Transportation Research Part E 102 (2017) 13–37

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

Responsive contingency planning of capacitated supply networks under disruption risks

Anubhuti Parajuli [a], Onur Kuzgunkaya [a,*], Navneet Vidyarthi [b]

[a] Department of Mechanical and Industrial Engineering, Faculty of Engineering and Computer Science, Concordia University, Montreal, QC H3G 1M8, Canada
[b] Department of Supply Chain and Business Technology Management, John Molson School of Business, Concordia University, Montreal, QC H3G 1M8, Canada

ARTICLE INFO

Article history:
Received 5 October 2016
Received in revised form 6 March 2017
Accepted 30 March 2017

Keywords:
Disruption
Interdiction
Fortification
Contingency planning
Response speed
Capacity backup
Tri-level optimization
Implicit enumeration

ABSTRACT

This paper prescribes an appropriate risk mitigation model for a capacitated supply chain subject to premeditated attacks on facilities. It presents a unified approach to responsive contingency planning and optimizing protection of supply facilities utilizing a game-theoretic framework of attack and defense which involves multi level optimization. Gradual capacity backups is proposed for the first time in the context of protection under intentional attacks in a capacitated supply system. A recursive tree search algorithm is proposed to solve the tri-level optimization problem. Computational efficiency of the algorithm is demonstrated and managerial insights are presented.

1. Introduction

Over the past decade, several natural and man-made disaster events have affected supply chains world-wide. In October 2011, production at several computer manufacturing companies 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 and Sodhi, 2014). The terrorist attacks of September 2001 led Ford to shutdown operations in five of its US assembly plants due to parts shortages. Such incidents of catastrophic events have increased in the recent years, spawning interest towards finding appropriate risk management strategies that help maintain supply chain performance amidst disruptions.

There are several risk management strategies through which disruption risks of supply chains can be managed, such as inventory, backup capacity, routine sourcing from multiple suppliers, contingent rerouting, demand management, facility location, and facility hardening. Tomlin (2006) categorizes inventory and sourcing as proactive mitigation strategies which can be taken ahead of disruption occurrences and rerouting and demand management as contingent or reactive strategies which are taken after disruption occurrences. Facility location is a proactive strategy appropriate while designing new supply chains but it can be very costly to relocate facilities in supply chains that have been operational, for which hardening or protection of existing facilities is a preferred strategy. Combining proactive and reactive strategies improve supply chain

* Corresponding author.
E-mail addresses: a_paraju@encs.concordia.ca (A. Parajuli), onurk@encs.concordia.ca (O. Kuzgunkaya), navneetv@jmsb.concordia.ca (N. Vidyarthi).

http://dx.doi.org/10.1016/j.tre.2017.03.010
1366-5545/

responsiveness to disaster events and are effective in minimizing both the short term economic impacts and long term losses of market shares from major disruptions.

Responsive contingency planning in supply chain risk management involves identification of appropriate risk management strategies ahead of disruptions, so that such strategies could be implemented in the event of disruptions to improve supply flows in the post disaster situation. It is thus a combination of both proactive and reactive risk management strategies. Holding strategic inventory ahead of disruptions and recovering disrupted flows through these inventories can be considered a responsive contingency approach, as supply chains can react to disruptions faster through such provisions. However, holding inventory for longer periods is cost prohibitive, and therefore is not an appropriate strategy under major disruptions that affect supply chains for longer periods (Hopp et al., 2012). To be of practical relevance, risk management strategies should investigate the trade-off of investments for capability improvements and risk reduction (Nooraie and Parast, 2016).

One of the cost effective and appropriate strategies for managing major disruption risks is contingent capacity adjustment through backup production. It involves utilizing backup capacity for extra production during disruptions so as to recover the lost supply. Unlike strategic inventory which needs to be held even when disasters have not occurred, back up production can be started only when such events occur and the disrupted supplies can be recovered from such backup productions. Risk mitigation through contingent capacity adjustment requires pro-active selection and planning of a backup capacity which improve post-disruption contingency operations. A major challenge in this approach 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 et al., 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 together constitute the design parameters of a capacity backup risk mitigation plan. Appropriate selection of these parameters ensure effective contingent re-routing operations during disruptions with adequate protection of supplies and faster recovery from major disruptions.

Contingent capacity adjustments through gradually available capacity has traditionally been used for managing risks under major disruptions from random natural events. However, to the best of our knowledge, this idea has not been applied for managing risks from intentional attacks, which is fundamentally different from managing risks from natural disasters. Intentional attacks, such as the terror attacks are deliberately planned to create maximum damages to the system. In such attacks, the attacker is capable of adjusting attacks to circumvent against protection measures. Unlike intentional attacks, natural disasters do not have the capability of adjusting damage levels. These fundamental differences mean that strategies that are developed for managing disruption risks due to natural disasters are inadequate under intentional attacks. Therefore, it merits to incorporate gradual availability of backup capacities when planning for protection against intentional attacks.

In this paper, we present model for optimizing protection in a capacitated supply network subject to intentional attacks by incorporating gradually available backup capacity for recovering supply flows. The proposed model is a unified framework for responsive contingency planning and protection of critical facilities. Building on a game theoretic modeling framework to capture the elements of dependence between attack and defense (Brown et al., 2006), the mathematical model is formulated as a tri-level mixed integer optimization. This model is used to identify the optimal allocation of protection budget for the necessary levels of backup capacity volumes and response speeds at the different supply facilities prior to attacks. A solution algorithm based on implicit enumeration of defense strategies is proposed to arrive at decisions involving (i) which facilities to protect and back up, (ii) what should be the volume of a backup capacity, and (iii) what should be the response speed of capacity additions.

The remainder of this paper is organized as follows: Section 2 discusses relevant literature in protection planning. Section 3 details the formulation of the mathematical model. Section 4 presents the proposed solution methodology. In Section 5, numerical study is conducted and results are discussed. Section 6 offers some concluding remarks.

2. Literature review

Recent literature on supply chain risk management have focused on managing disruption risks through resiliency improvements (Sawik, 2013; Torabi et al., 2015; Fahimnia and Jabbarzadeh, 2016; Lam and Bai, 2016). In resilient systems, the supply chain infrastructures perform cost efficiently both during normal operations and during disruptions. A resilient supply chain can survive, adapt and grow in the face of change and uncertainty (Fiksel, 2006). We focus our literature review on research involving resilient supply chain design for protection against supply flow losses caused due to facility failures. The related literature intersect research domains in supply chain risk management, critical infrastructure protection and location sciences.

The relevant literature can be classified into two major streams of research. The first stream of research deals with protection through strategic locations of facilities to create an intrinsically resilient supply chain network which can perform at low cost both normally and during disruptions. These models rely on contingent rerouting of disrupted demands to the facilities or routine sourcing from multiple facilities for mitigating disruption risks. Readers are referred to Snyder and Daskin (2005), Berman et al. (2007), Cui et al. (2010), Lim et al. (2010), Peng et al. (2011), Aksen and Aras (2012) and references

therein. Few models in this category have integrated location decisions with inventory decisions (Qi et al., 2010; Jeon et al., 2008; Mak and Shen, 2012). Review of literature involving reliable supply chain design models can be found in Snyder (2006), Klibi et al. (2010) and Snyder et al. (2016).

