, C_{\text{staging}} = 50, L = 11, N = 154$).

	System	By patient type		
		1	2	3
W	360	135	160	65
Total risk	10.1724	3.3949	5.2479	1.5296
Average risk	0.0283	0.0251	0.0328	0.0235
Threat risk	8.7015	2.5206	4.6995	1.4814
Transportation risk	1.5083	0.8923	0.5666	0.0495
Building clearance time	32	16	25	32
Evacuation duration	44	27	39	44
Team utilization	64.12%			
Fleet utilization	73.56%			
Staging area utilization	28.48%			
Loading area utilization	99.17%			
Solution time (s)	0.83			

The minimum required resource levels identified for the above problem instances can be too high in practice and the reasonable compromise in risk for more realistic resource levels should be determined. We assume that 30 teams and 70 ambulances are available at the beginning of evacuation. The results based on the optimal solution for the problem instance described above are shown in Table 5.6. Both the teams and the fleet have high utilization levels in this instance whereas the staging area has a very low utilization. The loading capacity has a 56.25% utilization and increasing L reduces this utilization while the team utilization remains as 83.60%, because the number of teams is the most restricting resource level. Therefore, increasing the number of available teams provides the greatest risk reduction. This problem instance exhibits the importance of determining the feasible range of resource levels for the actual evacuation network in order to implement the optimization-based recommendations.

Table 5.6: Optimal results for the case study ($M = 30, C_{\text{staging}} = 50, L = 10, N = 70$).

	System	By patient type		
		1	2	3
W	360	135	160	65
Total risk	25.4860	19.4725	5.5945	0.4190
Average risk	0.0708	0.1442	0.0350	0.0064
Threat risk	24.1320	18.7145	5.0478	0.3698
Transportation risk	1.5083	0.8923	0.5666	0.0495
Building clearance time	63	63	36	26
Evacuation duration	74	74	46	36
Team utilization	83.60%			
Fleet utilization	93.37%			
Staging area utilization	0.05%			
Loading area utilization	56.25%			
Solution time (s)	7.94			

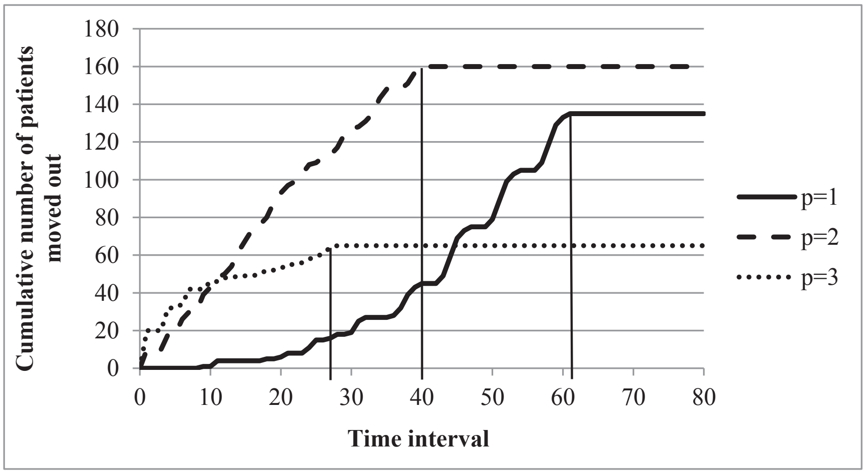
5.3.2 Patient prioritization

A significant question to ask about the optimal evacuation plan is whether there is an underlying patient prioritization scheme suggested by the risk-minimizing model. If there is such a scheme, is it logical and practical? Patient priorities depend greatly on the problem parameters and it is not possible to generalize a certain pattern for even simplified cases. For example, consider only two patient types: critical and non-critical. The greater vulnerability of the critical patients is reflected by higher threat risk and transportation risk parameters than the risk parameters for the non-critical patients. The critical patients also would require more time and effort to move out of the building. If the receiving beds are distributed over the evacuation network such that the distance to receiving hospitals with adequate capacity is the same for both patient types, the teams and fleet are not restricting resources, and the only restriction is the loading capacity, one might expect the optimal solution to be evacuating the critical patients to the closest hospitals before the non-critical patients. However, the risk-minimizing solution for such an instance can actually evacuate the non-critical patients before the critical ones in order to utilize the resources more by moving the non-critical patients out of the building earlier. In a real evacuation network, the order of patient types evacuated would change based on several parameters such as the distance to receiving beds, the number of available vehicles, or the number of teams.

