A Facility Reliability Problem: Formulation, Properties and Algorithm

Michael Lim[1], Mark S. Daskin[2,3], Achal Bassamboo[3], Sunil Chopra[3,2]

[1] Department of Business Administration,
University of Illinois at Urbana–Champaign, Champaign, IL 61820
mlim@illinois.edu

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

[3] Department of Managerial Economics and Decision Sciences,
Northwestern University, Evanston, IL 60208
a-bassamboo@kellogg.northwestern.edu; s-chopra@kellogg.northwestern.edu

---

Having a robustly designed supply chain network is one of the most effective ways to hedge against network disruptions since contingency plans in the event of a disruption are often significantly limited. In this paper, we study the facility reliability problem: how to design a reliable supply chain network in the presence of random facility disruptions with the option of hardening selected facilities. We consider a facility location problem incorporating two types of facilities, one that is unreliable and another that is reliable (which is not subject to disruption, but is more expensive). We formulate this as a mixed integer programming model and develop a Lagrangian Relaxation-based solution algorithm. We derive structural properties of the problem and show that for some values of the disruption probability, the problem reduces to the classical uncapacitated fixed charge location problem. In addition, we show that the proposed solution algorithm is not only capable of solving large-scale problems, but is also computationally effective.

Keywords: Supply chain disruption; Facility location; Network design

---

1. Introduction

Each year companies face numerous unexpected events in their supply chains. While companies manage to survive from risks arising at the operational level, many suffer heavily when longer term disruptions impact their supply chain networks. Supply chain disruptions are entirely different from operational level mishaps (machine failures or short-term supply-demand imbalances) since they completely block the flow of the network for a significant amount of time. To hedge against such events, providing a robustly designed network is of the utmost importance since contingency plans in the event of a major disruption are significantly limited. Imagine a supply chain network that is designed to minimize the operational cost without considering any disruption scenarios. Although it may be very difficult to exactly quantify the damage of supply chain disruptions, it is not impossible to infer the range of their magnitude. Yet, managers tend to underestimate (if not completely ignore) the impact of supply chain disruptions, deceived by their small probability of occurrence.

Recent incidents such as the SARS outbreak in Asia or the 9/11 terrorist attacks illustrate that today's supply chains operate in a highly uncertain environment and the consequences of a disruption can be devastating. Motorola and Honda incurred substantial supply chain delays after the outbreak of SARS (Businessweek, 2003); Ford suffered from part shortages due to the border closing after the 9/11 (Associated Press 2001). According to Swiss Re (2008), annual losses due to natural hazards have increased dramatically in the past 30 years despite large fluctuations as shown in Figure 1. Not surprisingly, the frequency of man-made disasters is also on the rise, due to political issues such as military conflicts and terrorist actions (Coleman 2006, Swiss Re 2009). These examples collectively highlight the importance of protecting supply chain networks against disruptions.

In this paper, we study *facility location and demand assignment decisions* incorporating facility disruptions. Facility location problems have been actively studied spanning numerous application areas in both public and private sectors (Drezner 2002). However, classical facility location problems implicitly assume that all facilities are perfectly reliable and derive the optimal facility locations under this ideal situation. We provide strategies for designing a robust supply chain network with the option of *facility hardening* in the presence of *random* facility disruptions. Here, not all facilities are assumed to be reliable. We consider two types

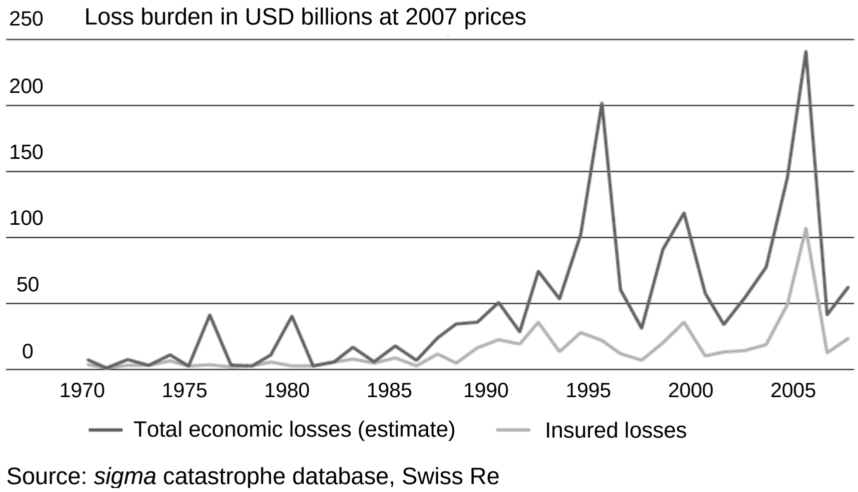

Figure 1: Insured losses attributable to natural hazards over the past years

of facilities: regular facilities, hereafter referred to as *unreliable* facilities, which are subject to random disruptions, and *reliable* facilities that are "*hardened*" as a result of a substantial investment. We assume reliable facilities are not subject to disruption.

We propose a facility reliability problem (FRP) incorporating the option of facility hardening to hedge against the risk of facility disruptions. The FRP is formulated as a mixed integer programming model and we propose a Lagrangian relaxation algorithm as a solution method. By solving the model, we determine the optimal number and locations of both types of facilities, as well as the assignments of demands to facilities. The main contributions of our paper are as follows:

(1) By formulating and solving the FRP, we study the relationship between the facility decisions (e.g., facility locations, hardening investment) and some key factors (e.g., disruption probability and customer demand) in the presence of random disruptions.

(2) We analytically show that under certain conditions, the FRP reduces to one of the well-studied classical facility location problems.

(3) We propose a Lagrangian Relaxation-based solution algorithm that is fast and effective for solving the FRP. For the examples analyzed, the proposed algorithm provides near-optimal solutions quickly (with an optimality gap below 0.1%) and handles larger problem

instances compared to CPLEX.

The remainder of this paper is organized as follows. In Section 2, the relevant literature is reviewed. In Section 3, the FRP is introduced along with the mathematical formulation and the solution algorithm. In Section 4, we study some structural properties of the model. Numerical results and computational tests are presented in Section 5. Finally, discussions and conclusions of the study comprise Section 6.

2. Literature Review

Increasing the reliability of facility network systems has been mainly studied by extending one or more of the classical location problems. Emergency service systems are one of the most important applications in the field. As such, maximizing the *coverage* was often the primary objective. Daskin (1983), ReVelle and Hogan (1989) and Batta et al. (1989) focus on maximizing the expected demand coverage while Hogan and ReVelle (1986) and Batta and Mannur (1990) focus on individual demand coverage with some degree of redundancy. Other works such as Drezner (1987) and Berman et al. (2007) focus on minimizing the demand-weighted transportation cost by modifying the p-median problem (pMP). Drezner (1987) provides a simple extension along with a heuristic solution method. Berman et al. (2007) analyze structural and algorithmic aspects of the pMP considering possible facility failures. They show that it is beneficial to centralize or even co-locate the facilities (for the extreme case) as the probability of facility failures increases.

In this paper, we extend the uncapacitated fixed charge location problem (UFLP): the decision maker seeks to locate a set of facilities that minimizes the sum of the fixed facility costs and the transportation costs. Several papers introduce reliability models extending the UFLP (as well as the pMP problem) in a context similar to ours. Snyder and Daskin (2005) introduces a multi-objective model from which the trade-off curve between the operational cost and the expected failure cost is derived. They demonstrate that the expected failure cost can be reduced significantly with minimal increases in the operating cost under normal circumstances. In Snyder and Daskin (2006), the concept of stochastic p-robustness is introduced where the relative regret is always less than p for any possible scenario. With this, the decision maker can minimize the total cost at a desired level of system performance. Zhan et al. (2007) and Cui et al. (2008) point out that site-specific failure probabilities may impact the choice of facility locations significantly and extend the literature by incorporating such characteristics. Zhan et al. (2007) provide a genetic algorithm and Cui et al. (2008) provide a Lagrangian relaxation algorithm as a solution method. We also allow site-specific failure probabilities in our model.

Our paper differs from these papers in two ways. First, we consider an option of "facility hardening;" hence another set of decisions is made. The notion of facility hardening implies various protection plans ranging from physical facility protection to exogenous outsourcing contracts. Using a mixture of reliable and unreliable facilities is akin to problems studied in Tomlin (2006) and Chopra et al. (2007) in the supplier risk management literature. These studies show that using reliable facilities can be valuable (even though they are more expensive) and effective in hedging against disruption risks. Second, we require each demand to have a "backup assignment" to a reliable facility whereas the earlier research allows a cascading strategy to the next available facility. Our approach significantly reduces the operational complexity of a firm since the hierarchical structure between the facilities is clear. Also, our approach allows us to study the problem analytically. A close similarity can also be found to Pirkul and Schilling (1988) and Pirkul (1989) where they designate the primary and secondary facilities for each demand node. However, our approach differs from these papers in a number of ways: we consider facility hardening decisions and provide structural properties of the model.

