Models to Improve Efficiency of Disaster Relief Operations

By

Merve Ozen

A Dissertation Submitted in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
Industrial and Systems Engineering

at the
University of Wisconsin-Madison
2018

Date of final oral examination: March 12, 2018

The dissertation is approved by the following members of the Final Oral Committee:

Ananth Krishnamurthy, Professor, Industrial and Systems Engineering
Jeffrey Linderoth, Professor, Industrial and Systems Engineering
Shiyu Zhou, Professor, Industrial and Systems Engineering
Gregory DeCroix, Professor, School of Business
Qunying Huang, Assistant Professor, Department of Geography

Abstract

A disaster is a sudden, calamitous event that disrupts the functioning of a community and causes human, material and economic damage that exceeds the society's ability to cope using its own resources. Each year there are about 500 disasters directly affecting approximately 200 million people across the world (Wassenhove, 2006). The unpredictability of disasters and the challenges in the response to these disasters has led to an increase in research on humanitarian logistics and disaster operations management. Disaster operations management research encompasses four phases: mitigation, preparation, immediate response, and reconstruction phase. Mitigation phase focuses on reducing possible risk or impact of a disaster, preparation phase involves planning to cope with the disaster consequences, immediate response phase focuses on operations to help relieve the immediate impact of the disaster and the reconstruction phase encompasses activities to help restore the community to its original state. This thesis focuses on the challenges faced in the *immediate response phase*.

The immediate response phase is the most challenging phase where relief operations including search and rescue efforts, relief item distribution to victims, and provision of medical aid take place. Relief item distribution to disaster victims include providing food, water, blankets, tents for the victims that have suffered from the disaster. This key activity has an important role in relieving human suffering caused by the disaster. Our interviews with Salvation Army, Red Cross, South East Wisconsin Citizens and Organizations Active in Disasters (COAD) practitioners and study of the literature suggest, there are opportunities for increasing the efficiency of relief distribution. Prior research on immediate response has largely focused on the logistics component of procuring and transporting the required items to the

disaster region. In comparison, there has been very little work on how to efficiently distribute the relief supplies once the items arrive at the disaster site. This thesis focuses specifically on the problems related to the efficient distribution of relief items once they arrive at the disaster site, which occurs at the very end of the "last mile". We define efficiency in relief item distribution as a combination of the following metrics; the reduction in waiting time experienced by the victims to receive aid, maximizing the total number of victims that are served, and the effective usage of available space and volunteer capacity.

In this research, we first focus our attention on analyzing relief center (RC) designs, to control the crowd and minimize the waiting time of victims that queue up to receive aid. A relief center is composed of points of distribution, each staffed with volunteers distributing relief items. We model the flow of victims through the relief center and effects of congestion using finite capacity, state dependent queuing networks. We provide analytic expressions to estimate the performance measures of this queuing system and compute the expected deprivation for the target community. Using this model, we first analyze RC operations based on current guidelines for relief centers. Then we propose alternative relief center designs based on the observed opportunities for improvement.

Next, we focus on problems across relief centers caused by imbalances in work load due to supply and demand uncertainties. We model a network of relief centers distributing aid in an affected region, as a generalized queuing network (G-network). The G-network structure allows to model probabilistic flow of victims between RCs. G-networks with single probabilistic victim movements and signals have been shown to have product form results, however for complex cases including batch transfers

product form solutions are not guaranteed. We derive conditions under which G-networks with batch transfer have product form solutions. We leverage this result to develop product form approximations that apply to settings with varying batch sizes and imbalances in queue utilization. We use these approximations to estimate the performance of the distribution operations.

Next, we focus on quantifying material convergence, a problem that commonly occurs in the immediate response phase. It refers to the convergence of relief related items, from many sources, including purchased goods, pre-positioned stock, and in-kind donations into the disaster affected region, in very high volumes, over a short period of time. We study how relief agencies should effectively manage available resources to maximize relief distribution to victims under material convergence. To quantify material convergence, we model both solicited and unsolicited donation arrival-sorting processes as a queuing process. We use transient analysis to quantify the level of material convergence and evaluate the impact of resource allocation decisions on relief item output. We provide insights on the dependent relationship between convergence and resource allocation through numerical studies based on Goodwill and Salvation Army data and inputs (Ozen and Krishnamurthy (2017)).

Next, we focus our attention to conducting a transient analysis of performance for relief centers. We relax the assumptions related to the probability distributions governing arrival and service processes, and analyze the relief distribution operations using a transient model instead of a steady state model. We use discrete event simulation to investigate the effects these modeling assumptions had both on the estimated performance measures and on the insights for relief center designs.

Finally, we conduct a simulation study to search for the optimal relief center design that minimizes victim waiting times and maximizes the throughput from a relief center. We enumerate several design options for a set of resource constraints and use discrete event simulation to evaluate each design. We study both the optimal design, as well as the performance difference between designs. We provide insights on key design parameters to achieve optimal performance during relief distribution.

This research provides contributions to both theory and practice. First, we propose queuing models to represent the inherent queuing effects that impact the efficiency of disaster relief operations. We develop efficient solution methodologies to solve these queuing models and determine operational insights. Finally, we use discrete event simulation models to investigate issues that are difficult to evaluate using queuing models. Where possible, we use insights from interviews with the Salvation Army, FEMA, International Federation of Red Cross and Red Crescent Societies (IFRC), National Voluntary Organizations Active in Disasters (NVOAD) and South East Wisconsin Citizens and Organizations Active in Disasters (COAD) and data from disaster response efforts to 2015 Nepal earthquake.

Acknowledgments

I would like to acknowledge and thank my research advisor, Ananth Krishnamurthy and all the committee members, Jeffrey Linderoth, Shiyu Zhou, Gregory DeCroix and Qunying Huang for their guidance, feedback and valuable comments. In addition to the support of the faculty, I was very fortune to be part of the Center for Quick Response Manufacturing and have colleagues that supported me including Tugce Martagan, Ashesh Sinha, and Yasemin Limon. Their guidance and friendship provided me with tremendous help over the years.

I have the greatest gratitude and love for my parents, Umit Ozen and Suat Ozen. They showed me the value of hard work, they gave me an equal voice from a very young age and supported all my aspirations in life with many sacrifices. I love both of you with all my heart and thank you for standing by me through the ups and downs of life and the PhD program.

I feel so lucky to have my partner, Kevin Barnett who was the greatest support system I could have ever asked for. He was right beside me through it all and he has always been my rock! You bring joy to my life and make me happy every single day! I deeply love and appreciate you, beyond words can explain.

I dedicate this thesis to everybody that has ever been affected by adversity, a personal or a natural disaster. My greatest hope is my thesis and my future work touches lives to make them better. I am reminded everyday that being kind to each other is the most important of all, and I thank everybody that has been kind to me along the way and made this thesis possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Each year disasters effect millions of people worldwide and effective disaster operations management is essential to minimize the suffering of communities impacted by these disasters. Disaster operations management (DOM) can be categorized into 4 phases: mitigation, preparedness, response and recovery (McLoughlin, 1985). The mitigation phase focuses on eliminating or reducing the possible impacts and risks of disasters, which happens before a disaster occurs. The preparedness phase also happens prior to the disaster, and focuses on finding ways to effectively cope with the consequences of the disaster. It includes actions such as setting coordination structures, planning evacuation routes and pre-positioning relief items. The response phase focuses on saving lives and providing relief during the immediate aftermath of the disaster. The preparedness and response phases are tightly coupled, however the response phase is more complex as demand, supply and resource availability changes dynamically. The final phase of DOM is the recovery phase, which can last years depending on the level of the catastrophe. In this phase the community slowly returns back to a stage where it can sustain itself. This thesis focuses on the response phase (See Figure 1.1).

Figure 1.1: Phases of Disaster Operations Management

Prior work on DOM has a heavy focus on the preparedness stage. The important operational questions in the preparedness phase include how to plan for critical supply distribution, how to plan for evacuation activities, where to position relief organization facilities, and how to coordinate efforts during the response. Once the disaster hits, the response phase needs to bridge the gap between the planning and real time needs and effective resource allocation, donations management, coordination and crowd control become priorities.

The rest of the chapter is organized as follows. In Section 1.1 we discuss challenges of the disaster response phase, identify research needs and summarize the research questions that are the focus of this thesis. In Section 1.3 we discuss our research approach and highlight our research contributions.

1.1 The Disaster Response Phase

The response phase is the most challenging phase of disaster operations management. In this phase, search and rescue operations, first aid medical attention, distribution of critical supplies and restoration of critical infrastructure activities take place. This thesis focuses on the distribution of critical relief supplies to victims.

The relief supplies go through a multi-tier supply chain (the relief distribution supply chain) where they are procured, transported through multiple channels and distributed at the disaster site. Figure 1.2 shows the relief chain and the associated problems with each tier of the supply chain.

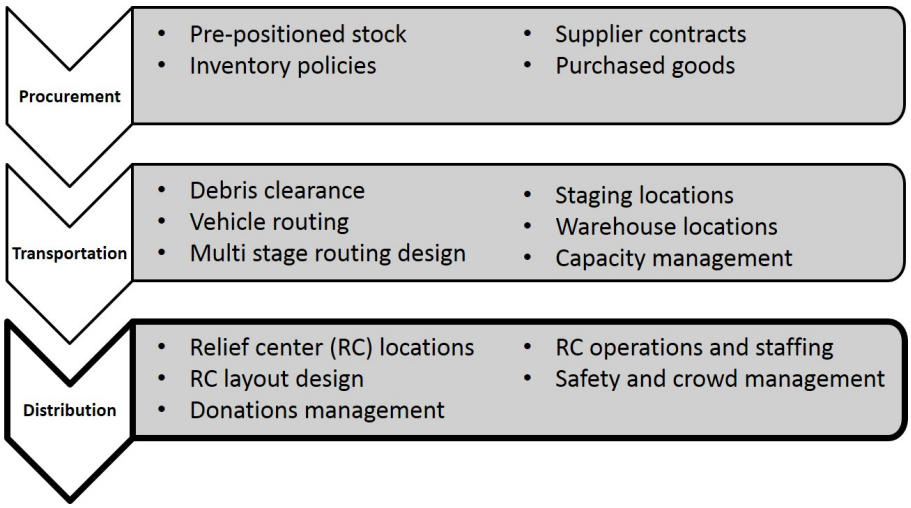

Figure 1.2: The Relief Distribution Supply Chain

The relief distribution supply chain has 6 key differences when compared to commercial supply chains:

1. The primary objective of commercial operations is to maximize profits (or minimize costs) while the objective of relief distribution supply chain is to minimize human deprivation.

2. The number of steps in the supply chain and the number of parties involved is higher for the relief distribution supply chain.

3. In commercial operations, resources are constrained (since they have associated costs), in disaster operations resources are constrained by cost and limited in availability.

4. Commercial operations have fewer decision makers than disaster operations.

5. The transportation needs of commercial logistics are considerably less while disaster relief items can block international airports due to the sudden and high volume of need.

6. Commercial logistics operate when all infrastructure is intact and operational, for disaster relief operations this is not always the case.

In the light of these key differences, the models addressing relief distribution supply chain problems should take into account appropriate objective functions, limited resources, congestion created by crowds and materials and damaged infrastructures. This is challenging for modeling due to the need to consider many aspects including (i) uncertainty in demand, supply, capacity, (ii) the many players (agencies, governments, companies, individuals) involved in the effort, (iii) modeling resource allocation and coordination aspects, (iv) modeling effects of crowds and congestion (both for materials and victims) and (v) modeling the dynamically changing nature of the response environment. In addition to constructing a versatile model, there is also the challenge of providing analytical solutions for these models. Therefore research on new solution methodologies is a pressing research need.

1.2 Challenges in Disaster Relief Distribution

One of the first steps in relief distribution is to set up the relief centers and deliver aid. Relief centers (RC) are setup by relief agencies, where victims arrive and wait in line to receive the supplies they need. The RCs are usually setup in large, open areas

to allow for crowd access, while the particular setup of an RC can vary depending on available staff, number of items being distributed and the expected crowd size and safety concerns. Figure 1.3 shows a relief center and the queuing and crowding that is often observed during relief distribution efforts. As seen from Figure 1.3, setting up the relief center so that crowds are effectively managed to ensure timely distribution and safety is a major challenge. In addition, even when the queues are orderly, waiting times can still be of major concern. Therefore, utilizing the RC setup and staffing so that the waiting times are minimized is a challenge faced during every relief distribution. Other operational problems are related to, managing a network of relief centers, and managing the in-kind donations in a disaster affected region. For all these problems, estimating the relief center performance (the expected throughput and victim waiting times) accurately under different decision variables is critical and can be challenging given the dynamically changing variables.

Figure 1.3: Distribution via Relief Centers

In this thesis we aim to model these challenges and provide insights on best practices and trade-offs for managing both the relief items and their distribution. Next we discuss each particular challenge in more detail and summarize the research questions this thesis focuses on.

Challenge 1: RC Layout and Congestion

To further understand the specific problems faced during relief distribution on the field, we interviewed practitioners and organized working group sessions with representatives from the Salvation Army, International Federation of Red Cross and Red Crescent Societies, Wisconsin Division of Emergency Management, Dodge County Office of Emergency Management, Federal Emergency Management Agency (FEMA) and the Disaster Management Center at UW Madison. The interviews revealed that established relief agencies including FEMA, Red Cross and the Salvation Army all have guidelines and procedures for setting up relief centers and distributing relief (Federal Emergency Management Agency (2008)), USAID (United States Agency for International Development (2005)) and IFRC (International Federation of Red Cross and Red Crescent Societies (2008)). However, they emphasized that even though the guidelines present a starting point, there are limited decision support tools to help agencies quickly adapt to the unique needs of the situation. Moreover, for the disasters where the affected population is large, how best to control the crowd and increase efficiency of distribution operations simultaneously becomes a pressing question. This leads to the following research questions:

RQ1: How can an RC be setup to control effects of crowding, increase throughput and decrease victim waiting times?

RQ2: Can operational strategies like triaging help improve the RC performance?

Uncontrolled Crowd

Figure 1.4: From the Salvation Army Port-au-Prince Relief Distribution Efforts

Figure 1.4 shows the effects of crowding at the Salvation Army relief center distribution area. To answer these research questions we analyze different relief center designs including the current practice design adapted from FEMA documentation and proposed designs to disperse and control crowds. In our models, we explicitly consider how excessive crowds slow down the movement of victims and the distribution efficiency, as well as create bottlenecks in the distribution. We discuss the details of the model, our solution methodology and the associated numerical experiments in Chapter 3.

Challenge 2: Dynamic Changes in Victim Needs and Behavior

The disaster relief efforts in a given region involves multiple relief agencies and multiple relief centers distributing various relief supplies. The displaced victim population in the region, based on their needs can wait at any RC to receive the supplies they need. Therefore, the relief distribution in a given region can be represented as a network of RCs. In this network, both the victim's needs and the disaster

environment conditions can dynamically change. This creates victim mobility in the network and varies arrival rates experienced by different RCs, which can lead to varying degrees of congestion in the network, effecting the performance of the overall relief distribution.

Efficiency of the overall distribution network is critical for minimizing the deprivation experienced by victims. In order to be efficient, the execution of response activities need to be agile and respond well to the inherent variability of the disaster aftermath conditions. Operations research (OR) models can help identify opportunities for improvement in efficiency of response activities, as well as serve as decision support tools. The success of such models rely on their ability to capture the described stochastic nature of disaster response environment, which includes various victim behavior and mobility patterns. Example scenarios that can cause victims to move or transfer between relief centers are:

- Increase in demand: An example would be the weather getting colder can increase the overall demand for blankets, causing more victims to move to RCs that are distributing blankets. Multiple demand shifts for multiple items can be experienced simultaneously and over different periods of time.

- Changes in victim needs: A victim waiting in line to receive water, can quit this queue and transfer to another RC to wait for receiving tarpaulins. This type of behavior usually happens due to different family members queuing to receive different supplies.

- Victim perception: Victims can have different perceptions of supply availability, congestion or waiting times at RCs and move between queues due to their perception. This does not always coincide with reality. However, it can create

significant victim mobility.

These probabilistic victim behaviors, coupled with multiple relief centers distributing multiple items, with varying levels of capacity, make operating the relief distribution efforts a challenging issue. These challenges lead to the following research questions:

RQ3: What is the effect of victim movement between relief centers on the performance of the relief distribution network?

To answer this question, we develop a framework that can take into account, probabilistic changes in demand for items and probabilistic victim movements using a generalized queuing network. We describe the details of the model, the proposed solution methodology as well as the numerical experiments in Chapter 4.

Challenge 3: Material Convergence

Immediately after the disaster; relief items, from many sources such as, purchased goods, pre-positioned stock, and in-kind donations flow into the disaster affected region, in high volumes and over a short period of time. This phenomenon is known as material convergence. Among the sources that contribute to the material flow, in-kind donations make up the bulk of the supplies received and handled during an emergency. When these donations include high proportions of low-priority items, they can exasperate the complexity of donations management and divert critical resources away. We will refer to this problem of high volumes of incoming donations coupled with high percentages of low-priority items as the material convergence problem.

Solicited Donations (Haiti Earthquake Response)

Unsolicited Donations (Hurricane Sandy Response)

Figure 1.5: Solicited versus Unsolicited Donations

Relief agencies including Salvation Army and the Red Cross purchase items or solicit donations of specific items to meet the needs of victims. Unsolicited donations from individuals or groups also get sent to the disaster affected region. These unsolicited donations can contain high-priority (HP) items that are urgently needed for distribution, low-priority (LP) items which are not urgent but can be useful for relief agencies and non-priority (NP) items which are not at all needed. In many disasters, the percentage of NP items for unsolicited donations can exceed 50%.

Figure 1.5 shows examples of solicited and unsolicited donations from the Haiti earthquake and hurricane Sandy. As seen from these pictures, sorting unsolicited donations is often a more complex activity with less return. However, dedicating resources to sort unsolicited donations is usually a must due to the high levels of demand that solicited donations and purchased goods alone can not satisfy. Therefore, the decision of how to allocate resources between solicited and unsolicited donations is a big challenge. This decision becomes even more challenging as the arrival rates for both types of donations change over time and the expected levels of high-priority good percentages for unsolicited donations vary.

The challenges identified led to the following research questions:

RQ4: How can we quantify material convergence?

RQ5: How does the mix of high priority and non priority items and arrival patterns impact material convergence and volunteer assignment decisions?

To answer these questions we develop analytical queuing models to quantify material convergence. This is a challenging analytical problem due to its transient nature. Thousands of donations of either kind can arrive by the hour and pile up, and the arrival rates, available resources and the allocation of resources change over time. To solve this problem we use numerical methods that increase the solution efficiency. We discuss the details of the model and the solution approach in detail in Chapter 5.

Challenge 4: Transient Analysis of Relief Center Performance

The disaster relief distribution setup and operations take place during a time of great urgency, uncertainty and need. Hence, both accuracy of evaluating options for relief center designs and doing so in a timely fashion is critical. Analytical models, such as the one proposed in Chapter 3 are very computationally efficient compared to extensive simulation models. However, to get analytical results, these models rely on key assumptions. The model developed in Chapter 3, assumes exponential distributed inter-arrival and service times and the performance measure estimates reflect steady state measures. We conduct a transient analysis of relief center performance that relaxes these assumptions and answers the following research question:

RQ6: What additional insights on RC performance can be obtained using transient

analysis that relaxes the Markovian assumptions?

To answer this research question, we model the relief center designs using Arena® discrete event simulation. We describe the details of the simulation models and numerical experiments investigating transient behavior of RC operations in Chapter 6.

Challenge 5: Optimizing the Relief Center Designs

Disaster relief distribution is a critical activity during disaster response that helps reduce the suffering imposed on communities. Increasing the relief distribution efficiency and quantifying impacts of crowds and changes in relief center design is an important first step. However, using the available resources to achieve the best possible relief distribution performance is a critical next step.

In relief center design, the availability of resources such as the number of volunteers and the size of the relief center area may be difficult to control. However, given these resource constraints, the challenge is setting up the relief center for optimal performance. This challenge leads us to the following research question:

RQ7: What is the optimal relief center design under resource constraints?

To answer this research question, we enumerate several relief center designs given constraints on resources. We then use Arena® discrete event simulation to evaluate the designs in the design space and find the optimal relief center design. We describe the details of constructing the design space, the simulation models and numerical

experiments in Chapter 7.

1.3 Research Approach and Contributions

Our research approach aims to integrate theory and practice by grounding the research problems in challenges faced by the disaster responders in the field, developing stochastic models, and integrating disaster response data with stochastic models to provide managerial insights for disaster response managers.

The research questions in this work were all motivated by discussions, interviews and working group sessions conducted with Red Cross, Salvation Army, National Voluntary Organizations Active in Disaster, Goodwill Industries, the Disaster Management Center and Wisconsin Division of Emergency Management. The practitioners we worked with had experience responding to disasters both nationally and internationally including the Haiti earthquake, Hurricane Sandy and Hurricane Andrew.

To answer the research questions, we developed analytical and simulation models that captured the effects of crowds, queuing of victims, dynamically changing victim needs, uncertainties in demand and supply, convergence of materials, limited capacity and deprivation experienced by victims. We collected data on Nepal earthquake relief distribution and Goodwill Industries donations management process. We conducted numerical experiments using the developed models and the data collected to gain insights into improving relief distribution operations. Through our practically grounded research approach, this work makes several contributions both

to the theory and practice. We hope that the findings of this research will impact how disaster response activities are planned and conducted.

In Chapter 3, we model relief center setup and operation practices. We model each point of distribution (POD) and the access walkways to the PODs within the RC as finite capacity queues with state dependent service rates. We derive analytic solutions to estimate the performance measures of the queuing network, taking into account blocking probabilities and state dependent service rates that model congestion. We analyze the trade-off between the length of queues victims need to go through and the congestion of the queues (RQ 1). We observe that congestion is a key driver in increasing victim waiting times and propose relief center designs that leverage crowd dissipation strategies, including alternative uses for the triage queue at the RC (RQ 2). We also compare the current practice layout used by FEMA and alternative designs leveraging the insights obtained from the model.

In Chapter 4, we model the relief distribution effort in a disaster affected region as a generalized queuing network (G-network), where we model the probabilistic victim movements between points of distribution and agency interventions that can redirect victims between RCs using routing probabilities and signal entities. To estimate performance measures for the relief centers under these dynamic changes, we provide a new product form result for G-networks with full batch transfer of victims via signals. We also propose product form approximations for G-networks with partial batch transfer of victims. We apply the proposed G-network model and quantify the impact of victim mobility on relief distribution performance (RQ 3).

In Chapter 5, we model the donation arrival-sorting process as multi-server

queues and analyze the transient evolution of the material convergence phenomenon. To our knowledge this work is the first in the literature to aim at quantifying material convergence (RQ 4). We use data from Goodwill to understand the donation sorting process and conduct numerical studies on volunteer assignment decisions under varying levels of high-priority percentage of unsolicited donations and changing arrival patterns (RQ 5).

In Chapter 6, we use discrete event simulation to model relief center designs. We use the simulation models to relax assumptions made in analytical models. In specific, we investigate the effects of Markovian assumptions and steady state measures on performance measure estimates and insights related to relief center setup and operations (RQ 6).

In Chapter 7, we focus on optimizing the relief center design for both throughput and victim waiting time measures under resource constraints. For a given level of available space and volunteers, we enumerate all possible relief center designs and exhaustively analyze each design using discrete event simulation. We comment on the optimal design (RQ 7) and the variation in performance between different designs.

The analytic models and results from this research can be used to create decision support tools that can evaluate, compare, optimize relief distribution related decisions.

Chapter 2

Literature Review

In this chapter we present an overview of the literature related to disaster relief operations. In the literature review, we discuss prior work in relief center design, relief distribution networks and coordination, and material convergence areas and use our discussion to position the contributions of this research.

2.1 Related Work on Disaster Preparedness

Disaster preparedness involves: (i) planning for relief item distribution, (ii) planning for a possible evacuation, (iii) deciding on facility and warehouse locations and minimizing transportation costs between facilities. Preparedness efforts start locally in counties and states but extent nationally and internationally. In the United States, Department of Homeland Security (DHS) prepared the National Preparedness Goal Report Department of Homeland Security (2011) to identify core capabilities required for responding to a disaster. Each year the National Preparedness Report is published to assess the progress on the goals set for preparednessDepartment of Homeland Security (2012). Technical reports by Royal et al. (2014); Executive Of-

fice of the President of the United States: Subcommittee on Disaster Reduction (2005) provide more detailed information on specific technology and strategy needs. Next we survey the preparedness literature in detail.

(i) Planning for Relief Item Distribution: Efficient distribution of relief items require pre-positioning supplies, choosing key suppliers, reserving supply capacity and planning transportation capacity and transportation routes. Balcik and Ak (2014) discuss supplier selection and supplier aggrements for relief supplies. Akkihal (2006) investigates inventory pre-positioning of non-consumable items for humanitarian operations, Wang et al. (2015a) looks at pre-purchasing decisions using an options contract framework for humanitarian supply chains. Adida et al. (2011) focus specifically on stockpiling at hospitals for disaster planning and Duran et al. (2011) discuss pre-positioning of relief items for CARE International. Acimovic and Goentzel (2016) investigate stock piling decisions for non-food items including blankets, buckets, mosquito nets based on stochastic optimization models to investigate the effect of making combined stock-piling decisions. Sodhi and Tang (2014) focus on integrating local micro-retailers while solving the pre-positioning problem to support local economies affected by disasters.

Other papers study the inventory management problem related to disasters. Clay (2007) discusses the challenges related to inventory management following a disaster. Das and Hanaoka (2014) and Ozbay and Ozguven (2008) present stochastic inventory replenishment models for inventory control. Natarajan and Swaminathan (2014) construct a multi-period stochastic inventory model with financial constraints to investigate the impact funding uncertainty on relief item inventories. Wu et al. (2010) combine inventory management with a demand forecasting approach to im-

prove the forecasts using similar case based reasoning to improve the emergency supplies management system. Xu et al. (2010), Deqiang et al. (2011), Zhao and Cao (2015), Sheu (2010) and Li (2010) discuss forecasting approaches applicable in a disaster setting. Taskin and Lodree (2011) use forecasting data for hurricanes to manage emergency supply.

(ii) Evacuation Planning: Issues related to evacuation of victims has been the subject of some studies. Tanaka et al. (2006) analyze traffic congestion effects for hurricane evacuation, Fang et al. (2011) analyze possible evacuation routes for a stadium and Hobeika et al. (1994) propose an evacuation decision support tool for evacuation planning in nuclear power station regions. Uno and Kashiyama (2008) integrate geographical information into a simulation model to investigate both damage and appropriate evacuation routes. Chiu et al. (2007) use a dynamic traffic flow optimization model to establish an evacuation network. Simonovic and Ahmad (2005) simulate a flood emergency evacuation taking into account certain human behavior characteristics and evaluates effectiveness of different flood emergency management procedures. Pidd et al. (1996) also use a simulation approach and describes a spatial decision support system to be used by emergency planners. Chen et al. (2012) develop a model for assessing risks associated with the evacuation process including both pre and post disaster factors. Aksu and Ozdamar (2014) focus on road restoration efforts and use a mathematical model to maximize network access for efficient evacuation of victims. Lo et al. (2006) use a game theory approach to model exit choice of victims evacuating a building.

(iii) Facility Location and Transportation: Transporting supplies from the stock points to the last mile distribution points usually involves a multi-tier supply

chain which includes facilities (warehouse, staging area) and transportation between these facilities and the final distribution points.Jia et al. (2007) provide a survey for facility location problems and proposes a model specifically for large scale emergencies. The proposed model can be cast as a generalization of the covering, P-median, and P-center models taking into account characteristics of large-scale emergencies. Balcik and Beamon (2008) analyze a distribution network design problem, which combines facility location and stock pre-positioning decisions. Hong et al. (2015) use a stochastic model to determine the size and the location of the response facilities and the inventory levels of relief supplies at each facility. Cho et al. (2014) look specifically at deciding locations for trauma centers and helicopters using a mathematical modeling approach and propose a specialized algorithm. Iakovou et al. (1996) focus on finding the optimal location for cleanup equipment for responding to possible oil spills using a linear integer programming model while Batta and Mannur (1990) study a covering-location problem for ambulances and fire trucks. Barbarosoglu and Arda (2004) model the transportation of first-aid commodities during a disaster as a multi-commodity, multi-modal two-stage stochastic program. Campbell et al. (2008) analyze routing problems relevant to disaster relief operations using the objective of minimizing arrival times. Manoj et al. (2016) combine decision for locating staging areas, assigning inventory to staging locations and routing trucks from staging areas to distribution sites. Han et al. (2011) use an optimization method to decide on warehouse locations, fleet routing and scheduling.

2.2 Related Work on Disaster Response

Wright et al. (2006) provides a survey of operations research (OR) literature on

disasters and emergency management, highlighting the need for more research in the disaster response phase. Larson et al. (2006) discuss lessons learned from 5 major emergencies and highlight the needs for more research related to emergencies. Larson (2004) provides a discussion on OR models in emergency management and homeland security related problems. Altay and Green (2006) and Galindo and Batta (2013) provide literature surveys in disaster operations management and highlight the need for models that take into account unique characteristics of the disaster environment. They also state that aid collection, allocation, and distribution problems as areas that need attention from the OR community. Simpson and Hancock (2009) survey emergency response articles and highlight appropriate performance measures and volunteer management for disaster operations as future research directions. Kovacs and Spens (2007) and Jiang et al. (2012) survey logistics issues in humanitarian relief operations and emphasize the need for additional research on stochastic models that use appropriate objective functions. Caunhye et al. (2012) surveys the optimization models in disaster operations, particularly related to relief distribution.

Next we discuss the literature related to (i) relief distribution at relief centers, (ii) relief distribution networks and coordination and (iii) the material convergence problem.

(i) Relief Distribution at the Disaster Site: Once the relief items are procured and transported, they are distributed to the victims at the affected region via distribution areas staffed by volunteers called relief centers (RC). The location of the RCs, the design of the RCs, its staffing and operational policies are all problems that require attention to make this process efficient. To the best of our knowledge,

Roy et al. (2011) is the only prior work in the literature that focus on relief supply distribution at the disaster site. The authors compare various relief center designs using queuing models.

(ii) Relief Distribution Networks and Coordination: To our knowledge, there is no prior work that considers a network of relief centers for relief item distribution. The closest line of work is related to coordination of the immediate response activities. Coordination both at the strategic level and the operational level has been identified as an area that needs further research (Simpson and Hancock (2009)). The literature has analyzed costs and benefits of inter agency coordination, coordinating procurement of supplies and coordinating transportation of supplies (Balcik et al. (2010)), Ergun et al. (2014), Tomasini and Van Wassenhove (2009), Altay and Pal (2014), Schulz and Blecken (2010), Wang et al. (2015b)). To our knowledge, operational coordination and strategies to deal with changing victim need and mobility for disaster relief distribution has received very little attention in the literature.

(iii) Material Convergence: Prior literature on humanitarian logistics and disaster operations management has recognized material convergence as a catastrophic and understudied problem in disaster response. Many case studies and empirical papers document the occurrence of material convergence and the strain it created on critical resources following disasters. Arnette and Zobel (2015) document material convergence during the 2013 flooding in Colorado. Thomas and Fritz (2006) document the convergence issues in the 2004 earthquake and tsunami in Southeast Asia. Holguin-Veras et al. (2007) discuss material convergence related problems in Hurricane Katrina. Holguin-Veras et al. (2012b) identify material convergence issues that occurred during the 2011 Tohoku earthquake in Japan. Holguin-Veras et al. (2014)

provide examples of material convergence problems that arose during 1953 Arkansas Tornado, 2004 Hurricane Charley, 2001 World Trade Center, 2010 Colombia Floods, 2011 Joplin Tornado, 2012 Hurricane Sandy and many more. Holguin-Veras et al. (2012a) provide an excellent discussion on material convergence and state that there is no established procedure for simultaneously handling the material convergence problem and expediting the flow of high-priority items to disaster survivors. They emphasize that this is mainly due to the lack of analytic models for material convergence in the literature. They specifically highlight the need for dynamic models to allocate resources to manage material convergence. To our knowledge, the closest line of research related to quantifying material convergence has been by Martelo (2011) where they use a mathematical programming model to determine optimal allocation of resources. However, the paper assumes that there is no accumulation of flows between time periods and therefore does not quantify the actual material convergence. In addition, the optimization model is based on estimating many cost parameters like, volunteer allocation and volunteer re-allocation costs, opportunity cost of rejecting donations and the delay cost, that might be hard to estimate in practice.

Other papers in the literature focus on the donors and donor types. For instance, Ülkü et al. (2015) and Destro and Holguin-Veras (2010) both analyze the socio-economic characteristics of donors to understand the expected donation flows more accurately. Stapleton et al. (2010) provide an overview of types of donors and discuss the pros and cons of different types of donations. Islam et al. (2013) discuss the use of donation registry or donation portals for matching donors with actual need. Ozpolat and Rilling (2015) and United States Agency for International Development (2016) focus on donor education to raise awareness regarding unsolicited

donations.

In addition, there have been studies related to the relief chain, donations and material convergence. Gatignon et al. (2010) compare centralized and decentralized supply chain operations to increase responsiveness and state reduced material convergence as a positive effect of decentralization. Aflaki and Pedraza-Martinez (2016) analyze the effects of different funding strategies and specifically focus on earmarked financial donations.

2.3 Perspective on the Literature

In disaster response literature, distributing relief items at the disaster site received minimal attention. Most of the research related to disaster relief items, focuses on the preparedness stage and studies problems related to pre-positioning supplies, supplier agreements, supply chain design, facility location problems and inventory management problems. The problems that disaster response agencies face in the field, while managing and distributing the relief supplies that arrive is a major gap in the literature. In addition, many surveys of disaster operations management (DOM) literature identified quantitative models, stochastic and dynamic features in modeling and appropriate objective functions as major needs for all DOM research. We aim to fill these gaps in the literature.

We first focus our attention on relief center design at the disaster site. To our knowledge there is only one prior work (Roy et al. (2011)) highlighting the need to understand trade-offs in relief center design. Our work differs from the prior

literature in the following ways: (1) We analyze the current practice on relief center design and propose alternate RC designs to improve operations, taking into account the area and design guidelines provided by FEMA; (2) We model the effects of congestion caused by crowding via state dependent queues and take into account blocking between the queues and (3) We explore ways to use triaging to improve RC performance.

Next, we consider a network of relief centers and propose a model that takes into account dynamically changing victim routing, victim mobility and external interventions that can act as operational coordination actions to balance network work load. To our knowledge, there is no prior research that considers a relief distribution network. Moreover, many prior studies highlight the need for dynamic and flexible modeling techniques for disaster response problems. We aim to fill this gap in the literature. We use a generalized queuing network with probabilistic transfers and batch movement of victims to model the dynamic aspects. In this research, we also contribute to the queuing literature by providing a new product form result and a product form approximation for cases where the product form result does not hold.

We then focus our attention to the material convergence problem, which directly affects the relief centers' access to supplies in a timely manner to distribute to the victims. The prior literature has very minimal quantitative models that aim to capture effects of material convergence. In this paper we aim to fill this gap in the literature by quantifying material convergence and providing an analytic model for resource allocation decisions under material convergence. We develop a multi-server transient queuing model representing the donations and apply numerical techniques to efficiently solve the model. We capture the relationship between material con-

vergence levels and resource allocation decisions, under varying donation arrival patters and high-priority item percentages. Our model allows the decision maker to evaluate multiple decisions (optimal and sub-optimal) and observe the trade-offs. We believe that this model can serve as a basis for decision support tools in the field.

Next, we analyze the impact of assumptions related to exponential inter-arrival and service rates, stationary arrival patterns, and steady state analysis used in the modeling of relief centers. We use discrete event simulation to model relief center layouts and relax the assumptions made previously. To our knowledge, there is only one prior work (Roy et al. (2011)) on relief center designs. This work, as well as our relief center model in Chapter 3, use these assumptions to obtain analytic results and expressions. Hence, this research contributes to understanding the impact of prior assumptions on the insights related to relief center design and operations.

Finally, we focus our attention on optimizing relief center layouts given constraints on resources. We use an exhaustive enumeration approach, coupled with simulation analysis to find the optimal relief center design considering both the throughput and the victim waiting time performance measures. The disaster response literature has used both simulation and optimization tools to model different problems. However, to our knowledge this is the first work to optimize relief center layouts.

Chapter 3

Relief Center Setup and Operations

3.1 Introduction

At the disaster site, the relief items are distributed to victims through temporary structures setup by the humanitarian organizations called relief centers (RCs). RCs are usually set up in large open areas (parking lots, play grounds) for ease of access and safety reasons. RCs usually have multiple points of distribution (PODs) each staffed with volunteers distributing the supplies. Each POD in a relief center is accessed via an entry walkway and exited via an exit walkway. In case of multiple PODs, there exits connecting walkways between the PODs. The disaster victims arrive to the RC and collect the supplies they need from the PODs. During the immediate aftermath of a disaster, demand for relief items is high, congestion is heavy due to crowds and resources are limited. Hence, victims might spend large amounts of time waiting in congested queues to receive critical relief supplies, or not receive the items at all.

Crowding experienced during distribution is a significant problem that can hinder the efficiency of relief distribution. In the case of Salvation Army, 2010 Haiti earthquake response, relief distribution efforts in Port-au-Prince came to halt due to the crowds converging around the relief center. One of the Salvation Army responders remembers it as: "Too many people gathered too fast. People were crushed and we had to shut down the distribution (Ozen and Krishnamurthy (2017))." After several days a new relief center design was implemented where the crowd was more strictly controlled, and victims were accepted one by one to the relief center. As in the example of Haiti relief distribution, relief center design and staffing changes can help manage the crowds better, and decrease victim waiting times.

Even though effects of congestion has been documented by relief agencies, there has been very limited research on modeling and quantifying this phenomenon for disaster relief distribution. Previous research on relief distribution focused on procuring and transporting the relief supplies to the disaster affected region, and lacked focus on the distribution of the supplies at the disaster site. The current relief distribution practice is based on prior experience of relief agencies. To understand the current practice and the challenges faced, we interviewed practitioners and organized working group sessions with representatives from the Salvation Army, Red Cross, South East Wisconsin Citizens and Organizations Active in Disasters (COAD), and Federal Emergency Management Agency (FEMA). The interviews revealed that relief agencies including FEMA, Red Cross and the Salvation Army have guidelines for setting up RCs and distributing relief (Federal Emergency Management Agency (2008)), USAID (United States Agency for International Development (2005)) and IFRC (International Federation of Red Cross and Red Crescent Societies (2008)).

However, they emphasized that the guidelines merely provide a starting point and there is currently no tool that can be used to evaluate alternative relief center designs based on the unique needs of a disaster. In this paper, we aim to address this need to quantify the impact of various RC design strategies. In particular, we aim to answer the following research questions: (1) How can an RC be setup to control effects of crowding, increase throughput and decrease waiting times? (2) What are some operational strategies that can help improve RC performance such as triaging?

To answer these research questions, we model each relief center as a finite capacity queuing network, where both the walkways and the PODs are represented as individual queues. To take into account effects of congestion, we model walkways within the RC as queues with state dependent (level of congestion) service rates. We model the PODs in the RC, as queues with service rates that are independent of congestion. To estimate performance measures of the queuing network model, we propose a solution approach based on decomposition. Using our solution methodology, we conduct numerical experiments where we compare different layouts and investigate the components of a layout that can reduce effects of crowding. We use data from the 2015 Nepal earthquake to estimate victim arrival rates and relief item distribution rates to test our model. Our main findings are: (1) given the same resources, how a relief center is setup has a significant impact on relief distribution performance, (2) dissipating the crowd by leveraging multiple distribution points within the RC decreases waiting times considerably, (3) triage queues can be leveraged to balance the workload in the RC to alleviate bottlenecks, (4) limiting number of victims allowed within the RC at once can decrease victim waiting times, however if decreased beyond a certain level, it can impair the throughput of the RC. All of our findings point to the importance of decision support models that can help practition-

ers decide on RC design strategies to serve more victims in a shorter amount of time.

The rest of this chapter is organized as follows. Section 3.2 provides a literature review, Section 3.3 describes the current and proposed RC layouts and models the associated queuing networks. Section 3.4 describes the solution methodology including the decomposition algorithm and the analytic formulas used for solving the decomposed queues. Section 3.5 explains the case study application of the Nepal earthquake and discusses the numerical studies and Section 3.6 concludes the chapter.

3.2 Literature Review

The work in this chapter is related to two main literature streams: relief center design and queuing network approximations.

Relief center design: To the best of our knowledge, Roy et al. (2011) and Ozen and Krishnamurthy (2013) are the only prior work in the literature that focus on relief supply distribution at the disaster site. These papers both highlight the importance of relief center design and emphasize the impact large crowds can have on relief distribution performance. Roy et al. (2011) compare different layouts using a network of $M/G/C/C$ queues. Our work differs from Roy et al. (2011) by using a different queuing approach, where the blocking relationship between queues in the network is captured explicitly. Therefore, the model proposed in this paper can provide more realistic performance measure estimates. This work also differs from Ozen and Krishnamurthy (2013) by modeling effects of crowding using state dependent queues. Due to these enhancements, new analytic formulas are derived in this paper

to solve the resulting queuing network. In addition to the mentioned modeling and theoretical differences, this work differs from the prior studies by evaluating layouts that are currently used in the field based on FEMA guidelines and uses Nepal earthquake data to bridge the gap between theory and practice of relief distribution.

Queuing network approximations: From a queuing network perspective, we model victims arriving and waiting at an RC to receive relief items as a network of finite capacity queues. Due to propagation of blocking in finite capacity networks, exact analytic solutions are hard to obtain and prior studies have focused on decomposition methods for approximate analysis. These decomposition methods are based on the idea of decomposing the system down to two-queue pairs, determining sub-system parameters and ensuring consistency and conservation of flow. Dallery and Frein (1993); Gershwin (1987); Brandwajn and Jow (1988); Altiok (1982) provide three popular decomposition methods for analyzing these queuing networks. Subsequently other studies have extended these methods to analyze more complex systems such as those with unreliable servers (Bihan and Dallery, 2000), closed-loop networks (Maggio et al., 2009), rework loops (Li, 2004) and assembly-type systems (Gershwin and Burman, 2000). Our work contributes to this line of literature by providing a solution methodology for queuing networks with blocking and state dependent service rates.

Like the prior studies, we also employ a decomposition approach to solve the queuing network model of the RC. In addition, we use state dependent service rates to capture the decreasing speed of crowd movement, resulting from increasing levels of congestion in the queues. Since existing methods cannot be applied directly to solve these networks, we derive new analytic expressions to characterize the system

of queues and obtain estimates of throughput and system time measures. Prior studies by Kerbache and Smith (1987); Smith (1991); Cheah and Smith (1994), use state dependent M/G/C/C queues to model congestion effects. The solution methodology is based on an "expansion method" that adds an artificial node to the network to hold the blocked victims and re-introduce them to the network via a feedback loop. The procedure proposed in this paper provides a more direct formulation of congestion effects and is more accurate under high utilization levels (common for disaster response).

3.3 The Relief Center Model

Relief centers have points of distribution (PODs), triage queues, volunteers and inventory and replenishment areas. The efficiency of relief center operations is impacted by the following criteria:

(i) Number of PODs: The number of PODs to open in a relief center can depend on number available volunteers, volume of relief supplies available at the RC, the area of the relief center and the number of victims the RC is serving.

(ii) Item assignment to PODs: The decision on how to assign the relief supplies to the points of distribution, will depend on the types of items and the total number of items being distributed, as well as the packaging of the delivered items. As an example, tarpaulins are large and bulky, hence it might be preferable to distribute them separately. Some items such as cleaning supplies come as a set and are distributed together. The assignment decision of items to PODs, can also impact the time it takes to distribute the items. If an RC decides to distribute tarpaulins

with cleaning sets at a POD, the time it takes to distribute items from that POD will increase (to the summation of both service times), compared to distributing these items separately. We will refer to this phenomenon as the additive service time assumption.

(iii) Physical capacity of the RC: The term physical RC capacity, refers to the number of victims that can be physically present in the RC at a time. Note that this is different than the RC throughput. This physical capacity will depend on the RC area, the number of victims the RC is serving, the number of available volunteers and the level of available supply. The decision parameter here is the number of victims allowed to queue within the RC at once. Note that, increasing the allowed RC capacity will increase the congestion within the relief center, and can increase waiting times of victims and decrease throughput. We investigate this trade-off using our model.

(iv) Triage queue: The triage queue is used to check victim information prior to item distribution. Note that triage, as it is used in relief distribution, is different from how it is used in medical emergency response (Iserson and Moskop (2007); Ergun et al. (2014)). In medical emergency response the aim is to prioritize victims according to the urgency of the problem. In contrast, at an RC, the triage is used to check information and route the victims to the correct locations to minimize congestion. How much service should be provided at the triage queue is an important decision parameter for relief center design.

(v) Walkways in the RC: The victims access the relief center using walkways, move between PODs via walkways and exits the RC via walkways. When congestion

in a walkway increases, the rate at which victims can move decreases significantly. This phenomenon is known as effects of crowding and has been quantified via empirical studies (Tregenza (1976)). The number, size and length of the walkways will all impact how fast victims can move within the relief center. Moreover, expanding the allowed physical capacity of the RC by accepting more victims to access the RC at once, will increase crowding and slow down movement of victims. In fact, this was the reason why the Salvation Army Port-au-Prince distribution came to a halt.

In order to capture the effects of these design criteria on the performance of the RC, we model the relief centers as a network of finite capacity queues where walkways, points of distribution and the triage are all represented as separate queues in the network. The number of PODs and the number of walkways will impact the number of queues in the network. The assignment of items to PODs and the tasks associated with triage will impact the distribution speed at these queues. The walkway queue service rates are dependent on the congestion in the relief center and are modeled using state dependent finite capacity queues. Next we discuss two layouts (one adapted from FEMA and one alternative layout we designed) using these decision criteria.

The Guide to Points of Distribution by FEMA and US Army Corps of Engineers (Federal Emergency Management Agency, 2008) provides one of the most detailed guidelines in terms of the current practice related to the setup of the relief center, the total required RC area, and the sizing of the pedestrian walkways. Figure 3.1 summarizes the design and essential parameters provided by FEMA (Federal Emergency Management Agency (2008)).

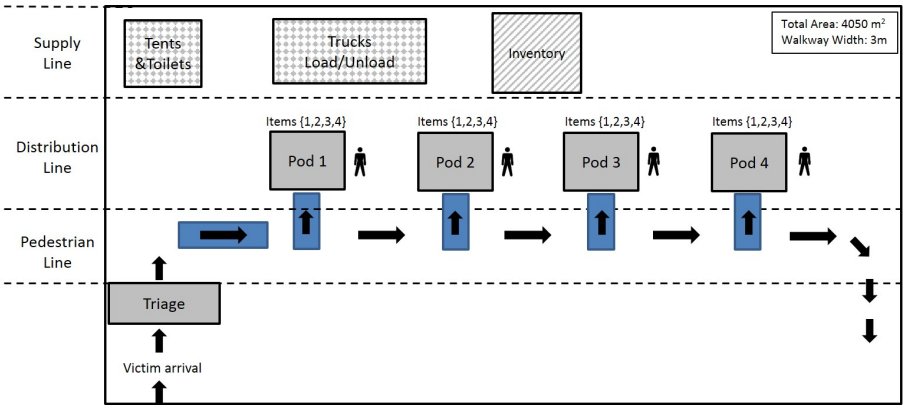

Figure 3.1: RC Layout: Current Practice (Federal Emergency Management Agency, 2008)

In this RC layout, there are 4 PODs staffed with a volunteer each, all PODs distribute all four of the available items and all victims access the RC using the same walkway. The triage queue is used to check victim information and track family affiliations prior item distribution. Note that triage as it is used in relief distribution is different from how it is used in medical emergency response (Iserson and Moskop (2007); Ergun et al. (2014)) where the aim is to prioritize victims according to urgency of treatment needed. In contrast, at an RC, the triage is used to check information and route the victims to the correct locations to minimize congestion. In the aftermath of a disaster, the number of victims that need supplies is often higher than the available distribution capacity. In a design where all victims converge to queue along the same entry walkway, congestion will increase and slow down the movement of the victims. Moreover, due to all PODs distributing all items, the service time at each POD is equal, and higher compared to only distributing a subset of items at a POD.