According to the results of the case study presented above with 30 teams and 70 vehicles, the order in which patients are assisted out of the building is shown by the vertical lines in Figure 5.7 where the cumulative number of patients of each type moved out is plotted. Although they are the most critical patients, patients of type 1 are completely moved out of the building the latest, at time interval 61. First, all of the least critical patients are moved by time interval 28 followed by type 2 patients completely moved out by time interval 40. This is an interesting result, because there are hospital evacuation incidents reported in the literature that suggest assisting the least critical patients first during building evacuation. This can be considered analogous to a job scheduling problem where the objective is minimizing the sum of completion times, or equivalently, the sum of flow times for jobs. Each team is analogous to a machine and the task of moving a patient from its ward to the staging area is a job. If there was a single team available and if the objective was minimizing the building clearance time instead of minimizing risk, moving patients in the shortest processing time (SPT) order would be optimal. When there are multiple teams, the patients prioritized based on the SPT order would be assigned to the earliest available team to minimize the sum of the times to reach the staging area. However, in our problem, the optimal solution is not limited to the building evacuation process and depends on the subsequent transportation plan as well.

It is possible to observe an opposite priority order of patient types to be evacuated when

Figure 5.7: Cumulative number of patients of each type (135, 160, and 65 patients) moved out of the building over time.



the desirable conditions for evacuating the most critical patients first are met. For example, if all the available beds for critical patients are at the closest receiving hospital and the available beds for the other patient types are at much further receiving hospitals, the optimal solution suggests that the most critical patients are evacuated first and the least critical patients are evacuated last. Moreover, if the threat risk of waiting inside the building and at the staging area are defined separately for each patient type, this may suggest a different prioritization scheme to minimize the total risk. If the critical patients are safer inside the hospital building than they would be outside at the staging area, they may not be the first patients evacuated even if they can be transferred to the closest receiving hospital, because there may not be enough vehicles to meet the flow of critical patients to the staging area. Several aspects of the problem should be considered in investigating the patient prioritization behavior of this integrated building evacuation and transportation model. Therefore, specific patient prioritization schemes can only be generalized for a very restricted set of problem instances.

5.4 Conclusions

The evacuation of the hospital building imposes restrictions on the transportation of patients to alternative facilities for continued care. In this chapter, we include the building evacuation phase in the model formulation instead of relying only on a loading capacity parameter that is an aggregation of all limitations that precede the transportation phase. The BETM integrating building evacuation with transportation is an integer program that is tractable due to its totally unimodular coefficient matrix structure and the integer-valued demand and supply levels. The utilization of teams of staff assisting patients out of the building, staging area capacity, loading area capacity, and fleet of vehicles is used in analyzing the resource requirements of the system. A risk target driven approach is described to determine the optimal resource levels and to guide the decision maker in the investment in additional resources. Due to the interactions among these reusable resources, investing in a single resource affects the minimum requirements for the other resources. Therefore, a combination of perturbations should be considered in sensitivity analysis to gain a complete understanding of the system.

The proposed model successfully captures the interactions between the hospital building evacuation and the transportation phases. Although this model defines the restrictions of building evacuation in terms of not only the loading capacity, but also the teams of staff available, we believe that a more detailed representation of the evacuation operations inside the building would be a valuable extension of this study.

Chapter 6

Conclusions and Directions for Future Research

In this dissertation, we have addressed the modeling and analysis of hospital evacuation logistics. Emergency management has become increasingly important as the number and magnitude of disasters increased in the past few decades. It is especially a challenging problem for hospitals due to the specific medical requirements of the evacuee population consisting of patients with various levels of criticality. Hospitals must develop emergency preparedness and evacuation plans in order to satisfy the accreditation requirements, but more importantly, they must develop operational evacuation plans that are strategically flexible in order to improve the safety of their patients, staff, and visitors. Hospital evacuation incidents reported in the literature show that even the hospitals with a written plan can find their plans futile in guiding them through evacuation in a particular disaster scenario. Therefore, hospital evacuations demand a more thorough analysis to produce functional plans. The scarcity of research on mathematical modeling for hospital evacuations motivates the work in this dissertation on developing an optimization-based decision support system.

