ISSN 1748-7595 (Online)

Kent Business School

Working Paper Series

Analysis of Facility Protection Strategies Against Uncertain Numbers of Attacks: The Stochastic R-Interdiction Median Problem with Fortification

Maria Paola Scaparra
Kent Business School

Federico Liberatore
Kent Business School

Mark Daskin
Northwestern University

Working Paper No.176
December 2008

University of Kent

Analysis of facility protection strategies against uncertain numbers of attacks: the Stochastic R-Interdiction Median Problem with Fortification

Federico Liberatore[*], Maria P. Scaparra[†] and Mark S. Daskin[‡]

September 3, 2008

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 problem instances considered. Extensive computational testing shows that both the optimal algorithm and the heuristics are very successful at solving the problem. A comparison of the results obtained by the two methods is provided as is a discussion of the importance of recognizing the stochastic nature of the number of possible attacks.

Keywords. Logistics, protection planning, combinatorial optimization, stochastic modeling.

1 Introduction

Today more than ever the protection of infrastructure has became 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 case of intentional attacks, but also when natural catastrophes occur.

[*]Kent Business School, University of Kent, CT2 7PE Canterbury, Kent, UK. fl51@kent.ac.uk

[†]Kent Business School, University of Kent, CT2 7PE Canterbury, Kent, UK. m.p.scaparra@kent.ac.uk

[‡]Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA. m-daskin@northwestern.edu

A variety of quantitative approaches have recently been developed to identify cost-effective ways of increasing the 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 [15] 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 [22] 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 at al. [14] 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 [23] 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 [13] 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. Since interdiction problems identify critical facilities they can be used to develop design protection strategies. Smith et al. exploit this idea in [21] where they extend their previous work [13] 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 for explicitly modeling protection efforts has been acknowledged in several recent works such as [6], [16] and [8]. Most studies in this area use a game theoretic approach and formulate protection problems as bilevel defender-attacker models. Brown et al. [6] provide an excellent introduction to bilevel and trilevel 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. [16] 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. 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 [20] is based on a reformulation of the problem as a maximal covering model with precedence constraints. The dimension of the model is reduced using a linear interpolation search procedure that exploits properties of the coverage function. The second approach [19] is a tree search algorithm that takes advantage of the bi-level formulation of the problem. Zhuang and Bier [24] 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. Finally, 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.

In this paper we present the stochastic R-interdiction median problem with fortification, an extension of the r-interdiction median problem with fortification were the number of possible losses or attacks is uncertain. We formulate the problem and, as the dimension of the problem grows very quickly with respect to the parameters, we propose some reduction methodologies based on the computation of upper and lower bounds. The proposed reductions allow us to solve to optimality problem instances of significant size. We also introduce three heuristic-concentration type rules and, based on them, two heuristic algorithms that are able to find the optimal solution for almost all the problems solved.

The reminder of the paper is organized as follows. In the next section we formulate the stochastic R-interdiction median problem with fortification as a bi-level integer program. In Section 3 we present the deterministic equivalent reformulation as a max-covering problem. The definition of the lower and the upper bounds to the problem are the topic of Sections 4 and 5 respectively. The resulting reduced model is shown in Section 6. The heuristic-concentration type rules are explained

in Section 7. The algorithms proposed have been tested on some geographical data sets and the results are displayed and discussed in Section 8. An extensive evaluation of the solutions obtained and of the importance of utilizing the stochastic R-interdiction median problem with fortification is the topic of Section 9. Finally, the paper concludes with a brief summary of the main contributions of this work and some ideas for future research.

2 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. 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:

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

Finally, 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:

$$T_{ij} = \{k \in F \,|k \neq j \text{ and } d_{ik} > d_{ij}\}, \; \forall i \in N, \; \forall j \in F.$$

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:

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

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

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

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

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

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

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

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

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

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

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

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 upper level program is to minimize the expected cost of the worst-case interdictions across all the values that r can have.

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 [20]. 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 [20] 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 [20] 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.

3 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 showed in Algorithm 1. Note that:

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

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
    $\bar{j} = \text{argmin}_{j \in F/I_h}\{d_{ij}\}$ //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 [20] 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 inderdiction pattern when exactly r facilities are interdicted.

The MCP is formulated as follows:

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

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

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

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

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

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

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 relative 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 $r = 1, \ldots, R$ at the same time and take into account the probabilities p_r, $\forall r = 1, \ldots, R$ in the objective function. The resulting formulation, called S-MCP, is:

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

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

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

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

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

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

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, the program will find the optimal fortification set $\mathbf{z}^\star$ that minimizes this expected cost.

The model has $R + P + \sum_{r=1}^{R} |H_r| = R + P + \sum_{r=1}^{R} \begin{pmatrix} P \\ r \end{pmatrix}$ variables and $1 + 2\sum_{r=1}^{R} |H_r| = 1 + 2\sum_{r=1}^{R} \begin{pmatrix} P \\ r \end{pmatrix}$ 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, in particular, of 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.

4 A Lower Bound

In the previous section we showed that the dimension of the stochastic model grows very quickly. By calculating a lower bound to 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, $\overline{\mathbf{z}}^r$, for a given 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 to 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^r 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:

$$\overline{H}_r = H_r \setminus \left\{ h \,\middle|\, c_h^r < \overline{W}_r \right\},\ r = 1, \ldots, R.$$

5 An Upper Bound

We can take advantage of the solutions obtained during the calculation of the lower bound (Section 4) 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 $\overline{z}^r$ and objective function $Z_m^{RIM}(\overline{z}^r)$. Recall that the RIM for a particular value of m (the number of inderdicted sites) will give the optimal interdictions of m sites given that the sites in fortification pattern $\overline{z}^r$ 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 $\overline{\mathbf{z}}^r$:

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

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

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

$\widetilde{Z}$ 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 $W_r^\star$; 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:

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

Hence:

$$W_r^\star \leq \frac{\widetilde{Z} - \sum_{\substack{m=1,\\ m\neq r}}^{R} p_m \overline{W}_m}{p_r}.$$

Therefore:

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

is an upper bound for $W_r^\star$.

