Responsive Contingency Planning for Supply Chain Disruption Risk Mitigation

Anubhuti Parajuli

A Thesis

In the Department

Of

Mechanical, Industrial and Aerospace Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Industrial Engineering) at

Concordia University

Montreal, Quebec, Canada

July 2018

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Anubhuti Parajuli

Entitled: Responsive Contingency Planning for Supply Chain Disruption Risk Mitigation

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Industrial Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Andrea Schiffauerova

______________________________ External Examiner
Dr. Fantahun Defersha

______________________________ External to Program
Dr. Chun Wang

______________________________ Examiner
Dr. Ming Yuan Chen

______________________________ Examiner
Dr. Ali Akgunduz

______________________________ Thesis Co-Supervisor
Dr. Onur Kuzgunkaya

______________________________ Thesis Co-Supervisor
Dr. Navneet Vidyarthi

Approved by ______________________________
Dr. Ali Dolatabadi, Graduate Program Director

Monday, October 15, 2018 ______________________________
Dr. Amir Asif, Dean
Gina Cody School of Engineering and Computer Science

Abstract

Responsive Contingency Planning for Supply Chain Disruption Risk Mitigation

Anubhuti Parajuli, PhD

Concordia University, 2018

Contingent sourcing from a backup resource is an effective risk mitigation strategy under major disruptions. The production volumes and speeds of the backup resource are important protection design considerations, as they affect recovery. The objective of this dissertation is to show that cost-effective protection of existing supply networks from major disruptions result from planning appropriate volume and response speeds of a backup production facility prior to the disruptive event by considering operational aspects such as congestion that may occur at facilities. Contingency strategy are more responsive and disruption recovery periods can be shortened through such prior planning.

The dissertation focuses on disruption risk arising from intelligent or pre-meditated attacks on supply facilities. An intelligent attacker has the capability to create worst case loss depending on the protection strategy of a given network. Since the attacker seeks the maximum loss and the designer tries to identify the protection scheme which minimizes this maximum loss, there exists an interdependence between attack and protection decisions. Ignoring this characteristic leads to suboptimal mitigation solutions under such disruptions. We therefore develop a mathematical model which utilizes a game theoretic framework of attack and defense involving nested optimization problems. The model is used to decide optimal selection of backup production volume and the response speeds, the facilities to build such capability within the available budget.

The reallocation of demands from a disrupted facility to an undisrupted facility in a contingency strategy leads to congestion of the undisrupted facility, which may result in longer lead times and reduced throughput during disruption periods, thereby limiting the effectiveness of a contingency strategy. In the second part of the dissertation, we therefore analyze congestion effects in responsive contingency planning. The congestion cost function is modeled and integrated into the mathematical model of responsive contingency planning developed in the first part of the dissertation.

The main contribution of this dissertation is that a decision tool has been developed to plan protection of an existing supply networks considering backup sourcing through gradual capacity acquisition. The solution methodology involving recursive search tree has been implemented which allows exploring protection solutions under a given budget of protection and multiple combinations of response speeds and production capacities of a backup facility. The results and analysis demonstrate the value of planning for responsive contingency in supply chains subject to risks of major disruptions and provide insights to aid managerial decision making.

Acknowledgements

The academic journey I started some years ago at Concordia University has finally come to an end. In retrospect, I have many people to thank who have helped me accomplish this journey.

The deepest gratitude go to my supervisors Dr Onur Kuzgunkaya and Dr Navneet Vidyarthi for their patience, supervision and support. Your beliefs in me and standing by it until the end has helped me in this accomplishment, Thank you!

I am grateful to the members of my examination committee for their constructive feedbacks and suggestions. Thank you as well, the staffs of the MIAE department at Concordia University, I have received a lot of assistance from all of you whenever required. Thanks are due to the staffs and faculties at JMSB, where I had numerous opportunities to become a Teaching Assistant.

My office mates at EV 13.119, it was great knowing you all and sharing experiences, Thank you! My colleagues from Nepal, occasional coffee chats with you during breaks would remain unforgettable, Thanks!

I am thankful to Srijana, my better half, for her unconditional support, patience and solidarity in this journey. To the twins Ayushi and Arya, you have been the true joys throughout this journey, Thank you girls!

Above all, I am thankful to the god which I believe resides inside every human being - that inner power which keeps me modest and upbeat amidst difficulties. Thank you Lord!

Dedication

To my beloved parents,

To my better half,

To my twin daughters, Ayushi and Arya

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols

I	Set of customers (indexed by i)
J	Set of facilities (indexed by j)
B	Total available budget of protection in monetary units
c_{jl}	Unit cost of protection of a facility j at level l
m_{tl}	A multiplier representing the proportion of extra capacity available each time period t for every level of protection l
d_{ij}	Distance from customer i to facility j
h_{it}	Demand of customer i in time period t
v_{jt}	Base capacity of facility j in time period t
a_{jl}	Maximum additional capacity that can be added at facility j at protection level l
ß$_i$	Unit cost of lost sales for unserved demand of customer i
r	Allowed number of facilities to be interdicted
x_{ijt}	Units of demands from customer zone i served by facility j in time period t
s_j	Interdiction variable -1 if facility j is interdicted and 0 otherwise
z_{jl}	Protection variable -1 if facility j is protected at level l, 0 otherwise
u_{it}	Unmet demand units of customer i in time period t
w,q	Indices used to represent the multiplier and the power factor in the power law convex congestion cost function
k	Index representing flows where tangent line (linear) approximation of congestion costs are made

Chapter One: **Introduction**

1.1 Overview and Motivation

Supply chain disruptions are random occurrences of major discontinuities in supply chain operations as a result of natural disasters or intentional and unintentional human actions. Disruptions due to operational contingencies such as machine breakdowns and transportation/delivery delays are more likely to occur but have less severe economic impacts. Disruptions due to natural disasters (e.g. earthquakes, hurricanes, storms, etc.) or the ones due to intentional or unintentional human actions (e.g. terrorist attacks, strikes, fire, etc.), on the other hand, often lead to severe economic impacts to the supply chains, even though their occurrences are rare.

Several instances of high impact supply chain disruptions in the recent years come to notice: The Taiwan earthquake in February 2016 affected global electronic supply chain with an estimated 25 billion dollars revenue impact. In October 2011, production at several computer manufacturers in Asia were halted by catastrophic flooding of hard disk supply facilities located at major cities of Thailand. The 2010 eruption of a volcano in Iceland disrupted millions of air travelers and affected time-sensitive air shipments (Chopra & Sodhi, 2014). A destructive earthquake hit Japan in 2011. Toyota Motor Company halted production at twelve of its assembly plants, resulting in a production loss of 140,000 vehicles. In 2005, the U.S. Gulf Coast was hit by hurricane Katrina. Several warehouses and manufacturing plants were shut down in the aftermaths, and a severe disruption to the crude oil production occurred in the Gulf of Mexico amounting nearly 1.4 million barrels a day (The Economist, 2005). The air traffic suspension after September 2001 terrorist attacks in the US, led to a disruption of material flow into Ford's assembly plants causing intermittent shutdowns of its five US plants resulting in a 13% decline in its fourth-quarter production (Vakharia and Yenipazarli, 2009). In 1998, due to the parts shortages following the labor strikes in two of its US manufacturing plants, General

Motors (GM) closed its 26 assembly plants resulting in a production deficit amounting to $809 million in quarterly loss (Simison, 1998).

As the above examples illustrate, disruptions result in serious financial consequences for supply chain firms. When high impact disruptions occur, these firms may face several periods of reduced production output, critical parts shortages, or inefficient goods distribution. When response and recovery mechanisms are not adequate, disruption periods are prolonged and the impacts cascade through one echelon of the supply chain to the others. For instance, it took six months for Evonik, an automotive resin manufacturer in Marl, Germany to restore its production operations after the devastating explosion in one of its manufacturing plant in March 2012. As a result of which, the downstream production facilities of Ford and other automakers were consequently severely disrupted during this time (Simchi-Levi, Schmidt and Wei, 2014).

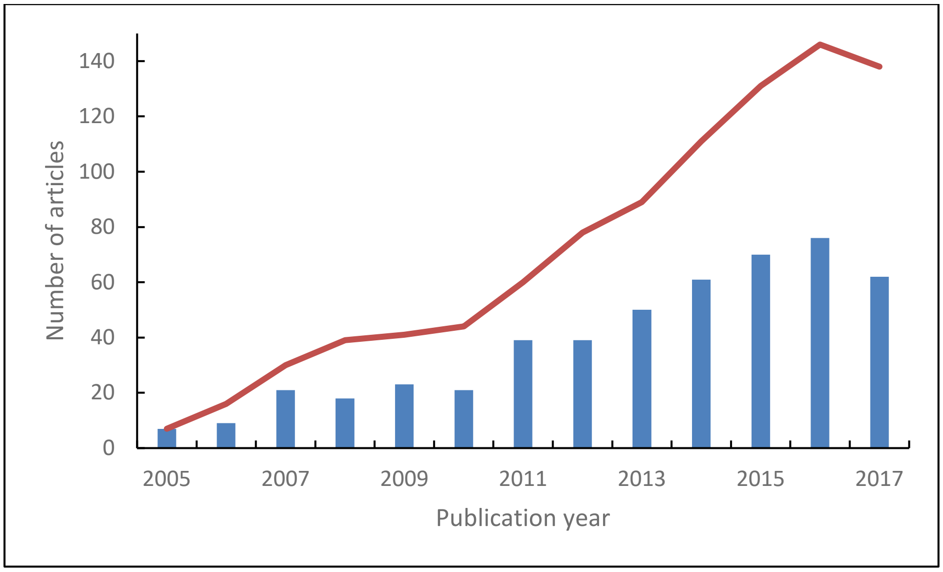


Figure 1.1: Cumulative and yearly peer-reviewed journal and review articles on supply chain disruption

(Source: Scopus database search with keywords “supply chain” and “disruption”)

Full recovery from high impact disruptive events is often difficult with firms eventually losing their market competitiveness and customers apart from the immediate revenue losses during downtimes. Empirical studies have shown that significant drop in sales along with diminishing stock returns and shareholder wealth for many years may be expected following major disruptions (Hendricks and Singhal, 2005). On the other hand, there are evidences where firms have gained competitive business advantage through better mitigation and management of disruption risks. For example, in 2001, shortages of semiconductor chips due to fire at Philips semiconductor plant in New Mexico (USA) resulted in Ericsson losing a significant market ($ 400 million in lost sales) as compared to its competitor Nokia because of its inferior mitigation planning (Latour, 2001).

Disruption risks have also grown in the recent years due to competitive business practices focusing on lower costs and leaner supply chains. The practices of just-in-time delivery, reduction of product life cycle, growth of global distribution channels and outsourcing make supply chains more complex and interdependent but leave little margin of error in operations. Risk management activities are however costly and firms need to investigate the trade-off of investments for capability improvements and risk reduction (Nooraie & Parast, 2016). These practical issues in managing risks, the unavoidability of disruptive events, and their impacts on supply chain operations, provide rationale for the study of supply chain disruption risk mitigation and management. A growing interest in this area is evident from the search we conducted with keywords “supply chain” and “disruption” utilizing Scopus database. The search results indicate an increasing trend in the journal articles and review publications dealing on supply chain disruptions, especially in the last decade (Figure 1.1). Neglected in the past by many firms, systematic planning and management of disruption risks due to major events such as terror attacks or catastrophic natural disasters is nowadays recognized as an important part of their business plans (Simchi-Levi, Snyder and Watson., 2002; Chopra and Sodhi, 2004; Sheffi, 2005).

Increased threat of terrorism worldwide and more sophisticated and well-devised techniques adopted by the adversaries suggest that new and effective mechanisms are required to ensure security and resilience of critical systems. The US Department of Homeland Security has identified 17 sectors of critical infrastructure where investments in structural and operational resilience should be made for vulnerability mitigation against acts of terrorism, other man-made threats and natural disasters. As terrorism or natural catastrophes cannot be prevented altogether, the quest for more effective risk mitigation strategies continues to motivate research in this direction.

1.2 Risk Mitigation Strategies

Strategies for mitigation and management of major disruption risks can be broadly classified as: a) proactive or preventive strategy and b) reactive or recovery strategy. The proactive or preventive risk mitigation strategies focus on appropriate plans and best course of actions which are adopted ahead of disruption occurrences so that the system is reliable, secured and suffers minimum loss during disruption periods. Such strategies are similar to increasing the mean time to failure (MTTF) of the machines or infrastructure systems. Preventive risk mitigation strategies follow practices of protecting supply flows through proactive redundancy measures such as acquiring redundant suppliers, inventory backup or protection (hardening) of supply facilities.

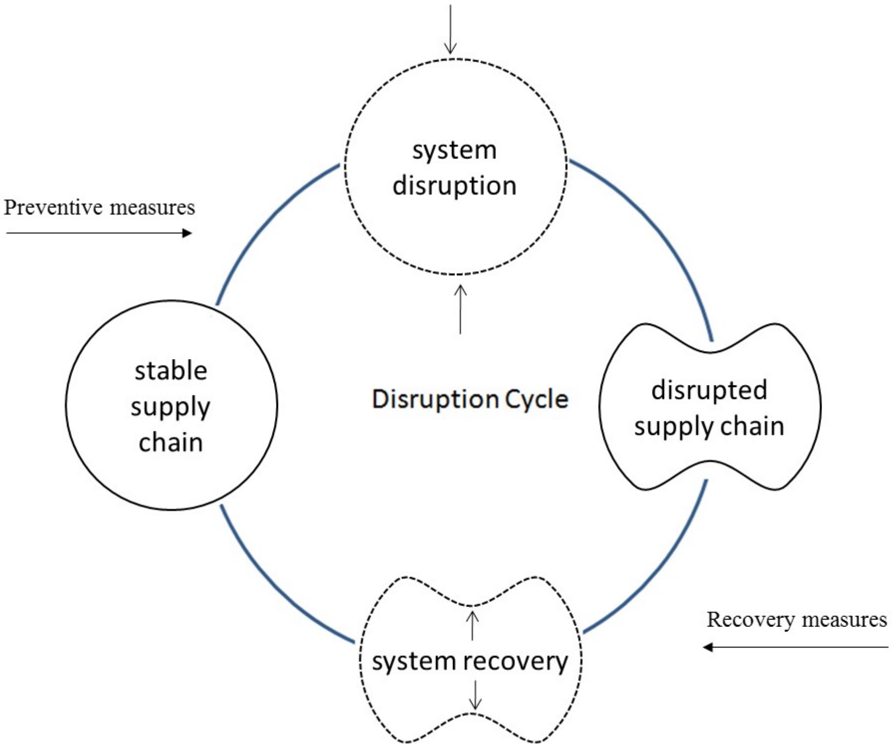


Figure 1.2: Disruption cycle with preventive and recovery stages

The reactive or recovery strategy is more concerned with the plans and course of actions following disruptions. The reactive strategy can be compared to reducing the mean time to repair (MTTR) of a failed machine or an infrastructure system. The recovery strategies ensure that the system transition from disrupted state to a stable state is fast, whereas preventive strategies ensure that it stays in the stable state for longer times, i.e., longer transition to disrupted states. This relation is depicted in Figure 1.2.

A common way to evaluate risk mitigation strategies is based on the measures of *system reliability, robustness, responsiveness* and *resilience*. These terms have been described in literatures with varying degrees of similarity. In this dissertation, we adopt the following definitions: *Reliability* is the ability of a supply chain to operate effectively even when parts of this system is disrupted (Snyder, 2005). *Robustness* is the ability of a supply chain to perform effectively over all possible future disruption scenarios including some worst scenarios (Klibi, Martel and Guitoni, 2010). *Responsiveness* is how quickly the supply chain can react/respond to disruptions (Klibi, *et. al*, 2010). *Resilience* is the ability of a supply chain to quickly recover from disruption (Klibi *et. al*, 2010). Resilience and responsiveness have often been used interchangeably because having a resilient supply chain equates to building its responsiveness capability, either through flexibility or redundancy measures.

From these definitions, it can be stated that reliability and robustness relate to preventive or proactive mitigation, whereas resilience and responsiveness relate to reactive or recovery mitigation strategy. As post-disruption recourse actions are limited, supply chain resilience and responsiveness cannot however, be enhanced without pro-active planning and positioning of recovery and contingent mechanisms. Such mechanisms may include provisions of inventory, back up production, redundant supplier, hardening (structural protection) of facilities etc. In other words, strategic and tactical level decisions such as where to hold inventory and its exact amount, where to have back up production and its volume or speed, which additional supplier to source from, which facilities to protect etc. should be implemented ahead of disruptions. The contingency strategies are more effective and response to disaster events are faster through such provisions.

This dissertation studies a responsive contingency planning problem in supply chain risk management, which involves pro-active protection plans and actions to enhance the effectiveness of a reactive or contingent operations enabling a reliable, robust, responsive and resilient supply flows capable of handling major disruptions.

1.3 Scope and Objectives

This research focuses on responsive contingency planning for mitigation and management of supply chain disruption risks. Achieving a faster and more effective post-disruption recovery operation is a major objective of such contingency planning. For example, holding strategic inventory ahead of disruptions and recovering disrupted flows through these inventories can be considered a responsive contingency approach of risk management as supply chains can react to disruptions faster through such provisions. Holding inventory for longer periods is however, cost prohibitive and is not appropriate for handling major disruptions that last for an extended period (Hopp, Iravani & Liu , 2012). Contingent capacity management through back up productions is a strategy that can be utilized under such disruptions. This strategy is more cost effective than strategic inventory since it does not lead to the accumulation of inventory because backup production can only be initiated after disaster occurrence. Capacity of a production facility can be contingently adjusted (ramped up) to partially recover the lost capacities due to disruptions or to partially/completely meet the re-routed demands from the failed facilities. This is especially facilitated in modern flexible or reconfigurable manufacturing system which can make quick capacity changeovers to adjust to the fluctuating demands (Putnik, Sluga, Elmaraghy, Teti, Koren, Tolio and Hon, 2013).

A major challenge in contingent backup capacity management is in having the desired units of backup production available within a short response time so as to improve disruption recovery speeds. Response time is dependent on the manufacturing system structure of a backup production facility. A scalable facility is able to quickly ramp up capacities in small increments, whereas a facility relying on dedicated equipment to reduce production cost will have a slower response time (Nejad, Niroomand, & Kuzgunkaya, 2014). Response speed is related to response time and determines how fast a facility can reach its desired level of production. Response speed and back up capacity volumes are critical decision factors in the selection and design of a backup production facility. Congestion is another factor that may affect capacity availability of a backup production facility during disruptions. The demands originally filled by a disrupted facility is shifted to a backup facility under a contingency strategy. This may create demand overload at the backup production facility despite its fast ramp-up characteristics. Consequently congestion of the backup facility result due to queuing which may affect the lead time and service levels of the supply system.

In this research, we study the problem of responsive contingency planning in supply chains in which post disruption recovery operations can be enhanced through pro-active protection of selected facilities and provisions of gradual backup production capacities in these facilities. System disruption is realized as intentional attacks on network facilities from a terrorist or an intelligent adversary, and therefore the analysis relate to this type of disruption. The main objective is to develop a decision tool for strategic and tactical decisions involving system security and backup capacity management along with operational decisions of a recourse action for handling major disruptions from intentional facility attacks or worst case scenarios. The specific goals can be stated as follows:

1. Determine which facilities in the existing supply network to secure and build the backup production capability

2. Determine the appropriate level of responsiveness of a contingency strategy through the selection of appropriate response speeds and capacity volumes of a backup production facility

3. Investigate the impacts of operational characteristic such as congestion on the contingent allocation and protection decisions.

1.4 Contributions

This dissertation makes several research contributions:

First, it proposes a modeling construct which is a unified framework for disruption recovery and infrastructure security planning in order to achieve a more reliable, robust, responsive and resilient supply chain design. By doing this, it extends the scope of protection models by adding recovery component to the earlier models which are mostly featured on decisions of security or facility hardening.

Second, it extends the scope of implicit enumeration algorithm as a solution methodology by enabling search tree branching based on multiple levels of protection. In earlier applications of this technique, this has been limited to a single level. Consideration of multi levels of protection increases the size of the search tree, nevertheless, it enhances the applicability of this technique to solving more complex problems.

Third, by considering the protection problem under budget constraint, it is demonstrated how the limited budget can be best utilized on security and backup production capability. As no distinction is made between the security (hardening) and backup capacity budgets, the decisions to add backup capacity to a facility also ensures its security. In other words the costs of security are assumed to be lumped into the costs of a backup capacity. Such an assumption is reasonable when contingent backup planning is prioritized over the security of the facility itself, since it does not restrict the investment to security at the expense of backup capacity investments.

Fourth, using different network topologies based on initial base capacity distributions, it is investigated whether a centralized or dispersed backup capacity is appropriate for a given network. It is shown that dispersed backups contribute more to risk diversification and loss mitigation in non-identical capacity network than in networks with identical distribution of initial capacity.

Fifth, congestion impacts on protection decisions are evaluated by explicit modeling of non-linear congestion costs in the objective function of the proposed tri-level game theoretic model. Piecewise linearization of the congestion cost function is developed to reformulate and solve this model as a linear problem.

1.5 Thesis Outline

The remainder of this thesis is organized as follows: Chapter 2 presents literature study. The relevant literature intersecting different research domains are developed and common features with existing gaps are identified. In Chapter 3, we study the responsive contingency planning problem. Section 3.2 formulates the problem and the key decisions to be made. Section 3.3 and 3.4 are concerned with the model formulation and solution methodology development respectively. Section 3.5 presents the numerical results and computational efficiency of the proposed algorithm. This Chapter ends with a summary in Section 3.6.

In Chapter 4, we analyze the congestion effects in responsive contingency planning. We present the congestion cost function and its linearization technique and the revised tri-level model under congestion effects along with the solution methodology in Section 4.2. In Section 4.3 results and analysis for this part of the dissertation is presented. Chapter concludes with a summary in Section 4.4. The Chapter 5 of the dissertation presents the conclusion and the future research avenues.

Chapter Two: **Literature Review**

2.1 Introduction

The related literature to this dissertation intersects research domains in critical infrastructure protection, location planning and supply chain risk management. The articles covered have a focus towards one or more risk mitigation aspects including reliable, robust, responsive or resilient system designs. Specifically, we review reliable facility location models in location planning domain, the interdiction-fortification models in critical infrastructure protection planning domain and contingent planning models in supply chain risk management. The focus of the literature study is to identify the protection or risk mitigation aspects considered and their modeling and solution approaches.

2.2 Strategic Design Models of Protection

The supply network vulnerability to disruption can be mitigated by considering risks during initial network designs. A large reduction in risk can generally be achieved through a relatively small increase in the costs of facility location when their disruption probabilities are accounted in the network design stages (Snyder and Daskin, 2007). The objective in such models is to achieve a reliable system that can perform at low cost both during disruptions and normal times.

The underlying facility location problem in most strategic design models is formulated either as a *p* median (Lee 2001, Snyder and Daskin 2005, Berman, Krass and Menezes 2007, Li, Zeng and Savachkin 2013) or a fixed charge location problem (Snyder and Daskin 2005, Lim, Daskin, Bassambo and Chopra 2010, Cui, Ouyang and Shen 2010, Li and Ouyang 2010, Shen, Zhan and Zhang 2011, Aboolian, Cui and Shen 2012, Li *et al.* 2013). In *p* median formulation, the number of facilities to be located is known (=*p*) and there is no fixed set up costs involved in locating these facilities. The expected total costs are formulated as expected costs of transportation, i.e., the expected costs of serving customers (e.g. retailers) through facilities (e.g. warehouses) measured as demand weighted distances from the customers to the facilities. In fixed charge location problems, number of facilities to be located is an endogenous decision as fixed costs are involved in locating facilities. The expected total costs are obtained as the sum of total fixed costs of facility location (independent of disruptions) and expected transportation costs under disruption risks.

The basic strategic design models involve decisions of opening a set of facilities, all of which have chances of being disrupted, i.e., all facilities unreliable (Lee 2001, Berman *et al.* 2007 and Berman *et al.* 2009). Some models have considered simultaneous locations of reliable and unreliable facilities (Snyder and Daskin 2005, Lim *et al.* 2010, Cui *et al.* 2010, Li and Ouyang 2010, Shen *et al.* 2011, Aboolian *et al.* 2012). Reliable facilities in these models are the ones that never get disrupted but their fixed costs of locations are higher than the unreliable facilities, otherwise there is no incentive in opening of unreliable facilities. Li *et al.* (2013) considers the problem of locating a set of unreliable facilities in which some unreliable facilities can be fortified (protected against disruption) utilizing a limited fortification budget. Therefore the decisions involve which facilities to open and which among them to fortify.

Strategic design models are based on the optimization of some measures of central tendency such as expected costs, expected profits, etc. which require explicit incorporation of disruption probabilities. Therefore, an implicit assumption in these models is that facility disruption probability are known or can be readily estimated. It is either incorporated as a scenario probability (q) or as individual facility failure probability *(p*) parameter. When scenario probability is utilized, disruptions are modeled as explicit scenarios, each scenario consist of a set of facility disruptions and a known probability. When disruptions are modeled with explicit scenarios, assignment of demands is based on these scenarios. Each scenario specifies which facility will be disrupted and which will be operational in the planning horizon. If a facility fails in a given scenario, demand is routed to the nearest other operational facility. Under scenario probabilities, the problem is formulated as a two stage stochastic programming model, where location decisions are made at the first stage before knowing which scenario will occur and the assignment of demands is made at the second stage after random disruptions occur. The solution of this is based on standard methods in stochastic linear programming (Higle, 2005).

When individual facility failure probability is known, specific demand assignment rules are defined to route demands from a disrupted facility to the nearest undisrupted facility. Snyder and Daskin (2007) consider a "level assignment" rule, in which facilities are arranged at several levels in increasing order of distance from the demands. If nearest facility to the demand fails, then reassignment is to the next nearest facility at a higher level. This assignment strategy allows allocation of demands to a primary facility under normal conditions, and to a set of backup facilities when the primary facility is disrupted. However, this approach of modeling disruption makes the model complex and intractable when disruption probabilities are non-uniform and the number of facilities are large. Therefore most models are formulated on the assumptions of identical disruption probability, i.e. all facilities have the same disruption probability. This assumption is relaxed in Cui *et al.* (2010) and Qi *et al.* (2010). Li and Ouyang (2010) extends this by considering correlation effects of facility disruption. The model considers that the probability of disruption of a facility is affected by the disruption of a nearby facility.

