Computers & Operations Research 38 (2011) 357–366

Contents lists available at ScienceDirect

Computers & Operations Research

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

Analysis of facility protection strategies against an uncertain number of attacks: The stochastic R-interdiction median problem with fortification

Federico Liberatore [a,*], Maria P. Scaparra [a], Mark S. Daskin [b]

[a] Kent Business School, University of Kent, CT2 7PE Canterbury, Kent, UK
[b] Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109-2117, USA

ARTICLE INFO

Available online 16 June 2010

Keywords:
Fortification
Interdiction
Stochastic programming
Location analysis
Supply chains

ABSTRACT

We present the Stochastic R-Interdiction Median Problem with Fortification (S-RIMF). This model optimally allocates defensive resources among facilities to minimize the worst-case impact of an intentional disruption. Since the extent of terrorist attacks and malicious actions is uncertain, the problem deals with a random number of possible losses. A max-covering type formulation for the S-RIMF is developed. Since the problem size grows very rapidly with the problem inputs, we propose pre-processing techniques based on the computation of valid lower and upper bounds to expedite the solution of instances of realistic size. We also present heuristic approaches based on heuristic concentration-type rules. The heuristics are able to find an optimal solution for almost all the problem instances considered. Extensive computational testing shows that both the optimal algorithm and the heuristics are very successful at solving the problem. Finally, a discussion of the importance of recognizing the stochastic nature of the number of possible attacks is provided.

1. Introduction

Today more than ever the protection of infrastructure has become very important. Recent events have brought this issue to the forefront of public concern. In fact, identifying critical system components and planning the strengthening of their security and protection are certainly key elements for the sustainability and efficiency of service systems not only in the case of intentional attacks, but also when natural catastrophes occur.

This paper extends the *r*-interdiction median problem with fortification originally formulated by Church and Scaparra [8]. In that model, P facilities have already been located. The objective of the defender is to identify the location of Q of the P facilities to defend against an attacker who will attack and completely destroy R of the remaining facilities. The attacker's objective is to maximize the demand weighted average distance between demand nodes and the nearest of the remaining operational P–R facilities. Clearly, the defender wants to minimize the damage inflicted by the attacker by protecting or defending an appropriate selection of the original P facilities. One serious limitation of the work of Church and Scaparra is that the model assumes that the defender knows with certainty how many attacks will be sustained. In fact, most attackers (e.g., Al Qaeda) do not publicly announce and advertise their capabilities. Therefore, this paper extends Church and Scaparra's earlier work by introducing uncertainty, expressed as the probability distribution over the number of facilities to be attacked, in the number of attacks that will be sustained by the defender. We formulate the problem as a bi-level mixed integer programming model (Section 3) which is then restructured as a stochastic maximum covering problem (Section 4). Since the dimension of the problem grows quickly with respect to the number of facilities, the number of protected sites and the maximum number of attacks, we propose reduction rules based on the computation of upper and lower bounds on the objective function (Sections 5 and 6). These reductions allow us to optimally solve problem instances of significant size. We also introduce and test three heuristic-concentration rules which are combined to form two heuristic algorithms (Section 8). Computational tests show that these algorithms are very effective at solving the problem (Section 9). Finally, we use the model to illustrate the importance of recognizing uncertainty in the number of attacks. We also show numerically how the exact form of the probability distribution does not have a great impact on the quality of the solution obtained in terms of robustness (Section 10).

2. Background

A variety of quantitative approaches have recently been developed to identify cost-effective ways of increasing the

* Corresponding author.
E-mail addresses: fl51@kent.ac.uk (F. Liberatore), m.p.scaparra@kent.ac.uk (M.P. Scaparra), msdaskin@umich.edu (M.S. Daskin).

0305-0548/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cor.2010.06.002

robustness of infrastructure systems to external disruptions. A line of research focuses on the study of protection strategies which entail a full re-design of the networks so that they are intrinsically robust to attacks. For example, O'Hanley and Church [16] develop a resilient design problem for a coverage-type service system. The objective of the problem is to optimally locate a set of facilities to maximize a combination of initial demand coverage and the minimum coverage level following the loss of one or more facilities. The authors propose an approach based on the successive use of super-valid inequalities. Snyder and Daskin [23] extend the classical p-median and uncapacitated fixed-charge location problems to take into account possible failures of the facilities. The goal of the resulting reliability models is to choose facility locations that are both inexpensive and reliable as it also considers the expected transportation cost after possible facility failures. The programs are solved using an optimal Lagrangian relaxation algorithm. A similar problem has been addressed by Berman et al. [4] who develop a more general model where the facility disruption probabilities are not identical. The authors propose several exact and heuristic solution approaches and analyze the impact of the disruption probabilities on the centralization and co-location of the facilities. Finally, Lim et al. [15] study the design of robust supply systems where both reliable and unreliable facilities can be located with different levels of investment. They propose a continuous model which provides valuable insights about the relationship between the failure probabilities and the optimal number of non-hardened sites. The robustness of these insights is then validated through the use of a mixed integer program.

A different line of research dealing with security issues focuses on the identification of critical components through the use of interdiction models. Interdiction models, which were first introduced by Wollmer [24] in 1964, have been extensively studied over the past few decades, especially within the context of network flow problems. The analysis of network interdiction models has been performed with respect to different reliability measures, such as connectivity, distance (or cost) and capacity. A survey of these models can be found in [9]. More recently, Grubesic and Murray in [11] have addressed the problem of losing critical infrastructure elements that are geographically connected and explore the topological complexities associated with network interconnections. Bell in [3] illustrates a game between a router and a virtual network tester. The originality of the problem is that the router wants to minimize the cost of the flow of packets or vehicles in the network while the virtual network tester strikes the link to maximize the cost of the trip. Therefore the method proposed identifies the components of the networks whose disruption would damage performance the most. Lim and Smith [14] consider a network interdiction problem on a multicommodity flow network. An attacker disables some of the arcs, according to an interdiction budget, with the objective of minimizing the maximum shipping profit. The authors consider both the cases where the interdiction is discrete—either an arc is safe or destroyed—and continuous—the attack may reduce the capacity of the arcs partially. Interdiction problems identify critical facilities they can be used to develop protection strategies. Smith et al. exploit this idea in [22] where they extend their previous work [14] by adding an additional design layer. The resulting problem is a three-level, two-player game in which a designer first constructs a network. Next, as in the previous work, an interdictor destroys a set of arcs and finally the designer decides the set of flows through the network. Interdiction problems within the location analysis framework have been studied by Church et al. [9] who consider the problem of identifying the most critical facilities in supply systems with different service protocols. They propose two different models: the r-interdiction median problem and the r-interdiction covering problem, which are based on the p-median problem and on the max covering problem, respectively.