A second stream of research focuses on developing models for managing risks of supply chains that are already operational and for which relocation of facilities is cost prohibitive. The mathematical models in this category are commonly known as fortification models. These models do not necessarily involve facility location decisions. The models follow the development of interdiction models where the focus is to assess vulnerability of the system. Interdiction models identify critical facilities (facility interdiction models) or network arcs (network interdiction models) whose loss can create the greatest loss of system efficiency (Church and Scaparra, 2007a; Cormican et al., 1998; Israeli and Wood, 2002; Wollmer, 1964; Wood, 1993). A survey of interdiction models can be found in Church et al. (2004). The shortcomings of interdiction models is that protection plans that rely on solutions provided by interdiction models are suboptimal because of the interdependence between attack and defense (Brown et al., 2006). This shortcoming is addressed through fortification models which explicitly include protection decisions in mathematical models of interdictions. These models are commonly prescribed in a game theoretic framework to capture the elements of dependence and often cast as bilevel 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. Several variants to this basic fortification model have subsequently been studied: probabilistic number of attacks rather than deterministic (Liberatore et al., 2011; Liberatore and Scaparra, 2011), shortest path network systems (Cappanera and Scaparra, 2011), facility recovery time and multiple disruption considerations (Losada et al., 2012), cascading effects of disruptions (Liberatore et al., 2012), and partial disruptions (Liberatore et al., 2012; Aksen et al., 2014). In Table 1 we summarize these articles on the basis of key features.

Table 1 highlights the fact that a majority of models developed for protecting supply chain from disruptions have ignored capacity issues by assuming that uncapacitated facilities. This drawback was addressed by Scaparra and Church (2012), when they presented a model for optimizing protection of capacitated facilities. In this model, the protection decision involved which facilities to secure or harden under the possibility that the attacker can attack a finite number of facilities i.e., when the number of facility that will be attacked is deterministic and known to the defense planner. Liberatore et al. (2012) also considered capacitated facilities but allowed partial disruption of facilities in their model inorder to represent disruptions that create ripple effects in a region of attack. It was assumed that attack of one facility would also lead to some percentage of capacity loss at a neighbourhood facility. Similar to Liberatore et al. (2012), the protection in their model also implied facility hardening. In both of these models, protection was a means to secure capacities of critical facilities, i.e., those facilities whose capacity losses would be most detrimental to supply chain operations. The models were used to solve a strategic level decision problem involving which facilities to protect from being attacked while the operational level demand assignments were executed in an optimal manner. These models can be considered as static models of protection since they do not incorporate temporal aspect of recovery. This feature was included in Losada et al. (2012), where protection of a facility implied reducing the recovery time of facilities rather than preventing it from failure due to attacks. However, their model considered uncapacitated facilities.

Table 1
Summary of features in relevant literature.

Articles	Facility features		Protection Features				
	Capacitated	Uncapacitated	Inventory Backup	Capacity Backup	Facility Hardening	Facility location	Recovery time reduction
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)		√			√	√	
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. (2012)		√					√
Mak and Shen (2012)		√	√			√	
Scaparra and Church (2012)	√				√		
Li et al. (2013)		√			√	√	

In all of the discussed models, the contingent capacity adjustments has been completely ignored. Through this mechanism, the total capacity can be increased temporarily by acquiring contingent capacities. This helps in partial or complete recovery of lost capacities due to disruptions. Contingent capacity adjustment requires tactical level planning involving decisions of quantities of contingent backup capacities and resource acquisitions. It is especially useful under major disruptions where holding inventory is cost prohibitive. Contingent adjustment of capacities is greatly facilitated in modern reconfigurable manufacturing systems. Such systems have better scalability than dedicated manufacturing systems through their modular structure which allow ease of reconfiguration and volume flexibility.

The only two articles that consider contingent capacity adjustment for designing protection are due to Aksen et al. (2010) and Aksen and Aras (2012) in which supply flows of a capacitated network are protected with flexible capacity expansion of facilities. These models however, assume that such capacity expansions occur instantaneously. The temporal aspect of such capacity addition which involves response times has been ignored and hence the solutions provided are based on overestimations of actual available capacity.

Response time is critical in the design of supply chains that are responsive to disruptions and rely on contingent capacity adjustments through capacity backups. It affects the volume of back-up capacity that can be utilized during disruption recovery and has been dealt in several literature on supply chain risk management. Tomlin (2006) and Hopp and Yin (2006) assume that the total backup capacity would be available only after the response time. Klibi and Martel (2012) consider gradual recovery of lost capacity based on the intensity of disruption and the time to recovery. Nejad et al. (2014), additionally consider congestion effect on partial capacity availability at the back up supplier within the response time during which primary supplier is disrupted. Schmitt (2011) considers response time for risk mitigation through capacity backups and safety stock in a multiechelon supply chain where disruptions can happen at any stage of the supply chain. Additionally, Wang and Koren (2012) present several backup facility configurations affecting the supply chain responsiveness levels in a cost and response time tradeoff analysis. In a similar study, Niroomand et al. (2012) propose a capacity planning model with a volume flexible backup facility where desired capacities are only partially available during the response time. These and most other articles in supply chain risk management literature have dealt with disruptions which are either operational or the ones caused due to random acts of nature.

To the best of our knowledge, there have been no article to date which consider gradual capacity availability in a contingent capacity adjustment plan for managing risks from intentional attacks. This is despite the fact that the modeling paradigms for intentional attacks are different from that of natural disasters. Our paper serves to fill this literature gap. We consider protection of supply chains for disruption risk management through capacity backups as in Aksen et al. (2010) but with gradual capacity availability. Utilizing a game-theoretic framework of attack and defense, we solve the defender's problem of deploying an appropriate level of response speed in addition to the capacity volumes to protect or partially recover supply flow losses during disruptions caused due to intentional attacks on facilities.

3. Problem formulation

Consider a network of multiple capacitated facilities supplying a set of customers with a single product. Let I be the set of customer zones. Each customer zone $i \in I$ 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 $i \in I$ from facility $j \in J$. 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 $t \in T$. 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.

Disruption to the system is modeled as an attack on facilities by an intelligent attacker who has prior information of the system and can adjust his attacks to harm the system maximally. A worst-case disruption scenario always results from such attacks if none of the system components (facilities) are protected. In a worst-case attack, the system operating costs are maximized, either because the customer needs to be assigned to more distant facilities, or due to lost sales incurred owing to inadequate system capacity, or both.

The system planner with limited protection budget B intends to protect/harden facilities to mitigate the impact of a worst-case disruption of a capacitated supply network. The impacts are measured in terms of increases in the operational costs as a result of disruption. A protection cost c_{jl} is involved in fortifying a facility j with backup capacity at a certain level l. The levels of backup capacity are determined by the nominal volume and response speeds of such capacity additions. It is further assumed that fortification of facilities is a part of backup capacity additions. In other words, the decisions to add capacity backups on a facility also implies its fortification. The protection decisions therefore involve 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.

The notations of the model are listed as follows:

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 that can be added 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 protection planning problem for managing disruptions risks of intentional attacks is formulated as a tri-level optimization model within a game-theoretic framework. This kind of modeling approach is more appropriate to assess the impacts of disruptions due to intentional attacks and to design protection against such impacts as it captures the underlying interdependent objectives among players. The model conceptually involves a sequential game among three players at different levels of hierarchy: (i) supply chain planner (*system defender*) at the top level determines the facilities to be protected in order to minimize the supply chain losses due to attacks; (ii) the interdictor (*attacker*), at the middle level, identifies the set of facilities that can be attacked to create the maximum loss in the supply chain; and (iii) supply chain operator (*system user*) identifies the most cost effective way of operating the system post attacks. The decision framework for the proposed protection planning problem is depicted in Fig. 1.