We can calculate the upper bound of the relative worst-case interdiction pattern $\widetilde{W}_r$ 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:

$$\widetilde{H}_r = \left\{ h \,\middle|\, c_h > \widetilde{W}_r \right\},\ r = 1, \ldots, R.$$

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

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

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

6 The Reduced Stochastic Max-Covering Model

The resulting reduced model RS-MCP is:

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

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

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

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

$$W_r \geq c_h \left(1 - y_h\right)\ \forall h \in \overline{H}_r \setminus \widetilde{H}_r,\ r = 1, \ldots, R \tag{28}$$

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

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

Computational experiments (see Section 8) proved that, thanks to the pre-processing on the interdiction patterns and the bounds, the solution time can be drastically reduced.

7 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 and competitive heuristics for the S-MCP. The counterexamples that show that the rules do not preserve the optimality are discussed in detail in Section 8.

All the rules are based on an idea similar to *heuristic concentration* [18]. As Rosing and Hodgson illustrate in [17] 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 interdicitions; i.e. $F^r = \left\{ j \,|\bar{z}_j^r = 1 \right\}$. 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

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

and by setting

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

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

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

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

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

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

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

and fixing the remaining variables

$$z_j = 0 \; \forall j \in F / \left\{ F^{AS} \cup F^{RS} \right\}.$$

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 4 and in Section 5 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 can not be less than the value found using any of the rules.

8 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 optimize 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 three different data sets: London, randompoints3 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 United States according to the 2000 census [1]. The third data set - randompoints3 - has been generated specifically to find counterexamples to the heuristic rules. 250 points were generated with integer-valued X and Y coordinates uniformly distributed between 0 and 249 (inclusive). Care was taken to ensure that no two points shared the same coordinates. Demands were also integer valued and were uniformly distributed between 100 and 105 inclusive.

The tests have been run over a wide number of combination 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 hour and a physical memory limit of 1 GB.

Two different probability distributions p_r have been employed:

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

and

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

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 behaviour of terrorist attacks that generally tend to be focused on a small number of targets. Since both 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:

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

The priority coefficients ϕ_j are subsequently provided to CPLEX.

Bounds and rules: what we achieved. Tables 1, 2 and 3 display the value of the optimal solution and the computation time of the algorithms for the data sets London, randompoints3 and USCities respectively. The first three columns are the parameters P, Q and R. The fourth column contains the objective value $Z^\star$ 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 suboptimal solution. A dash indicates that the algorithm has not been able to complete the optimization because the instance exceeded either the time or the memory limit. In the last row are shown the average solution times for RS-MCP and the heuristic algorithms.

The most noticeable observation is that S-MCP was able to solve within one hour 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 the 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.

Heur1 and Heur2 could not find the optimal solution only in three cases. Table 4 shows detailed information about these instances. The first four columns provide the data sets and the parameters P, Q and R. The last three columns are respectively the optimal solution value $Z^\star$, the solution value found by the heuristics and the corresponding gap. The heuristic algorithms found the same solution in all three instances. The gap expresses in percentage terms how far off the heuristic objective function value is from the optimum, and is calculated as follows:

$$\text{GAP} = 100 \cdot \frac{\text{Heur} - Z^\star}{Z^\star}.$$

The results shown in Table 4 confirm the effectiveness of the heuristic rules illustrated: in fact, although the algorithms could not reach the optimum in three instances, they found solutions close to optimality as the gap is very small in all cases. Furthermore, by comparing the average solution times it can be noticed that both the heuristics improved the average solution time when compared to RS-MCP, and Heur2 is slightly slower than Heur1. 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. At any rate, both the algorithms proved to be useful since on average they provided significant dimension and computation time reductions, especially on the data set USCities. Furthermore, the reductions are more significant for higher values of P. (Although the solution times are calculated considering

the average time over five executions, they can still be biased by memory swapping time, time to access the physical memory and other details related to the operating system. Thus differences of tenths of a second are completely unimportant. For the sake of correctness we decided to leave the solution times unaltered anyway. This explains why for some instances Heur2 seems to outperform 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 5 shows 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 earch data set. The last row of the table shows 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. The trend of the reduction of patterns when only the bounds are used is less predictable. In the data set London and USCities the percentage of remaining patterns is generally higher for higher P and R values, whereas in the data set randompoints3 there is no evident trend. For this data set, the reduction provided by the bounds is very effective in the instances with parameters $P = 50\, R = 5\, Q = 5$ and $P = 50\, R = 8\, Q = 5$. In these two cases, only 0.006% and 0.028% of the total number of patterns can not be fixed by using the bounds. Using the heuristic rule Always Fortified does not give any further contribution.

9 Solution Analysis

In this section we provide some insights gleaned from the analysis of the solutions of the S-RIMF. Some considerations of how the solution time is affected by the parameters P, Q and R is the topic of the first subsection. The next subsection presents the counterexample that proved that the heuristic rules introduced in Section 7 do not preserve the optimality of the solutions. The third subsection explores the benefits of optimizing the stochastic problem as compared to solving a deterministic problem with a number of interdictions r equal to the expected value of the number used to model the behavior of the attacker. Finally, the section concludes with an analysis of how sensitive the optimal solutions are to accurate estimations of the probabilities and provides a graphical comparison of the solutions to one problem instance.

Exponential regression. By applying exponential regression to the results shown in Table 1, 2 and 3 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 6 shows the values of the parameters and the correlation coefficient R^2. 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 8.56. 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, takes relatively small values. The 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.

Counterexamples to the heuristic rules. In this subsection we present the computational tests that disproved the optimality of the conjectures that led to the heuristic rules introduced in Section 7. The counter-example to the rule *Always Fortified* has been found in the solution of an instance of the data set USCities with parameters $P = 40$, $Q = 6$, $R = 2$ and the increasing probability distribution. The optimal fortification sets F^r and the corresponding Always Fortified set are showed in Table 7. The Always Fortified set contains the facilities 1, 3, 6, 23 and 154. As showed in Table 8, the optimal fortification set corresponding to the instance does not include facility 23. In fact, the best solution found after imposing the fortification of the facilities of F^{AF} correspond to F^2, and is a suboptimal solution. The difference in percentage between the values of the two solutions is 0.137%.