The identification of critical system components is only the first step towards the development of sound and economically efficient fortification strategies. The need to model protection efforts explicitly has been acknowledged in several recent works such as [6,17,8]. Most studies in this area use a game theoretic approach and formulate protection problems as bi-level defender-attacker models. Brown et al. [6] provide an excellent introduction to bi-level and tri-level problems that involve the presence of an intelligent attacker and a defender. They also describe some applications to electric power grids, subways, airports and other critical infrastructure. Qiao et al. [17] develop a max–min model to allocate a security budget to a water supply network so as to make the water infrastructure more resilient to physical attacks. Zhuang and Bier [25] formulate basic equilibrium models for both sequential and simultaneous games between an attacker and a defender. They also provide interesting insights related to the effects of risk attitudes on the attacker and defender decisions, and to the issue of balancing protection from terrorism and from natural disasters. In [2], Azaiez and Bier consider a problem where the defender's objective is to maximize the minimum expected cost of a feasible attack, subject to a budget constraint on the defensive investment. Bier [5] discusses the policy implications of a game-theoretic model of security investment where the attacker's goals are uncertain. Interestingly, one of Bier's conclusions is that, as counterintuitive as it may seem, it is preferable to announce which targets have been defended so that the attention of the attacker can be diverted toward less damaging objectives.

2.1. Fortification models in location analysis

Church and Scaparra [8] extend their previous interdiction models [9] to explicitly include protection decisions. In two subsequent works, the authors develop two different solution approaches for the resulting r-interdiction median problem with fortification. The first approach [21] is based on a reformulation of the problem as a maximal covering model with precedence constraints. The reformulation is obtained by explicitly enumerating all the possible ways of interdicting a subset of size r of the facilities operating within the system. Each combination is referred by the authors as an *interdiction pattern*. The cost associated with each interdiction pattern corresponds to the final cost of the underlying median system after the facilities in the set are interdicted; i.e., the cost of assigning each customer to the closest non-interdicted facility. The model allocates protection resources so as to thwart (or cover) the most disruptive interdiction patterns; i.e., the patterns with the highest associated cost. The dimension of the model is reduced using a linear interpolation search procedure that exploits properties of the coverage function, to fix and remove unnecessary patterns and variables. The interpolation search procedure iteratively calculates a lower bound on the maximum number of the most disruptive interdiction patterns that can be covered by solving a set covering problem at each iteration. The second approach [20] is a tree search algorithm that takes advantage of the bi-level formulation of the problem. Both methods proved quite effective in solving large instances of this deterministic facility protection model with modest computing effort. In this paper we present the stochastic R-interdiction median problem with fortification, an extension of the r-interdiction median problem with fortification where the number of possible losses or attacks is uncertain.

3. The stochastic *R*-interdiction median problem with fortification

Let N, indexed by i, represent the set of customers. Every customer i is characterized by a demand a_i. Let P represent the number of operating facilities in the system and let F, indexed by j, denote the set of facilities. The distance between facility j and customer i is d_{ij}. Fortification resources are limited and it is possible to protect exactly Q facilities. There is no certainty about the number of interdictions that will take place. We assume that the attacker would be able to interdict at most R facilities and that protected facilities cannot be interdicted. Thus, it necessarily follows that $R \le P-Q$. We associate with each $r=1,\ldots,R$ a probability p_r that gives the likelihood that the attacker will be able to interdict exactly r facilities. These probabilities must sum to 1:

$$\sum_{r=1}^{R} p_r = 1$$

Although an empirical estimation of the probabilities p_r is very difficult due to a general lack of sufficient historical data, the results of our model do not seem to be overly sensitive to the probability vector components. This issue is further investigated in Section 10, where we analyze two antithetical cases that model two possible realistic situations: a decreasing probability function that assigns higher probabilities to lower values of r, and an increasing probability function which indicates that more emphasis is placed on countering scenarios with a large number of losses. From the analysis of the solutions found with different probability distributions, the decision maker can choose the protection strategy that seems most adequate for the system under study.

Let the set T_{ij}, $\forall i \in N$, $\forall j \in F$ is the set of existing facilities (not including j) that are farther than j is from demand i:

$$T_{ij} = \{k \in F \mid k \ne j \text{ and } d_{ik} > d_{ij}\}, \quad \forall i \in N,\ \forall j \in F$$

The problem can be represented as a competitive discrete bi-level problem: an *interdictor*, corresponding to the lower level program, wants to destroy facilities to do as much damage as possible to the system while a defender, corresponding to the upper level program, decides which facilities to protect to minimize the damage resulting from the attack.

The decision variables are the following:

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

$$s_j = \begin{cases} 1 & \text{if a facility located at } j \text{ is eliminated by interdiction} \\ 0 & \text{otherwise} \end{cases}$$

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

The bi-level formulation of the stochastic r-interdiction median problem with fortification (S-RIMF) is

$$\min Z^{\star} = \sum_{r=1}^{R} p_r W_r(\mathbf{z}) \tag{1}$$

$$\sum_{j \in F} z_j = Q \tag{2}$$

$$z_j \in \{0,1\} \quad \forall j \in F \tag{3}$$

where for each $r=1,\ldots,R$, $W_r(\mathbf{z})$ is the solution to the following optimization problem:

$$W_r(\mathbf{z}) = \max \sum_{i \in N} a_i d_{ij} x_{ij} \tag{4}$$

$$\sum_{j \in F} x_{ij} = 1 \quad \forall i \in N \tag{5}$$

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

$$\sum_{k \in T_{ij}} x_{ik} \le s_j \quad \forall i \in N,\ \forall j \in F \tag{7}$$

$$1 - s_j \ge z_j \quad \forall j \in F \tag{8}$$

$$s_j \in \{0,1\} \quad \forall j \in F \tag{9}$$

$$x_{ij} \in \{0,1\} \quad \forall i \in N,\ \forall j \in F \tag{10}$$

Note that each lower level problem is an r-interdiction median problem (RIM) with the additional conditional constraint (8) that forbids the interdiction of fortified facilities. The mathematical programming model of RIM was first introduced in [9].

The objective function of the leader problem (1) minimizes the expected worst-case demand weighted distance where the expectation is taken over all possible values of the number of attacks. Constraint (2) stipulates that the defender can only protect Q sites. The interdictor's objective (4) is to maximize the demand weighted distance that results from r attacks on unprotected sites. Constraint (5) states that every demand node must be assigned. Constraint (6) permits exactly r attacks. Constraint (7) is the closest assignment constraint that says that for any facility site j, demands can only be assigned to locations further than j if site j is interdicted. Constraint (8) allows the interdiction of undefended sites only. Constraints (3), (9) and (10) are standard binary constraints on the key decision variables.

The deterministic version of the bi-level problem where r is fixed can be reformulated as a single level mixed-integer program and solved using the solution approach presented in [21] (for a description of the methodology, please refer to Section 2.1). The resulting model is a max-covering problem with precedence constraints. Although this reformulation requires enumerating all possible ways of losing r out of P facilities, the method described in [21] is quite efficient and can solve problem instances of considerable size. Unfortunately, this kind of reformulation cannot be adapted in a straightforward way to the stochastic version of the problem. In this paper we present an alternative max-covering formulation that can be easily adjusted to model the stochastic problem. Moreover the solution approach in [21] was tailored to the particular structure of the max-covering formulation with precedence constraints and, hence, cannot be applied to our new formulation. In this paper, we also propose a novel solution approach to solve the max-covering formulation of the S-RIMF.

4. A stochastic max-covering type formulation

Let H_r, indexed by h, be the set of all the interdiction patterns in which the attacker interdicts exactly r facilities. Each pattern h has an interdiction set I_h and a cost c_h associated with it. The cost c_h is calculated by assigning every customer $i \in N$ to the closest non-interdicted facility $j \in F/I_h$, as shown in Algorithm 1.