The design models utilizing probabilistic disruptions and central tendency measures focus on reliability improvements. Their solutions lead to network designs that may not perform adequately under extreme conditions imposed by a major disruptive event. More robust strategies and risk-averse approaches are required to manage major disruptions. A few authors have considered robustness in strategic location of facilities so that the network is protected from worst case losses under a major disruption (O'Hanley and Church 2011, Peng, Snyder, Lim and Liu 2011, Aksen and Aras 2012). Peng *et al.* (2011) considers opening facilities in the network so that the performance under disruption scenario does not deviate much from its performance under normal (non-disruption) scenario. They apply a criteria called *p*-robustness (Snyder and Daskin 2006), to optimize performance (costs) subject to a constraint requiring relative regret in each disruption scenario to be no more than a worst accepted *(p)* level. O'Hanley and Church (2011) utilizes a maximal covering formulation to locate facilities, which simultaneously maximizes pre-disruption coverage and post disruption worst-case coverage. The model is presented in a bi-level framework in which the post disruption worst-case coverage is modeled at the lower level in the form of an interdictor who attempts to minimize the coverage through attacks on located facilities. At the upper level facility location decisions are made to simultaneously maximize the pre and post disruption coverages. Similar to O'Hanley and Church (2011), Aksen and Aras (2012) also utilize a bi-level framework, with lower level used to model the worst case loss. However, the underlying formulation is a fixed charge location problem unlike the maximum covering formulation in O'Hanley and Church (2011). Their model additionally incorporates fortification and capacity expansion decisions integrated with the facility location decisions. Strategic design models involving protection or fortification decisions have been studied most recently in Bricha and Nourelfath (2013) and Jalali, Seifbarghy and Niaki (2018). These models are different to other models in that they apply a different concept of contest success function (Hausken, 2011) to model independent facility disruptions. While both apply game theoretic modeling approach to seek optimal protection strategy in an uncapacitated fixed charge location supply network, the location decisions are made ahead of protection decisions in Bricha and Nourelfath (2013) and are based on expected costs (utilities) measures (risk neutral), Jalali *et al.* (2018) considers protection decisions in the design stage by integrating the two decisions together and takes a risk averse approach.

It is important to note that while facility location is a major decision in all strategic design models, their applicability is limited to the planning of new networks. In existing networks facility relocation is not a viable option due to associated costs. Another stream of research focuses on protection of systems that are already existing and for which relocating facilities is cost prohibitive. This research stream specifically addresses network and infrastructure security and vulnerability mitigation and models are commonly referred as interdiction and fortification models. The next section discusses interdiction and fortification models of protection.

2.3 Interdiction and Fortification Models of Protection

Interdiction can be defined as a deliberate or intentional attack on the critical elements of a network system to disrupt or deteriorate its performance. Interdiction models have been applied in the literature to assess system vulnerability to disruptions or to identify the most critical elements of the system, whose loss deteriorate the performance the most. Fortification should be understood as a mechanism to enhance protection of such critical system components so that system disruption can be controlled. Fortification of facilities or infrastructure may involve investments for structural reinforcements, for example, seismic designs to protect against earthquakes, structural barriers to control flood, etc. It can also be achieved through some redundancy in the system, for example strategic stocks or inventory, backup resource, multiple sourcing, offshoring business, etc. The mathematical models that apply protection of the system and its components through these fortification measures are commonly known as fortification models.

2.3.1 Interdiction models

Interdiction models identify critical facilities (facility interdiction models) or network arcs (network interdiction models) whose disruption can create the greatest loss of system efficiency. Network interdiction were the earliest problems studied and involve removal of arcs from a network commonly involving objectives of minimizing maximum flow between origin and destination (Wollmer, 1964) and maximizing the shortest paths between supply and demands (Fulkerson and Harding 1977, Israeli and Wood 2002), etc. A survey of network interdiction models and their variants can be found in Church, Scaparra, & Middleton (2004).

In the context of facilities, the interdiction models are concerned with the identification of critical facilities whose failures represent worst-case system losses (Church *et al.* 2004; Church and Scaparra 2007a, 2007b; Losada *et al.* 2010a; O'Hanley and Church 2011), Losada *et al.* 2012a). The models have been based on two classical facility location models: a) *p* median b) max covering. The interdiction models with max covering type formulation (*r* interdiction covering problem) involve decisions of identifying a subset (*r*) of existing facilities whose interdiction/disruption can result in a maximal coverage loss. In *p* median based interdiction models (*r* interdiction median (*r*IM) models) the decisions involve removing/disrupting *r* number of facilities from the existing set of facilities so as to maximize the total cost of reassignment of demands, i.e., the demand weighted distances to the operational facilities post interdiction.

Church *et al.* (2004) presents MIP formulation of both *r* interdiction covering and the median models. The parameter *r* is deterministic in these models. Church and Scaparra (2007a) and Losada *et al.* (2012a) extended the study by considering *r* as probabilistic. Losada *et al.* (2010a) introduced recovery time aspects in the *r*IM model. Their study demonstrate that worst case losses may be underestimated if recovery times are ignored by the models, especially if the impacts associated with prolonged disruption are significant.

2.3.2 Fortification models

The worst case losses from a major disruption can be avoided through protection of critical facilities identified by the interdiction model. While planning defense under intentional attacks (e.g. terror attacks) however, the solutions provided by the interdiction models are not always reliable. This is because intelligent adversaries can adjust their actions to circumvent the defender’s strategy. This relationship between attack and defense needs to be accounted when planning protection against such disruptions (Brown, Carlyle, Salmeron and Wood 2006).

The shortcoming of the interdiction model is addressed through fortification models which integrate protection decisions into the mathematical models of interdictions. The models take risk averse approach to risk management as fortifications imply protection against possible worst-case loss. These models are commonly prescribed in a game theoretic framework to capture the elements of dependence and often cast as bi-level optimization models involving attack and defense. One of the first models in this direction was proposed by Church and Scaparra (2007b) and is called the *r*-interdiction median model with fortification (*r*IMF) in which *r* represents the number of facilities that can be interdicted or attacked. At the upper level of this bi-level model, the protection decision involves which of the *q* facilities to protect by optimal allocation of limited protection budget, and at a lower level the decision involves identification of the *r* most critical facilities of an existing *p* median network.

The mathematical formulation of the *r*IMF model (Church and Scaparra, 2007b) is illustrated here as the model developed in this dissertation is based on similar formulation:

$$[\boldsymbol{r}\textbf{IMF}]\text{: minimize } H(z) \tag{2.1}$$

subject to

$$\sum_{j\in J} z_j = q \tag{2.2}$$

$$z_j \in \{0,1\} \quad \forall j \in J \tag{2.3}$$

where $H(z)$ represents the *r*IM model, i.e., the lower level interdiction problem represented as:

$$[\boldsymbol{r}\textbf{IM}]\text{: } H(z) = \text{maximize} \sum_{i\in I} a_i d_{ij} x_{ij} \tag{2.4}$$

$$1 - s_j \geq z_j \qquad \forall j \in J \tag{2.5}$$

$$\sum_{j\in J} x_{ij} = 1 \qquad \forall i \in I \tag{2.6}$$

$$\sum_{j\in J} s_j = r \tag{2.7}$$

$$\sum_{k\in T_{ij}} x_{ik} \leq s_j \qquad \forall i \in I, j \in J \tag{2.8}$$

$$s_j \in \{0,1\} \qquad \forall j \in J \tag{2.9}$$

$$x_{ij} \in \{0,1\} \qquad \forall i \in I, j \in J \tag{2.10}$$

(notations: *I*- set of demand nodes indexed by *i, J*- set of existing facilities indexed by *j*, a_i- demand of node i, d_{ij}- distance between node *i* and *j*, *T*$_{ij}$ – set of facilities other than the closest facility to *i* , *r*- number of facilities to be interdicted)

The decisions variables are defined as:

$$z_j = \begin{cases} 1 & \text{if facility } j \text{ is fortified} \\ 0 & \text{otherwise} \end{cases}$$

$$s_j = \begin{cases} 1 & \text{if facility } j \text{ is interdicted} \\ 0 & \text{otherwise} \end{cases}$$

$$x_{ij} = \begin{cases} 1 & \text{if demand } i \text{ is assigned to facility } j \\ 0 & \text{otherwise} \end{cases}$$

The worst case loss is represented in the objective function of the lower level *r*IM problem (2.4) which maximizes the weighted distance of demands to facilities through the interdiction of *r* unprotected facilities. Corresponding to this, the objective function of the *r*IMF problem (2.1) at the upper level minimizes this loss by optimally protecting *q* facilities. The integrality of protection variables, interdiction variables and the demand assignment variables are represented in (2.3), (2.9) and (2.10) respectively. Constraint (2.2) states the cardinality of protected facilities. Constraint (2.5) links the upper level problem to the lower level problem and prohibits the attack of protected facilities. Constraint (2.6) states that every demand should be assigned to exactly one facility. Constraint (2.7) restricts the number of attack to a maximum of *r* facilities. Constraint (2.8) restrict the allocation of demand to a farther facility if the closer one is not interdicted.

Several variants to the basic fortification model (*r*IMF) have subsequently been studied. Liberatore, Scaparra, & Daskin (2011) and Liberatore & Scaparra (2011) considered uncertainty in the number of possible facility attacks (i.e., *r* is probabilistic). In Liberatore *et al.* (2011) the bi-level *r*IMF with probabilistic *r* is reformulated into a max-covering type problem. The problem is reduced to a single level which minimizes the expected worst case coverage across all possible values of *r*. Considering the difficulty associated with properly estimating the probability of the extent of attacks, Liberatore & Scaparra (2011) utilize a worst case regret as a performance measure rather than the expected costs measures utilized in Liberatore *et al.* (2011). The regrets considered represent the maximum deviations of a stochastic solution with unknown number of attacks to the solutions if these probabilities were known or deterministic. Losada, Scaparra, & O'Hanley (2012b) further introduce temporal dimensions to the basic *r*IMF model through facility recovery time and multiple disruption considerations. Partial disruptions were introduced through modeling correlation effects of a facility loss (attack) in Liberatore, Scaparra, & Daskin (2012) . The percentage of capacity lost by a facility from disruption of a neighbouring facility was represented in a correlation matrix which was utilized in determining facility capacities after disruptions. Partial disruptions were also modeled in Aksen, Akca & Aras (2014).

It is observed that a majority of the protection models have assumed that facilities are uncapacitated. Further, most of the models can be considered as static models of protection since they do not incorporate temporal aspect of recovery. Table 2.1 summarizes relevant protection planning articles on the features of facility (capacitated or uncapacitated), the risk tolerance considered (risk averse/risk neutral) and the implied mitigation strategy (preventive/recovery). The strategic design models are largely based on risk neutral approach. Such models take a preventive risk mitigation focus as their decisions involve facility location or facility location with some redundancy placement decisions e.g. inventory, backup supplier, structural reinforcements etc. so that the supply network is inherently reliable (Berman *et al.* 2007, Cui *et al.* 2010, Lim *et al.* 2010, Qi *et al.* 2010, Li and Ouyang 2010, Aboolian *et al.* 2013, Li *et al.* 2013, Bricha and Nourelfath 2013). A majority of models that take risk averse approaches also have a preventive mitigation focus and ignore temporal aspect of recovery and contingent mechanisms (Church and Scaparra 2007, Scaparra and Church 2008, Aksen *et al.* 2010, Liberatore *et al.* 2011, Liberarotore and Scaparra 2011, Aksen and Aras 2012, Liberatore *et al.* 2012, Scaparra and Church 2012, Jalili *et al.* 2018). Although under a major disaster, these models lead to a more robust supply flows than the risk neutral models because of their worst case loss considerations, they are still inadequate from a resilient and responsive design point of view due to the exclusion of temporal aspect of recovery in these models.

Losada *et al.* (2012b) were the first to introduce recovery time dimension in a protection model. The protection is implied as a decision to invest or allocate available budget into facilities for reducing their recovery times following disruption. A drawback of this model is that it does not explicitly model contingent mechanisms or how the recovery can be enhanced. As well, the model is less realistic as it considers uncapacitated facilities. Aksen, Piyade, & Aras (2010) and Aksen & Aras (2012) incorporate contingent mechanism in a capacitated model which involves capacity expansion decisions for contingent rerouting of demands originally handled by a disrupted facility. However, they assume that such capacity expansions occur instantaneously. The response time in building these capacities are not considered. Therefore their model overestimates actual available capacity during disruption periods.

Although lacking in the domains of reliable facility location or critical infrastructure protection planning literature, response and recovery aspects of contingent planning have been considered in supply chain risk management literature. The relevant literature are discussed in the next section.

Table 2.1 Relevant Literature Classification

Articles	Facility Features		Risk Features		Mitigation Strategy				
					Preventive				
	Capacitated	Uncapacitated	Risk Neutral	Risk Averse	Inventory Backup	Capacity Backup	Facility Hardening	Facility location	Recovery
Snyder and Daskin (2005)		√	√					√	
Berman *et al*. (2007)		√	√					√	
Church and Scaparra (2007b)		√		√			√		
Jeon *et al*. (2008)		√	√		√			√	
Scaparra and Church (2008a, 2008b)		√		√			√		
Aksen *et al*. (2010)		√		√		√	√		
Cui *et al*. (2010)		√	√					√	
Lim *et al.* (2010)		√	√				√	√	
Li and Ouyang (2010)		√	√		√			√	
Qi *et al.* (2010)		√	√		√			√	
Liberatore *et al.* (2011)		√		√			√		
Liberatore and Scaparra (2011)		√		√			√		
O'Hanley and Church (2011)		√						√	
Peng *et al*. (2011)		√	√					√	
Aksen and Aras (2012)	√			√		√	√	√	
Liberatore *et al*. (2012)	√			√			√		
Losada *et al.* (2012b)		√		√					√
Mak and Shen (2012)		√	√		√			√	
Scaparra and Church (2012)	√			√			√		
Aboolian *et al.* (2012)		√	√					√	
Bricha and Nourelfath (2013)		√	√				√	√	
Li *et al.* (2013)		√	√				√	√	
Jalali *et al.* (2018)		√		√			√	√	

2.4 Contingent Planning Models of Protection

Several contingent mechanisms of disruption exist, such as inventory, dual sourcing, contingent re-routing and backup capacity. The seminal work by Tomlin (2006) considers backup capacity and inventory measures for handling demand fluctuations and random disruptions. Their model identifies inventory as an appropriate mechanism for frequent and short disruptions of the main supplier while dual sourcing is optimal for rare and long or major disruptions. Whenever a backup supplier has a flexible capacity, contingent re-routing is a preferable mechanism. In considering the flexible backup supplier, it is assumed that the whole backup capacity is available only after a response time and there is no supply from the backup capacity during the response time. Hopp and Yin (2006) and Schmitt (2011) also make the same assumption as Tomlin on backup capacity availability during the response time. On a different premise, Klibi and Martel (2012) propose a discrete stepwise function to represent the gradual capacity availability based on the intensity of disruption and time to recovery. A two echelon supply chain with production facilities involving dedicated manufacturing system (DMS) and a reconfigurable manufacturing systems (RMS) is considered in Niroomand *et al.* (2012). The RMS is considered as a volume flexible backup resource with partial availability of the capacity within the response time.

Response speed of backup production is a critical decision determining the effectiveness of protection strategy utilizing contingent capacity adjustments. Wang & Koren (2012) present several backup facility configurations affecting the supply chain responsiveness levels in a cost and response time trade-off analysis. In a serial configuration, the added capacity can only become available after completing the reconfiguration process of all stages. This makes the transition to required capacity slow despite the low cost of reconfiguration and the use of simpler machine structure. On the other hand, in a pure parallel configuration, each machine could go under a reconfiguration process independently, which leads to a faster transition speeds. A higher configuration cost may result with parallel configuration since machine structure is more complex with each machine required to perform all steps.

2.5 Conclusion

The literature review suggests that there have been quite limited study related to protection planning of supply chain considering the response and recovery mechanisms. Furthermore, most of the study have considered uncapacitated systems and ignore congestion effects due to overflow during recovery. Models based on such assumptions are less practical and lead to unrealistic protection solutions although they are mathematically more tractable. The challenges in solving capacitated and congested network models have yet not been addressed adequately in the literature. The fortification models with responsiveness consideration under a capacitated facilities and flow congestion effects are non-existent, despite the importance of such models in managing major disruption risks.

Relying on the interdiction-fortification framework discussed above, in this dissertation we therefore develop a new mathematical model for contingency planning under major disruption. In the first part of the dissertation we introduce *r*IMF type formulation of the protection model in which contingency strategy is realized through a backup production with response time considerations (Chapter 3). In the second part of the dissertation (Chapter 4), we demonstrate that congestion related costs influence protection decisions and reformulate the model to incorporate congestion costs and analyze its impact on protection decisions.

Chapter Three: **Responsive Contingency Planning under Disruption**

3.1 Introduction

This chapter introduces the responsive contingency planning model developed for managing disruption risks of a capacitated supply network. The disruption contingencies can be implemented more effectively when proper planning is done and appropriate mechanisms are identified ahead of the disaster events. Contingent capacity adjustments through back up production is proposed as mechanism to enhance recovery. The level of responsiveness of this contingent mechanism relies on optimal selection of production volumes and response speeds of a backup production. The appropriate level of responsiveness need to be determined ahead of disruption so that backup production capability can be designed accordingly and implemented during disaster periods.

The developed model provides a unified framework for planning security and recovery in a supply network subject to premeditated attacks on its facilities. Supply chain strategic and tactical level decision problems involving facility security and backup capacity management are solved while the operational level demand assignments are executed in an optimal manner. Relying on a game theoretic modeling framework to capture the elements of dependence between attacks and defense observed in intentional attacks (Brown *et al.*, 2006), the mathematical model is formulated as a tri-level mixed integer optimization problem. A solution algorithm based on implicit enumeration of defense strategies is proposed to arrive at decisions involving (i) which facilities to protect with backup production capability (ii) what should be the volume of a backup production (iii) what should be the response speed of a backup production. The material used in this chapter is from Parajuli *et al.* (2017).

The remainder of this chapter is organized as follows: Section 3.2 and Section 3.3 respectively present the problem and the model formulation. Section 3. 4 discusses the solution methodology. In Section 3.5 numerical results and analysis is presented. The chapter ends with a summary in Section 3.6.

3.2 Problem Formulation

Consider a network of multiple capacitated facilities supplying a set of customers with a single product (Figure 3.1). Let I be the set of customer zones. Each customer zone $i \epsilon I$ has a specific product demand h_{it} in time period t in the planning horizon T. The demands are satisfied from the existing set of facilities J, each characterized by a maximum supply capacity v_j . Let d_{ij} represent the distances involved in transporting a unit demand to customer zone $i \epsilon I$ from facility $j \epsilon J$. These distances are proxies for unit costs of transportation. We assume that the facilities are subject to disruptions in which all of its existing capacity is lost for the entire recovery period lasting a finite number of time periods $t \epsilon T$. If a facility is disrupted, the demands of the customer zones it originally served are rerouted to the next nearest operational facility with adequate capacity to accommodate such demands. Demands are split among neighboring facilities if a single facility is not capable of fulfilling all of these demands. Unmet demands due to inadequate supply capacity are considered lost in the system and incur the cost of lost sales.

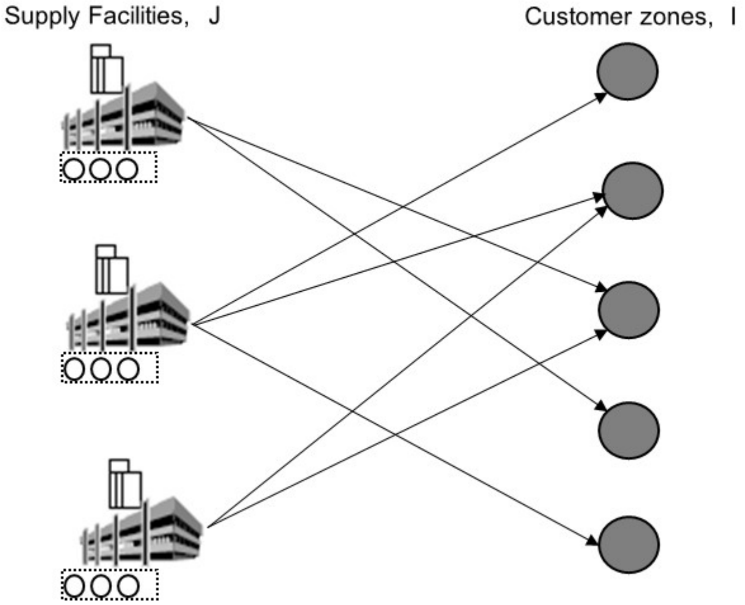


Figure 3.1: Single echelon supply network problem illustration

Disruption to the system is modeled as an attack on facilities of the supply network by an intelligent attacker who has prior information of the system and is capable of causing maximum damage to the system (worst-case). A worst-case system disruption always results from attacks if none of the system components (facilities) are protected because the attacker can deploy his budget to attack the most vulnerable sets of facilities. System operating costs are maximized through such attacks, either because the customer demands need to be assigned to more distant facilities, or due to lost sales incurred owing to inadequate system capacity, or both.

To counter attacks and the possible losses, the system planner with limited protection budget *B* designs protection of the network through building security and enhancing recovery. A protection cost c_{jl} is involved in securing (fortifying) a facility *j* with backup production capability at a certain level *l*. The levels of backup capability are determined by the nominal volume of production capacities and response speeds of such capacity additions. It is further assumed that fortification of facilities is linked to backup production decisions. In other words, the decisions to add capacity backups on a facility also implies its fortification. The protection design therefore involves identification of the facilities to be fortified, selecting the optimal volume and response speeds of capacity backups in fortified facilities, and the contingent re-assignment of customer demands to the surviving facilities in order to thwart the impacts of a worst-case disruption.

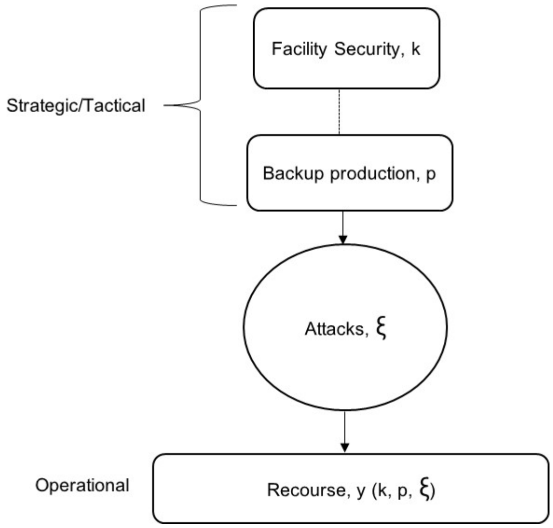


Figure 3.2: Decision stages of the proposed problem

As demonstrated in Figure 3.2, the protection design relates to solving strategic to tactical level decisions which include securing facilities and deciding on their backup production capabilities for contingent capacity adjustments during disruptions. The operational decision is the recourse action or the contingent allocation of demands which is a function of facility security and the backup production capability, since these two factors together determine the effectiveness of a contingency strategy.

3.3 Tri level D-A-D model for responsive contingency planning

The responsive contingency planning problem for managing disruptions risks of intentional attacks is formulated as a tri-level optimization model within a game-theoretic framework. The model conceptually involves a sequential game amongst three players at different levels of hierarchy: i) supply chain planner (*system defender, D*) at the top level determines the facilities to be protected with back up production capability to minimize the worst-case losses due to disruptions; ii) the interdictor (*attacker, A*), at the middle level, identifies the set of facilities that can be attacked to create the worst-case losses from disruptions; and iii) supply chain operator (*system user, D*) identifies the most cost effective way of operation post attacks. Figure 3.3 provides a schematic of the modeling framework of the proposed protection planning problem.