Table 9 displays the optimal fortification sets F^r and the associated worst-case interdiction sets I^r corresponding to the solutions of the MCP on the data set randompoints3 with parameters $P = 30$, $Q = 4$ and $r = 1, \ldots, 4$. Table 10 shows the fortification sets and the solution values obtained solving S-MCP on the same instance, using the monotonically decreasing probability distribution and parameter $R = 4$. The two columns are, respectively, the results related to the S-MCP and a heuristic algorithm that employs only the heuristic rule *Action Set*. By comparing the solution values it can be easily seen that the solution found with the heuristic concentration approach is suboptimal. In fact, facility 115 is contained only in the optimal fortification set but not in the Action Set. Interestingly this facility is in the interdiction set for $r = 4$.

Tables 11 and 12 show the results for instance randompoints3 with parameters $P = 30$, $Q = 3$, $R = 3$ and the monotonically decreasing probability distribution. This provides a counterexample to the third heuristic rule: *Reaction Set*. In this instance, facility 115 is in the optimal fortification but is not present in either the Action Set or in the Reaction Set.

How important is to model uncertainty in the number of attacks? In the S-MCP the impact of the attacks is evaluated using the expected cost objective function (18). In this subsection we investigate the importance, in terms of solution cost, of solving the stochastic problem compared to using the optimal fortification set of the deterministic problem where the number of attacks is exactly the expected number of attacks of the stochastic case. When $R = 4$, the expected value of the number of attacks is: $E(R) = 3$ for the increasing probability case, and $E(R) = 2$ for the decreasing

probability case. It is possible to solve the r-Interdiction Median Problem with Fortification with a number of attacks r equal to the expected number of interdictions, and subsequently to evaluate the optimal fortification set $F_{E(R)}$ by using the stochastic objective function. The resulting solution value is $\widetilde{Z}_{E(R)}$. Tables 13 and 14 show a comparison between the optimal stochastic fortification set $F^\star$ and the corresponding fortification set $F_{E(R)}$ using the USCities data set and, respectively, the increasing probability and the decreasing probability distributions. The first three columns show the parameters P, Q and R used in the instances. The next columns present: the optimal stochastic objective function value $Z^\star$, the optimal stochastic fortification set $F^\star$, the stochastic objective function value $\widetilde{Z}_{E(R)}$ and the corresponding fortification set $F_{E(R)}$. To highlight the differences between the two fortification sets, the sites fortified in the expected value solution, $F_{E(R)}$, that do not appear in the optimal fortification set, $F^\star$, are shown in *italics*. The last column displays the gap between $Z^\star$ and $\widetilde{Z}_{E(R)}$, and is calculated as follows:

$$\text{GAP} = \frac{\widetilde{Z}_{E(R)} - Z^\star}{Z^\star}.$$

The last two rows show the average and maximum gap. The results suggest that modeling the uncertainty in the number of attacks is rather important. The observed gaps were as high as 13.86% when increasing probabilities were used, and as high as 7.44% when decreasing probabilities were used.

How sensitive is the solution to the accurate estimation of the attack probabilities? In this work, we used two probability distributions (monotonically increasing (31) and monotonically decreasing (32) in the number of attacks) to model two antithetical offensive behaviors. We now want to analyze the possible impact on the solution costs of misestimating the attack probabilities p_r. To this end, we considered the fortification sets obtained by optimizing the S-MCP for a certain probability function. We then calculated how the fortification sets would perform if they were used for an actual protection plan with a different distribution. Table 15 displays the results of the tests done on the data set USCities. The tests have been run on the same combination of parameters explained in the first subsection of Section 8. The columns determine the probability distribution used for the optimization, while the rows define which probability may take place in the "real world". Every cell shows two values: the first one is the average solution cost in percentage, and the second one is the maximum percentage solution cost found. These results suggest that, for the instances considered, the probabilities do not seem to have a great influence on the final cost of the network after the attacks take place. In fact, the highest percentage cost is obtained when a protection plan obtained through the use of a decreasing probability is used to counter attacks in an increasing probability environment. This maximum cost is only 105.14% of the optimal solution. Also the averages are very close to the optimality: 100.60% when optimizing using the decreasing probability, and 100.38% when optimizing using the increasing probability.

A visual insight. Figures 1 and 2 display the graphical representations of the fortification sets and the interdiction sets associated to the solutions to the instance of the data set USCities with parameters $P = 40$, $Q = 4$ and $R = 3$. Each figure shows a political map of the United States of America. In each map the demands are represented as small diamonds and are connected to the

corresponding facility by a line. In the map are also reported the names of the fortified facilities and, in a box, those of the interdicted ones. Figure 1 contains the solutions associated with the increasing probability function, whereas the solution corresponding to the decreasing probabilities are in Figure 2. The graphs are sorted from the top to bottom by increasing number of interdictions.

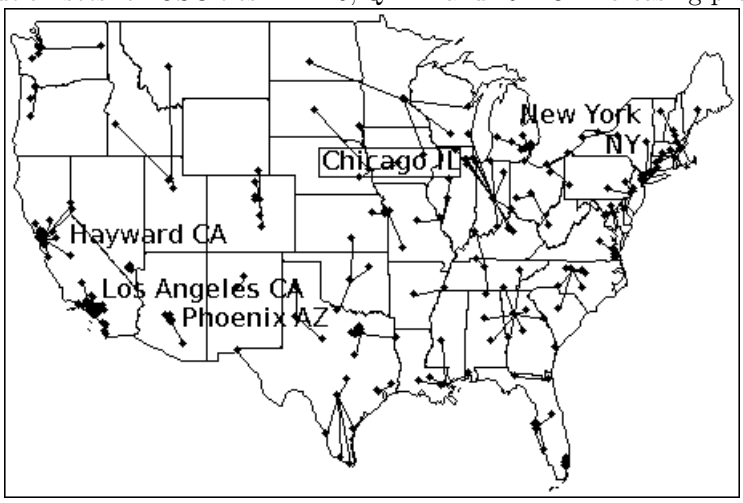
Note that the two protection plans differ by one facility (out of four). When the increasing probabilities are used, New York, NY, Los Angeles, CA, Phoenix, AZ and Hayward, CA, are protected. On the other hand, when the behavior of the attacker is modeled using the decreasing probability function, Los Angeles, CA, is replaced by Chicago, IL.