Note that

$$|I_h| = r, \quad \forall h \in H_r,\ r = 1,\ldots,R$$

Algorithm 1. Calculation of cost c_h associated to pattern h

input: N, F, $a_i \forall i \in N$, $d_{ij} \forall i \in N, j \in F$, h, I_h
output: c_h
begin
  c_h=:0
  for $i \in N$ do //loop over demands
    $\bar{j} = \mathrm{argmin}_{j \in F/I_h}\{d_{ij}\}$ //get index of closest non-interdicted site
    $c_h = c_h + a_i d_{i\bar{j}}$ //update cost
  done
end

Scaparra and Church in [21] proved that the worst-case interdiction will occur for an interdiction pattern attacking solely unprotected sites. We call an interdiction pattern h *covered* if any of the facilities $j \in I_h$ is fortified. Under this assumption, for each possible value of r the worst-case interdiction pattern in response to a given fortification strategy is the uncovered interdiction pattern with the highest cost.

We can now introduce the new max-covering formulation (MCP) of the deterministic version of the bi-level problem where r is fixed. The formulation requires the following additional decision variables:

$$y_h = \begin{cases} 1 & \text{if the interdiction pattern } h \text{ is covered} \\ 0 & \text{otherwise} \end{cases}$$

$\overline{W}_r$: cost of the worst-case interdiction pattern when exactly r facilities are interdicted. The MCP is formulated as follows:

$$\min \overline{W}_r \tag{11}$$

$$\sum_{j \in F} z_j = Q \tag{12}$$

$$\sum_{j \in I_h} z_j \geq y_h \quad \forall h \in H_r \tag{13}$$

$$\overline{W}_r \geq c_h(1-y_h) \quad \forall h \in H_r \tag{14}$$

$$z_j \in \{0,1\} \quad \forall j \in F \tag{15}$$

$$0 \leq y_h \leq 1 \quad \forall h \in H_r \tag{16}$$

The cardinality constraint (12) requires the number of fortifications to be exactly Q. Since there is no benefit to fortifying less than Q facilities the cardinality constraint can be relaxed and expressed as an inequality, rather than an equality:

$$\sum_{j \in F} z_j \leq Q \tag{17}$$

Constraints (13) are standard covering constraints. To cover an interdiction pattern h, at least one of the facilities in the interdiction set I_h must be fortified. Constraints (14) are min–max constraints and assure that the cost $\overline{W}_r$ of the worst-case pattern is the cost of the most expensive uncovered interdiction pattern. The objective (11) is to minimize this cost. Lastly, constraints (15) and (16) impose the conditions of integrality and non-negativity over the relevant variables. The solution to this program provides the set of optimal fortifications, $\overline{\mathbf{z}}^r$, for a given r.

To extend the MCP to the stochastic case, it is sufficient to optimize over all the possible values of r=1,…,R at the same time and to take into account the probabilities p_r, $\forall r = 1,\ldots,R$ in the objective function. The resulting formulation, called S-MCP, is

$$\min Z^* = \sum_{r=1}^{R} p_r W_r \tag{18}$$

$$\sum_{j \in F} z_j \leq Q \tag{19}$$

$$\sum_{j \in I_h} z_j \geq y_h \quad \forall h \in H_r,\ \forall r = 1,\ldots,R \tag{20}$$

$$W_r \geq c_h(1-y_h) \quad \forall h \in H_r,\ \forall r = 1,\ldots,R \tag{21}$$

$$z_j \in \{0,1\} \quad \forall j \in F \tag{22}$$

$$0 \leq y_h \leq 1 \quad \forall h \in H_r,\ \forall r = 1,\ldots,R \tag{23}$$

The objective of this program (18) is to minimize the weighted sum of the costs W_r associated with the worst-case interdiction patterns for every feasible value of r. Since the weights are represented by the probabilities p_r, the program minimizes the expected cost of the worst-case interdiction pattern across all the possible values of r. When solved to optimality—i.e. by using a generic MIP solver as CPLEX [13]—the program will find the optimal fortification set $\mathbf{z}^*$ that minimizes this expected cost.

The model has $R+P+\sum_{r=1}^{R}|H_r| = R+P+\sum_{r=1}^{R}\binom{P}{r}$ variables and $1+2\sum_{r=1}^{R}|H_r| = 1+2\sum_{r=1}^{R}\binom{P}{r}$ constraints. It is straightforward to see that the program grows linearly with respect to the number of interdiction patterns which, in turn, grows exponentially with respect to P and R. Thus for high value of P and R the problem can easily become intractable. In the next section, we show how the dimension of the program can be significantly reduced by removing some of the interdiction patterns and by fixing some variables to their optimal values.

5. A lower bound

In the previous section we showed that the dimension of the stochastic model grows very quickly. By calculating a lower bound on the optimal objective value for every occurrence of r (i.e. for every possible scenario) it is possible to remove from the model many interdiction patterns without affecting the solution.

To this end we solve independently R MCPs, one for each possible value of r. The solution of this program provides the set of optimal fortifications for a given r. In this and the next section, we refer to these variables as $\overline{\mathbf{z}}^r$, to distinguish between the optimal solution to the stochastic program ($\mathbf{z}$), and those of the deterministic problems ($\overline{\mathbf{z}}^r$). Since in this program we optimize only for a single scenario, the value of the optimal solution to this problem $\overline{W}_r$ is a lower bound on W_r:

$$W_r \geq \overline{W}_r$$

Once we know all the optimal values $\overline{W}_r$, $1 \leq r \leq R$, we can compute a lower bound to the value of the stochastic max-covering problem:

$$\overline{Z} = \sum_{r=1}^{R} p_r \overline{W}_r$$

This approach corresponds to the resolution of the Wait-and-See problem, as shown in [12].

As already stated, each $\overline{W}_r$ is a lower bound for the corresponding W_r. Therefore we can remove from the original stochastic problem all the $h \in H_r$ such that the relative cost c_h is less than $\overline{W}_r$, since those patterns will not affect the value of the optimal solution. The resulting reduced sets of interdiction patterns are

$$\overline{H}_r = H_r \backslash \{h | c_h < \overline{W}_r\}, \quad r = 1, \ldots, R$$

6. An upper bound

We can take advantage of the solutions obtained during the calculation of the lower bound (Section 5) to get a useful upper bound on the solution value of the S-MCP. We apply every fortification set $\overline{\mathbf{z}}^r$ to the stochastic r-interdiction median problem with fortification and calculate the objective value. To do so we need to solve R independent RIM problems with the additional constraint (8) that forbids the interdiction of fortified facilities $\overline{\mathbf{z}}^r$ and objective function $Z_m^{RIM}(\overline{\mathbf{z}}^r)$. Recall that the RIM for a particular value of m (the number of interdicted sites) will give the optimal interdictions of m sites given that the sites in fortification pattern $\overline{\mathbf{z}}^r$ are fortified. Once we know the values $Z_m^{RIM}(\overline{\mathbf{z}}^r)$, $m = 1,\ldots,R$ of the solutions to the RIM programs, we can calculate the value of the stochastic r-interdiction median problem with fortification relative to $\overline{\mathbf{z}}^r$:

$$\tilde{Z}_r = \sum_{m=1}^{R} p_m Z_m^{RIM}(\overline{\mathbf{z}}^r)$$

The upper bound $\tilde{Z}$ is given by the best (lowest) $\tilde{Z}_r$, $r = 1,\ldots,R$:

$$\tilde{Z} = \min_{r=1,\ldots,R} \tilde{Z}_r$$

$\tilde{Z}$ can be used by the branch-and-bound optimization algorithm of CPLEX to reduce the optimization time. It can also be used to calculate an upper bound on each W_r^*; i.e. on the value of W_r, $r = 1,\ldots,R$ in the optimal solution to S-MCP. To this end, consider the following sequence of inequalities:

$$\tilde{Z} \geq Z^* = \sum_{m=1}^{R} p_m W_m^* \geq \sum_{\substack{m=1,\\ m \neq r}}^{R} p_m \overline{W}_m + p_r W_r^*$$