Based on these observations and the decision criteria discussed previously, we design an alternative RC layout shown in Figure 3.2. In this alternative layout, we

keep the overall RC area, the total number of PODs, the total number of volunteers and the walkway widths equal to the FEMA guidelines. This allows us to separate design decisions from effects of resource availability. In the alternative design, we change the item assignment to PODs and increase the number of walkways in the RC. Our aim here is to: (1) decrease the service time at each POD by distributing a subset of items (due to the additive service time assumption), and (2) create alternative routes for victims which can "thin" the crowd within the RC, and as a result can increase the service rate of the walkways (note the initial division of crowd to two). Note that, this design may effectively spread the crowd, however, it also results in victims going through a larger network of queues. We aim to investigate the trade-off between victims going through more queues which are less congested, and victims going through less queues which are more congested. Moreover, in our numerical studies we vary the physical relief center capacity and work assigned to triage queue and investigate the impact these decisions have on both RC layouts. In particular we aim to investigate the impact of the following design decisions: (i) What is the impact of crowd dissipating layout designs on victim waiting times and RC throughput? (ii) What is the impact of item assignment to PODs and number of walkways in an RC design on performance? (iii) What is the impact of the physical RC capacity on performance? (iv) How can triaging be utilized to improve RC performance? Note that other alternative designs than the ones we discuss in this Chapter are possible. Our aim is not to present an optimal layout, it is to emphasize non-obvious trade-offs in RC design and demonstrate the impact decision support models can have on relief distribution efforts. Our methodology can be used to evaluate a broad class of layouts in addition to the ones shown in Figures 3.1 and 3.2. In Section 3.5 we show how our method can be used to evaluate other layout designs.

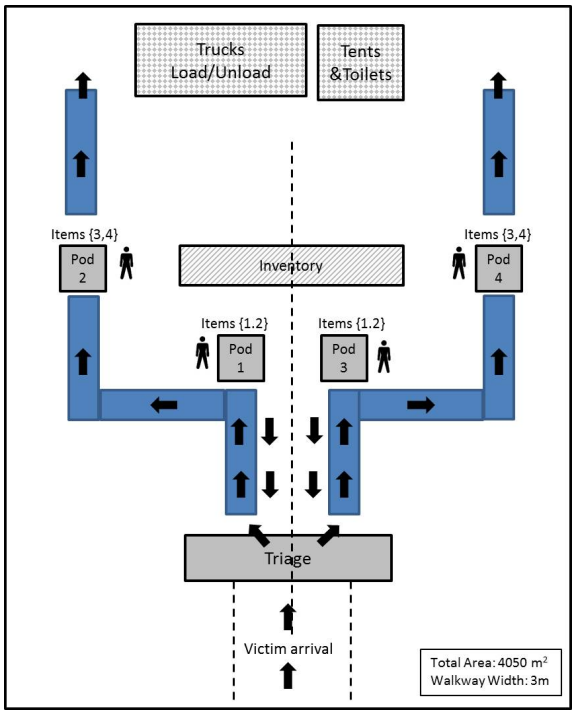

Figure 3.2: RC Layout: Proposed Layout

3.3.1 Queuing Model Representation of Relief Centers

We model the RC where victims wait in line to receive supplies from one or more PODs as a queuing network of finite capacity queues. We explicitly model both the walkways that provide access to the PODs and the PODs themselves as queues in the network. We model the walkways as finite capacity queues with state dependent service rates, where the rate of service is dependent on the speed at which pedestrians are able to move on the walkway. As the number of victims on the walkway (congestion) increases, the movement speed will slow down. We use empirical studies by Tregenza (1976) to model this relationship. We model the PODs as finite capacity queues with steady service rates where the service time only depends on the items that are being distributed.

Let the set $i = 1, 2, .., i, ..N$ represent the queues in the RC, and K_i represent the finite capacity of each queue. Also let $\mu_i(k_i)$ and μ_i represent the exponential service rate parameters for walkways and pods respectively, where $k_i = 1, 2, ..., k_i, .., K_i$. For analytic tractability, we assume that the external arrival rate of victims to the relief center follows a Poisson process with parameter λ. Although, the arrival rate to the RC is the same regardless of the design, the arrival rate to a particular POD depends on the set of items available at that POD and the demand for that subset of items. For the layout in Figure 3.2, a victim needing all items will go through both PODs 1 and 2, while a victim needing only item 1 will only visit POD 1. The queuing model therefore allows for converging and diverging flow of victims while keeping the overall flow of victims consistent.

Figure 3.3 depicts the resulting queuing network for the RC layout shown in Figure 3.1 representing the current practice and the flow corresponding to a victim visiting POD 1. Victims visiting PODs 2, 3 or 4 will only experience higher congestion and waiting times at the RC. Figure 3.4 depicts the resulting queuing network from Figure 3.2 for a victim requiring all items. By leveraging the symmetry of the RC design we only model pods 1 and 2.

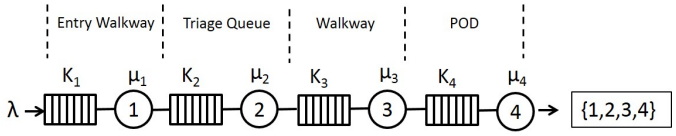

Figure 3.3: Current Practice: Queuing Network Representation

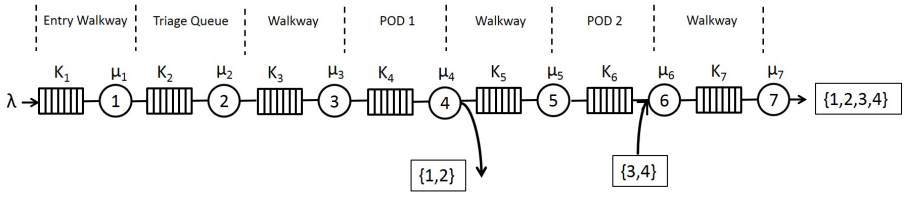

Figure 3.4: Proposed Design: Queuing Network Representation

Note that in the layout design given by Figure 3.1 (queuing network representation shown in Figure 3.3) all victims, irrespective of the set of items they require, use the same walkway to access the RC and the same long walkway is used to access each POD. In this design, even though victims go through less number of queues in total, the queues they go through are more congested. Moreover, a victim needing a single item or all items will go through the same network of queues. In the proposed design (see Figure 3.2 for layout and Figure 3.4 for queuing network), there are two entry points to the RC following the triage queue aiming to divide the crowd from the beginning. In addition, item assignment to each POD is varied so that a victim needing a single item need not go through all queues in the RC and a victim needing all items will go through less congested queues. Moreover, item assignment to PODs, not only creates alternative routes for victims but also can decrease the service time at the PODs. By distributing fewer number of items the time it takes for distribution can be shortened, resulting in faster queues. In Section 3.5 we consider various item assignments to analyze the impact of both varied routing and service rate balancing can have on victim suffering.

It is possible to analyze the layout in Figure 3.2 as a multi-class queuing network. However, in our solution approach we use input aggregation used by Whitt (1983) and calculate the arrival rate and service rate at each queue for a typical aggregate victim. See Whitt (1983) for the details and accuracy of the aggregation approach.

Next we discuss the methodology proposed to solve the resulting queuing networks. Using this methodology we estimate the performance measures (average time a victim going through the RC experiences and throughput of the RC) and we determine them using a decomposition methodology since exact analytic solutions are hard to obtain. The overall performance measure we use (victim deprivation objective) combines these two measures.

3.4 Performance Evaluation

The RC designs result in open queuing networks with finite capacity queues and blocking after service. Let k_i denote the number of victims at queue i and let $P = (k_1, k_2, ...k_i, ..., k_N)$ represent the steady state probability for the network. To estimate the network performance measures we first need to solve this queuing network for its steady state probabilities, which we do so following a three step process. (1) We first decompose the network of queues into 2-queue subsystems, each characterized with an effective arrival rate and a 2-phase Coxian effective service rate. (2) We then analyze each subsystem in isolation using the analytic formulas derived in Section 3.4.1. (3) Next, we use a recursive algorithm (See Section 3.4.2), to link back the network and capture the inter-dependencies via blocking probability estimations to calculate accurate performance measures for the network.

The two main assumptions made during decomposition are: (i) a queue may only be blocked by its immediate successor queue and (ii) the arrival rate to each queue is Poisson. Each individual queue obtained from the decomposition is defined by its effective arrival rate and the effective service rate. The effective arrival rate

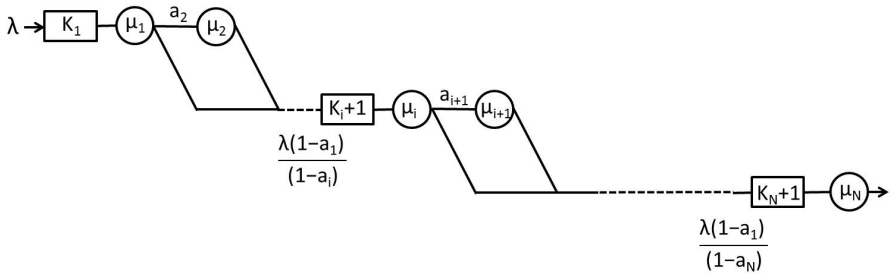

Figure 3.5: Queuing Network after Decomposition

must ensure that the system throughput stays constant while the effective service rate must account for the additional time spent at service due to the victim being served possibly getting blocked after service (See Figure 3.5).

Blocking in the system occurs after service completion at queue i when the downstream queue $i+1$ is at capacity, although the downstream queue being at capacity does not always imply blocking. To distinguish the two states, we increase the queue capacities by 1 to represent the blocked victim in queue. Then the probability of blocking for queue i can be written as $a_i = P(k_i = K_i + 1)$ for $i = 2, 3, ..., N$. Next we describe the effective arrival and service rate calculations.

The external arrival rate to the network is known and is assumed to be Poisson with parameter λ. Then the throughput of the network can be written as $TH_1 = \lambda * (1 - a_1)$, where a_1 is the probability that queue 1 is at capacity. This implies that the RC is unable to provide service to $\lambda * a_1$ victims. Then, the effective arrival rate to each subsequent queue is defined as $\lambda_i = TH_1/(1 - a_i), i = 2, ..., N$, where $a_i, i = 2, ..., N$ is the probability queue i is full.

The service time at each queue is assumed to be exponential with a fixed rate for the distribution pods and state dependent exponential for queues representing RC walkways. Due to the memory-less property of the exponential distribution, the time a station is blocked can also be characterized by the exponential distribution (See Figure 3.6). This implies that effective service time at queue i has a 2-phase Coxian distribution, where the first phase represents the actual service time and the second phase represents the time spent blocked, and a_{i+1} represents the probability of blocking. Note that all queues in the system besides the final queue can be characterized as an $M/C_2(k_i)/1/K_i$. Note that a queue with service time that is independent of the state k_i, is a special case of a queue with state dependent service times. Since the last queue can not be blocked it is characterized as an $M/M(k_i)/1/K_i$ queue. Next we derive the formulas required to solve $M/C_2(k_i)/1/K_i$ and $M/M(k_i)/1/K_i$ queues in isolation.

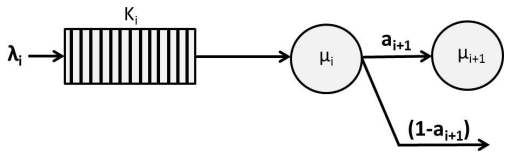

Figure 3.6: Model of a Queue in Isolation

3.4.1 Analysis of Queues in Isolation

We defined each queue in the network as a standalone queue with Poisson arrivals with rate $\lambda_i = TH_1/(1-a_i)$ and a two-phase Coxian distributed service time with parameters μ_i and $(\mu_{i+1} * a_{i+1})$ characterizing the first and second phases of the service time respectively. To calculate the steady state probabilities of the $M/C_2(k_i)/1/K$ queue in isolation we analyze the Markov Chain embedded at departure epochs. Let $\pi(k_i), k_i = 0, 1, ..., K_i - 1$ denote the steady state probabilities of the embedded

Markov chain and let $P(k_i), k_i = 0, 1, 2, ..., K_i$ represent the steady state probabilities of the $M/C_2(k_i)/1/K_i$ queue. Note that for simplicity the subscript i representing a specific queue in the network will be suppressed throughout the remainder of this section.

Let X_k denote the state of the Markov chain representing the number of victims present in the queue right after the k^{th} departure and let $q_{k,n}$ represent the probability that, following a departure leaving the queue in state k, exactly n victims arrive during the service time. Then, the transition probability matrix for the Markov chain embedded at departure instances can be written as:

$$\mathbb{Q} = \begin{bmatrix} q_{0,0} & q_{0,1} & q_{0,2} & \cdots & q_{0,K-2} & r_{0,K-1} \\ q_{1,0} & q_{1,1} & q_{1,2} & \cdots & q_{1,K-2} & r_{1,K-1} \\ 0 & q_{2,0} & q_{2,1} & \cdots & q_{2,K-3} & r_{2,K-2} \\ 0 & & & & \vdots & \vdots \\ 0 & & & \ddots & & \\ 0 & & & & q_{K-1,0} & r_{K-1,1} \end{bmatrix}$$

To estimate $q_{k,n}$, we repeatedly use the memory-less property of the exponential distribution along with the probability representing the minimum of two exponential random variables. Note that $\mu_d(K_d + 1)$ represents the service time of the downstream queue when it blocks its upstream queue.

For $k \geq 0$ and $n = 0$:

$$q_{k,n} = (1 - a)\left(\frac{\mu(k)}{\lambda + \mu(k)}\right) + a\left(\frac{\mu(k)}{\lambda + \mu(k)}\right)\left(\frac{\mu_d(K_d + 1)}{\lambda + \mu_d(K_d + 1)}\right) \tag{3.1}$$

For $k \geq 0$ and $n = 1$:

$$\begin{aligned} q_{k,n} &= (1-a)\left(\frac{\lambda}{\lambda+\mu(k)}\right)\left(\frac{\mu(k+1)}{\lambda+\mu(k+1)}\right) \\ &+ a\left[\left(\frac{\lambda}{\lambda+\mu(k)}\frac{\mu(k+1)}{\lambda+\mu(k+1)}\frac{\mu_d(K_d+1)}{\lambda+\mu_d(K_d+1)}\right)\right] \\ &+ \left(\frac{\mu(k)}{\lambda+\mu(k)}\frac{\lambda}{\lambda+\mu_d(K_d+1)}\frac{\mu_d(K_d+1)}{\lambda+\mu_d(K_d+1)}\right) \end{aligned} \tag{3.2}$$

For $k \geq 0$ and $n > 1$:

$$\begin{aligned} q_{k,n} &= (1-a)\prod_{j=k}^{k+n-1}\frac{\lambda}{\lambda+\mu(j)}\frac{\mu(n+k)}{\lambda+\mu(n+k)} \\ &+ a\sum_{b=1}^{n}\left[\left(\prod_{j=k}^{k+b-1}\frac{\lambda}{\lambda+\mu(j)}\right)\frac{\mu(k+b)}{\lambda+\mu(k+b)}\right] \\ &* \left(\prod_{j=k+b}^{k+n-1}\frac{\lambda}{\lambda+\mu_d(K_d+1)}\right)\frac{\mu_d(K_d+1)}{\lambda+\mu_d(K_d+1)} \\ &+ a\left[\left(\prod_{j=k}^{k+n-1}\frac{\lambda}{\lambda+\mu(j)}\right)\frac{\mu(n+k)}{\lambda+\mu(n+k)}\frac{\mu_d(K_d+1)}{\lambda+\mu_d(K_d+1)}\right] \\ &+ a\left[\frac{\mu(k)}{\lambda+\mu(k)}\left(\prod_{j=k}^{k+n-1}\frac{\lambda}{\lambda+\mu_d(K_d+1)}\right)\frac{\mu_d(K_d+1)}{\lambda+\mu_d(K_d+1)}\right] \end{aligned} \tag{3.3}$$

For $k \geq 0$ and $n \geq 0$:

$$r_{k,n} = q_{n,k} + q_{k,n+1} + q_{k,n+2} + q_{k,n+3} + ... \tag{3.4}$$

Now that the matrix $\mathbb{Q}$ is completely defined, we solve the equation $\pi = \pi\mathbb{Q}$ to determine the steady state probabilities of the Markov chain, π.

$$\begin{aligned} \pi(k) &= \frac{P(k)}{1-P(K)} \quad for \quad k = 0, 1, 2, \ldots, K-1 \\ \lambda(1-P(K)) &= \sum_{i=1}^{K} \mu_{eff}(i) P(i) \end{aligned} \tag{3.5}$$

Then using $\pi(k), k = 1, 2, ..K-1$ we get the steady state probabilities $P(k)$ using Equations 3.5 where $\mu_{eff}(k)$ is defined as $\left(\frac{1}{\mu(k)} + \frac{a_d}{\mu_d(K_d+1)}\right)^{-1}$. Then we can use Equations 3.6 to get the steady state probabilities from departure epoch probabilities.

$$\begin{aligned} P(k) &= \frac{\pi(k)}{\pi(0)+\rho}, \quad k = 0, 1, 2, \ldots, K-1 \\ P(K) &= \frac{\rho - 1 + P(0)}{\rho} \end{aligned} \tag{3.6}$$

Finally, we can calculate the throughput, average number of victims in each queue and the average time spent at each queue using Equations 3.7 respectively for each queue i. Note that, the overall waiting time $\bar{W}_i$ of a victim is obtained by summing over all waiting times of the queues the victim goes through. The overall throughput of the RC is equal to $\lambda * (1 - a_1)$ due to the conservation of mass.

$$\begin{aligned} TH_i &= \lambda_i * (1 - a_i) \\ L_i &= \sum_{j=0}^{K_i} P_i(j) j, i = 1, 2..., N \\ \bar{W}_i &= L_i / TH, i = 1, 2..., N \end{aligned} \tag{3.7}$$

Recall that every queue except the last queue in the queuing network in Figure 3.3 and 3.4 are $M/C_2(k)/1/K$ queues. The last queue in each of these networks is an $M/M(k)/1/K$ queue. Next we describe how we compute the steady state probabilities for these queues.

The steady state probabilities of the $M/M(k)/1/K$ queue can be obtained by constructing and solving the balance equations given by Equation 3.8.

$$\begin{aligned} P(0)\lambda &= P(1)\mu(1) \\ P(1)\lambda + P(1)\mu(1) &= P(0)\lambda + P(2)\mu(2) \\ P(2)\lambda + P(1)\mu(1) &= P(1)\lambda + P(3)\mu(3) \\ &\vdots \\ P(K-1)\lambda &= P(K)\mu(K) \end{aligned} \tag{3.8}$$

Solving the above equations, we can obtain:

$$P(0) = \frac{1}{1 + \sum_{n=1}^{K} \frac{\lambda^n}{\mu(1)\ldots\mu(n)}} \tag{3.9}$$

$$P(k) = P(0)\frac{\lambda^n}{\mu(1)\ldots\mu(k)} \tag{3.10}$$

We then use Equation 3.7 to obtain the throughput (TH) and total time (W) performance measures.

3.4.2 Iterative Algorithm to Link Queues

The steady state probabilities of any queue i in the network is dependent on the probability that their downstream queue $(i+1)$ is full and the probability that the first queue is full (since it governs the overall system throughput). Formally, the probability a_i is a function of a_1 and a_{i+1}, $\forall i = 1, 2, 3, .., N-1$ and a_N is solely dependent on a_1. Therefore, to accurately estimate the steady state probabilities of each queue, we solve each queue in isolation assuming starting blocking probabilities (using equations derived in Section 3.4) and iterate until all blocking probability estimates converge. The iterative algorithm is summarized below:

Inputs: λ, K_i, $\mu_i(k_i)$, for all $i = 1, 2, ..N$, initial estimates $a_i^{(0)}, i = 1, 2, ..N$ and ϵ

Step 1: For $m = 1, 2, ..., M$ execute steps 2-4.

Step 2: For queue $i = N$:

2.1 Analyze queue N as an $M/M(k_N)/1/K_N$ queue with arrival rate λ_N, service rate $\mu_N(k_N)$ and capacity K_N and obtain steady state probabilities $P_N(k_N = j), j = 1, 2, \ldots, K_N$.

2.2 Update arrival rate λ_N with revised estimate of $a_N = P_N(k_N = K_N)$.

2.3 Iterate until estimate of $P_N(k_N = K_N)$ converges.

2.4 Let $a_N^{(m)}$ correspond to this estimate of $P_N(k_N = K_N)$.

Step 3: For queues $i = N-1, N-2, \ldots, 2$:

3.1 Analyze queue i as an $M/C_2(k)/1/K_i$ queue with arrival rate λ_i, service rate $\mu_i(k_i)$ and capacity $K_i + 1$ and obtain steady state probabilities $P_i(k_i = j), j = 1, 2, .., K_i + 1$.

3.2 Update arrival rate λ_i with revised estimate of $a_i = P_i(k_i = K_i + 1)$.

3.3 Iterate until estimate of $P_i(k_i = K_i + 1)$ converges.

3.4 Let $a_i^{(m)}$ correspond to this estimate of $P_i(k_i = K_i + 1)$.

Step 4: For queue $i = 1$:

4.1 Analyze queue 1 as an $M/C_2(k)/1/K_1$ queue with arrival rate λ, service rate $\mu_1(k_1)$ and capacity $K_1 + 1$ and obtain steady state probabilities $P_1(k_1 = j), j = 1, 2, .., K_1 + 1$.

4.2 Compute the absolute difference $\delta = |a_1^{(m)} - P_1(k_1 = K_1 + 1)|$.

4.3 If $\delta \leq \epsilon$, go to step 5.

4.4 If $\delta > \epsilon$, set $m = m + 1$ and repeat steps 2-4.

Step 5: Record steady state probabilities $P_i(k_i), k_i = 1, 2, ..., K_i + 1, \forall i$.

Step 6: Compute performance measures using Equation 3.7.

For a given value of a_1, we start with the last queue N. We guess an initial value for a_N and solve this queue for its steady state probabilities and obtain a revised estimate of $a_N = P_N(n_N = K_N + 1)$. We iterate over values of a_N until the estimate of a_N is consistent with λ_N. Using the solution for queue N, we analyze queue $N - 1$. This process continues until we get a revised estimate of a_1. If the error between the current and previous estimate of a_1 is greater than ϵ, we repeat the solution process for queue N through queue 1, else we exit the loops. Note that for the search for a_1 we use a grid search while the search for all other a_i follows a fixed point iteration algorithm (Conte and de Boor (1980)). In the global loop we apply a grid search over values of a_1 as it allows us to solve for cases where the external arrival rate greatly exceeds the achievable throughput of the network.

3.4.3 Measuring Performance of a Relief Center

The two main RC performance measures are; the average time to service a victim and the hourly throughput of the RC because minimizing suffering of victims requires both reaching to all victims and reaching them as quickly as possible. Therefore, maximizing both measures is the most preferred outcome, however there are decisions that require trade-offs between these two measures. Moreover, the evaluation of the throughput measure is dependent on the total population the RC needs to serve. The throughput measure may be critical if the target population is large and many victims face the risk of being denied service. Therefore a single objective function capturing these trade-offs is needed. There has been many disaster response research on appropriate humanitarian objective functions such as minimizing waiting time, minimizing unsatisfied demand and ensuring equity. Among these, the victim deprivation objective proposed by Holguin-Veras et al. (2013) is a good candidate for evaluating relief centers because it quantifies the willingness to pay for a critical relief supply in terms of the time for which the victim was deprived of the item. The objective function takes into account: (1) the average waiting time experienced by victims ($\bar{W}$) and (2) the hourly rate at which the RC serves victims (TH), and (3) if supply (or capacity) is less than demand, the number of victims the RC was unable to serve. Equation 3.11 gives the objective function. The parameters used are: P the effected population in need of items, ω per hour cost of a volunteer, n the number of volunteers and τ the time period in hours.

The objective function is written as:

$$\gamma(\bar{W}, TH, \omega, \tau) = \phi(\bar{W}, TH, \tau) + \varphi(\bar{W}, TH, \tau) + \kappa(\omega, \tau, n) \tag{3.11}$$

The components of the objective function $\phi(\bar{W}, TH)$, $\varphi(\bar{W}, TH)$ and $\kappa(\omega, \tau, n)$ represent the deprivation function for victims (measured by waiting time (W)) that receive service at the RC (charecterized by throughput (TH)), the deprivation function for victims that could not receive service at the RC and the labor cost function respectively. The logistics cost for the RC is assumed to be the staffing cost alone since we assume that the items to be distributed are already in the affected area at the time of RC setup. The fucntions ϕ and φ have the exponential structure following the willingness to pay function (WTP) reported in Holguin-Veras et al. (2013), however in Section 3.5.6 we perform sensitivity analysis to understand the effects of this particular parameter choice on RC decisions. Note that the objective function uses the waiting time and throughput of the RC as input, which are the outputs of the particular RC design and operation strategy and are obtained using algorithm described in Section 3.4.

Deprivation function for victims that receive service at the RC:

$$\phi(\bar{W}, TH, \tau) = \left[\exp(a + (b * \bar{W})) - \exp(a)\right] * TH * \tau \tag{3.12}$$

Deprivation function for victims that do not receive service at the RC:

$$\varphi(\bar{W}, TH, \tau) = [\exp(a + (b * \tau)) - \exp(a)] * ((P - (TH * \tau))) \tag{3.13}$$

Labor cost function at the RC:

$$\kappa(\omega, \tau, n) = \omega * n * \tau \tag{3.14}$$

3.5 Case Study: Nepal Earthquake (2015)

In this section, we discuss the application of our model using the 2015 Nepal earthquake as a case study. The Nepal earthquake is a good case for analyzing relief distribution because it resulted in the displacement of three million people, making relief item distribution a top priority during the immediate response. Moreover, due to inadequate infrastructure and limited number of volunteers and agencies, the efficient distribution of relief was a significant challenge. Using Nepal data from Humanitarian Data Exchange Database (2016), we estimate arrival and service rates for relief centers and use our model to evaluate the RC designs given in Figures 3.1 and 3.2. In particular we quantify impact of congestion on RC performance, analyze effects of the proposed crowd dissipation strategies and evaluate balancing RC workload using the triage queue. We use the multi-criteria deprivation objective function to evaluate the RC designs and operational strategies. Sections 3.5.2, 3.5.3 and 3.5.4 present numerical analysis that answer these research questions.

3.5.1 Parameter Estimation

To solve the queuing model and estimate RC performance measures, arrival rate to the RC and service rate at the pods are required input parameters. We use data from the Humanitarian Data Exchange Database (HDX), which is an open platform for humanitarian data sharing (https://data.humdata.org/) to estimate these key parameters. Details on the items distributed are obtained from (Meta Data Source: https://data.hdx.rwlabs.org/dataset/scnepal-agency-data) and the demographics and casualty information is obtained from (Meta Data Source: https://data.hdx.rwlabs.org/dataset/official-figures-for-casualties-and-damage). The distributed supplies data set provides information on types of items distributed, the distribut-

ing agency, number of households served and the dates the distribution took place. The data set is prepared by the Global Shelter Cluster(GSC) which is a coordinating agency that brings together 35 global partners to coordinate and respond to disasters and conflict situations. The GSC is co-chaired by the International Federation of Red Cross (IFRC) and the United Nations High Commissioner for Refugees (UNHCR). The demographics and casualty data set provides population, number of households, total number of deaths and injured victims per district. The data set is provided by the Nepal United Nations Office for the Coordination of Humanitarian Affairs (OCHA) with data contributions from the Nepal Ministry of Home Affairs and the Nepali Police.

The data set on distributed supplies contains relief centers that were active for 7 months, distributing different types of items. Our study focuses on the items that support the immediate needs of victims following an earthquake such as tarpaulins, blankets, kitchen supplies (for water and food) and hygiene kits. Therefore we limit the analysis time frame to the first 3 weeks. Note that the distribution of perishable and repeat items such as water and food are not part of this analysis or data set. Next we describe the process used to estimate the arrival and service rates from the raw data. Note that as described in Section 3.4, these form crucial inputs to the queuing model.

Arrival Rate Estimation: The arrival rate per relief center per day is estimated for each district as the total survived population divided by the total relief center days of active item distribution. To calculate the total survived population we use the demographics information of each district and assume all earthquake survivors needed critical relief supplies. This gives us the total demand rate. Estimating the

Table 3.1: Nepal Data: Arrival Rate Estimation

District	Arrival/RC day	Arrival/RC hour
Bhaktapur	925	58
Dhading	465	29
Dolakha	461	29
Gorkha	168	10
Kathmandu	3,158	197
Lalitpur	643	40
Nuwakot	1,535	96
Ramechhap	368	23
Rasuwa	245	15
Sindhupalchok	138	9

arrival rate per hour per RC was challenging. To do so, we used the total RC days parameter for each district, which refers to the total number of days an RC was open, summed over all RCs. This parameter allows us to convert total demand to an hourly arrival rate for an RC, by assuming that the distribution was carried out for 16 hours each day.

While the demand per district is closely related to its population, the total RC days varies between districts due to the accessibility of the district, and the differences in the number of agencies responding to needs in a district. Hence the arrival rate to an RC vary between different districts. Table 3.1 provides the resulting arrival rate per RC per hour data. In our analysis we first discuss insights with respect to Bhaktapur since its arrival rate represents the average. Subsequently in Section 3.12 we discuss insights for Nuwakot and Kathmandu to capture the effects of increasing arrival rates.

Service Time Estimation: There are no standard times provided for disaster relief item distribution. Therefore, we estimate the time it takes to distribute an

Table 3.2: Items Required per Household

Item	Conversion Parameter
Tarpaulin	1/Household
Blanket	2/Household
Kitchen Sets	1/Household
Sleeping Mat	5/Household

item of a particular type using the data set of distributed supplies. The data set contains the number of households served in a given day for an agency, district, item triplet. In our analysis, we distinguish item types but we do not distinguish districts or agencies since the time to distribute a tarpaulin will be the same regardless of location or agency. Finally, we estimate number of households served per day for a given item type. The number of items required of each type of item per household is different for different items. For example every household (made up of 5 people on average) needs 1 tarpaulin while needing 2.5 blankets (2 per person). To calculate the distribution time per item based on our estimate of number of households served per day for a given item type, we need number of items (for a given item type) a household needs. We obtain this information from Red Cross Nepal Emergency Appeal Operation Update Reports as shown in Table 3.2. Then we calculate the time it takes to distribute a specific item. Table 3.3 provides the average service rates per item calculated over the 10 most effected districts.

Table 3.3: Service Rates for Items Distributed

Item	Average Distribution Rate
Tarpaulin	10.7/hr
Blanket	31.6/hr
Kitchen Sets	10.2/hr
Sleeping Mat	20.2/hr

3.5.2 Comparing RC Layout Designs

In the first set of experiments we compare layouts in Figure 3.1 (representing current practice) and Figure 3.2 (representing alternative design) and two alternatives to evaluate the separate impact of item assignment decisions and crowd dissipation strategies. Layout alternative 1 is the queuing representation of current practice adapted from FEMA guidelines, as previously given by Figures 3.1 and 3.3. Layout alternative 4 leverages the multi-access points to the RC idea, however distributes all items from all PODs as described in FEMA guidelines. Layout alternatives 1 and 4 are shown side by side in Figure 3.7. Both these layouts result in a 4 queue network composed of the entry walkway to triage, the triage queue, entry walkway to POD 1 and service queue for POD 1.

Layout alternatives 2 and 3 represent the queuing network representation of the alternative design, where in alternative 2, POD 1 distributes tarpaulins and blankets and POD 2 distributes kitchen sets and sleeping mats. Whereas in design 3, POD 1 distributes tarpaulins and kitchen sets while POD 2 distributes blankets and sleeping mats. Layout alternatives 2 and 3 are shown side by side in Figure 3.8. Both these layouts result in a 7 queue network composed of the entry walkway

to triage, the triage queue, entry walkway to POD 1, service queue for POD 1, connecting walkway to POD 2, entry walkway to POD 2 and service queue for POD 2. Note that in all 4 designs, the total area of the RC and the walkway widths for pedestrians are the same, as specified by FEMA documentation.

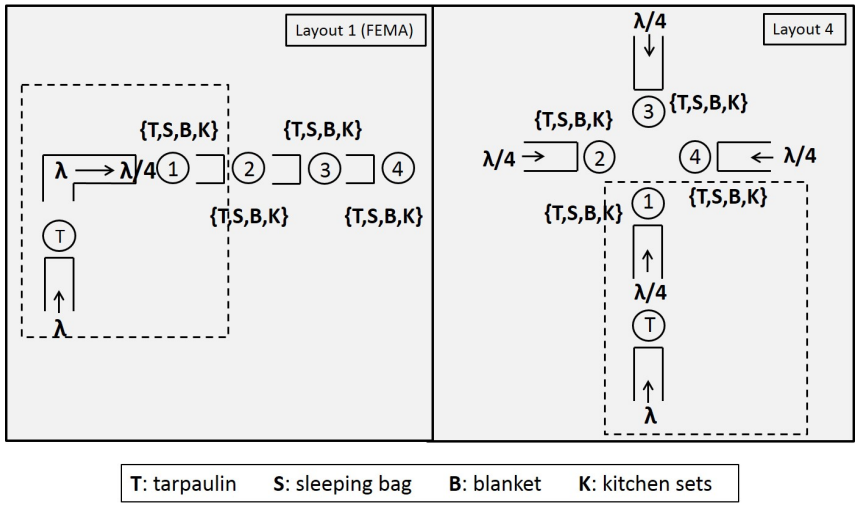

Figure 3.7: Queuing Network Representation of Layout Alternatives

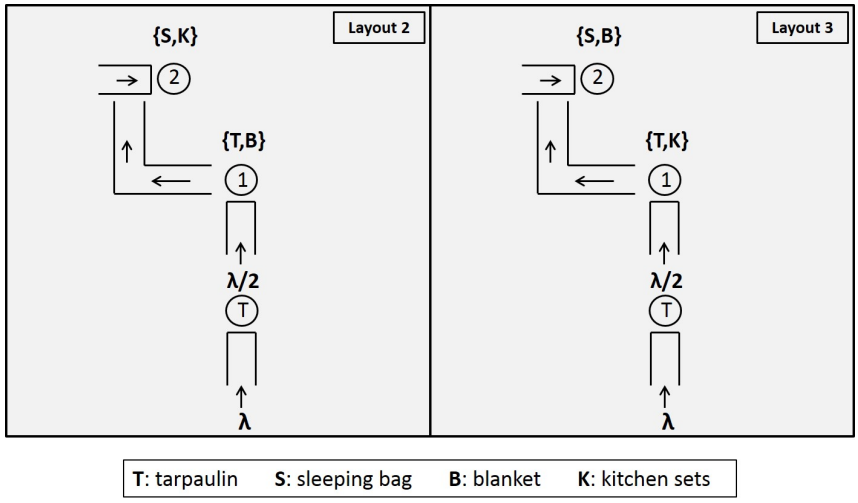

Figure 3.8: Queuing Network Representation of Layout Alternatives

Comparing layouts 1 and 4, we aim to investigate the impact of dissipating the crowd, on RC performance. In layout design 1, the entry walkway will be congested

due to the single access point. In layout 4, we remedy this by creating multi-access points in the layout. Next, we compare layouts 1 and 2. In layout 2, we assign a subset of items to each POD, and dissipate the crowd entering the RC. By comparing these two layouts, we aim to compare the combined impact of item assignment and crowd dissipation. Assigning a subset of items to each POD, is expected to decrease the distribution time. Note that in layout 2, victims will be going through less congested queues and receive faster service at the PODs. However, they will have to go through a higher number of queues (7 queues as opposed to 4). Through numerical experiments, we aim to compare the impact between going through a shorter network of queues, and going through a larger network with less congested, faster queues. Finally, we compare layouts 2 and 3 to investigate the impact of item assignment alone on RC performance. In layout 3, distributing tarpaulins and kitchen sets together results in less victims arriving to this queue (see Table 3.2). This results in lower blocking in the system, and therefore a higher throughput. However, a higher throughput, increases the number of victims present in the RC at once (hence increases the congestion), which in turn, creates higher victim waiting times.

Table 3.4 provides the parameters for the numerical experiement. We use the arrival rate of Bhaktapur (58/hour) as a representation of the average arrival rate. Note that the overall arrival to the RC is the same for all designs. However, the arrival rate to each queue may differ depending on the demand for items being distributed at a particular POD. To calculate arrival rate per POD, we use the data provided in Table 3.2. We use the distribution rate per item as given in Table 3.3 and assume the rates are additive. Additivity is a reasonable assumption (if not necessary) for items like tarpaulins and kitchen sets where handling takes considerable

time. Note that for PODs at which not all arriving victims need all items, we use a weighted average to calculate the service rate. For the walkway service rates we use one of the empirical curves that charecterize the exponential relationship between crowd density and walking speed given by Tregenza (1976). The exact exponential decay function used in these experiments is given in Table 3.4. To compute the suffering objective, we assume 16 hours of operation per day and for victims unable to receive service at the RC, we assume a 24 hour daily deprivation. For the labor cost we assume \$10/hour per volunteer. Note that since the number of volunteers in all layout designs are the same, the layout design evaluations are insensitive to this assumption.

Table 3.4: Numerical Experiment 1 Parameters

Parameter	Value
Arrival Rate	$\lambda = 58$/hour (for Bhaktapur)
Service Rates	$\mu_T = 10.7/hr$, $\mu_B = 31.6/hr$, $\mu_K = 10.2/hr$, $\mu_S = 20.2/hr$
Queue Capacity	$N_i = 50, \quad \forall i$
Walkway Service Rate	$\mu_w(k_i) = 60 * \exp((-c/d) * (k_i)^{1/4})$, $b = 40/60,\ c = -\log(b),\ d = (N_i)^{1/4}$

Table 3.5: Numerical Experiment 1 Results

Layout Design	$\bar{W}$	TH	$\phi(\bar{W}, TH)$	$\varphi(\bar{W}, TH)$	$\kappa(\omega, \tau, n)$	Total
Layout 1	6.63 hrs	27.14/hr	\$2,294	\$25,744	\$800	\$28,838
Layout 2	2.79 hrs	33.64/hr	\$936	\$18,424	\$800	\$20,160
Layout 3	3.59 hrs	46.86/hr	\$1,763	\$3,536	\$800	\$6,099
Layout 4	6.39 hrs	27.14/hr	\$2,176	\$25,744	\$800	\$28,720

Table 3.5 provides the results for the four different RC designs. In the table we

report the average time experienced by a victim ($\bar{W}$), the total hourly throughput (TH) of the relief center, and the suffering objective. Note that in the Appendix we validate our solution approach by comparing results obtained via the decomposition algorithm with result from simulation.

Comparing the results for layouts 1 and 4, we observe that dissipating the crowd by creating multiple access points, resulted in an improvement in victim waiting times. Here it is important to note that, the waiting time estimate for layout 1 is for victims going through POD 1 alone. Victims going through PODs 2,3 and 4 in layout 1 will experience higher waiting times, especially under congestion. Therefore, we conclude that when congestion is high, dissipating the crowd can have a considerable impact on victim waiting times. Since the item assignment to PODs is the same for layouts 1 and 4, the bottleneck service rate remains the same. And that is why, the throughput for both layouts is equal. Next, we compare results for layouts 1 and 2, and observe a significant improvement for both victim waiting times and the RC throughput. This shows that changing item assignment to PODs combined with crowd dissipation strategies (under the additive service time assumption) can drastically improve RC performance. This is because in layout 2, congestion in walkways is lower (crowd dissipation) and service time at the PODs is lower (item assignment). Hence, all queues in the network can flow faster, resulting in increased throughput and decreased waiting times. Finally, we compare the results for layouts 2 and 3. We observe that, item assignment can also impact both the throughput and service time measures.

The results obtained in this experiment set lead us to several insights. First, going through a larger number of queues can decrease the time victims spend in the

RC. This is a counter intuitive result, since waiting in more lines should result in an increase in time spent waiting. However, our numerical experiments show that, if creating more queues in the RC layout can be leveraged to decrease the congestion and balance service times, larger networks can be preferred. Secondly, the decision of how to assign items to PODs can have a significant impact on performance. Most often, having all items available at each POD is easier to setup. However, our results show that it is important to think about whether this decision is creating significantly higher service times at the PODs. This may not be the case for smaller items or bundled items where the additive service time assumption does not hold. However, quantifying the impact of assignment decision for cases where it will matter, can increase the awareness of practitioners making these decisions. Lastly, some decisions can have an opposite impact on throughput and waiting time measures. As an example, the throughput improvement in layout 3, comes at the expense of an increase in waiting time. In such cases, having a performance measure such as the suffering objective, that encompasses multiple criteria will be highly useful. Comparing the suffering function objective for layouts 2 and 3, we observe the cost of layout 3 is much lower, even though it creates a higher waiting time for victims at the RC. This is because, victims receiving their supplies waiting an additional hour has less weight compared to 13 victims (every hour) not receiving items at all. This showcases the need for humanitarian multi-criteria objective functions.

We would also like to note that in these experiments, we did not consider the impact of number of available volunteers or available relief center area on RC design decisions. If there are limited number of volunteers or if the area is limited which limits the number of PODs the RC can operate, distributing all items from a single point may be the only option. However, using models like the one proposed in

this paper, can estimate the impact these limitations will have on expected waiting times, and when options present themselves, these models can help make more efficient decisions.

3.5.3 Effect of Queue Capacity

The physical size of the RC and the walkways are given. However, the number of victims allowed to queue inside the RC is a decision parameter. We investigate the effectiveness of this decision called allowed physical queue capacity $K_i, i = 1, 2, ..., N$, as an alternative congestion control strategy. Since the actual queue capacity (for walkways) is dependent on the size of the walkway, which we do not change for these experiments, by varying K_i between 25 and 75 for all 4 layout designs, we impact the congestion of queues in the RC. We keep the rest of the parameters constant as given in Table 3.4. Figure 3.9 presents the resulting changes in RC throughput and average service time while Table 3.6 reports the suffering objective results.

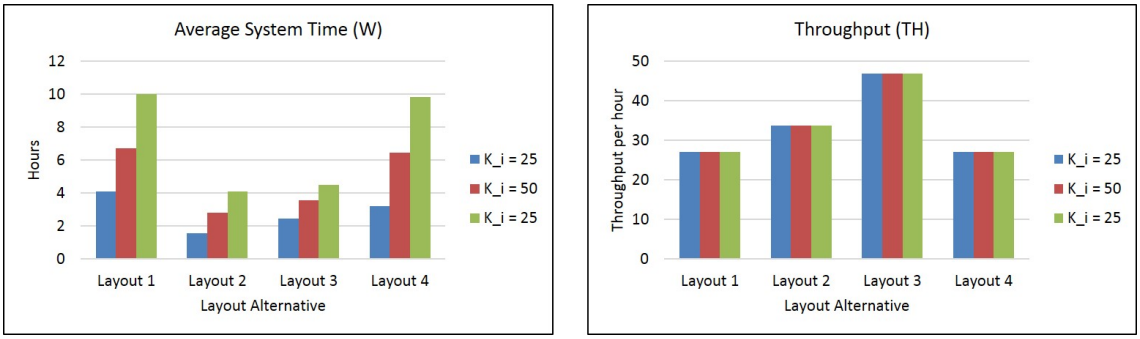

Figure 3.9: Increasing Queue Capacity Effects of the RC

Table 3.6: Numerical Experiment 2: Suffering Objective Results

Layout Design	$N = 25$	$N = 50$	$N = 75$
Layout 1	\$27,734	\$28,838	\$30,777
Layout 2	\$19,713	\$20,160	\$20,676
Layout 3	\$5,473	\$6,099	\$6,711
Layout 4	\$27,436	\$28,720	\$30,670

We observe that, by limiting the number of victims allowed to wait at each queue, we can control congestion and decrease waiting times for victims. The throughput is unaffected due to being bounded by the service rate at the distribution PODs. (Service rate for Tarpaulins is 10.7 per hour while the arrival rate to the RC is 58 victims per hour). Note that, the suffering function captures the cost of victims not served by the RC for all cases. For a setting where the bottleneck is the walkways instead of the distribution PODs, we would expect the throughput to improve as well. An example of this strategy from practice is, after the Nepal Earthquake, Red Cross relief distribution volunteers gated some of their relief centers and allowed victims to enter the RC one by one. On the other hand there were many other RCs which were flooded with victims. Through decision support models, this impact could be quantified and can help humanitarian organizations make informed decisions. By quantifying the effects of allowed capacity with these experiments, we conclude that the best practice is to keep the walkway density between 0.5-0.7 victims/sqm ($N = 25$ for walkway sizes in this experiment).

The key insight to take away from this experiment is the fact that congestion has a high impact on victim waiting times. And when possible, controlling the victim flow within the RC can help manage high waiting times. However, it is important

to note that, controlling the victim flow in practice will depend on how well the RC can be monitored, the available staff to gate the RC and the particular RC layout. Moreover, the extent at which crowding happens given a victim arrival rate, will depend on the overall available RC area and walkway widths. However, given these limitations, models that can quantify impact of decisions under many different scenarios will be helpful. Next, we investigate how best to leverage the triage queue to improve performance.

3.5.4 Effect of Triage

The triage queue is modeled in all designs and its primary function is to check identifications and route victims to correct PODs. In our next experiment, we investigate whether the triage queue can also be used to balance the queuing network and alleviate bottlenecks. To balance the work load in the relief center network, we look at transferring a portion of the work performed at the PODs to the triage queue. In this scenario, the triage queue will keep the victim routing function but might also hand out a common item. We use RC layout design 2, and assume work transfer from the bottleneck queue (K,S) to the triage queue, in increasing percentages.

We observe that work transfer whenever possible can be a very effective strategy to both increase the throughput and decrease the average system time. However, it is crucial to be aware of the risk of making triage the bottleneck in this process. Deciding the amount of work to transfer (or initially assign to triage) is not trivial, mainly because it is not only dependent on the rate of service, but also dependent on the rates of arrival to each queue, and the particular RC design (area and length of queues). Hence, models and fast solution algorithms such as those proposed in Section 5.3, are valuable tools to relief agencies. Figure 3.10 presents the changes

in throughput and average system time as a result of percentage work transfer from service queue distributing kitchen supplies and sleeping mats to the triage queue. The results of the experiments reveal an interesting queuing behavior. We observe that the effects of moving work to triage queue can be characterized in 3 zones. Zone 1 (in this experiment until 25%) increases throughput while decreasing system time. This is due to the common impact of bottleneck service rate increase and decreased blocking. As the transferred work amount increases in Zone 2, the throughput decreases. This decrease in throughput, decreases system congestion, resulting in lower waiting times. Finally in Zone 3, the triage becomes a major bottleneck, hurting both throughput and system time measures. Such threshold behavior is useful to know when making decisions on site to improve RC performance.

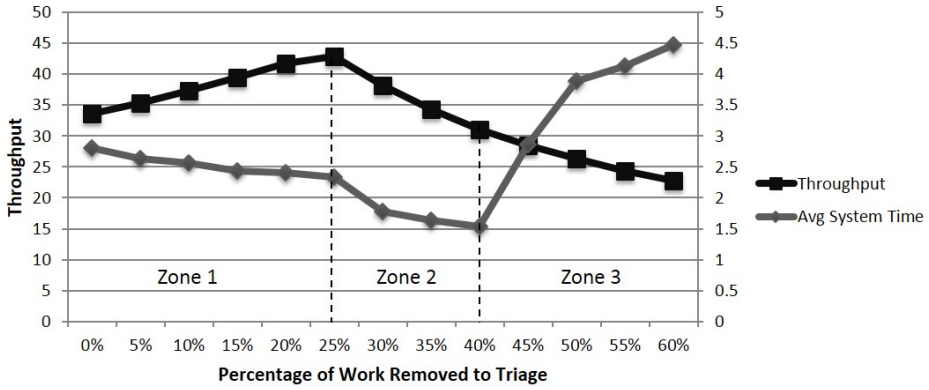

Figure 3.10: Work Transfer to Triage Queue: RC Performance Results

The main limitation for moving work to the triage queue is that, the optimal work distribution might not always be feasible. As an example, moving 20% of tarpaulin distribution time, to the triage queue is not an option. However, distributing water in addition to checking IDs can be practical, can eliminate the need for an additional queue and can speed distribution.