Church and Scaparra (2007) and Scaparra and Church (2008) extend the pMP to the case in which the decision maker tries to find the optimal number of facilities q to protect to hedge against a predetermined number, r, of facility disruptions. However, the facility disruption in these papers is caused by an intelligent adversary who tries to maximize the damage to the system. The objective of the model differs from the objective in our model - their model minimizes against the worst case scenario rather than the expected failure case - thus the protection policy significantly differs as a result.

Finally, we note that the FRP is also formulated in a *continuous model* in Lim et al. (2009). In that paper, the focus is on deriving managerial implications. The discrete model and the continuous model are compared using a case example. For more details on reliability considerations in facility location models, please refer to Snyder et al. (2006).

3. The Facility Reliability Model (FRP)

3.1 Problem Formulation

The facility reliability problem (FRP) extends the uncapacitated fixed charge location problem (UFLP) taking random facility disruptions into account. The objective of the problem is to minimize the total facility fixed cost and the *expected* transportation cost by properly locating reliable and unreliable facilities. We assume that each node $j \in N$, where N is a set of all nodes, is a demand node and a candidate facility site. At each node, we may locate either an unreliable facility at a cost of f_j^U which may fail with probability q_j $(0 < q_j < 1)$ or a reliable facility at a cost of f_j^R that does not fail. For the reliability premium, we assume $f_j^R > f_j^U$. In keeping with the UFLP, the decision maker determines the optimal number and location of both types of facilities while serving all demand nodes with probability 1 utilizing primary and backup assignments. A demand node i is served primarily by "any" type of facility as a primary assignment and by the closest "reliable" facility as a backup assignment in case the primarily assigned facility fails. The transportation cost per unit demand from a demand at node i to a facility at node j is given by d_{ij}^P and d_{ij}^B for a primary assignment and backup assignment, respectively. We allow $d_{ij}^B \geq d_{ij}^P$ to capture the penalty cost (emergency cost) associated with utilizing the backup source. When the closest facility from a certain demand node is a reliable facility, the primary assignment and the backup assignment are identical and we assume the unit transportation cost is always d_{ij}^P. To incorporate this into our model, we introduce d_{ij}^S $(= d_{ij}^B - d_{ij}^P \geq 0)$ as the unit savings that will be used to adjust the objective function by subtracting the over-levied travel cost as a savings. Facilities are assumed to have ample capacity to cover all the demands assigned to them and each node i has a demand of h_i. The complete notation for the FRP is summarized below.

Inputs:

N = set of all nodes

h_i = demand at node $i \in N$

f_j^U = fixed cost of locating an unreliable facility which is subject to failure at node $j \in N$

f_j^R = fixed cost of locating a reliable facility which does not fail at node $j \in N$

q_j = probability that an unreliable facility at $j \in N$ will be in the failure state

d_{ij}^P = unit transportation cost for a primary assignment from demand node $i \in N$ to a facility at $j \in N$

d_{ij}^B = unit transportation cost for a backup assignment from demand node $i \in N$ to a reliable facility at $j \in N$

d_{ij}^S = unit savings cost when demand node $i \in N$ is assigned to a reliable facility at $j \in N$ as both the primary and backup facility

Decision Variables:

X_j^U = 1 if an unreliable facility is located at candidate site j; 0 if not

X_j^R = 1 if a reliable facility is located at candidate site j; 0 if not

Y_{ij}^P = 1 if demands at i are assigned to a facility at j as the primary site; 0 if not

Y_{ij}^B = 1 if demands at i are assigned to a facility at j as the backup site; 0 if not

Y_{ij}^S = 1 if demands at i are assigned to a facility at j as the primary and backup site; 0 if not.

With this notation, we formulate the problem as follows:

$$
\begin{aligned}
\text{[FRP]}\quad \underset{X,Y}{\text{minimize}}\quad & \sum_{j\in N} f_j^U X_j^U + \sum_{j\in N} f_j^R X_j^R \\
& + \sum_{i\in N}\sum_{j\in N}(1-q_j)h_i d_{ij}^P Y_{ij}^P + \sum_{i\in N}\sum_{j\in N} q_j h_i d_{ij}^B Y_{ij}^B - \sum_{i\in N}\sum_{j\in N} q_j h_i d_{ij}^S Y_{ij}^S \qquad (1)
\end{aligned}
$$

subject to

$$\sum_{j\in N} Y_{ij}^P = 1, \qquad \sum_{j\in N} Y_{ij}^B = 1 \qquad \forall i \in N \tag{2}$$

$$Y_{ij}^P \leq X_j^R + X_j^U, \qquad Y_{ij}^B \leq X_j^R \qquad \forall i,j \in N \tag{3}$$

$$Y_{ij}^S \leq Y_{ij}^B, \qquad Y_{ij}^S \leq Y_{ij}^P \qquad \forall i,j \in N \tag{4}$$

$$X_j^R + X_j^U \leq 1 \qquad \forall j \in N \tag{5}$$

$$\sum_{j\in N} X_j^R \geq 1 \tag{6}$$

$$X_j^U,\, X_j^R \in \{0,1\} \qquad \forall j \in N \tag{7}$$

$$Y_{ij}^P,\, Y_{ij}^B,\, Y_{ij}^S \in \{0,1\} \quad \forall i,j \in N. \tag{8}$$

The objective function (1) consists of five terms. The first two are the fixed facility location terms for unreliable and reliable facilities, respectively. The third term represents the expected transportation cost when demands are served by their primary facility (which may be a reliable or an unreliable facility). The fourth term represents the transportation cost when facilities are served by their backup facility. The fifth term adjusts for the overestimate of using d_{ij}^B instead of d_{ij}^P, when a demand node is assigned to a reliable facility as both the primary and backup assignments.

Constraints (2) ensure that each demand is assigned to a primary and a backup facility respectively. Constraints (3) state that the primary assignment must be to an open facility and that the backup assignment must be to a reliable facility. Constraints (4) state that the savings associated with any assignment can only be realized if the demand node is assigned to the same facility as its primary and its backup facility. Constraints (5) state that we cannot locate both an unreliable and a reliable facility at any candidate site. Constraint (6) ensures that we locate at least one reliable facility. Technically, this is a redundant constraint as it is implied by (2) and (3). However, we keep this constraint for later use. Constraints (8) are the integrality constraints.

3.2 Solution Method: Lagrangian Relaxation Algorithm

The FRP can be solved using a standard optimization solver such as CPLEX, but the computation time and resource usage grows drastically as the size of the problem increases. Thus, to solve this problem effectively, we develop an algorithm based on Lagrangian relaxation (LR). A comparison between CPLEX and the LR algorithm is in §5.2.

We relax constraints (2) with Lagrange multipliers λ and μ, respectively, and obtain the following problem:

$$\begin{aligned} \max_{\lambda,\mu} \min_{X,Y} \quad & \sum_{j\in N} f_j^U X_j^U + \sum_{j\in N} f_j^R X_j^R + \sum_{i\in N}\sum_{j\in N} \big((1-q_j)h_i d_{ij}^P - \lambda_i\big) Y_{ij}^P \\ & + \sum_{i\in N}\sum_{j\in N} \big(q_j h_i d_{ij}^B - \mu_i\big) Y_{ij}^B - \sum_{i\in N}\sum_{j\in N} \big(q_j h_i d_{ij}^S\big) Y_{ij}^S + \sum_{i\in N} \lambda_i + \sum_{i\in N} \mu_i \\ \text{subject to } & (3)-(8). \end{aligned} \tag{9}$$

We try to minimize the objective function with respect to the original decision variables X and Y while maximizing it with respect to λ and μ. The objective function (1) can be rewritten as follows:

$$\begin{aligned} \max_{\lambda,\mu} \min_{X,Y} \quad & \sum_{j\in N} f_j^U X_j^U + \sum_{j\in N} f_j^R X_j^R + \sum_{i\in N}\sum_{j\in N} \alpha_{ij} Y_{ij}^P \\ & + \sum_{i\in N}\sum_{j\in N} \beta_{ij} Y_{ij}^B - \sum_{i\in N}\sum_{j\in N} \gamma_{ij} Y_{ij}^S + \sum_{i\in N} \lambda_i + \sum_{i\in N} \mu_i \end{aligned} \tag{10}$$

where $\alpha_{ij} = (1-q_j)h_i d_{ij}^P - \lambda_i$, $\beta_{ij} = q_j h_i d_{ij}^B - \mu_i$, $\gamma_{ij} = q_j h_i d_{ij}^S$.