Hence

$$W_r^* \leq \frac{\tilde{Z} - \sum_{m=1,\, m \neq r}^{R} p_m \overline{W}_m}{p_r}$$

Therefore

$$\tilde{W}_r = \frac{\tilde{Z} - \sum_{m=1,\, m \neq r}^{R} p_m \overline{W}_m}{p_r}$$

is an upper bound for W_r^*.

We can calculate the upper bound of the relative worst-case interdiction pattern $\tilde{W}_r$ for every scenario $r = 1,\ldots,R$. These upper bounds are very useful, since they can be used to fix some variables and reduce the number of constraints in the stochastic model. Let us define the following set of fixed interdiction patterns as

$$\tilde{H}_r = \{h | c_h > \tilde{W}_r\}, \quad r = 1, \ldots, R$$

In the original model we can fix to 1 the variables y_h corresponding to interdiction patterns belonging to $\tilde{H}_r$. In fact, any pattern $h \in \tilde{H}_r$ must be covered in an optimal solution to (18)–(23). It follows that the constraints (20) relative to these patterns become

$$\sum_{j \in I_h} z_j \geq 1 \quad \forall h \in \tilde{H}_r,\ r = 1, \ldots, R$$

and that the associated constraints (21) can be removed from the model.

7. The reduced stochastic max-covering model

The resulting reduced model RS-MCP is

$$\min Z^* = \sum_{r=1}^{R} p_r W_r \tag{24}$$

$$\sum_{j \in F} z_j \leq Q \tag{25}$$

$$\sum_{j \in I_h} z_j \geq 1 \quad \forall h \in \tilde{H}_r,\ r = 1, \ldots, R \tag{26}$$

$$\sum_{j \in I_h} z_j \geq y_h \quad \forall h \in \overline{H}_r \backslash \tilde{H}_r,\ r = 1, \ldots, R \tag{27}$$

$$W_r \geq c_h(1 - y_h) \quad \forall h \in \overline{H}_r \backslash \tilde{H}_r,\ r = 1, \ldots, R \tag{28}$$

$$z_j \in \{0,1\} \quad \forall j \in F \tag{29}$$

$$0 \leq y_h \leq 1 \quad \forall h \in \overline{H}_r \backslash \tilde{H}_r,\ r = 1, \ldots, R \tag{30}$$

Computational experiments (see Section 9) demonstrated that, thanks to the pre-processing on the interdiction patterns and the bounds, the solution time can be drastically reduced.

8. A heuristic approach

In this section we discuss three heuristic rules that can be used to reduce the search space and speed up the computation. Although the use of these rules within a solution algorithm does not guarantee that the optimality of the solution is preserved, they can be used to develop efficient heuristics for the S-MCP.

All the rules are based on an idea similar to *heuristic concentration* [19]. As Rosing and Hodgson illustrate in [18] the heuristic concentration procedure is based on a two-stage approach:

1. Multiple runs of some heuristic for the problem are used to produce a relatively small *concentration set*,
2. The problem is solved—either heuristically or to optimality—over the reduced search space of the concentration set

The rules used in our work generate different concentration sets which are subsequently used to fix some variables.

First rule: always fortified. Let F^r be the set of facilities which are fortified in the optimal solution to the deterministic MCP with exactly r interdictions; i.e. $F^r = \{j | \overline{z}_j^r = 1\}$. By solving R deterministic MCPs, one for each value of r, we obtain R different fortification sets. If these sets have some facilities in common, it is reasonable to think that the optimal fortification set of the S-MCP, $\mathbf{z}^\star$, may contain them. Then let

$$F^{AF} = \left\{ \bigcap_{r=1}^{R} F^r \right\}$$

By substituting the original cardinality constraint (19) in the formulation of the S-MCP with

$$\sum_{j \in F/F^{AF}} z_j \leq Q - |F^{AF}|$$

and by setting

$$z_j = 1 \quad \forall j \in F^{AF}$$

we can impose the fortification of the facilities contained in the set F^{AF}.

Second rule: action set. Let F^{AS}, referred to as the action set, be the set of facilities built from the union of the fortification sets F^r resulting from the optimization of R independent MCPs:

$$F^{AS} := \left\{ \bigcup_{r=1}^{R} F^r \right\}$$

To limit the search space to the facilities identified by the action set, we need to substitute the cardinality constraint (19) with the following

$$\sum_{j \in F^{AS}} z_j \le Q$$

and to fix the decision variables corresponding to the facilities not included in the action set

$$z_j = 0 \quad \forall j \in F/F^{AS}$$

Third rule: reaction set. The third rule was born from the idea of extending the action set to include also the facilities belonging to the worst-case interdiction patterns. Let I^r be the interdiction set I_h such that $c_h = \overline{W}_r$; i.e. I^r is the worst-case interdiction pattern in response to the optimal fortifications $\overline{z}_j^r$ of the deterministic MCP with exactly r interdictions. The *reaction set* F^{RS} is defined as the set of facilities which are interdicted in response to the optimal fortification strategy:

$$F^{RS} := \left\{ \bigcup_{r=1}^{R} I^r \right\}$$

As with the second rule, we can reduce the search space by replacing constraint (19) with

$$\sum_{j \in F^{AS} \cup F^{RS}} z_j \le Q$$

and fixing the remaining variables

$$z_j = 0 \quad \forall j \in F/\{F^{AS} \cup F^{RS}\}$$

The three conjectures illustrated can be combined to develop different variants of heuristic concentration type solution approaches. The lower bound and the upper bound illustrated, respectively, in Section 5 and in Section 6 can still be applied when using these rules. In fact, a lower bound to the optimal solution is also a lower bound for any primal heuristic. The upper bound is the best solution to S-MCP among all the possible solutions to the deterministic MCPs. All the rules optimize over a set of facilities that contains all the fortification sets F^r. Therefore, since the upper bound is calculated using only facilities included in the fortification sets F^r, its value cannot be less than the value found using any of the rules.

9. Computational tests

In this section we present the computational tests that have been run to evaluate the performance of the methodologies presented in this paper. We first provide some generic information about how the experiments were conducted, including information on the different variants of the algorithms tested, the data sets used in the experiments, and the parameter settings. We then describe the branching priority strategy that has been adopted in the branch-and-cut MIP solver. The last subsection deals with the analysis of the computational times of the tests.