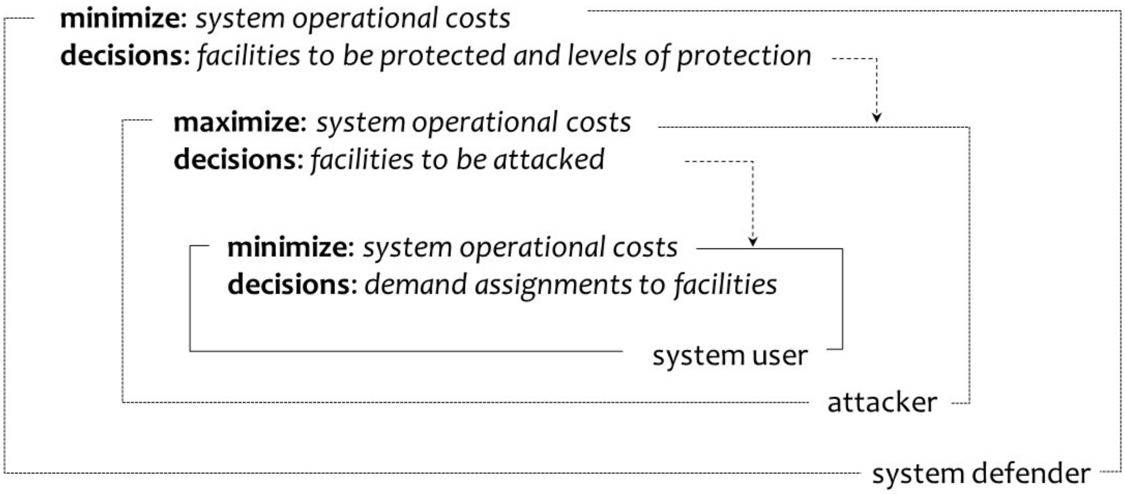

minimize: system operational costs
decisions: facilities to be protected and levels of protection
maximize: system operational costs
decisions: facilities to be attacked
minimize: system operational costs
decisions: demand assignments to facilities
system user
attacker
system defender

Figure 3.3. Tri-level D-A-D model framework

The notations of the model are listed in Table 3.1

Table 3.1 Notations used in Responsive Contingency Planning Model Formulation

Sets and Parameters:	
I	set of customers (indexed by *i*)
J	set of facilities (indexed by *j*)
B	total fortification budget
c_{jl}	cost of protection of facility *j* at level *l*
m_{tl}	a multiplier representing the proportion of extra capacity available each time period during the response time and after, based on selected response speeds of facilities
d_{ij}	distance from customer zone *i* to facility *j*
h_{it}	demand of customer *i* in time period *t*
v_{jt}	base supply capacity of facility *j* in time period t
a_{jl}	maximum additional capacity at facility *j* for corresponding fortification level *l*
β_i	unit cost of lost sales for unserved demand from customer *i*
r	number of (facility) interdictions
Decision variables:	
x_{ijt}	demand quantities from customer zone *i* served by facility *j* in time period *t*
s_j	1 if facility *j* is interdicted and 0 otherwise
z_{jl}	1 if facility *j* is fortified with capacity backups at level *l*, 0 otherwise
u_{it}	total unmet demand of customer *i* in time period *t*

The decisions made at the upper levels are parameterized at the lower levels. Mathematically, this nested decision framework is represented as a hierarchical mixed integer optimization problem as follows.

$$\textbf{[DLP]:} \min_{z} \; H(z) \tag{3.1}$$

subject to:

$$\sum_{j\in J}\sum_{l\in L} c_{jl} z_{jl} \le B \tag{3.2}$$

$$\sum_{l\in L} z_{jl} \le 1 \qquad \forall j \in J \tag{3.3}$$

$$z_{jl} \in \{0,1\} \qquad \forall j \in J, l \in L \tag{3.4}$$

where,

[ALP]: $H(z) = \max_{s} G(s,z)$ (3.5)

subject to:

$$\sum_{j \in J} s_j = r \qquad (3.6)$$

$$s_j + \sum_{l \in L} z_{jl} \leq 1 \qquad \forall j \in J \qquad (3.7)$$

$$s_j \in \{0,1\} \qquad \forall j \in J \qquad (3.8)$$

where,

[ULP]: $G(s,z) = \min_{x,u} \left(\sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} d_{ij} x_{ijt} + \sum_{i \in I} \sum_{t=1}^{T} \beta_i u_{it} \right)$ (3.9)

subject to:

$$\sum_{j \in J} x_{ijt} + u_{it} = h_{it} \qquad \forall i \in I, t = 1 \ldots T \qquad (3.10)$$

$$\sum_{i \in I} x_{ijt} \leq (1 - s_j) \left(v_{jt} + \sum_{l \in L} m_{tl} a_{jl} z_{jl} \right) \qquad \forall j \in J, t = 1 \ldots T \qquad (3.11)$$

$$x_{ijt} \geq 0 \qquad \forall i \in I, j \in J, t = 1 \ldots T \qquad (3.12)$$

$$u_{it} \geq 0 \qquad \forall i \in I, t = 1 \ldots T \qquad (3.13)$$

The decision framework comprises three optimization problems. The system planner or the defender level problem (DLP) is represented by Equations (3.1) – (3.4). This part of the problem models the defense of the supply network from worst-case attacks. The decisions at this level are represented by binary variables z_{jl} which is 1 if a facility j is protected at level l and 0 otherwise (constraint 4). Levels of protection ($l = 1 \ldots L$) represent a selected combination of capacity and response speeds that define backup production capability and are therefore dependent on the available sizes (volumes) of capacity and response speeds for capacity backups; for instance, 2 levels of capacity and 2 levels of response speed will result in four levels of protection ($L = 4$).

The problem represented by Equations (3.5) – (3.8) is associated with the attacker level problem (ALP). The decisions at this level are represented by binary variables s_j which are set to 1 if the facility is attacked and 0 otherwise. The ALP is concerned with the maximization of system loss by controlling the variables s_j. Finally the problem represented in Equations (3.9)–(3.13) is associated with the user level problem (ULP) where the decisions are represented in the non-negative flow variables x_{ijt} and a dummy variable representing the unmet demands u_{it}.

The decisions made at the DLP problem are parameterized in the ALP problem. Similarly, the decisions made at the DLP and ALP problems are parameterized in the ULP problem. The objective of the system planner is to protect the system by minimizing the maximum system operational cost the attacker can create (Equation 3.1). The vector of protection strategy, $Z = (z_{11}, z_{21}, \ldots z_{JL})$, corresponds to a vector of investments costs $C = (c_{11}, c_{21}, \ldots. c_{JL})$. Hence the system planner is constrained by protection budget B available to him. Further, a facility can only be protected at one level.

The objective of the attacker contradicts to that of the system planner. The attacker targets a set of unprotected facilities in order to maximally raise the system operational costs through his attacks (Equation 3.5). Constraint (3.6) defines the number of facilities that can be simultaneously attacked. The DLP problem is linked to the ALP problem through constraint (3.7). It prohibits the attack on protected facilities.

Following protection and attacks, the system operator seeks a minimum cost assignment of demands to the remaining supply facilities. This is represented in the objective function of the ULP (Equation 3.9). This objective function comprises two terms: the first term represents the transportation (flow) costs of the demands that are met (*cFlow*), and the second term represents the cost of lost sales (*cLS*) if the system capacity is inadequate to completely fill the demands. Any demands that cannot be met in a given time period due to insufficient system capacity are accounted as lost sales units in constraint (3.10). Constraint (3.11) specifies that the total demands handled by each facility in each time period cannot exceed the available total capacity of that facility. The assumption here is that a facility if protected has its total capacity equal to its base (original) capacity plus the backup capacity, whereas if a facility is attacked, it loses all of its base capacity and no demands can be assigned to it. Gradual capacity addition in each time period is reflected in the parameter m_{tl}, which is the proportion of selected capacity size (volume) that can be added in time period t for a selected response speed or a selected level of protection.

3.4 Solution Methodology

In order to identify facilities to be protected and to obtain the optimal capacity and response speeds of a backup resource in a contingency strategy, a tree search algorithm is presented. This algorithm explores optimal solution of the proposed model through a binary tree search procedure.

3.4.1 Background

Game theoretic attack-defense modeling framework involving bi-level optimization has been widely applied for optimizing protection of critical infrastructure with applications in supply chains, telecommunications, electric power grids, railways and pipeline networks, etc. This modeling framework is suitable for solving resource allocation problems in order to counter strategic risks such as malicious attacks (Golany, Kaplan, Marmur and Rothblum, 2009). Bi-level optimization problems are however, difficult to solve especially when they involve integer decisions at both levels (Moore and Bard, 1990). This is because of the nested structure which makes the solution of the lower level problem a function of the upper level problem and the solution of the upper level problem a function of the lower level problem. Church and Scaparra (2007b) applied the bi-level optimization framework to plan for fortification of supply facilities which would minimize the worst case losses due to attacks on a finite number of facilities of the supply chain network. The difficulty in solving the bi-level problem is handled in this model through reformulation into more tractable single level mixed integer linear programming (MIP) problem which is solved using the general purpose commercial MIP solver. The limitation of the approach is that only a small sized instances can be solved using this approach since it requires seeking integer decisions of protection and demand allocations through explicit enumeration of attack scenarios. Scaparra and Church (2008b) develop an alternative MIP formulation and exploit the mathematical structure of the reformulated problem to obtain the lower and upper bounds which is used to reduce the size of the original model. Decomposition techniques involving cutting plane algorithms such as Benders decomposition (Losada *et al.*, 2012b) and duality techniques (Wood, 1993) which involves taking dual of the inner problem to formulate it in a nested min-min or max-max structure are other common approaches of solving these type of problems.

One of the widely used approach for handling these problems is due to Scaparra and Church (2008a) who apply implicit enumeration algorithm tailored to the bilevel structure of their interdiction fortification problem. The conjecture of this algorithm is that at least one of the candidate facilities in the worst case attack should be protected for minimizing the impacts of such attacks. Implicit enumeration algorithm utilizes this conjecture in a recursive search tree to find the optimal protection strategy. The main advantage of this approach is that it does not face the size restrictions or complicated reformulations as in the previous algorithms of Church and Scaparra (2007b) and Scaparra and Church (2008b). This approach has subsequently been used in Aksen *et al.* (2010), Cappanera & Scaparra (2011), Scaparra & Church (2012) and Liberatore *et al.* (2011, 2012).

Our solution methodology is an extension of the implicit enumeration algorithm (IE) in Scaparra & Church (2008a). The search process in our algorithm is more extended than that of Scaparra & Church due to the considerations of different levels of facility protection. A more important difference is that while Scaparra & Church solve a mixed integer problem (MIP) at every child nodes of the search tree to identify attacked facilities at the lower level, we solve all of the problems as LP. This is possible in our algorithm because we arrive at the solution of the attacker's problem (ALP) by independently solving all of the lower level problems (ULP) for each attack scenario. This means attack decisions are inputs to our ULP problem, as a result of which it involves only continuous decision variables. Given the fact that computational effort of the IE approach depends on the difficulty of solving the MIP in the lower level interdiction problem (Scaparra & Church, 2008a), our approach does not encounter similar difficulty other than the need to enumerate feasible attack patterns at each node. This however does not severely deteriorate the computational performance of our algorithm. The computational performance is reported in Section 3.5.6. In the following sections we provide the details of our solution methodology.

3.4.2 Algorithm description

In the proposed algorithm, the search tree starts by creating a root node where facilities involved in worst-case attack of *r* facilities is identified. The worst case attack is the attack on facilities when there is no protection/fortification involved and the damage to the system is maximum. In other words, at the root node we obtain solutions for the attacker level problem (ALP) when the system is totally unprotected and hence the attacker is able to create maximum loss in the system through the attacks on selected facilities.

A pseudo code for implementing root node algorithm is presented in Table 3.2. At its initialization, all the protection variables z_{jl} in the root node are set to zero and all the combinations of attack scenarios involving *r* facility attacks are enumerated. Every attack scenario *p* involves a vector of facility attacks (s_j). For every attack scenario *p*, the user level problem (ULP) can be solved using a commercial LP solver. Note that solving ULP does not involve any integer decisions as the attack and protection decisions are parameterized at this level. After solving ULP for all attack scenarios, the corresponding values of total transportation costs and the total lost sales costs are normalized with respect to their ranges. Steps 6-10 in Table 2 illustrate normalization of costs after ULP is solved for all attack scenarios. Normalization scales the two cost components (cost of lost sales and the transportation or the flow costs) between 0 and 1, so that any decision bias due to scale differences in their absolute values are avoided.

In this algorithm, the ALP problem is not solved explicitly but its solution is obtained by solving several instances of ULP problems, which are linear programming problems (LP) with all non-integer decision variables. The ALP solution is obtained through a sequence of steps: a) solution of multiple instances of ULP (as many as the number of feasible attack scenarios) b) normalization of costs across all scenarios and c) identification of attack scenario with highest normalized total cost. The attack scenario leading to maximum normalized total costs is considered the worst-case attack scenario, and hence the ULP solutions for this scenario are also the solutions of the ALP problem for this node.

Table 3.2 Root Node Algorithm

Pseudo-code*: Solving attacker problem (ALP) at root node*

```
$\forall\, j, l \;\; z_{jl} \leftarrow 0$
enumerate $P = {}^{n}C_{r}$ attack combinations
for $p = 1 \;\ldots.., P$ \\ attack scenarios
  Solve ULP for $(S^{p}, cLS^{p}, cFlow^{p})$
end for
$mincLS \leftarrow$ minimum$(cLS^{1}, \ldots.\, cLS^{P})$; $minFlow \leftarrow$ minimum$(cFlow^{1}, \ldots.\, cFlow^{P})$
$maxcLS \leftarrow$ maximum$(cLS^{1}, \ldots.\, cLS^{P})$; $maxFlow \leftarrow$ maximum$(cFlow^{1}, \ldots.\, cFlow^{P})$
for $p = 1 \ldots P$
  $normcLS^{p} \leftarrow \left(\frac{cLS^{P} - mincLS}{maxcLS - mincLS}\right)$; $normcFlow^{p} \leftarrow \left(\frac{cFlow^{P} - mincFlow}{maxFlow - minFlow}\right)$ \\ normalization of costs
  $normTotCost^{p} \;\leftarrow normcLS^{p} + normcFlow^{p}$ \\ normalized total cost
end for
$maxnormTotCost \leftarrow$ maximum$(normTotCost^{1}, \ldots.\, normTotCost^{P})$
$S^{*} \leftarrow \{\, S^{p}\!: normTotCost^{p} = maxnormTotCost\}$;
$cLS^{*} \leftarrow \{\, cLS^{p}\!: normTotCost^{p} = maxnormTotCost\}$;
$cFlow^{*} \leftarrow \{\, cFlow^{p}\!: normTotCost^{p} = maxnormTotCost\}$;
return $(S^{*},\; cLS^{*},\; cFlow^{*})$ \\ optimal ALP solution at the root node
```

The facilities attacked in the worst-case attack scenario (S^*) obtained as a solution to the ALP problem in the root node, should be the constituents for protection if worst-case loss is to be avoided. This is as per the observation made in Scaparra & Church (2008). The rationale for this is that for avoiding worst-case loss, one of the facilities from the attacked member set in the worst-case attack has to be protected. If this is not the case, then the attacker is always free to attack facilities in this set and create maximum loss. The enumeration tree therefore proceeds from the root node by binary branching on protection variables, which are one of the facilities j from among the candidate sets (S^*) identified in the root node solution. The flow chart of the search tree is illustrated in Figure 3.4.

In the left branch of the enumeration tree the decision is to protect facility j at a selected level of protection. The facility to protect is selected arbitrarily from the candidate sets in the root node. As protection costs vary depending on the selected protection level, it is necessary to compute the remaining budget at each node before branching from it. If the remaining budget is inadequate to protect any of the candidate facilities, then this node is fathomed and becomes a leaf node (i.e., a node without any child node). If the budget is adequate for protecting a selected facility, sub-branches are created along this branch for each allowable protection level. The sub-branch is pruned if this level of protection cannot be achieved. For example, if there are four levels of protection available, depending on the budget available there can be up to four sub-branches each leading to a child node. In each of these sub-branches, the corresponding z_{jl} variable is set to 1, which indicates that the selected facility j is protected at a level l along that branch. The ULP is then solved at each of these child nodes by iteratively calling a commercial LP solver for each feasible attack scenario. Feasible scenarios are all the different combinations of attack scenarios involving r facility attacks from a set of n unprotected facilities. Once the ULP is solved for all feasible attack scenarios, the two cost components are normalized and the worst-case attack scenario is identified as the one leading to the maximum total normalized costs. Identification of the worst case attack scenario provides new sets of candidate facilities to be protected in the next stage.

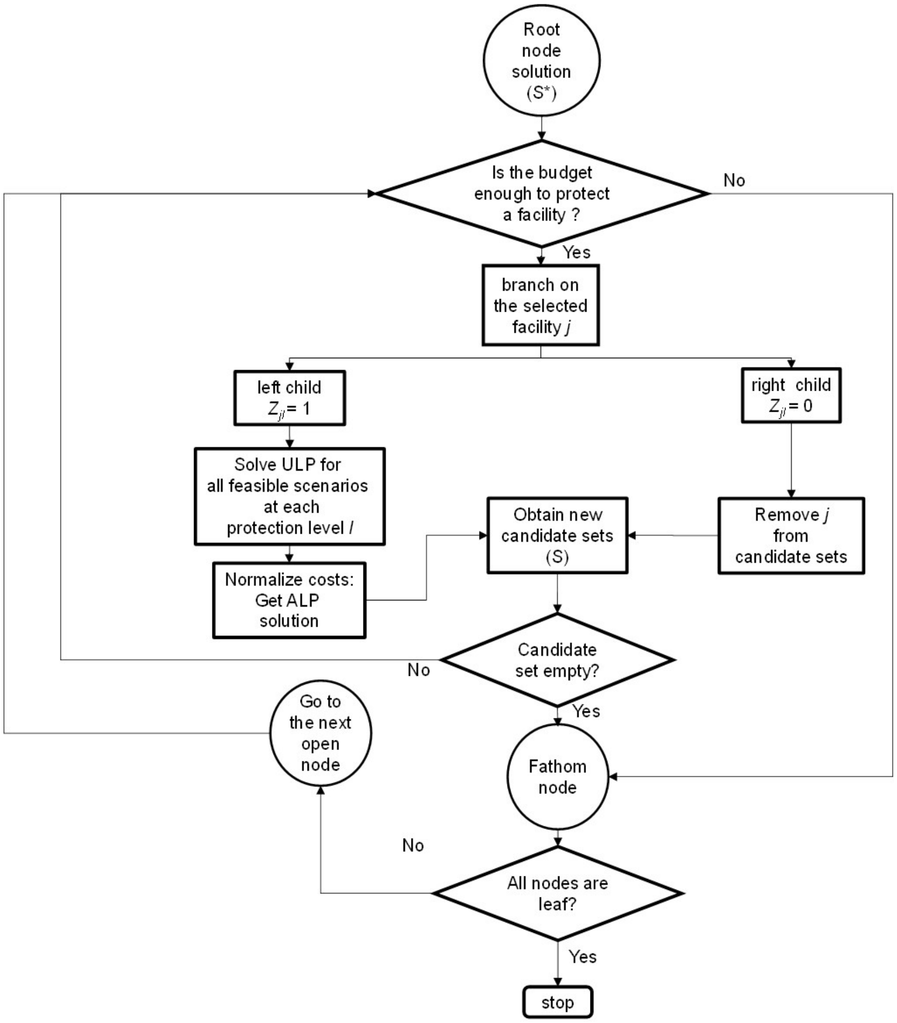

Root node solution (S*)
Is the budget enough to protect a facility ?
No
Yes
branch on the selected facility j
left child Z_jl = 1
right child Z_jl = 0
Solve ULP for all feasible scenarios at each protection level l
Normalize costs: Get ALP solution
Obtain new candidate sets (S)
Remove j from candidate sets
Candidate set empty?
No
Yes
Go to the next open node
Fathom node
No
All nodes are leaf?
Yes
stop

Figure 3.4: Flowchart of the search tree

The search tree progresses according to a depth-first strategy: at every node, arbitrarily selecting facilities to protect from the candidate sets; creating new child nodes for every level of protections that the budget allows; and fathoming those nodes with inadequate budget for further protection. At every unfathomed child node that follows, the LP solver is called iteratively for solving ULP with all feasible attack scenarios, taking into account the facilities protected until this node. After solving ULP for all attack scenarios, costs are normalized and solutions to the ALP problem is obtained. The size of the feasible attack scenarios for solving ULP reduces with the depth of the search tree. This is because more number of facilities are protected further down the tree and therefore cannot be attacked.

In the right branch to every parent node, a child node is reached by first setting all protection variables corresponding to facility *j* protected on the left branch to zero (i.e., $z_{jl} = 0$ for every *l*) and updating the candidate sets for protection by eliminating facility *j* from it. If the updated set is empty, this child node becomes a leaf node. Otherwise, branching from this node is continued in the aforementioned manner. The tree search terminates when there are no nodes remaining for further branching in any of the branches (i.e., all nodes are leaf nodes). This can happen for two reasons: either the candidate sets for protection are empty, or the budget is inadequate for further protection of a candidate facility identified in the parent node.

At the termination of the search tree, the costs of lost sales (*cLS*) and the transportation costs (*cFlow*) obtained as an optimal ALP solution at each leaf nodes are normalized to scale these costs between 0 and 1 by comparing all leaf node solutions (refer to Section 4.2.2). The leaf node with smallest normalized total cost is selected as the optimal solution to the tri-level problem. The optimal sets of facilities to be protected and their corresponding levels of backup capacity and response speeds are obtained by backtracking the path from that node to the root node.

3.4.3 Illustration of the proposed methodology

The proposed methodology is illustrated by generating a binary tree to solve a simple problem with five facilities, and ten demand zones. The five facilities are located in five states in the US (NY, CA, IL, TX and PA). Among these facilities, two facilities are to be interdicted by the attacker ($r = 2$). Two levels of capacity volumes (high, low) and two levels of response speed levels (high, low) are considered, the combination of which leads to four different levels (*l*) of protection. The levels of protection are designated as: level 1—high capacity volume with high response speeds; level 2—high capacity volume with low response speeds; level 3—low capacity volume with high response speeds; and level 4—low capacity volume with low response speeds.

The available protection budget (B) is assumed to be just sufficient to allow capacity additions up to two facilities when the lowest levels of available capacity volume and response speeds are selected. The costs of protection have been considered to be independent of facilities but are dependent on the selected volume and response speeds of capacity backups.

The illustrative problem highlights the branching and pruning rules, cost data normalization and identification of optimal defense strategy through backtracking.

3.4.3.1 Branching, pruning and nodes traversal

The enumeration tree corresponding to this illustrative problem is depicted in Figures (3.5a) – (3.5c). Every node of this enumeration tree is characterized by the following facility and costs data:

- Candidate sets of facilities for protection in the next stage (***S***): this includes a set of facilities attacked in the optimal attack, which maximizes the value of total normalized costs.
- Cost of lost sales (***cLS***), which is the absolute numerical value of total lost sales costs (i.e., computation of expression $\sum_{i \in I} \sum_{t=1}^{T} \beta_i u_{it}$ under the optimal attack scenario for this node).
- Flow costs (***cFlow***), which is the absolute numerical value of transportation (flow) costs (i.e., computation of expression $\sum_{i \in I} \sum_{j \in J} \sum_{t=1}^{T} x_{ijt} d_{ij}$ under the optimal attack scenario for this node).

Additionally, for creating new branches and for progression of the search tree, the algorithm needs to keep track of the following information at every node along every branch:

- Sets of facilities protected up to the current node with their levels of protection (i.e., list of z_{jl} variables that are set to 1 in this branch until the progression to this node).
- Remaining budget after protection of facilities until this node in this branch.

At the root node of the tree in Figure (3.5a), the attacker problem is solved without any facilities being protected by the defender. The worst-case attack plan is obtained for $S^{*}=\{CA, TX\}$. Among these two facilities, the facility at *CA* is arbitrarily selected for protection. The available budget is enough to protect *CA* at any of the 2, 3, or 4 levels of protection except level 1 for which the budget falls short. So three new nodes *B*, *C* and *D* are created in this branch corresponding to protection of *CA* at levels $l = 2$, 3 and 4 respectively. Note that node A is never reached due to insufficient budget for $l = 1$ level of protection, and this branch is therefore pruned (shown with a zigzagged line). The attacker level problem is solved at each of the nodes *B*, *C*, and *D* given that the facility at *CA* is protected at level 2 for node *B*, at level 3 for node *C* and at level 4 for node *D*. This gives rise to $S = \{IL, TX\}$ for each of the nodes *B*, *C* and *D*.

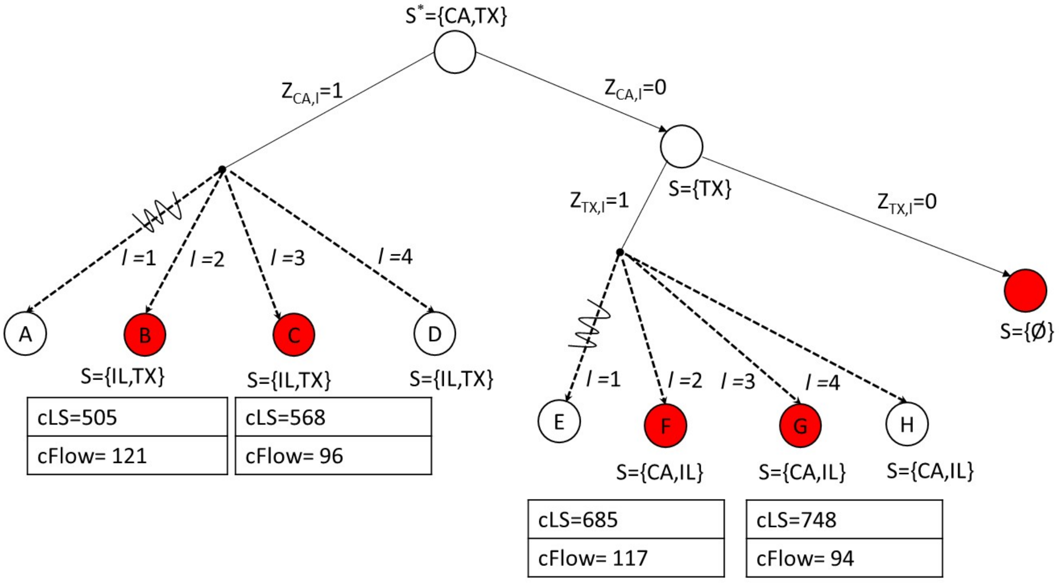


Figure 3.5a: Root node and initial branching of the search tree

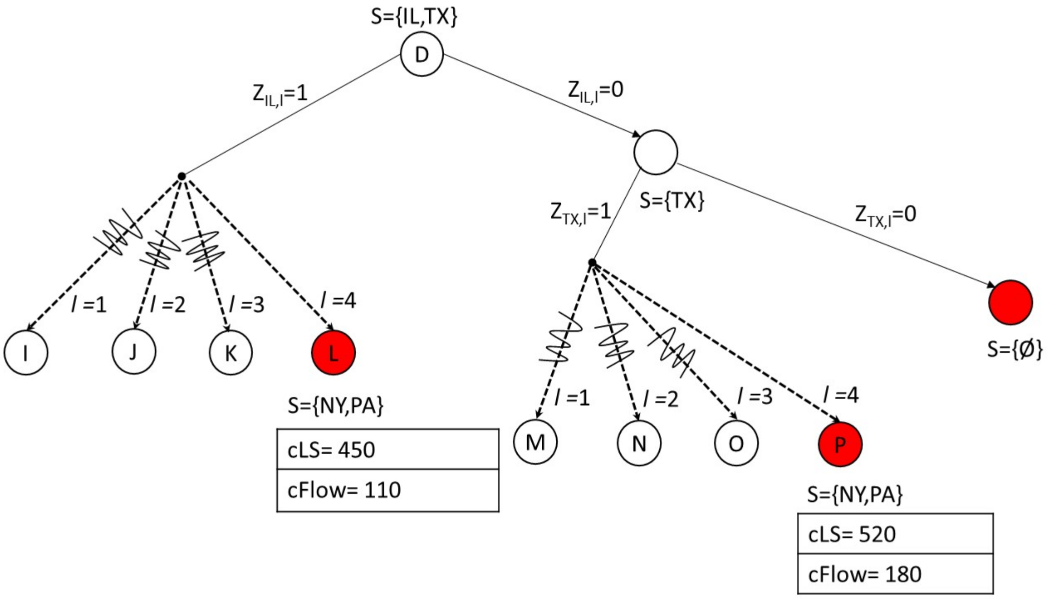


Figure 3.5b: Continuation of search tree from left branch of the root node

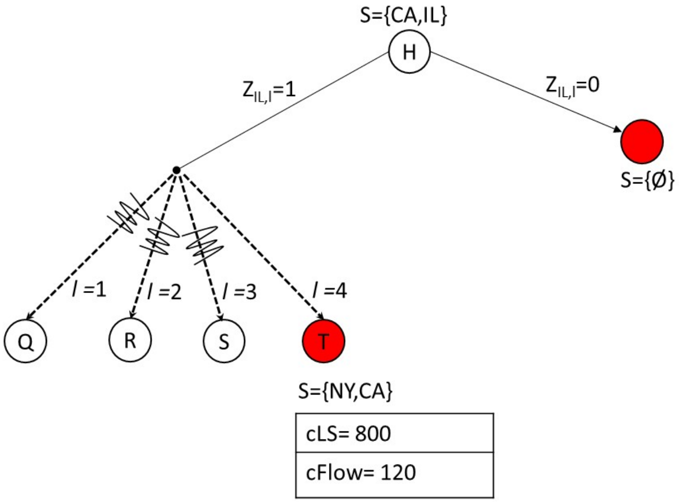


Figure 3.5c: Continuation of search tree from right branch of the root node

The other child node obtained from the root node corresponds to the branch $Z_{CA,l} = 0$ (i.e., the facility at *CA* is not protected at any level *l* which leaves $S = \{TX\}$. Since no budget is used so far in this branch, the available budget is adequate to protect *TX* at levels 2, 3, 4 but not 1. Hence the branching is continued from this node. On the left branch corresponding to $Z_{TX,l} = 1$, three new nodes *F*, *G* and *H* are created corresponding to the three levels of protection that can be attained. Node *E* is never reached due to insufficient budget for protection of the *TX* facility at level $l = 1$, hence this branch is pruned. The right branch from this node corresponds to $Z_{TX,l} = 0$ (i.e., the facility at Texas is not protected at any level and thus this branch leads to a node with $S = \{Ø\}$. Hence this node is fathomed.