This dissertation aims to develop decision support models for the logistical hospital evacuation problems with the consideration of threat and transportation risks involved. In Chapter 3, we have developed an integer programming model for the evacuation transportation problem that minimizes the expected evacuation risk subject to limited resources and patient care requirements. The model generates evacuation plans by scheduling the transportation of patients along with the allocation of evacuation resources, including a limited fleet of vehicles and receiving beds, to patients. The combination of threat risk, that depends on the type of disaster, and transportation risk, that depends on the patient-vehicle assignment and the travel time, is considered as the cost for each evacuation decision. These risk parameters are used to determine the optimal transportation plan given a set of limited resources. We demonstrated the structural properties of the model that enable computational efficiency for solving large-scale problems for both strategic planning and operations purposes. We, then, used a case study based on the evacuation of a large hospital for numerical analysis and showed that risk-minimizing solutions outperform solutions that minimize the evacuation duration in terms of both risk and duration. We illustrated the patient prioritization behavior of the model given different threat risk scenarios and different fleet scenarios and observed that the order of evacuation for patient types and the assignment of patient types to vehicle types are consistent with the relative cumulative risks of each decision. We also demonstrated the robustness of the risk-minimizing solutions under risk uncertainty. Being the most difficult parameters to estimate, risk parameters of forecasted functional forms are used to generate the optimal solutions and these solutions are evaluated using the actual functional forms. We observed that even the optimal solutions based on inaccurate forecasted risks do not deviate significantly from the optimal solution based on actual risks and still outperform duration-minimizing solutions.

In Chapter 4, we have expanded the transportation problem to multiple evacuating hospitals where resource management and equity issues arise. We developed an integer programming model that minimizes the total evacuation risk for multiple hospitals assuming the resources are centralized and distributed by a central authority, although evacuating hospitals may be in different hospital management groups. We demonstrated that the risk-minimizing utilitarian solutions are not necessarily equitable even under conditions that allow for a perfectly equitable resource allocation. In addition to the centralized strategy, we adopted the decentralized resource management strategy to determine the evacuation performance of each hospital management group by itself without sharing its set of resources with other groups. In order to produce evacuation plans with equitable resource allocation, we proposed an equity modeling framework that is based on welfare economics. The social welfare function depends on utility functions for the players among which resources are distributed and the non-negative inequity aversion parameter, α. Since they refer to explicable risk-aversion behavior of the decision maker, we especially considered three specific values of α that represent utilitarian, proportional fairness, and max-min fairness criteria. We focused on equity from the perspectives of evacuating hospitals and individual patients and define their utility functions. At the hospital-level, we demonstrated that the utility function that is complementary to risk outperforms the benchmark-based utility function that takes the decentralized performance as basis, because it minimizes the range of average patient risks for hospitals as the decision maker gets more inequity-averse. At the patient-level, we used the utility function complementary to risk and showed that the cost of equity among individual evacuees is higher than the cost of equity among hospitals.

In Chapter 5, we developed an integer programming formulation that integrates the building evacuation and transportation problems. This model is structurally complicated due to the dependencies of decision variables representing the movement of patients to the staging area and the transportation of patients to receiving hospitals. We presented the structural properties of the model that allow for its tractability. Then, we investigated the properties of the optimal plans in terms of resource utilization, optimal resource levels, and patient prioritization. We observed that sensitivity analysis of the model to changes in resource levels should consider a combination of resource level perturbations to comprehend the interactions between different resource groups in the system. We also observed that the model does not necessarily follow a specific priority rule among patient types and illustrated different patient prioritization schemes that appear in optimal solutions.

We have addressed hospital evacuation problems with logistical challenges in this dissertation. Our research contributes to a better understanding of the complexities of hospital evacuations and the optimal evacuation transportation plans. The research in this dissertation can be expanded in various directions. By the nature of the problem, several inputs to the model, that have been assumed to be deterministic in this dissertation, can be stochastic, including, but not limited to, risk parameters, capacity of receiving facilities, travel times, and fleet size. Therefore, we suggest studies of hospital evacuations with stochastic elements. We also note that simulation studies can be carried out for validation of hospital evacuation models. Although the models we have proposed have desirable structures that allow for tractability, additional interventions and solution algorithms can be necessary as the problem size is increased and new features are added to the models. Finally, we note that the evacuation operations inside the hospital building can be extremely challenging due to structural restrictions and we suggest that studies focusing on the hospital building evacuation will significantly contribute to this area of research.