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.

$$[\mathbf{DLP}] : \min_z H(z) \tag{1}$$

subject to:

$$\sum_{j\in J}\sum_{l\in L} c_{jl} z_{jl} \leqslant B \tag{2}$$

$$\sum_{l\in L} z_{jl} \leqslant 1 \quad \forall j \in J \tag{3}$$

$$z_{jl} \in \{0,1\} \quad \forall j \in J, l \in L \tag{4}$$

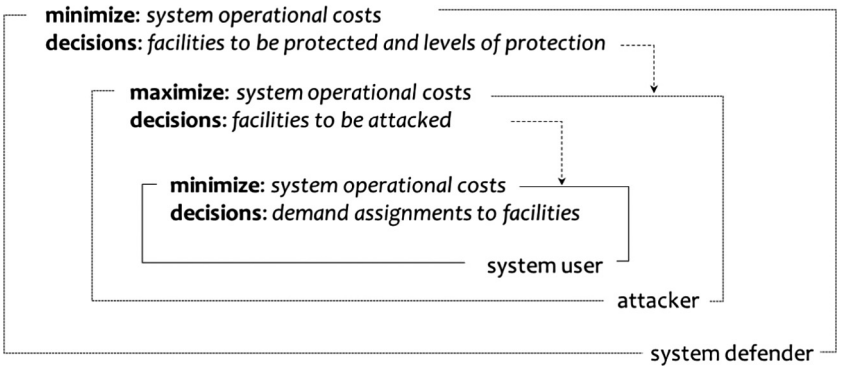

Fig. 1. Decision framework.

where,

$$[\mathbf{ALP}] : H(z) = \max_s G(s,z) \tag{5}$$

subject to:

$$\sum_{j\in J} s_j = r \tag{6}$$

$$s_j + \sum_{l\in L} z_{jl} \leqslant 1 \quad \forall j \in J \tag{7}$$

$$s_j \in \{0,1\} \quad \forall j \in J \tag{8}$$

where,

$$[\mathbf{ULP}] : G(s,z) = \min_{x,u} \left(\sum_{i\in I}\sum_{j\in J}\sum_{t=1}^{T} x_{ijt} d_{ij} + \sum_{i\in I}\sum_{t=1}^{T} \beta_i u_{it} \right) \tag{9}$$

subject to:

$$\sum_{j\in J} x_{ijt} + u_{it} = h_{it} \quad \forall i \in I, t = 1 \ldots T \tag{10}$$

$$\sum_{i\in I} x_{ijt} \leqslant (1 - s_j)\left(v_{jt} + \sum_{l\in L} m_{tl} a_{jl} z_{jl} \right) \quad \forall j \in J, t = 1 \ldots T \tag{11}$$

$$x_{ijt} \geqslant 0 \quad \forall i \in I, j \in J, t = 1 \ldots T \tag{12}$$

$$u_{it} \geqslant 0 \quad \forall i \in I, t = 1 \ldots T \tag{13}$$

The decision framework comprises three optimization problems. The system planner or the defender level problem (DLP) is represented by Eqs. (1)–(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 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 Eqs. (5)–(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 Eqs. (9)–(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 (Eq. (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 (Eq. (5)). Constraint 6 defines the number of facilities that can be simultaneously attacked. The DLP problem is linked to the ALP problem through constraint 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 (Eq. (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 (constraint 10). Constraint 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.

4. Solution methodology

In order to identify facilities to be protected and to obtain the optimal capacity and response speeds of a backup capacity for contingent capacity adjustments, a recursive binary tree search algorithm is presented. The algorithm utilizes the attacker-defender game theoretic framework of the proposed model to develop a binary search tree and explore the optimal solution.

4.1. Background

Game theoretic attacker-defender modeling framework involves optimization at two levels, one involving attack and the other involving defense. These models are called bi-level optimization models and have 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 inorder to counter strategic risks such as malicious attacks (Golany et al., 2009). It is mainly because this modeling framework captures the underlying interdependence between attack and defense. 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., 2012) 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 and Scaparra (2011), Scaparra and Church (2012) and Liberatore et al. (2011, 2012).

Our solution methodology is an extension of the implicit enumeration algorithm (IE) in Scaparra and Church (2008a). An important difference is that while Scaparra and Church (2008a) 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 linear program. 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 in the ULP problem, which makes its solution simpler involving only real 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 and Church, 2008a), our approach does not encounter similar difficulty other than the need to enumerate feasible attack patterns at each node. However, this does not deteriorate the computational performance of our algorithm. The computational performance is reported in Section 5. In the following sections, we provide the details of our solution methodology.

4.2. General description of the proposed algorithm

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 his attacks on selected facilities.

A pseudo code of the root node algorithm is presented in Table 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

Table 2
Root node algorithm.

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

1. $\forall j, lz_{jl} \leftarrow 0$
2. enumerate $P = {}^{n}C_{r}$ attack combinations
3. for $p = 1 \ldots, P$ \\attack scenarios
4. Solve ULP for $(S^p, cLS^p, cFlow^p)$
5. end for
6. $mincLS \leftarrow \text{minimum}(cLS^1, \ldots .cLS^P)$; $minFlow \leftarrow \text{minimum}(cFlow^1, \ldots .cFlow^P)$
$maxcLS \leftarrow \text{maximum}(cLS^1, \ldots .cLS^P)$; $maxFlow \leftarrow \text{maximum}(cFlow^1, \ldots .cFlow^P)$
7. for $p = 1 \ldots P$
8. $normcLS^p \leftarrow \left(\frac{cLS^p - mincLS}{maxcLS - mincLS}\right)$; $normcFlow^p \leftarrow \left(\frac{cFlow^p - mincFlow}{maxFlow - minFlow}\right)$ \\ normalization of costs
9. $normTotCost^p \leftarrow normcLS^p + normcFlow^p$ \\ normalized total cost
end for
10. $maxnormTotCost \leftarrow \text{maximum}(normTotCost^1, \ldots .normTotCost^P)$
11. $S^* \leftarrow \{S^p : normTotCost^p = maxnormTotCost\}$;
$cLS^* \leftarrow \{cLS^p : normTotCost^p = maxnormTotCost\}$;
$cFlow^* \leftarrow \{cFlow^p : normTotCost^p = maxnormTotCost\}$;
12. return $(S^*, cLS^*, cFlow^*)$ \\ optimal ALP solution at the root node

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.

The facilties 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 and 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 Fig. 2.

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.