3.2.1 Algorithm Outline

The Lagrangian relaxation algorithm improves the lower and upper bound by constantly adjusting the Lagrange multipliers until either the gap between the lower and upper bound goes below a prespecified level or another termination criterion is attained. Notice that relaxing (2) allows the solution to be feasible even without having any reliable facility if (6) did not exist. We incorporate these constraints into the procedure for solving the FRP by using the algorithm below.

```
Initialize Lagrange multipliers
While (Termination condition not met) Do
{
    Increment iteration counter
    Compute the lower bound
    Compute the upper bound
    (Optional) Local improvement
    Check termination condition
    Update Lagrange multipliers
}
```

Now we present each step of the algorithm in detail.

3.2.2 Computation of the Lower Bound

We compute the lower bound first. For fixed values of the Lagrange multipliers λ and μ, (10) provides a lower bound on the objective function (1). We solve (10) by using the following procedure:

Step 0. Initialize all decision variables, $X_j^U, X_j^R, Y_{ij}^P, Y_{ij}^B, Y_{ij}^S$, to 0.

Step 1. Determine the value of locating an unreliable facility (i.e., setting $X_j^U = 1$) at site j for every such site. This will allow us to set $Y_{ij}^P = 1$ by the first constraint of (3) if doing so is advantageous to the Lagrangian objective function (10), which is minimized with respect to the original decision variables. We will want to set $Y_{ij}^P = 1$, if $\alpha_{ij} = (1 - q_j)h_i d_{ij}^P - \lambda_i < 0$. Thus, the value of setting $X_j^P = 1$, exclusive of the fixed facility costs, is $V_j^U = \sum_{i \in N} \min(0, \alpha_{ij}) = \sum_{i \in N} \min(0, (1 - q_j)h_i d_{ij}^P - \lambda_i)$.

Step 2. Determine the value of locating a reliable facility at a node (i.e., setting $X_j^R = 1$) in a similar manner. This is somewhat more complicated than step 1 since doing so impacts constraints, (3), and by implication constraints (4). Once we set $X_j^R = 1$, we will want to set Y_{ij}^P, Y_{ij}^B and Y_{ij}^S

to minimize the sum of the corresponding terms of the Lagrangian function. In other words, we will set:

We compute the lower bound first. For fixed values of the Lagrange multipliers λ and μ, (10) provides a lower bound on the objective function (1). We solve (10) by using the following procedure:

$$
\begin{array}{ll}
Y_{ij}^P = 1 & \text{if } \alpha_{ij} = \min(\alpha_{ij}, \beta_{ij}, \alpha_{ij} + \beta_{ij} - \gamma_{ij}) < 0, \\
Y_{ij}^B = 1 & \text{if } \beta_{ij} = \min(\alpha_{ij}, \beta_{ij}, \alpha_{ij} + \beta_{ij} - \gamma_{ij}) < 0, \\
Y_{ij}^P = Y_{ij}^B = Y_{ij}^S = 1 & \text{if } \alpha_{ij} + \beta_{ij} - \gamma_{ij} = \min(\alpha_{ij}, \beta_{ij}, \alpha_{ij} + \beta_{ij} - \gamma_{ij}) < 0, \\
Y_{ij}^P = Y_{ij}^B = Y_{ij}^S = 0 & \text{otherwise.}
\end{array}
$$

Step 3. From these values we now compute $V_j^U = \sum_{i \in N} \min(0, \alpha_{ij}, \beta_{ij}, \alpha_{ij} + \beta_{ij} - \gamma_{ij})$. With the values of V_j^U and V_j^R in hand, we can begin to compute optimal values for the location variables. Ignoring constraint (6) for the moment, we set:

$$
\begin{array}{ll}
X_j^U = 1 & \text{if } V_j^U + f_j^U = \min(V_j^U + f_j^U, V_j^R + f_j^R) < 0, \\
X_j^R = 1 & \text{if } (V_j^R + f_j^R < V_j^U + f_j^U) \text{ and } (V_j^R + f_j^R < 0), \\
X_j^U = X_j^R = 0 & \text{otherwise.}
\end{array}
$$

Step 4. If, after setting the location variables in this manner, we have $\sum_{j \in N} X_j^R \geq 1$, then the lower bound problem is solved in terms of the location variables. If $\sum_{j \in N} X_j^R = 0$, we need to modify the location decisions as follows. Let $N_1 = \{j | X_j^U = 1\}$ and $N_0 = \{j | X_j^U = 1\}$. Now let $V_1 = \min_{j \in N_1} \left\{V_j^R + f_j^R - (V_j^U + f_j^U)\right\}$ and $j_1 = \arg\min_{j \in N_1} \left\{V_j^R + f_j^R - (V_j^U + f_j^U)\right\}$. Similarly, let $V_0 = \min_{j \in N_0} \{V_j^R + f_j^R\}$ and $j_0 = \arg\min_{j \in N_0} \{V_j^R + f_j^R\}$. Finally, we set:

$$
\begin{array}{ll}
X_{j_1}^R = 1 \text{ and } X_{j_1}^U = 0 & \text{if } V_1 < V_0, \\
X_{j_0}^R = 1 & \text{otherwise.}
\end{array}
$$

Once we have the values of the location variables, we simply need to extract the allocation variables using the relations above.

3.2.3 Computation of the Upper Bound

We now compute the upper bound by using the following procedure:

Step 0. Recall the location variables, X_j^U, X_j^R from the lower bound solution. The primary assignment Y_{ij}^P for each demand node is made to the closest open reliable or unreliable facility.

Step 1. The backup assignment Y_{ij}^B for each node is made to the closest open reliable facility. If the assignments are the same, the savings term activates ($Y_{ij}^S = 1$) as well.

Using the decision variables above, we compute the upper bound for the objective function (1). The location variables, X_j^U, X_j^R are not necessarily the optimal locations of the original FRP, hence this solution is clearly an upper bound of the FRP.

3.2.4 Lagrange Multiplier Initialization / Update

The Lagrange multipliers are initialized and updated by the following procedure:

Initialization. We initialize the Lagrange multipliers, λ and μ, by using the following values:

$$\lambda_i^{(0)} = \frac{10}{|N|}\Big\{\sum_{j\in N} f_j^U + h_i\Big\},$$
$$\mu_i^{(0)} = \frac{10}{|N|}\Big\{\sum_{j\in N} f_j^R + h_i\Big\}.$$

Update. The Lagrange multipliers are updated at each iteration using standard subgradient optimization as described in Fisher (1981, 1985) and Daskin (1995) with the search direction modification proposed by Crowder (1976). The following formulae update λ and μ at iteration n:

$$\lambda_i^{(n+1)} = \max\{0, \lambda_i^{(n)} + t^{(n)} \cdot d_i^{\lambda(n)}\},$$
$$\mu_i^{(n+1)} = \max\{0, \mu_i^{(n)} + t^{(n)} \cdot d_i^{\mu(n)}\} \tag{11}$$

where $d_i^{\lambda(n)}$ and $d_i^{\mu(n)}$ represent the direction to move respectively for λ and μ at the n^{th} iteration with the step size of $t^{(n)}$. In particular:

Step 0. We first compute the updated direction at n^{th} iteration as

$$d_i^{\lambda(n)} = \Big(1 - \sum_{j\in N} Y_{ij}^P\Big) + C \cdot d_i^{\lambda(n-1)},$$
$$d_i^{\mu(n)} = \Big(1 - \sum_{j\in N} Y_{ij}^B\Big) + C \cdot d_i^{\mu(n-1)}$$

where C is a Crowder damping constant. We typically set $C = 0.3$ and $d_i^{\mu(0)} = d_i^{\lambda(0)} = 0$ $(\forall i)$.

Step 1. We next compute the stepsize $t^{(n)}$ as

$$t^{(n)} = \alpha^{(n)} \frac{(UB - \mathcal{L}^{(n)})}{\sum_{i \in N} \left(d_i^{\mu(n)}\right)^2 + \sum_{i \in N} \left(d_i^{\lambda(n)}\right)^2}$$

where $\alpha^{(n)}$ is a constant at iteration n, initialized to 2.0 and halved when 24 consecutive iterations fail to improve the lower bound. UB denotes the best (lowest) known upper bound through iteration n and $\mathcal{L}^{(n)}$ denotes the value of the lower bound found at iteration n.

Step 2. Finally, we update λ and μ by using (11) above.