From observing the interdiction plans, we can infer some interesting considerations. When the decreasing probability function is used, and therefore higher emphasis is given to a small number of attacks, the interdicted facilities tend to be clustered. In fact the assaulted facilities are grouped in the Northwest (Seattle and Portland) when $r = 2$, and Southern California when $r = 3$. Moreover, the optimal interdiction sets for a given r are not subsets of the optimal interdictions for higher values of r: there are no overlaps between the interdiction sets. When the increasing probabilities are used, and thus higher emphasis is given to the scenario with $r = 3$, the fortification set includes Los Angeles, CA, instead of Chicago, IL. Los Angeles, CA, was one of the three cities in California interdicted when $r = 3$ with decreasing probabilities. In this way, the protection plan counters the most probable worst-case attack. As Chicago, IL, is no longer protected, it becomes the most critical facility for $r = 1$. It is also in the interdiction set for $r = 2$ with Dallas (TX). When $r = 3$ the most vulnerable area is Texas (Dallas, San Antonio and Huston). When the increasing probability function is used, the interdiction patterns for different r values overlap.

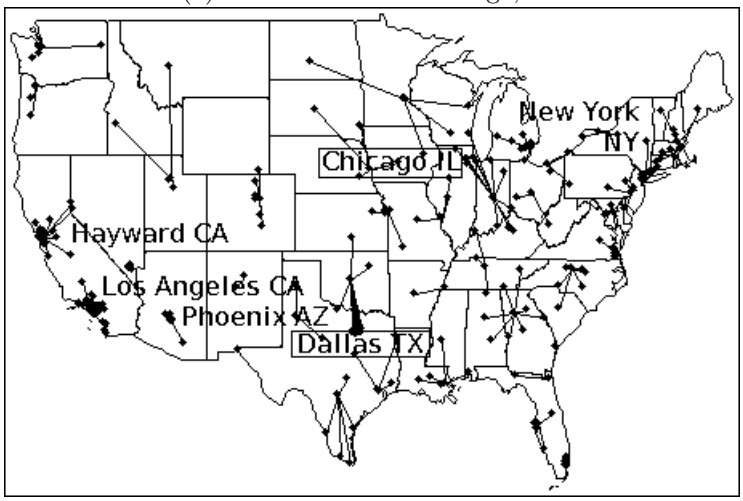
10 Conclusions

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 [20]. 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 one hour. Third, we proposed three rules that can be used to develop heuristics for the problem and derived two related algorithms. By finding counterexamples we showed that none of these rules preserves the optimality of the solutions. Nonetheless, in our tests both heuristic algorithms are faster than the optimal algorithm - especially for bigger instances - and find the optimal solution about 97% of the time. Furthermore, these rules can be applied to any stochastic problem with one stochastic variable that can take a finite number of discrete values. The last part of this paper analyzed the stochastic nature of problem. This analysis shows that the impact of taking into account the uncertainty in the number of attacks in the optimization process may be substantial. Despite that, the solutions obtained are not very sensitive to specific the probability distribution used.

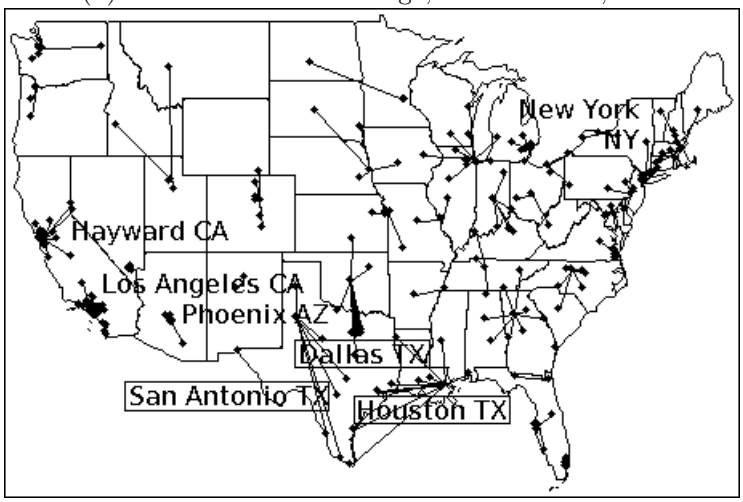
Figure 1: Solution sets for USCities $P = 40$, $Q = 4$ and $R = 3$. Increasing probability case.



(a) 1 interdiction in Chicago, IL.

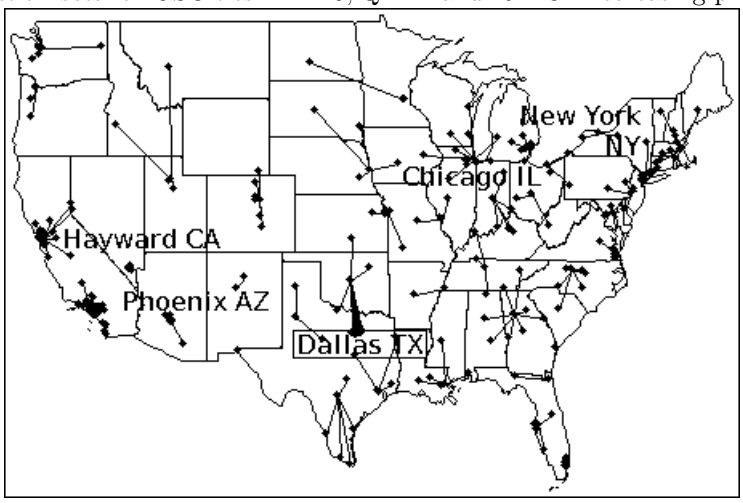


(b) 2 interdictions in Chicago, IL and Dallas, TX.