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

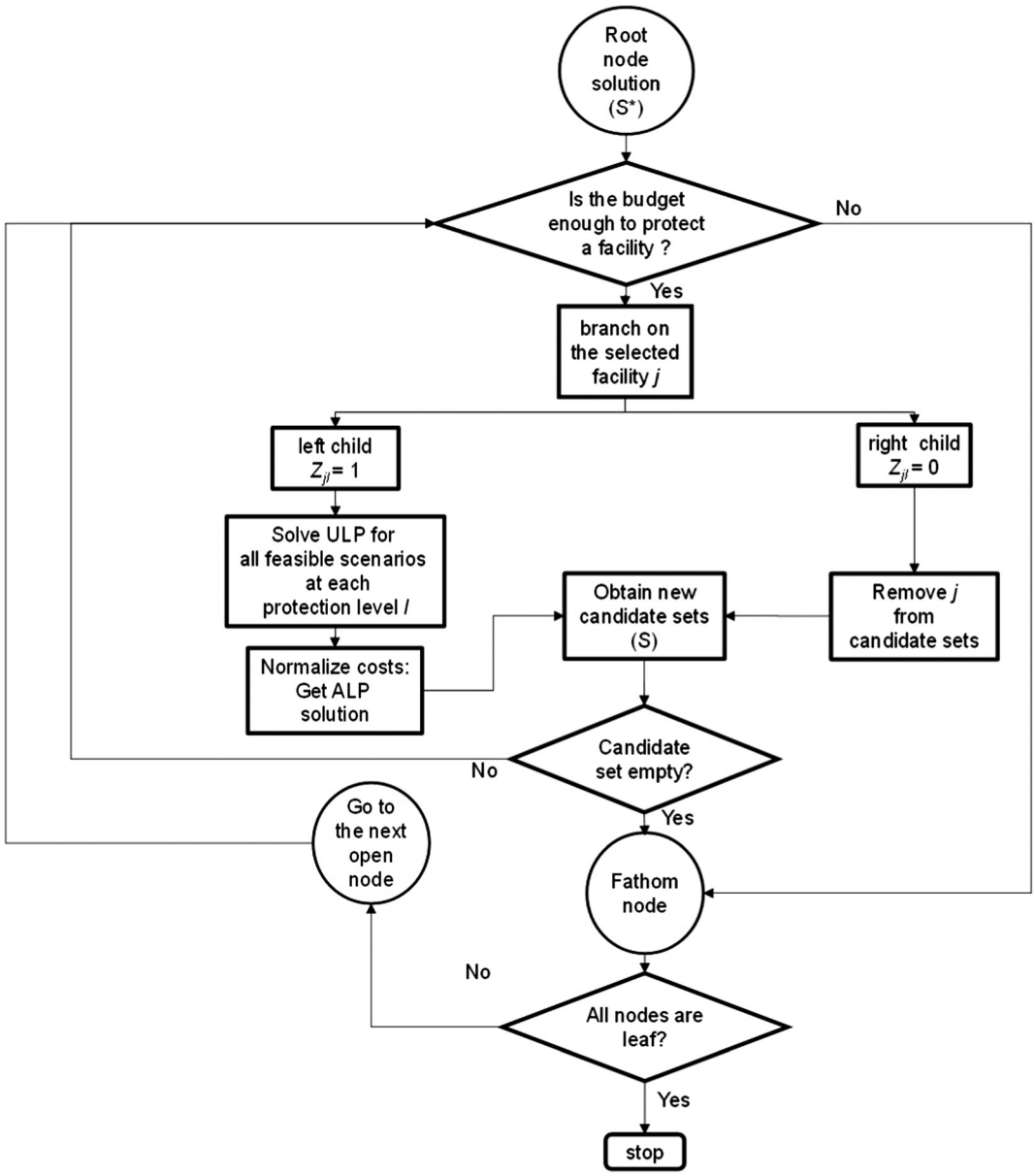

Fig. 2. Flowchart of the search tree.

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 4.3.2). 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.

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.

4.3.1. Branching, pruning and nodes traversal

The enumeration tree corresponding to this illustrative problem is depicted in Figs. 3a–3c. 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 Fig. 3a, 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*.

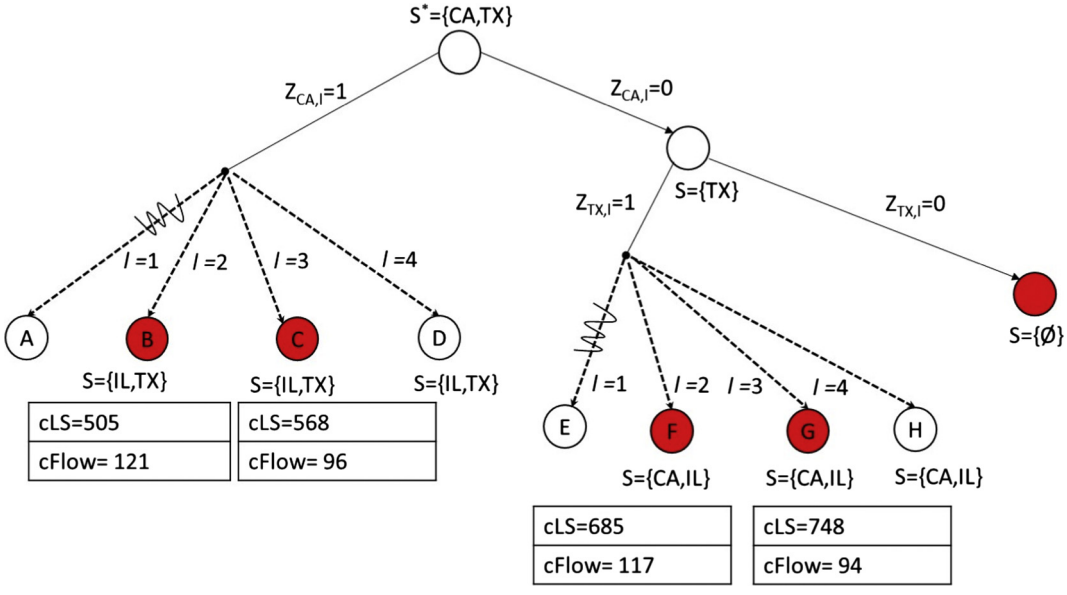

Fig. 3a. Root node and initial branching of the search tree.

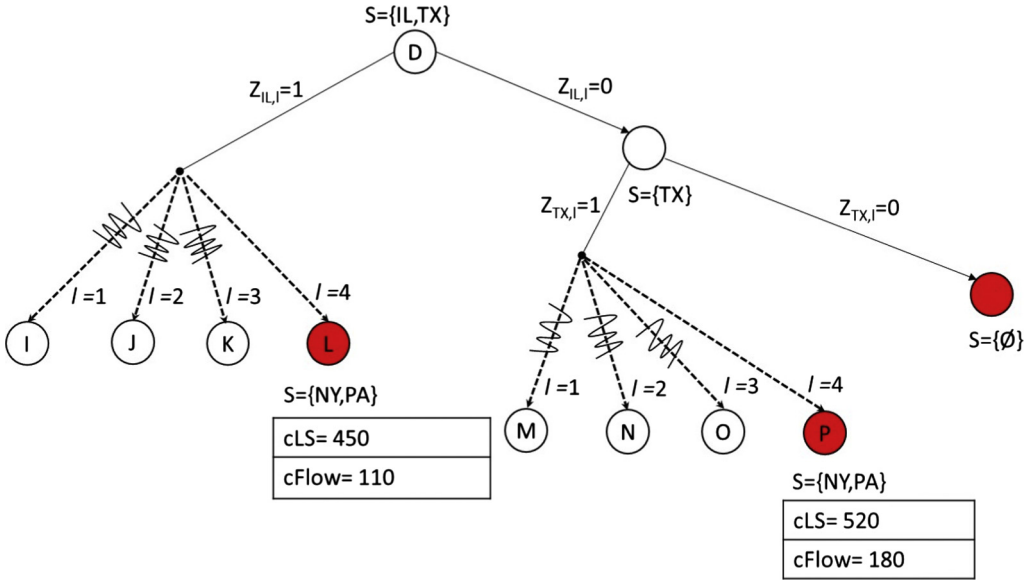

Fig. 3b. Continuation of search tree from left branch of the root node.

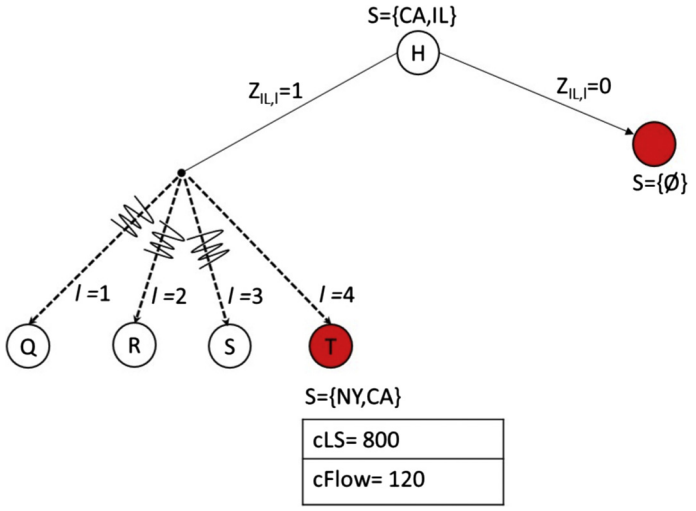

Fig. 3c. Continuation of search tree from right branch of the root node.

The other child node obtained from the root node corresponds to the branch $Z_{CA,I} = 0$ (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 $Z_{TX,I} = 1$, 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 $Z_{TX,I} = 0$ (i.e., the facility at Texas is not protected at any level and thus this branch leads to a node with $S = \{\emptyset\}$. Hence this node is fathomed.

The remaining budget at nodes *B*, *C*, *F* and *G* in Fig. 3a 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 $l = 4$. Figs. 3b and 3c 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 $Z_{IL,I} = 0$, 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 $S = \{\emptyset\}$ 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 $S = \{\emptyset\}$, 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.