3.2.5 Termination Conditions

The algorithm terminates when any of the following conditions are met:

1. Optimality Gap: When the gap between the lower and upper bound gets close to zero; i.e., $\frac{(UB - \mathcal{L}^{(n)})}{\mathcal{L}^{(n)}} < \epsilon$ where we typically prespecified the tolerance value ϵ to 0.00001.

2. Maximum Iteration: When the iteration number reaches a prespecified number; i.e., $n > n_{\max}$ where we typically set the iteration limit to 10,000.

3. Stepsize Constant: When the stepsize constant $\alpha^{(n)}$ gets close to zero; i.e., $\alpha^{(n)} < \alpha_{\min}$ where we typically set the stepsize constant limit to 0.0001.

3.2.6 (Optional) Local Improvements

Four local heuristic improvement (LHI) algorithms can be applied to the solution to enhance the performance of the algorithm. These heuristics typically work well in the early iterations until the solution becomes stable.

LHI 1. We consider changing an unreliable site to a reliable site. This incurs the incremental hardening cost, but may reduce the transport costs to some nodes sufficiently since the customers at those nodes no longer need to travel to a more remote backup or reliable facility in the event that the facility at node j fails. Also, this hardened site may reduce the backup transportation cost for some demand nodes which had been assigned to a more remote backup facility.

LHI 2. Similarly, we consider changing a reliable facility into an unreliable facility, provided there is at least one other reliable facility in the system. These two improvement algorithms are very fast since, for each facility, there is at most one other objective function evaluation.

LHI 3. We consider an exchange algorithm, in which each unreliable facility in turn is removed and replaced by an unreliable facility at a node that does not have a facility.

LHI 4. Lastly, we consider an exchange algorithm for the reliable facilities. The last two improvement algorithms are more intensive as they require $O(|J|)$ objective function evaluations for every facility, where J is the set of candidate locations.

4. Structural Properties

In this section, we develop important structural properties of the facility reliability problem (FRP). We show that for extreme values of the disruption probability, the FRP is reducible to the uncapacitated fixed charge location problem (UFLP). Consequently, we derive threshold bounds on the disruption probability that satisfy such conditions. We begin with the case when the disruption probabilities are large.

Theorem 1. *There exists a threshold disruption probability $\bar{q}_{th}\,(< 1)$ for the FRP such that if all $q_j \geq \bar{q}_{th}$, then it is optimal to deploy only reliable facilities. Further, the optimal location of these facilities coincides with the optimal solution to the UFLP with facility costs f_j^R and distances d_{ij}^P.*

Proof. Let $\mathbf{q}$ be the disruption probability vector. Consider a feasible solution $S : (\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S)$ which contains *at least one unreliable facility* where vectors $\mathbf{X}^U, \mathbf{X}^R$ and $\mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S$ represent the optimal locations and assignments for such a solution, respectively. Now consider another solution $\tilde{S} : (\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B)$. Denoting $Z^{\text{FRP}}(S)$ as the total cost of the FRP with the configuration S, we know the following holds for any $\mathbf{q}$:

$$
\begin{aligned}
& Z^{\text{FRP}}(\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S) - Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B) \\
&= \sum_{j\in N} f_j^U X_j^U + \sum_{i\in N}\sum_{j\in N}(1-q_j)h_i d_{ij}^P Y_{ij}^P + \sum_{i\in N}\sum_{j\in N} q_j h_i d_{ij}^B Y_{ij}^B - \sum_{i\in N}\sum_{j\in N} q_j h_i d_{ij}^S Y_{ij}^S - \sum_{i\in N}\sum_{j\in N} h_i d_{ij}^P Y_{ij}^B \\
&= \sum_{j\in N} f_j^U X_j^U + \sum_{i\in N}\sum_{j\in N} q_j h_i (d_{ij}^B - d_{ij}^P)(Y_{ij}^B - Y_{ij}^S) - \sum_{i\in N}\sum_{j\in N}(1-q_j)h_i d_{ij}^P (Y_{ij}^B - Y_{ij}^P) \\
&\geq \sum_{j\in N} f_j^U X_j^U - \sum_{i\in N}\sum_{j\in N}(1-q_j)h_i d_{ij}^P (Y_{ij}^B - Y_{ij}^P).
\end{aligned}
$$

The inequality holds since $Y_{ij}^B \geq Y_{ij}^S$ and $d_{ij}^B \geq d_{ij}^P$. Since $\sum_{j\in N} f_j^U X_j^U > 0$, we can find a $\mathbf{q}$ that satisfies $Z^{\text{FRP}}(\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S) \geq Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B)$ for q_j's that are sufficiently close to 1. That is, for a given configuration S, there always exists a threshold probability $\bar{q}(S)$ such that if $q_j \geq \bar{q}(S)\ \forall j$, then it is optimal to use only reliable facilities. Then, we can find a threshold probability $\bar{q}_{th}$ by taking the supremum over all possible $\bar{q}(S)$'s; i.e., $\sup_{S\in\Omega} \bar{q}(S) =: \bar{q}_{th}$ where Ω represents all possible configurations for a given network. There is only finite number of possible configurations in Ω, thus we know that $\bar{q}_{th} < 1$. Therefore, we conclude that if all $q_j \geq \bar{q}_{th}\ \forall j$, then it is optimal to deploy only reliable facilities.

Now, let $Z^{\text{UFLP}}(\mathbf{X}^R, \mathbf{Y}^B)$ be the total cost for the UFLP with the facility cost of f_j^R and the distance of d_{ij}^P where $\mathbf{X}^R, \mathbf{Y}^B$ are the location and assignment vectors. Note that this is equivalent to $Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B)$. Then, it follows that

$$\begin{aligned} Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B) &= Z^{\text{UFLP}}(\mathbf{X}^R, \mathbf{Y}^B) \\ &\geq Z^{\text{UFLP}}(\mathbf{X}^*, \mathbf{Y}^*) \ = \ Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^*, \mathbf{Y}^*, \mathbf{Y}^*, \mathbf{Y}^*) \end{aligned} \tag{12}$$

where $\mathbf{X}^*$ and $\mathbf{Y}^*$ is the optimal solution to the UFLP with the facility cost of f_j^R and the distance of d_{ij}^P. From the first statement, we know that all facilities should be hardened if $q_j \geq \bar{q}_{th}$ and under this condition, (12) implies that the optimal solution of the FRP coincides with the optimal solution to the UFLP with facility costs f_j^R and distances d_{ij}^P. ■

In the proof of Theorem 1, we notice from (12) that $\mathbf{q}$ satisfies $Z^{\text{FRP}}(\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S) \geq Z^{\text{FRP}}(\mathbf{0}, \mathbf{X}^R, \mathbf{Y}^B, \mathbf{Y}^B, \mathbf{Y}^B)$ if $\sum_{j\in N} f_j^U X_j^U - \sum_{i\in N}\sum_{j\in N}(1-q_j)h_i d_{ij}^P(Y_{ij}^B - Y_{ij}^P) \geq 0$. We can see that the worst bound would correspond to the case where all q_j's are equal. A $\bar{q}(S) = q_j$ that satisfies this relationship for configuration S can then be derived as

$$\begin{aligned} &\sum_{j\in N} f_j^U X_j^U - (1-\bar{q}(S))\sum_{i\in N}\sum_{j\in N} h_i d_{ij}^P (Y_{ij}^B - Y_{ij}^P) \geq 0 \\ \Longrightarrow \quad &\bar{q}(S) \geq 1 - \frac{\sum_{j\in N} f_j^U X_j^U}{\sum_{i\in N}\sum_{j\in N} h_i d_{ij}^P (Y_{ij}^B - Y_{ij}^P)}. \end{aligned} \tag{13}$$

Since $\sup_{S\in\Omega} \bar{q}(S) =: \bar{q}_{th}$, we construct an upper bound on $\bar{q}_{th}$ in the following corollary.

Corollary 1. *The threshold disruption probability $\bar{q}_{th}$ is bounded above (upper bound) by the value* *of* $\max\left[0,\, 1 - \frac{f^U_{\min}}{\sum_{i\in N} h_i(d^P_{\max i} - d^P_{\min i})}\right]$, *where* $d^P_{\max i} = \max_j\{d^P_{ij}\}$ *and* $d^P_{\min i} = \min_j\{d^P_{ij}\}$.