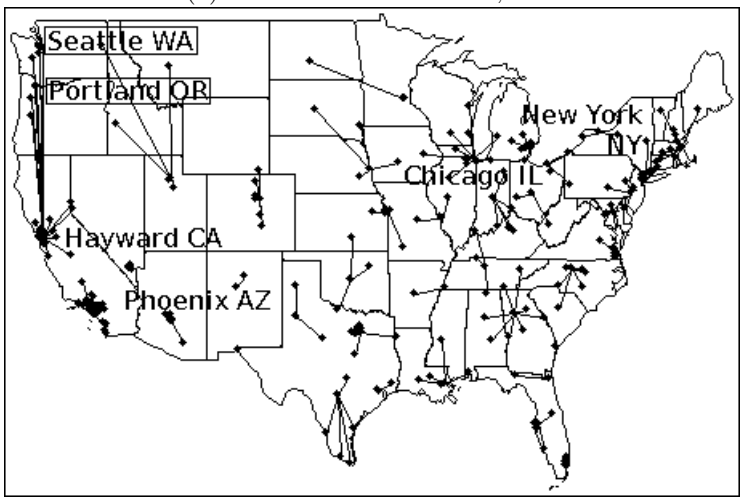


(c) 3 interdictions in Dallas, TX, San Antonio, TX and Houston, TX.

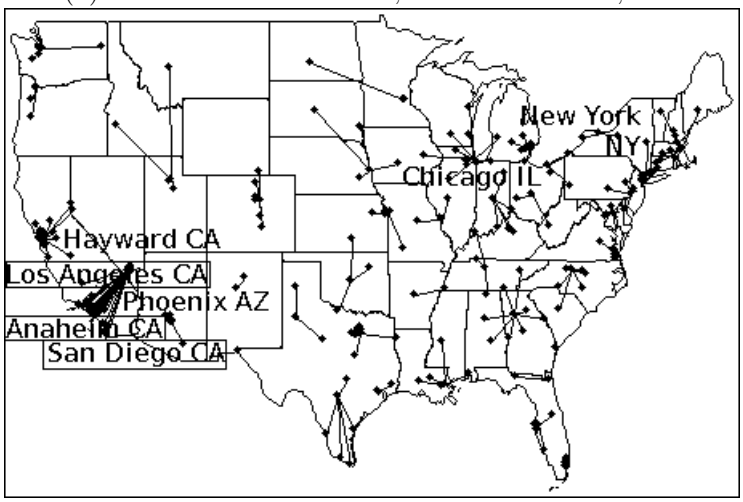
Figure 2: Solution sets for USCities $P = 40$, $Q = 4$ and $R = 3$. Decreasing probability case.



(a) 1 interdiction in Dallas, TX.



(b) 2 interdictions in Seattle, WA and Portland, OR.



(c) 3 interdictions in Los Angeles, CA, San Diego, CA, and Anaheim, CA.

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 attacks 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 models for fortification and interdiction problems.

Acknowledgment

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] M.N. Azaiez and V. M. Bier. Optimal resource allocation for security in reliability systems. 181:773–786, 2007.

[3] M.G.H. Bell. The use of game theory to measure the vulnerability of stochastic networks. *IEEE Transactions on Reliability*, 52:63–68, 2003.

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

[5] V.M. Bier. Choosing what to protect. *Rick Analysis*, 27:606–620, 2007.

[6] G. Brown, M. Carlyle, J. Salmerón, and K. Wood. Analyzing the vulnerability of critical infrastructure to attack , and planning defenses. *INFORMS Tutorials in Operations Research*, pages 102–123, 2005.

[7] G. Chen, M.S. Daskin, Z.J.M. Shen, and S. Uryasev. The α -reliable mean-excess regret model for stochastic facility location modeling. *Naval Research Logistics*, 53:617–626, 2006.

[8] R.L. Church and M.P. Scaparra. Protecting critical assets: the r-interdiction median problem with fortification. *Geographical Analysis*, 39:129–146, 2006.

[9] R.L. Church, M.P. Scaparra, and R.S. Middleton. Identifying critical infrastructure: the median and covering facility interdiction problem. *Annals of the Association of American Geographers*, 94:491–502, 2004.

[10] M.F. Goodchild and V.T. Noronha. Location-allocation for small computers. *Operational Geographer*, 7:10, 1985.

[11] T.H. Grubesic and A.T. Murray. Vital nodes, interconnected infrastructures, and the geographies of network survivability. *Annals of the Association of American Geographers*, 96:64–83, 2006.

[12] J.L. Higle. Stochastic programming: optimization when uncertainty matters. *INFORMS Tutorials in Operations Research*, 2005.

[13] C. Lim and J.C. Smith. Algorithms for discrete and continuous multicommodity flow network interdiction problems. *IIE Transactions*, 39:15–26, 2007.

[14] M. Lim, M. Daskin, S. Chopra, and A. Bassamboo. Managing risks of facility disruptions. Northwestern University Working Paper, 2008.

[15] J.R. O'Hanley and R.L. Church. Designing robust coverage networks to hedge against worst-case facility losses. Kent Business School Working Paper no. 173, 2008.

[16] J. Qiao, D. Jeong, M. Lawley, J. P. Richard, D. M. Abraham, and Y. Yih. Allocating security resources to a water supply network. *IIE Transactions*, 39:95–109, 2007.

[17] K.E. Rosing and M. John Hodgson. Heuristic concentration for the p-median: an example demonstrating how and why it works. *Computers & Operations Research*, 29:1317–1330, 2002.

[18] K.E. Rosing and C.S. Re Velle. Heuristic concentration: two stage solution construction. *European Journal of Operational Research*, 97:75–86, 1997.

[19] M.P. Scaparra and R.L. Church. A bilevel mixed-integer program for critical infrastructure protection planning. *Computers & Operations Research*, 35:1905–1923, 2008.

[20] M.P. Scaparra and R.L. Church. An exact solution approach for the interdiction median problem with fortification. *European Journal of Operational Research*, 189:76–92, 2008.

[21] J.C. Smith, C. Lim, and F. Sudargho. Survivable network design under optimal and heuristic interdiction scenarios. *Journal of global optimization*, 38:181–199, 2007.

[22] L.V. Snyder and M.S. Daskin. Reliability models for facility location: the expected failure cost case. *Transportation Science*, 39(3):400–416, 2005.

[23] R. Wollmer. Removing arcs from a network. *Operations Research*, 12:934–940, 1964.

[24] J. Zhuang and V. M. Bier. Balancing terrorism and natural disasters - defensive strategy with endogenous attacker effort. 55(5):976–991, 2007.