Bibliography

Abkowitz, M. and Cheng, P. (1988). "Developing a risk/cost framework for routing truck movements of hazardous materials." *Accident Analysis and Prevention*, **20 (1)**, pp. 39–51.

Allison, P. (1978). "Measures of inequality." *American Sociological Review*, **43**, pp. 865–880.

Altay, N. and Green III, W.G. (2006). "OR/MS research in disaster operations management." *European Journal of Operational Research*, **175 (1)**, pp. 475–493.

Argon, N., Ziya, S., and Righter, R. (2008). "Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents." *Probabability in the Engineering and Informational Sciences*, **22 (3)**, pp. 301–332.

Atkinson, A. (1970). "On the measurement of inequality." *Journal of Economic Theory*, **2 (3)**, pp. 244–263.

Auf der Heide, E. (2003). "Convergence behavior in disasters." *Annals of Emergency Medicine*, **41 (4)**, pp. 463–466.

Augustine, J. and Schoettmer, J. (2005). "Evacuation of a rural community hospital: Lessons learned from an unplanned event." *Disaster Management Response*, **3**, pp. 68–72.

Avi-Itzhak, B., Levy, H., and Raz, D. (2008). "Quantifying fairness in queuing systems: Principles, approaches, and applicability." *Probabability in the Engineering and Informational Sciences*, **22 (4)**, pp. 495–517.

Bagaria, J., Heggie, C., Abrahams, J., and Murray, V. (2009). "Evacuation and sheltering of hospitals in emergencies: A review of international experience." *Prehospital Disaster Medicine*, **24 (5)**, pp. 461–467.

Balcik, B., Iravani, S.M.R., and Smilowitz, K. (2010). "A review of equity in nonprofit and public sector: A vehicle routing perspective." In J.J. Cochran, L.A. Cox, P. Keskinocak, J.P. Kharoufeh, and J.C. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, Inc.

Barbera, J.A. and Mcintyre, A.G. (2007). *Medical Surge Capacity Handbook: A Management System for Integrating Medical and Health Resources During Large-Scale Emergencies*. U.S. Department of Health and Human Services. [Online; accessed June 20, 2013]. URL `https://www.phe.gov/Preparedness/planning/mscc/handbook/Documents/mscc080626.pdf`

Bazaraa, M., Jarvis, J., and Sherali, H. (2005). *Linear Programming and Network Flows*. John Wiley & Sons, 3rd edition.

Bertsimas, D., Farias, V., and Trichakis, N. (2012). "On the efficiency–fairness trade–off." *Management Science*, **58 (12)**, pp. 2234–2250.

Bertsimas, D., Farias, V., and Trichakis, N. (2013). "Fairness, efficiency, and flexibility in organ allocation for kidney transplantation." *Operations Research*, **61 (1)**, pp. 73–87.

Bish, D., Agca, E., and Glick, R. (2011). "Decision support for hospital evacuation and emergency response." *Annals of Operations Research*, pp. 1–18. [Online first on August 24, 2011].
URL `http://dx.doi.org/10.1007/s10479-011-0943-y`

Broz, D., Levin, E.C., Mucha, A.P., Pelzel, D., Wong, W., Persky, V.W., and Hershow, R.C. (2009). "Lessons learned from Chicago's emergency response to mass evacuations caused by Hurricane Katrina." *American Journal of Public Health*, **99 (8)**, pp. 1496 – 1504.

Burgess, J. (1999). "Hospital evacuations due to hazardous materials incidents." *American Journal of Emergency Medicine*, **17 (1)**, pp. 50–52.

Camion, P. (1965). "Characterization of totally unimodular matrices." *Proceedings of the American Mathematical Society*, **16 (5)**, pp. 1068–1073.

Chalmet, L., Francis, R., and Saunders, P. (1982). "Network models for building evacuation." *Management Science*, **28 (1)**, pp. 86–105.