The remaining budget at nodes *B*, *C*, *F* and *G* in Figure 3.5a are insufficient to protect further facilities. So these nodes become leaf nodes at this stage. For nodes *D* and *H*, the remaining budget allows further protection of a facility at $l = 4$. Figures 3.5b and 3.5c show the continuation of branching from these nodes. At node *D*, the facility at *IL* is arbitrarily selected for protection along the left branch. The available budget only allows a level $l = 4$ protection for this facility, which leads to the node *L* where the attacker problem can be solved by setting $Z_{CA,4} = Z_{IL,4} = 1$. This leads to $S = \{NY,PA\}$; however, the remaining budget is insufficient for further protection of a facility. Hence node L becomes a leaf node. The right branch from node *D* corresponds to $Z_{IL,l} = 0$, which leads to a node with $S = \{TX\}$. Continuing branching on this node leads to a leaf node P with $S = \{NY, PA\}$ on the left branch and a fathomed node with $S = \{Ø\}$ on the right branch.

The node *H* with $S = \{CA, IL\}$ is reached from the path with $Z_{CA,l} = 0$. Hence the protection of *CA* is not allowed in node *H* or any other child node along this path. This leaves *IL* as the only candidate for branching from node *H*. The available budget is sufficient to protect IL at level $l = 4$. The left branch from this node, corresponding to protection of *IL*, leads to node *T* with $S = \{NY, CA\}$. Node *T* becomes a leaf node as the available budget becomes insufficient for further protection at this stage. The right branch from node H leads to another fathomed node with $S = \{Ø\}$, because both *CA* and *IL* cannot be protected in this branch.

At its termination this enumeration tree results in seven leaf nodes—*B*, *C*, *F*, *G*, *L*, *P* and *T*—each with a unique values of *S, cLS* and *cFlow* obtained as ALP solution in these nodes.

3.4.3.2 Normalization of costs data at the leaf nodes

In order to obtain the optimal solution of the defender, costs data for the seven sets of leaf nodes are normalized with respect to the range of values obtained. Normalized values of lost sales costs ($normcLS_n$), flow costs ($normcFlow_n$) and the total costs ($normTotal_n$) for each node n in the set of all the leaf nodes N are obtained using the formulas in Equations (3.14)–(3.16).

$$normcLS_n = \left(\frac{cLS_n \text{ - } \min\limits_{n\in N}\left\{cLS_n\right\}}{\max\limits_{n\in N}\left\{cLS_n\right\} \text{ - } \min\limits_{n\in N}\left\{cLS_n\right\}} \right) \tag{3.14}$$

$$normcFlow_n = \left(\frac{cFlow_n - \min\limits_{n\in N}\left\{cFlow_n\right\}}{\max\limits_{n\in N}\left\{cFlow_n\right\} - \min\limits_{n\in N}\left\{cFlow_n\right\}} \right) \tag{3.15}$$

$$normTotal_n = normcLS_n + normcFlow_n \tag{3.16}$$

The normalized costs at the leaf nodes for this example are presented in Table 3.3.

Table 3.3 Values of normalized costs obtained at the leaf nodes of the IE tree

Nodes	***cLS***	***cFlow***	**norm*cLS***	**norm*cFlow***	**normTotal**
B	505	121	0.16	0.31	0.47
C	568	96	0.34	0.02	0.36
F	685	117	0.67	0.27	0.94
G	748	94	0.85	0.00	0.85
L	**450**	**110**	**0.00**	**0.19**	**0.19**
P	520	180	0.20	1.00	1.20
T	800	120	1.00	0.30	1.30
Min	450	94	0	0	0.19
Max	800	180	1	1	1.30

3.4.3.3 Selection of optimal defense strategy

The minimum normalized total costs is obtained for the leaf node *L* (Table 3.3). Therefore this node is selected as the optimal solution for the defender. Backtracking the tree from this node to the root node, we obtain facilities at *IL* and *CA* as optimal sets of facilities that can be protected within the available budget. Further, it can be observed that the protection budget is best utilized by adding low level of capacity and low response speeds to both the facilities (i.e., both facilities receive level $l = 4$ of protection. The optimal defense strategy is thus to protect two facilities (*IL* and *CA*) with low levels of capacity and response speeds. As a result of this protection, the attacker will target to attack facilities at *NY* and *PA* which will lead to a maximum system operational costs.

3.5 Results and Analysis

This section reports and discusses computational results obtained with the proposed solution methodology. The algorithm was coded in Java and all the problem instances were solved using ILOG CPLEX 12.6 solver (using Concert Technology) on a Dell Latitude E5430 station with an Intel Core i5-3340M processor at 2.7 GHz and 8 GB of RAM running Windows 7 operating system.

3.5.1 Problem instance generation

The test problems for numerical analysis are derived from the data of the largest metropolitan areas (by population) according to the US Census Bureau for 2000 (Daskin, 2004). The demands are proxy to population and are obtained by dividing the population of the cities by 10^3 rounded to the nearest integer. The original network is constructed first by ranking customer zones on the basis of its population (demands) size and opening of J facilities in these zones in the order of their ranking. In the problems considered, the demands and facility base capacities are held constant for every time periods. The unit costs of transportation from a customer demand zone i to facility location j is considered to be proportional to the distances and is presented in Appendix 1. The unit cost of lost sales are set at 2% higher than the maximum distances of all facility-demand pairs, calculated as: 1.02* *max* (d_{ij}). This ensures that lost sales are incurred only if system capacity is inadequate to handle all of the demands.

3.5.1.1 Facility base capacity

Two different supply networks are considered which have the same total system capacity but differ in their distributions of initial (base) facility capacities. In the first network, every facility has the same initial capacity every time period. This is computed by dividing the sum of total demands and built-in slack (idle capacity) by the total number of facilities (3.17). This relation means the original network is always capable of meeting the demands fully if all of the facilities are functioning. The parameter α in (3.17) represents the idle capacity of the system (i.e., system capacity in excess of total demands).

$$v_{jt} = \left(\frac{(1+\alpha)\sum_{i=1}^{I} h_{it}}{J} \right) \qquad \forall\, j \in J,\ t \in T \qquad (3.17)$$

The second network (network 2) has non-identical facility capacities but the same total system capacity as of network 1. Initial facility capacities for this network are assigned such that every open facility is able to completely satisfy demand of its nearest customer zone. For every time period *t*, the base facility capacity for a facility *j* in this network is therefore computed as total demand of its nearest customer zone with a finite increment λ_t (3.18).

$$v_{jt} = \left(h_{kt} : k, i \in I,\ d_{kj} = \min_i d_{ij} \right) + \lambda_t \quad \forall\, j \in J,\ t \in T \qquad (3.18)$$

where λ_t is calculated as:

$$\lambda_t = \left(\frac{(1+\alpha)\sum_{i=1}^{I} h_{it} - \sum_{j=1}^{J} \left(h_{kt} : k, i \in I,\ d_{kj} = \min_i d_{ij} \right)}{J} \right) \qquad (3.19)$$

Any remaining system capacity after satisfying the nearest demands to every facility are evenly distributed amongst all open facilities. This equally distributed remaining system capacity is represented by parameter λ_t in Equation (3.19).

3.5.1.2 Backup capacity volumes and response speeds

It is assumed that finite capacity sizes are available for backup capacity additions at low (1000 units) and high levels (2000 units). The amount of these backup capacities available during the recovery phase depends on response speeds of capacity additions. Capacity additions at higher speeds is assumed to take two time periods, whereas at slower speeds three time periods are required to add the same amount of capacity. The proportion of capacity m_{tl} that can be added each time period at high and low response speeds therefore varies. A time horizon of four time periods is considered, for which the proportion of capacities that can be added each time period is shown in Table 3.4.

Table 3.4 Proportions of capacity added each time period at high and low response speeds

	T1	**T2**	**T3**	**T4**
Hi	1/2	1	1	1
Lo	1/3	2/3	1	1

3.5.1.3 Capacity addition costs

Unit cost of capacity additions is assumed to have a linear relationship with response time. These costs are set at 1 and 2/3 monetary units respectively for the high and low response speeds, corresponding to response times of two and three time periods. Adding more capacities faster would therefore ensure a highest level of protection but would also lead to higher costs of protection, and vice versa. Costs of protection at several combinations of capacity and response speeds (from high to low) are therefore obtained by multiplying the extra capacity sizes (units) by unit costs of adding capacity at selected response speeds.

3.5.2 Flow re-allocation vs contingent rerouting operation

Supply flow re-allocation is a contingency operation that allows entire supply flows to be re-allocated with an objective of minimizing total operational costs under disruptions. This is different to contingent re-routing operation where only the disrupted supply flows are routed to the surviving facilities, other network flows remaining undisturbed. We demonstrate the comparative effectiveness of the two mechanisms for risk mitigation utilizing a small network involving seven facilities, ten demand zones and attack of a single facility. This small network size is chosen for the ease of mapping supply flows graphically so that effect of the two different contingency operations can be demonstrated clearly. The demands, facility capacities (non-identical) and distance data are obtained as described in Section 5.1 and are provided in Appendix 1.

Figure 3.6a shows the supply flow configuration of an original network under consideration when there is no attack and no protection. This configuration is obtained by solving the ULP problem by setting all the protection and interdiction variables to zero. In this configuration the demands of each customer zone is exactly met from one or more nearest facilities. With respect to this original network we consider two cases of contingency operations when one of the facilities (NY) is attacked.

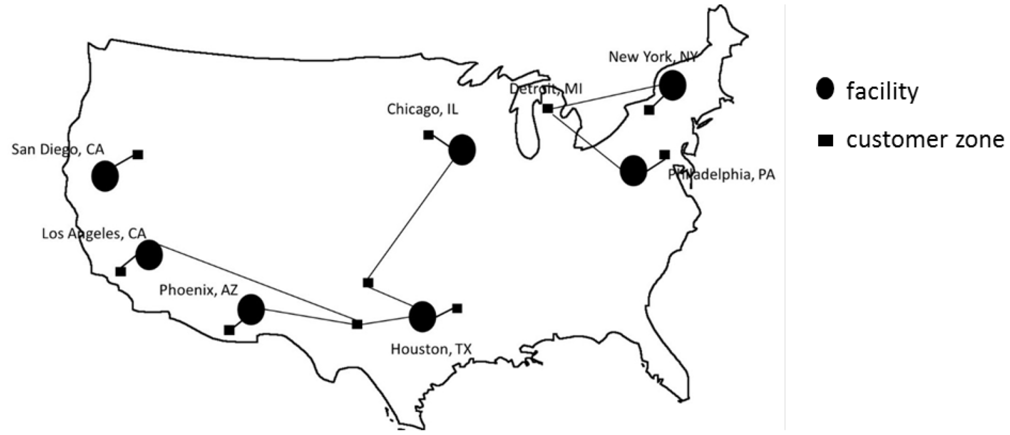


Figure 3.6a: Supply flows of an original network

Figure 3.6b shows the configuration when a facility at New York, NY is attacked, and the contingency operation involves redesign of network through supply re-allocations (Case I). New assignments from that of the original network are shown as dotted lines. This configuration can be obtained by solving the ULP problem by setting the interdiction variable corresponding to the attacked facility NY to 1 i.e., $s_{NY} = 1$, while keeping all other protection and interdiction variables as zero.

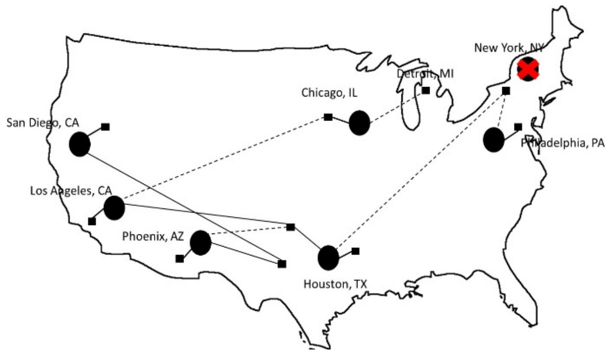


Figure 3.6b: Supply flows of a redesigned network with flow re-allocations

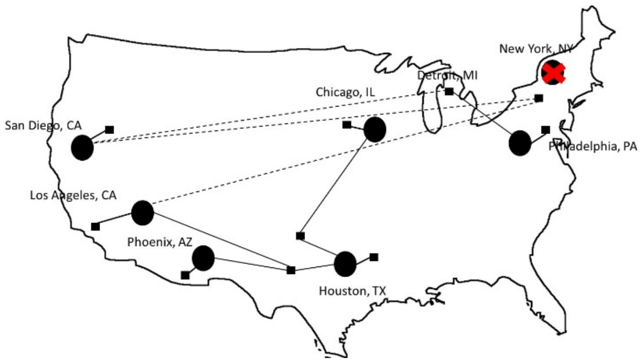


Figure 3.6c: Supply flows of a network with contingently rerouted flows after disruption

The configuration in Figure 3.6c results when the facility at NY is attacked ($s_{NY} = 1$), but the contingency plan involves only rerouting of supply originally handled by NY (Case II).

This requirement means only the demands of New York, NY and Detroit, MI customer zones can be reassigned. The new assignments of this network are shown as dotted lines (Detroit demand is contingently assigned to the San Diego, CA facility while NY demand is partly fulfilled from facilities at San Diego, CA and Los Angeles, CA.

Table 3.5 Operational costs under two different contingency operations

Cases	*cLS*	*cFlow*
Case I: Flow Re-allocation	73,555,000	12,037,000
Case II: Contingent Rerouting	73,555,000	17,860,000

Table 3.5 provides operational costs involving lost sales costs (*cLS*) and the transportation costs (*cFlow*) under the two cases of contingency operations illustrated through Figure 3.6b and Figure 3.6c. As can be observed in these results, contingency operation that relies on redesigned network through supply re-allocations is more cost effective than the operation that relies only on re-routing of disrupted flows. In this example, redesign through re-allocation of flows yielded approximately 32 % more reduction in system operational costs than re-routing of disrupted flows only.

Re-allocations allow flow exchanges which lead to a new optimal flow configuration under disruptions. For instance, in the network of Figure 3.6b, after attack of NY, a facility at Philadelphia, PA starts to partially serve demands of NY and therefore the PA facility no longer serves the demands of Detroit, MI which is now served by a facility at Chicago, IL. This exchange of flows will result in lower total network costs of transportation as compared to the contingent re-routing approach which allow only demands originally handled by NY facility to be re-routed. Note that such an exchange would be unnecessary when one ignores facility capacity limits (uncapacitated) since any amount of disrupted flows can then be contingently rerouted to the next nearest facility. Therefore, contingent rerouting, which is an appropriate recourse solution for an uncapacitated system, may not be appropriate for a capacitated system. Redesign of network flows through supply re-allocations is a more effective flow mechanism than contingent re-routing in capacitated systems.

The effectiveness of such contingency operations can be further enhanced through capacity backup provisions. This is because more supply flows can be recovered through backup productions and contingent capacity adjustments. Since response speeds impact the available capacity during recovery, appropriate selection of response speeds are necessary in planning such contingency strategies. Under a limited budget of protection, the main trade-off of response speed is with capacity volumes. At slower speeds, transitions to the desired capacities are slower and disruption impacts are prolonged even though the costs of such response speeds are low. At higher speeds desired capacities can be achieved faster. This however raises protection costs.

In the following sections we investigate optimal protection strategies for risk mitigation with respect to attack, protection budget and backup capacity features. Two types of networks with different initial capacity layouts are studied, the first network has identical distribution of initial capacities, while the second network is more generic with varying initial facility capacities. The network analyzed consists of 15 customers and 10 facilities for which the input parameters are derived as explained in Section 3.5.1. The two networks are equivalent in terms of the total system capacity, total demands and capacity slacks, hence the results are comparable.

3.5.3 Protection of networks with identical facility capacities

In identical capacity networks, losing any facility will result in the same units of capacity loss. The disruption risks of such networks can be considered to be more evenly distributed among facilities than similar networks with non-identical capacity. The two networks are therefore amenable to different protection strategies. Table 3.6 summarizes the optimal protection strategies for different levels of attacks and protection budgets (expressed in monetary units) for the network with identical capacity. Under the different combinations of protection budget and attack levels, it displays the optimal sets of facilities protected with levels of protection (Z_{jl}), facility sets that would be attacked as a consequence of this protection strategy (*Sj*), the units of lost sales (*uLS*), flow units (*uFlow*), total cost of lost sales (*cLS*), total flow costs (*cFlow*) and average flow distances/costs (*đ*).

The total operational costs of this network under varying budget levels are plotted for different attack levels in Figure 3.7. The total operational costs tend to grow when attacker capability is raised. Under this network, the attacker chooses to attack facilities that raise transportation costs the most. This rule is specific to networks where every facility has the same amount of initial capacity and penalties per unit lost sales are uniform and independent of customer locations or their demand sizes. This is because total cost of lost sales will remain unaffected irrespective of which sets of facilities in the network are attacked as long as the number of attacks are the same. The attacker can thus maximize benefits by attacking facilities that raise transportation costs the most. Consequently, protection efforts are concentrated in securing such facilties.

Table 3.6 Results of protection on an identical capacity network

Budget Level (*B*)	**Attack Level (*r*)**	Z_{jl}	S_j	***uLS***	***uFlow***	***cLS***	***cFlow***	***đ***
B = 0	1	-	1	6472	107280	16,982,500	36,205,500	337
	2	-	1,5	18392	95360	48,260,600	35,287,200	370
	3	-	1,5,10	30312	83440	79,538,700	29,737,700	356
	4	-	1,3,5,10	42232	71520	110,817,000	24,835,900	347
B = 670	1	1(4)	5	3472	110280	9,110,530	35,302,100	320
	2	1(4)	5,10	15392	98360	40,388,600	29,752,600	302
	3	1(4)	3,5,10	27312	86440	71,666,700	24,850,800	287
	4	1(4)	2,3,5,7	39232	74520	102,945,000	20,637,800	277
B = 1340	1	1(4), 5(4)	10	1236	112516	3,243,260	29,343,400	261
	2	1(4), 5(4)	3,10	12392	101360	32,516,600	25,545,400	252
	3	1(4), 5(4)	2,3,7	24312	89440	63,794,700	21,332,400	239
	4	1(4), 2(4)	3,5,8,10	36232	77520	95,072,800	18,633,200	240
B = 2010	1	1(2), 5(4)	10	628	113124	1,647,870	25,034,500	221
	2	1(2), 5(4)	3,10	9392	104360	24,644,600	25,548,400	245
	3	1(2), 2(4), 5(4)	3,8,10	21312	92440	55,922,700	19,327,900	209
	4	1(2), 2(4), 5(4)	3,6,7,10	33232	80520	87,200,800	15,809,400	196
B = 2680	1	1(2), 5(2)	2	298	113454	78,1952	22,928,200	202
	2	1(4),2(4),3(4),5(4)	7,10	6392	107360	16,772,600	24,074,800	224
	3	1(4),2(4),3(4),5(4)	6,7,10	18312	95440	48,050,700	17,571,600	184
	4	1(4),2(4),3(4),5(4)	4,8,9,10	30232	83520	79,328,800	9,688,960	116

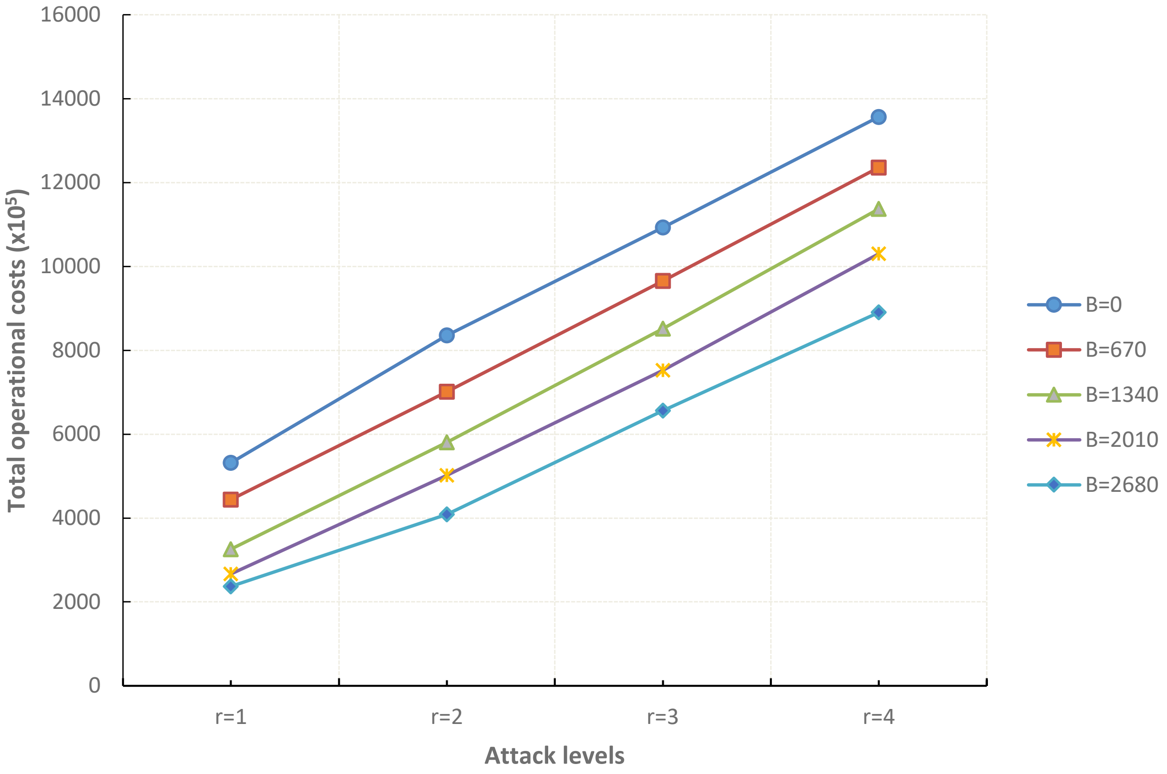


Figure 3.7: Total operational costs for network with identical facility capacities

.

3.5.4 Protection of a network with non-identical facility capacities

In networks with non-identical facility capacities, protection is managed through securing both high capacity facilities and facilities that are critical to minimizing average flow distances. Optimal protection strategies of the identical capacity network cannot therefore be substituted for protecting these networks. The optimal protection strategies for non-identical capacity network are analyzed with respect to different combinations of protection budget and attack levels and the results are summarized in Table 3.7.