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 Eqs. (14)–(16).

$$normcLS_n = \left(\frac{cLS_n - \min_{n\in N}\{cLS_n\}}{\max_{n\in N}\{cLS_n\} - \min_{n\in N}\{cLS_n\}}\right) \tag{14}$$

$$normcFlow_n = \left(\frac{cFlow_n - \min_{n\in N}\{cFlow_n\}}{\max_{n\in N}\{cFlow_n\} - \min_{n\in N}\{cFlow_n\}}\right) \tag{15}$$

$$normTotal_n = normcLS_n + normcFlow_n \tag{16}$$

Table 3 illustrates the computations of normalized costs at the leaf nodes for this example.

4.3.3. Selection of optimal defense strategy

The minimum normalized total costs is obtained for the leaf node L (Table 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 direct his target to attack facilities at NY and PA through which he can maximize system operational costs.

5. Numerical results

This section reports and discusses computational results obtained with the proposed solution methodology. The algorithm was coded in C++ 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 i5-3340 M processor at 2.7 GHz and 8 GB of RAM running Windows 7 operating system.

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 per 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 A. The unit cost of lost sales are set at 2% higher than the maximum distances

Table 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

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.

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 (Eq. (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 Eq. (17) represents the idle capacity of the system (i.e., system capacity in excess of total demands).

$$v_{jt} = \left(\frac{(1+\alpha)\sum_{i=1}^{I} h_{it}}{J}\right) \quad \forall j \in J, t \in T \tag{17}$$

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 (Eq. (17)).

$$v_{jt} = (h_{kt} : k, i \in I,\ d_{kj} = \min_i d_{ij}) + \lambda_t \quad \forall j \in J, t \in T \tag{18}$$

where λ_t is calculated as:

$$\lambda_t = \left(\frac{(1+\alpha)\sum_{i=1}^{I} h_{it} - \sum_{j=1}^{J}(h_{kt} : k, i \in I,\ d_{kj} = \min_i d_{ij})}{J}\right) \tag{19}$$

Any remaining system capacity after satisfying the nearest demands to every facility are evenly distributed among all open facilities. This equally distributed remaining system capacity is represented by parameter λ_t in Eq. (18). Its computation is shown in Eq. (19).

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 4.

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.

5.2. Analysis of results

The initial experiment involves a small network ($I = 10, J = 7$) and a single facility attack ($r = 1$). This small network size is chosen for the ease of mapping supply flows graphically so that effect of different contingency operations under disruption can be demonstrated clearly. Specifically, we look into the impact of reallocation of entire network flows under disruption and compare it with contingent re-routing of only the disrupted flows to highlight the effectiveness of each operation for risk mitigation. The demands, facility capacities (non-identical) and distance data are obtained as described in Section 5.1 and are provided in Appendix A.

Table 4
Proportions of capacity added each time period at high and low response speeds.

	T1	T2	T3	T4
Hi	½	1	1	1
Lo	1/3	2/3	1	1

5.2.1. Impact of supply flow re-allocation for risk mitigation effectivness

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. In this experiment we demonstrate the effectiveness of one over the other in mitigation of disruption risks by considering a network involving seven facilities, ten demand zones and attack of a single facility. Fig. 4a 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.

Fig. 4b 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 all protection variables z_{jl} to zero and all interdiction variables to zero except s_{NY} which is set at 1.

The configuration in Fig. 4c results when the facility at NY is attacked 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 5 provides operational costs involving lost sales costs (cLS) and the transportation costs (cFlow) under the two cases of contingency operations illustrated through Figs. 4b and 4c. 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 on re-routing of disrupted flows only. 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.

Contingency operations involving network redesign through re-allocations are necessary when facilities operate within capacity limits. Re-allocations allow flow exchanges in order to reduce overall transportation costs. For instance, in the network of Fig. 4b, 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. Note that such an exchange would be unnecessary if facilities are uncapacitated as the routed demand can always be fully met by a single facility. Contingent rerouting, which is an appropriate recourse solution for an uncapacitated system, therefore may not be appropriate for a capacitated system. Redesign of network flows through supply re-allocations is a more effective contingency approach 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 for contingency operations relying on capacity backups. Under a limited budget of protection, the main tradeoff 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

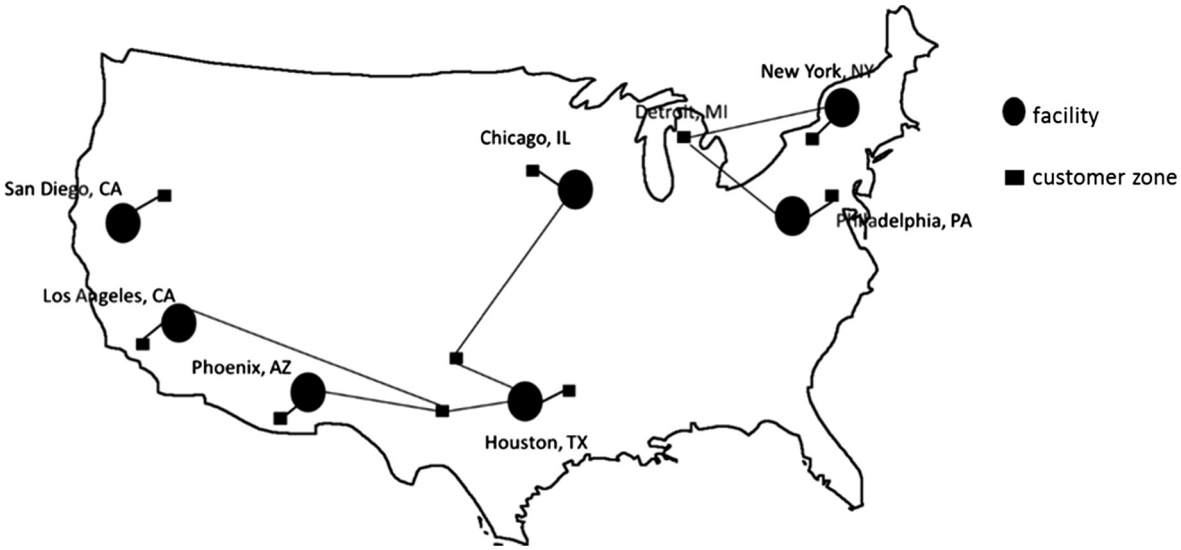

Fig. 4a. Supply flows of an original network.

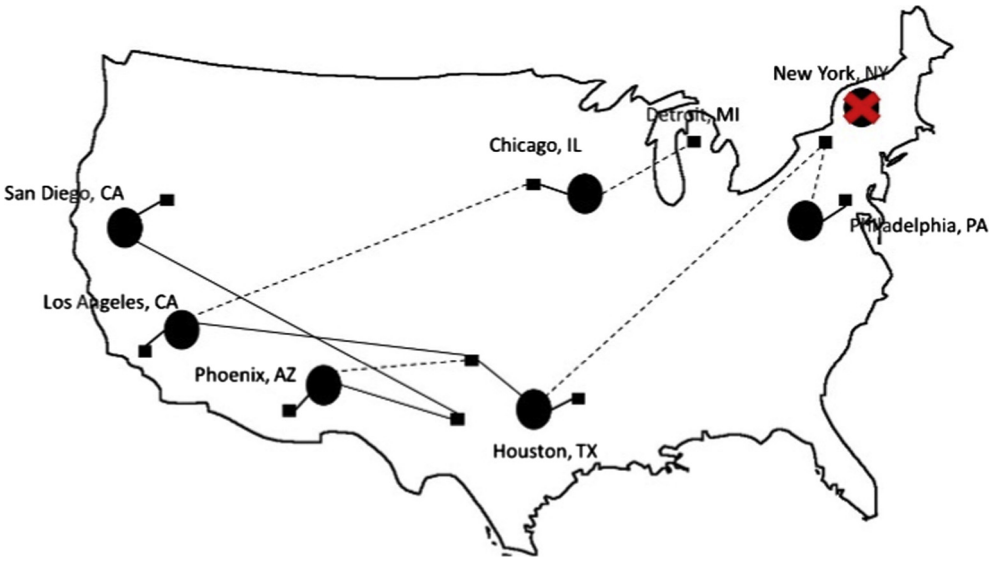

Fig. 4b. Supply flows of a redesigned network with flow re-allocations.