Tools and experiments: In this work we developed an algorithm that incorporates all the functionality described in this paper. The algorithm has been implemented in C++ and compiled using Microsoft Visual C++ .NET 2003. To solve the MIP problems we used the generic MIP solver CPLEX 9.1. The bounds and the rules can be easily turned on and off using compiler directives. From the original algorithm we obtained four different optimization programs: one that solves the stochastic max-covering formulation (S-MCP), one that solves the reduced formulation using the bounds (RS-MCP), a heuristic that exploits the rules Always Fortified and Action Set (Heur1) and, finally, another heuristic that makes use of all the rules (Heur2). Both the heuristics also use the bounds.

The tests have been run on a computer equipped with an Intel Core 2 CPU 6700 @ 2.66 GHz, 2 GB of RAM and Windows XP Professional operating system. All the programs are single-threaded, use only one processor at a time and use the same configuration of the CPLEX parameters. We tested the algorithms using two different data sets: London and USCities. The data set London (Ontario) has 150 nodes and was first introduced in [10] while the data set USCities contains the 263 largest cities in the contiguous United States according to the 2000 census [1].

The tests have been run over a wide number of combinations of the parameters P, Q and R. P takes on values of 40, 50 and 60. The set of facilities considered correspond to the optimal solutions to the P-Median Problem. The value of R is kept relatively small, ranging between 2 and 5. Q was set to a proportional value of P: 10%, 15% and 20% (rounded up to the next integer when fractional). Every test has a computational time limit of 1 h and a physical memory limit of 1 GB.

Three different probability distributions p_r have been employed

$$p_r = 2\frac{r}{R(R+1)}, \quad r = 1, \ldots, R \tag{31}$$

$$p_r = 2\frac{R-r+1}{R(R+1)}, \quad r = 1, \ldots, R \tag{32}$$

$$p_r = \frac{1}{R}, \quad r = 1, \ldots, R \tag{33}$$

The function (31) is monotonically increasing. With this choice, higher probabilities are associated with higher values of r, to indicate that more emphasis is placed on countering scenarios with a large number of attacks. The second function (32) is monotonically decreasing and consequently assigns higher probabilities to lower values of r. This distribution has been chosen to model the average behavior of terrorist attacks that generally tend to be focused on a small number of targets. The last one (33) is a uniform probability function, which assigns the same probability to all the possible outcomes. Since all the distributions produce a similar trend in the solution time of the algorithms, for the sake of brevity only the results corresponding to the first probability function are reported.

Branching priority heuristic: To calculate the bounds and the concentration sets for the rules it is necessary to solve R independent MCPs, one for each possible value of r. Every problem returns the corresponding optimal fortification sets F^r and the interdiction set I^r. Following the same idea of the heuristic rules, it is reasonable to assume that a facility that appears frequently in the deterministic fortification sets has a higher probability of appearing in the optimal solution of the stochastic problem. Therefore the information provided by the fortification sets can be exploited to produce a heuristic ordering for the branching variables. To each facility $j \in F$ we associate a priority value ϕ_j that represents the number of times that the corresponding facility is fortified in the optimal solution of each MCP:

$$\phi_j = \sum_{r=1}^{R} \bar{z}_j^r \quad \forall j \in F$$

The priority coefficients ϕ_j are subsequently provided to CPLEX that assigns higher branching priority to the variables with higher priority coefficient.

Bounds and rules: what we achieved. Tables 1 and 2 display the value of the optimal solution and the computation time of the algorithms for the data sets London and USCities, respectively. The first three columns are the parameters P, Q and R. The fourth column contains the objective value Z^* of the optimal solution for each instance. Next we have the average computational times (expressed in seconds) of the algorithms calculated over five independent runs of each algorithm. An asterisk is placed next to the solution time when the algorithm found a sub-optimal solution. A dash indicates that the algorithm was not able to complete the optimization because the instance exceeded either the time or the memory limit. In the last row the average solution times for RS-MCP and the heuristic algorithms are shown. Note that the solution times for RS-MCP include the time necessary to complete the bounding procedure: (1) solving the R MCPs by means of reformulation as described in [21] (for a description of the methodology, please refer to Section 2.1), (2) calculating the bounds, and (3) identifying the reduced sets of interdiction patterns $\overline{H}_r$ and the set of fixed interdiction patterns $\tilde{H}_r$. Table 3 presents which percentage of the total RS-MCP solution time is dedicated to the bounding procedure, and which percentage is needed to optimize the reduced program for each of the data sets considered.

The most noticeable observation is that S-MCP was able to solve within 1 h only a very limited number of instances while, thanks to the new bounds and the reductions, RS-MCP could solve almost all the instances. Moreover, RS-MCP was able to find the optimal solution in only a fraction of the time required by the S-MCP formulation when the latter could find the solution. Only two instances are still unsolved by all four algorithms: London and USCities with parameters $P=60$, $Q=12$ and $R=5$ where the algorithms were interrupted because of the memory limit. Therefore the reductions proved to be very useful and effective. By using the bounds it is indeed possible to solve instances of realistic dimension in a very short time.

Table 1
Data set London—Algorithms solution time comparison.

P	Q	R	Z^*	Time (s)			
				S-MCP	RS-MCP	Heur1	Heur2
40	4	2	73,953.98	0.31	0.03	0.04	0.06
40	4	3	77,944.60	115.00	0.13	0.13	0.13
40	4	4	82,217.60	–	0.56	0.54	0.59
40	4	5	86,396.73	–	4.07	3.89	4.30
40	6	2	73,781.74	1.58	0.03	0.03	0.03
40	6	3	77,685.54	462.70	0.11	0.11	0.12
40	6	4	81,640.34	–	0.55	0.52	0.57
40	6	5	85,522.47	–	5.54	5.36	5.97
40	8	2	73,401.42	2.70	0.03	0.03	0.03
40	8	3	77,076.18	794.40	0.11	0.11	0.12
40	8	4	80,963.72	–	1.25	1.19	1.33
40	8	5	84,697.62	–	27.55	20.36	23.17
50	5	2	58,529.73	1.80	0.03	0.03	0.03
50	5	3	61,915.36	1,007.00	0.13	0.12	0.12
50	5	4	65,155.98	–	1.14	1.10	1.11
50	5	5	68,335.60	–	16.85	16.43	16.50
50	8	2	57,753.83	3.30	0.04	0.04	0.04
50	8	3	60,726.85	2,781.00	0.13	0.13	0.13
50	8	4	63,757.20	–	1.58	1.50	1.49
50	8	5	66,875.69	–	40.36	39.55	39.59
50	10	2	57,283.29	8.41	0.05	0.04	0.04
50	10	3	59,772.23	3,333.00	0.17	0.17	0.17
50	10	4	63,065.32	–	2.49	2.12	2.12
50	10	5	66,035.96	–	718.12	167.48	170.46
60	6	2	45,365.71	4.59	0.03	0.05	0.03
60	6	3	48,207.24	–	0.17	0.17	0.17
60	6	4	50,797.93	–	2.13	2.09	2.08
60	6	5	53,403.87	–	443.84	443.00	443.14
60	9	2	44,776.85	14.31	0.44	0.05	0.03
60	9	3	47,383.84	–	0.19	0.19	0.18
60	9	4	50,014.09	–	4.38	2.37	2.38
60	9	5	52,377.06	–	266.42	232.34	232.32
60	12	2	44,225.01	27.89	0.40	0.05	0.04
60	12	3	46,830.21	–	0.36	0.28	0.28
60	12	4	48,902.84	–	7.66	7.00	7.00
60	12	5	–	–	–	–	–
Average			–	–	44.20	27.10	27.31

Table 2
Data set USCities—Algorithms solution time comparison.