Table 3.7 Results of protection on a non-identical capacity network

Budget Level (*B*)	**Attack Level (*r*)**	Z_{jl}	S_j	***uLS***	***uFlow***	***cLS***	***cFlow***	***đ***
B = 0	1	-	1	28952	84800	75,970,000	10,345,600	122
	2	-	1,3	42952	70800	112,706,000	9,063,540	128
	3	-	1,3,5	51352	62400	134,748,000	8,877,430	142
	4	-	1,2,3,5	68552	45200	179,880,000	7,051,860	156
B = 670	1	1(4)	2	8752	105000	22,965,200	19,282,600	184
	2	1(4)	2,7	15952	97800	41,858,000	18,617,100	190
	3	1(4)	2,6,7	23552	90200	61,800,400	16,103,600	179
	4	1(4)	2,3,6,7	37552	76200	98,536,400	10,179,400	134
B = 1340	1	1(4), 2(4)	3	2552	111200	6,696,450	16,507,100	148
	2	1(4), 2(4)	3, 10	8552	105200	22,440,400	15,235,200	145
	3	1(4), 2(4)	3, 5, 9	16952	96800	44,482,000	12,055,600	125
	4	1(4), 2(4)	3, 4, 8, 9	26952	86800	70,722,000	9,332,140	108
B = 2010	1	1(4), 2(4), 3(4)	4	248	113504	650,752	10,625,300	94
	2	1(4), 2(4), 3(4)	4,6	3552	110200	9,320,450	14,643,800	133
	3	1(4), 2(4), 10(4)	3,4,8	17152	96600	45,006,800	10,538,600	109
	4	1(4), 2(4), 3(4)	4,6,8,9	17552	96200	46,056,400	12,017,000	125
B = 2680	1	1(4), 2(2),3(4)	4	0	113752	0	10,622,800	93
	2	1(4), 2(4),10(2)	3,4	6952	106800	18,242,000	12,085,600	113
	3	1(4), 2(4),10(4),8(4)	3,4,9	13752	100000	36,085,200	11,235,500	112
	4	1(4), 2(4), 3(4), 4(4)	6,7,8,9	11352	102400	29,787,600	12,803,900	125

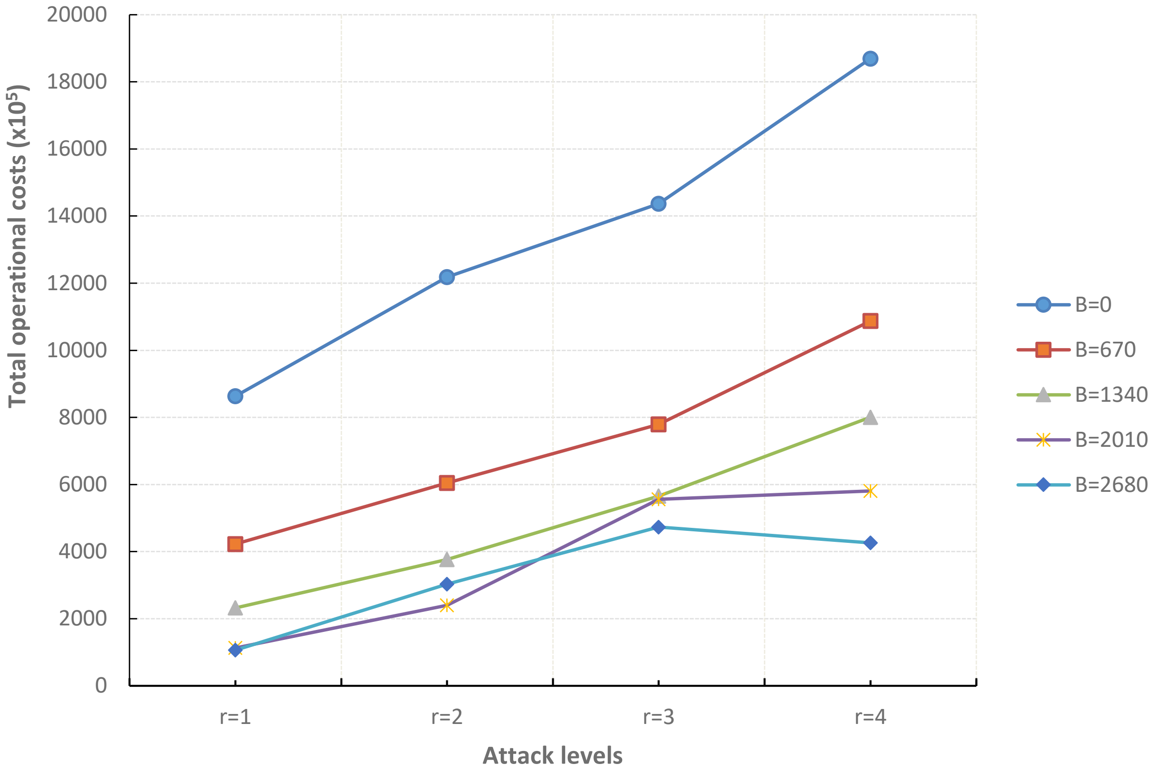


Figure 3.8: Total operational costs for network with non-identical facility capacities

The operational costs increase when attack levels are higher but decrease when the protection budget is raised. The total operational costs of this network is plotted against available protection budget and attack levels in Figure 3.8. Although the decreasing trend of operational cost is observed for raising protection, this result is less intuitive in the specific case involving two facility attacks. The total operational costs seems to grow under protection involving a higher budget level (B = 2680) than with a lower budget level (B = 2010). This solution is however superior with respect to both average flow distances and lost sales than other solutions at this budget level, which results due to the selection of optimal strategy based on normalized total costs. This can be explained by looking at two competing feasible strategies obtained for this problem as follows: Consider two feasible protection strategies A and B with a budget level of B = 2680. Table 3.8 provides the total cost in absolute values obtained with a feasible strategy A, involving four facility protection at the lowest levels of capacity and response speeds. The total cost obtained by selecting this strategy is the minimum among all feasible strategies that can be obtained at this budget level. Strategy *B,* however, results in the minimum normalized total costs with respect to all the feasible protection strategies at this budget level, and hence is selected as an optimal strategy, even though its total cost in absolute values is higher than that obtained for strategy *A*. Note that normalization is done to uniformly scale the two cost components (between 0 and 1), and hence to avoid any dominance of one cost component (higher values) over the other when protection decisions are made.

Table 3.8 Illustration of results of two competing strategies in a non-identical capacity network

Strategy	Z_{jl}	S_j	*cLS*	*cFlow*	*cTotal*	$\bar{d}$
A	1(4)+2(4)+3(4)+6(4)	4, 8	5,500,000	15,500,000	21,000,000	139
B	1(4)+2(4)+10(2)	3, 4	18,200,000	12,100,000	30,300,000	113

It can further be observed from Table 3.8 that selecting strategy *B* reduces the average flow distances but increases lost sales units (i.e., fewer demand units are satisfied in this strategy as compared to strategy *A*). This will increase the total cost of disruption (*cTotal*), since unit costs of lost sales are higher compared to unit flow costs. Nevertheless, by avoiding the dominance of the lost sales cost component, strategy *B* leads to a lower transportation (flow) costs for customers who are served. If the strategy selection were based on absolute total costs rather than normalized costs, the resulting solution would lead to higher costs of serving customers due to increased flow distances. Selecting strategies based on normalized costs reduces such tendencies.

Comparing experimental results of the two networks, it can be observed that the total operational costs of identical capacity networks are lower than that of a network with non-identical capacities under no budget of protection (B=0). These results suggests that identical capacity network is more cost effective to operate if protection budget is non-existent. Raising the level of protection however, marginal reductions in operational costs for this network is much lower than that can be achieved from such protections in a non-identical capacity network. This effect can be observed by comparing the graphs in Figure 3.7 and Figure 3.8. These results indicate protection strategies are dependent on network characteristic of initial facility layout. In distance based network designs, transportation costs can be lowered if high capacity facilities are located in densely populated areas and low capacity facilities in less populated areas. Such high capacity facilities are strong candidates for attacks, particularly if post disruption demand allocations significantly raise travel costs. Protection strategies of non-identical capacity networks seek for trade-off solutions that balance the loss of high capacity facility against the loss of travel distances.

3.5.5 Centralized vs distributed backups

An important issue in contingent capacity adjustments through capacity backups is whether such backup capacities should be confined to a fewer facilities (centralized) or distributed over many to enable cost effective contingency strategies. The observation made from the above results suggests that one of the determinants is the capability of the attacker or the size of the attacks (*r*). A more offensive attacker is able to strike more facilities causing larger capacity losses. Under this condition, it is important for the defense planner to fortify as many facilities as possible so that maximum amount of existing capacities are preserved. This is true especially when available backup capacity volumes for contingent capacity adjustments are low as compared to capacities that may be lost from attacks. Planning denfense against more capable attacker therefore necessitates spreading out protection budget over many facilities. This means contingent capacity adjustments are done at several facilities by utilizing low capacities and slower speeds rather than confining such adjustments to fewer facilities with high capacities and higher speeds.

Under a less offensive attack, recovering capacities faster is a more important priority than recovering more units of capacities since less capacities are lost from attacks. Contingent capacity adjustments can be done by utilizing higher response speeds and capacities of backups which centralizes backup over few facilities. These inferences can be drawn from the above results. For instance, when r = 1, 2, and under a budget level of B = 2680, Table 3.7 results show that, an optimal protection plan involves adding backups at three facilities; whereas when attacker capability is increased further (r = 3, 4) , it is necessary to spread backups to four facilities.

The decision to centralize or distribute backup is also influenced by the available sizes (volumes) of backup capacities. If the range of available sizes differ widely, recovery through high volume and high response speed become more efficient, which result in centralized backups. This is illustrated in the following example. Consider protection in an identical facility capacity network under a protection budget of B = 2680 and attack of two facilities (r = 2). The identical capacity network is chosen for this illustration to avoid any influence on the results due to initial capacity variations. Table 3.9 shows the results of optimal protection under two different settings of available backup capacity for this problem. Under the first setting the range of available volumes is small (high volume =2000 units, low volume=1000 units). The optimal protection plan obtained under this backup capacity availability involves protecting four facilities at low capacity and response speeds, i.e. distributed backup.

Table 3.9 Effect of varying backup capacity sizes on optimal protection

Backup capacity sizes	Z_{jl}	S_j	***uLS***	***uFlow***	***cLS***	***cFlow***	***đ***
H: 2000 units L: 1000 units	1(4), 2(4), 3(4), 5(4)	7,10	6392	107360	16,772,600	24,074,800	224
H: 2500 units L: 500 units	1(1), 5(4)	2,7	8142	105610	21,364,600	25,497,000	241

Under the second setting the range of available volumes is increased (high volume =2500, low volume=500). As a result of this variation, the optimal protection changes from a distributed backup plan of the initial setting to a centralized one in which only two facilities are protected by utilizing higher volumes and higher available response speeds.

The above results have demonstrated that decisions to centralize or distribute backups are dependent on how the network capacities are affected due to attacks as well as what capacity sizes are available for contingent adjustments. When planning defense under more offensive attacks, it is generally preferable to distribute the available backup, unless the low volume and speeds of contingent capacity adjustment through such plans outweigh the benefits gained from securing more units of existing capacities.

3.5.6 Algorithm performance

Computational performance and robustness of the proposed algorithm is evaluated under larger networks. The number of customer zones (demands) and facilities were varied to study the effect of these variations on algorithm performance. The experiments were conducted for all combination of five level of attacks (r = 1 to 5) , four levels of budget (B=670, 1340, 2010, 2680), three levels of demands (I=25, 35, 50) and three levels of facility (J=10,15,20). The demands, distance and facility capacities were derived as explained in Section 5.1 and are provided in Appendix. All of the experiments were conducted under identical settings. The computational results are summarized in Table 3.10. All problem instances provided in this Table could be solved to optimality in a reasonable amount of computational time.

Increasing the size of the number of facilities (J) and the number of demands (I) both increase the size of the ULP problem to be solved at each iteration because it increases the number of variables and constraints of the problem. The depth and the breadth of the binary search tree is independent of these parameters but increases with the increase of attack level r and the budget level B. The computational time of the algorithm is therefore sensitive to these parameters.

Table 3.10 Computational performance of the algorithm on larger networks

Budget	Problem instance	*Sj*			*Zjl*			Total Cost(x10^6)			CPU time(s)		
		I=25	I=35	I=50	I=25	I=35	I=50	I=25	I=35	I=50	I=25	I=35	I=50
B=670	10.1	2	2	2	1(4)	1(4)	1(4)	24.00	38.20	73.96	0.10	0.20	0.30
	10.2	2,7	2,7	2,7	1(4)	1(4)	5(4)	41.10	66.20	111.28	0.20	0.25	0.3
	10.3	2,3,4	1,3,10	2,6,7	1(4)	5(4)	1(4)	43.80	145.40	147.71	0.25	0.35	1
	10.4	2,4,6,7	1,2,3,7	2,3,6,7	1(4)	5(4)	1(4)	90.20	129.77	196.59	1	1	2
	10.5	1,4,6,7,9	1,2,3,7,10	2,3,6,7,10	1(4)	5(4)	1(4)	115.00	235.70	229.92	1	1	2
	15.1	2	2	2	1(4)	1(4)	1(4)	10.80	22.20	52.20	0.15	0.25	0.35
	15.2	2,7	2,3	2,7	1(4)	1(4)	5(4)	20.50	29.66	79.80	0.25	0.3	1
	15.3	2,6,7	1,3,10	2,6,7	1(4)	5(4)	5(4)	34.80	111.40	107.90	1	3	4
	15.4	2,3,4,6	1,3,10,12	2,7,11,13	1(4)	5(4)	5(4)	38.20	133.00	124.80	6	11	14
	15.5	2,3,4,6,7	1,3,10,12,15	2,6,7,11,13	1(4)	5(4)	5(4)	66.20	154.20	153.20	17	24	33
	20.1	2	2	1	1(4)	1(4)	2(4)	13.40	29.60	84.90	0.3	0.35	0.4
	20.2	2,7	1,3	2,11	1(4)	2(4)	7(4)	24.50	46.28	73.50	0.35	1	2
	20.3	2,6,7	2,3,4	7,11,13	1(4)	1(4)	2(4)	36.90	54.90	63.30	5	9	15
	20.4	1,2,3,7	1,3,5,18	6,7,11,13	5(4)	2(4)	2(4)	82.50	110.30	84.60	29	47	69
	20.5	1,2,3,6,7	2,6,11,13,16	2,6,11,13,16	5(4)	7(4)	7(4)	104.20	79.20	123.50	112	171	234
B=1340	10.1	3	3	3	1(4),2(4)	1(4),2(4)	1(4),2(4)	21.80	28.20	53.15	0.2	0.3	0.5
	10.2	3,10	3,4	3,10	1(4),2(4)	1(4),2(4)	1(4),2(4)	29.40	37.71	86.48	0.25	0.35	1
	10.3	2,3,4	2,4,6	3,4,8	1(4),5(4)	1(4),7(4)	1(4),2(4)	40.20	83.30	129.04	1	2	3
	10.4	2,4,7,9	2,4,6,9	3,4,8,9	1(4),6(4)	1(4),7(4)	1(4),2(4)	71.30	115.00	166.14	3	4	5
	10.5	2,4,7,8,9	2,4,6,8,9	3,4,8,9,10	1(4),6(4)	1(4),7(4)	1(4),2(4)	102.70	151.60	199.47	3	4	5
	15.1	3	4	3	1(4),2(4)	1(4),2(4)	1(4),2(4)	10.30	17.10	32.40	0.3	0.35	0.6
	15.2	4,9	4,9	3,4	1(4),2(4)	1(4),2(4)	1(4),2(4)	17.20	26.80	65.20	0.4	1	3
	15.3	4,8,9	3,4,5	2,11,13	1(4),2(4)	1(4),2(4)	1(4),7(4)	26.20	40.61	89.60	6	10	18
	15.4	2,3,4,5	3,4,8,9	3,4,8,9	1(4),6(4)	1(4),2(4)	1(4),2(4)	25.98	64.90	118.80	26	45	64
	15.5	3,4,5,8,9	3,4,5,8,9	3,4,8,9,12	1(4),2(4)	1(4),2(4)	1(4),2(4)	46.78	90.80	140.90	71	110	163
	20.1	6	6	1	1(4),2(4)	1(4),2(4)	2(3)	6.76	13.60	83.90	0.3	0.35	0.45
	20.2	4,9	6,7	6,7	1(4),2(4)	1(4),2(4)	2(4),11(4)	10.10	20.80	42.40	2	3	7
	20.3	4,8,9	3,4,5	6,7,13	1(4),2(4)	1(4),2(4)	2(4),11(4)	16.40	17.19	58.00	20	32	52
	20.4	4,8,9,16	3,4,6,7	2,6,13,16	1(4),2(4)	1(4),2(4)	7(4), 11(4)	22.70	43.30	96.90	124	195	270
	20.5	2,3,4,5,6	3,4,6,8,9	2,4,8,11,13	1(4),7(4)	1(4),2(4)	6(4),7(4)	34.10	61.40	122.40	540	753	1095
B=2010	10.1	6	6	7	1(4),2(4),3(4)	1(4),2(4),3(4)	1(4),2(4),3(4)	15.60	22.90	40.73	0.3	0.4	0.5
	10.2	6,7	3,4	3,4	1(4),2(4),3(4)	1(4),2(4),5(4)	1(4),2(4),10(4)	21.00	34.08	83.83	1	2	3
	10.3	2,4,5	3,4,5	3,4,8	1(4),3(4),6(4)	1(4),2(4),10(4)	1(3),2(4)	24.87	64.10	128.02	4	7	8
	10.4	4,6,8,9	4,6,8,9	3,4,8,9	1(4),2(4),3(4)	1(4),2(4),3(4)	1(4),2(4),10(4)	42.17	85.10	158.64	8	15	17
	10.5	2,3,5,6,10	3,5,8,9,10	3,5,6,7,10	1(4),4(4),7(4)	1(4),2(4),4(4)	1(4),2(4),8(4)	99.50	123.00	186.64	9	13	17
	15.1	6	3	4	1(4),2(4),3(4)	1(4),2(4),4(4)	1(4),2(4),3(4)	9.10	15.90	26.78	0.35	0.45	0.6
	15.2	3,12	3,5	3,5	1(4),2(4),4(4)	1(4),2(4),4(4)	1(4),2(4),4(4)	14.70	21.30	53.60	2	4	10
	15.3	3,10,12	3,5,6	3,5,10	1(4),2(4),8(4)	1(4),2(4),4(4)	1(4),2(4),4(4)	20.20	25.58	77.20	20	43	65
	15.4	2,3,5,6	3,4,5,10	3,4,5,9	1(4),4(4),7(4)	1(4),2(4),9(4)	1(4),2(4),8(4)	22.87	57.60	111.70	91	150	247
	15.5	3,5,6,8,9	3,6,7,8,9	3,4,9,10,12	1(4),2(4),4(4)	1(4),2(4),3(4)	1(4),2(4),8(4)	28.30	68.30	129.30	297	461	743
	20.1	3	4	3	1(4),2(4),6(4)	1(4),2(4),6(4)	1(4),2(2)	6.57	12.20	25.15	0.4	0.45	1
	20.2	3,19	4,9	3,4	1(4),2(4),9(4)	1(4),2(4),6(4)	1(4),2(4),7(4)	9.48	18.20	46.70	6	10	24
	20.3	7,11,13	4,8,9	3,4,8	1(4),2(4),9(4)	1(4),2(2)	1(4),2(2)	13.30	26.40	68.50	76	106	195
	20.4	6,7,11,13	3,5,6,7	3,4,8,9	1(4),2(4),9(4)	1(4),2(4),4(4)	1(4),2(4),6(4)	21.50	26.01	88.70	612	893	1073
	20.5	2,3,4,5,9	3,4,6,7,9	4,7,8,9,11	1(4),6(4),7(4)	1(4),2(4),8(4)	1(4),2(4),6(4)	21.89	54.40	94.30	2245	3509	4367
B=2680	10.1	10	7	4	1(4),2(4),3(4),6(4)	1(4),2(4),3(4),6(4)	1 (4), 2 (3), 3 (3)	15.00	21.40	34.57	0.4	0.5	0.65
	10.2	4,9	4,8	3,4	1(4),2(4),3(4),6(4)	1(4),2(4),3(4),6(4)	1(4),2(4),10(2)	20.70	31.10	76.60	2	5	7
	10.3	6,7,8	5,6,7	4,6,7	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),4(4)	3(4),1(4),2(4),8(4)	22.90	38.36	98.30	24	20	25
	10.4	5,6,7,8	3,5,6,7	3,4,6,9	1(4),2(4),3(4),4(4)	1(4),2(4),4(4),10(4)	1(4),2(4), 8(4),10(4)	24.46	82.00	149.00	28	33	42
	10.5	5,6,7,8,9	5,6,7,8,9	5,6,7,8,9	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),4(4)	41.10	92.60	156.64	29	33	45
	15.1	4	6	6	1(4),2(4),3(4),6(4)	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),4(4)	8.95	15.30	23.70	0.45	0.55	0.65
	15.2	8,9	8,9	3,5	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),4(4)	1(2),2(4), 4(4)	11.60	20.30	45.90	7	11	25
	15.3	6,8,9	3,5,7	3,5,10	1(4),2(4),3(4),4(4)	1(4),2(4),4(4),6(4)	1(3),2(4),4(4)	14.70	21.70	75.90	75	119	246
	15.4	2,3,5,8	6,7,8,9	4,6,7,9	1(4),4(4),6(4),7(4)	1(4),2(4),3(4),4(4)	1(4),2(4),3(4),8(4)	18.31	33.72	88.90	350	510	1091
	15.5	3,5,6,7,8	4,5,6,8,9	3,5,8,9,10	1(2),2(4),4(4)	1(4),2(4),3(4),7(4)	1(4),2(4),4(4),12(4)	18.36	60.30	119.20	1042	1747	3130
	20.1	14	4	7	1(4),2(4),3(4),6(4)	1(4),2(4),6(2)	2(4),1(4),3(4),6(4)	6.33	11.70	24.00	0.5	0.65	0.7
	20.2	6,7	7,11	3,4	1(4),2(4),3(4),9(4)	1(4),2(4),4(4),6(4)	2(4),13(4), 6(4), 1(4)	8.86	16.10	39.90	16	26	55
	20.3	3,12,19	6,7,13	3,6,7	1(3),2(3),4(4)	1(4),2(4),4(4),11(4)	1(4),2(2),4(4)	13.20	24.10	57.40	196	339	568
	20.4	3,10,12,19	3,5,7,8	3,4,8,9	1(4),2(4),6(4),8(4)	1(4),2(4),4(4),6(4)	2(4),6(2),1(4)	17.10	17.25	82.50	2166	4131	4311
	20.5	2,3,5,8,9	4,5,7,8,9	3,4,8,9,16	1(4),4(4),6(4),7(4)	1(4),2(4),3(4),6(4)	1(4),2(4),19(4),6(4)	15.29	31.66	98.40	8550	16541	17227

3.5.6.1 Effect of raising budget levels (B) on computational time under different network sizes

The average CPU time under varying levels of protection budget with respect to demand sizes (*I*) and number of facilities (*J*) are shown in Table 3.11a and Table 3.11b respectively. The average CPU times are computed from the results shown in Table 3.10. The values in Table 3.11a represents the mean computational time across all attack levels and facility levels for each settings of budget level and demand level. The values of Table 3.11b is the mean computational time across all attack levels and all demand levels for each setting of budget level and facility level. As expected, the CPU times increase when budget levels are increased. This is because more number of facilities can be protected at higher budgets which will increase the depth of the search tree and consequently the number of nodes at which the ULP needs to be solved. Comparing the values across the two tables, shows that changing the number of facilities has more influence on computational effort than changing the number of demand. For the same level of protection, it is observed in these results that a network with 20 facilities raises the average CPU time by almost seven folds than the network with 15 facilities. Although changing the number of demands, also raises the CPU time, it is less dominant than changing the number of facilities.

Table 3.11 Variation of CPU time with respect to budget levels under different demand and facility levels

	(a) Number of Demand Zones			(b) Number of Facilities		
Budget Level	I=25	I=35	I=50	J=10	J=15	J=20
B=670	11	18	25	1	8	46
B=1340	53	77	112	2	34	206
B=2010	225	348	451	7	142	874
B=2680	832	1568	1785	20	557	3609

3.5.6.2 Effect of raising attack levels (*r*) on computational time under different network sizes

The variations in average CPU time with respect to the number of attacks and the demand levels is depicted in Table 3.12a. The averages are obtained from Table 3.10 by computing the mean CPU time across all levels of budget and number of facilities. The results show that the computational effort is increased when attack levels are raised. At lower level of attacks, computations are quite fast irrespective of the demand levels.

Table 3.12 Variation of CPU time with respect to attack levels under different demand and facility levels

	(a) Number of Demand Zones			(b) Number of Facilities		
Attack levels	I=25	I=35	I=50	J=10	J=15	J=20
r=1	0.31	0.40	0.56	0.36	0.42	0.49
r=2	3.12	5.33	11.53	1.86	5.41	12.70
r=3	35.69	57.53	100.00	7.97	50.83	134.42
r=4	287.00	502.92	600.42	13.25	217.08	1160.00
r=5	1076.33	1947.25	2255.08	13.50	653.17	4612.00

Table 3.12b demonstrates the effect of varying attack levels when the average CPU time is computed across all demand levels and budget levels. Comparing results in Table 3.12a and 3.12b, it can be seen that neither the increase in number of facilities nor the number of demands significantly affect the solution times if the attack levels are low. When attack levels increase, solution times increase. Higher facility levels are quite dominant in raising computational effort of the algorithm than raising the demand levels. This is similar to observation made in Section 5.3.1 under varying budget levels.

3.6 Chapter Summary

In this Chapter, we examined the problem of responsive contingency planning under a major disruption from intentional attack on supply facilities. The tri-level game theoretic model was developed in which facility security and disruption recovery aspects were integrated. The contingent mechanism to enhance recovery was considered through backup production, the appropriate volume and response speeds for which were involved as model decisions. We further developed implicit enumeration algorithm utilizing the structure of the problem which enabled solving the proposed model for developing protection strategies for risk mitigation. Through a small illustrative example involving single facility attack, we first demonstrated the demand routing characteristic optimal under a capacitated system. In particular, the importance of flow re-allocations rather than contingent re-routing in a capacitated network was demonstrated. As well, it was shown that instantaneous capacity addition assumptions leads to capacity overestimations and inappropriate mitigation solutions and inadequate recourse actions during major disruptions.

We presented two hypothetical networks with different initial capacity distributions to demonstrate that network topologies affect its vulnerability to disruptions and contingency strategies of protection can be different under varying network configurations. It was demonstrated that identical capacity distributions have risk diversification effect, and are effective for controlling losses under a major disruptions. However, the relative efficiency improvements through protection are lower for such networks than networks with non-identical capacity distributions in its facilities. Through this analysis, we also observed that planning against a more capable adversary (higher number of attacks), protection needs to be dispersed or decentralized in order to spread out risk.

We further examined the computational performance of the developed algorithm with respect to the variations in the budget of protection, number of attacks, number of facilities and number of demands. It is observed that computational time increase with the increase of budget level, attack levels, number of facilities or the number of demands. Raising the protection budget or the attack level lead to increase in the size of the search tree which affect the computational time. As well, increasing number of facilities and the number of demands both raise computation time but the effect of raising the number of facilities is higher than the effect of raising the number of demands. The solution time for problem size involving up to fifty demands, twenty facilities, four levels of protection budget and five levels of attacks have been reported. Since protection budget are normally tight and simultaneous attacks on multiple facilities are limited to a few facilities, the test problem sizes are practical. The optimal solution with the largest values for each parameter (I = 50, J = 20, r =5 and B=2680) was obtained in less than 5 hours of CPU time. Considering the fact that it is a strategic long term decision, the CPU time is acceptable.