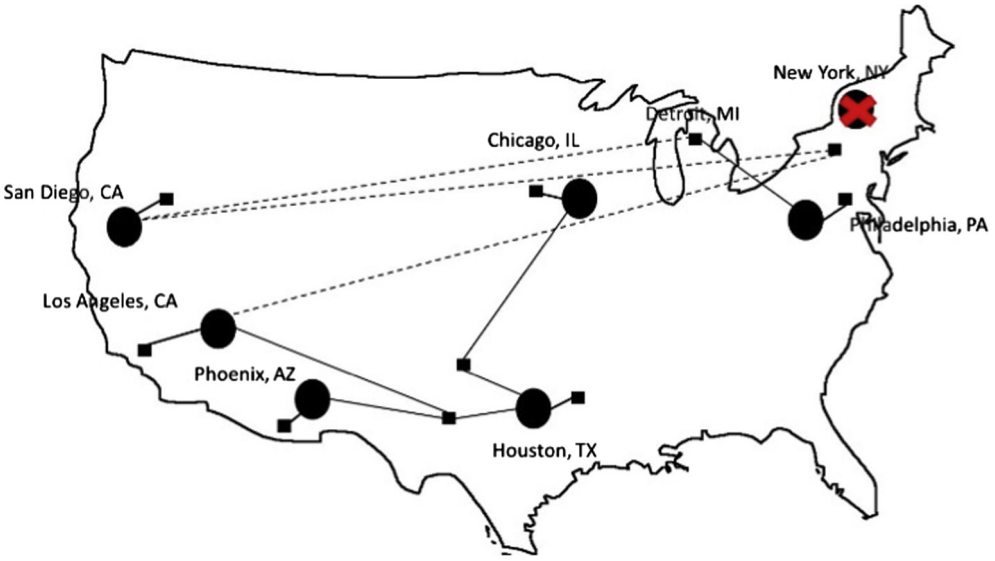

Fig. 4c. Supply flows of a network with contingently rerouted flows after disruption.

Table 5
Operational costs under two different contingency operations.

Case	*cLS*	*cFlow*
Case I	73,555,000	12,037,000
Case II	73,555,000	17,860,000

achieved faster. This however raises protection costs. In the following sections we investigate optimal strategies for risk mitigation with respect to attack, protection budget, network structure and backup capacity features.

5.2.2. Impact of network initial capacity layout

The empirical analysis is performed for a supply network involving 15 customers and 10 facilities for which the input parameters are derived as explained in Section 5.1. We analyze the protection strategies and performance levels under two types of networks, the first network has identical distribution of initial capacities, while the second network is more

generic with varying initial facility capacities. The capacities are computed as outlined in Section 5.1.1. The two networks are equivalent in terms of the total system capacity, total demands and capacity slacks, hence the results are comparable.

5.2.2.1. Network with identical facility capacities. In identical capacity networks, loosing 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 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 Fig. 5. 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 his benefits by attacking facilities that raise transportation costs the most. Consequently, protection efforts are concentrated in securing such facilties.

5.2.2.2. 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 7.

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 Fig. 6. 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 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 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	107,280	16,982,500	36,205,500	337
	2	–	1,5	18,392	95,360	48,260,600	35,287,200	370
	3	–	1,5,10	30,312	83,440	79,538,700	29,737,700	356
	4	–	1,3,5,10	42,232	71,520	110,817,000	24,835,900	347
$B = 670$	1	1(4)	5	3472	110,280	9,110,530	35,302,100	320
	2	1(4)	5,10	15,392	98,360	40,388,600	29,752,600	302
	3	1(4)	3,5,10	27,312	86,440	71,666,700	24,850,800	287
	4	1(4)	2,3,5,7	39,232	74,520	102,945,000	20,637,800	277
$B = 1340$	1	1(4), 5(4)	10	1236	112,516	3,243,260	29,343,400	261
	2	1(4), 5(4)	3,10	12,392	101,360	32,516,600	25,545,400	252
	3	1(4), 5(4)	2,3,7	24,312	89,440	63,794,700	21,332,400	239
	4	1(4), 2(4)	3,5,8,10	36,232	77,520	95,072,800	18,633,200	240
$B = 2010$	1	1(2), 5(4)	10	628	113,124	1,647,870	25,034,500	221
	2	1(2), 5(4)	3,10	9392	104,360	24,644,600	25,548,400	245
	3	1(2), 2(4), 5(4)	3,8,10	21,312	92,440	55,922,700	19,327,900	209
	4	1(2), 2(4), 5(4)	3,6,7,10	33,232	80,520	87,200,800	15,809,400	196
$B = 2680$	1	1(2), 5(2)	2	298	113,454	78,1952	22,928,200	202
	2	1(4),2(4),3(4),5(4)	7,10	6392	107,360	16,772,600	24,074,800	224
	3	1(4),2(4),3(4),5(4)	6,7,10	18,312	95,440	48,050,700	17,571,600	184
	4	1(4),2(4),3(4),5(4)	4,8,9,10	30,232	83,520	79,328,800	9,688,960	116

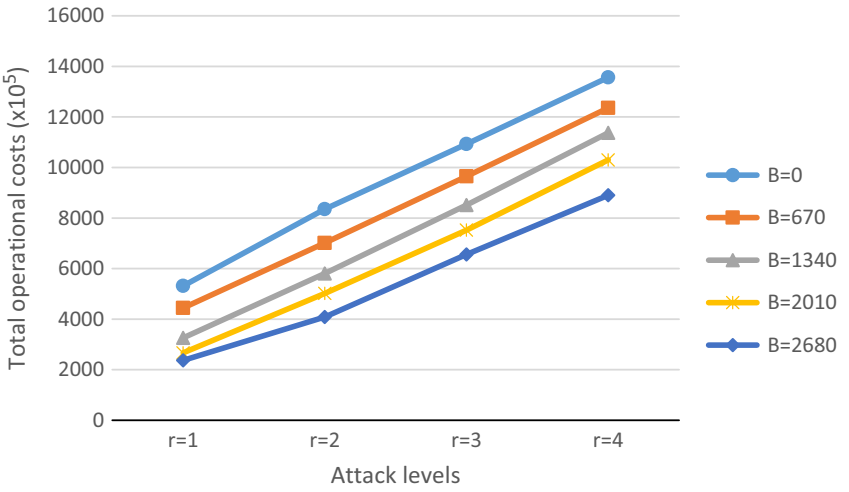

Fig. 5. Total operational costs for network with identical facility capacities.