Chavez, C. and Binder, B. (1995). "A hospital as victim and responder: The Sepulveda VA Medical Center and the Northridge Earthquake." *The Journal of Emergency Medicine*, **14 (4)**, pp. 445–454.

Chen, L. and Miller-Hooks, E. (2008). "The building evacuation problem with shared information." *Naval Research Logistics*, **55 (4)**, pp. 363–376.

Childers, A., Visagamurthy, K., and Taaffe, K. (2009). "Prioritizing patients for evacuation from a health-care facility." *Transportation Research Record: Journal of the Transportation Research Board*, **2137 (1)**, pp. 38–45.

Choi, W., Hamacher, H., and Tufecki, S. (1988). "Modeling of building evacuation problems by network flows with side constraints." *European Journal of Operational Research*, **35**, pp. 98–110.

Cocanour, C.S., Allen, S.J., Mazabob, J., Sparks, J.W., Fischer, C.P., Romans, J., and Lally, K.P. (2002). "Lessons learned from the evacuation of an urban teaching hospital." *Archives of Surgery*, **137**, pp. 1141–1145.

Cone, D.C., Weir, S.D., and Bogucki, S. (2003). "Convergent volunteerism." *Annals of Emergency Medicine*, **41 (4)**, pp. 457–462.

Correa, J., Schulz, A., and Stier-Moses, N. (2007). "Fast, fair, and efficient flows in networks." *Operations Research*, **55 (2)**, pp. 215–225.

CRED (2013). "Disaster trends." Collaborating Centre for Research on the Epidemiology of Disasters. [Online; accessed June 20, 2013].
URL `http://www.emdat.be/disaster-trends`

Cyganik, K. (2001). "Disaster preparedness in Virginia Hospital Center-Arlington after September 11, 2001." *Disaster Management and Response*, **1 (3)**, pp. 80–86.

Dalton, H. (1920). "The measurement of inequity of incomes." *The Economic Journal*, **30**, pp. 348–361.

de Silva, F., Eglese, R., and Pidd., M. (1996). "A simulation model for emergency evacuation." *European Journal of Operational Research*, **90 (3)**, pp. 413–419.

Duanmu, J., Taaffe, K., and Chowdury, M. (2010). "Minimizing patient transport times during mass population evacuations." *Transportation Research Record: Journal of the Transportationn Research Board*, **1 (2196)**, pp. 150–158.

FEMA (2013). "Disaster declarations by year." Federal Emergency Management Agency. [Online; accessed June 20, 2013].
URL `http://www.fema.gov/disasters/grid/year`

Fink, S. (2009, August 27). "The deadly choices at Memorial." [Online; accessed November 20, 2012].
URL `http://www.propublica.org/article/the-deadly-choices-at-memorial-826`

Ford, L.R. and Fulkerson, D.R. (1962). *Flows in Networks.* Princeton University Press, Princeton, NJ.

GAO (2006). "Disaster preparedness: Limitations in federal evacuation assistance for health facilities should be addressed." *Technical Report GAO-06-826*, Government Accountability Office, Washington, DC. [Online; accessed February 15, 2012].
URL `http://www.gao.gov/new.items/d06826.pdf`

Golmohammadi, D. and Shimshak, D. (2011). "Estimation of the evacuation time in an emergency situation in hospitals." *Computers and Industrial Engineering*, **61 (4)**, pp. 1256–1267.

Gopalan, R., Kolluri, K.S., Batta, R., and Karwan, M.H. (1990). "Modeling Equity of Risk in the Transportation of Hazardous Materials." *Operations Research*, **38 (6)**, pp. 961–973.

Gray, B.H. and Hebert, K. (2007). "Hospitals in Hurricane Katrina: Challenges facing custodial institutions in a disaster." *Journal of Health Care for the Poor and Underserved*, **18 (2)**, pp. 283–298.

Green, W.I. (2002). "Four phases of emergency management." In Electronic Encyclopedia of Civil Defense and Emergency Management. [Online; accessed February 15, 2012]. URL `https://facultystaff.richmond.edu/~wgreen/Ecd4phases.htm`

Hamacher, H.W. and Tjandra, S.A. (2001). "Mathematical modeling of evacuation problems: A state of the art." In M. Schreckenberg and S.D. Sharma, editors, *Pedestrian and Evacuation Dynamics*. Springer, pp. 227–266.