Chapter Four: **Analyzing Congestion Effects in Responsive Contingency Planning**

4.1 Background

The level of responsiveness of a backup supplier is one of the important considerations in a contingency strategy relying on backup sourcing. The supply network may incur significant short term losses through lost sales if backup production is not adequately responsive to fill customer demands. In the long term, the incurred losses may result in losing market share to the competitors. However, despite the responsiveness of a backup supplier, rerouting the affected demand at the beginning of the response period will create an overflow when the backup capacity is not fully available. As a result of this overflow, the congestion effects build up which would increase lead time and lower throughput during the response period.

Congestion in the system is a general consequence of a mismatch between demand arrival rates and production rates at facilities. Contingency strategies relying on backup production should consider this effect in order to get a better representation of the available system capacity during disruption periods. This will enable proper identification of appropriate levels of response speed and production volumes at backup facilities necessary for optimal protection design and flow recovery.

The literature deals with congestion effects through its implicit modeling using queueing systems and capacity constraints or through explicit modeling by including it in the objective function of an optimization model. In the implicit approach, congestion effects are incorporated by introducing expected queue lengths (WIP) or waiting times derived using queueing models as constraints. This approach is taken in most capacity planning and strategic design models involving facility location decisions. In capacity planning, Bitran and Tirupati (1989) introduced congestion effects for the first time by developing a model to minimize total capacity costs in which WIP level was computed using GI/G/m queueing system and a capacity constraint was imposed to limit it to a target level. Marianov and Serra (2003) apply this approach in a hub and spoke network topology where the hubs are modeled as M/D/c queueing system and congestion is captured using a probabilistic capacity constraint that limits the queue length at hub facilities.

Congestion effects are modeled with explicit incorporation of congestion related costs in the objective function of the optimization models in many articles. Amiri (1997) considered a service facility location problem in which service facilities are modeled as M/M/1 queue. The model incorporates the cost of waiting time in the objective function of a cost minimization problem which include transportation costs and fixed costs of facility location. Wang, Batta and Rump (2002) also model service facilities as M/M/1 queueing system in the immobile server facility location problem but they additionally include a constraint which bounds the waiting time at each facility to a desired limit. This model therefore allows a better control of congestion than in Amiri (1997). The M/M/1 queue is also applied in Zhang, Berman and Verter (2009) to deal with congestion effects in preventive healthcare facilities. Vidyarthi and Kuzgunkaya (2015) study similar problem using a more general M/G/1 queues in order to deal with situations with different coefficient of variations in service times. Clearing function is utilized to model congestion effects on available capacity (throughput) in a few articles. Clearing function is introduced by Karmarkar (1989) and provide throughputs as a function of WIP and the service rates at facilities. Kim (2012) utilizes this function to estimate the capacity levels of facilities under congestion. The WIP is computed by considering facilities as a GI/G/1 queueing models. The model involves a facility location problem which simultaneously minimizes congestion cost (the order waiting cost due to congestion), the fixed costs and the transportation costs. Nejad *et al.* (2014) also utilize clearing function to model congestion in a responsive contingency planning problem involving backup supplier. The WIP for the clearing function is derived using a M/G/1 queueing system. The appropriate level of capacity of a backup supplier is estimated using this clearing function.

The application of queueing models facilitate congestion effect analysis under demand variability and stochastic processing times. When this is not the case, congestion effects can be handled using appropriate other functions that model flows to the facilities. Elhedhli and Hu (2005) explicitly model congestion costs on the objective function using a convex cost function that increases exponentially as more flows are routed to facilities (hubs). Camargo, Miranda, Ferreira and Luna (2009) also propose a generalized convex cost function to model congestion under a deterministic demand at the hubs. In a closely related article involving competitive facility location in a distribution netwrok, Konur and Geunes (2011, 2012) also introduce convex cost functions in the objective function of their model to capture traffic congestion cost on the distribution network link.

Most of the relevant articles consider handling congestion under a common problem of matching supply with demand, despite the different approaches in modeling. With the exception of Nejad *et al.* (2014) none of the aforementioned articles focus on the management of major disruptions. In this Chapter we focus on the management of major disruption risks under congestion effects. In particular, we investigate the congestion impacts on responsive contingency planning under the risk of a major supply disruption. The remainder of the Chapter is organized as follows: Section 4.2 discusses the congestion cost function utilized in this work, its linearization technique and the re-formulation of the tri-level model and its solution. Section 4.3 presents results and analysis. Chapter summary is presented in Section 4.4.

4.2 Congestion integrated responsive contingency planning model

The responsive contingency planning model under congestion applies the same framework as the model presented in Chapter 3. The overall problem is presented in a game theoretic defender-attacker-defender principle but in this case, the recourse plan involves consideration of congestion related costs. In other words, within the D-A-D framework developed in Chapter 3, in the model presented here, the user level problem is solved with an additional cost term related to congestion in the objective function. We first present the congestion cost function and its linearization.

4.2.1 Congestion cost function

Congestion affects supply flows of a facility, the cause and effect of which can be represented as shown in Figure 4.1. When more flows are directed to a facility, congestion levels grow, whereas increased congestion leads to reduced throughput. Therefore as more flows are directed to a facility, the congestion costs tend to grow. This relationship can be represented using a convex cost function as provided in (4.1):

$$f(\phi) = w\phi^q \tag{4.1}$$

where ϕ is the amount of flow to a facility and w and q are positive constants with $q \geq 1$.

The congestion cost function (4.1) does not consider the capacity limit on facilities, however it implies that congestion costs increase in flow volume at an increasing rate and reflects the nature of facility congestion in supply systems. Similar function is applied in Konur and Geunes (2011, 2012) to represent traffic congestion in a distribution network design of a competitive supply chain. Elhedhli and Hu (2005) also apply this kind of function for computing congestion costs in a hub location problem. The rationale of using such convex cost functions is consistent with the study of Weisbrod, Vary and Treyz (2001) which investigates the sensitivity of congestion by industry sectors. Their study mentions that higher level of congestion are associated with companies with volume flows or higher shipping levels.

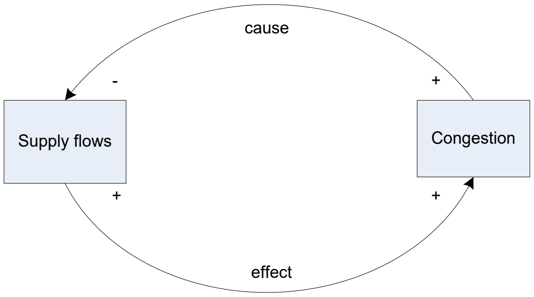


Figure 4.1: Representing congestion with a cause and effect relationship

Using the notations in Table 3.1, the total flow directed to a facility *j* at each time period *t* can be expressed as:

$$\phi_{jt} = \sum_{i \in I} x_{ijt} \tag{4.2}$$

Using relations (4.1) and (4.2) the congestion cost function for every facility *j* at each time period *t* can therefore be expressed as:

$$f(\phi_{jt}) = w\phi_{jt}^{q} = w\left(\sum_{i \in I} x_{ijt}\right)^{q} \tag{4.3}$$

This function (4.3) is non-linear and convex. The linear approximation of congestion cost function $f(\phi_{jt})$ near a given flow $\phi_{jt}{}^{k}$ is obtained from the equation of the tangent line:

$$f(\phi_{jt}{}^{k}) + f'(\phi_{jt}{}^{k})(\phi_{jt} - \phi_{jt}{}^{k}) = w(1-q)\left(\phi_{jt}{}^{k}\right)^{q} + wq\left(\phi_{jt}{}^{k}\right)^{q-1}\phi_{jt} \tag{4.4}$$

The linear approximation of the congestion cost function $f(\phi_{jt})$ corresponds to the maximum of a set of piecewise linear and tangent hyperplanes.

$$f\left(\phi_{jt}\right) \approx \max_{k \in K_j}\left(w(1-q)\left(\sum_{i} x_{ijt}{}^{k}\right)^{q} + wq\left(\sum_{i} x_{ijt}{}^{k}\right)^{q-1}\left(\sum_{i} x_{ijt}\right)\right) \tag{4.5}$$

Here *kϵK_j* represents the infinite sets of points where tangent line equations can be written for estimating the function.

4.2.2 Model formulation

Using the same set of notations as in Section 3.3, we formulate the tri-level responsive contingency planning model under congestion. In this model structure, the user level problem corresponds to decisions where the system user (defender) chooses allocations for each demand facility pair and the corresponding lost sales units to solve the underlying cost minimization problem. The attacker level corresponds to the worst case attack where the attacker selects an attack scenario *s* from a feasible attack scenario set *S* to maximize costs, while the defender level problem involves defender's choice of a protection solution *z* from a feasible solution set *Z* to minimize the costs. The model therefore seeks an optimal protection z ϵ Z by solving the nested optimization problem (4.6)

$$\min_{z\in Z}\max_{s\in S}\min_{x,u\in X,U} Costs(z,s,x,u) \tag{4.6}$$

This nested structure is similar to the problem formulated in Chapter 3, except the additional cost term introduced due to congestion which affects the user level problem. The defender level and the attacker level problems are therefore structurally similar and identical to the formulations of a responsive contingency planning model developed in Section 3.3. We rewrite here the formulations of these two levels and develop a modified formulation for the user level problem.

4.2.2.1 Defender level problem

This layer of problem is concerned with optimal utilization of available budget to secure facilities and determine the level of responsiveness and volume of backup production. The model is written as follows:

$$\min_{z} \; H(z) \tag{4.7}$$

s.t.

$$\sum_{j\in J}\sum_{l\in L} c_{jl} z_{jl} \le B \tag{4.8}$$

$$\sum_{l\in L} z_{jl} \le 1 \qquad \forall j \in J \tag{4.9}$$

$$z_{jl} \in \{0\,,\,1\} \qquad \forall j \in J, l \in L \tag{4.10}$$

The defender objective function (4.7) minimizes the maximum network costs due to attacks by selecting facilities to protect with production backups. Constraint (4.8) mentions that protection budget cannot be exceeded. Constraint (4.9) ensures that every facility is protected with only one level of protection. Constraint (4.10) sets a binary restriction on protection variable.

4.2.2.2 Attacker level problem

The assumption made at this level is that the attacker knows the congestion related costs that the network may suffer under an optimal contingency strategy following attacks. Since the congestion costs are accounted at the user level, the attacker problem under congestion is structurally similar to the problem developed in Section 3.3. It involves attacker's objective of creating maximum network operation costs through the selection of optimal attack scenario s, given the protection decisions z of the defender. The model is written as follows:

$$H(z) = \max_{s} G(s,z) \tag{4.11}$$

s.t.

$$\sum_{j\in J} s_j = r \tag{4.12}$$

$$s_j + \sum_{l\in L} z_{jl} \le 1 \qquad \forall j \in J \tag{4.13}$$

$$s_j \in \{0, 1\} \qquad \forall j \in J \qquad (4.14)$$

Objective function (4.11) maximizes the network operation cost for the system user. Constraint (4.12) represents that attacker has capability of attacking only a finite number (*r*) of facilities. Constraint (4.13) prohibits attacks of protected facilities. Constraint (4.14) imposes binary restriction on the attack variable.

4.2.2.3 User level problem

The user level problem allocates demands to facilities in order to minimize the total costs of network operation, given the attack and protection solutions. The total costs include the transportation costs, the lost sales costs and the congestion costs. The model can be written as follows:

$$G(s,z) = \min_{x,u}\left(\sum_{i\in I}\sum_{j\in J}\sum_{t=1}^{T} d_{ij}x_{ijt} + \sum_{i\in I}\sum_{t=1}^{T} \beta_i u_{it} + \sum_{j\in J}\sum_{t=1}^{T} w\left(\sum_{i\in I} x_{ijt}\right)^q \right) \qquad (4.15)$$

s.t.

$$\sum_{j\in J} x_{ijt} + u_{it} = h_{it} \qquad \forall i \in I, t = 1...T \qquad (4.16)$$

$$\sum_{i\in I} x_{ijt} \le (1 - s_j)\left(v_{jt} + \sum_{l\in L} m_{tl}a_{jl}z_{jl}\right) \qquad \forall j \in J, t = 1...T \qquad (4.17)$$

$$x_{ijt} \ge 0 \qquad \forall i \in I, j \in J, t = 1...T \qquad (4.18)$$

$$u_{it} \ge 0 \qquad \forall i \in I, t = 1...T \qquad (4.19)$$

Objective function (4.15) minimizes the transportation costs, lost sales costs and the congestion costs of network operation given the protected and attacked facilities of the network. Congestion cost is expressed in the third term $\sum_{j\in J}\sum_{t=1}^{T} w\left(\sum_{i\in I} x_{ijt}\right)^q$, which is a convex and non-linear function, which is further explained in Section 4.2. Constraint (4.16) ensures that unmet demands are accounted as lost sales so that lost sales penalty can be applied. Constraint (4.17) ensures that the flows (allocated units) are no higher than the facility capacity including the base capacity and the backup production if that facility is protected. It also ensures that there is no allocation to an attacked facility. Constraint (4.18) and (4.19) are non-negativity constraints for allocation and lost sales decisions.

The compact version of the overall tri-level protection design model with congestion can be expressed as follows:

$$\min_z\left(\max_s\left(\min_{x,u}\sum_{i\in I}\sum_{j\in J}\sum_{t=1}^{T}d_{ij}x_{ijt}+\sum_{i\in I}\sum_{t=1}^{T}\beta_i u_{it}+\sum_{j\in J}\sum_{t=1}^{T}w\left(\sum_{i\in I}x_{ijt}\right)^q\right)\right) \qquad (4.20)$$

s.t. (4.8)-(4.10), (4.12)-(4.14), (4.16)-(4.19)

As the congestion cost term (third term) expressed in (4.20) is non-linear, it is substituted with a linear approximation function presented in Section 4.2. Using the linearized congestion cost function (4.5), the tri level model (4.20) is written as:

$$\min_z\left(\max_s\min_{x,u}\left(\sum_{i\in I}\sum_{j\in J}\sum_{t=1}^{T}d_{ij}x_{ijt}+\sum_{i\in I}\sum_{t=1}^{T}\beta_i u_{it}+\sum_{j\in J}\sum_{t=1}^{T}\max_k\left(w(1-q)\left(\sum_i x_{ijt}{}^k\right)^q+wq\left(\sum_i x_{ijt}{}^k\right)^{q-1}\left(\sum_i x_{ijt}\right)\right)\right)\right)$$

(4.21)

s.t. (4.8)-(4.10), (4.12)-(4.14), (4.16)-(4.19)

Replacing the congestion term in (4.21) with η_{jt} and adding constraint (4.23), the tri-level model (4.21) is equivalent to:

$$\min_z\left(\max_s\min_{x,u}\left(\sum_{i\in I}\sum_{j\in J}\sum_{t=1}^{T}d_{ij}x_{ijt}+\sum_{i\in I}\sum_{t=1}^{T}\beta_i u_{it}+w\sum_{j\in J}\sum_{t=1}^{T}\eta_{jt}\right)\right) \qquad (4.22)$$

s.t. (4.8)-(4.10), (4.12)-(4.14), (4.16)-(4.19)

$$\eta_{jt} - q\left(\sum_i x_{ijt}^k\right)^{q-1}\left(\sum_i x_{ijt}\right) \geq (1-q)\left(\sum_i x_{ijt}^k\right)^q \quad \forall j \in J, k \in K_j, t = 1...T \qquad (4.23)$$

We further note that it is necessary to ensure high costs of congestion do not cause the system to incur lost sales at the expense of unutilized capacity (in order to reduce total network costs). To this end, the constraint (4.24) is introduced in the user level problem to bound lost sales .Constraint (4.24) implies an upper bound of zero for lost sale units if total system capacity is higher than total demands and a positive difference if total demands are higher than total system capacity.

$$\sum_{i \in I} u_{it} \leq \max\left(0, \sum_{i \in I} h_{it} - \sum_{j \in J}\left(1-s_j\right)\left(v_{jt} + \sum_l m_{tl} a_{jl} z_{jl}\right)\right) \qquad (4.24)$$

A big M variable and a binary variable y_t are introduced to linearize (4.24) with the following three constraints:

$$My_t \geq \sum_{i \in I} h_{it} - \sum_{j \in J}\left(1-s_j\right)\left(v_{jt} + \sum_l m_{tl} a_{jl} z_{jl}\right) \qquad (4.24a)$$

$$M(1-y_t) \geq \sum_{j \in J}\left(1-s_j\right)\left(v_{jt} + \sum_l m_{tl} a_{jl} z_{jl}\right) - \sum_{i \in I} h_{it} \qquad (4.24b)$$

$$\sum_{i \in I} u_{it} - \left(\sum_{i \in I} h_{it} - \sum_{j \in J}\left(1-s_j\right)\left(v_{jt} + \sum_l m_{tl} a_{jl} z_{jl}\right)\right)\left(y_t\right) \leq 0 \qquad (4.24c)$$

Adding these sets of constraints, the complete tri-level game theoretic responsive contingency planning model (MILP) under congestion can therefore be expressed as follows:

$$\text{[TLM]: } \min_z\left(\max_s \min_{x,u}\left(\sum_{i \in I}\sum_{j \in J}\sum_{t=1}^T d_{ij} x_{ijt} + \sum_{i \in I}\sum_{t=1}^T \beta_i u_{it} + w\sum_{j \in J}\sum_{t=1}^T \eta_{jt}\right)\right) \qquad (4.25)$$

s.t. (4.8)-(4.10), (4.12)-(4.14), (4.16)-(4.19), (4.23), (4.24a)-(4.24c)

4.2.3 Solution Methodology

Implicit enumeration algorithm is applied on a binary search tree to solve the tri-level responsive contingency planning model under congestion effects [TLM]. The algorithm is similar to the one developed in Chapter 3, but we remove the cost normalization procedure, so that the impacts of congestion on protection decisions and its trade-off with the other costs can be better understood, as this is the main objective in this part of the dissertation.

With these changes, the enumeration tree corresponding to an illustrative problem involving five facilities located in five states in the US (*NY, CA, IL, TX* and *PA*) and two facility attacks (r = 2) is depicted in Figure 4.2. Every node of this enumeration tree is now characterized by a) total network cost (*cTot*) involving the sum of transportation cost, lost sales cost and the congestion cost b) attacked facilities (S) , which are the candidate facilities for protection in the next stage and c) the remaining budget of protection (B_{rem}). Note that in this part of the problem, the attacker's objective is to maximize the total costs as opposed to the normalized total costs which we considered in Chapter 3. The search tree progression, branching, pruning and fathoming of nodes are however, similar to the one discussed in Section 3.4, so its detailed explanation is omitted in this section.

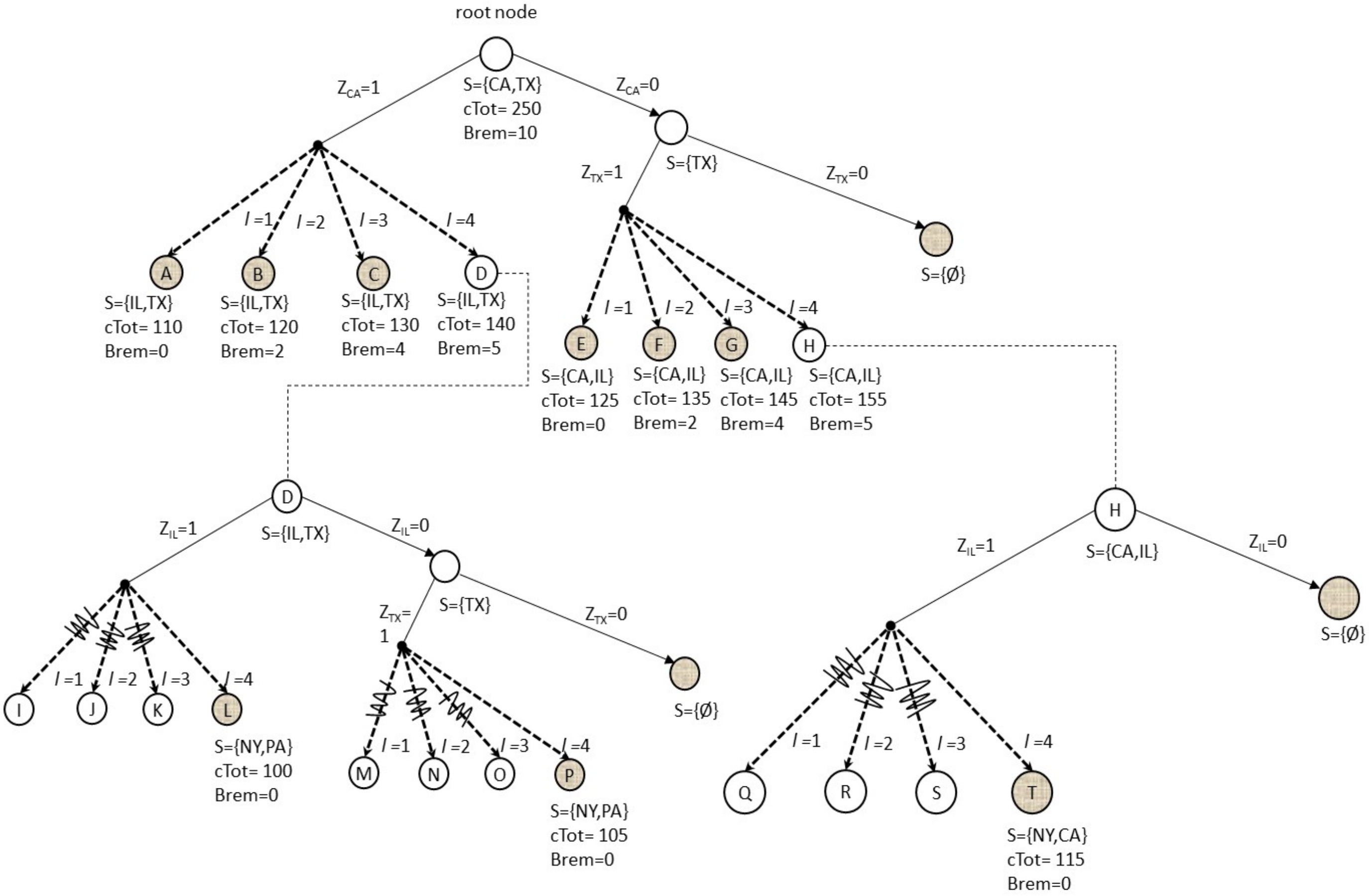


Figure 4.2: Congestion model solution binary search tree illustration

Considering two levels of capacity (high, low) and two levels of response speed (high, low) generates four different levels (*l*) of facility protection with backup production capability. In this example we assign ten units of budget (arbitrary) for protection and each level of protection consumes the budget as follows: level 1=10 units; level 2= 8 units; level 3= 6 units; level 4= 5 units. Under given budget and the costs of protection, this enumeration tree results in nine leaf nodes at its termination (A, *B*, *C*, E, *F*, *G*, *L*, *P* and *T*), each with a unique values of S *and cTot* obtained as solutions to the attacker problem. Backtracking from the leaf node with least total network cost, i.e. node L (with a *cTot*=100), we obtain optimal protection solution as securing both IL and CA with backup production at level 4, i.e., we select low volume capacity and low response speeds for backup production by securing these two facilities. Consequently, the attacker will interdict facilities NY and PA inorder to create the maximum possible cost to the system operator under this protection scenario.

4.3 Results and Analysis

In this section, computational results and managerial insights are presented. In this part of the dissertation, the experiments were solved using ILOG CPLEX 12.6 solver implemented using Java and Concert Technology on a Dell Latitude E5430 station with an Intel Core i5-3340 M processor at 2.7 GHz and 8 GB of RAM running Windows 7 operating system. The demands and distances data are derived from the US Census Bureau 2000 dataset (Daskin, 2004). The backup capacity volumes and response speeds and capacity addition costs are computed in a similar way as in Section 3.5 of Chapter 3. Facility base capacities are assumed identical and its computation also follows discussions on Section 3.5. The data are presented in Appendix A.

An illustrative example is presented using 15 demand nodes (*i*=15) and 10 facilities (*j*=10). The 10 facilities and the 15 demand nodes network used for this illustrative example is graphically represented in Figure 4.3 (small circles represent demand nodes and larger circles represent facilities; concentric circles represent the existence of both demands and facility at the same location). This network is constructed by ranking demand nodes by population size and opening of ten (*j*=10) facilities in the top ten demand zones. The network is utilized to demonstrate the flow allocation characteristic under congestion effects and highlight the significance of considering congestion in protection designs for disruption risk mitigations. We further investigate the trade-off between the congestion cost and other operational costs and analyze network performance under varying levels of attacks and congestion severity. The computational efficiency of the algorithm is tested on larger network sizes and the results are presented.

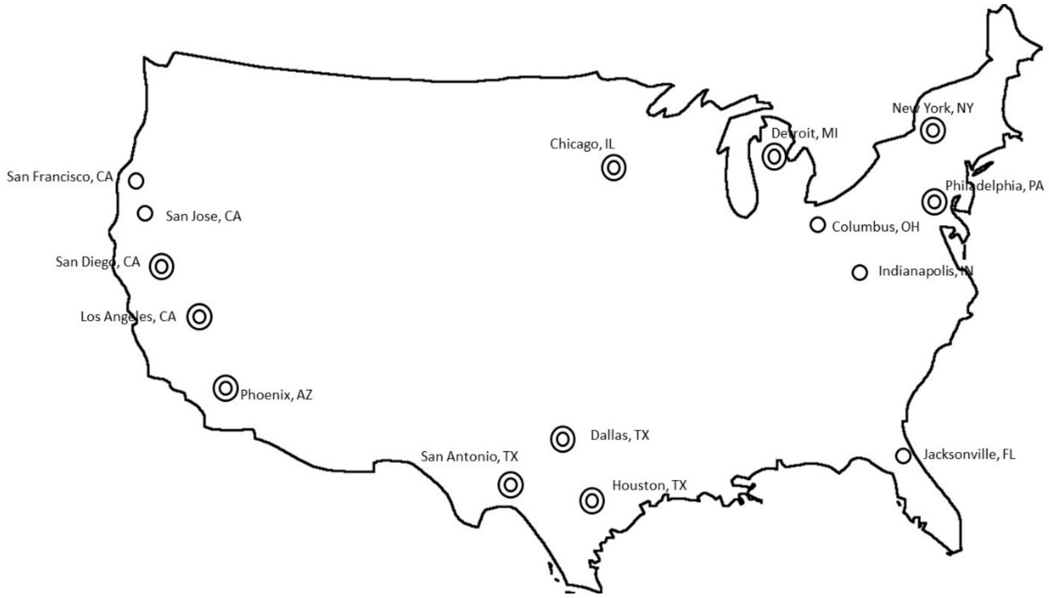


Figure 4.3: Demand and facility locations considered for congestion model analysis

4.3.1 Characteristic of supply flow allocations under congestion

The more demands are shifted to a facility, the higher is its congestion level. The allocations that are more balanced across facilities tend to lower costs of congestion. The supply flow allocation under congestion effect is demonstrated in Table 4.1. This Table lists the total flows allocated to facilities under both the traditional model which ignores congestion related costs and the model which incorporates congestion costs. The traditional model is based on the minimization of lost sales costs and the transportation costs in the ULP problem and ignores the impacts of congestion. The results are compared for varying congestion profiles i) low (*w*= 0.1, *q*=1.1) ii) medium (*w*=1, *q*=1.5) and iii) high (*w*=10, *q*=2.0), subject to a given budget (B=2680 monetary units) and attack level (*r*=1).