P	Q	R	Z^*	Time (s)			
				S-MCP	RS-MCP	Heur1	Heur2
40	4	2	3,383,035,777.67	0.27	0.03	0.03	0.03
40	4	3	4,019,125,339.00	41.10	0.11	0.09	0.10
40	4	4	4,476,779,495.50	–	0.67	0.59	0.60
40	4	5	4,866,585,540.33	–	4.81	4.41	4.48
40	6	2	3,286,216,658.33	0.39	0.04	0.03*	0.03*
40	6	3	3,760,914,991.00	79.93	0.13	0.11	0.11
40	6	4	4,109,323,708.20	–	0.94	0.83	0.84
40	6	5	4,448,644,420.20	–	14.00	10.22	10.35
40	8	2	3,120,104,252.33	0.48	0.04	0.03	0.04
40	8	3	3,450,922,837.50	116.60	0.14	0.13	0.13
40	8	4	3,782,176,529.30	–	2.12	1.79	1.77
40	8	5	4,059,997,875.00	–	37.75	29.68	29.77
50	5	2	2,428,424,600.33	0.61	0.05	0.03	0.05
50	5	3	2,795,869,185.00	291.10	0.16	0.14	0.16
50	5	4	3,063,770,060.60	–	1.42	1.34	1.35
50	5	5	3,365,535,167.40	–	15.23	14.34	14.36
50	8	2	2,286,362,847.33	0.91	0.05	0.03	0.04
50	8	3	2,687,197,980.83	1,471.00	0.19	0.16	0.17
50	8	4	2,931,244,316.40	–	1.95	1.78	1.78
50	8	5	3,165,552,184.60	–	77.89	42.80	42.81
50	10	2	2,228,271,926.00	1.48	0.05	0.03	0.05
50	10	3	2,446,813,824.17	1,296.00	0.23	0.20	0.22
50	10	4	2,594,055,200.10	–	3.66	3.38	3.43
50	10	5	2,763,407,634.73	–	321.76	318.46	318.78
60	6	2	1,889,173,705.67	1.20	0.04	0.04	0.03
60	6	3	2,119,472,868.67	507.50	2.05	0.23	0.23
60	6	4	2,439,201,722.40	–	3.15	2.79	2.72
60	6	5	2,706,738,479.53	–	38.84	35.41	35.36
60	9	2	1,801,763,821.00	1.81	0.34	0.03	0.03
60	9	3	1,970,406,925.50	–	1.61	0.27	0.27
60	9	4	2,234,785,151.30	–	21.02	3.54	3.52
60	9	5	2,418,063,692.67	–	498.24	79.44	79.16
60	12	2	1,696,507,683.67	1.97	0.41	0.03	0.04
60	12	3	1,841,699,671.83	–	1.43	0.34	0.34
60	12	4	2,007,633,404.70	–	9.69	8.61	8.60
60	12	5	–	–	–	–	–
Average			–	–	30.29	16.04	16.05

Table 3
Bounding and optimization total solution time average percentages (%).

Data set	Bounding	Optimization
London	82.26	17.74
randompoints3	81.22	18.78
USCities	74.41	25.59

Concerning the heuristic rules, Heur1 and Heur2 found a suboptimal solution in only one instance (i.e., data set USCities with P=40, 6, and R=2) with an optimality gap equal to 0.137%. As presented in Tables 1 and 2, both heuristics improved the average solution time when compared to RS-MCP, and Heur2 is slightly slower than Heur1. For both the algorithms, the improvement is higher for higher values of P. As expected, the use of the third rule—Reaction Set—is computationally more expensive because it extends the solution space to all the facilities included in the interdiction sets I^r and therefore the solution time increases. Interestingly, despite of this enlargement of the search area, for the tests reported in this paper Heur2 solved to optimality exactly the same number of instances solved to optimality by Heur1.

The first heuristic rule and the bounds can be used to fix the variables relative to some of the interdiction patterns $h \in H_r$ and remove the related constraints. Since the number of variables and constraints strongly depends on the number of interdiction patterns, it is desirable to reduce them as much as possible. Table 4 presents the percentages of interdiction patterns that are still free in the problem when only the bounds are used—column RS-MCP—and when the bounds are used in conjunction with the first heuristic rule—column Heur—for each data set. The last row of the table presents the average percentage of remaining free interdiction patterns for each data set. Using the bounds alone reduces the number of interdiction patterns by more than 75%. When the bounds are used in conjunction with the heuristic rule always fortified, the average number of remaining decision patterns drops to less than 0.2%. Moreover this configuration is more effective as the number of facilities P and the number of interdictions R grow. In the data sets considered, the percentage of remaining patterns is generally higher for higher P and R values. Using the heuristic rule reaction set does not give any further contribution.

Table 4
Comparison of the percentage (%) of free interdiction patterns.

P	Q	R	London		USCities	
			RS-MCP	Heur	RS-MCP	Heur
40	4	2	10.122	0.488	19.024	0.244
40	4	3	14.308	0.075	14.364	0.131
40	4	4	18.719	0.012	18.731	0.024
40	4	5	23.048	0.003	23.056	0.011
40	6	2	10.610	0.976	23.537	0.366
40	6	3	14.346	0.112	27.121	0.075
40	6	4	18.729	0.022	34.686	0.031
40	6	5	23.051	0.006	32.905	0.056
40	8	2	23.902	0.732	28.049	0.610
40	8	3	21.084	0.271	33.196	0.252
40	8	4	27.090	0.082	41.775	0.091
40	8	5	32.875	0.026	49.538	0.079
50	5	2	11.843	0.314	11.843	0.314
50	5	3	11.574	0.062	11.583	0.072
50	5	4	15.192	0.014	15.186	0.008
50	5	5	18.764	0.005	18.760	0.001
50	8	2	22.745	0.392	22.667	0.314
50	8	3	27.114	0.048	17.044	0.125
50	8	4	34.639	0.020	22.115	0.032
50	8	5	34.597	0.008	18.822	0.062
50	10	2	26.275	0.471	26.275	0.471
50	10	3	36.388	0.038	31.943	0.129
50	10	4	45.687	0.110	40.418	0.131
50	10	5	41.681	0.163	47.921	0.074
60	6	2	13.005	0.219	9.945	0.273
60	6	3	4.996	0.083	9.839	0.178
60	6	4	6.589	0.043	6.615	0.069
60	6	5	8.198	0.021	8.192	0.015
60	9	2	22.022	0.219	19.180	0.328
60	9	3	23.010	0.055	23.024	0.069
60	9	4	24.360	0.094	29.660	0.090
60	9	5	23.022	0.096	35.757	0.042
60	12	2	30.546	0.219	27.814	0.273
60	12	3	34.932	0.058	31.090	0.042
60	12	4	43.833	0.012	34.690	0.099
60	12	5	–	–	–	–
Average			22.826	0.159	24.753	0.148

10. Solution analysis

In this section we provide some insights gleaned from the analysis of the solutions of the S-RIMF. How the solution time is affected by the parameters P, Q and R is the topic of the first subsection. The next subsection explores the benefits of optimizing the stochastic problem as compared to solving a deterministic problem with some number of interdictions, and investigates the relation between solutions sensitivity and probability function used.