Hamacher, H. and Tufecki, S. (1987). "On the use of lexicographic min cost flows in evacuation modeling." *Naval Research Logistics*, **34 (4)**, pp. 487–503.

Hamilton, D.R., Gavagan, T.F., Smart, K.T., Upton, L.A., Havron, D.A., Weller, N.F., Shah, U.A., Fishkind, A., Persse, D., Shank, P., and Mattox, K. (2009). "Houston's medical disaster response to Hurricane Katrina: Part 1: The initial medical response from trauma service area Q." *Annals of Emergency Medicine*, **53 (4)**, pp. 505–514.

Han, L., Yuan, F., and Urbanik, T. (2007). "What is an efficient evacuation operation?" *ASCE Journal of Urban Planning and Development*, **133 (1)**, pp. 3–8.

Henry, S. (1980). "Mississauga Hospital: Largest evacuation in Canada's history." *Canadian Medical Association Journal*, **122**, pp. 582–586.

Hobeika, A. and Kim, C. (1998). "Comparison of traffic assignments in evacuation modeling." *IEEE Transactions on Engineering Management*, **45 (2)**, pp. 192–198.

Hooker, J. and Williams, H. (2012). "Combining equity and utilitarianism in a mathematical programming model." *Management Science*, **58 (9)**, pp. 1682–1693.

Jacobson, E., Argon, N.T., and Ziya, S. (2012). "Priority assignment in emergency response." *Operations Research*, **60 (4)**, pp. 813–832.

Jarvis, J. and Ratliff, H. (1982). "Some equivalent objectives for dynamic network flow problems." *Management Science*, **28 (1)**, pp. 106–109.

Kaji, A. and Lewis, R. (2006). "Hospital disaster preparedness in Los Angeles County." *Academic Emergency Medicine*, **13 (11)**, pp. 1198–1203.

Kalai, E. and Smorodinsky, M. (1975). "Other solutions to Nash's bargaining problem." *Econometrica*, **43 (3)**, pp. 513–518.

Kue, R., Brown, P., Ness, C., and Scheulen, J. (2011). "Adverse clinical events during intrahospital transport by a specialized team: A preliminary report." *American Journal of Critical Care*, **20 (2)**, pp. 153–161.

Leclerc, P.D., McLay, L.A., and Mayorga, M.E. (2012). "Modeling equity for allocating public resources." In M.P. Johnson, editor, *Community-Based Operations Research*, volume 167 of *International Series in Operations Research and Management Science*. Springer New York, pp. 97–118.

Løvås, G. (1995). "On performance measures for evacuation systems." *European Journal of Operational Research*, **85 (2)**, pp. 352–367.

Løvås, G. (1998). "Models of wayfinding in emergency evacuations." *European Journal of Operational Research*, **105 (3)**, pp. 371–389.

Maese, J. (2009). "Medical societys blueprint for a successful community response to community preparedness." *Prehospital Disaster Medicine*, **24 (1)**, pp. 73–75.

Mandell, M. (1991). "Modeling effectiveness-equity trade-offs in public service delivery systems." *Management Science*, **37 (4)**, pp. 467–482.

Marsh, M.T. and Schilling, D.A. (1994). "Equity measurement in facility location analysis: A review and framework." *European Journal of Operational Research*, **74 (1)**, pp. 1–17.

Maxwell, C. (1982). "Hospital organizational response to the nuclear accident at Three Mile Island: Implications for future-oriented disaster planning." *American Journal of Public Health*, **72 (3)**, pp. 275–279.

McGlown, K.J. (2001). "Evacuation of health care facilities: A new twist to a classic model." *Natural Hazards Review*, **2 (2)**, pp. 90–99.

Morton, D., Popova, E., Popova, I., and Zhong, M. (2003). "Optimizing benchmark-based utility functions." *Bulletin of the Czech Econometric Society*, **10 (18)**.

Nash, J. (1950). "The bargaining problem." *Econometrica*, **18 (2)**, pp. 155–162.

Palmer, D.J., Stephens, D., Fisher, D.A., Spain, B., Read, D.J., and Notaras, L. (2003). "The Bali bombing: the Royal Darwin Hospital response." *The Medical Journal of Australia*, **179 (7)**, pp. 358–361.