Table 4.1 Supply flow allocations of traditional and congestion model

Facilities	Traditional model	Congestion model		
		Low Congestion (w=0.1, q=1.1)	Medium Congestion (w=1, q=1.5)	High Congestion (w=10, q=2.0)
1.New York, NY	20652	20652	20652	12780
2.Los Angeles, CA	0	0	0	0
3.Chicago, IL	13652	13652	13652	12660
4.Houston, TX	10766	10766	10766	12540
5.Philadelphia,PA	16652	16652	16652	12780
6.Phoenix, AZ	13652	13652	13652	12660
7.San Diego, CA	13652	13652	13652	12660
8.Dallas, TX	8195	8195	7890	12540
9.San Antonio TX	6984	6984	7289	12472
10.Detroit, MI	9547	9547	9547	12660

Table 4.2 Comparison of solutions of traditional and congestion model

	Traditional model	Congestion model		
		Low Congestion (w=0.1, q=1.1)	Medium Congestion (w=1, q=1.5)	High Congestion (w=10, q=2.0)
Sj	2	2	2	2
Zjl	1(1), 5(4)	1(1), 5(4)	1(1), 5(4)	1(4), 3(4), 5(4), 10(4)
cTot	17,552,758	17,578,374	24,207,733	3,626,415,156
fRatio	2.96	2.96	2.83	1.02

It is observed from Table 4.1 that optimal supply allocations of a congestion model increasingly differ from the traditional model as congestion severity grows. For example, the allocated supplies of a traditional model and the congestion model are identical for all facilities under low levels of congestion, while they are different for two facilities, i.e., San Antonio, TX and Detroit, MA facilities under medium congestion level. Increasing congestion further (high congestion level), it is observed that the allocations of a congestion model and the traditional models differ for every surviving facilities.

Table 4.2, provides the optimal protection strategies (z_{jl}), resulting attacks (s_j), as well as total network costs (*cTot*) and maximum/minimum flow ratios (*fRatio*) obtained for this analysis under both traditional and congestion models. The flow ratio (*fRatio*) is a metric that is used to evaluate the balance of flows into facilities of the given network which affects the networks costs of congestion. A decreasing flow ratio with increasing congestion severity as demonstrated in this Table 4.2 indicate a more balanced flow allocations at higher congestion levels. As well, the optimal protection strategies under high congestion are observed to be different from traditional model while it is identical under low and medium levels of congestion. As balanced flow allocation ensure low costs of congestion, the protection design strategies that facilitate flow balance are generally preferable under high congestion severity. Therefore, designs which distribute available backup capacity in small units to many facilities, are desirable for flow balance and for minimizing the network congestion costs.

4.3.2 Value of incorporating congestion

In this section we investigate what compromise on the network costs will be made if the decision maker relies on conventional solutions and whether protection strategies and allocation decisions of a conventional model can be substituted for a model which considers congested network. The cost benefits signify the value of considering congestion in protection design models.

Under the same levels of budget of protection (B=2680) and attacks (r=1) we compute the congestion value index (φ) which indicates the relative change in the total network costs when the flow allocations and protection strategy of a traditional model ignoring congestion is utilized for designing protection of a network which can be congested. It can be expressed as:

$$\varphi = (cTot_{ts} - cTot_{cs})/cTot_{cs} \tag{4.26}$$

The term $cTot_{cs}$ in the above expression represents the total costs (optimal objective function value) obtained from the model that considers congestion. The term $cTot_{ts}$ represents the total cost obtained when the protection strategy and allocations of a traditional model is substituted for the problem involving congestion. In Table 4.3 we summarize the results for different congestion profiles.

Table 4.3 Variations in total costs under congestion model and traditional model for varying congestion profiles

Congestion profiles	*cTot_{ts}*	*cTot_{cs}*	*φ (%)*
w=0.1, q=1.1	17,578,374	17,578,374	0
w=1, q=1.1	17,808,919	17,808,919	0
w=10, q=1.1	20,114,369	20,114,369	0
w=0.1, q=1.5	18,218,284	18,218,284	0
w=1, q=1.5	24,208,019	24,207,733	0.002
w=10, q=1.5	84,105,367	83,917,369	0.224
w=0.1, q=2.0	57,452,123	57,051,093	0.703
w=1, q=2.0	416,546,404	389,161,124	7.037
w=10, q=2.0	4,007,489,218	3,626,415,156	10.508

In this illustrative example, the given network may incur up to 10% less in total costs by considering congestion effects. The higher the congestion severity, greater is the congestion value index. Therefore, when congestion associated costs are high, it may be cost effective to rely on protection designs and models that take congestion effects into account, whereas decision makers may rely on traditional models under low congestion severity.

4.3.3 Congestion cost trade-off analysis

A trade-off among congestion costs, costs of transportation and cost of lost sales is involved when designing protection under congestion. As the choice of parameters *w* and *q* affect congestion costs, we analyze the cost trade-offs with respect to these parameter variations. The protection strategies and the resulting costs under varying values of congestion parameters (*w*=0.1, 1, 10 and *q*=1.1, 1.5, 2.0, 2.5) are summarized in Table 4.4. The operational costs and congestion cost variations under these congestion profiles are graphically demonstrated in Figure 4.4a-f.

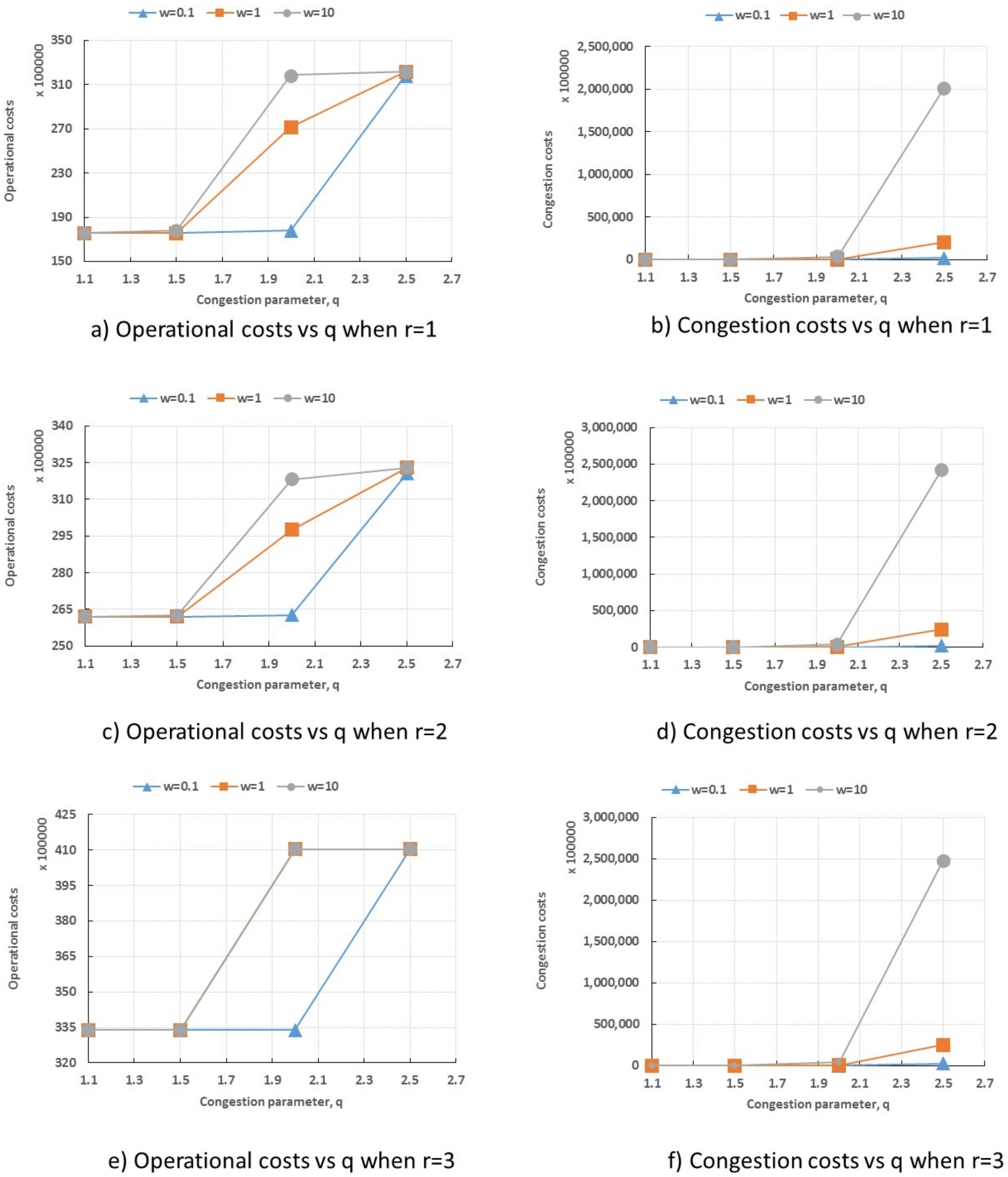
Congestion cost trades off with transportation costs if the network incurs no lost sales. This is the case when attacks are low (i.e., *r*=1, *r*=2 in Table 4.4) and where capacity losses can be fully compensated, either from the slack system capacity or through the available backup protection. The transportation costs are more dominant than the congestion costs for smaller *w* and *q*, i.e., low congestion severity. Under low congestion severity, the optimal protection strategies and flow allocations of a congestion model may not differ from a traditional model and both models can result in same total transportation costs. However, as can be observed in Table 4.4, as congestion severity increases (increasing *w* and *q*), the flow allocations of a congestion model start to deviate from the traditional model. As congestion costs become more dominant, the model tries to balance flows in order to reduce this cost, which leads to a lower flow ratio than obtained in a traditional model. Since some demands are shifted to facilities other than their closest ones, the network cost of transportation increases.

The congestion cost trade-off is with both the cost of lost sales and the transportation costs under increased attack levels. As more capacities are lost under increased attacks, the surviving facilities tend to be fully utilized and cost of lost sales may be incurred due to unmet demands. Reductions of congestion costs through demand shifts or network flow balance are impossible under this scenario. Recovering lost capacities through a high volume backup protection, i.e., centralized protection, is not a preferred protection strategy of such networks due to the risks of increased congestion costs for its unbalanced flow distribution. A simultaneous reduction of congestion costs and the lost sales can be generally achieved through a decentralized backup strategy. Such a strategy will yield low congestion costs for its more balanced flow distribution amongst the protected facilities.

As demonstrated in Table 4.4, a tendency towards decentralizing protection is therefore observed under higher attack levels and a limited protection budget. However, under high congestion severity (high *w* and *q*), marginal increase in congestion costs for unit flow recovered are much higher than decrease of lost sales costs through recovery. Under this condition, the congestion model may prescribe a protection strategy where recovered capacity volumes are lower compared to a decentralized protection. Such a strategy obviously raises the cost of lost sales but it is the dominance of congestion costs which makes this strategy favorable over a decentralized protection strategy. In the illustrative example, this scenario arises typically at congestion parameter settings of w=10 and q greater than or equal to 2. As demonstrated in Table 4.4, under w=10 and q=2.5 and r=3, the cost of lost sales of the selected strategy (z_{jl}: 1(3), 5(4), 6(4)) was about 40% higher than what can be achieved under a decentralized protection strategy (z_{jl}: 1(4), 2(4), 3(4), 5(4)) at this attack level. Nevertheless, this strategy was still selected for its lower congestion cost (about 5% lower than of the decentralized strategy).

Table 4.4 Cost trade-offs in optimal protection design

r	w	q	Sj	Zjl	cTot	cFlow	cLS	cCong	fRatio
1	0	-	2	1(1),5(4)	17,552,758	17,552,758	0	0	2.96
1	0.1	1.1	2	1(1),5(4)	17,578,374	17,552,758	0	25,616	2.96
1	0.1	1.5	2	1(1),5(4)	18,218,284	17,552,758	0	665,526	2.96
1	0.1	2	3	1(2),5(2)	57,051,093	17,797,897	0	39,253,196	2.53
1	0.1	2.5	2	1(4),3(4),5(4),10(4)	2,052,546,943	31,881,848	0	2,020,665,095	1.02
1	1	1.1	2	1(1),5(4)	17,808,919	17,552,758	0	256,161	2.96
1	1	1.5	2	1(1),5(4)	24,207,733	17,553,673	0	6,654,061	2.83
1	1	2	2	1(4),3(4),5(4),10(4)	389,161,124	27,192,724	0	361,968,400	1.26
1	1	2.5	2	1(4),3(4),5(4),10(4)	20,236,919,203	32,218,681	0	20,204,700,523	1.01
1	10	1.1	2	1(1),5(4)	20,114,369	17,552,758	0	2,561,611	2.96
1	10	1.5	2	1(1),5(4)	83,917,369	17,782,906	0	66,134,463	2.53
1	10	2	2	1(4),3(4),5(4),10(4)	3,626,415,156	31,882,196	0	3,594,532,960	1.02
1	10	2.5	2	1(4),3(4),5(4),10(4)	202,079,223,907	32,218,681	0	202,047,005,227	1.01
2	0	-	6,7	1(4),2(4),3(4),5(4)	26,194,341	26,194,341	0	0	1.45
2	0.1	1.1	6,7	1(4),2(4),3(4),5(4)	26,220,147	26,194,341	0	25,806	1.45
2	0.1	1.5	6,7	1(4),2(4),3(4),5(4)	26,879,754	26,194,341	0	685,413	1.45
2	0.1	2	6,7	1(4),2(4),5(4),10(4)	67,641,122	26,270,375	0	41,370,747	1.44
2	0.1	2.5	3,8	1(4),2(4),5(4),10(4)	2,450,288,382	32,055,666	0	2,418,232,716	1.09
2	1	1.1	6,7	1(4),2(4),3(4),5(4)	26,452,401	26,194,341	0	258,060	1.45
2	1	1.5	6,7	1(4),2(4),3(4),5(4)	33,048,473	26,194,341	0	6,854,132	1.45
2	1	2	3,8	1(4),2(4),5(4),10(4)	435,928,468	29,762,852	0	406,165,616	1.19
2	1	2.5	3,8	1(4),2(4),5(4),10(4)	24,213,087,019	32,289,098	0	24,180,797,921	1.09
2	10	1.1	6,7	1(4),2(4),3(4),5(4)	28,774,942	26,194,341	0	2,580,601	1.45
2	10	1.5	6,7	1(4),2(4),5(4),10(4)	94,671,602	26,251,052	0	68,420,550	1.44
2	10	2	3,8	1(4),2(4),5(4),10(4)	4,082,226,728	31,822,248	0	4,050,404,480	1.09
2	10	2.5	3,8	1(4),2(4),5(4),10(4)	241,840,268,307	32,289,098	0	241,807,979,209	1.09
3	0	-	6,7,10	1(4),2(4),3(4),5(4)	33,395,008	25,969,408	7,425,600	0	1.22
3	0.1	1.1	6,7,10	1(4),2(4),3(4),5(4)	33,419,578	25,969,408	7,425,600	24,570	1.22
3	0.1	1.5	6,7,10	1(4),2(4),3(4),5(4)	34,064,760	25,969,408	7,425,600	669,752	1.22
3	0.1	2	6,7,10	1(4),2(4),3(4),5(4)	75,225,321	25,969,408	7,425,600	41,830,313	1.22
3	0.1	2.5	3,8,10	1(3),5(4),6(4)	2,515,715,051	30,609,728	10,425,600	2,474,679,723	1.26
3	1	1.1	6,7,10	1(4),2(4),3(4),5(4)	33,640,711	25,969,408	7,425,600	245,703	1.22
3	1	1.5	6,7,10	1(4),2(4),3(4),5(4)	40,092,529	25,969,408	7,425,600	6,697,521	1.22
3	1	2	3,8,10	1(3),5(4),6(4)	440,407,860	30,609,728	10,425,600	399,372,532	1.26
3	1	2.5	3,8,10	1(3),5(4),6(4)	24,787,832,559	30,609,728	10,425,600	24,746,797,231	1.26
3	10	1.1	6,7,10	1(4),2(4),3(4),5(4)	35,852,035	25,969,408	7,425,600	2,457,027	1.22
3	10	1.5	6,7,10	1(4),2(4),3(4),5(4)	100,370,218	25,969,408	7,425,600	66,975,210	1.22
3	10	2	3,8,10	1(3),5(4),6(4)	4,034,760,648	30,609,728	10,425,600	3,993,725,320	1.26
3	10	2.5	3,8,10	1(3),5(4),6(4)	247,509,007,637	30,609,728	10,425,600	247,467,972,309	1.26
				min	17,552,758	17,552,758	0	0	1.01
				avg	19,542,093,340	26,063,887	2,788,320	19,513,241,133	1.55
				max	247,509,007,637	32,289,098	10,425,600	247,467,972,309	2.96



a) Operational costs vs q when r=1

b) Congestion costs vs q when r=1

c) Operational costs vs q when r=2

d) Congestion costs vs q when r=2

e) Operational costs vs q when r=3

f) Congestion costs vs q when r=3

Figure 4.4: Operational costs and Congestion costs under varying attack levels and congestion profile settings

4.3.4 Computational efficiency

The computational efficiency of the algorithm is assessed under a larger network involving 50 demand nodes derived from the US Census Bureau dataset (see Appendix). Table 4.5 summarizes the CPU times obtained under different parameter settings of attack levels *(r*), number of facilities (*J*), and the congestion profile parameters (*w, q*) under a finite budget level (B). These results demonstrate that under a finite budget of protection, the CPU times generally increase with both the increase in the number of attacks and the number of facilities. These results are consistent with the observations made in Chapter 3.

We further observe that the computational efficiency is affected by incorporating congestion term in the protection design model. The linear approximations of convex non-linear congestion functions result in the addition of a set of constraints in the model. When the desired approximation error is low, interval granularity (*K*) is high i.e., the congestion cost function is approximated using an increased number of tangent hyperplanes. This has the effect of increasing problem size because of the new constraints that are added in the model. Table 4.6 highlights the CPU time and the relative error of congestion approximation under different interval granularity (K) for different problem combinations and a finite budget of protection (*B*=2010 monetary units). It can be observed that as K increases, there is more accuracy in linear approximation of the congestion cost function, but it significantly deteriorates the CPU time.

Table 4.5 CPU time variations under different parameter combinations

r	w	q	CPU time (s)		
			j=5	j=10	j=15
1	0.1	1.1	11	59	176
1	0.1	1.5	11	58	175
1	0.1	2	12	60	164
1	0.1	2.5	14	63	198
1	1	1.1	14	82	257
1	1	1.5	14	57	207
1	1	2	8	46	158
1	1	2.5	14	63	186
1	10	1.1	10	59	177
1	10	1.5	11	60	174
1	10	2	8	46	180
1	10	2.5	13	65	186
2	0.1	1.1	31	517	2544
2	0.1	1.5	28	542	2919
2	0.1	2	23	425	2170
2	0.1	2.5	32	586	2770
2	1	1.1	29	608	3685
2	1	1.5	27	516	2791
2	1	2	23	395	2332
2	1	2.5	32	637	2420
2	10	1.1	31	523	2550
2	10	1.5	33	529	2604
2	10	2	24	437	2285
2	10	2.5	37	551	2767
3	0.1	1.1	20	1953	19536
3	0.1	1.5	21	1953	24461
3	0.1	2	13	1394	14969
3	0.1	2.5	22	2169	18687
3	1	1.1	20	25	2752
3	1	1.5	21	2186	22126
3	1	2	13	1461	14679
3	1	2.5	24	2138	20917
3	10	1.1	20	2164	19342
3	10	1.5	21	2039	19628
3	10	2	13	1460	15185
3	10	2.5	22	2360	20791
		Min	8	25	158
		Avg	20	785	6865
		Max	37	2360	24461

Table 4.6 Congestion approximation errors and CPU times

Instances	CPU time (s)				Congestion approximation error (%)			
(I_J_r)	K=25	K=50	K=75	K=100	K=25	K=50	K=75	K=100
50_5_1	3	7	11	13	0.012	0.007	0.006	0.002
50_5_2	8	18	27	37	0.008	0.007	0.003	0.001
50_5_3	7	14	18	22	0.020	0.007	0.006	0.001
50_10_1	16	36	48	65	0.148	0.039	0.017	0.009
50_10_2	140	304	435	551	0.046	0.024	0.004	0.001
50_10_3	535	1176	1686	2360	0.048	0.021	0.003	0.001
50_15_1	41	96	150	186	0.241	0.059	0.027	0.016
50_15_2	559	1273	1995	2767	0.205	0.055	0.023	0.013
50_15_3	4262	9592	14593	20791	0.039	0.030	0.011	0.002

As can be observed in the above results, the average computational time of the test instances on a 50 node network across all combinations was 6865 seconds. A maximum time of 24461 seconds resulted under 3 facility attacks and 15 facilities. Although the CPU times are high, considering that the protection design problem presented here is a strategic decision problem the computational efforts of the proposed solution algorithm is acceptable. We remark that the need to incorporate more elements of resilient design in an already difficult nested optimization modeling framework, poses additional computational burden. In this analysis we limited the attacks to a few facilities considering that simultaneous attacks on many facilities is rare. Also we have limited the budget of protection which limits the depth of our search tree. Nevertheless, the proposed methodology has extended the scope of implicit enumeration algorithm and the bi-level fortification interdiction problems.

4.4 Chapter Summary

In this Chapter, we extended the responsive contingency planning model to incorporate congestion effects. The incorporation of congestion makes the model more robust to operate under realistic situations. A non-linear congestion cost function was developed to model congestion related costs due to increased flow volume on facilities. This convex non-linear function was then linearized applying piecewise linearization technique in which a set of tangent hyperplanes determined the linear congestion costs under varying flow levels at facilities. The incorporation of linear congestion cost function in the responsive contingency model enabled its formulation as a MILP. The tri-level responsive contingent planning model under congestion was then solved applying implicit enumeration technique. Computational efficiency of the algorithm was demonstrated using a 50 node (demand) network under varying number of facilities and attack levels. The impact of interval granularity (K) for linear approximation of the congestion cost function was analyzed as this significantly affected the computation time. It is demonstrated that as K increases, the relative errors of linear approximations are low but CPU time grows significantly.

This chapter has demonstrated that congestion affects design decisions of protection. We have demonstrated that the supply flow allocation which is more balanced is optimal under congestion effects as it lower over utilization of a single facility (and hence congestion). This allocation requirement implies that traditional model solutions cannot be relied when congestion effects are significant. We demonstrate the value of congestion through empirical study comparing network total costs of a congestion incorporated model to a conventional (congestion impacts ignored) model. It was observed that when supply network's congestion associated costs are high, it is better to rely on protection designs and models that take this factor into account, even though one may substitute traditional solutions when congestion effects are low or negligible.

The trade-off of congestion costs with the operational cost of transportation and lost sales were investigated and the optimal protection strategies under varying congestion severity were analyzed. The results demonstrate that decentralized protection is generally a preferred strategy under congestion conditions. However, when congestion severity is very high, the marginal increase of congestion cost for unit flow recovery through production backup is much higher than the decrease in the cost of lost sales. Under this condition it may be appropriate to centralize protection in order to keep the total network costs low.

Chapter Five: **Conclusion and Future works**

5.1 Summary

This dissertation proposes new mathematical models and appropriate mechanisms to address supply chain disruption risk mitigation and management problem. It focuses on enhancing supply chain responsiveness under disruption through contingency planning involving backup sourcing. In the first part of this dissertation, a game theoretic mathematical model is developed for generating contingency strategy involving appropriate volume and response speeds of a backup resource and protection of critical supply facilities. The major contribution of this model is in having a proper representation of the available backup resource capacity during disruption periods which enable more effective risk mitigation solutions. An illustrative example involving single facility attack is used to demonstrate that instantaneous production capacity assumptions lead to capacity overestimations and inappropriate mitigation solutions and inadequate recourse actions during major disruptions. Furthermore, through this example the demand routing characteristic optimal under a capacitated system is demonstrated. In particular, the importance of flow re-allocations rather than contingent re-routing in a capacitated network is established.

The developed model is tested on two hypothetical networks with different initial capacity distributions to demonstrate that network topologies affect its vulnerability to disruptions and contingency strategies of protection can be different under varying network configurations. It is demonstrated that identical capacity distributions result in risk diversification effect, and are effective for controlling losses under a major disruptions. However, the relative efficiency improvements through protection are lower for such networks than networks with non-identical capacity distributions in its facilities. Through this analysis, it is observed that planning against a more capable adversary (higher number of attacks), necessitate the need to disperse or decentralize protection in order to spread out risk.

A solution methodology based on implicit enumeration algorithm utilizing the structure of the problem is developed in this part of the dissertation for solving the proposed model. The methodology extends the scope of implicit enumeration algorithm in handling multi-level of protections. The computational performance of the developed algorithm is assessed with respect to the variations in the budget of protection, number of attacks, number of facilities and number of demands. It is observed that computational time increase with the increase of budget level, attack levels, number of facilities or the number of demands. Raising the protection budget or the attack level lead to increase in the size of the search tree which affect the computational time. As well, increasing number of facilities and the number of demands both raise computation time but the effect of raising the number of facilities is higher than the effect of raising the number of demands.