Proof. Note that Y^B_{ij} is the backup assignment to the closest reliable facility from demand i while Y^P_{ij} is to the closest facility (of either type) to demand i. Thus, for every i, $0 \le h_i\big(\sum_{j\in N} d^P_{ij}(Y^B_{ij} - Y^P_{ij})\big) \le h_i\big(d^P_{\max i} - d^P_{\min i}\big)$ holds. Consequently, $0 \le \sum_{i\in N}\sum_{j\in N} h_i d^P_{ij}(Y^B_{ij} - Y^P_{ij})$ $\le \sum_{i\in N} h_i\big(d^P_{\max i} - d^P_{\min i}\big)$ holds. Hence, it follows that

$$
\begin{aligned}
&\frac{1}{\sum_{i\in N} h_i\big(d^P_{\max i} - d^P_{\min i}\big)} \le \frac{1}{\sum_{i\in N}\sum_{j\in N} h_i d^P_{ij}(Y^B_{ij} - Y^P_{ij})}\\
\Longrightarrow\ &\frac{f^U_{\min}}{\sum_{i\in N} h_i\big(d^P_{\max i} - d^P_{\min i}\big)} \le \frac{\sum_{j\in N} f^U_j X^U_j}{\sum_{i\in N}\sum_{j\in N} h_i d^P_{ij}(Y^B_{ij} - Y^P_{ij})}\\
\Longrightarrow\ & 1 - \frac{\sum_{j\in N} f^U_j X^U_j}{\sum_{i\in N}\sum_{j\in N} h_i d^P_{ij}(Y^B_{ij} - Y^P_{ij})} \le 1 - \frac{f^U_{\min}}{\sum_{i\in N} h_i\big(d^P_{\max i} - d^P_{\min i}\big)}
\end{aligned}
\tag{14}
$$

where $f^U_{\min}$ represents the unreliable facility with the lowest fixed cost. By (13) and (14), we have an upper bound on $\bar{q}_{th}$ in $1 - \frac{f^U_{\min}}{\sum_{i\in N} h_i\big(d^P_{\max i} - d^P_{\min i}\big)}$. Note that this bound may become negative when the fixed cost is much greater than transportation cost (since the FRP requires having at least one reliable facility), hence we take the maximum between 0 and $1 - \frac{f^U_{\min}}{\sum_{i\in N} h_i(d^P_{\max i} - d^P_{\min i})}$.
■

To summarize, the above theorem and corollary suggest that when the disruption probabilities are larger than the threshold ($q_j \ge \bar{q}_{th}$), it is optimal to have only reliable facilities. Once it becomes advantageous to harden all the facilities, the FRP is solved in a risk-free environment, thus its optimal facility location becomes identical to that of the UFLP.

Next, we consider the case when the disruption probabilities are very small.

Theorem 2. *If the facility hardening costs are identical for all sites ($f^R_j - f^U_j = \delta\, \forall j$), there exists a threshold disruption probability $\underline{q}_{th}$ (> 0) for the FRP such that if all $q_j \le \underline{q}_{th}$, then it is optimal to harden exactly one of the facilities. Further, the optimal locations of the facilities (including both reliable and unreliable) coincide with the optimal locations to the UFLP with facility costs f^U_j and distances d^P_{ij}.*

Proof. Consider a feasible solution $S : (\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S)$ which contains *at least two reliable facilities*. Say $\sum_{j\in N} X^R_j$ is $\kappa\,(\ge 2)$. Also, consider another feasible solution which converts all the reliable facilities from S into unreliable facilities except for one site. (Recall, the smallest number of reliable facilities is one for the FRP.) More precisely, let $\tilde{S} : (\tilde{\mathbf{X}}^U, \tilde{\mathbf{X}}^R, \tilde{\mathbf{Y}}^P, \tilde{\mathbf{Y}}^B, \tilde{\mathbf{Y}}^S)$ be as follows:

$\tilde{\mathbf{X}}^R$ = a unit vector with k^{th} element being 1 where $k \in \{j \mid X_j^R = 1\}$

(convert all the facilities from $\mathbf{X}^R$ to unreliable ones except for one)

$\tilde{\mathbf{X}}^U$ $= \mathbf{X}^U + \mathbf{X}^R - \tilde{\mathbf{X}}^R$ (existing unreliable facilities, $\mathbf{X}^U$, plus the converted facilities from $\mathbf{X}^U$)

$\tilde{\mathbf{Y}}^P$ = Each demand is assigned to its closest facility; i.e., $\tilde{\mathbf{Y}}^P = \mathbf{Y}^P$

$\tilde{\mathbf{Y}}^B$ = All demands are assigned to the k^{th} facility; i.e., $\tilde{Y}_{ij}^B = 1$ if $j = k$; 0 if not

$\tilde{\mathbf{Y}}^S$ = Demands whose primary and backup assignments are identical; i.e., $\tilde{Y}_{ij}^S = 1$ if $j = k$; 0 if not, hence $\tilde{\mathbf{Y}}^S = \tilde{\mathbf{Y}}^B$.

Then, for any disruption probability $\mathbf{q}$, we know the following:

$$
\begin{aligned}
& Z^{\text{FRP}}(\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S) - Z^{\text{FRP}}(\tilde{\mathbf{X}}^U, \tilde{\mathbf{X}}^R, \tilde{\mathbf{Y}}^P, \tilde{\mathbf{Y}}^B, \tilde{\mathbf{Y}}^S) \\
&= \sum_{j \in N} f_j^U (X_j^U - \tilde{X}_j^U) + \sum_{j \in N} f_j^R (X_j^R - \tilde{X}_j^R) + \sum_{i \in N} \sum_{j \in N} q_j h_i d_{ij}^B (Y_{ij}^B - \tilde{Y}_{ij}^B) - \sum_{i \in N} \sum_{j \in N} q_j h_i d_{ij}^S (Y_{ij}^S - \tilde{Y}_{ij}^S) \\
&= (\kappa - 1)\delta + \sum_{i \in N} \sum_{j \in N} q_j h_i d_{ij}^B (Y_{ij}^B - \tilde{Y}_{ij}^B) - \sum_{i \in N} \sum_{j \in N} q_j h_i d_{ij}^S (Y_{ij}^S - \tilde{Y}_{ij}^S).
\end{aligned}
$$

Since $\delta > 0$, we find a $\mathbf{q}$ that satisfies $Z^{\text{FRP}}(\mathbf{X}^U, \mathbf{X}^R, \mathbf{Y}^P, \mathbf{Y}^B, \mathbf{Y}^S) \geq Z^{\text{FRP}}(\tilde{\mathbf{X}}^U, \tilde{\mathbf{X}}^R, \tilde{\mathbf{Y}}^P, \tilde{\mathbf{Y}}^B, \tilde{\mathbf{Y}}^S)$ for q_j's that are sufficiently close to 0. That is, for a given configuration S, there always exists a threshold probability $\underline{q}_1(S)$ such that if $q_j \leq \underline{q}_1(S)\ \forall j$, then it is optimal to harden exactly one of the facilities.

Now, let $Z^{\text{UFLP}}(\bar{\mathbf{X}}, \bar{\mathbf{Y}})$ be the total cost of solving a UFLP with facility costs f_j^U and distances d_{ij}^P where $\bar{\mathbf{X}} = \tilde{\mathbf{X}}^U + \tilde{\mathbf{X}}^R$ and $\bar{\mathbf{Y}} = \tilde{\mathbf{Y}}^P$. Then, we can find a $\mathbf{q}$ that satisfies the following:

$$
\begin{aligned}
& Z^{\text{FRP}}(\tilde{\mathbf{X}}^U, \tilde{\mathbf{X}}^R, \tilde{\mathbf{Y}}^P, \tilde{\mathbf{Y}}^B, \tilde{\mathbf{Y}}^S) \\
&= \ Z^{\text{UFLP}}(\bar{\mathbf{X}}, \bar{\mathbf{Y}}) + \delta - \sum_{i \in N} \sum_{j \in N} q_j h_i \left[d_{ij}^P \tilde{Y}_{ij}^P - d_{ij}^B \tilde{Y}_{ij}^B + d_{ij}^S \tilde{Y}_{ij}^S \right] \\
&\geq \ Z^{\text{UFLP}}(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*) + \delta - \sum_{i \in N} \sum_{j \in N} q_j h_i \left[d_{ij}^P \bar{Y}_{ij}^{P*} - d_{ij}^B \bar{Y}_{ij}^{B*} + d_{ij}^S \bar{Y}_{ij}^{S*} \right] \\
&= \ Z^{\text{FRP}}(\bar{\mathbf{X}}^{U*}, \bar{\mathbf{X}}^{R*}, \bar{\mathbf{Y}}^{P*}, \bar{\mathbf{Y}}^{B*}, \bar{\mathbf{Y}}^{S*}) \qquad (15)
\end{aligned}
$$

where $\bar{\mathbf{X}}^*$ and $\bar{\mathbf{Y}}^*$ is the optimal solution to the UFLP with facility costs f_j^U and distances d_{ij}^P. The configuration $(\bar{\mathbf{X}}^{U*}, \bar{\mathbf{X}}^{R*}, \bar{\mathbf{Y}}^{P*}, \bar{\mathbf{Y}}^{B*}, \bar{\mathbf{Y}}^{S*})$ represents the optimal solution to the FRP with one facility hardened from the configuration $(\bar{\mathbf{X}}^*, \bar{\mathbf{Y}}^*)$ for the UFLP. Note that these configurations have exactly the same facility locations. Hence, there exists a threshold probability $\underline{q}_2(S)$ such that if $q_j \leq \underline{q}_2(S)\ \forall j$, then the optimal location for the FRP coincides with the optimal location for the UFLP.