Pigou, A. (1912). *Wealth and Welfare*. Macmillan, London.

Pratt, J. (1964). "Risk aversion in the small and in the large." *Econometrica*, **32 (1/2)**, pp. 122–136.

Quarantelli, E.L. (1980). "Evacuation behavior and problems : Findings and implications from the research literature." *Technical report*. Disaster Research Center. Ohio State University, Columbus, Ohio.

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press, Cambridge, MA.

Savas, E. (1978). "On equity in providing public services." *Management Science*, **24 (8)**, pp. 800–808.

Schrijver, A. (1998). *Theory of Linear and Integer Programming*. John Wiley & Sons, Inc., New York, NY, USA.

Schultz, C., Koenig, K., and Auf der Heide, E. (2005). "Benchmarking for hospital evacuation: A critical data collection tool." *Prehospital Disaster Medicine*, **20 (5)**, pp. 331–342.

Schultz, C., Koenig, K., and Lewis, R. (2003). "Implications of hospital evacuation after the Northridge, California, Earthquake." *The New England Journal of Medicine*, **348 (14)**, pp. 1349–1355.

Schwartz, J. and Winship, C. (1980). *The Welfare Approach to Measuring Inequality*. Jossey-Bass, San Francisco.

Sen, A. (1973). *On Economic Inequality*. Clarendon Press, Oxford, England.

Sen, A. and Foster, J. (1971). *On Economic Inequality*. Oxford University Press, New York.

Sheffi, Y., Mahmassani, H., and Powell, W.B. (1982). "A transportation network evacuation model." *Transportation Research Part A*, **16 (3)**, pp. 209–218.

Sorensen, J. (1987). "When shall we leave? Factors affecting the timing of evacuation departures." *International Journal of Mass Emergencies and Disasters*, **9 (2)**, pp. 153–165.

Sorensen, J. and Mileti, D. (1987). "Decision-making uncertainties in emergency warning system organization." *International Journal of Mass Emergencies and Disasters*, **5 (1)**, pp. 33–62.

Sternberg, E. and Lee, G.C. (2009). "New York City's healthcare transportation during disaster: A preparedness framework for a wicked problem." *Prehospital Disaster Medicine*, **24 (2)**, pp. 95–107.

Sternberg, E., Lee, G., and Huard, D. (2004). "Counting crises: US hospital evacuations, 1971–1999." *Prehospital Disaster Medicine*, **19 (2)**, pp. 150–157.

Taaffe, K., Johnson, M., and Steinmann, D. (2006). "Improving hospital evacuation using simulation." In L. Perrone, F. Wieland, J. Liu, B. Lawson, D. Nicol, and R. Fujimoto, editors, *Proceedings of the 2006 Winter Simulation Conference*. pp. 509–515.

Taaffe, K., Kohl, R., and Kimbler, D. (2005). "Hospital evacuations: Issues and complexities." In M. Kuhl, N. Steiger, F. Armstrong, and J. Joines, editors, *Proceedings of the 2005 Winter Simulation Conference*. pp. 943–950.

Tayfur, E. and Taaffe, K. (2009). "A model for allocating resources during hospital evacuations." *Computers and Industrial Engineering*, **57 (4)**, pp. 1313–1323.

Toth, P. and Vigo, D., editors (2002). *The Vehicle Routing Problem*. Society for Industrial and Applied Mathematics.

Tufekci, S. (1995). "An integrated emergency management decision support system for hurricane emergencies." *Safety Science*, **20 (1)**, pp. 39–48.

Van Wassenhove, L. (2006). "Blackett Memorial Lecture Humanitarian Aid Logistics: Supply chain management in high gear." *Journal of the Operational Research Society*, **57**, pp. 475–489.

Vogt, B.M. (1991). "Issues in nursing home evacuations." *International Journal of Mass Emergencies and Disasters*, **9 (2)**, pp. 247–265.

Young, P. (1995). *Equity: In Theory and Practice*. Princeton University Press, Princeton, NJ.

Zenios, S.A. (2005). "Models for kidney allocation." In M.L. Brandeau, F. Sainfort, and W.P. Pierskalla, editors, *Operations Research and Health Care*, volume 70 of *International Series in Operations Research and Management Science*. Springer US, pp. 537–554.