Figure 3.11 provides the suffering objective results for this experiment. Due to the limited available capacity to serve the affected crowd, the current objective puts more emphasis on higher throughput. Therefore, the decrease in throughput in zone 2 is penalized, even with the decrease in waiting time. Appendix 3 in Section 3.5.6 provides a sensitivity analysis on willingness to pay function parameters used in the suffering level calculations, and concludes that the results obtained are not very sensitive to these parameters.

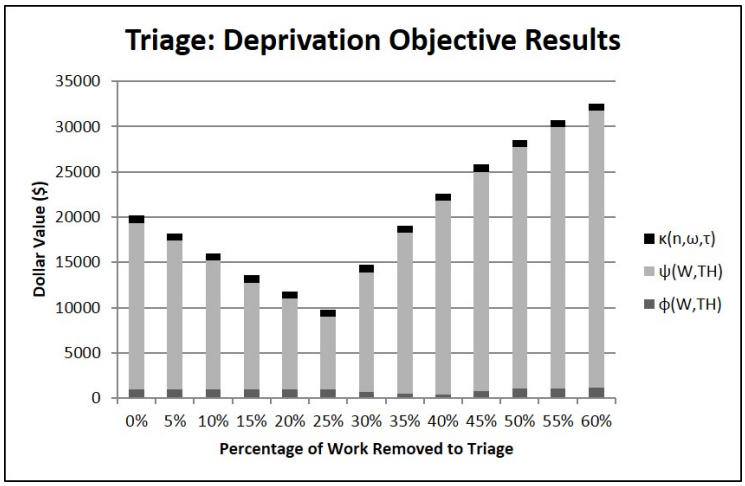

Figure 3.11: Numerical Experiment 3: Suffering Objective Results

3.5.5 Performance Under Increasing Arrival Rates

In previous experiments we used the arrival rate observed at Bhaktapur. However there were districts observing much higher arrival rates due to higher earthquake damage and less available resources. In this section we run experiments for Nuwakot and Kathmandu to understand how the RC performance results would change under heavier arrival rates. Figure 3.12 shows the system time results for Nuwakot and Kathmandu for all 4 layout designs shown in Figure 3.7.

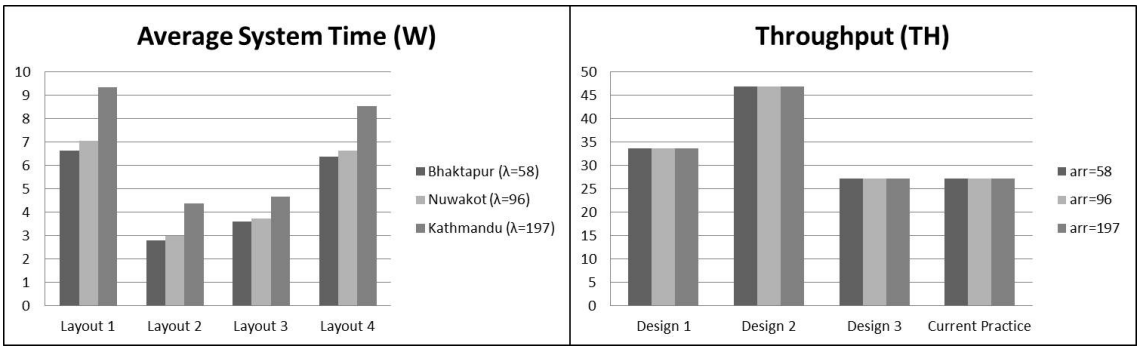

Figure 3.12: Average System Time: Increasing Rates of Arrival

We observe that the throughput remains the same with increased arrivals due to the limit on capacity (note that the service PODs are the bottleneck). Due to finite capacity queues, arrivals beyond capacity are lost, keeping the throughput measure steady. Figure 3.12 shows the average system time change for increasing arrival rates. Increasing rates of arrival, increases the congestion causing the system to be slower, resulting in a much lower performance. Strategies involving allowed queue capacity and using more efficient layouts leveraging crowd dissipation becomes increasingly important for cases such as Kathmandu and Nuwakot.

3.5.6 Impact of Deprivation Function Parameters

In the deprivation objective function, we use a willingness to pay (WTP) function to express waiting (deprivation) time (both for served and not served victims) in monetary terms so that system time, throughput and logistics cost can be expressed with a single measure (dollar value). However, it is important to note that, the resulting measure is dependent on the form of the particular function used for willingness to pay. The function we used for our numerical experiments are based on the willingness to pay function for water defined in Holguin-Veras et al. (2013). Given that the relief items in our data set are tarpaulins and kitchen sets, which are less

sensitive to deprivation we ran experiments to investigate how the objective function results change for a WTP function that is less steep. For the new willingness to pay function we decreased parameter (b) to obtain a less steep exponential curve. Figure 3.13 graphs both functions g_1 and g_2, where g_1 represents the function from Holguin-Veras et al. (2013) and function g_2 represents a more gradual incline.

Figure 3.13: Willingness to Pay Functions

Our numerical studies show that, the only differences we observed happened for the triage experiments. Table 3.7 compares the objective function values for triage experiments for both willingness to pay functions. The steepness of the function used effects the trade off between throughput and average system time. This phenomenon changes the ranking of 5% and 35% work removal to triage scenarios. Under the second exponential function (g_2) the decrease in system time out weighs the loss in throughput.

We conclude that the decisions are not very sensitive to the parameters of the WTP function. What is more important is to use an objective function that takes into account both victim deprivation and labor costs. Moreover the multi-criteria

used by the deprivation objective allows for easier decision making for relief agencies. If future empirical research provides item specific WTP parameters, we expect that these improved estimates will only provide more robust decisions.

Table 3.7: Sensitivity Analysis: Deprivation Objective

Work Transfer to Triage	Function g_1		Function g_2	
	Objective Value	Ranking	Objective Value	Ranking
0%	$20,163.9	8	$ 1,664.3	8
5%	$ 18,187.4	6	$ 1,585.9	7
10%	$ 15,986.0	5	$ 1,501.7	5
15%	$ 13,571.9	3	$1,407.2	3
20%	$11,783.2	2	$1,311.2	2
25%	$9,764.7	1	$ 1,261.8	1
30%	$ 14,715.1	4	$1,422.1	4
35%	$ 19,033.5	7	$1,578.7	6
40%	$ 22,601.7	9	$1,710.0	9
45%	$ 25,796.0	10	$1,872.2	10
50%	$28,550.3	11	$2,002.1	11
55%	$30,711.8	12	$2,085.2	12
60%	$ 32,575.0	13	$2,159.4	13

3.6 Conclusions

In this Chapter, we present a modeling and solution approach for relief centers that distribute aid to victims at the disaster site. We model the relief distribution operations through a relief center using a finite capacity queuing network, which explicitly models effects of crowding via state dependent service rates. To solve this queuing network, we derive new analytic formulas for steady state probabilities of

state dependent Coxian queues, and estimate the throughput of the RC and the waiting time of victims. We use this solution methodology and Nepal earthquake data, to numerically investigate four layout designs. Using these layout designs we analyze critical decision variables that can impact RC performance including; item assignment decisions to points of distribution in the RC, the effects of crowding on RC performance, the impact of physical RC capacity, and the impact of alternative uses for the triage queue. We emphasize that, the methodology of this Chapter is not specific to the layouts discussed here, and it can be used to evaluate the performance of any relief center design.

Our findings from the analysis conducted has significant practical impacts. Firstly, we find that, given a certain level of resource availability (available area, number of volunteers, number of PODs), how the relief center is setup can have a significant impact on performance. Our conversations with practitioners regarding the current practice suggests that, there is a lack of awareness with respect to the impact certain decisions can have of relief distribution performance. We believe studies like this one, that quantify trade-offs and effects of design decisions can impact the way practitioners think about RC setup and operations.

We suggest that practitioners think about ways to dissipate the crowd within the RC area. We find that, even when crowd dissipating designs create more queues for victims to go through, the decrease in congestion can actually shorten the victim waiting times despite the longer route. We also suggest to be mindful of the types of items an RC will be distributing and adjusting the design of the RC accordingly. When an RC is distributing large items such as tarpaulins, one-stop-shop approaches can be more harmful than helpful for performance. We also suggest to

limit physical RC capacity. This can be counter intuitive when lots of victims are waiting to access the relief center. However, due to increasing congestion, allowing as many victims as possible to enter the RC at once will increase the waiting time of victims, not decrease it. Finally, we suggest practitioners to look for ways to balance the workload between the PODs in the RC. We performed numerical experiments using the triage queue to balance the workload, however this is not the only way. The findings apply to balancing the service time between PODs as well.

We believe that the findings of this work will: (1) create awareness for the impact design decisions have on relief distribution performance among practitioners, (2) highlight the need for models that can quantify impact of RC design decisions and (3) highlight some of the essential trade-offs (such as impact of crowd dissipation, impact of item assignment, impact of allowed RC capacity) in RC design. This study also contributes to the body of work that can increase the efficiency of relief distribution during disaster response and relieve some of the suffering caused by disasters.

3.7 Appendix: Validation of the Model

We provide simulation results for the experiments to validate our mathematical model and solution approach in Table 3.8. We use Arena® simulation (version 15.0) of the queuing model for various cases where each run was for 100, 16 hour days and the results show the average of 10 replications. The error for throughput on average was between 0-2% and the errors for system time was on average 1-10%.

Table 3.8: Accuracy of Algorithm

Layout Design	Model Results		Simulation Results	
	$\bar{W}$	TH	$\bar{W}$	TH
Layout 1	6.63 hrs	27.14/hr	6.90 hrs	26.74/hr
Layout 2	2.79 hrs	33.64/hr	3.70 hrs	46.54/hr
Layout 3	3.59 hrs	46.86/hr	3.07 hrs	33.29/hr
Layout 4	6.39 hrs	27.14/hr	6.46 hrs	26.8/hr

Chapter 4

G-network Models for Disaster Relief Distribution

4.1 Introduction

In a disaster affected region, multiple relief centers (RCs) are setup, by a variety of relief agencies, to distribute a variety of supplies. The goal of the relief distribution efforts is to minimize the deprivation experienced by the disaster victims. To do so, the distribution activities should be efficient. In order to be efficient, the execution of response activities need to be agile and adapt to the inherent variability of the disaster aftermath conditions. Operations research models can help identify opportunities for improvement in efficiency of response activities, as well as provide decision support to enable efficient response to changing needs. In this Chapter we model the relief distribution operations as a generalized queuing network (G-network). The contributions of this research are twofold. First, it contributes to the disaster response literature by providing a flexible, evaluative model to investigate relief distribution efficiency. The G-network framework provides a flexible struc-

ture to model multiple RCs, changes in demand for items, victim needs, and victim movement within the network. Changes in demand for items is a frequent occurrence because the number of victims that need a particular item (ex: blankets) can increase due to changing conditions (ex:weather). Changes in victim needs refers to the change in the types of items a particular victim needs. This often happens because victims waiting in line at an RC collect items as a family, or they may obtain items via trading with other victims. Changes in victim movement, refer to victims jockeying between queues based on their perception of shorter queues or more availability of supplies at another RC.

The second main contribution of this research is that it contributes to the queuing literature by proving a new product form result for G-networks and proposing an efficient product form approximation for G-networks based on this result. In addition to providing a flexible framework, G-networks also have attractive solution efficiencies. In particular, the existence of product form solutions is of interest, as it can enable rapid performance evaluation. This becomes especially useful for large relief distribution networks where alternative methods like solving a Markov chain or a simulation model become computationally challenging. However, in the literature, there is no study that investigates product form result for G-networks with batch departures and batch transfers. In this chapter, we show that product form for such networks exist under certain conditions and approximations can be used for cases where such conditions are not met.

The rest of this chapter is organized as follows. Section 4.2 provides a literature review of G-networks, Section 5.2 describes the G-network model of relief centers. Section 4.4 investigates the product form results while Section 4.5 proposes product

form approximations for cases where exact results fail to hold. Section 5.4 discusses numerical experiments that demonstrate the use of the proposed model.

4.2 G-network Literature Review and Gaps

The G-network framework can be a very useful way of modeling the queuing behavior during relief distribution due to the flexibility it provides in modeling and the ease of performance estimation using product form results. G-networks differ from other queuing networks such as BCMP and Jackson networks (Jackson (1963) and Baskett et al. (1975)), as it allows for negative and/or signal entity arrivals. In a G-network, the signal entities are assumed to arrive to the system according to a Poisson process in addition to the regular customers. An arriving signal can (a) remove work from a queue or (b) transfer work between queues in the network. We use routing parameters and signals of the G-network to model the dynamic changes in victim movement between RCs (due to changes in demand and in individual needs) and interventions by the relief agencies that aim at improving performance.

The articles Bocharov and Vishnevskii (2003) and Artalejo (2000) provide detailed surveys on G-network models. Gelenbe et al. (1991) and Gelenbe (1991) introduced the first product form results for G-networks, followed by Gelenbe (1993b) analyzing G-networks with signals. The G-network theory has later been extended to cover multi-class arrivals (Gelenbe and Labed, 1998), generally distributed service times (Harrison and Pitel, 1996), state dependent service times (Bocharov et al., 2004a,b) and tandem queues with blocking (Gomez-Corral, 2002). In this paper, the G-network model of relief centers allows single or batch transfer of victims between queues in the network triggered by a signal arrival. The literature has shown exis-

tence of product form results for networks where signals only cause batch removal of customers from the network (Gelenbe (1993a)), networks with batch service (Chao and Pinedo (1995)) and networks with batch arrival and batch service (Miyazawa and Taylor (1997)). To our knowledge there is no prior work investigating existence of product form solution for a generalized queuing network with signals causing batch transfers of customers between queues in the network. This chapter attempts to fill this gap.

4.3 The G-network Model of Relief Distribution

At a disaster affected region, multiple agencies distribute aid by setting up one or more relief centers (RCs), each distributing a set of supplies. In this paper we model each RC as a single server queue and the network of RCs as a G-network.

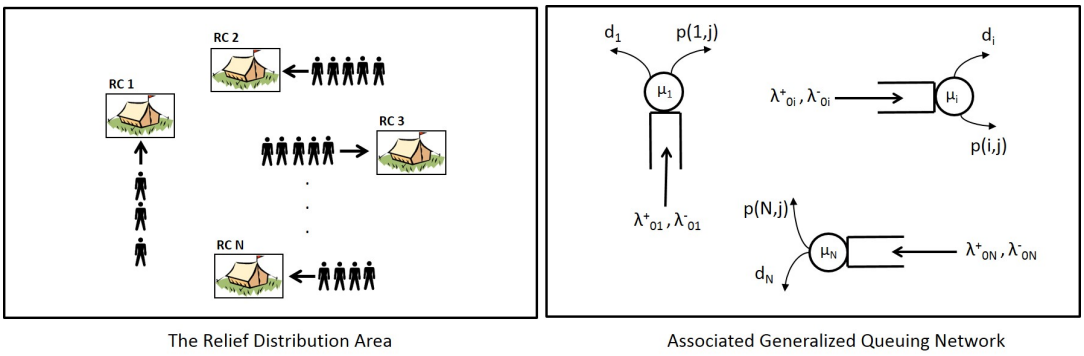

Figure 4.1: The G-network Representation of Relief Distribution Efforts

The G-network has 2 types of entities (victims and signals). We assume that victims and signals arrive to any queue i, $i = 1, 2, ..., N$ in the network from outside according to a Poisson process with parameter λ_{0i}^{+} and λ_{0i}^{-} respectively. While the arrival of a victim increases the queue length by one, the arrival of a signal simply triggers movement and hence does not add to queue length. An arriving

victim queues to eventually receive service at an RC. We assume that this service time is characterized by an exponential distribution with parameter μ_i. Upon service completion, victims can either depart the network (with probability $d(i)$), or move to another queue j, $j \neq i$ for additional needs (with probability $p^+(i,j)$), where $d(i) = 1 - \sum_{j=1}^{N} p^+(i,j)$. An arriving signal to queue i can either trigger a movement of a batch of victims from queue i to another queue j or force a batch of victims to depart the system altogether, with probabilities $q(i,j)$ and $D(i)$ respectively where $D(i) = 1 - \sum_{s=1}^{N}(q(i,j))$. Note that for the relief distribution network modeling, we assume $\sum_{s=1}^{N}(q(i,j)) = 1$ or equivalently $D(i) = 0, \forall i$. This is based on the assumption that no victim will leave the system without receiving any supplies or be denied service. The batch size of victims impacted by a signal is random and drawn from a probability distribution defined by $P\{B_i = s\} = \pi_{is}$, $s = 0, 1, 2, ..., \infty$. We assume that a signal arriving to a queue that has less victims than the batch size B_i will have no effect on queue i. We will refer to this assumption as the *full batch transfer* assumption. Figure 4.1 summarizes the probabilistic routing for victims and the effect of signals.

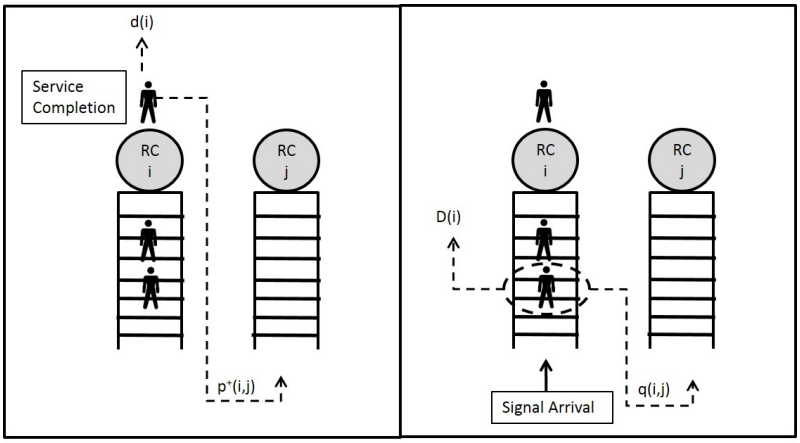

Figure 4.2: Service Completion Routing and Signal Routing Probabilities

The described framework can be used to model a variety of scenarios that occur in relief distribution including; (i) increase in demand for a particular item, (ii) change in the need of individual victims, (iii) victim jockeying between RCs. We use the setting in Figure 4.2 to explain each scenario in detail.

(i) Increase in Demand: Assume items $\{Blanket, Water\}$ are being distributed at RC i and items $\{Tarpaulin, KitchenSets\}$ are being distributed at RC j. Changes in the area conditions, such as an earthquake aftershock damaging more buildings and increasing the likelihood for more victims needing tarpaulins, can be modeled by an increase in the $p^{+}(1,2)$ probability.

(ii) Changes in Individual Victim Needs: Often victims queue at an RC to collect items for their families. As family needs change or as different family members collect or trade items, the needs of a victim in queue can vary with time. For instance, a victim waiting at RC i to collect blankets and water, might quit RC i and moves to queue at RC j to collect a kitchen set instead. This behavior can be modeled by external signal arrivals moving victims to another RC.

(iii) Victim Jockeying: Jockeying happens when a victim waiting in queue at one RC, moves to another RC to collect the same item, expecting a shorter waiting time. This often happens due to the perception of shorter queues, better supply availability, or lower congestion at an another RC. This perception does not always coincide with reality, but it can create significant victim movement. This behavior can be modeled using signals, that cause victims to move between RCs prior to service.

The ability to include such behaviors in models of relief distribution is important to provide realistic estimates of RC performance. However, G-networks with signals that cause batch transfer of victims are not guaranteed to have product form results. In Section 4.4 we show product form result for G-networks with full batch transfer of victims. Then in Section 4.5 we relax the full batch assumption and investigate the accuracy of product form approximation for cases where transfer of partial batches are allowed. Finally in Section 5.4 we conduct numerical experiments to model scenarios (i), (ii) and (iii) described above. We investigate the effects of changes in demand, changes in victim needs, and victim jockeying on RC performance. The numerical experiments demonstrate the importance of modeling the dynamic aspects of the disaster environment and provide high level insights for practice.

4.4 Product Form Result for G-Networks with Full Batch Transfer

In this section, we derive product form results for a G-network with signals that trigger a full batch transfer. Consider the G-network shown in Figure 4.1, with N RCs, each modeled as an individual queue in the network.

Based on the assumptions (arrival of victims, signals and the service process), the time evolution of this G-network can be characterized as a time-homogeneous Continuous Time Markov Chain. We represent the state of the system by $\{k(t) : t \geq 0\}$ where $k(t) = (k_1(t), ..., k_N(t)) \in S$, where each $k_i(t),\ i = 1, 2, ..., N$ represents the number of victims at queue i, at time t and S represents the state space. Based on the vector $k(t)$, we define the vectors $k_i^+(t)$, $k_i^-(t)$,$k_{i,j}^{+-}(t)$, $k_{i,j}^{++B_j}(t)$ and $k_{i,j,m}^{++B_j-B_j}(t)$ as follows:

$$
\begin{aligned}
k_i^+(t) &= (k_1(t), \ldots, k_i(t)+1, \ldots, k_n(t)) \\
k_i^-(t) &= (k_1(t), \ldots, k_i(t)-1, \ldots, k_n(t)) \\
k_{i,j}^{+-}(t) &= (k_1(t), \ldots, k_i(t)+1, \ldots, k_j(t)-1, \ldots, k_n(t)) \\
k_{i,j}^{++B_j}(t) &= (k_1(t), \ldots, k_i(t)+1, \ldots, k_j(t)+B_j, \ldots, k_n(t)), B_j \geq 1 \\
k_{i,j,m}^{++B_j-B_j}(t) &= (k_1(t), \ldots, k_i(t)+1, \ldots, k_j(t)+B_j, \ldots, k_m(t)-B_j, \ldots), B_j \geq 1
\end{aligned}
\tag{4.1}
$$

To determine the steady state probabilities of this Markov chain, we first define the effective rate of victim arrivals λ_i^+ and the effective rate of signal arrivals λ_i^- to each queue i. Let q_i be the utilization of queue i, then Equation 4.2 provides a function for q_i.

$$
q_i = \frac{\lambda_i^+}{\mu_i + \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^{s-1}} \tag{4.2}
$$

Equations 4.3 and 4.4 provide the associated non-linear equations that define the effective rate of arrivals for victims and signals to each queue $i, i = 1, ..., N$. Effective signal arrival rate is equal to the external arrival rate, while the effective victim arrivals to queue i is the sum of external arrivals and batch transfers from other queues j to queue i.

$$\begin{aligned}
\lambda_i^+ &= \lambda_{0i}^+ + \sum_j \mu_j q_j p^+(j,i) \\
&+ \sum_j \lambda_{0j}^- q(j,i) \sum_{s=1}^{\infty} \pi_{is} s q_j^s \\
&= \lambda_{0i}^+ + \sum_j \mu_j q_j p^+(j,i) + \sum_j \lambda_j^- q(j,i) \sum_{s=1}^{\infty} \pi_{is} s q_j^s \qquad (4.3) \\
\lambda_i^- &= \lambda_{0i}^- \qquad (4.4)
\end{aligned}$$

The Chapman-Kolmogorov equations for state k is given by Equation 5. The left hand side represents the rate out from the state k and the right hand side represents the rate in to state k from all other possible states. The indicator functions, $1_{[X]} = 1$ if X is true and 0 otherwise, characterize the balance equations for different subsets of the state space.

$$\begin{aligned}
p(k) \sum_i \left[\lambda_{0i}^+ + \lambda_{0i}^- 1_{[k_i \geq s]} + \mu_i 1_{[k_i > 0]}\right] &= \sum_i p(k_i^+) \mu_i d(i) \qquad (4.5) \\
&+ \sum_i p(k_i^-) \lambda_{0i}^+ 1_{[k_i > 0]} \\
&+ \sum_i \lambda_{0i}^- D(i) \sum_{s=1}^{\infty} \pi_{is} p(k_i^{+s}) \\
&+ \sum_i \sum_j \left[p(k_{ij}^{+-}) \mu_i p^+(i,j) 1_{[k_j > 0]}\right] \\
&+ \sum_i \sum_j \sum_{s=1}^{\infty} \pi_{is} \left[p(k_{ij}^{+s-s}) \lambda_{0i}^- q(i,j) 1_{[k_j \geq s]}\right]
\end{aligned}$$

We next present an important product form result for this Markov chain.

Theorem 1. *For a G-network with victims and signals that cause a full batch of victims, the steady state distribution $p(k)$ can be represented as the product of the marginal probabilities, i.e. $p(k) = \prod_{i=1}^{N} p(k_i)$ where $p(k_i) = (1 - q_i)(q_i^{k_i})$, $k_i \geq 0$ if $q_i = q_j < 1, for\, distinct\, pairs\, (i, j)$.*

Proof. We show this result under the assumption that $P\{B_i = s\} = \pi_{is}$ and $\sum_j q(i, j) = 1, \forall\ (i, j),\ i \neq j$. To prove the product form result holds, we show that the solution to Equation 5 can be expressed as the product of marginal probabilities. Next, we partition the state space S, into 3 subsets S_1, S_2, S_3 (see Figure 4.3).

1. $S_1 = \{k_1(t), ..., k_i(t), ..., k_N(t)\} : k_i(t) \geq s \quad \forall i, \quad i = 1, 2, ..., N$
2. $S_2 = \{k_1(t), ..., k_i(t), ..., k_j(t), ..., k_N(t)\}$: $k_i(t) < s$, for at least one i and $k_i(t) \neq 0\ \forall i \quad i, j \in \{1, 2, .., N\}$
3. $S_3 = \{k_1(t), ..., k_i(t), ..., k_j(t), ..., k_N(t)\} : k_i(t) = 0$, for at least one i, $i \in \{1, 2, .., N\}$

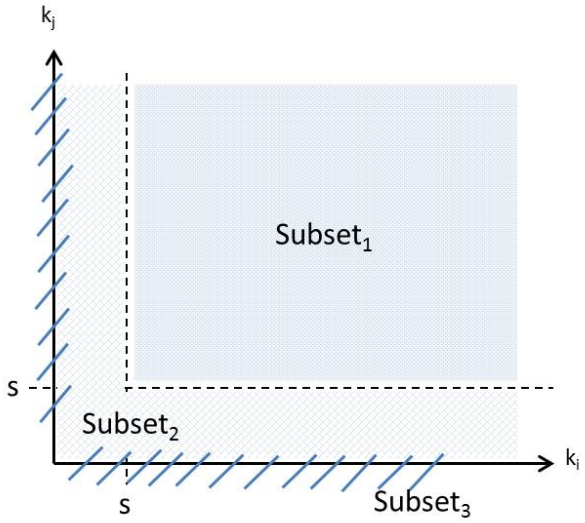

Figure 4.3: Subsets of the State Space

Then, using proof by substitution, we show that the product form result holds for each subset and for each state k in each subset of the state space. We provide the detailed steps of the proof for S_1 below. The details of the proof for states in S_2 and S_3 are given in the Appendix.

Proof for state $k \in S_1$: We prove the product form result by substitution and using the definitions of q_i, λ_i^+ and λ_i^-. The balance equations in Equation 5 take the form given in Equation 4.6 for the states in S_1.

$$
\begin{aligned}
p(k)\sum_i \left[\lambda_{0i}^+ + \lambda_{0i}^- + \mu_i\right] &= \sum_i p(k_i^+)\mu_i d(i) \\
&+ \sum_i p(k_i^-)\lambda_{0i}^+ \\
&+ \sum_i \lambda_{0i}^- D(i) \sum_{s=1}^{\infty} \pi_{is} p(k_i^{+s}) \\
&+ \sum_i \sum_j \left[p(k_{ij}^{+-})\mu_i p^+(i,j)\right] \\
&+ \sum_i \sum_j \sum_{s=1}^{\infty} \pi_{is} p(k_{ij}^{+s-s})\lambda_{0i}^- q(i,j) \qquad (4.6)
\end{aligned}
$$

Assuming the product form result holds for all states in S_1, these balance equations reduce to Equation 4.7 below.

$$
\begin{aligned}
\sum_i \left[\lambda_{0i}^+ + \lambda_{0i}^- + \mu_i\right] &= \sum_i \mu_i q_i d(i) \\
&+ \sum_i \frac{\lambda_{0i}^+}{q_i} \\
&+ \sum_i \lambda_{0i}^- D(i) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&+ \sum_i \sum_j \frac{\mu_i q_i p^+(i,j)}{q_j} \\
&+ \sum_i \sum_j \lambda_{0i}^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{q_i^s}{q_j^s} \qquad (4.7)
\end{aligned}
$$

$$
\sum_i \frac{\lambda_{0i}^+}{q_i} + \sum_i \sum_j \frac{\mu_j q_j p^+(j,i)}{q_i} = \sum_i \frac{\lambda_i^+}{q_i} - \sum_i \sum_j \sum_{s=1}^{\infty} \pi_{js} \frac{\lambda_j^- q(j,i) s q_j^s}{q_i} \qquad (4.8)
$$

We simplify the balance equations (Equation 4.7) by substituting Equation 4.8 obtained from Equation 4.3. Then we use Equation 4.8 and the definition of λ_i^- given by Equation 4.4 to obtain Equation 4.9.

$$\begin{aligned}
\sum_i \left[\lambda_{0i}^+ + \lambda_{0i}^- + \mu_i\right] &= \sum_i \mu_i q_i d(i) \\
&+ \sum_i \lambda_i^- D(i) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&+ \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{q_i^s}{q_j^s} \\
&- \sum_i \sum_j \sum_{s=1}^{\infty} \pi_{js} \frac{\lambda_j^- q(j,i) s q_j^s}{q_i} \\
&+ \sum_i \frac{\lambda_i^+}{q_i} \qquad (4.9)
\end{aligned}$$

Next we insert definitions for λ_i^+ and λ_i^- (Equations 4.3 and 4.4) into the left hand side (LHS) terms and $D(i)$ and $d(i)$ on the right hand side (RHS), to obtain Equation 4.10.

$$\begin{aligned}
\sum_i \lambda_i^+ - \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s + \sum_i \mu_i + \sum_i \lambda_i^- &= \sum_i \mu_i q_i \\
&+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&+ \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{q_i^s}{q_j^s} \\
&- \sum_i \sum_j \frac{\lambda_j^- q(j,i) \sum_{s=1}^{\infty} \pi_{js} s q_j^s}{q_i} \\
&+ \sum_i \frac{\lambda_i^+}{q_i} \qquad (4.10)
\end{aligned}$$

Next, we use the definition of q_i (from Equation 4.2) to rewrite λ_i^+ on both the LHS and the RHS. We also interchange i and j in the 5th term, using the fact that both i and j are in the set $\{1, 2, ..., N\}$. Then the balance equation (Equation 4.10) can be written as Equation 4.11.

$$\begin{aligned}
0 &= \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{q_i^s}{q_j^s} + \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^{s-1} \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{s q_i^s}{q_j} + \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s + \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s \\
&- \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^s - \sum_i \lambda_i^- \qquad (4.11)
\end{aligned}$$

Combining common terms we get Equation 4.12 and further simplify to get

Equation 4.13.

$$\begin{aligned} 0 &= \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \left[\frac{q_i^s}{q_j^s} - \frac{sq_i^s}{q_j} - q_i^s + sq_i^s \right] \\ &+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} \left[sq_i^{s-1} + q_i^s - sq_i^s - 1 \right] \end{aligned} \tag{4.12}$$

$$0 = \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} \left[sq_i^{s-1} - sq_i^{s-1} + q_i^s - q_i^s - sq_i^s + sq_i^s - 1 + 1 \right] \tag{4.13}$$

It can be seen that the equation holds for balance networks where $q_i = q_j$, $\forall (i,j) \in N$ given that $\sum_j q(i,j) = 1$. This completes the proof for state subset S_1. For subsets S_2 and S_3 the proof starts by writing the balance equations for a state in the set and using a similar process. The details of the proof are provided in the Appendix.

□

4.5 Product Form Approximations for Unbalanced Networks and Partial Batch Transfer

Theorem 1 states that the steady state probabilities of a G-network can be expressed in product form under assumptions of full batch transfer and balanced utilizations. However in many practical cases, the utilization of queues in the network need not be balanced, and partial batch transfers may occur.

However, the product form result fails to hold for unbalanced networks and

when the full batch assumption is relaxed. However, our observations indicate that the product of the marginal probabilities can provide a good approximation for the steady state probabilities of these queuing networks. Table 4.1 summarizes the product form results for G-networks with signals and batch transfer. We numerically analyze the accuracy of the product form approximation for unbalanced networks ($q_i \neq q_j$) both under the full batch assumption and when its relaxed. We investigate how the accuracy of the approximation changes when the difference of utilization between the queues increases.

Table 4.1: Product Form Results Summary

	Balanced: $q_i = q_j$, $i \neq j$	Unbalanced: $q_i \neq q_j, i \neq j$
Full Batch	Product form holds	Product form does not hold
Partial Batch	Product form does not hold	Product form does not hold

4.5.1 Product Form Approximation: Full Batch Transfer in Unbalanced Networks

We numerically investigate the accuracy of the product form result for unbalanced networks where the utilization of queues in the network are not equal. Let $p(k)$ represent the actual steady state probability vector of the unbalanced G-network with full batch transfer of victims and let $\hat{p}(k)$ represent the estimated steady state probability vector calculated using the product of marginal probabilities $\hat{p}(k) = \prod_{i=1}^{N} \hat{p}(k_i)$ *where* $\hat{p}(k_i) = (1 - q_i)(q_i^{k_i})$, $k_i \geq 0$ and $q_i < 1, \forall i$, where $i = 1, 2, ..., N$.

Once we obtain $p(k)$ and $\hat{p}(k)$ we compute the throughput (TH) and waiting

time (W) performance measures using Equation 4.16. We compare these performance measures obtained via the product form approximation ($T\hat{H}_i, \hat{W}_i$) and by solving the Markov chain directly (TH_i, W_i) to assess accuracy of the approximation. We compute both the absolute and percentage errors as defined by Equations 4.17 and 4.18 respectively.

$$\begin{aligned} TH_i &= \mu_i q_i, \forall i = 1, .., N && (4.14) \\ L_i &= \sum_{n \in S} P_i(k_i) k_i && (4.15) \\ \bar{W}_i &= L_i / TH_i && (4.16) \end{aligned}$$

$$\begin{aligned} E_{a,i}(TH) &= |TH_i - T\hat{H}_i|, i = 1, 2, .., N \\ E_{a,i}(W) &= |W_i - \hat{W}_i|, i = 1, 2, .., N \end{aligned} \quad (4.17)$$

$$\begin{aligned} E_{p,i}(TH) &= \frac{|TH_i - T\hat{H}_i|}{TH_i}, i = 1, 2, .., N \\ E_{p,i}(W) &= \frac{|W_i - \hat{W}_i|}{W_i}, i = 1, 2, .., N \end{aligned} \quad (4.18)$$

To investigate the accuracy of the approximation for unbalanced networks with full batch transfers, we analyze a two queue network where each queue has one volunteer distributing supplies at the rate of 32 items per hour. We vary the overall arrival rate as $\lambda = 53, 50, 47, 44$ victims per hour and use the parameter $\beta = 0.60, 0.56$ to

divide the arrivals between the two queues in the network. This way we can vary the utilization difference between the two queues. The experiment sets obtained by λ, β combinations are summarized in Table 4.2. In each experiment set, we introduce signal arrivals to queue 1 and vary the signal arrival rate and the batch size the signal affects as $(\lambda_1^-, B_1) = (1,2),\ (1,4),\ (1,6),\ (1,8),\ (2,1),\ (4,1),\ (6,1),\ (8,1),\ (2,2)$ $(2,3),\ (3,2)$. In total we analyze 88 experiments and report on the errors observed.

Table 4.2: Full Batch Transfer in Unbalanced Networks

	$\beta = 0.60$	$\beta = 0.56$
$\lambda = 53/hr$	Experiment Set 1 $(\rho_1, \rho_2) = (0.99, 0.66)$	Experiment Set 2 $(\rho_1, \rho_2) = (0.93, 0.73)$
$\lambda = 50/hr$	Experiment Set 3 $(\rho_1, \rho_2) = (0.94, 0.63)$	Experiment Set 4 $(\rho_1, \rho_2) = (0.88, 0.69)$
$\lambda = 47/hr$	Experiment Set 5 $(\rho_1, \rho_2) = (0.88, 0.59)$	Experiment Set 6 $(\rho_1, \rho_2) = (0.82, 0.65)$
$\lambda = 44/hr$	Experiment Set 7 $(\rho_1, \rho_2) = (0.83, 0.55)$	Experiment Set 8 $(\rho_1, \rho_2) = (0.77, 0.61)$

The errors from the 8 experiment sets are reported in Table 4.3, where throughput (TH_i) is reported as victims per hour and the waiting time (W_i) is reported as hours.

Table 4.3: Full Batch Transfer in Unbalanced Networks

	Absolute Error Range				Percent Error Range			
	$TH1$	$TH2$	$W1$	$W2$	$TH1$	$TH2$	$W1$	$W2$
Experiment Set 1	(0, 0.26)	(0, 0.26)	(0, 0.03)	(0, 0.04)	(0%, 1%)	(0%, 1%)	(0%, 15%)	(0%, 24%)
Experiment Set 2	(0, 0.22)	(0, 0.22)	(0, 0.02)	(0, 0.03)	(0%, 1%)	(0%, 1%)	(0%, 11%)	(0%, 15%)
Experiment Set 3	(0, 0.24)	(0, 0.24)	(0, 0.02)	(0, 0.03)	(0%, 0.8%)	(0%, 1%)	(0%, 12%)	(0%, 20%)
Experiment Set 4	(0, 0.11)	(0, 0.11)	(0, 0.01)	(0, 0.01)	(0%, 0.4%)	(0%, 0.4%)	(0%, 7%)	(0%, 12%)
Experiment Set 5	(0, 0.18)	(0, 0.18)	(0, 0.01)	(0, 0.01)	(0%, 1%)	(0%, 1%)	(0%, 9%)	(0%, 16%)
Experiment Set 6	(0, 0.12)	(0, 0.12)	(0, 0)	(0, 0.01)	(0%, 1%)	(0%, 1%)	(0%, 7%)	(0%, 10%)
Experiment Set 7	(0, 0.13)	(0, 0.13)	(0, 0)	(0, 0.01)	(0%, 1%)	(0%, 1%)	(0%, 7%)	(0%, 12%)
Experiment Set 8	(0, 0.08)	(0, 0.08)	(0, 0)	(0, 0)	(0%, 0%)	(0%, 0%)	(0%, 5%)	(0%, 8%)

Note that, we report the minimum and maximum errors measured, both in absolute and percentage terms as defined by Equations 4.17 and 4.18. From these results, we observe that the percentage errors are between 0 and 20%, while the absolute errors are small throughout all experiments. We also observe that the errors increase as the batch size increases for a given signal arrival rate and as the signal arrival rate increases, for a given batch size. Finally, we observe that as utilization of queues get closer the errors decrease. Over all, we conclude that for instances where solution speed has priority over high accuracy, the product form approximation for unbalanced networks can perform well. We provide the detailed results of all 8 experiment sets in the Appendix.

4.5.2 Product Form Approximation: Partial Batch Transfer in Unbalanced Networks

In this section, in addition to unequal utilization, we also relax the full batch transfer assumption and numerically investigate the errors for performance measures. The experiment parameters remain the same as given in Table 4.2. The errors from the 8 experiment sets are reported in Table 4.4.

From these results, we observe that the percentage errors are between 0 and 28%, while the absolute errors are small throughout all experiments. We observe that the error percentages are higher for the partial batch transfer case. However, the behavior of the errors (increase with higher batch size, increase with higher signal rate, decrease for closer utilization levels) remain the same. We provide the detailed results of all 8 experiment sets in Appendix 2.

Table 4.4: Partial Batch Transfer in Unbalanced Networks

	Absolute Error Range				Percent Error Range			
	$TH1$	$TH2$	$W1$	$W2$	$TH1$	$TH2$	$W1$	$W2$
Experiment Set 9	(0, 0)	(0, 0)	(0, 0)	(0, 0.05)	(0%, 0%)	(0%, 0%)	(0%, 1%)	(0%, 28%)
Experiment Set 10	(0, 0)	(0, 0)	(0, 0)	(0, 0.05)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 21%)
Experiment Set 11	(0, 0)	(0, 0)	(0, 0)	(0, 0.04)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 25%)
Experiment Set 12	(0, 0)	(0, 0)	(0, 0)	(0, 0.02)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 15%)
Experiment Set 13	(0, 0)	(0, 0)	(0, 0)	(0, 0.02)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 22%)
Experiment Set 14	(0, 0)	(0, 0)	(0, 0)	(0, 0.02)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 17%)
Experiment Set 15	(0, 0)	(0, 0)	(0, 0)	(0, 0.02)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 19%)
Experiment Set 16	(0, 0)	(0, 0)	(0, 0)	(0, 0.01)	(0%, 0%)	(0%, 0%)	(0%, 0%)	(0%, 14%)

We conclude that the product form result can be used as an approximation method to calculate performance measures of G-networks with batch transfers and departures. Next, we apply the G-network model using the product form approximation, to model scenarios (i) demand increase, (ii) changes in individual needs and (iii) jockeying in Section 5.4.

4.6 Performance Evaluation Using a Case Study

In this section we use the G-network formulation to evaluate the relief distribution efforts that took place during the 2015 Nepal earthquake. We first introduce the case and discuss the data analysis efforts to estimate arrival and service rate parameters at relief centers in Nepal. We summarize the parameter settings in Section 4.6.2. Next, we analyze scenarios where the demand for items increase (see Section 4.6.3), individual victim needs change (see Section 4.6.4), and victims jockey (see Section 4.6.5. Finally, in Section 4.6.6 we analyze a scenario where item assignment to RCs change.

4.6.1 Disaster Scenario Based on 2015 Nepal Earthquake

During the aftermath of the Nepal earthquake relief efforts had to support three million displaced people, making relief item distribution a top priority during the immediate response. The data we analyzed included all relief efforts. However, for the purposes of this chapter, we focus on the Thumi, a small district of the Gorkha municipality.

The arrival rate of victims to the RCs and service rate distribution of items are required input parameters to the G-network. To estimate the service rates, we use data

from the Humanitarian Data Exchange Database (HDX:https://data.humdata.org/). We use the distributed supplies data set (Meta Data Source: https://data.hdx.rwlabs.org/dataset/scnepal-agency-data) to extract data on types of items distributed, the distributing agency, number of households served and the dates the distribution took place. The data set is prepared by the Global Shelter Cluster (GSC) which is a coordinating agency that brings together 35 global partners to coordinate and respond to disasters and conflict situations. The GSC is co-chaired by the International Federation of Red Cross (IFRC) and the United Nations High Commissioner for Refugees (UNHCR).

Service Time Estimation: We estimate the time it takes to distribute an item of a particular type using the data set of distributed supplies. The data set contains the number of households served in a given day for an agency, district, item triplet. From this data, we estimate number of households served per day for a given item type. To calculate the distribution time per item, in addition to our estimate of number of households served per day for a given item type, we use data on the number of items (for a given item type) needed by a household, for example 2.5 blankets per household. Using this additional data, we calculate the time it takes to distribute a specific item. We use calculated maximum, minimum and average values obtained by analyzing 10 most affected districts in Nepal. Table 4.5 shows the obtained service rate estimates. In our numerical experiments we use the maximum service rates observed, since the distribution rates can be skewed down due to supply availability in regions.

Table 4.5: Service Rate per Item

Item	Minimum	Average	Maximum
Blanket	14.8/hr	31.6/hr	59.5/hr
Tarpaulin	4.9/hr	10.7/hr	19.1/hr
Cash Voucher	1/hr	29/hr	64.3/hr

Arrival Rate Estimation: To estimate the arrival rates of victims in need of various supplies, we use the demographics and casualty data set (Meta Data Source: https://data.hdx.rwlabs.org) which provides population, number of households, total number of deaths and injured victims per district data. This data is provided by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) in Nepal with data contributions from the Nepal Ministry of Home Affairs and the Nepali Police. To determine the population in need of supplies from the general population we also use the affected districts and displaced population data provided by the Center for Disaster Management and Risk Reduction Technology (CEDIM Forensic Disaster Analysis Group and South Asia Institute (SAI) (2015)).

From these data sets we know that Thumi has a population of 4500 people, making up 900 households, most of which were badly damaged or destroyed. We focus our analysis to the 5 days immediately following the disaster and we assume 8 hours of distribution time per day. Given the whole population was impacted, we calculate an hourly victim arrival rate of $\lambda = 112$ per hour.

4.6.2 Thumi Distribution Base Case

From 2015 Nepal earthquake data, we have estimated victim arrival and item distribution rates. Therefore, the known parameters are, the total arrival rate of victims, the set of items distributed in Thumi, and the distribution rate for these items. Based on these parameters, and assuming single staffing for distribution at each RC, we construct the base case with 5 relief centers, distributing tarpaulins, blankets and cash vouchers as shown in Figure 4.4. We use the maximum distribution rate for each item given by Table 4.5 as service rate at the RCs and we use 112/hr as the overall arrival rate λ to the network.

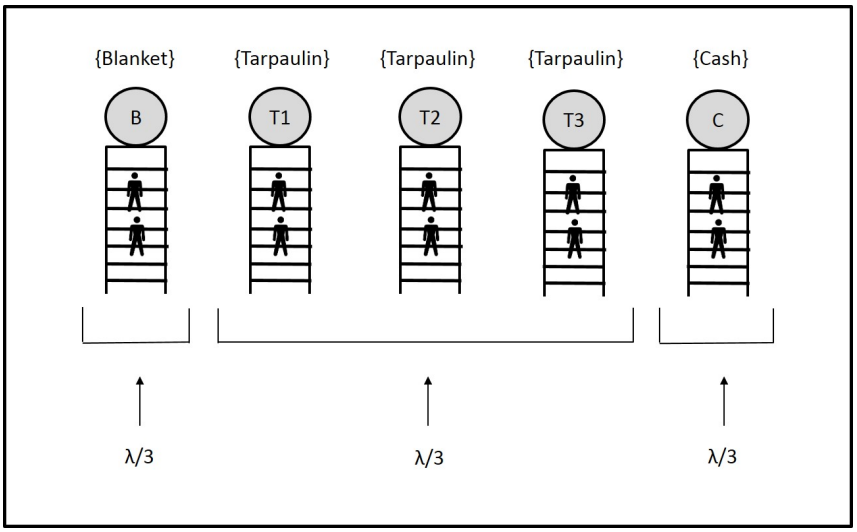

Figure 4.4: Thumi Relief Distribution: Base Case

For the base case, we assume that the arriving victims chose to queue for items blanket (B), tarpaulin (T) or cash (C) with equal probability, and they move to queue for another item following service completion based on the $p^{+}(i, j)$ probabilities given in Table 4.6. In this table B, T1, T2, T3 and C refer to the item being distributed and correspond to relief centers 1, 2, 3 and 4 respectively as shown in Figure 4.4. Note that if the same item is available at multiple RCs (as in the case of

tarpaulins), choosing any of these RCs is equally likely. Hence, the probability of a victim collecting a tarpaulin following the blanket queue is equally divided between all three tarpaulin queues. Also note that, the probabilities in Table 4.6 were chosen to have roughly 80% utilization for all queues in the network.

Table 4.6: Victim Routing Probabilities Between RCs

	RC1: B	RC2: T1	RC3: T2	RC4: T3	RC5: C
RC1: B	0	0.066	0.066	0.066	0.1
RC2: T1	0.2	0	0	0	0.1
RC3: T2	0.2	0	0	0	0.1
RC4: T3	0.2	0	0	0	0.1
RC5: C	0.05	0.0166	0.0166	0.0166	0

Note that based on the $p^+(i,j)$ probabilities, there are 7 combinations of items a victim can collect. Using the routing probabilities given in Table 4.6 we can compute the expected number of victims that belong to each one of these categories. Note that the probabilities given in Table 4.6 are routing dependent while the expectations given in Table 4.7 show the total number of victims that collected the given subset of items, regardless of the sequence in which they collected them. Computing the expected demand for each combination of items will help update the service and arrival rates to each RC, if the item assignment to relief centers changes.