Table 1: Data set London - Algorithms solution time comparison.

				Time (s)			
P	Q	R	$Z^\star$	S-MCP	RS-MCP	Heur1	Heur2
40	4	2	73953.98	0.31	0.03	0.04	0.06
40	4	3	77944.60	115.00	0.13	0.13	0.13
40	4	4	82217.60	-	0.56	0.54	0.59
40	4	5	86396.73	-	4.07	3.89	4.30
40	6	2	73781.74	1.58	0.03	0.03	0.03
40	6	3	77685.54	462.70	0.11	0.11	0.12
40	6	4	81640.34	-	0.55	0.52	0.57
40	6	5	85522.47	-	5.54	5.36	5.97
40	8	2	73401.42	2.70	0.03	0.03	0.03
40	8	3	77076.18	794.40	0.11	0.11	0.12
40	8	4	80963.72	-	1.25	1.19	1.33
40	8	5	84697.62	-	27.55	20.36	23.17
50	5	2	58529.73	1.80	0.03	0.03	0.03
50	5	3	61915.36	1007.00	0.13	0.12	0.12
50	5	4	65155.98	-	1.14	1.10	1.11
50	5	5	68335.60	-	16.85	16.43	16.50
50	8	2	57753.83	3.30	0.04	0.04	0.04
50	8	3	60726.85	2781.00	0.13	0.13	0.13
50	8	4	63757.20	-	1.58	1.50	1.49
50	8	5	66875.69	-	40.36	39.55	39.59
50	10	2	57283.29	8.41	0.05	0.04	0.04
50	10	3	59772.23	3333.00	0.17	0.17	0.17
50	10	4	63065.32	-	2.49	2.12	2.12
50	10	5	66035.96	-	718.12	167.48	170.46
60	6	2	45365.71	4.59	0.03	0.05	0.03
60	6	3	48207.24	-	0.17	0.17	0.17
60	6	4	50797.93	-	2.13	2.09	2.08
60	6	5	53403.87	-	443.84	443.00	443.14
60	9	2	44776.85	14.31	0.44	0.05	0.03
60	9	3	47383.84	-	0.19	0.19	0.18
60	9	4	50014.09	-	4.38	2.37	2.38
60	9	5	52377.06	-	266.42	232.34	232.32
60	12	2	44225.01	27.89	0.40	0.05	0.04
60	12	3	46830.21	-	0.36	0.28	0.28
60	12	4	48902.84	-	7.66	7.00	7.00
60	12	5	-	-	-	-	-
Average			-	-	44.20	27.10	27.31

Table 2: Data set randompoints3 - Algorithms solution time comparison.

				Time (s)			
P	Q	R	$Z^\star$	S-MCP	RS-MCP	Heur1	Heur2
40	4	2	322444.00	0.72	0.04	0.03	0.03
40	4	3	336768.50	181.60	0.14	0.13	0.14
40	4	4	354644.50	-	0.81	0.78	0.81
40	4	5	372024.33	-	12.24	11.99	12.45
40	6	2	320528.00	1.64	0.03	0.04	0.03
40	6	3	335450.67	593.90	0.13	0.11	0.13
40	6	4	351375.60	-	0.79	0.74	0.76
40	6	5	365763.80	-	13.16	12.71	13.18
40	8	2	318973.67	13.69	0.04	0.04	0.03
40	8	3	331940.00	720.20	0.11	0.11	0.11
40	8	4	347074.20	-	0.72	0.64	0.67
40	8	5	364153.93	-	64.28	41.82	43.42
50	5	2	266237.00	2.34	0.03	0.03*	0.03*
50	5	3	278181.33	1435.00	0.16	0.14	0.14
50	5	4	290231.60	-	2.71	2.64	2.64
50	5	5	303348.40	-	56.32	56.33	56.35
50	8	2	265093.00	18.74	0.03	0.03	0.04
50	8	3	275298.17	-	0.16	0.14	0.14
50	8	4	286547.90	-	2.68	2.52	2.52
50	8	5	299296.67	-	74.16	68.56	68.46
50	10	2	263700.33	24.14	0.03	0.03	0.04
50	10	3	274328.33	-	0.19	0.17	0.16
50	10	4	284899.30	-	2.68	2.68	2.64
50	10	5	296662.93	-	233.06	227.22*	227.40*
60	6	2	225900.33	18.55	0.03	0.04	0.04
60	6	3	235483.17	-	0.24	0.24	0.24
60	6	4	243768.20	-	3.06	3.00	3.00
60	6	5	252756.07	-	1323.80	1319.60	1319.00
60	9	2	224423.33	46.01	0.05	0.07	0.04
60	9	3	232261.33	-	0.25	0.25	0.25
60	9	4	241140.40	-	11.00	9.88	9.88
60	9	5	250045.27	-	715.64	619.82	618.94
60	12	2	222201.33	58.42	0.05	0.04	0.05
60	12	3	229879.83	-	0.29	0.25	0.25
60	12	4	237647.20	-	12.73	9.94	9.92
60	12	5	244826.80	-	900.98	788.90	787.70
Average			-	-	95.36	88.38	88.38

Table 3: Data set USCities - Algorithms solution time comparison.