From the first and second statements, let $\underline{q}(S) = \underline{q}_1(S) \cup \underline{q}_2(S)$. Then, we can find a threshold probability $\underline{q}_{th}$ as $\inf_{S\in\Omega} \underline{q}(S) =: \underline{q}_{th}$ for any given configuration $S \in \Omega$. There are only a finite number of possible configurations in Ω, thus we know that $\underline{q}_{th} > 0$. Therefore, if $q_j \leq \underline{q}_{th}$, then it is optimal to have only one reliable facility, and (15) implies that the optimal locations for the FRP coincide with the optimal locations for the UFLP with facility costs f_j^R and distances d_{ij}^P. ■

After solving the UFLP, we can decide which facility to harden by solving a suitably modified 1-median location problem. Since the facility hardening cost is constant, we find one facility to harden from among the opened facilities which minimizes the total transportation cost for the case when disruption occurs. Letting N' be the set which contains all the opened sites from Theorem 2, we summarize this in the following corollary.

Corollary 2. *The one facility that should be hardened from the UFLP in Theorem 2 can be determined by solving the following 1-median location problem:*

$$\begin{aligned}
\underset{X,Y}{\textit{mininize}} \quad & \sum_{i\in N}\sum_{j\in N'} h_i d'_{ij} Y_{ij}^B && (16)\\
& \textit{where } d'_{ij} = q_j d_{ij}^P \textit{ if facility } j \textit{ is the closest to demand } i;\ q_j d_{ij}^B \textit{ if not} \\
\textit{subject to} \quad & \sum_{j\in N'} Y_{ij}^B = 1 \quad \forall i \in N \\
& \sum_{j\in N'} X_j^R = 1 \\
& Y_{ij}^B \leq X_j^R \qquad \forall i \in N,\, \forall j \in N' \\
& Y_{ij}^B \in \{0,1\} \quad \forall i \in N,\, \forall j \in N', \qquad X_j^R \in \{0,1\} \quad \forall j \in N'.
\end{aligned}$$

Proof. Recall that the optimal solution of $\mathbf{X}^U, \mathbf{Y}^P$ from Theorem 2 is a minimizer of the UFLP, $\sum_{j\in N} f_j^U X_j^U + \sum_{i\in N}\sum_{j\in N} h_i d_{ij}^P Y_{ij}^P$. This minimizes $\sum_{j\in N} f_j^U X_j^U + \sum_{i\in N}\sum_{j\in N}(1-q_j) h_i d_{ij}^P Y_{ij}^P$

as well under $q_j \leq \underline{q}_{th}$. Also, the facility hardening cost, $\sum_{j \in N'} f_j^R X_j^R = \delta$, is a constant. From Theorem 2, we know $q_j \leq \underline{q}_{th}$ so that it is optimal to decide $\mathbf{X}^U, \mathbf{Y}^P$ first and then decide $\mathbf{X}^R, \mathbf{Y}^B$ (with fixed $\mathbf{X}^U, \mathbf{Y}^P$). Hence, reducing the solution set of $\mathbf{X}^R, \mathbf{Y}^B$ from N to N' does not affect the solution. With the definition of d'_{ij}, the objective function $\sum_{i \in N} \sum_{j \in N'} h_i d'_{ij} Y_{ij}^B$ is equivalent to $\sum_{i \in N} \sum_{j \in N} q_j h_i (d_{ij}^B Y_{ij}^B - d_{ij}^S Y_{ij}^S)$ and this is minimized by (16). Note that the objective function of FRP, (1), is the sum of these terms, thus the solution of (16) provides the global optimum. ■

Finally, we note that deriving a simple lower bound on $\underline{q}_{th}$ is difficult because of the condition on $\underline{q}_2(S)$. The second and third equation from (15) have different facility configurations and this makes it analytically challenging to calculate a good closed form bound on the disruption probability.

5. Numerical Results

5.1 Case Example

Case Input. To analyze the problem and to gain insights from the model, we employ a data set of 263 nodes representing the largest cities in the contiguous 48 states in the United States. The disruption probability q_j was assumed to occur independently and was randomly generated from $U \sim [0.025, 0.075]$. The cost of an unreliable facility f_j^U was determined by a fixed cost plus a variable cost which was a function of the population in each node; specifically, $f_j^U = 500,000 + 1.7h_j$. The hardening cost for each site was determined as a linear function of the disruption probability; i.e., $\delta_j = (f_j^R - f_j^U) = 5,000,000\, q_j$. Thus, cities facing higher disruption probabilities will be more costly to harden. In this setting, the hardening cost was roughly 25% of the fixed cost of the unreliable facility on average. The distance between any two cities was calculated as the great circle distance based on the longitude and latitude of the cities. Then, we multiplied the distance by $c = 0.002$ to get the transportation cost per unit distance, d_{ij}^P. For simplicity, the backup traveling cost was set to $d_{ij}^B = 1.25\, d_{ij}^P$.

Solution. The solution algorithm was coded in C++ and was run on an IBM workstation with 2.4GHz quad core processor and 9GB of virtual RAM (4GB of physical memory and 5GB of swap space). A solution with 0.0006% optimality gap (the percentage difference between the upper bound and lower bound) was found in 10.21 second after 304 iterations. Figure 2 shows the locations of the facilities and the assignments of demand nodes to the facilities for the case example. The solution

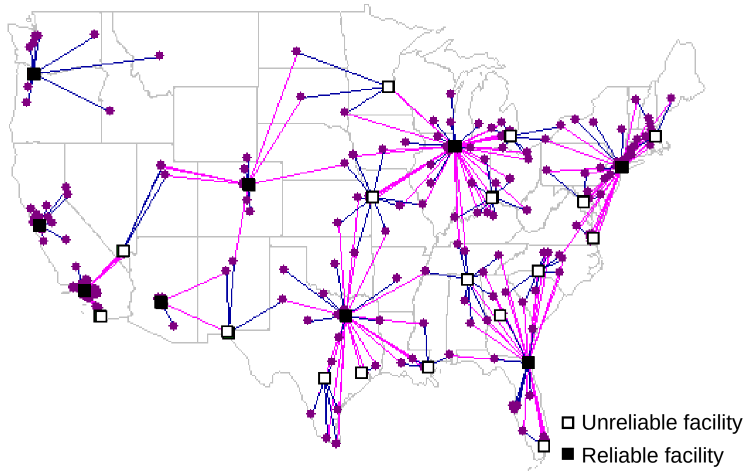

Figure 2: The solution to the case example

located 16 unreliable facilities (Alexandria, VA; Cambridge, MA; Clarksville, TN; Clearwater, FL; Dayton, OH; Kansas City, KS; Livonia, MI; New Haven, CT; Olympia, WA; Pasadena, TX; Pembroke Pines, FL; Saint Paul, MN; Santa Fe, NM; Springfield, IL; Westminster, CO; Winston-Salem, NC) and 9 reliable facilities (Grand Prairie, TX; Hampston, VA; Macon, GA; Naperville, IL; Norwalk, CA; Santa Clara, CA; Tempe, AZ; Trenton, NJ; West Valley City, UT). Note that demand nodes whose primary facility is a reliable site are shown with only a single assignment, while demand nodes served primarily by an unreliable facility are depicted with two assignments. The objective function value was $45,878,459 and breaks down as follows:

Table 1: Summary of the case solution

Components	Unreliable Facilities	Reliable Facilities	Both Facilities
Number of facilities	11	4	25
Fixed cost	$11,565,900	$8,399,130	$19,965,030
Assignments	Primary Assignments	Backup Assignments	Both Assignments
Exp. Transport'n Cost	$22,793,600	$3,228,700	$26,022,300
Subtotal	$34,359,500	$11,627,830	-
Expected Savings			$108,871
Total		$45,878,459	

Modified Case Examples. We also numerically confirmed Theorems 1 and 2 by changing the disruption probabilities. Figure 3(a) corresponds to the case in which q is large (Theorem 1) and Figure 3(b) corresponds to the case in which q is small (Theorem 2). Here, we set $q_j = q = 0.2$ and

0.01, respectively, and the other parameters were unchanged. In each case, we can observe that the number of reliable facilities is pushed to the extremes (either harden all or just one) depending on the values of the disruption probabilities.