In the second part of the dissertation, we analyze the impact of congestion on responsive contingency planning against major disruptions. In many supply systems, congestion related costs are severe and disruptions tend to raise such costs. Congestion effects need to be considered during the planning stages. We therefore extend the responsive contingency planning model developed in the first part of the dissertation to incorporate congestion effects. Incorporation of congestion effects makes the proposed model even more robust to handling major disruptions. Through empirical study we demonstrate that congestion affects design decisions of protection. A significant savings in network costs can result by relying on models that explicitly incorporate congestion effects than the models that ignore congestion. A centralized protection which adds high volume back up capacity faster at higher response speeds is generally not desirable under congestion of the network, even though such protection plans can be utilized when there is no congestion.

5.2 Limitations and Future directions

This dissertation leads to interesting future avenues which include the extension of the proposed model, improvements on solution methods and study of related problems. An obvious extension to the proposed model would be to relax some of the assumptions. For example, the considered problem assumes that facilities protected, i.e., where backup production is planned, never loses its base capacity. In practice, achieving facilities completely immune to disruptions is difficult. Relaxing this assumption will make our model more realistic. Further, focusing more on the recovery aspect or the contingent mechanism of protection, the model lumps the cost of protection into the cost of backup selection. Although this assumption is not restrictive, it may be possible to segregate budget into hardening or security of facilities and backup up production and let the model decide where to invest on security and where to invest on backup production. Few other model extensions can be: a) integrating location decisions with protection decisions for design of new networks b) partial interdictions (i.e. attack not 100% successful) than complete interdiction c) probabilistic disruption considerations rather than deterministic

In the solution methodology, the proposed approach has extended the scope of implicit enumeration approach in handling multiple levels of protection as earlier works considered protection at single level only. Although computation times grow, this approach can be utilized in solving larger problem instances since the responsive capacity planning problem dealt in this dissertation is of strategic nature. When decision makers are more risk averse, they tend to implement solutions that are robust under more severe attacks and larger network topologies involving larger protection budgets. Since the search tree grows with these considerations, it may be worthwhile to develop reduction rules to reduce the size of the search tree. The use of heuristic rules within the implicit enumeration scheme would reduce the computation time and therefore enhance the applicability of this methodology.

A related future work to this dissertation can involve areas such as emergency relief, and crime protection, infectious disease spread-out protection, etc. In these areas, response and recovery aspects are critical. The decisions may involve how to preposition shelters, inventories, etc or how to deploy security forces so that disruption responses are efficient. The backup production volume and the response speed factors under these circumstances can be viewed as sets of activities each involving a response time, and therefore decisions would involve selection of appropriate activities and their proper speeds of response. The model developed in this dissertation can therefore be successfully adapted in such application areas.

References

1. Aboolian, R., Cui, T., & Shen, Z. J. M. (2012). An efficient approach for solving reliable facility location models. *INFORMS Journal on Computing*, *25*(4), 720-729.

2. Aksen, D., Akca, S. S., & Aras, N. (2014). A bilevel partial interdiction problem with capacitated facilities and demand outsourcing. *Computers & Operations Research, 41*, 346-358.

3. Aksen, D., & Aras, N. (2012). A bilevel fixed charge location model for facilities under imminent attack. *Computers & Operations Research, 39*(7), 1364-1381.

4. Aksen, D., Piyade, N., & Aras, N. (2010). The budget constrained r-interdiction median problem with capacity expansion. *Central European Journal of Operations Research, 18*(3), 269-291.

5. Amiri, A. (1997). Solution procedures for the service system design problem. *Computers & operations research*, *24*(1), 49-60.

6. Berman, O., Krass, D., & Menezes, M. B. (2007). Facility reliability issues in network p-median problems: strategic centralization and co-location effects. *Operations Research*, *55*(2), 332-350.

7. Berman, O., Krass, D., & Menezes, M. B. (2009). Locating facilities in the presence of disruptions and incomplete information. *Decision Sciences*, *40*(4), 845-868.

8. Bitran, G. R., & Tirupati, D. (1989). Tradeoff curves, targeting and balancing in manufacturing queueing networks. *Operations Research*, *37*(4), 547-564.

9. Bricha, N., & Nourelfath, M. (2013). Critical supply network protection against intentional attacks: A game-theoretical model. *Reliability Engineering & System Safety*, *119*, 1-10.

10. Brown, G., Carlyle, M., Salmerón, J., & Wood, K. (2006). Defending critical infrastructure. *Interfaces, 36*(6), 530-544.

11. Camargo, R. S., Miranda Jr, G., Ferreira, R. P. M., & Luna, H. P. (2009). Multiple allocation hub-and-spoke network design under hub congestion. *Computers & Operations Research*, *36*(12), 3097-3106.

12. Cappanera, P., & Scaparra, M. P. (2011). Optimal allocation of protective resources in shortest-path networks. *Transportation Science, 45*(1), 64-80.

13. Chopra, S., & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. *MIT Sloan management review*, *46*(1), 53.

14. Chopra, S., & Sodhi, M. S. (2014). Reducing the risk of supply chain disruptions. *MIT Sloan Management Review*, *55*(3), 73.

15. Church, R. L., Scaparra, M. P., & Middleton, R. S. (2004). Identifying critical infrastructure: The median and covering facility interdiction problems. *Annals of the Association of American Geographers, 94*(3), 491-502.

16. Church, R. L., & Scaparra, M. P. (2007a). Analysis of facility systems' reliability when subject to attack or a natural disaster. *Critical infrastructure* (pp. 221-241). Berlin Heidelberg: Springer.

17. Church, R. L., & Scaparra, M. P. (2007b). Protecting critical assets: The r-interdiction median problem with fortification. *Geographical Analysis, 39*(2), 129-146.

18. Cui, T., Ouyang, Y., & Shen, Z.M. (2010). Reliable facility location design under the risk of disruptions. *Operations Research, 58*(4-part-1), 998-1011.

19. Daskin, M. S. (2004). SITATION—facility location software. *Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.*

20. Elhedhli, S., & Hu, F. X. (2005). Hub-and-spoke network design with congestion. *Computers & Operations Research*, *32*(6), 1615-1632.

21. Fulkerson, D. R., & Harding, G. C. (1977). Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming*, *13*(1), 116-118.

22. Golany, B., Kaplan, E. H., Marmur, A., & Rothblum, U. G. (2009). Nature plays with dice–terrorists do not: Allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Researc*h, 192(1), 198-208.

23. Hausken, K. (2011). Protecting complex infrastructures against multiple strategic attackers. *International Journal of Systems Science*, *42*(1), 11-29.

24. Hendricks, K. B., & Singhal, V. R. (2005). An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production and Operations management*, *14*(1), 35-52.

25. Higle, J. L. (2005). Stochastic programming: Optimization when uncertainty matters. *Tutorials in operations research*, *3053*.

26. Hopp, W. J., & Yin, Z. (2006). Protecting supply chain networks against catastrophic failures. *Working Paper, Dept. of Industrial Engineering and Management Science, Northwestern University, Evanston, IL.*

27. Hopp, W. J., Iravani, S. M., & Liu, Z. (2012). Mitigating the impact of disruptions in supply chains. In *Supply Chain Disruptions* (pp. 21-49). Springer, London.

28. Israeli, E., & Wood, R. K. (2002). Shortest-path network interdiction. *Networks*, *40*(2), 97-111.

29. Jalali, S., Seifbarghy, M., & Niaki, S. T. A. (2018). A risk-averse location-protection problem under intentional facility disruptions: A modified hybrid decomposition algorithm. *Transportation Research Part E: Logistics and Transportation Review*, *114*, 196-219.

30. Karmarkar, U. S. (1989). Capacity loading and release planning with work-in-progress (WIP) and leadtimes. *Journal of Manufacturing and Operations Management*, *2*(105-123).

31. Kim, S. (2013). A column generation heuristic for congested facility location problem with clearing functions. *Journal of the Operational Research Society*, *64*(12), 1780-1789.

32. Klibi, W., Martel, A., & Guitouni, A. (2010). The design of robust value-creating supply chain networks: A critical review. *European Journal of Operational Research, 203*(2), 283-293.

33. Klibi, W., & Martel, A. (2012). Modeling approaches for the design of resilient supply networks under disruptions. *International Journal of Production Economics, 135*(2), 882-898.

34. Konur, D., & Geunes, J. (2011). Analysis of traffic congestion costs in a competitive supply chain. *Transportation Research Part E: Logistics and Transportation Review*, *47*(1), 1-17.

35. Konur, D., & Geunes, J. (2012). Competitive multi-facility location games with non-identical firms and convex traffic congestion costs. *Transportation Research Part E: Logistics and Transportation Review*, *48*(1), 373-385.

36. Latour, A. (2001). Trial by fire: A blaze in Albuquerque sets off major crisis for cell-phone giants. *Wall Street Journal*, *1*(29), 2001.

37. Lee, S. D. (2001). On solving unreliable planar location problems. *Computers & Operations Research*, *28*(4), 329-344.

38. Li, X., & Ouyang, Y. (2010). A continuum approximation approach to reliable facility location design under correlated probabilistic disruptions. *Transportation research part B: methodological*, *44*(4), 535-548.

39. Li, Q., Zeng, B., & Savachkin, A. (2013). Reliable facility location design under disruptions. *Computers & Operations Research*, *40*(4), 901-909.

40. Liberatore, F., & Scaparra, M. P. (2011). Optimizing protection strategies for supply chains: comparing classic decision-making criteria in an uncertain environment. *Annals of the Association of American Geographers*, *101*(6), 1241-1258.

41. Liberatore, F., Scaparra, M. P., & Daskin, M. S. (2011). Analysis of facility protection strategies against an uncertain number of attacks: The stochastic R-interdiction median problem with fortification. *Computers & Operations Research, 38*(1), 357-66.

42. Liberatore, F., Scaparra, M. P., & Daskin, M. S. (2012). Hedging against disruptions with ripple effects in location analysis. *Omega, 40*(1), 21-30.

43. Lim, M., Daskin, M. S., Bassamboo, A., & Chopra, S. (2010). A facility reliability problem: Formulation, properties, and algorithm. *Naval Research Logistics, 57*(1), 58-70.

44. Losada, C., Scaparra, M. P., & Church, R. L. (2010 a). Interdiction of p-median systems with facility recovery time and disruptions frequency: analysis of resiliency. KBS Working Paper no. 220, University of Kent, UK.

45. Losada, C., Scaparra, M. P., & Church, R. L. (2010b). On a bi-level formulation to protect uncapacitated p-median systems with facility recovery time and frequent disruptions. *Electronic Notes in Discrete Mathematics*, *36*, 591-598.

46. Losada, C., Scaparra, M. P., Church, R. L., & Daskin, M. S. (2012a). The stochastic interdiction median problem with disruption intensity levels. *Annals of Operations Research, 201*, 345-65.

47. Losada, C., Scaparra, M. P., & O'Hanley, J. R. (2012b). Optimizing system resilience: A facility protection model with recovery time. *European Journal of Operational Research, 217*(3), 519-30.

48. Marianov, V., & Serra, D. (2003). Location models for airline hubs behaving as M/D/c queues. *Computers & Operations Research*, *30*(7), 983-1003.

49. Moore, J. T., & Bard, J. F. (1990). The mixed integer linear bilevel programming problem. Operations research, 38(5), 911-921.

50. Nejad, A. E., Niroomand, I., & Kuzgunkaya, O. (2014). Responsive contingency planning in supply risk management by considering congestion effects. *Omega, 48*, 19-35.

51. Niroomand, I., Kuzgunkaya, O., & Bulgak, A. A. (2012). Impact of reconfiguration characteristics for capacity investment strategies in manufacturing systems. *International Journal of Production Economics, 139*(1), 288-301.

52. Nooraie, S. V., & Parast, M. M. (2016). Mitigating supply chain disruptions through the assessment of trade-offs among risks, costs and investments in capabilities. *International Journal of Production Economics*, *171*, 8-21.

53. O'Hanley, J. R., & Church, R. L. (2011). Designing robust coverage networks to hedge against worst-case facility losses. *European Journal of Operational Research*, *209*(1), 23-36.

54. Parajuli, A., Kuzgunkaya, O., & Vidyarthi, N. (2017). Responsive contingency planning of capacitated supply networks under disruption risks. *Transportation Research Part E: Logistics and Transportation Review*, *102*, 13-37.

55. Peng, P., Snyder, L. V., Lim, A., & Liu, Z. (2011). Reliable logistics networks design with facility disruptions. *Transportation Research Part B: Methodological*, *45*(8), 1190-1211.

56. Putnik, G., Sluga, A., ElMaraghy, H., Teti, R., Koren, Y., Tolio, T., & Hon, B. (2013). Scalability in manufacturing systems design and operation: State-of-the-art and future developments roadmap. CIRP Annals, 62(2), 751-774.

57. Qi, L., Shen, Z. J. M., & Snyder, L. V. (2010). The effect of supply disruptions on supply chain design decisions. *Transportation Science*, *44*(2), 274-289.

58. Scaparra, M. P., & Church, R. L. (2008a). A bilevel mixed-integer program for critical infrastructure protection planning. Computers & Operations Research, 35(6), 1905-1923.

59. Scaparra, M. P., & Church, R. L. (2008b). An exact solution approach for the interdiction median problem with fortification. *European Journal of Operational Research, 189*(1), 76-92.

60. Scaparra, M. P., & Church, R. L. (2012). Protecting supply systems to mitigate potential disaster: A model to fortify capacitated facilities. *International Regional Science Review, 35*(2), 188-210.

61. Schmitt, A. J. (2011). Strategies for customer service level protection under multi-echelon supply chain disruption risk. *Transportation Research Part B: Methodological, 45*(8), 1266-1283.

62. Sheffi, Y., & Rice Jr, J. B. (2005). A supply chain view of the resilient enterprise. *MIT Sloan management review*, *47*(1), 41.

63. Shen, Z. J. M., Zhan, R. L., & Zhang, J. (2011). The reliable facility location problem: Formulations, heuristics, and approximation algorithms. *INFORMS Journal on Computing*, *23*(3), 470-482.

64. Simchi-Levi, D., Snyder, L., & Watson, M. (2002). Strategies for uncertain times. *Supply Chain Management Review*, *6*(1), 11-12.

65. Simchi-Levi, D., Schmidt, W., & Wei, Y. (2014). From superstorms to factory fires: Managing unpredictable supply chain disruptions. *Harvard Business Review*, *92*(1-2), 96-101.

66. Simison, R. L. (1998). GM contains its quarterly loss at $809 million. *Wall Street Journal*, *24*.

67. Snyder, L. V., & Daskin, M. S. (2005). Reliability models for facility location: the expected failure cost case. *Transportation Science*, *39*(3), 400-416.

68. Snyder, L. V., & Daskin, M. S. (2006). Stochastic p-robust location problems. *IIE Transactions*, *38*(11), 971-985.

69. Snyder, L. V., & Daskin, M. S. (2007). Models for reliable supply chain network design. In *Critical Infrastructure* (pp. 257-289). Springer, Berlin, Heidelberg.

70. Tomlin, B. (2006). On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science, 52*(5), 639-657.

71. Unit, E. I. (2005). Business 2010–Embracing the challenge of change. *The Economist*.

72. Vakharia, A. J., & Yenipazarli, A. (2009). Managing supply chain disruptions. *Foundations and Trends® in Technology, Information and Operations Management*, *2*(4), 243-325.

73. Vidyarthi, N., & Kuzgunkaya, O. (2015). The impact of directed choice on the design of preventive healthcare facility network under congestion. *Health care management science*, *18*(4), 459-474.

74. Wang, Q., Batta, R., & Rump, C. M. (2002). Algorithms for a facility location problem with stochastic customer demand and immobile servers. *Annals of operations Research*, *111*(1-4), 17-34.

75. Wang, W., & Koren, Y. (2012). Scalability planning for reconfigurable manufacturing systems. *Journal of Manufacturing Systems*, *31*(2), 83-91.

76. Weisbrod, G., Vary, D., & Treyz, G. (2001). Economic implications of congestion. NCHRP Report #463. National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.

77. Wollmer, R. (1964). Removing arcs from a network. *Operations Research, 12*, 934-40.

78. Wood, R. K. (1993). Deterministic network interdiction. *Mathematical and Computer Modeling, 17*, 1-18.

79. Zhang, Y., Berman, O., & Verter, V. (2009). Incorporating congestion in preventive healthcare facility network design. *European Journal of Operational Research*, *198*(3), 922-935.

APPENDIX A: DISTANCE, DEMANDS AND FACILITY BASE CAPACITY DATA INPUTS

A.1. Distance and demand quantities data inputs (Chapter 3 and Chapter 4)

		New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego	Dallas	San Antonio	Detroit	San Jose	Indianapolis	San Francisco	Jacksonville	Columbus	Austin	Memphis	Baltimore	Milwaukee	Boston
City	Demand	NY	CA	IL	TX	PA	AZ	CA	TX	TX	MI	CA	IN	CA	FL	OH	TX	TN	MD	WI	MA
New York NY	8104	1	2458	718	1421	78	2142	2429	1372	1584	489	2554	647	2573	836	480	1513	956	171	740	191
Los Angeles CA	3805	2458	1	1747	1381	2399	367	116	1250	1211	1985	294	1815	340	2154	1983	1235	1612	2324	1746	2601
Chicago IL	2935	718	1747	1	939	667	1446	1726	800	1049	238	1836	164	1855	864	277	975	482	607	86	854
Houston TX	2029	1421	1381	939	1	1345	1015	1300	225	189	1108	1605	868	1644	822	995	147	485	1253	1007	1607
Philadelphia PA	1523	78	2399	667	1345	1	2079	2367	1300	1509	446	2501	585	2521	764	417	1438	882	93	697	268
Phoenix AZ	1396	2142	367	1446	1015	2079	1	298	887	847	1683	609	1495	652	1792	1663	869	1262	1999	1457	2296
San Diego CA	1260	2429	116	1726	1300	2367	298	1	1182	1125	1963	409	1783	456	2087	1951	1154	1559	2290	1730	2578
Dallas TX	1234	1372	1250	800	225	1300	887	1182	1	253	998	1449	764	1485	906	913	181	420	1211	856	1550
San Antonio TX	1189	1584	1211	1049	189	1509	847	1125	253	1	1238	1448	1000	1488	1011	1141	75	632	1418	1107	1767
Detroit MI	960	489	1985	238	1108	446	1683	1963	998	1238	1	2069	240	2087	837	166	1164	626	401	252	617
San Jose CA	922	2554	294	1836	1605	2501	609	409	1449	1448	2069	1	1926	47	2340	2090	1462	1775	2433	1821	2681
Indianapolis IN	804	647	1815	164	868	585	1495	1783	764	1000	240	1926	1	1948	701	169	926	387	510	246	808
San Francisco CA	800	2573	340	1855	1644	2521	652	456	1485	1488	2087	47	1948	1	2373	2112	1501	1805	2455	1838	2699
Jacksonville FL	762	836	2154	864	822	764	1792	2087	906	1011	837	2340	701	2373	1	672	960	588	683	947	1019
Columbus OH	715	480	1983	277	995	417	1663	1951	913	1141	166	2090	169	2112	672	1	1067	512	343	334	644
Austin TX	682	1513	1235	975	147	1438	869	1154	181	75	1164	1462	926	1501	960	1067	1	559	1347	1034	1695
Memphis TN	666	956	1612	482	485	882	1262	1559	420	632	626	1775	387	1805	588	512	559	1	792	561	1137
Baltimore MD	664	171	2324	607	1253	93	1999	2290	1211	1418	401	2433	510	2455	683	343	1347	792	1	645	360
Milwaukee WI	605	740	1746	86	1007	697	1457	1730	856	1107	252	1821	246	1838	947	334	1034	561	645	1	862
Boston MA	598	191	2601	854	1607	268	2296	2578	1550	1767	617	2681	808	2699	1019	644	1695	1137	360	862	1
El Paso TX	585	1899	712	1243	672	1831	348	629	569	500	1472	954	1259	995	1468	1423	526	973	1746	1271	2066
Nashville TN	584	761	1786	395	667	687	1442	1739	615	824	473	1934	252	1962	501	335	752	196	597	481	944
Denver CO	583	1626	843	910	876	1571	587	835	661	799	1146	933	993	956	1461	1158	768	876	1501	904	1762
Seattle WA	578	2409	957	1734	1891	2375	1111	1060	1683	1786	1932	714	1870	681	2454	2010	1770	1869	2330	1686	2490
Washington DC	578	204	2304	597	1221	127	1977	2269	1182	1387	399	2417	492	2439	649	328	1317	763	35	639	394
Charlotte NC	560	533	2126	590	928	456	1779	2076	928	1105	512	2274	431	2301	341	352	1039	519	367	665	723
Fort Worth TX	555	1403	1218	822	237	1330	854	1150	34	238	1024	1419	790	1455	938	941	171	451	1241	876	1580
Portland OR	542	2444	822	1753	1834	2406	1002	928	1633	1717	1961	572	1882	538	2437	2029	1707	1850	2356	1712	2536
Las Vegas NV	518	2236	232	1523	1231	2178	257	260	1076	1075	1760	375	1596	414	1972	1763	1088	1414	2105	1519	2376
Tucson AZ	514	2120	455	1438	933	2054	116	367	824	759	1673	715	1474	759	1727	1641	787	1216	1973	1456	2280
Oklahoma City OK	514	1328	1189	689	413	1260	840	1137	191	420	909	1357	689	1389	985	852	357	424	1176	732	1495
New Orleans LA	489	1159	1686	824	328	1081	1320	1612	445	517	932	1893	705	1930	495	790	468	349	989	905	1349
Cleveland OH	481	408	2054	311	1115	358	1744	2028	1024	1257	96	2146	263	2166	771	124	1183	631	307	340	552
Long Beach CA	475	2454	27	1745	1364	2394	351	90	1236	1193	1982	320	1810	366	2141	1978	1218	1602	2319	1745	2599
Albuquerque NM	455	1813	674	1123	752	1749	330	624	588	615	1359	862	1166	898	1485	1334	614	939	1670	1138	1969
KS City MO	448	1096	1365	408	649	1035	1047	1333	455	705	640	1483	451	1508	950	619	636	374	960	438	1249
Fresno CA	440	2460	201	1743	1487	2405	491	313	1333	1329	1977	120	1827	162	2227	1993	1343	1664	2336	1731	2592
VA Beach VA	437	295	2375	715	1215	232	2037	2332	1206	1391	542	2504	588	2528	549	439	1325	790	180	768	471
Atlanta GA	436	749	1941	587	701	671	1587	1884	717	881	601	2107	428	2137	287	438	817	333	579	671	939
Sacramento CA	419	2505	353	1787	1604	2453	630	467	1439	1454	2018	91	1883	77	2321	2046	1463	1749	2388	1768	2628
Mesa AZ	419	2128	387	1434	995	2064	21	316	868	827	1671	630	1481	673	1773	1649	849	1245	1985	1446	2283
Oakland CA	411	2561	331	1844	1632	2509	641	447	1472	1476	2076	39	1936	13	2361	2100	1489	1792	2443	1827	2687
Tulsa OK	399	1228	1277	593	441	1160	933	1229	236	485	810	1435	589	1466	916	752	416	340	1077	639	1395
Omaha NE	394	1150	1317	433	796	1097	1029	1300	588	828	669	1405	528	1425	1101	689	764	536	1030	431	1286
Minneapolis MN	392	1022	1527	354	1057	984	1274	1526	862	1109	539	1570	510	1584	1192	626	1042	703	938	295	1124
Colorado Springs CO	379	1635	826	922	825	1577	549	809	614	742	1159	935	996	961	1436	1163	713	854	1505	922	1776
Miami FL	375	1091	2345	1189	968	1027	1979	2267	1108	1149	1160	2557	1027	2593	327	996	1114	870	958	1273	1259
Saint Louis MO	353	878	1593	260	680	813	1266	1557	546	792	456	1717	233	1742	754	399	717	245	734	328	1040
Wichita KS	349	1267	1203	588	559	1204	875	1165	341	573	820	1341	620	1369	1031	788	511	445	1125	617	1424
Santa Ana CA	348	2441	40	1732	1348	2381	335	78	1220	1177	1970	333	1797	379	2125	1965	1202	1587	2305	1733	2586

A.2. Facility base capacity in each time periods for 7 cities and 10 demand zone problem in Chapter 3

	New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego
	NY	CA	IL	TX	PA	AZ	CA
Non-identical	8800	4500	3600	2700	2200	2100	2000
Identical	3700	3700	3700	3700	3700	3700	3700

A.3. Facility base capacity in each time periods for 10 cities and 15 demand zones problem in Chapter 3

	New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego	Dallas	San Antonio	Detroit
	NY	CA	IL	TX	PA	AZ	CA	TX	TX	MI
Non-identical	8600	4300	3500	2600	2100	1900	1800	1800	1700	1500
Identical	2980	2980	2980	2980	2980	2980	2980	2980	2980	2980

A.4.. Facility base capacity for larger networks involving 50,35, and 25 demand zones in Chapter 3

	New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego	Dallas	San Antonio	Detroit	San Jose	Indianapolis	San Francisco	Jacksonville	Columbus	Austin	Memphis	Baltimore	Milwaukee	Boston
	NY	CA	IL	TX	PA	AZ	CA	TX	TX	MI	CA	IN	CA	FL	OH	TX	TN	MD	WI	MA
10 facility problem	10400	6100	5300	4400	3900	3700	3600	3600	3500	3300										
15 facility problem	9400	5100	4200	3300	2800	2700	2600	2500	2500	2300	2200	2100	2100	2000	2000					
20 facility problem	8900	4600	3800	2800	2300	2200	2100	2000	2000	1800	1700	1600	1600	1600	1500	1500	1500	1500	1400	1400

A.5. Facility base capacity (identical) each time period for 10 cities 15 demand zone problem in Chapter 4

	New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego	Dallas	San Antonio	Detroit
	NY	CA	IL	TX	PA	AZ	CA	TX	TX	MI
Identical	3400	3400	3400	3400	3400	3400	3400	3400	3400	3400

A.6. Facility base capacity for 50 demand zone problem in Chapter 4

	New York	Los Angeles	Chicago	Houston	Philadelphia	Phoenix	San Diego	Dallas	San Antonio	Detroit	San Jose	Indianapolis	San Francisco	Jacksonville	Columbus
	NY	CA	IL	TX	PA	AZ	CA	TX	TX	MI	CA	IN	CA	FL	OH
5 facility problem	10950	10950	10950	10950	10950										
10 facility problem	5475	5475	5475	5475	5475	5475	5475	5475	5475	5475					
15 facility problem	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650	3650