				Time (s)			
P	Q	R	$Z^\star$	S-MCP	RS-MCP	Heur1	Heur2
40	4	2	3383035777.67	0.27	0.03	0.03	0.03
40	4	3	4019125339.00	41.10	0.11	0.09	0.10
40	4	4	4476779495.50	-	0.67	0.59	0.60
40	4	5	4866585540.33	-	4.81	4.41	4.48
40	6	2	3286216658.33	0.39	0.04	0.03*	0.03*
40	6	3	3760914991.00	79.93	0.13	0.11	0.11
40	6	4	4109323708.20	-	0.94	0.83	0.84
40	6	5	4448644420.20	-	14.00	10.22	10.35
40	8	2	3120104252.33	0.48	0.04	0.03	0.04
40	8	3	3450922837.50	116.60	0.14	0.13	0.13
40	8	4	3782176529.30	-	2.12	1.79	1.77
40	8	5	4059997875.00	-	37.75	29.68	29.77
50	5	2	2428424600.33	0.61	0.05	0.03	0.05
50	5	3	2795869185.00	291.10	0.16	0.14	0.16
50	5	4	3063770060.60	-	1.42	1.34	1.35
50	5	5	3365535167.40	-	15.23	14.34	14.36
50	8	2	2286362847.33	0.91	0.05	0.03	0.04
50	8	3	2687197980.83	1471.00	0.19	0.16	0.17
50	8	4	2931244316.40	-	1.95	1.78	1.78
50	8	5	3165552184.60	-	77.89	42.80	42.81
50	10	2	2228271926.00	1.48	0.05	0.03	0.05
50	10	3	2446813824.17	1296.00	0.23	0.20	0.22
50	10	4	2594055200.10	-	3.66	3.38	3.43
50	10	5	2763407634.73	-	321.76	318.46	318.78
60	6	2	1889173705.67	1.20	0.04	0.04	0.03
60	6	3	2119472868.67	507.50	2.05	0.23	0.23
60	6	4	2439201722.40	-	3.15	2.79	2.72
60	6	5	2706738479.53	-	38.84	35.41	35.36
60	9	2	1801763821.00	1.81	0.34	0.03	0.03
60	9	3	1970406925.50	-	1.61	0.27	0.27
60	9	4	2234785151.30	-	21.02	3.54	3.52
60	9	5	2418063692.67	-	498.24	79.44	79.16
60	12	2	1696507683.67	1.97	0.41	0.03	0.04
60	12	3	1841699671.83	-	1.43	0.34	0.34
60	12	4	2007633404.70	-	9.69	8.61	8.60
60	12	5	-	-	-	-	-
Average			-	-	30.29	16.04	16.05

Table 4: Optimal and heuristic solutions GAP.

Data Set	P	Q	R	$Z^\star$	Heur	GAP
randompoints3	50	5	2	266237.00	266261.33	0.009%
randompoints3	50	10	5	296662.93	296676.87	0.005%
USCities	40	6	2	3286216658.33	3290718099.00	0.137%

Table 5: Comparison of the percentage of free interdiction patterns.

			London		USCities		randompoints3	
P	Q	R	RS-MCP	Heur	RS-MCP	Heur	RS-MCP	Heur
40	4	2	10.122%	0.488%	19.024%	0.244%	14.634%	0.366%
40	4	3	14.308%	0.075%	14.364%	0.131%	14.327%	0.093%
40	4	4	18.719%	0.012%	18.731%	0.024%	9.752%	0.035%
40	4	5	23.048%	0.003%	23.056%	0.011%	12.132%	0.006%
40	6	2	10.610%	0.976%	23.537%	0.366%	19.268%	0.488%
40	6	3	14.346%	0.112%	27.121%	0.075%	20.981%	0.168%
40	6	4	18.729%	0.022%	34.686%	0.031%	27.027%	0.020%
40	6	5	23.051%	0.006%	32.905%	0.056%	23.053%	0.009%
40	8	2	23.902%	0.732%	28.049%	0.610%	23.902%	0.732%
40	8	3	21.084%	0.271%	33.196%	0.252%	27.168%	0.121%
40	8	4	27.090%	0.082%	41.775%	0.091%	27.065%	0.058%
40	8	5	32.875%	0.026%	49.538%	0.079%	23.081%	0.036%
50	5	2	11.843%	0.314%	11.843%	0.314%	15.451%	0.235%
50	5	3	11.574%	0.062%	11.583%	0.072%	16.958%	0.038%
50	5	4	15.192%	0.014%	15.186%	0.008%	15.191%	0.013%
50	5	5	18.764%	0.005%	18.760%	0.001%	0.006%	0.006%
50	8	2	22.745%	0.392%	22.667%	0.314%	19.294%	0.471%
50	8	3	27.114%	0.048%	17.044%	0.125%	22.175%	0.072%
50	8	4	34.639%	0.020%	22.115%	0.032%	22.109%	0.026%
50	8	5	34.597%	0.008%	18.822%	0.062%	0.028%	0.028%
50	10	2	26.275%	0.471%	26.275%	0.471%	22.902%	0.549%
50	10	3	36.388%	0.038%	31.943%	0.129%	11.851%	0.340%
50	10	4	45.687%	0.110%	40.418%	0.131%	7.913%	0.090%
50	10	5	41.681%	0.163%	47.921%	0.074%	9.811%	0.042%
60	6	2	13.005%	0.219%	9.945%	0.273%	9.945%	0.273%
60	6	3	4.996%	0.083%	9.839%	0.178%	4.990%	0.078%
60	6	4	6.589%	0.043%	6.615%	0.069%	6.558%	0.012%
60	6	5	8.198%	0.021%	8.192%	0.015%	8.189%	0.012%
60	9	2	22.022%	0.219%	19.180%	0.328%	16.230%	0.383%
60	9	3	23.010%	0.055%	23.024%	0.069%	14.358%	0.108%
60	9	4	24.360%	0.094%	29.660%	0.090%	18.705%	0.037%
60	9	5	23.022%	0.096%	35.757%	0.042%	8.212%	0.035%
60	12	2	30.546%	0.219%	27.814%	0.273%	27.814%	0.273%
60	12	3	34.932%	0.058%	31.090%	0.042%	31.218%	0.169%
60	12	4	43.833%	0.012%	34.690%	0.099%	34.683%	0.093%
60	12	5	-	-	-	-	41.477%	0.039%
Average			22.826%	0.159%	24.753%	0.148%	17.457%	0.154%

Table 6: Coefficients of exponential regression .