Table 4.7: Expected Demand for Each Victim Category

Victim Category	Expected Demand
$\{B\}$	26.6 victims/hr
$\{T\}$	25.6 victims/hr
$\{C\}$	33.4 victims/hr
$\{B,T\}$	15.2 victims/hr
$\{B,C\}$	5.7 victims/hr
$\{T,C\}$	5.7 victims/hr
$\{B,T,C\}$	2.6 victims/hr

Table 4.8 summarizes the results and reports on the expected throughput (TH), queue length (L) and waiting time (W) at each RC. Note that, all queues are utilized around 80% and the time spent in each queue is related to its utilization and service time, causing a higher waiting time at tarpaulin distributing RCs.

Table 4.8: Results: Base Case

RC	Item	Utilization	Effective Arrivals (victims/hr)	TH	L	W (hrs)
1	B	0.85	50.44	50.44	5.57	0.11
2	T1	0.88	16.73	16.73	7.36	0.44
3	T2	0.88	16.73	16.73	7.36	0.44
4	T3	0.88	16.73	16.73	7.36	0.44
5	C	0.75	48.06	48.06	2.96	0.06

4.6.3 Demand Increase for Blankets

Having established the base case and its performance, we next model the scenario where demand for blankets increase. To model this demand increase, we increase

the $p^+(i,B), i = T1, T2, T3, C$ probabilities (see Table 4.9), which represent the probability of victims receiving service at other RCs (T1, T2, T3, C), moving to the blanket RC (B) upon service completion. Note that the change in $p^+(i,j)$ probabilities impact arrivals at every node in the network. Therefore any change in the effective arrival rate to queue i, will change the effective arrival rate to all queues connected to queue i in the network. In this particular case, an increase in blanket demand will increase utilization, throughput and waiting time for all other queues where $p^+(B,j), j = T1, T2, T3, C$ is nonzero. Since we are modeling a demand increase for blankets alone, this should not impact the performance of the other RCs. Hence, to keep the performance of all other queues the same, we adjust the $p^+(B,j)$ probabilities using Equation 4.19. Note that the derivation of this equation uses the definition of the effective victim arrival equation (Equation 4.3) and adjusts the $p^+(B,j)$ probabilities based on the change in the effective victim arrival rate to the blanket queue. The adjusted $p^+(B,j)$ probabilities are shown in Table 4.9.

$$(p^+(B,i))^{new} = \frac{(\lambda_B^+)^{old}(p^+(B,i))^{old}(\mu_B + \lambda_B^- \sum_{s=1}^{\infty} \pi_{B,s} s (q_B^{s-1})^{new})}{\mu_B + (\sum_{s=1}^{\infty} \pi_{B,s} s (q_B^{s-1})^{old})}, \forall i \qquad (4.19)$$

Table 4.9: Victim Routing Probabilities Between RCs: Blanket Demand Increase

	B	T1	T2	T3	C
B	0	0.0597	0.0597	0.0597	0.0905
T1	**0.25**	0	0	0	0.1
T2	**0.25**	0	0	0	0.1
T3	**0.25**	0	0	0	0.1
C	**0.1**	0.0166	0.0166	0.0166	0

Table 4.10 shows the results of the performance measures following the increase in $p^+(i, B)$ probabilities for all RCs. We observe that, increasing $p^+(i, B)$ probabilities increase the throughput from the blanket RC, the utilization of the blanket RC and the waiting time at the blanket RC. Due to the adjusted $p^+(B, i)$ probabilities, other queues in the network remain unaffected.

Table 4.10: Demand Increase for Blankets: Results

RC	Item	Utilization	Effective Arrivals	TH	L	W
1	B	0.93	55.33	55.33	13.27	0.24
2	T1	0.88	16.73	16.73	7.36	0.44
3	T2	0.88	16.73	16.73	7.36	0.44
4	T3	0.88	16.73	16.73	7.36	0.44
5	C	0.75	48.06	48.06	2.96	0.06

4.6.4 Changing Victim Needs

Next we model dynamically changing victim needs, using the signal entities of the G-network. Signals arriving to RC i, coupled with their frequency and batch size parameters, represent how often and how many victims experience a change in the

items they need, and the probability $q(i,j)$ represents the probability of these victims moving to RC j as a result. In this experiment set, we specifically model the scenario where, some victims decide to collect cash vouchers instead of tarpaulins. These victims who are in line at the tarpaulin RC might move to the cash RC because they can utilize the cash to purchase tarpaulins and they expect a shorter waiting time at the cash distributing RC. We model this behavior by introducing signal arrivals to the queues $T1, T2, T3$. We assume that the signal arrival rate and the batch size for each tarpaulin queue is equal. We vary the signal arrival rates and the batch size the signals affect, to investigate the impact of both the signal arrivals and batch size on the performance measures. Note that the external victim arrival rates and the service rates remain the same as given in the base case. Table 4.11 summarizes the signal and batch size parameters.

Table 4.11: Changing Victim Needs: Parameters

Experiment	$\lambda^-_{T1} = \lambda^-_{T2} = \lambda^-_{T3}$	$\pi_{T1,s} = \pi_{T2,s} = \pi_{T3,s}$	$q(i,j)$
1	1/hr	1	$q(i,C) = 1, i = T1, T2, T3$
2	1/hr	2	$q(i,C) = 1, i = T1, T2, T3$
3	1/hr	3	$q(i,C) = 1, i = T1, T2, T3$
4	1/hr	4	$q(i,C) = 1, i = T1, T2, T3$
5	2/hr	1	$q(i,C) = 1, i = T1, T2, T3$
6	3/hr	1	$q(i,C) = 1, i = T1, T2, T3$
7	4/hr	1	$q(i,C) = 1, i = T1, T2, T3$

Table 4.22 in Appendix 4 provides the performance measures for all RCs and compares the results of all 7 experiments. We observe that, both increasing signal frequency and increasing batch size, increases the number of victims that move from the tarpaulin RCs to the cash RC. As more victims move to collect cash vouchers

instead of tarpaulins, the utilization and throughput of relief centers T1, T2 and T3 decrease. However, the effective arrival rate and the average queue length of relief centers T1, T2, T3 remain the same. This is because the victims that join this queue that later change their minds (or needs), still add to the congestion experienced in these queues. Hence, the movement in the network triggered by victims changing their minds, not only fails to relieve the waiting time experienced at the tarpaulin RCs, but also significantly increases the waiting time at the cash RC. The impact of such victim behavior can be more pronounced for scenarios with higher demand or stricter resources. From these results, we conclude that victim mobility due to changing victim needs can significantly impact the relief distribution efficiency for the whole network. Hence, it is critical for relief agencies to inform the crowd in need, track the distribution and control victim movement to the extent possible.

Note that, the change in utilization of the cash distributing RC will impact effective arrival rate to the tarpaulin RCs T1, T2 and T3. Hence, we adjusted $p^+(C,i),\ i = T1, T2, T3$ probabilities using Equation 4.19 for queue (C) such that the effective arrival rate to tarpaulin queues will remain the same. In addition, the change in utilization for both the tarpaulin and cash distributing RCs will impact the effective arrival rate to the blanket distributing RC. Therefore, we also adjusted $p^+(C,B)$ and $p^+(i,B),\ i = T1, T2, T3$ probabilities such that the performance of the blanket RC remains unaffected.

4.6.5 Victim Jockeying

In Section 4.6.2 we assumed an equal distribution of victims between tarpaulin RCs, $T1, T2$ and $T3$, without any victim switching between the tarpaulin RCs. However,

due to varying victim perception, victims often jockey between queues that distribute the same relief item. In this numerical experiment, we model a scenario in which victims at RCs $T1$ and $T3$, jockey and move to RC $T2$. We model this scenario, using signal entities arriving to RCs $T1$ and $T3$ to move victims of batch size s, to RC $T2$. We vary the signal arrival rates and the batch size the signals affect, to investigate the impact of both the signal arrivals and batch size on the performance measures. Note that the external victim arrival rates and the service rates remain the same as the base case. Table 4.12 summarizes the signal and batch size parameters.

Table 4.12: Victim Jockeying: Parameters

Experiment	$\lambda^{-}_{T1} = \lambda^{-}_{T3}$	$\pi_{T1,s} = \pi_{T3,s}$	$q(i,j)$
1	1/hr	1	$q(i,T2) = 1, i = T1, T3$
2	1/hr	2	$q(i,T2) = 1, i = T1, T3$
3	1/hr	3	$q(i,T2) = 1, i = T1, T3$
4	2/hr	1	$q(i,T2) = 1, i = T1, T3$
5	2/hr	2	$q(i,T2) = 1, i = T1, T3$
6	2/hr	3	$q(i,T2) = 1, i = T1, T3$
7	3/hr	1	$q(i,T2) = 1, i = T1, T3$
8	3/hr	2	$q(i,T2) = 1, i = T1, T3$
9	3/hr	3	$q(i,T2) = 1, i = T1, T3$

Table 4.23 in Appendix 3 shows the detailed performance measure results for all 9 experiments. First we observe that, both increasing signal frequency and increasing batch size, increase the number of victims that jockey from RCs $T1$ and $T3$ to RC $T2$. In this set of experiments, we again observe that increasing signal frequency is more effective in transferring victims due to the associated probability of having batch size number of victims in queue. Next, we observe that as more victims jockey

from RCs $T1, T3$ to RC $T2$, the utilization and throughput of relief centers $T1$, $T3$ are decreasing while the average queue length stays the same. This is because the victims that jockey still add to the congestion experienced in RCs $T1$ and $T3$. As more victims move to RC $T2$, the utilization, waiting time and queue length at $T2$ all increase. This increase is observed to be steep and starting from experiment 2, RC $T2$ is over utilized. The observed increase in utilization for RC $T2$ is due to: (i) the tarpaulin RC $T2$ has a high utilization prior to jockeying and (ii) the service time to distribute tarpaulins is high. Note that, the over utilized RCs are marked by an asterisk in Table 4.23 and the performance measures for these cases can not be calculated. We conclude that, victim jockeying can significantly decrease system performance and create higher waiting times for victims.

4.6.6 Changing Item Assignment to Relief Centers

In this experiment set, we investigate the effects of changing item assignment at the RCs on performance measures. The assignment decisions are based on various factors such as, types of items being distributed by each agency, coordination level between agencies and available staff for distribution. We model the scenario where all items are available at all RCs. The overall arrival rate of 112 victims per hour and an equal distribution of victims between all RCs applies to this scenario as well. Figure 4.5 shows the relief distribution setting.

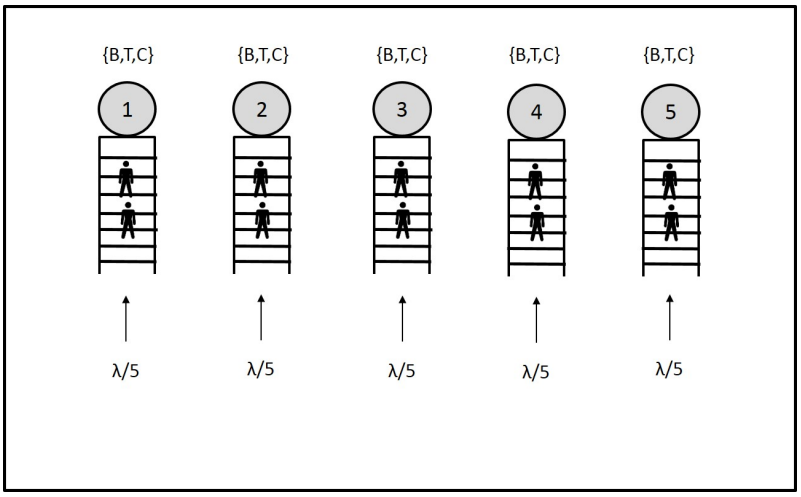

Figure 4.5: Item Assignment Change

Depending on the subset of items each victim collects at each RC, the average service time at the RCs need to be updated. To estimate the new service times, we assume that each victim category given in Table 4.7 is equally represented at each RC and we take the weighted average and estimate the service rate as 27.6 victims per hour at each RC. Note that, since all items are available at all RCs, we assume that $p^+(i,j) = 0, \forall(i,j)$.

Table 4.13 shows the resulting performance measures for all RCs. The results show that making all items available in all RCs improve the waiting times at the tarpaulin RCs and increase the waiting times for blanket and cash RCs, with the decrease at tarpaulin RCs being more significant. Overall, we observe that the performance of the network improves. However, it is important to note that this result is dependent on the assumption that all victim categories are equally distributed between all RCs. If more victims that need all items (or higher service time items like tarpaulins) converge to a single RC, the waiting times at that RC can increase significantly. Moreover, for different demand parameters, such as all victims need-

ing all items, this design can again result in high waiting times. Therefore, it is important for practitioners to estimate performance measures using models such as the one we propose before deciding on the relief distribution setup. Lastly it is important to note that, even for cases where having all items available at all RCs can provide a better performance, such a configuration might not be feasible. This is due to different relief agencies sourcing different supplies and limited coordination capabilities in the field.

Table 4.13: Results: Changing Item Assignment

RC	Utilization	Effective Arrivals	TH	L	W
1	0.81	22.40	22.40	4.31	0.19
2	0.81	22.40	22.40	4.31	0.19
3	0.81	22.40	22.40	4.31	0.19
4	0.81	22.40	22.40	4.31	0.19
5	0.81	22.40	22.40	4.31	0.19

4.7 Conclusions

In this Chapter we model the relief distribution in a disaster affected area as a G-network where we model victim mobility due to changes in need, changes in demand and changes in perception. To solve the G-network model, we investigate the existence of product form solution for G-networks with batch transfer and departure of victims. The product form results are of interest in this setting due to the computational advantages they can provide, especially for large networks. We show a new product form result for G-networks with signals and full batch transfers for balanced networks. For the conditions under which the product form does not analytically hold, we investigate the accuracy of the product form approximation to

estimate the performance measures of the G-network. We numerically conclude that the product form approximation can provide good performance estimations where solution speed is important.

We apply the analytic results to analyze relief distribution performance using a case study. We investigate the impacts of victim mobility under the three scenarios; changes in need, changes in demand and changes in perception.

From our numerical studies we learn the following insights: (i) Victim mobility has a significant impact on performance measure estimates, (ii) The extent of the impact of victim mobility is difficult to estimate without models, due to the probabilistic nature of victim demand and victim mobility, (iii) Decisions such as number of relief centers to open and the assignment of items to relief centers need to depend on demand estimates as well as expected victim movements. These insights all highlight the need for integrating victim movement and victim behavior aspects to models for relief distribution.

We believe that this research will motivate future work on extending operations research tools and models to include victim mobility aspects for disaster response support.

4.8 Appendix 1

4.8.1 Product Form Proof For Full Batch Transfer For State Group S_2

Recall that state group S_2 is defined as $S_2 = \{k_1(t), ..., k_i(t), ..., k_j(t), ..., k_N(t)\}$: $k_i(t) < s$, for at least one i and $k_i(t) \neq 0 \ \forall i \quad i, j \in \{1, 2, .., N\}$. The proof below considers the case where $k_i(t) < s, \quad \forall i$. The proof where $k_i(t) < s$ for some but not all i is a combination of the proof for this special case and the proof for states in S_1.

Proof. For the states in S_1 the balance equations given by Equation 5 take the form given in 4.20 following the substitution of product form $p(k) = \prod_{i=1}^{N} p(k_i)$ where $p(k_i) = (1 - q_i)(q_i^{k_i})$.

$$\begin{aligned} \sum_i \left[\lambda_{0i}^+ + \mu_i\right] &= \sum_i \mu_i q_i d(i) \\ &+ \sum_i \frac{\lambda_{0i}^+}{q_i} \\ &+ \sum_i \lambda_{0i}^- D(i) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\ &+ \sum_i \sum_j \frac{\mu_i q_i p^+(i,j)}{q_j} \end{aligned} \tag{4.20}$$

We simplify the balance equations by using Equations 4.21 and obtain the balance equations given by 4.22.

$$\sum_i \frac{\lambda_{0i}^+}{q_i} + \sum_i \sum_j \frac{\mu_j q_j p^+(j,i)}{q_i} = \sum_i \frac{\lambda_i^+}{q_i} - \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{s q_i^s}{q_j} \tag{4.21}$$

$$\begin{aligned}\sum_i \left[\lambda_{0i}^+ + \mu_i\right] &= \sum_i \mu_i q_i d(i) \\ &+ \sum_i \lambda_i^- D(i) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\ &+ \sum_i \frac{\lambda_i^+}{q_i} \\ &- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \frac{\pi_{is} s q_i^s}{q_j} \end{aligned} \tag{4.22}$$

Next we insert the definition for λ_i^+ given by 4.3 on both the LHS and the RHS while inserting the definitions of $D(i)$ and $d(i)$ to the RHS, which yields Equation 4.23.

$$\begin{aligned}\sum_i \lambda_i^+ - \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s + \sum_i \mu_i &= \sum_i \mu_i q_i \\ &+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} q_i^s \\ &- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\ &- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{s q_i^s}{q_j} \\ &+ \sum_i \frac{\lambda_i^+}{q_i} \end{aligned} \tag{4.23}$$

Next we use the definition of q_i in Equation 4.2 to rewrite λ_i^+ on both the LHS and the RHS. Then the balance equations can be written as Equation 4.24.

$$
\begin{aligned}
0 &= \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^{s-1} \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \frac{s q_i^s}{q_j} \\
&+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&+ \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s \\
&- \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^s
\end{aligned} \tag{4.24}
$$

Rearranging the terms we get Equation 4.25 where it can be seen that the equation holds for balance networks where $q_i = q_j, \forall (i,j) \in N$ given that $\sum_j q(i,j) = 1 \, \forall i \in N$.

$$
\begin{aligned}
0 &= \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \left[-\frac{s q_i^s}{q_j} - q_i^s + s q_i^s \right] \\
&+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} \left[s q_i^{s-1} + q_i^s - s q_i^s \right]
\end{aligned} \tag{4.25}
$$

This proves the product form result for states in S_2.

□

4.8.2 Product Form Proof for Full Batch Departure For State Group S_3

Recall that state group S_3 is defined as $S_3 = \{k_1(t), ..., k_i(t), ..., k_j(t), ..., k_N(t)\}$: $k_i(t) = 0$, for at least one i, $i \in \{1, 2, .., N\}$. The proof below considers the case where $k_i(t) = 0, \quad \forall i$. The proof where $k_i(t) = 0$ for some but not all i is a combination of the proof for this special case and the proof for states in S_1 and S_2.

Proof. For the states in S_3 the balance equations given by Equation 5 take the form given in 4.26 following the substitution of product form $p(k) = \prod_{i=1}^{N} p(k_i)$ where $p(k_i) = (1 - q_i)(q_i^{k_i})$.

$$\begin{aligned} \sum_i \lambda_{0i}^{+} &= \sum_i \mu_i q_i d(i) \\ &+ \sum_i \lambda_{0i}^{-} D(i) \sum_{s=1}^{\infty} \pi_{is} q_i^s \end{aligned} \tag{4.26}$$

Next we insert the definition for λ_i^+ given by 4.3 on the LHS terms and insert the definitions of $D(i)$ and $d(i)$ to the RHS terms yielding Equation 4.27.

$$\begin{aligned} \sum_i \lambda_i^{+} - \sum_i \sum_j \lambda_i^{-} q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s &= \sum_i \mu_i q_i \\ &+ \sum_i \lambda_i^{-} \sum_{s=1}^{\infty} \pi_{is} q_i^s \\ &- \sum_i \sum_j \lambda_i^{-} q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s \end{aligned} \tag{4.27}$$

Next we use the definition of q_i in Equation 4.2 to rewrite λ_i^+ on the LHS and the RHS. Then the balance equations can be written as Equation 4.28.

$$
\begin{aligned}
0 &= \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&- \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} q_i^s \\
&+ \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} s q_i^s \\
&- \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} s q_i^s
\end{aligned} \tag{4.28}
$$

Rearranging the terms we get Equation 4.29 where it can be seen that the equation holds for balance networks where $q_i = q_j, \forall (i,j) \in N$ given that $\sum_j q(i,j) = 1 \, \forall i \in N$.

$$
\begin{aligned}
0 &= \sum_i \sum_j \lambda_i^- q(i,j) \sum_{s=1}^{\infty} \pi_{is} \left[s q_i^s - q_i^s \right] \\
&+ \sum_i \lambda_i^- \sum_{s=1}^{\infty} \pi_{is} \left[q_i^s - s q_i^s \right]
\end{aligned} \tag{4.29}
$$

This proves the product form result for states in S_3.

□

Table 4.14: Experiment Results for Full Batch Transfer: 53 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	29.98	23.02	0.50	0.12	29.98	23.02	0.50	0.11	0%	0%	0%	4%
A,2		1	4	28.93	24.07	0.31	0.15	28.90	24.10	0.32	0.13	0%	0%	5%	13%
A,3		1	6	28.57	24.43	0.25	0.17	28.44	24.56	0.28	0.13	0%	1%	11%	20%
A,4		1	8	28.56	24.44	0.24	0.18	28.30	24.70	0.27	0.14	1%	1%	15%	24%
B,5		2	1	29.93	23.07	0.48	0.11	29.93	23.07	0.48	0.11	0%	0%	0%	0%
B,6		4	1	28.27	24.73	0.27	0.14	28.27	24.73	0.27	0.14	0%	0%	0%	0%
B,7		6	1	26.78	26.22	0.19	0.17	26.78	26.22	0.19	0.17	0%	0%	0%	0%
B,8		8	1	25.44	27.56	0.15	0.23	25.44	27.56	0.15	0.23	0%	0%	0%	0%
C,9		2	2	28.44	24.56	0.28	0.14	28.44	24.56	0.28	0.13	0%	0%	0%	6%
C,10		2	3	27.56	25.44	0.22	0.17	27.53	25.47	0.22	0.15	0%	0%	4%	12%
C,11		3	2	27.11	25.89	0.20	0.18	27.11	25.89	0.20	0.16	0%	0%	0%	7%
A,12	$\beta = 0.56$	1	2	28.04	24.96	0.25	0.15	28.04	24.96	0.25	0.14	0%	0%	0%	3%
A,13		1	4	27.28	25.72	0.20	0.18	27.25	25.75	0.21	0.16	0%	0%	5%	10%
A,14		1	6	27.15	25.85	0.18	0.19	27.02	25.98	0.20	0.17	0%	0%	9%	14%
A,15		1	8	27.25	25.75	0.18	0.20	27.02	25.98	0.20	0.17	1%	1%	11%	15%
B,16		2	1	27.93	25.07	0.25	0.14	27.93	25.07	0.25	0.14	0%	0%	0%	0%
B,17		4	1	26.38	26.62	0.18	0.19	26.38	26.62	0.18	0.19	0%	0%	0%	0%
B,18		6	1	24.99	28.01	0.14	0.25	24.99	28.01	0.14	0.25	0%	0%	0%	0%
B,19		8	1	23.74	29.26	0.12	0.36	23.74	29.26	0.12	0.36	0%	0%	0%	0%
C,20		2	2	26.63	26.37	0.19	0.19	26.63	26.37	0.19	0.18	0%	0%	0%	5%
C,21		2	3	25.96	27.04	0.16	0.22	25.93	27.07	0.16	0.20	0%	0%	3%	9%
C,22		3	2	25.41	27.59	0.15	0.24	25.41	27.59	0.15	0.23	0%	0%	0%	6%

Table 4.15: Experiment Results for Full Batch Transfer: 50 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	28.30	21.70	0.30	0.10	28.30	21.70	0.30	0.10	0.00%	0.00%	0.02%	3.51%
A,2		1	4	27.46	22.54	0.23	0.13	27.43	22.57	0.24	0.11	0.12%	0.14%	4.95%	12.22%
A,3		1	6	27.26	22.74	0.20	0.14	27.14	22.86	0.22	0.11	0.47%	0.56%	9.82%	17.86%
A,4		1	8	27.34	22.66	0.20	0.15	27.10	22.90	0.22	0.11	0.89%	1.07%	12.23%	20.81%
B,5		2	1	28.21	21.79	0.30	0.10	28.21	21.79	0.30	0.10	0.00%	0.00%	0.01%	0.00%
B,6		4	1	26.63	23.37	0.20	0.12	26.63	23.37	0.20	0.12	0.00%	0.00%	0.00%	0.00%
B,7		6	1	25.21	24.79	0.16	0.15	25.21	24.79	0.16	0.15	0.00%	0.00%	0.00%	0.00%
B,8		8	1	23.94	26.06	0.13	0.18	23.94	26.06	0.13	0.18	0.00%	0.00%	0.00%	0.00%
C,9		2	2	26.86	23.14	0.21	0.13	26.86	23.14	0.21	0.12	0.00%	0.00%	0.00%	5.60%
C,10		2	3	26.12	23.88	0.18	0.15	26.10	23.90	0.18	0.13	0.11%	0.12%	3.33%	11.17%
C,11		3	2	25.60	24.40	0.17	0.15	25.60	24.40	0.17	0.14	0.00%	0.00%	0.00%	6.84%
A,12	$\beta = 0.56$	1	2	26.46	23.54	0.19	0.13	26.46	23.54	0.19	0.12	0.00%	0.00%	0.00%	2.81%
A,13		1	4	25.87	24.13	0.17	0.15	25.84	24.16	0.17	0.13	0.12%	0.13%	4.22%	8.99%
A,14		1	6	25.84	24.16	0.16	0.16	25.73	24.27	0.17	0.14	0.43%	0.46%	7.79%	12.30%
A,15		1	8	26.08	23.92	0.15	0.15	25.90	24.10	0.16	0.13	1%	1%	9%	13%
B,16		2	1	26.33	23.67	0.19	0.13	26.33	23.67	0.19	0.13	0.00%	0.00%	0.00%	0.00%
B,17		4	1	24.85	25.15	0.15	0.15	24.85	25.15	0.15	0.15	0.00%	0.00%	0.00%	0.00%
B,18		6	1	23.53	26.47	0.12	0.19	23.53	26.47	0.12	0.19	0.00%	0.00%	0.00%	0.00%
B,19		8	1	22.40	27.60	0.10	0.23	22.40	27.60	0.10	0.23	0.00%	0.00%	0.00%	0.00%
C,20		2	2	25.14	24.86	0.15	0.16	25.14	24.86	0.15	0.15	0.00%	0.00%	0.00%	4.53%
C,21		2	3	24.59	25.41	0.14	0.18	24.57	25.43	0.14	0.16	0.11%	0.11%	3.00%	8.63%
C,22		3	2	23.99	26.01	0.13	0.19	23.99	26.01	0.13	0.18	0.00%	0.00%	0.00%	5.56%

Table 4.16: Experiment Results for Full Batch Transfer: 47 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	26.67	20.33	0.19	0.09	26.67	20.33	0.19	0.09	0%	0%	0%	3%
A,2		1	4	26.10	20.90	0.16	0.10	26.06	20.94	0.17	0.09	0%	0%	4%	10%
A,3		1	6	26.07	20.93	0.15	0.11	25.96	21.04	0.17	0.09	0%	1%	8%	15%
A,4		1	8	26.22	20.78	0.15	0.11	26.04	20.96	0.17	0.09	1%	1%	9%	16%
B,5		2	1	26.54	20.46	0.18	0.09	26.54	20.46	0.18	0.09	0%	0%	0%	0%
B,6		4	1	25.07	21.93	0.14	0.10	25.07	21.93	0.14	0.10	0%	0%	0%	0%
B,7		6	1	23.75	23.25	0.12	0.11	23.75	23.25	0.12	0.11	0%	0%	0%	0%
B,8		8	1	22.56	24.44	0.11	0.13	22.56	24.44	0.11	0.13	0%	0%	0%	0%
C,9		2	2	25.36	21.64	0.15	0.10	25.36	21.64	0.15	0.10	0%	0%	0%	5%
C,10		2	3	24.82	22.18	0.13	0.11	24.79	22.21	0.14	0.10	0%	0%	3%	10%
C,11		3	2	24.21	22.79	0.13	0.12	24.21	22.79	0.13	0.11	0%	0%	0%	6%
A,12	$\beta = 0.56$	1	2	24.93	22.07	0.14	0.10	24.93	22.07	0.14	0.10	0%	0%	0%	3%
A,13		1	4	24.55	22.45	0.13	0.11	24.52	22.48	0.13	0.11	0%	0%	3%	8%
A,14		1	6	24.62	22.38	0.13	0.12	24.54	22.46	0.13	0.10	0%	0%	6%	10%
A,15		1	8	24.80	22.20	0.13	0.12	24.67	22.33	0.14	0.10	1%	1%	7%	10%
B,16		2	1	24.77	22.23	0.14	0.10	24.77	22.23	0.14	0.10	0%	0%	0%	0%
B,17		4	1	23.40	23.60	0.12	0.12	23.40	23.60	0.12	0.12	0%	0%	0%	0%
B,18		6	1	22.16	24.84	0.10	0.14	22.16	24.84	0.10	0.14	0%	0%	0%	0%
B,19		8	1	21.06	25.94	0.09	0.17	21.06	25.94	0.09	0.17	0%	0%	0%	0%
C,20		2	2	23.74	23.26	0.12	0.12	23.74	23.26	0.12	0.11	0%	0%	0%	4%
C,21		2	3	23.34	23.66	0.11	0.13	23.32	23.68	0.12	0.12	0%	0%	3%	8%
C,22		3	2	22.69	24.31	0.11	0.14	22.69	24.31	0.11	0.13	0%	0%	0%	5%

Table 4.17: Experiment Results for Full Batch Transfer: 44 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	25.01	18.99	0.14	0.08	25.01	18.99	0.14	0.08	0%	0%	0%	3%
A,2		1	4	24.62	19.38	0.13	0.09	24.59	19.41	0.13	0.08	0%	0%	4%	9%
A,3		1	6	24.69	19.31	0.13	0.09	24.60	19.40	0.14	0.08	0%	0%	6%	12%
A,4		1	8	24.87	19.13	0.13	0.09	24.73	19.27	0.14	0.08	1%	1%	7%	12%
B,5		2	1	24.85	19.15	0.14	0.08	24.85	19.15	0.14	0.08	0%	0%	0%	0%
B,6		4	1	23.47	20.53	0.12	0.09	23.47	20.53	0.12	0.09	0%	0%	0%	0%
B,7		6	1	22.23	21.77	0.10	0.10	22.23	21.77	0.10	0.10	0%	0%	0%	0%
B,8		8	1	21.12	22.88	0.09	0.11	21.12	22.88	0.09	0.11	0%	0%	0%	0%
C,9		2	2	23.81	20.19	0.12	0.09	23.81	20.19	0.12	0.08	0%	0%	0%	5%
C,10		2	3	23.41	20.59	0.11	0.10	23.38	20.62	0.12	0.09	0%	0%	3%	9%
C,11		3	2	22.75	21.25	0.11	0.10	22.75	21.25	0.11	0.09	0%	0%	0%	6%
A,12	$\beta = 0.56$	1	2	23.38	20.62	0.12	0.09	23.38	20.62	0.12	0.09	0%	0%	0%	2%
A,13		1	4	23.13	20.87	0.11	0.10	23.10	20.90	0.11	0.09	0%	0%	3%	7%
A,14		1	6	23.25	20.75	0.11	0.10	23.19	20.81	0.11	0.09	0%	0%	5%	8%
A,15		1	8	23.43	20.57	0.11	0.10	23.35	20.65	0.12	0.09	0%	0%	5%	8%
B,16		2	1	23.19	20.81	0.11	0.09	23.19	20.81	0.11	0.09	0%	0%	0%	0%
B,17		4	1	21.90	22.10	0.10	0.10	21.90	22.10	0.10	0.10	0%	0%	0%	0%
B,18		6	1	20.75	23.25	0.09	0.11	20.75	23.25	0.09	0.11	0%	0%	0%	0%
B,19		8	1	19.71	24.29	0.08	0.13	19.71	24.29	0.08	0.13	0%	0%	0%	0%
C,20		2	2	22.28	21.72	0.10	0.10	22.28	21.72	0.10	0.10	0%	0%	0%	4%
C,21		2	3	22.00	22.00	0.10	0.11	21.97	22.03	0.10	0.10	0%	0%	2%	7%
C,22		3	2	21.31	22.69	0.09	0.11	21.31	22.69	0.09	0.11	0%	0%	0%	5%

Table 4.18: Experiment Results for Partial Batch Transfer: 53 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	29.98	23.01	0.49	0.12	29.98	23.02	0.50	0.11	0%	0%	1%	4%
A,2		1	4	28.72	24.28	0.31	0.15	28.72	24.28	0.31	0.13	0%	0%	0%	14%
A,3		1	6	27.96	25.04	0.25	0.18	27.96	25.04	0.25	0.14	0%	0%	0%	22%
A,4		1	8	27.50	25.50	0.22	0.21	27.50	25.50	0.22	0.15	0%	0%	0%	28%
B,5		2	1	29.93	23.07	0.48	0.11	29.93	23.07	0.48	0.11	0%	0%	1%	0%
B,6		4	1	28.27	24.73	0.27	0.14	28.27	24.73	0.27	0.14	0%	0%	0%	0%
B,7		6	1	26.78	26.22	0.19	0.17	26.78	26.22	0.19	0.17	0%	0%	0%	0%
B,8		8	1	25.44	27.56	0.15	0.23	25.44	27.56	0.15	0.23	0%	0%	0%	0%
C,9		2	2	28.44	24.56	0.28	0.14	28.44	24.56	0.28	0.13	0%	0%	0%	6%
C,10		2	3	27.37	25.63	0.22	0.18	27.37	25.63	0.22	0.16	0%	0%	0%	12%
C,11		3	2	27.11	25.89	0.20	0.18	27.11	25.89	0.20	0.16	0%	0%	0%	7%
A,12	$\beta = 0.56$	1	2	28.04	24.96	0.25	0.15	28.04	24.96	0.25	0.14	0%	0%	0%	3%
A,13		1	4	27.01	25.99	0.20	0.19	27.01	25.99	0.20	0.17	0%	0%	0%	11%
A,14		1	6	26.44	26.56	0.18	0.22	26.44	26.56	0.18	0.18	0%	0%	0%	17%
A,15		1	8	26.12	26.88	0.17	0.25	26.12	26.88	0.17	0.20	0%	0%	0%	21%
B,16		2	1	27.93	25.07	0.25	0.14	27.93	25.07	0.25	0.14	0%	0%	0%	0%
B,17		4	1	26.38	26.62	0.18	0.19	26.38	26.62	0.18	0.19	0%	0%	0%	0%
B,18		6	1	24.99	28.01	0.14	0.25	24.99	28.01	0.14	0.25	0%	0%	0%	0%
B,19		8	1	23.74	29.26	0.12	0.36	23.74	29.26	0.12	0.36	0%	0%	0%	0%
C,20		2	2	26.63	26.37	0.19	0.19	26.63	26.37	0.19	0.18	0%	0%	0%	5%
C,21		2	3	25.74	27.26	0.16	0.23	25.74	27.26	0.16	0.21	0%	0%	0%	10%
C,22		3	2	25.41	27.59	0.15	0.24	25.41	27.59	0.15	0.23	0%	0%	0%	6%

Table 4.19: Experiment Results for Partial Batch Transfer: 50 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	28.30	21.70	0.30	0.10	28.30	21.70	0.30	0.10	0.00%	0.00%	0.02%	3.51%
A,2		1	4	27.21	22.79	0.23	0.13	27.21	22.79	0.23	0.11	0.00%	0.00%	0.00%	13.11%
A,3		1	6	26.58	23.42	0.20	0.15	26.58	23.42	0.20	0.12	0.00%	0.00%	0.00%	20.63%
A,4		1	8	26.22	23.78	0.19	0.17	26.22	23.78	0.19	0.13	0.00%	0.00%	0.00%	25.58%
B,5		2	1	28.21	21.79	0.30	0.10	28.21	21.79	0.30	0.10	0.00%	0.00%	0.01%	0.00%
B,6		4	1	26.63	23.37	0.20	0.12	26.63	23.37	0.20	0.12	0.00%	0.00%	0.00%	0.00%
B,7		6	1	25.21	24.79	0.16	0.15	25.21	24.79	0.16	0.15	0.00%	0.00%	0.00%	0.00%
B,8		8	1	23.94	26.06	0.13	0.18	23.94	26.06	0.13	0.18	0.00%	0.00%	0.00%	0.00%
C,9		2	2	26.86	23.14	0.21	0.13	26.86	23.14	0.21	0.12	0.00%	0.00%	0.00%	5.60%
C,10		2	3	25.91	24.09	0.18	0.15	25.91	24.09	0.18	0.13	0.00%	0.00%	0.00%	11.71%
C,11		3	2	25.60	24.40	0.17	0.15	25.60	24.40	0.17	0.14	0.00%	0.00%	0.00%	6.84%
A,12	$\beta = 0.56$	1	2	26.46	23.54	0.19	0.13	26.46	23.54	0.19	0.12	0.00%	0.00%	0.00%	2.81%
A,13		1	4	25.58	24.42	0.17	0.16	25.58	24.42	0.17	0.14	0.00%	0.00%	0.00%	10.15%
A,14		1	6	25.11	24.89	0.15	0.18	25.11	24.89	0.15	0.15	0.00%	0.00%	0.00%	15.76%
A,15		1	8	24.95	25.05	0.14	0.18	24.95	25.05	0.14	0.14	0%	0%	0%	19%
B,16		2	1	26.33	23.67	0.19	0.13	26.33	23.67	0.19	0.13	0.00%	0.00%	0.00%	0.00%
B,17		4	1	24.85	25.15	0.15	0.15	24.85	25.15	0.15	0.15	0.00%	0.00%	0.00%	0.00%
B,18		6	1	23.53	26.47	0.12	0.19	23.53	26.47	0.12	0.19	0.00%	0.00%	0.00%	0.00%
B,19		8	1	22.40	27.60	0.10	0.23	22.40	27.60	0.10	0.23	0.00%	0.00%	0.00%	0.00%
C,20		2	2	25.14	24.86	0.15	0.16	25.14	24.86	0.15	0.15	0.00%	0.00%	0.00%	4.53%
C,21		2	3	24.35	25.65	0.14	0.19	24.35	25.65	0.14	0.17	0.00%	0.00%	0.00%	9.35%
C,22		3	2	23.99	26.01	0.13	0.19	23.99	26.01	0.13	0.18	0.00%	0.00%	0.00%	5.56%

Table 4.20: Experiment Results for Partial Batch Transfer: 47 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	26.67	20.33	0.19	0.09	26.67	20.33	0.19	0.09	0%	0%	0%	3%
A,2		1	4	25.80	21.20	0.16	0.10	25.80	21.20	0.16	0.09	0%	0%	0%	12%
A,3		1	6	25.33	21.67	0.15	0.12	25.33	21.67	0.15	0.10	0%	0%	0%	18%
A,4		1	8	25.09	21.91	0.14	0.13	25.09	21.91	0.14	0.10	0%	0%	0%	22%
B,5		2	1	26.54	20.46	0.18	0.09	26.54	20.46	0.18	0.09	0%	0%	0%	0%
B,6		4	1	25.07	21.93	0.14	0.10	25.07	21.93	0.14	0.10	0%	0%	0%	0%
B,7		6	1	23.75	23.25	0.12	0.11	23.75	23.25	0.12	0.11	0%	0%	0%	0%
B,8		8	1	22.56	24.44	0.11	0.13	22.56	24.44	0.11	0.13	0%	0%	0%	0%
C,9		2	2	25.36	21.64	0.15	0.10	25.36	21.64	0.15	0.10	0%	0%	0%	5%
C,10		2	3	24.58	22.42	0.13	0.12	24.58	22.42	0.13	0.10	0%	0%	0%	11%
C,11		3	2	24.21	22.79	0.13	0.12	24.21	22.79	0.13	0.11	0%	0%	0%	6%
A,12	$\beta = 0.56$	1	2	24.93	22.07	0.14	0.10	24.93	22.07	0.14	0.10	0%	0%	0%	3%
A,13		1	4	24.23	22.77	0.13	0.12	24.23	22.77	0.13	0.11	0%	0%	0%	9%
A,14		1	6	23.89	23.11	0.12	0.13	23.89	23.11	0.12	0.11	0%	0%	0%	14%
A,15		1	8	23.72	23.28	0.12	0.14	23.72	23.28	0.12	0.11	0%	0%	0%	17%
A,16		2	1	24.77	22.23	0.14	0.10	24.77	22.23	0.14	0.10	0%	0%	0%	0%
A,17		4	1	23.40	23.60	0.12	0.12	23.40	23.60	0.12	0.12	0%	0%	0%	0%
A,18		6	1	22.16	24.84	0.10	0.14	22.16	24.84	0.10	0.14	0%	0%	0%	0%
A,19		8	1	21.06	25.94	0.09	0.17	21.06	25.94	0.09	0.17	0%	0%	0%	0%
A,20		2	2	23.74	23.26	0.12	0.12	23.74	23.26	0.12	0.11	0%	0%	0%	4%
A,21		2	3	23.09	23.91	0.11	0.14	23.09	23.91	0.11	0.12	0%	0%	0%	9%
A,22		3	2	22.69	24.31	0.11	0.14	22.69	24.31	0.11	0.13	0%	0%	0%	5%

Table 4.21: Experiment Results for Partial Batch Transfer: 44 victims/hour

Case No	β	Variables		Markov Chain Result				G-Network Result				Absolute Error Percentage (E_p)			
		λ_1^-	B_1	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$	TH_1	TH_2	$\bar{W}_1$	$\bar{W}_2$
A,1	$\beta = 0.6$	1	2	25.01	18.99	0.14	0.08	25.01	18.99	0.14	0.08	0%	0%	0%	3%
A,2		1	4	24.29	19.71	0.13	0.09	24.29	19.71	0.13	0.08	0%	0%	0%	10%
A,3		1	6	23.95	20.05	0.12	0.10	23.95	20.05	0.12	0.08	0%	0%	0%	16%
A,4		1	8	23.78	20.22	0.12	0.11	23.78	20.22	0.12	0.08	0%	0%	0%	19%
B,5		2	1	24.85	19.15	0.14	0.08	24.85	19.15	0.14	0.08	0%	0%	0%	0%
B,6		4	1	23.47	20.53	0.12	0.09	23.47	20.53	0.12	0.09	0%	0%	0%	0%
B,7		6	1	22.23	21.77	0.10	0.10	22.23	21.77	0.10	0.10	0%	0%	0%	0%
B,8		8	1	21.12	22.88	0.09	0.11	21.12	22.88	0.09	0.11	0%	0%	0%	0%
C,9		2	2	23.81	20.19	0.12	0.09	23.81	20.19	0.12	0.08	0%	0%	0%	5%
C,10		2	3	23.15	20.85	0.11	0.10	23.15	20.85	0.11	0.09	0%	0%	0%	10%
C,11		3	2	22.75	21.25	0.11	0.10	22.75	21.25	0.11	0.09	0%	0%	0%	6%
A,12	$\beta = 0.56$	1	2	23.38	20.62	0.12	0.09	23.38	20.62	0.12	0.09	0%	0%	0%	2%
A,13		1	4	22.80	21.20	0.11	0.10	22.80	21.20	0.11	0.09	0%	0%	0%	8%
A,14		1	6	22.55	21.45	0.11	0.11	22.55	21.45	0.11	0.09	0%	0%	0%	12%
A,15		1	8	22.43	21.57	0.10	0.11	22.43	21.57	0.10	0.10	0%	0%	0%	14%
B,16		2	1	23.19	20.81	0.11	0.09	23.19	20.81	0.11	0.09	0%	0%	0%	0%
B,17		4	1	21.90	22.10	0.10	0.10	21.90	22.10	0.10	0.10	0%	0%	0%	0%
B,18		6	1	20.75	23.25	0.09	0.11	20.75	23.25	0.09	0.11	0%	0%	0%	0%
B,19		8	1	19.71	24.29	0.08	0.13	19.71	24.29	0.08	0.13	0%	0%	0%	0%
C,20		2	2	22.28	21.72	0.10	0.10	22.28	21.72	0.10	0.10	0%	0%	0%	4%
C,21		2	3	21.73	22.27	0.10	0.11	21.73	22.27	0.10	0.10	0%	0%	0%	8%
C,22		3	2	21.31	22.69	0.09	0.11	21.31	22.69	0.09	0.11	0%	0%	0%	5%

Table 4.22: Change in Victim Needs (Tarpaulin to Cash): Results

Experiment	RC	Item	Utilization	Effective Arrivals	TH	L	W
1	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.832	16.73	15.89	7.05	0.44
	3	T2	0.832	16.73	15.89	7.05	0.44
	4	T3	0.832	16.73	15.89	7.05	0.44
	5	C	0.786	50.55	50.55	3.67	0.07
2	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.807	16.73	15.41	7.05	0.45
	3	T2	0.807	16.73	15.41	7.05	0.45
	4	T3	0.807	16.73	15.41	7.05	0.45
	5	C	0.808	51.97	51.97	4.21	0.08
3	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.796	16.73	15.20	7.05	0.46
	3	T2	0.796	16.73	15.20	7.05	0.46
	4	T3	0.796	16.73	15.20	7.05	0.46
	5	C	0.818	52.60	52.60	4.49	0.085
4	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.793	16.73	15.14	7.05	0.46
	3	T2	0.793	16.73	15.14	7.05	0.46
	4	T3	0.793	16.73	15.14	7.05	0.46
	5	C	0.821	52.80	52.80	4.59	0.087
5	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.792	16.73	15.12	7.05	0.46
	3	T2	0.792	16.73	15.12	7.05	0.46
	4	T3	0.792	16.73	15.12	7.05	0.46
	5	C	0.821	52.81	52.81	4.59	0.087
6	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.757	16.73	14.45	7.05	0.48
	3	T2	0.757	16.73	14.45	7.05	0.48
	4	T3	0.757	16.73	14.45	7.05	0.48
	5	C	0.853	54.87	54.87	5.81	0.10
7	1	B	0.850	50.44	50.44	5.57	0.11
	2	T1	0.724	16.73	13.82	7.05	0.51
	3	T2	0.724	16.73	13.82	7.05	0.51
	4	T3	0.724	16.73	13.82	7.05	0.51
	5	C	0.882	56.75	56.75	7.51	0.13

Table 4.23: Victim Jockeying: Results

Experiment	RC	Utilization	Effective Arrivals	TH	L	W
1	B	0.85	50.44	50.44	5.57	0.11
	T1	0.84	16.73	15.89	7.36	0.46
	T2	0.97	18.40	18.40	30.64	1.67
	T3	0.84	16.73	15.89	7.36	0.46
	C	0.75	48.06	48.06	2.96	0.06
2	B	0.85	50.44	50.44	5.57	0.11
	T1	0.81	16.73	15.41	7.36	0.48
	T2	1.02*	19.36*	N/A	N/A	N/A
	T3	0.81	16.73	15.41	7.36	0.48
	C	0.75	48.06	48.06	2.96	0.06
3	B	0.85	50.44	50.44	5.57	0.11
	T1	0.80	16.73	15.19	7.36	0.48
	T2	1.04*	19.79*	N/A	N/A	N/A
	T3	0.80	16.73	15.19	7.36	0.48
	C	0.75	48.06	48.06	2.96	0.06
4	B	0.85	50.44	50.44	5.57	0.11
	T1	0.80	16.73	15.13	7.36	0.49
	T2	1.05*	19.91*	N/A	N/A	N/A
	T3	0.80	16.73	15.13	7.36	0.49
	C	0.75	48.06	48.06	2.96	0.06
5	B	0.85	50.44	50.44	5.57	0.11
	T1	0.76	16.73	14.42	7.36	0.51
	T2	1.12*	21.34*	N/A	N/A	N/A
	T3	0.76	16.73	14.42	7.36	0.51
	C	0.75	48.06	48.06	2.96	0.06
6	B	0.85	50.44	50.44	5.57	0.11
	T1	0.75	16.73	14.21	7.36	0.52
	T2	1.14*	21.75*	N/A	N/A	N/A

	T3	0.75	16.73	14.21	7.36	0.52
	C	0.75	48.06	48.06	2.96	0.06
7	B	0.85	50.44	50.44	5.57	0.11
	T1	0.76	16.73	14.45	7.36	0.51
	T2	1.12*	21.29*	N/A	N/A	N/A
	T3	0.76	16.73	14.45	7.36	0.51
	C	0.75	48.06	48.06	2.96	0.06
8	B	0.85	50.44	50.44	5.57	0.11
	T1	0.72	16.73	13.64	7.36	0.54
	T2	1.21*	22.91*	N/A	N/A	N/A
	T3	0.72	16.73	13.64	7.36	0.54
	C	0.75	48.06	48.06	2.96	0.06
9	B	0.85	50.44	50.44	5.57	0.11
	T1	0.71	16.73	13.50	7.36	0.55
	T2	1.22*	23.18*	N/A	N/A	N/A
	T3	0.71	16.73	13.50	7.36	0.55
	C	0.75	48.06	48.06	2.96	0.06

Chapter 5

Resource Allocation Models for Material Convergence

5.1 The Material Convergence Problem

Immediately after the disaster relief items, from purchased goods, pre-positioned stock, and in-kind donations flow into the disaster affected region, in high volumes and over a short period of time. The arriving in-kind donations can be classidied into three groups as *(i)* high-priority (HP), *(ii)* low-priority (LP) and *(iii)* non-priority (NP) items (Pan American Health Organization and World Health Organization (2001)). The HP items refer to the supplies needed for immediate distribution, the LP items refer to supplies that are not of immediate need but can be used in later stages and NP items refer to supplies that are not at all needed and should not have been sent to the disaster region in the first place. Past disasters indicate that in many cases the NP items exceeded 50% of all donated items (Arnette and Zobel (2015)). The high proportions of non-urgent items can exasperate the complexity of donations management and divert critical resources away from high-priority ac-

tivities. This problem of high volumes of incoming donations coupled with high percentages of non-priority items is often referred to as the material convergence problem (Holguin-Veras et al. (2012a)).