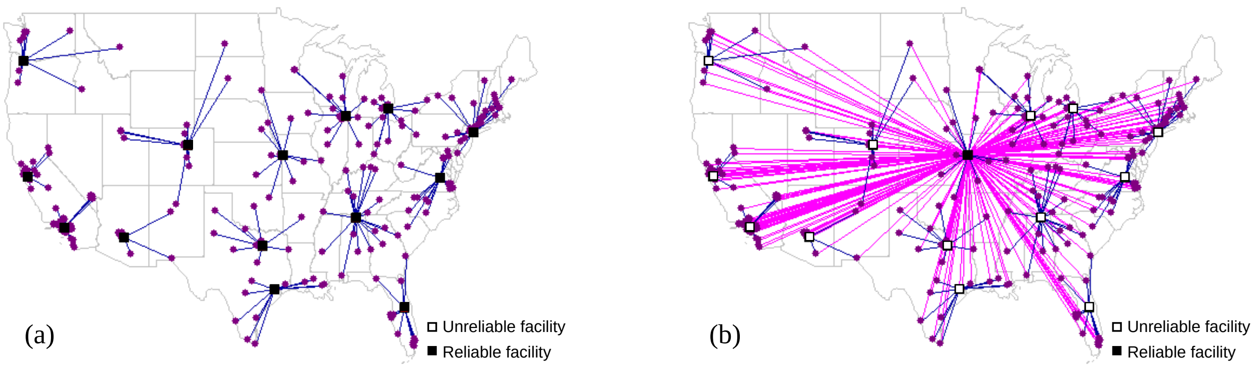

Figure 3: The solutions to the modified case examples: (a) when $q_j = q = 0.2$, (b) when $q_j = q = 0.01$

5.2 Comparison with CPLEX Solver

To compare the performance of CPLEX and the LR algorithm, we applied both algorithms to instances of different sizes ranging from 100 nodes to 1000 nodes. For each instance size, we randomly generated 5 different sets of demands. For a problem with $|N|$ demand nodes, the demand of node i in instance k, h_i^k, was drawn from a uniform distribution with $h_{low} = h(N)$ and $h_{hi} = h(1)$, where $h(i)$ is the true population of the i^{th} most populous county in the United States. The fixed cost of a reliable facility was \$1,000,000 while the fixed cost of an unreliable site was \$500,000 regardless of their locations. The disruption probability was 0.05 for all sites in all runs. The Great circle distances were multiplied by $c = 0.001$ to obtain the primary unit transportation cost, d_{ij}^P. The backup unit transportation cost was calculated as $d_{ij}^B = 1.5\, d_{ij}^P$. The CPLEX code was written by AMPL (Fourer et al., 2002) in a straightforward manner without using any decomposition technique and only the pure CPU time for the CPLEX solver was measured. For the LR algorithm, none of the optional local improvement heuristics were used. The average of 5 replication for this experiment is summarized in Table 2. The quality of LR algorithm solution is presented in two measures: solution performance and optimality gap. The solution performance represents the percentage difference between the CPLEX optimal solution and the LR algorithm solution. LR optimality gap represents the gap between the lower and upper bound of the LR algorithm. Each observation for CPU time and memory usage is graphically shown in Figure 4.

Table 2: Summary of algorithm comparison between CPLEX and LR

		100	200	300	400	500	600	700	800	900	1000
CPU time:	CPLEX	2	8	22	54	99	227	449	887	N/A	N/A
(sec.)	LR	1	5	13	33	54	78	92	108	116	127
Memory:	CPLEX	57	201	435	887	1296	1878	2699	4175	N/A	N/A
(MB)	LR	44	49	53	56	59	61	64	66	69	71
LR solution quality:											
- Solution performance (%)		-	-	-	-	-	0.002	0.005	0.012	N/A	N/A
- Optimality gap (%)		-	-	-	0.002	0.007	0.011	0.018	0.034	0.022	0.040

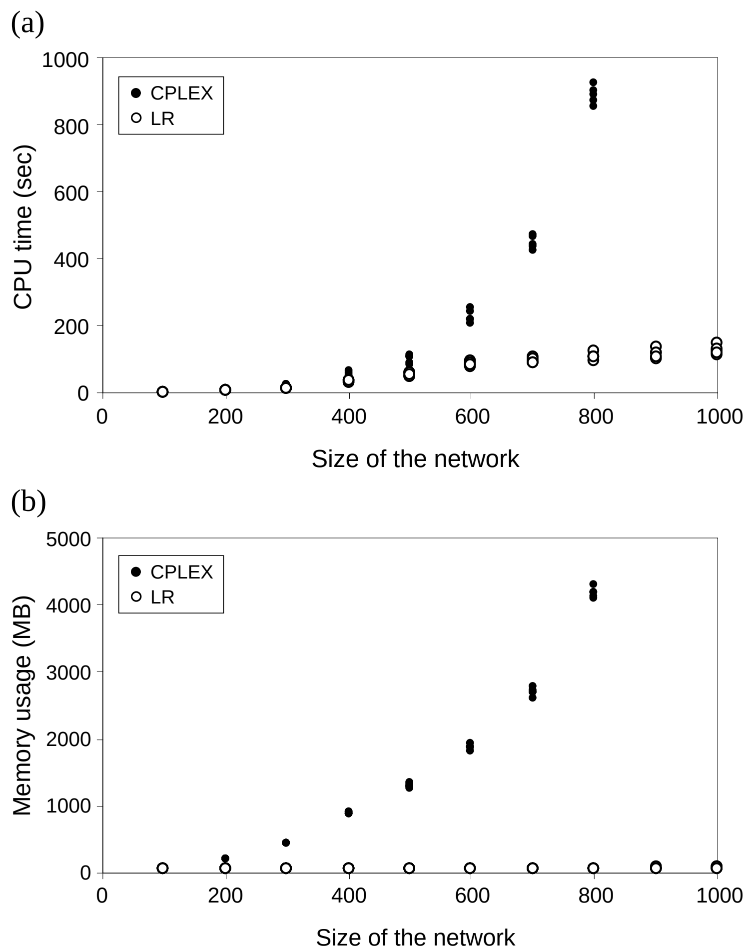

Figure 4: CPLEX and LR algorithm performance comparison on: (a) CPU time (sec.), (b) memory usage (MB)

The LR algorithm solved the problem quickly and effectively obtaining near-optimal solutions.

(In many instances, LR obtained the optimal solution as well). On the other hand, the CPU time and the memory usage for CPLEX increased rapidly with the size of the problem. We estimated the CPU time as a function of the problem size for both algorithms and found:

- CPLEX CPU time $= 1.8927e^{0.0082|N|}$ with $R^2 = 95.1\%$,
- LR CPU time $= 9.9741e^{0.0020|N|}$ with $R^2 = 84.4\%$.

Note that the LR regression time includes the data from the next section in which $|N|$ extends to 3000 nodes. The CPU time for both algorithms grew exponentially with the size of the problem. However, we observe that the CPLEX time increased more rapidly. More importantly, we observe that the LR algorithm uses computer memory far more efficiently and has a competitive edge over CPLEX for solving large scale problems. The memory usage for CPLEX increased drastically, and as such the server was not able to handle the computations beyond 800 nodes. In contrast, the LR algorithm required much less memory.

5.3 Algorithm Performance for the Larger Instances

We applied the LR algorithm to even larger instances ranging from 500 nodes to 3000 nodes on the same data set. While the county population was used for the demand h_i, all other conditions remained the same as in the the previous tests. The LR algorithm also proved to be quite effective for large instances; it solved a 3,000 node instance in less than 40 minutes as shown in Table 3. Also, the optimality gaps in all instances were still less than 0.1%. Table 4 in Appendix A summarizes the detailed statistics. In all cases, all of the reliable nodes were located at one of the largest 100 nodes (in terms of the county population). Half or more of the unreliable sites were also located in the top 100 nodes. Finally, we note that while the number of reliable and unreliable sites remains the same for instances with 2000 or more nodes, the selected locations differ (except for the 3000 node instance).

Table 3: Algorithm performance for larger instances

Number of nodes	500	1000	1500	2000	2500	3000
CPU Time (sec.)	54	123	258	661	1,693	2,301
Optimality gap (%)	0.010%	0.035%	0.007%	0.007%	0.053%	0.043%

6. Discussion and Conclusions

In this paper, we presented a facility reliability problem in the presence of random facility disruptions. We formulated the problem of determining how many hardened and non-hardened sites to use as a mixed integer programming problem. We also outlined a solution algorithm based on Lagrangian relaxation. Computational studies showed that the algorithm is fast (in terms of CPU time) and efficient (in terms of memory usage) with a small optimality gap. We also showed that this algorithm has a significant competitive edge over CPLEX particularly for solving large-scale problems.