Exponential regression: By applying exponential regression to the results presented in Tables 1 and 2 it is possible to determine how each parameter affects the computation time. For the algorithms RS-MCP, Heur1 and Heur2 we calculated the parameters of the exponential regression function:

$$y = \alpha \cdot P^{\beta} \cdot Q^{\gamma} \cdot R^{\delta}$$

where y represents the solution time. Table 5 presents the values of the parameters and the correlation coefficient R^2 (Note that here R^2 is the coefficient of determination commonly used in statistical models and not the square of the number of interdictions.) The standard errors for the estimated values are reported in parentheses (Note: because of the exponential nature of the regression function used, the standard error associated to the coefficient α should be compared to $\ln\alpha$ and not α). All the functions have a very high correlation coefficient, higher than 0.80. The exponents β, γ, and δ can be used to identify the parameters that have more impact on the solution time. The computational time is mostly influenced by the number of facilities P; in fact the β coefficients vary between 6.94 and 7.41. Furthermore, the p-value for β in each regression is less than 0.001, indicating that the estimated exponents are significantly different from 0. On the other hand, the number of fortifications Q affects the solution time the least as the associated coefficient, γ,

Table 5
Coefficients of exponential regression.

Algorithm	Instance	α	β	γ	δ	R^2
RS-MCP	London	5.32 E-11 (5.45)	7.10 (0.77)	1.20 (1.52)	3.27 (0.65)	0.81
	USCities	1.10 E-11 (4.08)	6.94 (0.49)	1.46 (0.58)	3.68 (1.14)	0.88
Heur1	London	6.01 E-10 (4.70)	7.24 (0.56)	0.54 (0.67)	2.88 (1.31)	0.85
	USCities	5.00 E-08 (3.59)	7.41 (0.43)	1.05 (0.51)	1.44 (1.00)	0.91
Heur2	London	3.83 E-09 (4.74)	7.38 (0.56)	0.53 (0.67)	2.37 (1.32)	0.85
	USCities	1.07 E-07 (3.67)	7.21 (0.44)	1.13 (0.53)	1.28 (1.02)	0.90

Table 6
Cross-comparison of RIMF and S-RIMF optimal protection plans for the instance of the data set USCities with parameters $P=60$, $Q=6$, and $R=5$. Objective function values.

Used for protection planning	Actual number of losses					Fortification set
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	
$r=1$	1,641,449,239	2,287,026,236	3,652,251,174	3,920,371,221	4,238,742,456	1 3 4 6 8 29
$r=2$	1,689,643,011	1,988,939,053	2,537,232,890	2,918,795,003	3,266,213,753	1 3 4 6 8 52
$r=3$	1,730,495,481	2,025,328,338	2,316,986,763	2,918,795,003	3,246,842,713	1 2 6 8 19 56
$r=4$	1,906,782,446	2,234,830,156	2,522,025,396	2,821,321,438	3,272,179,361	1 2 3 6 14 37
$r=5$	1,730,495,481	2,017,690,721	2,537,232,890	2,865,280,600	3,152,475,840	1 3 6 8 40 52
Increasing probability	1,730,495,481	2,017,690,721	2,537,232,890	2,865,280,600	3,152,475,840	1 2 3 6 8 40
Decreasing probability	1,730,495,481	2,017,690,721	2,316,986,763	2,918,795,003	3,246,842,713	1 2 3 6 8 14
Uniform probability	1,730,495,481	2,017,690,721	2,316,986,763	2,918,795,003	3,246,842,713	1 2 3 6 8 56

Table 7
Cross-comparison of RIMF and S-RIMF optimal protection plans for the instance of the data set USCities with parameters $P=60$, $Q=6$, and $R=5$. Relative increase in percentage (%).

Used for protection planning	Actual number of losses					MAX	AVERAGE
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$		
$r=1$	0.00	14.99	57.63	38.96	34.46	57.63	29.21
$r=2$	2.94	0.00	9.51	3.45	3.61	9.51	3.90
$r=3$	5.42	1.83	0.00	3.45	2.99	5.42	2.74
$r=4$	16.16	12.36	8.85	0.00	3.80	16.16	8.23
$r=5$	5.42	1.45	9.51	1.56	0.00	9.51	3.59
Increasing probability	5.42	1.45	9.51	1.56	0.00	9.51	3.59
Decreasing probability	5.42	1.45	0.00	3.45	2.99	5.42	2.66
Uniform probability	5.42	1.45	0.00	3.45	2.99	5.42	2.66

takes relatively small values. The maximum number of interdictions R has a considerable impact on the solution time of the exact method (δ varies between 3 and 4) while its impact on the heuristic time is somewhat lower.

Solution quality and robustness: In this subsection we investigate the importance of considering uncertainty in the number of attacks as compared to using the solution of the deterministic MCP when the behavior of the attacker is unknown, which is generally the case.

The analysis consists of the following steps:

- For any given instance, we find the optimal protection plan for the S-RIMF with each of the probability functions, and for the deterministic MCP problem for different values of r, namely $r=1,\ldots,R$.
- We then evaluate the cost of each protection plan for a specific realization of the parameter r; i.e., the cost of the plan when a specific number of losses occurs in practice. Table 6 presents an example for the instance of the data set USCities with parameters $P=60$, $Q=6$, and $R=5$. The first column represents the protection plan used. The subsequent columns show the solution value for each protection plan when the actual number of losses is the one specified in the header of the column (scenarios). The last column shows the optimal fortification set for each program solved. Note that from the data in this last column we can infer that these fortification/interdiction problems often have multiple optimal solutions. In fact, despite the fact that the fortification sets of MCP with $r=5$ and S-RIMF with the increasing probability function differ by one facility, the solution costs are the same across all the scenarios. The same is true for S-RIMF with the decreasing and uniform probability functions.
- To compare the performances of the solutions, for each protection plan and each value of r, we calculate the percentage cost increase from the optimal solution. Table 7 contains the percentage cost increases relative to Table 6. The last two columns illustrate, respectively, the maximum and the average percentage cost increase. In the example, the analysis suggests that the best protection strategies for this problem are obtained with the optimal solution to the S-RIMF using the decreasing and uniform probability functions. The S-RIMF with the increasing probability function performs well too. On the other hand, when using the deterministic model it is difficult to have a clear idea of the relation between solution quality and the number of interdictions. In fact, the plan computed with $r=3$ works fairly well, followed by the plans obtained with $r=2$ or 5. However, the plan with $r=4$ yields very high percentage cost increases, with a maximum percentage cost increase as high as 16%. The stochastic model, instead, seems to have a more robust behavior.

Table 8
Average results for cross-comparison analysis for the data set USCities and London. Relative increase in percentage (%).