In order to further understand the problems created by material convergence and low-priority goods, we interviewed practitioners from the Salvation Army and the National Voluntary Organizations Active in Disaster (NVOAD). The National VOAD is a coordinating body with 62 national members including the Salvation Army, the Red Cross, Federal Emergency Management Agency (FEMA) and many others. The interviews conducted, as well as the research on prior disasters showed the following as the top problems created by material convergence: (i) Diversion of volunteer capacity, (ii) Diversion of transportation capacity, (iii) Diversion of storage space, (iv) Requirement of waste management/recycling activities, (v) Increased complexity of coordination and decision making.

For example, the Salvation Army director who was a part of the response to the Haiti Earthquake, stated that there had to be full time dedicated people at the airport to manage and move arriving items. Many of the arriving donations were unnecessary items, that consumed volunteer capacity and created storage space issues at the airport. Another Salvation Army officer involved in the response efforts related to Hurricane Sandy, noted that the Yankee stadium was half full of in-kind, unsolicited donations. Roughly 30 Salvation Army trucks were dedicated to collecting unsolicited donations left in trash bags along the streets and this collection effort went on for three weeks. Once the donated items of bags were off the streets, they had to be sorted. The sorting process revealed that a high percentage of donated items were not usable for the disaster response purposes which required more

volunteer capacity to adequately recycle the donations. It was noted that the New York City Department of Sanitation provided the planning and resources required for the waste management stage.

Our interviews with NVOAD revealed that the unsolicited donations created donations management problems in almost every disaster. One of our interviews was with the director of NVOAD who has been involved in many response efforts. He stated that, the first in-kind donations are usually from the local community and they start to pile up hours after the disaster. The media has major impact on the timing and volume of arriving donations, underlining that excessive media coverage can result in increased material convergence. Although, this did not mean in-kind donations were not needed, saying: "In fact lots of donations are needed, if they are the right product."

The problems created by material convergence and the high volumes of low-priority in-kind donations have been recognized as a major problem during disaster response by the prior literature. Many papers focused on case studies and interviews documenting the extent of the problems created by material convergence, often times referring to it as a "second tier disaster" (Islam et al. (2013)). Even though the prior literature recognized the extent and importance of the material convergence problem, there is very little work in the literature that tries to quantify material convergence using an analytical model.

This chapter models and quantifies material convergence caused by two streams of donations arrivals, representing the solicited and unsolicited donations respectively. We use a transient queuing model to quantify material convergence and

throughput of high priority goods based on volunteer assignment decisions. In particular, we aim to answer the following research questions: (1) How can we quantify material convergence?, (2) How does the mix of high-priority and non-priority items and arrival patterns impact material convergence and volunteer assignment decisions?

To answer these research questions we model the material convergence process using a multi-server transient queuing model where the decision is to allocate available servers to between the solicited and unsolicited donation arrivals. We conclude that the evolution of queue length can be used to quantify the level of material convergence. We further conclude that, while allocating resources one should take into account high priority item percentage, the current material convergence level and the arrival rates together.

The rest of this chapter is organized as follows. Section 5.2 provides the details of the queuing model used to study the material convergence problem. Section 5.3 discusses the solution approach and Section 5.4 discusses numerical experiments and managerial insights on effective resource assignment to address material convergence. Finally, Section 5.5 concludes the chapter.

5.2 The Analytical Model for Donations

In the aftermath of a disaster, both solicited and unsolicited donations arrive to the disaster affected region, packed in boxes, pallets or crates. Prior to distribution these goods need to be unpacked. The unsolicited goods, in addition to being un-

packed, need to be sorted to separate the HP items from the NP items, and this is often a time-consuming task. We model the arrival and sorting of solicited and unsolicited donations using separate multi-server queues. We analyze the problem over T time periods, $t = 1, 2, ..., T$, and t' denotes the duration of each time period. We assume that the arrival process for both the solicited and unsolicited donations follow a time dependent Poisson process with parameters $\lambda_{s,t}$ and $\lambda_{u,t}$ respectively, where $t = 1, 2, ..., T$. We assume that there are a fixed number of total available volunteers during each period, C_t, that can be assigned to either queue, and the decision variables $c_{s,t}$ and $c_{u,t}$, represent the particular assignment of volunteers for period t, to sort the solicited and unsolicited items respectively, i.e. $C_t = c_{s,t} + c_{u,t}$. We assume that the service time for a volunteer to sort the arriving items follow an exponential distribution with parameter $\mu_{s,t}$ and $\mu_{u,t}$ for solicited and unsolicited items respectively and that $\mu_{u,t}^{-1} \geq \mu_{s,t}^{-1}$ due to the more involved process needed to sort the unsolicited items. Figure 5.1 shows both the solicited and unsolicited donation queues for sorting. Let $\rho_{i,t} = \frac{\lambda_{i,t}}{\mu_{i,t} c_{i,t}}, \quad i = s, u, \quad t = 1, .., T$ denote the utilization of queue i where $\rho_{i,t} \geq 0, \ \forall (i, t)$. Note that we permit $\rho_{i,t} \geq 0$; i.e queues are allowed to be unstable. We assume that solicited donations received contain only high-priority items ready for distribution, whereas for unsolicited donations received, the fraction of high-priority items in the arriving donations at the time t is given by h_t where $0 \leq h_t \leq 1$.

Figure 5.1: Donation Sorting Process

We analyze the queues over T time periods and quantify the material convergence. The main performance measures of interest are the throughput and queue build up at each time period. Next we describe the solution methodology.

5.3 Solution Methodology

We solve the queuing model for each period $t \in T$ that represent a t' length of interval. The state for each queue i corresponds to the number of donations in each queue i. Let $S_{s,t} = 0, 1, 2, ...$ and $S_{u,t} = 0, 1, 2, ...$ represent the state space of each queue for period $t \in T$. Then the evolution of the queues can be represented as a continuous time Markov processes (CTMC) denoted by $X_{s,t}(t'), t' \geq 0$ and $X_{u,t}(t'), t' \geq 0$, on state spaces $S_{s,t}$ and $S_{u,t}$. Let $L_{s,t}$ and $L_{u,t}$, $t = 1, 2, ..., T$ denote the queue length at each queue at the end of time period t. Then $L_{i,t}$ is the starting state for queue i at time period $t + 1$.

The stochastic evolution of each queue $i,\ i = s, u$, is characterized by the CTMC and its generator matrix $Q_{i,t}$ where $q_{i,t,j,k}$ represent the rate of transition from state j to state k and is given by Equation 5.1, where $i = s, u \quad j, k \in S_i$ and $t \in T$ and $q_{i,t,j,j} = -\sum_{j \neq k} q_{i,t,j,k}$.

$$q_{i,t,j,k} = \begin{cases} \frac{\lambda_{i,t}}{\lambda_{i,t}+(\mu_{i,t}*c_{i,t})}, & \text{if } k = j+1 \\ \frac{\mu_{i,t}*min(j,c_{i,t})}{\lambda_{i,t}+(\mu_{i,t}*c_{i,t})}, & \text{if } k = j-1 \end{cases} \tag{5.1}$$

Let $P_{i,t}(t')$ represent the state probability matrix and let $p_{i,t,j,k}(t')$ represent the transition probability starting from state j at time $t = 0$, to state k at time t, after a length of t' hours, where $p_{i,t,j,k}(t') = P_{i,t}(X_{i,t}(t') = k | X_{i,t}(0) = j), \quad j, k \in S_{i,t} \quad i = u, s\,, \quad t \in T$. Note that we assume without loss of generality that, each period t starts at zero and is of length t' hours. Then $P_{i,t}(t')$ satisfies the Chapman-Kolmogorov forward and backward differential equations as given by Equations 5.2 and 5.3. It is well known that the solution to the Chapman-Kolmogorov differential Equations 5.2 and 5.3 is given by Equation 5.4.

$$P'_{i,t}(t') = P_{i,t}(t') * Q_{i,t}, \quad i = u, s \text{ and } t \in T \tag{5.2}$$

$$P'_{i,t}(t') = Q_{i,t} * P_{i,t}(t'), \quad i = u, s \text{ and } t \in T \tag{5.3}$$

$$P_{i,t}(t') = e^{Q_{i,t}t'}, \quad i = u, s \text{ and } t \in T \tag{5.4}$$

Theoretically, the matrix exponential solution $e^{Q_{i,t}t'}$ can be defined by the series given by Equation 5.5. However, computing the solution using this series approach creates efficiency and accuracy problems in estimating the exponential of matrix $Q_{i,t}$, especially if $Q_{i,t}$ has a high dimension.

$$e^{Q_{i,t}t'} = \sum_{k=0}^{\infty} \frac{1}{k!}(Q_{i,t}t')^k \tag{5.5}$$

Moler and Loan (2003) provide numerical examples showcasing the efficiency and accuracy problems even for a 2 by 2 matrix and they also provide a survey of methods to compute the exponential of a matrix which involve approximation theory, differential equations, matrix eigenvalues and the matrix characteristic polynomial. All mentioned methods are effective for small enough matrices. For a typical disaster relief distribution, hourly arrival rates ($\lambda_{s,t}$, $\lambda_{u,t}$) can exceed 1000 items per hour and the size of the $Q_{i,t}$ matrix can increase to tens of thousands, which makes the computation of the matrix exponential very difficult. Hence other methods such as the uniformization technique, Runge-Kutta methods and the predictor-corrector methods (Maron (1987); Arsham et al. (1983)) are needed to obtain $P_{i,t}(t')$ probabilities.

The uniformization method is especially useful in our setting due to three reasons: (i) it has been shown to provide greater accuracy at a lower computational cost (Reibman and Trivedi (1988)), (ii) it allows for further improvements on computational time by allowing to leverage the form of the generator matrix and (iii) it has a clear probabilistic interpretation that translates well to the original problem (Gross and Miller (1984)). Next we discuss the uniformization method and its application to our problem setting.

5.3.1 Uniformization and the SERT Method

To apply the uniformization method, we first approximate the $M/M/C$ queues using a finite state space for numerical tractability. We truncate the infinite state space at

a large $N_{i,t}$ (discussed in Section 5.4). The truncation introduces an approximation error in the calculations. This error can be bound by collapsing all the truncated states into one absorbing state (states $A_{s,t} = N_{s,t} + 1$ and $A_{u,t} = N_{u,t} + 1$). The probability that the system is in the absorbing state will be the bound on the error introduced by truncating the infinite state space (Gross and Miller (1982)). We further report on the choice of $N_{i,t}$ and the associated truncation error in Section 5.4.

The uniformization method is based on the subordination of a Markov chain to a Poisson process. The probabilistic process of the CTMC $X_{i,t}(t'), t' \geq 0$, $i = u, s$, $t \in T$, on countable state space $S_{i,t} = \{0, 1, 2..., N_{i,t}, A_{i,t}\}$ is first represented using a DTMC, $Y_{i,t,n}, n = 0, 1, 2, ...$ on $S_{i,t}$ with transition probability matrix $\bar{P_{i,t}} = Q_{i,t}/\Lambda_{i,t} + I$ where $\Lambda_{i,t} = sup_{j \in S_{i,t}} q_{i,t,j,j}$ and a Poisson process $\{\bar{N_{i,t}}(t'), t' \geq 0\}$ with rate $\Lambda_{i,t}$, $i = s, u$, $t \in T$. The Poisson process and the DTMC are independent of each other. Note that the initial distribution of the CTMC and the constructed DTMC are the same. Then it can be shown that $\{Y_{i,t,\bar{N}_i(t')}, t' \geq 0\}$ and $\{X_{i,t}(t'), t' \geq 0\}$ are probabilistically identical (Heyman and Sobel (1982)).

This construction provides a meaningful and computationally efficient formula for the transient probabilities of the CTMC model. The number of jumps the DTMC goes through between the discrete observation epochs is modeled by the Poisson process and the multi-step transition probabilities of the DTMC identifies what state the CTMC evolves to as a result of the transition. To get the transient probabilities for $X_{i,t}(t'), t' \geq 0$, we condition on the number of arrivals in $[0, t']$ and we use the law of total probability as shown in Equation 5.6.

$$
\begin{aligned}
p_{i,t,j,k}(t) &= P\{X_{i,t}(t') = k | X_{i,t}(0) = j\} \\
&= P\{Y_{i,t\bar{N}_{i,t}(t')} = k | X_{i,t}(0) = j\} \\
&= \sum_{n=0}^{\infty} P\{Y_{i,t\bar{N}_{i,t}(t')} = k | \bar{N}_{i,t}(t') = n, X_{i,t}(0) = j\} P\{Y_{i,t\bar{N}_{i,t}(t')} = n\} \\
&= \sum_{n=0}^{\infty} P\{Y_{i,t,n} = k | X_{i,t}(0) = j\} e^{\Lambda_{i,t}t'} (\Lambda_{i,t}t')^n / n! \qquad (5.6)
\end{aligned}
$$

Finally we obtain Equation 5.7 for the state probability vector of the CTMC represented by $p_{i,t}(t')$, where $\bar{p}_{i,t}(n) = \bar{p}_{i,t}(0)\bar{P}_{i,t}^n$ represent the state probabilities of the DTMC. Note that the initial state probabilities are assumed to be the same for both the DTMC and the CTMC, $p_{i,t}(0) = \bar{p}_{i,t}(0)$, $i = u, s$, $t \in T$.

$$
p_{i,t}(t') = \sum_{n=0}^{\infty} \bar{p}_{i,t}(n) e^{\Lambda_{i,t}t'} (\Lambda_{i,t}t')^n / n! \quad i = u, s, \quad t \in T \qquad (5.7)
$$

Note that, Equation 5.7 also involves the computation of an infinite series. For numerical computation, we truncate the series at $n = K$. We can bound the error this truncation creates, by letting K satisfy Equation 5.8 (Fox and Glynn (1988)). Note that including this truncation, the solution approach now has two sources of error; the truncation of the state space of the CTMC and the truncation of the Poisson process in the uniformization. The first error is estimated by adding an absorbing state to the CTMC and the latter is bounded by choosing an appropriate ϵ. We report on both of these errors in Section 5.4.

$$
1 - e^{\Lambda_{i,t}t'} \sum_{n=0}^{K} (\Lambda_{i,t}t')^n / n! \leq \epsilon, \; i = u, s, \; t \in T \qquad (5.8)
$$

Computationally the uniformization approach is more efficient and stable than computing the matrix exponential since it requires less terms until the series converge and avoids problems due to the truncation of arithmetic. However, the size of the probability transition matrix, $\bar{P}_{i,t}$ for the DTMC and computing its powers ($\bar{P}^n_{i,t}$) can be computationally cumbersome for our setting. To increase the computational efficiency we use the SERT modeling approach (Gross and Miller (1984)). This approach represents the transitions of the matrix by storing the state space, types of events, transition probabilities and target states for the transitions as separate vectors. Specifically, S - denotes the state space, E - denotes the set of events, R (or P) - denotes the rate (or probability) vectors and, T - denotes the target state vector. This approach also allows us to take advantage of the sparsity of the $\bar{P}_{i,t}$ matrix and increase computational efficiency.

$S_{i,t}$ = state space

$E_{i,t}$ = set of types of events

$\bar{P}_{i,t}$ = set of probability ($\bar{P}_{i,t}$) vectors (one for each element of $E_{i,t}$, each with $|S_{i,t}|$ components)

$T_{i,t}$ = set of target state vectors (one for each element of $E_{i,t}$, each with $|S_{i,t}|$ components)

For an arrival rate $\lambda_{i,t}$, service rate $\mu_{i,t}$ and volunteer allocation of $c_{i,t}$ for each $i = u, s$ and $t = 1, 2, .., T$ we can define the SERT sets as follows:

$$
\begin{aligned}
S_{i,t} &= \{0, 1, 2, ..., N_{i,t}, A_{i,t}\} \text{ where } A_{i,t} = N_{i,t} + 1 \text{ represents the absorbing state} \\
E_{i,t} &= \{a, d, n\} \text{ where a, d, n represent arrival, departure and null events} \\
\bar{P}^a_{i,t} &= \{\frac{\lambda_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, \frac{\lambda_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, ..., 0\} \\
\bar{P}^d_{i,t} &= \{\frac{0}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, \frac{\mu_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, \frac{2\mu_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, ..., \frac{c_{i,t}\mu_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, 0\} \\
\bar{P}^n_{i,t} &= \{1 - \frac{\lambda_{i,t}}{\lambda_{i,t} + (c_{i,t}mu_{i,t})}, 1 - \frac{\lambda_{i,t} + \mu_{i,t}}{\lambda_{i,t} + (c_{i,t}\mu_{i,t})}, ..., 1\} \\
T^a_{i,t} &= \{1, 2, ..., N_{i,t}, A_{i,t}, A_{i,t}\} \\
T^d_{i,t} &= \{0, 0, 1, ..., N_{i,t} - 1, A_{i,t}\} \\
T^n_{i,t} &= \{0, 1, 2, ..., N_{i,t}, A_{i,t}\} \qquad (5.9)
\end{aligned}
$$

In this representation, $P^a_{i,t}$ represents the set of probabilities where the event is an arrival instead of a departure or null event (meaning the state remains the same) and the $T^a_{i,t}$ vector represents the target state the DTMC will transition to upon an arrival event. Each entry in both vectors correspond to an element of the state space. The exact same structure applies to the departure and null events. An example is, the first entry of the $P^d_{i,t}$ vector is 0 representing the probability of having a departure when the system is at state 0 (empty) being 0 and the first entry of the vector $T^d_{i,t}$ is also 0 representing the state we transition to as state 0, since there was no change to the number in queue. The above sets fully define the DTMC.

The SERT approach reduces computational complexity involved in calculating the DTMC probability transitions in two ways: (1) instead of using matrix multiplications to compute all state transition probabilities ($\bar{P}^n_{i,t}$), it only computes the transition probabilities from the given starting state vector and (2) it only takes

into account positive entries in $\bar{P}_{i,t}$ (represented as events) therefore minimizing multiplications by 0. This becomes especially effective for sparse matrices such as $\bar{P}_{i,t}$. For example if $\bar{P}_{i,t}$ was a 4 by 4 matrix, calculating $\bar{P}^3_{i,t}$ would require 128 multiplications. This number goes down to 20 multiplications by using the SERT representation.

Using the sets defined by Equation 5.9, we can calculate the state probability vector of the DTMC, $\bar{p}_{i,t}(n)$, using Equation 5.10. Note that Equation 5.10 calculates the state probabilities individually for all states $s \in S_{i,t}$. Note that the third subscript s refers to the s-th element of that vector.

$$\begin{aligned} \bar{p}_{i,t,s}(n+1) &= \bar{p}_{i,t,s}(n)P^n_{i,t,s} \\ \bar{p}_{\{t^a_{i,t,s} \in T^a_{i,t}\}}(n+1) &= \bar{p}_{i,t,s}(n)P^a_{i,t,s} \\ \bar{p}_{\{t^d_{i,t,s} \in T^d_{i,t}\}}(n+1) &= \bar{p}_{i,t,s}(n)P^d_{i,t,s} \end{aligned} \tag{5.10}$$

Substituting solutions of Equation 5.10 in Equation 5.7, we can obtain the state probabilities, $p_{i,t,j,k}(t)$, that denote the length of queue i at time t.

Once we obtain the probabilities $p_{i,t,j,k}(t)$, we calculate the expected throughput from each donation queue i, for time interval t, given a specific volunteer assignment $c_{i,t}$. This proves to be challenging since the throughput $TH_{i,t}$ is path dependent. As an example, if the queue i transitions from state $j = 1$ at $t = 0$ to $k = 2$ at time t, it could follow many paths such as, (i) exactly one arrival happening between time $t = 0$ and $t = t'$ and no service completion, or (ii) exactly two arrivals and one departure happening between time $t = 0$ and $t = t'$, or (iii) exactly three arrivals and

two departures happening between time $t = 0$ and $t = t'$, or (iv) exactly n arrivals and (n-1) departures happening between time $t = 0$ and $t = t'$ and so on. Note that the throughput from queue i between $t = 0$ and $t = t'$ depends on the number of departures from the queue in this time period. Therefore we use Algorithm 1 to compute the expected throughput, $TH_{i,t}$ for each queue i during time period t. In the inner most loop, the algorithm first calculates $TH'_{j,k}$, which is the expected throughput of each possible transition. To calculate $TH'_{j,k}$, the algorithm loops through possible paths (truncated at $M_{i,t}$), given an initial state $j = a_{i,t}$ and for each path, conditions the calculations based on queue length increasing or decreasing. It calculates the probability of each path, using the probability distribution of the arrival process alone (Poisson) since the path probability is conditioned over the particular state transition. Here, $V_{i,t}(t'), t' \geq 0$ represents the random variable for the Poisson process with rate $(\lambda_{i,t}t')$. Then in the outer loop, the algorithm loops through all possible transitions in the state space. Finally in the outer most loop, the algorithm computes the overall expected throughput for the queue.

The performance measure we use is the total high-priority item throughput for each $t \in T$ and we track the length of each queue i as a measure of material convergence. Both the material convergence and the queue throughput evolve over time. The connection between time periods are made by updating the starting state $a_{i,t}$ for each time interval. The decision parameter at each interval is the volunteer assignment to each queue $c_{i,t}, i = s, u$. The arrival rates, service rates, volunteer assignment and the current material convergence level impact the throughput measure. Moreover, the volunteer assignment decision and the level of material convergence depend on each other as well. In the next section, we numerically investigate the relationship between material convergence, volunteer assignment and throughput.

Algorithm 1 Estimating Expected Throughput

For any (i,t) pair where $i = u, s \quad t = 1, 2, .., T$
$a_{i,t} \leftarrow$ starting state
$k \leftarrow k$ *in state space* $S_{i,t}$
$M_{i,t} \leftarrow$ *sum limit; greater than size of state space* $N_{i,t}$
for $k = 0 : N_{i,t}$ **do**
  if $k \geq a_{i,t}$ **then**
    $dif = k - a_{i,t}$
    $TH'_{a_{i,t},k} = [\sum_{j=0}^{M_{i,t}} P\{V_{i,t}(t') = dif + j\} * j]$
  else if $k < a_{i,t}$ **then**
    $dif = a_{i,t} - k$
    $TH'_{a_{i,t},k} = [\sum_{j=0}^{M_{i,t}} P\{V_{i,t}(t') = j\} * (dif + j)]$
  end if
end for
$TH_{i,t} = \sum_{k=0}^{N_{i,t}} TH_{a_{i,t},k} * p_{i,t,a_{i,t},k}(t)$

5.4 Numerical Experiments

We present two sets of numerical experiments to investigate: (1) effects of material convergence and (2) effects of HP item percentage of unsolicited donations on resource allocation decisions. In experiment set 1 we focus on quantifying material convergence and investigating the relationship between material convergence and volunteer assignment. In experiment set 2, we add the consideration of HP percentage for unsolicited donations and observe how the volunteer assignment strategy changes as the unsolicited donations have less and less high-priority goods.

5.4.1 Experiment Set 1: Effect of Volunteer Assignment

The parameter setup for the numerical experiments are based on our collaboration with Goodwill and the Salvation Army, both of which carry out voucher programs to distribute household and clothing type of relief items for disaster relief. Accord-

ing to data from Goodwill stores, the sorting rate of donations vary between 150 to 240 items per hour. We base the service rate used in our models for unsolicited donation sorting, on this data obtained from Goodwill and use an average of 200 items per hour per volunteer for unsolicited donations and 300 per hour per volunteer for solicited donations. The sorting process during the disaster response period is very similar to sorting mixed, unsolicited items at Goodwill. However, the arrival rate of donations during disaster response is much higher than what is observed day to day at a Goodwill store. Therefore, we vary the arrival rate in our numerical experiments based on the expected utilization levels during the disaster aftermath. It is expected that the donation arrival rates will vary from disaster to disaster and even hourly. For more details on interviews with Goodwill and Salvation Army, we refer the reader to the technical report, Ozen and Krishnamurthy (2017).

Experiment cases 1, 2 and 3 represent different levels of arrival rates. In case 1 we picked an arrival rate such that the utilization of both queues can be kept below 1 with the total number of available volunteers. In case 2, we picked an arrival rate to represent a scenario where no allocation of resources will make the utilization of either queue less than 1. Finally in case 3, we picked an arrival rate such that, depending on the allocation of resources, one of the queues can have a utilization less than 1. In all three cases, arrivals to both queues are assumed to be from a Poisson process with $\lambda_{s,t} = \lambda_{u,t},\ \forall t$. We assume a total of $C_t = 10,\ \forall t$ volunteers are available, all unsolicited donations are high-priority, $h_t = 100\%,\ \forall t$ and all items have equal priority.

We investigate the evolution of material convergence and resource assignment decisions over 12 time intervals, i.e $t = 1, 2, ..., T$ where $T = 12$, and each time

interval represents an 8 hour period, i.e $t' = 8$. We assume the arrival and service rates are equal for each time period, i.e $\lambda_{s,t} = \lambda_{u,t},\ \mu_{s,t} = \mu_{u,t}, t = 1, 2, ..., 12,$ to clearly see the relationship between material convergence and resource allocation. We investigate the impact of high priority item percentage in the experiment set 2.

Also note that we choose the truncation point $N_{i,t} = a+(\lambda_{i,t} * t'), \quad i = s, u,\ t \in T$ where a represents the starting state and is equal to the average queue length of queue i during the previous time period ($a = \bar{W}_{i,t-1}$) while t' represents the number of hours in a given time interval. The truncation point used aims to make the capacity large enough to accommodate the total number of arrivals and to make sure the probability of an arrival finding the queue full is minimal.

Table 5.1: Experiment 1: Parameters

Parameters	Case 1	Case 2	Case 3
$\lambda_{s,t}$, $\forall t$	1000/hr	3000/hr	1250/hr
$\lambda_{u,t}$, $\forall t$	1000/hr	3000/hr	1250/hr
$\mu_{s,t}$, $\forall t$	300/hr	300/hr	300/hr
$\mu_{u,t}$, $\forall t$	200/hr	200/hr	200/hr
C_t, $\forall t$	10	10	10
Average System Utilization	80 %	240 %	100 %
h_t, $\forall t$	1	1	1
Priority	Equal	Equal	Equal
Demand Pattern	Steady	Steady	Steady

In each time interval, we search over all possible volunteer allocation decisions and choose the allocation that provides the optimal high-priority item throughput. It is important to note that this allocation may not reflect the global optimum solution for the resource allocation problem. The approach used to get these results, evaluates the overall throughput, given the current queue length, for all possible allocation decisions in the given time interval and hence represent a local optimum

decision for the overall performance. We will call this policy the single-period optimal policy (SOP).

Tables 5.2, 5.3, 5.4 show the results resource allocation, throughput and material convergence (queue length) results for cases 1, 2 and 3 respectively. We observe that for case 1, the allocation of $c_{u,t} = 6$ and $c_{s,t} = 4$ is optimal for all t, since it makes the utilization of both queues less than 1. Material convergence does not happen since both queues are balanced and almost all arrivals are serviced.

Table 5.2: Case 1 Material Convergence: System Utilization Below 1

	Resources		Utilization		Queue Length		Throughput		
t	$c_{u,t}$	$c_{s,t}$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	6	4	0.83	0.83	7.93	6.62	7,988.10	7,989.42	15,977.52
2	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
3	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
4	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
5	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
6	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
7	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
8	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
9	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
10	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
11	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30
12	6	4	0.83	0.83	7.93	6.62	7,995.51	7,995.79	15,991.30

For case 2, we observe a very different behavior. No allocation of available re-

sources make the utilization of either queue less than 1 and therefore most resources (all volunteers) are allocated to the faster queue (solicited donations). Material convergence is observed for both queues, with the unsolicited queue lengths exceeding 40,000 within the first three days. The overall throughput is equal to the throughput of the solicited donations queue. Note that, in Table 5.3 queue length for the unsolicited donations queue is unreported due to computational time and storage space problems.

Table 5.3: Case 2 Material Convergence: System Utilization Above 1

	Resources		Utilization		Queue Length		Throughput		
t	$c_{u,t}$	$c_{s,t}$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	0	10	∞	1.00	0.0	0.0	0.0	23,807.67	23,807.67
2	0	10	∞	1.00	23,988.10	180.51	0.0	23,935.05	23,935.05
3	0	10	∞	1.00	46,364.00	233.60	0.0	23,954.00	23,954.00
4	0	10	∞	1.00	68,728.91	267.58	0.0	23,962.95	23,962.95
5	0	10	∞	1.00	*	292.61	0.0	23,968.20	23,968.20
6	0	10	∞	1.00	*	312.34	0.0	23,971.54	23,971.54
7	0	10	∞	1.00	*	327.99	0.0	23,973.96	23,973.96
8	0	10	∞	1.00	*	341.56	0.0	23,975.82	23,975.82
9	0	10	∞	1.00	*	353.69	0.0	23,977.24	23,977.24
10	0	10	∞	1.00	*	364.27	0.0	23,978.31	23,978.31
11	0	10	∞	1.00	*	373.19	0.0	23,979.19	23,979.19
12	0	10	∞	1.00	*	381.31	0.0	22,387.84	22,387.84

Case 3 represents the scenario where the overall system utilization is equal to

1. Table 5.4 shows the SOP resource allocation, the material convergence (queue lengths) and the throughput results. The first observation we make based on the results is, neither the queue lengths nor the local optimum allocation decisions converge. The second observation is the interaction between material convergence and resource allocation decisions. Given empty queues at the beginning, we would expect a $(c_{u,t}, c_{s,t}) = (6, 4)$ assignment throughout the time horizon, since this allocation would make both the queues to be utilized 104%, and result in neither queue having a utilization way over 1. However, as the system evolves , we observe material convergence increasing in both queues. When the convergence at the solicited donations queue hits the 1200 mark at $t = 4$, it makes sense to move an additional resource to this queue since it has a faster service rate and the convergence of materials creates a sense of increasing arrival rate (hence increasing utilization). We observe that this decision decreases the material convergence at the solicited donations queue down to 7 by the next time interval. Then over the next three time intervals we observe the material convergence increase in both queues. Then at time interval $t = 8$ the material convergence at the solicited donations queue hits 1200 again and results in a volunteer to be moved to the solicited donations queue. The reason behind this behavior repeating is due to the material convergence increasing over time and warranting additional resources periodically.

Table 5.4: Case 3 Material Convergence: System Utilization at 1

	Resources		Utilization		Queue Length		Throughput		
t	$c_{u,t}$	$c_{s,t}$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	6	4	1.04	1.04	0.0	0.0	9,568.22	9,569.58	19,137.80
2	6	4	1.04	1.04	426.81	425.44	9,595.25	9,595.25	19,190.49
3	6	4	1.04	1.04	826.56	824.56	9,595.25	9,595.25	19,190.49
4	5	5	1.25	0.83	1,226.36	1,224.36	7,996.18	11,210.23	19,206.41
5	6	4	1.04	1.04	3,224.27	7.26	9,595.25	9,575.87	19,171.11
6	6	4	1.04	1.04	3,622.17	426.15	9,595.25	9,595.25	19,190.49
7	6	4	1.04	1.04	4,019.97	825.56	9,595.25	9,595.25	19,190.49
8	5	5	1.25	0.83	4,417.78	1,225.36	7,996.18	11,211.23	19,207.41
9	6	4	1.04	1.04	6,414.69	7.26	9,595.25	9,575.87	19,171.11
10	6	4	1.04	1.04	6,811.58	426.15	9,595.25	9,595.25	19,190.49
11	6	4	1.04	1.04	7,208.39	825.56	9,595.25	9,595.25	19,190.49
12	5	5	1.25	0.83	7,604.19	1,225.36	7,996.18	11,211.23	19,207.41

Figures 5.2 and 5.3 compare the material convergence for unsolicited and solicited donations respectively, for all 3 cases. Comparing the material convergence levels, we observe that the convergence experienced at the unsolicited donations queue is higher than the solicited donations in all cases. This is because more resources are allocated to the solicited donations queue due to having a higher service rate and equivalent arrival rate. For the solicited donations queue, we observe the oscillating behavior for case 3, because the allocation of resources shifts based on convergence. For case 1, both queues have a utilization well below 1, and hence no major convergence issues are observed. For case 2, both queues have a utilization equal to or above 1, and hence we observe a steadily increasing convergence for both queues.

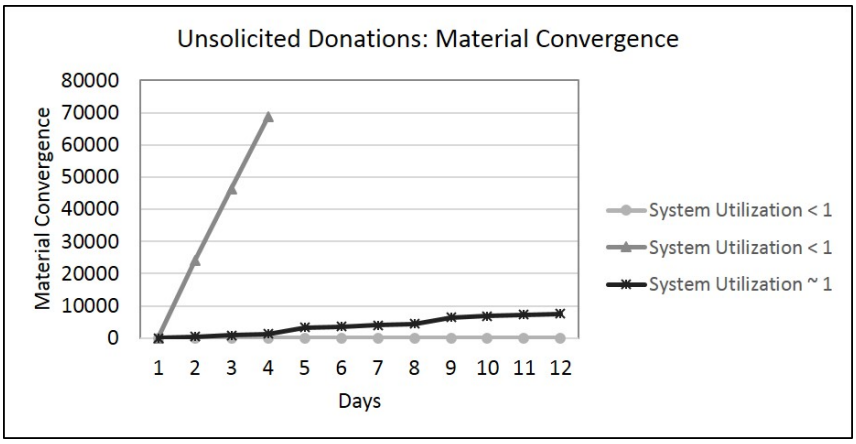

Figure 5.2: Material Convergence for Unsolicited Donations

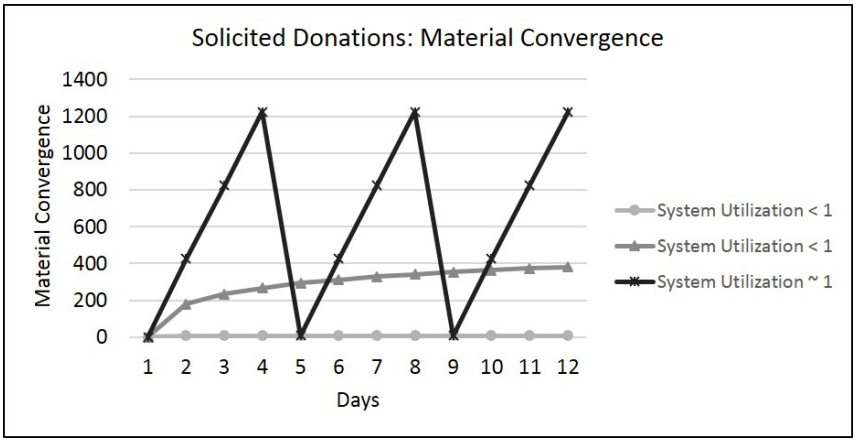

Figure 5.3: Material Convergence for Solicited Donations

In Table 5.5 we compare the throughput results of optimally allocating resources in each interval to a stationary policy that is based on the arrival and service rates alone. Given the unsolicited donations sorting rate of 200 per hour and the solicited donations sorting rate of 300 per hour, the stationary policy makes the assignment of $c_{u,t} = 6$ and $c_{u,t} = 4$, which makes utilization of both queues 1.04. The stationary policy does not take into account the effect of material convergence on resource allocation and throughput. We observe that the SOP policy allocation performs better from a total high-priority item perspective, since it takes into account the combined

impact of material convergence and volunteer allocation on throughput. Moreover, it is important to note the difference in the material convergence evolution for the SOP and stationary policies as it is given in Figure 5.5 and Table 5.5. In the SOP policy, the solicited donations queue length does not grow beyond 1300. This is due to the pattern of allocation where every 4 time intervals, an additional resource is allocated to the solicited donations queue due to the effect of increasing material convergence. The material convergence at the unsolicited queue keeps increasing since its utilization never falls below 1. For the stationary policy, we observe that the material convergence in both queues increase steadily over time. This is because the initial allocation makes the utilization of both queues over 1, for all time periods.

Table 5.5: Comparison of SOP and Stationary Policy

	SOP Policy			Stationary Policy		
t	$L_{u,t}$	$L_{s,t}$	$\sum_i TH_i, t$	$L_{u,t}$	$L_{s,t}$	$\sum_i TH_i, t$
1	0.0	0.0	19,137.80	0.0	0.0	19,135.80
2	426.81	425.44	19,190.49	426.81	425.44	19,190.49
3	826.56	824.56	19,190.49	826.56	824.56	19,190.49
4	1,226.36	1,224.36	19,206.41	1,226.36	1,224.36	19,190.49
5	3,224.27	7.26	19,171.11	1,625.17	1,623.17	19,190.49
6	3,622.17	426.15	19,190.49	2,023.97	2,021.97	19,190.49
7	4,019.97	825.56	19,190.49	2,422.77	2,420.77	19,190.49
8	4,417.78	1,225.36	19,207.41	2,422.77	2,420.77	19,190.49
9	6,414.69	7.26	19,171.11	2,821.57	2,819.57	19,190.49
10	6,811.58	426.15	19,190.49	3,220.37	3,218.37	19,190.49
11	7,208.39	825.56	19,190.49	3,618.17	3,616.17	19,190.49
12	7,604.19	1,225.36	19,207.41	4,015.98	4,013.98	19,190.49
Total			230,244.21			230,231.22

Material convergence and resource allocation decisions are interdependent due to the change in the probability that the queue is empty. When material convergence increases, the probability that the server of this queue will ever be idle decreases significantly. This is why the SOP moves resources to the solicited donations queue

when the material convergence increases in that queue, which results in a higher total throughput. We observe the same trade-off at multiple points in time as the queue lengths grow.

One of the important takeaways from experiment set 1 is that the service rate, the arrival rate and the level of convergence effect the resource allocation decisions even for the most simple case where $h_t = 100\%$, $\forall t \in T$, item priorities are equal and arrival rates are stable. Therefore, it is crucial to understand the effect of material convergence and provide practitioners with analytic tools to evaluate the impact of possible decisions.

5.4.2 Experiment Set 2: Effect of HP Item Percentage

In experiment set 2, we compare 4 different cases where the high priority (HP) item percentage is 0.8, 0.6, 0.4 and 0.2. The rest of the parameters are same as experiment 1. Table 5.6 summarizes the parameters. We discuss the evolution of volunteer assignments, the material convergence at each queue and the throughput of the system for each case.

Table 5.6: Experiment 2: Parameters

Parameters	Case 1	Case 2	Case 3	Case 4
h_t, $\forall t$	0.8	0.6	0.4	0.2
$\lambda_{s,t}$, $\forall t$	1250/hr			
$\lambda_{u,t}$, $\forall t$	1250/hr			
$\mu_{s,t}$, $\forall t$	300/hr			
$\mu_{u,t}$, $\forall t$	200/hr			
C_t, $\forall t$	10			

In Tables 5.7, 5.8, 5.9 and 5.10 we report the resource allocation, material con-

vergence and throughput results for $h_t = 0.8, 0.6, 0.4, 0.2$ cases respectively. First observation is the difference in the volunteer assignment decisions (See Table 5.11 for a comparison). In particular the frequency of an additional resource allocation to the solicited donations queue increases as the HP item percentage of the unsolicited donations increase. This is due to the material convergence (queue length) at the solicited queue becoming more and more important, due to the HP item throughput from the unsolicited queue decreasing. Next, we compare the material convergence behavior and observe that as the HP item percentage decreases for the unsolicited donations, the material convergence in this queue increases significantly. In the $h_t = 0.2$ cases, this behavior is at its most extreme. For $h_t = 0.2$ the unsolicited donations pile up significantly, while the solicited donations have an average of 7 items in queue. At this point, the solicited donations queue is the top priority and resources are allocated such that the utilization of this queue is at 80%. Since going below 80% utilization does not result in a notable increase in the throughput, the remaining resources are allocated to the unsolicited queue. Finally, we compare the HP item throughput from each queue and the total expected HP item throughput (See Figure 5.4 for an overall throughput comparison). We observe that there is an increase in solicited donations queue throughput when h_t drops from 0.8 to 0.6. This is due to more resources being allocated to the solicited donations queue. However, as h_t drops further and more resources are allocated to the solicited donations queue, we observe that the throughput of the solicited donations queue remain at the same level. This is because allocating more and more resources to the solicited donations queue has diminishing returns on throughput.

Table 5.7: $h_t = 0.8$ Results: Resource Allocation, Queue Lengths and Throughput

	Resource Allocation	Utilization		Queue Length		Throughput		
t	$(c_{u,t}\ ,\ c_{s,t})$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	(6,4)	1.04	1.04	0.0	0.0	7,654.57	9,569.58	17,224.15
2	(6,4)	1.04	1.04	426.81	425.44	7,676.20	9,595.25	17,271.44
3	(6,4)	1.04	1.04	826.56	824.56	7,676.20	9,595.25	17,271.44
4	(5,5)	1.25	0.83	1,226.36	1,224.36	6,396.94	11,209.23	17,606.17
5	(6,4)	1.04	1.04	3,224.27	7.26	7,676.20	9,575.87	17,252.06
6	(6,4)	1.04	1.04	3,622.17	426.15	7,676.20	9,595.25	17,271.44
7	(6,4)	1.04	1.04	4,019.97	825.56	7,676.20	9,595.25	17,271.44
8	(5,5)	1.25	0.83	4,417.78	1,225.36	6,396.94	11,211.23	17,608.17
9	(6,4)	1.04	1.04	6,414.69	7.26	7,676.20	9,575.87	17,252.06
10	(6,4)	1.04	1.04	6,811.58	426.15	7,676.20	9,595.25	17,271.44
11	(6,4)	1.04	1.04	7,208.39	825.56	7,676.20	9,595.25	17,271.44
12	(5,5)	1.25	0.83	7,604.19	1,225.36	6,396.94	11,211.23	17,608.17
Total						88,254.98	119,924.48	208,179.47

Table 5.8: $h_t = 0.6$ Results: Resource Allocation, Queue Lengths and Throughput

	Resource Allocation	Utilization		Queue Length		Throughput		
t	$(c_{u,t}, c_{s,t})$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	(6,4)	1.04	1.04	0.0	0.0	5,740.93	9,569.58	15,310.51
2	(6,4)	1.04	1.04	426.81	425.44	5,757.15	9,595.25	15,352.39
3	(5,5)	1.25	0.83	826.56	824.56	4,797.71	10,811.43	15,609.13
4	(6,4)	1.04	1.04	2,825.46	7.26	5,757.15	9,575.87	15,333.01
5	(6,4)	1.04	1.04	3,223.37	426.15	5,757.15	9,595.25	15,352.39
6	(5,5)	1.25	0.83	3,621.17	825.56	4,797.71	10,812.43	15,610.13
7	(6,4)	1.04	1.04	5,618.08	7.26	5,757.15	9,575.87	15,333.01
8	(6,4)	1.04	1.04	6,014.98	426.15	5,757.15	9,595.25	15,352.39
9	(5,5)	1.25	0.83	6,411.78	825.56	4,797.71	10,812.43	15,610.13
10	(6,4)	1.04	1.04	8,407.70	7.26	5,757.15	9,575.87	15,333.01
11	(6,4)	1.04	1.04	8,803.59	426.15	5,757.15	9,595.25	15,352.39
12	(5,5)	1.25	0.83	9,199.39	825.56	4,797.71	10,812.43	15,610.13
Total						65,231.80	119,926.87	185,158.67

Table 5.9: $h_t = 0.4$ Results: Resource Allocation, Queue Lengths and Throughput

	Resource Allocation	Utilization		Queue Length		Throughput		
t	$(c_{u,t}, c_{s,t})$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	(6,4)	1.04	1.04	0.0	0.0	3,827.29	9,569.58	13,396.87
2	(5,5)	1.25	0.83	426.81	425.44	3,198.47	10,411.62	13,610.09
3	(6,4)	1.04	1.04	2,425.66	7.26	3,838.10	9,575.87	13,413.96
4	(5,5)	1.25	0.83	2,824.57	426.15	3,198.47	10,412.62	13,611.09
5	(6,4)	1.04	1.04	4,822.48	7.26	3,838.10	9,575.87	13,413.96
6	(5,5)	1.25	0.83	5,219.38	426.15	3,198.47	10,412.62	13,611.09
7	(6,4)	1.04	1.04	7,215.29	7.26	3,838.10	9,575.87	13,413.96
8	(5,5)	1.25	0.83	7,611.19	426.15	3,198.47	10,412.62	13,611.09
9	(6,4)	1.04	1.04	9,606.11	7.26	3,838.10	9,575.87	13,413.96
10	(5,5)	1.25	0.83	10,001.00	426.15	3,198.47	10,412.62	13,611.09
11	(6,4)	1.04	1.04	11,994.92	7.26	3,838.10	9,575.87	13,413.96
12	(5,5)	1.25	0.83	12,388.81	426.15	3,198.47	10,412.62	13,611.09
Total						19,189.65	119,924.67	139,114.32

Table 5.10: $h_t = 0.2$ Results: Resource Allocation, Queue Lengths and Throughput

	Resource Allocation	Utilization		Queue Length		Throughput		
t	$(c_{u,t}, c_{s,t})$	$\rho_{u,t}$	$\rho_{s,t}$	$L_{u,t}$	$L_{s,t}$	$TH_{u,t}$	$TH_{s,t}$	$\sum_i TH_i, t$
1	(5,5)	1.25	0.83	0.0	0.0	1,598.06	9,987.83	11,585.89
2	(5,5)	1.25	0.83	2,004.77	7.26	1,599.24	9,994.26	11,593.49
3	(5,5)	1.25	0.83	4,002.88	7.26	1,599.24	9,994.26	11,593.49
4	(5,5)	1.25	0.83	5,999.89	7.26	1,599.24	9,994.26	11,593.49
5	(5,5)	1.25	0.83	7,995.90	7.26	1,599.24	9,994.26	11,593.49
6	(5,5)	1.25	0.83	9,990.92	7.26	1,599.24	9,994.26	11,593.49
7	(5,5)	1.25	0.83	11,984.93	7.26	1,599.24	9,994.26	11,593.49
8	(5,5)	1.25	0.83	13,977.94	7.26	1,599.24	9,994.26	11,593.49
9	(5,5)	1.25	0.83	15,969.96	7.26	1,599.24	9,994.26	11,593.49
10	(5,5)	1.25	0.83	17,960.97	7.26	1,599.24	9,994.26	11,593.49
11	(5,5)	1.25	0.83	19,950.99	7.26	1,599.24	9,994.26	11,593.49
12	(5,5)	1.25	0.83	21,940.00	7.26	1,599.24	9,994.26	11,593.49

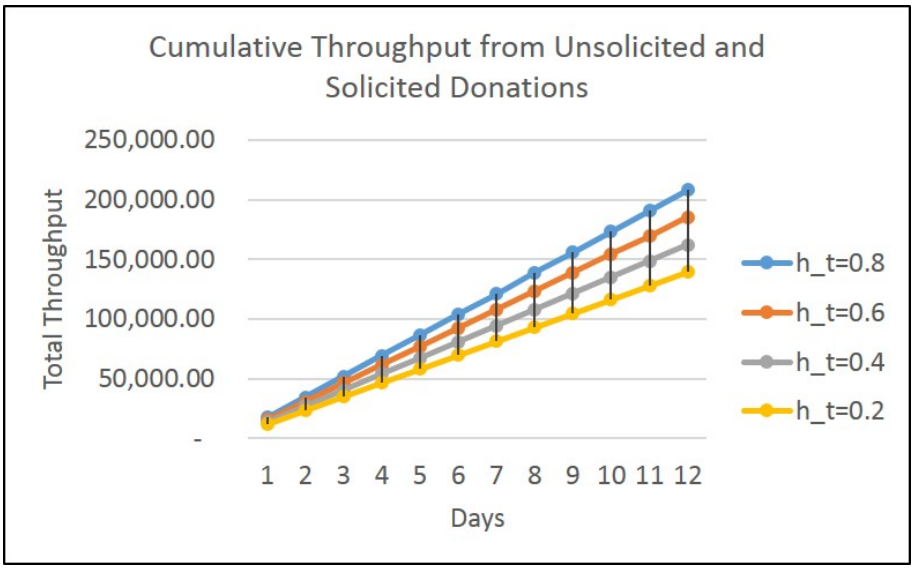

Figure 5.4: Cumulative Throughput from All Donations

Figure 5.5 and 5.6 show the material convergence at the unsolicited and solicited donation queues respectively. There are two take-aways from these graphs: first, the material convergence at the unsolicited donations queue steadily increases over time for any HP percentage level. This happens because the system prioritizes the solicited donations queue since it is faster, has the same arrival rate and has equal or higher proportion of high-priority items. The second important take away is the material convergence pattern observed for the solicited donations queue. When the queue length of solicited donations hit a certain level (determined by the HP percentage of the unsolicited donations), this queue warrants an additional volunteer. The required queue length is lower for lower HP percentage.