Our findings pose new questions and motivate additional future research. First, we have ignored capacity issues in this study. While uncapacitated models provide valuable insights and guidelines at the strategic level, exploring the effect of capacity issues would be interesting. We believe this study provides a good starting point for research in this direction. Second, we do not account for partial facility disruptions as well as partial hardening options (with less investment). While this problem may be more realistic, some of the concepts in the problem have to be redefined such as what backup assignments mean, how many backups are enough if no facility is totally reliable, and so on. Lastly, it would be desirable to extend the model to cases in which the disruption probabilities are not independent, but rather are correlated. This would enable us to construct a network that is more robust against the worst case scenario.

Acknowledgments

The research was supported in part by the National Science Foundation (DMI-0457503) and the Office of Naval Research (N00014-05-1-0190). This support is gratefully appreciated. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the Office of Naval Research.

References

[1] AP (Associated Press), Parts shortages cause Ford shutdown, Sep 14, 2001.

[2] R. Batta, J.M. Dolan, and N.N. Krishnamurthy, The maximal expected covering location problem - revisited, Transport Sci 23(4) (1989), 277–287.

[3] R. Batta, and N.R. Mannur, Covering-location models for emergency situations that require multiple response units, Manage Sci 36(1) (1990), 16–23.

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

[5] Businessweek, SARS: The Sequel? Oct 27, 2003.

[6] S. Chopra, G. Reinhardt, and U. Mohan, The importance of decoupling recurrent and disruption risks in a supply chain, Nav Res Log 54(5) (2007), 544–555.

[7] R.L. Church, and M.P. Scaparra, Protecting critical assets: The r-interdiction median problem with fortification, Geogr Anal 39(2) (2007), 129–146.

[8] L. Coleman, Frequency of man-made disasters in the 20th century, J Conting Crisis Manage 14(1) (2006), 3–11.

[9] H. Crowder, "Computational improvements for subgradient optimization," Symposia Mathematica 19, Academic Press, New York, 1976, pp. 357–371.

[10] T. Cui, Y. Ouyang, and Z.-J.M. Shen, Reliable facility location design under the risk of disruptions. Working paper, Department of Industrial Engineering and Operations Research, University of California at Berkeley, 2008.

[11] M.S. Daskin, A maximum expected covering location model: Formulation, properties and heuristic solution, Transport Sci 17(1) (1983), 48–70.

[12] M.S. Daskin, Network and discrete location: Models, algorithms and applications, John Wiley and Sons, Inc., New York, 1995.

[13] Z. Drezner, Heuristic solution methods for two location problems with unreliable facilities, J Oper Res Soc 38(6) (1987), 509–514.

[14] Z. Drezner, and H.W. Hamacher, Facility location: Application and theory, Springer, Berlin, Germany, 2002.

[15] M.L. Fisher, The Lagrangian relaxation method for solving integer programming problems, Manage Sci 27(1) (1981), 1–18.

[16] M.L. Fisher, An applications oriented guide to Lagrangian relaxation, Interfaces 15(2) (1985), 10–21.

[17] R. Fourer, D.M. Gay, and B.W. Kernighan, AMPL: A modeling language for mathematical programming, Duxbury Press, California, 2002.

[18] K. Hogan, and C. ReVelle, Concepts and applications of backup coverage, Manage Sci 32(11) (1986), 1434–1444.

[19] M. Lim, M.S. Daskin, A. Bassamboo, and S. Chopra, Facility location decisions in supply chain networks with random disruption and imperfect information. Working paper, Department of Industrial Engineering and Management Sciences, Northwestern University, 2009.

[20] H. Pirkul, The uncapacitated facility location problem with primary and secondary facility requirements, IIE Trans 21(4) (1989), 337–348.

[21] H. Pirkul, and D.A. Schilling, The siting of emergency service facilities with workload capacities and backup service, Manage Sci 34(7) (1988), 896–908.

[22] C. ReVelle, and K. Hogan, The maximum availability location problem, Transport Sci 23(3) (1989), 192–200.

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

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

[25] L.V. Snyder, and M.S. Daskin, Stochastic p-robust location problems, IIE Trans 38(11) (2006), 971–985.

[26] L.V. Snyder, P.M. Scaparra, M.S. Daskin, and R.L. Church, Planning for disruptions in supply chain networks, TutORials in Operations Research, INFORMS, 2006.

[27] Swiss Re, Focus report: Disaster risk financing, 2008. Available online via <http://www.preventionweb.net/files/3224_3224SwissRePubl08FRDisasterriskfinancing.pdf>.

[28] Swiss Re, Sigma report: Natural catastrophes and Man-made disasters, 2009. Available online via <http://www.swissre.com/resources/dd6346004d4e9669ac76eecedd316cf3-sigma2_2009_e.pdf>.

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

[30] R.L. Zhan, Z.-J.M. Shen, and M.S. Daskin. System reliability with location-specific failure probabilities. Working paper, Department of Industrial Engineering and Operations Research, University of California at Berkeley, 2007.

Appendix

Appendix A: Summary of FRP solution for larger instances

Table 4: Summary of FRP solution for larger size of instances (full version)

Number of nodes	100	250	500	1000	1500	2000	2500	3000
Demand (Population)	118,984,152	171,100,691	208,462,630	242,353,743	259,782,739	270,168,542	276,441,828	279,411,542
Demand share (%)	42.6%	61.2%	74.6%	86.7%	92.9%	96.6%	98.9%	99.9%
Upper bound	18,964,459	25,896,416	30,761,118	35,695,827	38,240,640	39,829,714	40,851,890	41,486,978
Lower bound	18,964,393	25,884,330	30,757,975	35,683,167	38,237,982	39,826,931	40,830,200	41,468,958
Optimality gap (%)	0.000%	0.047%	0.010%	0.035%	0.007%	0.007%	0.053%	0.043%
Iterations	325	471	519	505	466	596	1,055	676
Time (sec.)	15	27	54	123	258	661	1,693	2,301
	Unreliable facilities							
Number of unrel. facilities	10	15	18	18	20	21	21	21
Ordered by the size of the population	2 3 10 11 14 18 20 42 47 87	2 3 10 14 18 20 31 33 42 69 90 92 99 173 249	2 3 14 18 20 29 31 33 42 69 87 99 129 173 231 249 284 430	2 3 14 18 20 29 31 55 99 129 166 173 215 238 262 284 519 815	2 3 14 18 23 29 31 33 37 43 55 99 129 173 238 262 284 475 519 815	2 3 14 18 23 29 31 33 37 43 50 92 99 129 238 262 284 475 519 815 1812	2 3 14 18 23 29 31 33 37 43 50 92 99 129 238 284 475 519 815 868 1812	2 3 14 18 23 29 31 33 37 43 50 92 99 129 279 284 475 519 815 868 1812
Unrel. facilities in top 100 nodes (%)	100.0%	86.7%	66.7%	50.0%	60.0%	61.9%	61.9%	61.9%
	Reliable facilities							
Number of rel. facilities	5	5	5	6	6	6	6	6
Ordered by the size of the population	0 1 9 16 34	0 1 9 34 86	0 1 9 86 262	0 1 9 69 86 87	0 1 9 53 69 86	0 1 9 53 69 86	0 1 9 53 69 86	0 1 9 53 69 86
Rel. facilities in top 100 nodes (%)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Cost structure							
Unrel. facility cost (%)	26.4%	29.0%	29.3%	25.2%	26.2%	26.4%	25.7%	25.3%
Rel. facility cost (%)	26.4%	19.3%	16.3%	16.8%	15.7%	15.1%	14.7%	14.5%
Primary trans. cost (%)	38.8%	41.5%	43.3%	47.6%	47.6%	47.9%	48.8%	49.4%
Backup trans. cost (%)	8.9%	10.6%	11.6%	10.8%	11.0%	11.1%	11.2%	11.2%
Cost savings (%)	-0.5%	-0.4%	-0.4%	-0.4%	-0.4%	-0.4%	-0.4%	-0.4%
Total trans. cost (%)	47.3%	51.7%	54.5%	58.0%	58.2%	58.6%	59.6%	60.2%
Total cost (%)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Cities ordered in population size: 0 New York, NY; 1 Los Angeles, CA; 2 Chicago, IL; 3 Houston, TX; 4 Philadelphia, PA; 5 Phoenix, AZ; 6 San Diego, CA; 7 Dallas, TX; 8 San Antonio, TX; 9 Detroit, MI; 10 San Jose, CA; and so on.