	USCities		London	
	MAX	AVERAGE	MAX	AVERAGE
$r=1$	45.16	25.27	9.60	6.21
$r=2$	16.19	8.11	2.61	1.17
$r=3$	7.80	3.43	2.64	1.15
$r=4$	7.92	3.53	1.28	0.59
$r=5$	8.78	3.53	1.42	0.71
Increasing probability	7.15	2.92	1.18	0.50
Decreasing probability	5.64	2.73	1.24	0.51
Uniform probability	6.58	2.70	1.09	0.48

We performed the analysis for the instances solved in Section 9 with $R=5$ (e.g., 16 instances overall). The results are presented in Table 8. The analysis shows that, for the instances considered, the solutions to the S-RIMF problem provide the best results in the face of an uncertain number of attacks. In fact, the S-RIMF solutions are always better than the deterministic ones in terms of both the maximum and average percentage cost increase. We recognize that the differences are quite small, especially for the cases with a high number of interdicted facilities; i.e., $r \geq 3$. Nevertheless, we believe that even minimal improvements can justify the solution of a more complex and sophisticated model considering that these fortification problems are strategic in nature and that small percentage improvements may result in substantial cost savings.

Another important considerations is that there is no clear dominance among the stochastic models with different probabilities. Besides, the solutions found by using the three probability distributions have similar cost increases. These results suggest that, for the instances considered, the solutions are robust to misestimations of the probability functions. In fact, the choice of the probabilities does not seem to have a massive influence on the final cost of the system as long as all the possible outcomes are taken into considerations; i.e., a sufficiently large probability p_r is assigned to every value of the parameter r.

11. Conclusions

Building upon the preliminary research of Scaparra and Church [8], our study took a step forward in the development of the interdiction problems with fortification. First, we extended the r-interdiction median problem with fortification to the stochastic case using a max-covering formulation that requires neither precedence constraints nor ordering of the interdiction patterns as required by Scaparra and Church [21]. Second, we developed bounds that exploit the stochastic nature of the problem to reduce the dimensionality of the model. The resulting reduced formulation was extensively tested and the experiments demonstrated that the bounds found are extremely effective: they drastically reduced the dimension of the solution space and the computational times, and allowed us to solve instances of realistic size within 1 h. Third, we proposed three rules that can be used to develop heuristics for the problem and derived two related algorithms. The heuristic algorithms reduced the solution time by approximately 40% and were able to find the optimal solution in all but one instance, with a solution gap equal to only 0.137%. The major contribution of the heuristic rules lies in their versatility. In fact, they can be applied to any stochastic fortification/interdiction problem that cannot be efficiently reformulated as a single-level program (e.g., because of the complexity of the underlying model), or for which no effective solution methodology (such as the bound based approach presented in this paper) can be devised. The last part of this paper analyzed the stochastic nature of problem. This analysis shows that, for the instances considered, taking into account the uncertainty in the number of attacks in the optimization process leads to good solution regardless of the real number of interdictions that take place. Although the improvement from the solutions of the deterministic model is sometimes small (ranging between less than 1% in the worst case and 40% in the best case), it is still advisable to use the stochastic model rather than the deterministic one as its solutions always perform better. Finally, the solutions obtained seem to be quite robust to possible misestimations of the probability distributions.

Since interdiction problems with fortification are a very recent field of research, there are plenty of research opportunities to be pursued. An interesting variation, that is currently being explored, is to minimize the amount of resources used to protect the facilities while keeping the impact of the disruptions on the performances of the system under a given percentage from the optimal state. In addition, we plan to explore objective functions other than the expected cost objective analyzed in this paper, including the conditional value at risk objective studied in [7]. We hope that this work will be a useful source of ideas for future research on stochastic problems and will contribute further in the development and solution of more complex and more realistic models for fortification and interdiction problems.

Acknowledgements

This research was supported by EPSRC Grant EP/E048552/1. The research of the third author was also supported by a National Science Foundation Grant (DMI-0457503). All the supports are gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Census 2000 gateway. 〈http://www.census.gov/main/www/cen2000.html〉.
[2] Azaiez MN, Bier VM. Optimal resource allocation for security in reliability systems. European Journal of Operational Research 2007;181:773–86.
[3] Bell MGH. The use of game theory to measure the vulnerability of stochastic networks. IEEE Transactions on Reliability 2003;52:63–8.
[4] Berman O, Krass D, Menezes MBC. Facility reliability issues in network p-median problems: strategic centralization and co-location effects. Operations Research 2007;55(2):332–50.
[5] Bier VM. Choosing what to protect. Rick Analysis 2007;27:606–20.
[6] Brown G, Carlyle M, Salmerón J, Wood K. Analyzing the vulnerability of critical infrastructure to attack and planning defenses. In: INFORMS tutorials in operations research, 2005. p. 102–23.
[7] Chen G, Daskin MS, Shen ZJM, Uryasev S. The $\alpha-$reliable mean-excess regret model for stochastic facility location modeling. Naval Research Logistics 2006;53:617–26.
[8] Church RL, Scaparra MP. Protecting critical assets: the r-interdiction median problem with fortification. Geographical Analysis 2006;39:129–46.
[9] Church RL, Scaparra MP, Middleton RS. Identifying critical infrastructure: the median and covering facility interdiction problem. Annals of the Association of American Geographers 2004;94:491–502.
[10] Goodchild MF, Noronha VT. Location-allocation for small computers. Operational Geographer 1985;7:10.
[11] Grubesic TH, Murray AT. Vital nodes, interconnected infrastructures, and the geographies of network survivability. Annals of the Association of American Geographers 2006;96:64–83.
[12] Higle JL. Stochastic programming: optimization when uncertainty matters. In: INFORMS tutorials in operations research, 2005.
[13] ILOG Inc. Cplex version 11.1. Mountain View, CA, 2008.
[14] Lim C, Smith JC. Algorithms for discrete and continuous multicommodity flow network interdiction problems. IIE Transactions 2007;39:15–26.
[15] Lim M, Daskin M, Chopra S, Bassamboo A. Managing risks of facility disruptions. Northwestern University Working Paper, 2008.
[16] O'Hanley JR, Church RL. Designing robust coverage networks to hedge against worst-case facility losses. Kent Business School Working Paper no. 173, 2008.
[17] Qiao J, Jeong D, Lawley M, Richard JP, Abraham DM, Yih Y. Allocating security resources to a water supply network. IIE Transactions 2007;39:95–109.
[18] Rosing KE, John Hodgson M. Heuristic concentration for the p-median: an example demonstrating how and why it works. Computers & Operations Research 2002;29:1317–30.
[19] Rosing KE, Re Velle CS. Heuristic concentration: two stage solution construction. European Journal of Operational Research 1997;97:75–86.
[20] Scaparra MP, Church RL. A bilevel mixed-integer program for critical infrastructure protection planning. Computers & Operations Research 2008;35:1905–23.
[21] Scaparra MP, Church RL. An exact solution approach for the interdiction median problem with fortification. European Journal of Operational Research 2008;189:76–92.
[22] Smith JC, Lim C, Sudargho F. Survivable network design under optimal and heuristic interdiction scenarios. Journal of Global Optimization 2007;38:181–99.
[23] Snyder LV, Daskin MS. Reliability models for facility location: the expected failure cost case. Transportation Science 2005;39(3):400–16.
[24] Wollmer R. Removing arcs from a network. Operations Research 1964;12:934–40.
[25] Zhuang J, Bier VM. Balancing terrorism and natural disasters—defensive strategy with endogenous attacker effort. Operations Research 2007;55(5):976–91.