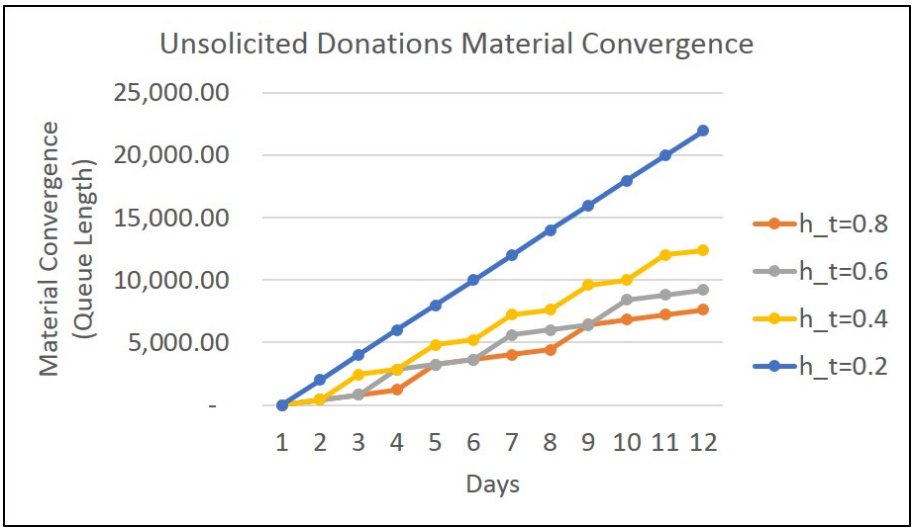

Figure 5.5: Unsolicited Donations Material Convergence

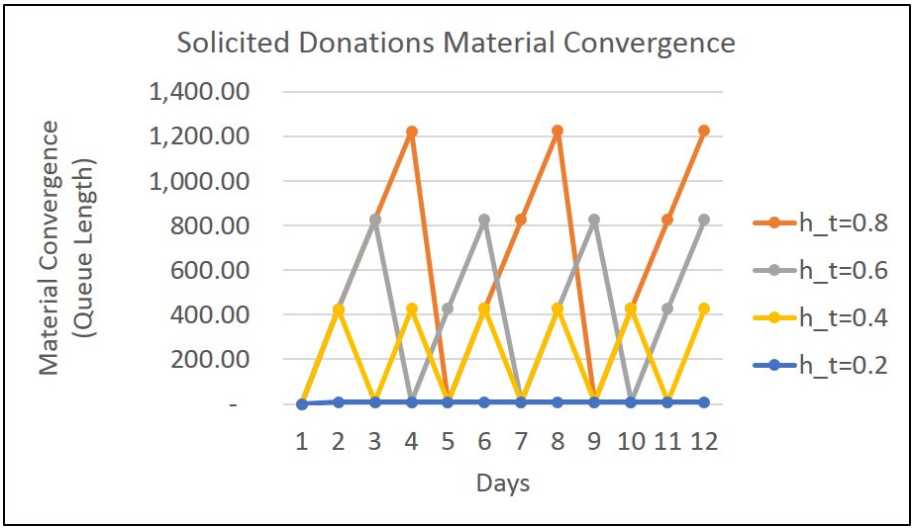

Figure 5.6: Solicited Donations Material Convergence

Next we compare the local optimum allocation patterns with the stationary policy of $(c_{u,t}, c_{s,t}) = (6, 4),\ \forall t$. Table 5.11 compares the overall HP item throughput. Here we observe that, the stationary allocation becomes less and less effective as HP% increases. This is because the stationary allocation only considers the initial arrival and service rates and disregards the effects of both the HP% and material convergence.

Table 5.11: Stationary Policy Comparison

Case	SOP Policy $\sum_i TH_i, t$	Stationary Policy $\sum_i TH_i, t$
$h_t = 0.8$	208,179.00	207,208.24
$h_t = 0.6$	185,158.00	184,185.25
$h_t = 0.4$	162,132.00	161,162.27
$h_t = 0.2$	139,114.00	138,139.28

5.5 Conclusions

This chapter quantifies material convergence of both solicited and unsolicited in-kind donations using a transient queuing model. Convergence of donated items is a critical and common phenomenon recorded in almost all major disasters. Often the material convergence happens in such high volumes that donations block international airports, fill warehouses and stadiums and divert significant resources from other disaster response tasks. In addition to the complexities created by volume, unsolicited donations containing useless items exasperate the problem greatly. These items that have no use for the relief efforts take away significant volunteer capacity.

We model the donation arrivals and sorting for both solicited and unsolicited donations as separate multi-server transient queues over multiple periods. We discuss efficient ways to solve this model for high disaster response parameters. We estimate model parameters based on our interviews and data from the Salvation Army and Goodwill. In our first numerical experiment set we investigate the relationship between material convergence and resource allocation decisions. We observe that material convergence and resource allocation decisions are inter-dependent. The numerical experiments showed the interaction between material convergence and allocation of resources. Next we investigate the effects of varying high-priority goods percentage of unsolicited donations and whether there is a point under which re-

source allocation to unsolicited donations is not warranted.

In our numerical studies, we investigate the relationships between resource allocation decision, material convergence and high-priority item percentage of unsolicited donations. We use high-priority item throughput maximization as a performance measure in our experiments and we observe that some decisions that increase the HP throughput can result in very high material convergence levels. It is important to note that material convergence increasing to very high levels can be an undesired result in practice. Hence, policies that maximize HP item throughput with no regard to material convergence may not be optimal. Using models like the one proposed in this chapter can help decision makers observe such trade-offs and make decisions accordingly.

For future studies, we will also investigate how resource allocation decisions are impacted if the arrival rate of solicited and unsolicited donations vary between periods. We will model different patterns (both increasing, both decreasing and opposite) to observe how the relationship impacts resource allocations.

Chapter 6

Transient Analysis of Relief Center Performance

6.1 Introduction

At the disaster site, the relief items are distributed to victims through relief centers (RCs) with multiple points of distribution (PODs), each staffed with volunteers distributing relief supplies. These RCs can be setup and staffed in various ways and the RC design and staffing can have a big impact on performance, measured by victim waiting times and RC throughput. In Chapter 3, we introduced a queuing model and a solution algorithm that can analyticaly solve for the performance measures for any layout design.

The development of an analytical solution to the queuing network model of an RC is very practically beneficial, since it allows for the fast and computationally efficient evaluation of alternative designs. However, assumptions related to Poisson arrivals and exponential service can be restrictive. Moreover, the analytical

formulas provide the steady state system measures. Therefore, in this chapter we develop simulation models to answer the following research questions: (i) What is the impact of non-stationary arrivals on RC performance?, (ii) What is the impact of relaxing assumptions related to Poisson arrivals and exponential service on RC performance?, (iii) What are insights obtained from transient analysis of RC performance?

To answer these research questions, we build simulation models of each RC design. In specific we use Arena®, a discrete event simulation software (version 15.0) to relax the Markovian assumptions, model non-stationary arrival patterns and analyze the transient behavior of the relief distribution operations.

The rest of the chapter is organized as follows. Section 6.2 reviews literature on simulation models for disaster response operations. Section 6.3 describes the relief center designs and the simulation models representing the different relief center designs. Next, in Section 7.4 we validate the simulation models and investigate the impact of relaxing stationary arrivals and Markovian assumptions. Finally, Section 6.5 concludes this chapter.

6.2 Literature Review

Simulation models are widely used in the disaster response literature to model victim (or agency) behavior, understand complex systems and use simulation-optimization models to optimize performance for a variety of disaster response operations.

Simulation-Optimization Models: Yang et al. (2013) use simulation optimiza-

tion techniques to analyze the resource allocation problem under equity. Zou et al. (2005) propose a simulation-optimization model for evacuation plans for hurricanes in Maryland. Niessner et al. (2017) use a simulation-optimization approach to analyze the operational policy of field hospitals following mass casualty events, while Ahmed and Alkhamis (2009) use simulation- optimization to design a decision support tool for an emergency department. Min et al. (2005) propose an analysis framework to combine simulation, system dynamics, functional models, and nonlinear optimization algorithms to analyze inter-dependencies of national infrastructures. In this study, we do not integrate simulation and optimization models and use discrete event simulation solely.

Agent Based Simulation Models: Aros and Gibbons (2018) use agent based simulation to analyze the effect of different communication lines between agencies on response times. Y.Wang et al. (2012) use agen-based simulation to model coordination between multiple response agencies during mass casulty incidents. Nagarajan et al. (2012) use agent based simulation to model the impact of warning channels on evacuation success. Chen and Zhan (2008) use agent based modeling to evaluate evacuation strategies. Fikar et al. (2017) use agent based simulation optimization to select optimal distribution points for disaster relief. Hawe et al. (2015) use agent-based simulation for resource allocation between two disaster sites. Agent based simulations are mainly used to integrate entity characteristics into the model. In this study we are interested in the system-wide behavior, hence we do not use an agent based simulation model.

Discrete Event Simulation Models: Pitana and Kobayashi (2009) use discrete event simulation to evaluate scenarios related to ship evacuation for tsunami prepa-

ration. Hobeika et al. (1994) propose a simulation based decision support model for developing evacuation plans for nuclear power stations. Chiu et al. (2008) use simulation to evaluate traffic evacuation strategies. Albores and Shaw (2008) use discrete event simulation to understand resource positioning requirements for different incident and locations. de Silva and Eglese (2000) propose a simulation model connected to a geographical information system for enhanced evacuation modeling. Reshetin and Regens (2003) use simulation to model a bio-terrorism attack and possible dispersion scenarios, and Das et al. (2008) present a large-scale simulation model for the stochastic propagation of an influenza pandemic. Jain and McLean (2003) propose a framework for integration of modeling, simulation, and visualization tools for emergency response. This branch of the literature is the most relevant to our study. The main difference between our work and the other discrete event simulation papers in the literature is the focus on relief center design in this work. To our knowledge, there are no prior simulation papers on relief center design.

6.3 System Description

The first layout we consider has been adapted from The Guide to Points of Distribution by FEMA and US Army Corps of Engineers (Federal Emergency Management Agency, 2008) as a representation of current practice for relief center setup (See Figure 6.1). Note that in this layout, the whole crowd accesses the RC from the a single point of entry, all 4 points of distribution (POD) within the RC distribute all four of the available items and the PODs are aligned in a linear fashion.

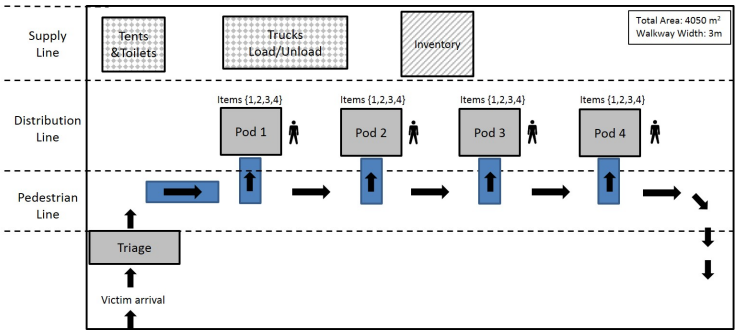

Figure 6.1: RC Layout: Current Practice (Federal Emergency Management Agency, 2008)

The second layout we consider is an alternative design we propose (shown in Figure 6.2) where we aim to dissipate the crowd by creating multiple access points and variability in routings via item assignment. Note that item assignment can: (1) create different routing options to dissipate the crowd, (2) influence arrival rate to a POD based on the assigned items and their associated demand, (3) balance service times at the PODS. In this design, the overall area, the number of distribution PODs and the width of the pedestrian walkways are kept the same as prescribed in Federal Emergency Management Agency (2008).

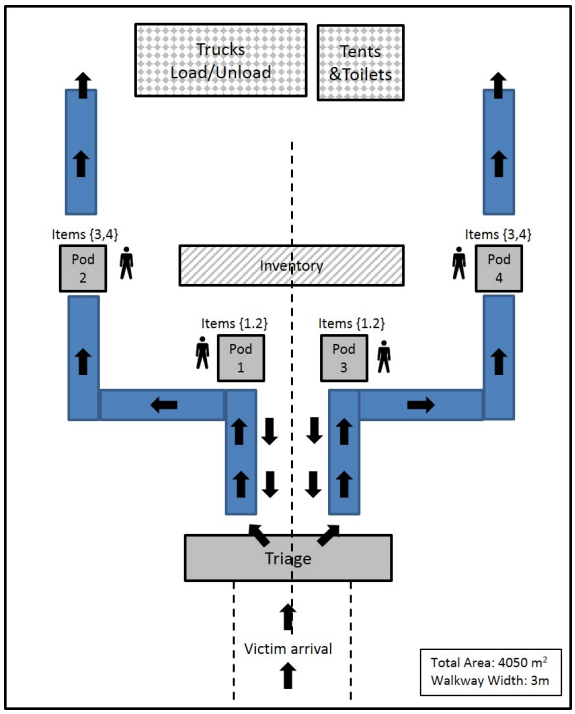

Figure 6.2: Proposed RC Layout

We model each RC design as a queuing network where victims arrive, queue and receive their supplies. Each POD in a relief center is accessed via an entry walkway and exited via an exit walkway. In case of multiple PODs, there exists connecting walkways between the PODs. The walkways (entry, exit, connecting) and the points of distribution within the relief center design are all modeled as finite capacity queues. The service rate of a walkway queue represents the rate at which victims can move within the walkway, while the service rate of a POD represents the rate at which the relief items can be distributed. Note that, the rate at which victims can move within the walkway decrease as congestion increases. This phenomenon is known as the "crowd effect" and has been document via empirical studies by Tregenza (1976). On the other hand, the distribution rate of each relief item is only dependent on its type and is unaffected by the level of crowding. If multiple items are being distributed from a POD, then the time it takes to distribute these

items is the sum of distribution time for each item at the POD. We will refer to this assumption as the "additive service time" assumption.

Figure 6.3 shows the three layouts we analyze in this Chapter. We assume that the four items being distributed from the relief center are blankets (B), tarpaulins (T), kitchen sets (K) and sleeping mats (S). Layout 1 represents the current practice adapted from FEMA, where all items are available at all four PODs and all victims access the RC from the same entry point. Layouts 2 and 3 both represent the alternative layouts, where there are multiple entry points and the crowd can be dissipated based on the items each victim needs. The difference between layout alternative 2 and 3 is the item assignment to PODs.

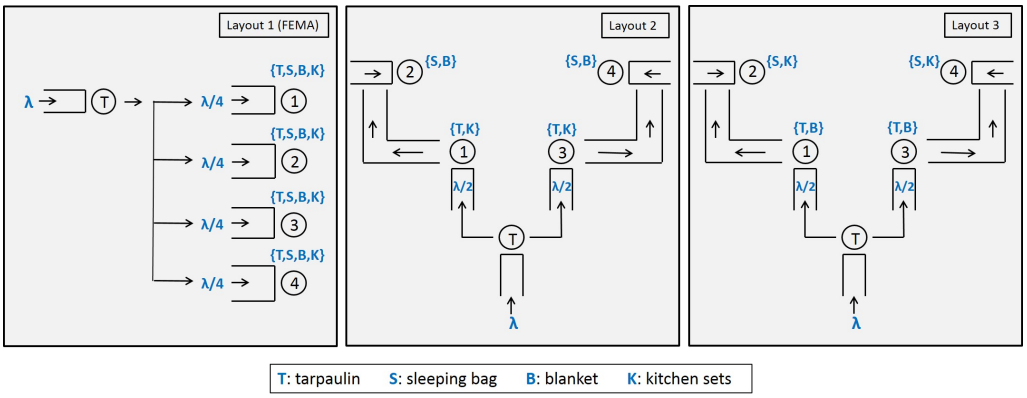

Figure 6.3: Layout Alternatives

Figure 6.4 shows the queuing network representation associated with all three layouts. Note that, in layout 1, a victim who needs all 4 items will need to go through 4 queues and in layouts 2 and 3, a victim who needs all four items will need to go through 7 queues.

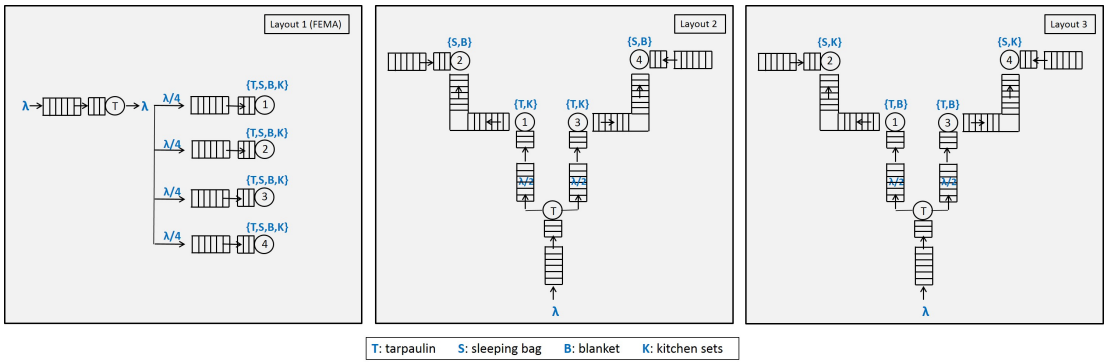

Figure 6.4: Queuing Network Model of the Layout Alternatives

Next, we describe the simulation models for the three layouts previously described. Figure 6.5 depicts the simulation model for RC layout 1, and Figure 6.6 depicts the simulation model for RC layout 2 and 3. For layouts 2 and 3, the logic of the model is similar, however the parameters representing demand and service time experienced at each POD changes in parallel to the item assignments. We next discuss the modeling logic in detail.

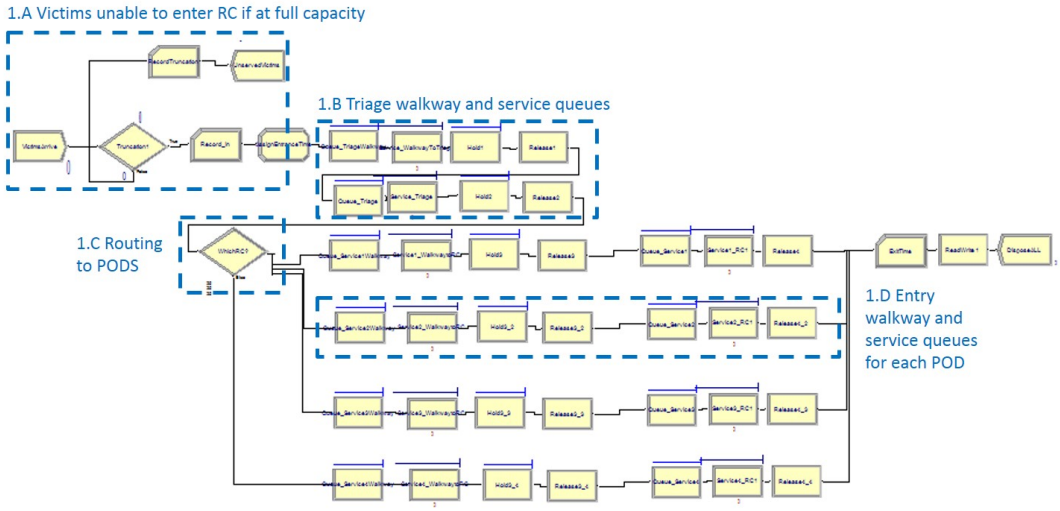

Figure 6.5: Arena Simulation Model of Layout 1

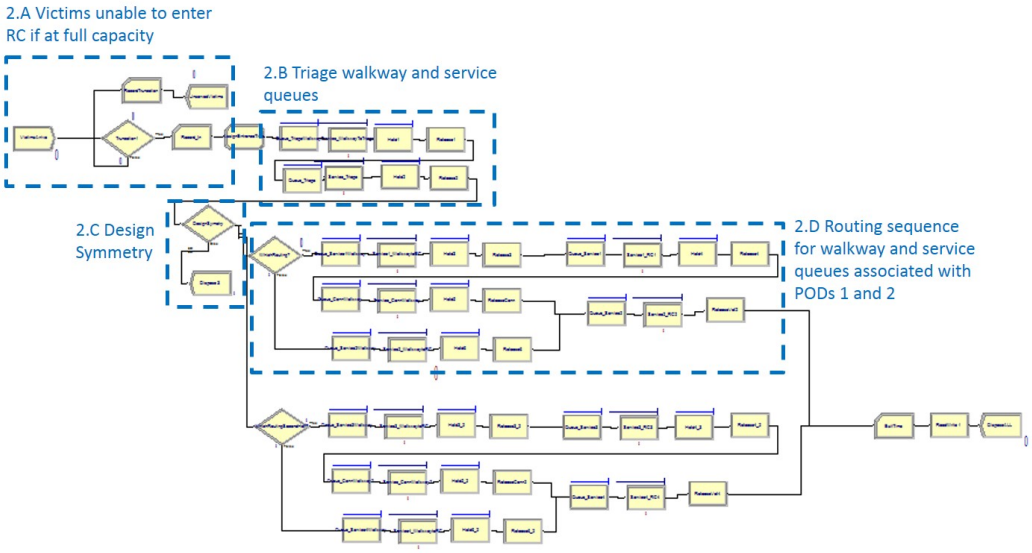

Figure 6.6: Arena Simulation Model of Layouts 2 and 3

6.3.1 Modeling the Queues in the Network

There are two types of queues in the network, walkway queues with state dependent service rates and distribution queues with state independent service rates. Let $i \in N$ represent the queue in the network, let K_i represent the capacity of queue i and let $\mu_i(k_i)$ represent the hourly service rate at queue i.

Distribution queues: The distribution queues model the relief item distribution rate at each POD. Each POD is assumed to be staffed with a single volunteer, where the distribution rate is only dependent on the subset of items being distributed at the given POD. To model the queuing behavior of queue i, we use a Hold-Process-Hold-Release logic in Arena®, where the first hold module assures that no more than one victim is served at a time and the second hold module assures that queue $i+1$ has capacity before the victim is released to the next queue. If the downstream queue is full at the time of service completion, then the victim is blocked in ser-

vice and no other victim can be served until the downstream queue has a service completion (blocking after service). Note that, since all queues have finite capacity, blocking can propagate through the network, starting from the end of the network towards the beginning. This blocking logic is critical in determining the system throughput. If an arrival to the system finds the first queue full, than the arrival is rejected. This is ensured via the decide modules shown as 1.A and 2.A in both Figures 6.5 and 6.6. The process module is modeled as a seize-delay type, where the volunteer resource is seized throughout the service time. The releasing of the volunteer is controlled by a separate release module at the end of the sequence to ensure the service gets blocked if the downstream queue is at capacity.

Walkway queues: The walkway queues are modeled based on the same Hold-Process-Hold-Release logic with one difference being the state dependent service rates to model the effects of congestion. According to empirical studies by Tregenza (1976), the traveling speed $V(k_i)$ at which victims can move on the walkway, exponentially decreases with the number of occupants (k_i) on the walkway, for any i representing a walkway queue. Let L represent the walkway length and let W represent the walkway width. The empirical studies report the following parameters:

- The victim movement comes to a halt when the crowd density on the walkway approaches 5 people per square meter, resulting in an effective capacity of $C = \lfloor 5LW \rfloor$.
- The average walking speed is $(A = 1.5m/s)$.
- The speed for the density of $a = 2$ victims per square meter is $V_a = 0.64m/s$.
- The speed for the density of $b = 4$ victims per square meter is $V_b = 0.25m/s$.

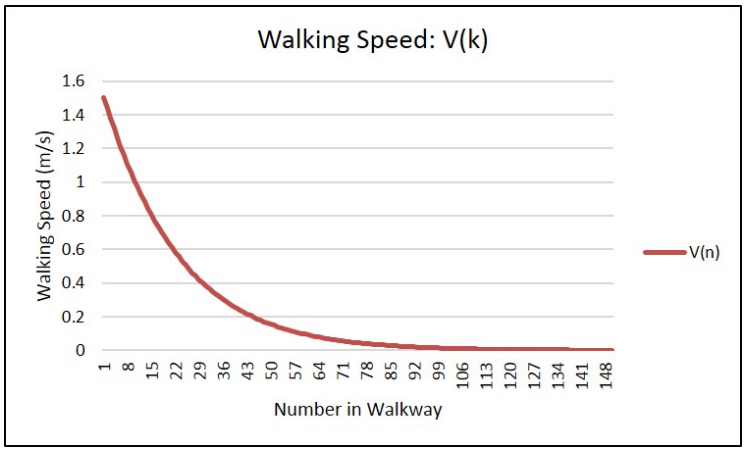

Figure 6.7: Walking Speed Changes by Number of Victims on Walkway

Based on these empirical parameters Cheah and Smith (1994), Smith (1994) characterize the walking speed as given by Equation 6.1, while Figure 6.7 shows the graphical representation of the speed and crowd relationship.

$$\begin{aligned} V(k) &= A \exp\left\{-\left(\frac{k-1}{\beta}\right)^{\gamma}\right\} \\ \beta &= \frac{a-1}{\left[ln\left(\frac{A}{V_a}\right)\right]^{(1/\gamma)}} \\ \gamma &= \frac{ln\left[\frac{ln(V_a/A)}{ln(V_b/A)}\right]}{ln\left(\frac{a-1}{b-1}\right)} \end{aligned} \tag{6.1}$$

This relationship between walking speed and congestion have been modeled via different approaches in the literature. These approaches mainly fall into three categories: (1) Microscopic models, which take into account dynamics between pedestrians, (2) Macroscopic models that focus on the behavior of the crowd as a whole, and (3) Mesoscopic models that combine the properties of both macroscopic and microscopic modeling techniques. See Xiaoping et al. (2009) for a discussion on these categories. Modeling walkway pedestrian flow via queuing falls under the macro-

scopic models category. The queuing literature has also taken varying approaches in modeling pedestrian flow on a walkway. Cruz et al. (2003), Cheah and Smith (1994), Smith (1991) models the pedestrian flow via a state dependent $M/G/C/C$ queues, while Woensel et al. (2005) model it via state dependent $M/M/1$, $M/G/1$, $G/G/1$ queues.

In this chapter we use a state dependent $M/M(k)/1/K$ queue to model the walkways. The state is the number of victims present in the walkway, which determines the current walkway speed $V(k)$ (as given by Equation 6.1), which is then used to calculate the service rate based on the walkway length. In the $M/M(k)/1/K$ model of pedestrian flow, the queue length performance measure is anologous to the number of people on the walkway and the waiting time performance measure is anologous to the pedestrian walking speed. Hence, we would expect to see a similar relationship to the one observed by Tregenza (1976) between these two measures. Figure 6.8 shows the graph of the relationship between the queue length and the waiting time obtained from solving the state dependent $M/M(k)/1/K$ model. Note that steady state equations were used to obtain the graph and both $\bar{W}$ and $\bar{L}$ are outputs for different arrival rates. As seen, the $M/M(k)/1/K$ model represents the crowd density and pedestrian movement relationship accurately.

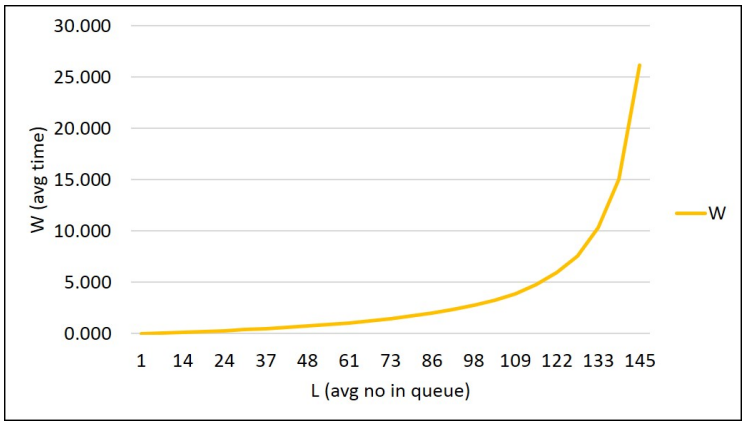

Figure 6.8: Effects of Congestion: $M/M(k)/1/K$ Queuing Model

6.3.2 Modeling Victim Routing

Victim routing is modeled by decide modules in Arena®, using the n-way by chance option. Below we describe the probabilistic routing for each layout design:

In layout 1, all PODs distribute all available items. Hence, victims choose one of the four PODs with equal probability. This is modeled using a 4-way by chance decide module in Arena® with equal probabilities of $p = 0.25$. In layouts 2 and 3, the design is symmetric. Hence, following the triage queue, the victims are divided into two streams using a 2-way by chance decide module in Arena® with equal probabilities of $p = 0.5$. In these designs, the first and second POD in the routing distributes different items. We utilize a second decide module, and route the victims probabilistically based on the demand for each item pair.

6.3.3 Model Outputs

There are two main performance measures we are interested in: (1) the average waiting time of victims to receive supplies at the RC, and (2) the throughput of the RC. For all models, we record the time stamp of entry, when a victim first enters the relief center using a record module. Next, we record the time stamp of the victim exiting the relief center at the end of the network and record the total time spent at the RC for each victim. We then utilize a read/write module to export all data for victim waiting times as a csv file for analysis.

In our models, we allow for arrival rates higher than the capacity. Hence, in certain cases, the RC is not able to serve all arrivals. Therefore, in addition to the throughput, which is defined as the number of victims that completed service at the RC within a given time period τ, we also track the number of victims that arrived and could not get into the RC, as well as the total number of arrivals. To obtain these statistics, we use the statistics module combined with a record module to define the described counters. Lastly, we track the probability of an arriving victim finding the first queue full (or the probability of an arriving victim being denied service), represented as a_1.

6.4 Design of Experiments

In this section, we present numerical experiments to answer our research questions. We first describe the parameter estimation based on Nepal earthquake data and validate the simulation models. In Section 6.4.1 we present the input parameters and validate the simulation models. Then, in Section 6.4.2 we analyze the effect of non-stationary arrival schedules on RC performance for the 3 different layout

configurations. Lastly, in Section 6.4.3 we relax the Markovian assumptions and experiment using different probability distributions for both the inter-arrival and service times.

6.4.1 Inputs and Validation

The arrival rate to the relief center and service rate at the PODs are the input parameters we estimate using Nepal earthquake relief distribution data. This data was obtained from the Humanitarian Data Exchange Database (an open platform for humanitarian data sharing: https://data.humdata.org/), which was prepared by contributions from the International Federation of Red Cross (IFRC), the United Nations High Commissioner for Refugees (UNHCR), Nepal United Nations Office for the Coordination of Humanitarian Affairs (OCHA), the Nepal Ministry of Home Affairs, and the Nepali Police. In Table 6.1 we summarize the arrival and service rate parameters estimated using the data obtained. The details of data analysis are given in Chapter 3.

Table 6.1: Parameters Used in Simulations

Parameter	Value
Arrival Rate	$\lambda = 58/\text{hour}$
Service Rate by Item	$\mu_T = 10.7/hr$, $\mu_B = 31.6/hr$, $\mu_K = 10.2/hr$, $\mu_S = 20.2/hr$
Service Rate for Walkways	Exp with parameter $= (216 * \exp(-0.024 * k))$
Queue Capacity	$K_i = 50, \quad \forall i$

Note that the service rate function for walkways in Table 6.1 is based on the formulas given in Section 6.3.1. For the service rate at PODs, we use the distribution

rate per item as given in Table 3.3 and assume the rates are additive. Additivity is a reasonable assumption (if not necessary) for items like tarpaulins and kitchen sets where handling takes considerable time. Note that for PODs at which not all arriving victims need all items, we use a weighted average to calculate the service rate.

Also note that, the overall arrival to the RC is the same for all designs. However, the arrival rate to each queue may differ depending on the demand for items being distributed at a particular POD. To calculate arrival rate per POD we use the data provided in Table 3.2 coupled with the assumption that the average household size is 5 victims per household. Based on this data, we assume that one fifth of all arriving victims need tarpaulins and kitchen sets, two fifths need blankets and all arriving victims need sleeping mats.

Table 6.2: Items Required per Household

Item	Conversion Parameter
Tarpaulin	1/Household
Blanket	2/Household
Kitchen Sets	1/Household
Sleeping Mat	5/Household

Number of Replications: The minimum number of replications needed is based on the desired confidence interval range and level (Kelton et al. (2004)). The formula used to calculate the replication number is given by Equation 6.2, where n represents the number of replications, s represents the sample standard deviation, and h represents the desired confidence interval range. In our numerical experiments we use a 95% confidence interval level and set $h = 2$ hours for the average waiting time

measure. We first take a sample run and calculate the sample standard deviation. Then we use the formula to compute the minimum number of replications.

$$n = t^2_{n-1,1-\alpha/2}\frac{s^2}{h^2} \tag{6.2}$$

The Warm-up Period: For transient analysis, we do not use a warm-up period and the terminating condition used is the length of the distribution period. For validation, we do a steady state analysis, for which we need to choose an appropriate warm-up length such that the system will reach its steady state. The time the system reaches steady state can vary based on the design and parameters used. Hence, for each experiment, we take a long sample run (200 days) and plot both the throughout and average waiting time measures over time to decide the required warm-up period for each experiment.

We validate the simulation model of all three layouts by comparing the simulation results to the analytical results obtained under assumptions of exponential service rates, Poisson arrivals and steady state estimates. To obtain steady state performance measures, the simulations were run for a 100 replications, each replication for 20 days with a 10 day warm-up period. From the results, we conclude that the simulation models are a valid representation of the RC layouts. Appendix 1 in section 6.6 provides a detailed discussion of numerical studies related to model validation.

6.4.2 Impact of Non-Stationary Arrivals

In this section we investigate the impact of non-stationary arrivals on the performance of all three layouts. The analytical model assumed Poisson arrivals with a

steady arrival rate of 58 victims per hour. However, during the disaster aftermath arrival rates of victims can vary daily for the following reasons:

- **The time elapsed after the disaster:** Immediately after a disaster, the majority of the efforts are usually focused on search and rescue operations and relief distribution can take lower priority in the first few days of disaster response.
- **Demand peaks:** Victim arrival rates can vary and peak at certain times. This can have many reasons including disaster aftershocks, changes in environment and number of RCs currently open.

In addition to modeling non-stationary arrivals, we also analyze the transient behavior of relief distribution at the RCs. We specifically focus on the first 10 days after the disaster, since this is the critical period for relief distribution efforts. We use a terminating simulation model, where the model terminates after 10, 16 hour distribution days with no warm up period. We analyze three different non-stationary schedules as listed in Table 6.3. We keep the total victim arrival over the 10 days equal for all schedules, but vary the daily rates. To model these schedules, we use the schedule module in Arena®.

Table 6.3: Arrival Schedules

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
Schedule 1 (per hr)	29	29	58	58	87	116	87	58	29	29
Schedule 2 (per hr)	14.5	29	145	145	101.5	58	29	29	14.5	14.5
Schedule 3 (per hr)	29	29	116	29	29	29	174	87	29	29

In schedule 1, we model an arrival pattern where fewer arrivals happen the first 2 days while the search and rescue operations are going on. Then for the next 6 days, the arrival rates are much higher as the victims are looking to collect the supplies

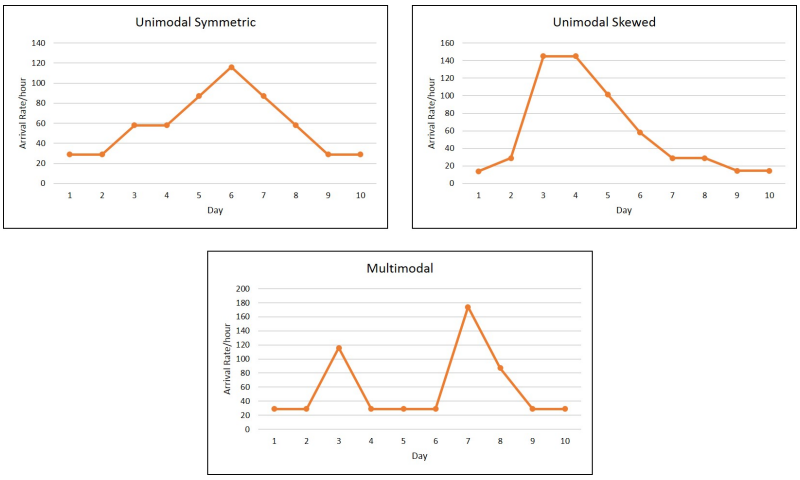

Figure 6.9: Arrival Schedules

they need. Then, towards the end of the period, the arrival rates are lower again. In schedule 2, we model a similar pattern as schedule 1, however the peak rates experienced in schedule 2 are considerably higher. In schedule 3, we model a multi-peak arrival pattern, where unforeseen events such as disaster aftershocks or changes in the disaster aftermath conditions cause peaks in victim arrivals. Schedule 3 also assumes lower arrival rates for the beginning and end of the 10 day period. Figure 6.9 compares the three arrival schedules.

Figure 6.10 compares the daily average throughput and average waiting time measures for all three layouts under all arrival schedules. Tables 6.9, 6.10 and 6.11 in the Appendix 3 in section 6.8 provide the detailed results.

For layout 1 we make the following observations:

- The maximum throughput for layout 1 (due to the bottleneck rate) is an average of 27 per hour. As long as the minimum daily arrival rate for a

schedule is greater than or equal to 27, there will be no throughput loss. In addition, since 27 per hour represents the maximum capacity, arrival rates much higher will not result in a throughput increase. As a result, schedules 1 and 3 do not impact the throughput measure, while schedule 2 results in a loss in throughput.

- The average waiting time is lower for the first few days of both for the stationary and non-stationary arrival schedules due to the system being empty at the beginning of the simulation.

- The average waiting time is lower for the non-stationary schedules due to the lower arrival rate days having lower waiting times due to lower congestion. Note that the higher arrival rate days, cannot create higher congestion due to the finite capacity of the network. However, this comes at the expense of loss in throughput, resulting in 59, 340 and 25 victims not being able to get service from this RC layout under schedules 1, 2 and 3 respectively over the course of the 10 days.

For layout 2 we make the following observations:

- The maximum throughput for layout 2 (due to the bottleneck rate) is an average of 47 per hour. Note that, this is a much higher maximum throughput than layout 1. Since all 3 of the non-stationary schedules have days with arrival rates smaller than 47 per hour, under all three schedules, the average throughput from layout 2 decreases. Among all arrival schedules, schedule 2 results in the highest throughput loss.

- The average waiting time is lower for the first few days for both for the stationary and non-stationary arrival schedules due to the system being empty

at the beginning of the simulation.

- The average waiting time is lower for all the non-stationary schedules due to the lower arrival rate days having lower waiting times due to lower congestion. Note that the higher arrival rate days, can not create higher congestion due to the finite capacity of the network. However, this comes at the expense of loss in throughput, resulting in 875, 2098 and 1343 victims not being able to get service from this RC layout under schedules 1, 2 and 3 respectively over the course of the 10 days.

For layout 3 we make the following observations:

- The maximum throughput for layout 3 (due to the bottleneck rate) is an average of 33.6 per hour. Note that, this is higher than layout alternative 1 and a lower than layout alternative 2. Since all 3 of the non-stationary schedules have days with arrival rates smaller than 33.6 per hour, under all three schedules, the average throughput from layout alternative 3 decreases. Among all arrival schedules, schedule 2 results in the highest throughput loss for layout 3.

- The average waiting time is lower for the first few days of both for the stationary and non-stationary arrival schedules due to the system being empty at the beginning of the simulation.

- The average waiting time is lower for all the non-stationary schedules due to the lower arrival rate days having lower waiting times due to lower congestion. Note that the higher arrival rate days, can not create higher congestion due to the finite capacity of the network. However, this comes at the expense of loss in throughput, resulting in 162, 668 and 134 victims not being able to get

service from this RC layout under schedules 1, 2 and 3 respectively over the course of the 10 days.

We conclude that, non-stationary arrivals do impact the average throughput and average waiting time measures. However, under a given arrival schedule, dispersing the crowd and leveraging item assignment to PODs still improves both performance measures.

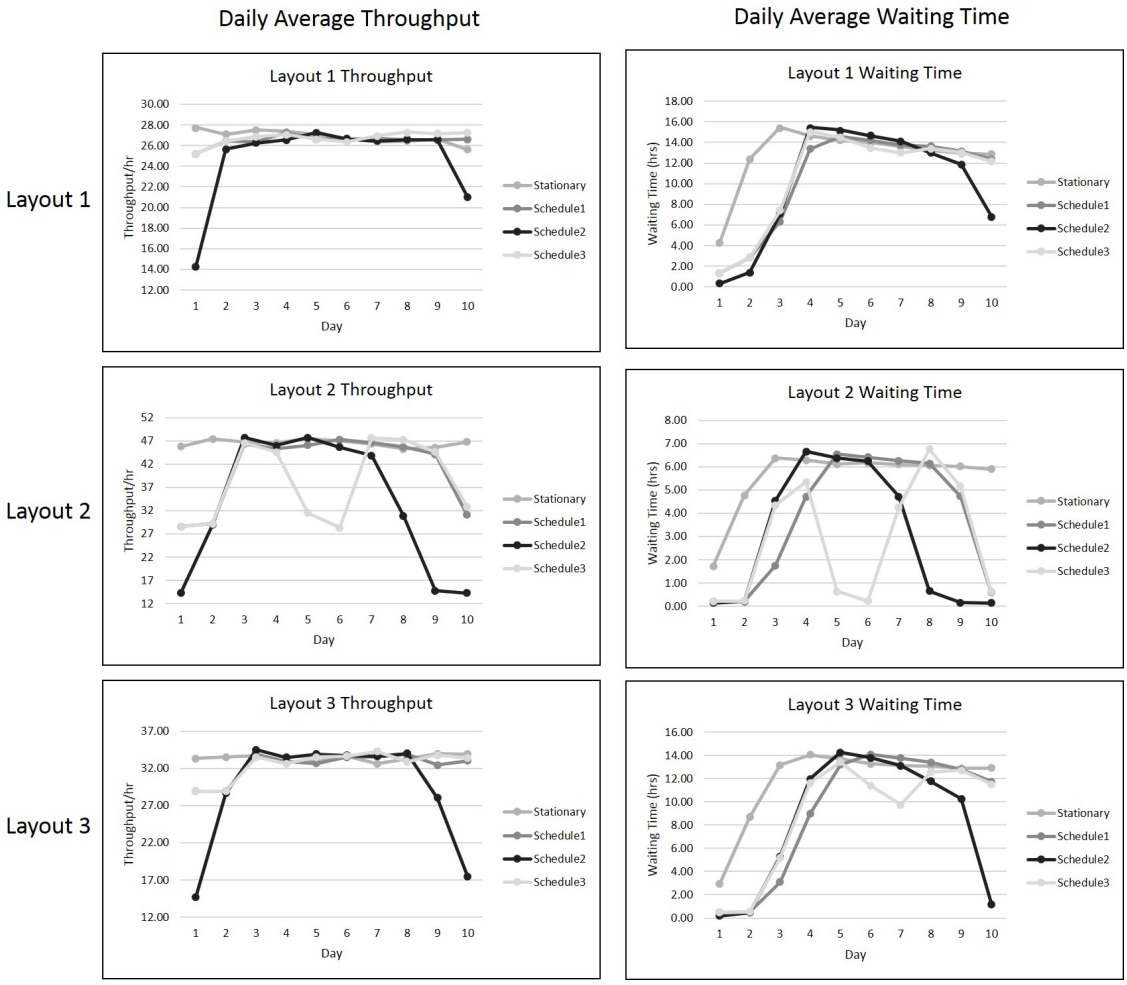

Figure 6.10: Performance Comparison: Non-stationary Arrivals

6.4.3 Impact of Markovian Assumptions

In this section, we investigate the impact of the Markovian assumptions (exponentially distributed service and inter-arrival times) on the results and insights obtained from the analytical modeling of relief center designs. We relax the Markovian assumption for both the service time distribution at the PODs and the inter-arrival time distribution of the victims.

We use Gamma, Weibull, Lognormal and Hyperexponential distributions to model the service and inter-arrival times. For all cases, we equate the first moments (the mean) of the distributions. Then we choose the second moments (variance) to consider low-medium-high variability cases. Hopp and Spearman (2008) defines low variability distribution as having a coefficient of variation (CV) less than or equal to 3/4, medium variability as having a CV between 3/4 and 4/3 and high variability as having a CV greater than or equal to 4/3. As it follows, the Markovian model represents a medium variability case. To model low variability, we use Gamma and Weibull distributions with $CV = 1/3$ and to model high variability we use Log-normal and Hyper-exponential distributions with $CV = 4/3$. In section 6.7 (Appendix 2) , we provide the formulas for mean and variance for all four distributions, which we used to obtain the parameters to define the probability distributions. Tables 6.4 and 6.5 below summarize these parameters.

Table 6.4: Gamma and Weibull Distribution Parameters ($CV = 1/3$)

	Gamma Distribution		Weibull Distribution	
	Shape Parameter	Scale Parameter	Shape Parameter	Scale Parameter
Inter-arrival Time	$\alpha = 9$	$\beta = 0.001915$	$k = 3.3035$	$\lambda = 0.0192$
Service at Triage	$\alpha = 9$	$\beta = 0.000926$	$k = 3.3035$	$\lambda = 0.0093$
Service: Layout 1 PODs	$\alpha = 9$	$\beta = 0.016388$	$k = 3.3035$	$\lambda = 0.1644$
Service: Layout 2 PODs 1&3	$\alpha = 9$	$\beta = 0.021164$	$k = 3.3035$	$\lambda = 0.2123$
Service: Layout 2 PODs 2&4	$\alpha = 9$	$\beta = 0.004744$	$k = 3.3035$	$\lambda = 0.0476$
Service: Layout 3 PODs 1&3	$\alpha = 9$	$\beta = 0.010915$	$k = 3.3035$	$\lambda = 0.1095$
Service: Layout 3 PODs 2&4	$\alpha = 9$	$\beta = 0.006606$	$k = 3.3035$	$\lambda = 0.0663$

Table 6.5: LogNormal and HyperExponential Distribution Parameters ($CV = 4/3$)

	LogNormal Distribution		Hyper Exponential Distribution	
	Mean	Variance	(p_1, p_2)	(λ_1, λ_2)
Inter-arrival Time	$\mu = 0.017241$	$\sigma = 0.022988$	(0.5,0.5)	(154.0953,35.7229)
Service at Triage	$\mu = 0.008333$	$\sigma = 0.01111$	(0.5,0.5)	(318.8179,73.9094)
Service: Layout 1 PODs	$\mu = 0.147492$	$\sigma = 0.196656$	(0.5,0.5)	(18.0132,4.1759)
Service: Layout 2 PODs 1&3	$\mu = 0.190476$	$\sigma = 0.253968$	(0.5,0.5)	(13.9483,3.2335)
Service: Layout 2 PODs 2&4	$\mu = 0.042698$	$\sigma = 0.056931$	(0.5,0.5)	(62.2226,14.4247)
Service: Layout 3 PODs 1&3	$\mu = 0.098231$	$\sigma = 0.130975$	(0.5,0.5)	(27.0464,6.2700)
Service: Layout 3 PODs 2&4	$\mu = 0.059453$	$\sigma = 0.079271$	(0.5,0.5)	(44.6876,10.3596)

We run the Arena® models for each layout alternative, for each distribution specified in Tables 6.4 and 6.5. The experiments are run for a 100 replications, each replication is run as a transient model terminating at 10 days, each day composed of 16 hours. Figure 6.11 compares the average throughput and waiting time performance measures from the simulation experiments, for Exponential, Gamma, Weibull, Lognormal and Hyper-exponential distributions used for inter-arrival and service times for layouts 1, 2 and 3 respectively. The details of the numerical results can also be found in Tables 6.12, 6.13 and 6.14 in the Appendix. From these results we conclude that both the performance measures themselves, as well as their comparisons are robust to the chosen probability distribution and the level of variability.