Table 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	28,952	84,800	75,970,000	10,345,600	122
	2	–	1,3	42,952	70,800	112,706,000	9,063,540	128
	3	–	1,3,5	51,352	62,400	134,748,000	8,877,430	142
	4	–	1,2,3,5	68,552	45,200	179,880,000	7,051,860	156
B = 670	1	1(4)	2	8752	105,000	22,965,200	19,282,600	184
	2	1(4)	2,7	15,952	97,800	41,858,000	18,617,100	190
	3	1(4)	2,6,7	23,552	90,200	61,800,400	16,103,600	179
	4	1(4)	2,3,6,7	37,552	76,200	98,536,400	10,179,400	134
B = 1340	1	1(4), 2(4)	3	2552	111,200	6,696,450	16,507,100	148
	2	1(4), 2(4)	3, 10	8552	105,200	22,440,400	15,235,200	145
	3	1(4), 2(4)	3, 5, 9	16,952	96,800	44,482,000	12,055,600	125
	4	1(4), 2(4)	3, 4, 8, 9	26,952	86,800	70,722,000	9,332,140	108
B = 2010	1	1(4), 2(4), 3(4)	4	248	113,504	650,752	10,625,300	94
	2	1(4), 2(4), 3(4)	4,6	3552	110,200	9,320,450	14,643,800	133
	3	1(4), 2(4), 10(4)	3,4,8	17,152	96,600	45,006,800	10,538,600	109
	4	1(4), 2(4), 3(4)	4,6,8,9	17,552	96,200	46,056,400	12,017,000	125
B = 2680	1	1(4), 2(2),3(4)	4	0	113,752	0	10,622,800	93
	2	1(4), 2(4),10(2)	3,4	6952	106,800	18,242,000	12,085,600	113
	3	1(4), 2(4),10(4),8(4)	3,4,9	13,752	100,000	36,085,200	11,235,500	112
	4	1(4), 2(4), 3(4), 4(4)	6,7,8,9	11,352	102,400	29,787,600	12,803,900	125

It can further be observed from Table 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 Figs. 5 and 6. 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-

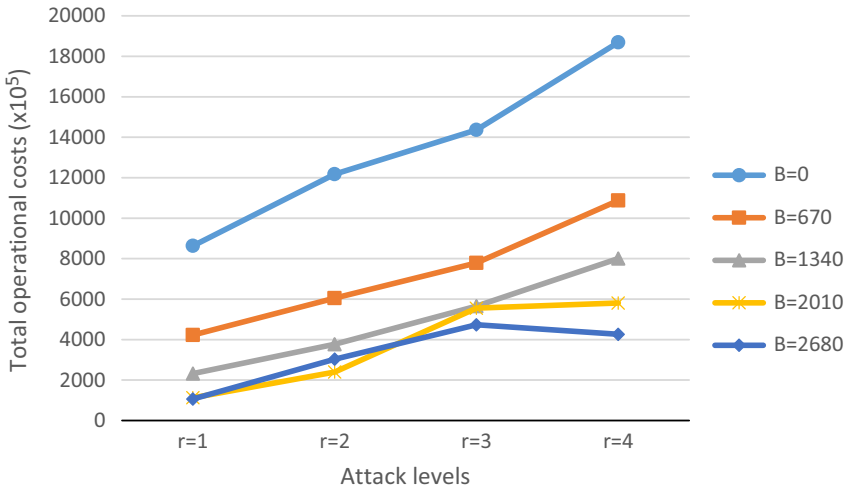

Fig. 6. Total operational costs for network with non-identical facility capacities.

Table 8
Illustration of results of two competing strategies in a non-identical capacity network.

Strategy	Z_{jl}	S_j	*cLS*	*cFlow*	*cTotal*	đ
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

identical capacity networks seek for tradeoff solutions that balance the loss of high capacity facility against the loss of travel distances.

5.2.3. 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 make contingency operations more cost effective. 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 defense 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 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 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.

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.

Table 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	107,360	16,772,600	24,074,800	224
H: 2500 units L: 500 units	1(1), 5(4)	2,7	8142	105,610	21,364,600	25,497,000	241

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.

5.3. 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, 2, 3, 4, 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 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.

5.3.1. 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 11a and b respectively. The average CPU times are computed from the results shown in Table 10. The values in Table 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 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 sevenfolds 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.

5.3.2. 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 12a. The averages are obtained from Table 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 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 12a and b, 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.

In this Section, we examined the computational performance of the 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 facilties are limited to a few facilities, the test problem sizes

Table 10
Computational performance of the algorithm on larger networks.

Budget	Problem instance	*Sj*			Z_{jl}			Total cost ($\times 10^6$)			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

Table 10 (*continued*)

Budget	Problem instance	*Sj*			Z_{jl}			Total cost ($\times 10^6$)			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
	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	16,541	17,227

Table 11
Variation of CPU time with respect to budget levels under different demand and facility levels.

Budget level	(a) Number of demand zones			(b) Number of facilities		
	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

Table 12
Variation of CPU time with respect to attack levels under different demand and facility levels.

Attack levels	(a) Number of demand zones			(b) Number of facilities		
	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

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 h of CPU time. Considering the fact that it is a strategic long term decision, the CPU time is acceptable.

6. Conclusion

In this paper, we address the problem of protecting capacitated supply network from catastrophic disruptions due to intentional attacks on supply facilities. Capacity backup is proposed as a solution to hedge disruption risks. The contribution of the model is the consideration of gradual capacity availability for capacity backups in defense planning against intentional attacks. The mathematical model involves mixed integer hierarchical optimization which is solved using implicit enumeration of protection strategies. Through the appropriate selection of response speeds, the model provides more accurate estimates of backup capacities available for system recovery under a budgeted protection plan. The selection of response speeds and volumes of backup capacity is evaluated, in order to increase resilience of supply chains through effective contingency operations during disruption periods.

Using the single facility attack example, we first demonstrate that rerouting of only the disrupted supply flows is an inadequate contingency operation for capacitated systems which require flow redistributions to minimize total operational costs under disruptions. Secondly, this example is used to demonstrate that ignoring response times leads to inaccurate estimations of available backup capacity during initial recovery periods and hence imperfect mitigation solutions.

We study network vulnerability to attacks using two different hypothetical networks with different initial capacity distributions. The results show that both initial capacity distributions and the available sizes of backups affect the costs of contingency operations during attacks. A supply network designed with low initial transportation costs are generally more vulnerable to disruptions. Characteristic to such networks are non-identical capacity networks with unevenly distributed initial system capacity. If sufficient budget can be allocated to secure critical facilities in this network, relative efficiency improvements are higher compared to networks with identical capacity distributions, which are efficient to operate under low or nonexistent protection budgets. Investigation of the two networks also demonstrate that protection need to be dispersed (decentralized) when planning against larger number of facility attacks so as to spread out risk. In centralized protection, high volume backups are added at faster speeds to a limited number of facilities. These decisions are also dependent on the available range of backup capacity. If the range is wide, the results demonstrate that protection should be centralized to enable faster recovery of lost capacities.

The computational performance of the proposed algorithm is tested on a larger problem instances involving up to 50 demand zones and 20 facilities involving 5 attacks and four levels of protection budget. The experimental results show that the algorithm is able to solve such problems to optimality in a reasonable amount of computation time. For larger networks involving more attacks and protection levels, in future research, one can explore heuristics to reduce the size of the search tree. This paper, considers all of the system parameters, including the attack levels, as deterministic. The methodology can be extended to incorporate a probabilistic attack. This will help in defense planning when attacker capability is unknown.

Acknowledgements

This research supported by the Discovery Grant# RGPIN-2015-03795 from National Science and Engineering Research Council of Canada. Their support is highly acknowledged.

Appendix A. Distances, demands and facility base capacities data inputs used for numerical examples

A.1. Distances and demand quantities data inputs

City	Demand	New YorkNY	Los AngelesCA	ChicagoIL	HoustonTX	PhiladelphiaPA	PhoenixAZ	San DiegoCA	DallasTX	San AntonioTX	DetroitMI	San JoseCA	IndianapolisIN	San FranciscoCA	JacksonvilleFL	ColumbusOH	AustinTX	MemphisTN	BaltimoreMD	MilwaukeeWI	BostonMA
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 10 demand zones problem

	New YorkNY	Los AngelesCA	ChicagoIL	HoustonTX	PhiladelphiaPA	PhoenixAZ	San DiegoCA
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 15 demand zones problem

	New YorkNY	Los AngelesCA	ChicagoIL	HoustonTX	PhiladelphiaPA	PhoenixAZ	San DiegoCA	DallasTX	San AntonioTX	DetroitMI
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.