Algorithm	Instance	α	β	γ	δ	R^2
	London	5.32 E-11 (5.45)	7.10 (0.77)	1.20 (1.52)	3.27 (0.65)	0.81
RS-MCP	randompoints3	7.84 E-12 (5.22)	8.56 (0.62)	0.51 (0.74)	3.72 (1.47)	0.86
	USCities	1.10 E-11 (4.08)	6.94 (0.49)	1.46 (0.58)	3.68 (1.14)	0.88
	London	6.01 E-10 (4.70)	7.24 (0.56)	0.54 (0.67)	2.88 (1.31)	0.85
Heur1	randompoints3	4.47 E-12 (5.25)	8.41 (0.62)	0.38 (0.74)	3.96 (1.48)	0.86
	USCities	5.00 E-08 (3.59)	7.41 (0.43)	1.05 (0.51)	1.44 (1.00)	0.91
	London	3.83 E-09 (4.74)	7.38 (0.56)	0.53 (0.67)	2.37 (1.32)	0.85
Heur2	randompoints3	4.86 E-12 (5.16)	8.46 (0.61)	0.39 (0.73)	3.92 (1.46)	0.87
	USCities	1.07 E-07 (3.67)	7.21 (0.44)	1.13 (0.53)	1.28 (1.02)	0.90

Table 7: Fortication sets and interdiction sets for MCP, data set USCities, parameters $P = 40$, $Q = 6$ and $r = 1, 2$.

r	F^r
1	1 3 6 8 23 154
2	1 3 6 23 28 154
F^{AF}	1 3 6 23 154

Table 8: Optimal solution for S-MCP and solution found using the Always Fortified rule on data set USCities and parameters $P = 40$, $Q = 6$ and $R = 2$.

	S-MCP	Heur	GAP
Fortification Set	1 3 6 8 24 154	1 3 6 23 28 154	-
Solution Value	3286216658.33	3290718099	0.137%

Table 9: Fortication sets and interdiction sets for MCP, data set randompoint3, parameters $P = 30$, $Q = 4$ and $r = 1, \ldots, 4$.

r	F^r	I^r
1	191 220 221 238	45
2	103 191 220 238	95 110
3	31 95 103 191	78 201 238
4	31 191 218 220	45 103 *115* 211

Table 10: Optimal solution for S-MCP and solution found using the Action Set rule on data set randompoint3 and parameters $P = 30$, $Q = 4$ and $R = 4$.

	S-MCP	Heur	GAP
Fortification Set	*115* 220 238 191	31 103 191 238	-
Solution Value	419738.70	420799.81	0.253%

Table 11: Fortication sets and interdiction sets for MCP, data set randompoint3, parameters $P = 30$, $Q = 3$ and $r = 1, \ldots, 3$.

r	F^r	I^r
1	191 221 238	220
2	103 191 238	220 221
3	31 103 191	51 55 110

Table 12: Optimal solution for S-MCP and solution found using the Action Set and Reaction Set rules on data set randompoint3 and parameters $P = 30$, $Q = 3$ and $R = 3$.

	S-MCP	Heur	GAP
Fortification Set	*115* 191 238	31 103 191	-
Solution Value	410457.17	411363.01	0.220%

Table 13: Comparison between optimal stochastic fortification set and optimal deterministic fortification set with expected number of attacks. Increasing probability case, USCities data set.

P	Q	R	$Z^\star$	$F^\star$	$\widetilde{Z}_{E(R)}$	$F_{E(R)}$	GAP
40	4	4	4476779495.50	1 8 52 154	4476779495.50	1 8 52 154	0.00%
40	6	4	4109323708.20	1 3 6 8 52 154	4109323708.20	1 3 6 8 52 154	0.00%
40	8	4	3782176529.30	1 2 3 4 6 8 28 154	3782176529.30	1 2 3 4 6 8 *24* 154	0.00%
50	5	4	3063770060.60	1 2 8 11 23	3097793925.60	1 *7* 8 11 23	1.11%
50	8	4	2931244316.40	1 8 14 23 24 30 37 52	2931244316.40	1 8 14 23 24 30 37 52	0.00%
50	10	4	2594055200.10	1 2 3 4 6 8 11 23 28 72	2620421806.90	1 2 3 4 6 8 11 *14* 23 *24*	1.02%
60	6	4	2439201722.40	1 2 6 8 19 56	2439201722.40	1 2 6 8 19 56	0.00%
60	9	4	2234785151.30	1 2 3 4 6 8 11 14 24	2355379896.80	1 2 3 4 6 8 *23* 24 *72*	5.40%
60	12	4	2007633404.70	1 3 4 6 8 11 23 24 29 52 71 72	2285839324.60	1 3 4 *5* 6 8 23 24 29 *32* 52 *56*	13.86%
						Average GAP	2.38%
						Max GAP	13.86%

Table 14: Comparison between optimal stochastic fortification set and optimal deterministic fortification set with expected number of attacks. Decreasing probability case, USCities data set.

P	Q	R	$Z^\star$	$F^\star$	$\widetilde{Z}_{E(R)}$	$F_{E(R)}$	GAP
40	4	4	3783835491.50	1 8 52 154	3858183410.20	1 *3 6* 154	1.96%
40	6	4	3552307233.80	1 3 6 8 52 154	3816640454.80	1 3 6 *23 24* 154	7.44%
40	8	4	3343016490.70	1 2 3 4 6 8 28 154	3343016490.70	1 2 3 4 6 8 28 154	0.00%
50	5	4	2616890238.90	1 2 8 11 23	2637828393.90	1 2 8 23 *30*	0.80%
50	8	4	2533408624.80	1 8 14 23 24 30 37 52	2561511577.00	1 *2 3 4 6* 8 23 *28*	1.11%
50	10	4	2355532806.40	1 2 3 4 6 8 11 23 28 72	2495408601.80	1 2 3 4 6 8 11 23 *24 32*	5.94%
60	6	4	2052782261.60	1 2 6 8 19 56	2071864998.60	1 *3 4* 6 8 *52*	0.93%
60	9	4	1940845535.50	1 2 3 4 6 8 23 29 72	1940845535.50	1 2 3 4 6 8 23 29 72	0.00%
60	12	4	1789395827.50	1 3 4 6 8 11 23 24 29 52 71 72	1846319529.70	1 *2* 3 4 6 8 *28* 29 *32 46* 71 72	3.18%
						Average GAP	2.37%
						Max GAP	7.44%

Table 15: Evaluation of the protection plan in the USCities data set: average and maximum solution cost in percentage.

		Used for protection plan	
		Increasing probability	Decreasing probability
Real world	Increasing probability	100% / 100%	100.60% / 105.14%
	Decreasing probability	100.38% / 102.00%	100% / 100%

http://www.kent.ac.uk/kbs/research-information/index.htm