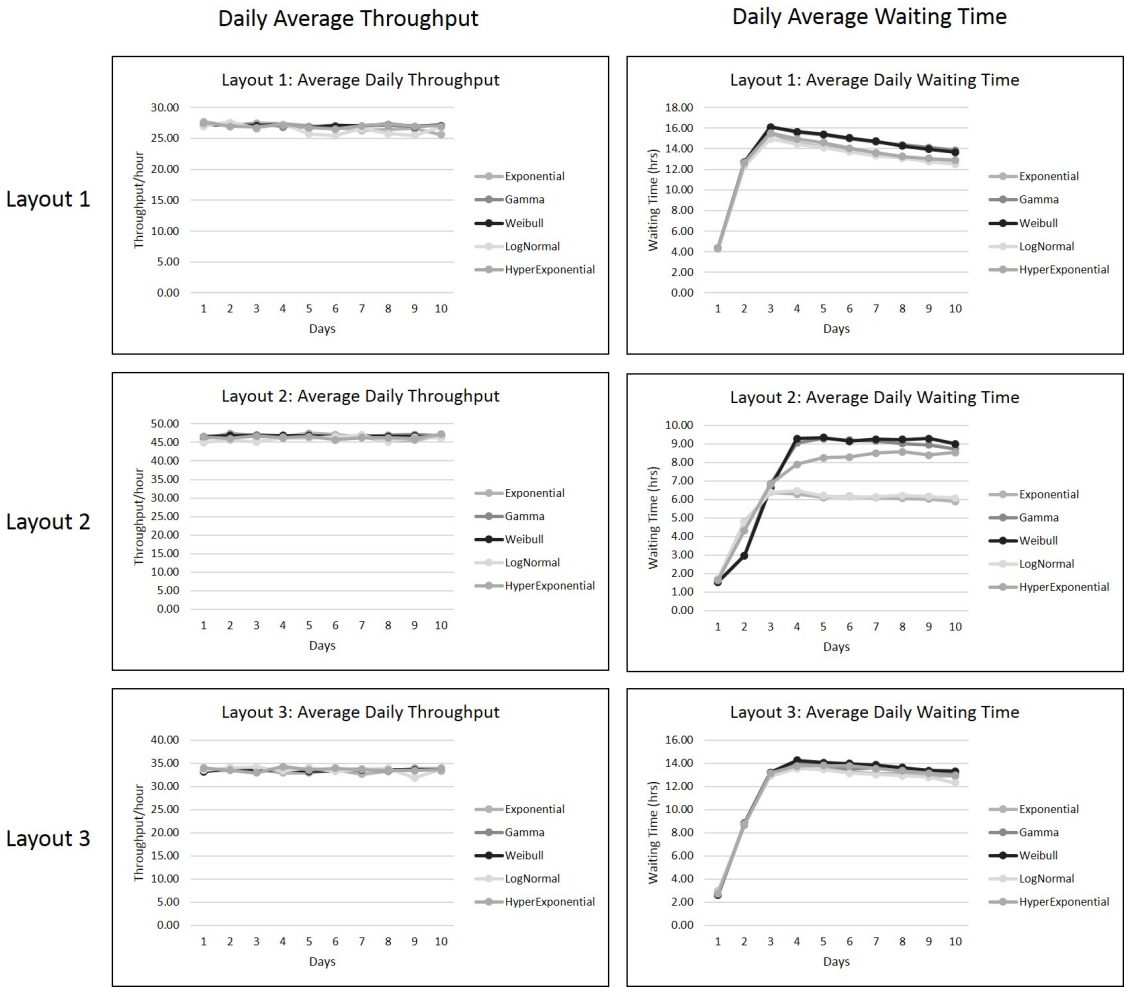

Figure 6.11: Performance Comparison: Inter-arrival and Service Time Distributions

6.5 Conclusions

In this chapter we analyze the relief center designs introduced in Chapter 3 and investigate the impact of the Markovian assumptions as well as the steady state analysis. We model the three relief center designs using discrete event simulation model built in Arena®.

We first relax the stationary arrivals assumption and model three different sched-

ules of arrivals as non-stationary Poisson arrivals to the relief centers. We conclude that varying arrival rates over time impact the throughput and waiting time measures of the relief centers. However, crowd dissipation strategies and item assignment to RCs remain to be impactful strategies in improving RC performance.

Next, we relax the Markovian assumption for both the inter-arrival time and the service time distributions. We use Gamma, Weibull, Log-Normal and Hyper-Exponential distributions with equal mean under high and low variance compared to the exponential distribution. We conclude that both the average throughput and the average waiting time performance measures are very robust to the chosen probability distribution.

6.6 Appendix 1: Numerical Details of Validation

This appendix discusses numerical details of the validation of the simulation model. We validate the simulation model of all three layout alternatives by comparing the performance measures obtained from simulation models with the analytical model results.

It is important to note that the analytical model assumes utilization of all queues in the network to be below 1. However, the input parameters we use for the numerical studies have an arrival rate higher than the network capacity. To handle such cases, the analytical model's outer loop runs on a grid search and picks the largest a_1 (the probability that the first queue is full) such that $\rho < 1$ for all queues in the network. Following this, the analytical model estimates the performance measures based on an arrival rate of $\lambda(1 - a_1)$ to the system, where λ represents the external arrival rate. To model this behavior in the simulation models, we use the a_1 probability as an input to the simulation model and truncate λa_1 of all arrivals to the system. Table 6.6 compares the performance measure results from these two models and as seen from the results, the performance of both systems is equivalent.

Table 6.6: Comparison of Simulation and Analytical Model

Layout Design	Analytical Model			Simulation Results		
	$\bar{W}$	TH	a_1	$\bar{W}$	TH	a_1
Layout 1	6.39 hrs	27.14/hr	0.53	6.32 hrs	27/hr	0.53
Layout 2	3.59 hrs	46.86/hr	0.19	3.87 hrs	46/hr	0.19
Layout 3	2.79 hrs	33.64/hr	0.42	3.24 hrs	33/hr	0.42

We next update the simulation model such that we do not dictate the a_1 pa-

rameter. Instead, we let all arrivals into the system and only if the first queue is full ($k_1 = K_1$), the system will not be able to accept the arriving victim. We provide the results of this simulation model in Table 6.7 based on 100 replications of 20 days each with a 10 day warm-up period. Note that the a_1 parameters are close, however slightly lower than the previous case, putting the system utilization close to 1 ($\rho \cong 1$). This results in a much higher waiting time estimate due to the exponential relationship between utilization and waiting time. The analytical model shows numerical issues $\rho \cong 1$ and shows convergence problems. Hence, the analytical model can only provide accurate results when $\rho < 1$, which is ensured by the outer loop that searches for a_1.

Table 6.7: Simulation Model Results: Free Truncation

Layout Design	Simulation Results		
	$\bar{W}$	TH	a_1
Layout 1	12.3 hrs	26/hr	0.49
Layout 2	5.9 hrs	46/hr	0.16
Layout 3	12.09 hrs	33/hr	0.38

6.7 Appendix 2: Details on the Probability Distributions

Table 6.8: Summary of Probability Distributions

Distribution	Parameters	Probability Distribution Function	Range	Mean	Variance
Exponential	β	$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$	$[0, +\infty)$	β	β^2
Gamma	scale:β, shape:α	$f(x) = \begin{cases} \frac{\beta^{-\alpha} x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$	$[0, +\infty)$	$\alpha\beta$	$\alpha\beta^2$
Log-Normal	mean:μ_l, std=σ_l	$f(x) = \begin{cases} \frac{1}{\sigma x \sqrt{2\pi}} e^{-(ln(x)-\mu)^2/(2\sigma)^2} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$	$[0, +\infty)$	$\mu_l = e^{\mu+\sigma^2/2}$	$\sigma_l^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$
Hyper-exponential	Y_i $exp(\lambda_i)$, p_i	$f_X(x) = \sum_{i=1}^{n} f_{Y_i}(x) p_i$	$f_X(x) = \sum_{i=1}^{n} f_{Y_i}(x) p_i$	$\sum_{i=1}^{n} \frac{p_i}{\lambda_i}$	$\left[\sum_{i=1}^{n} \frac{p_i}{\lambda_i}\right]^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_i \left(\frac{1}{\lambda_i} - \frac{1}{\lambda_j}\right)^2$
Weibull	shape: α, scale:β	$f(x) = \begin{cases} \alpha\beta^{-\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$	$[0, +\infty)$	$\beta/\alpha\Gamma(1/\alpha)$	$\frac{\beta^2}{\alpha}\left[2\Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha}\left[\Gamma\left(\frac{1}{\alpha}\right)\right]^2\right]$

6.8 Appendix 3: Detailed Numerical Results for Non-Stationary Arrivals

Table 6.9: Non-Stationary Poisson Arrivals: Layout 1

	Stationary			Schedule 1			Schedule 2			Schedule 3		
Day	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$
1	58	27.77	4.28	29	25.18	1.33	14	14.26	0.33	29	25.18	1.33
2	58	27.09	12.38	29	26.44	2.83	29	25.65	1.38	29	26.44	2.83
3	58	27.55	15.45	58	26.38	6.33	145	26.25	7.06	116	26.85	7.39
4	58	27.36	14.60	58	27.29	13.37	145	26.53	15.44	29	27.04	15.07
5	58	27.06	14.19	87	26.59	14.56	101.5	27.28	15.19	29	26.61	14.46
6	58	26.71	13.97	116	26.59	14.19	58	26.65	14.63	29	26.35	13.47
7	58	26.37	13.60	87	26.69	13.74	29	26.46	14.11	174	26.94	12.99
8	58	26.43	13.21	58	26.58	13.62	29	26.56	12.99	87	27.32	13.41
9	58	26.61	12.96	29	26.54	13.13	14.5	26.58	11.87	29	27.14	13.01
10	58	25.62	12.85	29	26.59	12.38	14.5	20.99	6.77	29	27.24	12.14

Table 6.10: Non-Stationary Poisson Arrivals: Layout 2

	Stationary			Schedule 1			Schedule 2			Schedule 3		
Day	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$	λ	$T\bar{H}$	$\bar{W}$
1	58.00	45.81	1.72	29.00	28.59	0.21	14.00	14.26	0.13	29.00	28.59	0.21
2	58.00	47.44	4.78	29.00	29.18	0.22	29.00	28.96	0.21	29.00	29.18	0.22
3	58.00	46.77	6.38	58.00	46.37	1.74	145.00	47.75	4.55	116.00	46.56	4.35
4	58.00	46.59	6.28	58.00	45.35	4.71	145.00	46.02	6.66	29.00	44.58	5.35
5	58.00	47.42	6.11	87.00	46.08	6.55	101.50	47.70	6.38	29.00	31.52	0.64
6	58.00	47.11	6.17	116.00	47.27	6.42	58.00	45.63	6.25	29.00	28.32	0.22
7	58.00	46.37	6.10	87.00	46.64	6.26	29.00	43.86	4.72	174.00	47.68	4.24
8	58.00	45.27	6.06	58.00	45.73	6.15	29.00	30.88	0.65	87.00	47.27	6.77
9	58.00	45.61	6.02	29.00	44.18	4.74	14.50	14.74	0.15	29.00	44.64	5.16
10	58.00	46.81	5.90	29.00	31.10	0.61	14.50	14.24	0.13	29.00	32.90	0.64

Table 6.11: Non-Stationary Poisson Arrivals: Layout 3

	Stationary			Schedule 1			Schedule 2			Schedule 3		
Day	λ	$\bar{TH}$	$\bar{W}$	λ	$\bar{TH}$	$\bar{W}$	λ	$\bar{TH}$	$\bar{W}$	λ	$\bar{TH}$	$\bar{W}$
1	58.00	33.36	2.96	29.00	28.91	0.48	14.00	14.68	0.18	29.00	28.91	0.48
2	58.00	33.51	8.68	29.00	28.91	0.53	29.00	28.68	0.47	29.00	28.91	0.53
3	58.00	33.74	13.19	58.00	33.92	3.09	145.00	34.49	5.29	116.00	33.51	5.17
4	58.00	32.96	14.06	58.00	32.92	8.97	145.00	33.48	11.97	29.00	32.62	11.62
5	58.00	32.93	13.71	87.00	32.67	13.22	101.50	33.91	14.26	29.00	33.54	13.42
6	58.00	33.61	13.29	116.00	33.53	14.08	58.00	33.77	13.81	29.00	33.66	11.41
7	58.00	32.62	13.12	87.00	33.58	13.77	29.00	33.67	13.11	174.00	34.29	9.75
8	58.00	33.31	13.08	58.00	33.87	13.40	29.00	34.03	11.79	87.00	32.88	12.58
9	58.00	33.96	12.90	29.00	32.48	12.80	14.50	28.02	10.26	29.00	33.80	12.72
10	58.00	33.93	12.93	29.00	33.03	11.75	14.50	17.41	1.16	29.00	33.41	11.51

6.9 Appendix 4: Detailed Numerical Results for Performance under Different Distributions

Table 6.12: Layout 1: Comparison of Performance under Different Distributions

Day	Exponential		Gamma		Weibull		LogNormal		Hyper-Exp	
	$\bar{W}$	$\bar{TH}$	$\bar{W}$	$\bar{TH}$	$\bar{W}$	$\bar{TH}$	$\bar{W}$	$\bar{TH}$	$\bar{W}$	$\bar{TH}$
1	4.28	27.77	4.34	27.18	4.33	27.21	4.28	27.00	4.33	27.57
2	12.38	27.09	12.69	27.15	12.70	27.19	12.31	27.71	12.63	26.98
3	15.45	27.55	16.13	27.28	16.13	27.11	14.95	26.60	15.52	26.83
4	14.60	27.36	15.60	26.95	15.68	27.28	14.47	27.38	14.95	27.28
5	14.19	27.06	15.36	26.87	15.39	26.76	14.10	25.70	14.58	26.81
6	13.97	26.71	14.97	27.15	15.04	27.03	13.68	25.45	14.06	26.56
7	13.60	26.37	14.65	27.01	14.71	27.07	13.29	26.60	13.59	27.10
8	13.21	26.43	14.36	27.35	14.27	27.23	13.07	25.78	13.24	27.29
9	12.96	26.61	14.11	26.97	13.96	26.95	12.71	25.49	13.06	26.98
10	12.85	25.62	13.84	27.14	13.67	27.15	12.52	26.89	12.88	27.03

Table 6.13: Layout 2: Comparison of Performance under Different Distributions

Day	Exponential		Gamma		Weibull		LogNormal		Hyper-Exp	
	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$
1	1.72	45.81	1.54	46.36	1.55	46.55	1.72	44.94	1.65	46.45
2	4.78	47.44	2.98	46.91	2.98	46.83	4.74	45.59	4.32	46.03
3	6.38	46.77	6.84	46.54	6.62	46.92	6.38	45.11	6.83	46.73
4	6.28	46.59	9.05	46.51	9.29	46.83	6.46	46.27	7.90	46.23
5	6.11	47.42	9.29	46.76	9.34	46.84	6.21	45.99	8.25	46.59
6	6.17	47.11	9.20	46.63	9.14	46.64	6.13	46.63	8.28	45.63
7	6.10	46.37	9.16	46.35	9.25	46.72	6.15	47.03	8.51	46.33
8	6.06	45.27	9.03	46.89	9.24	46.59	6.22	45.11	8.57	46.19
9	6.02	45.61	8.95	47.11	9.29	46.64	6.16	46.31	8.39	46.00
10	5.90	46.81	8.74	46.80	9.00	46.36	6.08	46.21	8.54	47.20

Table 6.14: Layout 3: Comparison of Performance under Different Distributions

Day	Exponential		Gamma		Weibull		LogNormal		Hyper-Exp	
	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$	$\bar{W}$	$T\bar{H}$
1	2.96	33.36	2.66	33.15	2.64	33.23	2.88	33.51	2.75	33.99
2	8.68	33.51	8.83	33.89	8.80	33.83	8.74	34.06	8.68	33.51
3	13.19	33.74	13.19	33.73	13.19	33.31	12.89	34.12	13.14	32.90
4	14.06	32.96	14.05	33.73	14.26	33.58	13.55	33.38	13.78	34.26
5	13.71	32.93	13.86	33.56	14.05	33.27	13.45	34.04	13.79	33.58
6	13.29	33.61	13.51	33.66	13.97	33.36	13.15	33.23	13.76	33.90
7	13.12	32.62	13.62	33.71	13.87	33.60	13.03	33.81	13.54	33.67
8	13.08	33.31	13.40	33.34	13.64	33.63	12.90	33.93	13.33	33.49
9	12.90	33.96	13.19	33.71	13.36	33.57	12.81	31.87	13.15	33.40
10	12.93	33.93	13.12	33.42	13.31	33.59	12.32	33.56	12.95	33.68

Chapter 7

Simulation Study: Optimal Relief Center Layout

7.1 Introduction

Relief centers (RCs) are temporary structures setup following a disaster, where victims arrive, queue and collect critical relief supplies. RCs can have varying sizes, number of volunteers and number of points of distribution (PODs). In Chapters 3 and 6, we analyzed the RC layout used in the field by FEMA and proposed alternative layouts to improve the RC performance and obtain insights. In this chapter, we aim to identify the optimal layout design through a simulation study.

We model the RC layouts as a queuing network, where both the walkways used to access and move between PODs, and the PODs themselves are modeled as finite capacity queues. The analyses and results from Chapters 3 and 6 shed light into the possible ways to improve the performance of a relief center without additional resources (space or volunteers). In this chapter we use simulation studies to answer

the following research questions: (1) What is the optimal RC layout design for a given set of resource constraints?, (2) What is the performance difference between the optimal layout and the rest of the alternatives?

The rest of the chapter is organized as follows. Section 7.2 provides a literature review. Section 7.3 discusses relief center designs and enumerates all alternatives in the design space. Section 7.4 leverages results from the literature to eliminate dominated configurations and presents the numerical results obtained from simulation studies. Finally, Section 7.6 concludes the chapter.

7.2 Literature Review

Recall that in Chapter 3, the relief center was modeled as a tandem queuing network with finite buffers and state dependent service rates(See Figure 7.1). In this section we review the queuing literature on tandem lines. Dallery and Gershwin (1992) provide a detailed summary of models and results for tandem lines. Buzacott and Shanthikumar (1992) review queuing models used in manufacturing systems design and connect the theoretical findings to design problems. Papadopoulos and Heavey (1996) provide a survey of the literature on tandem queuing networks and categorize the literature based on their modeling assumptions. Li et al. (2009) survey analytical models in throughput analysis of production systems including tandem lines.

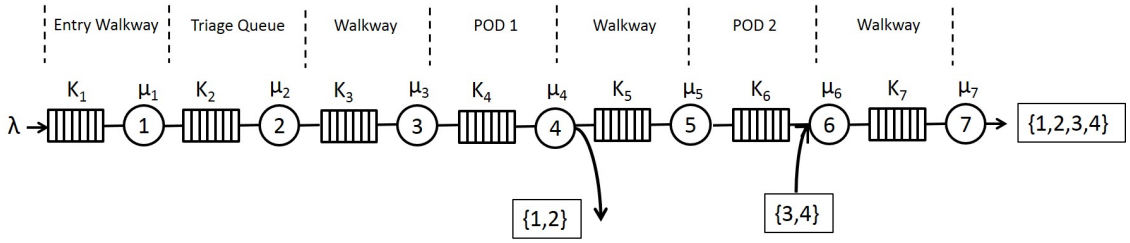

Figure 7.1: Tandem Queuing Network Model of an Relief Center

The following papers in the literature provide important results that can be used in the optimal design of tandem lines. Hillier and Boling (1966) introduce what is known as the "bowl phenomenon", which is a useful way of improving tandem line throughput by unbalancing the line such that the fastest station is placed in the middle. Hillier and Boling (1979) extend the initial result and conclude that for longer lines the improvement obtained in throughput from unbalancing the line is greater and increasing buffers decrease the average imbalance needed for the optimal work allocation. Iyama and Ito (1987) analyze the design of a tandem line with multiple servers of unequal operation rates to maximize system throughput. They analyze an exact Markov model and an approximation and they conclude that multi-server systems preserve the bowl phenomenon. Other papers to date extend this phenomenon to different settings including to longer tandem lines, unbalanced lines, effect of buffer capacities and effects of variance and McNamara et al. (2016) review the related literature to date. However, none of the results in this branch of the literature are directly applicable to the queuing network model of a relief center, due to the state dependent queues in the network.

Other papers in the literature focus on developing analytical solution methodologies (exact or approximate) for the optimal allocation of work, buffers and servers to improve performance measures of a tandem queuing system. Tcha et al. (1992) an-

alyze a server allocation problem for a tandem line with finite buffers and blocking. They propose an algorithm for server assignment based on increasing the throughput upper bound. Hillier and So (1996) analyze a multi-server tandem line with exponential service times and investigate the allocation of work and servers simultaneously. Vuuren et al. (2005) study multi-server tandem queues with finite buffer and blocking and propose an approximation method based on decomposition to estimate performance measures. Li et al. (2010) provide characterizations of production system problems and provides insights on performance analysis, lean buffer allocation and reliability of production lines. Fleuren et al. (2014) present an approximation method for flow lines with multiple unreliable servers and finite buffers.

One of the most relevant work in the literature is Buzacott and Shanthikumar (1992). They provide analytical results for optimal allocation of buffer and workload in a multi-stage, finite capacity queuing network with exponential service times. However, these analytical results are not directly applicable to the models we consider due to having state dependent service queues as part of the network. Calabrese (1992) analyzes the optimal workload allocation in open multi-server Jackson networks. They provide useful optimality conditions. However, these results are not directly applicable to the RC queuing model either due to the state dependent model used for the RCs being more general. In addition to these papers, Mandelbaum and Reiman (1998) and Argon and Andradottir (2017) analyze pooling service stations in tandem queuing networks. They provide results on work allocation and performance improvement due to pooling. Pooling is related to our discussion since we consider relief center designs where a subset of distribution points are pooled together. However, due to having state dependent queues in the queuing network representation of the RC layouts, these results do not directly apply to our settings

either. We conclude that among many related articles we surveyed, very few results are directly applicable to the tandem line model of a relief center.

Some papers integrate an optimization model to the queuing network analysis. Smith et al. (2009) analyze queuing networks with finite buffers, multiple servers and general time distributions and model the problem of optimally allocating servers. Woensel et al. (2010) formulate a non-linear program to jointly optimize buffer and server allocation in networks with varying topologies and with general service times.

7.3 The Design Space for Relief Center Layouts

In this section we discuss the design space under volunteer and area constraints and enumerate all possible relief center designs. We start by introducing the components of a relief center layout that need to be decided prior to setup:

- The size (area) of the relief center
- Number of points of distribution (POD)
- Assignment of items to the points of distribution

The size of the RC area is dependent on the number of victims the RC is planned to serve, as well as the availability of space. The maximum number of points of distribution is equal to the number of volunteers available. The number of PODs to setup is related to the decision of item assignment to PODs, which will also be impacted by the types of items being distributed and the demand for each type of item. Often, the availability of space and the availability of volunteers will be dictated by the disaster aftermath conditions. Hence, the relief organizations have the most control over decisions related to the number of PODs and the assignment

of items to PODs.

For the design space we consider in this chapter, to be consistent with our assumptions in chapter 3, we assume that a fixed total area is available, 4 volunteers are present and 4 items (tarpaulins, blankets, kitchen sets, sleeping mats) are being distributed from the RC. Lastly, we assume that all victims requiring service from this RC, need all four of the items. Given these assumptions, we focus on searching the optimal number of PODs and item assignment combination. Note that the optimal buffer allocation is not considered in this chapter. Next, we describe the process of enumeration for all possible designs in the design space.

Within the design space, we consider all possibilities of number of distribution points, several variations of item assignment to distribution points, and several variations of volunteer assignments. Figure 7.2 provides a graphical representation of the enumeration logic, which first divides the design space into the number of PODs a victim would need to visit to collect all available items. Since there are four available volunteers, it is possible to assign items such that items are distributed among 1, 2, 3, or 4 PODs. Correspondingly we define four sets of designs, group A, B, C and D respectively.

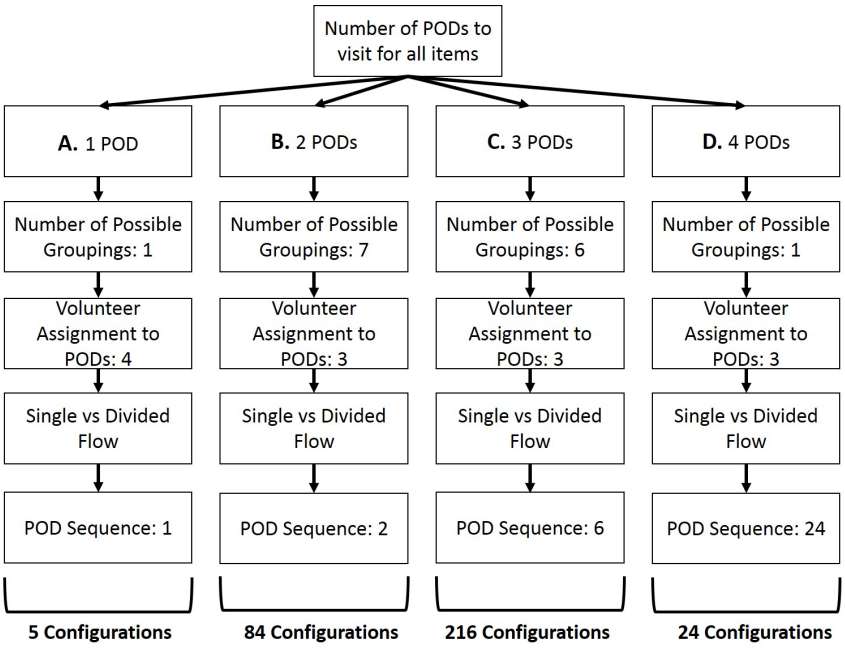

Figure 7.2: The Design Space

Group A: In this set of layout designs, the number of PODs any individual victim has to go through is one. Hence, all PODs in this group carry all available items. Based on possible volunteer assignments, there can be 4 configurations in this group as listed below. For all four configurations, each POD has its own queue characterized by λ/n, where λ is the total external arrival rate and n is the total number of PODs. See Figure 7.3, Group A-1 through Group A-4 for these configurations. In addition to these separate flow configurations, one can also design the system such that there is a common queue where all servers pull from, as shown in Figure 7.3, configuration Group A-5.

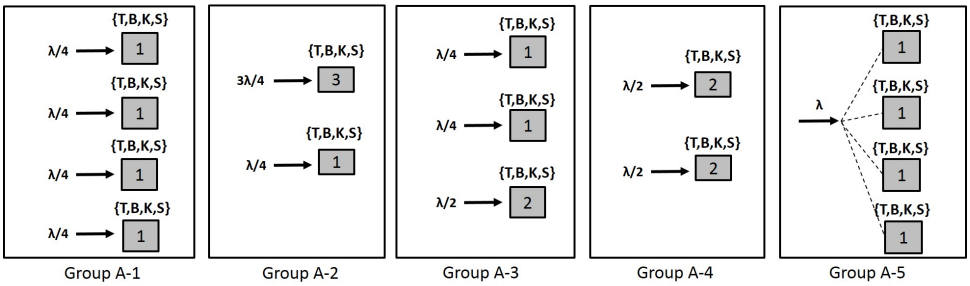

Figure 7.3: Configurations: Group A

Group B: In this set of designs, the number of PODs an individual victim has to go through is 2. Hence, the four available items need to be allocated between two PODs. The possible item assignments for these two PODs are given in Table 7.1 and there a total of seven (7) item assignment possibilities.

Table 7.1: Group B: Possible Item Assignment to PODs

	POD 1	POD 2
1	$\{T\}$	$\{B, K, S\}$
2	$\{B\}$	$\{T, K, S\}$
3	$\{K\}$	$\{T, B, S\}$
4	$\{S\}$	$\{T, B, K\}$
5	$\{T, B\}$	$\{K, S\}$
6	$\{T, K\}$	$\{B, S\}$
7	$\{T, S\}$	$\{B, K\}$

For each item assignment there are three possible volunteer assignments denoted by (2,2), (3,1) or (1,3), where the first entry denotes the number of volunteers assigned to POD 1, and the second entry denotes the number of volunteers assigned to POD 2. This makes up three (3) possible volunteer assignments for each item assignment option. For the multi-server assignments, we can design the system either

with separating arrival streams or as a common queue, creating two (2) options for each item-server assignment pair. See Figure 7.4 for a depiction of the difference in queuing between completely pooled and completely separated arrival streams. This Figure also graphically summarizes the configurations in Group B. Note that configurations in between, where one POD is pooled and the other is not, is also possible. Lastly, the sequence in which the victims flow through the two PODs can vary, and with two PODs the number of possible sequences is two (2). Therefore, the total number of configurations in group B is 84 (7x3x2x2).

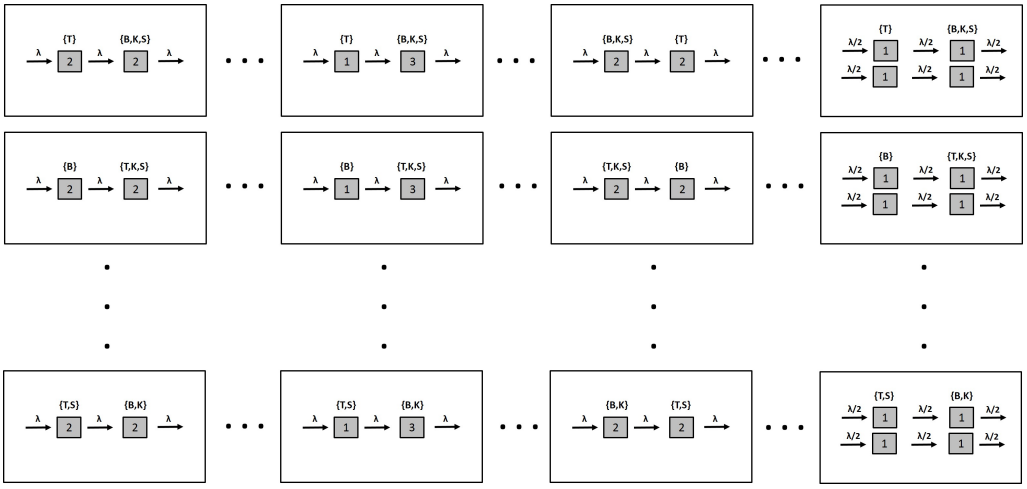

Figure 7.4: Configurations: Group B

Group C: In this set of layout designs, the number of PODs an individual victim has to go through is three. Hence, the four available items need to be allocated between three PODs. The possible item assignments are given in Table 7.2.

Table 7.2: Group C: Possible Item Assignment to PODs

	POD 1	POD 2	POD 3
1	$\{T\}$	$\{B\}$	$\{K, S\}$
2	$\{T\}$	$\{K\}$	$\{B, S\}$
3	$\{T\}$	$\{S\}$	$\{B, K\}$
4	$\{B\}$	$\{K\}$	$\{T, S\}$
5	$\{B\}$	$\{S\}$	$\{T, K\}$
6	$\{K\}$	$\{S\}$	$\{T, B\}$

For each item assignment there are three possible volunteer assignments as (1,2,1), (2,1,1) or (1,1,2), where the first, second and third entry denote the number of volunteers assigned to POD 1, POD 2 and POD 3 respectively. This makes up three (3) possible volunteer assignments for each item assignment option. For the multi-server assignments, we can design the system either with separate arrival streams or as a common queue, creating two (2) options for each item-server assignment pair. Lastly, the sequence in which the victims flow through the two PODs can be designed in 3! ways. Therefore, the total number of configurations in group C is 216 (6x3x2x6). Figure 7.5 summarizes all configurations in group C.

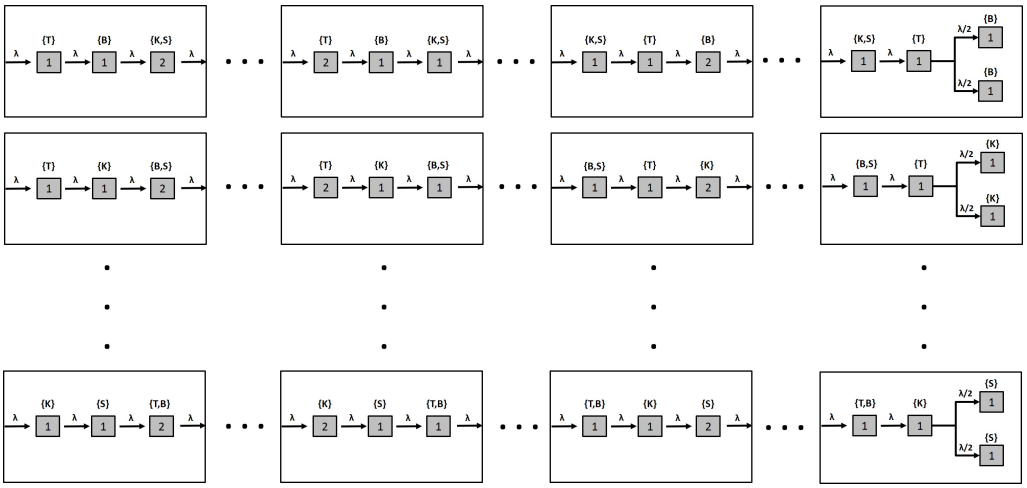

Figure 7.5: Configurations: Group C

Group D: The number of PODs an individual victim has to go through is four. Since, the four available items can only be allocated with a single item per POD, the only assignment configuration is $\{T\}$ & $\{B\}$ & $\{K\}$ & $\{S\}$. And since every POD needs to be staffed, the only possible volunteer assignment is (1,1,1,1), where the first, second, third and fourth entry denote the number of volunteers assigned to POD 1, POD 2, POD 3 and POD 4 respectively. However, the sequence in which the victims flow through the network can be designed in 4! ways. Therefore, the total number of configurations in group D is 24 (1x1x24). Figure 7.6 summarizes all configurations in group D.

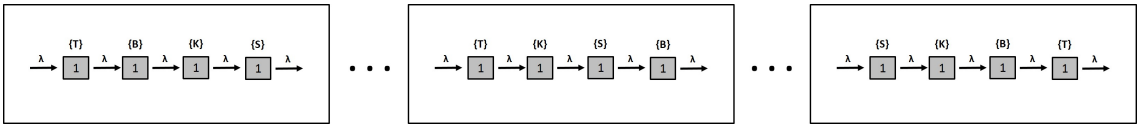

Figure 7.6: Configurations: Group D

Combining all configurations in Groups A, B, C and D, the total number of configurations to be evaluated is 329.

7.4 Simulation Design

In this section, we first describe the modeling of the RC layouts in Arena® and present the parameter setting used. Next, we discuss how to leverage the known results from the queuing literature to reduce the number of configurations to simulate. Finally, we present the numerical results and discuss the insights.

7.4.1 Simulation Model of the Relief Center Configurations

All the relief center designs are modeled as finite capacity queuing networks with state dependent (walkways) and state independent (PODs) queues. Next we describe the modeling of both types of queues in Arena®.

Modeling the Queues in the Network: Let $i \in N$ represent a queue in the network, let K_i represent the capacity of queue i and let $\mu_i(k_i)$ represent the hourly service rate at queue i. To model the queuing behavior at queue i, we use a Hold-Process-Hold-Release logic in Arena®, where the first hold module assures that the server is idle and not blocked prior to a victim moving into service and the second hold module assures that queue $i+1$ has capacity before the victim is released to the next queue (blocking after service). If queue i is blocked, then the server i stays blocked until a service completion at the downstream queue. Note that, since all queues have finite capacity, blocking can propagate through the network. This logic is critical in determining the system throughput, which is equal to $\lambda(1-a_1)$, where λ is the external arrival rate to the network and a_1 is the probability that the first queue in the network is at capacity. The only difference between the modeling of walkway and distribution queues is the service rate. We assume that the distribution PODs have an exponential service time with rate μ_i, while the walkway's have an

exponential service time with rate $\mu_i(k_i)$. The rate of the walkway service decreases as the number in queue (k_i) increases. The mathematical relationship between the number of people on the walkway and the service rate has been captured by Tregenza (1976) and is modeled using the Equations given in Chapter 6, Section 6.3.1.

Victim Routing: In this chapter, we assume that all victims arriving to the RC need all items being distributed at the RC. Hence, the exact routing will depend on the layout configuration (specifically item assignment decision to PODs).

Model Outputs: There are two main performance measures of interest: (1) the average waiting time of victims to receive supplies at the RC, and (2) the average throughput of the RC. For all models, we record the time stamp of entry, when a victim first enters the relief center using a record module. Next, we record the time stamp of the victim exiting the relief center at the end of the network and record the total time spent at the RC for each victim. We then utilize a read/write module to export all data for victim waiting times as a .csv file for analysis.

Modeling Separate versus Common Arrival Streams: For the PODs with multiple servers in all designs, we differentiate between separate and common arrival flows. Figure 7.7 shows the Arena® model of the separate arrival streams for two different PODs with single servers each. Note that each POD is accessed by a separate walkway. Since these two PODs are identical, the arriving victims can pick either POD with 50% probability. Figure 7.8 shows the Arena® model for the common arrival case. In this case, there is a single POD with 2 servers. The POD is accessed via a single walkway and a victim gets served by one of the victims based on whichever server is idle. Note that both the walkways and the PODs in Figure

7.7 is pooled in the case of Figure 7.8. We assume that for pooled walkways, the size and capacity of the walkway increases to the sum of the walkways pooled, and the curve corresponding to state dependent service time is adjusted appropriately to reflect the same relationship between crowd density and service rate.

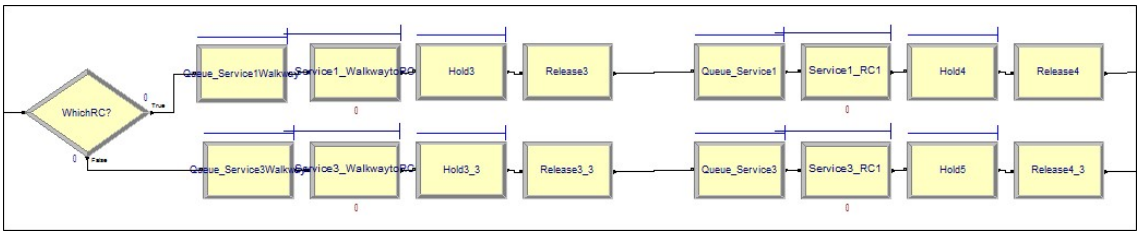

Figure 7.7: Separate Arrival Streams

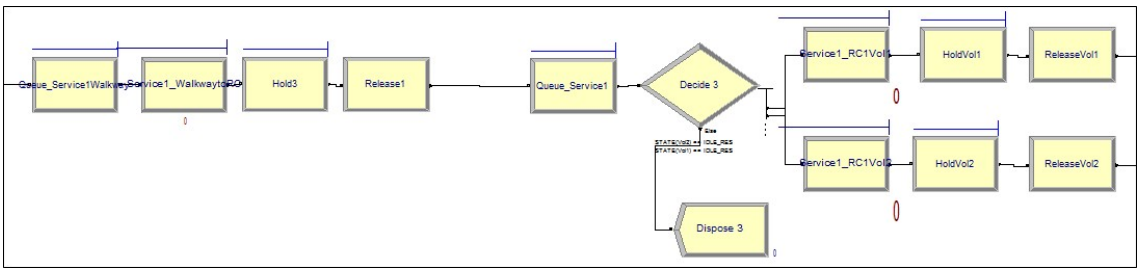

Figure 7.8: Common Arrival Stream for 2 Servers

7.4.2 The Simulation Parameters

The arrival rate to the RC and service rate at the PODs are the input parameters required for the simulation studies. We use parameters estimated from the Nepal earthquake relief distribution data. This data was obtained from the Humanitarian Data Exchange Database (an open platform for humanitarian data sharing: https://data.humdata.org/), which was prepared by contributions from the International Federation of Red Cross (IFRC), the United Nations High Commissioner for Refugees (UNHCR), Nepal United Nations Office for the Coordination of Humanitarian Affairs (OCHA), the Nepal Ministry of Home Affairs, and the Nepali Police.

In Table 7.3 we summarize the arrival and service rate parameters estimated using the data obtained. The details of data analysis are given in Chapter 3.

Note that we assume that the service rates are additive and all victims need all items. For example, if blankets and tarpaulins are distributed together at a POD, we sum the individual service times to estimate the service time at this POD. Also note that, the walkway service rate function differs based on the buffer size (number of walkways pooled). This is because the walkway speed is dependent on the number of victims present on the walkway, through the increase in density, which is a function of both the buffer space and occupants.

Table 7.3: Parameters Used in Simulations

Parameter	Value
Arrival Rate	$\lambda = 58/\text{hour}$
Service Rate by Item	$\mu_T = 10.7/hr$, $\mu_B = 31.6/hr$, $\mu_K = 10.2/hr$, $\mu_S = 20.2/hr$
Walkway Service Rate for $K_i = 50$	$(216 * \exp(-0.024 * k))$
Walkway Service Rate for $K_i = 100$	$(216 * \exp(-0.012 * k))$
Walkway Service Rate for $K_i = 150$	$(216 * \exp(-0.008 * k))$
Walkway Service Rate for $K_i = 200$	$(216 * \exp(-0.006 * k))$
Single Queue Capacity	$K_i = 50, \quad \forall i$

We note the following characteristics of all layout designs due to the parameter setting:

- In all designs the POD service rates are the bottleneck. This is because the buffer size of walkways are chosen such that the walkway speed never falls

below a certain level. For effects of lower & higher buffer levels on both walkway speed and RC performance refer to Chapter 3.

- In all of our settings the external arrival rate is higher than the maximum service rate of the PODs. Consequently, in all designs one of the PODs will be the rate determining step (the bottleneck).
- In all designs the bottleneck service rate will dictate the throughput.

7.4.3 Leveraging Results from the Queuing Literature

In this section we discuss if and when we can leverage the existing results in the literature to determine dominated designs in the design space. As presented in Section 7.2, the queuing literature has many results on optimal work and server allocation in tandem lines. Below we highlight the most relevant work and if and how the results can help us limit the number of designs to evaluate.

For a given item assignment decision, Smith et al. (2009) provide a simple and useful way of allocating servers. They show that Equation 7.1 can be used to bound the optimal server allocation in finite queuing networks with varying topologies. In this equation, λ_i represents the arrival rate to node i, c_i^* represents the optimal server allocation, μ_i is the service rate at node i, and γ is a constant showing the grade of service. Hence, the optimal server allocation is directly dependent on the utilization of the queue in a given setting.

$$\left\lceil \frac{\lambda_i}{\mu_i} \right\rceil \leq c_i^* \leq \left\lceil \frac{\lambda_i}{\mu_i} + \gamma \sqrt{\frac{\lambda_i}{\mu_i}} \right\rceil \tag{7.1}$$

Using this result and by considering the utilization of each queue in a design, we can eliminate some of the design configurations as being less efficient than other configurations.

Table 7.4 describes the reduced scenarios for group A. In the table, the first column enumerates the five configurations. In each configuration, all PODs carry all items and the associated service rate at each POD is determined as 3.67/hour. The server assignment column in the table shows the number of servers assigned to each POD in the configuration and the last column specifies the queuing discipline.

Table 7.5 describes the reduced scenarios for group B. In the table, the first column enumerates the configurations. In all configurations there are two PODs and the item assignment for each scenario is given in column four. Column five shows the service rate at each POD in the given configuration and column six shows the server assignment for each POD. The last column specifies the queuing discipline.

Table 7.6 describes the reduced scenarios for group C. In the table, the first column enumerates the configurations. In all configurations there are three PODs and the item assignment for each scenario is given in column four. Column five shows the service rate at each POD in the given configuration and column six shows the server assignment for each POD. The last column specifies the queuing discipline.

Table 7.7 describes the reduced scenarios for group D. In the table, we show one configurations since there is one item and server assignment option. Column five shows the service rate at each POD in the given configuration and column six shows the server assignment for each POD. The last column specifies the queuing discipline.

Note that in the tables showing the reduced configurations, we did not enumerate alternative scenarios based on the sequence of PODs. Including the sequence variation, for each configuration given in Table 7.5, there can be 2 alternatives for each configuration considering sequence of PODs in the network, while for configurations given in Table 7.6 there can be 6 alternatives for each and for configurations given in Table 7.7 there can be 24 alternatives.

Table 7.4: Summary of Layout Configurations: Group A

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	ueue Type
1	A	4	$\{T, B, K, S\}$	3.67/hr	(1,1,1,1)	Separate
2	A	3	$\{T, B, K, S\}$	3.67/hr	(2,1,1)	Separate
3	A	2	$\{T, B, K, S\}$	3.67/hr	(2,2)	Separate
4	A	2	$\{T, B, K, S\}$	3.67/hr	(3,1)	Separate
5	A	1	$\{T, B, K, S\}$	3.67/hr	(4)	Common

Table 7.5: Summary of Layout Configurations: Group B

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type
6	B	2	$\{B, K, S\}$ & $\{T\}$	5.58/hr & 10.7/hr	(2,2)	Common
7	B	2	$\{B, K, S\}$ & $\{T\}$	5.58/hr & 10.7/hr	(2,2)	Separate
8	B	2	$\{T, B, S\}$ & $\{K\}$	5.73/hr & 10.2/hr	(2,2)	Common
9	B	2	$\{T, B, S\}$ & $\{K\}$	5.73/hr & 10.2/hr	(2,2)	Separate
10	B	2	$\{K, S\}$ & $\{T, B\}$	6.67/hr & 8.01/hr	(2,2)	Common
11	B	2	$\{K, S\}$ & $\{T, B\}$	6.67/hr & 8.01/hr	(2,2)	Separate
12	B	2	$\{T, S\}$ & $\{B, K\}$	7.00/hr & 7.72/hr	(2,2)	Common
13	B	2	$\{T, S\}$ & $\{B, K\}$	7.00/hr & 7.72/hr	(2,2)	Separate
14	B	2	$\{T, K, S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Common
15	B	2	$\{T, K, S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Separate-partial
16	B	2	$\{T, K, S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Separate-full
17	B	2	$\{T, B, K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Common
18	B	2	$\{T, B, K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Separate-partial
19	B	2	$\{T, B, K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Separate-full
20	B	2	$\{B, S\}$ & $\{T, K\}$	12.34/hr & 5.22/hr	(1,3)	Common
21	B	2	$\{B, S\}$ & $\{T, K\}$	12.34/hr & 5.22/hr	(1,3)	Separate-partial
22	B	2	$\{B, S\}$ & $\{T, K\}$	12.34/hr & 5.22/hr	(1,3)	Separate-full

Table 7.6: Summary of Layout Configurations: Group C

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type
23	C	3	$\{T\}$ & $\{B\}$ & $\{K, S\}$	10.7/hr & 31.6/hr & 6.7/hr	(1,1,2)	Common
24	C	3	$\{T\}$ & $\{B\}$ & $\{K, S\}$	10.7/hr & 31.6/hr & 6.7/hr	(1,1,2)	Separate
25	C	3	$\{T\}$ & $\{B, S\}$ & $\{K\}$	10.7/hr & 12.3/hr & 10.2/hr	(1,1,2)	Common
26	C	3	$\{T\}$ & $\{B, S\}$ & $\{K\}$	10.7/hr & 12.3/hr & 10.2/hr	(1,1,2)	Separate
27	C	3	$\{T\}$ & $\{S\}$ & $\{B, K\}$	10.7/hr & 20.2/hr & 7.7/hr	(1,1,2)	Common
28	C	3	$\{T\}$ & $\{S\}$ & $\{B, K\}$	10.7/hr & 20.2/hr & 7.7/hr	(1,1,2)	Separate
29	C	3	$\{K\}$ & $\{B\}$ & $\{T, S\}$	10.2/hr & 31.6/hr & 7/hr	(1,1,2)	Common
30	C	3	$\{K\}$ & $\{B\}$ & $\{T, S\}$	10.2/hr & 31.6/hr & 7/hr	(1,1,2)	Separate
31	C	3	$\{T, K\}$ & $\{B\}$ & $\{S\}$	5.2/hr & 31.6/hr & 20.2/hr	(2,1,1)	Common
32	C	3	$\{T, K\}$ & $\{B\}$ & $\{S\}$	5.2/hr & 31.6/hr & 20.2/hr	(2,1,1)	Separate
33	C	3	$\{K\}$ & $\{S\}$ & $\{T, B\}$	10.2/hr & 20.2/hr & 8.01/hr	(1,1,2)	Common
34	C	3	$\{K\}$ & $\{S\}$ & $\{T, B\}$	10.2/hr & 20.2/hr & 8.01/hr	(1,1,2)	Separate

Table 7.7: Summary of Layout Configurations: Group D

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type
35	D	4	$\{K\}$ & $\{T\}$ $\{B\}$ & $\{S\}$	10.2/hr & 10.7/hr 31.6/hr & 20.2/hr	(1,1,1,1)	Common

7.5 Simulation Results

In our numerical studies, we evaluate all the configurations detailed in Section 7.4.3. For all models, a sample run of 20 days was evaluated first to determine the warm-up period length and the number of replications required.