	New YorkNY	Los AngelesCA	ChicagoIL	HoustonTX	PhiladelphiaPA	PhoenixAZ	San DiegoCA	DallasTX	San AntonioTX	DetroitMI	San JoseCA	IndianapolisIN	San FranciscoCA	JacksonvilleFL	ColumbusOH	AustinTX	MemphisTN	BaltimoreMD	MilwaukeeWI	BostonMA
10 facility problem	10,400	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

References

Aksen, D., Akca, S.S., Aras, N., 2014. A bilevel partial interdiction problem with capacitated facilities and demand outsourcing. Comput. Oper. Res. 41, 346–358.

Aksen, D., Piyade, N., Aras, N., 2010. The budget constrained r-interdiction median problem with capacity expansion. CEJOR 18 (3), 269–291.

Aksen, D., Aras, N., 2012. A bilevel fixed charge location model for facilities under imminent attack. Comput. Oper. Res. 39 (7), 1364–1381.

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

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

Cappanera, P., Scaparra, M.P., 2011. Optimal allocation of protective resources in shortest-path networks. Transport. Sci. 45 (1), 64–80.

Chopra, S., Sodhi, M.S., 2014. Reducing the risk of supply chain disruptions. MIT Sloan Manage. Rev. 55 (3), 73.

Church, R.L., Scaparra, M.P., 2007a. Analysis of Facility Systems' Reliability When Subject to Attack or a Natural Disaster. Critical Infrastructure. Springer, Berlin Heidelberg (pp. 221–241).

Church, R.L., Scaparra, M.P., 2007b. Protecting critical assets: the r-interdiction median problem with fortification. Geograph. Anal. 39 (2), 129–146.

Church, R.L., Scaparra, M.P., Middleton, R.S., 2004. Identifying critical infrastructure: the median and covering facility interdiction problems. Ann. Assoc. Am. Geogr. 94 (3), 491–502.

Cormican, K.J., Morton, D.P., Wood, R.K., 1998. Stochastic network interdiction. Operat. Res. 46 (2), 184–197.

Cui, T., Ouyang, Y., Shen, Z.M., 2010. Reliable facility location design under the risk of disruptions. Oper. Res. 58 (4-part-1), 998–1011.

Daskin, M.S., 2004. SITATION—Facility Location Software. Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.

Fahimnia, B., Jabbarzadeh, A., 2016. Marrying supply chain sustainability and resilience: a match made in heaven. Transport. Res. Part E: Logist. Transport. Rev. 91, 306–324.

Fiksel, J., 2006. Sustainability and resilience: toward a systems approach. Sustain.: Sci., Pract., Pol. 2 (2).

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. Eur. J. Oper. Res. 192 (1), 198–208.

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.

Hopp, W.J., Iravani, S.M., Liu, Z., 2012. Mitigating the impact of disruptions in supply chains. In: Supply Chain Disruptions. Springer, London, pp. 21–49.

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

Jeon, H.-M., Snyder, L.V., Shen, Z.-J.M., 2008. Location-Inventory Models With Supply Disruptions, Technical Report. Lehigh University, Bethlehem, PA.

Klibi, W., Martel, A., 2012. Modeling approaches for the design of resilient supply networks under disruptions. Int. J. Prod. Econ. 135 (2), 882–898.

Klibi, W., Martel, A., Guitouni, A., 2010. The design of robust value-creating supply chain networks: a critical review. Eur. J. Oper. Res. 203 (2), 283–293.

Lam, J.S.L., Bai, X., 2016. A quality function deployment approach to improve maritime supply chain resilience. Transport. Res. Part E: Logist. Transport. Rev. 92, 16–27.

Li, Q., Zeng, B., Savachkin, A., 2013. Reliable facility location design under disruptions. Comput. Oper. Res. 40 (4), 901–909.

Liberatore, F., Scaparra, M.P., 2011. Optimizing protection strategies for supply chains: comparing classic decision-making criteria in an uncertain environment. Ann. Assoc. Am. Geogr. 101 (6), 1241–1258.

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. Comput. Oper. Res. 38 (1), 357–366.

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

Lim, M., Daskin, M.S., Bassamboo, A., Chopra, S., 2010. A facility reliability problem: formulation, properties, and algorithm. Naval Res. Logist. 57 (1), 58–70.

Losada, C., Scaparra, M.P., O'Hanley, J.R., 2012. Optimizing system resilience: a facility protection model with recovery time. Eur. J. Oper. Res. 217 (3), 519–530.

Mak, H.Y., Shen, Z.J., 2012. Risk diversification and risk pooling in supply chain design. IIE Trans. 44 (8), 603–621.

Moore, J.T., Bard, J.F., 1990. The mixed integer linear bilevel programming problem. Operat. Res. 38 (5), 911–921.

Nejad, A.E., Niroomand, I., Kuzgunkaya, O., 2014. Responsive contingency planning in supply risk management by considering congestion effects. Omega 48, 19–35.

Niroomand, I., Kuzgunkaya, O., Bulgak, A.A., 2012. Impact of reconfiguration characteristics for capacity investment strategies in manufacturing systems. Int. J. Prod. Econ. 139 (1), 288–301.

Nooraie, S.V., Parast, M.M., 2016. Mitigating supply chain disruptions through the assessment of trade-offs among risks, costs and investments in capabilities. Int. J. Prod. Econ. 171, 8–21.

O'Hanley, J.R., Church, R.L., 2011. Designing robust coverage networks to hedge against worst-case facility losses. Eur. J. Oper. Res. 209 (1), 23–36.

Peng, P., Snyder, L.V., Lim, A., Liu, Z., 2011. Reliable logistics networks design with facility disruptions. Transport. Res. Part B: Methodol. 45 (8), 1190–1211.

Qi, L., Shen, Z.J.M., Snyder, L.V., 2010. The effect of supply disruptions on supply chain design decisions. Transport. Sci. 44 (2), 274–289.

Sawik, T., 2013. Selection of resilient supply portfolio under disruption risks. Omega 41 (2), 259–269.

Scaparra, M.P., Church, R.L., 2008a. A bilevel mixed-integer program for critical infrastructure protection planning. Comput. Oper. Res. 35 (6), 1905–1923.

Scaparra, M.P., Church, R.L., 2008b. An exact solution approach for the interdiction median problem with fortification. Eur. J. Oper. Res. 189 (1), 76–92.

Scaparra, M.P., Church, R.L., 2012. Protecting supply systems to mitigate potential disaster: a model to fortify capacitated facilities. Int. Reg. Sci. Rev. 35 (2), 188–210.

Schmitt, A.J., 2011. Strategies for customer service level protection under multi-echelon supply chain disruption risk. Transport. Res. Part B: Methodol. 45 (8), 1266–1283.

Snyder, L.V., 2006. Facility location under uncertainty: a review. IIE Trans. 38 (7), 547–564.

Snyder, L.V., Daskin, M.S., 2005. Reliability models for facility location: the expected failure cost case. Transport. Sci. 39 (3), 400–416.

Snyder, L.V., Atan, Z., Peng, P., Rong, Y., Schmitt, A.J., Sinsoysal, B., 2016. OR/MS models for supply chain disruptions: a review. IIE Trans. 48 (2), 89–109.

Tomlin, B., 2006. On the value of mitigation and contingency strategies for managing supply chain disruption risks. Manage. Sci. 52 (5), 639–657.

Torabi, S.A., Baghersad, M., Mansouri, S.A., 2015. Resilient supplier selection and order allocation under operational and disruption risks. Transport. Res. Part E: Logist. Transport. Rev. 79, 22–48.

Wang, W., Koren, Y., 2012. Scalability planning for reconfigurable manufacturing systems. J. Manufact. Syst. 31 (2), 83–91.

Wollmer, R., 1964. Removing arcs from a network. Operat. Res. 12, 934–940.

Wood, R.K., 1993. Deterministic network interdiction. Math. Comput. Model. 17, 1–18.