After determining the warm-up period, we next determine the number of replications. In Equation 7.2 n represents the number of replications, s represents the sample standard deviation, and h represents the desired confidence interval range. In our numerical experiments we use a 95% confidence interval level and set $h = 2$ hours for the average waiting time measure. The number of replications requirement ranged between 3 and 75 for different designs. Hence, we run each experiment for 100 replications in all the simulations.

$$n = t^2_{n-1,1-\alpha/2}\frac{s^2}{h^2} \tag{7.2}$$

Next, we present the results of the configurations listed in Tables 7.4, 7.5, 7.6, 7.7 and we discuss the impact of sequence of PODs on the performance measures. Tables 7.8, 7.9, 7.10 and 7.11 show the average throughput (in victims per hour) and average waiting time (in hours) measures for configurations in Groups A, B, C and D respectively. Next we discuss the results for each group.

Insight for Group A: Firstly, with respect to the throughput performance measure, the configurations in Group A perform the best. However, they also result in a high average waiting time for victims. This is because the the service rate at the PODs are higher due to the additive service times. Thus, even though the network

is shorter, the queues are slower, causing higher waiting times. Note that in all configurations, the throughput measure is more consistent compared to the waiting time measure. This is because the throughput measure is tightly correlated with the combined effect of service time and server assignment, which is steady in all Group A configurations. Lastly, the lower the number of PODs (and for combined configurations) the waiting time is higher. This is bacause when we combine the PODs, we also combine the buffer space, and greater buffer space results in a higher average number in queue, causing victims to wait behind more individuals prior to service.

Insight for Group B: Looking at the combined results for both the RC throughput and victim waiting time measures, we conclude that configurations 20, 21, 22 are the best performing designs both within group B and overall. They all have a relatively high throughput for a considerably low average victim waiting time. Comparing these configurations with the other in the group, we first note that the throughput is again driven by the combination of service rate and item assignment for the slowest queue in the network. Hence, the throughput measure does not vary much based on queue type. However, the waiting time does depend on the queue type. This is because we combine the buffer space for the common queue, creating a possibility for a higher average number in queue, which in turn can increase the average waiting time metric. We observe that balanced networks perform the best and the queuing discipline (hence the buffer space) matter the least. This is because with balanced service rates at the PODs, blocking is minimal (note that for blocking to happen the downstream queue need to be full). For the networks with greater imbalances in service rates (ex: see rows 14 and 15), the common queuing discipline where we pool the buffers increase the average number in queue in drive the waiting times up.

Table 7.8: Group A: Results

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type	$\bar{TH}$	$\bar{W}$
1	A	4	$\{T, B, K, S\}$	3.67/hr	(1,1,1,1)	Separate	14.432	16.995
2	A	3	$\{T, B, K, S\}$	3.67/hr	(2,1,1)	Separate	14.590	19.360
3	A	2	$\{T, B, K, S\}$	3.67/hr	(2,2)	Separate	14.640	22.865
4	A	2	$\{T, B, K, S\}$	3.67/hr	(3,1)	Separate	14.652	22.275
5	A	1	$\{T, B, K, S\}$	3.67/hr	(4)	Common	14.690	27.543

Table 7.9: Group B: Results

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type	TH	W
6	B	2	$\{B,K,S\}$ & $\{T\}$	5.58/hr & 10.7/hr	(2,2)	Common	11.128	18.351
7	B	2	$\{B,K,S\}$ & $\{T\}$	5.58/hr & 10.7/hr	(2,2)	Separate	10.087	17.594
8	B	2	$\{T,B,S\}$ & $\{K\}$	5.73/hr & 10.2/hr	(2,2)	Common	11.447	17.865
9	B	2	$\{T,B,S\}$ & $\{K\}$	5.73/hr & 10.2/hr	(2,2)	Separate	10.707	16.903
10	B	2	$\{K,S\}$ & $\{T,B\}$	6.67/hr & 8.01/hr	(2,2)	Common	13.340	15.610
11	B	2	$\{K,S\}$ & $\{T,B\}$	6.67/hr & 8.01/hr	(2,2)	Separate	11.832	15.179
12	B	2	$\{T,S\}$ & $\{B,K\}$	7.00/hr & 7.72/hr	(2,2)	Common	13.968	15.223
13	B	2	$\{T,S\}$ & $\{B,K\}$	7.00/hr & 7.72/hr	(2,2)	Separate	11.585	15.231
14	B	2	$\{T,K,S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Common	12.432	24.496
15	B	2	$\{T,K,S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Separate-partial	12.285	17.117
16	B	2	$\{T,K,S\}$ & $\{B\}$	4.15/hr & 31.6/hr	(3,1)	Separate-full	12.316	9.168
17	B	2	$\{T,B,K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Common	13.418	22.788
18	B	2	$\{T,B,K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Separate-partial	13.328	16.566
19	B	2	$\{T,B,K\}$ & $\{S\}$	4.48/hr & 20.2/hr	(3,1)	Separate-full	13.240	14.936
20	B	2	$\{B,S\}$ & $\{T,K\}$	12.34/hr & 5.22/hr	(1,3)	Common	12.343	8.624
21	B	2	$\{B,S\}$ & $\{T,K\}$	12.34/hr & 5.22/hr	(1,3)	Separate-partial	12.351	8.913
22	B	2	$\{B,S\}$ & $\{T,K\}$	12.34/hr & 5.22/hr	(1,3)	Separate-full	12.316	9.168

Insight for Group C: The throughput measure for configuration in group C is again driven by the combination of service rate and item assignment for the slowest queue in the network. Since the slowest service rate for these configurations is slower than configurations in group B, we observe lower average throughput. Also note that, the configurations in group C have a more balanced work load assignment, resulting in a waiting time measure less effected by the queue type. Note that rows 31 and 32 are the only configurations in group C with higher waiting time measure due to having more buffer space at the beginning of the network, increasing the average number in queue.

Insight for Group D: The throughput measure for the configuration in group D is again driven by the combination of service rate and item assignment for the slowest queue in the network. Since the slowest service rate for these configurations is slower than configurations in group B, we observe lower average throughput. One important result to note here is that the waiting time measure is lower compared to the results of group A. This comparison lead us to the conclusion that a longer network of queues is not necessarily harmful when the queues are faster.

Table 7.10: Group C: Results

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type	$\bar{TH}$	$\bar{W}$
23	C	3	$\{T\}$ & $\{B\}$ & $\{K,S\}$	10.7/hr & 31.6/hr & 6.7/hr	(1,1,2)	Common	10.717	9.957
24	C	3	$\{T\}$ & $\{B\}$ & $\{K,S\}$	10.7/hr & 31.6/hr & 6.7/hr	(1,1,2)	Separate	10.692	10.317
25	C	3	$\{T\}$ & $\{B,S\}$ & $\{K\}$	10.7/hr & 12.3/hr & 10.2/hr	(1,1,2)	Common	10.664	10.297
26	C	3	$\{T\}$ & $\{B,S\}$ & $\{K\}$	10.7/hr & 12.3/hr & 10.2/hr	(1,1,2)	Separate	10.690	10.344
27	C	3	$\{T\}$ & $\{S\}$ & $\{B,K\}$	10.7/hr & 20.2/hr & 7.7/hr	(1,1,2)	Common	10.679	9.895
28	C	3	$\{T\}$ & $\{S\}$ & $\{B,K\}$	10.7/hr & 20.2/hr & 7.7/hr	(1,1,2)	Separate	10.720	10.034
29	C	3	$\{K\}$ & $\{B\}$ & $\{T,S\}$	10.2/hr & 31.6/hr & 7/hr	(1,1,2)	Common	10.210	10.333
30	C	3	$\{K\}$ & $\{B\}$ & $\{T,S\}$	10.2/hr & 31.6/hr & 7/hr	(1,1,2)	Separate	10.190	10.571
31	C	3	$\{T,K\}$ & $\{B\}$ & $\{S\}$	5.2/hr & 31.6/hr & 20.2/hr	(2,1,1)	Common	10.365	19.720
32	C	3	$\{T,K\}$ & $\{B\}$ & $\{S\}$	5.2/hr & 31.6/hr & 20.2/hr	(2,1,1)	Separate	10.303	15.452
33	C	3	$\{K\}$ & $\{S\}$ & $\{T,B\}$	10.2/hr & 20.2/hr & 8.01/hr	(1,1,2)	Common	10.198	10.300
34	C	3	$\{K\}$ & $\{S\}$ & $\{T,B\}$	10.2/hr & 20.2/hr & 8.01/hr	(1,1,2)	Separate	10.204	10.426

Table 7.11: Group D: Results

Number	Group	Number of PODs	Item Assignment	Service Rate	Server Assignment	Queue Type	$\bar{TH}$	$\bar{W}$
35	D	4	$\{K\}$ & $\{T\}$ $\{B\}$ & $\{S\}$	10.2/hr & 10.7/hr 31.6/hr & 20.2/hr	(1,1,1,1)	Common	10.166	11.968

Next, we focus on investigating the impact of sequence of PODs on the performance measures of a relief center design. Table 7.12 compares all sequence combinations for a select set of configurations (configuration number 6, 7 and 23). We specifically chose configurations 6 and 7 to show the effects of sequence on different queue types (common versus separate), and we chose configuration 23 since configurations is group C have more sequence combinations. Note that in Table 7.12 the numbering enumerates the sequence combinations for each configuration. First, we conclude that the sequence of PODs for a given configuration does not have a considerable impact on throughput but can greatly change the waiting time experienced. Secondly, the average waiting time is the shortest when the slowest POD is at the beginning of the route. This happens since it causes the majority of the arrivals that are above capacity, to be truncated at the beginning of the network, instead of increasing the work in process, hence congestion. This effect is doubled since all configurations have state dependent queues in the network, and these queues slow down as congestion increases. As a result, the waiting time measure suffers even further. It is important to note that due to the way the capacity of the walkways were set, the walkway speed can not fall below a minimum bound. Hence, in none of the configurations considered in this chapter, the bottleneck is the walkways. This is why the sequence only impacts the waiting time measure. However, if the capacity of the walkways are not setup adequately and can allow congestion to increase steeply, the sequence and its effects on increasing congestion can lead to both performance measures to suffer greatly.

Table 7.12: Impact of Sequence

Number	Group	Item Assignment	Queue Type	$T\bar{H}$	$\bar{W}$
6.1	B	$\{B, K, S\}$ & $\{T\}$	Common	11.128	18.351
6.2	B	$\{T\}$ & $\{B, K, S\}$	Common	11.175	36.294
7.1	B	$\{B, K, S\}$ & $\{T\}$	Separate	10.087	17.594
7.2	B	$\{T\}$ & $\{B, K, S\}$	Separate	10.304	32.595
23.1	C	$\{T\}$ & $\{B\}$ & $\{K, S\}$	Common	10.717	9.957
23.2	C	$\{T\}$ & $\{K, S\}$ & $\{B\}$	Common	10.696	10.324
23.3	C	$\{B\}$ & $\{T\}$ & $\{K, S\}$	Common	10.684	19.813
23.4	C	$\{B\}$ & $\{K, S\}$ & $\{T\}$	Common	10.734	31.140
23.5	C	$\{K, S\}$ & $\{B\}$ & $\{T\}$	Common	10.701	31.405
23.6	C	$\{K, S\}$ & $\{T\}$ & $\{B\}$	Common	10.690	21.748

7.6 Conclusions

In this chapter, we determine the optimal layout configuration under constraints related to the pre-determined number of volunteers and types items being distributed. We exhaustively enumerate the number of possible relief center layout designs. We consider all possible combinations of number of points of distribution (PODs) in an RC, the assignment of items to the PODs, the assignment of volunteers to the PODs and the sequence of the PODs in the constructed layout. In total we find 329 possible designs.

We then discuss the known results in the literature for tandem queuing networks and eliminate designs that are dominated based on server assignment for a

given item assignment. We highlight that, even though the queuing literature has many results on optimizing tandem queuing networks for maximizing the system throughput, we find many of them are not directly applicable to a network with state dependent queues.

We evaluate the remaining configurations in the design space using Arena® simulation. From our results we conclude that: (1) The throughput of the RC is governed by the bottleneck rate, which is determined by the combination of the item and volunteer assignment decision, (2) Placing the slowest POD to the beginning of the network , limits the congestion within the RC and results in the lowest waiting time measures, (3) State dependent queues increase the impact of congestion (measured by average queue lengths) on waiting times and for designs with higher buffer space can lead to a significant decrease in throughput as well.

The main takeaways from this study for the practitioners are, first to assign items to PODs in conjunction with available volunteer capacity and to consider the bottleneck rate of service as a good indicator of RC throughput. Secondly, it is critical for practitioners to be aware of the effects of increased congestion within the relief center. Hence designs that limit congestion (by limiting buffer space, sequence of PODs or volunteer assignment) will perform the best.

Chapter 8

Conclusions and Future Research Directions

In this chapter, we summarize our main conclusions and discuss future research directions.

8.1 Conclusions

In this thesis, we focus on improving relief distribution operations during the immediate response phase of a disaster to help relieve the human suffering inflicted by a disaster. The operations research literature on relief distribution has mostly focused on getting the relief items to the disaster effected region. However, challenges faced once the relief items arrive at the disaster site has received little attention. This thesis aims to fill this gap.

The challenges this thesis focuses on were all motivated by the needs in the field. These needs were identified by a series of interviews with Salvation Army,

Red Cross, South East Wisconsin Citizens and Organizations Active in Disasters (COAD) practitioners.

The first challenge we focus on is how to set up a relief center for the most effective distribution of relief items to victims. We define effectiveness by increasing the number of victims served and decreasing the average victim waiting time. We model the relief distribution operations through a relief center using a finite capacity state dependent queuing network. To estimate the performance of a relief center using this model, we derive new analytic formulas for steady state probabilities of state dependent Coxian queues. We then apply the proposed model to Nepal earthquake data. Our analysis of relief centers lead us to the following insights: (1) Strategies that dissipate the crowds in the design of a relief center can greatly improve the efficiency of relief distribution, (2) Operational strategies including work assignment within the relief center, limiting queue lengths at the relief center and alternative uses of a triage queue can also increase relief distribution efficiency.

Next, we model a network of relief centers distributing aid in a given area as a generalized queuing network (G-network). We specifically model increase in demand for items, jockeying of victims and changing victim needs. We use a G-network model, (i) since it allows for a flexible probabilistic structure to model such victim behavior and (2) G-networks have been shown to have product form results, which can be of great computational advantage. In this chapter, we first prove a new product form result for G-networks with batch transfers under certain conditions. Then, we relax these conditions and propose a product form approximation. Finally, we use the model and solution methodology on a case study and obtain the following insights: (1) Victim mobility has a significant impact on performance measure es-

timates, (2) Decisions such as number of relief centers to open and the assignment of items to relief centers need to depend on demand estimates as well as expected victim movements.

Next, we focus our attention on the problem of material convergence and how to effectively allocate resources to the solicited and unsolicited in-kind donations to obtain the maximum number of high-priority items. We model the donation arrival and sorting process for both solicited and unsolicited donations as separate multi-server transient queues over multiple periods. In the model, the evolution of queue lengths represent the material convergence levels. Material convergence is a frequently observed phenomenon, however, to our knowledge this work is the first in the literature to propose a model to quantify it. Through numerical experiments we arrive at the following insights: (1) Material convergence and resource allocation decisions are inter-dependent, (2) Decisions that increase the high priority (HP) item throughput can result in very high material convergence levels, hence, policies that maximize HP item throughput with no regard to material convergence may not be optimal.

Next, we focus on analyzing the transient behavior of relief center design and operations using a simulation study. The queuing models previously used to estimate relief center performance, analyze steady state performance and assume Markovian inter-arrival and service times. In Chapter 6, we aim to relax these assumptions and analyze the transient behavior. We conclude that: (1) Non-stationary arrival rates impact the throughput and waiting time measures of the relief centers, (2) In the transient analysis, crowd dissipation strategies and item assignment to RCs remain to be effective strategies in improving RC performance, (3) Both the average

throughput and the average waiting time performance measures are very robust to the probability distribution chosen to represent the inter-arrival and service time distributions.

Finally, we focus on determining the optimal layout configuration for a relief center under resource constraints using discrete event simulation. We again model the relief center as a finite capacity, state dependent queuing network. Then, we enumerate layout designs in the design space and leverage known queuing network results to eliminate less efficient layout designs. Finally, we simulate the remaining designs and reach the following conclusions: (1) Assign items to the points of distribution in conjunction with available volunteer capacity. (2) Designs that limit congestion (by limiting buffer space, sequence of PODs or volunteer assignment) perform the best.

We hope that the insights obtained in this thesis will increase our ability to more effectively respond to disasters and the models in this thesis will enrich the literature and create future research directions.

8.2 Future Research Directions

Optimal Buffer Allocation in State Dependent Queuing Networks: In Chapter 7 we model the relief center layout as a finite capacity, state dependent queuing network. We then survey the related literature on optimal buffer, work and server allocation results in tandem lines to limit our design space. We find that the literature on tandem lines is rich, however very few results are applicable to a network with state dependent queues. Hence, we foresee the extension of analytic

results on optimal buffer and work load allocation in tandem lines with state dependent service rates, as an important future research direction.

We believe such results can be utilized in disaster response literature to optimize relief center layouts and model relief distribution operations. In addition, these results can also transcend this application and be used in many manufacturing system design and optimization problems.

Using Crowd Sourced Data to Manage Material Convergence: In Chapter 5 we model the material convergence problem that occurs in almost all disasters due to the convergence of in-kind donations that contain non-priority supplies in large volumes over a short period of time. We believe it would be worthwhile to extend this model so that crowd sourced data can be integrated with the model. In specific, crowd sourced data can be used to track the donation arrivals and feed the analytic model. In turn, the model will output resource assignment decisions.

The level of donations sent to the disaster region is closely related to social media and news cycles calling for supplies (Yates and Paquette (2011); Gao et al. (2011)). Moreover, many crowd sourced donations are organized via social media (during Haiti earthquake response, approximately 2.3 million tweets included the word "Haiti" or "Red Cross" between January 12 and January 14) (The New Media Index (2017)). We believe this data can be mined to estimate the expected arrival rate of donations. Using real time data mining, the model developed can be integrated to become a real time decision support tool to manage material convergence.

Bibliography

Acimovic, J. and Goentzel, J. (2016). Models and metrics to assess humanitarian response activity. *Journal of Operations Management*, 45:11–29.

Adida, E., DeLaurentis, P.-C. C., and Lawley, M. A. (2011). Hospital stockpiling for disaster planning. *IIE Transactions*, 43(5):348–362.

Aflaki, A. and Pedraza-Martinez, A. J. (2016). Humanitarian funding in a multi-donor market with donation uncertainty. *Production and Operations Management*, 25(7):1274–1291.

Ahmed, M. A. and Alkhamis, T. M. (2009). Simulation optimization for an emergency department healthcare unit in kuwait. *European Journal of Operations Research*, 198(3):936–942.

Akkihal, A. R. (2006). Inventory pre-positioning for humanitarian operations. Master's thesis, MIT, Boston.

Aksu, D. T. and Ozdamar, L. (2014). A mathematical model for post-disaster road restoration: Enabling accessibility and evacuation. *Transportation Research Part E: Logistics and Transportation Review*, 61(1):56–67.

Albores, P. and Shaw, D. A. (2008). Government preparedness: Using simulation to

prepare for a terrorist attack. *Computers and Operations Research*, 35(6):1924–1943.

Altay, N. and Green, W. G. (2006). Or/ms research in disaster operations management. *European Journal of Operational Research*, 175(1):475–493.

Altay, N. and Pal, R. (2014). Information diffusion among agents: Implications for humanitarian operations. *Production and Operations Management*, 23(6):1015–1027.

Altiok, T. (1982). Approximate analysis of exponential tandem queues with blocking. *European Journal of Operational Research*, 11(4):390–398.

Argon, N. T. and Andradottir, S. (2017). Pooling in tandem queuing networks with non-colloborative servers. *Queuing Systems*, 87(1):345–377.

Arnette, A. N. and Zobel, C. W. (2015). An empirical investigation of the material convergence problem. *Quick Response Grant Report Series*, 253.

Aros, S. K. and Gibbons, D. E. (2018). Exploring communication media options in an inter-organizational disaster response coordination network using agent-based simulation. *European Journal of Operational Research*.

Arsham, H., Balana, A. R., and Gross, D. (1983). Numerical methods for transient solutions of machine repair problems. *Computers and Industrial Engineering*, 7(2):149–157.

Artalejo, J. (2000). G-networks: A versatile approach for work removal in queuing networks. *European Journal of Operational Research*, 126(1):233–249.

Balcik, B. and Ak, D. (2014). Supplier selection for framework agreements in humanitarian relief. *Production and Operations Management*, 23(6):1028–1041.

Balcik, B. and Beamon, B. M. (2008). Facility location in humanitarian relief. *International Journal of Logistics: Research and Applications*, 11(2):101–121.

Balcik, B., Beamon, B. M., Krejci, C. C., Muramatsu, K. M., and Ramirez, M. (2010). Coordination in humanitarian relief chains: Practices, challenges and opportunities. *International Journal of Production Economics*, 126(1):22–34.

Barbarosoglu, G. and Arda, Y. (2004). A two-stage stochastic programming framework for transportation planning in disaster response. *Journal of the Operational Research Society*, 55(1):43–53.

Baskett, F., Chandy, K. M., Muntz, R., and Palacios, F. (1975). Open, closed and mixed networks of queues with different classes of customers. *Journal of ACM*, 22(2):248–260.

Batta, R. and Mannur, N. (1990). Covering location models for emergency situations that require multiple response units. *Management Science*, 36(1):16–23.

Bihan, H. L. and Dallery, Y. (2000). A robust decomposition method for the analysis of production lines with unreliable machines and finite buffers. *Annals of Operations Research*, 93(1):265–297.

Bocharov, P., Dapice, C., Gavrilov, E., and Pechinkin, A. (2004a). Product form solution for a g-networks with dependent service. *RAIRO Operations Research*, 38(1):105–119.

Bocharov, P. and Vishnevskii, V. M. (2003). G-networks: Development of the theory of multiplicative networks. *Automation and Remote Control*, 64(5):714–739.

Bocharov, P. P., Dapice, C., Gavrilov, E., and Pechinkin, A. (2004b). Decomposition

of queuing networks with dependent service and negative customers. *Automation and Remote Control*, 65(1):86–103.

Brandwajn, A. and Jow, Y. (1988). An approximation method for tandem queues with blocking. *Operations Research*, (36):73–83.

Buzacott, J. A. and Shanthikumar, J. G. (1992). Design of manufacturing systems using queueing models. *Queuing Systems*, 12(1):135–214.

Calabrese, J. M. (1992). Optimal workload allocation in open networks of multi-server queues. *Management Science*, 38(12):1792–1802.

Campbell, A. M., Vandenbussche, D., and Hermann, W. (2008). Routing for relief efforts. *Transportation Science*, 42(2):127–145.

Caunhye, A. M., Nie, X., and Pokharel, S. (2012). Optimization models in emergency logistics: A literature review. *Socio-Economic Planning Sciences*, 46(1):4–13.

CEDIM Forensic Disaster Analysis Group and South Asia Institute (SAI) (2015). Shelter response and vulnerability of displaced populations in the april 25, 2015 nepal earthquake. Technical report.

Chao, X. and Pinedo, M. (1995). Networks of queues with batch services, signals and product form solutions. *Operations Research Letters*, 17(1):237–242.

Cheah, J. and Smith, J. M. (1994). Generalized m/g/c/c state dependent queuing models and pedestrian traffic flows. *Queuing Systems*, 15:365–386.

Chen, X., Kwan, M., Li, Q., and Chen, J. (2012). A model for evacuation risk assessment with consideration of pre- and post-disaster factors. *Computers, Environment and Urban Systems*, 36(3):207–217.

Chen, X. and Zhan, F. B. (2008). Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies. *Journal of the Operational Research Society*, 59(1):25–33.

Chiu, Y., Zheng, H., Villalobos, J., and Gautam, B. (2007). Modeling no-notice mass evacuation using a dynamic traffic flow optimization model. *IIE Transactions*, 39(1):83–94.

Chiu, Y., Zheng, H., Villalobos, J. A., Peacock, W., and Henk, R. (2008). Evaluating regional contra-flow and phased evacuation strategies for texas using a large-scale dynamic traffic simulation and assignment approach. *Journal of Homeland Security and Emergency Management*, 5(1).

Cho, S.-h., Jang, H., Lee, T., and Turner, J. (2014). Simultaneous location of trauma centers and helicopters for emergency medical service planning. *Operations Research*, 62(4):751–771.

Clay, D. (2007). Issues in managing disaster relief inventories. *International Journal of Production Economics*, 108:228–235.

Conte, S. D. and de Boor, C. (1980). *Elementary Numerical Analysis*. McGraw-Hill, New York.

Cruz, F. R. B., Smith, J. M., and Medeiros, R. (2003). An m/g/c/c state dependent network simulation model. *Computers and Operations Research*, 32(4):919–941.

Dallery, Y. and Frein, Y. (1993). On decomposition methods for tandem queuing networks with blocking. *Operations Research*, 41(2):386–399.

Dallery, Y. and Gershwin, S. B. (1992). Manufacturing flow line systems: a review of models and analytical results. *Queuing Systems*, 12(1):3–94.

Das, R. and Hanaoka, S. (2014). Relief inventory modelling with stochastic lead-time and demand. *European Journal of Operational Research*, 235:616–623.

Das, T. K., Savachkin, A. A., and Zhu, Y. (2008). A large-scale simulation model of pandemic influenza outbreaks for development of dynamic mitigation strategies. *IIE Transactions*, 40(9):893–905.

de Silva, F. N. and Eglese, R. W. (2000). Integrating simulation modelling and gis: Spatial decision support systems for evacuation planning. *The Journal of the Operational Research Society*, 51(4):423–430.

Department of Homeland Security (2011). National preparedness goal. Technical report.

Department of Homeland Security (2012). National preparedness report. Technical report.

Deqiang, F., Yun, L., and Changbing, L. (2011). Forecasting the demand of emergency supplies: Based on the CBR theory and BP neural network. In Kaminishi, K., Duysters, G., and de Hoyos, A., editors, *Proceedings of the 8th International Conference on Innovation and Management*, pages 700–704.

Destro, L. and Holguin-Veras, J. (2010). Estimating material convergence: Flow of donations for hurricane Katrina. In *Proceedings of the 90th Annual Meeting of the Transportation Research Board*.

Duran, S., Gutierrez, M. A., and Keskinocak, P. (2011). Pre-positioning of emergency items worldwide for care international. *Interfaces*, 41(3):223–237.

Ergun, O., Gui, L., Heier Stamm, J. L., Keskinocak, P., and Swann, J. (2014). Im-

proving humanitarian operations through technology-enabled collaboration. *Production and Operations Management*, 23(6):1002–1014.

Executive Office of the President of the United States: Subcommittee on Disaster Reduction (2005). Grand challenges for disaster reduction. Technical Report 1.

Fang, Z., Zong, X., Li, Q., Li, Q., and Xiong, S. (2011). Hierarchical multi-objective evacuation routing in stadium using ant colony optimization approach. *Journal of Transport Geography*, 19(1):443–451.

Federal Emergency Management Agency (2008). *IS-26 Guide to Points of Distribution*. FEMA.

Fikar, C., Hirsch, P., and Nolz, P. C. (2017). Agent-based simulation optimization for dynamic disaster relief distribution. *Central European Journal of Operations Research*, pages 1–20.

Fleuren, S., Bierbooms, R., and Adan, I. (2014). Performance analysis of exponential multi-server production lines with fluid flow and finite buffers. *Stochastic Models*, 30(1):469–493.

Fox, B. L. and Glynn, P. W. (1988). Computing poisson probabilities. *Communications of the ACM*, 31(4):440–445.

Galindo, G. and Batta, R. (2013). Review of recent developments in or/ms research in disaster operations management. *European Journal of Operational Research*, 230(1):201–211.

Gao, H., Barbier, G., and Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, 26(3):10–14.

Gatignon, A., Vassenhove, L. N. V., and Charles, A. (2010). The yogyakarta earthquake: Humanitarian relief through ifrc's decentralized supply chain. *International Journal of Production Economics*, (126):102–110.

Gelenbe, E. (1991). Product-form queuing networks with negative and positive customers. *Journal of Applied Probability*, 28(3):656–663.

Gelenbe, E. (1993a). G-networks with signals and batch removal. *Probability in the Engineering and Informational Sciences*, 7(3):335–342.

Gelenbe, E. (1993b). G-networks with triggered customer movement. *Journal of Applied Probability*, 30(3):742–748.

Gelenbe, E., Glynn, P., and Sigman, K. (1991). Queues with negative arrivals. *Journal of Applied Probability*, 28(1):245–250.

Gelenbe, E. and Labed, A. (1998). G-networks with multiple classes of signals and positive customers. *European Journal of Operational Research*, 108(1):293–305.

Gershwin, S. B. (1987). An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking. *Operations Research*, 35(2):291–305.

Gershwin, S. B. and Burman, M. H. (2000). A decomposition method for analyzing inhomogeneous assembly/disassembly systems. *Annals of Operations Research*, 93(1):91–115.

Ghurye, J. and Krings, G. (2016). A framework to model human behavior at large scale during natural disasters. In *2016 17th IEEE International Conference on Mobile Data Management*.

Gomez-Corral, A. (2002). On a tandem g-network with blocking. *Advances in Applied Probability*, 34(3):626–661.

Gross, D. and Miller, D. R. (1982). The randomization technique as a modeling tool and solution procedure for transient markov processes, technical paper serial t-467. Technical report, The George Washington University, Institute for Management Science and Engineering.

Gross, D. and Miller, D. R. (1984). The randomization technique as a modeling tool and solution procedure for transient markov processes. *Operations Research*, 32(2):343–361.

Han, Y., Guan, X., and Shi, L. (2011). Optimization based method for supply location selection and routing in large scale emergency material delivery. *IEEE Transactions on Automation Science and Engineering*, 8(4):683–693.

Harrison, P. G. and Pitel, E. (1996). The m/g/1 queue with negative customers. *Advances in Applied Probability*, 28(2):540–566.

Hawe, G. I., Coates, G., Wilson, D. T., and Crouch, R. S. (2015). Agent-based simulation of emergency response to plan the allocation of resources for a hypothetical two-site major incident. *Engineering Applications of Artificial Intelligence*, 46:336–345.

Heyman, D. P. and Sobel, M. J. (1982). *Stochastic Models in Operations Research*. McGraw-Hill, New York.

Hillier, F. S. and Boling, R. W. (1966). The effect of some design factors on the efficiency of production lines with variable operation times. *Journal of Industrial Engineering*, 17(12):651–658.

Hillier, F. S. and Boling, R. W. (1979). On the optimal allocation of work in symmetrically unbalanced production line systems with variable operation times. *Management Science*, 25(8):721–728.

Hillier, F. S. and So, K. C. (1996). On the simultaneous optimization of server and work allocations in production line systems with variable processing times. *Operations Research*, 44(3):435–443.

Hobeika, A. G., Kim, S., and Beckwith, R. E. (1994). A decision support system for developing evacuation plans around nuclear power stations. *Interfaces*, 24(5):22–35.

Holguin-Veras, J., Jaller, M., Wassenhove, L. N. V., Perez, N., and Wachtendorf, T. (2012a). On the unique features of post-disaster humanitarian logistics. *Journal of Operations Management*, 30(1):494–506.

Holguin-Veras, J., Jaller, M., Wassenhove, L. N. V., Perez, N., and Wachtendorf, T. (2014). Material convergence: Important and understudied disaster phenomenon. *Natural Hazards Review*, 3:494–506.

Holguin-Veras, J., Perez, N., Jaller, M., Wassenhove, L. N. V., and Aros-Vera, F. (2013). On the appropriate objective function for post-disaster humanitarian logistics models. 31(1):262–280.

Holguin-Veras, J., Perez, N., Ukkusuri, S., Wachtendorf, T., and Brown, B. (2007). Emergency logistics issues affecting the response to Katrina. *Transportation Research Record: Journal of the Transportation Research Board*, (2022):76–82.

Holguin-Veras, J., Taniguchi, E., Ferreira, F., Jaller, M., and Thompson, R. G. (2012b). The Tohoku disasters: Preliminary findings concerning the post disaster

humanitarian logistics response. In *Proceedings of the 2012 Annual Meeting of the Transportation Research Board*.

Hong, X., Lejeune, M., and Noyan, N. (2015). Stochastic network design for disaster preparedness. *IIE Transactions*, 47(4):329–357.

Hopp, W. J. and Spearman, M. L. (2008). *Factory Physics*. McGraw-Hill, Boston.

Humanitarian Data Exchange Database (2016). https://data.humdata.org/.

Iakovou, E., Ip, C. M., Douligeris, C., and Korde, A. (1996). Optimal location and capacity of emergency cleanup equipment for oil spill response. *European Journal of Operational Research*, 96(1):72–80.

International Federation of Red Cross and Red Crescent Societies (2008). Guidelines for assessment in emergencies. Technical report.

International Federation of Red Cross and Red Crescent Societies (2015). Emergency appeal operation update 2. Technical report.

Iserson, K. and Moskop, J. (2007). Triage in medicine, part i: Concept, history and types. *Annals of Emergency Medicine*, 49(3):275–281.

Islam, M. M., Vate, J. V., Heggestuen, J., Nordenson, A., and Dolan, K. (2013). Transforming in-kind giving in disaster response: A case for on-line donation registry with retailers. In *Proceedings of the 2013 IEEE Global Humanitarian Technology Conference*.

Iyama, T. and Ito, S. (1987). The maximum production rate for an unbalanced multi-server flow line system with finite buffer storage. *International Journal of Production Research*, 25(8):1157–1170.

Jackson, J. R. (1963). Jobshop-like queuing systems. *Management Science*, 10(1):131–142.

Jain, S. and McLean, C. (2003). A framework for modeling and simulation for emergency response. In *Proceedings of the 2003 Winter Simulation Conference*.

Jia, H., Ordonez, F., and Dessouky, M. (2007). A modeling framework for facility location of medical services for large-scale emergencies. *IIE Transactions*, 39(1):41–55.

Jiang, Y., Yuan, Y., Huang, K., and Zhao, L. (2012). Logistics for large-scale disaster response: Achievements and challenges. In *Proceedings of the 45th Annual Hawaii International Conference on System Sciences*, pages 1277–1285.

Kelton, W. D., Sadowski, R. P., and Sturrock, D. T. (2004). *Simulation with Arena*. McGraw Hill, New York.

Kerbache, L. and Smith, J. M. (1987). The generalized expansion method for open finite queuing networks. *European Journal of Operations Research*, 32:448–461.

Kovacs, G. and Spens, K. M. (2007). Humanitarian logistics in disaster relief operations. *Emerald*, 37(2):99–114.

Larson, R. C. (2004). O.r. models for homeland security. *OR/MS Today*, (10):1–10.

Larson, R. C., Metzger, M. D., and Cahn, M. F. (2006). Responding to emergencies: Lessons learned and the need for analysis. *Interfaces*, 36(6):486–501.

Law, A. (2007). *Simulation Modeling and Analysis*. McGraw Hill, New York.

Li, J. (2004). Performance analysis of production systems with rework loops. *IIE Transactions*, 36(8):755–765.

Li, J., Blumenfeld, D. E., Huang, N., and Alden, J. M. (2009). Throughput analysis of production systems: recent advances and future topics. *International Journal of Production Research*, 47(14):3823–3851.

Li, J., Meerkov, S., and Zhang, L. (2010). Production systems engineering: Problems, solutions, and applications. *Annual Reviews in Control*, 34(1):73–88.

Li, R. (2010). Research on emergency supplies demand forecasting model. In *2nd International Conference on Industrial Mechatronics and Automation*.

Lo, S. M., Huang, H. C., Wang, P., and Yuen, K. K. (2006). A game theory based exit selection model for evacuation. *Fire Safety Journal*, 41(5):364–369.

Maggio, N., Matta, A., Gershwin, S., and Tolio, T. (2009). A decomposition approximation for three-machine closed-loop production systems with unreliable machines, finite buffers and a fixed population. *IIE Transactions*, 41(6):562–574.

Mandelbaum, A. and Reiman, M. I. (1998). On pooling in queuing networks. *Management Science*, 44(7):971–981.

Manoj, V., Subodha, K., and Sushil, G. (2016). An integrated logistic model for predictable disasters. *Production and Operations Management*, 25(5):791–811.

Maron, M. J. (1987). *Numerical Analysis: A Practical Approach*. Macmillan Publishers Limited.

Martelo, M. A. J. (2011). *Resource Allocation Problems During Disasters: Points of Distribution Planning and Material Convergence Control*. PhD thesis, Rensselaer Polytechnic Institute.

McLoughlin, D. (1985). A framework for integrated emergency management. *Public Administration Review*, 45(1):165–172.

McNamara, T., Shaaban, S., and Hudson, S. (2016). Fifty years of the bowl phenomenon. *Journal of Manufacturing Systems*, 41(1):1–7.

Min, H. J., Beyeler, W., Brown, T., Son, Y. J., and Jones, A. T. (2005). Toward modeling and simulation of critical national infrastructure interdependencies. *IIE Transactions*, 39(1):57–71.

Miyazawa, M. and Taylor, P. G. (1997). A geometric product form distribution for a queuing network with non-standard batch arrivals and batch transfers. *Advances in Applied Probability*, 29(2):523–544.

Moler, C. and Loan, C. V. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45(1).

Nagarajan, M., Shaw, D., and Albores, P. (2012). Disseminating a warning message to evacuate: A simulation study of the behaviour of neighbours. *European Journal of Operational Research*, 220:810–819.

Natarajan, K. V. and Swaminathan, J. M. (2014). Inventory management in humanitarian operations: Impact of amount, schedule and uncertainty in funding. *Manufacturing and Service Operations Management*, 16(4):595–603.

Niessner, H., Rauner, M. S., and Gutjahr, W. J. (2017). A dynamic simulation-optimization approach for managing mass casulty incidents. *Operations Research for Healthcare*.

Ozbay, K. and Ozguven, E. E. (2008). Stochastic humanitarian inventory control model for disaster planning. *Transportation Research Record*, 2022:63–75.

Ozen, M. and Krishnamurthy, A. (2013). Models to evaluate layout and staffing policies at relief centers. In Krishnamurthy, A. and Chan, W., editors, *Proceedings*

of the 2013 Industrial and Systems Engineering Research Conference, San Juan, Puerto Rico.

Ozen, M. and Krishnamurthy, A. (2017). In-kind donations for disaster response: Interviews with practitioners. Technical report, University of Wisconsin Madison.

Ozpolat, K. and Rilling, J. (2015). Engaging donors in smart compassion: Usaid cidi's greatest good calculator. *Journal of Humanitarian Logistics and Supply Chain Management*, 5(1).

Pan American Health Organization and World Health Organization (2001). Humanitarian supply management and logistics in the health sector.

Papadopoulos, H. T. and Heavey, C. (1996). Queueing theory in manufacturing systems analysis and design: A classification of models for production and transfer lines. *European Journal of Operations Research*, 92(1):1–27.

Pel, A. J., Bliemer, M. C. J., and Hoogendoorn, S. P. (2012). A review on travel behaviour modelling in dynamic traffic simulation models for evacuations. *Transportation*, 39(1):97–123.

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

Pitana, T. and Kobayashi, E. (2009). Optimization of ship evacuation procedures as part of tsunami preparation. *Jorunal of Simulation*, 3:235–247.

Reibman, A. and Trivedi, K. (1988). Numerical transient analysis of markov models. *Computational Operations Research*, 15(1):19–36.

Reshetin, V. P. and Regens, J. (2003). Simulation modeling of anthrax spore dispersion in a bioterrorism incident. *Risk Analysis*, 6(23):1135–1145.

Roy, D., Krishnamurthy, A., and Bhat, S. (2011). *Humanitarian and relief logistics: Research issues, case studies and future trends*, chapter Performance trade-offs in layouts for distribution of supplies at relief centers, pages 21–30. Springer Series Operations Research/Computer Science(ORCS).

Royal, M., Jennings, D., Altmire, B., Dugan, E., Jones, K., and Reichow, S. (2014). Project responder 4: 2014 national technology plan for emergency response to catastrophic incidents. Technical Report 1.

Schulz, S. F. and Blecken, A. (2010). Horizontal cooperation in disaster relief logistics: benefits and impediments. *International Journal of Physical Distribution & Logistics Management*, 40(8/9):636–656.

Sheu, J. B. (2010). Dynamic relief-demand management for emergency logistics operations under large-scale disasters. *Transportation Research Part E: Logistics and Transportation Review*, 46(1):1–17.

Simonovic, S. P. and Ahmad, S. (2005). Computer-based model for flood evacuation emergency planning. *Natural Hazards*, 34(1):25–51.

Simpson, N. and Hancock, P. (2009). Fifty years of operational research and emergency response. *Journal of the Operational Research Society*, 60(1):126–139.

Smith, J. M. (1991). State dependent queuing models in emergency evacuation networks. *Transportation Research B*, 25B(6):373–389.

Smith, J. M. (1994). Application of state-dependent queues to pedestrian/vehicular network design. *Operations Research*, 42(3):414–427.

Smith, J. M., Cruz, F. R. B., and Woensel, T. V. (2009). Optimal server allocation

in general, finite, multi-server queuing networks. *Applied Stochastic Models in Business and Industry*, 26(1):705–736.

Sodhi, M. S. and Tang, C. S. (2014). Buttressing supply chains against floods in asia for humanitarian relief and economic recovery. *Production and Operations Management*, 23(6):938–950.

Song, X., Zhang, Q., Sekimoto, Y., and Shibasaki, R. (2014). Prediction of human emergency behavior and their mobility following large-scale disaster. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, volume 1.

Stapleton, O., Wassenhove, L. N. V., and Tomasini, R. (2010). The challanges of matching corporate donations to humanitarian needs and the role of brokers. *Supply Chain Forum*, 11(3):42–53.

Tanaka, K., Nagatani, T., and Hanaura, H. (2006). Traffic congestion and dispersion in hurricane evacuation. *Physica A*, 376(1):617–627.

Taskin, S. and Lodree, E. (2011). A bayesian decision model with hurricane forecast updates for emergency supplies inventory management. *Journal of the Operational Research Society*, 62:1098–1108.

Tcha, D., Lee, W., and Yamazaki, G. (1992). Server assignment for multi-stage production systems with finite buffers. *International Journal of Production Research*, 30(7):1637–1653.

The New Media Index ("2010 (accessed June, 2017)"). *Social Media Aid the Haiti Relief Effort*. http://www.journalism.org/2010/01/21/social-media-aid-haiti-relief-effort/.

Thomas, A. and Fritz, L. (2006). Disaster relief, Inc. *Harvard Business Review*.

Tomasini, R. M. and Van Wassenhove, L. N. (2009). From preparedness to partnerships: case study research on humanitarian logistics. *International Transactions in Operational Research*, 16(5):549–559.

Tregenza, P. (1976). *The Design of Interior Circulation*. Van Nostrand Reinhold, New York.

Ülkü, M. A., Bell, K. M., and Wilson, S. G. (2015). Modeling the impact of donor behavior on humanitarian aid operations. *Annals of Operations Research*, (1):153–168.

United States Agency for International Development (2005). Field operations guide for disaster assessment and response. Technical report.

United States Agency for International Development (2016). Guidelines for effective international disaster donations. Technical report.

Uno, K. and Kashiyama, K. (2008). Development of simulation system for the disaster evacuation based on multi-agent model using GIS. *Tsinghua Science and Technology*, 13(1):348–353.

Vuuren, M., Adan, I., and Resing-Sassen, S. A. E. (2005). Performance analysis of multi-server tandem queues with finite buffers and blocking. *OR Spectrum*, 27(1):315–338.

Wang, X., Li, F., Liang, L., Huang, Z., and Ashley, A. (2015a). Pre-purchasing with option contract and coordination in a relief supply chain. *International Journal of Production Economics*, 167:170–176.

Wang, X., Li, F., Liang, L., Huang, Z., and Ashley, A. (2015b). Pre-purchasing with option contract and coordination in a relief supply chain. *International Journal of Production Economics*, 167:170–176.

Wassenhove, L. N. V. (2006). Humanitarian aid logistics: Supply chain management in high gear. *Journal of the Operational Research Society*, 57(5):475–489.

Whitt, W. (1983). The queuing network analyzer. *The Bell System Technical Journal*, 62(9):2779–2815.

Woensel, T. V., Andriansyah, R., Cruz, F. R. B., Smith, J. M., and Kerbache, L. (2010). Buffer and server allocation in general multi-server queueing networks. *International Transactions in Operational Research*, 17(1):257–286.

Woensel, T. V., Wuyts, B., and Vandaele, N. (2005). Validating state-dependent queueing models for uninterrupted traffic flows using simulation. *A Quarterly Journal of Operations Research*, pages 159–174.

Wright, P. D., Liberatore, M. J., and Nydick, R. L. (2006). A survey of operations research models and applications in homeland security. *European Journal of Operational Research*, 36(6):514–529.

Wu, S., Ru, Y., and Li, H. (2010). A study on inventory management method in emergency logistics based on natural disasters. In *2010 International Conference on E-Product E-Service and E-Entertainment, ICEEE2010*, pages 1–4.

Xiaoping, Z., Tingkuan, Z., and Mengting, L. (2009). Modeling crowd evacuation of a building based on seven methodological approaches. *Building and Environment*, 44(1):437–445.

Xu, X., Qi, Y., and Hua, Z. (2010). Forecasting demand of commodities after natural disasters. *Expert Systems with Applications*, 37:4313–4317.

Yang, M., Allen, T. T., Fry, M. J., and Kelton, W. D. (2013). The call for equity: simulation optimization models to minimize the range of waiting times. *IIE Transactions*, 45(1):781–795.

Yates, D. and Paquette, S. (2011). Emergency knowledge management and social media technologies: A case study of the 2010 haitian earthquake. *International Journal of Information Management*, 31(1):6–13.

Y.Wang, Luangkesorn, L., and Shuman, L. (2012). Modeling emergency medical response to a mass casulty incident using agent based simulation. *Socio-Economic Planning Sciences*, 46(3):281–290.

Zhao, J. and Cao, C. (2015). Review of relief demand forecasting problem in emergency logistic system. *Journal of Service Science and Management*, 8(February):92–98.

Zou, N., Yeh, S., Chang, G., Marquess, A., and Zezeski, M. (2005). Simulation-based emergency evacuation system for ocean city, maryland, during hurricanes. *Transportation Research Record: Journal of the Transportation Research Board*, 1922:138–148.